

# Minimal parameter implicit solvent model for DFT calculations

Jacek Dziedzic<sup>1</sup>, James Womack<sup>1</sup>, Lucian Anton<sup>2</sup>, Chris-Kriton Skylaris<sup>1</sup>

with previous work by

Hatem H. Helal<sup>3</sup>, Arash A. Mostofi<sup>4</sup>, and Mike Payne<sup>3</sup>

<sup>1</sup> UNIVERSITY OF  
**Southampton**  
School of Chemistry

<sup>2</sup>  Science & Technology  
Facilities Council

<sup>3</sup>  UNIVERSITY OF  
CAMBRIDGE

<sup>4</sup> **Imperial College**  
London

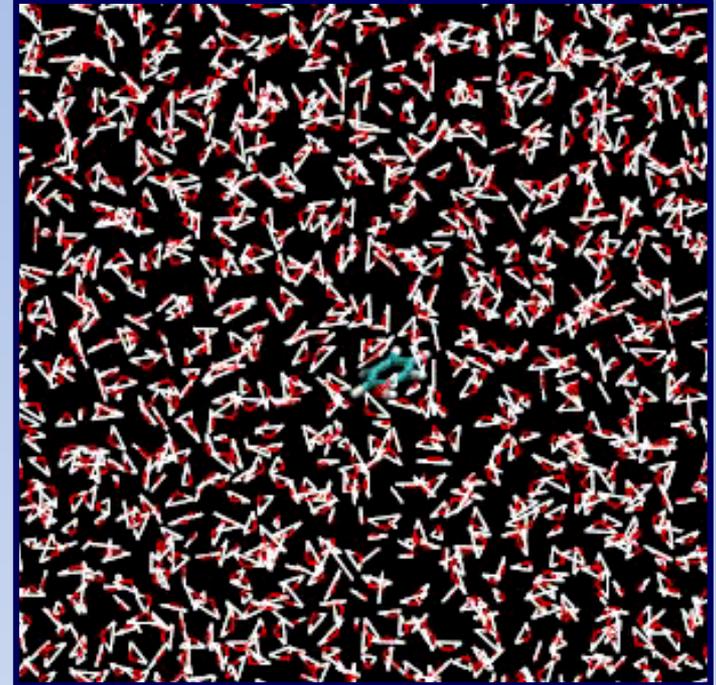
# Why bother?

- Many important biochemical reactions occur in aqueous solution.
- Performing calculations *in vacuo* often leads to greatly inaccurate results. Especially sensitive properties and phenomena include, among others:
  - energy differences between molecular conformers [1],
  - rates of reactions [2],
  - tautomeric equilibria [1],
  - molecular (esp. protein-protein) associations [1,3],
  - **protein structures** [4],
  - **ligand binding free energies** [5].
- Thus it is crucial to include the solvent environment in simulations of biological molecules.

