Association for the International Development of Academic and Scientific Collaboration

LECTURE

INTRODUCTION TO COMPUTATIONAL MODELING OF MOLECULES AND PERIODIC STRUCTURES

Dr. Stevan Armaković
AIDASCO Board President

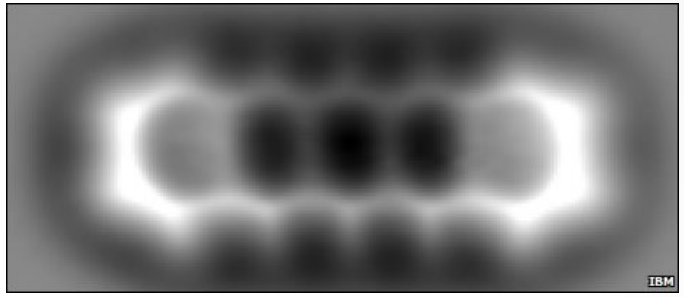
Assistant Professor
University of Novi Sad, Faculty of Sciences, Department of Physics

www.aidasco.org
www.armakovic.com
www.impactfactors.online

Can we see molecules?

2009. year

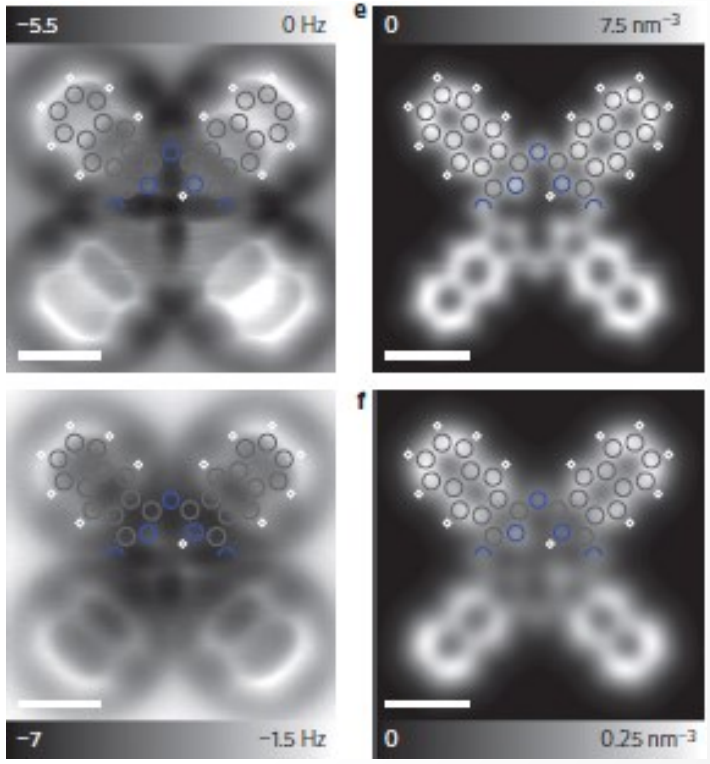
- First image of a molecule - pentacene
- Atomic force microscopy
- IBM Research Zurich



Science 325 (2009) 1110

2012. year

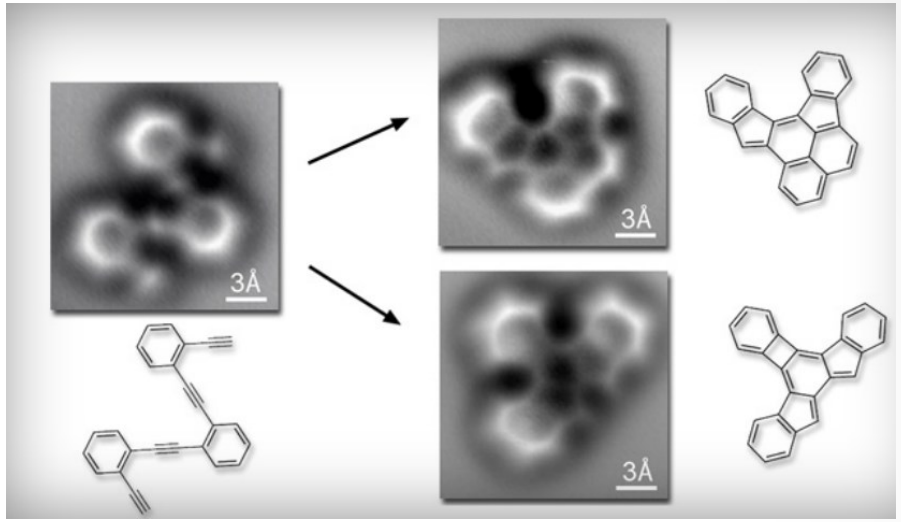
- First images of molecules which indicate the charge distribution within molecules
- Atomic Force Microscopy
- IBM Research Zurich



Nature Nanotechnology 7 (2012) 227

2013. year

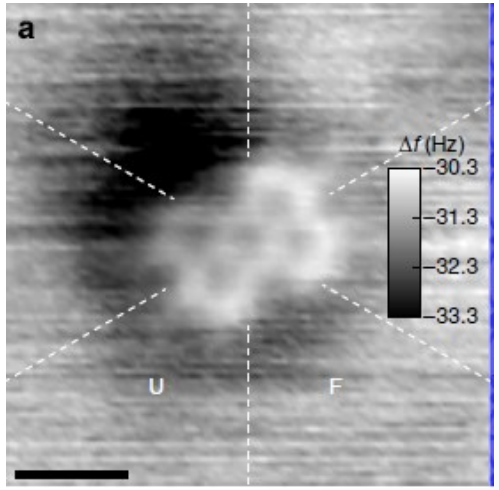
- First images of molecules, indicating the formation of bonds between atoms
- Atomic Force Microscopy
- Lawrence Berkley National Laboratory



Science 340 (2013) 1434

2015. year

- First image of PTCDA molecule
- Atomic force microscopy at room temperature
- IBM Research Zurich



Nature Communications 6 (2015) 7766

The Chemical Structure of a Molecule Resolved by Atomic Force Microscopy

Leo Gross,[1]* Fabian Mohn,[1] Nikolaj Moll,[1] Peter Liljeroth,[1,2] Gerhard Meyer[1]

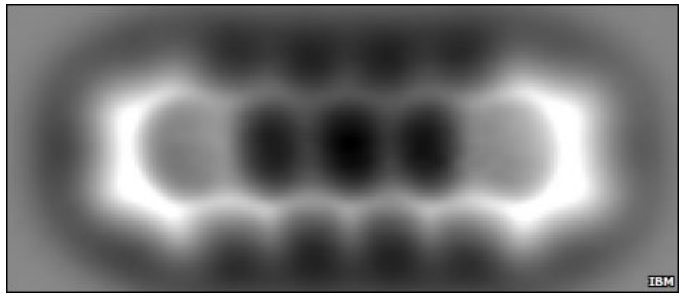
Resolving individual atoms has always been the ultimate goal of surface microscopy. The scanning tunneling microscope images atomic-scale features on surfaces, but resolving single atoms within an adsorbed molecule remains a great challenge because the tunneling current is primarily sensitive to the local electron density of states close to the Fermi level. We demonstrate imaging of molecules with unprecedented atomic resolution by probing the short-range chemical forces with use of noncontact atomic force microscopy. The key step is functionalizing the microscope's tip apex with suitable, atomically well-defined terminations, such as CO molecules. Our experimental findings are corroborated by ab initio density functional theory calculations. Comparison with theory shows that Pauli repulsion is the source of the atomic resolution, whereas van der Waals and electrostatic forces only add a diffuse attractive background.

Noncontact atomic force microscopy (NC-AFM), usually operated in frequency-modulation mode (*1*), has become an important tool in the characterization of nanostructures on the atomic scale. Recently, impressive progress has been made, including atomic resolution with chemical identification (*2*) and measurement of the magnetic exchange force with atomic resolution (*3*). Moreover, lateral (*4*) and vertical (*5*) manipulation of atoms and the determination of the vertical and lateral forces during such manipulations (*6*) have been demonstrated. Striking results have also been obtained in AFM investigations of single molecules. For example, atomic resolution was achieved on single-walled carbon nanotubes (SWNTs) (*7*, *8*), and the force needed to switch a molecular conformation was measured (*9*).

However, the complete chemical structure of an individual molecule has so far not been imaged with atomic resolution. The reasons that make AFM investigations on single molecules so challenging are the strong influence of the exact

[1]IBM Research, Zurich Research Laboratory, 8803 Rüschlikon, Switzerland. [2]Debye Institute for Nanomaterials Science, Utrecht University, Post Office Box 80000, 3508 TA Utrecht, Netherlands.

*To whom correspondence should be addressed. E-mail: lgr@zurich.ibm.com



nature nanotechnology

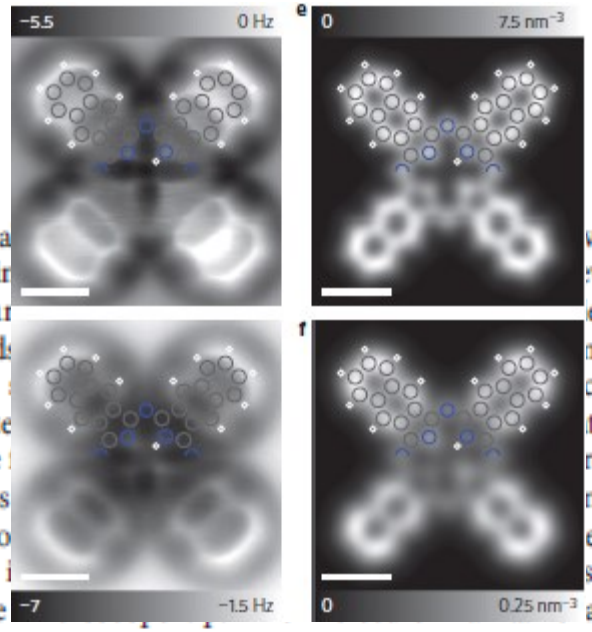
LETTERS

PUBLISHED ONLINE: 26 FEBRUARY 2012 | DOI: 10.1038/NNANO.2012.20

Imaging the charge distribution within a single molecule

Fabian Mohn*, Leo Gross, Nikolaj Moll and Gerhard Meyer

Scanning tunnelling microscopy and atomic force microscopy can be used to study the electronic and structural properties of surfaces, as well as molecules and nanostructures adsorbed on surfaces, with atomic precision[1–7], but they cannot directly probe the distribution of charge in these systems. However, another form of scanning probe microscopy, Kelvin probe force microscopy, can be used to measure the local contact potential difference between the scanning probe tip and the surface, a quantity that is closely related to the charge distribution on the surface[8–12]. Here, we use a combination of scanning tunnelling microscopy, atomic force microscopy and Kelvin probe force microscopy to examine naphthalocyanine molecules (which have been used as molecular switches[13]) on a thin insulating layer of NaCl on Cu(111). We show that Kelvin probe force microscopy can map the local contact potential difference of this system with submolecular resolution, and we use density functional theory calculations to verify that these maps reflect the intramolecular distribution of charge. This approach could help to provide fundamental insights into single-molecule switching and bond formation, processes that are usually accompanied by the redistribution of charge within or between molecules[14–16].



free-base naphtha vious
study with scannir veral
advantages for our ently
decouples the ads netal,
which prevents a ctron
states and facilitate tions
carried out for the re, by
comparing images nd to
different adsorptio e out
that the contrast i ssible
asymmetry of the a flat
adsorption geometry in both states, and the tautomerization switching involves only a minor structural change. This excludes spurious geometrical effects on the measured LCPD that might be expected for systems in which the switching of a molecule is accompanied by a major change in the molecular structure[12].

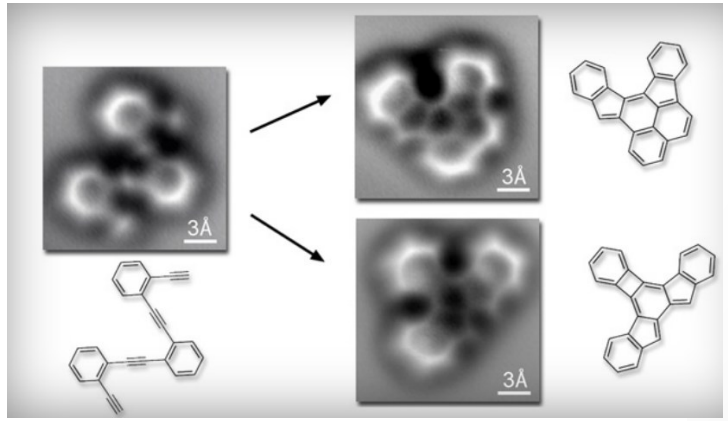
Figure 1a presents a constant-current STM image of a naphthalocyanine molecule on NaCl(2 ML)/Cu(111), recorded at a sample bias of $V = 0.25$ V. The molecule appears as a cross-like structure, with a central protrusion surrounded by four lobes. There is no apparent asymmetry between the lobes parallel (H-lobes) and perpendicular (N-lobes) to the two inner hydrogen atoms. This

Direct Imaging of Covalent Bond Structure in Single-Molecule Chemical Reactions

Dimas G. de Oteyza,[1,2]* Patrick Gorman,[3]* Yen-Chia Chen,[1,4]* Sebastian Wickenburg,[1,4] Alexander Riss,[1] Duncan J. Mowbray,[5,6] Grisha Etkin,[3] Zahra Pedramrazi,[1] Hsin-Zon Tsai,[1] Angel Rubio,[2,5,6] Michael F. Crommie,[1,4]† Felix R. Fischer[3,4]†

Observing the intricate chemical transformation of an individual molecule as it undergoes a complex reaction is a long-standing challenge in molecular imaging. Advances in scanning probe microscopy now provide the tools to visualize not only the frontier orbitals of chemical reaction partners and products, but their internal covalent bond configurations as well. We used noncontact atomic force microscopy to investigate reaction-induced changes in the detailed internal bond structure of individual oligo-(phenylene-1,2-ethynylenes) on a (100) oriented silver surface as they underwent a series of cyclization processes. Our images reveal the complex surface reaction mechanisms underlying thermally induced cyclization cascades of enediynes. Calculations using ab initio density functional theory provide additional support for the proposed reaction pathways.

