



# Hybrid QM/classical models in chemistry



## Benedetta Mennucci

email: benedetta.mennucci@unipi.it



http://molecolab.dcci.unipi.it

# Outline

#### • Methodological aspects

- $\circ$  Introduction
- Continuum models
- Atomistic models
- Self Consistent reaction field

#### • Molecular Properties of Solvated Systems

- Energy Derivatives
- Direct vs. indirect effects
- Bulk versus specific solvation
- Some examples

#### • Excited states in solvated systems

- Nonequilibrium
- State specific vs. linear response formulations
- $\circ$  Examples

2nd part

# The QM description of environment effects on molecular properties & spectroscopies

### Physical issue:

the solvation model should include all the main physical interactions giving rise to solvent-induced change on the property/ spectroscopic signal

Quantum-Mechanical issue:

the QM model should be able to describe the effects of solvent on the solute charge density and on its response to perturbations The Quantum Mechanical issue

We need a proper definition of molecular properties

We have to start from a quantity which is the basic element of any QM description:





The molecular property can be defined in terms of the change of the energy of the system with respect to a perturbation (an external electric or magnetic field, a geometrical deformation ...):

this definition is still valid for a solvated system

## General definition: property as derivative of the energy

The energy is expanded in a Taylor series in the perturbation strength  $\lambda$ 

$$\mathbf{E}(\lambda) = \mathbf{E}(0) + \frac{\partial \mathbf{E}}{\partial \lambda} \lambda + \frac{1}{2} \frac{\partial^2 \mathbf{E}}{\partial \lambda^2} \lambda^2 + \frac{1}{6} \frac{\partial^3 \mathbf{E}}{\partial \lambda^3} \lambda^3 + \dots$$

The nth-order property is the nth-order derivative of the energy

For example, by considering four types of perturbations: external electric (F) or magnetic field (B), nuclear magnetic moment (nuclear spin, I) and a change in the nuclear geometry (R).

Property 
$$\propto \frac{\partial^{n_F + n_B + n_I + n_R} E}{\partial F^{n_F} \partial B^{n_B} \partial I^{n_I} \partial R^{n_I}}$$

$n_{F}$	$n_B$	$n_{I}$	$n_R$	Property
0	0	0	0	Energy
1	0	0	0	Electric dipole moment
0	1	0	0	Magnetic dipole moment
0	0	1	0	Hyperfine coupling constant
0	0	0	1	Energy gradient
2	0	0	0	Electric polarizability
0	2	0	0	Magnetizability
0	0	2	0	Spin-spin coupling (for différent nuclei)
0	0	0	2	Harmonic vibrational frequencies
1	0	0	1	Infra-red absorption intensities
1	1	0	0	Circular, dichroism
0	1	1	0	Nuclear magnetic shielding
3	0	0	0	(first) Electric hyperpolarizability
0	3	0	0	(first) Hypermagnetizability
0	0	0	3	(cubic) Anharmonic corrections to vibrational frequencies
2	0	0	1	Raman intensities
2	1	0	0	Magnetic circular dichroism (Faraday effect)
1	0	0	2	Infra-red intensities for overtone and combination bands
4	0	0	0	(second) Electric hyperpolarizability
0	4	0	0	(second) Hypermagnetizability
0	0	0	4	(quartic) Anharmonic corrections to vibrational frequencies
2	0	0	2	Raman intensities for overtone and combination bands
2	2	0	0	Cotton-Mutton effect

## General definition: property as derivative of the energy

Now the energy contains the environment contributions

$$\hat{H}_{eff} = \hat{H}_{QM} + \hat{V}_{int} \qquad \hat{V}^{MM(Pol)} = \sum_{k} q_{k}^{MM} \hat{V}_{k} \left( -\sum_{i} \mu_{i}^{ind} \cdot \hat{E}_{i} \right)$$
$$\hat{V}^{ASC} = \sum_{i} q_{i}^{ASC} \hat{V}_{i}$$

Continuum models 
$$\left\langle \Psi \middle| \hat{H}^{eff} \middle| \Psi \right\rangle - \frac{1}{2} \left\langle \Psi \middle| \sum_{i} q_{i}^{ASC} \hat{V}_{i}^{QM} \middle| \Psi \right\rangle = E_{s} - \frac{1}{2} E_{int}$$

MM(Pol) models 
$$\left\langle \Psi \middle| \hat{H}^{eff} \middle| \Psi \right\rangle + E_{MM(pol)} = E_s + E_{MM(pol)}$$

As a result the derivatives (and the corresponding response properties) will be changed by the presence of the environment

## The environment effects on properties



## Solvent effects & properties: a simple example





ELECTRONIC EFFECTS

Indirect effects: solvent induced changes in the solute geometry Direct effects: solvent induced changes in the solute electronic charge distribution

## Solvent effects & properties: a simple example

Push-pull systems: asymmetric conjugated systems with donor and acceptor groups



## Direct effects on the electronic density



QM/PCM electron density localized into atomic charges



## Indirect effects on the conjugation

**BLA** (Bond Length Alternation)

It is an index of conjugation defined as the average difference in length between single & double bonds along a conjugation path.