# Explicit solvent approach

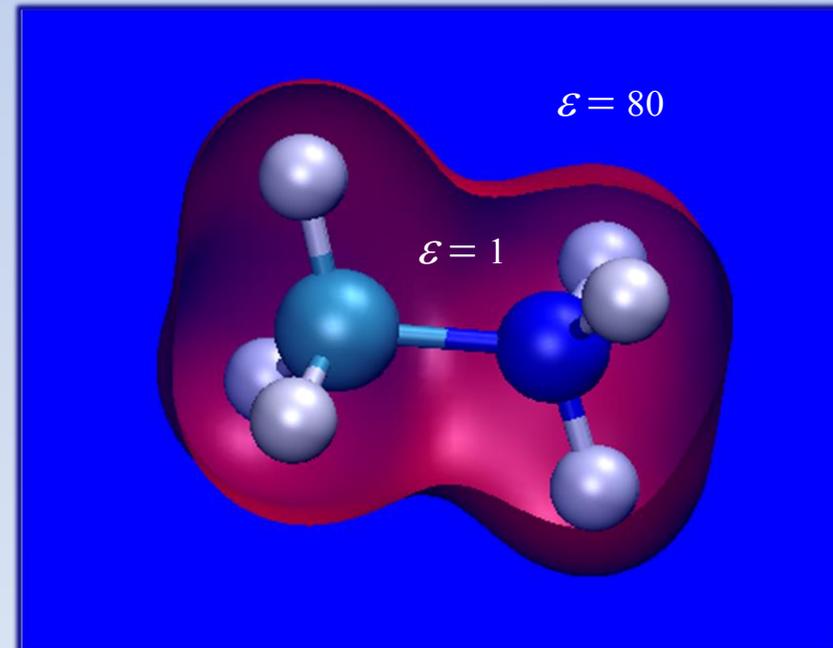
Animation by C. Pittock

- Introduce the solvent in molecular detail.
- Pros:
  - provides accurate treatment of solute-solvent interactions,
- Cons:
  - leads to an increase in system size, possibly by an order of magnitude [6],
  - must average out instantaneous interactions before the results become meaningful, ("integrate out" the degrees of freedom of the solvent) [4],
  - to avoid surface effects, PBCs must be introduced. This in turn requires large cells to avoid artificial interactions between replicated solutes.



# Implicit solvent approach

- Treat only the solute explicitly and embed it within a suitably defined cavity, the inside of which is inaccessible to the solvent.
- Replace the solvent with an *unstructured dielectric continuum*, only retaining its **average** effect on the solute [7].
- Pros:
  - no solvent atoms necessary,
  - eliminates the costly sampling of the solvent degrees of freedom [6],
  - faster sampling of solute conformations owing to absence of solvent friction [5].



# Two terms in free energy of solvation

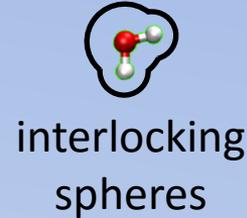
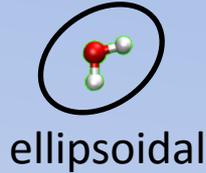
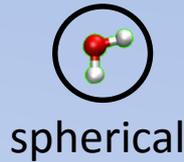
only the total can  
be obtained  
experimentally

$$\longrightarrow \Delta G_{sol} = \Delta G_{pol} + \Delta G_{npol}$$

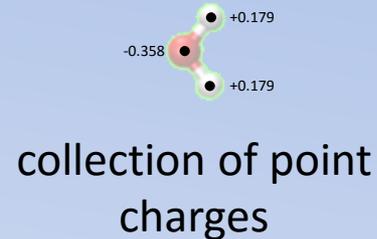
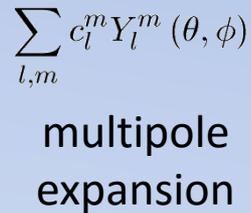
- The **electrostatic** or **polar term** describes the response of the solvent to the charge distribution of the solute [7].
- It is the difference between the electrostatic energy
$$\frac{1}{2} \int \rho(\mathbf{r}) \phi(\mathbf{r}) d\mathbf{r}$$
in solvent and in vacuum.
- The **nonpolar term** accounts for
  - the entropic cost of forming a cavity within the solvent (**cavitation energy**),
  - for the van der Waals interaction of the solute with the solvent [4] (**dispersion-repulsion energy**).
- Difficult to describe rigorously.
- A widely used approach is to represent it as a linear function of the molecular surface area [7]:  $\Delta G_{npol} = \gamma A_{SA}$ .

# Pick and mix to get a model

**cavity  
shape**



**charge  
representation  
(of the solute)**



**method of  
obtaining the  
reaction field of the  
dielectric**

analytical  
solution  
of NPE

discretizing the apparent  
surface charge (ASC) on  
the cavity surface

numerical solution of  
the NPE in the bulk

Generalized (nonhomogeneous)  
Poisson equation (NPE)

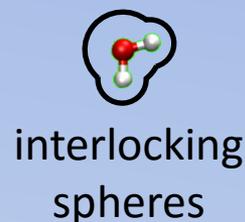
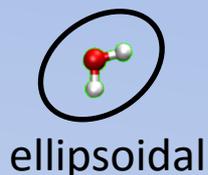
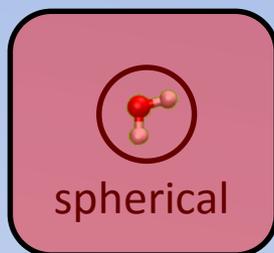
$$\nabla \epsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) = -4\pi \rho(\mathbf{r})$$

increasing level of sophistication

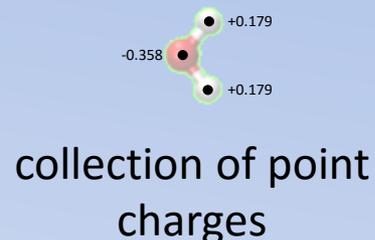
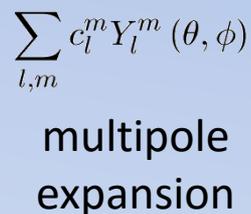