Understanding the microscopic rearrangements of matter that occur during chemical reactions is of great importance for catalytic mechanisms and may lead to greater efficiencies in industrially relevant processes (*1*, *2*). However, traditional chemical structure characterization methods are typically limited to



ARTICLE

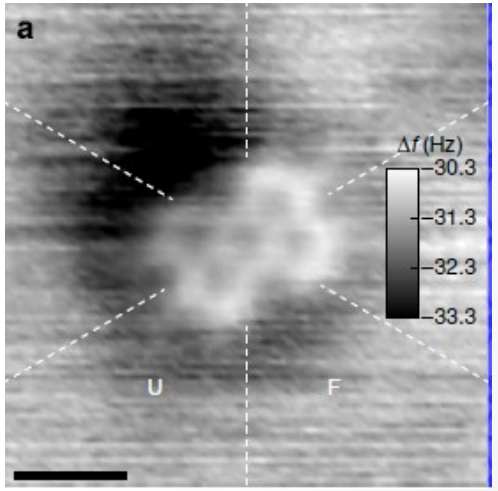
Received 28 Mar 2015 | Accepted 5 Jun 2015 | Published 16 Jul 2015

DOI: 10.1038/ncomms8766 OPEN

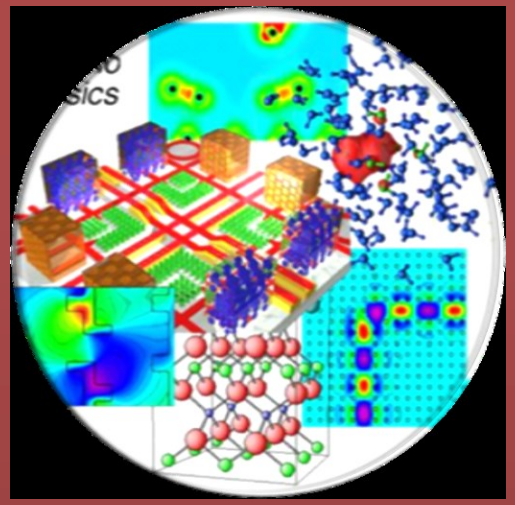
Chemical structure imaging of a single molecule by atomic force microscopy at room temperature

Kota Iwata[1], Shiro Yamazaki[2], Pingo Mutombo[3], Prokop Hapala[3], Martin Ondráček[3], Pavel Jelínek[3,1] & Yoshiaki Sugimoto[1,4]

Atomic force microscopy is capable of resolving the chemical structure of a single molecule on a surface. In previous research, such high resolution has only been obtained at low temperatures. Here we demonstrate that the chemical structure of a single molecule can be clearly revealed even at room temperature. 3,4,9,10-perylene tetracarboxylic dianhydride, which is strongly adsorbed onto a corner-hole site of a Si(111)-(7 × 7) surface in a bridge-like configuration is used for demonstration. Force spectroscopy combined with first-principle calculations clarifies that chemical structures can be resolved independent of tip reactivity. We show that the submolecular contrast over a central part of the molecule is achieved in the repulsive regime due to differences in the attractive van der Waals interaction and the Pauli repulsive interaction between different sites of the molecule.



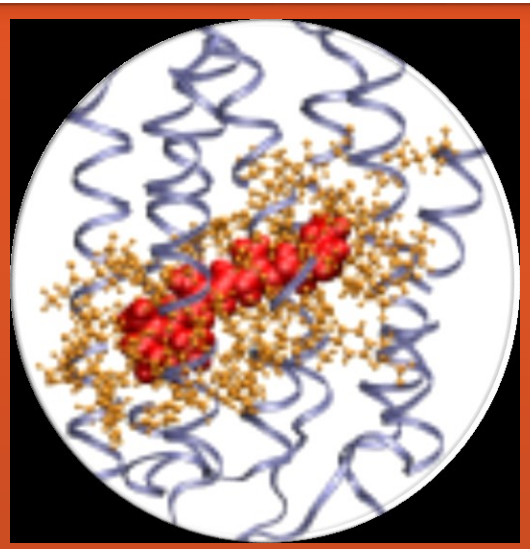
Theories and methods



Quantum-mechanical methods:

ab initio methods

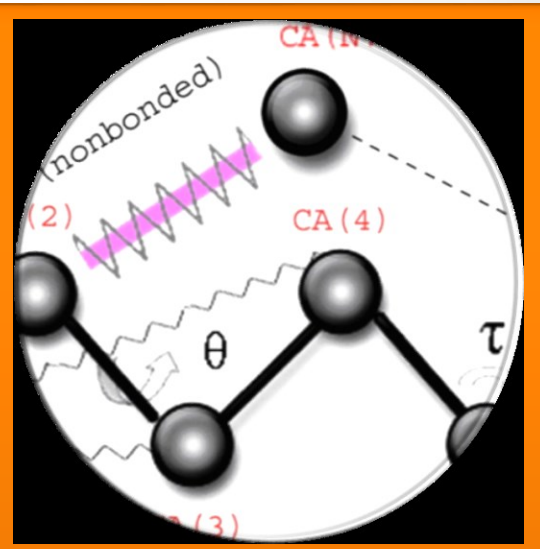
Molecule treated as a quantum mechanical object



Semi-empirical methods:

AM1, PM6, PM7, DFTB

Additional approximations and parametrizations



Molecular mechanics methods:

Charmm,OPLS…

Molecule treated as balls connected with springs

Quantum-mechanical methods

State of the quantum-mechanical object is described by the Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2}+V(x)\psi(x)=E\psi(x)$$

Hartree-Fock (HF) method

- Simplest ab initio method, "from first principles", calculations based on pure theory, without empirical approximations
- Method to determine the wave function and all properties that can be extracted from it
- Doesn't take into account the correlation between electrons

Density functional theory (DFT)

- Instead of wavefunction, electron density is used for determination of the ground state
- Structurally similar to HF methods, $\rho(r)$ instead of ψ
- Exact functional is unknown, so we have to use approximationi

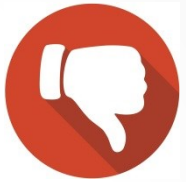
Post-HF

- Development of HF methods, to include the correlation between electrons
- Application of perturbation theory
- Demanding in terms of computational time

Semi-empiric methods

Solution of Schrodinger equation, but with empirical approximations

2-3 orders of magnitude faster than the standard DFT methods



Accuracy and efficiency depends on the parameters, which are not defined for all elements of PSE

NDDO/MNDO/CNDO

- Hartree-Fock based semi-empiric methods
- Based on the neglecting of diatomic differential overlap

Austin Method (AM1)

- MNDO-based method with improved nuclear-nuclear core repulsion
- Reparametrized, better than NDDO family of SE methods

PM3/PM6/PM7

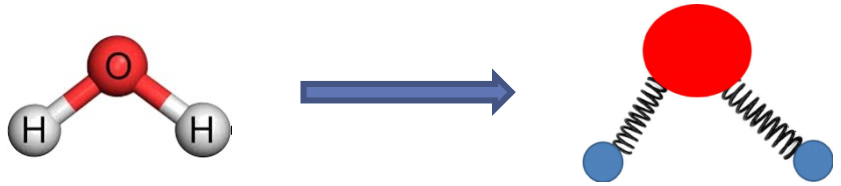
- Defined to reproduce large number of molecular properties
- Able to treat hydrogen bonds

DFTB

- DFT-based SE method, tight-binding approximation
- Taylor expansion of the energy with respect to a reference density

Molecular mechanics

Atoms in molecules are observed as balls connected with springs



$$E = \sum_{bonds} E_{stretch} + \sum_{angles} E_{bend} + \sum_{dihedrals} E_{torsion} + \sum_{pairs} E_{nonbond}$$

General form of equation for the potential energy of molecule

$$E_{stretch} = k_{stretch}(l - l_{eq})^2$$

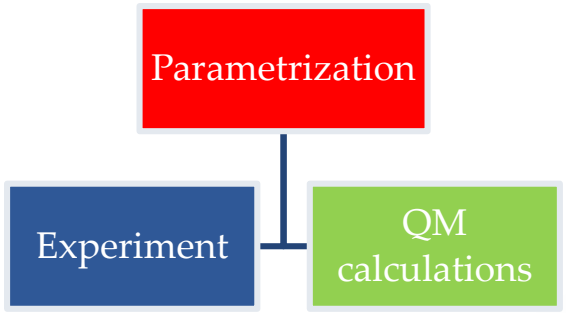
$$E_{bend} = k_{bend}(a - a_{eq})^2$$

$$E_{torsion} = k_o + \sum_{r=1}^{n} k_r[1 + \cos(r\theta)]$$

$$E_{nonbond} = k_{nb}\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$

Constants k have such values to reproduce the experimental results

FORCE FIELD

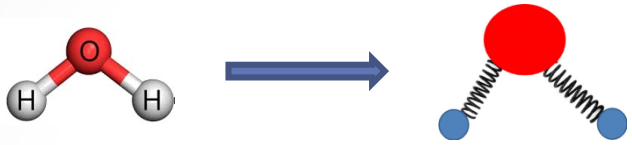


Fastest calculations

Accuracy depends on parametrization

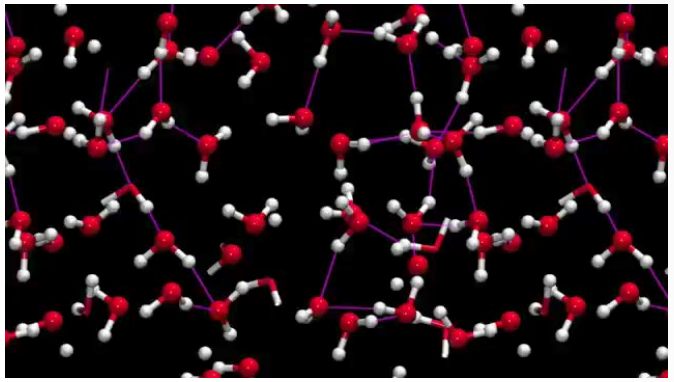
Molecular dynamics

Atoms in molecules are observed as balls connected with springs



$$E = \sum_{bonds} E_{stretch} + \sum_{angles} E_{bend} + \sum_{dihedrals} E_{torsion} + \sum_{pairs} E_{nonbond}$$

General form of equation for the potential energy of molecule



Time is divided in discrete steps, and for each step the Newton equation of motion are solved for each atom:

$$F_i = m_i \cdot a_i = m_i \cdot \frac{dv_i}{dt} = m \cdot \frac{d^2 r_i}{dt^2}$$

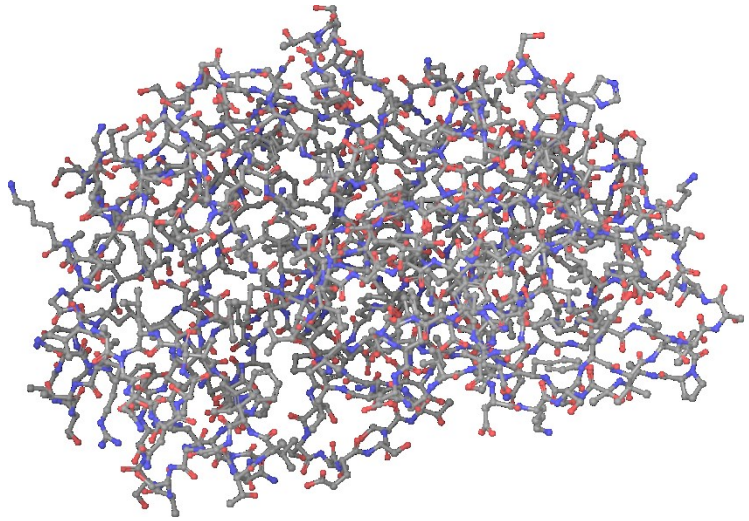
For each time interval we get the forces which act on atoms. In next step, atoms are slightly moved, and the whole procedure is repeated. In the end, we collect trajectories for each step, and analyze the motion of atoms.

