German Research School for Simulation Sciences

JÜLICH FORSCHUNGSZENTRUM

# Car-Parrinello Molecular Dynamics, QM/MM and applications to biophysical systems


Dr. Emiliano Ippoliti

German Research School for Simulation Sciences

and

Institute for Advanced Simulations - Computational Biomedicine (IAS-5)

Jülich (Germany)

German Research School for Simulation Sciences

JÜLICH FORSCHUNGSZENTRUM

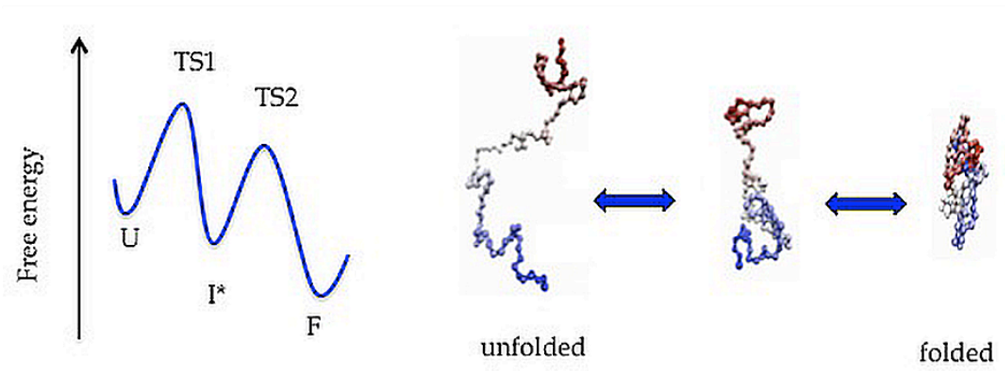
# Summary

- Introduction about MD
- *Ab initio* MD
- Car-Parrinello MD
- QM/MM within the Car-Parrinello MD
- Typical biological oriented applications

# Why dynamics?

What "minimum-energy chemistry" cannot predict:

- Finite temperature effects
- Environmental effects (entropic contributions, etc.)
- Processes changing in time



German Research School for Simulation Sciences

JÜLICH FORSCHUNGSZENTRUM

# Why not classical molecular dynamics?

Predefined fixed potentials either based on empirical data or on independent electronic structure calculations:

- Many different atom types give rise to a myriad of different interatomic interactions that have to be parametrized;
- Electronic structure and thus the bonding pattern changes qualitatively in the course of time.

# *Ab initio* molecular dynamics

- Electronic variables considered as active degrees of freedom
- Forces acting on the nuclei from the electronic structure calculations performed "on-the-fly" as the molecular dynamics trajectory is generated

Problem shifted from the level to selecting the model potential to the level of selecting a particular approximation for solving the Schrödinger equation.

German Research School for Simulation Sciences

JÜLICH FORSCHUNGSZENTRUM

# Limitations

- The correlation lengths and relaxation times that are accessible are much smaller than what is affordable via standard molecular dynamics.
- Tracing back the properties of a given system to a simple physical picture is much harder in *ab initio* molecular dynamics.

The bright side is that new phenomena, which were not foreseen before starting the simulation, can simply happen if necessary.

# Starting point

*Non-relativistic* quantum mechanics as formalized via time-dependent Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\Phi\left(\{\mathbf{r}_i\},\{\mathbf{R}_I\};t\right)=\mathcal{H}\Phi\left(\{\mathbf{r}_i\},\{\mathbf{R}_I\};t\right)$$

in its position representation in conjunction with the standard Hamiltonian:

$$\mathcal{H}=\underbrace{-\sum_I\frac{\hbar^2}{2M_I}\nabla_I^2-\sum_i\frac{\hbar^2}{2m_e}\nabla_i^2+\sum_{i<j}\frac{e^2}{|\mathbf{r}_i-\mathbf{r}_j|}}_{\mathcal{H}_e}\underbrace{-\sum_{i,I}\frac{e^2Z_I}{|\mathbf{R}_I-\mathbf{r}_i|}+\sum_{i<j}\frac{e^2Z_IZ_J}{|\mathbf{R}_I-\mathbf{R}_J|}}_{V_{n-e}}$$

for the electronic $\{\mathbf{r}_i\}$ and nuclear $\{\mathbf{R}_I\}$ degrees of freedom.

# Electronic and nuclear separation

This approximation must lead to a mean-field description of the coupled dynamics

"One determinant" or "single configuration" ansatz (which is *not equivalent to the Born-Oppenheimer approximation*):

$$\Phi\left(\{\mathbf{r}_i\},\{\mathbf{R}_I\};t\right) \approx \Psi\left(\{\mathbf{r}_i\};t\right)\Xi\left(\{\mathbf{R}_I\};t\right)\exp\left[\frac{i}{\hbar}\int_{t_0}^{t} dt' E'_e(t')\right]$$

where the phase factor

$$E'_e(t) = \int \Psi^*\left(\{\mathbf{r}_i\};t\right)\Xi^*\left(\{\mathbf{R}_I\};t\right)\mathcal{H}_e\Psi\left(\{\mathbf{r}_i\};t\right)\Xi\left(\{\mathbf{R}_I\};t\right)d\mathbf{r}d\mathbf{R}$$

is introduced for convenience such that the final equations will look simpler.

# Electronic and nuclear separation

Inserting this particular separation ansatz into time-dependent Schrödinger equation, multiplying from the left by $\Psi^*$ and $\Xi^*$, integrating over nuclear and electronic coordinates, respectively, and imposing the total energy conservation:

Both electrons and nuclei move quantum mechanically in time-dependent effective potential

$$i\hbar\frac{\partial\Psi}{\partial t} = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2\Psi + \left\{\int d\mathbf{R}\,\Xi^*\left(\{\mathbf{R}_I\};t\right)\mathrm{V}_{n-e}\left(\{\mathbf{r}_i\};\{\mathbf{R}_I\}\right)\Xi\left(\{\mathbf{R}_I\};t\right)\right\}\Psi$$

$$i\hbar\frac{\partial\Xi}{\partial t} = -\sum_i \frac{\hbar^2}{2M_I}\nabla_I^2\Xi + \left\{\int d\mathbf{r}\,\Psi^*\left(\{\mathbf{r}_i\};t\right)\mathcal{H}_e\left(\{\mathbf{r}_i\};\{\mathbf{R}_I\}\right)\Psi\left(\{\mathbf{r}_i\};t\right)\right\}\Xi$$

# Semiclassical approximation for nuclei

Polar representation of nuclear wavefunction:

$$\Xi\left(\left\{\mathbf{R}_I\right\};t\right) = A\left(\left\{\mathbf{R}_I\right\};t\right)\exp\left[iS\left(\left\{\mathbf{R}_I\right\};t\right)/\hbar\right]$$

The real and imaginary part of the equation for $\Xi$ becomes:

$$\frac{\partial S}{\partial t} + \sum_I \frac{1}{2M_I}\left(\nabla_I S\right)^2 + \int d\mathbf{r}\,\Psi^*\mathcal{H}_e\Psi = \hbar^2 \sum_I \frac{1}{2M_I}\frac{\nabla_I^2 A}{A} \quad \overset{\hbar\to 0}{\rightarrow} \quad 0$$

$$\frac{\partial A}{\partial t} + \sum_I \frac{1}{M_I}\left(\nabla_I A\right)\left(\nabla_I S\right) + \sum_I \frac{1}{2M_I} A\left(\nabla_I^2 S\right) = 0$$

TDSCF equations

# Semiclassical approximation for nuclei

$$\frac{\partial S}{\partial t}+\sum_I \frac{1}{2M_I}\left(\nabla_I S\right)^2+\int d\mathbf{r}\,\Psi^*\mathcal{H}_e\Psi=0$$

is isomorphic to the equation of motion in the Hamilton-Jacobi formulation of classical mechanics:

$$\frac{\partial S}{\partial t}+T\left(\{\mathbf{P}_I\}\right)+V\left(\{\mathbf{R}_I\}\right)=0$$

$$\mathbf{P}_I\equiv\nabla_I S$$

$$\dot{\mathbf{P}}_I=-\nabla_I\int d\mathbf{r}\,\Psi^*\mathcal{H}_e\Psi$$

Ehrenfest potential

$$M_I\ddot{\mathbf{R}}_I=-\nabla_I\left\langle\Psi\middle|\mathcal{H}_e\middle|\Psi\right\rangle$$

# Various ground state ($\Psi_0$) *ab initio* molecular dynamics

- **Ehrenfest** molecular dynamics:

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle$$

*Unitary propagation of* $\Psi_0$

$$i\hbar \frac{\partial \Psi_0}{\partial t} = \mathcal{H}_e \Psi_0$$ ← *Time-dependent* Schrödinger eq.

- **Born-Oppenheimer** molecular dynamics:

*No propagation of* $\Psi_0$

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \min_{\Psi_0} \{ \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle \}$$

⇨ *Min. at each time step*

$$E_0 \Psi_0 = \mathcal{H}_e \Psi_0$$ ← *Time-independent* Schrödinger eq.

- **Car-Parrinello** molecular dynamics:

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle + \sum_{ij} \Lambda_{ij} \frac{\partial}{\partial \mathbf{R}_I} \langle \psi_i | \psi_j \rangle$$

$$\mu \ddot{\psi}_i(t) = -\frac{\delta}{\delta \psi_i^*} \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle + \sum_j \Lambda_{ij} \psi_j$$

Electronic d.o.f. as (fictitious) dynamical variables

# Car-Parrinello molecular dynamics

Constant of motion:

$$E_{cons} = \sum_I \frac{1}{2} M_I \dot{\mathbf{R}}_I^2 + \sum_i \mu \langle \dot{\psi}_i | \dot{\psi}_j \rangle + \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle$$

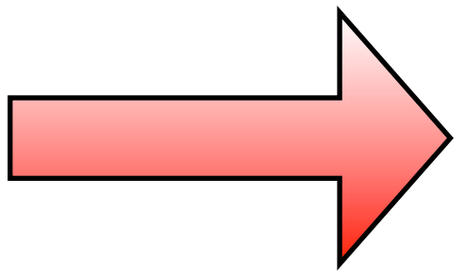
- Nuclei evolve in time
  at a certain (instantaneous) physical temperature $\propto \sum_I M_I \dot{\mathbf{R}}_I^2$
- Electronic degrees of freedom evolve
  at a "fictitious temperature" $\propto \sum_i \mu \langle \dot{\psi}_i | \dot{\psi}_j \rangle$

- "Low electronic temperature" or "cold electrons" means that the electrons are close to the minimum energy $\min_{\Psi_0} \{ \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle \}$ i.e. close to the Born–Oppenheimer surface.

# CPMD vs BOMD

- In CPMD electronic and ionic structure evolve *simultaneously* whereas in BOMD first the electronic structure is optimized, *then* the ions are move.

- Instantaneous value of $\langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle$ is not a minimum
- No need to optimize the orbitals at each step



Faster than BOMD and still accurate (i. e. stable)

# Adiabaticity condition

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle + \sum_{ij} \Lambda_{ij} \frac{\partial}{\partial \mathbf{R}_I} \langle \psi_i | \psi_j \rangle$$

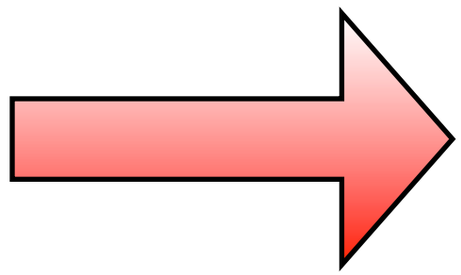
$$\mu \ddot{\psi}_i(t) = -\frac{\delta}{\delta \psi_i^*} \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle + \sum_j \Lambda_{ij} \psi_j$$

In CPMD an explicit electronic minimization at each time step, as done in BOMD, is not needed: after an initial standard electronic minimization, the fictitious dynamics of the electrons keeps them on the electronic ground state corresponding to each new ionic configuration visited along the dynamics, thus yielding accurate ionic forces.

In order to maintain this adiabaticity condition, it is necessary that the fictitious mass of the electrons $\mu$ is chosen small enough to avoid a significant energy transfer from the ionic to the electronic degrees of freedom.

# Adiabaticity condition

This small fictitious mass in turn requires that the equations of motion are integrated using a smaller time step $\tau$ (~0.1 fs) than the one (1–10 fs) commonly used in BOMD.



Trade off between small $\mu$ and large $\tau$

The merit of Car and Parrinello was to show that for typical physical system finding this trade-off is possible.

# Electronic structure method

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \left\langle \Psi_0 \middle| \mathcal{H}_e \middle| \Psi_0 \right\rangle + \sum_{ij} \Lambda_{ij} \frac{\partial}{\partial \mathbf{R}_I} \left\langle \psi_i \middle| \psi_j \right\rangle$$

$$\mu \ddot{\psi}_i(t) = -\frac{\delta}{\delta \psi_i^*} \left\langle \Psi_0 \middle| \mathcal{H}_e \middle| \Psi_0 \right\rangle + \sum_j \Lambda_{ij} \psi_j$$

- Hartree-Fock theory
- Generalized Valence Bond
- Complete active space SCF
- ...
- Density Functional Theory

# PW Basis Set

To calculate KS equations (and the rest) on a computer we need to expand KS orbitals on some **basis set**.

A traditional and historical choice in the framework of *ab initio* MD is the Plane Wave basis set:

$$\psi_i(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}}^{\mathbf{G}_{\max}} c_i(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}$$

- $\mathbf{G}$ = a reciprocal lattice vector in the periodic MD cell
- $\Omega$ = volume of the MD cell
- $\mathbf{G}_{\max}$= maximum length of $\mathbf{G}$ vectors in the expansion which determines the total number of plane waves $M \approx \frac{1}{4\pi^2}\Omega\mathbf{G}_{\max}^3$

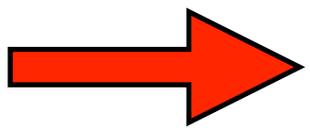
Depends on volume!

