ANALYSIS AND VISUALIZATION OF (LARGE) NETWORKS

Vladimir BATAGELJ

University of Ljubljana, FMF, Department of Mathematics
&
IMFM Ljubljana, Department of Theoretical Computer Science, SI-1000 Ljubljana
Jadranska 19, Slovenia
vladimir.batagelj@fmf.uni-lj.si

We shall start with a presentation of some typical and well-known real-life networks. After introducing the fundamental concepts of network analysis the following topics will be presented:

- network representations: graph, matrix;
- types of networks: undirected/directed networks, multi-relational networks, 2-mode networks, temporal networks;
- size and density: small, large and huge networks; sparse and dense networks;
- description of networks in Pajek input file;
- network layouts: automatic and manual drawing;
- connection with statistical packages (SPSS, R);
- paths in networks: the shortest path; diameter; k-neighbours; acyclic networks;
- centrality: degree, closeness, betweenness; hubs and authorities, clustering coefficient; small world and scale-free networks;
- weights and properties: line and vertex cuts, sub-networks;
- connectivity: weakly, strongly and bi-connected components;
- global and local views; contraction; extraction;
- cohesion: triads, cliques, rings, cores, islands;
- 2-mode networks: direct analysis of 2-mode networks; network multiplication; transforming 2-mode to 1-mode networks;
- clustering and blockmodeling.

In examples we shall use program Pajek.

INVESTIGATION OF THE PLANT HORMONE ACTIVITY BY THEORETICAL METHODS

Branimir BERTOŠA, Sanja TOMIĆ, Biserka KOJIĆ-PRODIĆ

Ruđer Bošković Institute, HR-10002 Zagreb, POB 180, Croatia
bbertosa@irb.hr

Plant hormones are a group of naturally occurring, organic substances which influence physiological processes in plants at concentrations far below those where either nutrients or vitamins would affect these processes. The first discovered class of plant hormones are auxins. They control and regulate plant growth and development through complex signalling pathways. A great amount of experimental data on auxin activity of different compounds enabled construction of different Quantitative Structure–Activity Relationship (QSAR) models and the crystal structure of potential auxin receptor enabled docking study in order to investigate and understand auxin activity at molecular level.

Molecular Interaction Fields (MIF) were calculated for 92 molecules for which experimental data on auxin activity were available and initial 3D–QSAR model was obtained. The biological activity of a compound depends on its ability to specifically bind to the target receptors, as well as on its ADME properties (Absorption, Distribution, Metabolism, Excretion). A challenge to prediction of biological activity is to consider both these types of contributions simultaneously in deriving quantitative models. In order to at least partly include the bioavailability, initial 3D–QSAR model was improved by considering also the lipophilicity of the compounds. Log*P* and log*D* values were calculated using standard programs, but also using our own models for log*P* prediction. Distributions of log*P* and log*D* values clearly show that auxin activity of the compounds is largely influenced by their lipophilicity. The final 3D-QSAR model,[1] which beside structural features, also partly considers bioavailability, enabled classification of 92 compounds in accord with the measured biological activities and the results provided explanation for different auxin activity of some structurally similar molecules.

In order to investigate auxin activity at molecular level, interactions of auxin-related compounds with Auxin Binding Protein 1 (ABP1), potential auxin receptor, were studied using Monte Carlo (MC) algorithm, Molecular Dynamics (MD) simulations and Random Acceleration Molecular Dynamic (RAMD) simulations. The results of simulations (altogether 70-80 ns of simulations were carried out) led to the model of ABP1 mechanism[2] which is in agreement with the experimental observations. According to the model ABP1 is an auxin receptor that can exist in two conformations that differ primarily in the position of the C-terminus and one of these conformations is stabilised by auxin binding. RAMD simulations revealed the routes used by auxin molecule to enter and to leave the ABP1 binding site. Proposed model of auxin induced ABP1 function also gives a possible explanation for the inhibitory behaviour of the antiauxin compounds. Further, the results of simulations point out to possible proton shuffle mechanism in which the zinc ion from the active site acts as a strong Lewis acid that helps deprotonation of coordinated water molecule and protonation of an auxin carboxyl group.

