

Universidade Federal da Paraíba

Departamento de Química

Laboratório de Química Quântica Computacional



Grupo de Química Computacional mais Oriental das Américas
L.U.C.E. - UFPB · João Pessoa · PB

Molecular Modeling of Biomolecules using Quantum Chemistry Methods

Gerd Bruno Rocha

https://github.com/RochaGerd/Chemistry_with_Python

gbr@academico.ufpb.br

www.quantum-chem.pro.br



EMMSB

2024

Table of Contents (1st Day)

- Methods in Computational Quantum Chemistry
 - Ab initio
 - Semiempirical
- Computational Details
 - HPC, linear scaling
- Some applications
 - Modeling biomolecular systems

Modeling chemical systems and proteins with MOPAC

1. Calculating ΔH_r for chemical reactions
2. Finding the native structure in a set of disordered proteins

Table of Contents (2nd Day)

Modeling proteins with MOPAC

1. Calculating ΔH_{bind} for an enzyme-ligand binding reactions

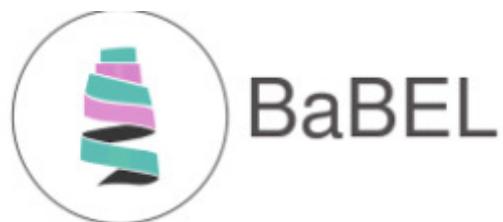
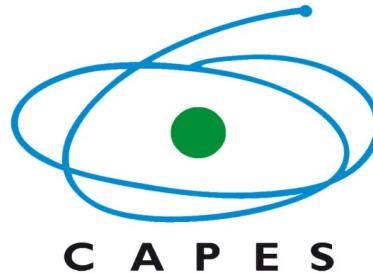
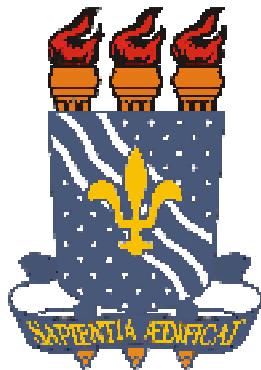
Table of Contents (3rd and 4th days)

Modeling enzymes with PRIMoRDiA

1. About PRIMoRDiA
2. Installing PRIMoRDiA
3. Tutorial: RDs for enzymes
4. Tutorial: RDs for an Enzyme Catalysis

https://github.com/bardenChem/XI_EMMSB_PRIMoRDiA.git

Acknowledgments



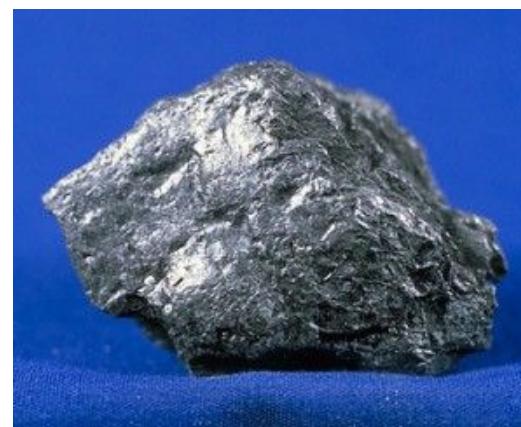
<http://babel.dcc.ufmg.br/> PETROBRAS



Let's start

Molecular Modeling of Biomolecules using Quantum Chemistry Methods

We live in a material world

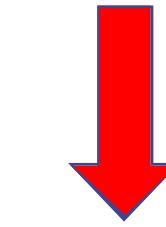


<https://pt.wikipedia.org/wiki/Carbono>

We live in a material world



- Macroscopically different materials
- Composition: carbon
- Macroscopic differences for materials: atomic composition, chemical bonds, intermolecular interactions and structural arrangements.



Properties
of the materials

$$= f$$

(composition, bonds, spatial-
structural arrangements)

Properties of materials = f (structures, bonds)

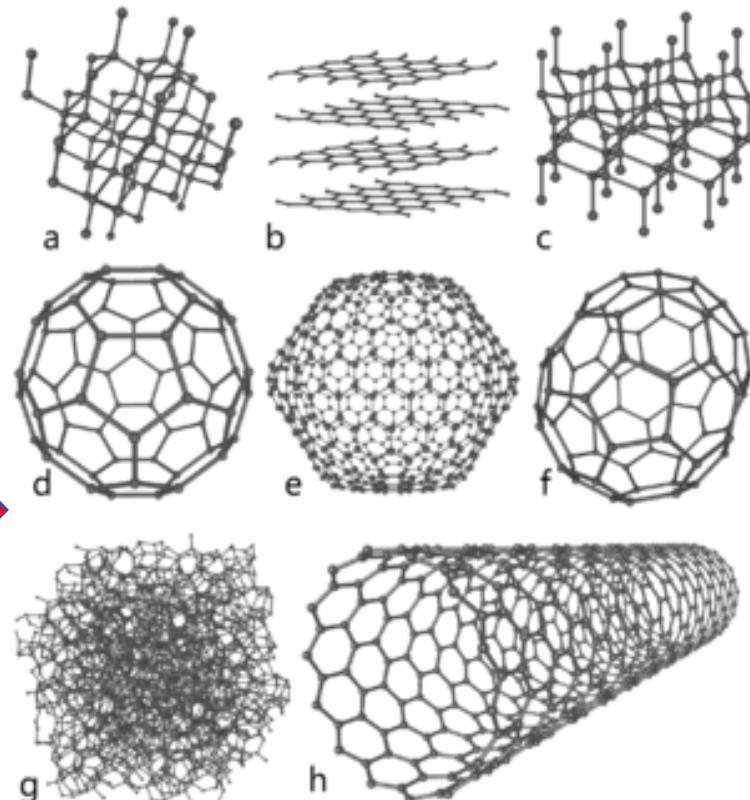
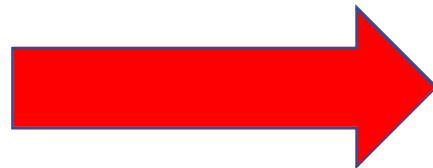
(a)



(b)



(h)



**These properties can be
measured and/or calculated**

https://en.wikipedia.org/wiki/Allotropes_of_carbon

How do we calculate the properties of materials?

- It depends on how we see the matter:
 - Classic interpretation (1600s - 1900s)



Dalton

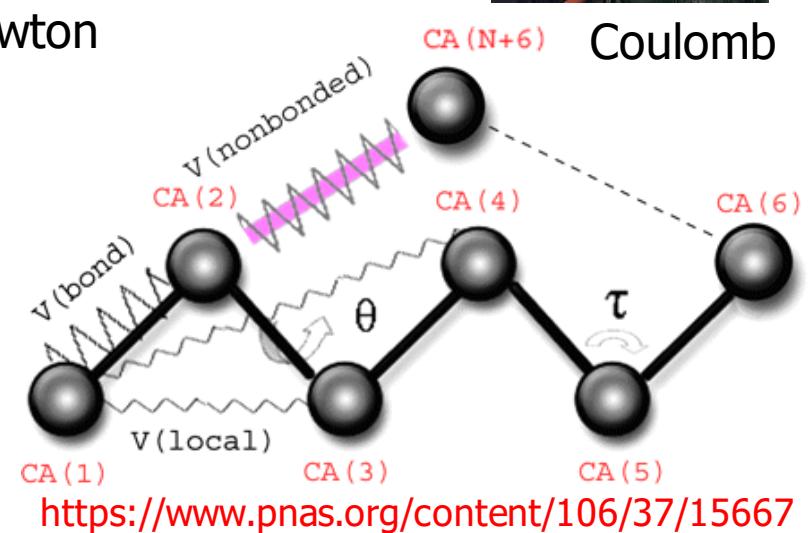


Newton



Coulomb

- Indivisible atoms
- Electrically neutral
- Chemical elements formed by different atoms



How do we calculate the properties of materials?

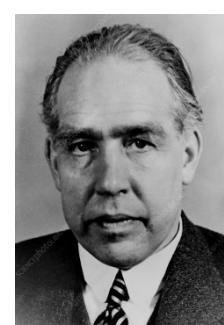
- It depends on how we see the matter:
 - Quantum interpretation (1900s – today)



Schrödinger



Born



Bohr



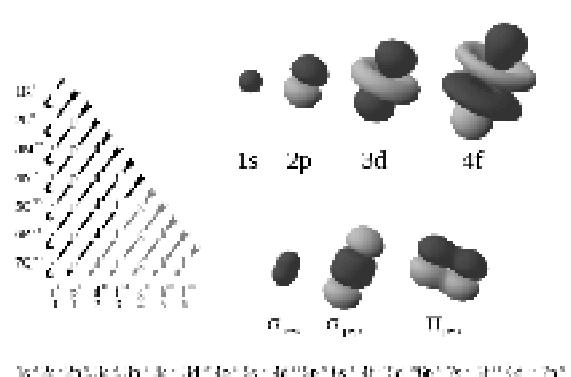
Heisenberg



de Broglie

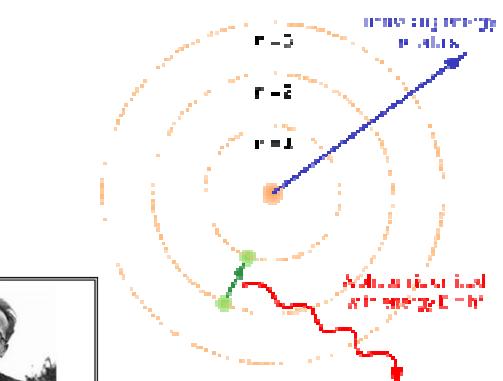
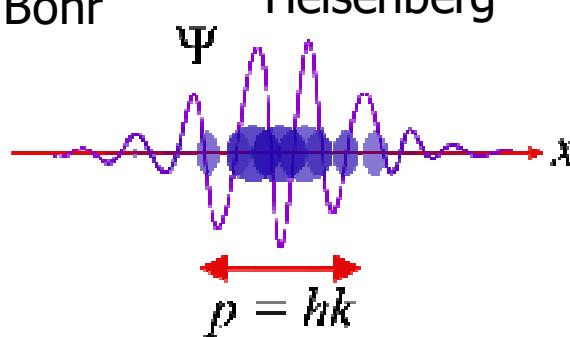


Pauli



$$\hat{H}\psi = E\psi$$

The Schrodinger equation was discovered in 1926 by Erwin Schrodinger, an Austrian theoretical physicist. It is an important equation that is fundamental to quantum mechanics.



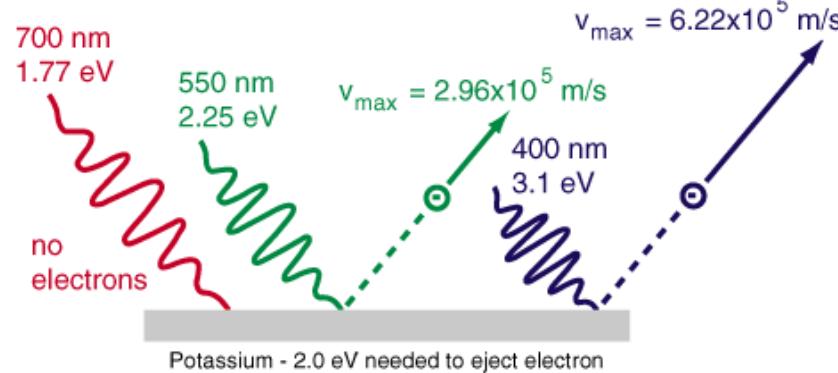
Quantum Mechanics in 10 min

Particle behavior for the light

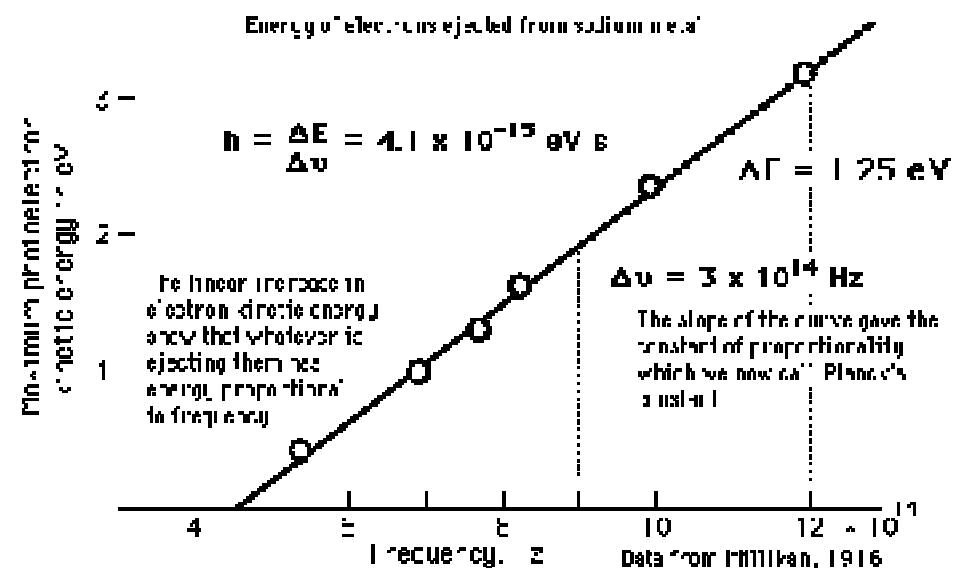
Could light show particle behavior?

The answer to this question was given by Einstein in 1905.

$$E_{\text{photon}} = h\nu$$



Photoelectric effect



Metal electrons are emitted with a specific kinetic energy. Einstein said that "light consists of a finite number of quanta of energy, located at points in space that move themselves without dividing, and can be absorbed or generated only as complete units." What Einstein referred to were photons.

Wave behavior for the matter

Could particles show wave behavior?

The photoelectric effect shows us that the electromagnetic radiation presents particle behavior

Could the matter, as we know it, have wave behavior? (Universal symmetry?)

The French physicist Louis De Broglie dedicated his PhD to answer this question



Louis de Broglie

$$\frac{h}{\lambda} = mv = p$$

Particle property

Wave property

A diagram showing the de Broglie equation $\frac{h}{\lambda} = mv = p$. A vertical blue arrow points upwards from the equation to the text "Particle property". Another vertical blue arrow points downwards from the equation to the text "Wave property".

Wave behavior for the matter

Results from de Broglie hypothesis:

Entities recognized as particles can behave as waves !!

We have a consistent mathematical expression that links properties of matter to electromagnetic radiation

$$\frac{h}{\lambda} = mv = p$$

Experimental observation: Electron diffraction
(Clinton Davisson and Lester Gerner, 1927).



Heisenberg's Uncertainty Principle

Classical mechanics says that a macroscopic particle has a definite trajectory (we can know its position and linear moment at any time)

(**motion equations !!**)



Werner Heisenberg |

How can we measure the position of a particle that behaves as wave?

- If we try to find the electron using a light of wavelength λ , the measurement accuracy is at most $+/- \lambda$;
- If λ is decreased to increase the accuracy of the measurement, we will have an uncertainty in the linear momentum, because the energy transferred to the electron will change its speed;
- The Heisenberg Principle of Uncertainty sets a limit to the precision for simultaneous measurements of the position and linear momentum of an atomic particle.

Heisenberg's Uncertainty Principle

$$\Delta p \cong \frac{h}{\lambda} \quad \Delta x \cong \lambda$$
$$\Delta x \Delta p \cong h$$

$$\Delta x \Delta p = \frac{\hbar}{2}$$

The measurement itself interferes with the position or velocity of the electron!

"On the mass scale of atomic particles, we can not determine exactly the position, direction of motion and velocity simultaneously"

For a quantum particle, the concept of trajectory does not exist

Classical vs Quantum

Classical Mechanics

In order to determine the state of the system, we must know the positions and momenta of all particles in each time t

- 6 variables for each particle: $V_x(t)$, $V_y(t)$, $V_z(t)$; $X(t)$, $Y(t)$, $Z(t)$
- We are able to predict the future and go back to the past
- It is a deterministic mechanic
- We have total knowledge for the state of the system

BUT....

de Broglie said \longrightarrow Electrons show wave behavior

Heisenberg said \longrightarrow For a quantum particle, the concept of trajectory is not valid.

We need a new mechanics!!

Wave function

Of course, the formulation of this new mechanics should include the quantization of energy and the wave behavior of the particles

If the particles behave as waves then there must be a wave function, a mathematical function, that describes the states of the quantum system

Such mathematical function was called the Wave Function

$$\Psi(x, y, z, t)$$

- We can not predict the future with total certainty
- Deterministic Mechanics only in Probabilities
- The full wave function is a complex function
- The wave function contains all information for the system
- We can obtain all information using operators

Schrödinger Equation

Erwin Schrödinger replaced the precise trajectory of the particle with a wave function $\Psi(x, y, z, t)$

Plane-wave equation (1-dimensional) :

$$\frac{\partial^2 \Psi(x, t)}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 \Psi(x, t)}{\partial t^2}$$

$$\Psi(x, t) = A e^{i\alpha}$$

$$\alpha = 2\pi \left(\frac{x}{\lambda} - vt \right)$$

Solution of the differential equation

Wave phase

We know that

$$E = h\nu$$

$$\lambda = \frac{h}{p}$$

$$\alpha = \left(\frac{xp_x - Et}{\hbar} \right)$$

Schrödinger Equation

Hamilton Formalism

Total energy conservation = Kinetic energy + potential energy

$$E_T = T + V$$

$$E_{Tot} = \frac{mv^2}{2} + V(x)$$

$$E_{Tot} = \frac{p_x^2}{2m} + V(x)$$

$$E_{Tot} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(\hat{x}) \equiv \hat{H}$$

\hat{H} is the **Hamiltonian Operator** in the Quantum Mechanics language

Time-dependent Schrödinger Equation

$$-\frac{\hbar}{i} \frac{\partial \Psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x, t)}{\partial x^2} + V(x, t) \Psi(x, t) = E_{Tot} \Psi(x, t)$$

Time-independent Schrödinger Equation

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

In a compact form : $\hat{H} \psi(x) = E \psi(x)$

E is the energy of the system. Solving the Schrödinger equation means obtaining ψ and **E**, that is, the wave functions and their corresponding energies;

These wave functions correspond to states of constant energy. Each wave function has its own energy, constant for that state;

We know that ψ contains all possible information about a system.

Quantum Chemistry

"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation."

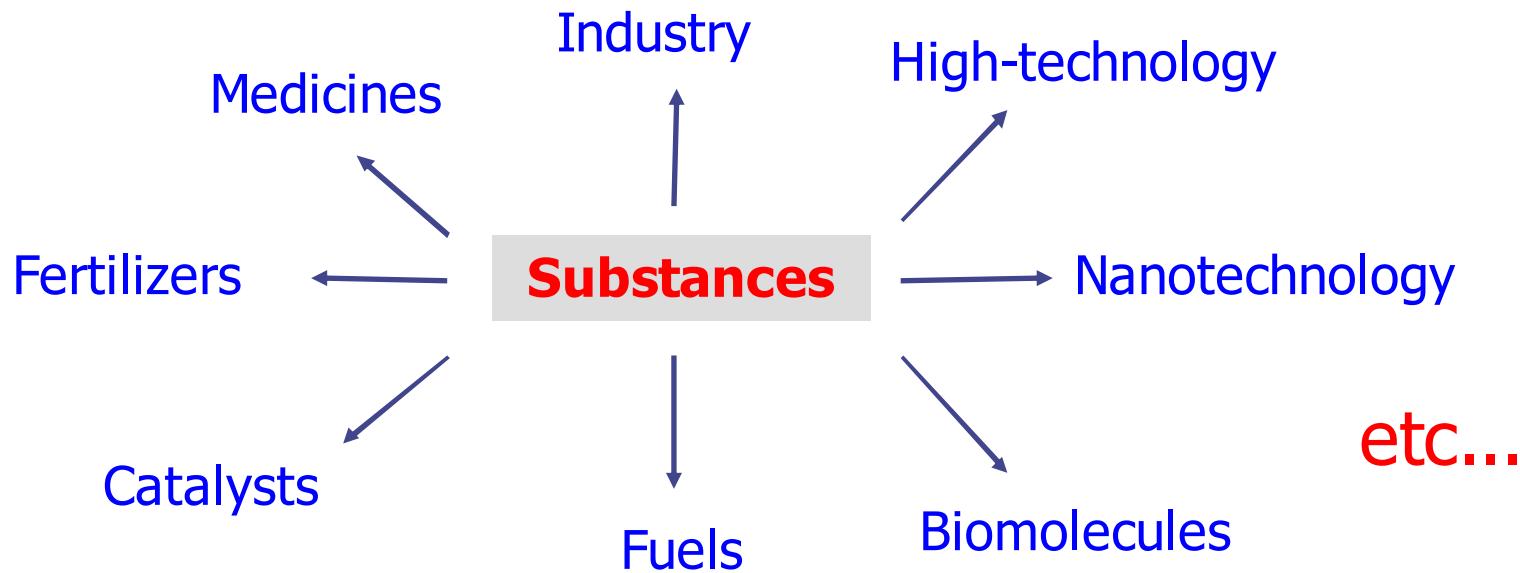
Paul Dirac, 1929.



<http://rspa.royalsocietypublishing.org/cgi/doi/10.1098/rspa.1929.0094>

**Molecular modeling for the
calculation of material properties.**

Modeling of materials: the art of the molecular and material scientists.



Molecular scientists model the materials so that they play important roles in the society

Molecular Modeling

Computational experiment on a
model system

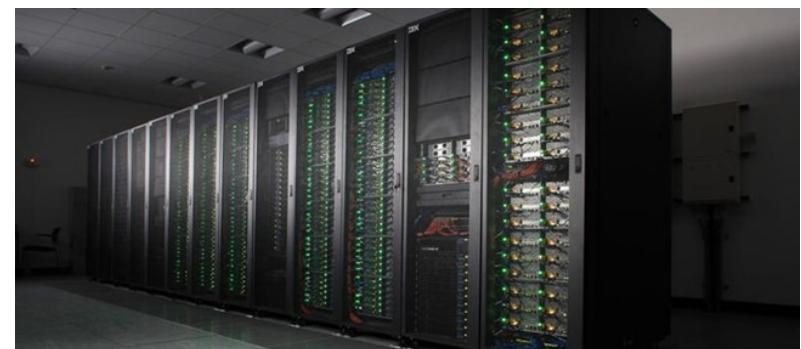
Molecular Modeling

IUPAC definition: Molecular modeling is the investigation of molecular structures and properties using computational chemistry and graphical visualization techniques in order to provide a plausible three-dimensional representation under a given set of circumstances.

<https://goldbook.iupac.org/html/M/MT06971.html>

- It requires an interaction between people from physics, chemistry, mathematics, computer science, statistic and biology.
 - model proposing
 - calculation of properties for molecules
 - synthesis of the designed materials.

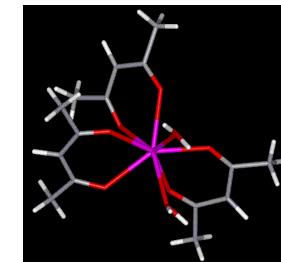
Computational experiment on a
model system



What are the methods for calculating the energy and properties of one molecular system ???

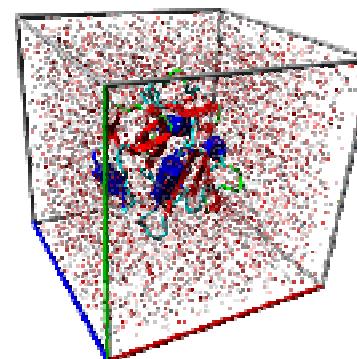
- **Classical Methods** → Molecular Mechanic

- **Quantum Methods** → { Ab initio
DFT
Semiempirical



What are the methods for calculating macroscopic properties of one ensemble of molecules ???

- Molecular dynamics
- Monte Carlo



Computational Chemistry Methods

Quantum Chemistry Methods try to solve:

$$\hat{H}\psi = E\psi$$

The Schrodinger equation was discovered in 1926 by Erwin Schrodinger, an Austrian theoretical physicist. It is an important equation that is fundamental to quantum mechanics.



Accuracy of the results

- ab initio

- Semiempirical

- Molecular Mechanic

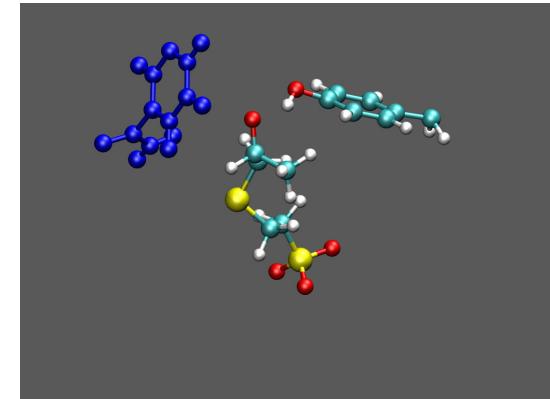
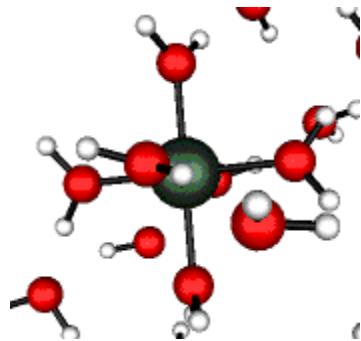
Non-empirical methods
that try to solve
Schrödinger equation

Quantum chemistry
methods that use
empirical approximations

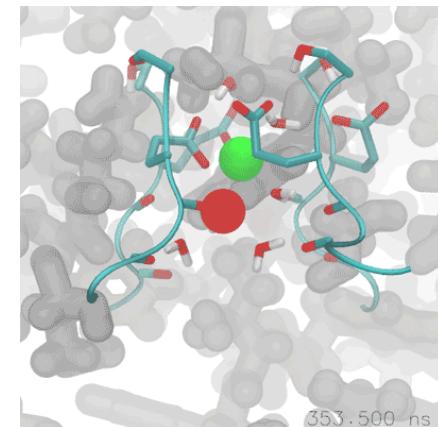
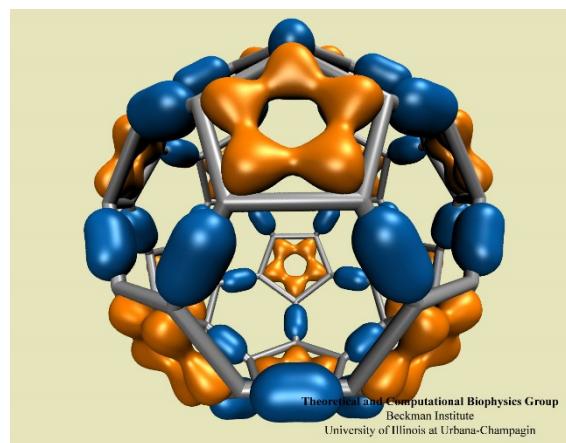
Purely empirical methods
that do not consider the
electrons

Size of the molecular systems

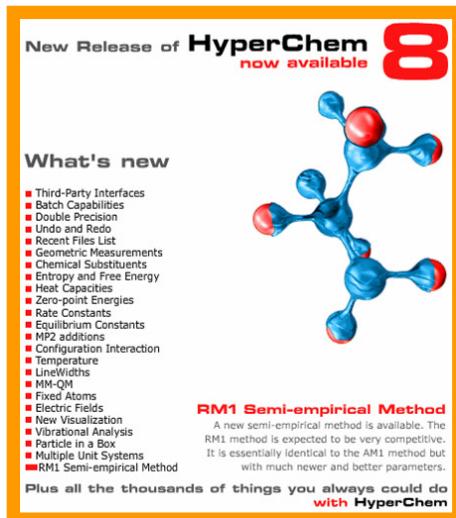
What can we do with these molecular modeling and simulation methods?



*Molecular modeling and
simulation methods*



SOFTWARES



VAMP

MOPAC®



PC GAMESS/FIREFLY



AMPAC 9 **NEW**

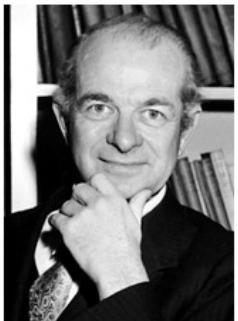
- https://en.wikipedia.org/wiki/List_of_quantum_chemistry_and_solid-state_physics_software
- <http://www.ccl.net/chemistry/links/software/index.shtml>
- <http://linux4chemistry.info/categories/index.html>
- <https://opensourcemolecularmodeling.github.io/>
- <https://doi.org/10.1016/j.jmgm.2016.07.008>



Expanding the limits of
computational chemistry

ORCA

Nobel Prizes in Theoretical Chemistry and Molecular Modeling



Linus Carl Pauling



Robert S. Mulliken



Gerhard Herzberg



William N. Lipscomb



Kenichi Fukui



Roald Hoffmann

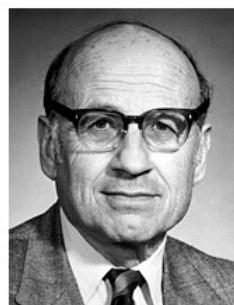
1954

1966

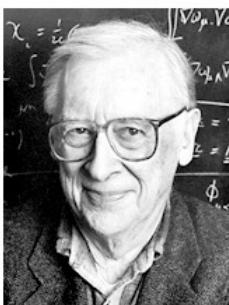
1971

1976

1981



Walter Kohn



John A. Pople

1998



Photo © Harvard University

Martin Karplus



Photo: S. Fisch

Michael Levitt



Photo: Wikimedia Commons

Arieh Warshel

2013

Quantum Chemistry Methods for Molecules

Schrödinger Equation for Molecules

The theory that describes the behavior of atomic and molecular systems is the quantum theory of matter. In this theory the main equation is the Schrödinger equation.

$$\hat{H}\Psi = E\Psi \longrightarrow \begin{array}{l} \text{Molecular wavefunction} \\ \text{Total energy} \\ \text{Hamiltonian of the system} \end{array}$$

```
graph LR; Eq["\u0302H\Psi = E\Psi"] --> H["Hamiltonian of the system"]; Eq --> E["Total energy"]; Eq --> Psi["Molecular wavefunction"]
```

Schrödinger Equation for Molecules

The theory that describes the behavior of atomic and molecular systems is the quantum theory of matter. In this theory the main equation is the Schrödinger equation.

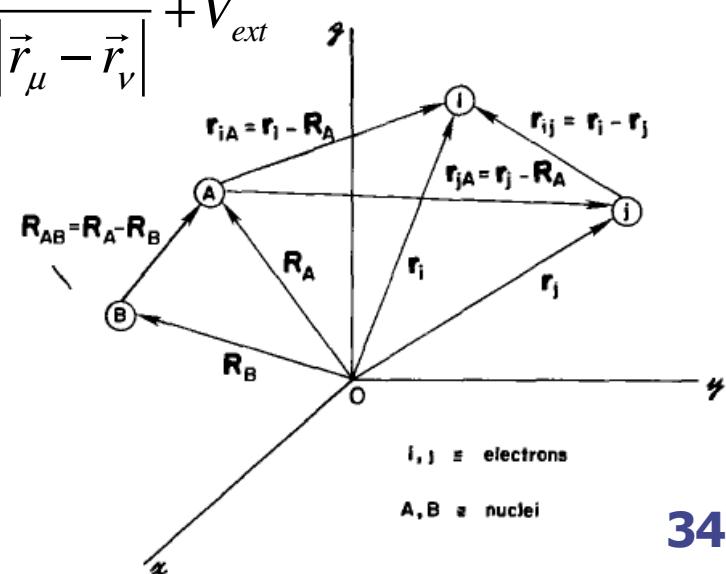
$$\hat{H}\Psi = E\Psi$$

N-nuclei and n-electrons molecular Hamiltonian:

$$\begin{aligned} \hat{H} = & \sum_{\lambda=1}^N \left(\frac{-\hbar^2}{2M_\lambda} \nabla_\lambda^2 \right) + \sum_{\lambda=1}^{N-1} \sum_{\sigma>\lambda}^N \frac{1}{4\pi\epsilon_0} \frac{Z_\lambda Z_\sigma}{|\vec{R}_\lambda - \vec{R}_\sigma|} e^2 + \sum_{\mu=1}^n \left(\frac{-\hbar^2}{2m_\mu} \nabla_\mu^2 \right) - \\ & \sum_{\lambda=1}^N \sum_{\mu=1}^n \frac{1}{4\pi\epsilon_0} \frac{Z_\lambda}{|\vec{R}_\lambda - \vec{r}_\mu|} e^2 + \sum_{\mu=1}^{n-1} \sum_{\nu>\mu}^n \frac{1}{4\pi\epsilon_0} \frac{e^2}{|\vec{r}_\mu - \vec{r}_\nu|} + V_{ext} \end{aligned}$$

Coordinates:

- $R_1 \rightarrow$ Nuclear coordinates
- $R_{12} \rightarrow$ Core-core correlation coordinate
- $r_1 \rightarrow$ Electron coordinate
- $\bar{r}_{11} \rightarrow$ Core-electron correlation coordinate
- $r_{12} \rightarrow$ Electron-electron correlation coordinate



Total molecular wave function

$$\Psi(\vec{R}_1, \dots, \vec{R}_N; R_{12}, R_{13}, \dots, R_{N-1N}; \vec{r}_1, \dots, \vec{r}_n; \bar{r}_{11}, \bar{r}_{12}, \dots, \bar{r}_{Nn}; r_{12}, r_{13}, \dots, r_{n-1n})$$

We cannot get closed expressions for the total molecular wave function because of the correlation coordinates mentioned before. So, we must apply strategies to get acceptable solutions.

Analytical solutions only for H atom and H_2^+ molecule !!



**Chemistry or
biochemistry!!!!**

Therefore, we need to do some approximations in both wave function and Hamiltonian to describe molecular systems by means of QM.

Born-Oppenheimer Approximation (1st)

$$\Psi_{BO} = \Psi_{Nuc}(\vec{R}_1, \dots, \vec{R}_N) \Psi_{ele}(\vec{r}_1, \dots, \vec{r}_n, r_{12}, \dots, r_{n-1n}; R_1, \dots, R_N)$$

Wave function without electron-nucleus correlation

- By applying BO approximation we are able to introduce concepts as molecular geometry and potential energy surface.

$$\hat{H}_{ele} \Psi_{ele}(r; R) = E(R_1, R_2, R_3, \dots, R_N) \Psi_{ele}(r; R)$$

$$\Psi_{ele}(\vec{r}_1, \dots, \vec{r}_n, r_{12}, \dots, r_{n-1n}; R_1, \dots, R_N)$$

is known as Hylleraas wavefunction (only for atoms or very small molecules).

Remove the explicit electron-electron correlation (2nd)

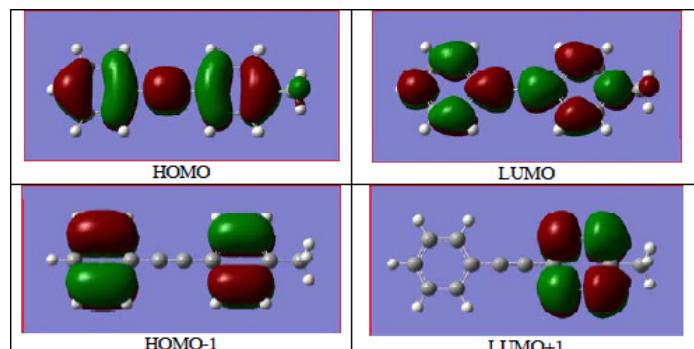
$$\Psi_{BO} = \Psi_N(\vec{R}_1, \dots, \vec{R}_N) \Psi_e(\vec{r}_1, \dots, \vec{r}_n; R_1, \dots, R_N)$$

All correlated QM methods: MPBT, CI, CC, MR-CI, MCSCF, CASSCF

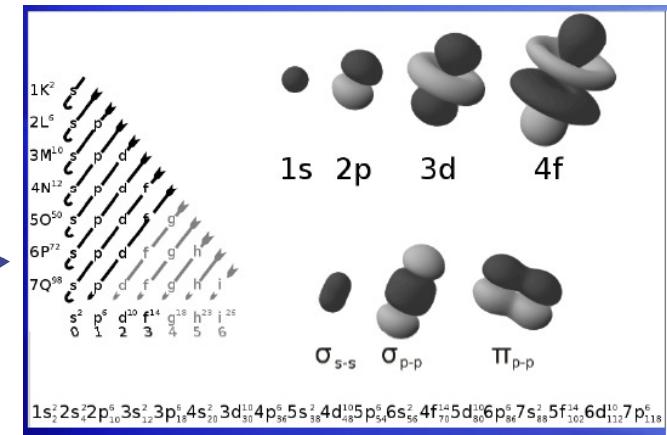
Hartree-Fock-Roothann (3rd)

- Molecular wavefunction

$$\Phi_{SD} = \begin{vmatrix} \phi_1(1) & \phi_2(1) & \cdots & \phi_N(1) \\ \phi_1(2) & \phi_2(2) & \cdots & \phi_N(2) \\ \cdots & \cdots & \cdots & \cdots \\ \phi_1(N) & \phi_2(N) & \cdots & \phi_N(N) \end{vmatrix}, \quad \langle \phi_i | \phi_j \rangle = \delta_{ij}$$



$$\phi_\mu = \sum_{i=1}^n c_{i\mu} \varphi_i$$



Molecular Orbital (MO-LCAO)

Single-determinant

Hartree-Fock

Many-determinants

post-Hartree-Fock methods

Linear combination of Slater determinants

$$\Psi = \sum \Phi_{SD}$$

Formalism of Hartree-Fock-Roothaan theory

Molecular wavefunction → Slater determinant

$$\Psi = (2n!)^{-\frac{1}{2}} \begin{vmatrix} \chi_i(x_1) & \chi_j(x_1) & \cdots & \chi_k(x_1) \\ \chi_i(x_2) & \chi_j(x_2) & \cdots & \chi_k(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i(x_{2n}) & \chi_j(x_{2n}) & \cdots & \chi_k(x_{2n}) \end{vmatrix}$$

Spin-orbital $\chi(x) = \psi(r)\alpha(w)$
or $\psi(r)\beta(w)$.

$$E = \sum_{\mu\nu} P_{\mu\nu} h_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} \left((\mu\nu|\lambda\sigma) - \frac{1}{2} (\mu\sigma|\nu\lambda) \right)$$

N^4
computationally expensive

For the electrons, we have to solve:

$$\sum_v (F_{\mu\nu} - \epsilon_i S_{\mu\nu}) c_v^{(i)} = 0 \therefore \mu = 1, 2, 3, \dots$$

Generalized eigenvalues equation

Matrix form

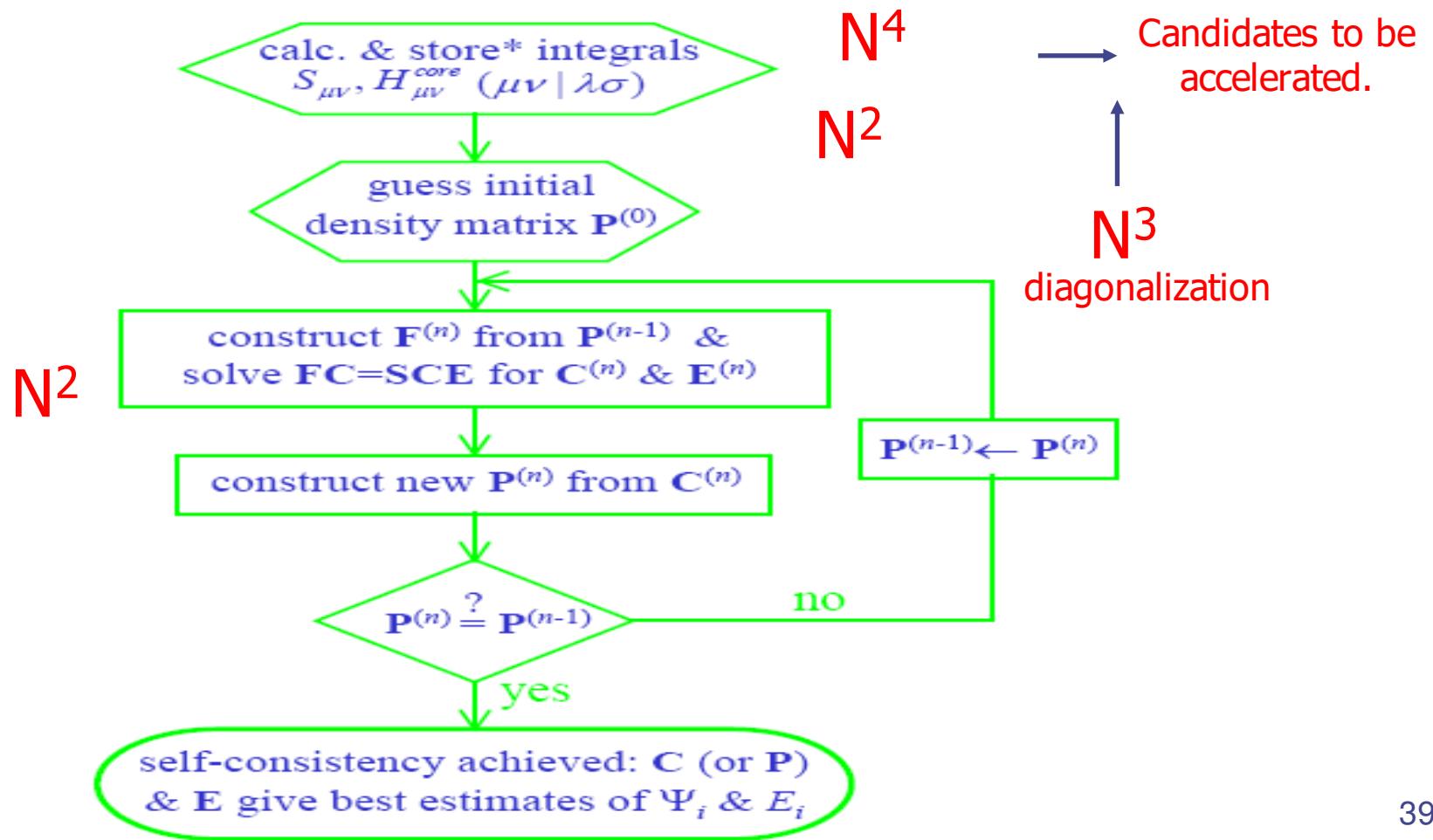
$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\epsilon}$$

$$F_{\mu\nu} = h_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma} \left((\mu\nu|\lambda\sigma) - \frac{1}{2} (\mu\sigma|\nu\lambda) \right)$$

This formulation is computationally expensive for large molecular systems

Where are the bottlenecks?