# PW Basis Set

$$\psi_i(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}}^{\mathbf{G}_{max}} c_i(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}$$

Benefits:

- No Basis Set Superposition Error (BSSE)
- Orthogonal basis set
- Independent of atomic positions



Hellmann-Feynman theorem

No Pulay forces

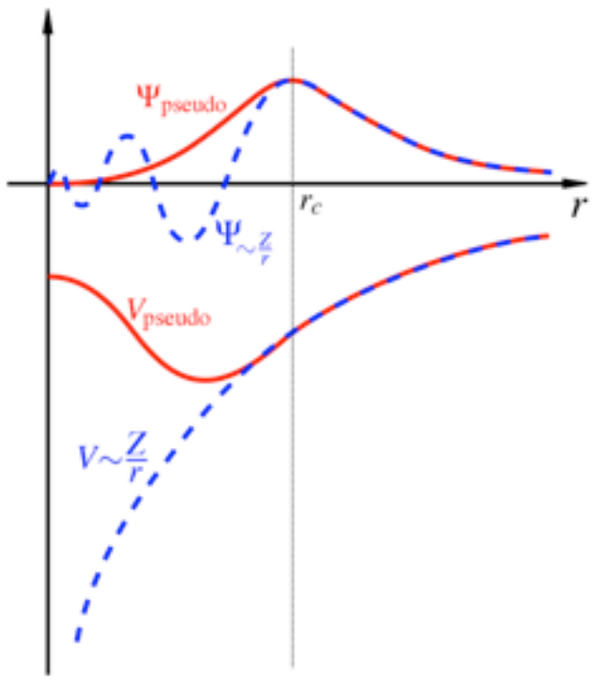
$$\nabla_I \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle = \langle \Psi_0 | \nabla_I \mathcal{H}_e | \Psi_0 \rangle + 2 \langle \nabla_I \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle$$

Drawbacks: naturally periodic, many functions needed to reach convergence

# Pseudopotentials

Idea: Replace electronic degrees of freedom in the Hamiltonian by an effective potential

Desirable properties: The potential should be additive and transferable



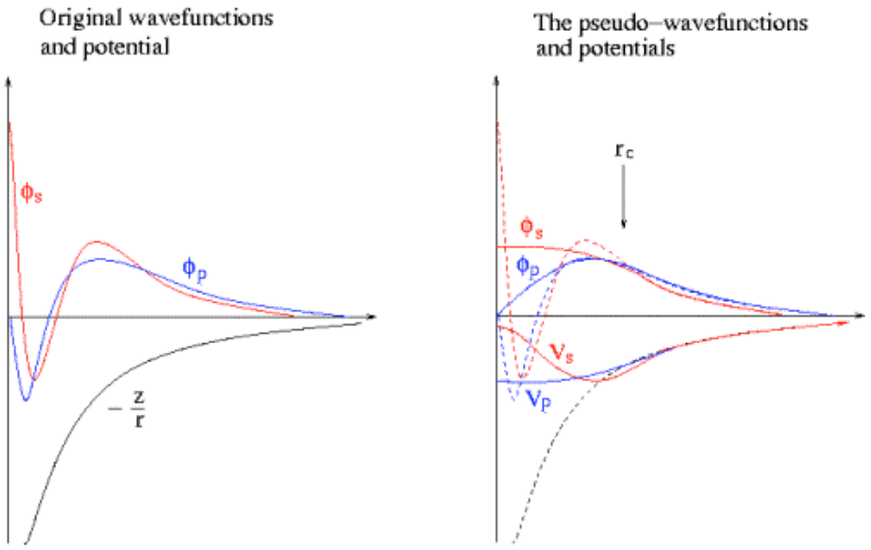

Atomic pseudopotentials

Remove only core electrons

# Atomic Pseudopotentials

Hypothesis: $\rho(\mathbf{r}) = \rho^{c}(\mathbf{r}) + \rho^{v}(\mathbf{r})$

$$\left\{-\frac{1}{2}\nabla^2 + \mathrm{V}_{\mathrm{PP}}(\mathbf{r}) + \int d\mathbf{r}' \frac{\rho^{V}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta E_{\mathrm{xc}}[\rho^{V}]}{\delta\rho^{V}(\mathbf{r})}\right\}\psi_i^{V}(\mathbf{r}) = \varepsilon_i \psi_i^{V}(\mathbf{r})$$




$$\mathrm{V}^{\mathrm{PP}}(\mathbf{r}) = \sum_{l} |l\rangle \mathrm{V}_{l}^{\mathrm{PP}}(\mathbf{r}) \langle l| \quad |l\rangle =$$ Angular momentum projection operator

Pseudopotentials have to be chosen such that the main properties of the atom are reproduced.

# Pseudopotentials

1. Reduction of basis set size → Effective speedup of calculation

2. Reduction of number of electrons → Only valence electron as degrees of freedom

3. Inclusion of relativistic effects or van der Waals corrections (bad described by DFT) → Partially included into effective potentials

# First principles QM/MM

- Combines the advantages of both the accuracy and the general applicability of QM methods with the efficiency of classical, force field-based MM methods.
- Reconciles the possibility of describing electronic processes such as chemical reactions, charge transfer or photoinduced electronic excitations with a proper description of extended environmental effects, e.g. arising from ambient solvents and/or biomolecular environments.
- The combination with statistical mechanics via MD simulations at finite temperature, provides a powerful approach to investigate dynamics, reactivity and thermodynamics of biological molecules, such as enzymes, photoreceptors, receptor/drug complexes.

# QM/MM Limitations

- Considerable computational resources required
- Therefore, affordable simulation times are often restricted to sub-nanosecond times scales. This time scale might not be sufficient to capture rare events or ensure adequate conformational sampling

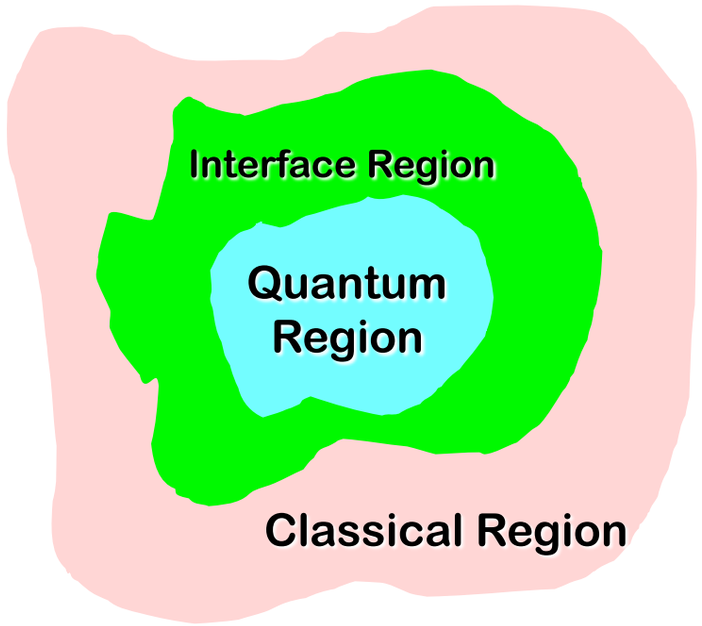
Coupling with enhance sampling methods, such as:

- Free-energy perturbation
(Zwanzig, R.W. J. Chem. Phys. 22, 1420-1426, 1954)
- Thermodynamic integration
(Kirkwood J. G. J. Chem. Phys., 3:300-313,1935)
- Umbrella sampling
(Torrie G.M.; J.P. Valleau J. Comp. Phys. 23, 187–199, 1977)
- Metadynamics
(Laio, A.; Parrinello, M. Proc. Nat. Acad. Sci. USA 99, 12562–12566, 2002)

# First principles QM/MM

The system is separated into two parts:

- One (the QM part) comprises the chemically/photophysically active region treated by computationally demanding electronic structure methods.
  (Typical size **~$10^2$ atoms**)




- The remainder (the MM part) is described efficiently at a lower level of theory by classical force fields.
  (Typical size **~$10^5$ atoms**) → PBC or continuum models

Special attention, obviously, has to be paid to the coupling of both regions (Interface region).

# CP Lagrangian

Car-Parrinello motion equations:

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle + \sum_{ij} \Lambda_{ij} \frac{\partial}{\partial \mathbf{R}_I} \langle \psi_i | \psi_j \rangle$$

$$\mu \ddot{\psi}_i(t) = -\frac{\delta}{\delta \psi_i^*} \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle + \sum_j \Lambda_{ij} \psi_j$$

can be derived from the Euler-Lagrange equations:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{R}}_I}\right) = \frac{\partial \mathcal{L}}{\partial \mathbf{R}_I}$$

$$\frac{d}{dt}\left(\frac{\delta \mathcal{L}}{\delta \dot{\psi}_i^*}\right) = \frac{\delta \mathcal{L}}{\delta \psi_i^*}$$

where the Lagrangian $\mathcal{L}$ is:

$$E^{KS}\left[\{\psi_i^V\}\right] = \sum_i \left\langle \psi_i^V \left| -\frac{1}{2}\nabla^2 \right| \psi_i^V \right\rangle + \int d\mathbf{r}\, V_{PP}(\mathbf{r}) \rho^V(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r}\, d\mathbf{r}' \frac{\rho^V(\mathbf{r}) \rho^V(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[\rho^V]$$

$$\mathcal{L}_{CP} = \sum_I \frac{1}{2} M_I \dot{\mathbf{R}}_I^2 + \sum_i \frac{1}{2} \mu_i \langle \dot{\psi}_i^V | \dot{\psi}_j^V \rangle - \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle + \sum_{i,j} \Lambda_{ij} \left( \int d\mathbf{r}\, \psi_i^{V*}(\mathbf{r}) \psi_j^V(\mathbf{r}) - \delta_{ij} \right)$$

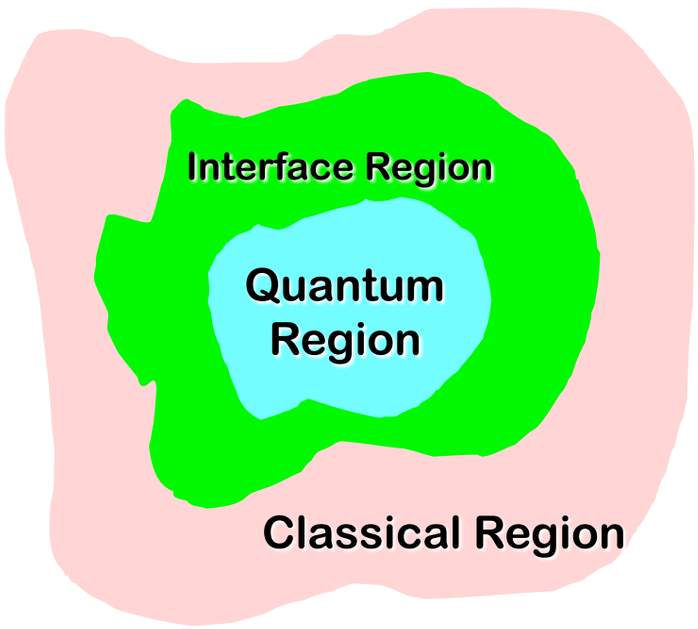
# QM/MM Lagrangian

$$E_{\mathrm{QM}} = E^{KS}\left[\left\{\psi_i^V\right\}\right]$$

$$\mathcal{L}_{\mathrm{CP/MM}} = \sum_I \frac{1}{2} M_I \dot{\mathbf{R}}_I^2 + \frac{1}{2}\mu \sum_i \left\langle \dot{\psi}_i^V \middle| \dot{\psi}_j^V \right\rangle - E_{\mathrm{QM}} - E_{\mathrm{QM/MM}} - E_{\mathrm{MM}} + \sum_{i,j} \Lambda_{ij}\left(\int d\mathbf{r}\, \psi_i^{V*}(\mathbf{r})\psi_i^V(\mathbf{r}) - \delta_{ij}\right)$$

$$E_{\mathrm{QM/MM}} = \sum_{J\in\mathrm{MM}} q_J \int d\mathbf{r} \frac{\rho_{\mathrm{QM}}^V(\mathbf{r}) + \rho_{\mathrm{QM}}^{ION}(\mathbf{r})}{\left|\mathbf{R}_J - \mathbf{r}\right|} + \frac{1}{2}\sum_{\substack{I\in\mathrm{QM}\\ J\in\mathrm{MM}}} 4\varepsilon_{IJ}\left[\left(\frac{\sigma_{IJ}}{r_{IJ}}\right)^{12} - \left(\frac{\sigma_{IJ}}{r_{IJ}}\right)^{6}\right] +$$

$$+ \sum_{Bonds} k_b\left(r_{IJ} - l_{0b}\right)^2 + \sum_{Angles} k_a\left(\theta_{I'J'K'} - \theta_{0a}\right)^2 + \sum_{Torsions} k_n\left[1 + \cos\left(n\omega_{I''J''K''L''} - \omega_{0n}\right)\right]$$




at least one of *I', J', K' and I", J", K", L"* is a MM atom

$E_{\mathrm{MM}} = \mathrm{V}\left(\mathbf{r}^N\right)$ from standard force field (Amber, Gromos, etc)

# Amber Force Field

Energy between covalently bonded atoms. Good approximation near the equilibrium bond length, but increasingly poor as atoms separate.

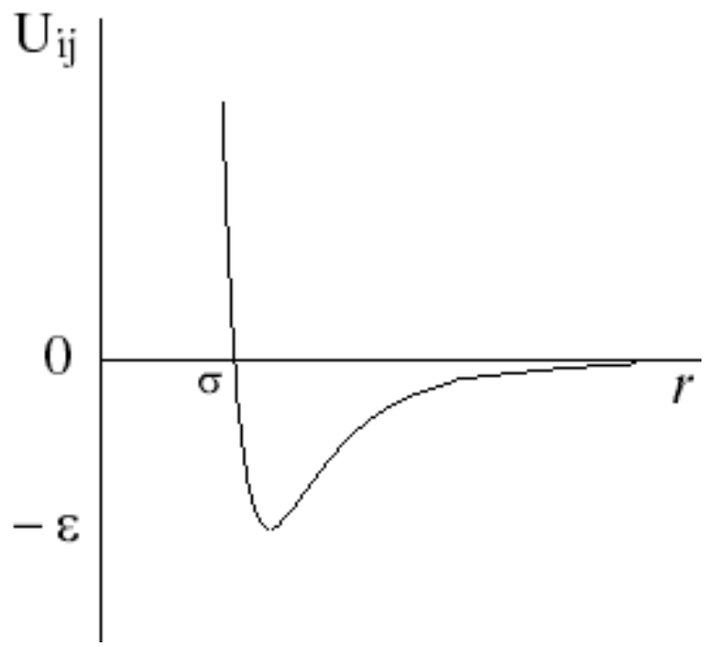
Energy due to the geometry of electron orbitals involved in covalent bonding.

Energy for twisting a bond due to bond order (e.g. double bonds) and neighboring bonds or lone pairs of electrons. A single bond may have more than one of these terms, such that the total torsional energy is expressed as a Fourier series.

$$V(\mathbf{r}^N) = \sum_{Bonds} \frac{1}{2} k_b (l - l_b)^2 + \sum_{Angles} \frac{1}{2} k_a (\theta - \theta_a)^2 + \sum_{Torsions} \frac{1}{2} V_n \left[1 + \cos(n\omega - \gamma)\right] +$$

$$+ \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} \left\{ \varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 2 \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right\}$$

Non-bonded energy between all atom pairs, which can be decomposed into van der Waals (first term of summation) and electrostatic (second term of summation) energies.