1. B. Bertoša, B. Kojić-Prodić, M. Ramek, S. Piperaki, A. Tsantili-Kakoulidou, R. Wade, S. Tomić, *J. Chem. Inf. Comput. Sci.* **43** (2003) 1532-1541.
2. B. Bertoša, B. Kojić-Prodić, R. C. Wade, S. Tomić, *Biophys. J.* **94** (2008) 27-37.

COMPUTER SIMULATIONS OF FREE ENERGY

Urban BREN

National Institute of Chemistry, SI-1001 Ljubljana, Hajdrihova 19, Slovenia
urban.bren@fov.uni-mb.si

Gibbs free energy G is the thermodynamic potential of the $\Delta(N, P, T)$ ensemble. Therefore, it dictates direction of processes at typical laboratory conditions of constant temperature T, pressure P, and number of particles N. Since similar conditions apply to living cells, free energy dictates direction of biochemical processes as well.

The first example to be presented is calculation of binding constants for possible inhibitors of a biological target. As we are dealing with macromolecules, empirical force fields have to be applied. I will present development of empirical force field parameters for small molecules. Free energy is a state function. Therefore, free energy differences depend only on the choice of starting and final states and are independent of the selection of the path connecting them. This feature is exploited in the construction of thermodynamic cycles – theoretical tools that express free-energy differences associated with complex realistic processes with the help of free-energy changes associated with simple non-physical processes. I will introduce two rigorous free-energy methods (Free Energy Perturbation FEP and Thermodynamic Integration TI) and three approximate free-energy methods (Slow Growth, Linear Response Approximation LRA and Linear Interaction Energy LIE). All of them are used in conjunction with molecular dynamics simulations. A discussion on free-energy decomposition will follow.

The second example to be presented is calculation of reaction rate constants which directly correspond to activation free energies. As we are dealing with chemical reactions, quantum calculations have to be applied. I will introduce the concept of the Potential of Mean Force PMF. Solvent effects have to be incorporated through implicit models (Solvent Reaction Field SCRF and Langevin Dipoles LD). Chemical reactions taking place in complex biochemical environment can be simulated with the help of mixed quantum-mechanical/molecular-mechanical QM/MM approaches. The oldest QM/MM method Empirical Valence Bond EVB and its application to enzymatic reactions will be discussed.

This course should provide basic introduction to computer simulations of free energy – a field of significant importance in synthetic, biological as well as medicinal chemistry.

COMPUTING IN DRUG DISCOVERY – *PART I*

Darko BUTINA

ChemoMine Consultancy Limited, 201 Icknield Way, Letchworth Garden City, Herts SG6 4TT, UK
darko.butina@chemomine.co.uk

Drug discovery is very complex, takes long time and costs huge amount of money. In the first part of this three parts workshop, we will dissect the main processes that are going on during drug discovery processes and the role of computing. We will discuss the role of combinatorial chemistry where millions of the structures could be generated in fully automated way and use of robotics, high throughput biology where thousands of biological datapoints could be obtain in few days. Was that success, did it brought more new drugs and in shorter time to the market and how did it affect computational side of the drug discovery. Has the emerging of new hardware and software that *understands* chemical structure started to change the basic principles of drug discovery?

COMPUTING IN DRUG DISCOVERY – *PART II*

Darko BUTINA

ChemoMine Consultancy Limited, 201 Icknield Way, Letchworth Garden City, Herts SG6 4TT, UK
darko.butina@chemomine.co.uk

The key to use of computing in drug discovery is the understanding of how the changes in chemical structure, that medicinal chemists are doing, is affecting target activity or some property of the molecules in question. To achieve that, one needs:

1. Way of describing chemical structure, *i.e.*, descriptors
2. Experimental data (activity or property) to model or simulate
3. Statistical tools to find the potential relationship between the chemical structure and the information from the experiments

Each of those three steps represents separate problem in itself. Chemical structure can be described in thousands of different ways and the chemical descriptors broadly fall within either 2D or 3D class. The quality and precision of experimental data can enormously vary depending on the experimental protocol, with by far the worst problem being whether experimental data measured are truly reflecting activity or property in humans. The last, but by no means least of the problem is to use an appropriate statistical approach to enable us to find a pattern(s) and some sort of correlation between the chemical structure and the biological event that we measure. One of the most important issue in building quantitative structure activity or property relationships (QSAR or QSPR) is design of the training and test sets to prevent over-fitting or over-learning models.