# Onsager's model (1936)

cavity  
shape



charge  
representation  
(of the solute)



method of  
obtaining the  
reaction field of the  
dielectric

analytical  
solution  
of NPE

discretizing the apparent  
surface charge (ASC) on  
the cavity surface

numerical solution of  
the NPE in the bulk

Generalized (nonhomogeneous)  
Poisson equation (NPE)

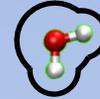
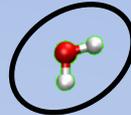
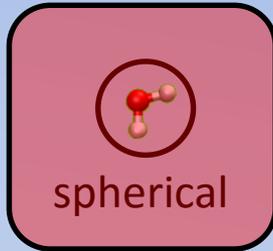
$$\nabla \epsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) = -4\pi \rho(\mathbf{r})$$

increasing level of sophistication

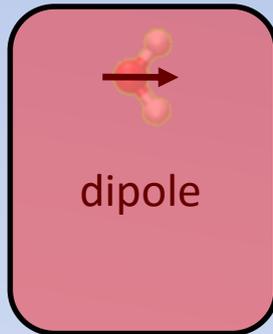


# Onsager's model (1936)

cavity shape



charge representation (of the solute)



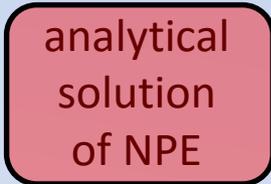
$$\sum_{l,m} c_l^m Y_l^m(\theta, \phi)$$

multipole expansion

**TOO SIMPLE**



method of obtaining the reaction field of the dielectric



discretizing the apparent surface charge (ASC) on the cavity surface

numerical solution of the NPE in the bulk

Generalized (nonhomogeneous) Poisson equation (NPE)

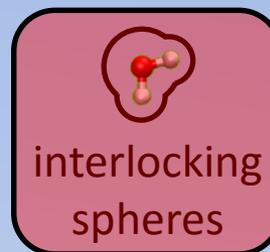
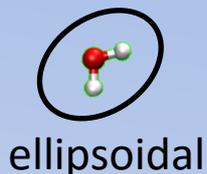
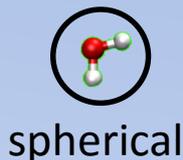
$$\nabla \cdot \epsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) = -4\pi \rho(\mathbf{r})$$

increasing level of sophistication



# Original PCM (1981)

cavity shape

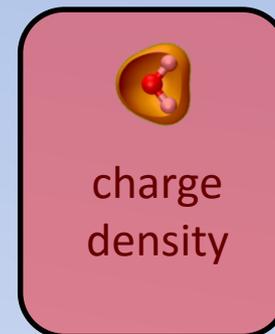
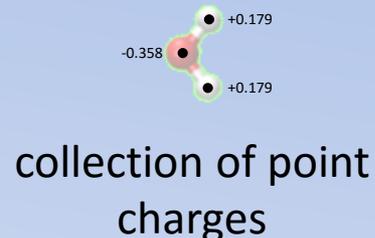


charge representation (of the solute)