# QM/MM in detail

$$E_{\mathrm{QM/MM}} = \sum_{J\in\mathrm{MM}} q_J \int d\mathbf{r} \frac{\rho^{V}_{\mathrm{QM}} + \rho^{ION}_{\mathrm{QM}}}{\left|\mathbf{R}_J - \mathbf{r}\right|} + \frac{1}{2}\sum_{\substack{I\in\mathrm{QM}\\ J\in\mathrm{MM}}} 4\varepsilon_{IJ}\left[\left(\frac{\sigma_{IJ}}{r_{IJ}}\right)^{12} - \left(\frac{\sigma_{IJ}}{r_{IJ}}\right)^{6}\right] + \ldots$$

Short-range problem: positively charged MM atoms can act as traps for the electron if the basis set is flexible enough: the Pauli repulsion from the electron clouds that would surround the classical atoms is absent, and therefore, the electron density is overpolarized by an incorrect purely attractive potential (*electron spill-out* problem).

Long-range problem: the first term in $E_{\mathrm{QM/MM}}$ is computationally expensive within a PW scheme (~ $N_r N_{\mathrm{MM}}$, $N_r$ =space grid points, $N_{\mathrm{MM}}$ =classical atoms). A straightforward computation would increase the computational cost by several orders of magnitude.

# QM/MM in detail

Laio A. *et al.* J. Chem Phys.. 116: p. 6941-6947 (2002)

$$E_{\mathrm{QM/MM}} = \sum_{J\in\mathrm{MM}} q_J \int d\mathbf{r} \frac{\rho^{V}_{\mathrm{QM}} + \rho^{ION}_{\mathrm{QM}}}{\left|\mathbf{R}_J - \mathbf{r}\right|} + \frac{1}{2}\sum_{\substack{I\in\mathrm{QM}\\ J\in\mathrm{MM}}} 4\varepsilon_{IJ}\left[\left(\frac{\sigma_{IJ}}{r_{IJ}}\right)^{12} - \left(\frac{\sigma_{IJ}}{r_{IJ}}\right)^{6}\right] + \ldots$$

Focus on reproducing pattern and energetics of hydrogen bonds

**Solution for the short-range problem**

$$\sum_{J\in\mathrm{MM}} q_J \int d\mathbf{r} \frac{\rho^{V}_{\mathrm{QM}}}{\left|\mathbf{R}_J - \mathbf{r}\right|} \rightarrow \sum_{J\in\mathrm{MM}} q_J \int d\mathbf{r} \frac{\mathbf{r}_{cJ}^4 - \mathbf{r}^4}{\mathbf{r}_{cJ}^5 - \mathbf{r}^5}\rho^{V}_{\mathrm{QM}}(\mathbf{r})$$

$\mathbf{r}_{cJ}$ ~ covalent radius of atom $J$

**Solution for the long-range problem**

$$\sum_{J\in\mathrm{MM}} q_J \int d\mathbf{r} \frac{\mathbf{r}_{cJ}^4 - \mathbf{r}^4}{\mathbf{r}_{cJ}^5 - \mathbf{r}^5}\rho^{V}_{\mathrm{QM}}(\mathbf{r}) \rightarrow \sum_{J\in\mathrm{NN}} q_J \int d\mathbf{r} \frac{\mathbf{r}_{cJ}^4 - \mathbf{r}^4}{\mathbf{r}_{cJ}^5 - \mathbf{r}^5}\rho^{V}_{\mathrm{QM}}(\mathbf{r}) + C\sum_{J\notin NN}\frac{q_J}{\mathbf{r}_J - \bar{\mathbf{r}}} + \sum_{\alpha=1}^{3} D^{\alpha}\sum_{J\notin NN} q_J \frac{\mathbf{r}_J^{\alpha} - \bar{\mathbf{r}}^{\alpha}}{\left|\mathbf{r}_J - \bar{\mathbf{r}}\right|^3} + \sum_{\alpha,\beta=1}^{3} Q^{\alpha\beta}\sum_{J\notin NN} q_J \frac{\left(\mathbf{r}_J^{\alpha} - \bar{\mathbf{r}}^{\alpha}\right)\left(\mathbf{r}_J^{\beta} - \bar{\mathbf{r}}^{\beta}\right)}{\left|\mathbf{r}_J - \bar{\mathbf{r}}\right|^5} + \ldots$$

Multipolar expansion of the full interaction up to quadrupole

NN atoms = MM (non-neutral) atoms belonging to *charge groups* with at least one atom inside a shell of ray RCUT_NN

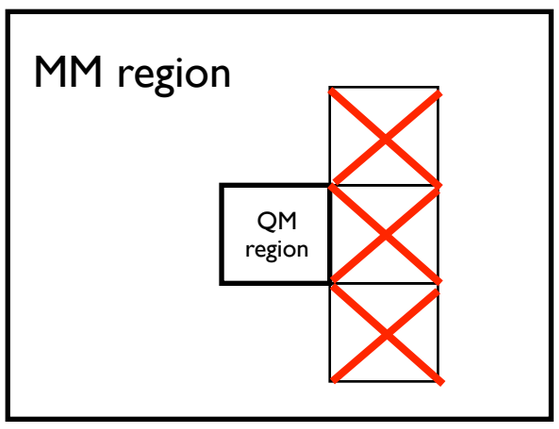
# Isolated System

We use PW to expand the wavefunctions:

$$\psi_i(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}}^{\mathbf{G}_{max}} c_i(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}$$

so, the QM charge density is intrinsically periodic!

This is not good in a QM/MM scheme where there is a QM box enbedded in a MM environment:




Decoupling scheme to remove the effect of the periodic images interactions: es. Martyna-Tuckerman

Martyna G.J. *et al.* J. Chem. Phys. 110(6): p. 2810-2821 (1999)

German Research School for Simulation Sciences

JÜLICH FORSCHUNGSZENTRUM

# Some examples

- Investigating proton dynamics
- QM/MM for optical properties
- Biological systems with metal ions
- QM/MM Force matching
- Formation of covalent bonds

# Proton dynamics: Importance in Biology

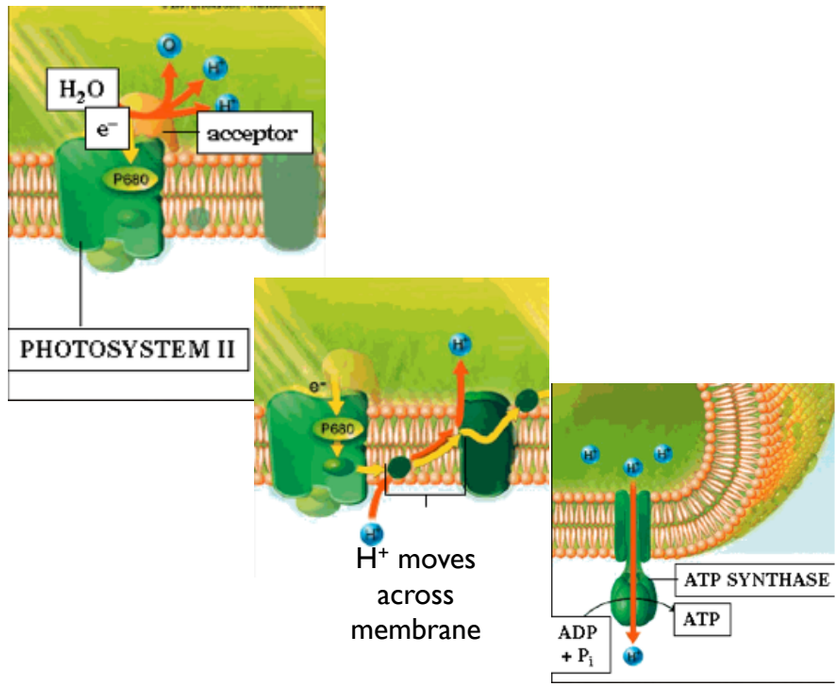
- Proton production and consumption processes play a pivotal role for the bioenergetics of all organisms.

- Most of these processes involve proton diffusion at cell membrane/water.

E.g. the synthesis of ATP relies on two types of membrane-bound enzymes:

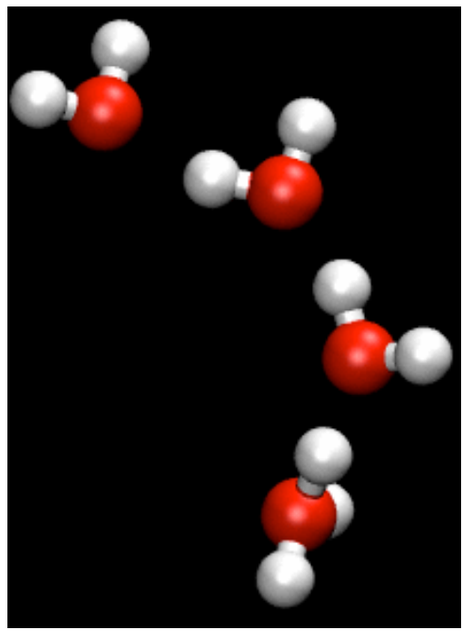
- Proton pumps creating transmembrane proton gradient

- ATP synthases consuming this transmembrane potential to drive ATP synthesis



# Proton dynamics: Grotthuss mechanism

An 'excess' proton or protonic defect diffuses through the hydrogen bond network of water molecules or other hydrogen-bonded liquids through the formation or cleavage of covalent bonds.



Proton-hopping mechanism

The Grotthuss mechanism, along with the relative lightness and small size of the proton, explains the unusually high diffusion rate of the proton relative to that of other common cations, which is due simply to random thermal motion, i.e. Brownian motion

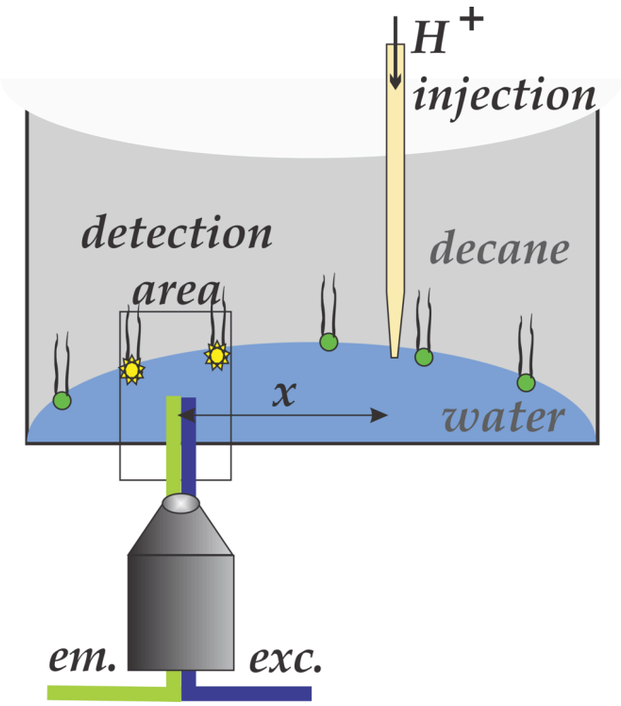
# Proton dynamics: Lateral diffusion along interfaces

- Direct experimental observation suggests fast ($10^{-4}$ $cm^2$ $s^{-1}$) and long-distance (~ 10 μm) proton surface diffusion along membrane interface.

- Surprisingly, such diffusion pattern is independent of the titratable head groups.

- Simulations performed on non-biological (hydrophobic) interfaces suggest that excess proton prefers the dielectric mismatched interfaces.

- Experiments on water/*n*-decane interface find that the fast and long-distance diffusion of the excess proton still remains



# Proton dynamics: CPMD on hydrophobic interface

German Research School for Simulation Sciences

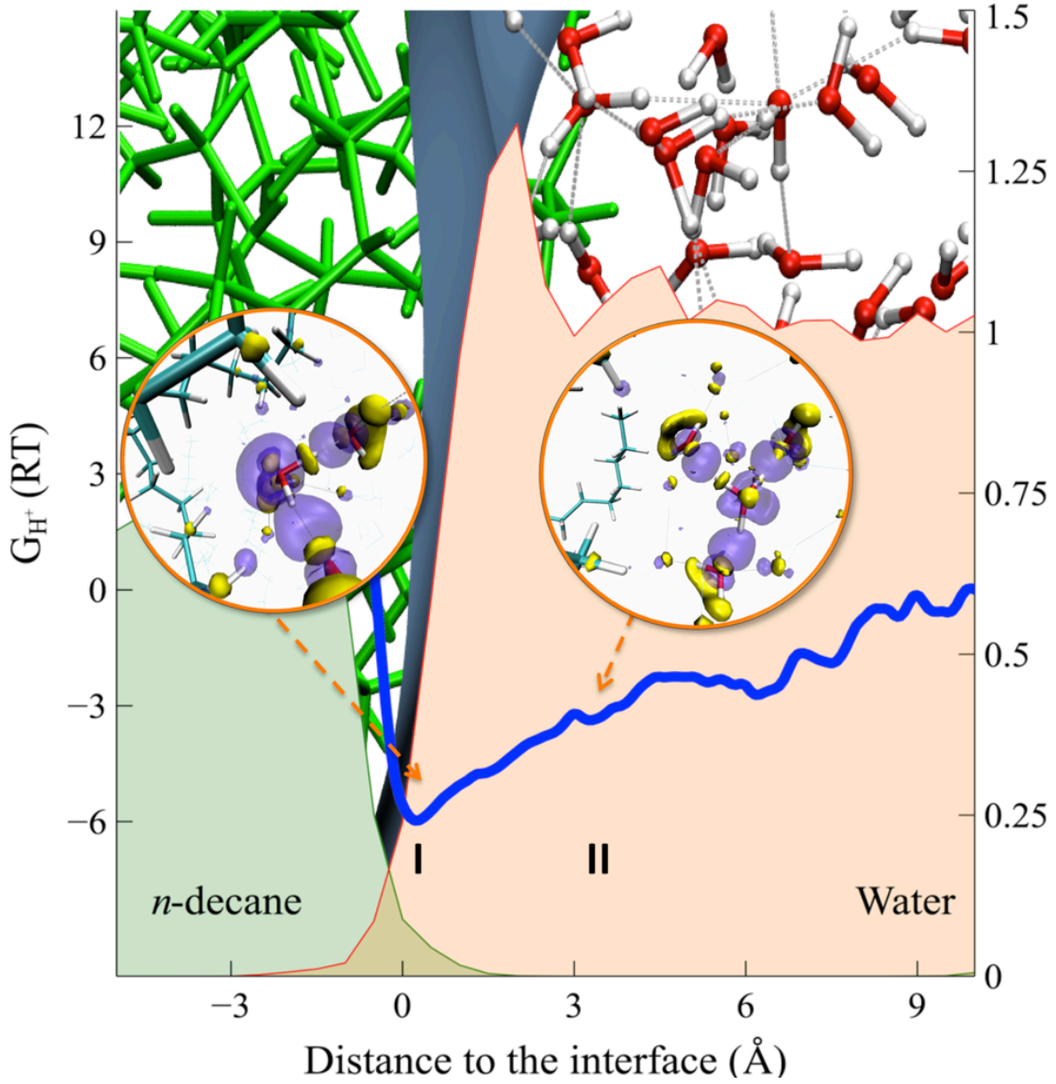
JÜLICH FORSCHUNGSZENTRUM

- The free energy profile $G_{H^+}$ is calculated by 75 ps-long CPMD simulations based on metadynamics.