Complexities for some parts of conventional single point energy Hartree-Fock calculation (SCF algorithm)

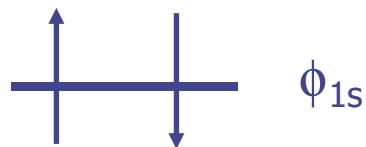


RHF and UHF wavefunctions

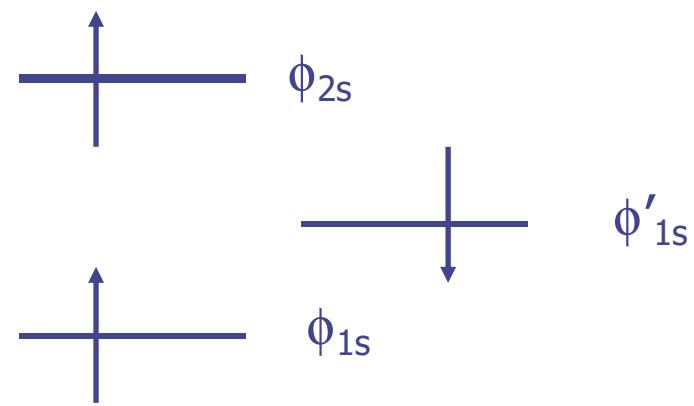
RHF → Restricted HF

UHF → Unrestricted HF

The ground state for the Li atom: $1s^2 2s^1$



$${}^2\Psi_{\text{RHF}} = \det[\phi_{1s}\alpha \phi_{1s}\beta \phi_{2s}\alpha]$$



$${}^2\Psi_{\text{UHF}} = \det[\phi_{1s}\alpha \phi'_{1s}\beta \phi_{2s}\alpha]$$

It is not eigenfunction of \hat{S}^2

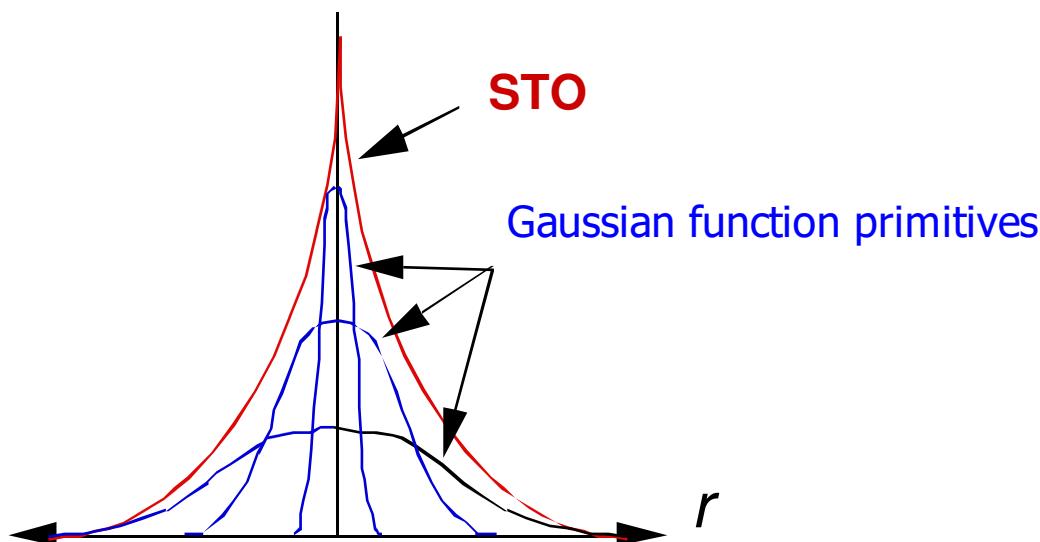
- RHF wavefunctions must be used only for closed-shell systems (e.g. organic molecules)
- For radicals and molecules containing metals we use UHF or ROHF wavefunctions.

Basis set - STO x Gaussians

- Slater functions;
 - Correct short-range and long-range behavior
- Gaussians functions;
 - Cusp in $r=0$? No, GTOs have zero slope at the nucleus.
 - But, there are analytical derivatives
- Linear combination of Gaussian functions : STO-3G
 - 3 Gaussian primitives to fit 1 STO.
 - 6 GTOs fit well 1 STO.

$$\phi_{1s}(\vec{r};\zeta_1) = \sqrt{\frac{\zeta_1^3}{\pi}} \exp(-\zeta_1 \vec{r})$$

$$g_s(\vec{r};\alpha) = \left(\frac{2\alpha}{\pi}\right)^{3/4} \exp(-\alpha \vec{r}^2)$$



Basis set functions

Cartesian GTOs → Boys 1950.

$$x_A^l y_A^m z_A^n e^{-\alpha r_A^2}$$

$1s$

$$\phi_{1s}^{GF}(\alpha, r_A) = \left(\frac{2\alpha}{\pi}\right)^{\frac{3}{4}} e^{-\alpha r_A^2}$$

$2p_x$

$$\phi_{2p_x}^{GF}(\alpha, r_A) = \left(\frac{128\alpha^5}{\pi^3}\right)^{\frac{1}{4}} x_A e^{-\alpha r_A^2}$$

$2p_y$

$$\phi_{2p_y}^{GF}(\alpha, r_A) = \left(\frac{128\alpha^5}{\pi^3}\right)^{\frac{1}{4}} y_A e^{-\alpha r_A^2}$$

$2p_z$

$$\phi_{2p_z}^{GF}(\alpha, r_A) = \left(\frac{128\alpha^5}{\pi^3}\right)^{\frac{1}{4}} z_A e^{-\alpha r_A^2}$$

$3d_{xy}$

$$\phi_{3d_{xy}}^{GF}(\alpha, r_A) = \left(\frac{2048\alpha^7}{\pi^3}\right)^{\frac{1}{4}} x_A y_A e^{-\alpha r_A^2}$$

etc.

Basis set functions (STO-NG)

Contracted GTO → GTO primitives expansion

$$\phi_{\mu}^{\text{CGF}}(r_A) = \sum_{p=1}^L d_{p\mu} \phi_p^{\text{GF}}(\alpha_p, r_A)$$

The coefficients are kept fixed during the calculation.

STO-NG

$$\phi_{1s}^{\text{CGF}}(\zeta = 1.0) = \sum_{i=1}^N d_{i, 1s} \phi_{i, 1s}^{\text{GF}}(\alpha_{i, 1s})$$

$$\phi_{2s}^{\text{CGF}}(\zeta = 1.0) = \sum_{i=1}^N d_{i, 2s} \phi_{i, 1s}^{\text{GF}}(\alpha_{i, 2sp})$$

$$\phi_{2p}^{\text{CGF}}(\zeta = 1.0) = \sum_{i=1}^N d_{i, 2p} \phi_{i, 2p}^{\text{GF}}(\alpha_{i, 2sp})$$

Basis set functions (split-valence)

They use more than one function for valence orbitals and only one function for inner orbitals. Ex: 3-21G, 6-31G, etc.

- 3-21G

Carbon \rightarrow (1s, 2s, 2p_x, 2p_y, 2p_z).

Inner orbitals (1s) \rightarrow CGF = d₁GF(1) + d₂GF(2) + d₃GF(3)

Valence Orbitals (2s, 2p_x, 2p_y, 2p_z)

Ex:

2s \rightarrow CGF₁ = c₁GF(4) + c₂GF(5)

CGF₂ = a₁GF(6), having smaller exponent (diffuse function).

The same is applied for the other valence orbitals.

Basis set functions (split-valence)

- 6-31G

Carbon → (1s, 2s, 2p_x, 2p_y, 2p_z).

Inner orbitals (1s) → CGF = d₁GF(1) + d₂GF(2) + d₃GF(3) +
d₄GF(4) + d₅GF(5) + d₆GF(6)

Valence orbitals (2s, 2p_x, 2p_y, 2p_z)

Ex:

2s → CGF₁ = c₁GF(7) + c₂GF(8) + c₃GF(9)
CGF₂ = a₁GF(6) , having smaller exponent (diffuse function).

The same is applied for the other valence orbitals

Basis set functions (polarization)

- Insert basis functions of higher angular momentum.
- Important to describe chemical bonding.
- 6-31G*

Carbon → (1s, 2s, 2p_x, 2p_y, 2p_z, 3d_{xy}, 3d_{xz}, 3d_{yz}, 3d_{zz}, 3d_{yy}, 3d_{xx})

- 6-31G**

Insert polarization function also for the Hydrogen atom

Carbon → (1s, 2s, 2p_x, 2p_y, 2p_z, 3d_{xy}, 3d_{xz}, 3d_{yz}, 3d_{zz}, 3d_{yy}, 3d_{xx})

Hydrogen → (1s, 2p_x, 2p_y, 2p_z)

Basis set functions (diffuse)

- Insert basis functions of the same angular momentum, but with different exponents (diffuse = small exponent).
- Important for anions, dimers formed by hydrogen bonding and atoms with lone pairs.

- 6-31+G

Carbon → (1s, 2s, 2s, 2p_x, 2p_y, 2p_z, 2p_x, 2p_y, 2p_z)

- 6-31++G

Insert diffuse basis function also for the Hydrogen atom

Carbon → (1s, 2s, 2s, 2p_x, 2p_y, 2p_z, 2p_x, 2p_y, 2p_z)

Hydrogen → (1s, 1s)

Pople basis set functions

- STO-NG
- 3-21G
- 4-31G
- 6-31G
- 6-311G

Limit

6-311++G**

Electron Correlation problem

- The Hartree-Fock method gets 99% of the total electronic energy.
- The electron correlation is defined by:

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}}$$

Methods to get E_{corr} :

- CI, MBPT, CC, MCSCF, MR-CI

Electron Correlation problem

How important is the electron correlation?

Let's consider the Helium Atom

Core (4u) + 2 electrons

The simplest atomic system with electron correlation.

$E = -2.90372$ a.u. (Exact value)

$E = -2.86168$ a.u. (Hartree-Fock limit)

$E_{\text{corr}} = 0.04204$ a.u. $\approx \mathbf{26.0 \text{ kcal.mol}^{-1}}$

Let's quantify how large is that error!

Electron Correlation problem

Chemical accuracy

- What is the error in the ΔG that changes K_{eq} in 10 times?
- What is the error in the $\Delta G^\#$ that changes k in 10 times?

1.4 kcal.mol⁻¹

Conclusion: The chemistry is in the electron correlation.

How to improve our predictions?

Estimates for Computational Demand

- Ab initio methods

Each new digit for energy requires 10,000 times more computation time.

1 min → 1 weeks → 200 years → 20 thousands centuries

100 kcal/mol 10 kcal/mol 1 kcal/mol 0.1 kcal/mol

Conclusion: we need fast computers and/or efficient numerical strategies.

Ab-initio

- Positive points:
 - They are methods of general use
 - They do not use empirical approximations
 - All chemical and physical phenomena can be modeled: chemical reactions, polarizabilities, spectra, optical properties, electronic properties, etc.
- Limitations:
 - Applied to small systems and / or few molecules
 - Demand a lot of computational resources

Semiempirical theory

Let's talk about the three main approximations that define the formalism of all semiempirical methods.

Approximations for the Semiempirical wavefunction (1st)

Simplifications in HFR formulation allows calculations for larger molecular systems.

In semiempirical methods the electronic configuration is split in two parts:

- inner shell electrons (core), Ψ_c
- outer shell electrons (valency), Ψ_v

$$|\Psi^S\rangle$$

It can be written as an anti-symmetrized product containing these two parts (core and valence electrons)

Experiments indicate that the inner shell electrons are relatively inert in typical chemical processes. So, we expect that the assumption of a fixed, Ψ_c , should not lead to significant errors.

$$H_{\mu\nu}^S = \left\langle \phi_\mu \left[-\frac{1}{2} \nabla^2 - \sum_A^M \frac{Z_A'}{\left| \vec{R}_A - \vec{r}_i \right|} \right] \phi_\nu \right\rangle \rightarrow \sum_v^{n_{val}} F_{\mu\nu} c_{vi} = \sum_v^{n_{val}} S_{\mu\nu} c_{vi} \mathcal{E}_i$$

The inner shell electrons are frozen or implicitly considered



smaller matrices and less equations to solve.

Zero Differential Overlap Approximation (ZDO) (2nd)

$$S_{\mu_A \nu_B} = \langle \varphi_{\mu_A} | \varphi_{\nu_B} \rangle = \delta_{\mu\nu} \delta_{AB}$$

1. ZDO approximation impacts in the overlap matrix. The overlap matrix becomes an identity matrix and consequently it can be eliminated in the generalized eigenvalues HFR equation;
2. ZDO approximation also impacts in the two-electron integrals. All three- and four-centre two-electron repulsion integrals are neglected and some of two-centre are also eliminated (non-Coulombic type);

$$(\mu\nu | \lambda\sigma) = (\mu\mu | \lambda\lambda) \delta_{\mu\nu} \delta_{\lambda\sigma} \longrightarrow$$

Only up to two-centre two-electron repulsion integrals need to be calculated.

Integral approximations (3rd)

Zero differential overlap - ZDO

$$\int \varphi_{\mu_A}^*(1) \varphi_{\nu_B}(1) d\tau_1 = 0 \quad \mu_A \neq \nu_B$$

Integral approximations

CNDO

$$(\mu_A \nu_B | \lambda_C \sigma_D) = \\ (\mu_A \mu_A | \lambda_C \lambda_C) \delta_{\mu\nu} \delta_{\lambda\sigma} \delta_{AB} \delta_{CD}$$

INDO

$$(\mu_A \nu_B | \lambda_C \sigma_D) = \\ (\mu_A \nu_A | \lambda_C \sigma_C) \delta_{\mu_A \nu_B} \delta_{\lambda_C \sigma_D}$$

NDDO

$$(\mu_A \nu_B | \lambda_C \sigma_D) = \\ (\mu_A \nu_A | \lambda_C \sigma_C) \delta_{AB} \delta_{CD}$$

Parameterization procedure:

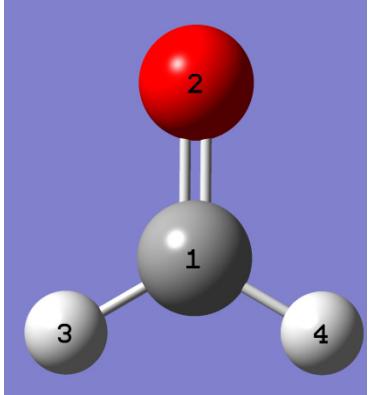


Adjusted using numerical procedures



Taken from experiments or high level calculations

Assembling the Fock Matrix for NDDO



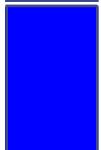
C ¹				O ²				H ³ H ⁴		
	s	p _x	p _y	p _z	s	p _x	p _y	p _z	s	s
C ¹	s	1								
	p _x	2	3							
	p _y	4	5	6						
	p _z	7	8	9	10					
O ²	s	11	12	13	14	15				
	p _x	16	17	18	19	20	21			
	p _y	22	23	24	25	26	27	28		
	p _z	29	30	31	32	33	34	35	36	
H ³	s	37	38	39	40	41	42	43	44	45
	s	46	47	48	49	50	51	52	53	54
H ⁴	s	46	47	48	49	50	51	52	53	55



$$F_{\mu\mu} = H_{\mu\mu} + \sum_v^A P_{vv} [(\mu\mu|vv) - \frac{1}{2}(\mu v|\mu v)] + \sum_B \sum_{\lambda\sigma} P_{\lambda\sigma} (\mu\mu|\lambda\sigma)$$



$$F_{\mu\nu} = H_{\mu\nu} + \frac{1}{2} P_{\mu\nu} [3(\mu\nu|\mu\nu) - (\mu\mu|vv)] + \sum_B \sum_{\lambda\sigma} P_{\lambda\sigma} (\mu\nu|\lambda\sigma) \quad \therefore \quad \mu, \nu \in A$$



$$F_{\mu\lambda} = H_{\mu\lambda} - \frac{1}{2} \sum_v^A \sum_{\sigma} P_{v\sigma} (\mu\nu|\lambda\sigma) \quad \therefore \quad \mu \in A \quad \lambda \in B$$

O(N²)

Parametrization procedure

A conception of a semiempirical method is composed by:

1. its theory
2. a parameterization procedure
3. the statistical validation.

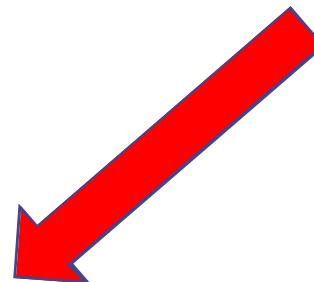
**Defines the predictive power
of the semiempirical method**

You have to decide about:

- Response function;
- Molecular data set;
- Properties that will be used;
- Numerical minimization algorithms;

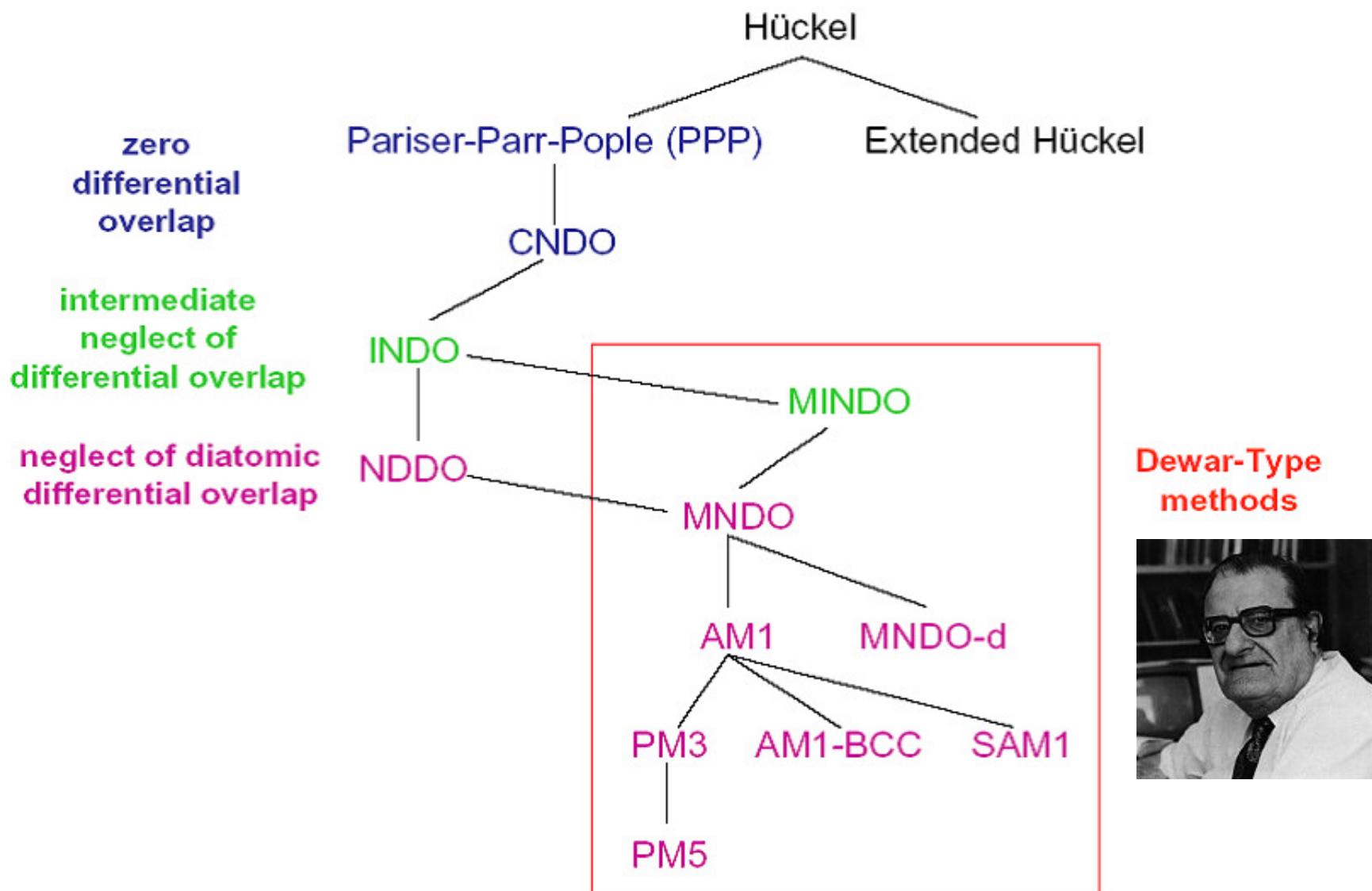
$$F^{resp} = \sum_{i=1}^n \left(X_i^{Calc} - X_i^{Exp} \right)^2 \cdot w_i^2$$

The parameters impact this part



Property	Weight
Heat of formation (ΔH_f)	1 kcal ⁻¹ mol
Ionization Potential (IP)	10 eV ⁻¹
Dipole Moment (μ)	20 D ⁻¹
Bond Length (R_{ab})	100 Å ⁻¹
Angle (θ_l)	2/3 degree ⁻¹
Torsion angle (θ_d)	1/3 degree ⁻¹

Semiempirical Hamiltonians



Dewar-type Methods → New directions for Semiempirical Methods

New Semiempirical methods

With emphasis on the study of biomolecular systems

Recife Model 1 (RM1)

Theory = NDDO/AM1

◆ Atoms (24)

- C, H, N, O ← **Organic Chemistry**
- P, S ← **Biochemistry**
- F, Cl, Br, I ← **Drugs**
- Lanthanides ← **Photonics and Medicine**

◆ 191 + 252 = 443 parameters



RM1: A Reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I

GERD B. ROCHA,¹ RICARDO O. FREIRE,¹ ALFREDO M. SIMAS,¹ JAMES J. P. STEWART²

¹Departamento de Química Fundamental, CCEN, UFPE, 50590-470, Recife PE, Brazil

²Stewart Computational Chemistry, 15210 Paddington, Colorado Springs Colorado 80921-2512

Received 25 October 2005; Accepted 27 December 2005

DOI 10.1002/jcc.20425

Published online in Wiley InterScience (www.interscience.wiley.com).

www.rm1.sparkle.pro.br

RM1 Semiempirical Molecular Orbital Model - Windows Internet Explorer
http://www.rm1.sparkle.pro.br/ Google

Novidades Perfil Email Fotos Calendário MSN Compartilhar Entrar

Favoritos Sites Sugeridos Obtenha mais completo... RM1 Semiempirical Molecular Orbital Model

中文 | 日本語 | 한국어 |ไทย

RM1 Semiempirical Molecular Orbital Model

Scholarly article

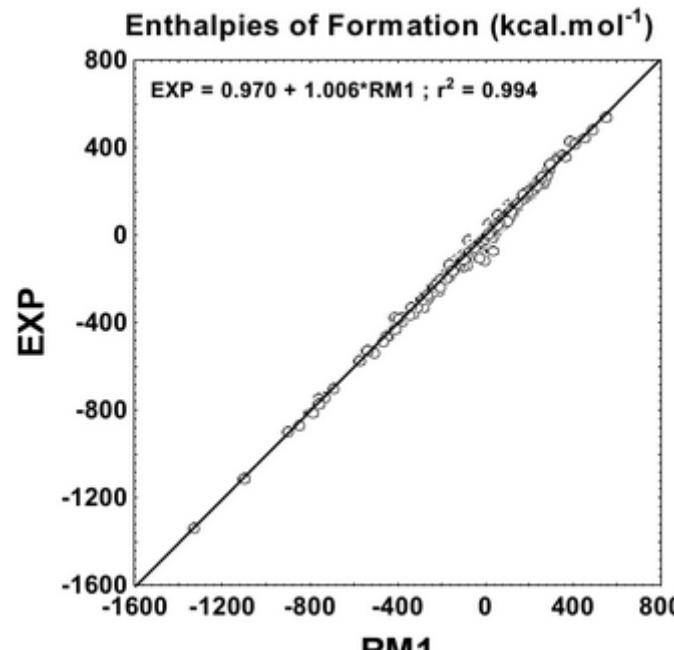
RM1: a Reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I

Gerd Bruno Rocha, Ricardo Oliveira Freire, Alfredo Mayall Simas*, and James J. P. Stewart.
Journal of Computational Chemistry 27(10), 1101-1111, 2006

Predictive power

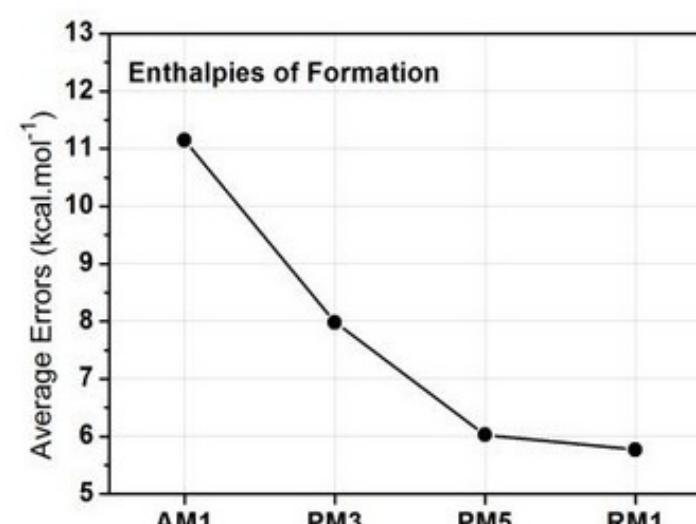
Enthalpies of Formation (kcal.mol⁻¹)

EXP = 0.970 + 1.006*RM1 ; r² = 0.994



Average Errors (kcal.mol⁻¹)

Enthalpies of Formation



Semiempirical methods X molecular force fields

11938

J. Phys. Chem. A 2009, 113, 11938–11948

Are Current Semiempirical Methods Better Than Force Fields? A Study from the Thermodynamics Perspective[†]

Gustavo de M. Seabra,[‡] Ross C. Walker,[§] and Adrian E. Roitberg^{**‡}

The semiempirical Hamiltonians MNDO, AM1, PM3, RM1, PDDG/MNDO, PDDG/PM3, and SCC-DFTB, when used as part of a hybrid QM/MM scheme for the simulation of biological molecules, were compared on their abilities to reproduce experimental ensemble averages at or near room temperatures for the model system alanine dipeptide in water. Free energy surfaces in the (ϕ , ψ) dihedral angle space, $^3J(\text{H}_\alpha, \text{H}_\beta)$ NMR dipolar coupling constants, basin populations, and peptide–water radial distribution functions (RDF) were calculated from replica exchange simulations and compared to both experiment and fully classical force field calculations using the Amber ff99SB force field. In contrast with the computational chemist's intuitive idea that the more expensive a method the better its accuracy, the ff99SB force field results were more accurate than most of the semiempirical methods, with the exception of RM1. None of the methods, however, was able to accurately reproduce the experimental data. Analysis of the results indicate that the specific QM/MM interactions have little influence on the sampling of free energy surfaces, and the differences are well explained simply by the intrinsic properties of the various QM methods.

Force fields are suitable to protein dynamics but the atomic charges are prefixed, and there is no explicit treatment of:

1. electrostatic polarization
2. charge transfer

Duan, L. L.; Mei, Y.; Zhang, D.; Zhang, Q. G.; Zhang, J. Z. H. Folding of a Helix at Room Temperature Is Critically Aided by Electrostatic Polarization of Intraprotein Hydrogen Bonds. *J. Am. Chem. Soc.* 2010, 132, 11159–11164.

Tong, Y.; Mei, Y.; Li, Y. L.; Ji, C. G.; Zhang, J. Z. H. Electrostatic Polarization Makes a Substantial Contribution to the Free Energy of Avidin-Biotin Binding. *J. Am. Chem. Soc.* 2010, 132, 5137–5142.

PM6 method

J Mol Model (2007) 13:1173–1213
DOI 10.1007/s00894-007-0233-4

ORIGINAL PAPER

Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements

James J. P. Stewart

Abstract Several modifications that have been made to the NDDO core–core interaction term and to the method of parameter optimization are described. These changes have resulted in a more complete parameter optimization, called PM6, which has, in turn, allowed 70 elements to be parameterized. The average unsigned error (AUE) between calculated and reference heats of formation for 4,492 species was $8.0 \text{ kcal mol}^{-1}$. For the subset of 1,373 compounds involving only the elements H, C, N, O, F, P, S, Cl, and Br, the PM6 AUE was $4.4 \text{ kcal mol}^{-1}$. The equivalent AUE for other methods were: RM1: 5.0, B3LYP 6–31G*: 5.2, PM5: 5.7, PM3: 6.3, HF 6–31G*: 7.4, and AM1: $10.0 \text{ kcal mol}^{-1}$. Several long-standing faults in AM1 and PM3 have been corrected and significant improvements have been made in the prediction of geometries.

PM6 Method

Comparison with RM1

In 2006, ten elements, H, C, N, O, F, P, S, Cl, Br, and I, that had been parameterized at the AM1 level were re-parameterized [35]; the result was a new method, RM1. No changes were made to the set of approximations used, so that, for example, P, S, Cl, Br, and I used only the *s-p* basis set.

That is, RM1 was functionally identical to AM1. A statistical analysis showed that RM1 was more accurate than any of the other NDDO methods, and therefore was the method of choice for modeling organic compounds.

An indication of the effect of the current changes to the set of approximations can be obtained by comparing the AUE for PM6 and RM1 in Tables 10, 11, 12, 13 and 14.

Voityuk reported the parameterization of molybdenum [14] at the AM1* level. These parameters were added to the standard AM1 parameters and were used in the analysis.

OM1 and OM2

ZDO approximation adds two deficiencies to the semiempirical methods:

- 1) Torsion angles for Conformational analysis
- 2) Ionization potentials

Theor Chem Acc (2000) 103:495–506
DOI 10.1007/s002149900083

Theoretical
Chemistry Accounts

Regular article

Orthogonalization corrections for semiempirical methods

Wolfgang Weber*, Walter Thiel**

Organisch-chemisches Institut, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Received: 30 March 1999 / Accepted: 29 July 1999 / Published online: 2 November 1999
© Springer-Verlag 2000

Basis set orthogonalization

In HFR approach, we solve this matrix-equation $\rightarrow \mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon}$

In ZDO semiempirical methods, we solve a direct eigenvalue equation.

$$\mathbf{F}^{\text{ZDO}} \approx \varphi \mathbf{F}$$

- Corrections $\longrightarrow {}^\varphi F = {}^\varphi H + {}^\varphi G$

$${}^\varphi H = S^{-\frac{1}{2}} H S^{-\frac{1}{2}}$$

$${}^\varphi G = S^{-\frac{1}{2}} G S^{-\frac{1}{2}}$$

$$S^{-\frac{1}{2}} = (I + X)^{-\frac{1}{2}} = I - \frac{1}{2} X + \frac{3}{8} X^2 - \frac{5}{16} X^3 + \dots$$

$$X = S - I$$

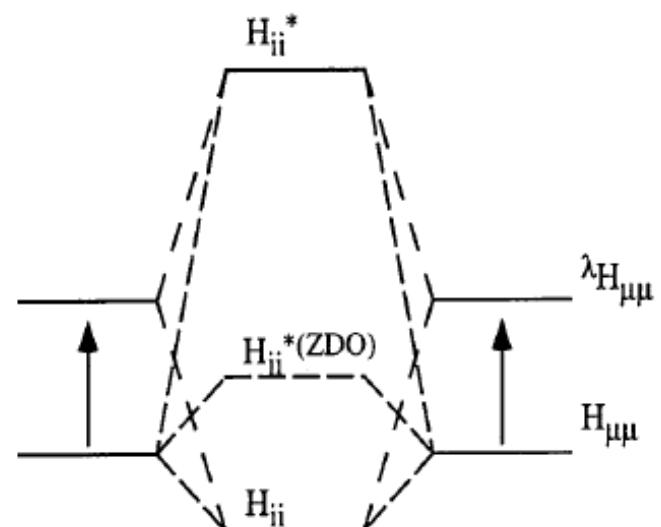
$$\longrightarrow {}^\varphi H = (I + X)^{-\frac{1}{2}} H (I + X)^{-\frac{1}{2}}$$

$${}^\varphi H = H - \frac{1}{2}(HX + XH) + \frac{3}{8}(X^2 H + HX^2) + \frac{1}{4} X H X + \dots$$

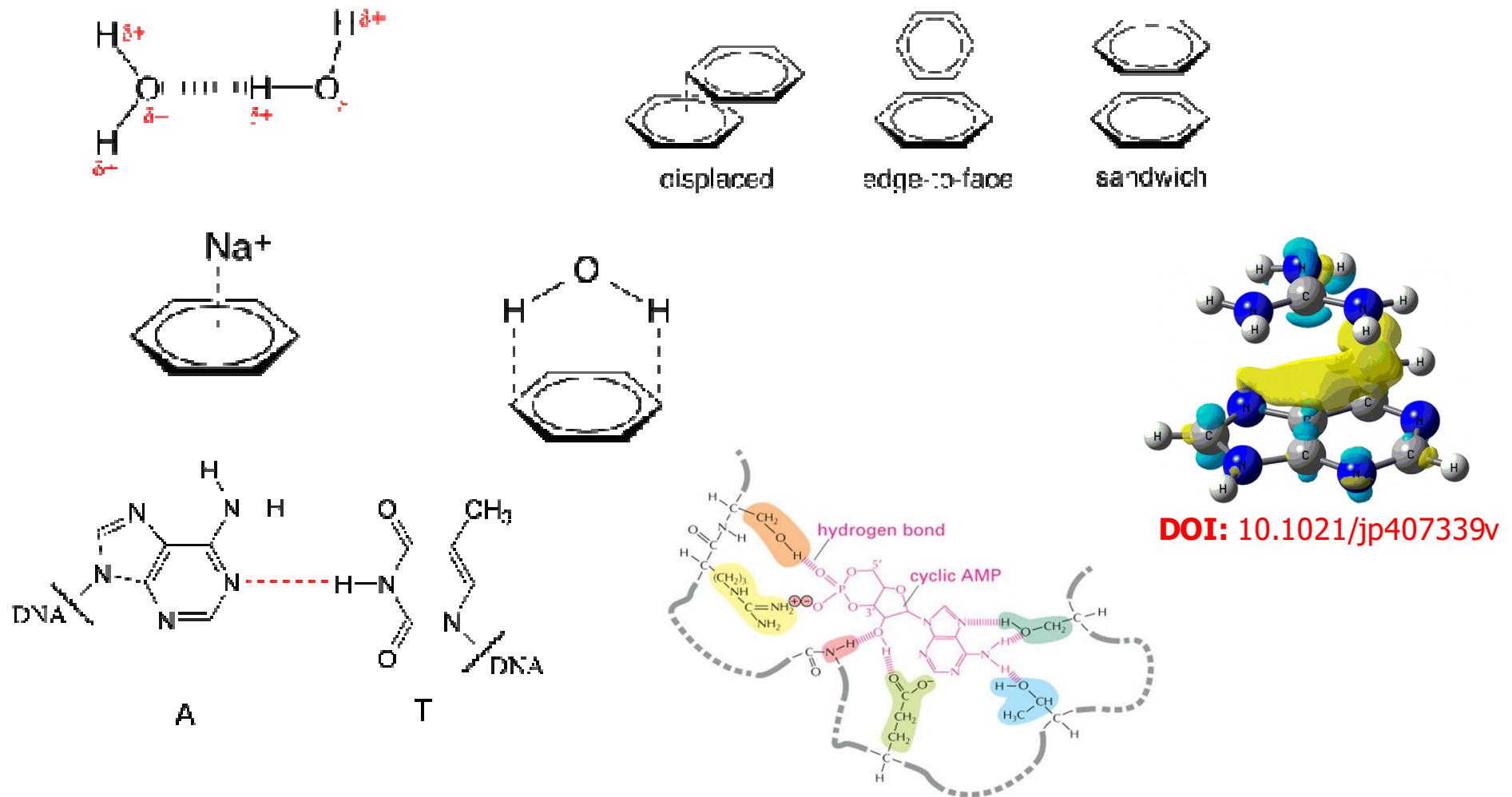
Ionization Potential

Conformational analysis

$${}^\varphi H_{\mu\lambda} = \beta_{\mu\lambda} - \frac{1}{2} \sum_{\rho}^C (S_{\mu\rho} \beta_{\rho\lambda} + \beta_{\mu\rho} S_{\rho\lambda}) + \frac{1}{8} \sum_{\rho}^C S_{\mu\rho} S_{\rho\lambda} (H_{\mu\mu} + H_{\lambda\lambda} + 2H_{\rho\rho})$$



Non-covalent interactions



DOI: [10.1021/jp407339v](https://doi.org/10.1021/jp407339v)

- Low-level quantum chemistry methods fail in describing non-covalent interactions: semiempirical methods, DFT, TB-DFT, RHF, etc.

DFT-D Methods

The main idea behind the DFT-D method [Grimme, S., J. Comput. Chem., 25 (2004) 1463.] is to add dispersion corrections after calculating the total converged energy.

$$E_{DFT-D} = E_{DFT} + E_{disp}$$

$$E_{disp} = -s_6 \sum_{i=1}^{N_{at}-1} \sum_{j=i+1}^{N_{at}} \frac{C_6^{ij}}{R_{ij}^6} f_{dump}(R_{ij}) \quad C_6^{ij} = 2 \cdot \frac{C_6^i C_6^j}{C_6^i + C_6^j}$$

N_{at} is the number of atoms, C_6 is the dispersion coefficients for the pair of atoms ij , s_6 is a global multiplicative parameter and R_{ij} is the interatomic distance. A function, f_{dump} , is added to avoid singularities.

$$f_{dump}(R_{ij}) = \frac{1}{1 + e^{-\alpha(\frac{R}{R_0} - 1)}}$$

This function introduces another adjustable parameter, α .

R_0 is obtained from the sum of the van der Waals radii for the atomic pair ij .

AM1-D and PM3-D Methods

PAPER

www.rsc.org/pccp | Physical Chemistry Chemical Physics

Semi-empirical molecular orbital methods including dispersion corrections
for the accurate prediction of the full range of intermolecular interactions
in biomolecules†

Jonathan P. McNamara and Ian H. Hillier*

Received 6th February 2007, Accepted 5th March 2007

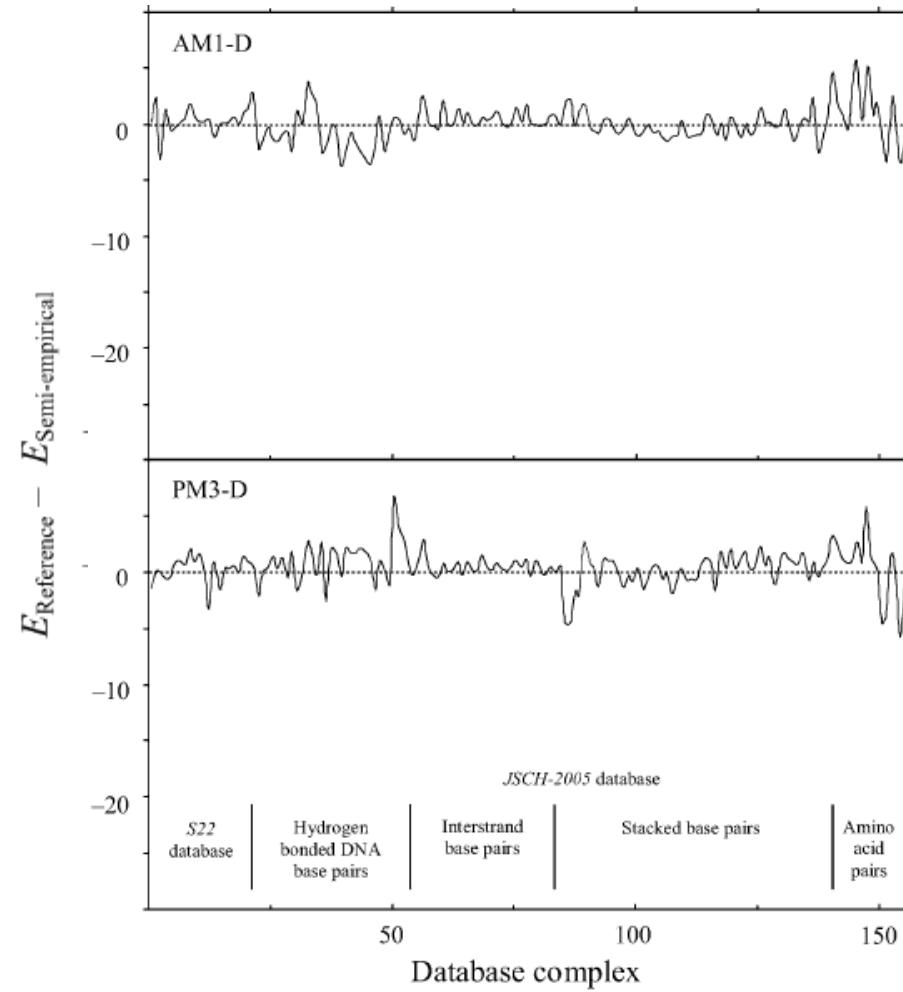
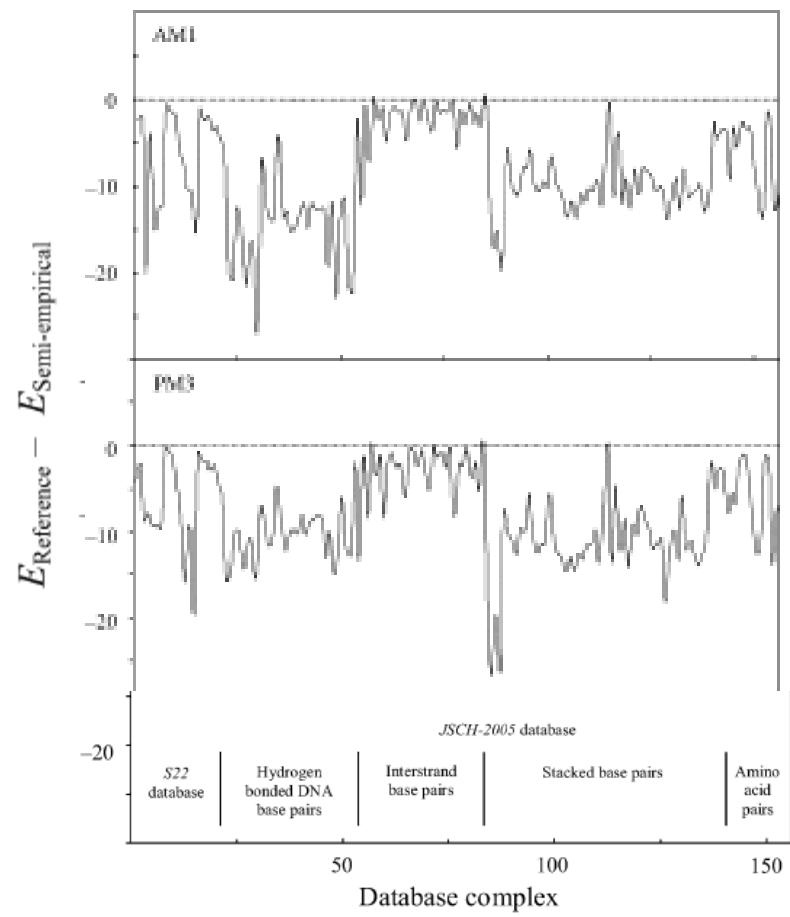
First published as an Advance Article on the web 22nd March 2007

DOI: 10.1039/b701890h

$$E_{\text{PM3-D}} = E_{\text{PM3}} + E_{\text{disp}}$$

1st Semiempirical version for DFT-D

SM methods AM1-D and PM3-D



Non-covalent corrections for many Semiempirical Hamiltonians

Advanced Corrections of Hydrogen Bonding and Dispersion for
Semiempirical Quantum Mechanical Methods

Jan Řezáč *† and Pavel Hobza †‡

† Institute of Organic Chemistry and Biochemistry,
Academy of Sciences of the Czech Republic and Center
for Biomolecules and Complex Molecular Systems, 166
10 Prague, Czech Republic

‡ Regional Centre of Advanced Technologies and
Materials, Department of Physical Chemistry, Palacky
University, 771 46 Olomouc, Czech Republic

J. Chem. Theory Comput., 2012, 8 (1), pp 141-151

DOI: 10.1021/ct200751e

$$E_{SM-D} = E_{SM} + E_{disp}$$

SM = AM1-D, PM3-D, RM1-D, PM6-DH2, PM6-D3,
PM6-DH+, OMx-D

PM7

- Corrected some NDDO errors;
 - For solids
- Special parameterizations for:
 - Solids
 - Crystals
 - Transition metals
 - non-covalent interactions;

J Mol Model (2013) 19:1–32
DOI 10.1007/s00894-012-1667-x

ORIGINAL PAPER

Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters

James J. P. Stewart

OMx Methods

OMx-D: semiempirical methods with orthogonalization and dispersion corrections. Implementation and biochemical application[†]

Tell Tuttle^{*a} and Walter Thiel^{*b}

Received 5th December 2007, Accepted 25th January 2008

First published as an Advance Article on the web 25th February 2008

DOI: 10.1039/b718795e

The semiempirical methods of the OMx family (orthogonalization models OM1, OM2, and OM3) are known to describe biochemical systems more accurately than standard semiempirical approaches such as AM1. We investigate the benefits of augmenting these methods with an empirical dispersion term (OMx-D) taken from recent density functional work, without modifying the standard OMx parameters. Significant improvements are achieved for non-covalent interactions, with mean unsigned errors of 1.41 kcal/mol (OM2-D) and 1.31 kcal/mol (OM3-D) for the binding energy of the complexes in the JSCH-2005 data base. This supports the use of these augmented methods in quantum mechanical/molecular mechanical (QM/MM) studies of biomolecules, for example during system preparation and equilibration. As an illustrative application, we present QM and QM/MM calculations on the binding between antibody 34E4 and a hapten, where OM3-D performs better than the methods without dispersion terms (AM1, OM3).