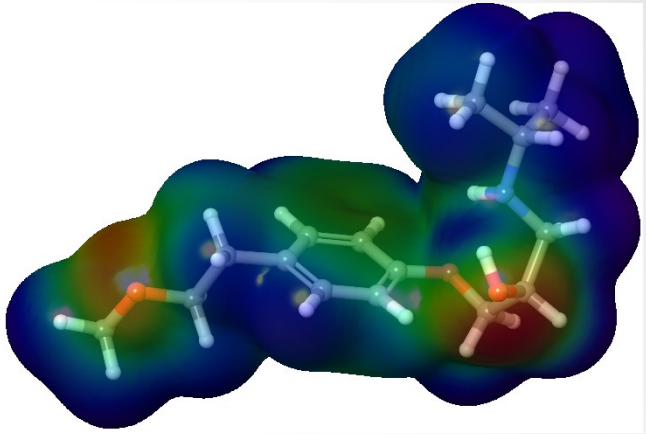
Speed of calculations and the ability to take into account large systems and temperature

Tačnost zavisi od kvaliteta parametrizacije

Basic protocol for computational modeling

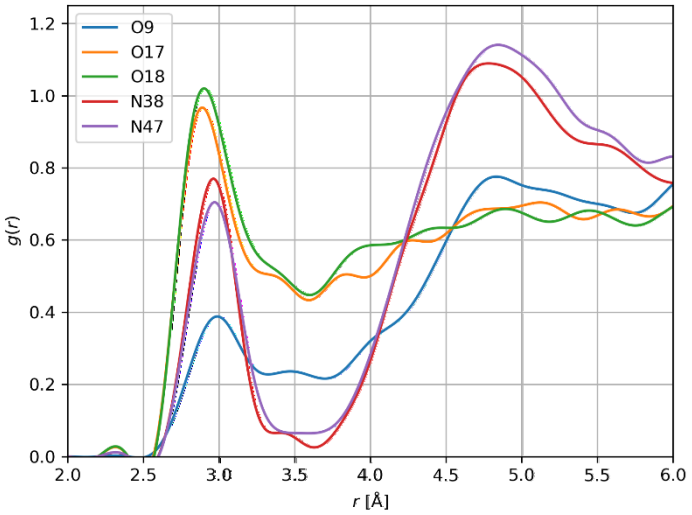
INPUT





OUTPUT

POST PROCESSING



Example

Input file

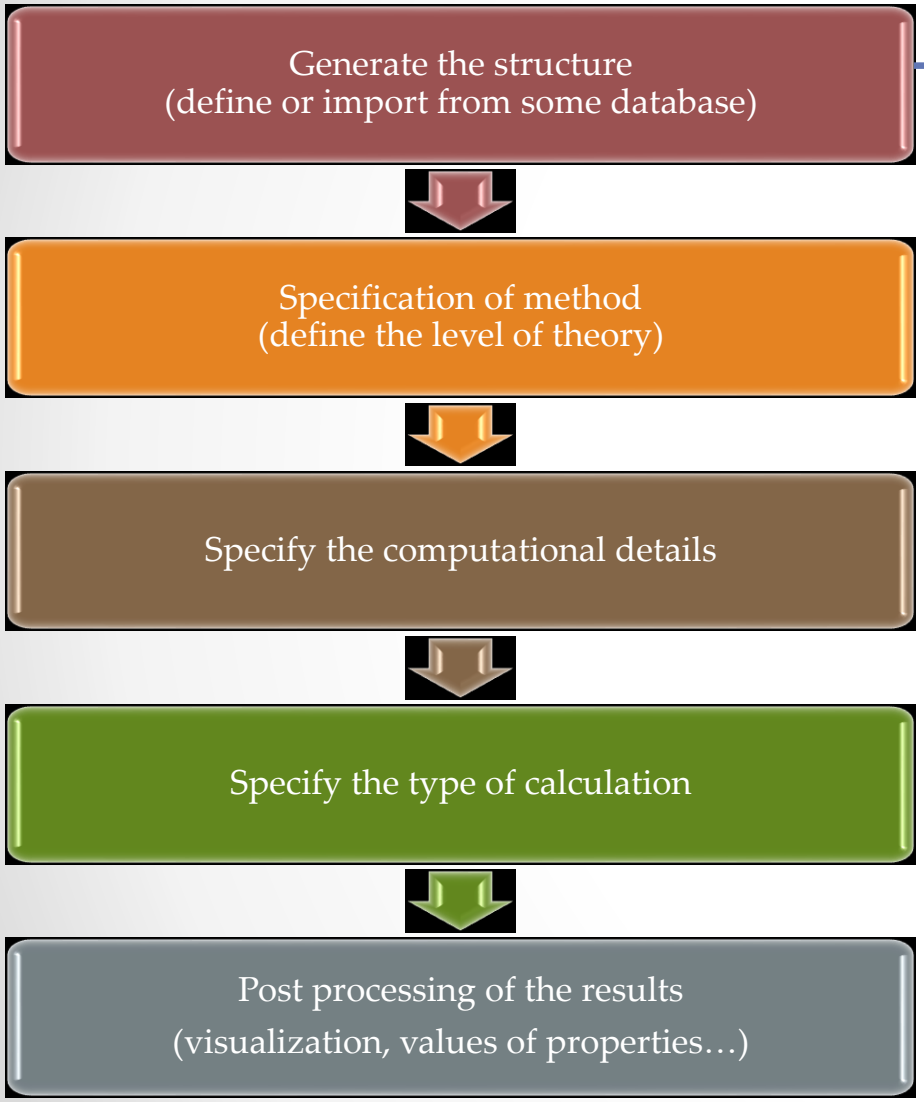
```
!   Example of input file for optimization of C24
 $BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 NPFUNC=1 $END
 $PCM SOLVNT=WATER $END
 $CONTRL SCFTYP=RHF RUNTYP=OPTIMIZE DFTTYP=B3LYP MOLPLT=.TRUE.
PLTORB=.TRUE. $END
 $SYSTEM PARALL=.TRUE. $END
 $GUESS PRTMO=.TRUE. $END
 $SCF DIRSCF=.TRUE. $END
 $STATPT OPTTOL=0.0001 NSTEP=20 STPT=.TRUE. NPRT=1 $END
 $DATA
C24_opt_pristupno_predavanje
C1
C    6.0   -0.56658    2.22823    0.38084
C    6.0   -2.21299    0.62344    0.38084
C    6.0   -1.64642   -1.60479    0.38084
C    6.0    0.56658   -2.22823    0.38084
C    6.0    2.21299   -0.62344    0.38084
C    6.0    1.64642    1.60479    0.38084
C    6.0    2.21299    0.62344   -0.38084
C    6.0    0.56658    2.22823   -0.38084
C    6.0   -1.64642    1.60479   -0.38084
C    6.0   -2.21299   -0.62344   -0.38084
... truncated
 $END
```

Generate this with Avogadro program

Avogadro: an open-source molecular builder and visualization tool. Version 1.XX. http://avogadro.cc/

Marcus D Hanwell, Donald E Curtis, David C Lonie, Tim Vandermeersch, Eva Zurek and Geoffrey R Hutchison; "Avogadro: An advanced semantic chemical editor, visualization, and analysis platform" Journal of Cheminformatics 2012, 4:17.

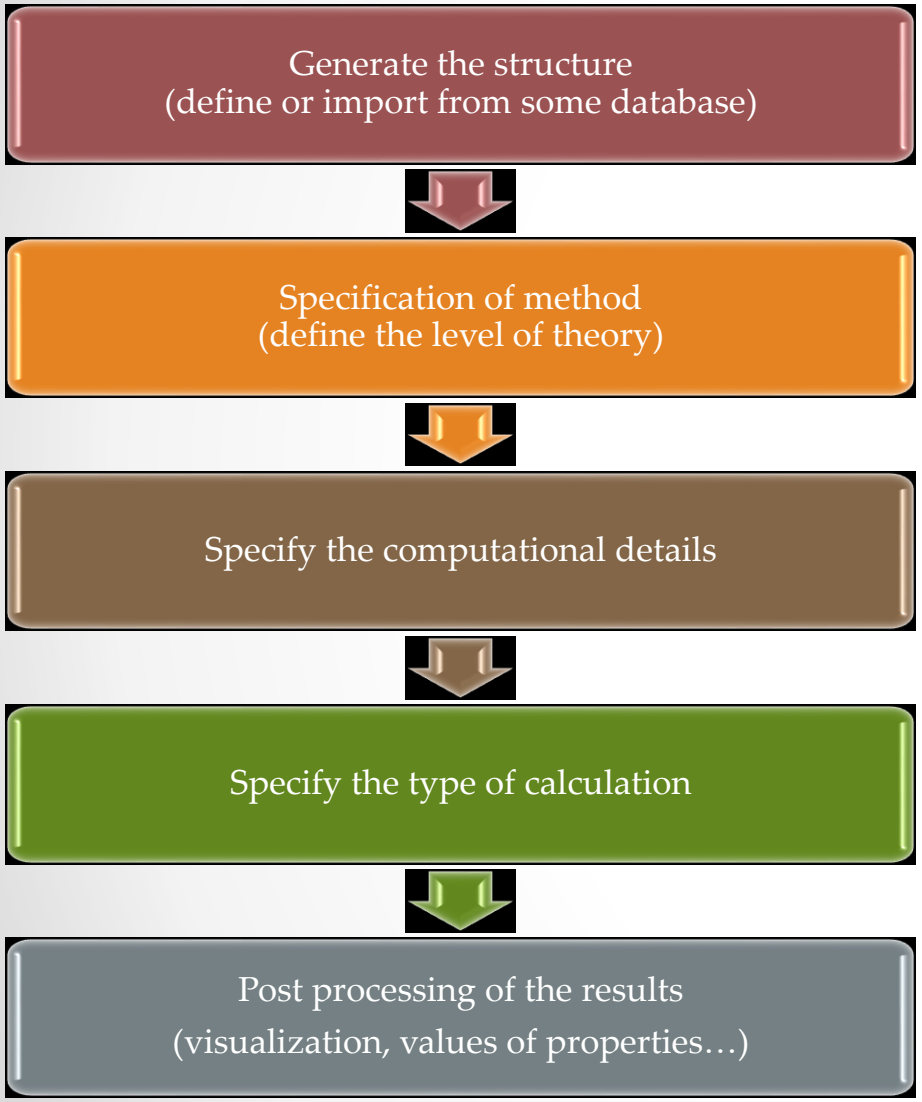
Basic protocol for computational modeling



File formats for structures

Molecules	Periodic structures	Proteins
• .xyz • .mol • .mol2 • .sdf	• .cif • .mmcif	• .pdb • .pdbqt • .ent

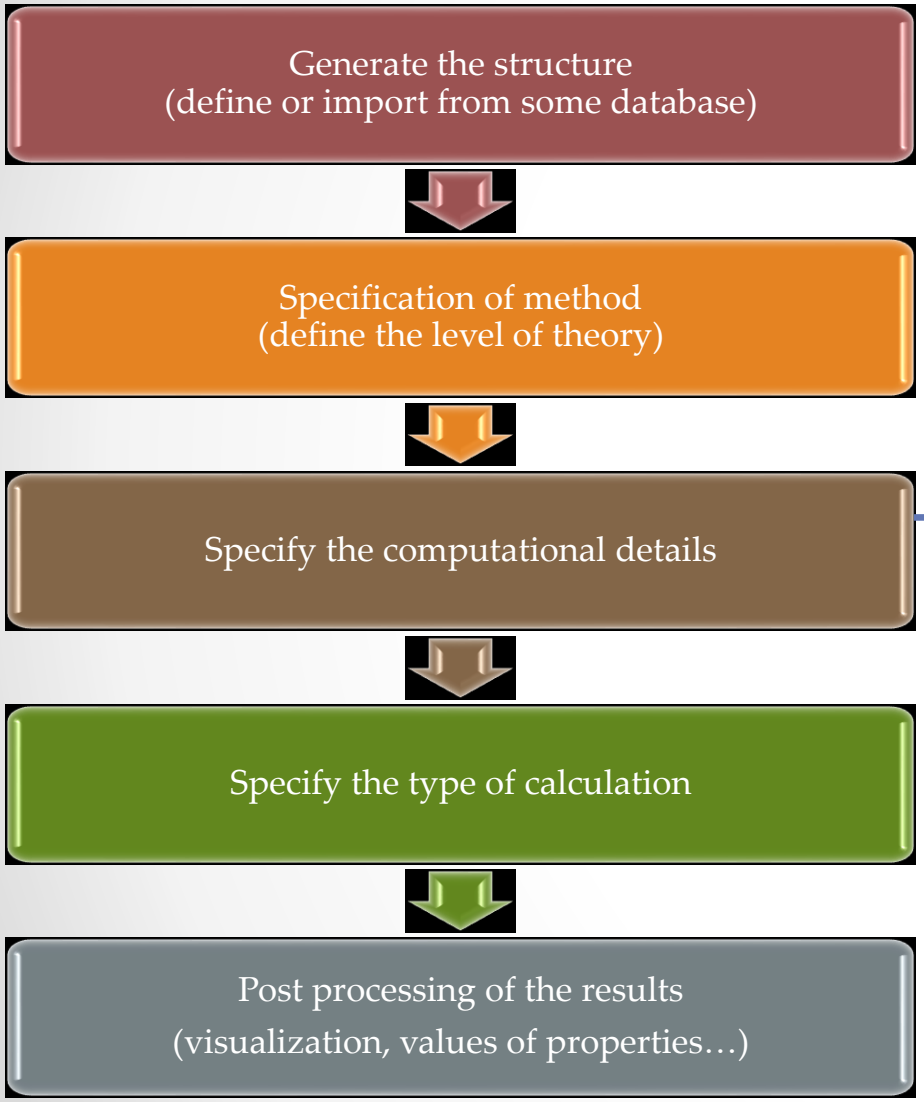
Basic protocol for computational modeling



Some possible methods

QM	Semi-empiric	MM/MD
• Hartree-Fock • DFT • Post HF	• NDDO • AM1 • PM3/6/7 • DFTB	• CHARMM • OPLS • AMBER • …