- A wide minimum of $G_{H^+}$ is located within 6 Å from the hydrophobic surface with a depth of 6 RT and an estimated statistical error of about 2 RT.

- Within this minimum, two populations of the excess proton (**I** and **II**) are observed.

- Population II represents the main species which diffuses fast along the interface, because the adjacent water molecules form more H-bonds on average, than the water molecules adjacent to population I.



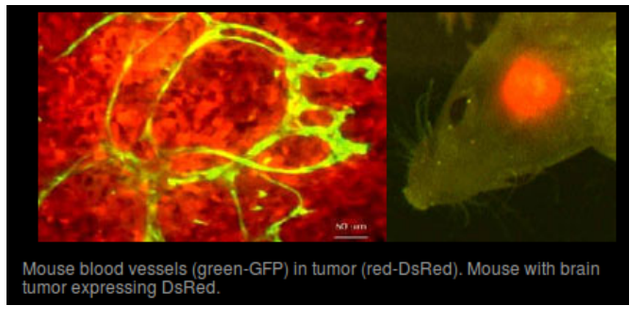

Zhang C. *et al.* Proc. Natl. Acad. Sci. USA 109: p. 9744-9749 (2012)

# Optical properties: Fluorescence probes

German Research School for Simulation Sciences

JÜLICH FORSCHUNGSZENTRUM

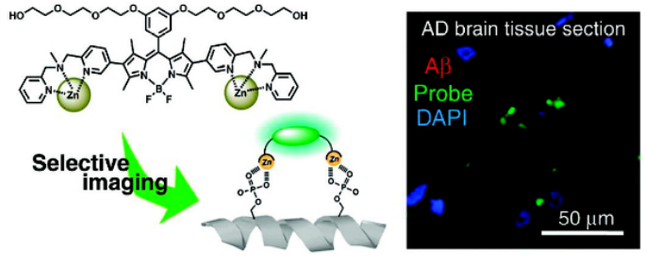
Fluorescence probes are everyday used in biological experiments

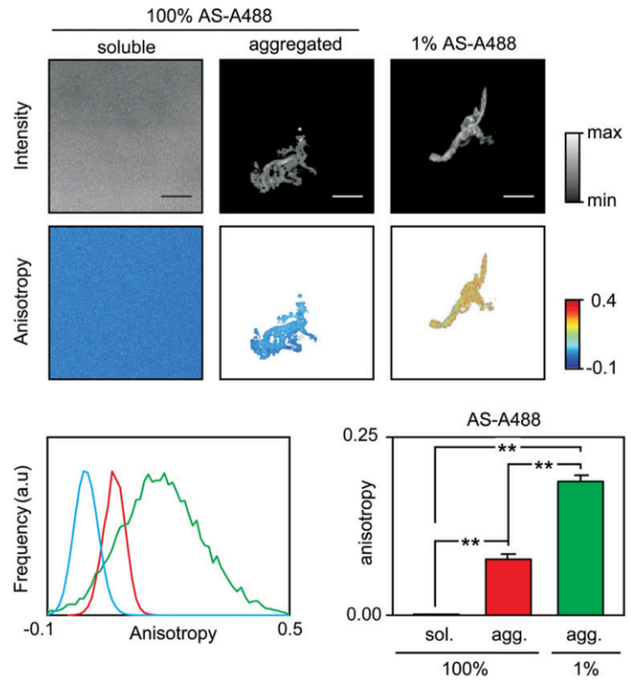



source: http://www.conncoll.edu/ccacad/zimmer/GFP-ww/GFP-1.htm

YET

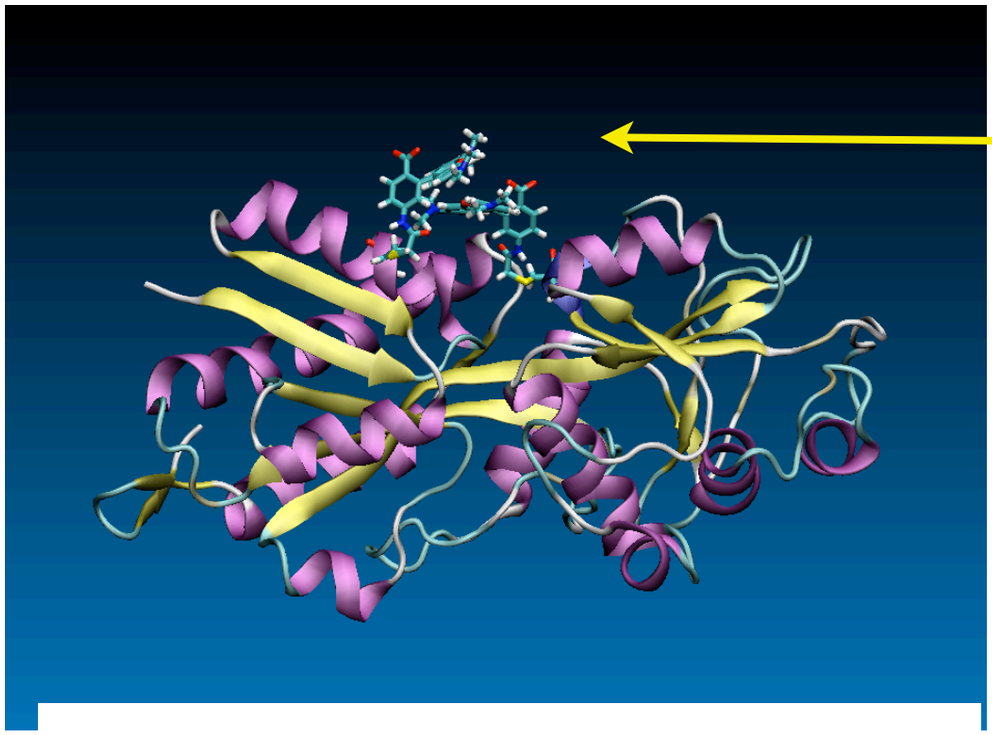
**No structural model of the probe-system interaction *in vivo*!**

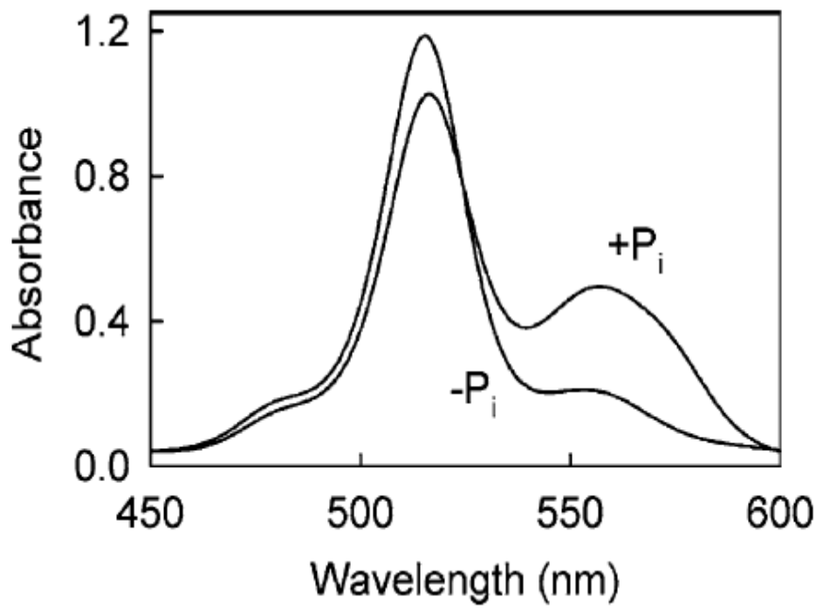


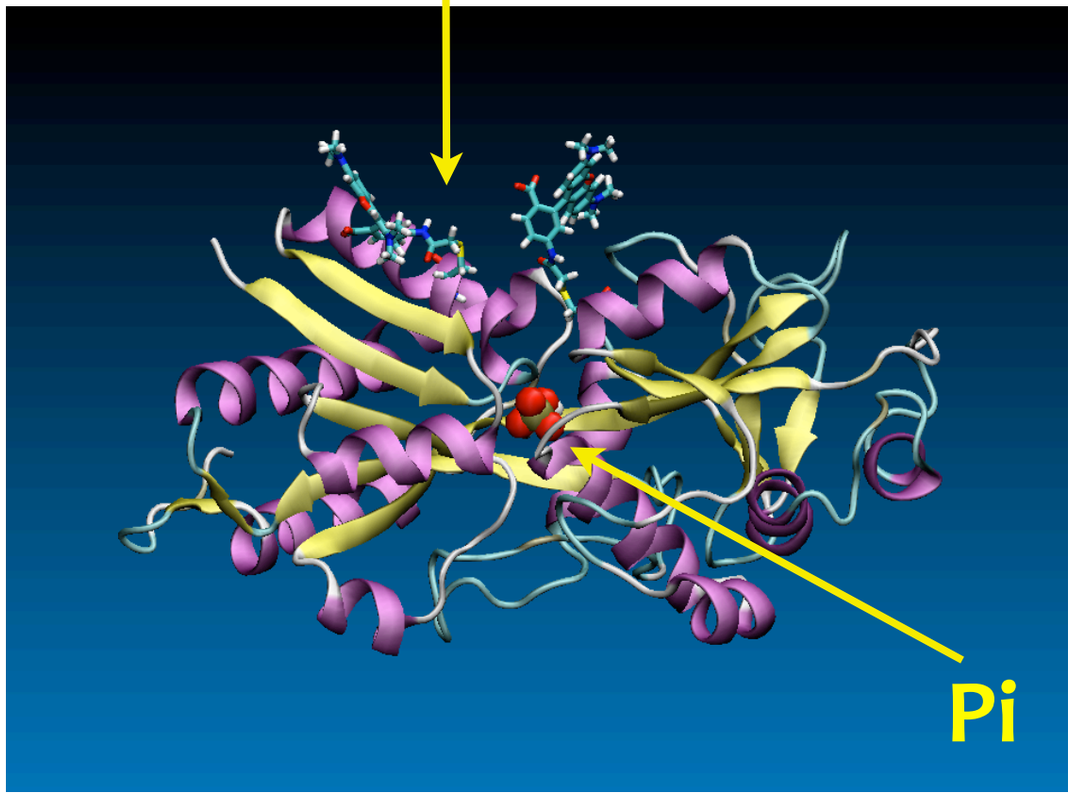

- Only coarse interpretation of the spectra
- No control of the effects of probe-macromolecules interactions *in vivo*

# Optical properties: 2-Rhodamines-PBP probe






**6IATR RHODAMINES: which conformation?**



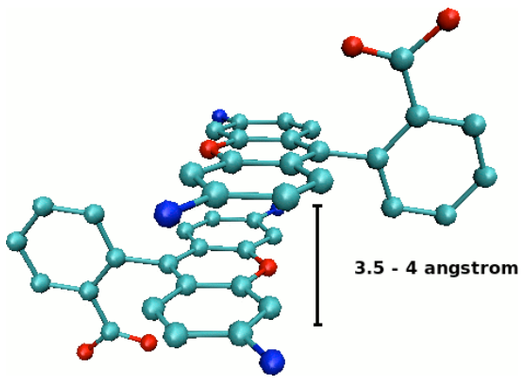
# Optical properties: Staking

German Research School for Simulation Sciences

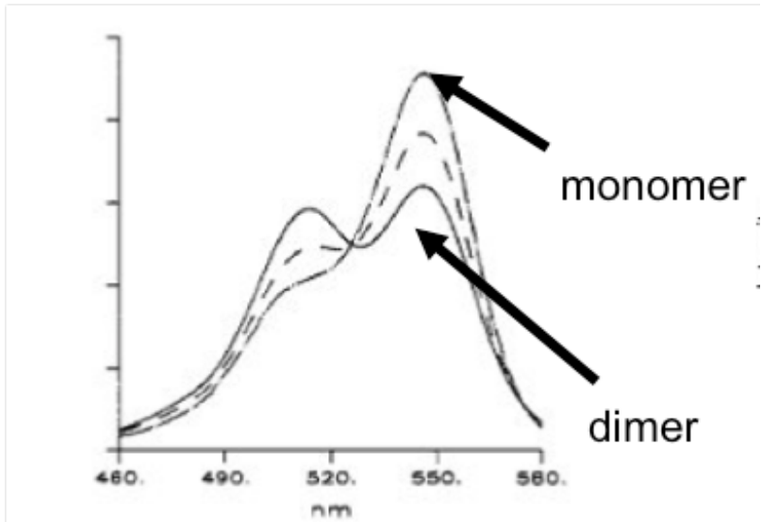
JÜLICH FORSCHUNGSZENTRUM

**In WATER SOLUTION**

Dimerization and optical properties:

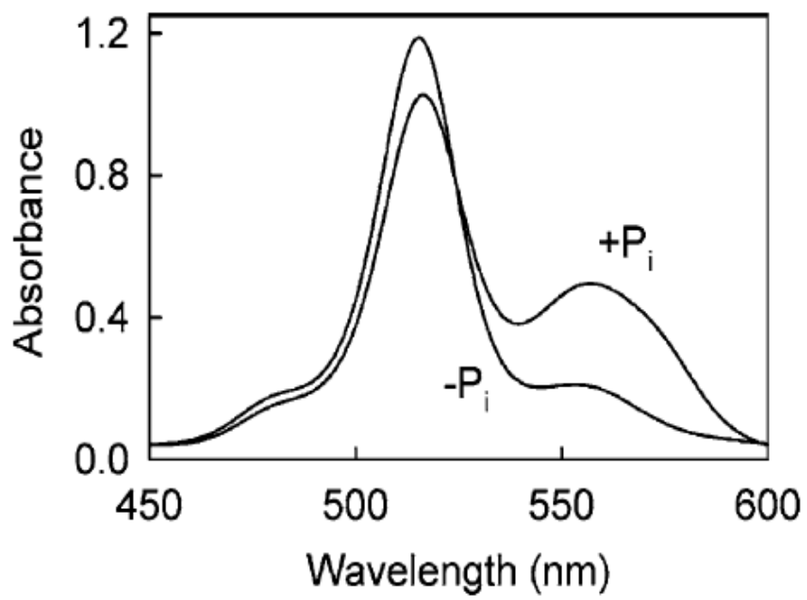



**Absorption spectrum:**

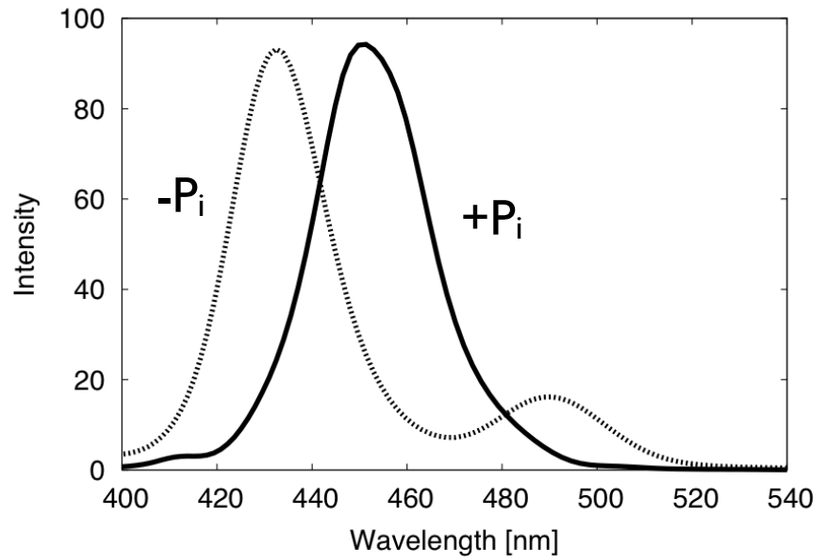



**In COMPLEX**

**EXP**




**QM/MM**




Gonçalves M.B. *et al.* PCCP 15: p. 2177-2183 (2013)

# Metal ions:
# Role in neurodegenerative diseases

German Research School for Simulation Sciences

JÜLICH FORSCHUNGSZENTRUM

- There is now increasing evidence that altered metal homeostasis may be involved in the progression of neurodegenerative diseases:

  - Sayre, L. M.; Perry, G.; Smith, M.A. Curr. Opin. Chem. Biol. **1999**, 3, 220-225.
  - Gaeta, A.; Hider, R. C. Br. J. Pharmacol. **2005**, 146, 1041-1059.
  - Molina-Holgado, F. *et.al.* Biometals **2007**, 20, 639-654.