COMPUTING IN DRUG DISCOVERY – *PART II*

Darko BUTINA

ChemoMine Consultancy Limited, 201 Icknield Way, Letchworth Garden City, Herts SG6 4TT, UK
darko.butina@chemomine.co.uk

The key to use of computing in drug discovery is the understanding of how the changes in chemical structure, that medicinal chemists are doing, is affecting target activity or some property of the molecules in question. To achieve that, one needs:

4. Way of describing chemical structure, *i.e.*, descriptors
5. Experimental data (activity or property) to model or simulate
6. Statistical tools to find the potential relationship between the chemical structure and the information from the experiments

Each of those three steps represents separate problem in itself. Chemical structure can be described in thousands of different ways and the chemical descriptors broadly fall within either 2D or 3D class. The quality and precision of experimental data can enormously vary depending on the experimental protocol, with by far the worst problem being whether experimental data measured are truly reflecting activity or property in humans. The last, but by no means least of the problem is to use an appropriate statistical approach to enable us to find a pattern(s) and some sort of correlation between the chemical structure and the biological event that we measure. One of the most important issue in building quantitative structure activity or property relationships (QSAR or QSPR) is design of the training and test sets to prevent over-fitting or over-learning models.

HARMONIUM ATOMS, WIGNER MOLECULES, COULOMB CRYSTALS, QUANTUM DOTS, AND COLD PLASMAS

Jerzy CIOSLOWSKI

Institute of Physics, University of Szczecin, Wielkopolska 15, 70-451 Szczecin, Poland
jerzy@kyoko.chem.fsu.edu

This course deals with various properies of systems described by Coulombic-harmonic potentials organized as follows:

I. Introduction
- *1. Pairwise-Additive Potentials*
- *2. The Coulombic-Harmonic Potential*
- *3. Harmonium Atoms*
- *4. Wigner Molecules*

II. Geometries, Energies, and Vibrations of Small Coulomb Crystals
- *1. Coulomb Crystals: Theory*
- *2. The Three- and Four-Particle Coulomb Crystals*
- *3. The Six-Particle Coulomb Crystals*
- *4. The Five and Seven-Particle Coulomb Crystals*
- *5. The Eight-Particle Coulomb Crystals*
- *6. Summary*

III. Properties of Coulomb Crystals: Rigorous Results
- *1. Pairwise-Additive Double-Power-Law Potentials*
- *2. Bounds for the Energy per Particle Pair*
- *3. Lower Bound for the Mean Reciprocal Radius of Coulomb Crystals*
- *4. Harmonic Vibrational Spectra of Coulomb Crystals*
- *5. Summary*

IV. The Shell Model of Coulomb Crystals

V. The Three-Electron Harmonium Atom
- *1. The Bosonic State*
- *2. The Quartet State*
- *3. The Doublet State*

VI. The Four-Electron Harmonium Atom
- *1. The Bosonic State*
- *2. The Fermionic States*
- *3. The Quintet State*
- *4. The Triplet State*
- *5. The Singlet State*

VII. Harmonium Atoms: A Summary
- *1. The Three-Electron Harmonium Atom*
- *2. The Four- Electron Harmonium Atom*

CLUJ-STYLE IN MODELING NANOSTRUCTURES

Mircea V. DIUDEA

Faculty of Chemistry & Chemical Engineering, Babes-Bolyai University, RO-400028 Cluj, Romania
diudea@chem.ubbcluj.ro

In modeling molecules, particularly nanostructures, scientists often use the embedding of a polygonal lattice in a given 3D surface *S*. Such a *combinatorial* surface is called a map.[1] The coordinates of the lattice points are obtained by partitioning *S*, either by dedicated algorithms or by simply drawing vertices and edges on display, by the aid of some builders to switch from 2D to 3D. Another way uses templates, or unit blocks with a prescribed spatial arrangement. This last technique is used in self assembling reactions, in experiments (or occurring *in vivo*).

Crystallography is another domain of great importance which searches infinite polyhedral coverings appearing in crystalline state of matter.