Basic protocol for computational modeling



Some computational details

QM

- Functional
- Basis sets
- Convergence criteria

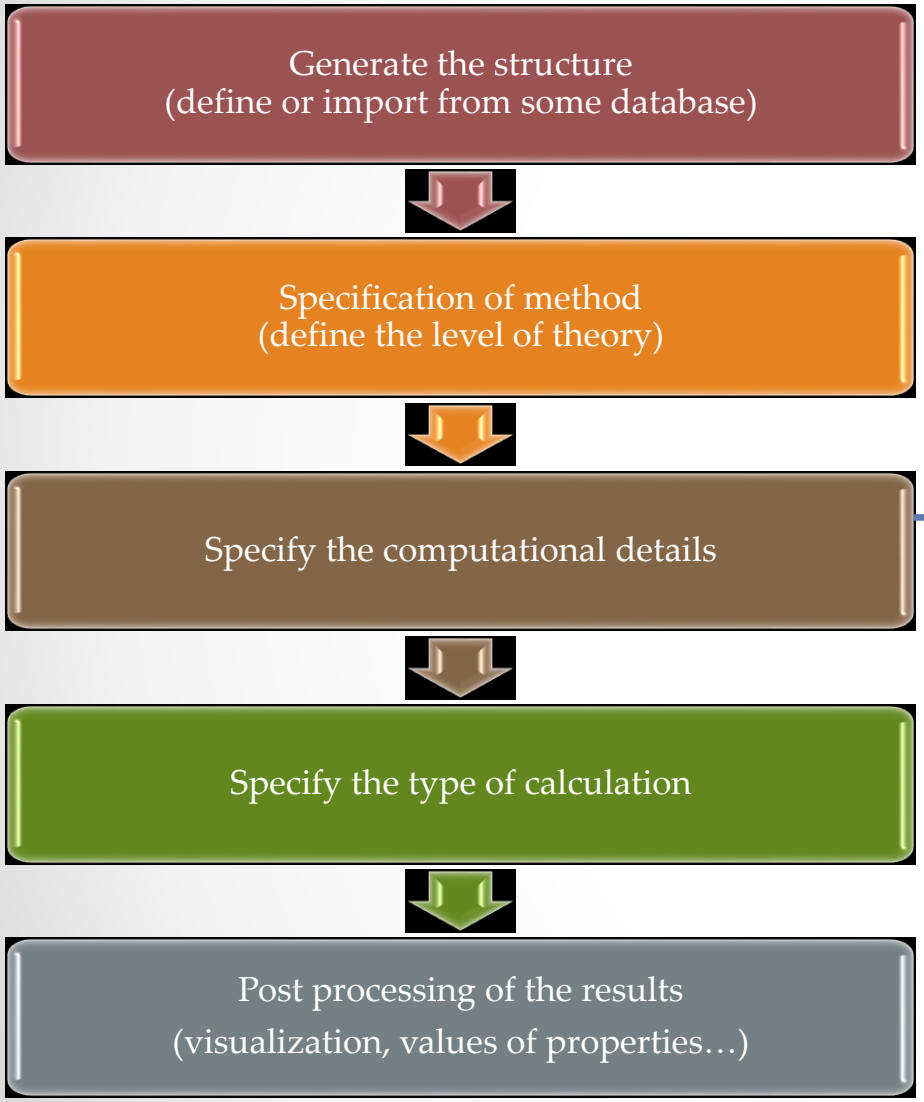
Semi-empiric

- Selection of parameters for atom

MM/MD

- Force field
- Simulation time
- Time step
- Temperature

Basic protocol for computational modeling



Some computational details

QM

- [redacted]
- Basis sets
- Convergence criteria

Semi-empiric

- Selection of parameters for atom

MM/MD

- Force field
- Simulation time
- Time step
- Temperature

Local density approximation

- SVWN5

Gradient-corrected

- BLYP
- PBE
- BP86
- OLYP

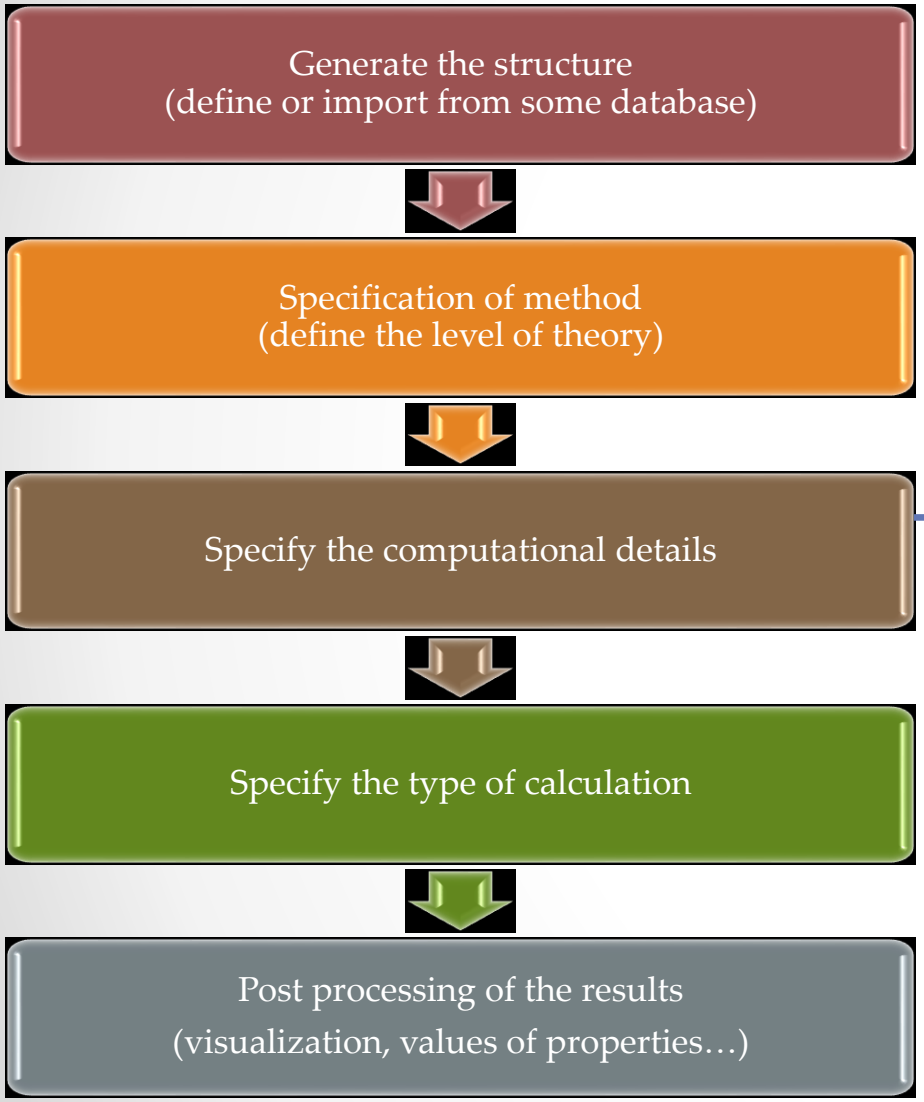
Hybrid

- PBE0
- B3LYP
- PW6B95
- M06-2X

Long range corrected

- CAM-B3LYP
- ωB97
- ωB97x
- ωPBE

Basic protocol for computational modeling



Some computational details

QM

- Functional
- [redacted]
- Convergence criteria

Semi-empiric

- Selection of parameters for atom

MM/MD

- Force field
- Simulation time
- Time step
- Temperature

Pople

- 6-31G
- 6-311G
- 6-311G-3df-3pd

Dunning

- cc-pVDZ
- cc-pVDZ-(f)
- cc-pVTZ
- cc-pVTZ-(f)

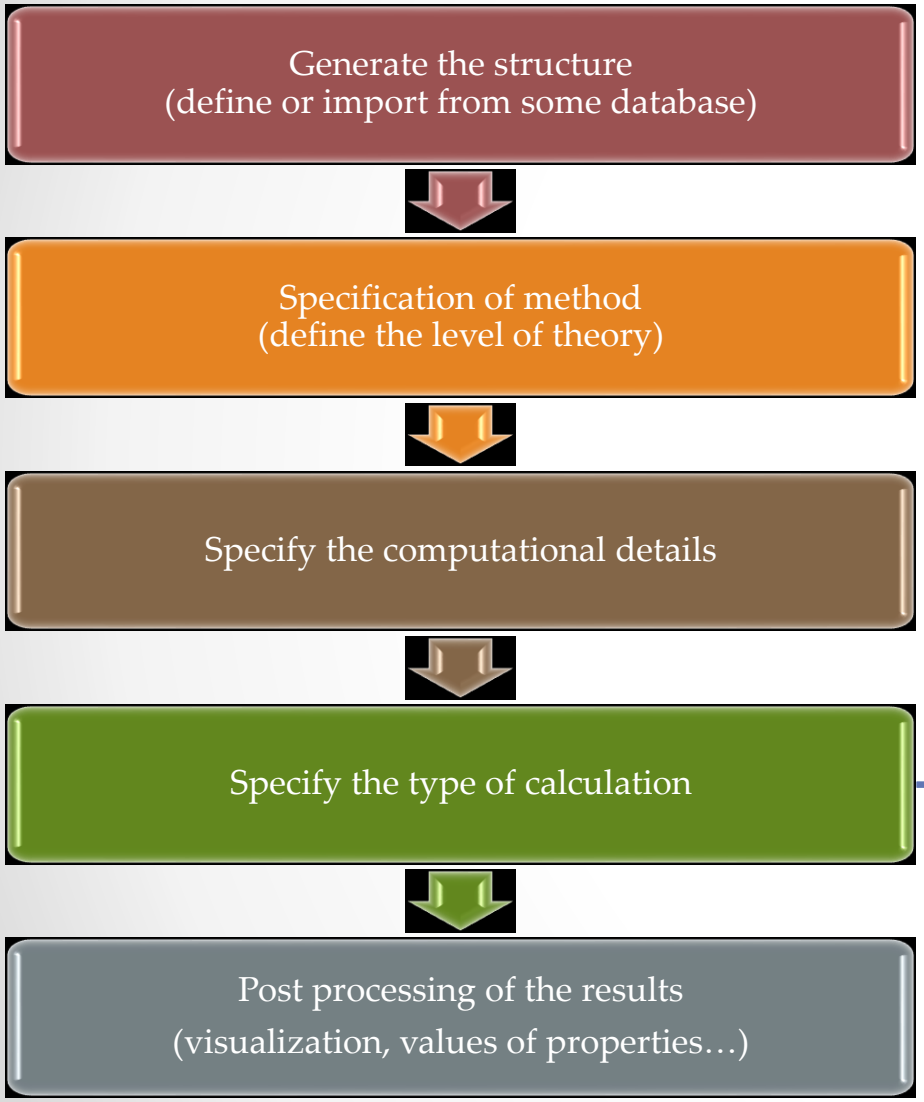
Effective core potentials (ECP)

- LACVP
- LACV3P
- CSDZ
- ERMLER2

Ostali

- MIDIX
- TZV
- Rappoport
- D95

Basic protocol for computational modeling



Some types of calculations

QM

- Geometrical optimization
- Single point energy
- Property calculation

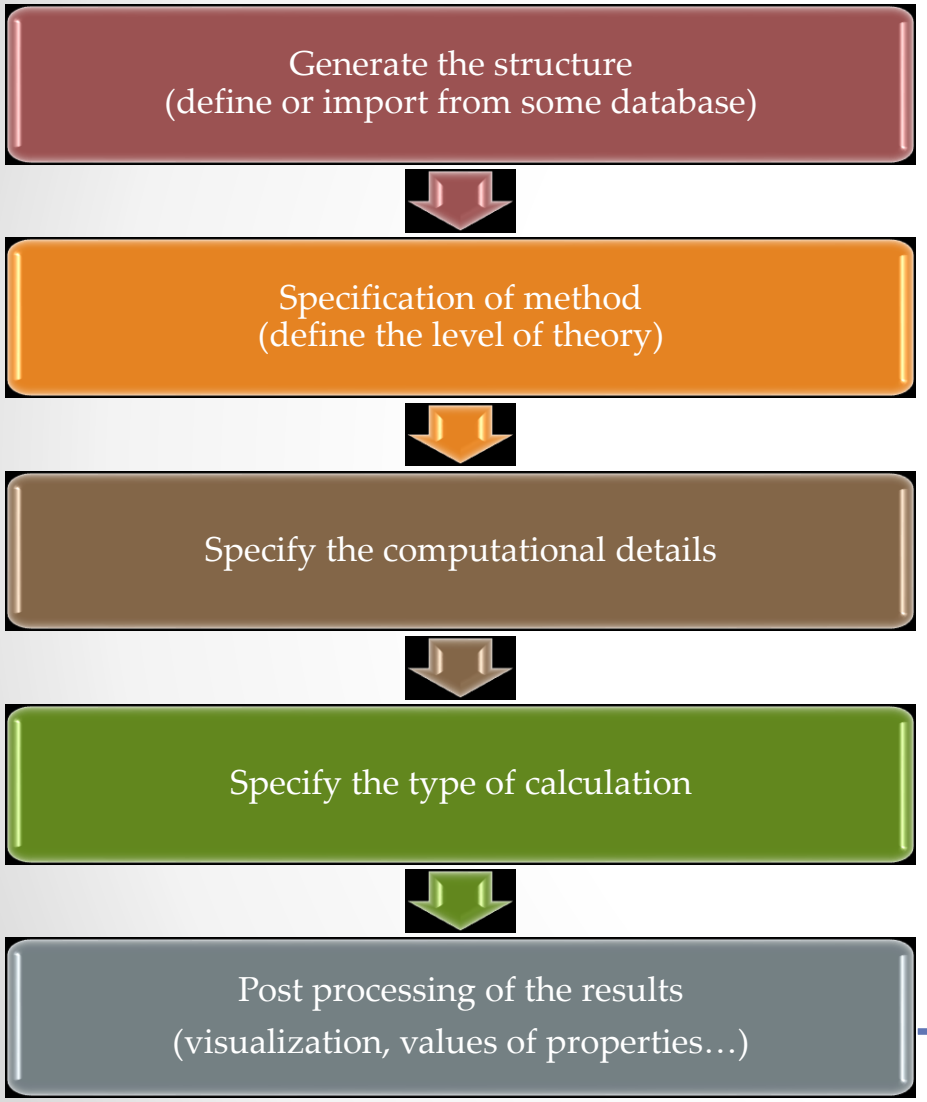
Semi-empiric

- Geometrical optimization
- Single point energy
- Property calculation

MM/MD

- Minimization
- Dynamics
- Temperature dependence

Basic protocol for computational modeling



Some examples of properties obtained after post procession of output file

Charge distribution

- Molecular electrostatic potential
- Atomic charges

Spectroscopy

- IR
- NMR
- UV
- Raman...