$$\sum_{l,m} c_l^m Y_l^m(\theta, \phi)$$

multipole expansion



method of obtaining the reaction field of the dielectric

analytical solution of NPE

discretizing the apparent surface charge (ASC) on the cavity surface

numerical solution of the NPE in the bulk

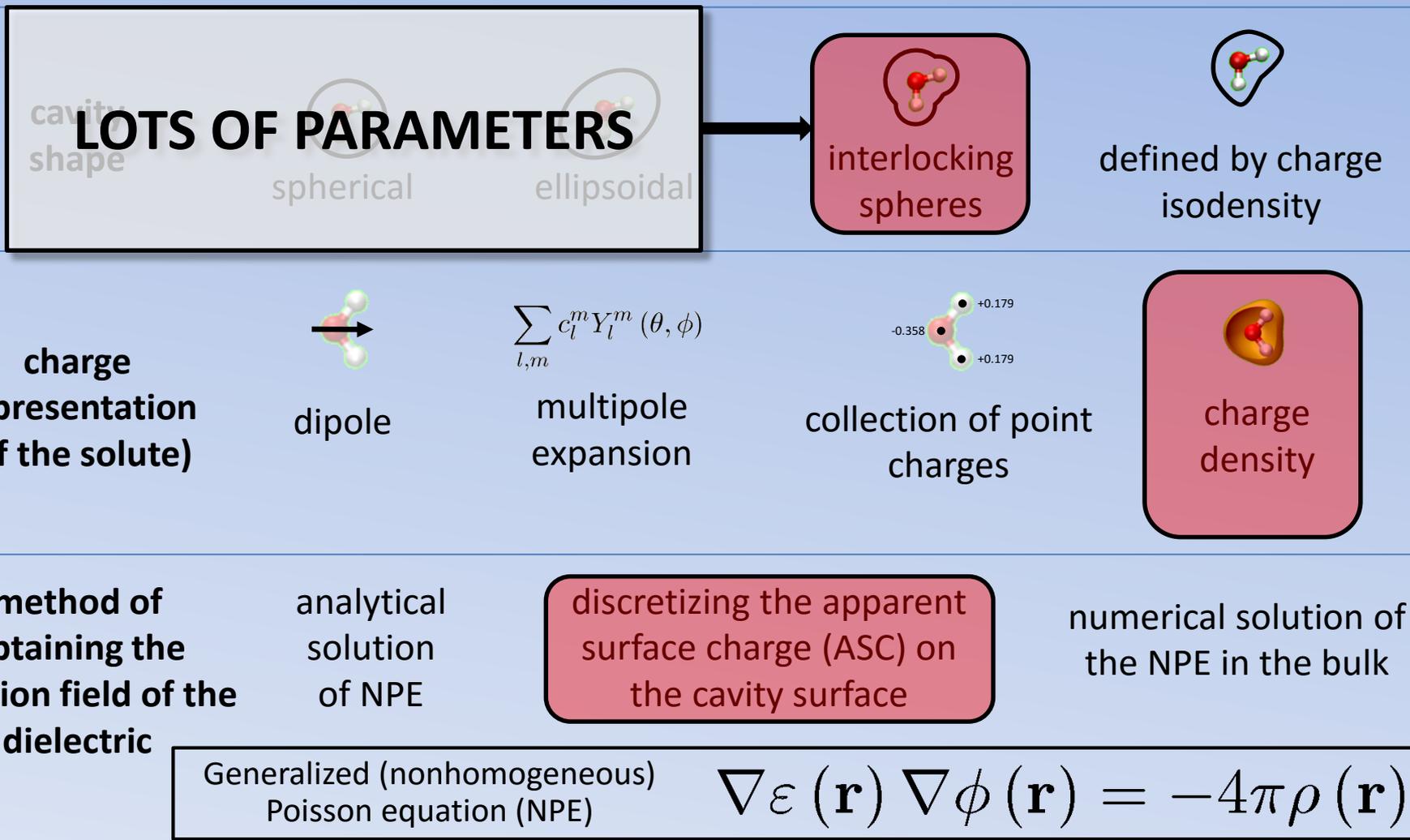
Generalized (nonhomogeneous) Poisson equation (NPE)

$$\nabla \epsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) = -4\pi \rho(\mathbf{r})$$

increasing level of sophistication



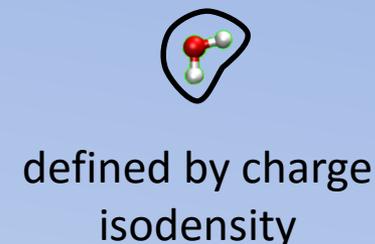
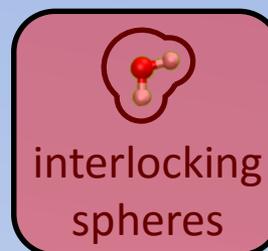
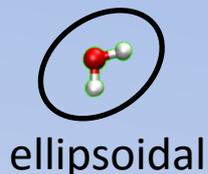
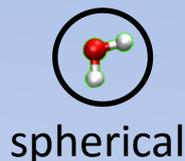
# Original PCM (1981)