It is well established that covering/tessellating/tiling of fullerenes (nanostructures, in general) dictates the stability and reactivity of these molecules. Covering modification is one of the ways in understanding chemical reactions occurring in carbon nanostructures, which finally provide functionalized materials.

TOPO GROUP CLUJ has developed four main software programs dedicated to polyhedral tessellation and embedment in surfaces of various genera, either as finite or infinite lattices: TORUS,[2] CageVersatile (CVNET),[3] JSCHEM[4] and OMEGA.[5]

Covering the cylinder or the torus by hexagons is most often achieved by the *graphite zone-folding* procedure. [6,7] The method defines an equivalent planar parallelogram on the graphite sheet and identifies a pair of opposite sides to form a tube. Finally the two ends of the tube are glued in order to form a torus.

A second procedure uses the so-called *topological coordinates*, extracted from the adjacency matrix eigenvectors.[8,9]

Construction of polyhex (6,3) nanotubes and tori from square tiled (4,4) objects, is a *third main route*. [10-12] We implemented the cutting and graphite zone-folding procedures in our TORUS and JSCHEM original programs.[13,14]

Some geometrical-topological transformations, called operations on maps,[1,15,16] are used to relate parents and transformed associate graphs of fullerenes (in general, nanostructures). In this respect, operations such as: dualization *Du*, medial *Me*, truncation *Tr*, polygonal P_r capping or Snub *Sn*, are well known. Our original software CageVersatile (CVNET),[17] enables such operations and proved to be a useful tool.

The idea of increasing aromaticity/stability of fullerenes tessellated by disjoint flowers *DFws*, is rooted in the classical texts of Clar,[18,19] which postulated *disjoint benzenoid rings* as a criterion for the full aromatic conjugation.[20] Molecular structures, showing such fully *resonant sextets* are expected to be extremely stable, according to the VB theory.

Composite and generalized map operations or even sequences of such operations can be used to change the initial covering of a nanostructure, which eventually lead to non-classical *circulene-disjoint* coverings,[11,12] which could revive the attempts of *artificial* synthesis of fullerenes by wet chemistry.

A polygonal motif, encountered within a 3D network, is called a *tiling*[21,22] and it fills the space by tiles sharing faces. The operations providing such units are the same as for the maps but working now on *all rings* of a 3D lattice. We name these *operations on nets N*, which is a combinatorial representation of a 3D domain. Such operations are also available by our CVNET software.

There exist negatively curved graphitic structures,[11,12] which can be described as infinite triply periodic minimal surfaces *TPMSs*, also called schwarzites, in the honor of H. A. Schwarz,[23,24] who first investigated, in the early 19th century, the differential geometry of this class of structures. Schwarzites may appear during the fullerene and nanotube synthesis, as random (amorphous) spongy carbon materials.[25,26]

Spongy carbon and also nano-diamond can be modeled, by our CVNET software, as finite and infinite structures and nanotube junctions as well.[12]

Resuming, nanostructures of dimensionality from zero (fullerenes) - to three (spongy carbon) can be designed by the software programs developed at the TOPO GROUP CLUJ and further addressed to energetic characterization by comercial quantum chemistry software.