Dynamics

- Radial distribution functions
- Thermophysical properties
- Glass transition temperature
- etc

The tricky questions

- **What is the best level of theory?**
- **What is the best density functional?**
- **What is the best force field?**
- **What is the best computational modeling program?**

One answer: "It depends ☺"

Use QM methods for:

- Small or medium sized molecules
- Unit cells or slabs with up to 100 atoms
- Obtaining accurate properties

Use SE methods for:

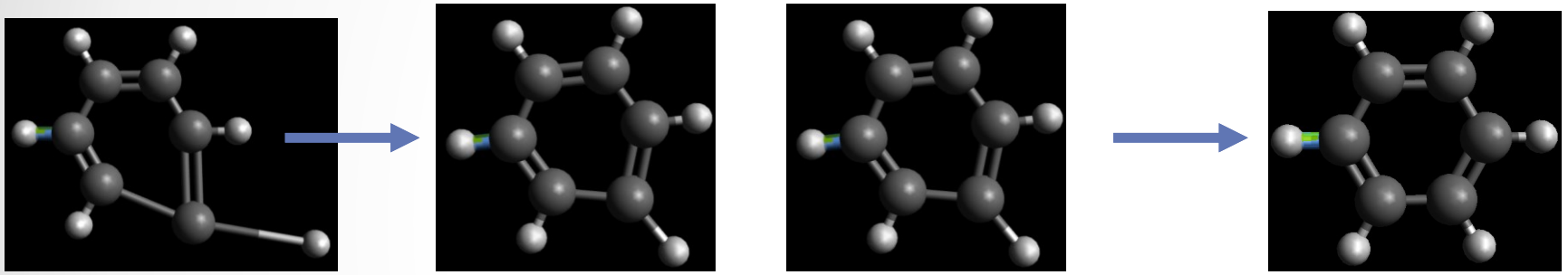
- Large molecules or unit cells
- Particularly useful for organic molecules and periodic structures
- Pre-screening of selected properties

Use MM methods for:

- Very large structures, such as proteins
- Explicit taking into account of solvent
- Taking into account the effects of temperature

Combine levels of theory

To improve efficiency or when you have to investigate a very large structure



Charge, spectroscopy, reactivity…

MM — Pre-optimize (and pre-screen for selected properties)

QM — Re-optimize selected candidates at QM level of theory

QM — Re-calculate properties at even higher QM level of theory

Combine levels of theory

When you have to investigate a very large number of structures

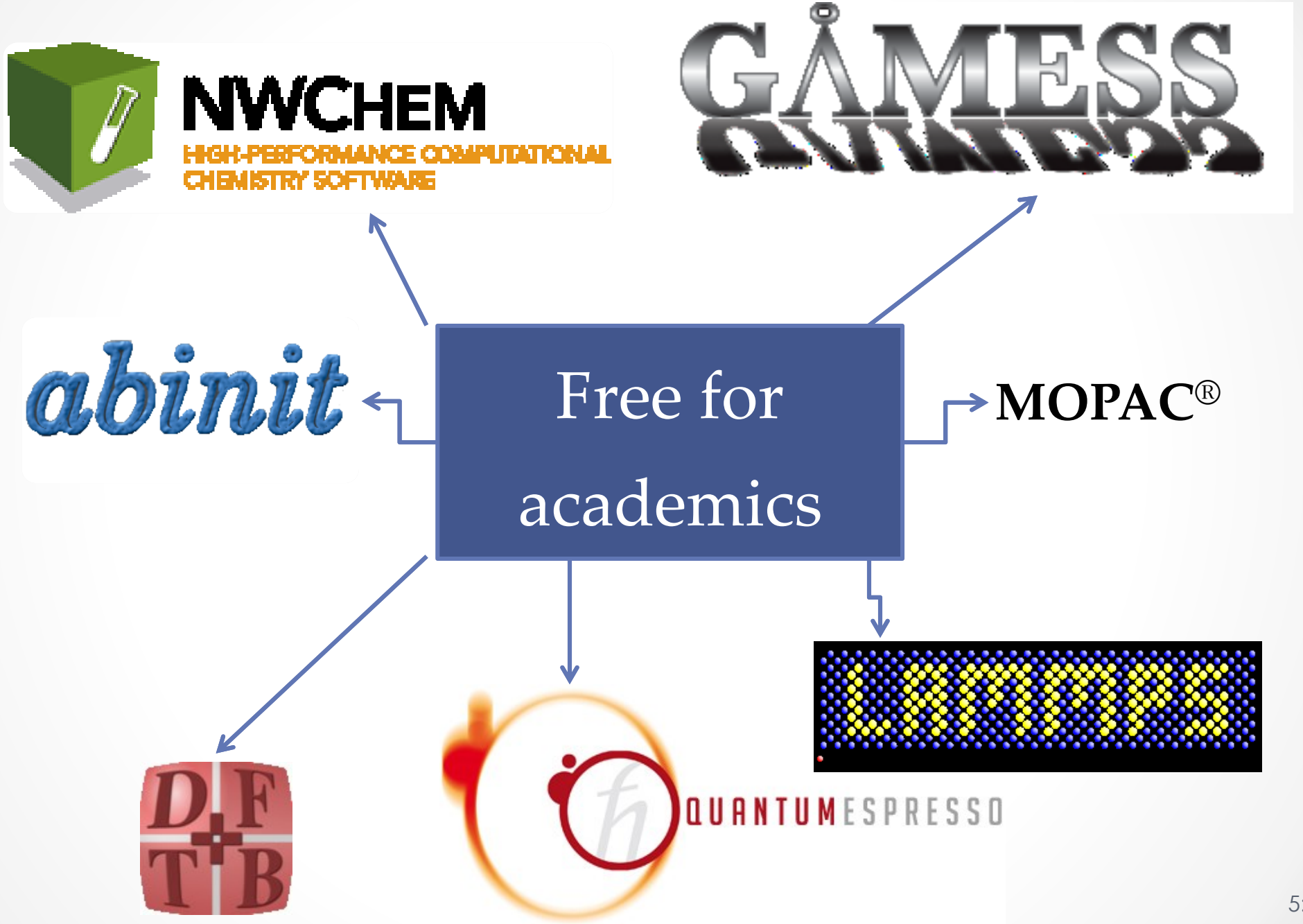
Consider a molecule with many possible (for example hundreds of) conformations

This will optimize structures at lower level of theory and provide reasonable starting points for higher level of theory

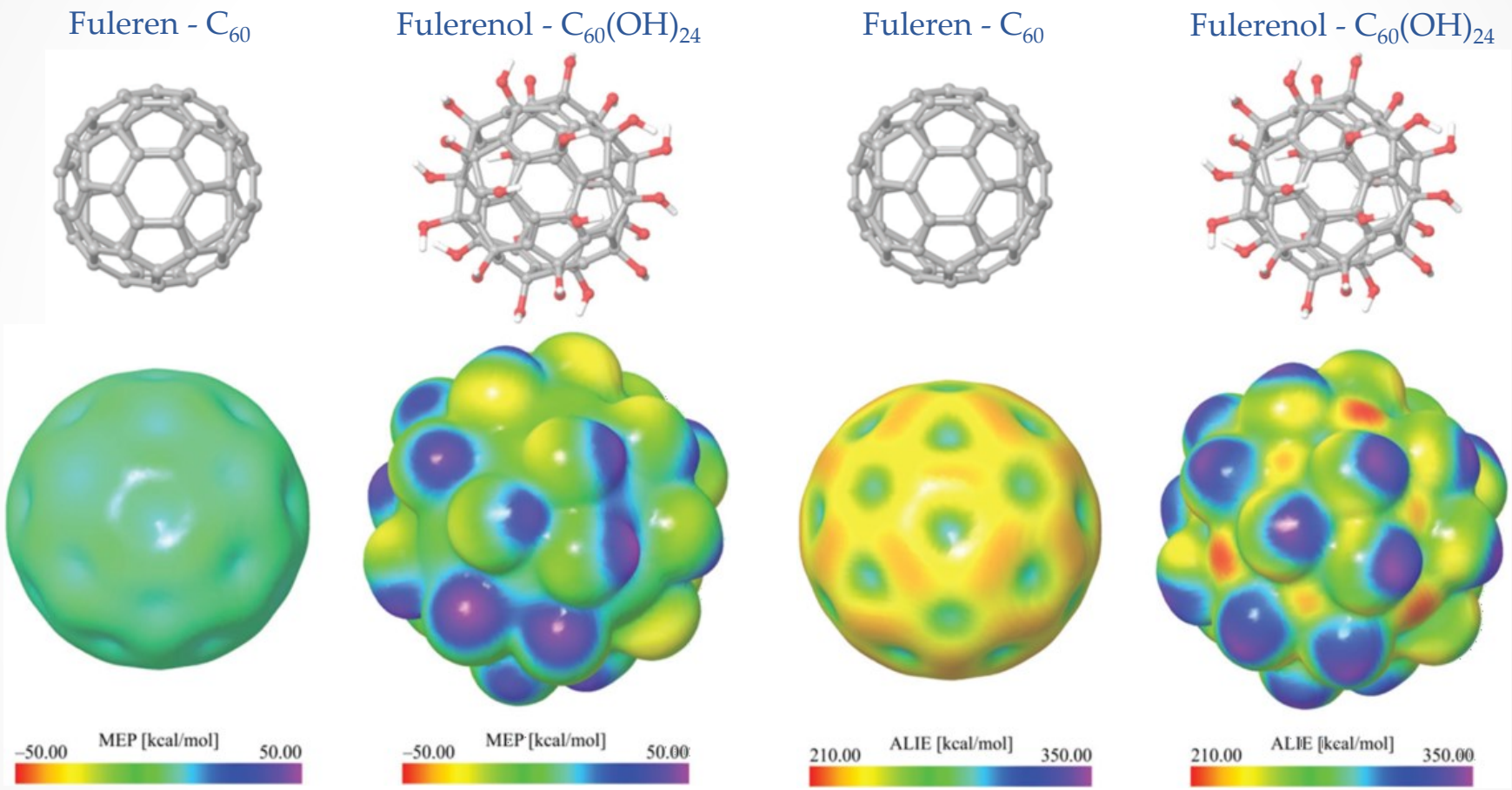
Re-optimization at SE level will improve the accuracy of structures and enable narrowing the number of relevant conformations

Re-optimization of selected candidates at QM level will provide the lowest energy conformations

Calculate relevant properties at higher QM levels of theory to get the best results


NWCHEM
HIGH-PERFORMANCE COMPUTATIONAL CHEMISTRY SOFTWARE
GAMESS
abinit
Free for academics
MOPAC®
DFTB
QUANTUMESPRESSO

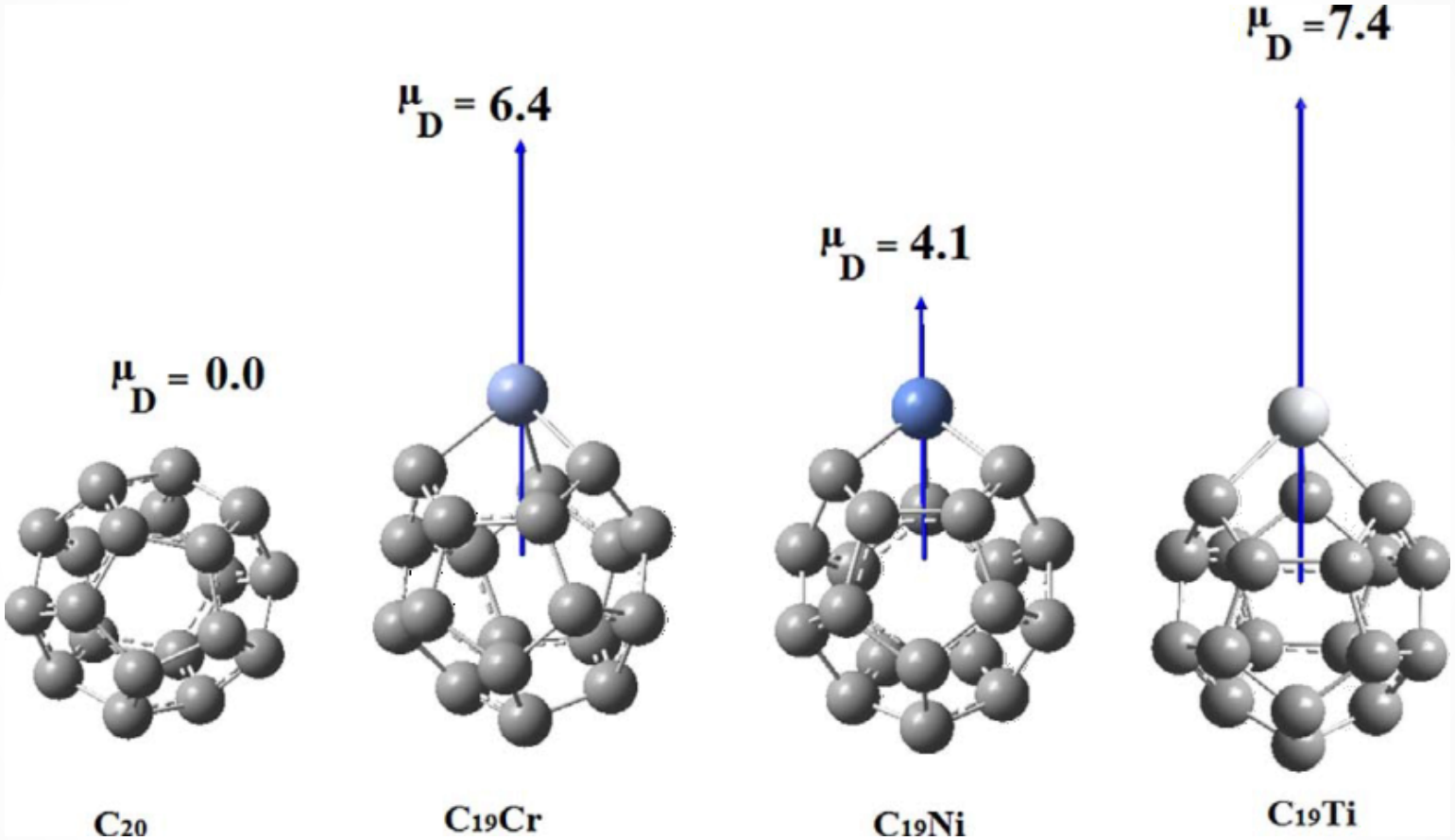
Some examples of what can be obtained by DFT



The influence of attached OH groups to reactivity of fullerene C60
- MEP and ALIE surfaces -