increasing level of sophistication

# Classical methods (MMPBSA, DelPhi, APBS, ...)

cavity  
shape

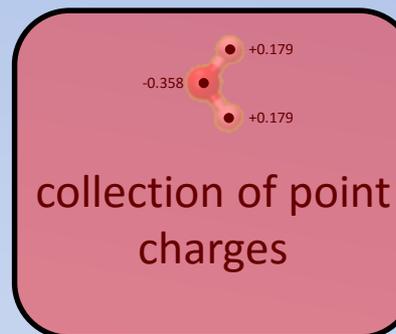


charge  
representation  
(of the solute)



$$\sum_{l,m} c_l^m Y_l^m(\theta, \phi)$$

multipole  
expansion



method of  
obtaining the  
reaction field of the  
dielectric

analytical  
solution  
of NPE

discretizing the apparent  
surface charge (ASC) on  
the cavity surface

numerical solution of  
the NPE in the bulk

Generalized (nonhomogeneous)  
Poisson equation (NPE)

$$\nabla \epsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) = -4\pi \rho(\mathbf{r})$$

increasing level of sophistication



# Classical methods (MMPBSA, DelPhi, APBS, ...)

cavity shape

**LOTS OF PARAMETERS**

spherical      ellipsoidal

interlocking spheres




defined by charge isodensity

charge representation (of the solute)

**LOTS OF PARAMETERS**

dipole expansion

$$\sum_{l,m} c_l^m Y_l^m(\theta, \phi)$$

collection of point charges




charge density

method of obtaining the reaction field of the dielectric

analytical solution of NPE

discretizing the apparent surface charge (ASC) on the cavity surface

numerical solution of the NPE in the bulk

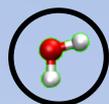
Generalized (nonhomogeneous) Poisson equation (NPE)

$$\nabla \epsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) = -4\pi \rho(\mathbf{r})$$

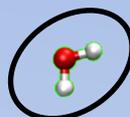


# Model of Fattebert and Gygi (2002)

cavity  
shape



spherical



ellipsoidal



interlocking  
spheres



defined by charge  
isodensity

charge  
representation  
(of the solute)



dipole

$$\sum_{l,m} c_l^m Y_l^m(\theta, \phi)$$

multipole  
expansion



collection of point  
charges



charge  
density

method of  
obtaining the  
reaction field of the  
dielectric

analytical  
solution  
of NPE

discretizing the apparent  
surface charge (ASC) on  
the cavity surface

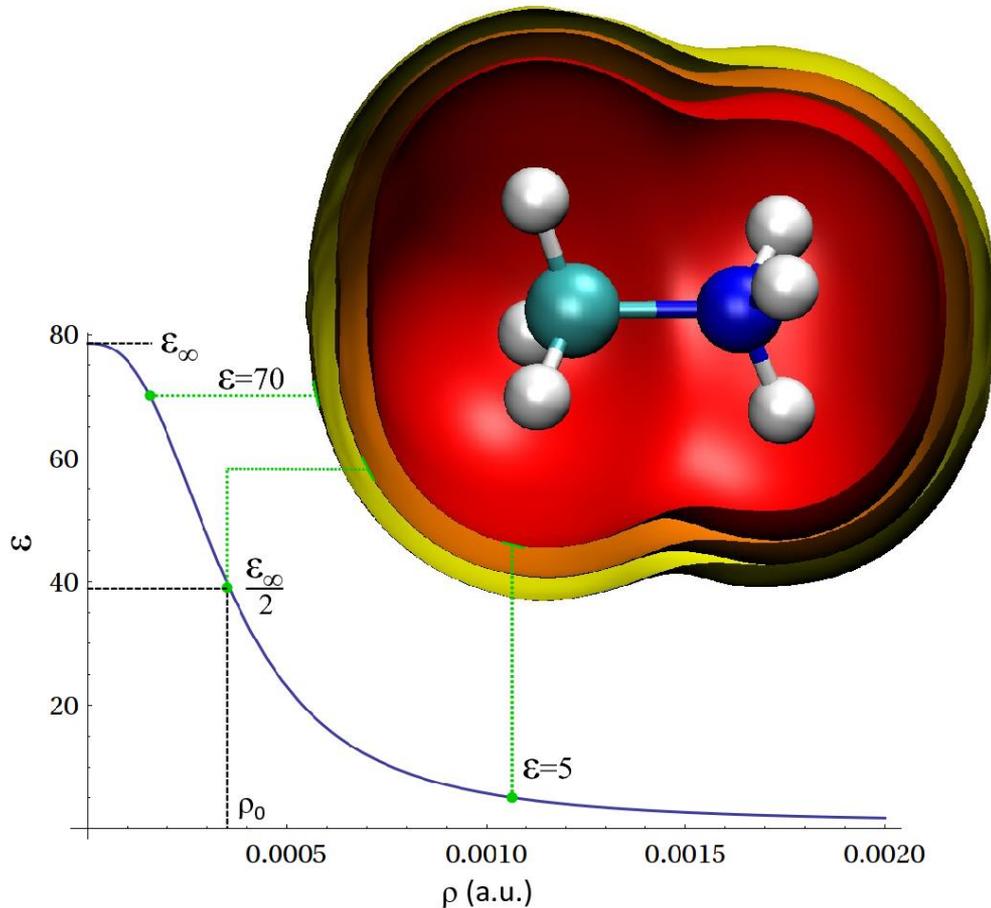
numerical solution of  
the NPE in the bulk