1. T. Pisanski, M. Randić, in: *Geometry at Work*, (C.A. Gorini, ed.) M. A. A. Notes **53** (2000) 174-194.
2. M.V. Diudea, B. Parv, O. Ursu, TORUS 3.0, Babes-Bolyai University, Cluj, 2003.
3. M. Stefu, M.V. Diudea, *CageVersatile* 1.5, Babes-Bolyai University, Cluj, 2005.
4. Cs.L. Nagy, M.V. Diudea, *JSChem* 3.5, Babes-Bolyai University, Cluj, 2005.
5. S. Cigher, M.V. Diudea, *Omega* 1.0, Babes-Bolyai University, Cluj, 2006.
6. A. Ceulemans, L.F. Chibotaru, P.W. Fowler, *Phys. Rev. Lett*. **80** (1998) 1861-1864.
7. E.C. Kirby, R.B. Mallion, P. Pollak, *J. Chem. Soc. Faraday Trans*. **89** (1993)1945-1953.
8. A. Graovac, D. Plavšić, M. Kaufman, T. Pisanski, E.C. Kirby, *J. Chem. Phys*. **113** (2000)1925-1931.
9. I. Laszlo, A. Rassat, P.W. Fowler, A. Graovac, *Chem. Phys. Lett*. **342** (2001) 369-374.
10. M.V. Diudea, *Bull. Chem. Soc. Jpn.* **75** (2002) 487-492.
11. M.V. Diudea, ed., *Nanostructures, Novel Architecture*, Nova, N.Y., 2005.
12. M.V. Diudea, Cs.L. Nagy, *Periodic Nanostructures*, Springer, 2007.
13. M.V. Diudea, B. Parv, O. Ursu, TORUS 3.0 Babes-Bolyai University, Cluj, 2003.
14. Cs.L. Nagy, M.V. Diudea, *JSChem* 3.5, Babes-Bolyai University, Cluj, 2005.
15. M.V. Diudea, P.E. John, A. Graovac, M. Primorac, T. Pisanski, *Croat. Chem. Acta* **76** (2003) 153-159.
16. M.V. Diudea, M. Ştefu, P.E. John, A. Graovac, *Croat. Chem. Acta* **79** (2006) 355-362.
17. M. Stefu, M. V. Diudea, *CageVersatile*, Babes-Bolyai University, Cluj, 2005.
18. E. Clar, *Polycyclic Hydrocarbons*, Academic Press, London, 1964.
19. E. Clar, *The Aromatic Sextet*, Wiley, N.Y., 1972.
20. P.W. Fowler, T. Pisanski, *J. Chem. Soc. Faraday Trans*. **90** (1994) 2865-2871.
21. O. Delgado-Friedrichs, M. O'Keeffe, *J. Solid State Chem*. **178** (2005) 2480-2485.
22. V.A. Blatov, L. Carlucci, G. Ciani, D.M. Proserpio, *Cryst. Eng. Comm*. **6** (2004) 377-395.
23. H.A. Schwarz, Über Minimalflächen, *Monatsber. Berlin Akad*., **1865**.
24. H.A. Schwarz, *Gesammelte Matematische Abhandlungen*, Springer, Berlin, **1890**.
25. E. Barborini, P. Piseri, P. Milani, G. Benedek, C. Ducati, J. Robertson, *Appl. Phys. Lett*. **81** (2002) 3359-3361.
26. G. Benedek, H. Vahedi-Tafreshi, E. Barborini, P. Piseri, P. Milani, C. Ducati, J. Robertson, *Diam. Relat. Mater*.**12** (2003) 768–773.

THEORETICAL ASPECTS OF FULLERENE CHEMISTRY

Patrick W. FOWLER

Department of Chemistry, University of Sheffield, Sheffield S3 7HF UK
P.W.Fowler@sheffield.ac.uk

Fullerene chemistry and physics is an area where three-way interaction between chemistry, mathematics and even computer science is particularly strong. In these lectures, some applications of chemical graph theory will be discussed for these chemically important polyhedral all-carbon molecules, with emphasis on the way that trends that can be obtained with the simplest of theoretical tools.

Systematics of fullerene physics and chemistry can be developed using a mixture of symmetry, combinatoric and graph theoretical arguments. Adjacency matrices and their eigenvalues (at least partially) characterise stability and reactivity of the unsaturated fullerene carbon frameworks, through the Hückel model. Isomer counts, magic numbers in electronic structure, energetics and patterns of addition chemistry can all be understood with the help of symmetry/graph theory based models. Discussion of the specifics of fullerenes can be used as a way of learning more about these simple but powerful techniques. All the ideas and models used will be introduced *from scratch*, so that no specialised background knowledge will be assumed.

The main reference used will be *An Atlas of Fullerenes* by P.W. Fowler and D.E. Manolopoulos, which was first issued by Oxford University Press in 1995 and is now (from 2006) available in Dover paperbacks (ISBN 0-486-45362-6).