## Direct vs indirect effects: the NMR evidence



## Direct vs indirect effects: the conformational aspect

Measured IR spectrum in water





N-Methyl Acetylproline Amide (NAP) in water:

a simple model for peptides

# Conformational analysis: three main conformers NH<sub>2</sub> H<sub>3</sub>C H<sub>2</sub>Ċ H<sub>2</sub>

Boltzmann Populations (%)

	In gas-phase B3LYP/6-311++G**
I	99
2	Ι
3	-

Conformer I dominates in gasphase due to the presence of a stabilizing intramolecular H-bond.

C. Cappelli and B. Mennucci J. Phys. Chem. B 2008, 112, 3441.

## Conformational analysis: three main conformers



PCM cavities



Boltzmann Populations (%)

2

	In gas-phase B3LYP/6-311++G**	In water: PCM B3LYP/ 6-311++G**
I	99	4
2	I	28
3	-	68

In water the other two conformers become important: they present a better interaction with the solvent (polar groups are more exposed to the solvent)

## Infrared spectra of single conformers: direct effects



IR



Both position and intensity of peaks change passing from gas-phase to water

## Averaged Infrared spectra: direct vs indirect effects



L-AP in  $D_2O$  (a) 400 U  $10^{\circ}C$  400 0  $10^{\circ}C$   $70^{\circ}C$  0 1550 1600 1600 1650 1700 1750Frequency (cm<sup>-1</sup>)

EXP

800

B3LYP/6-311++G\*\*

Exp taken from: Oh, K.-I.; Han, J.; Lee, K.-K.; Hahn, S.; Han, H.; Cho, M. J. Phys. Chem. B **2006**, 110, 13335– 13365.

# Bulk versus specific solvation



How does solvation at the solute surface differ from the bulk?
 Are there local rigid structures of solvent at the solute surface?
 What are the time scales for solvent dynamics at the solute surface?

## Solvation shells

#### The minimum model:

**solute** surrounded by some explicit **solvent** molecules

(the supermolecule)

The different components of the supermolecule can be described at the same level (QM) or using an hybrid approach:

a better level for the solute and a more approximated level for the solvent molecules (semiempirical or MM)

## The Supermolecule

#### How many explicit solvent molecules are needed?

An example: N-methyl acetamide in water





3 possible H-bonds: 2 on the O(C) and I on the H(N)

## A more refined analysis

A radial distribution function (RDF) measures the probability of finding an atom in function of the distance from another given atom compared to the case of the ideal gas



It can be obtained from a classical molecular dynamics

## Back to the N-methyl-acetamide in water



Integrating the radial distribution function gives the number of particles surrounding the central particle



Hydrogen bonding sites

O-Hw: a very well-defined first peak centered at 1.8 Å that integrates to 2.3. On average, two hydrogen atoms of two water molecules hydrate the carbonyl group by means of hydrogen bonding.

(N)H-OW: a much less structured solvent. The first peak, is centered at 2.0 Å, the integration number up to the first minimum is 1.1.

# Radial & Spatial distribution functions

#### N-methyl-acetamide (NMA) in water





Hydrogen bonding sites

#### The 3D equivalent Spatial Distribution Function (SDF)



## Which configuration?

#### From QM geometry optimization: the most stable configuration

Proper description for strongly interacting solute-solvent systems giving rise to stable clusters (example: strong H-bonded clusters)

#### ♦ From MD simulations:

More general than the QM optimization.

Better for weaker solute-solvent interactions described by a more dynamic situation.

Main limitation: strongly dependent on the quality of the MD

From MD simulations: how does it work?

First we select a large enough set of uncorrelated snapshots from the MD simulation

Then we define solute-solvent clusters on the basis of a cutoff distance  $(r_{cut})$ : each cluster includes all solvent molecules inside the sphere of radius=  $r_{cut}$ The value used for  $r_{cut}$  has to be chosen so the represent the first solvation shell.

Finally, to achieve a statistically meaningful description, we average on all the clusters

Calculations of the property of interest have to be repeated for all the clusters so to obtain a correct average value

Advantage:

Proper description of weak solute-solvent specific interactions which cannot be represented by a single configuration obtained from a QM geometry optimization

Disadvantage:

Quite demanding from a computational point of view





## How can we include long-range effects?

#### > Enlarging the dimension of the supermolecule:



Problems:

- 1. The accuracy of the QM level has to be largely reduced, or an hybrid QM/MM has to be introduced
- 2. We need many clusters to get a correct statistical picture

## The Supermolecule

## How can we include long-range effects?

Adding an "external" continuum:
the solvated supermolecule



The main problems disappear:

- 1. We do not need to reduce the accuracy of the QM level, or to shift to an hybrid QM/MM method
- 2. The statistical representativity is automatically satisfied by using the continuum description in terms of the solvent bulk properties.