Generalized (nonhomogeneous)  
Poisson equation (NPE)

$$\nabla \epsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) = -4\pi \rho(\mathbf{r})$$

increasing level of sophistication

# Density-dependent cavity of Fattebert and Gygi

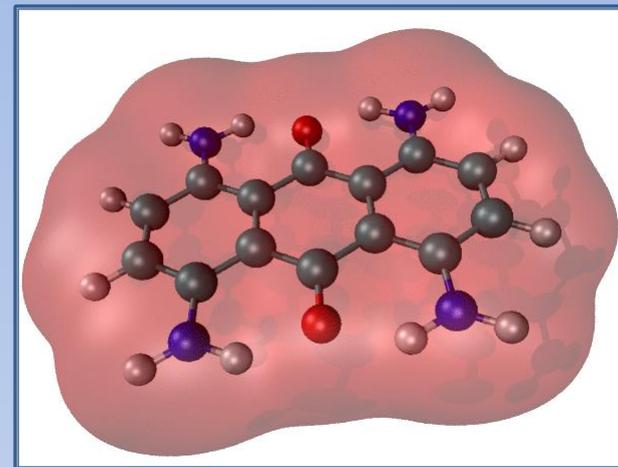


$$\nabla \epsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) = -4\pi \rho(\mathbf{r})$$

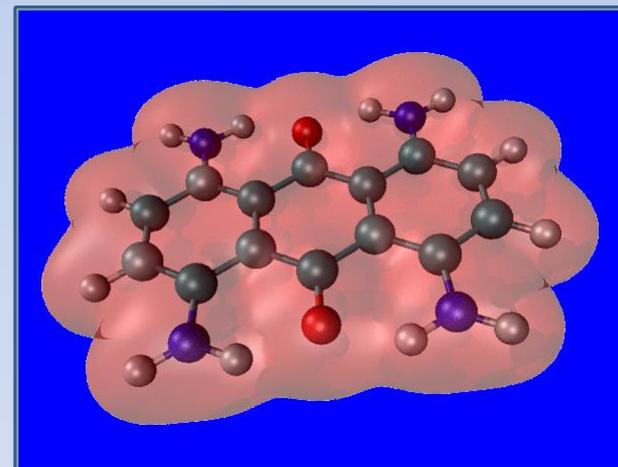
$$\epsilon(\rho(\mathbf{r})) = 1 + \frac{\epsilon_{\infty} - 1}{2} \left( 1 + \frac{1 - (\rho(\mathbf{r})/\rho_0)^{2\beta}}{1 + (\rho(\mathbf{r})/\rho_0)^{2\beta}} \right)$$

# The procedure

- Perform a calculation in vacuo to obtain  $E_{vac}$  and the charge density in vacuum.
- Start a calculation in solvent, using the charge density in vacuum as initial guess.
- Generate the cavity basing on current charge density.
- Solve  $\nabla \varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) = -4\pi\rho(\mathbf{r})$  with a multigrid solver to obtain  $\phi(\mathbf{r})$  in solvent. Use this in the electrostatic energy terms.
- Repeat until convergence in solvent.



First achieve self-consistency in vacuum.