SHAPE ANALYSIS OF CARBON NANOTUBE JUNCTIONS

Ante GRAOVAC[1,2], István LÁSZLÓ[3], Tomaž PISANSKI[4]

[1]*Faculty of Science, University of Split, HR-21000 Split, Nikole Tesle 12, Croatia*
[2]*Ruđer Bošković Institute, HR-10002 Zagreb, POB 180, Croatia*
graovac@irb.hr

[3]*Department of Theoretical Physics, Institute of Physics, Budapest University of Technology & Economics, H-1521 Budapest, Hungary*
laszlo@eik.bme.hu

[4]*Department of Theoretical Computer Science, Institute of Mathematics, Physics & Mechanics, University of Ljubljana, SI-1000 Ljubljana, Jadranska 19, Slovenia*
Tomaz.Pisanski@fmf.uni-lj.si

Eigenvectors of the adjacency matrix can produce satisfactory Descartes coordinates for some spherical, toroidal and planar structures where it is usually supposed that the molecules under study are spherical in two or three dimensional space. Namely the fullerenes are spheroidal and the tori can be imagined as the direct product of two circles. The question arises whether the method of the topological coordinates can be used for non-spherical structures as well? Here we will present a shape analysis of nanotube junctions in order to examine possibilities to extend the topological coordinates method to non-spherical structures. The **X**, **Y**, **Z** Descartes coordinates of the atoms were calculated as linear combinations of eigenvectors. We have obtained that the partial sums of eigenvectors generated plausible three-dimensional structures only if they contained all three bi-lobal eigenvectors of the Laplacian matrix. We have also found partial sums that produced satisfactory initial coordinates for molecular mechanics calculations.

DENSITY FUNCTIONAL THEORY METHODS

Mihai V. PUTZ

Laboratory of Computational & Structural Physical Chemistry, Chemistry Department, West University of Timişoara, RO-300115 Timişoara, Str. Pestalozzi No. 16a, Romania
mvputz@cbg.uvt.ro

Since the appearance of density functional theory (DFT) the quantum chemistry was reformulated in both its conceptual and computational sides to achieve more comprehension of structure and reactivity of electronic systems either in isolated or interacting states. In this context, the lectures will address general and specific topics so that the DFT formalisms together with its physical-chemical concepts equally will be presented at background and advanced levels. Actually, the present course is organized in three lectures as follows:

1. DFT Formalism
   - *The need of density over wave function*
   - *Thomas-Fermi foreground theory*
   - *Hohenberg-Kohn theorems*
   - *Chemical action concept*
   - *Kohn-Sham equations and the Schrödinger reminiscence*
   - *Exchange and correlation density functionals*

2. Density Functional Softness Theory
   - *Electronegativity χ and chemical hardness η concepts*
   - *Chemical reactivity equations*
   - *Softness hierarchy*
   - *Fukui functions*
   - *Systematic expressions for χ and η*
   - *Bosonic and fermionic atomic states and atomic periodicity*

3. Density Functional Chemical Reactivity
   - *Electronegativity equalization principle*
   - *Electronegativity fluctuation*
   - *Hard and soft acids and basis (HSAB) principle and its reliability through improved chemical schemes, density functional softness theory and atoms-in-molecules bosonic/fermionic states*
   - *Maximum hardness (MH) principle*
   - *Towards unification of HSAB and MH principles*
   - *Testing the quadratic $E = E(\chi, \eta)$ dependency at various density functional correlation levels*

Further perspectives are as well emphasized: from physical view in the line of a field theoretical development of the DFT formalisms, while from the chemical side the unification of the chemical reactivity principles in describing the chemical bond and bonding is finally advanced.