OMx Methods

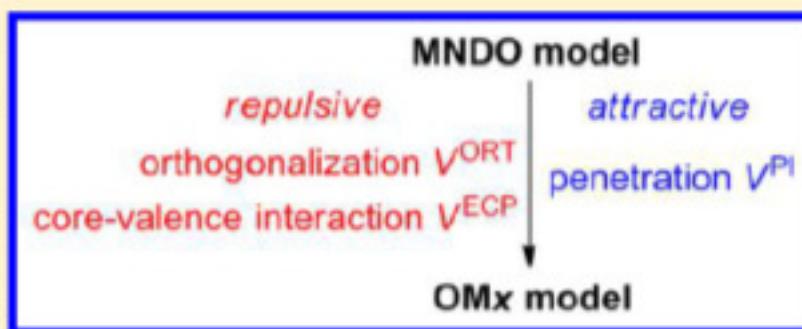
Semiempirical Quantum-Chemical Orthogonalization-Corrected Methods: Theory, Implementation, and Parameters

Pavlo O. Dral, Xin Wu, Lasse Spörkel, Axel Koslowski, Wolfgang Weber,[†] Rainer Steiger,[‡] Mirjam Scholten,[§] and Walter Thiel*

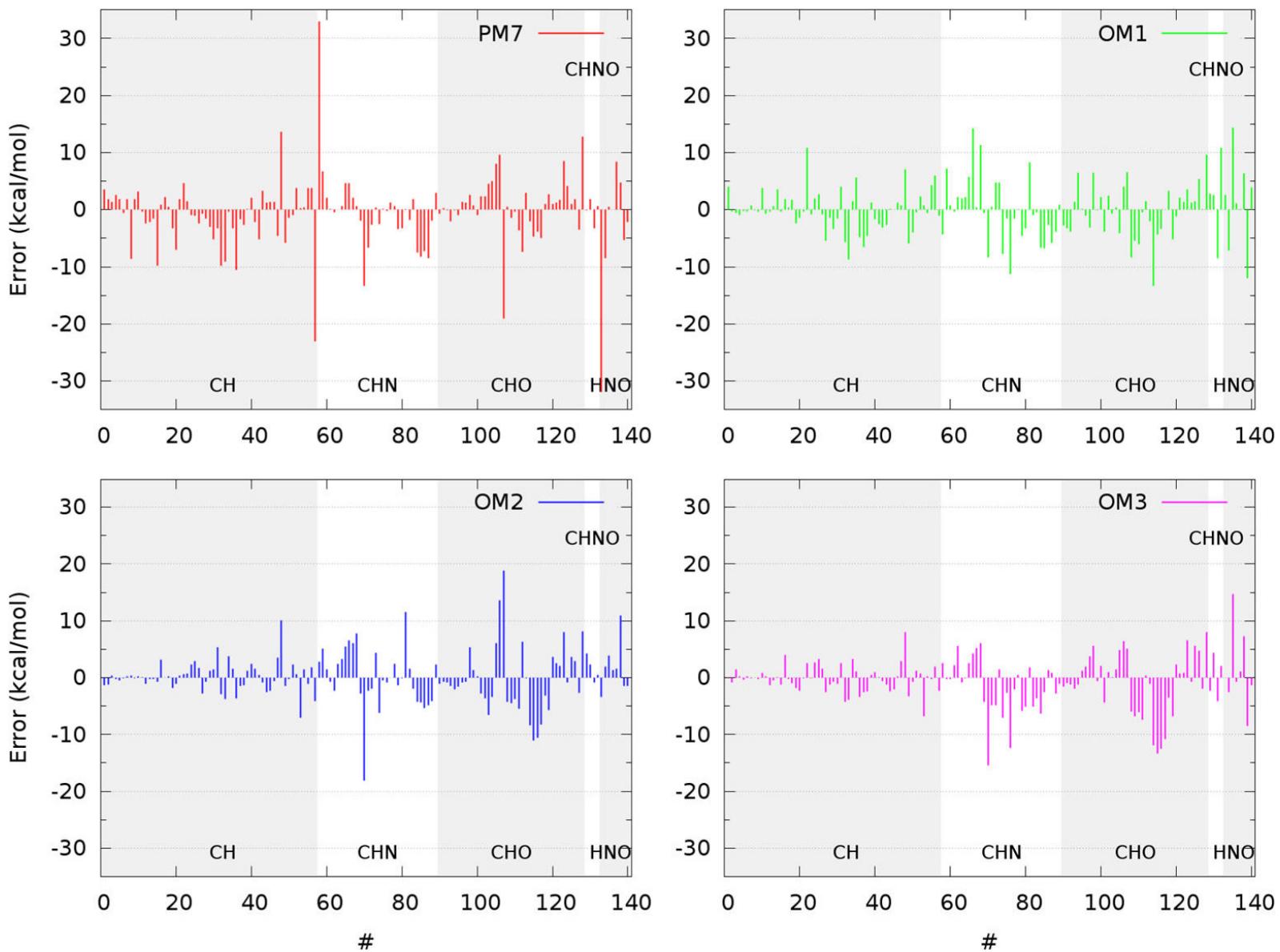
Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr, Germany

Supporting Information

ABSTRACT: Semiempirical orthogonalization-corrected methods (OM1, OM2, and OM3) go beyond the standard MNDO model by explicitly including additional interactions into the Fock matrix in an approximate manner (Pauli repulsion, penetration effects, and core–valence interactions), which yields systematic improvements both for ground-state and excited-state properties. In this Article, we describe the underlying theoretical formalism of the OM_x methods and their implementation in full detail, and we report all relevant OM_x parameters for hydrogen, carbon, nitrogen, oxygen, and fluorine. For a standard set of mostly organic molecules commonly used in semiempirical method development, the OM_x results are found to be superior to those from standard MNDO-type methods. Parametrized Grimme-type dispersion corrections can be added to OM2 and OM3 energies to provide a realistic treatment of noncovalent interaction energies, as demonstrated for the complexes in the S22 and S66×8 test sets.



OMx Methods



Semiempirical Quantum Mechanical Methods for Noncovalent Interactions for Chemical and Biochemical Applications

Anders S. Christensen,[†] Tomáš Kubář,[‡] Qiang Cui,^{*,†} and Marcus Elstner^{*,§}

[†]Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin—Madison, 1101 University Avenue, Madison, Wisconsin 53706, United States

[‡]Institute of Physical Chemistry & Center for Functional Nanostructures and [§]Institute of Physical Chemistry, Karlsruhe Institute of Technology, Kaiserstrasse 12, 76131 Karlsruhe, Germany

ABSTRACT: Semiempirical (SE) methods can be derived from either Hartree–Fock or density functional theory by applying systematic approximations, leading to efficient computational schemes that are several orders of magnitude faster than ab initio calculations. Such numerical efficiency, in combination with modern computational facilities and linear scaling algorithms, allows application of SE methods to very large molecular systems with extensive conformational sampling. To reliably model the structure, dynamics, and reactivity of biological and other soft matter systems, however, good accuracy for the description of noncovalent interactions is required. In this review, we analyze popular SE approaches in terms of their ability to model noncovalent interactions, especially in the context of describing biomolecules, water solution, and organic materials. We discuss the most significant errors and proposed correction schemes, and we review their performance using standard test sets of molecular systems for quantum chemical methods and several recent applications. The general goal is to highlight both the value and limitations of SE methods and stimulate further developments that allow them to effectively complement ab initio methods in the analysis of complex molecular systems.



Semiempirical quantum chemistry methods

- Positive points:
 - For most methods the parameters are atomic. Once an atom is parameterized, in principle, it can be used in any environment.
 - All chemical and physical phenomena can be modeled: chemical reactions, polarizabilities, spectra, optical properties, electronic properties, etc.
- Limitations:
 - Accuracy is insufficient.
 - Parameterization procedure is difficult
 - Most semiempirical methods do not have parameters for all the atoms in the periodic table.

**A break for drinking
water or coffee:**

5 min.

High Performance Computing for Semiempirical Methods

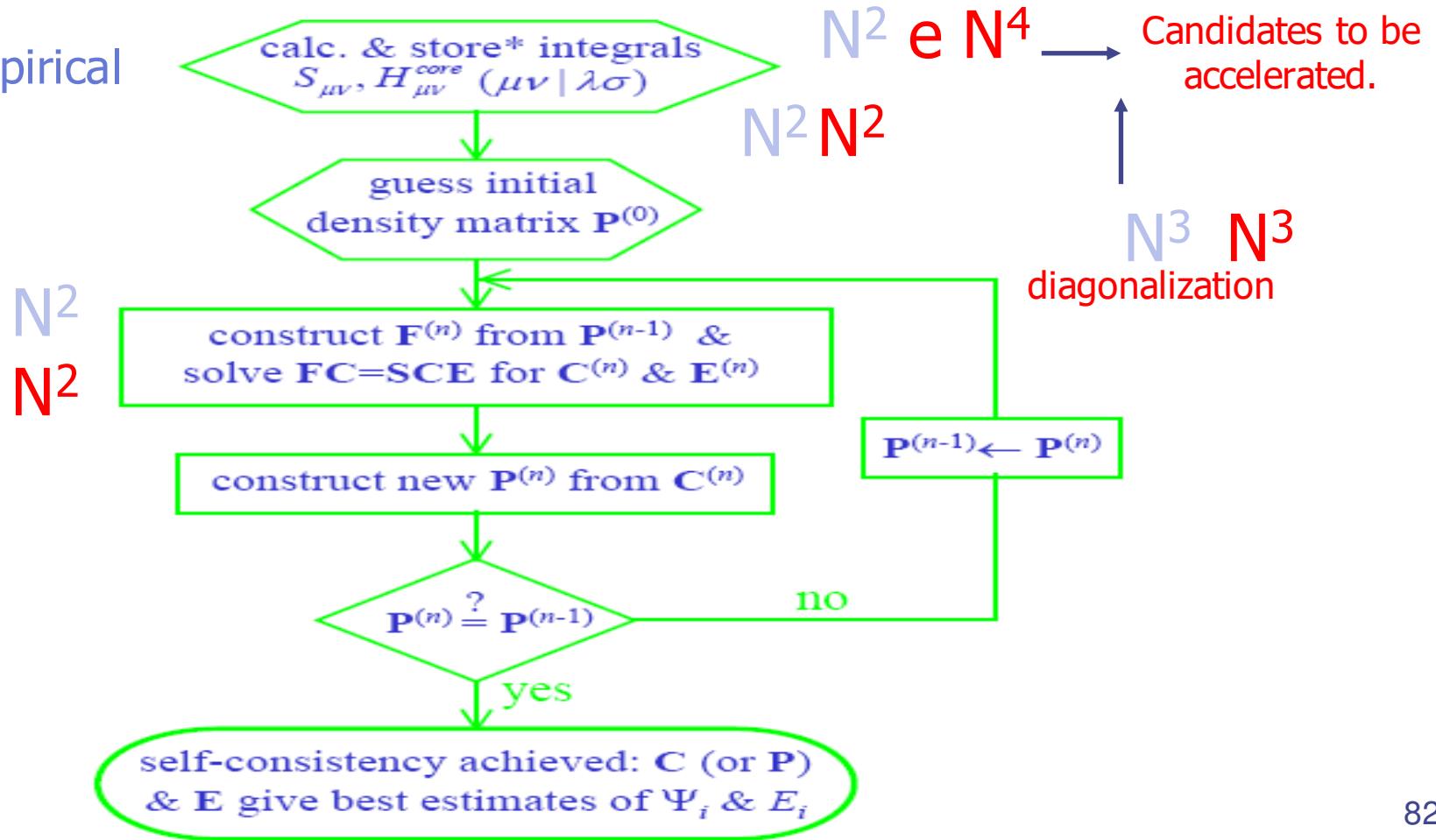
With emphasis on the study of biomolecular systems

Where are the bottlenecks?

Complexities for some parts of conventional single point energy calculation (SCF algorithm)

Ab initio

Semiempirical



Where are the bottlenecks?

Complexities

- Diagonalization (N^3)
- Matrix-matrix multiplication (N^3)
- Unitary transformations (N^3)
- Molecular surface calculations
- Finding the critical points of functions (N^3)
- Atomic Integrals calculations (N^4)
- CI calculations (N^7)

Many these procedures are called many times within iterative algorithms

A solution for these tasks:

High Performance Computing

High performance computing in chemistry

Important themes

- Parallelization of codes
- Efficient Algorithms
- Programming languages
- High performance architectures
- Benchmarks
- Numerical libraries
- Communication libraries



Clusters, GRIDs, Supercomputers, GPUs, etc.

PERSPECTIVE

www.rsc.org/pccp | Physical Chemistry Chemical Physics

Utilizing high performance computing for chemistry: parallel computational chemistry

Wibe A. de Jong,^{*a} Eric Bylaska,^a Niranjan Govind,^a Curtis L. Janssen,^c
Karol Kowalski,^a Thomas Müller,^d Ida M. B. Nielsen,^c Hubertus J. J. van Dam,^a
Valera Veryazov^b and Roland Lindh^{*e}

Received 10th February 2010, Accepted 4th May 2010

First published as an Advance Article on the web 8th June 2010

DOI: 10.1039/c002859b

Parallel hardware has become readily available to the computational chemistry research community. This perspective will review the current state of parallel computational chemistry software utilizing high-performance parallel computing platforms. Hardware and software trends and their effect on quantum chemistry methodologies, algorithms, and software development will also be discussed.

High performance computing in chemistry

NOTA TÉCNICA

PARALELIZAÇÃO EM QUÍMICA

Nelson Henrique Morgan

Instituto de Química - Universidade Estadual de Campinas - CP 6154 - 13083-970 - Campinas - SP

Recebido em 17/11/94; aceito em 9/2/95

Parallel computation applied to theoretical chemistry has increased very much in the last 15 years. In this work is presented results employing parallel algorithms in the calculation of electronic repulsion integrals. Although the improvement of the performance is almost obvious, it was observed that the random expansion of the number of CPUs does not necessarily provides the best performance of the calculations. General features about parallel computation and other *ab initio* calculations are presented.

Keywords: parallel algorithms; tools to parallelization; *ab initio* calculations.

QUÍMICA NOVA, 18(5) (1995)

High performance computing in chemistry

John von Neumann Institute for Computing



Tools for parallel quantum chemistry software

Thomas Steinke

published in

Modern Methods and Algorithms of Quantum Chemistry,
J. Grotendorst (Ed.), John von Neumann Institute for Computing,
Jülich, NIC Series, Vol. 1, ISBN 3-00-005618-1, pp. 49-67, 2000.

© 2000 by John von Neumann Institute for Computing
Permission to make digital or hard copies of portions of this work for
personal or classroom use is granted provided that the copies are not
made or distributed for profit or commercial advantage and that copies
bear this notice and the full citation on the first page. To copy otherwise
requires prior specific permission by the publisher mentioned above.

<http://www.fz-juelich.de/nic-series/>

High performance computing in chemistry

ORAL COMPREHENSIVE EXAM POSITION PAPER, DECEMBER 2015

1

HPC in Computational Chemistry: Bridging Quantum Mechanics, Molecular Dynamics, and Coarse-Grained Models

David Ozog
University of Oregon
ozog@cs.uoregon.edu

Abstract—The past several decades have witnessed tremendous strides in the capabilities of computational chemistry simulations, driven in large part by the extensive parallelism offered by powerful computer clusters and scalable parallel programming methods. However, cluster computing has also seen flattening processor clock frequencies, unsustainable increases in power requirements, and more complicated software that accommodates for the vast diversity of modern heterogeneous computer systems. As scientific methods for modeling and simulating atoms and molecules continue to evolve, software developers struggle to keep up. This position paper describes the primary challenges that face the computational chemistry community, along with recent solutions and techniques that circumvent these difficulties. In particular, I describe in detail the 3 primary models used to simulate atoms and molecules: quantum mechanics (QM), molecular mechanics (MM), and coarse-grained (CG) models. The research literature is rife with examples that utilize high performance computing (HPC) to scale these models to large and relevant chemical problems. However, the grand challenge lies in effectively bridging these scales, both spatially and temporally, to study richer chemical models that go beyond single-scale physics. This paper describes the state of the art in multiscale computational chemistry, with an eye toward improving developer productivity with upcoming exascale architectures, in which we require productive software environments, enhanced support for coupled scientific workflows, adaptive and introspective runtime systems, resilience to hardware failures, and extreme scalability.



High performance computing in chemistry

For a parallel computing environment, it is necessary to have a parallel computer or computers in parallel, a parallel operating system and a parallel programming language.

For a parallel system, these are the possibilities:

- Computers working in parallel;
- Each such a computer with multiple processors;
- In each processor, the execution units executing instructions in parallel.

High performance computing in chemistry

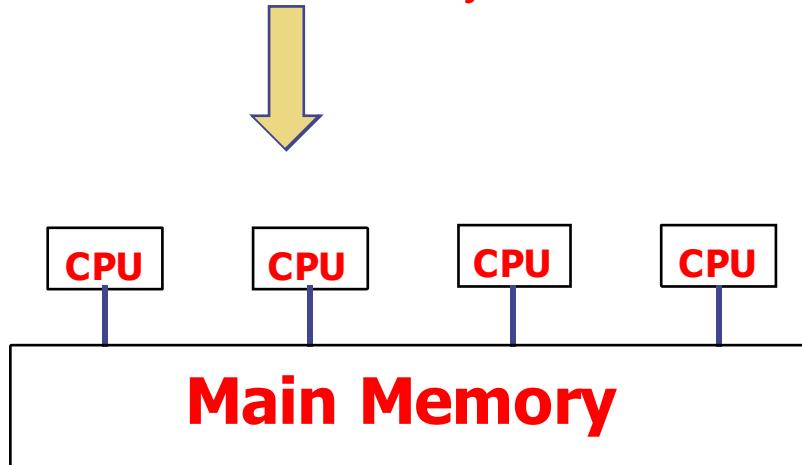
We can categorize the systems in parallel by the way they handle instructions and data.

- SISD (single instruction, single data): these are conventional serial computers on which a single instruction is executed and operates on a single data thread before the next instruction and the next data thread are found.
- SIMD (single instruction, multiple data): in this case, instructions are processed from a single thread, but they operate concurrently on multiple data blocks.
- MIMD (multiple instruction, multiple data): for this each processor works independently of the others with independent instructions and data. These are the computers (or supercomputers) that employ message passing (communication method used in parallel processing).

High performance computing in chemistry

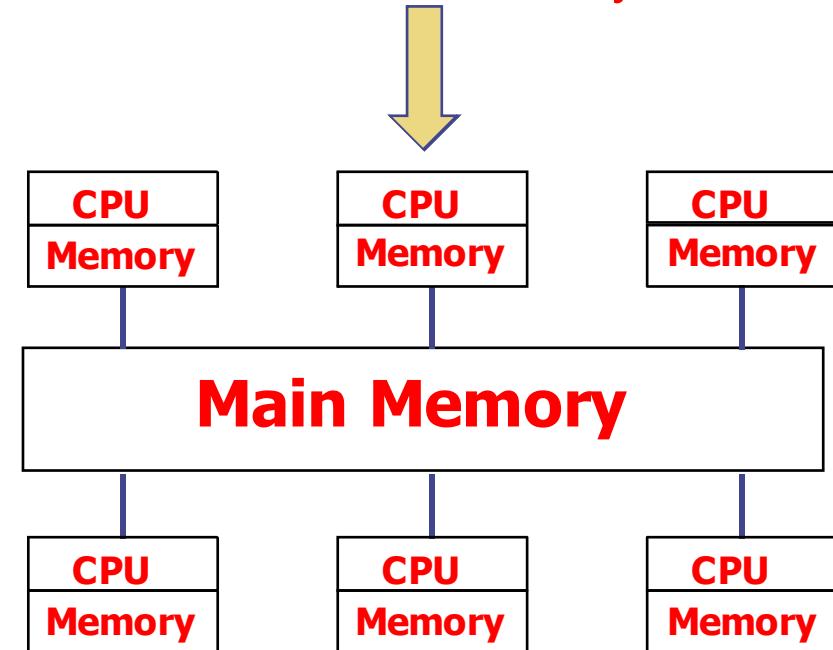
Regarding the use of memory, there are two types of parallelism:

Shared memory



PCs with some CPUs

Distributed memory.



Clusters and supercomputers

Graphical Processing Units



"A serious competitor for the multi-core CPU is represented by **graphical processing units (GPUs)**, which are graphic cards used for scientific computing. There are four basic things about GPUs. They are fast and will get a lot faster. They are cheap, measured on a performance-per-dollar basis. They use less power than CPUs when compared on a performance-per-watt basis. But the fourth thing is their limitations."

Accelerating computational chemistry using GPGPU

NOVEL
ARCHITECTURES

Graphical Processing Units for Quantum Chemistry

The authors provide a brief overview of electronic structure theory and detail their experiences implementing quantum chemistry methods on a graphical processing unit. They also analyze algorithm performance in terms of floating-point operations and memory bandwidth, and assess the adequacy of single-precision accuracy for quantum chemistry applications.

1521-9615/08/\$25.00 © 2008 IEEE
COPIED BY THE IEEE CS AND THE ACP

IVAN S. UFMITSYEV AND TODD J. MARTINEZ
University of Illinois at Urbana-Champaign

26 THIS ARTICLE HAS BEEN PEER-REVIEWED.

COMPUTING IN SCIENCE & ENGINEERING

Accelerating Molecular Dynamic Simulation on Graphics Processing Units

MARK S. FRIEDRICHSS,¹ PETER EASTMAN,¹ VISHAL VAIDYANATHAN,² MIKE HUSTON,³
SCOTT LEGRAND,⁴ ADAM L. BEBERG,⁵ DANIEL L. ENSIGN,² CHRISTOPHER M. BRUNS,¹ VIJAY S. PANDE^{2,5}

¹Department of Bioengineering, Stanford University, Stanford, California 94305

²Department of Chemistry, Stanford University, Stanford, California 94305

³Advanced Micro Devices, Sunnyvale, California

⁴NVIDIA Corporation, Santa Clara, California

⁵Department of Computer Science, Stanford University, Stanford, California 94305

Received 19 September 2008; Revised 12 December 2008; Accepted 13 December 2008

DOI 10.1002/jcc.21209

Published online in Wiley InterScience (www.interscience.wiley.com).

Abstract: We describe a complete implementation of all-atom protein molecular dynamics running entirely on a graphics processing unit (GPU), including all standard force field terms, integration, constraints, and implicit solvent. We discuss the design of our algorithms and important optimizations needed to fully take advantage of a GPU. We evaluate its performance, and show that it can be more than 700 times faster than a conventional implementation running on a single CPU core.

© 2009 Wiley Periodicals, Inc. *J Comput Chem* 00: 000–000, 2009

Today, we have ...

JCTC Journal of Chemical Theory and Computation

Subscriber access provided by UNIV FED DA PARAIBA UFPB

Article

Accelerating Density Functional Calculations with Graphics Processing Unit

Koji Yasuda

J. Chem. Theory Comput., 2008, 4 (8), 1230–1236 • DOI: 10.1021/ct8001046 • Publication Date (Web): 04 July 2008

Downloaded from <http://pubs.acs.org> on March 4, 2009

Ab Initio Quantum Chemistry for Protein Structures

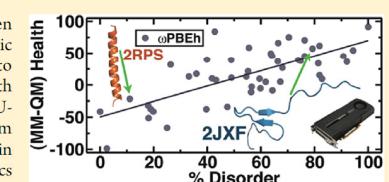
Heather J. Kulik,^{†,‡} Nathan Luehr,^{†,‡} Ivan S. Ufimtsev,^{†,‡} and Todd J. Martinez^{*,†,‡}

[†]Department of Chemistry and PULSE Institute, Stanford University, Stanford, California, 94305, United States

[‡]SLAC National Accelerator Laboratory, Menlo Park, California 94025, United States

 Supporting Information

ABSTRACT: Structural properties of over 55 small proteins have been determined using both density-based and wave-function-based electronic structure methods in order to assess the ability of ab initio “force fields” to retain the properties described by experimental structures measured with crystallography or nuclear magnetic resonance. The efficiency of the GPU-based quantum chemistry algorithms implemented in our TeraChem program enables us to carry out systematic optimization of ab initio protein structures, which we compare against experimental and molecular mechanics force field references. We show that the quality of the ab initio optimized structures, as judged by conventional protein health metrics, increases with increasing basis set size. On the other hand, there is little evidence for a significant improvement of predicted structures using density functional theory as compared to Hartree–Fock methods. Although occasional pathologies of minimal basis sets are observed, these are easily alleviated with even the smallest double- ζ basis sets.



and many others ...

GPGPU in the top500.org

TOP500 LIST - NOVEMBER 2020

R_{max} and R_{peak} values are in TFlops. For more details about other fields, check the [TOP500 description](#).

R_{peak} values are calculated using the advertised clock rate of the CPU. For the efficiency of the systems you should take into account the Turbo CPU clock rate where it applies.

← 1-100 101-200 201-300 301-400 401-500 →

Rank	System	Cores	Rmax [TFlop/s]	Rpeak [TFlop/s]	Power (kW)
1	Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442,010.0	537,212.0	29,899
2	Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148,600.0	200,794.9	10,096
3	Sierra - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
4	Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371

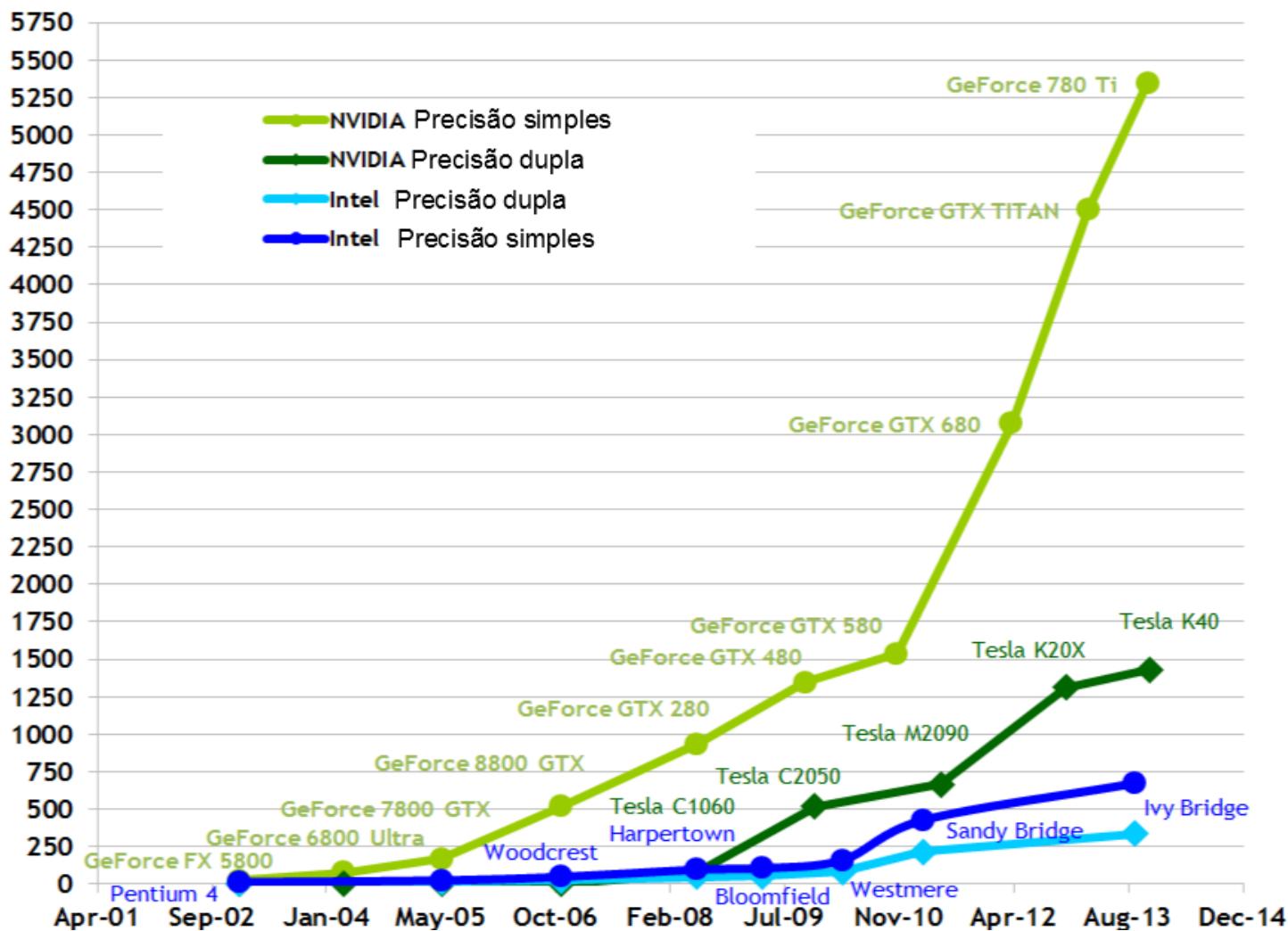
SUPERCOMPUTER FUGAKU - SUPERCOMPUTER FUGAKU, A64FX 48C 2.2GHZ, TOFU INTERCONNECT D

Site:	RIKEN Center for Computational Science
System URL:	https://www.r-ccs.riken.jp/en/fugaku/project
Manufacturer:	Fujitsu
Cores:	7,630,848
Memory:	5,087,232 GB
Processor:	A64FX 48C 2.2GHz
Interconnect:	Tofu interconnect D
Performance	
Linpack Performance (Rmax)	442,010 TFlop/s
Theoretical Peak (Rpeak)	537,212 TFlop/s
Nmax	21,288,960
HPCG [TFlop/s]	16,004.5

Linpack Performance ~ 442 petaflop/s
(quadrillion of calculations per second)

CPUs and GPUs Evolution

GFLOP/s teórico



GPUs Evolution

NVIDIA GeForce GTX 580

0.196 GFLOPS



2011

38x



NVIDIA TITAN V

7.450 TFLOPS



2017

http://www.nvidia.com/object/computational_chemistry.html

NAVIGATION: DOWNLOAD DRIVERS COOL STUFF SHOP PRODUCTS TECHNOLOGIES COMMUNITIES SUPPORT

TESLA

NVIDIA Home > Products > High Performance Computing > Vertical Industry Solutions Share this page

GPU COMPUTING SOLUTIONS

- Main
- What is GPU Computing?
- Why Choose Tesla
- Industry Software Solutions
- Tesla Workstation Solutions
- Tesla Data Center Solutions
- Tesla Bio Workbench
- Where to Buy
- Contact US
- Sign up for Tesla Alerts
- Fermi GPU Computing Architecture

SOFTWARE AND HARDWARE INFO

- Tesla Product Literature
- Tesla Software Features
- Software Development Tools
- CUDA Training and Consulting Services
- GPU Cloud Computing Service Providers

NEWS AND SUCCESS STORIES

Computational Chemistry

There are several ongoing projects on accelerating quantum chemistry codes using CUDA-enabled GPUs, including work on Gaussian and GAMESS. The charts below are representative results, followed by links to software and technical reports on CUDA acceleration of computational chemistry.

With the introduction of NVIDIA Tesla Bio Workbench, it provides bio-physicists and computational chemists the tools to push the boundaries of bio-chemical research, optimizing the scientific workflow and accelerating the pace of research. [Learn more](#).

Direct self-consistent field calculations

Molecule	1 Tesla 8-series GPU (sec)	GAMESS on Intel Pentium D (sec)
Valinomycin	8.1	12.5 min
Buckyball	5.7	5.5 min
Taxol	4.5	4.7 min
Cholesterol	1.2	1.1 min
Caffeine	0.2	4.4 sec

Time (log-scale)

Legend: █ 1 Tesla 8-series GPU █ GAMESS on Intel Pentium D (3.0 GHz)

Coulomb Potential Evaluation

System	1 Tesla 8-series GPU (sec)	Gaussian 03 on Intel Pentium (sec)
Valinomycin/PW91/6-31G	64.5	9.9 min
Valinomycin/LSDA/3-21G	36.1	4 min
Taxol/PW91/6-31G	32	8 min
Taxol/LSDA/3-21G	21.6	2.8 min

Time (sec)

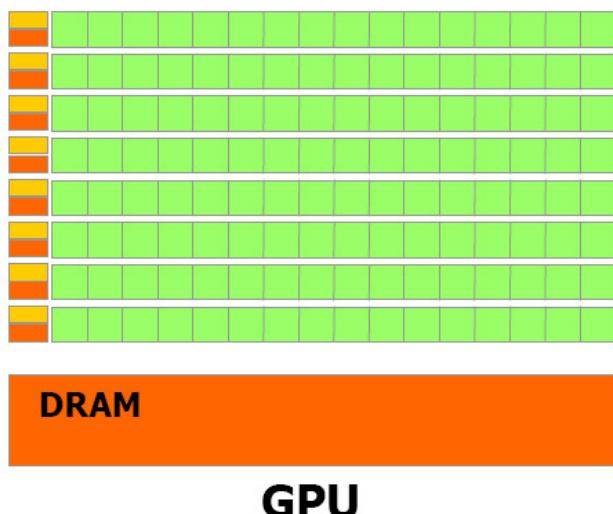
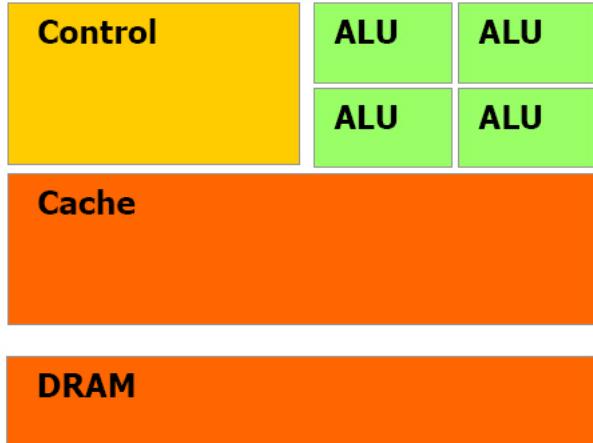
Legend: █ CUDA code on 1 Tesla 8-series GPU █ Gaussian 03 on Intel Pentium (2.4 GHz)

[Direct self-consistent field \(SCF\) calculations](#) Ufimtsev and Martinez

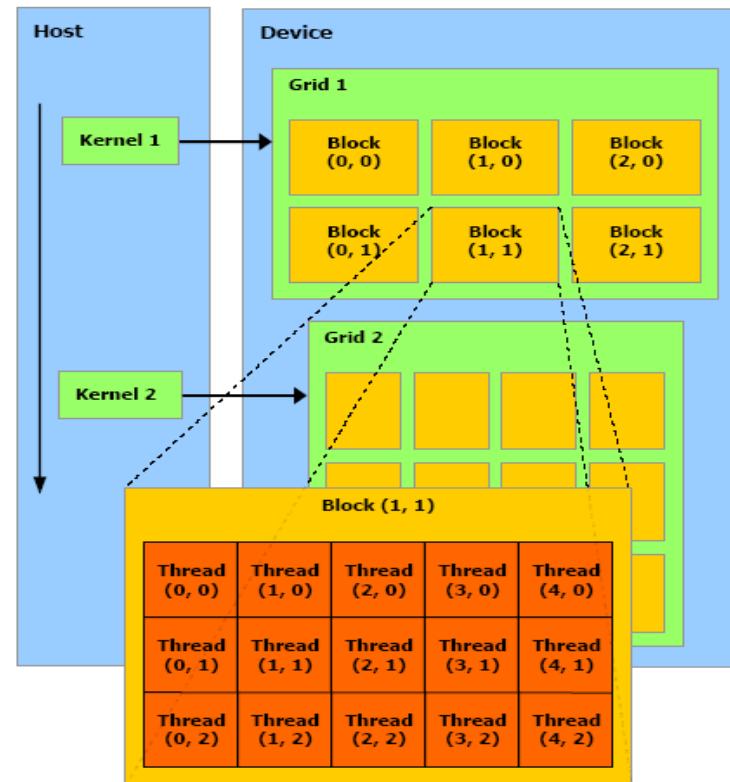
[Two-Electron integral evaluation](#) Koji Yasuda

CPU x GPU

GPUs have many transistors dedicated to data processing

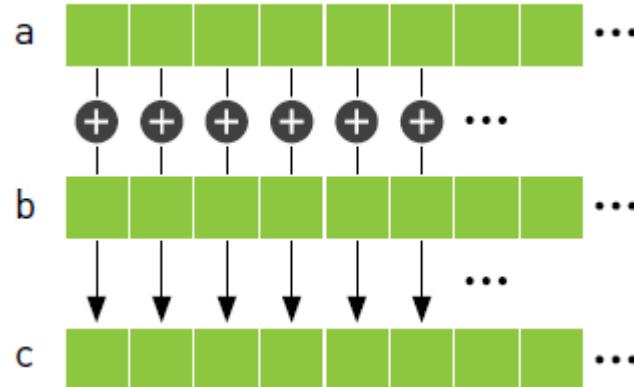


High parallelism



The host issues a succession of kernel invocations to the device. Each kernel is executed as a batch of threads organized as a grid of thread blocks

C++ code for GPUs



```
void add( int *a, int *b, int *c ) {  
    for (i=0; i < N; i++) {  
        c[i] = a[i] + b[i];  
    }  
  
__global__ void add( int *a, int *b, int *c ) {  
    int tid = blockIdx.x;      // handle the data at this index  
    if (tid < N)  
        c[tid] = a[tid] + b[tid];  
}
```

C++ code for summing two vectors in CPU

CUDA code for summing two vectors in GPU

CUDA by Example

JASON SANDERS

EDWARD KANDROT

<http://openmopac.net>

For the semiempirical calculations there is one traditional code, MOPAC

MOPAC®

[What MOPAC is](#)

MOPAC (Molecular Orbital PACkage) is a semiempirical quantum chemistry program based on Dewar and Thiel's NDDO approximation. Most users use MOPAC with a [Graphical User Interface](#).

[MOPAC2012](#)

[MOPAC2012](#) is MOPAC2009 plus the [PM7](#) and [PM7-TS](#) methods. If a bug is detected, please send a message by [E-mail](#) to MrMOPAC@OpenMOPAC.net, along with an example illustrating the bug. If you qualify for Academic not-for-profit use and already have a password for MOPAC 2009, go straight to [download](#), otherwise request a [password and download](#). For commercial and governmental prices, see [Prices](#).

Commercial users in Europe and America, please contact [CAChe Research](#) (see [Resellers](#))

Commercial users in Japan, please contact [Ryoka Systems](#)

Commercial users in Korea, please contact [KREIS I&C](#)

[PM7](#)

PM7 is a modified form of PM6. A few errors in NDDO theory that affect large systems have been removed. All atomic and diatomic parameters were re-optimized. Average errors in organic compounds have been reduced by ~10%, and errors in large organics and [solids](#) have been significantly reduced, see [PM7 Accuracy](#).

Our objective was to accelerate semiempirical calculations in MOPAC by using GPUs.

MOPAC calculation using CPU-GPU hybrid computational systems

In this paper we detail our strategies to get this:

GPU linear algebra libraries and GPGPU programming for accelerating MOPAC semiempirical quantum chemistry calculations

Autores Julio Daniel Carvalho Maia, Gabriel Aires Urquiza Carvalho, Carlos Peixoto Mangueira Jr, Sidney Ramos Santana, Lucidio Anjos Formiga Cabral, Gerd B Rocha

Data de 2012/9/11

publicação

Publicações Journal of chemical theory and computation

Volume 8

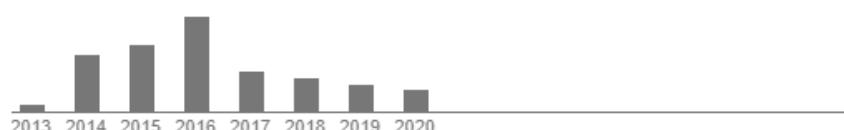
Edição 9

Páginas 3072-3081

Editora American Chemical Society

Descrição In this study, we present some modifications in the semiempirical quantum chemistry MOPAC2009 code that accelerate single-point energy calculations (1SCF) of medium-size (up to 2500 atoms) molecular systems using GPU coprocessors and multithreaded shared-memory CPUs. Our modifications consisted of using a combination of highly optimized linear algebra libraries for both CPU (LAPACK and BLAS from Intel MKL) and GPU (MAGMA and CUBLAS) to hasten time-consuming parts of MOPAC such as the pseudodiagonalization, full diagonalization, and density matrix assembling. We have shown that it is possible to obtain large speedups just by using CPU serial linear algebra libraries in the MOPAC code. As a special case, we show a speedup of up to 14 times for a methanol simulation box containing 2400 atoms and 4800 basis functions, with even greater gains in performance when using ...

Total de citações Citado por 266



J. Chem. Theory Comput., 2012, 8 (9), pp 3072–3081 DOI: 10.1021/ct3004645

Pseudo-diagonalization method (PD) in SCF

PD is based in:

"To have a self-consistent density matrix is sufficient to have annihilated all Fock matrix elements connecting the occupied and virtual molecular orbitals". (Stewart, J.J.P., JCC, 3, 2 227, 1982).