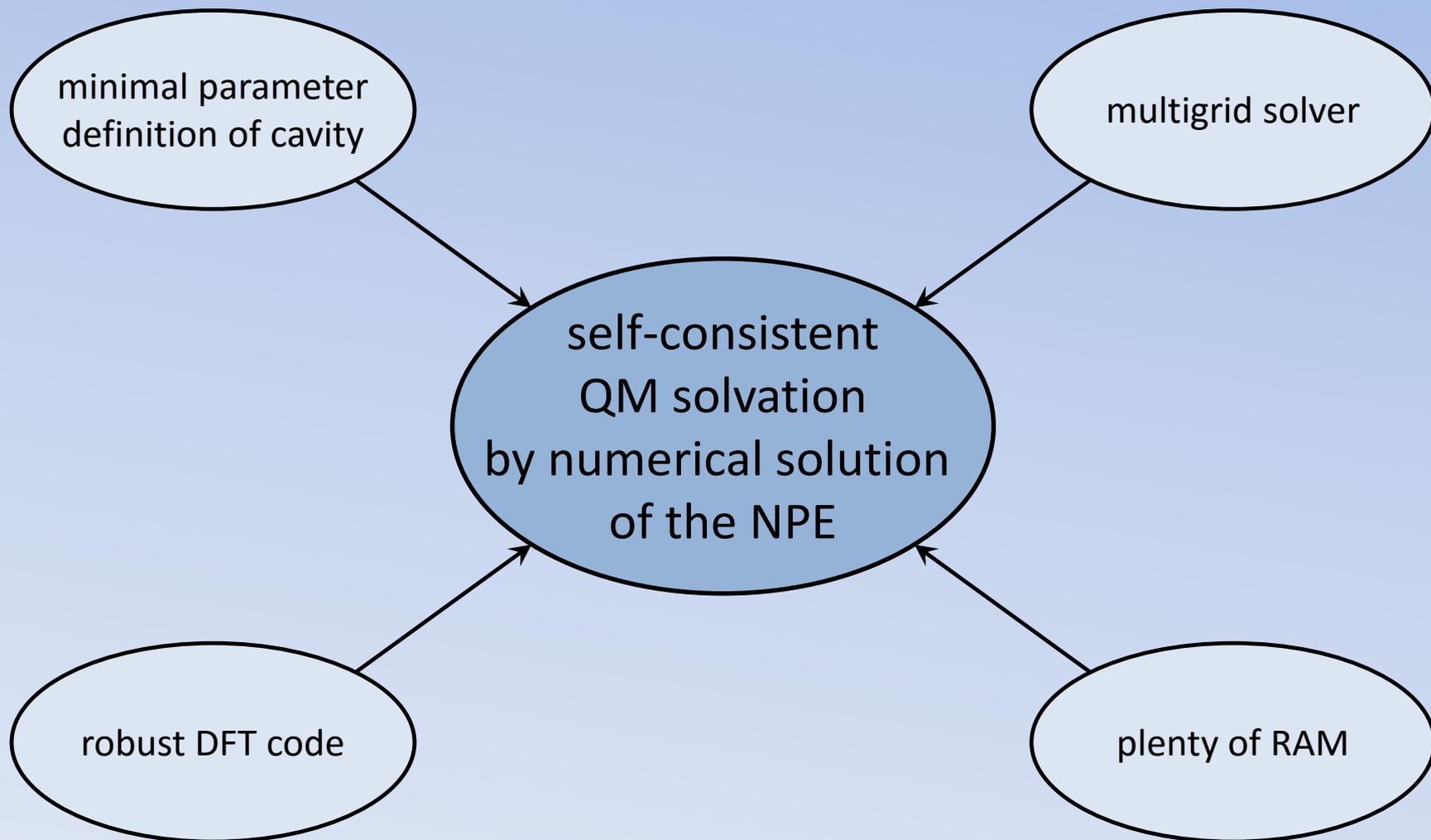
... then in solution.  
NB how the density hardly changes at all in solvent.

Once you have  $\rho(\mathbf{r})$  and  $\phi(\mathbf{r})$   
in vacuum and in solvent...