M.V. Putz, *Contributions within Density Functional Theory with Applications in Chemical Reactivity Theory and Electronegativity*, Dissertation (www.dissertation.com/book.php?method=ISBN&book=1581121849), Com, Parkland, Florida, 2003.
M.V. Putz, N. Russo, E. Sicilia, Atomic Radii Scale and Related Size Properties from Density Functional Electronegativity Formulation, *J. Phys. Chem.* **A107** (2003) 5461-5465.
M.V. Putz, N. Russo, E. Sicilia, On the Application of the HSAB Principle through the Use of Improved Computational Schemes for Chemical Hardness Evaluation, *J. Comp. Chem.* **25** (2004) 994-1003.
M.V. Putz, N. Russo, E. Sicilia, About the Mulliken Electronegativity in DFT, *Theoret. Chem. Acc.* **114** (2005) 38-45.
M.V. Putz, Systematic Formulation for Electronegativity and Hardness and Their Atomic Scales within Density Functional Softness Theory, *Int. J. Quantum Chem.* **106** (2006) 361-386.
M.V. Putz, Can Quantum-Mechanical Description of Chemical Bond Be Considered Complete?, Expert Commentary, in: *Quantum Chemistry Research Trends* (M.P. Kaisas, ed.), Nova Publishers, N.Y., 2007.
M.V. Putz, Unifying Absolute and Chemical Electronegativity and Hardness Density Functional Formulations through the Chemical Action Concept, in: *Progress in Quantum Chemistry Research* (E.O. Hoffman, ed.), Chapter 2, Nova Publishers, N.Y., 2007.
M.V. Putz, Semiclassical Electronegativity and Chemical Hardness, *J. Theor. Comp. Chem.* **6** (2007) 33-47.
M.V. Putz, A. Chiriac, Quantum Perspectives on the Nature of the Chemical Bond, in: *Advances in Quantum Chemical Bonding Structures*, (M.V. Putz, ed.), Chapter 1, Transworld Research Network, Kerala, India, 2007-2008, in press.
M.V. Putz, Density Functionals of Chemical Bonding, *Int. J. Mol. Sci.* (2008), (www.mdpi.org/ijms), in preparation.
M.V. Putz, Maximum Hardness Index of Quantum Acid-Base Bonding, *MATCH Commun. Math. Comput. Chem.* (2008-2009), accepted.

APPLICATIONS OF GROUP THEORY TO ATOMIC AND MOLECULAR SPECTROSCOPY

Hassan SABZYAN

Department of Chemistry, University of Isfahan, Isfahan 81746-73441, I. R. Iran
sabzyan@sci.ui.ac.ir

To chemists, it is always interesting to find and use new well-established and clearly worked out mathematical methods to solve unsolved chemical problems or to take easier approaches for solving problems already solved by other difficult, tedious and long (yet mathematical!) methods. Application of mathematical methods may also provide deeper physical intuitions and allow better understanding of the chemical and physical processes. In this course, we will review the results of Group Theory and their applications to the derivation of selection rules in spectroscopic transitions. The following topics will be covered.

1. Introduction

1.1. *Group Theory*

Definitions
Relations/Operations
Theorems
Subgroups and Classes
Products and Transformations

1.2. *Molecular Symmetry*

Symmetry Operations and Elements
Classification of Operations
Symmetry Point Groups
Matrix Representation
Reducible and Irreducible Representations
Transformations
Symmetry Subgroups and Classes
Examples

1.3. *Quantum Mechanics*

Schrödinger Equation, Eigenstates and Eigenfunctions
Symmetry of the Operators (e.g. Hamiltonian) and Its Implications
Group Theoretical Representation of Wavefunctions

1.4. *Light-Matter Interactions*

Electromagnetic Waves (Spectrum, Regions, Polarizations, Dichroism)
Perturbation Theory and Transition Dipole Moments
Selection Rules

1.5. *Questions and Answers*

2. Application of Group Theory to Spectroscopy

2.1. *General Procedures*

Building Up the Full Reducible Representation (RR) of a Molecule
Finding Irreducible Representations (IR) of RR Using Orthogonality Theorems
Decomposition of the IR into Translational, Rotional, and Vibrational Degrees of Freedom
Selection Rules for Individual Degrees of Freedom
Selection Rules for Overtones, Hot Bands and Combination Bands
Questions and Answers

2.2. *Case Study*

Examples of Free Molecules with Different Point Groups
Touching the Spherical and Infinite Point Groups
Examples of Interfacial and Adsorbed Molecules
Questions and Answers

OUR VIEW ON PROTEIN STRUCTURES IN THE CURRENT, POST-GENOMIC ERA

István SIMON

Institute of Enzymology, BRC, Hungarian Academy of Sciences, H-1518 Budapest, P.O.Box 7, Hungary
simon@enzim.hu