Equations for PD:

$$\mathcal{F}_{o-v} = C_o^T F C_v$$

$$C_i^{new} = cC_i - sC_\alpha \quad \text{and} \quad C_\alpha^{new} = sC_i + cC_\alpha$$

$$s = \frac{\mathcal{F}_{i\alpha}}{\varepsilon_i - \varepsilon_\alpha} \quad \text{and} \quad c = \sqrt{1 - s^2}$$

Our GPU implementation:

**CUBLAS Library
DGEMM**

**Plane rotation
of two vectors**

Equation for density matrix:

$$P = C_o^T C_o$$

Our GPU implementation:

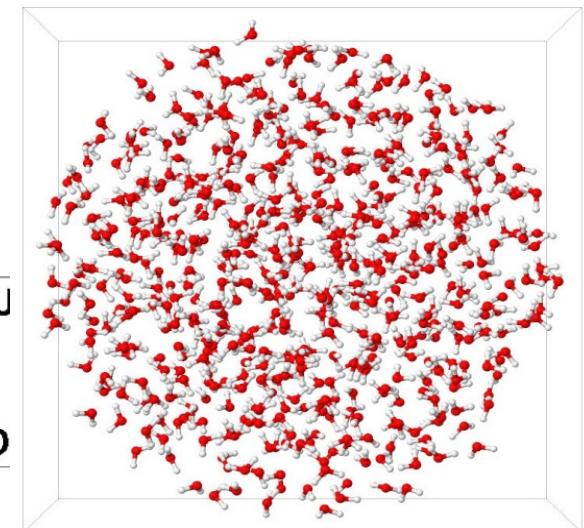
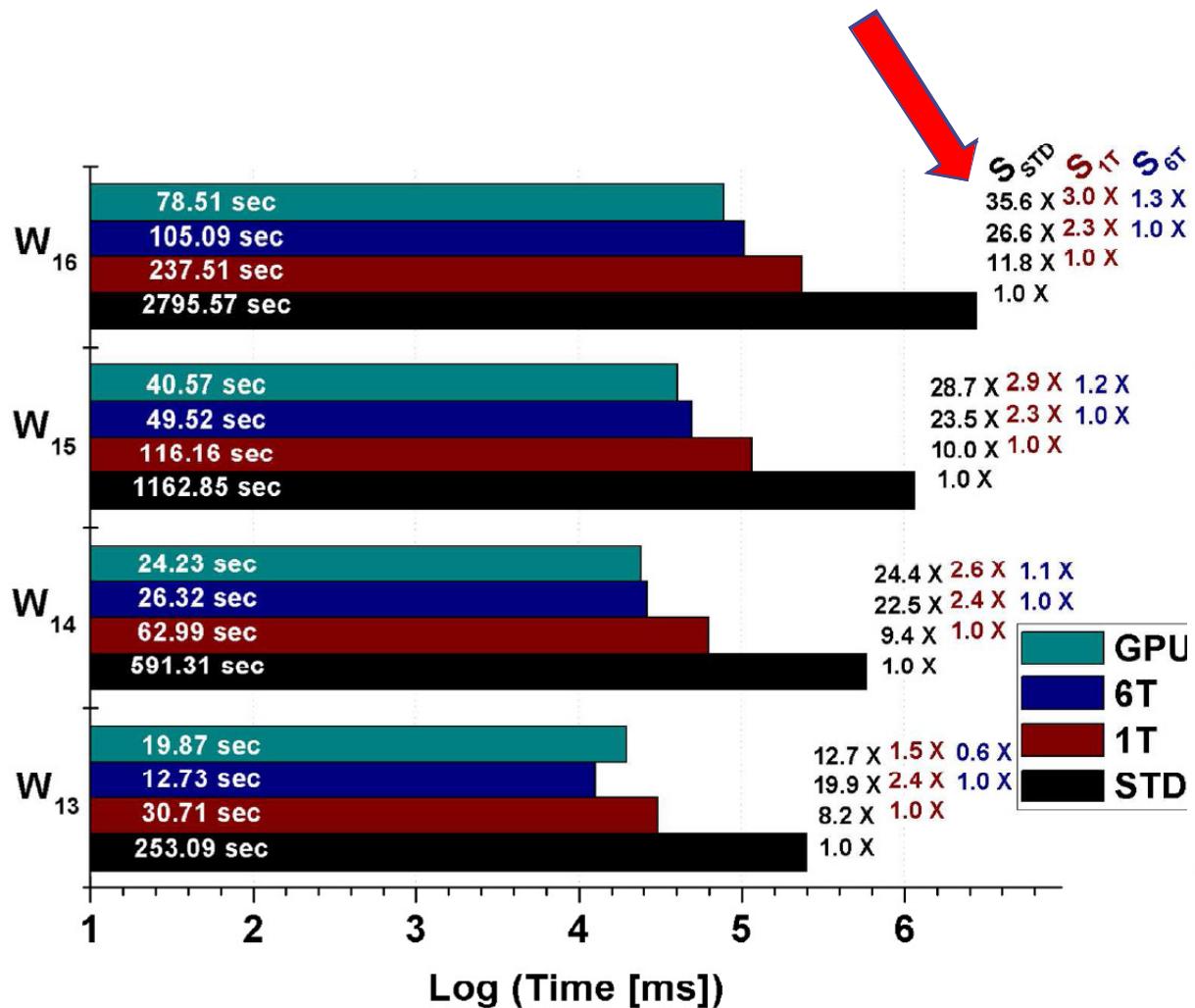
**CUBLAS Library
DSYRK**

Full diagonalization

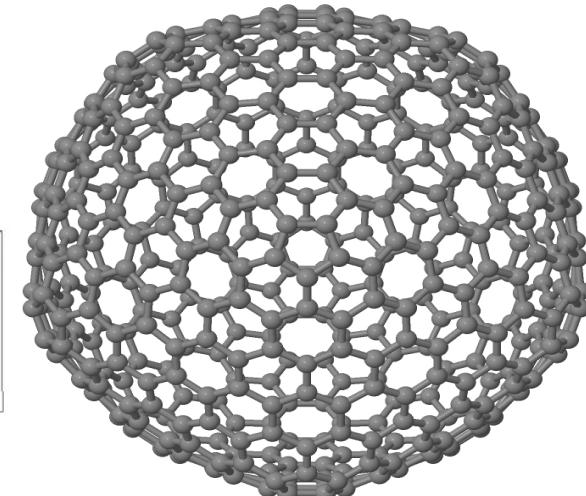
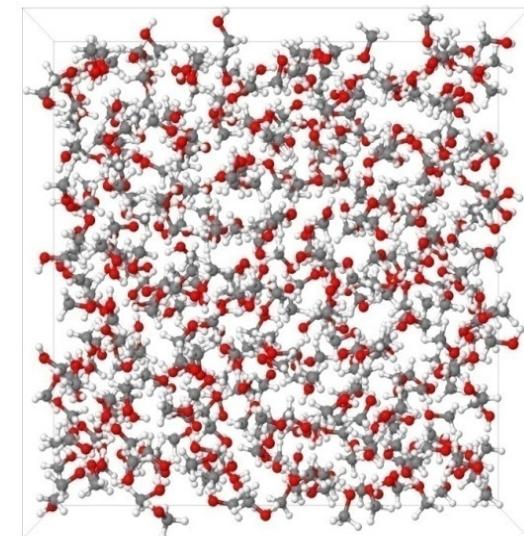
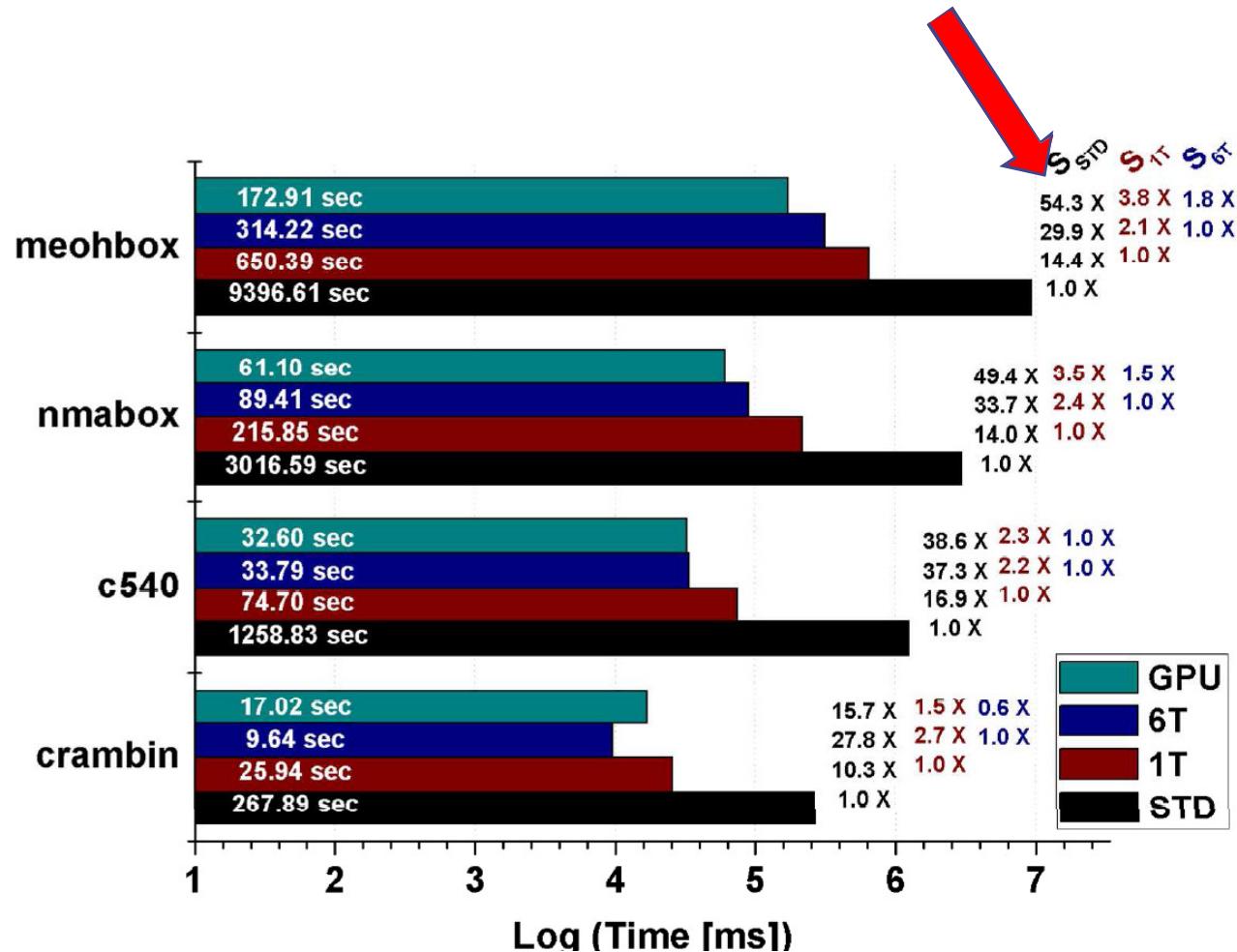
Our GPU implementation:

**MAGMA Library
and Intel MKL**

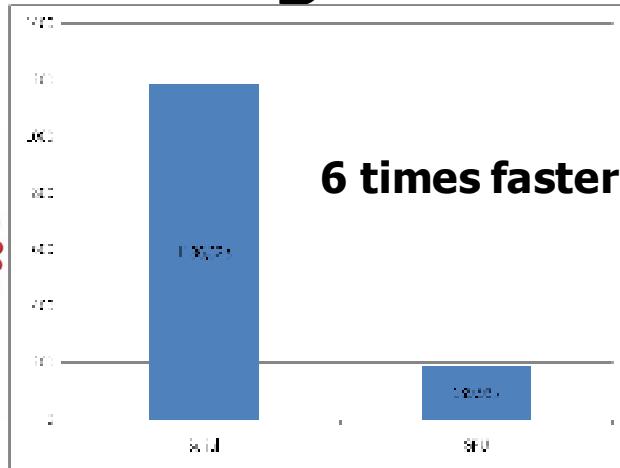
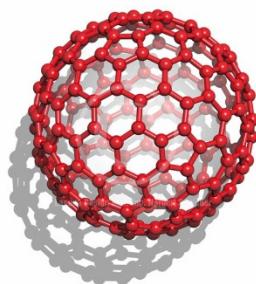
Speedups for Water clusters



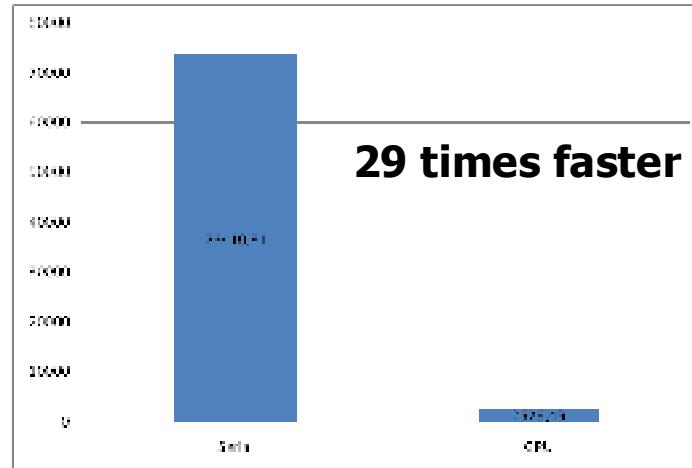
Speedups for simulation boxes, fullerenes and proteins



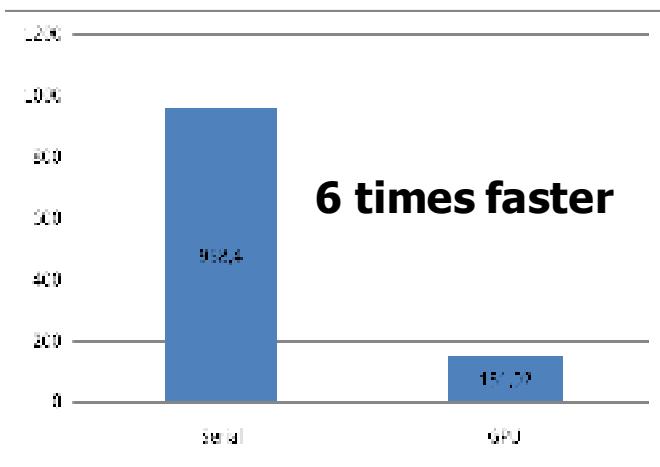
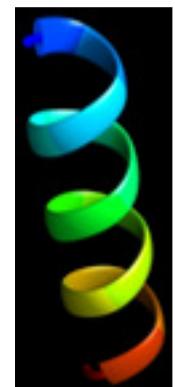
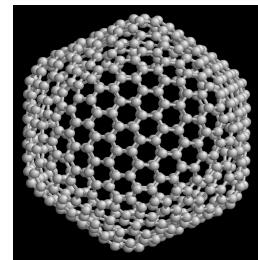
Geometry Optimization of large molecules on GPUs



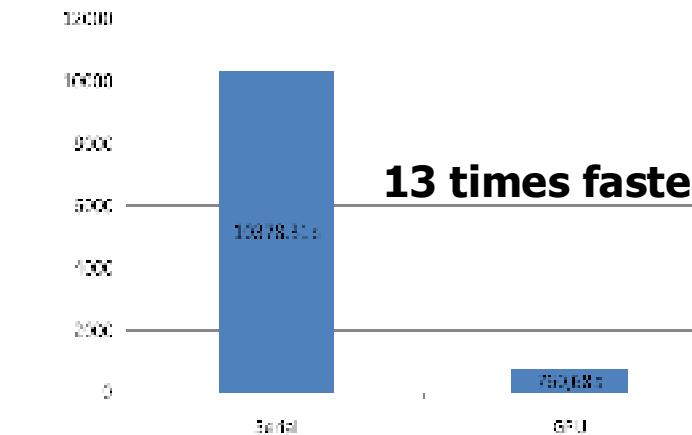
Times (s) for geometry optimization (only 100 cycles) of C₁₈₀.



Times (s) for geometry optimization (only 100 cycles) of C₅₄₀.



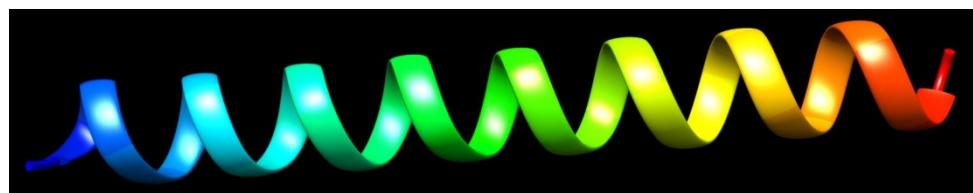
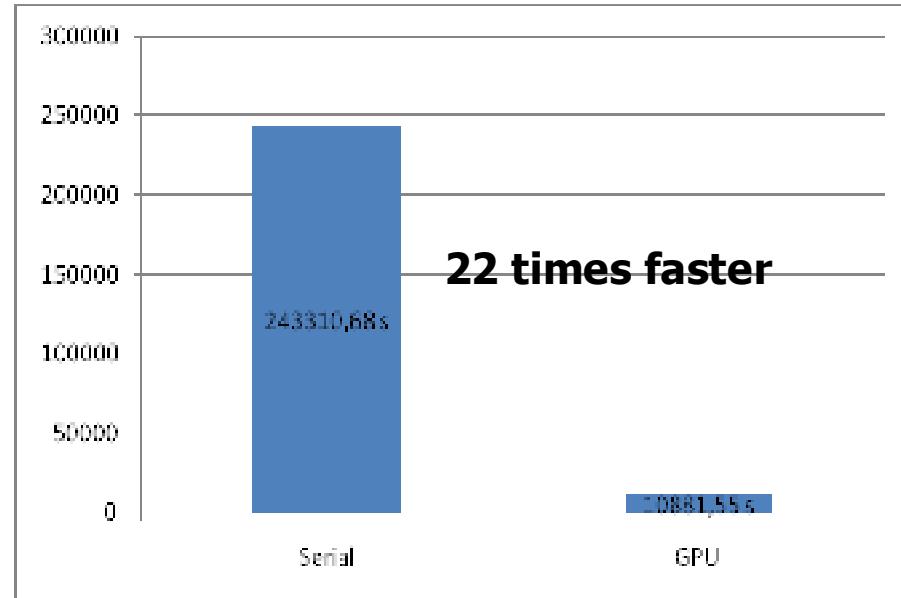
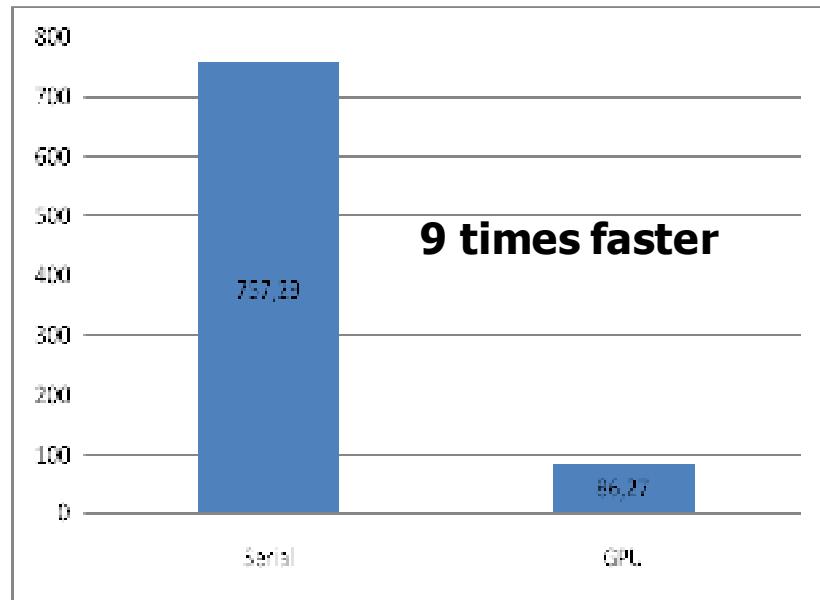
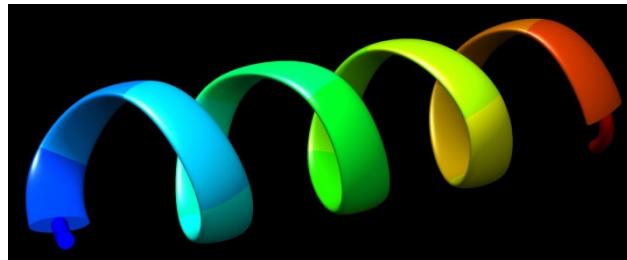
Times (s) for geometry optimization (only 100 cycles) of (Glu-Ala)₈.



Times (s) for geometry optimization (only 100 cycles) of (Glu-Ala)₁₆.



Vibrational frequencies of large molecules on GPUs



<http://openmopac.net>

MOPAC®

What MOPAC is

MOPAC (Molecular Orbital PACKage) is a semiempirical quantum chemistry program based on Dewar and Thiel's NDDO approximation. Most users use MOPAC with a [Graphical User Interface](#).

[MOPAC2012](#)

[MOPAC2012](#) is MOPAC2009 plus the [PM7](#) and [PM7-TS](#) methods. If a bug is detected, please send a message by [E-mail](#) to MrMOPAC@OpenMOPAC.net, along with an example illustrating the bug. If you qualify for Academic not-for-profit use and already have a password for MOPAC 2009, go straight to [download](#), otherwise request a [password](#) and [download](#). For commercial and governmental prices, see [Prices](#).

Commercial users in Europe and America, please contact [CAChe Research](#) (see [Resellers](#))

Commercial users in Japan, please contact [Ryoka Systems](#)

Commercial users in Korea, please contact [KREIS I&C](#)

[PM7](#)

PM7 is a modified form of PM6. A few errors in NDDO theory that affect large systems have been removed. All atomic and diatomic parameters were re-optimized. Average errors in organic compounds have been reduced by ~10%, and errors in large organics and [solids](#) have been significantly reduced, see [PM7 Accuracy](#).

Reducing Computation Time

References and Citations

Instructions and Manual

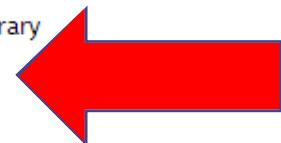
[Maintenance record](#)

[PM6](#)

PM6 is a re-parameterization of the NDDO method. Three modifications to the approximations were made, these mainly affect the way the core-core interaction was defined. Parameters were optimized for most elements, the exceptions being 12 of the lanthanides and all of the actinides. The lanthanides can be represented by sparkles. For details, see the on-line [PM6 journal article](#), and its [supplementary material](#).

Downloads

[Downloads](#) of source and binaries for the various older versions of MOPAC are available.



http://www.nvidia.com/object/computational_chemistry.html

Quantum Chemistry Applications



Application	Features Supported	GPU Perf	Release Status	Notes
MOPAC2012	Pseudodiagonalization, full diagonalization, and density matrix assembling	3.8-14X	Released MOPAC2013 available Q1 2014 Single GPU	Academic port. http://openmopac.net
NWChem	Triples part of Reg-CCSD(T), CCSD & EOMCCSD task schedulers	7-8X	Released, Version 6.3 Multiple GPUs	Development GPGPU benchmarks: www.nwchem-sw.org And http://www.olcf.ornl.gov/wp-content/training/electronic-structure-2012/Krishnamoorthy-ESCMA12.pdf
Octopus	Full GPU support for ground-state, real-time calculations; Kohn-Sham Hamiltonian, orthogonalization, subspace diagonalization, poisson solver, time propagation	1.5-8X	Released, Version 4.1.0	http://www.tddft.org/programs/octopus/

Structure			
Gaussian (In Development)	Predicts energies, molecular structures, and vibrational frequencies of molecular systems	Joint NVIDIA, PGI and Gaussian collaboration	Yes
GPAW	Real-space grid DFT code written in C and Python	Electrostatic poisson equation, orthonormalizing of vectors, residual minimization method (rmm-diis)	Yes
LATTE	Density matrix computations	CU_BLAS, SP2 Algorithm	Yes
MOLCAS	Methods for calculating general electronic structures in molecular systems in both ground and excited states	CU_BLAS	Single only Additional GPU support coming in Version 8
MOPAC2013	Semiempirical Quantum Chemistry	Pseudodiagonalization, full diagonalization, and density matrix assembling	Single only
NWChem	Calculations	Triples part of Reg-CCSD(T), CCSD and EOMCCSD task schedulers	Yes
Octopus	Used for ab initio virtual experimentation and quantum chemistry calculations	Full GPU support for ground-state, real-time calculations; Kohn-Sham Hamiltonian, orthogonalization, subspace diagonalization, poisson solver, time propagation	TBD
Q-CHEM	Computational chemistry package designed for HPC clusters	Various features including RI-MP2	TBD

Linear Scaling Algorithms

QC calculations Scaling

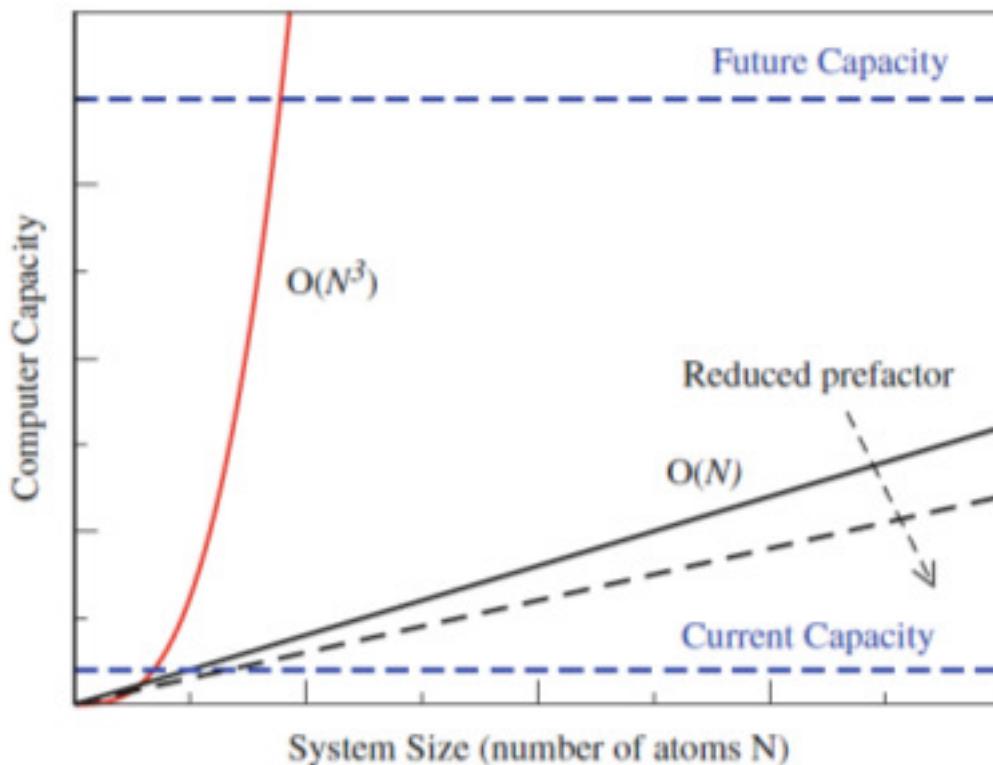
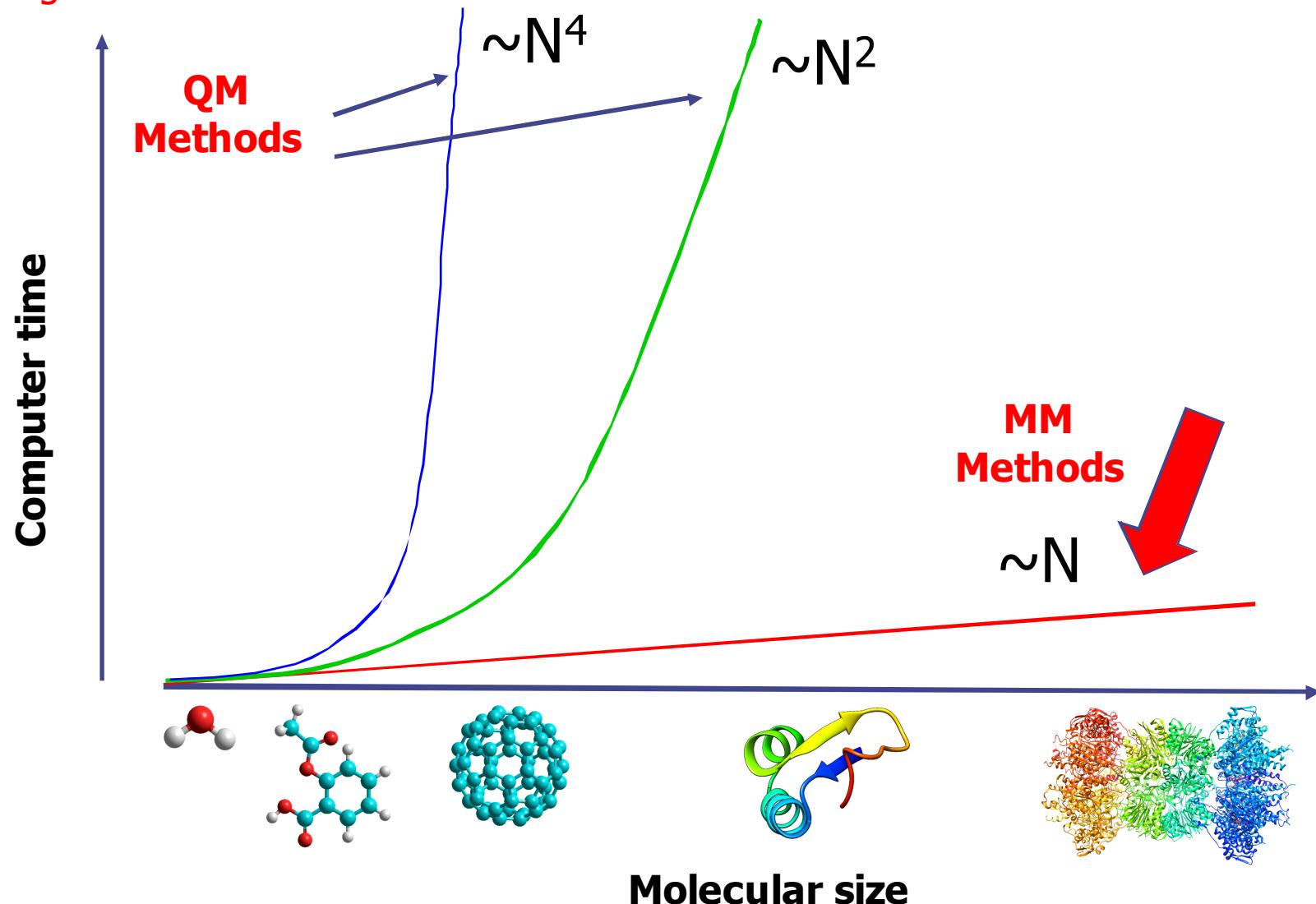


Figure 16-1. A computational forecast of reduced complexity $\mathcal{O}(N)$ algorithms for electronic structure calculations. Linear scaling electronic structure theory in large scale calculations will increase in importance as the computational capacity continues to improve in comparison to conventional $\mathcal{O}(N^3)$ methods

Complexity of algorithms in the molecular modeling theories

The computational complexity of the equations is a function of the number of particles, in general.



Quantum Chemical Calculation for Large Molecular Systems

Locality effects in quantum chemistry

Coulomb interaction dependence is r^{-1}

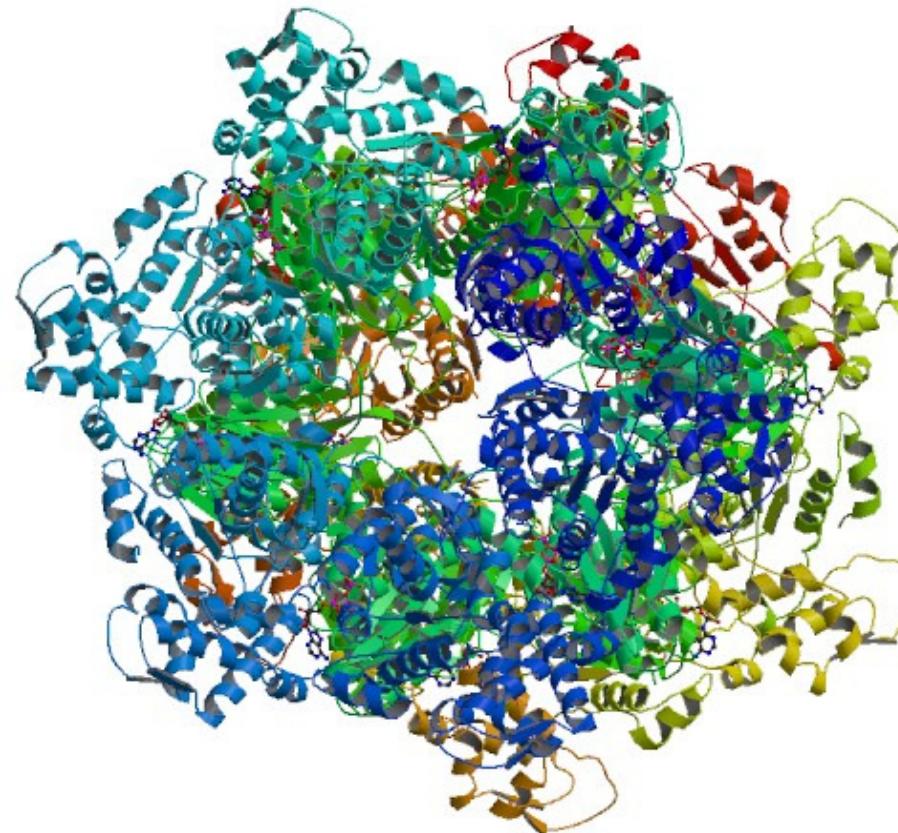
For the density matrix:

$$\rho(r,r') \sim \exp[-\sqrt{E_{\text{gap}}} |r-r'|]$$

- Usage of cutoffs
- Sparse matrices are generated (eg. Density and Fock)

Sparsity for the matrices in macromolecules

Crystal Structure of the HSLUV Protease-chaperone Complex,
91.510 atoms, PDB: 1G3I



Using a cutoff of 9Å, ~99% elements are zero and
only <1% are non-zero.

Sparse Matrices

$$A = \begin{pmatrix} 10 & 0 & 0 & 0 & -2 & 0 \\ 3 & 9 & 0 & 0 & 0 & 3 \\ 0 & 7 & 8 & 7 & 0 & 0 \\ 3 & 0 & 8 & 7 & 5 & 0 \\ 0 & 8 & 0 & 9 & 9 & 13 \\ 0 & 4 & 0 & 0 & 2 & -1 \end{pmatrix}$$

CSR format: Compressed Sparse

	val	10	-2	3	9	3	7	8	7	3 ... 9	13	4	2	-1
col_ind	0	4	0	1	5	1	2	3	0 ... 4	5	1	4	5	
row_ptr	0	2	5	8	12	16	19							.

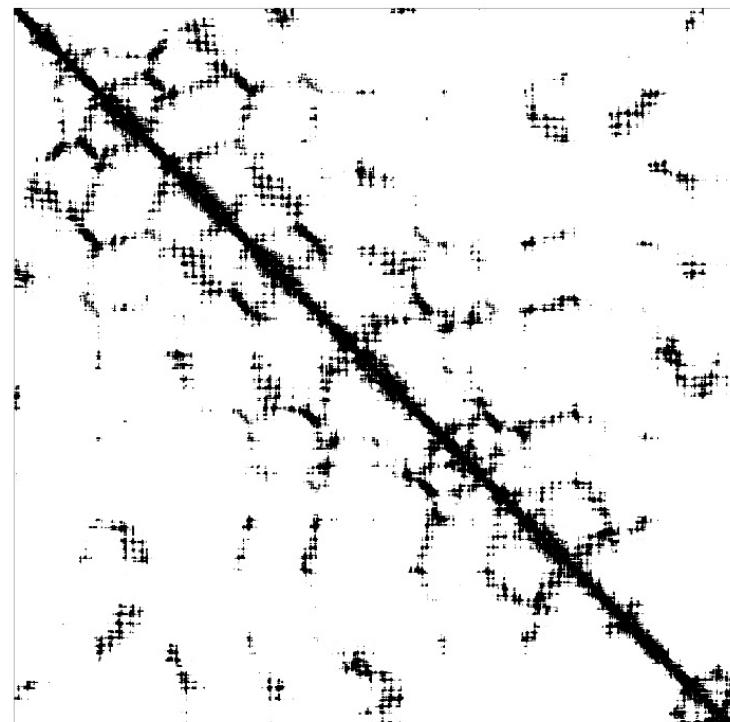
Sparsity pattern of Fock matrix for a protein containing 4626 atoms

Sparse libraries for CPU



A basic tool-kit for sparse matrix computations (Version 2)

Sparse BLAS from Intel MKL



Linear Scaling Strategies

Linear scaling algorithms in chemistry and physics are based on:

- Localization of molecular orbitals;
- Matrix density search and/or purification;
- Divide & Conquer;
- Fragmented molecular orbitals.
- Etc.

Localized Molecular Orbitals and MOZYME

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon}$$

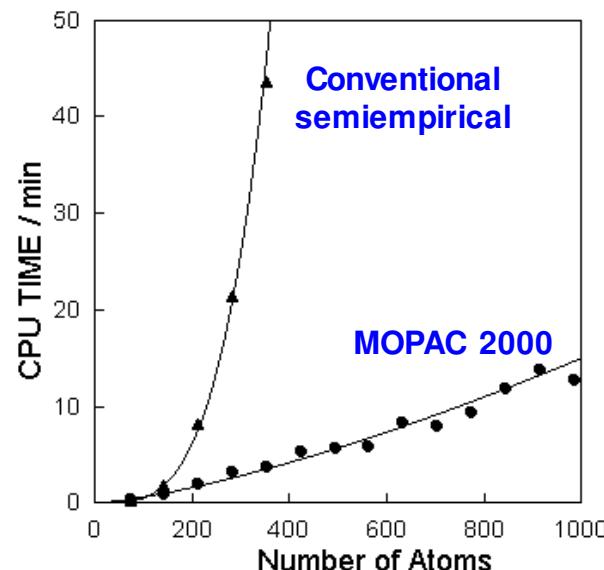
Canonical molecular orbitals are delocalized solutions to the HFR equation.

The canonical MOs are not suitable for the calculation of large molecular systems. →
They do not reflect the local nature of the problem.

To circumvent such difficulty we have to use localized molecular orbitals as initial guess (Lewis based MOs) to solve the HFR equations.

MOZYME

Application of Localized Molecular Orbitals to the Solution of Semiempirical Self-Consistent Field Equations. J. J. P. Stewart. Int. J. Quantum Chem., 58, 1996, 133-146.



Scales as N^2

LocalSCF – linear scaling method

JOURNAL OF CHEMICAL PHYSICS

VOLUME 121, NUMBER 3

15 JULY 2004

LocalSCF method for semiempirical quantum-chemical calculation of ultralarge biomolecules

N. A. Anikin

Quantum Biochemistry Group, Konstantina Fedina-3/24, 105215 Moscow, Russian Federation

V. M. Anisimov^{a)}

FQS Poland, Fujitsu, Starowislna 13-15, 31-422 Krakow, Poland

V. L. Bugaenko, V. V. Bobrikov, and A. M. Andreyev

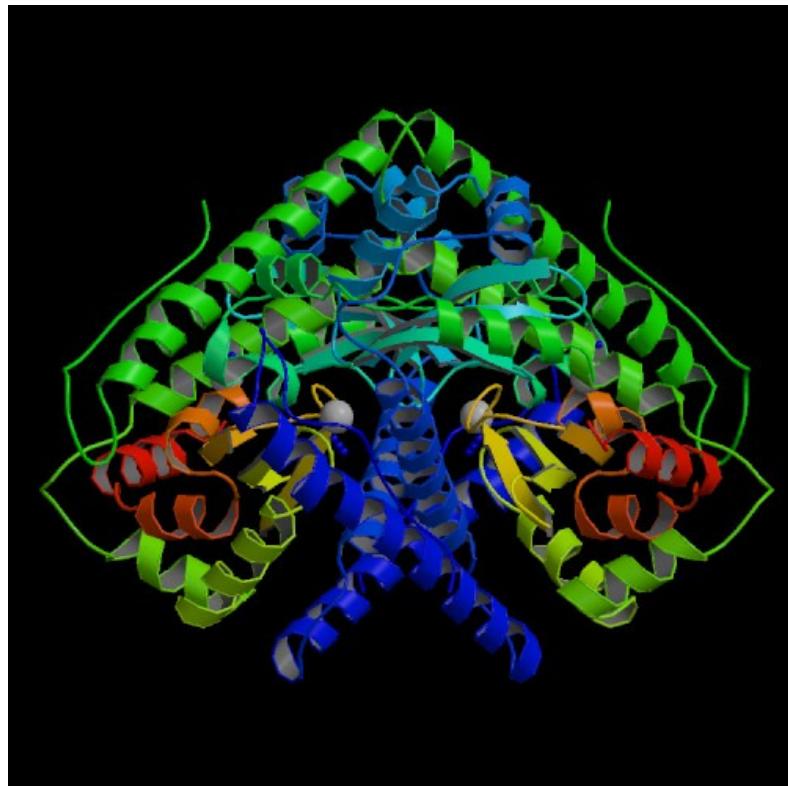
Quantum Biochemistry Group, Konstantina Fedina-3/24, 105215 Moscow, Russian Federation

(Received 26 January 2004; accepted 29 April 2004)

A linear-scaling semiempirical method, LocalSCF, has been proposed for the quantum-chemical calculations of ultralarge molecular systems by treating the large-scale molecular task as a variational problem. The method resolves the self-consistent field task through the finite atomic expansion of weakly nonorthogonal localized molecular orbitals. The inverse overlap matrix arising from the nonorthogonality of the localized orbitals is approximated by preserving the first-order perturbation term and applying the second-order correction by means of a penalty function. This allows for the separation of the orbital expansion procedure from the self-consistent field optimization of linear coefficients, thereby maintaining the localized molecular orbital size unchanged during the refinement of linear coefficients. Orbital normalization is preserved analytically by the variation of virtual degrees of freedom, which are orthogonal to the initial orbitals. Optimization of linear coefficients of localized orbitals is performed by a gradient procedure. The computer program running on a commodity personal computer was applied to the GroEL-GroES chaperonin complex containing 119 273 atoms. © 2004 American Institute of Physics. [DOI: 10.1063/1.1764496]

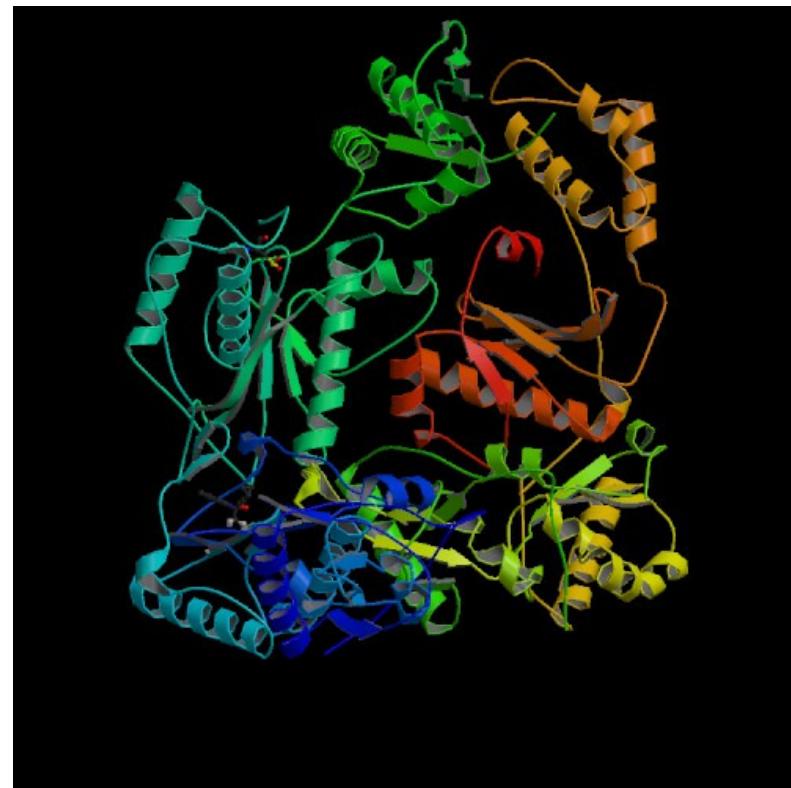
In order to combine the computational efficiency of orthogonal LMOs with the non-orthogonal aspect to better reflect chemical intuition, the authors used weakly non-orthogonal LMOs, where the non-orthogonal case was treated as an orthogonal perturbation.