## An example of bulk versus specific effects: <sup>15</sup>N nuclear shieldings of diazines in solution



The NMR signal for a particular nucleus depends on its molecular environment. This is why NMR is such a useful tool for structure determination... and solvent effects.

> Nuclear Shielding constant



Derivative with respect to the magnetic field and the nuclear magnetic moment

## An example of bulk versus specific effects: <sup>15</sup>N nuclear shieldings of diazines in solution



for but only with inclusion of bulk effects (PCM clusters) the observed solvent effect is reproduced Can we go beyond ground state processes?

# Vertical excitations & environment response



A non responsive environment does NOT ''see'' the transition & does NOT distinguish between different states

# Vertical excitations & environment

### response



## Nonequilibrium & solvation dynamics: the experimental evidence

In a polar solvent



## Nonequilibrium & solvent polarization



Time-dependent dielectric response

 $\epsilon(\omega)$  is the complex dielectric constant

 $\varepsilon(\omega) = \varepsilon'(\omega) + i\varepsilon''(\omega)$ 

<u>Real part  $\epsilon'(\omega)$ </u>: frequency-dependent dielectric constant, describes the component of the polarization in phase with the oscillating field

<u>Imaginary part (loss factor)  $\epsilon''(\omega)$ </u>: component of the polarization with a phase difference of  $\pi/2$  with respect to the field; gives rise to the loss of energy of the electric field in the medium.

At Low frequencies: for most polar solvents,  $\epsilon'(\omega)$  is equal to the static dielectric constant  $\epsilon$ , and  $\epsilon''(\omega)$  is zero. For very high frequencies,  $\epsilon'(\omega)$  reduces to the square of the refractive index:  $\epsilon'(\infty)=n^2$ 



Debye relaxation is the dielectric relaxation response of an ideal, noninteracting population of dipoles to an alternating external electric field. The Polarisation decays exponentially with a time  $\tau_D$ 

It is usually expressed in the complex permittivity  $\varepsilon$  as a function of the field's frequency



Separating the real and imaginary parts of the complex dielectric permittivity:

$$\varepsilon'(\omega) = \varepsilon(\infty) + \frac{\varepsilon(0) - \varepsilon(\infty)}{1 + \omega^2 \tau_D^2}$$
$$\varepsilon''(\omega) = \frac{(\varepsilon(0) - \varepsilon(\infty))\omega\tau_D}{1 + \omega^2 \tau_D^2}$$

## Nonequilibrium & solvent polarization

For fast processes:

partition of the solvent polarization into dynamic (electronic motions) & inertial (molecular and nuclear motions) contributions

$$\vec{P} \simeq \vec{P}_{dyn}(\mathcal{E}_{\infty}) + \vec{P}_{iner}$$

$$\uparrow \qquad \uparrow$$
Electronic Inertial polarization (only for **polar polarization solvents**)

## Nonequilibrium: the ASC picture



I. Upon excitation only the dynamic response of the solvent "readjusts" while the inertial one remains frozen

 $\mathbf{T}(\boldsymbol{\varepsilon}_{\infty})\mathbf{q}_{EX}^{dyn} = -\mathbf{R}\mathbf{V}_{EX}$  $\mathbf{q}_{GS}^{in} = \mathbf{q}_{GS}^{eq} - \mathbf{q}_{GS}^{dyn}$ 

2. Solute & solvent nuclear degrees of freedom relax towards a new EX equilibrium

 $\mathbf{T}(\boldsymbol{\varepsilon}_{0})\mathbf{q}_{EX}^{eq} = -\mathbf{R}\mathbf{V}_{EX}$ 

3. On emission, the situation I is reverted

4. Solute & solvent nuclear degrees of freedom relax towards the initial GS equilibrium

## Nonequilibrium: the results

Acrolein: solvatochromic shifts (acetonitrile-cyclohexane)

water (polar solvent)  $\epsilon$ =78.4,  $\epsilon_{\infty}$ =1.8 cyclohexane (apolar solvent)  $\epsilon$ = $\epsilon_{\infty}$ =2.0



EOM-CC/6-31+G(d). All values are in eV.

An equilibrium model gives a completely wrong picture due to the over stabilization of the excited state

# Nonequilibrium: the QM/MM picture

#### Within a QM/MM scheme

•Chromophore: QM

•Environment: MM

•nonequilibrium:

• fixed charges which represent the inertial component



Can we include also the dynamic part of polarization?