Our view on protein structure has been changed dramatically in the past two decades, mainly as a result of the genome projects. First, it became clear that proteins that were known twenty years ago did not provide a general representation of all proteins. Sequencing at DNA level instead of protein level resulted in new datasets that were independent from the possibility of protein isolation. The new sequencing technique resulted in a significant shift in residue composition of the protein databanks. Amino acid sequences of the *hard to isolate* transmembrane proteins were expected to considerably influence the residue distribution of sequence databases, but the effect was less apparent than expected. At that time, the number of transmembrane proteins were believed to be rather small relative to the number of water soluble globular proteins. By now, the reason for this discrepancy has become apparent: the presence of intrinsically unstructured/disordered proteins that contain very few hydrophobic residues, compensate for the effect of transmembrane proteins. The first surprising outcome of the genome projects was that the number of transmembrane proteins exceeds the quarter of the number of proteins encoded in the genomes. Later, we learned that even many of the *water soluble* proteins can not fold without appropriate template macromolecules, such as nucleic acids or proteins. Furthermore, there are proteins which do not fold at all and perform their function in an unstructured state. Besides unstructured proteins, various types of proteins become available that do not obey conventional definitions of protein structure.

It is worth to mention that these findings challenged one of the fundamental principles of protein science: the amino acid sequence of a polypeptide chain exclusively determines its 3D structure and native proteins can perform certain biochemical functions only when they adopt their unique structure. Apparently, the long-standing, unsolved basic question of how the amino acid sequence determines the 3D structure of a protein - the so-called protein folding problem - has become more complex. The sequence information has to dictate not only the structure of *traditional* uniquely folded water soluble proteins, but it also has to determine which (part of the) protein folds into a unique structure. The primary sequence also encodes the folding conditions, whether it occurs in homogeneous, anisotropic aqueous medium or in a water-membrane-water environment; or whether the folding requires a structured or unstructured template. The questions about principles governing the formation of protein structure or its absence under various circumstances became rather complicated.

The number of known protein sequences and structures has increased in the past twenty years by orders of magnitudes and extended from globular water soluble proteins to the various classes of the

protein kingdom. In principle, this could help us to find some basic principles of protein structure formation. In practice, the automatic methods developed for the genome projects currently produce data at a much higher rate than that of the complete analyses of these data. There is a real danger that in the near future we would face the same failure than the astrophysicists did a decade ago, when they realized that billions of dollars had been spent in vain for data collections, since there was capacity for analyzing only 10% of the data received from space ships, satellites and observation posts. The remaining 90% of the data have been stored in recorders, like magnetic tapes, which can not be read by current computers. Therefore, in the field of protein science there is an urgent necessity for high speed methods for data analyses and in particular, *in silico* methods for predicting the kinds of data, which were traditionally obtained by experimental investigation of individual proteins.

Of course, there is also an optimistic reading of the story. From the principles on which the *in silico* prediction methods are based and from the analyses of various classes of the protein kingdom new ideas started to emerge, related to protein structures. There is some indication that the structure determination occurs in a hierarchical manner *via* multiple steps. Most of these steps are governed by dramatically reduced sequence information, *i.e.*, information on amino acid composition of distinguished sets of residues, segments or set of segments. The detailed sequence information might be only used for the final fine tuning of the structure.

In my lectures, first I will talk about the traditional (water soluble, globular, stable, uniquely folded) proteins. Why crystallization were considered as the best way of isolation and purification. What we learned from X-ray diffraction and NMR spectroscopy about their structure. Why and how they adopt stable, uniquely folded structure. What we know about the hierarchy and the topology of these structures. Then, I will discuss the structure and topology of the transmembrane proteins. Some of the ways, the topology can be predicted and analyzed. This will be followed by a short introduction on unstructured proteins. Why they remained hidden for almost a century when the folded proteins were studied with various techniques. How we can identify them on the basis of their sequence. What we know about their function. Finally, I will discuss which kind of generalization can be taken based on our current knowledge of the various proteins types.

SOLID-STATE REACTION MECHANISMS

Hrvoj VANČIK

Department of Chemistry, Faculty of Science & Mathematics, University of Zagreb, HR-10000 Zagreb, Horvatovac 102 A, Croatia
vancik@chem.pmf.hr

Although the idea of the reaction mechanism is well established for the reactions in solutions, conceptualization of reaction mechanism in the solid-state becomes much more complex because of the following reasons:

(*i*) besides the chemical reaction, there is also the another process that include phase transformations;
(*ii*) the medium is anisotropic;
(*iii*) the mechanism strongly depends on the history of the preparation of the solid phase, and the experiments are not very strongly reproducible.

In this frame, the possibilities of the generalization of the solid-state reaction mechanism concept will be discussed.