Physical Chemistry Chemical Physics, 19 (2017) 135

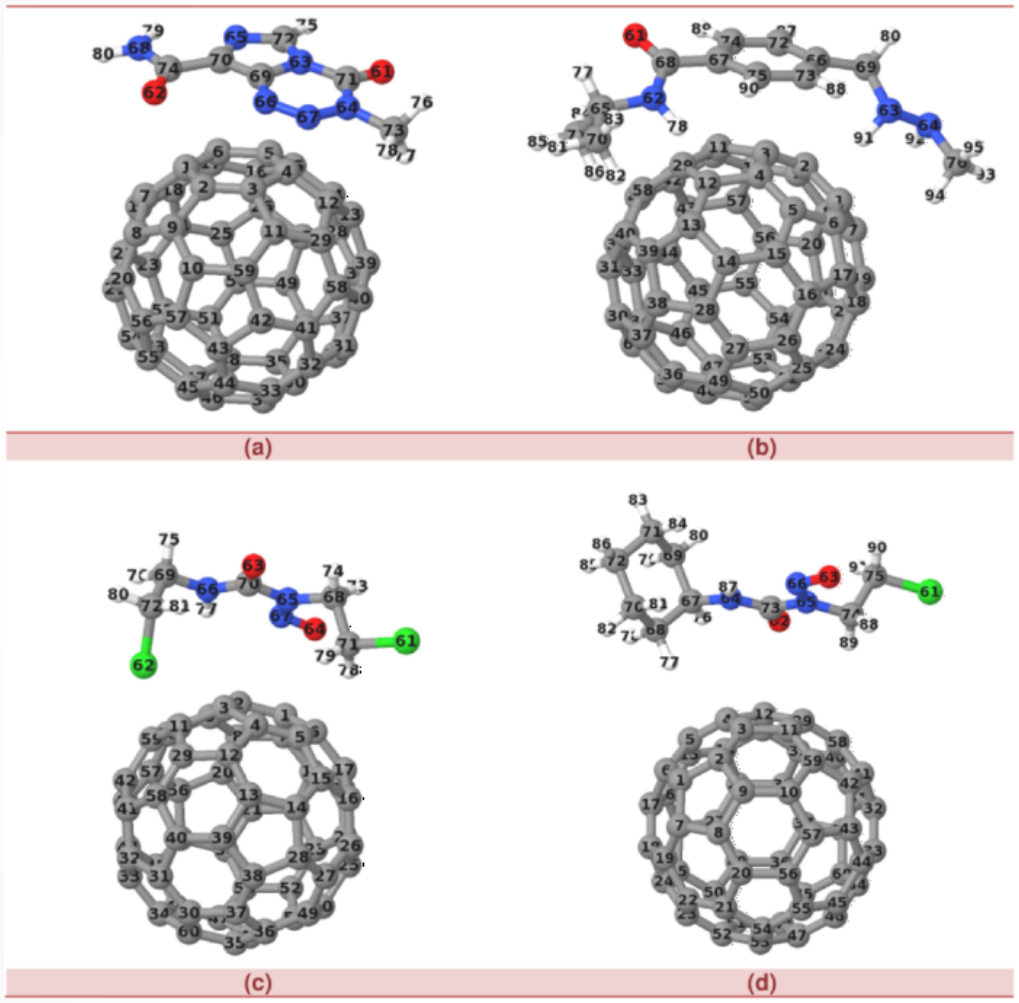
Some examples of what can be obtained by DFT



The influence of modifications of fullerene C20 to dipole moment

Materials Research Bulletin, 97 (2018) 399

Some examples of what can be obtained by DFT



Interaction of fullerene with some drug molecules

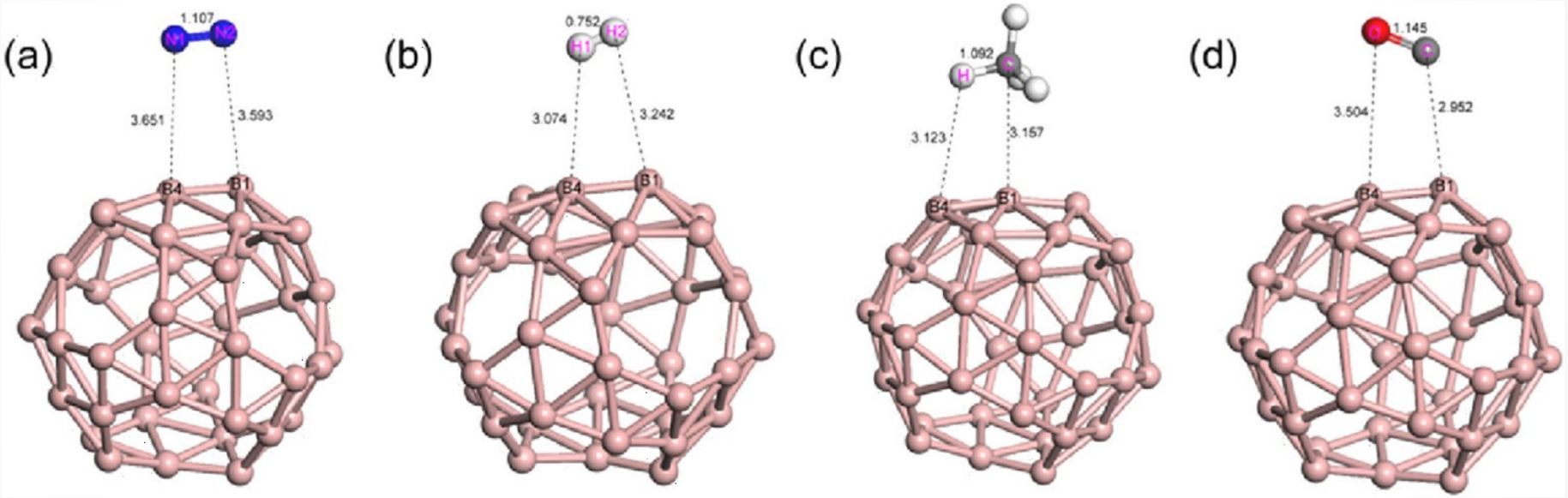
System	Binding energy [kcal/mol]
C60+TMZ	-5.17
C60+PCZ	-5.46
C60+BCNU	-2.52
C60+CCNU	-4.46

(a) Temozolomide (TMZ), (b) Procarbazine (PCZ), (c) Carmustine (BCNU) and (d) Lomustine (CCNU)

Journal Molecular Graphics and Modelling, 72 (2017) 187

Some examples of what can be obtained by DFT

Zaštita životne sredine i skladištenje energije



(a) N_2 (b) H_2 (c) CH_4 (d) CO

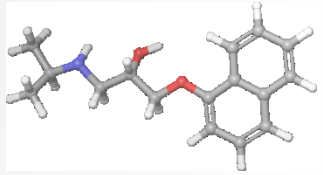
Interaction of B40 cages with selected molecules

Current Applied Physics, 15 (2015) 1084

QM vs. SE methods

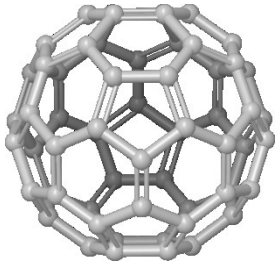
Time necessary to identify the ground state geometry at QM level
B3LYP

Propranolol



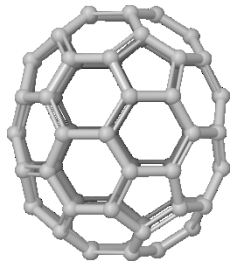
40 atoms, ~ few hours

C60

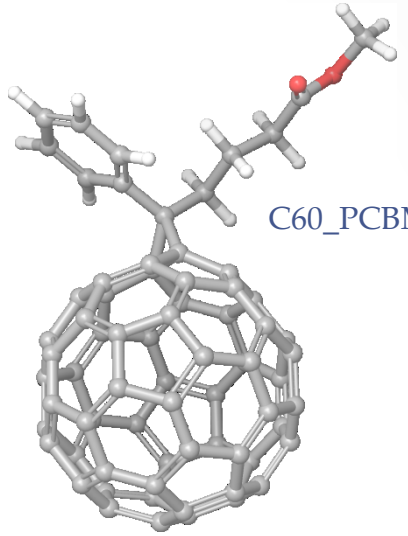


60 atoms, ~ 24 h

C70

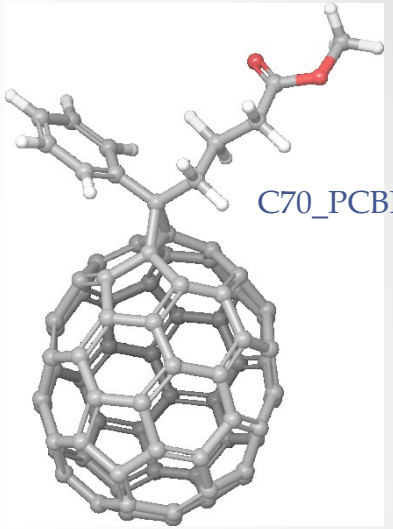


70 atoms, ~ 36 h



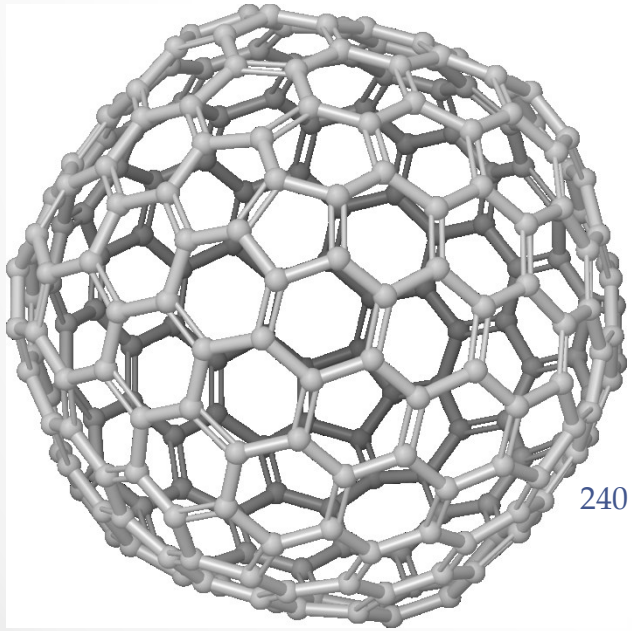
C60_PCBM

88 atoms, > 48 h



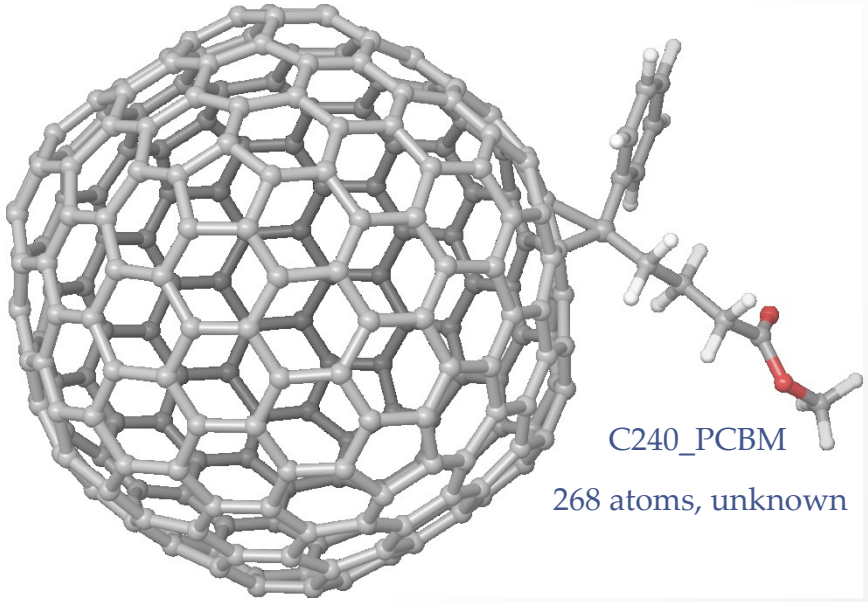
C70_PCBM

98 atoms, > 60 h



C240

240 atoms, unknown



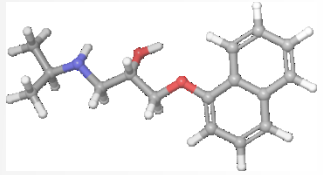
C240_PCBM

268 atoms, unknown

QM vs. SE methods

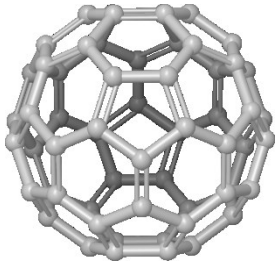
Time necessary to identify the ground state geometry at SE **method**
PM6 method