#### YES

but only if we use a polarizable embedding: the induced dipoles describe the dynamic polarization

# Excitation in solvated systems: which QM approach?

#### • State Specific (SS):

• The wavefunction of the excited state is explicitly calculated together with the energy

• CASSCF, CI, ....

"Natural" extension for polarizable models but computationally expensive

• Linear Response (LR)

• We obtain the excitation energies from the response of the ground state density to a perturbation

• ZINDO, TDDFT, EOM-CC, ...

Computationally efficient but which response of the polarizable environment?

## The State-Specific (SS) vertical excitation



#### A nonequilibrium scheme is used

The inertial part remains frozen in the GS

$$\mathbf{q}_{in}^{PCM}(\mathbf{P}_{GS}) \Leftrightarrow \mathbf{q}^{MM}$$

The dynamic part relaxes in the EXC  $\mathbf{q}_{dyn}^{PCM}(\varepsilon_{\infty}) \Leftrightarrow \boldsymbol{\mu}^{ind}(\alpha)$ 

## The DFT version of the linear response: TDDFT

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \boldsymbol{\omega} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}$$

$$\begin{split} A_{ia,jb} &= \delta_{ij} \delta_{ab} (\epsilon_a - \epsilon_i) + (ia|jb) + (ia|f_{\rm xc}|jb) \\ B_{ia,jb} &= (ia|bj) + (ia|f_{\rm xc}|bj) \end{split}$$

$$(ia \mid jb) + \left(ia \mid f_{xc} \mid jb\right) = \int d\vec{r} \int d\vec{r} \phi_i^*(\vec{r}) \phi_a(\vec{r}) \left[\frac{1}{\left|\vec{r} - \vec{r} \mid\right|} + \frac{\delta^2 E_{xc}}{\delta \rho(\vec{r}) \delta \rho(\vec{r}')}\right] \phi_j^*(\vec{r}') \phi_b(\vec{r}')$$

## The densities



The relaxation term requires to go beyond the LR approach (Lagrangian method)

## The embedded TDDFT

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \boldsymbol{\omega} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}$$

Orbitals & orbital energies are obtained in the presence of the environment@GS

$$\begin{bmatrix} A_{ia,bj} = \delta_{ab} \delta_{ij} (\varepsilon_a - \varepsilon_i) + K_{ia,bj} + C_{ai,bj}^{Env} \\ B_{ia,bj} = K_{ia,jb} + C_{ai,bj}^{Env} \end{bmatrix}$$

The dynamic response of the environment induced  $C_{\mu}$  by the "transition"

of  
ed 
$$C_{ia,jb}^{env} = \int d\vec{r} \phi_i^*(\vec{r}) \phi_a(\vec{r}) \sum_i$$

$$q_{i}^{ASC} \left[ \boldsymbol{\varepsilon}_{\infty}; \boldsymbol{\phi}_{j}^{*} \boldsymbol{\phi}_{b} \right] \frac{1}{\left| \vec{r} - \vec{s}_{i} \right|}$$
$$\boldsymbol{\mu}_{i}^{ind} \left[ \boldsymbol{\alpha}_{i}; \boldsymbol{\phi}_{j}^{*} \boldsymbol{\phi}_{b} \right] \cdot \frac{\vec{r} - r_{i}}{\left| \vec{r} - r_{i} \right|^{3}}$$

State-Specific correction to the energy

$$\omega_{0K} + \frac{1}{2} \sum_{i} \begin{cases} q_i^{ASC} [\varepsilon_{\infty}; \Delta \mathbf{P}_K] V^{\Delta P_K}(\vec{s}_i) \\ \mu_i^{ind} [\alpha; \Delta \mathbf{P}_K] \cdot E^{\Delta P_K}(\vec{s}_i) \end{cases}$$

## A very unique solvatochromic probe



A well-known probe with a strong negative solvatochromism





The excitation involves a Charge-Transfer from O to N-ring



HOMO

LUMO

# The modeling



gas-to-water shift (eV) SS/TDCAM-B3LYP/6-31+G(d) QM/PCM 0.79 QM/QM/PCM 1.17

Ехр	I.56

- A continuum only description (PCM): NOT enough
- Inclusion of two hydrogen bonded water molecules in the QM region (plus PCM): still NOT enough

- The orienting effects of the zwitterion GS go well beyond the first solvation shells: a large number of water molecules is needed (MMPol or MM)
- MM cannot stabilise the excited state: only the the SS-MMPol reproduces the correct shift

## Conclusions: which environment model?



# Unfortunately, not a unique solution!

The choice requires a preliminary analysis on:

- (i) the system of interest,
- (ii) the property/process of interest,
- (iii) the required accuracy,
- (iv) the computational cost

## Conclusions: some warnings





BLACKBOX model

NOT a BLACKBOX model

QM/MMpol



## Conclusions: some suggestions