LocalSCF – Tests



Superóxido Dismutase

6.254 átomos, 1AVM



Transcriptase reversa HIV-1

15.239 átomos, 1FK9

LocalSCF – Tests

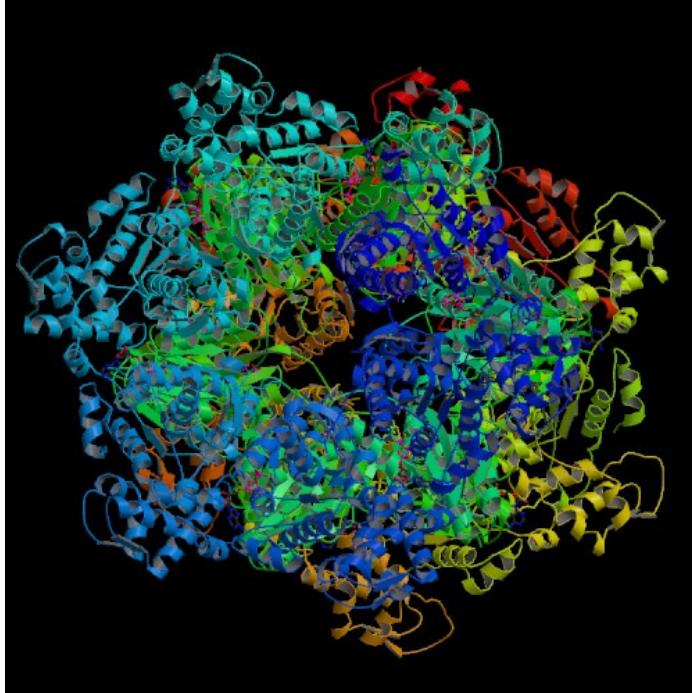


Anticorpo HIV-1 humano
20.462 átomos, 1HZH

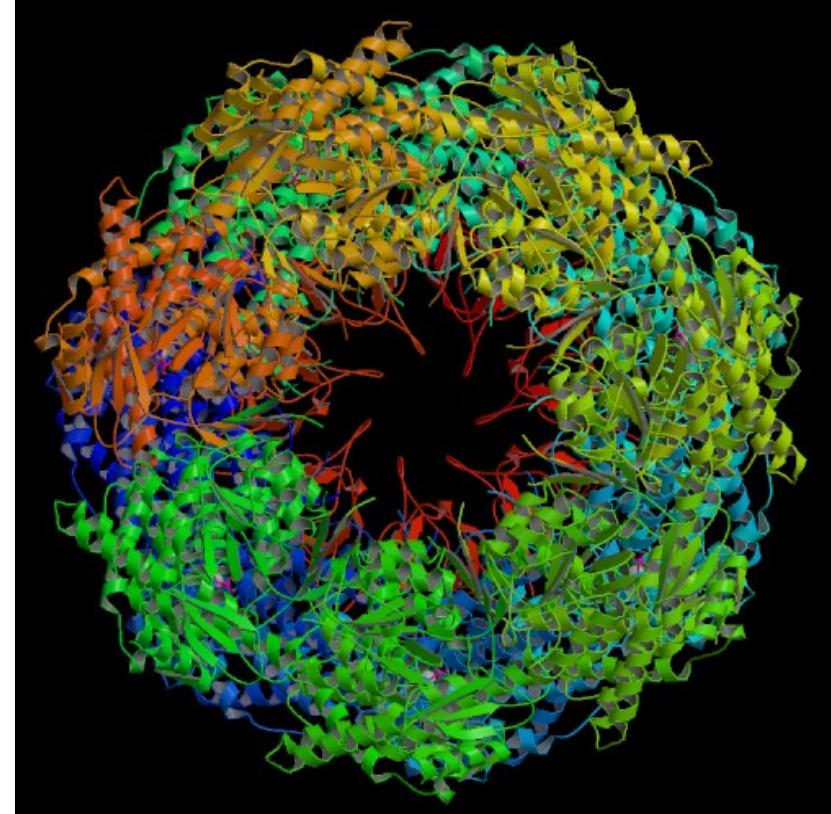


Polimerase RNA
42.470 átomos, 1I6V

LocalSCF – Tests



Complejo protease chaperona
91.510 átomos, 1G3I

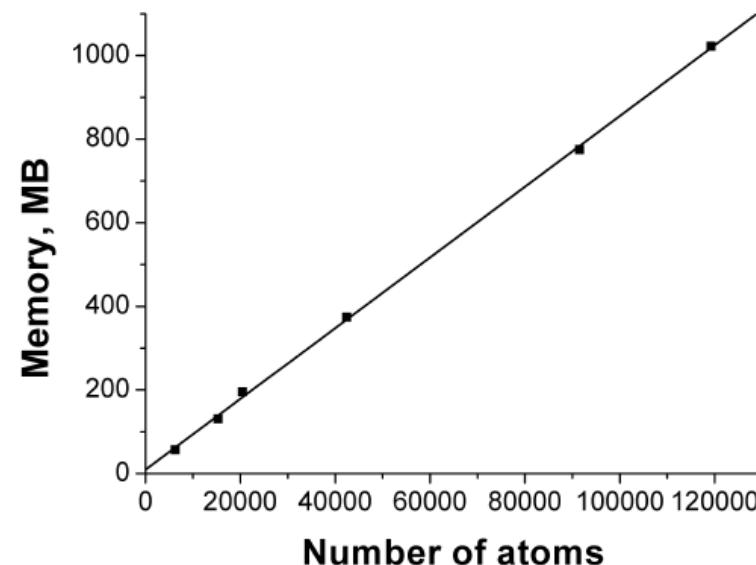
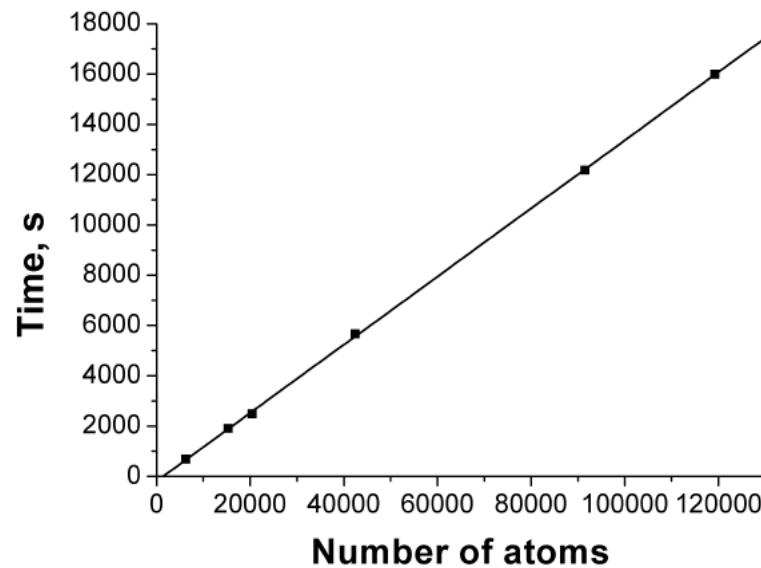


Complejo chaperonina GroEL/GroES
119.273 átomos, 1AON

LocalSCF Results

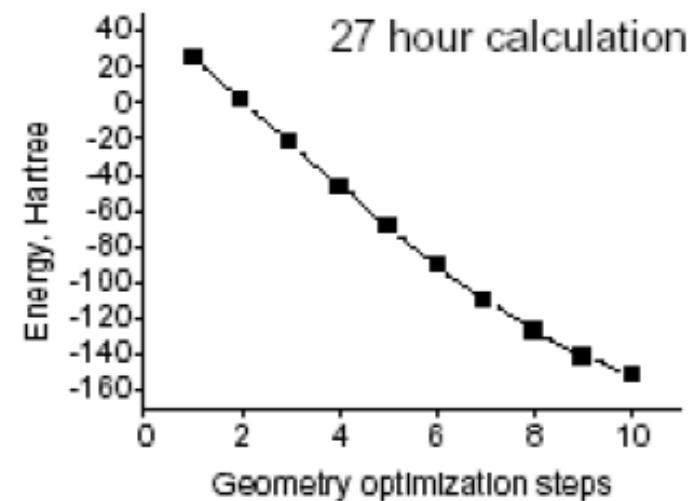
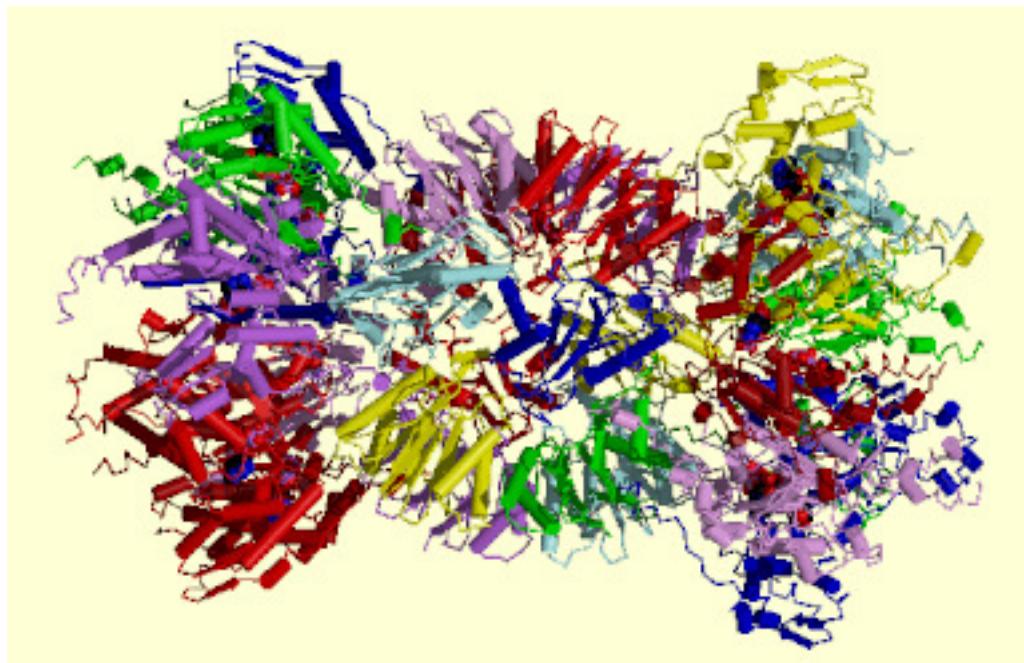
Table 1. Computer time and memory requirement of LocalSCF total energy calculation of proteins at the fixed geometry

Protein	N atoms	PDB id	CPU	Memory,
			Time, sec	Mb
Dismutase	6254	1AVM	684	57
HIV-1 Reverse Transcriptase	15329	1FK9	1911	131
HIV-1 Antibody	20462	1HZH	2489	196
RNA Polymerase	42470	1I6V	5670	374
Protease-Chaperone Complex	91510	1G3I	12174	774
GroEL-GroES Chaperonin	119273	1AON	15988	1022



LocalSCF – linear scaling method

Hsluv protease–chaperone complex (PDB ID: 1G3I)–91.509 átomos



Pentium IV – 2.4GHz – 1Gb RAM

Anikin, N.A., Anisimov, V.M., Bugaenko, V.L, Bobrikov, V.V, Andreyev, A.M., *J. Chem. Phys.* 2004, 121, 1266.

Charge Transfer Effects in the GroEL–GroES Chaperonin Tetramer in Solution

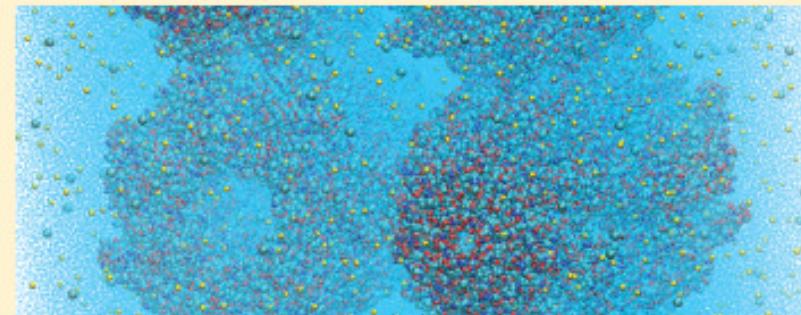
Victor M. Anisimov^{*,†} and Andrey A. Bliznyuk[‡]

[†]National Center for Supercomputing Applications, University of Illinois at Urbana—Champaign, 1205 West Clark Street, Urbana, Illinois 61801, United States

[‡]Australian National University, Supercomputer Facility, Leonard Huxley Bld. (#56), Canberra, ACT 0200, Australia

Supporting Information

ABSTRACT: In this work, we present the results of a large-scale, semiempirical LocalSCF quantum mechanical study of GroEL–GroES chaperonin in solution containing 2 481 723 atoms. We find that large biological systems exhibit strong quantum mechanical character, the extent of which was not previously known. Our data show that protein transfers -743 electron units of charge to solvent, which is not described by classical force fields. Contrary to the commonly held belief, which is based on classical mechanics, our computational data suggest that the quantum mechanical effects of charge transfer increase with the size of biological systems. We show that the neglect of charge transfer in classical force fields leads to significant error in the electrostatic potential of the macromolecule. These findings illustrate that a quantum mechanical framework is necessary for a realistic description of electrostatic interactions in large biological systems.



dx.doi.org/10.1021/jp211385e | J. Phys. Chem. B XXXX, XXX, XXX–XXX

PC Intel octa-core with 48GB de RAM.

GPU Algorithm for $P^2 = P \times P$

Original Paper | Published: 22 October 2020

GPU algorithms for density matrix methods on MOPAC: linear scaling electronic structure calculations for large molecular systems

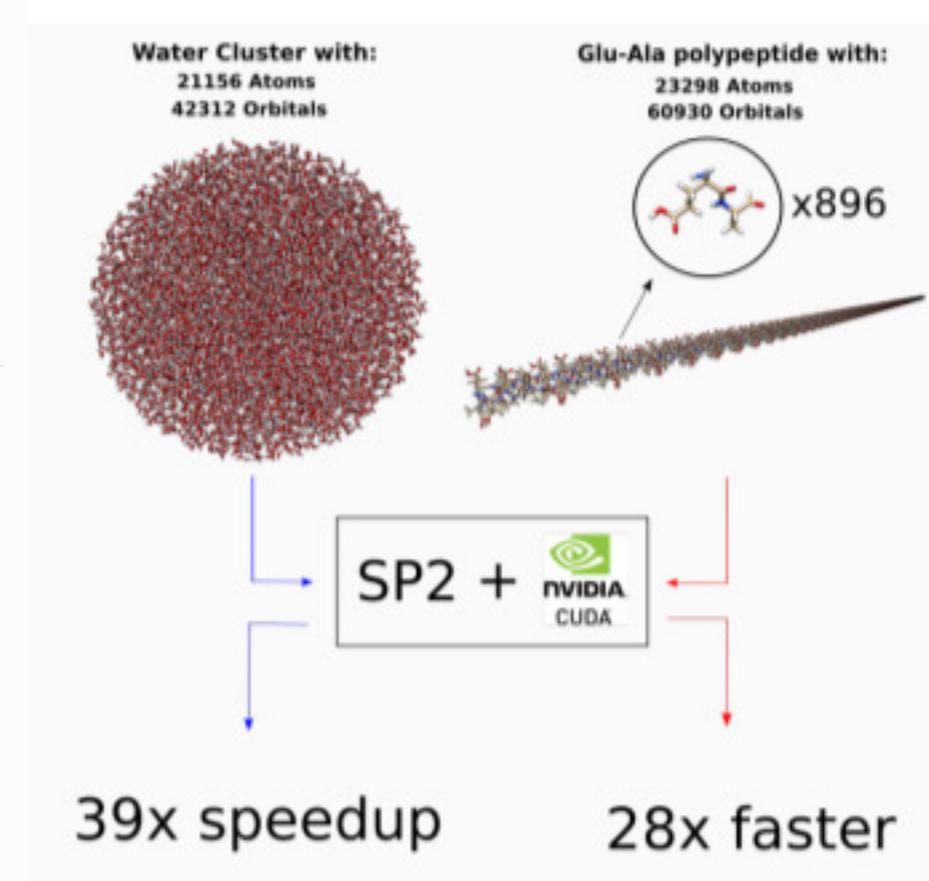
[Julio Daniel Carvalho Maia](#), [Lucidio dos Anjos Formiga Cabral](#) & [Gerd Bruno Rocha](#)✉

Journal of Molecular Modeling **26**, Article number: 313 (2020) | [Cite this article](#)

79 Accesses | [Metrics](#)

Abstract

Purification of the density matrix methods should be employed when dealing with complex chemical systems containing many atoms. The running times for these methods scale linearly with the number of atoms if we consider the sparsity from the density matrix. Since the efficiency expected from those methods is closely tied to the underlying parallel implementations of the linear algebra operations (e.g., $P^2 = P \times P$), we proposed a central processing unit (CPU) and graphics processing unit (GPU) parallel matrix-matrix multiplication in SVBR (symmetrical variable block row) format for energy calculations through the SP2 algorithm. This algorithm was inserted in MOPAC's MOZYME method, using the original LMO Fock matrix assembly, and the atomic integral calculation implemented on it. Correctness and performance tests show that the implemented SP2 is accurate and fast, as the GPU is able to achieve speedups up to 40 times for a water cluster system with 42,312 orbitals running in one NVIDIA K40 GPU card compared to the single-threaded version. The GPU-accelerated SP2 algorithm using the MOZYME LMO framework enables the calculations of semiempirical wavefunction with stricter SCF criteria for localized charged molecular systems, as well as the single-point energies of molecules with more than 100.000 LMO orbitals in less than 1 h.



Purification of Density-matrix Methods (PDM)

- **Objective:** To discover the density matrix (\mathbf{P}) without explicit knowledge of the molecular orbitals (\mathbf{C} , which is not a sparse matrix), making diagonalization of the Fock matrix (\mathbf{F}) unnecessary.

$$E = \text{Tr}(\mathbf{P} \cdot [\mathbf{H} + \mathbf{F}])$$

$$\mathbf{P} = \mathbf{C}_{\text{occ}} \times \mathbf{C}_{\text{occ}}^T$$

\mathbf{P} : is symmetrical ($\mathbf{P} = \mathbf{P}^T$), normalized for n_e , ($\text{Tr}(\mathbf{P}) = n_e/2$) and idempotent ($\mathbf{P}^2 = \mathbf{P}$).

Properties that are considered by using PDM strategy

- The density matrix also has to commute with the \mathbf{F} matrix on the SCF convergence limit.
- Any eigenvector from \mathbf{F} is also an eigenvector from \mathbf{P} .
- Eigenvalues should be **0** or **1**, and this statement forces that the first **nocc** eigenvectors from \mathbf{F} are also eigenvectors from \mathbf{P} with eigenvalue **1**. The other **n-occ** eigenvectors from \mathbf{P} have **0** as their corresponding eigenvalue.

Purification of Density-matrix Methods (PDM)

Strategy:

- From \mathbf{F} , you can find \mathbf{P} transforming the n_{occ} lowest eigenvalues to $\mathbf{1}$ and the highest eigenvalues to $\mathbf{0}$ by using purification expansions applied to the one-electron density matrix.

$$\rho = \theta[\mu\mathbf{I} - \mathbf{H}]$$

Heaviside step function

$$\theta[x] = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

To expand $\theta(\mu\mathbf{I} - \mathbf{H})$ by using fast convergent polynomial functions.

$$\theta[\mu\mathbf{I} - \mathbf{H}] = \lim_{i \rightarrow \infty} f_i [f_{i-1} [\dots f_0 [\mathbf{X}_0] \dots]]$$

$$\mathbf{X}_0 = (\epsilon_{\max}\mathbf{I} - \mathbf{H}) / (\epsilon_{\max} - \epsilon_{\min})$$

Effective Scaled Hamiltonian

Gershgorin Circle Theorem

$$\epsilon_{\min} = \min_i \left\{ \mathbf{H}_{ii} - \sum_{j \neq i} |\mathbf{H}_{ij}| \right\},$$

$$\epsilon_{\max} = \max_i \left\{ \mathbf{H}_{ii} + \sum_{j \neq i} |\mathbf{H}_{ij}| \right\}.$$

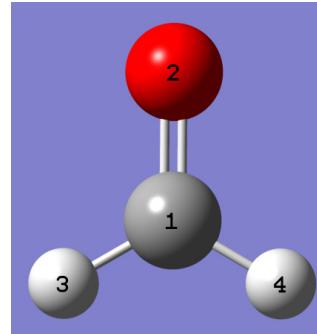
$$f_i[\mathbf{X}_i] = \begin{cases} \mathbf{X}_i^2 & \text{if } 2\text{Tr}[\mathbf{X}_i] \geq N_e \\ 2\mathbf{X}_i - \mathbf{X}_i^2 & \text{if } 2\text{Tr}[\mathbf{X}_i] < N_e \end{cases}$$

We need to evaluate $\mathbf{A} \cdot \mathbf{A}$ and $\text{Tr}[\mathbf{A}]$

Sparse Format in MOPAC (SVBR)

SVBR → Symmetrical Variable Block Row Sparse Format

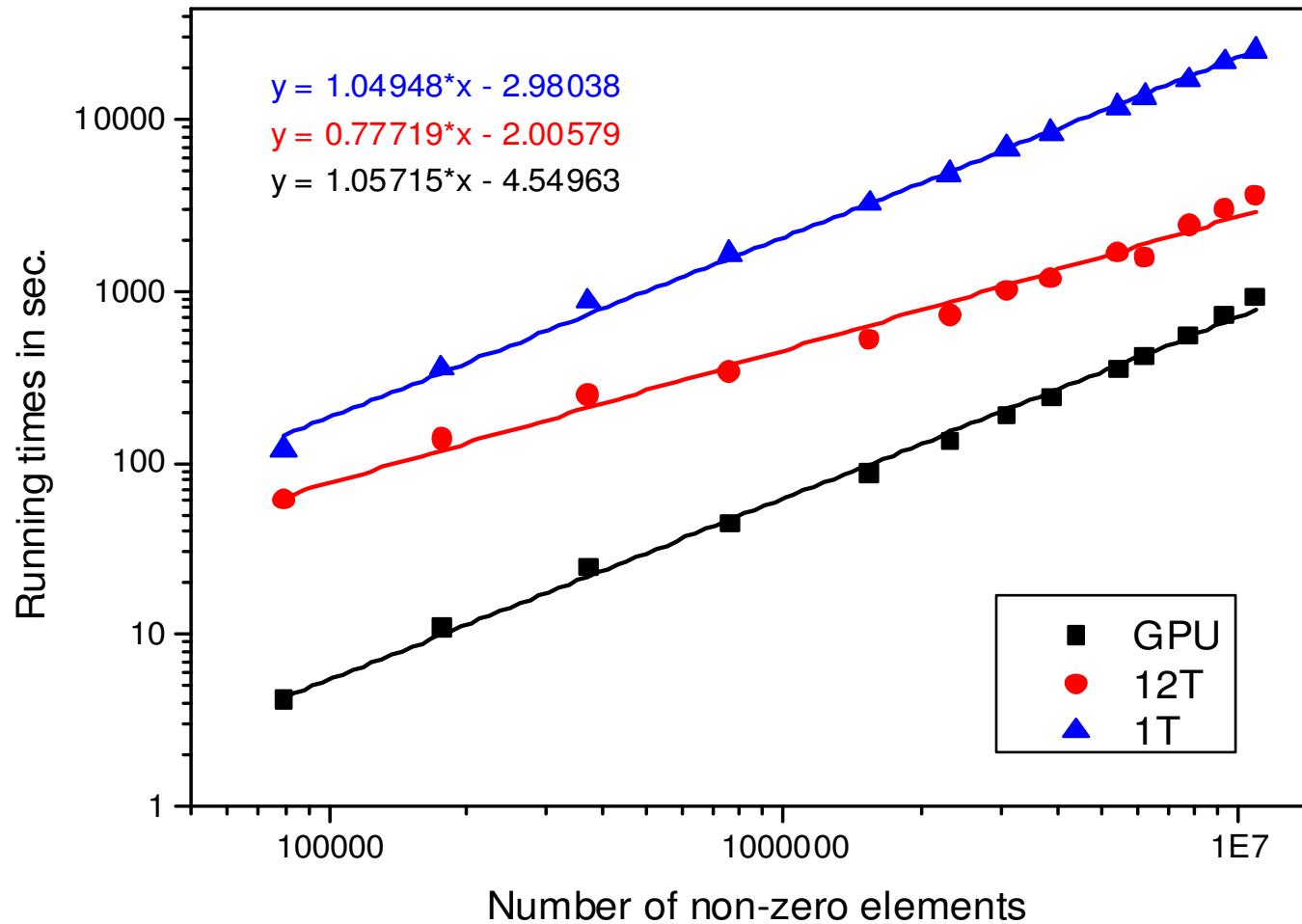
- Blocks that represent interatomic interactions beyond the pre-set cut-off radius are avoided during the calculation;



C ¹				O ²				H ³ H ⁴	
s	p _x	p _y	p _z	s	p _x	p _y	p _z	s	s
C ¹	1								
	2	3							
	4	5	6						
	7	8	9	10					
O ²	11	12	13	14	27				
	15	16	17	18	28	29			
	19	20	21	22	30	31	32		
	23	24	25	26	33	34	35	36	
H ³	37	38	39	40	41	42	43	44	45
	46	47	48	49	50	51	52	53	55
H ⁴									

We developed a parallel (CPU/GPU) algorithm to multiply two symmetrical sparse matrices and put the results in another sparse matrix with the same SVBR format → $P^2 = P \times P$.

Scaling

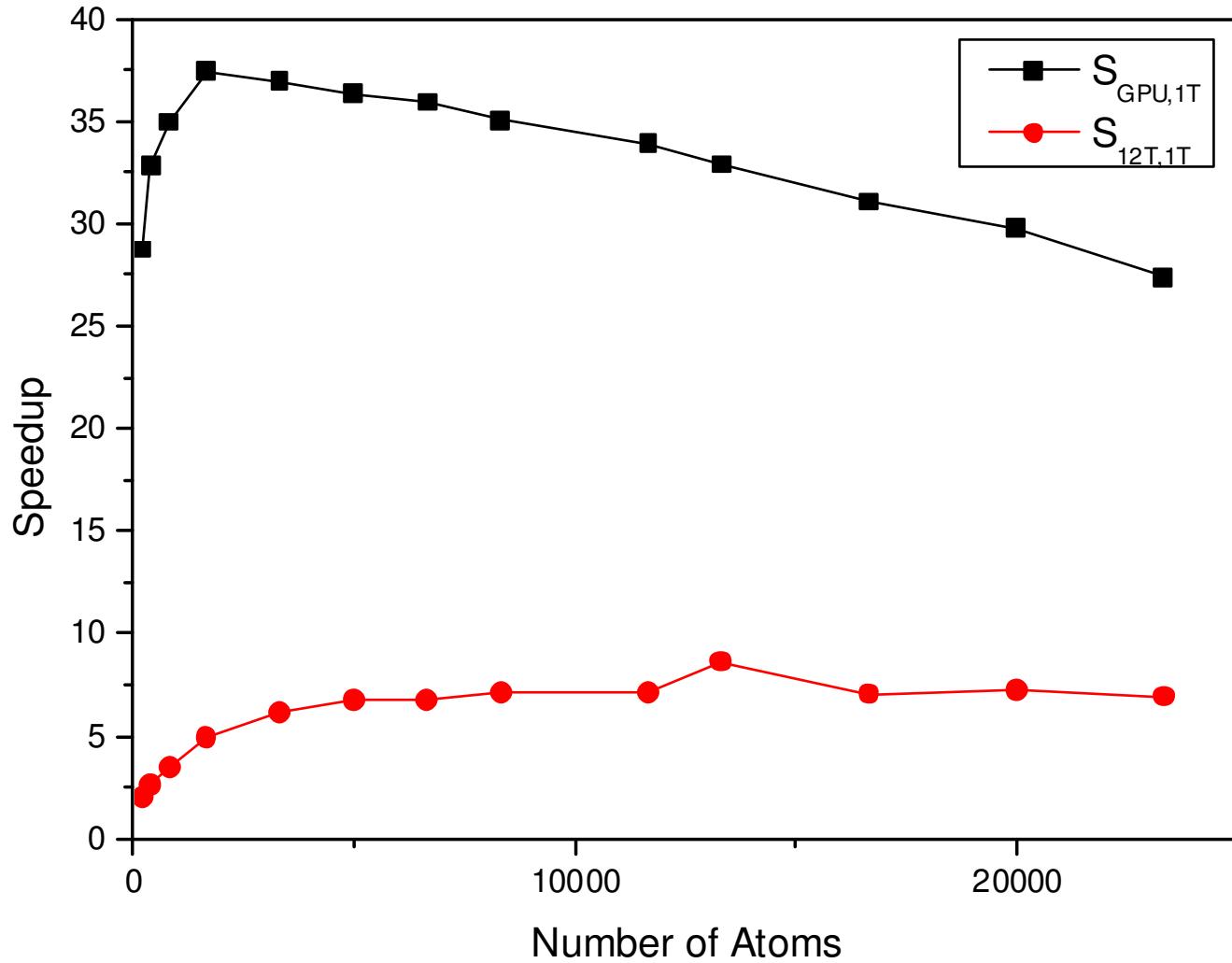


$(\text{Glu-Ala})_{16}$

The largest molecular system contains 23 thousand atoms totalizing 60k atomic orbitals

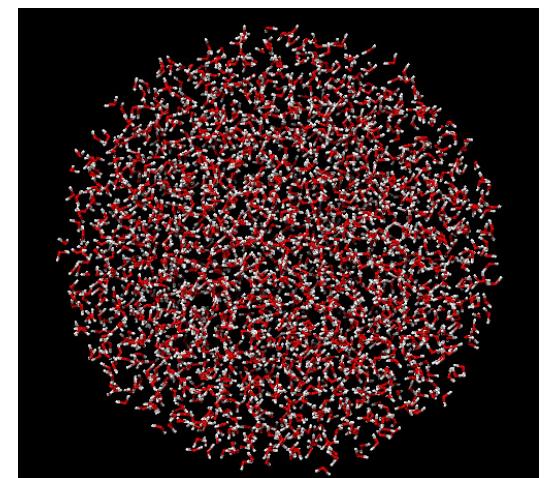
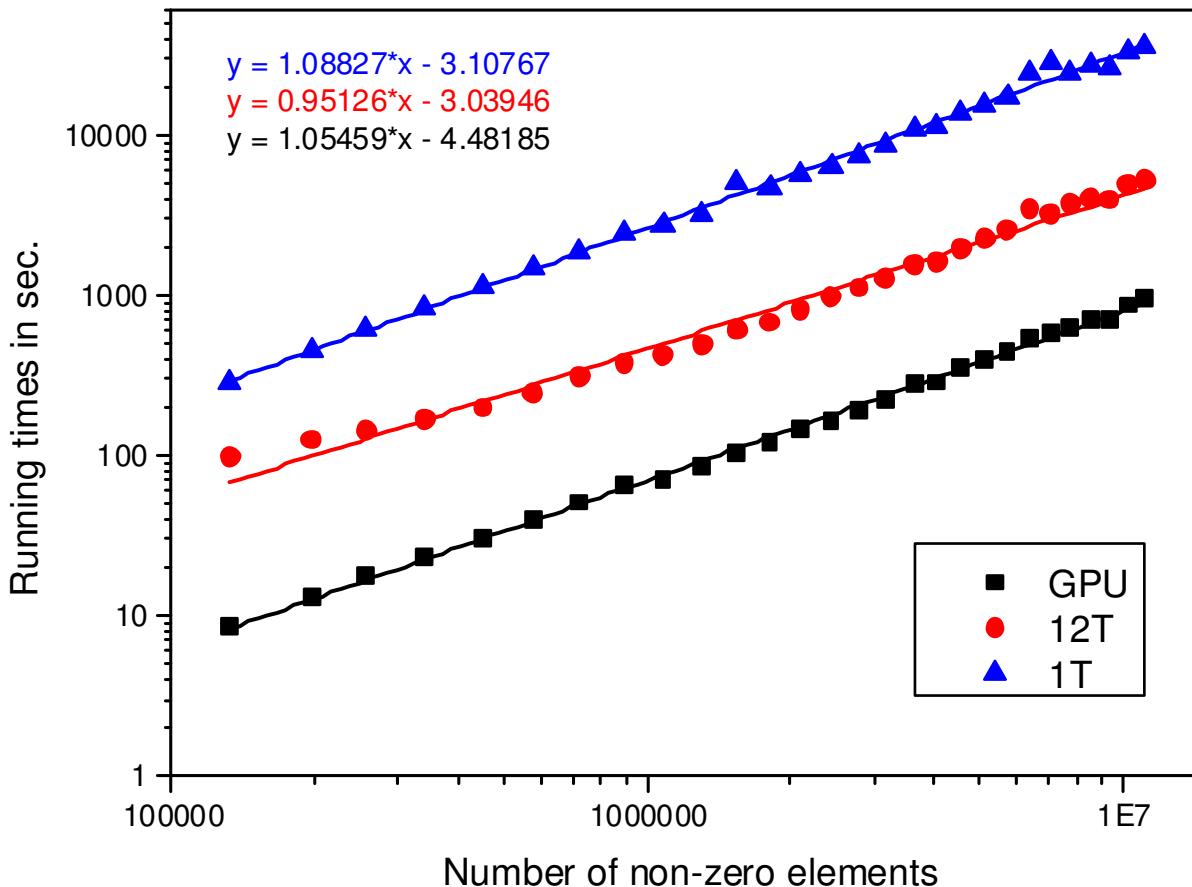
$(\text{Glu-Ala})_{896}$.

Speedup



(Glu-Ala)₁₆

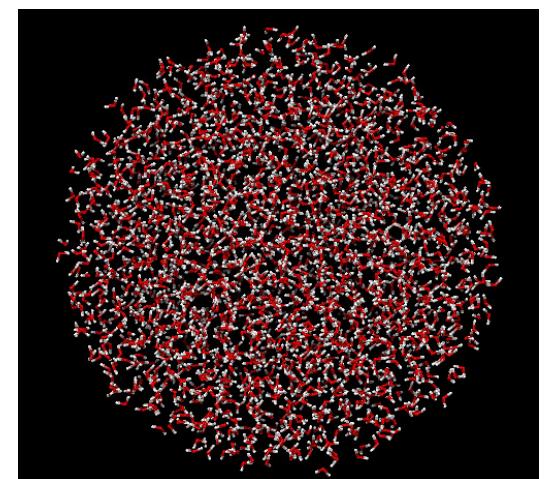
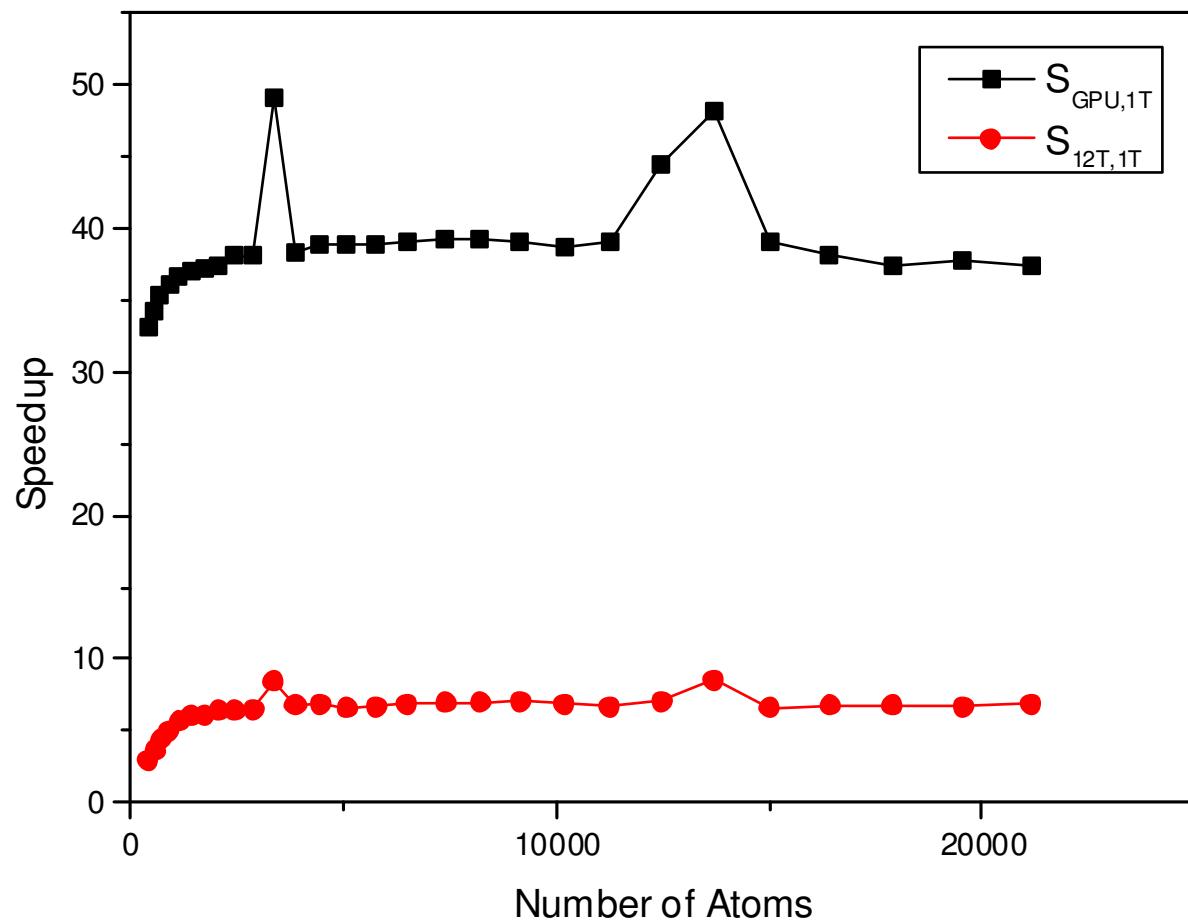
Scaling



Spherical water box of
25 Å of radius.

Largest molecular system has more than **20k atoms**, **W₃₇**, resulting in **42k basis set**.

Speedup



Spherical water box of
25 Å of radius.

Linear scaling strategies for QM calculations

Received: 21 March 2020 | Revised: 25 May 2020 | Accepted: 31 May 2020
DOI: 10.1002/wcms.1494

SOFTWARE FOCUS



TeraChem: A graphical processing unit-accelerated electronic structure package for large-scale ab initio molecular dynamics

Stefan Seritan^{1,2} | Christoph Bannwarth^{1,2} | Bryan S. Fales^{1,2} |
Edward G. Hohenstein^{1,2} | Christine M. Isborn³ |
Sara I. L. Kokkila-Schumacher⁴ | Xin Li⁵ | Fang Liu⁶ |
Nathan Luehr⁷ | James W. Snyder Jr.⁸ | Chenchen Song^{9,10} |
Alexey V. Titov¹¹ | Ivan S. Ufimtsev¹² | Lee-Ping Wang¹³ |
Todd J. Martinez^{1,2}

Cite this: *Phys. Chem. Chem. Phys.*, 2012, 14, 7562–7577

www.rsc.org/pccp

PERSPECTIVE

Exploring chemistry with the fragment molecular orbital method

Dmitri G. Fedorov,^{*a} Takeshi Nagata^{a,b} and Kazuo Kitaura^{a,b}

Received 29th November 2011, Accepted 21st February 2012

DOI: 10.1039/c2cp23784a

The fragment molecular orbital (FMO) method makes possible nearly linear scaling calculations of large molecular systems, such as water clusters, proteins and DNA. In particular, FMO has been widely used in biochemical applications involving protein-ligand binding and drug design. The method has been efficiently parallelized suitable for petascale computing. Many commonly used wave functions and solvent models have been interfaced with FMO. We review the historical background of FMO and summarize its method development and applications.

Available online at www.prace-ri.eu



Partnership for Advanced Computing in Europe

Million Atom KS-DFT with CP2K

Iain Bethune^{a*}, Adam Carter^a, Xu Guo^a, Paschalidis Korosoglou^{b,c}

^aEPCC, The University of Edinburgh, James Clark Maxwell Building, The King's Buildings, Edinburgh EH9 3JZ, United Kingdom

^bAUTH, Aristotle University of Thessaloniki, Thessaloniki 52124, Greece

^cGRNET, Greek Research & Technology Network, L. Mesogeion 56, Athens 11327, Greece

QM calculations for molecular sysmetms with over 1 million of atoms.

131

Charge Transfer Effects in the GroEL–GroES Chaperonin Tetramer in Solution

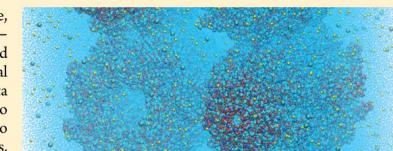
Victor M. Anisimov^{*,†} and Andrey A. Bliznyuk[‡]

^{*}National Center for Supercomputing Applications, University of Illinois at Urbana—Champaign, 1205 West Clark Street, Urbana, Illinois 61801, United States

[‡]Australian National University, Supercomputer Facility, Leonard Huxley Bld. (#56), Canberra, ACT 0200, Australia

Supporting Information

ABSTRACT: In this work, we present the results of a large-scale, semiempirical LocalSCF quantum mechanical study of GroEL–GroES chaperonin in solution containing 2 481 723 atoms. We find that large biological systems exhibit strong quantum mechanical character, the extent of which was not previously known. Our data show that protein transfers ~743 electron units of charge to solvent, which is not described by classical force fields. Contrary to the commonly held belief, which is based on classical mechanics, our computational data suggest that the quantum mechanical effects of charge transfer increase with the size of biological systems. We show that the neglect of charge transfer in classical force fields leads to significant error in the electrostatic potential of the macromolecule. These findings illustrate that a quantum mechanical framework is necessary for a realistic description of electrostatic interactions in large biological systems.



ErgoSCF

home page!

Ergo, a quantum chemistry program for large-scale self-consistent field



Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Linear Scaling Self-Consistent Field Calculations with Millions of Atoms in the Condensed Phase

Joost VandeVondele,^{*,†} Urban Borštník,^{‡,§} and Jiří Hutter[‡]

Linear scaling strategies for QM calculations

We are also able to get averages of many events !!!!

arXiv.org > physics > arXiv:2104.08245

Search...

Help | Advanced

Physics > Computational Physics

[Submitted on 16 Apr 2021 (v1), last revised 23 Apr 2021 (this version, v2)]

Enabling Electronic Structure-Based Ab-Initio Molecular Dynamics Simulations with Hundreds of Millions of Atoms

Robert Schade, Tobias Kenter, Hossam Elgabarty, Michael Lass, Ole Schütt, Alfio Lazzaro, Hans Pabst, Stephan Mohr, Jürg Hutter, Thomas D. Kühne, Christian Plessl

We push the boundaries of electronic structure-based ab-initio molecular dynamics (AIMD) beyond 100 million atoms. This scale is otherwise barely reachable with classical force-field methods or novel neural network and machine learning potentials. We achieve this breakthrough by combining innovations in linear-scaling AIMD, efficient and approximate sparse linear algebra, low and mixed-precision floating-point computation on GPUs, and a compensation scheme for the errors introduced by numerical approximations.