- Protein-metal interactions appear to play a critical role in protein aggregation involved in Parkinson's disease:

  - Paik, S. R.; Shin, H. J.; Lee, J. H.; Chang, C. S.; Kim, J. Biochem. J. **1999**, 340, 821-828.
  - Requena, J. R. *et al.* Proc. Natl. Acad. Sci. U.S.A. **2001**, 98, 7170-7175.
  - Atwood, C. S. *et al.* J. Biol. Chem. **1998**, 273, 12817-12826.

- and recently has been emphasized the role of copper, iron and zinc as contributors both to amyloid Aβ assembly *in vitro* and to the neuropathology of the Alzheimer´s disease:

  - Bush, A. I. *et al.* J. Biol. Chem. **1993**, 268, 16109-16112.
  - Miura, T.; Suzuki, K.; Kohata, N.; Takeuchi, H. Biochemistry **2000**, 39, 7024-7031.
  - Karr, J. W.; Szalai, V. A. Biochemistry **2008**, 47, 5006-5016.

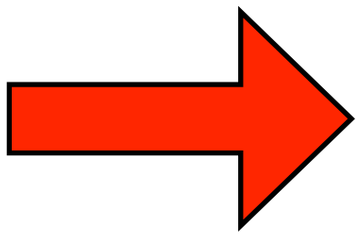
# Metal ions:
# Cu(II) in Neurodegenerative Diseases

- High levels of Cu (and also Zn and Fe) were found in and around amyloid plaques of Alzheimer´s disease brains.

- Elevated Cu concentrations have been reported in the cerebrospinal fluid of Parkinson's disease patients.

- Individuals with chronic industrial exposure to Cu (but also Mn and Fe) have an increased rate of Parkinson's disease

- Cu(II) ions (also Al and Fe but at higher concentrations) have been shown to bind α-synuclein and accelerate its fibrillation *in vitro*.



Need of structural characterization of
Cu(II) – α-synuclein interactions

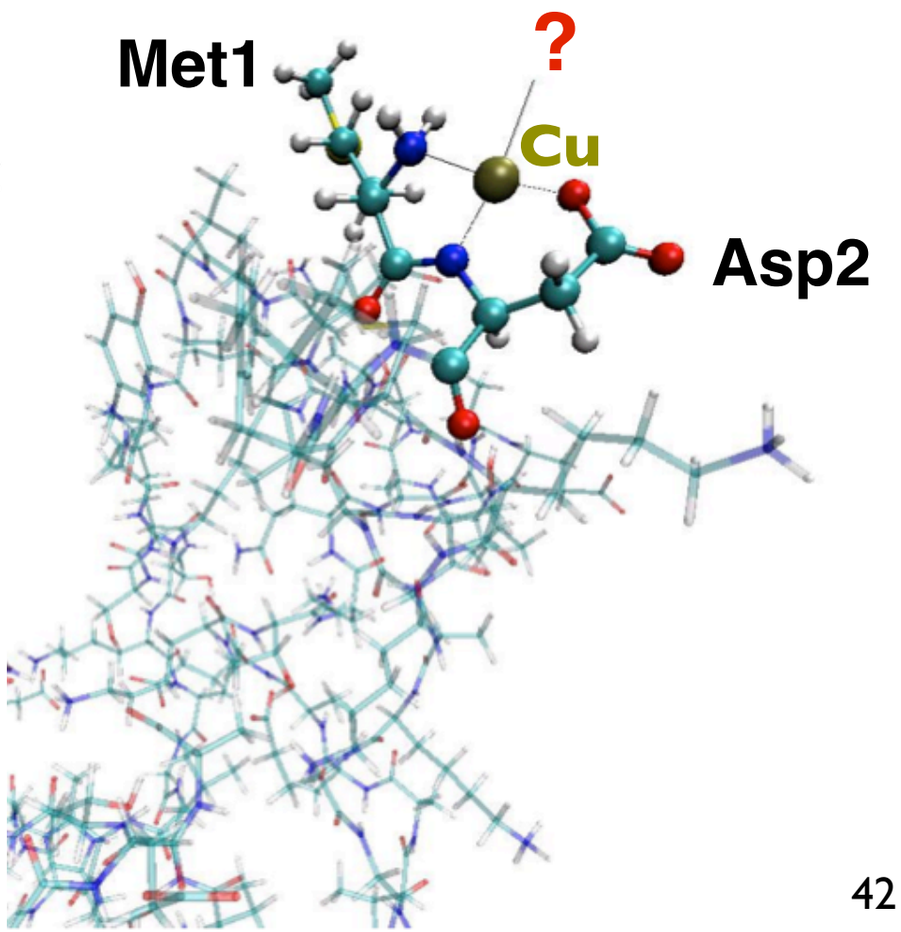
# Metal ions: Experimental evidence

Cu(II) binds specifically to α-synuclein and is effective in accelerating aggregation at physiologically relevant concentrations.

Rasia, R. M.; Bertoncini, C. W.; Marsh, D.; Hoyer, W.; Cherny, D.; Zweckstetter, M.; Griesinger, C.; Jovin, T.; Fernández, C. O. Proc. Natl. Acad. Sci. U.S.A. 2005, 102, 4294-4299

NMR mapping revealed the most affected regions are located at the N-terminus.

In particular the highest affinity binding site involves:

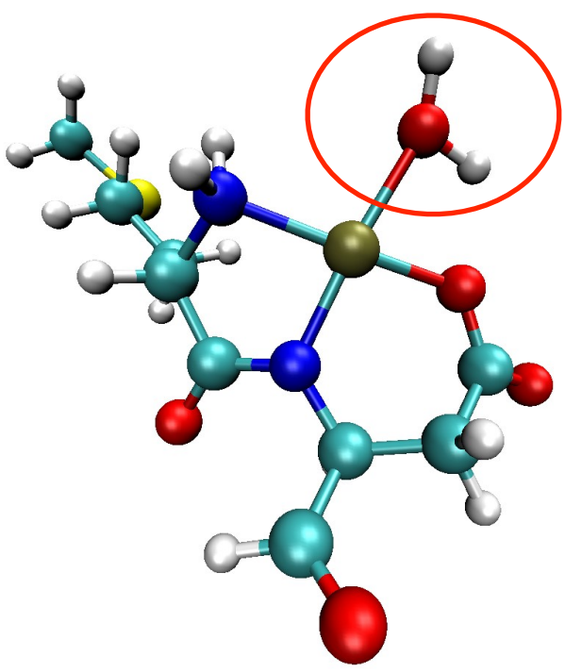



Unfortunately, due to the properties of Cu(II) binding to this site (long residence time of the metal ion in the site and therefore slow exchange), obtaining a structure based on NMR data has proven rather difficult up to now.

# Metal ions: Possible candidates

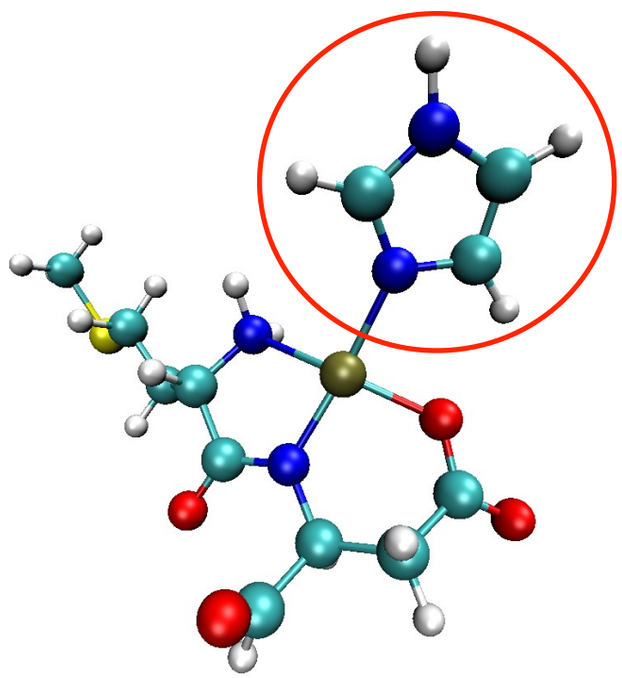
Experimental data from Fernández's Labs and free energy calculations on simple models, reduce the possible candidates to the fourth atom coordinating Cu(II), to two alternatives:

**Water molecule**



ΔG = -15.10227 Kcal/mol

**His50**

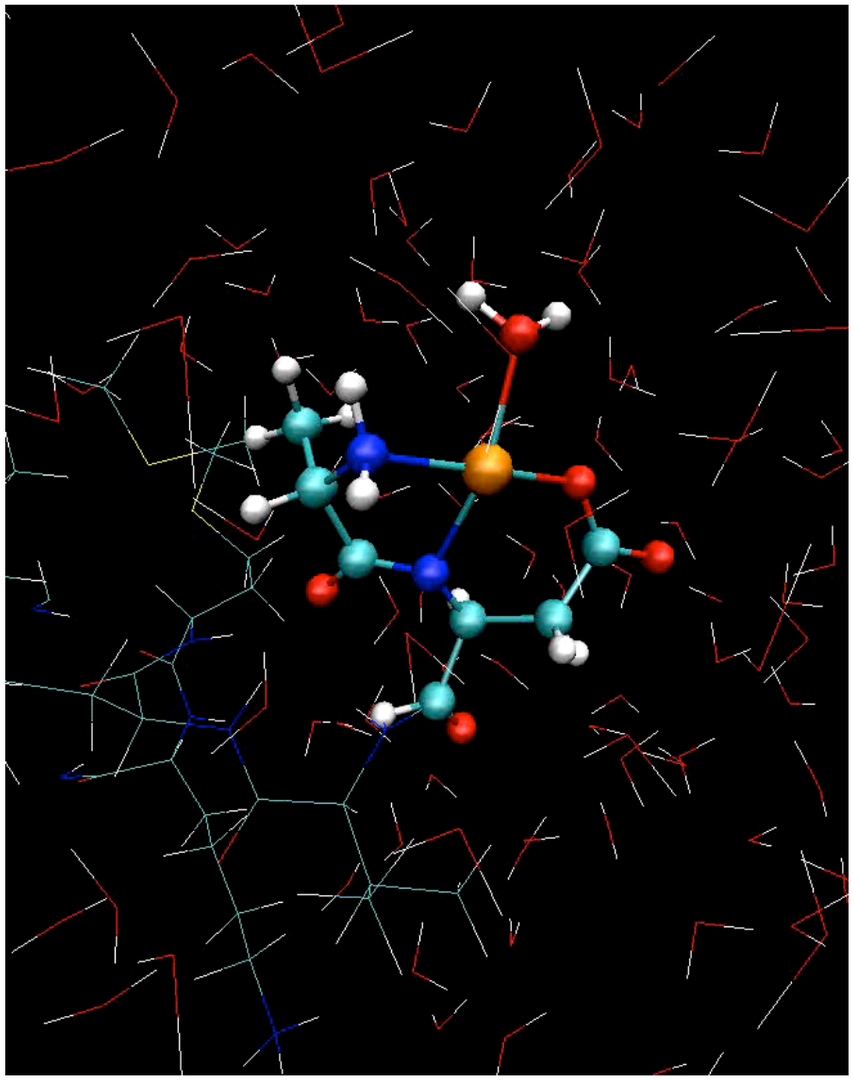


ΔG = -15.73982 Kcal/mol

# Metal ions:
# QM/MM Car-Parrinello MD results

German Research School for Simulation Sciences

JÜLICH FORSCHUNGSZENTRUM

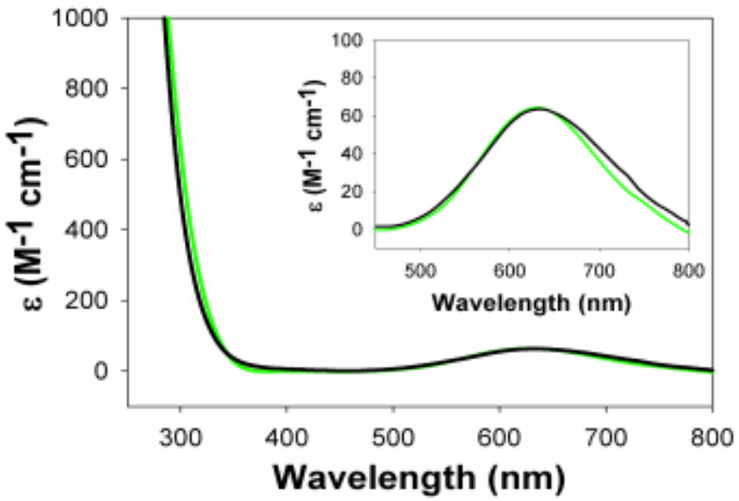


T = 310 K (Physiological conditions)