Propranolol



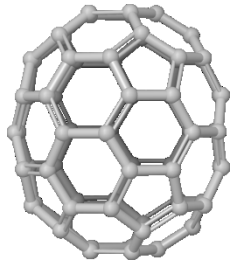
40 atoms, ~ few SECONDS

C60

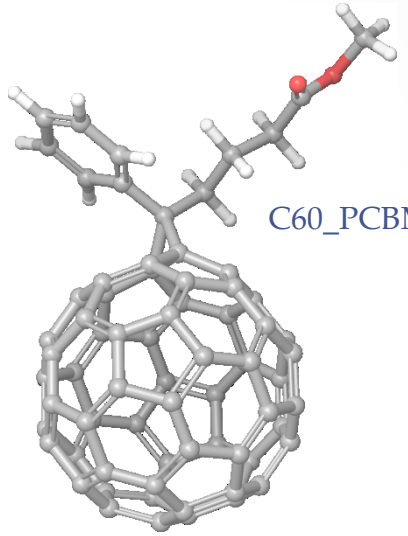


60 atoms, 26 s

C70

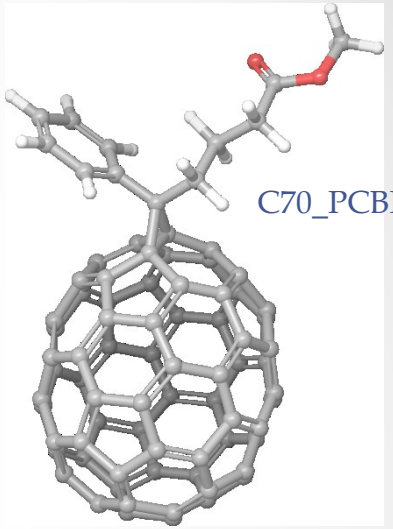


70 atoms, 32 s



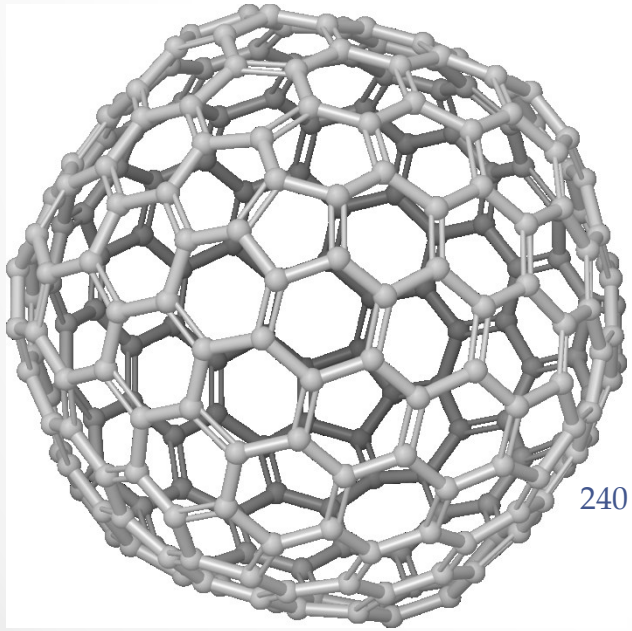
C60_PCBM

88 atoms, 38 s



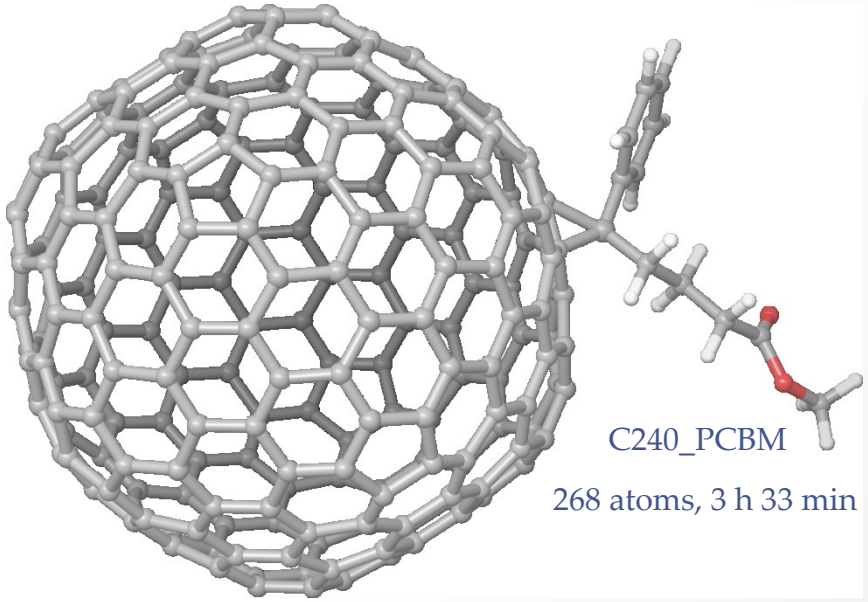
C70_PCBM

98 atoms, > 44 s



C240

240 atoms, 5 min 50 s



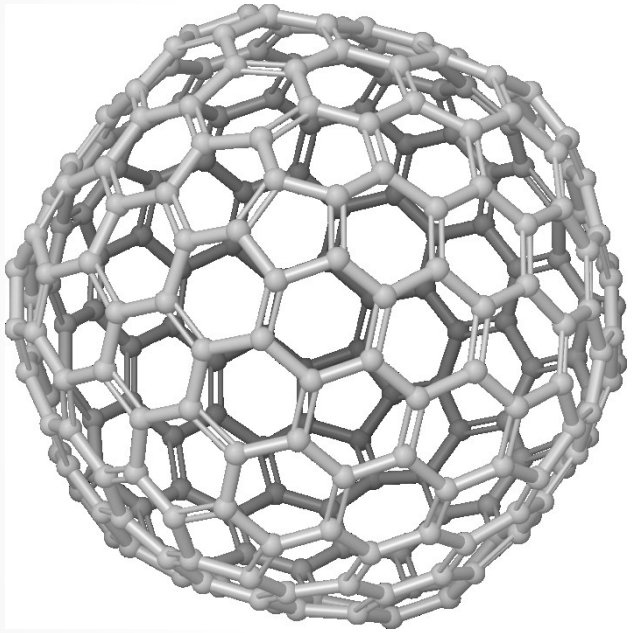
C240_PCBM

268 atoms, 3 h 33 min

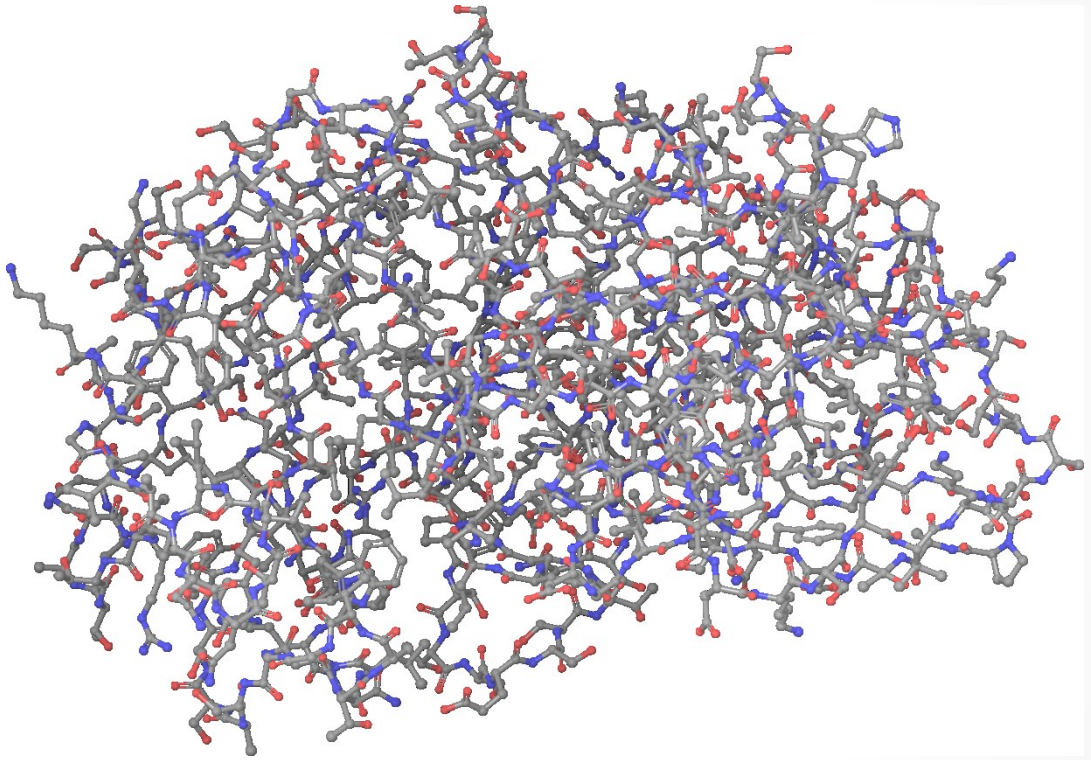
Šta za velike molekule?
– poluempirijski pristup–

Molecule	Time necessary to identify the ground state geometry	
	DFT/6-31G(d,p)	PM6
Propranolol	~ h	~ s
C60	~ 24 h	26 s
C70	~ 36 h	32 s
C60_PCBM	> 48 h	38 s
C70_PCBM	> 60 h	44 s
C240	Unknown	5 m 50 s
C240_PCBM	Unknown	3 h 33 m

What about really large molecules such as proteins

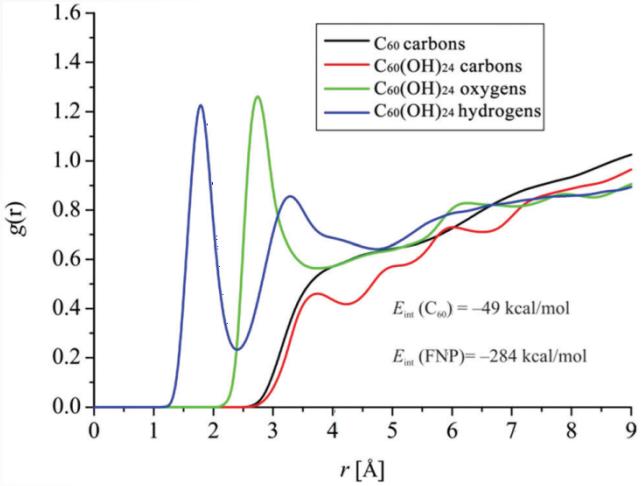


C240

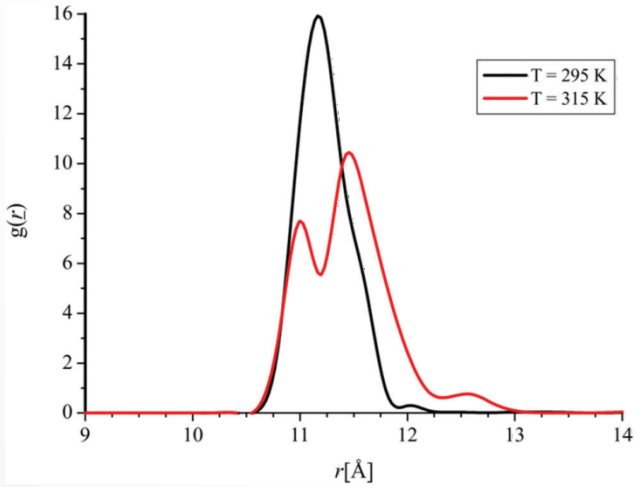


1WKR protein, 2520 atoma

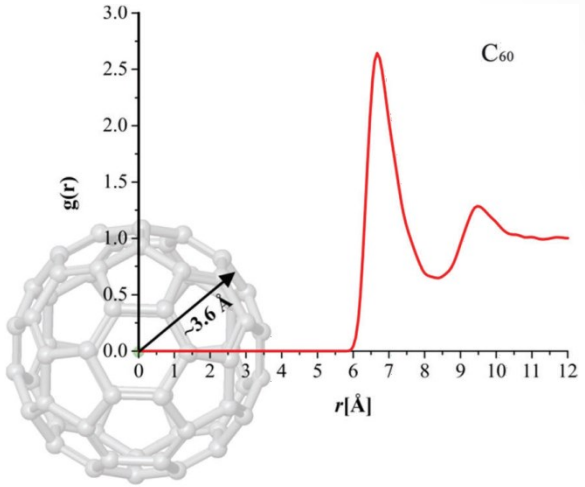
Some examples of what can be obtained with MD