The core of our work is the non-orthogonalized local submatrix (NOLSM) method, which scales very favorably to massively parallel computing systems and translates

Code	Year	Method	Basis	System	# Atoms	# Cores	Machine	Peak Performance	Efficiency
CPMD [23]	2005	DFT	PW	bulk SiC	1k	1.2k CPU	IBM p690	1.087 TFLOP/s	≈ 20%
Qbox [24]	2006	DFT	PW	bulk Mo	8*1k	128k CPU	IBM BlueGene/L	207.3 TFLOP/s	56.5%
LS3DF [25]	2009	DFT	PW	bulk ZnTeO	36k	147k CPU	Cray Jaguar	442 TFLOP/s	≈ 33%
CP2K [26]	2012	LS-DFT	GPW	bulk H ₂	1M	47k CPU	Cray XT5		
ONETEP [27]	2014	LS-DFT	NGWF	amyloid fibril trimer	42k	115k CPU	IBM BlueGene/Q		
RSDFD [28]	2014	DFT	RS-FD	Si nanowire	107k	664k CPU	K-Computer	5.48 PFLOP/s	51.67%
CP2K [29]	2016	SS-DFT	GPW	satellite tobacco mosaic virus	1M	20k CPU	Cray XC30		
LDC-DFT [30]	2014	SS-DFT	RMG-PW	bulk SiC	6.3M	786k CPU	IBM Blue Gene/Q	5.08 PFLOP/s	50.5%
OpenAtom [31]	2016	DFT	PW	periodic MOF	32*424	262k CPU	IBM BlueGene/Q		≈ 52%
MGmol [32]	2016	LS-DFT	FD	bulk H ₂ O	1.2M	1.6m CPU	IBM BlueGene/Q		≈ 39%
DFT-FE [33]	2019	DFT	FEM	Mg cluster	10.5k	159k CPU +22.8k GPUs	IBM Summit	46 PFLOP/s	27.8%
CONQUEST [34]	2020	LS-DFT	PAO	bulk Si	1M	200k CPU	K-Computer		
This work	2021	LS-DFT	GTO	bulk water	102M	18.4k CPU +1.5k GPUs	JUWELS Booster	206 PFLOP/s	43%
This work	2021	LS-DFT	GTO	HIV-1 capsid in solution	62.5M	18.4k CPU +1.5k GPUs	JUWELS Booster	324 PFLOP/s	67.7%

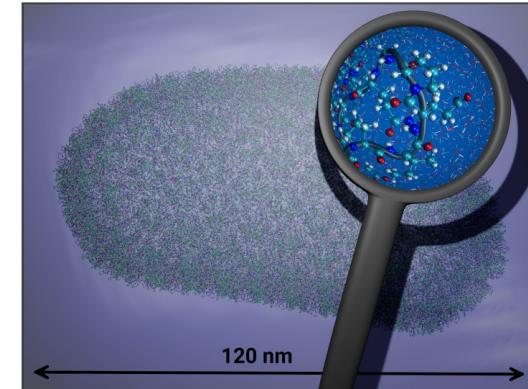
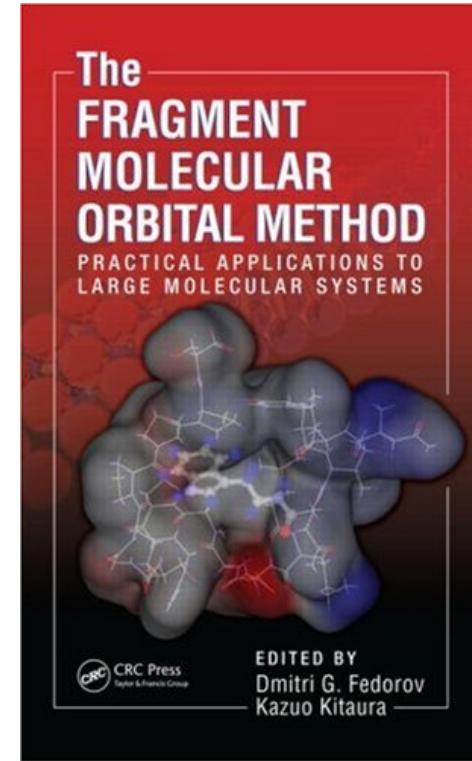
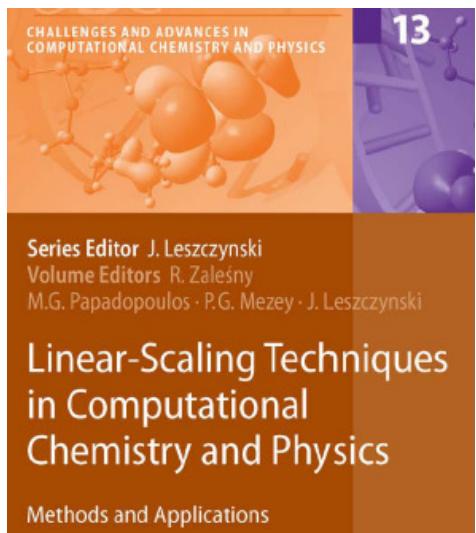


Figure 3. Graphical representation of the present HIV-1 capsid in aqueous solution containing more than 62.5 million atoms.

PDB 3J3Q

324 PFLOP/s in mixed FP16/FP32 precision corresponding to an efficiency of 67.7% when running on 1536 NVIDIA A100 GPUs.



COMPUTATIONAL
CHEMISTRY

10.1109/MCISE.2003.1208637

LINEAR SCALING ELECTRONIC STRUCTURE METHODS IN CHEMISTRY AND PHYSICS

Calculating the electronic structure of large atomistic systems requires algorithms that scale linearly with system size. Efficient implementations of these emerging algorithms provide scientists in various fields with powerful software tools to address challenging problems.

STEFAN GOEDECKER
CEA-Grenoble
GUSTAVO E. SCUSERIA
Rice University

Hybrid Methods

QM/MM

QM/MM Hybrid Methods

Theoretical Studies of Enzymic Reactions:

Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme

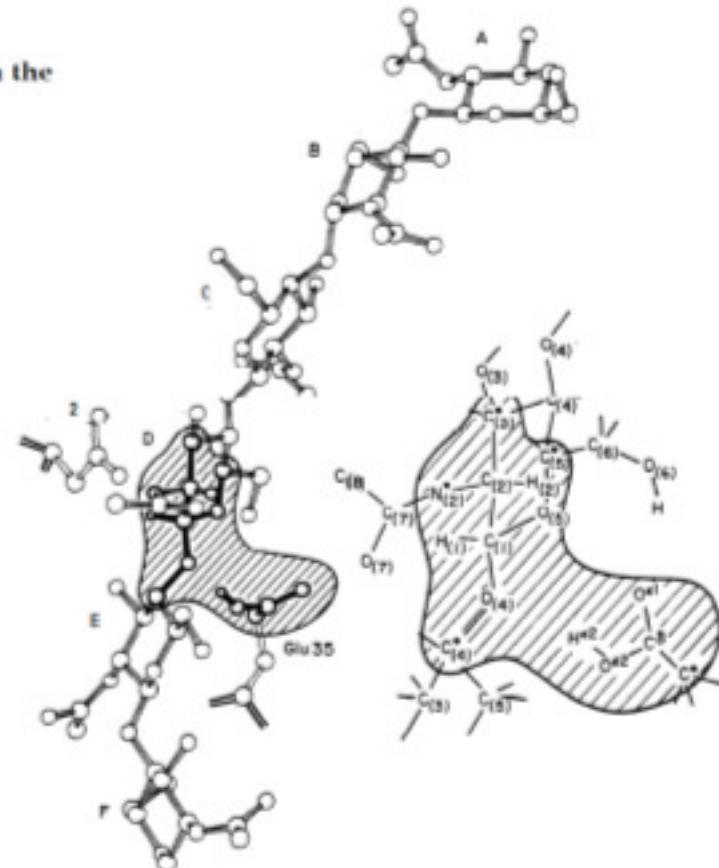
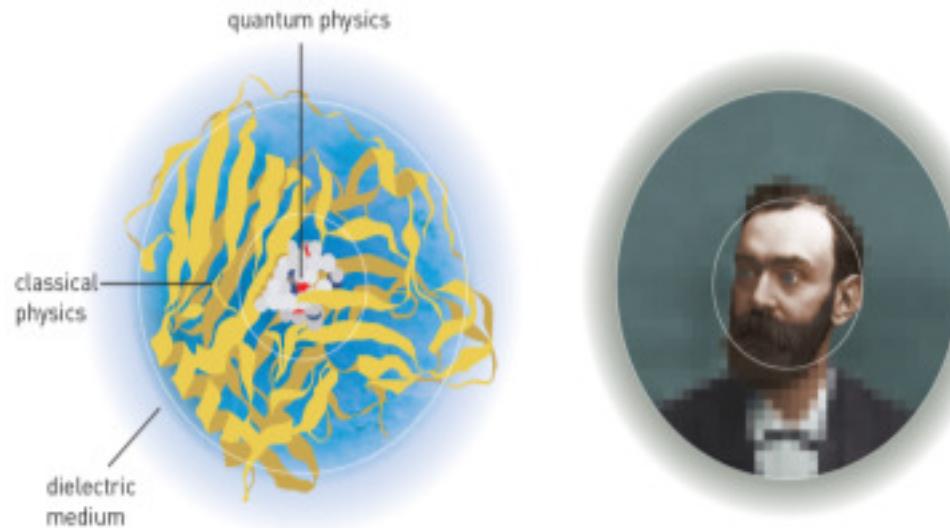
A. MARSHAL AND M. LEVITT

Medical Research Council Laboratory of Molecular Biology
Hills Road, Cambridge CB2 2QH, England

and

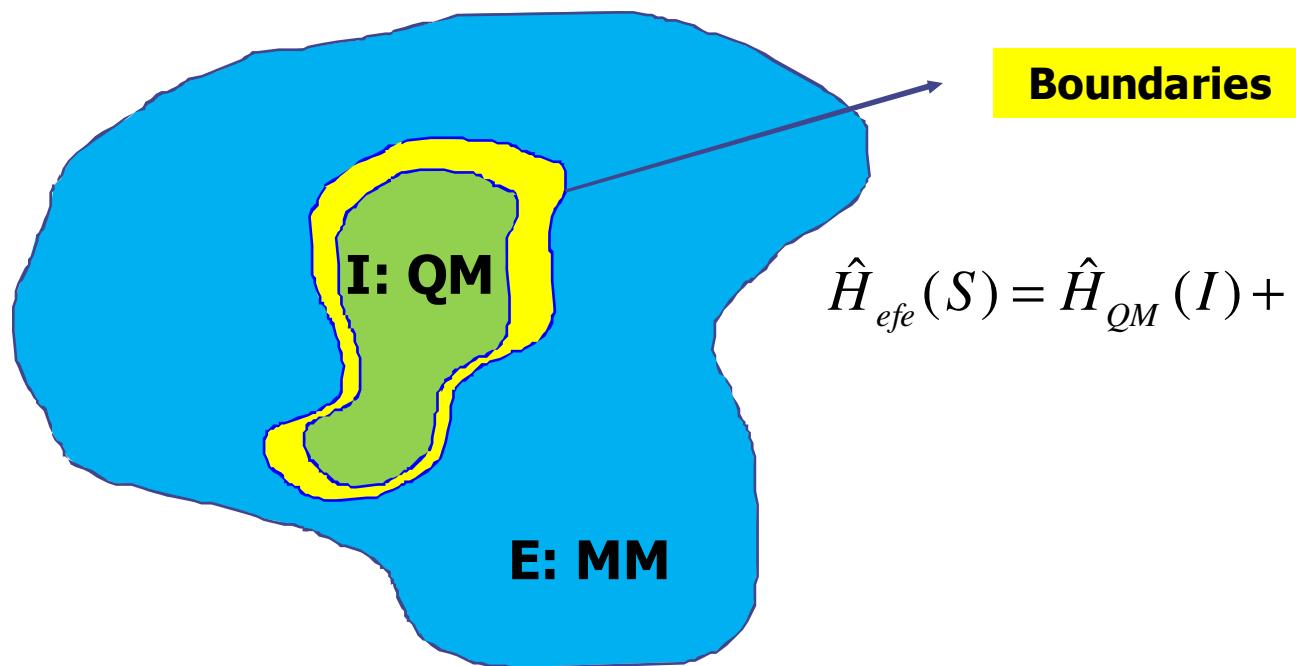
Department of Chemical Physics
The Weizmann Institute of Science
Rehovot, Israel

J. Mol. Biol. (1976) 103, 227-249



QM/MM Hybrid Methods

S: QM/MM



$$\hat{H}_{efe}(S) = \hat{H}_{QM}(I) + \hat{H}_{MM}(E) + \hat{H}_{QM/MM}(E, I)$$

$$E_{total}(S) = E_{MM}(S) + E_{QM}(I) - E_{MM}(I)$$

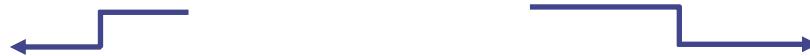
QM/MM subtractive coupling

$$E_{total}(S) = E_{QM}(I) + E_{MM}(E) + E_{QM/MM}(E, I)$$

QM/MM additive coupling

$$E_{QM/MM}(E, I) = E_{QM/MM}^{lig} + E_{QM/MM}^{vDW} + E_{QM/MM}^{elet}$$

For cutting
chemical bonds



- *Mechanical Embedding*
- *Electrostatic Embedding*
- *Polarization Embedding*

QM/MM: Subtractive coupling

ONIOM

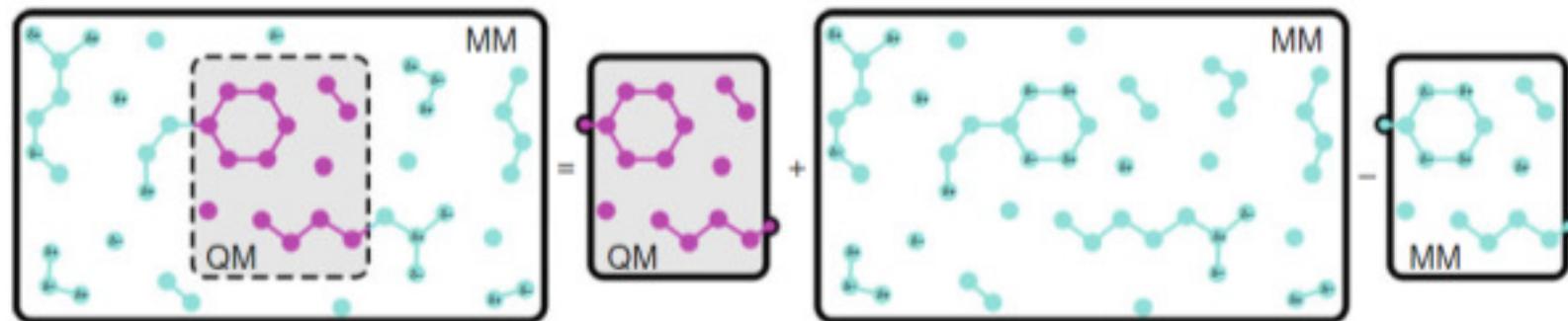


Fig. 2. Subtractive QM/MM coupling: The QM/MM energy of the total system (*left hand side of the equation*) is assumed to be equal to the energy of the isolated QM subsystem, evaluated at the QM level, plus the energy of the complete system evaluated at the MM level, minus the energy of the isolated QM subsystem, evaluated at the MM level. The last term is subtracted to correct for double counting of the contribution of the QM subsystem to the total energy. A prerequisite for the calculation is that a force field for the QM subsystem is available.

<https://www.ncbi.nlm.nih.gov/pubmed/23034745>

S. Dapprich; I. Komaromi; K.S. Byun; K. Morokuma & M.J. Frisch (1999). "A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives". *Journal of Molecular Structure: THEOCHEM*. 461-462: 1. doi:10.1016/S0166-1280(98)00475-8

<http://chem.wayne.edu/schlegel/Software/oniontoolTAO/TAOTutorial.html>

QM/MM: Additive coupling

Mechanical Embedding

1. The interaction is handled classically, partial charges are attributed to the QM atoms in order to interact with the partial charges of the MM part.
2. The QM region is not polarized
3. Simple implementation
4. Difficulty in obtaining good partial charges for QM atoms.

$$\hat{H}_{QM/MM} = \sum_{A < B} \sum_{a \in A}^{N_{MM-atoms}} \sum_{b \in B}^{N_{QM-atoms}} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ab}}{r_{ab}} \right)^{12} \left(\frac{\sigma_{ab}}{r_{ab}} \right)^6 \right] + \sum_{A < B} \sum_{a \in A} \sum_{b \in B} \frac{q_a q_b}{r_{ab}}$$

QM/MM: Additive coupling

Electrostatic Embedding

1. MM atoms appear as generating centers of electrostatic contribution to the Hamiltonian QM.
2. This contribution is added in the Hamiltonian QM, (1-electron part).
3. The QM region is polarized
4. Computationally more expensive.

$$\hat{H}_{QM/MM} = \sum_{A < B} \sum_{a \in A}^{N_{MM-atoms}} \sum_{b \in B}^{N_{QM-atoms}} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ab}}{r_{ab}} \right)^{12} \left(\frac{\sigma_{ab}}{r_{ab}} \right)^6 \right] + \sum_m \sum_b \frac{Z_m q_b}{r_{mb}} - \sum_i \sum_b \frac{q_b}{r_{ib}}$$

Classical term

Added to 1-electron elements of H_{QM} .

<https://www.ncbi.nlm.nih.gov/pubmed/23034745>

QM/MM: Additive coupling

Polarization Embedding

In the Polarization Embedding scheme, both regions can polarize each other. Thus, not only the QM region is polarized by MM atoms, but the QM region can also induce polarization in the MM system, using these approaches:

- charge-on-a-spring model
- the induced dipole model
- the fluctuating charge model

It needs self-consistency between the QM part and the MM part. This leads to high computational demand, so it is used more for systems with fewer atoms.

It is the most realistic of all QM/MM approaches

QM/MM: Cutting chemical bonds

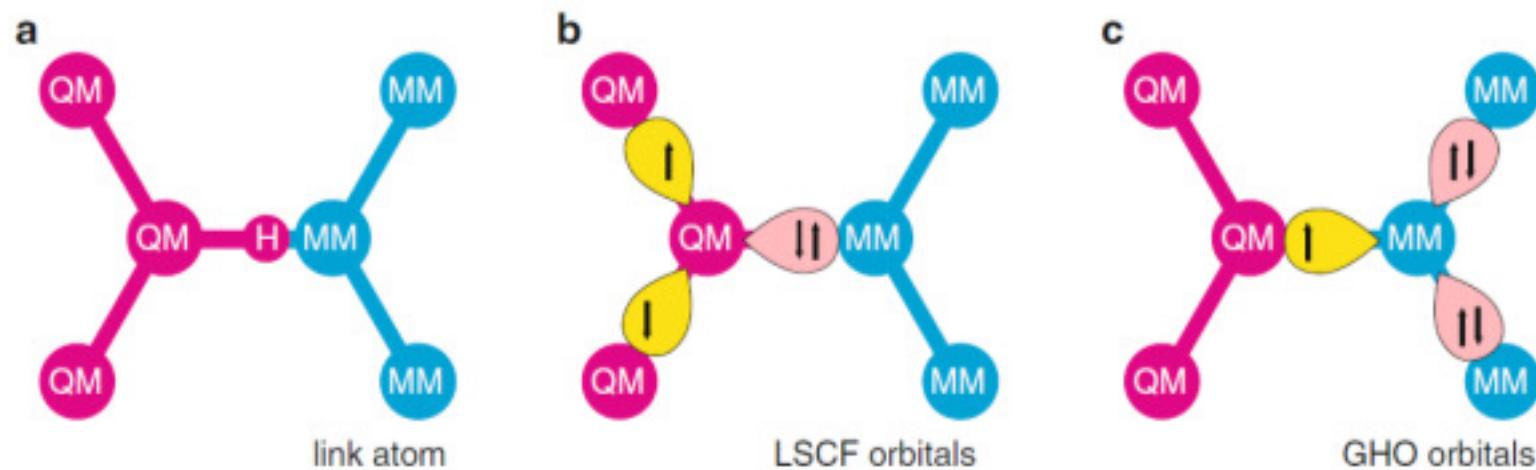


Fig. 5. Different approaches to cap the QM region: link atoms (a) and frozen orbitals (b,c). The hydrogen link atom (a) is placed at an appropriate distance along the QM/MM bond vector and is present only in the QM calculation. In the localized SCF method (b), a set of localized orbitals is placed on the QM atom. During the SCF iterations, the orbital pointing towards the MM atom is double-occupied and frozen, while the other orbitals are single-occupied and optimized. In the generalized hybrid orbital approach (c), a set of localized orbitals is placed on the MM atom. During the SCF interaction, the orbitals pointing towards the other MM atoms are double occupied and frozen, while the orbital pointing towards the QM atom is single-occupied and optimized.

NAMD goes quantum: an integrative suite for hybrid simulations

NIH CENTER FOR MACROMOLECULAR MODELING & BIOINFORMATICS | UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

THEORETICAL and COMPUTATIONAL BIOPHYSICS GROUP

Home Research Publications Software Instruction News Galleries Facilities About Us

Hybrid QM/MM NAMD

NAMD QM/MM interface extends existing NAMD features to the quantum mechanical level, presenting features that are not yet available in any QM/MM implementation. The first is the ability to execute multiple QM regions in parallel, thorough independent executions of your choice of quantum chemistry code. This allows one to account for multiple reaction centers that are known to work synergistically, for example, or even distant allosteric regulation sites and a reaction center. Investigation of processes occurring on a timescale usually not accessible by QM/MM methods can now be performed by a combination of temperature replica exchange molecular dynamics and QM/MM molecular dynamics. Another innovation comes from the fact that most QM/MM implementations have pre-defined QM and MM regions with no changes of atoms between these regions during a simulation, which would not efficiently allow the study of translocation of DNA in, e.g., nanosensors. Taking advantage of an easy-to-use Tcl based interface and capabilities integrated from VMD, NAMD QM/MM allows the update of the QM and MM regions at every step of the QM/MM simulation.

Recent News and Announcements

NAMD 2.12 Release

NAMD 2.12 is now available allowing hybrid QM/MM calculations. More information will be available soon.

VMD QM/MM Features

The new VMD, to be released in the Fall 2017, contains many QM/MM features, including new orbital representations and simulation setup and control through the QwikMD interface. More information will be available soon.

Availability

Hybrid QM/MM calculations are a new feature available in NAMD 2.12 and newer.

Download NAMD 2.12 for free here. We recommend the use of the nightly build version as the QM/MM interface is in constant development.

Hybrid QM/MM features are now available in VMD.

Download VMD 1.9.4 alpha version for free here. We recommend using the latest release, the QM/MM interface is in constant development.

How to cite

Available soon.

Introduction to hybrid QM/MM simulations

This brief introduction aims at providing some basic context to the following description of capabilities and commands available in NAMD's QM/MM interface.

- Division of Labor

The basic idea behind a hybrid QM/MM simulation in NAMD is to use a classical force field to treat the

```
graph LR; A[qmElecEmbed  
default ON] --- B[Electrostatic embedding]; A --- C[Mechanical embedding]
```

nature methods

Brief Communication | Published: 26 March 2018

NAMD goes quantum: an integrative suite for hybrid simulations

Marcelo C R Melo, Rafael C Bernardi, Till Rudack, Maximilian Scheurer, Christoph Riplinger, James C Phillips, Julio D C Maia, Gerd B Rocha, João V Ribeiro, John E Stone, Frank Neese, Klaus Schulten & Zaida Luthey-Schulenz

Nature Methods 15, 351–354 (2018) | Download Citation ↓

<http://www.ks.uiuc.edu/~rcbernardi/QMMM/>

<https://www.nature.com/articles/nmeth.4638>

NAMD goes quantum: an integrative suite for hybrid simulations

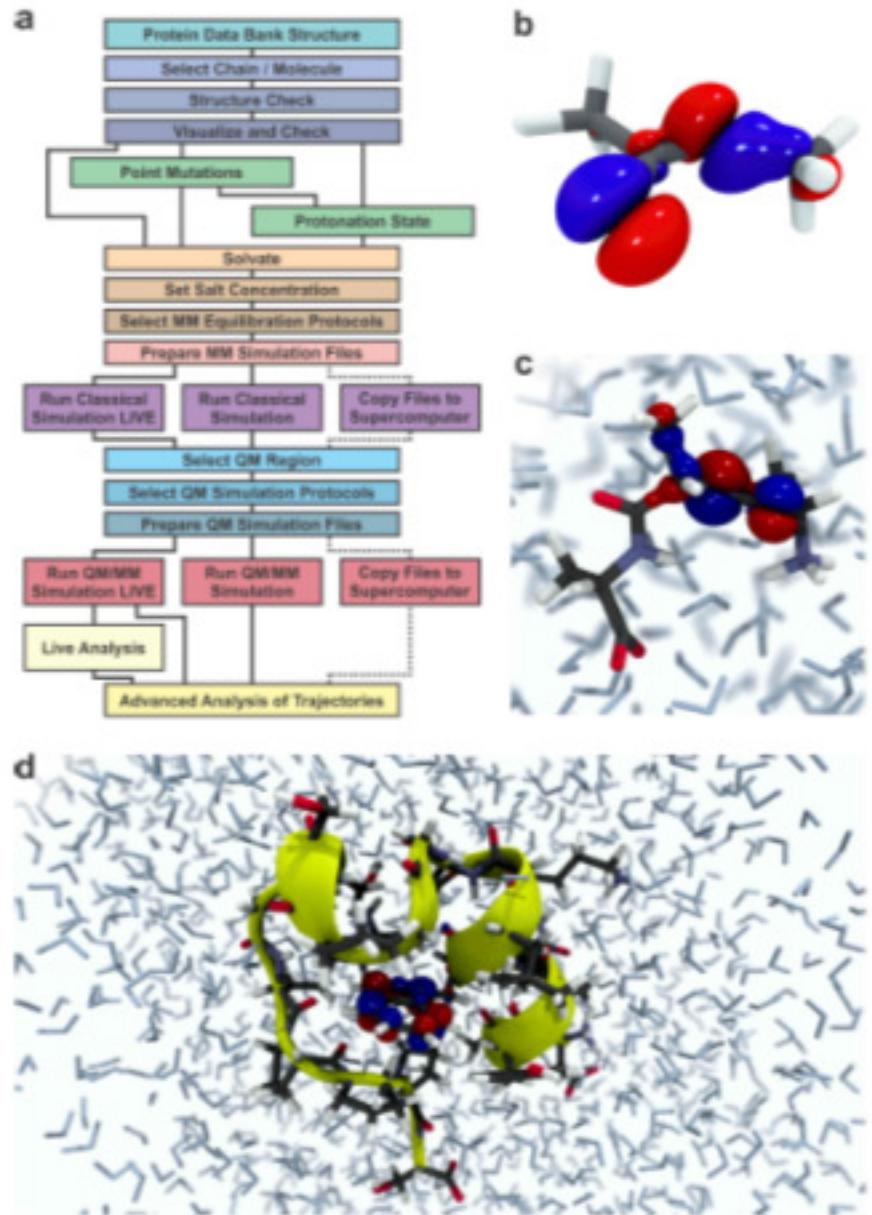
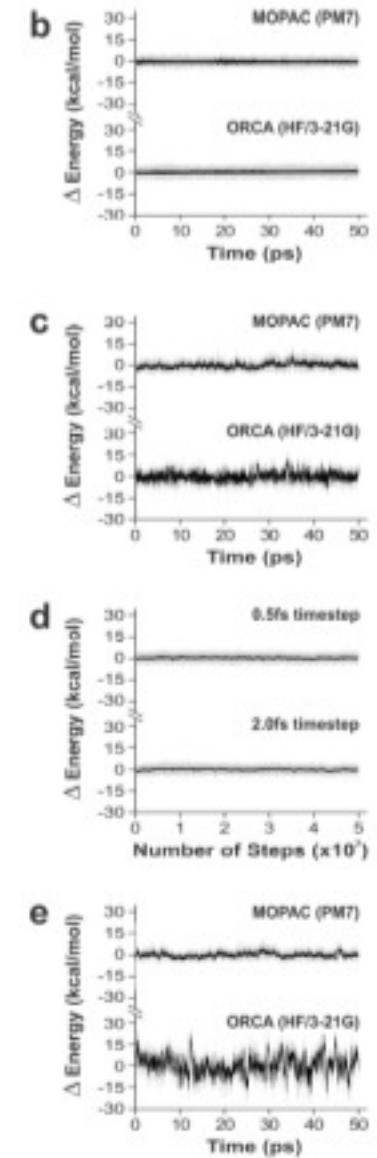
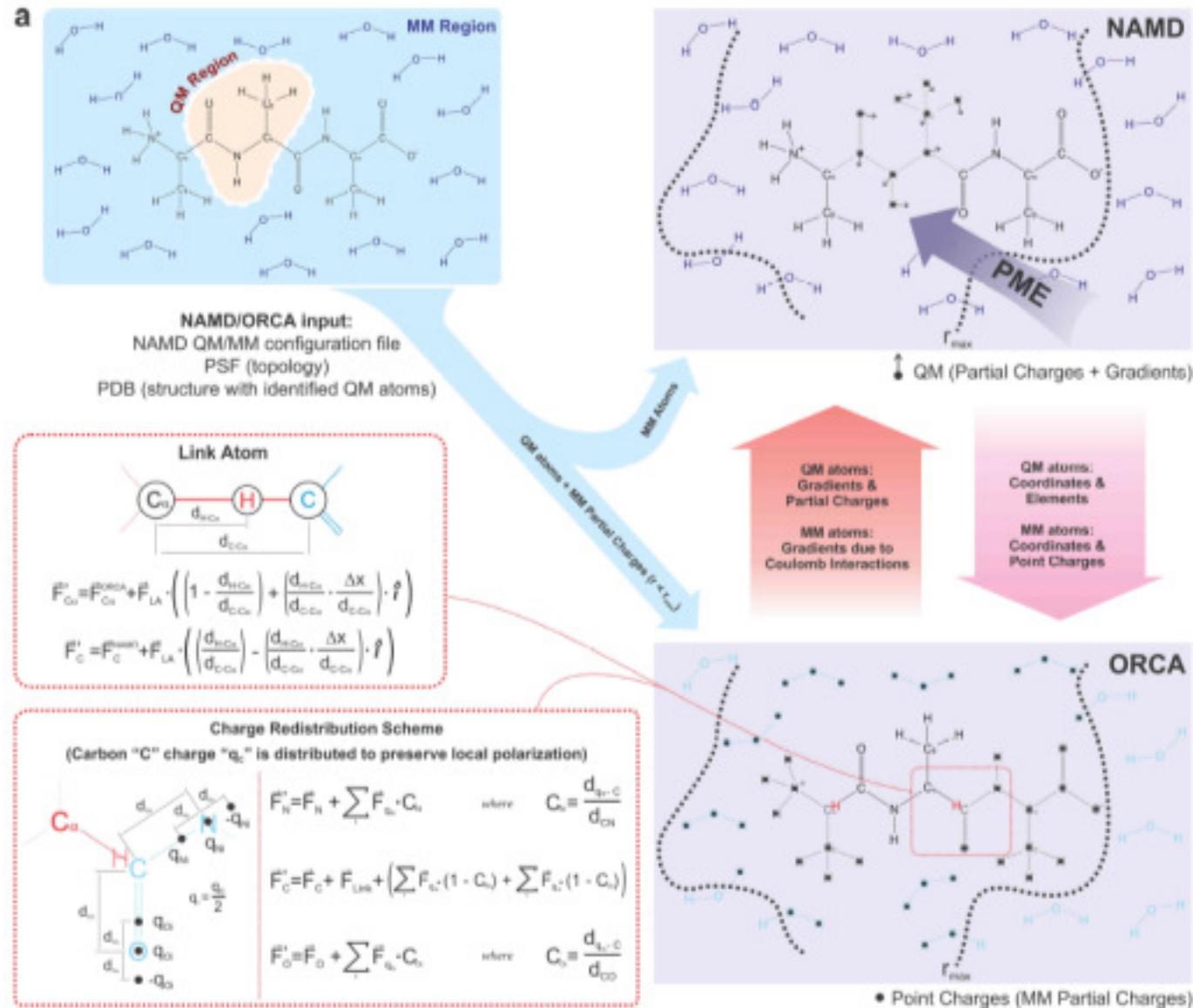


Figure 2: Hybrid QM/MM VMD features.
a) QwikMD provides a graphical user interface (GUI) in VMD for performing QM/MM simulations. The image shows the workflow to prepare, run, analyze, and visualize a hybrid QM/MM molecular dynamics simulation. b) Highest occupied molecular orbital of an alanine molecule in vacuum. c) Alanine's highest occupied molecular orbital in a solvated QM/MM tri-alanine. d) Trp-Cage protein highest occupied molecular orbital in water solution.

NAMD goes quantum: an integrative suite for hybrid simulations



NAMD goes quantum: an integrative suite for hybrid simulations

Supplementary Table 3: Benchmarks

System	Number of atoms		Average Performance (ps/day)					
	Full System	QM Region	PM7	PM7 (GPU)	PM7 (MOZ.)	HF (3-21G)	B3LYP (6-31G*)	
NMA	12	12	9,600	104	10,164	96	29	
NMA + H ₂ O	3648	12	4,800	104	4,800	77	27	
TriAla	33	10+2	9,600	104	10,164	90	30	
TriAla + H ₂ O	4115	10+2	8,228	104	8,228	87	28	
Active Site	263577	213+13	51	21	69	—	—	
TrpCage	8360	284	28	16	64	0.06	—	
BR [†]	80594	3663	—	—	0.89	—	—	
GluRS [‡]	263577	7684	—	—	0.45	—	—	

Benchmarks comparing average time for a NAMD timestep, using different theory levels, software implementations and system sizes. The Full System column indicates the total number of atoms in the simulated system, while the QM Region column indicates the number of atoms sent to the QM code for computation of forces, charges and energy. PM7, PM7 (GPU) and PM7 (MOZ.) columns reflect performance when NAMD used MOPAC with, respectively, multithreaded, multithreaded with GPU, and MOZYME implementations. HF and DFT columns indicate performance while using ORCA, and varying parameters. Both the NMA and TriAla systems were simulated with and without a water box (solvated systems are indicated in the table by "+ H₂O"). The GluRS system (composed of the GluRS:tRNA:Glu-AMP complex), TrpCage and the Bactheriorhodopsin system (which was embedded in a lipid bi-layer) were simulated in a water box. For GluRS, TrpCage and Bactheriorhodopsin systems, all amino acid residues were calculated quantum mechanically. The "Active Site" results represent the catalytic region of the GluRS:tRNA:Glu-AMP complex, containing the adenylate, the A₇₆ residue of the tRNA and several amino acids. The number of atoms in the QM region for the TriAla systems indicates the number of QM atoms in the peptide (only the inner residue) plus the number of link atoms (one per QM/MM bond). Analogously, for the Active Site system, QM Region column indicates the number of atoms form the protein, tRNA and Glu-AMP plus the number of link atoms. [†] Bactheriorhodopsin. [‡] GluRS:tRNA^{Glu} system with the entire protein and ligand in a single QM region.

NAMD goes quantum: an integrative suite for hybrid simulations

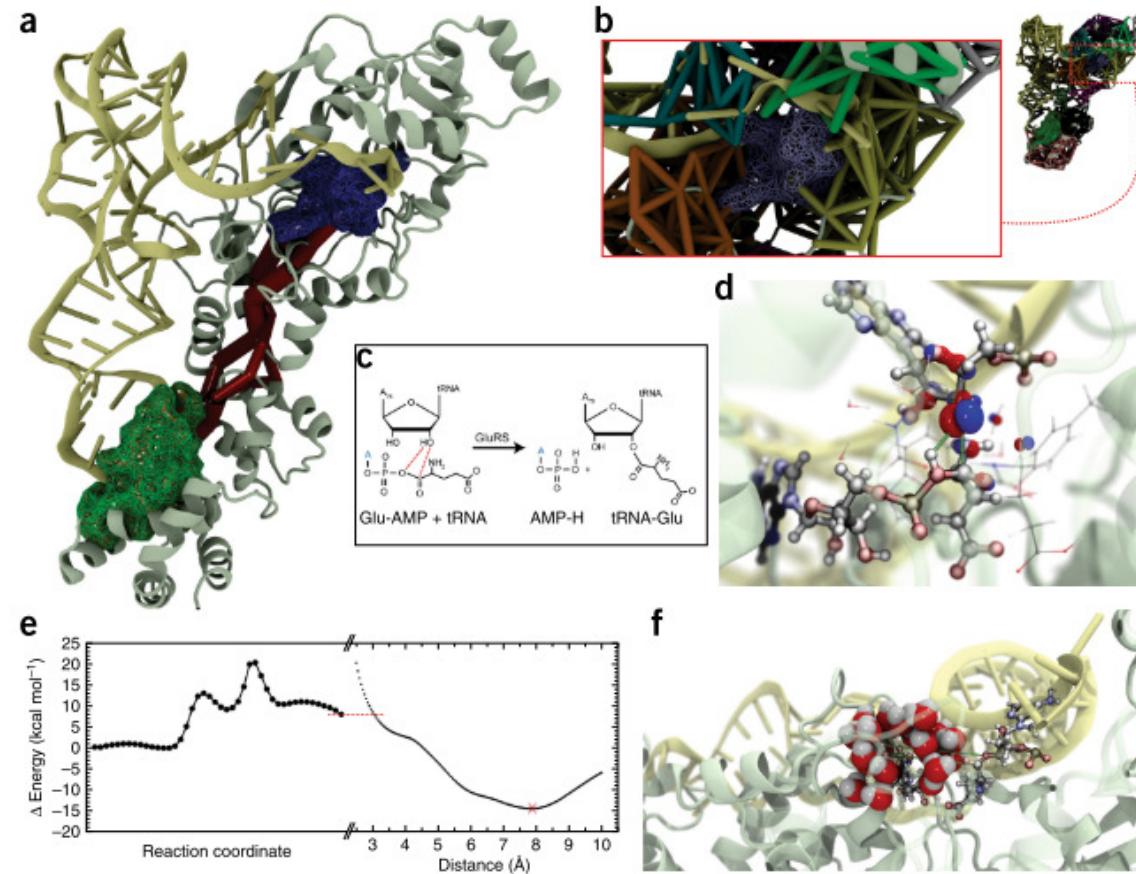


Figure 3 | Mechanism of glutamyl-tRNA synthetase. **(a)** The glutamyl-tRNA synthetase allosteric pathway (red). Two independent QM regions are highlighted, indicating the active site (blue) and the anticodon binding region (green). Yellow and white ribbons represent tRNA and synthetase, respectively. **(b)** Cross-correlation-based community analysis of the active site calculated from QM-MM trajectories. Color-coded as in **a**. **(c)** Reaction mechanism of glutamyl-tRNA synthetase. **(d)** Intermediate state of the glutamyl-tRNA synthetase reaction, showing the highest occupied molecular orbital. **(e)** Left, free energy profile of the most favorable glutamyl-tRNA synthetase reaction mechanism, calculated by eABF after a string-method path optimization. Both eABF and the string method were carried out with QM-MM MD simulations with NAMD-MOPAC and PM7. Right, free energy profile of the distancing and solvation of AMP calculated via ABF and classical MD simulations. “X” indicates the minimum energy state. The red dashed line indicates the connecting point between the eABF and ABF free energy profiles. $n = 1$ experiment each for eABF and ABF. **(f)** Snapshot of the minimum energy state (red X in **e**) during the release of the AMP, showing the solvation of the phosphate group. Color-coding in **d,f** defined as in **Figure 2**, except for gold element structures (phosphatase) and white and yellow ribbons (as in **a**).

NAMD goes quantum: an integrative suite for hybrid simulations

NAMD goes quantum: An integrative suite for QM/MM simulations

Marcelo C. R. Melo¹, Rafael C. Bernardi¹, Till Rudack^{1,2}, Maximilian Scheurer³, Christoph Riplinger⁴, James C. Phillips¹, Julio D. C. Maia⁵, Gerd B. Rocha⁵, João V. Ribeiro¹, John E. Stone¹, Frank Neese⁴, Klaus Schulten¹, Zaida Luthey-Schulten¹

1. University of Illinois at Urbana-Champaign, USA

2. Ruhr-Universität Bochum, Germany

3. Heidelberg University, Germany

4. Max-Planck Institut für Kohlenforschung, Germany

5. Federal University of Paraíba, Brazil

NIH Center for Macromolecular Modeling and Bioinformatics

<https://www.nature.com/articles/nmeth.4638>

Let's do some calculations with
MOPAC

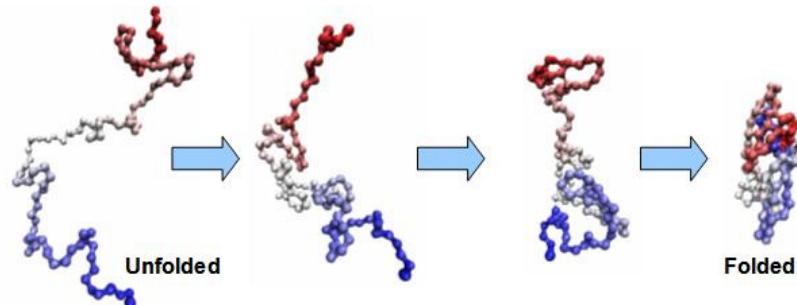
https://github.com/RochaGerd/Chemistry_with_Python

Applications of Quantum Chemistry Methods to Biomolecular systems

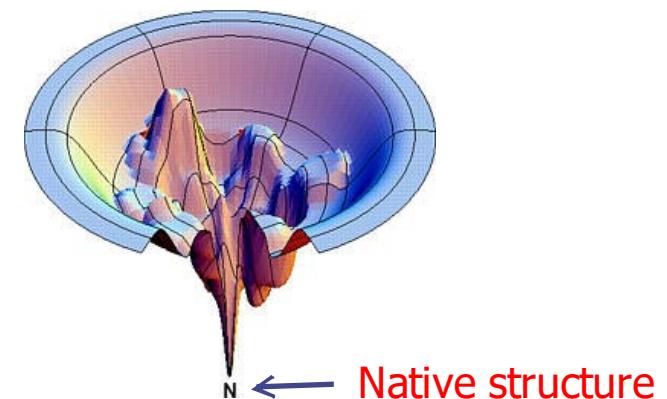
Protein folding Problem

In its most fundamental form, the folding problem can be separated into two distinct parts:

1. To find an efficient and accurate way of sampling the very large conformational space of a protein.
2. To find an energy function that can discriminate accurately between the native and non-native forms of the protein geometry.



Free energy



The accuracy of these functions is determined by their ability to correctly guess the native function from among the many other misfolded conformations.

Behind the folding funnel diagram

Martin Karplus

This Commentary clarifies the meaning of the funnel diagram, which has been widely cited in papers on protein folding. To aid in the analysis of the funnel diagram, this Commentary reviews historical approaches to understanding the mechanism of protein folding. The primary role of free energy in protein folding is discussed, and it is pointed out that the decrease in the configurational entropy as the native state is approached hinders folding, rather than guiding it. Diagrams are introduced that provide a less ambiguous representation of the factors governing the protein folding reaction than the funnel diagram.