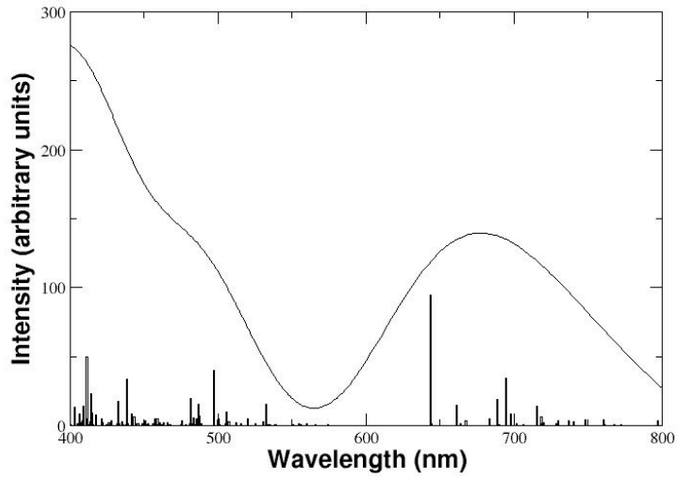
T = 300 K (Abs. Exp. conditions)

T = 77 K (EPR Exp. Conditions)

## Optical Absorption Spectrum




TDDFT Calculations

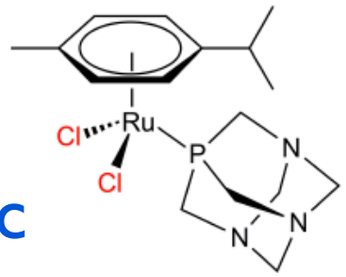



| EPR | Highest Affinity Site | Second Site | Lowest Affinity Site | Mean Theoretical Value |
|---|---|---|---|---|
| $g_z$ | **2.248** | 2.330 | 2.365 | **2.21±0.16** |
| $A_z$ $(10^4\ cm^{-1})$ | **191** | 163 | 152 | **187.9±3.6** |
| Possible Coordination | **$N_2O_2$** | $NO_3$ | $O_4$ | **$N_2O_2$** |

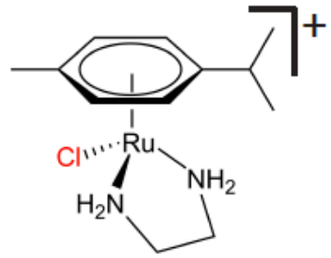
Binolfi A. *et al.* Inorganic Chemistry. 49: p. 10668-10679 (2010)

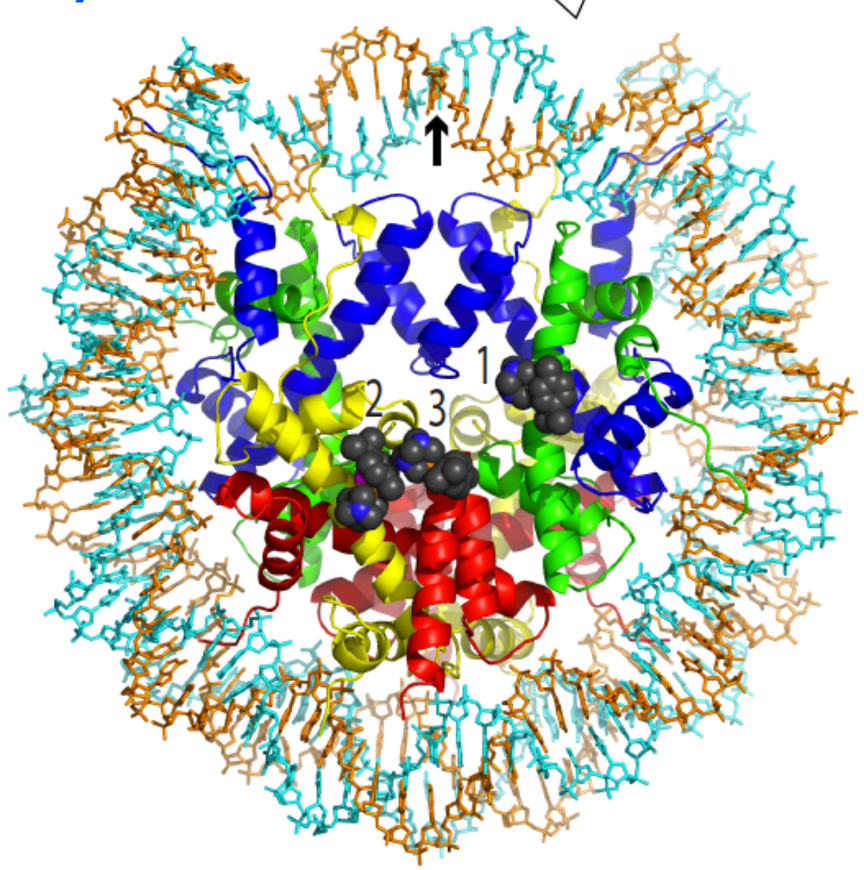
# Covalent bonds:
# Ruthenium-based anticancer drugs

RAPTA-C
non-cytotoxic

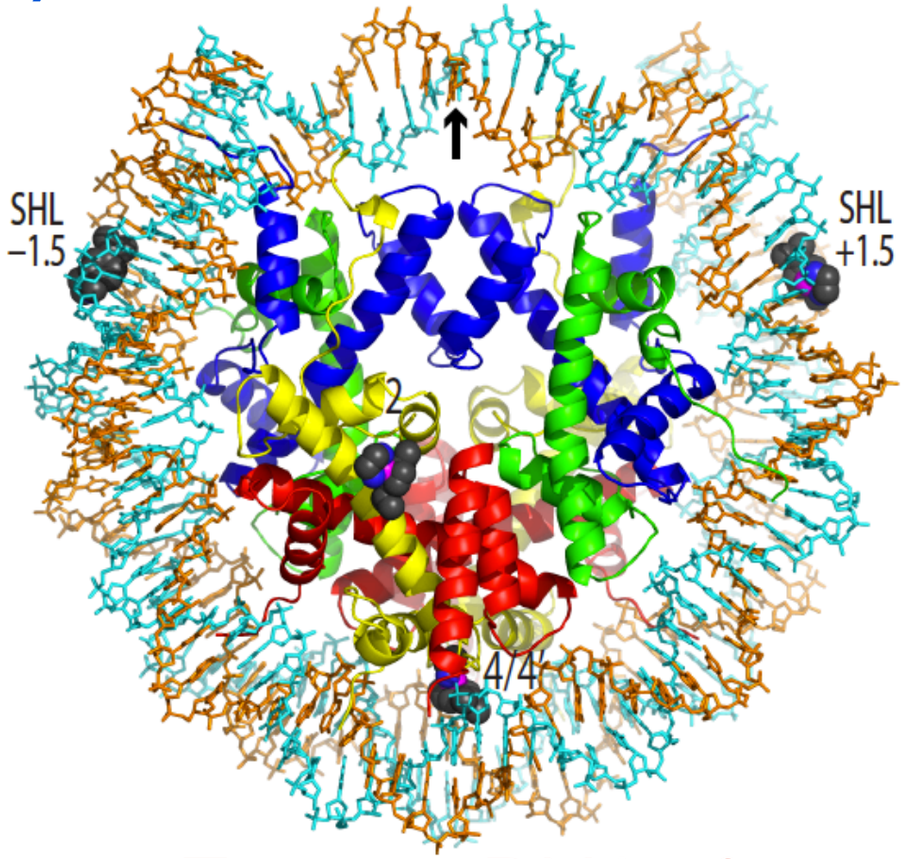



RAED-C
cytotoxic



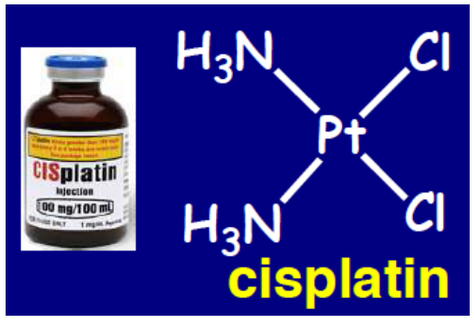

Nucleosome




Adducts with the histone proteins

Targeting DNA of Chromatin

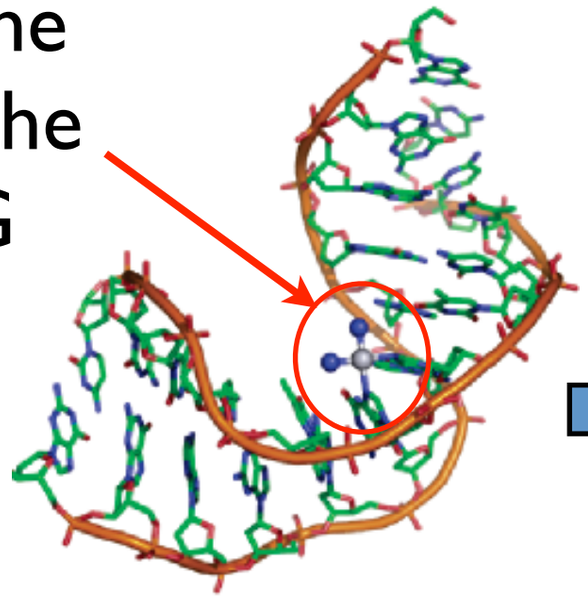
# Metal ions 2: Cisplatin




- One of the most widely used anticancer drug, active against testicular, ovarian, cervical, colorectal, small-cell lung, … cancers.

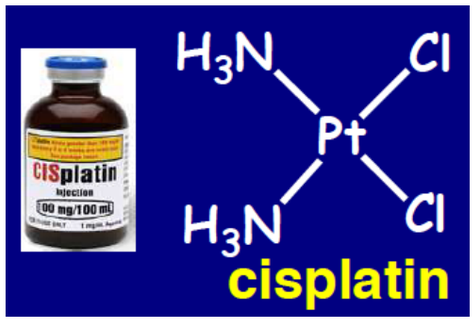
- Discovered in 1965 and approved for clinical use in 1978.

- Once injected, cisplatin enters the cell and binds preferentially to the N7 atoms of two neighboring G bases within a strand (1,2-intrastrand crosslink adduct).



Cell apoptosis

# Metal ions 2: Cisplatin resistance

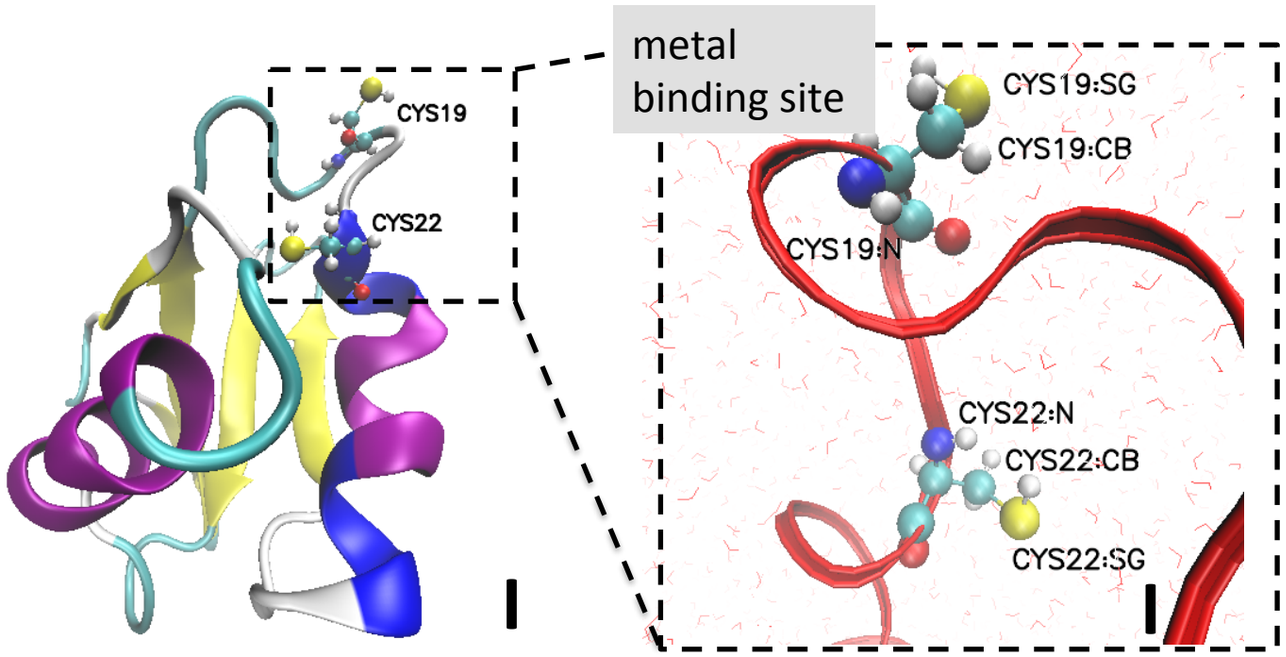



- Unfortunately, the efficacy of cisplatin-based therapies is strongly limited by the emergence of RESISTANCE mechanisms: only ~3% of cisplatin reaches the nucleus.

**RESISTANCE**:

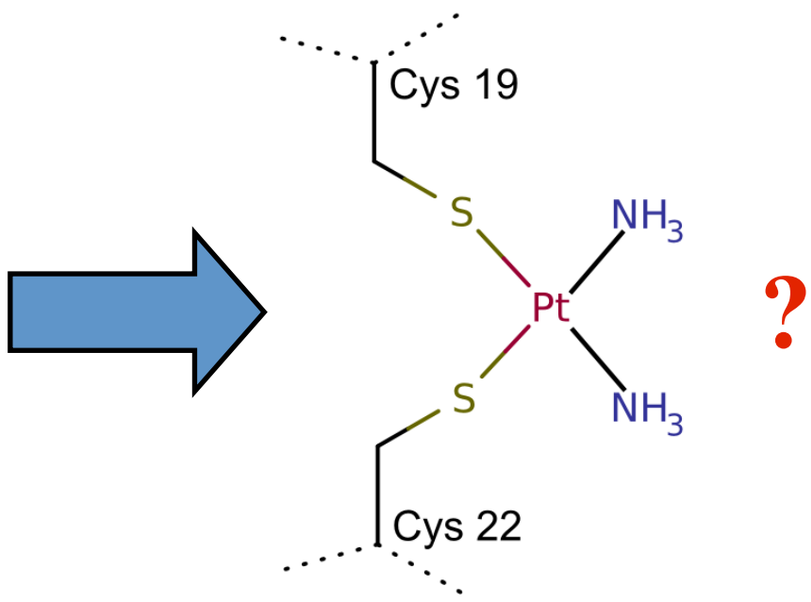
- Reduce drug accumulation:
  - Decrease of drug uptake
  - Increase of drug efflux and sequestration
- Cellular transport of cisplatin (copper transport proteins):
  - Ctr1 (uptake)
  - Atox1 (down regulation of Ctr1 and sequestration)
  - **ATP7A**/ATP7B (sequestration and efflux)

# Metal ions 2: Mnk1

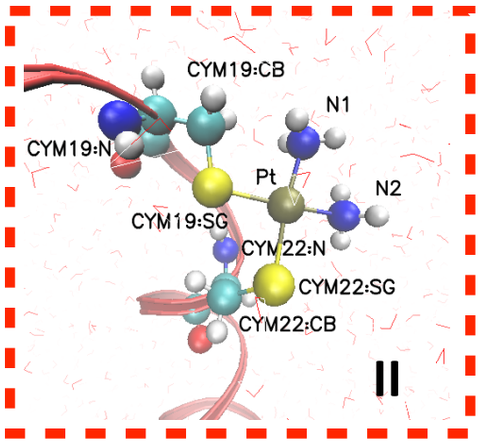



A quantitative 3D structural model of the platinated protein does not exist.