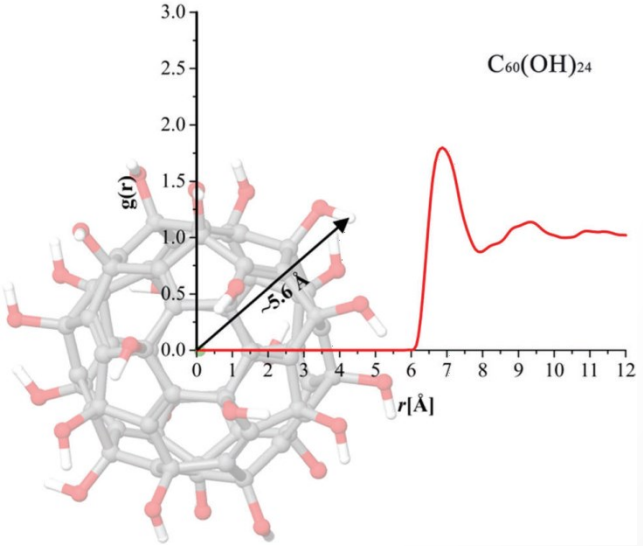


RDF fuleren i fulerenol u vodi



RDF fulerenola i vode na dve temperature

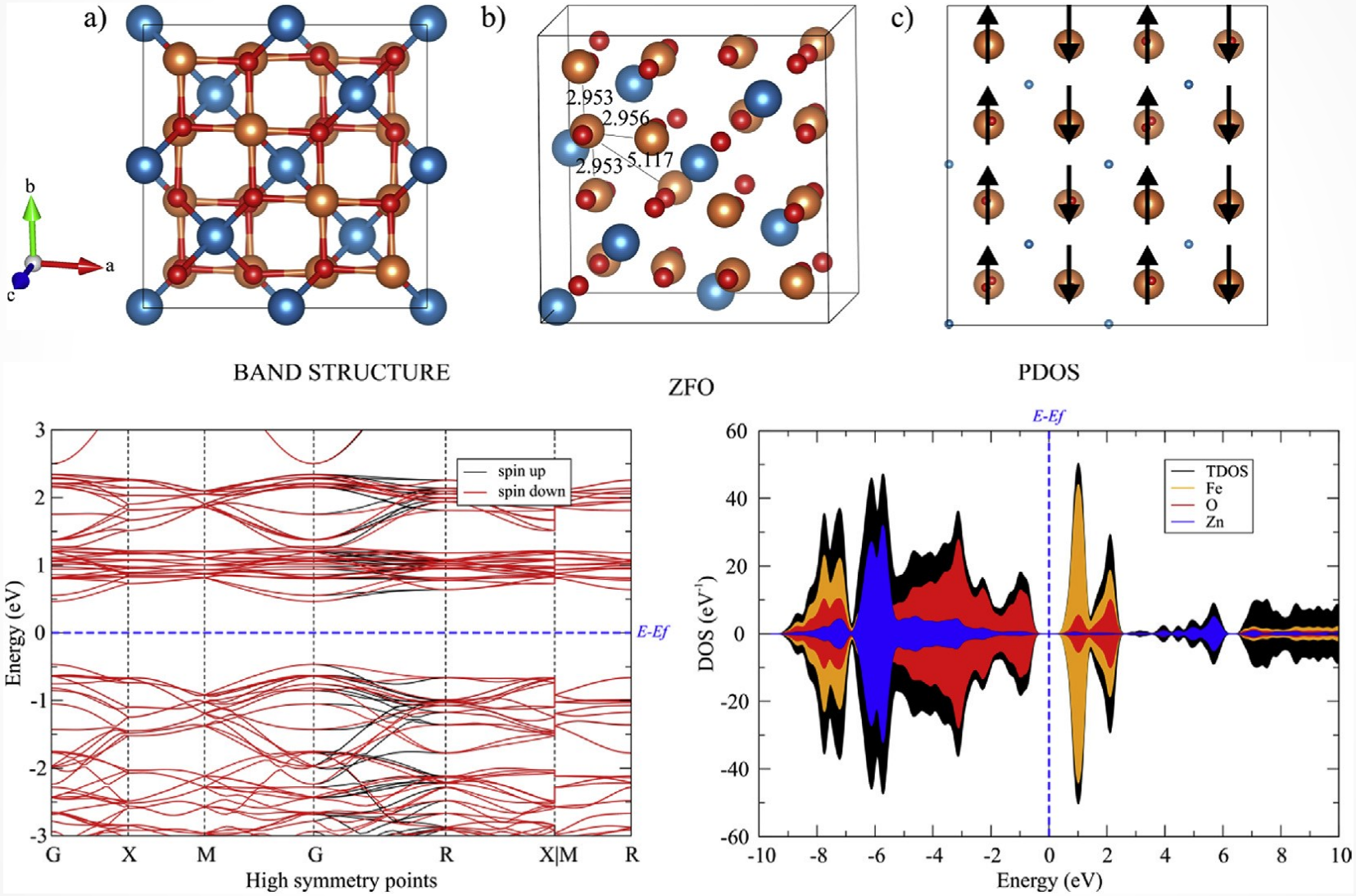




Interakcije fulerena i fulerenola sa vodom

Physical Chemistry Chemical Physics, 19 (2017) 135

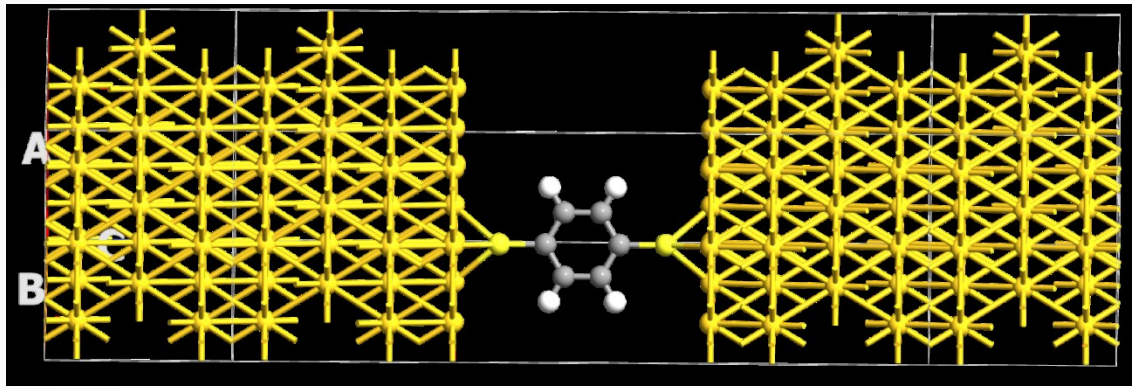
Investigation of electronic structure of crystals with DFT



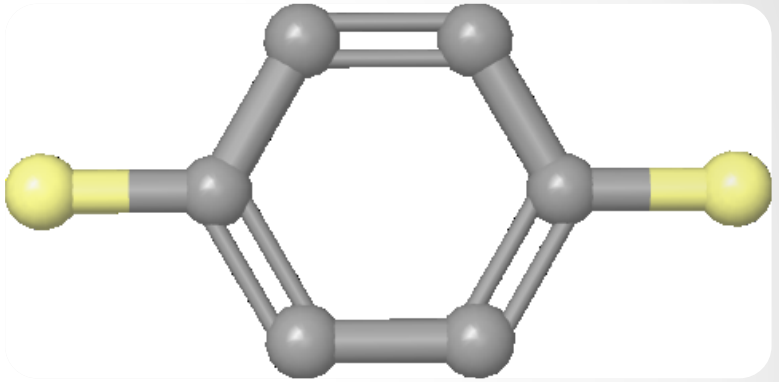
Journal of Alloys and Compounds 842 (2020) 155704

Investigation of transport properties with DFT(B)

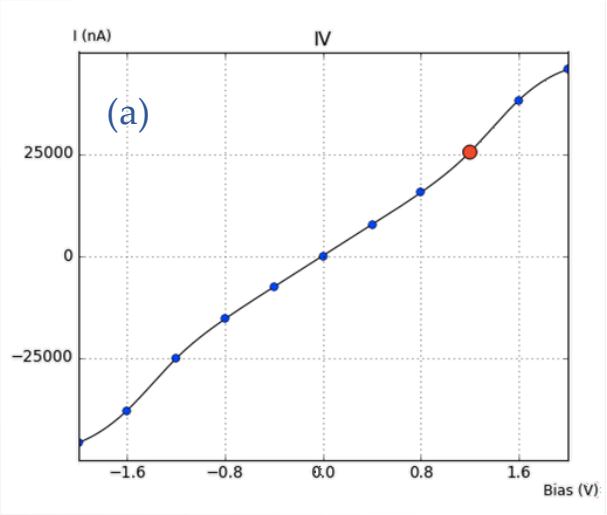
Di-tiol–benzene (DTB) is most frequently used (both exp and theo) molecule for study of electron transport through molecule

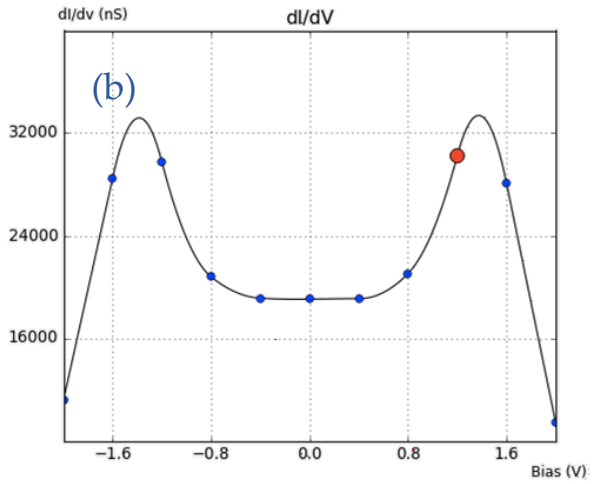


DTB molecule between Au (111) electrodes



DTB molecule



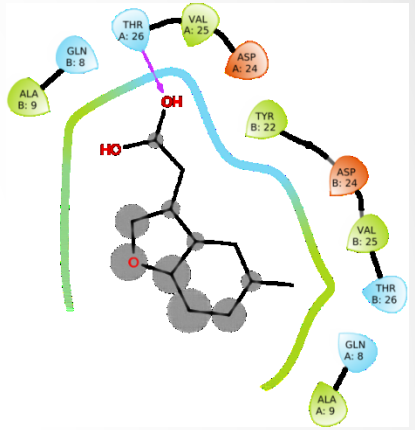


(a) I-V curve and
(b) Differential conductivity

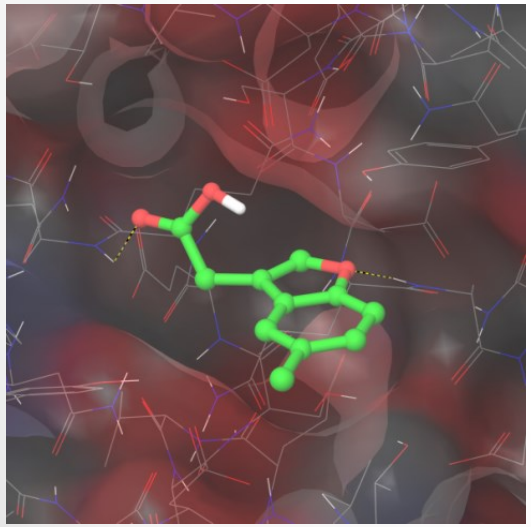
Computational Materials Science 27 (2003) 151–160

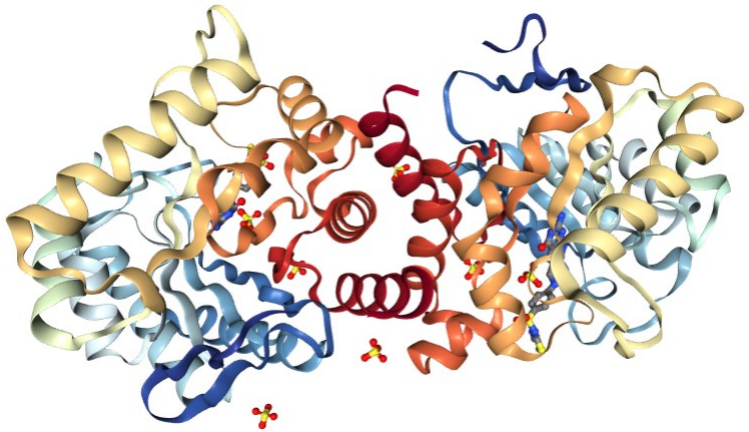
Molecular docking

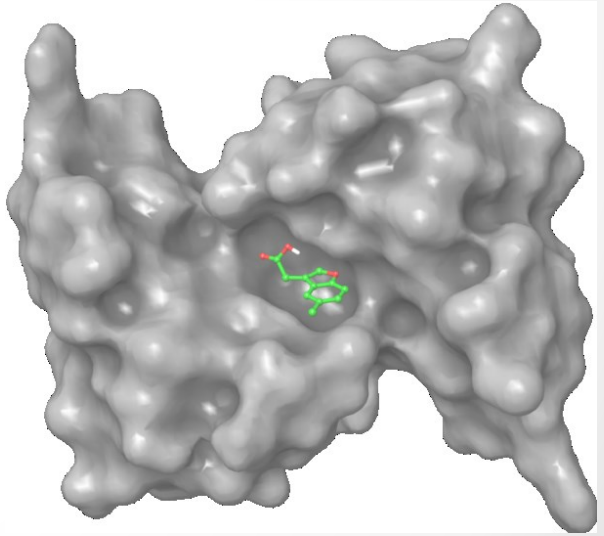
Possible orientations and binding sites of ligand to protein



- Crystal structure of protein
- Define the binding site or consider the whole protein (blind docking)
- Scoring function to classify the best docking modes

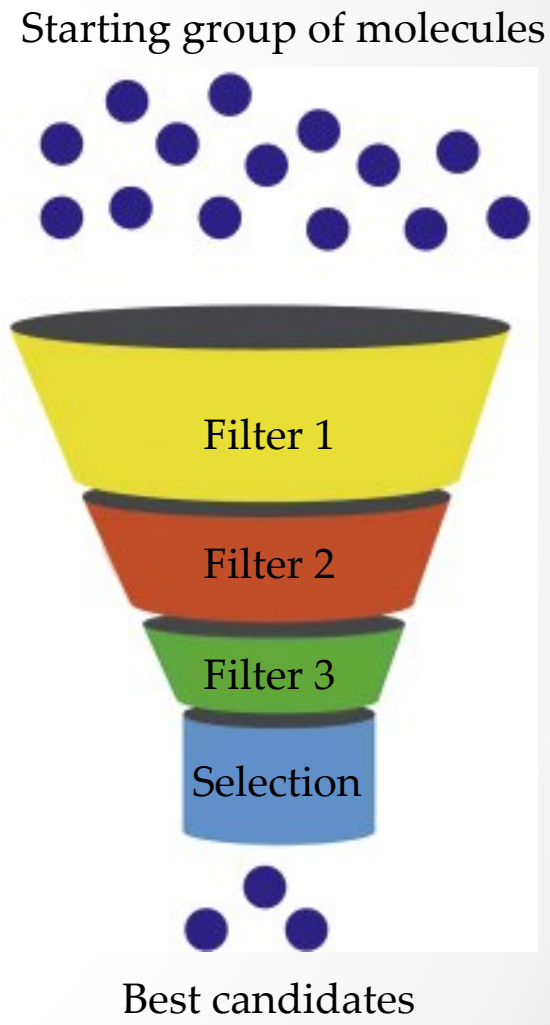






Drug likeness parameters

- Drug likeness parameters are quantities which indicate whether some molecule could possess suitable physical properties to be applied in drug formulation.
- In order to be considered as a good candidate for further investigation, molecule has to have suitable values of drug parameters
- To calculate drug likeness parameters, you need only a structure of a molecules
- Some examples of drug likeness parameters are: logP, number of rotabile bonds, number of acceptor/donors of hydrogen bond, mass, polar surface, etc
- The best known example of application of drug likeness parameters is the so called Lipinski rule of 5:
  - logP < 5
  - Number of donors of H bond < 5
  - Number of acceptors of H bond <10
  - Molecular mass < 500 Dalton
  - Molar refractivity > 40 i < 130
  - Polar surface < 140 A^2



Links

AIDASCO

Website: https://aidasco.org

LinkedIn: https://www.linkedin.com/company/aidasco/

Twitter: @aidasco_ngo

Facebook: https://www.facebook.com/Aidasco-110655184066851/

Instagram: https://www.instagram.com/aidasco2020/

Impact factors online

Website: https://impactfactors.online

LinkedIn: https://www.linkedin.com/company/impact-factors-online/

Twitter: @journalmetrics

Facebook: https://www.facebook.com/impactfactors.online/

Stevan Armaković

Website: https://armakovic.com

LinkedIn: https://www.linkedin.com/in/stevanarmakovic/

Twitter: @armakov

Facebook: https://www.facebook.com/stevan.armakovic

Comments & Questions session