New results on the Protein Folding problem

 **Zhang Lab**

Home Research Services Publications People Teaching Job Opening Labonly

Online Services

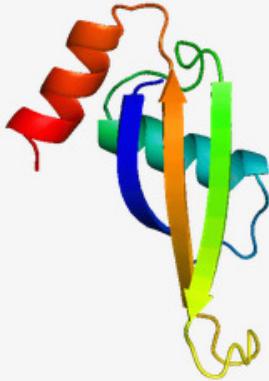
- I-TASSER
- QUARK
- LOMETS
- COACH
- COFACTOR
- MUSTER
- SEGMER
- FG-MD
- ModRefiner
- REMO
- SPRING
- COTH
- BSpred
- SVMSEQ
- ANGLOR
- BSP-SLIM
- SAXSTER

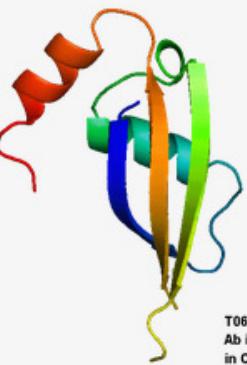
Yang Zhang's Research Group

[Department of Computational Medicine and Bioinformatics](#)
[Department of Biological Chemistry](#)
[Medical School](#)
[University of Michigan](#)
100 Washtenaw Avenue
Ann Arbor, MI 48109-2218
Phone: (734)647-1549
Fax: (734)615-6553
Email: yangzhanglab@umich.edu

We are interested in:

- [Protein Structure Prediction](#)
- [Structure-based Function Annotation](#)
- [Protein Design](#)
- [G Protein-Coupled Receptor](#)
- [Amyloid Diseases and Fiber Aggregation](#)
- [Protein-Protein Interactions](#)
- [RNA Alternative Splicing](#)
- [Ligand Screening and Structure-Based Drug Design](#)


Predicted model (2.66Å)


X-ray structure
T0604_1
Ab initio folding
in CASP9

We selected 33 decoy sets from Zhang Lab
I-TASSER algorithm

New results on the Protein Folding problem

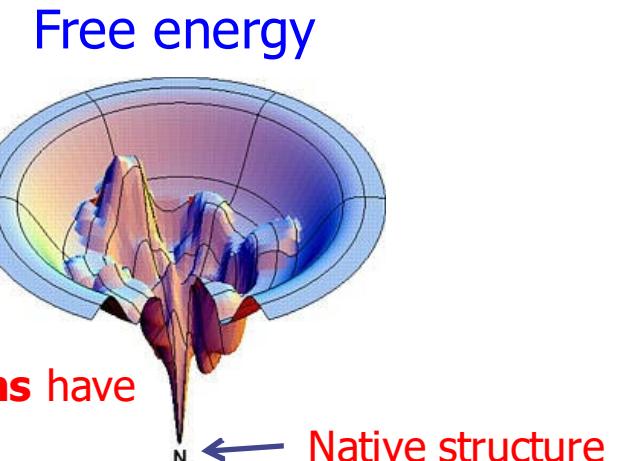
Table 1. Characterization of the decoy sets used in the study.

Set index	PDB	Description	Type	Size
1	1ABV	F1F0-ATP Synthase delta subunit n-term	NMR	528
2	1AF7	Chemotaxis receptor CheR	X-Ray	528
3	1AH9	IF1 translational initiation factor	NMR	511
4	1AOY	Arginine repressor n-term	NMR	530
5	1B4B	Arginine repressor oligomerization domain	X-Ray	461
6	1BM8	Mbp1 DNA-Binding domain	X-Ray	330
7	1CEW	Chicken egg white cystatin	X-Ray	453
8	1CQK	Antibody domain CH3	X-Ray	285
9	1DCJ	<i>E. coli</i> YHHP	NMR	526
10	1EGX	ENA-VASP Homology 1 domain	NMR	353
11	1FAD	FADD death domain	NMR	515
12	1FO5	reduced MJ0307	NMR	342
13	1G1C	I1 Domain from titin	X-Ray	308
14	1GJX	Lipoyl Domain from P64K	NMR	526
15	1GPT	Gamma 1-H Thionin	NMR	470
16	1ITP	<i>P. ostreatus</i> proteinase A inhibitor 1	NMR	527
17	1KJS	C5a C-terminus	NMR	549
18	1KVI	Reduced form of ATP7A	NMR	551
19	1MLA	Acyl carrier protein transacylase	X-Ray	336
20	1N0U	Yeast elongation factor 2	X-Ray	302
21	1NE3	Ribosomal protein S28E	NMR	566
22	1OF9	<i>E. histolytica</i> pore forming toxin	NMR	508
23	1R69	Phage 434 repressor N-term domain	X-Ray	292
24	1SHF	Human Fyn SH3 domain	X-Ray	537
25	1SRO	S1 RNA binding domain	NMR	516
26	1TFI	TFIIS nuclei acid-binding domain	NMR	340
27	1TIF	Translational factor IF3 N-term domain	X-Ray	543
28	1TIG	Translational factor IF3 C-term domain	X-Ray	565
29	1VCC	N-terminal fragment of vaccinia virus DNA topoisomerase I	X-Ray	552
30	256B	Cytochrome b562	X-Ray	507
31	2CR7	PAH domain of the mouse transcriptional repressor SIN3B	NMR	541
32	2F3N	Native Shank SAM domain	X-Ray	486
33	2PCY	Poplar apoplastocyanin	X-Ray	436

New results on the Protein Folding problem

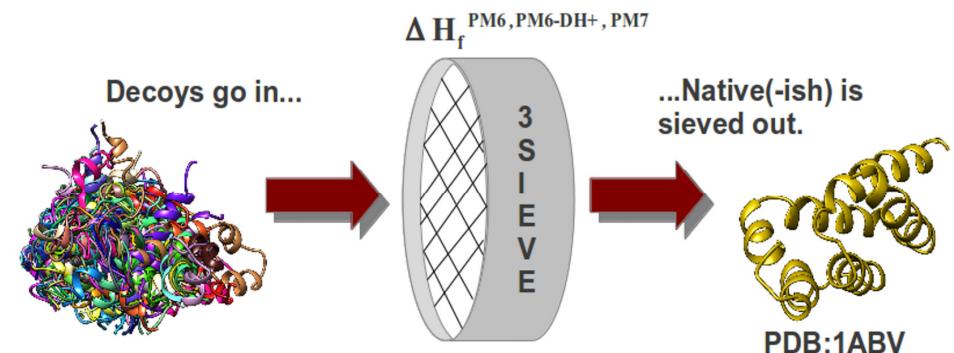
What we think about the problem.

- The native state must be located at a very deep minimum in the energy hypersurface and thus there is a **significant energy gap** between a native structure and an average non-native structure.
- This feature is expected in a good score function.**
- We assumed that a **random ensemble of conformations** have a **Gaussian distribution of enthalpies**.
- Certain conformations have extremely different enthalpy values. For these cases, we assumed that these specific conformations are outliers that differ from the population.



3SIEVE Algorithm

3Sigma Iterative
Enthalpy native-likeness Exposer



3SIEVE Algorithm

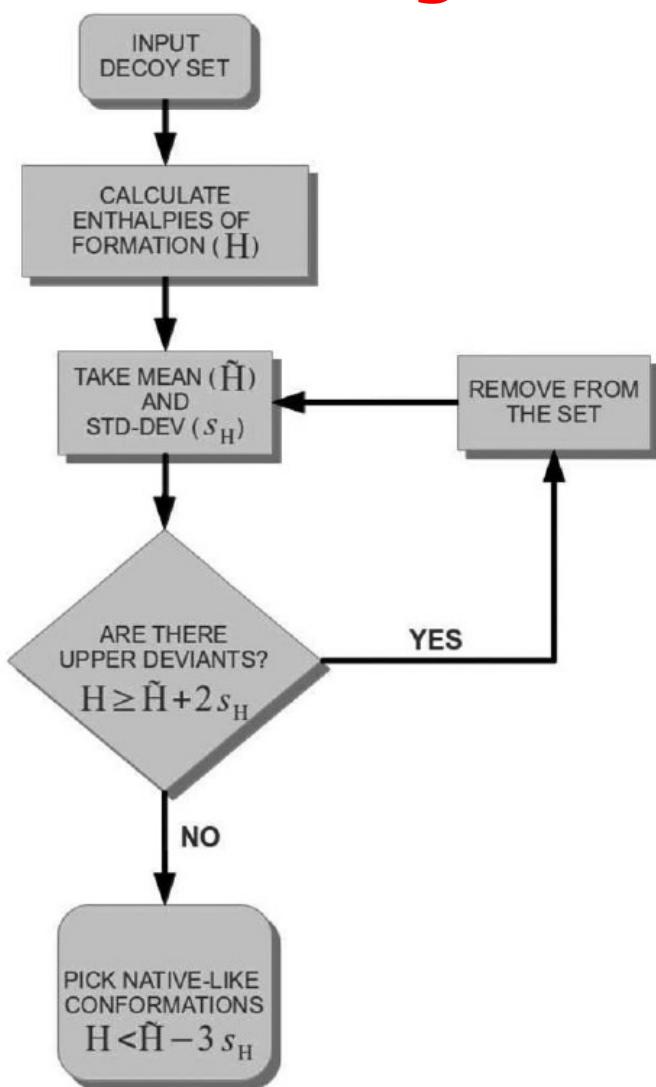


Figure 1. A schematic representation of the 3SIEVE algorithm.

Urquiza-Carvalho GA, Fragoso WD, Rocha GB, J Comput Chem. 2016, 37(21):1962-72, doi: 10.1002/jcc.24415.

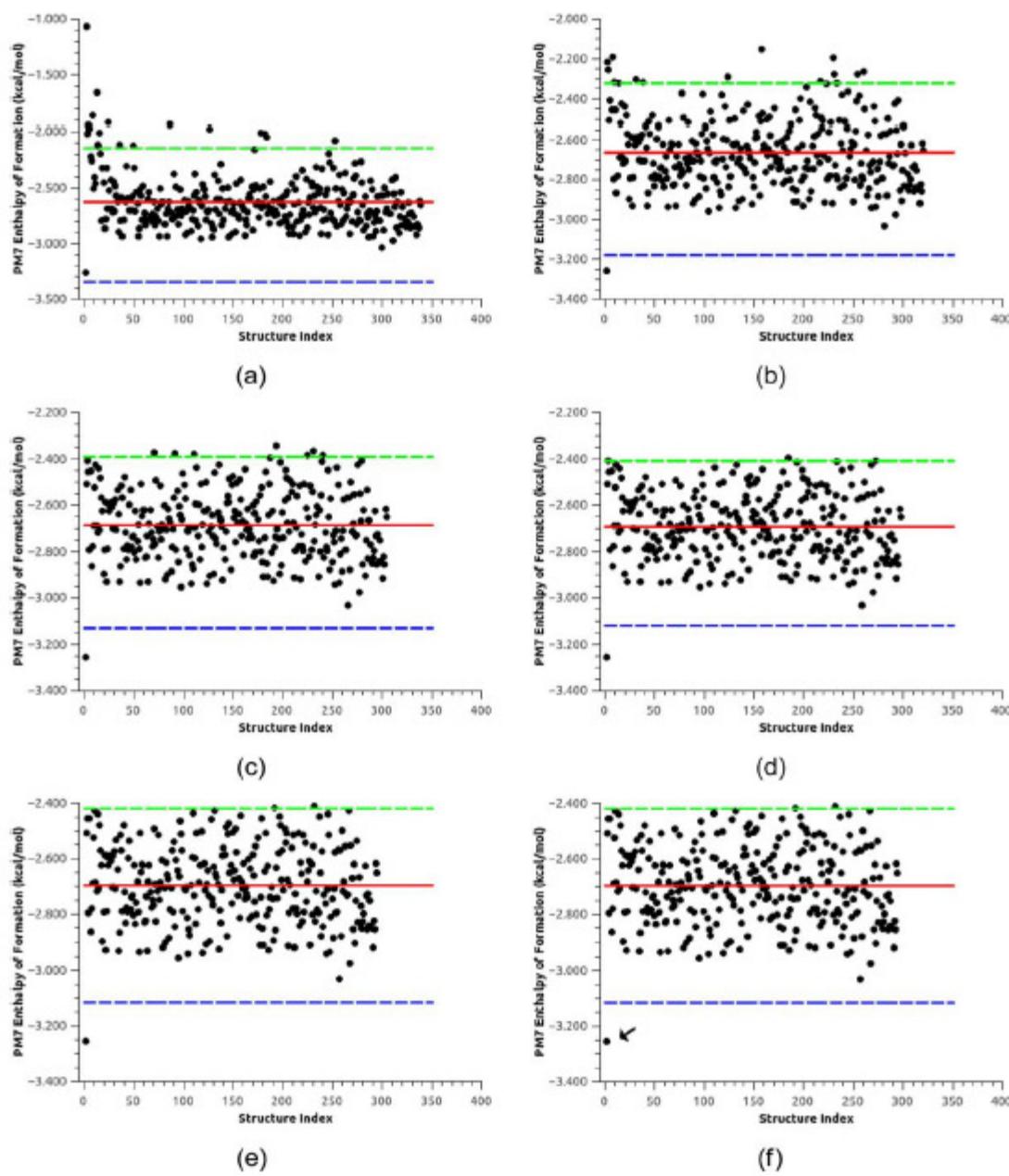


Figure 2. The iterative process illustrated by example with the 1TFI protein decoy set using the PM7 Hamiltonian. The abscissae of the graphs represent the index of the decoy in the set and the ordinates represent the heats of formation in kcal/mol. The dashed lines above and below represent the upward 2 and downward 3 standard deviation lines, respectively, and the solid line represents the mean of the set. The arrow in the last figure points to a native-like structure. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

New results on the Protein Folding problem

PM7 or PM6 or PM6-D3H4 or PM6-DH+ 1SCF COSMO

- 33 proteins
- Each containing about 450 decoys

Totalizing ~ 61k QM calculations

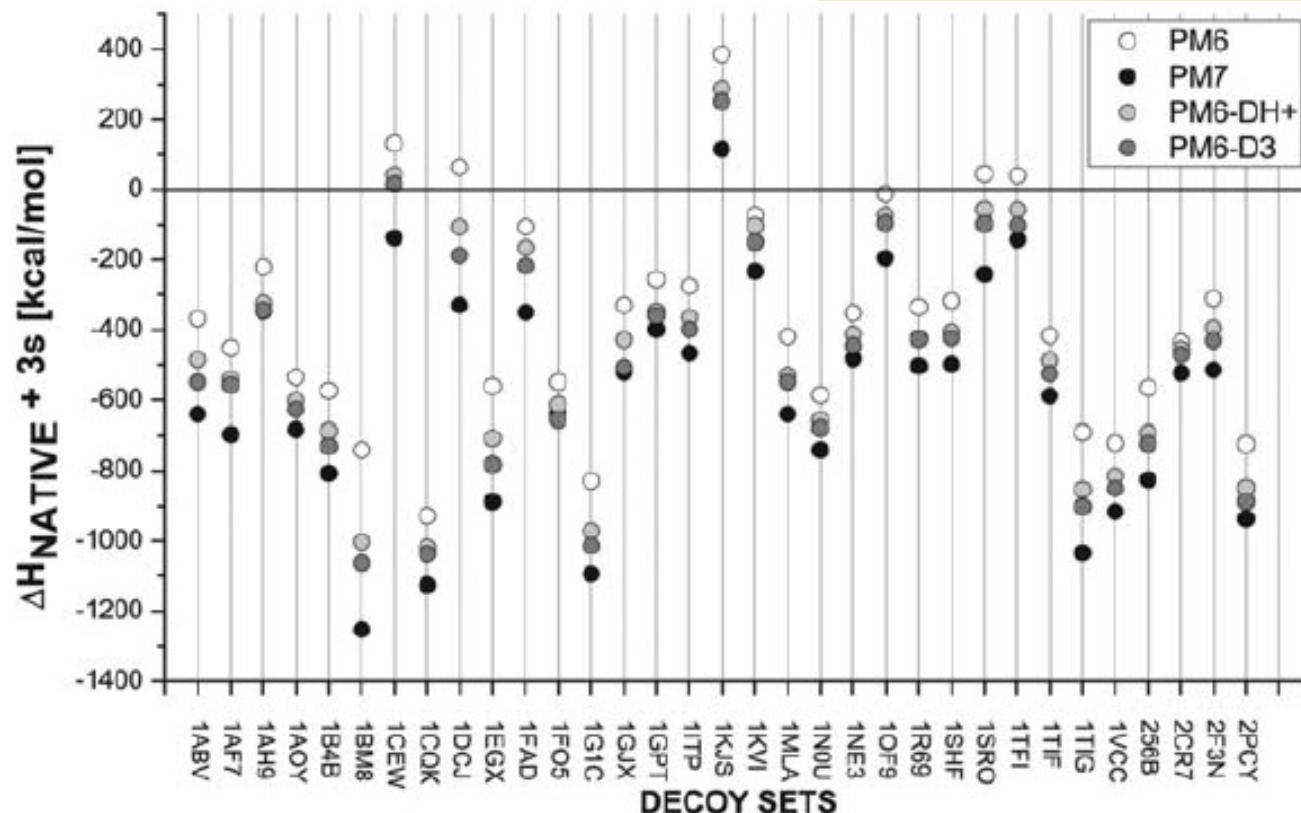


Figure 3. The distance between the native structure enthalpy and the $-3s$ line in each decoy set, for each method.

New results on the Protein Folding problem

FULL PAPER

WWW.C-CHEM.ORG

Journal of
**COMPUTATIONAL
CHEMISTRY**

Assessment of Semiempirical Enthalpy of Formation in Solution as an Effective Energy Function to Discriminate Native-like Structures in Protein Decoy Sets

Gabriel Aires Urquiza-Carvalho, Wallace Duarte Fragoso, and Gerd Bruno Rocha*

In this work, we tested the PM6, PM6-DH+, PM6-D3, and PM7 enthalpies of formation in aqueous solution as scoring functions across 33 decoy sets to discriminate native structures or good models in a decoy set. In each set these semiempirical quantum chemistry methods were compared according to enthalpic and geometric criteria. Enthalpically, we compared the methods according to how much lower was the enthalpy of each native, when compared with the mean enthalpy of its set. Geometrically, we compared the methods according to the fraction of native contacts (Q), which is a measure of geometric closeness between an arbitrary structure and the native. For each set and method,

the Q of the best decoy was compared with the Q_0 , which is the Q of the decoy closest to the native in the set. It was shown that the PM7 method is able to assign larger energy differences between the native structure and the decoys in a set, arguably because of a better description of dispersion interactions, however PM6-DH+ was slightly better than the rest at selecting geometrically good models in the absence of a native structure in the set. © 2016 The Authors. Journal of Computational Chemistry Published by Wiley Periodicals, Inc.

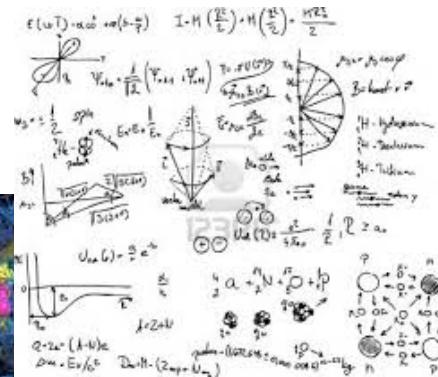
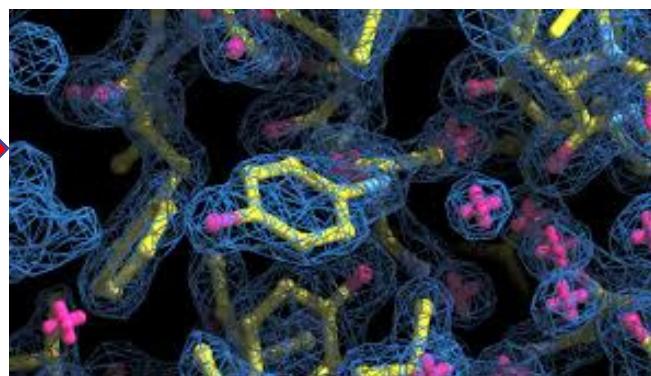
DOI: 10.1002/jcc.24415

Semiempirical calculations of Biomolecules

What are we able to do with a wavefunction
and the electron density spread over millions
of atoms?

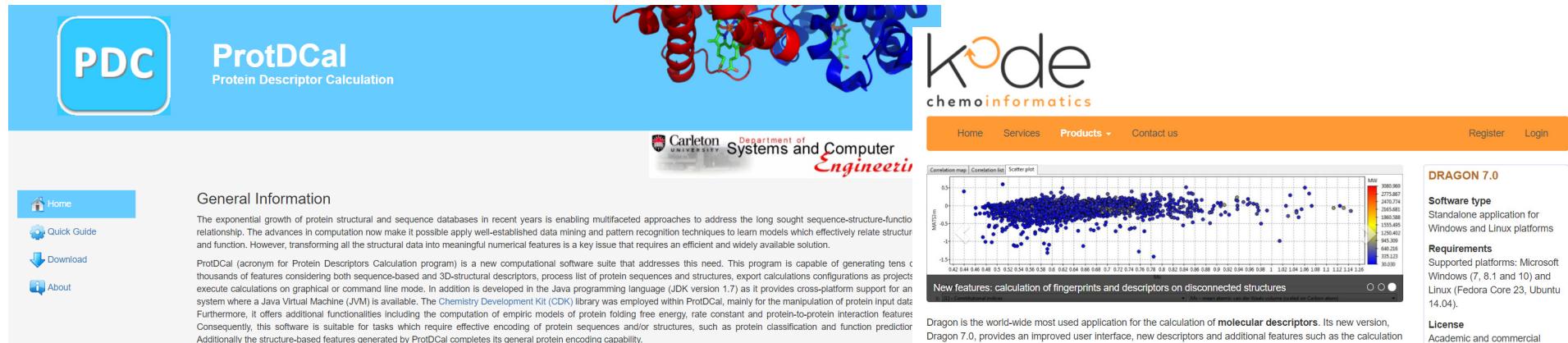
Molecular Descriptors for Biomolecules

In order to substantially contribute in areas such as biochemistry, molecular biology, or structural biology, a quantum-chemistry program should provide to these researches (non-specialists in quantum chemistry) the possibility of retrieving, in an agile and visual form, electron density descriptors that can carry information about the reactivity of biological systems and contribute to the characterization of important events in biology, such as: protein folding, electron transfer, signalizing, biochemical reactions, etc.

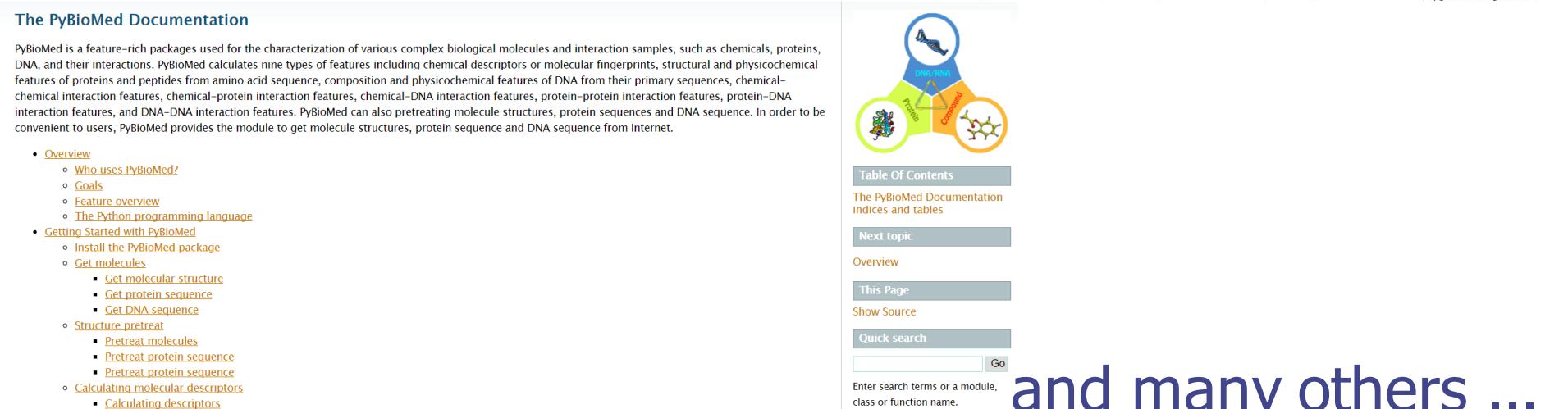


Molecular Descriptors for biomolecules

Molecular descriptors for biological systems are largely calculated using chemoinformatics and bioinformatics tools.



The screenshot shows the ProtDCal software interface. At the top left is the PDC logo and the text "ProtDCal Protein Descriptor Calculation". In the center is a 3D ribbon diagram of a protein molecule. To the right is the Kode chemoinformatics logo. Below these are navigation links: Home, Services, Products, Contact us, Register, and Login. On the left, there's a sidebar with links for Home, Quick Guide, Download, and About. The main content area has a section titled "General Information" with text about the exponential growth of protein databases and the capabilities of ProtDCal. It also mentions the Java Virtual Machine (JVM) and the Chemistry Development Kit (CDK) library.



The screenshot shows the PyBioMed Documentation website. At the top is a navigation bar with icons for Home, Services, Products, Contact us, Register, and Login. The main content area features a large scatter plot titled "Correlation map | Correlation list | Scatter plot" with a color scale for MW (Molecular Weight). Below the plot is a section titled "New features: calculation of fingerprints and descriptors on disconnected structures". To the right is a sidebar for "DRAGON 7.0" with sections for Software type (standalone application for Windows and Linux), Requirements (supported platforms: Microsoft Windows 7, 8.1 and 10, and Linux Fedora Core 23, Ubuntu 14.04), License (academic and commercial licenses), and Pricing (standard academic license fee is 900 €). The PyBioMed documentation page itself has a circular diagram at the top with segments for "DNA", "Protein", and "Chemical". Below it is a "Table of Contents" section with links to "The PyBioMed Documentation Indices and tables", "Next topic", "Overview", "This Page", "Show Source", and a "Quick search" bar. The search bar placeholder is "Enter search terms or a module, class or function name".

The PyBioMed Documentation

PyBioMed is a feature-rich package used for the characterization of various complex biological molecules and interaction samples, such as chemicals, proteins, DNA, and their interactions. PyBioMed calculates nine types of features including chemical descriptors or molecular fingerprints, structural and physicochemical features of proteins and peptides from amino acid sequence, composition and physicochemical features of DNA from their primary sequences, chemical-chemical interaction features, chemical-protein interaction features, chemical-DNA interaction features, protein-protein interaction features, protein-DNA interaction features, and DNA-DNA interaction features. PyBioMed can also pretreat molecule structures, protein sequences and DNA sequence. In order to be convenient to users, PyBioMed provides the module to get molecule structures, protein sequence and DNA sequence from Internet.

- [Overview](#)
 - [Who uses PyBioMed?](#)
 - [Goals](#)
 - [Feature overview](#)
 - [The Python programming language](#)
- [Getting Started with PyBioMed](#)
 - [Install the PyBioMed package](#)
 - [Get molecules](#)
 - [Get molecular structure](#)
 - [Get protein sequence](#)
 - [Get DNA sequence](#)
 - [Structure pretreat](#)
 - [Pretreat molecules](#)
 - [Pretreat protein sequence](#)
 - [Pretreat protein sequence](#)
 - [Calculating molecular descriptors](#)
 - [Calculating descriptors](#)

and many others ...

Molecular Descriptors for biomolecules

Explore omic applications



GENOMICS

Decipher your DNA sequences from whole genome sequencing, WES, de novo sequencing, SNP array, Rep-seq, explore metagenomics, synthetic biology, genome edition, phylogenomics or browse databases for genome annotation and more.



EPIGENOMICS

Study DNA modifications from BS-seq and DNA methylation arrays, and determine DNA-protein interactions from ChIP-seq, ChIP-on-chip, or Hi-C experiments and more.



TRANSCRIPTOMICS

Unravel your transcriptome from RNA-seq, gene expression array, or qPCR, detect low-level RNA with single cell RNA-seq, RNA-protein interactions with CLIP-seq and Ribo-seq, discover non-coding RNA from small RNA-seq, determine RNA structure and browse databases for ncRNA, RNA modification and transcription, and more.



PROTEOMICS

Determine protein sequences and expression from MS-based untargeted or targeted proteomics, analyze post-translational modifications, protein structure and NMR-based proteomics, study the immune system, explore flow-cytometry data, browse databases for sequences, PTM, structure, or immunology and more.



METABOLOMICS

Find the best tools to analyze metabolites data from MS-based untargeted or targeted metabolomics, NMR-based metabolomics, discovery and design of drugs, and browse chemoinformatics databases and more.



PHENOMICS

Characterize your phenotypes with laser scanning and cryo-electron microscopy, mass-spectrometry and SPIM mesoscopy, and magnetic resonance imaging or nuclear medicine imaging macroscopy and more.



Visual Guidance

SIB resources

External resources - (No support from the ExPASy Team)

Categories

proteomics

- protein sequences and identification
- proteomics experiment
- function analysis
- sequence sites, features and motifs
- protein modifications
- protein structure
- protein interactions
- similarity search/alignment

genomics

- structure analysis
- systems biology
- evolutionary biology
- population genetics
- transcriptomics
- biophysics
- imaging
- IT infrastructure
- medicinal chemistry
- glycomics

Databases

- UniProtKB • functional information on proteins • [more]
- UniProtKB/Swiss-Prot • protein sequence database • [more]
- STRING • protein-protein interactions • [more]
- SWISS-MODEL Repository • protein structure homology models • [more]
- PROSITE • protein domains and families • [more]
- ViralZone • portal to viral UniProtKB entries • [more]
- neXtProt • human proteins • [more]

Tools

- SWISS-MODEL Workspace • structure homology-modeling • [more]
- Vital-IT • life science informatics initiative • [more]
- SwissDock • protein ligand docking server • [more]

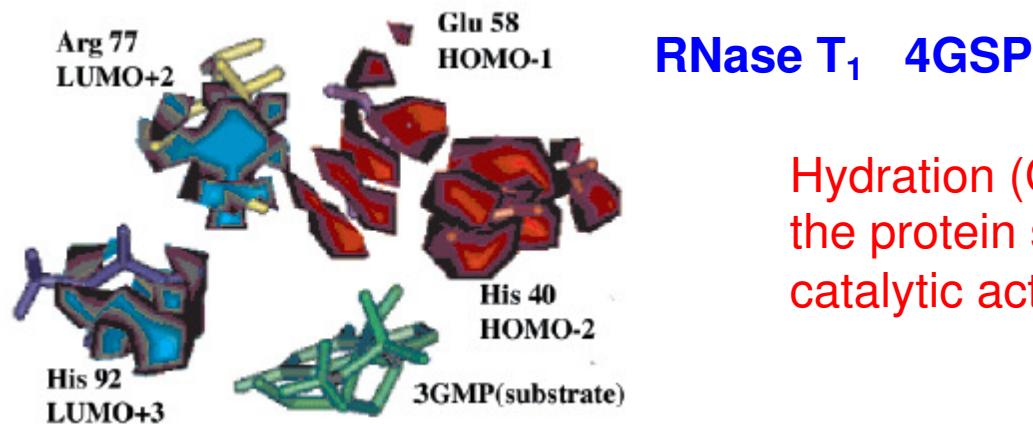
- 2ZIP • Prediction of leucine zipper domains • [more]
- 3of5 • find user-defined patterns in protein sequences • [more]
- AACompIdent • protein identification by aa composition • [more]
- AACompSim • amino acid composition comparison • [more]
- Agadir • Prediction of the helical content of peptides • [more]
- ALF • simulation of genome evolution • [more]
- Alignment tools • Four tools for multiple alignments • [more]
- APSSP • Advanced Protein Secondary Structure Prediction • [more]
- Ascalaph • Molecular modeling software • [more]
- big-PI • predict GPI modification sites • [more]
- Biochemical Pathways • Biochemical Pathways • [more]
- BLAST • sequence similarity search • [more]
- BLAST (UniProt) • BLAST search on the UniProt web site • [more]
- BLAST - NCBI • Biological sequence similarity search • [more]
- BLAST - PBLI • BLAST search on protein sequence databases • [more]
- Blast2Fasta • Blast to Fasta conversion • [more]
- boxshade • MSA pretty printer • [more]
- CRESSP • Protein secondary structure prediction • [more]



QM Molecular Descriptors for biomolecules

Frontier molecular orbitals are located in the active site

Effects of Hydration on the Electronic Structure of an Enzyme: Implications for the Catalytic Function, Kazuki Ohno, Narutoshi Kamiya,[†] Naoki Asakawa, Yoshio Inoue, and Minoru Sakurai*, J. Am. Chem. Soc. **2001**, *123*, 8161.



Hydration (COSMO) not only stabilizes the protein structure but also affects the catalytic action of the enzyme.

Table 1. Molecular Orbitals Localized on the Functionally Important Residues of RNase T₁

	His 40	Glu 58	Arg 77	His 92
in vacuum	HOMO-18	HOMO-10	LUMO+6	LUMO+3
in water	HOMO-2	HOMO-1	LUMO+2	LUMO+3

QM Molecular Descriptors for biomolecules



An insight into the general relationship between the three dimensional structures of enzymes and their electronic wave functions: Implication for the prediction of functional sites of enzymes

K. Fukushima,¹ M. Wada,² and M. Sakurai^{1*}

Table I
Distribution of Enzymes Among the Six EC Classes

This study		PDB	
No. of structures ^a	Ratio (%)	No. of structures ^b	Ratio (%)
19	16.9	3609	17.9
21	18.8	5346	26.6
49	43.6	8233	40.9
12	10.7	1409	7.0
8	8.0	891	4.4
3	2.7	625	3.1

^aThe data for the 112 enzymes studied here.

^bThe data for all the enzymes deposited in the PDB.

- Wavefunction calculation of 112 non-homologous and non-redundant enzymes from PDB database.

- These enzymes can be classified as:

- 1) Extremely localized Molecular Orbitals → Enzyme specificity
- 2) Highly delocalized Molecular Orbitals
- 3) Middle size Molecular Orbitals

These findings are only possible when electronic structure methods are applied to the problem.



Universidade Federal da Paraíba

Departamento de Química

Laboratório de Química Quântica Computacional
Prof. Gerd Bruno Rocha

(gbr@quimica.ufpb.br, gerd.rocha@gmail.com, gbr@acadêmico.ufpb.br)

www.quantum-chem.pro.br

Applications of Reactivity Descriptors for post-processing QM information of Biological Systems



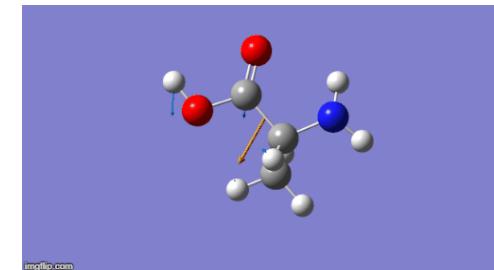
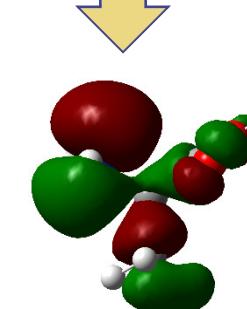
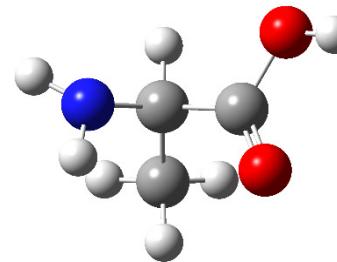
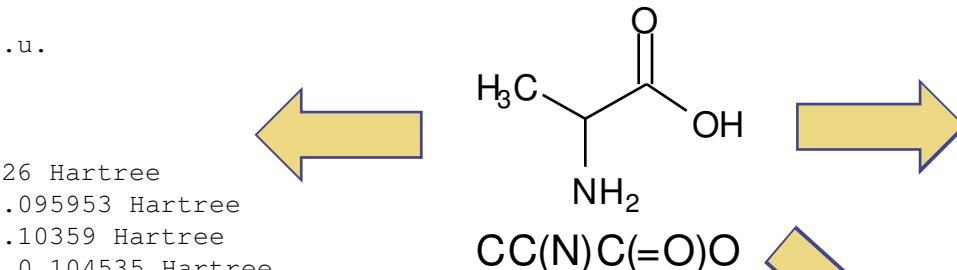
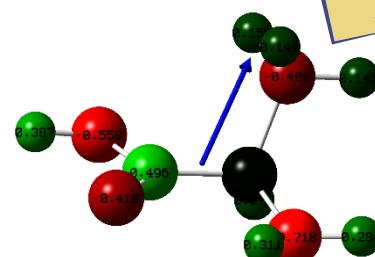
Practical QC calculations

- When I calculate a small molecule, what information do I get?

Dipole Moment = 1.8497674 Debye
Polarizability (?) = 42.452657 a.u.
Point Group =

Thermo Tab Data Section:

Electronic Energy (EE) = -323.6326 Hartree
Zero-point Energy Correction = 0.095953 Hartree
Thermal Correction to Energy = 0.10359 Hartree
Thermal Correction to Enthalpy = 0.104535 Hartree
Thermal Correction to Free Energy = 0.064588 Hartree
EE + Zero-point Energy = -323.53665 Hartree
EE + Thermal Energy Correction = -323.52901 Hartree
EE + Thermal Enthalpy Correction = -323.52806 Hartree
EE + Thermal Free Energy Correction = -323.56801 Hartree
E (Thermal) = 65.004 kcal/mol
Heat Capacity (C_v) = 27.006 cal/mol-kelvin
Entropy (S) = 84.074 cal/mol-kelvin



- From this information I am able to obtain some important chemical concepts.

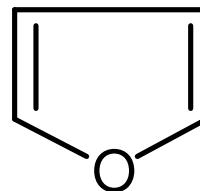
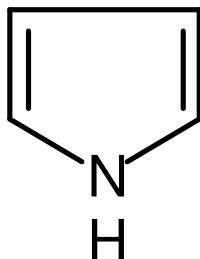
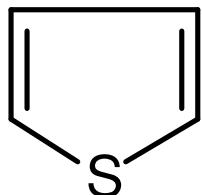
Some concepts in chemistry

As a general rule, chemical concepts are abstract:

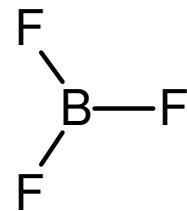
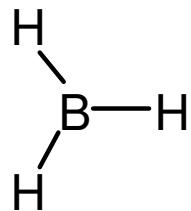
- Covalence;
- Hardness;
- Softness;
- Aromaticity;
- Electrophilicity;
- nucleophilicity, etc.

Some concepts in chemistry

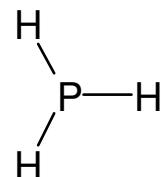
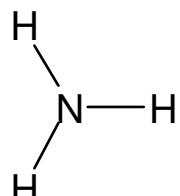
- Which heterocycle is the most aromatic?



- Which acid is the strongest?

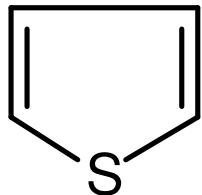


- Which base is the strongest?

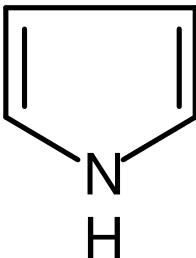


Some concepts in chemistry

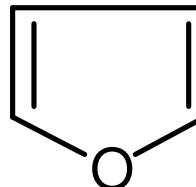
- Which heterocycle is the most aromatic?



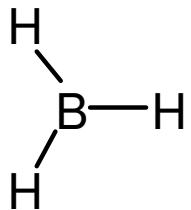
>



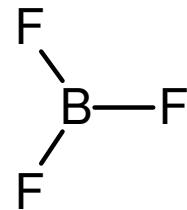
>



- Which acid is the strongest?

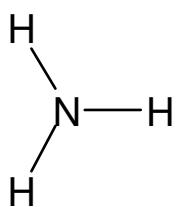


<

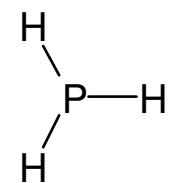


- With the chemical knowledge gained from many experiments, the chemists know how to answer these questions easily without having to do any calculations.

- Which base is the strongest?



>



- But is there a way to answer these questions using quantitative results?

Pearson theory and the Hard-Soft Acid-Base principle

- Chemical Softness is related to the reactivity of the system and the hardness to its stability;
- Soft-soft interactions: large electronic delocalization; polarization; orbital-controlled; formation of adducts; covalent character;
- Hard-hard interactions: Electrostatic attraction; long range; charge-controlled; ion formation;

Conceptual DFT

- Concepts of chemical reactivity from the variation of electronic energy

$$dE = \mu dN + \int \rho(r) d\nu(r) dr$$

$$\mu = \left(\frac{\partial E}{\partial N} \right)_v$$

Electon Chemical potential:

Electron transfer trend in an equilibrium system

$$\eta = 1/S = \left(\frac{\partial^2 E}{\partial N^2} \right)_v$$

Chemical hardness/softness:

They describe the resistance / ease in transferring electrons, polarizability and stability

$$f(r) = \left(\frac{\partial \rho(r)}{\partial N} \right)_v$$

Fukui function:

Variation of electronic density with variation in the number of electrons (Ex .: electrophilic, nucleophilic and radical attacks)

CDFT

Global and local descriptors

- Principle of Hard/Soft Acid-Base, and Frontier Molecular Orbital

Chemical Reactivity

$$\Delta E = \mu \Delta N + \frac{1}{2} \eta \Delta N^2$$

Stabilization energy

$$\omega = \frac{\mu^2}{2\eta}$$

Electrophilicity

$$\Delta N = -\frac{\mu_B^0 - \mu_A^0}{2(\eta_A^0 + \eta_B^0)}$$

Electron transfer from B to A

Regioselectivity

$$f^+(\mathbf{r}) = \left(\frac{\partial \rho(\mathbf{r})}{\partial N} \right)_{\nu(\mathbf{r})}^+$$

Electrophilic
attack

$$f^-(\mathbf{r}) = \left(\frac{\partial \rho(\mathbf{r})}{\partial N} \right)_{\nu(\mathbf{r})}^-$$

Nucleophilic
Attack

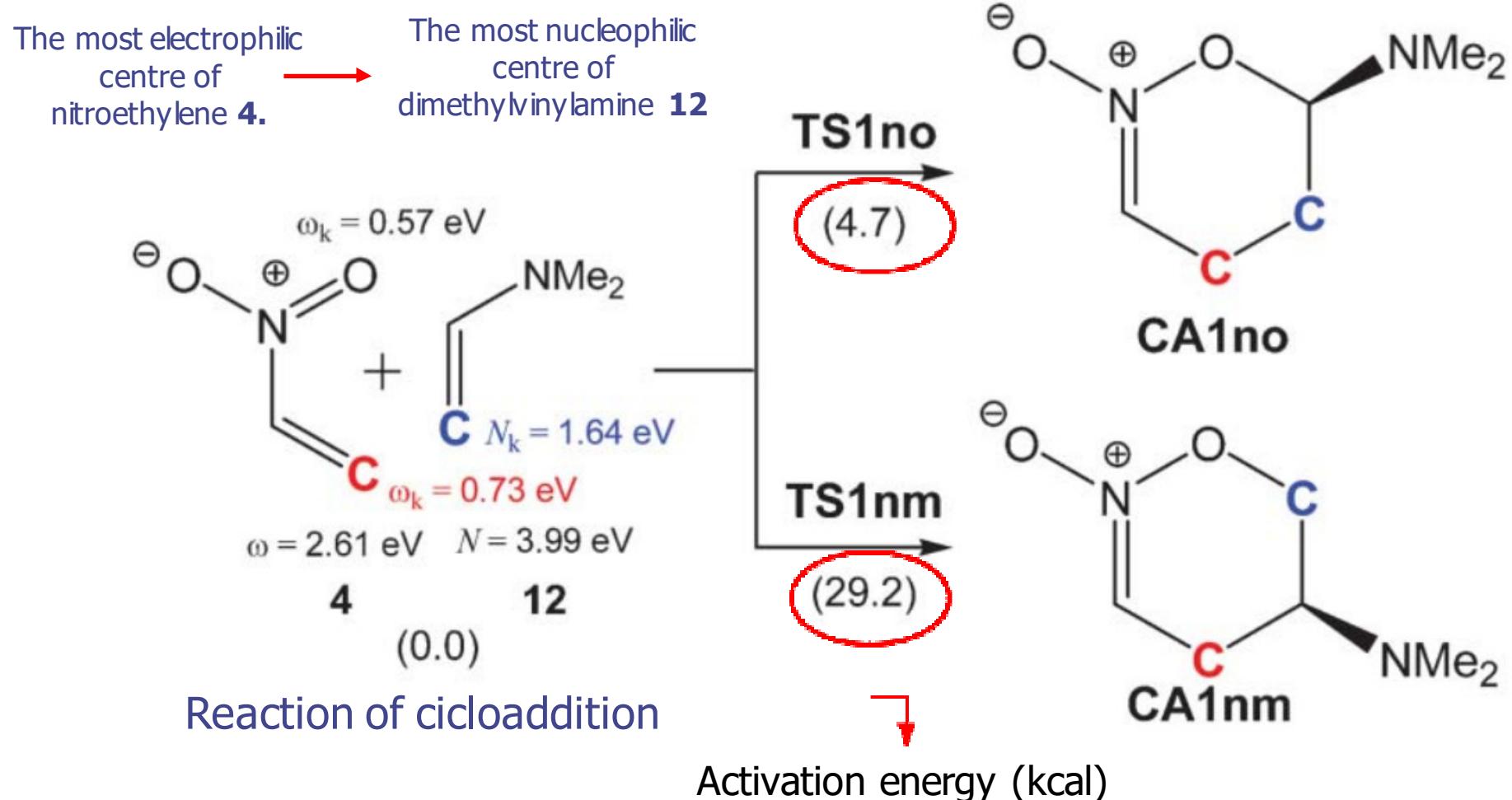
Condensed Fukui functions

$$f_A^+ = q_A(N+1) - q_A(N) \equiv q_{A,N_0+1} - q_{A,N_0}$$

$$f_A^- = q_A(N) - q_A(N-1) \equiv q_{A,N_0+1} - q_{A,N_0-1}$$

CDFT applications in organic chemistry

Regioselectivity in Diels-Alder reactions



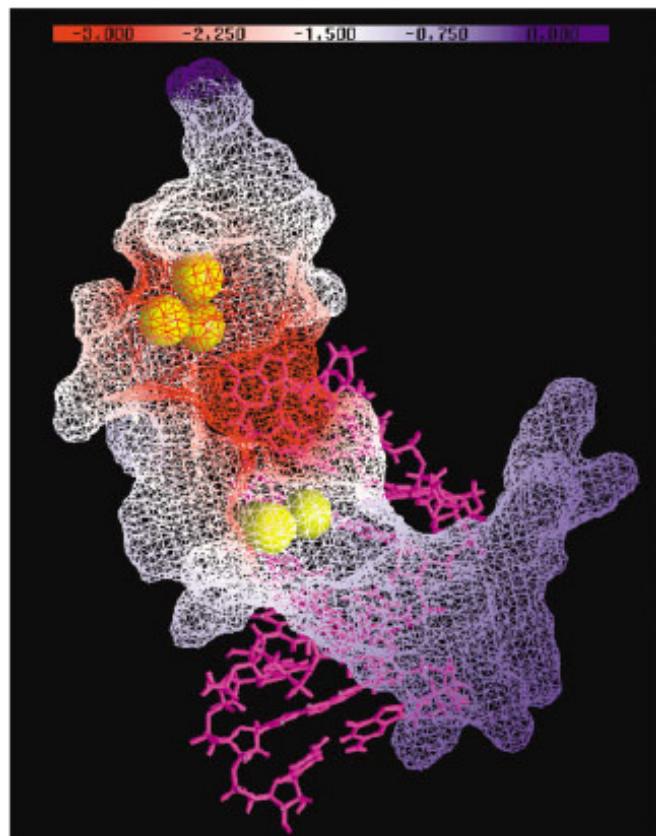
CDFT application for other molecular systems

The point is:

could the same concepts be used to describe details of the electronic structure of biomolecules and then link them to important biological events?