ESI-MS and $^{1}H,^{13}C$ and $^{1}H,^{15}N$ HSQC measurements suggest that in aqueous solution cisplatin loses Cl ligands and binds to the metal binding site of the first cythosolic domain of ATP7A, called **Mnk1**.



# Metal ions 2: 4 possible binding modes




## $^{13}C$ NMR Chemical Shift

| Residue name | Atom name | I | II | A | B | C | *Apo*-Mnk1 experimental value | Platinated Mnk1 experimental value |
|---|---|---|---|---|---|---|---|---|
| Cys19 | $C_\beta$ | 28.2(4.2) | 31.1(4.8) | 30.5(4.4) | 28.7(4.5) | 27.0(5.1) | 27.9 | 29.4 |
| Cys22 | $C_\beta$ | 30.1(4.6) | 31.4(5.0) | 37.0(5.0) ☹ | 30.4(4.9) | 25.9(4.3) ☹ | 25.6 | 28.5 |

Deviation larger than the mean error of the theory (~7.6 ppm)

Upfield while a downfield shift is observed experimentally

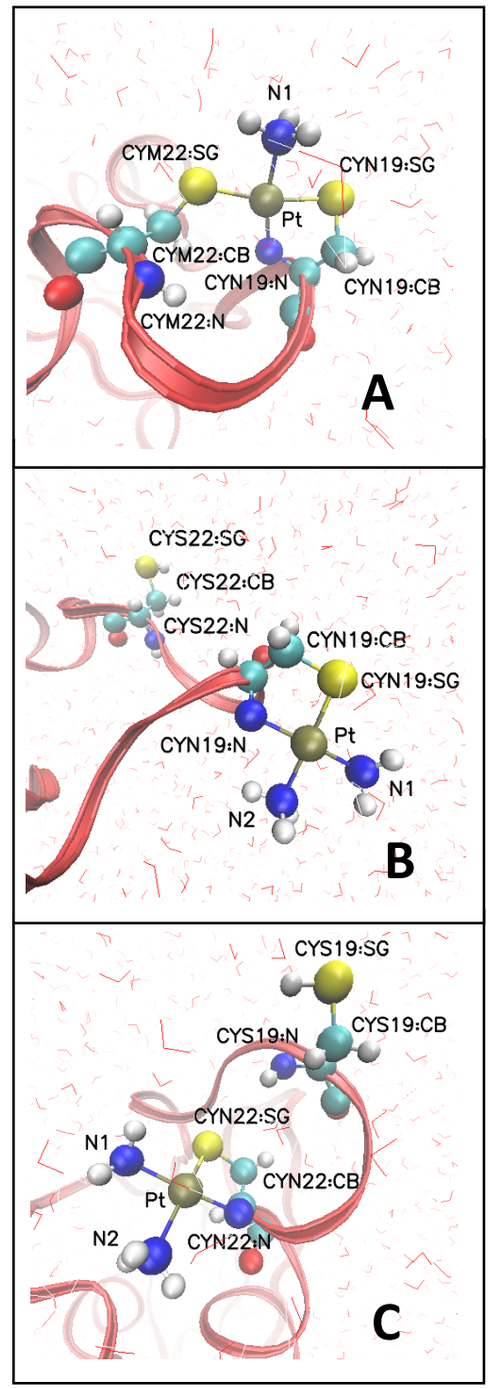
**Only model II is compatible with all experimental data**

## $^{15}N$ NMR Chemical Shift

| Atom name | Cisplatin | II | A | B | C | Atom name | Cisplatin experimental value | Mnk1-coordinated Cisplatin experimental value |
|---|---|---|---|---|---|---|---|---|
| N1 | ref | -57.2(11.7) | -74.7(11.3) | -78.1(9.9) | -72.9(10.2) | N* | -67.1, -67.1 | -35.8, -42.8 |
| N2 | ref | -48.2(10.8) | | -50.6(9.9) | -52.8(11.1) | | | |

☹

Deviation larger than the mean error of the theory (~13-16 ppm)




Calandrini V. *et al.* Dalton Transaction. In press (2014)

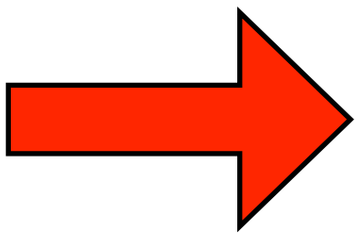
# Metal ions 2: CD spectra

- To cross-check the results we employed the experimental Circular Dichroism (CD) spectra information on Pt-Mnk1.

- CD spectra give information about secondary structure elements content.

Typical timescale ~ ns-μs

Far beyond CPMD simulations time windows!



**QM/MM Force matching**

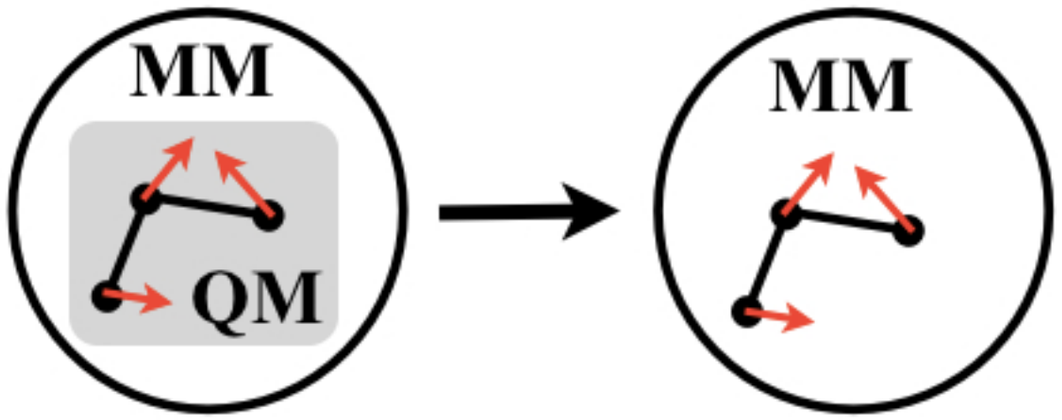
+ classical MD

# Metal ions 2: QM/MM Force matching

Parameterisation of biomolecular force fields based on QM/MM reference calculations:

A. FIT OF ATOMIC CHARGES: Derivation of a set of D-RESP charges $q_\alpha$ for the QM atoms by imposing they have to reproduce:

1. The **electric field** on a grid that is defined by the positions of the NN atoms.
2. The **potential** on a grid that is defined by the positions of the NN atoms.
3. A weak **restraint** to their respective Hirshfeld values.




$$\chi^2\left(\{q_\alpha\}\right)=\sum_{l=1}^{L}\Big[\sum_{\beta\in NN_l}\left(w^V\left(\mathrm{V}_{\beta l}^{\mathrm{MM}}-\mathrm{V}_{\beta l}^{\rho}\right)^2+w^E\left|\mathbf{E}_{\beta l}^{\mathrm{MM}}-\mathbf{E}_{\beta l}^{\rho}\right|^2\right)+$$

$$+\sum_{\alpha\in\mathrm{QM}}w^H\left(q_\alpha-q_{\alpha l}^H\right)^2\Big]+w^Q\left(Q^{\mathrm{tot}}-\sum_{\alpha\in\mathrm{QM}}q_\alpha\right)^2$$

To exclude chemically unreasonable and strongly conformation dependent solutions


Mauer P. *et al.* J. Chem. Theory Comput. 3: p. 628-639 (2007)

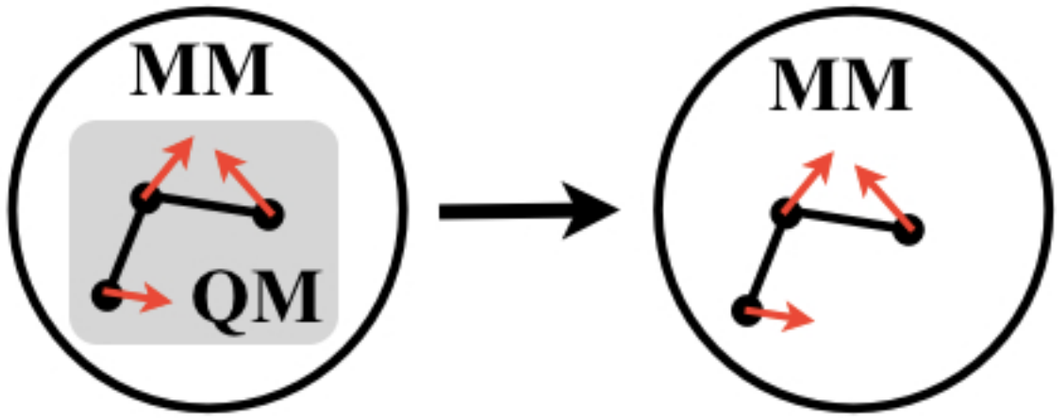
Doemer M. *et al.* J. Chem. Theory Comput. 10:p. 412-422 (2014)

# Metal ions 2: QM/MM Force matching

Parameterisation of biomolecular force fields based on QM/MM reference calculations:

B. FORCE MATCHING OF THE BONDED INTERACTIONS:

- After the set of charges has been determined, the classical **non-bonded forces** due to the electrostatics employing the new charges and the van der Waals interactions is calculated.
- Then, the non-bonded forces are **subtracted from the QM/MM reference forces** and the penalty function under variations of the bonded parameters $\tau_n$ minimized.

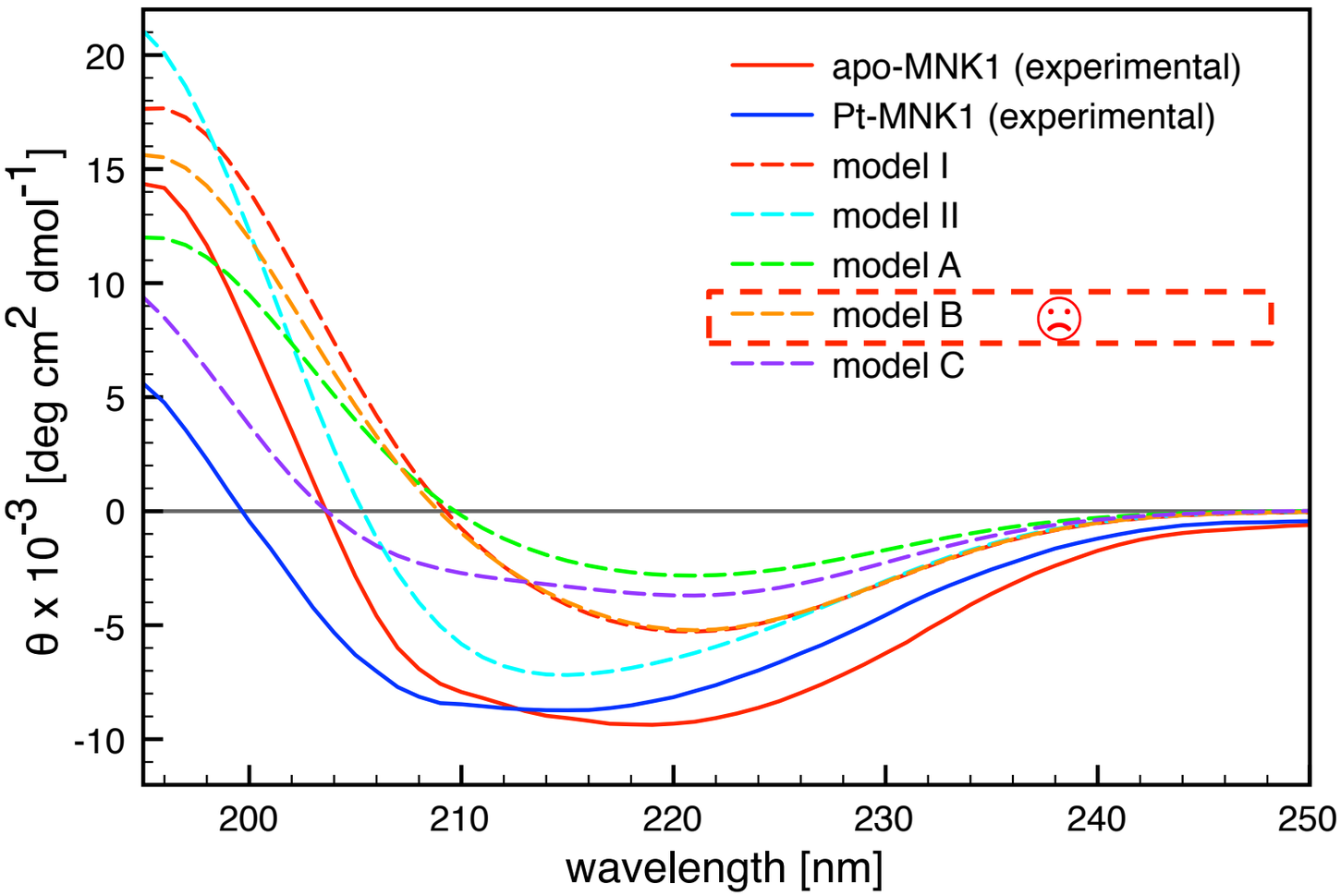



$$\sigma^2\left(\{\tau_n\}\right) = \sum_{l=1}^{L} \sum_{\alpha \in \mathrm{QM}} \left\| \mathbf{F}_{l\alpha}^{\mathrm{MM}_{bonded}} - \left( \mathbf{F}_{l\alpha}^{\mathrm{QM}} - \mathbf{F}_{l\alpha}^{\mathrm{MM}_{nonbonded}} \right) \right\|^2$$


Mauer P. *et al.* J. Chem. Theory Comput. 3: p. 628-639 (2007)

Doemer M. *et al.* J. Chem. Theory Comput. 10:p. 412-422 (2014)

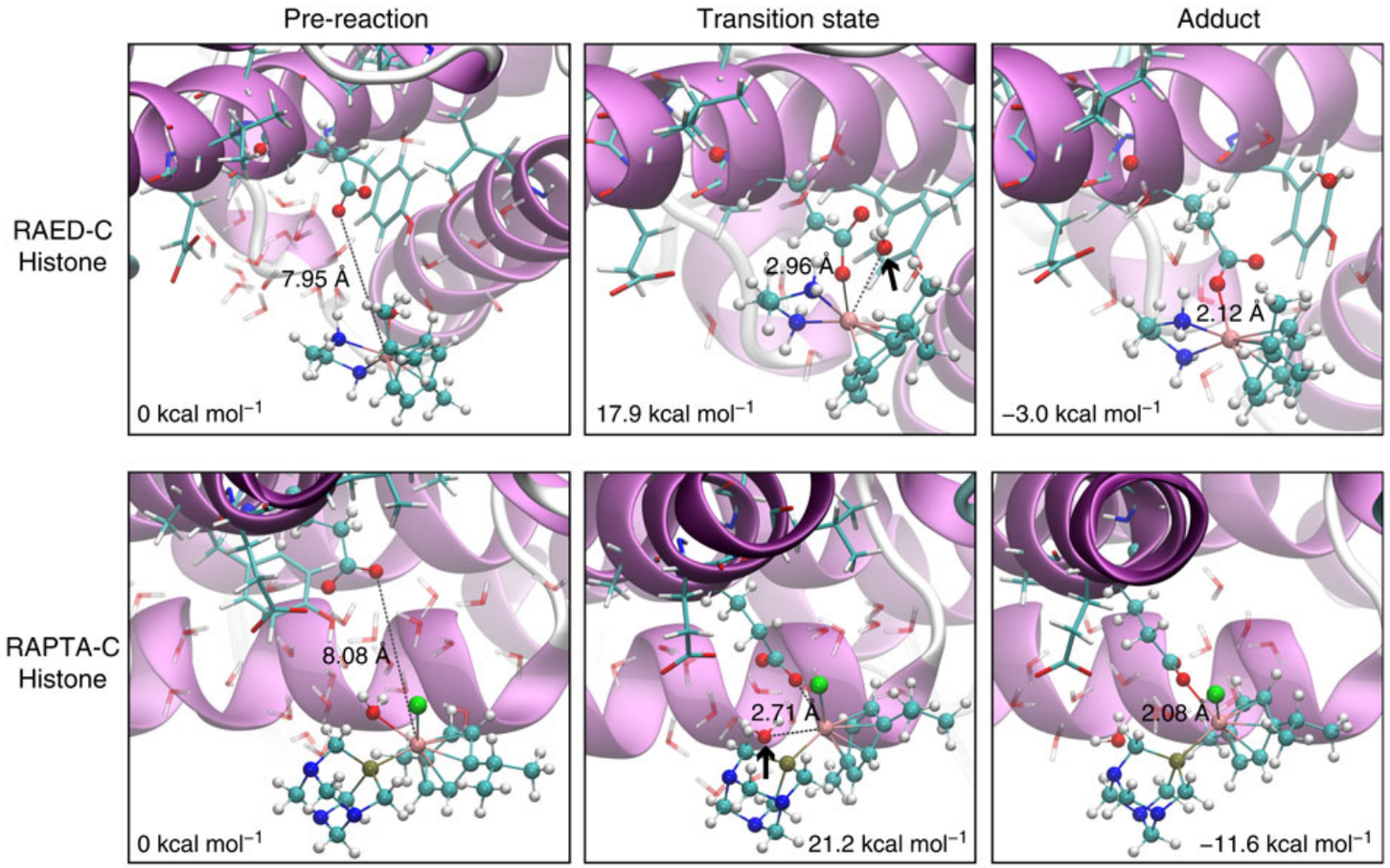
# Metal ions 2: CD spectra results

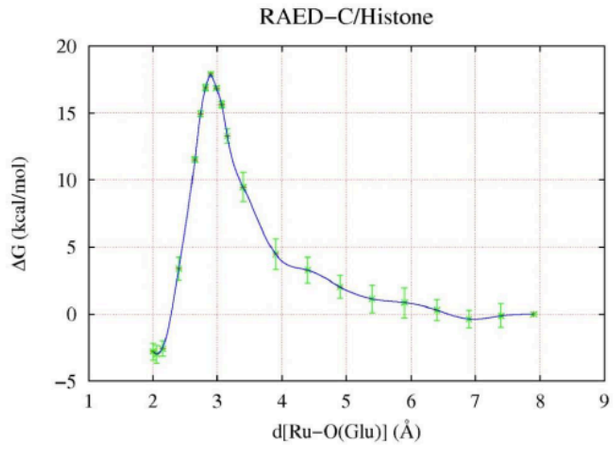


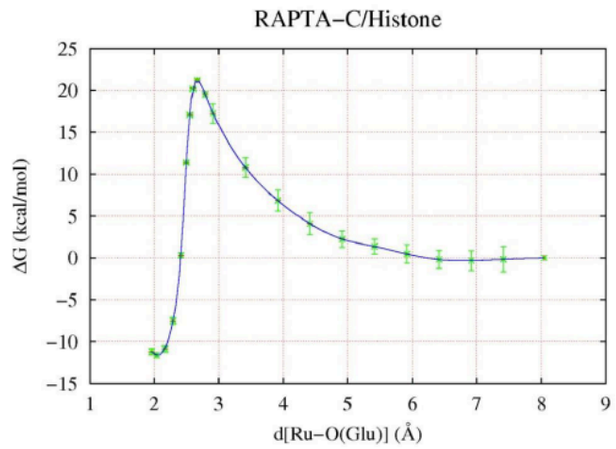

- Model B is very close to model I (apo). This is in disagreement with the experiment.
- Model II qualitatively reproduces the shift of the maximum negative ellipticity toward lower wavelengths, which indicates a loss of helical structure.
- Model I (apo) is in fair agreement with the experimental data.

# Covalent bonds: In Histone

- Similar free-energy barriers
- Binding thermodynamically favored



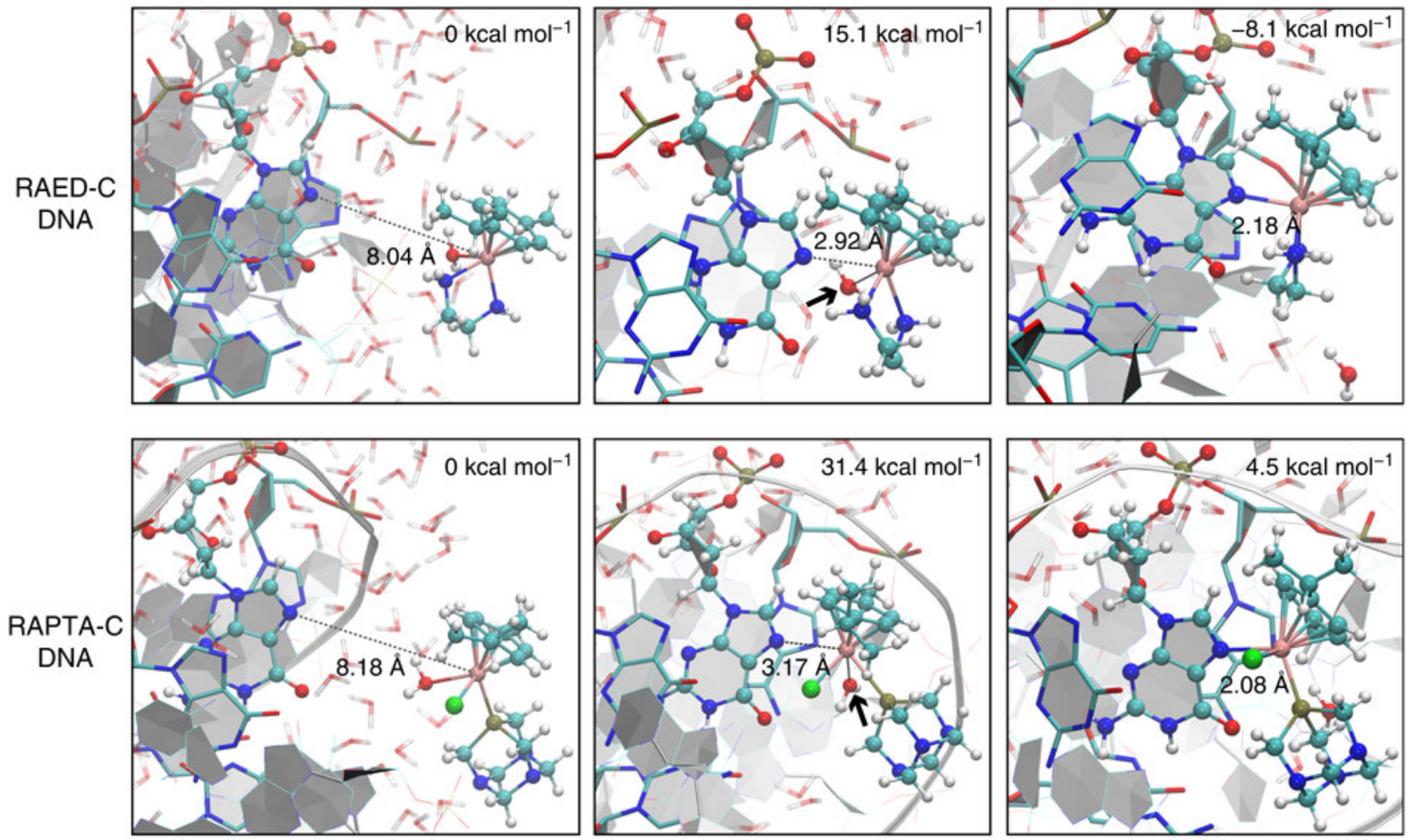



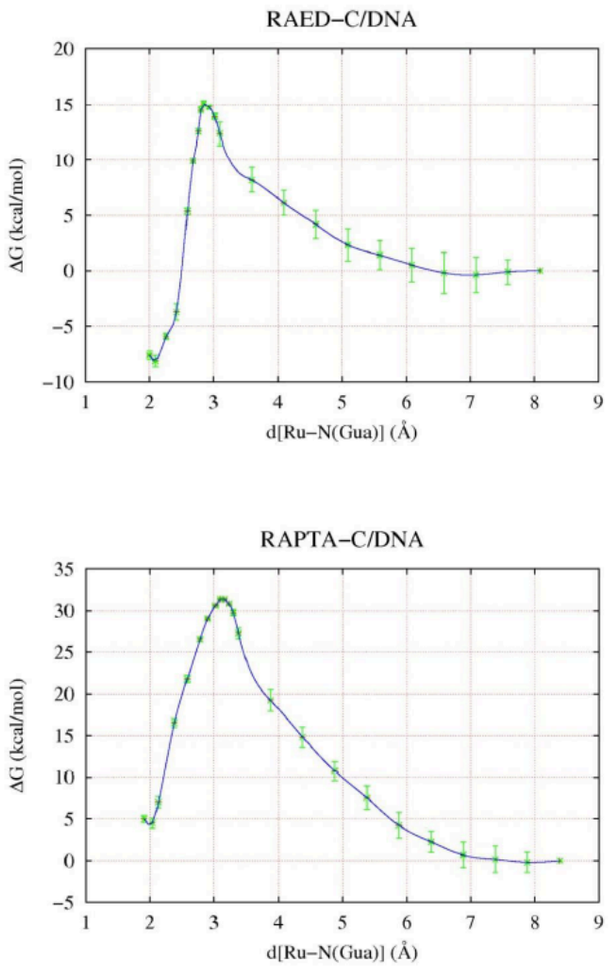

CPMD + TI study

Adhireksan Z. *et al.* Nature Comm. 5: p. 3462-639 (2014)

# Covalent bonds: in Chromatin

- RAPTA-C/DNA free-energy barrier very high
- RAPTA-C/DNA binding thermodynamically unfavored




CPMD + TI study

Adhireksan Z. *et al.* Nature Comm. 5: p. 3462-639 (2014)

# Covalent bonds:
# DNA or Protein binding?

## **RAED-C/histone vs <u>RAED-C/DNA</u>**

Second case thermodynamically (by about 3 kcal/mol) and kinetically favored

→ Preferential binding to DNA

## **<u>RAPTA-C/histone</u> vs RAPTA-C/DNA**

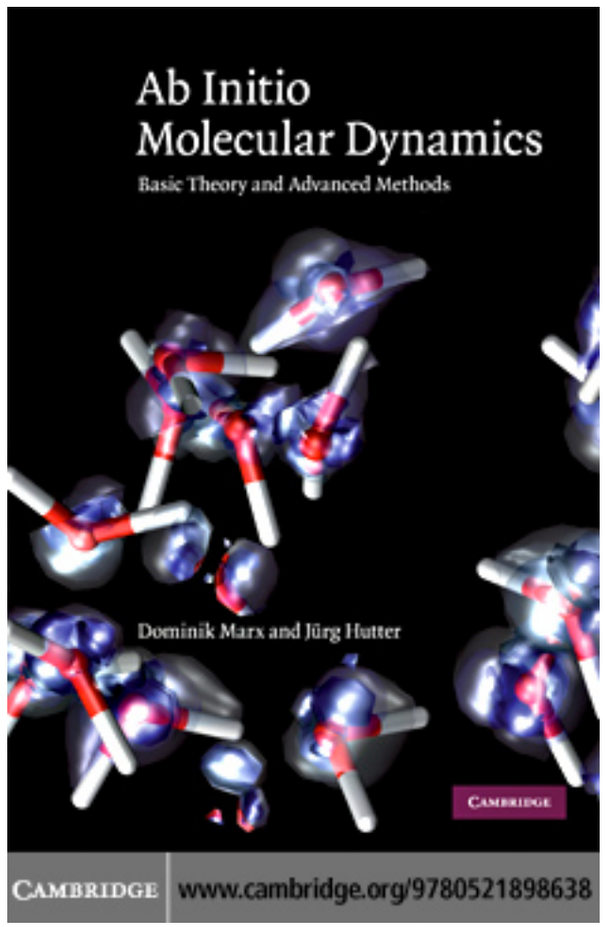
Second case thermodynamically unfavorable and kinetically forbidden (too high free energy barrier)

→ Exclusive binding to histone protein

Adhireksan Z. *et al.* Nature Comm. 5: p. 3462-639 (2014)

# Suggested readings

D. Marx & J. Hutter
**Ab Initio Molecular Dynamics**
Basic Theory and Advanced Methods
(2009)




D. Marx & J. Hutter
**Ab Initio Molecular Dynamics: Theory and Implementation**
in Modern Methods and Algorithms of Quantum Chemistry,
Proceedings, II Ed., J. Grotendorst (Ed.),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. 3, ISBN 3-00-005834-6, pp. 329-477
(2000)