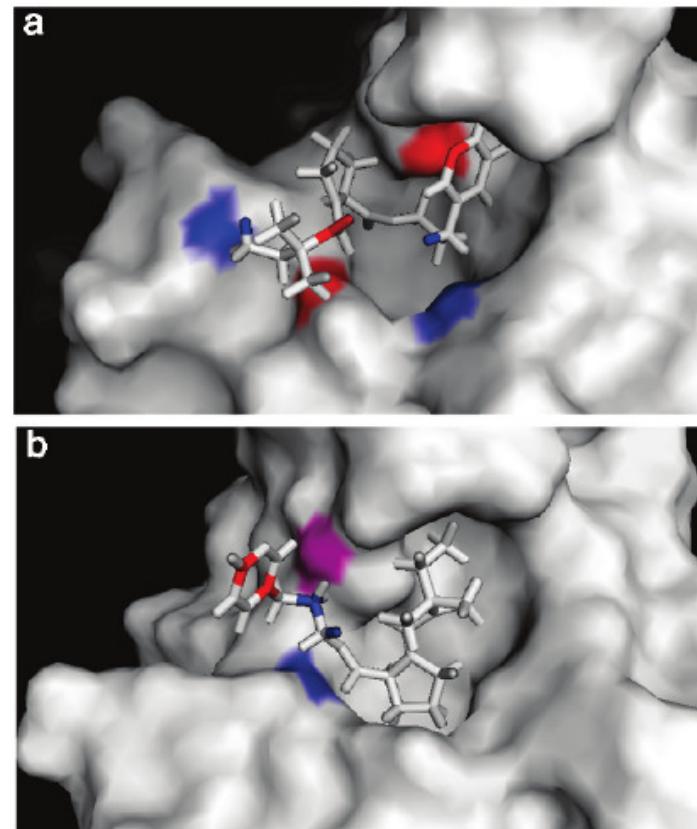
CDFT applications in biomolecules

<http://doi.wiley.com/10.1002/prot.20171>



Local hardness map for identification of enzyme active-sites.

<https://pubs.acs.org/doi/10.1021/ct9005085>



Hard-hard and soft-soft interactions for the refinement of docking poses.

Few programs allow the calculation of CDFT descriptors for large systems.

QM descriptors for biomolecules

Implementation :

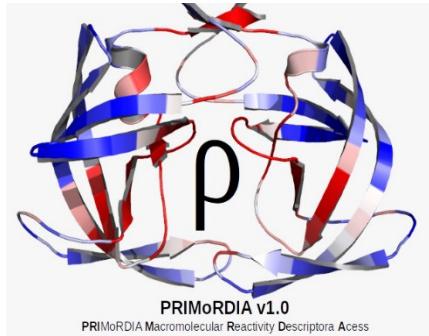
- Calculate many QM descriptors using linear scaling methods (MOZYME, FMO, LocalSCF, etc.)
- Automate the use of QM software for calculating the descriptors applying the methods and levels of theory of interest.

Application:

- Use the descriptors to predict reaction steps in enzymes, propose kinetic efficiency scales for ligands, etc.

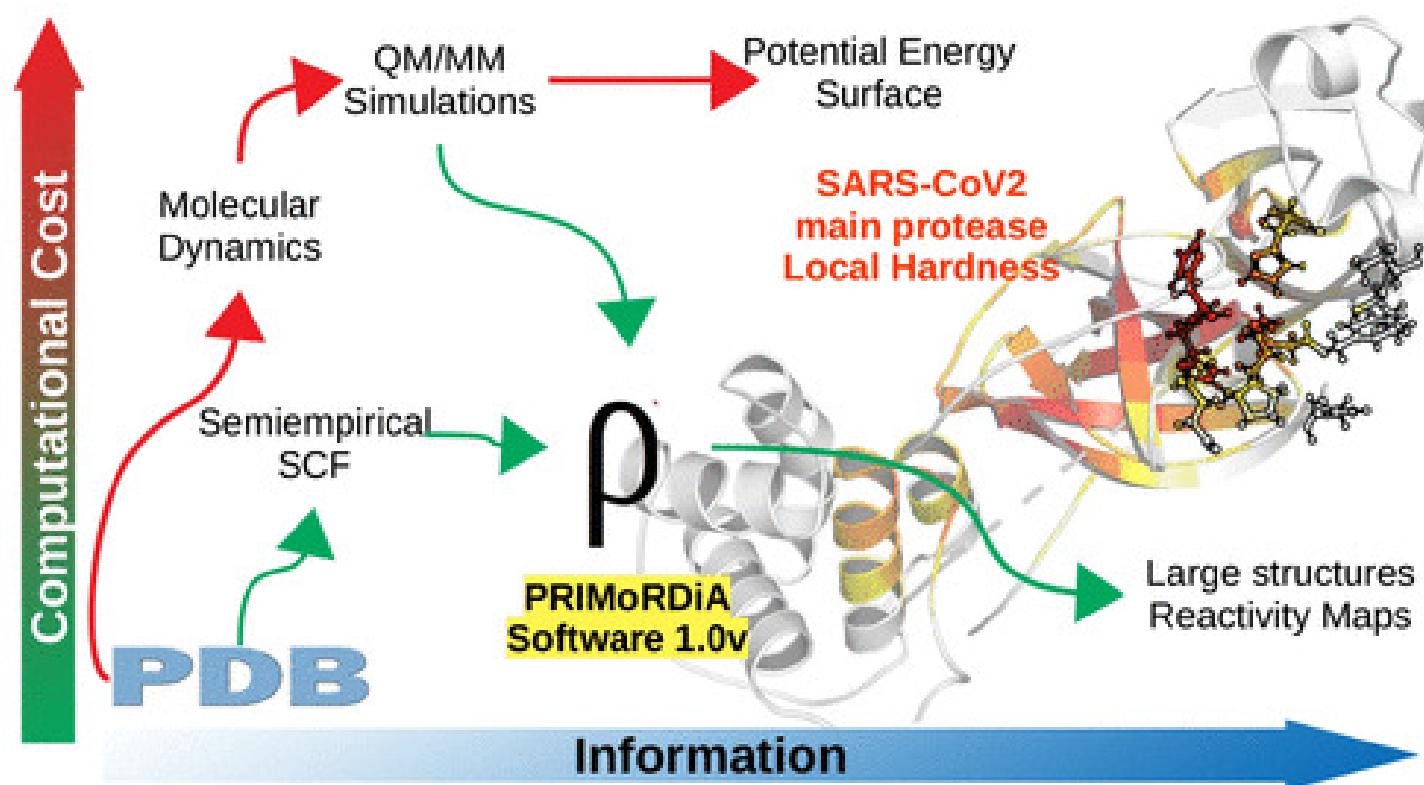
PRIMoRDiA:

Software for Quantum Chemical Reactivity Descriptors Calculation for Biomolecules



<https://pubs.acs.org/doi/10.1021/acs.jcim.0c00655>

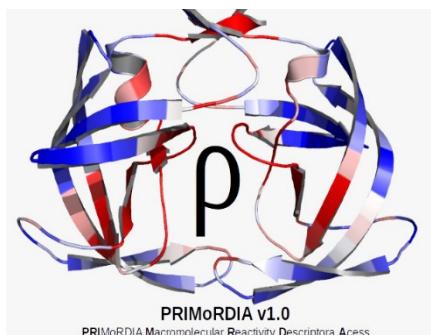
<https://github.com/igorChem/PRIMoRDiA1.0v>



PRI^MoRDⁱA:

Software for Quantum Chemical Reactivity Descriptors Calculation for Biomolecules

<https://github.com/igorChem/PRIMoRDIA1.0v>



PRIMoRDIA: A Software to Explore Reactivity and Electronic Structure in Large Biomolecules

Igor Barden Grillo, Gabriel A. Urquiza-Carvalho, and Gerd Bruno Rocha*

Cite this: *J. Chem. Inf. Model.* 2020, 60, 12, 5885–5890

Publication Date: November 13, 2020 ▾

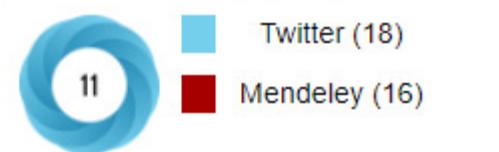
<https://doi.org/10.1021/acs.jcim.0c00655>

Copyright © 2020 American Chemical Society

[RIGHTS & PERMISSIONS](#)

Article Views Altmetric Citations

413 11 1



SI Supporting Info (1) ▾

Share Add to Export



[Read Online](#)

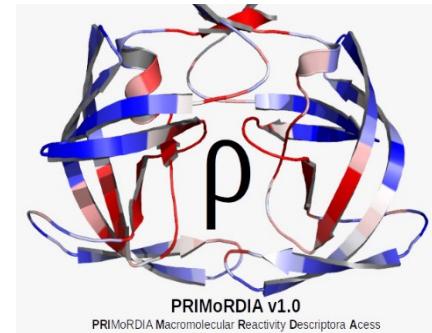
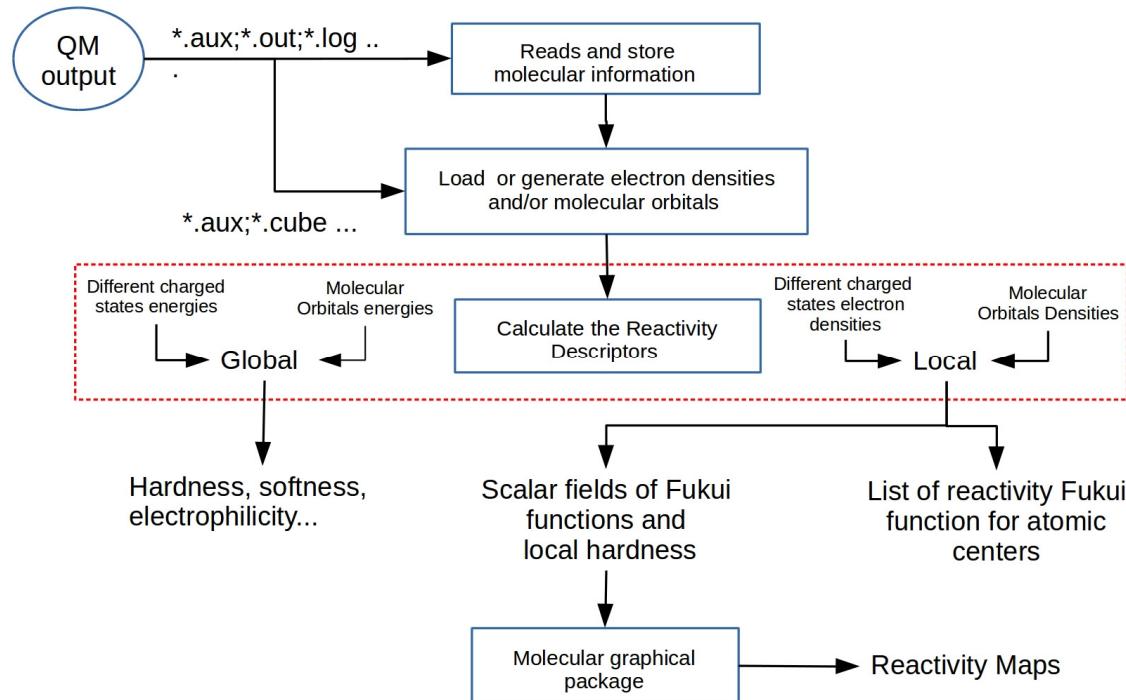
[PDF \(4 MB\)](#)

SUBJECTS:

Peptides and proteins, ▾

PRIMoRDIA:

Software for Quantum Chemical Reactivity Descriptors Calculation for Biomolecules



Library Features

- Secure and efficient dynamic memory allocation with smart pointers
- Secure and efficient arrays storage and access
- Multi-thread shared memory enabled
- Entirely written in C++

API capabilities

- RAM efficiency in big files loading
- Efficiency in electron density and molecular orbital calculation for proteins
- Several enabled methods and types of reactivity descriptors calculation
- Automation in getting the descriptors for several given QM outputs: MOPAC, G09, GAMESS, ORCA.
- HOMO-LUMO band reactivity maps screening

<https://github.com/igorChem/PRIMoRDIA1.0v>

<https://pubs.acs.org/doi/10.1021/acs.jcim.0c00655>

PRIMoRDiA: Software for Quantum Chemical Reactivity Descriptors Calculation for Biomolecules

igorChem / PRIMoRDiA1.0v

Watch 1 Star 1

Code Issues Pull requests Actions Projects Wiki Security Insights

master ▾ 2 branches 1 tag Go to file Code ▾

igorChem Merge branch 'master' of https://github.com/igorChem/PRIMoRDiA1.0v	921e73c on 3 Nov 2020	67 commits
data_test	file uploading	2 months ago
include	uploading source	4 months ago
src	adjusted cmakelists	2 months ago
userguide	removing tex files	2 months ago
CMakeLists.txt	adjusted cmakelists	2 months ago
LICENSE	Create LICENSE	4 months ago
PRIMoRDiA_1.0v_LINUX64	adjusted cmakelists	2 months ago
cover.png	Add files via upload	2 months ago
logo.png	logo	7 months ago
logo_primordia.pdf	logo	7 months ago

About

PRIMoRDiA 1.0v source repository, executable built for Linux x86_64 machines and user guide

Readme

MPL-2.0 License

Releases 1

v1.0 Latest on 28 Aug 2020

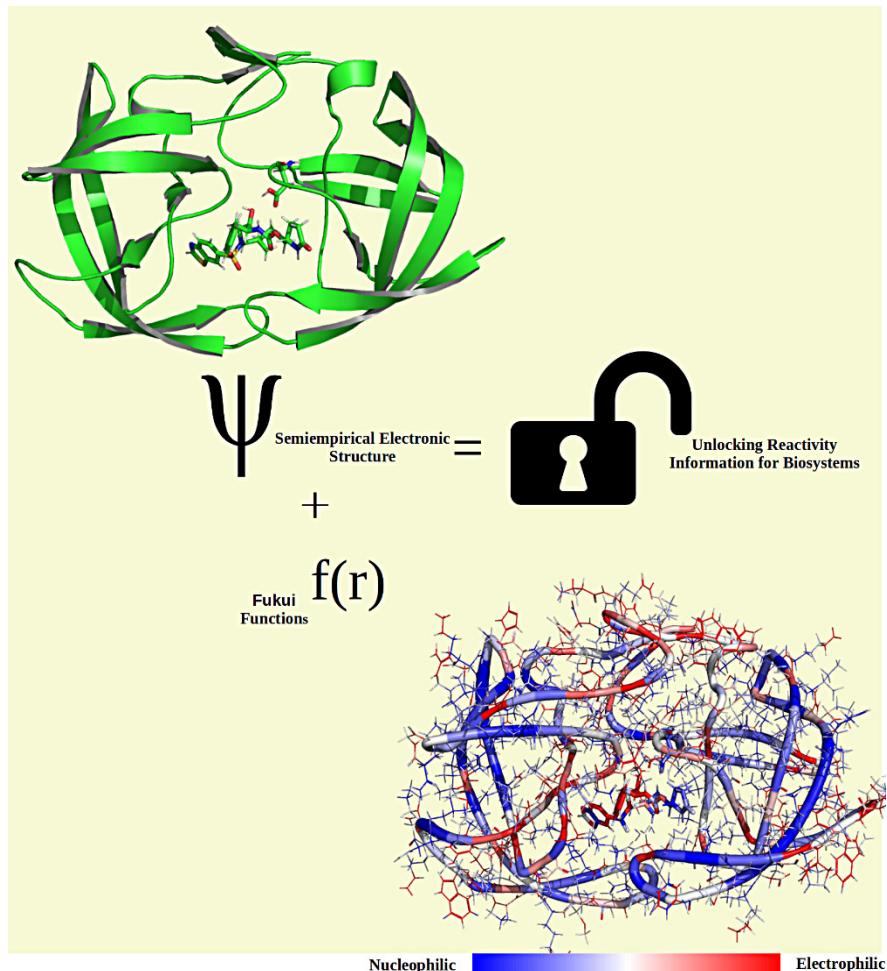
Packages

No packages published

<https://github.com/igorChem/PRIMoRDiA1.0v>

<https://pubs.acs.org/doi/10.1021/acs.jcim.0c00655>

Semiempirical methods do Fukui functions: Unlocking a modeling framework for biosystems



Received: 19 August 2019 | Revised: 15 October 2019 | Accepted: 1 January 2020
DOI: 10.1002/jcc.26148

FULL PAPER



Semiempirical methods do Fukui functions: Unlocking a modeling framework for biosystems

Igor Barden Grillo¹ | Gabriel A. Urquiza-Carvalho² | Elton José Ferreira Chaves³ | Gerd Bruno Rocha¹

¹Departamento de Química, Universidade Federal do Paraná, João Pessoa, Brazil

²Departamento de Química Fundamental, Universidade Federal do Paraná, Brazil

³Departamento de Biociências, Universidade Federal do Paraná, João Pessoa, Brazil

Correspondence:
Gerd Rocha, Departamento de Química,
Universidade Federal do Paraná, João Pessoa,
Brazil.
E-mail: gerdr@ufpr.br

Funding Information:
Conselho Nacional de Desenvolvimento
Científico e Tecnológico (CNPq), Grant/Award
Number: 309743/2017-4; Coordenação de
Aperfeiçoamento de Pessoal de Nível Superior
(CAPES), Grant/Award Number:
APQ-FPPE375/2018; Bioinformática Estrutural
de Peptídeos: Modelos, Algoritmos e
Aplicações Biotecnológicas (Editor: Biologia
Computacional 11/2013); Centro Nacional de
Processamento de Alto Desempenho em São
Paulo (CNAHAD-SP); Universidade Federal do
Pará (UFPA); Financiadora de Estudos e
Projetos; Programa de Apoio a Núcleos de
Excellência (PRONEX-FACEPE); Instituto
Nacional de Ciência e Tecnologia de
Biotecnologia para Medicamentos Integrados
(INCT-BMII).

Abstract

Obtaining reactivity information from the molecular electronic structure of a chemical system is a computationally intensive process. As a way of probing reactivity information around that, there exist electron density response variables, such as the Fukui functions (FFs), which are well-established descriptors that summarize the local susceptibility to react. These properties only require few single-point quantum chemical calculations, but even then, the intrinsic high cost and unfavorable computational complexity with respect to the number of atoms in the system makes this approach available only to small fragments and systems. In this study, we explore the computation of FFs, showing that semi-empirical quantum chemical methods can be used to obtain the reactivity information equivalent to that of a Density Functional Theory (DFT) functional, for the eight entire polypeptide chains. The combination of semiempirical methods with the frozen orbital approximation allows for the obtention of these reactivity descriptors for biological systems with reasonable accuracy and speed, unlocking the utilization of these methods for such systems. These results for the frozen orbital approximation can be additionally improved when other molecular orbitals from the frontier band are employed in the computation. We also show the potential of this computational protocol in the ligand–protein complexes of HIV-1 protease, predicting which of those ligands are active inhibitors.

KEY WORDS

Biological systems; Fukui functions; ligand–protein interactions; polypeptides reactivity; semiempirical methods

DOI: 10.1002/jcc.26148

Elucidating Enzymatic Catalysis Using Fast Quantum Chemical Descriptors



pubs.acs.org/jcim

Article

Elucidating Enzymatic Catalysis Using Fast Quantum Chemical Descriptors

Igor Barden Grillo, Gabriel A. Urquiza-Carvalho, José Fernando Ruggiero Bachega, and Gerd Bruno Rocha*



Cite This: <https://dx.doi.org/10.1021/acs.jcim.9b00860>



Read Online

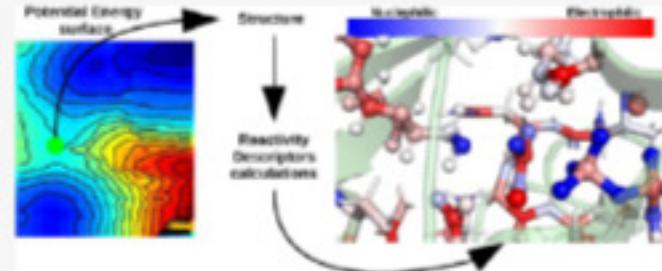
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: In general, computational simulations of enzymatic catalysis processes are thermodynamic and structural surveys to complement experimental studies, requiring high level computational methods to match accurate energy values. In the present work, we propose the usage of reactivity descriptors, theoretical quantities calculated from the electronic structure, to characterize enzymatic catalysis outlining its reaction profile using low-level computational methods, such as semiempirical Hamiltonians. We simulate three enzymatic reactions paths, one containing two reaction coordinates and without prior computational study performed, and calculate the reactivity descriptors for all obtained structures. We observed that the active site local hardness does not change substantially, even more so for the amino-acid residues that are said to stabilize the reaction structures. This corroborates with the theory that activation energy lowering is caused by the electrostatic environment of the active sites. Also, for the quantities describing the atom electrophilicity and nucleophilicity, we observed abrupt changes along the reaction coordinates, which also shows the enzyme participation as a reactant in the catalyzed reaction. We expect that such electronic structure analysis allows the expedient proposition and/or prediction of new mechanisms, providing chemical characterization of the enzyme active sites, thus hastening the process of transforming the resolved protein three-dimensional structures in catalytic information.



Theoretical characterization of the shikimate 5-dehydrogenase reaction from *Mycobacterium tuberculosis* by hybrid QC/MM simulations and quantum chemical descriptors

Journal of Molecular Modeling (2020) 26:297

<https://doi.org/10.1007/s00894-020-04536-9>

ORIGINAL PAPER



Theoretical characterization of the shikimate 5-dehydrogenase reaction from *Mycobacterium tuberculosis* by hybrid QC/MM simulations and quantum chemical descriptors

Igor Barden Grillo¹ · José Fernando Ruggiero Bachega^{2,3,4} · Luis Fernando S. M. Timmers^{2,5} · Rafael A. Caceres⁶ · Osmar Norberto de Souza^{2,7} · Martin J. Field^{8,9} · Gerd Bruno Rocha¹

Received: 2 June 2020 / Accepted: 7 September 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

In this study, we have investigated the enzyme shikimate 5-dehydrogenase from the causative agent of tuberculosis, *Mycobacterium tuberculosis*. We have employed a mixture of computational techniques, including molecular dynamics, hybrid quantum chemical/molecular mechanical potentials, relaxed surface scans, quantum chemical descriptors and free-energy simulations, to elucidate the enzyme's reaction pathway. Overall, we find a two-step mechanism, with a single transition state, that proceeds by an energetically uphill hydride transfer, followed by an energetically downhill proton transfer. Our mechanism and calculated free energy barrier for the reaction, 64.9 kJ mol^{-1} , are in good agreement with those predicted from experiment. An analysis of quantum chemical descriptors along the reaction pathway indicated a possibly important, yet currently unreported, role of the active site threonine residue, Thr65.

Thermochemical and Quantum Descriptor Calculations for Gaining Insight into Ricin Toxin A (RTA) Inhibitors



<http://pubs.acs.org/journal/acsofd>

Article

Thermochemical and Quantum Descriptor Calculations for Gaining Insight into Ricin Toxin A (RTA) Inhibitors

Acassio Rocha-Santos, Elton José Ferreira Chaves, Igor Barden Grillo, Amanara Souza de Freitas, Demétrius Antônio Machado Araújo, and Gerd Bruno Rocha*



Cite This: *ACS Omega* 2021, 6, 8764–8777



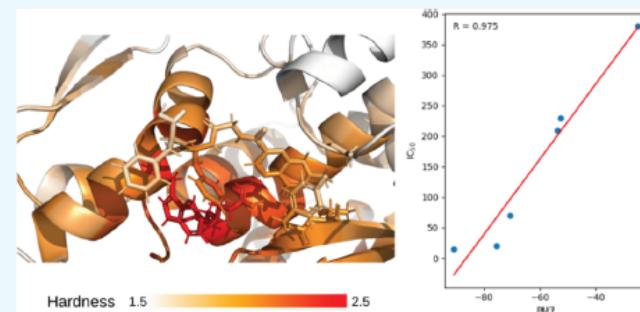
Read Online

ACCESS |

Metrics & More

Article Recommendations

ABSTRACT: In this work, we performed a study to assess the interactions between the ricin toxin A (RTA) subunit of ricin and some of its inhibitors using modern semiempirical quantum chemistry and ONIOM quantum mechanics/molecular mechanics (QM/MM) methods. Two approaches were followed (calculation of binding enthalpies, ΔH_{bind} , and reactivity quantum chemical descriptors) and compared with the respective half-maximal inhibitory concentration (IC_{50}) experimental data, to gain insight into RTA inhibitors and verify which quantum chemical method would better describe RTA–ligand interactions. The geometries for all RTA–ligand complexes were obtained after running classical molecular dynamics simulations in aqueous media. We found that single-point energy calculations of ΔH_{bind} with the PM6-DH+, PM6-D3H4, and PM7 semiempirical methods and ONIOM QM/MM presented a good correlation with the IC_{50} data. We also observed, however, that the correlation decreased significantly when we calculated ΔH_{bind} after full-atom geometry optimization with all semiempirical methods. Based on the results from reactivity descriptors calculations for the cases studied, we noted that both types of interactions, molecular overlap and electrostatic interactions, play significant roles in the overall affinity of these ligands for the RTA binding pocket.



A higher flexibility at the SARS-CoV-2 main protease active site compared to SARS-CoV and its potentialities for new inhibitor virtual screening targeting multi-conformers

JOURNAL OF BIOMOLECULAR STRUCTURE AND DYNAMICS
<https://doi.org/10.1080/07391102.2021.1924271>



Check for updates

A higher flexibility at the SARS-CoV-2 main protease active site compared to SARS-CoV and its potentialities for new inhibitor virtual screening targeting multi-conformers

Rafael E. O. Rocha^{a,b}, Elton J. F. Chaves^c , Pedro H. C. Fischer^d, Leon S. C. Costa^e, Igor Barden Grillo^c, Luiz E. G. da Cruz^c, Fabiana C. Guedes^f, Carlos H. da Silveira^f , Marcus T. Scotti^g , Alex D. Camargo^h, Karina S. Machado^h, Adriano V. Werhli^h, Rafaela S. Ferreira^a, Gerd B. Rocha^c and Leonardo H. F. de Lima^d

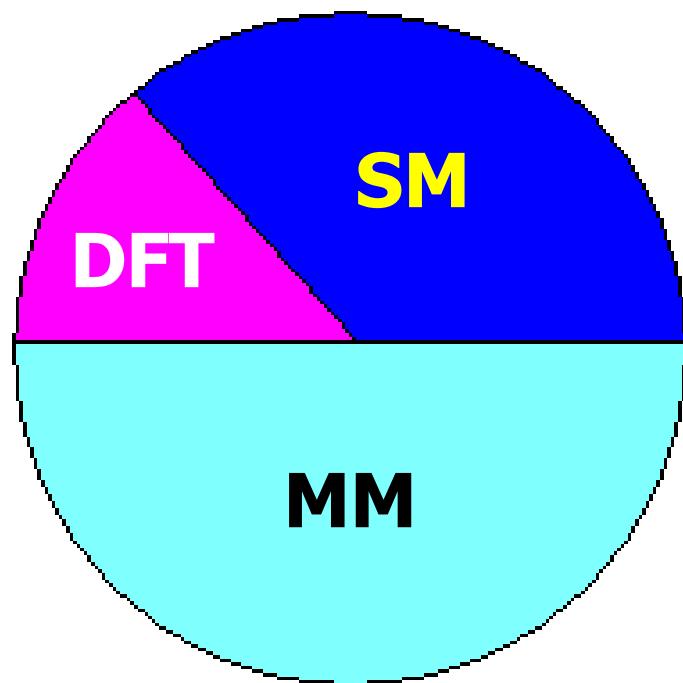
ABSTRACT

The main-protease (M^{Pro}) catalyzes a crucial step for the SARS-CoV-2 life cycle. The recent SARS-CoV-2 presents the main protease (M^{CoV2}_{pro}) with 12 mutations compared to SARS-CoV (M^{CoV1}_{pro}). Recent studies point out that these subtle differences lead to mobility variances at the active site loops with functional implications. We use metadynamics simulations and a sort of computational analysis to probe the dynamic, pharmacophoric and catalytic environment differences between the monomers of both enzymes. So, we verify how much intrinsic distinctions are preserved in the functional dimer of M^{CoV2}_{pro} as well as its implications for ligand accessibility and optimized drug screening. We find a significantly higher accessibility to open binding conformers in the M^{CoV2}_{pro} monomer compared to M^{CoV1}_{pro} . A higher hydration propensity for the M^{CoV2}_{pro} S2 loop with the A46S substitution seems to exercise a key role. Quantum calculations suggest that the wider conformations for M^{CoV2}_{pro} are less catalytically active in the monomer. However, the statistics for contacts involving the N-finger suggest higher maintenance of this activity at the dimer. Docking analyses suggest that the ability to vary the active site width can be important to improve the access of the ligand to the active site in different ways. So, we carry out a multiconformational virtual screening with different ligand bases. The results point to the importance of taking into account the protein conformational multiplicity for new promissors anti M^{CoV2}_{pro} ligands. We hope these results will be useful in prospecting, repurposing and/or designing new anti SARS-CoV-2 drugs.

The usage of computational chemistry methods

Scenario before 1990

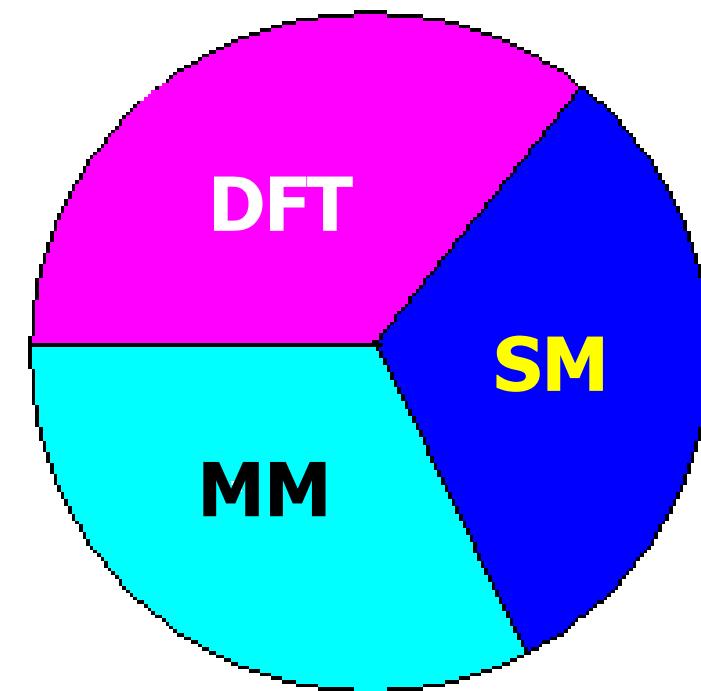
Organic Chemistry



Biochemistry

Scenario today

Organic chemistry

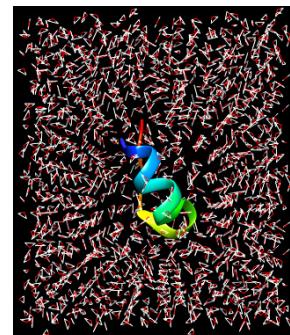


Biochemistry

New usage of quantum chemical methods

- Many calculations of many small molecules

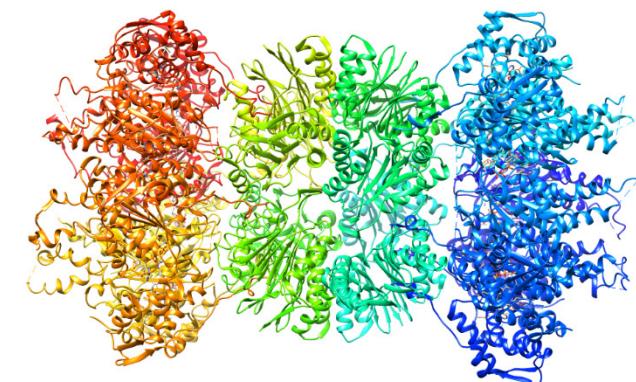
- Molecular dynamics
- Monte Carlo



- Giant molecules.

- **Molecular dynamics using force fields**

- 3D structure, motion and function



- **Molecular dynamics using quantum potentials**

- 3D structure, **electronic structure**, motion and function

Our projects ...

Theoretical and computational projects

- 1. Code development for calculation of large molecular systems;**
 - Parallel programming for Multi-CPU and Multi-GPU.
- 2. Parameterization of Semiempirical Methods;**
 - RM1, Sparkle model, etc.
- 3. Development of new Linear Scaling Methods**

Applications of these projects ...

- 1. *Drug Design* for Tropical diseases;**
 - QSAR, QSAR-4D, molecular dynamics, Docking, MMGBSA, Free energy calculations, etc.
- 2. Study of coordination complex (lanthanides, transition metals and Bi, Sb) in solution;**
 - Molecular dynamics, QM/MM and conformational analysis.
- 3. *Design* of new MRI contrast agents;**
 - Conformational analysis, Thermodynamic and kinetic quantities, etc.
- 4. *Design* of new molecules for nonlinear optics;**
 - DFT, ab initio and semiempirical methods.
- 5. Molecular modeling and simulation of biomolecules**
 - MD, QM, etc.

Pós-graduação no PPGQ da UFPB

Se você gosta de programar venha fazer parte do PPGQ e do nosso grupo de pesquisa para desenvolver projeto conjunto das pós-graduações de Química e de Informática na temática **“Computação Paralela Aplicada em Química Computacional: uso de arquiteturas paralelas híbridas GPU/multi-core no desenvolvimento de novos códigos de química quântica”**.

**Perfil exigido: ter domínio de alguma linguagem de programação
Fortran, C ou C++**



gbr@quimica.ufpb.br

gbr@academico.ufpb.br

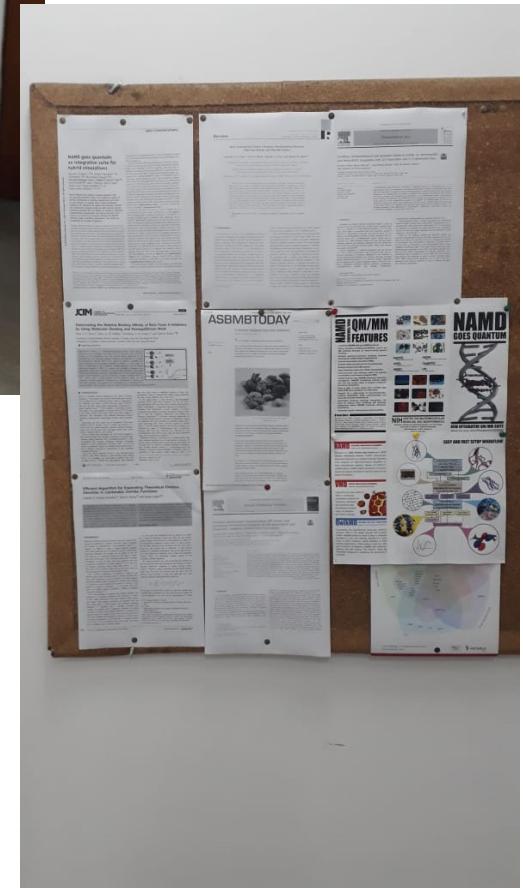
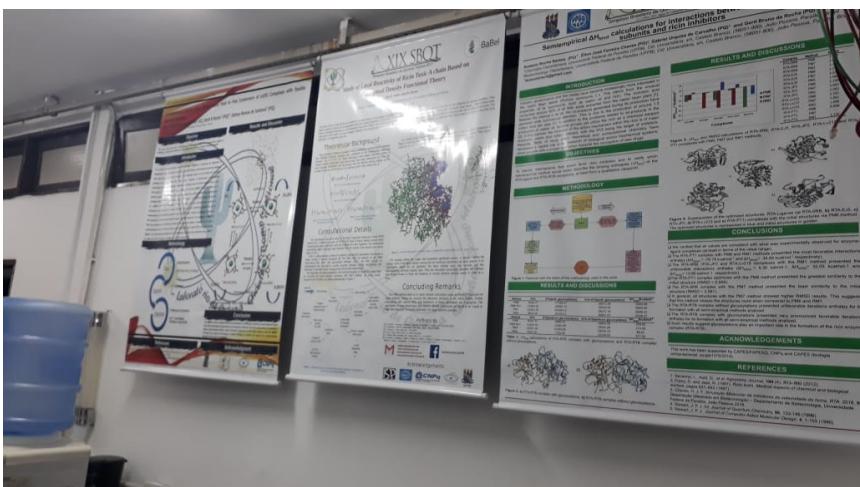
www.quantum-chem.pro.br

www.quimica.ufpb.br/posgrad

LQQC – Laboratório de Química Quântica Computacional



LQQC – Laboratório de Química Quântica Computacional



LQQC – Laboratório de Química Quântica Computacional



LQQC – Laboratório de Química Quântica Computacional



LQQC – Laboratório de Química Quântica Computacional



LQQC – Laboratório de Química Quântica Computacional

Infraestrutura

- **PCs**
- **Workstations**
- **Acesso a um Cluster (com GPUs) localizado no LMMRQ**
- **Licença de vários softwares de química e física**
- **Licença de vários programas matemáticos, estatísticos**
- **Sala de estudos**
- **Livros e manuais de programas**
- **Ambiente de pós-graduação**
- **Copa e banheiros**



Universidade Federal da Paraíba

Departamento de Química

Laboratório de Química Quântica Computacional

Prof. Gerd Bruno da Rocha

(gbr@quimica.ufpb.br, gerd.rocha@gmail.com)

www.quantum-chem.pro.br





Universidade Federal da Paraíba

Departamento de Química

Laboratório de Química Quântica Computacional
Prof. Gerd Bruno da Rocha

(gbr@quimica.ufpb.br, gerd.rocha@gmail.com)

www.quantum-chem.pro.br

The screenshot shows the homepage of the research group's website. At the top, there is a logo of three torches and the text: "Research Group on Quantum Chemistry Applied to Biological Systems", "Gerd Bruno Rocha", "DQ - CCEN - UFPB - Brazil". Below this is a navigation menu with links: HOME, RESEARCH, PUBLICATIONS, MEMBERS, COLLABORATORS, LINKS, ABOUT US, SPONSORS. The main content area contains several diagrams illustrating the QM/MM modeling approach. One diagram shows the division of space into QM and MM regions, with NAMD/ORCA input files and PSF topology. Another diagram shows the interaction between QM atoms (partial charges and gradients) and MM atoms (gradients due to Coulomb interactions). A third diagram shows the charge redistribution scheme where a carbon atom's charge is distributed to preserve local polarization. To the right, a blue box welcomes visitors and provides a brief overview of the group's focus on developing new quantum chemistry methods and high-performance computing for molecular modeling.

WELCOME TO:
the Group Webpage of Gerd Rocha

We are a Theoretical Quantum Chemistry Group in the Department of Chemistry at UFPB.

Our research focuses on the development of new quantum chemistry methods, high performance computing in chemistry and molecular modeling of organic, inorganic and biological systems.

More information can be found through the links in the upper menu.

THE END
Thank you for your
attention!

João Pessoa – PB - Brasil



<http://turismo.joaopessoa.pb.gov.br/>



João Pessoa – PB - Brasil

- it is also known as *the city where the sun rises first*, because it is the easternmost city in the Americas at $34^{\circ} 47' 38''$ W, $7^{\circ} 9' 28''$ S.



Sunrise at Jacaré Beach

Its easternmost point is known as Ponta do Seixas



João Pessoa – PB - Brasil



Beach avenue
at night

João Pessoa is the
second greenest city
in the world



João Pessoa – PB - Brasil



Red sand island

João Pessoa – PB - Brasil



The sea water temperature is around 27ºC in summer and spring.

The End