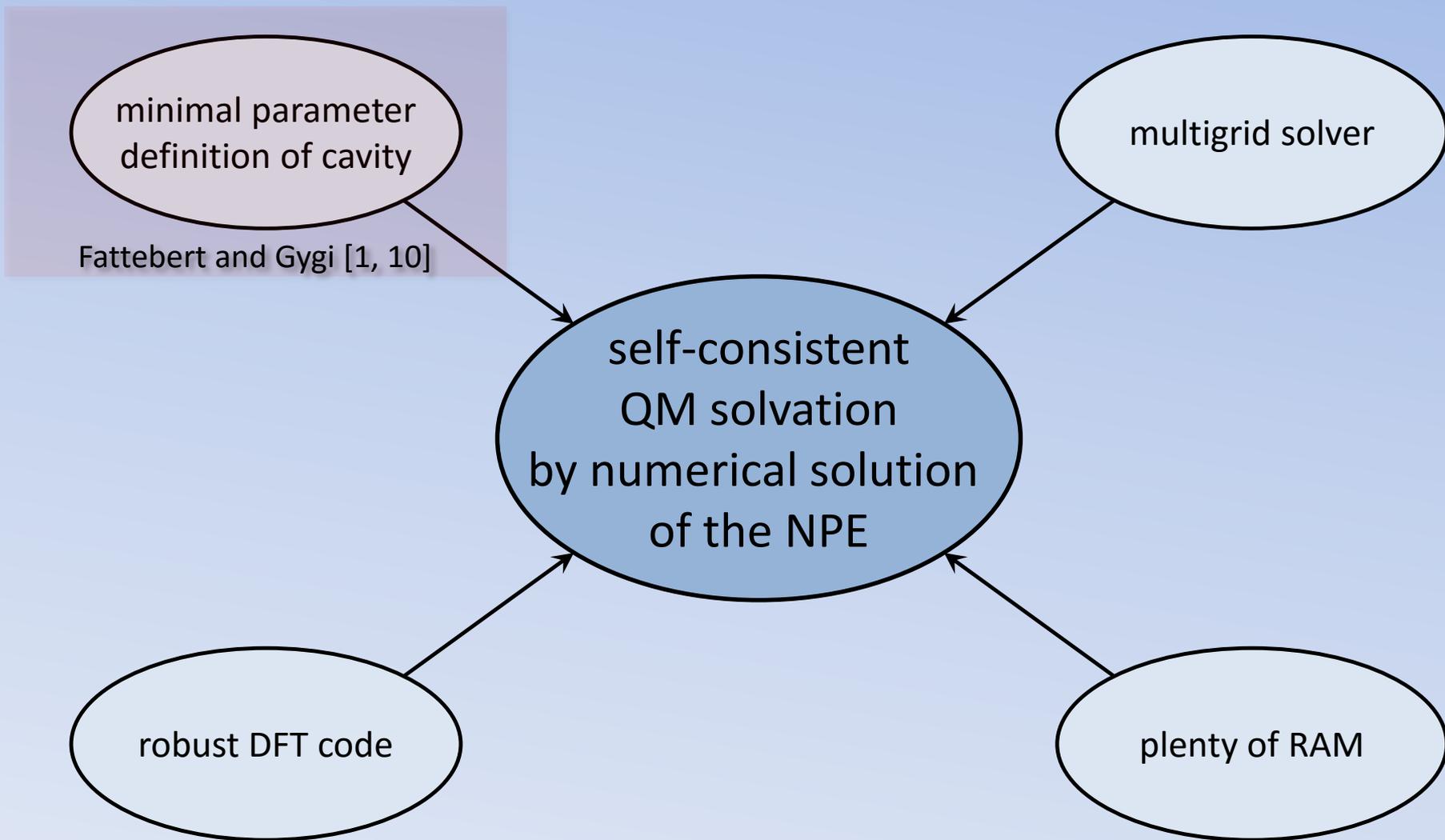
$$\Delta G_{sol} = \Delta G_{pol} + \Delta G_{npol}$$

- The **electrostatic** or **polar term** describes the response of the solvent to the charge distribution of the solute [7].
- It is the difference between the electrostatic energy
$$\frac{1}{2} \int \rho(\mathbf{r}) \phi(\mathbf{r}) d\mathbf{r}$$
in solvent and in vacuum.
- The **nonpolar term** accounts for
  - the entropic cost of forming a cavity within the solvent (**cavitation energy**),
  - for the van der Waals interaction of the solute with the solvent [4] (**dispersion-repulsion energy**).
- Difficult to describe rigorously.
- A widely used approach is to represent it as a linear function of the molecular surface area [7]:  $\Delta G_{npol} = \gamma A_{SA}$ .

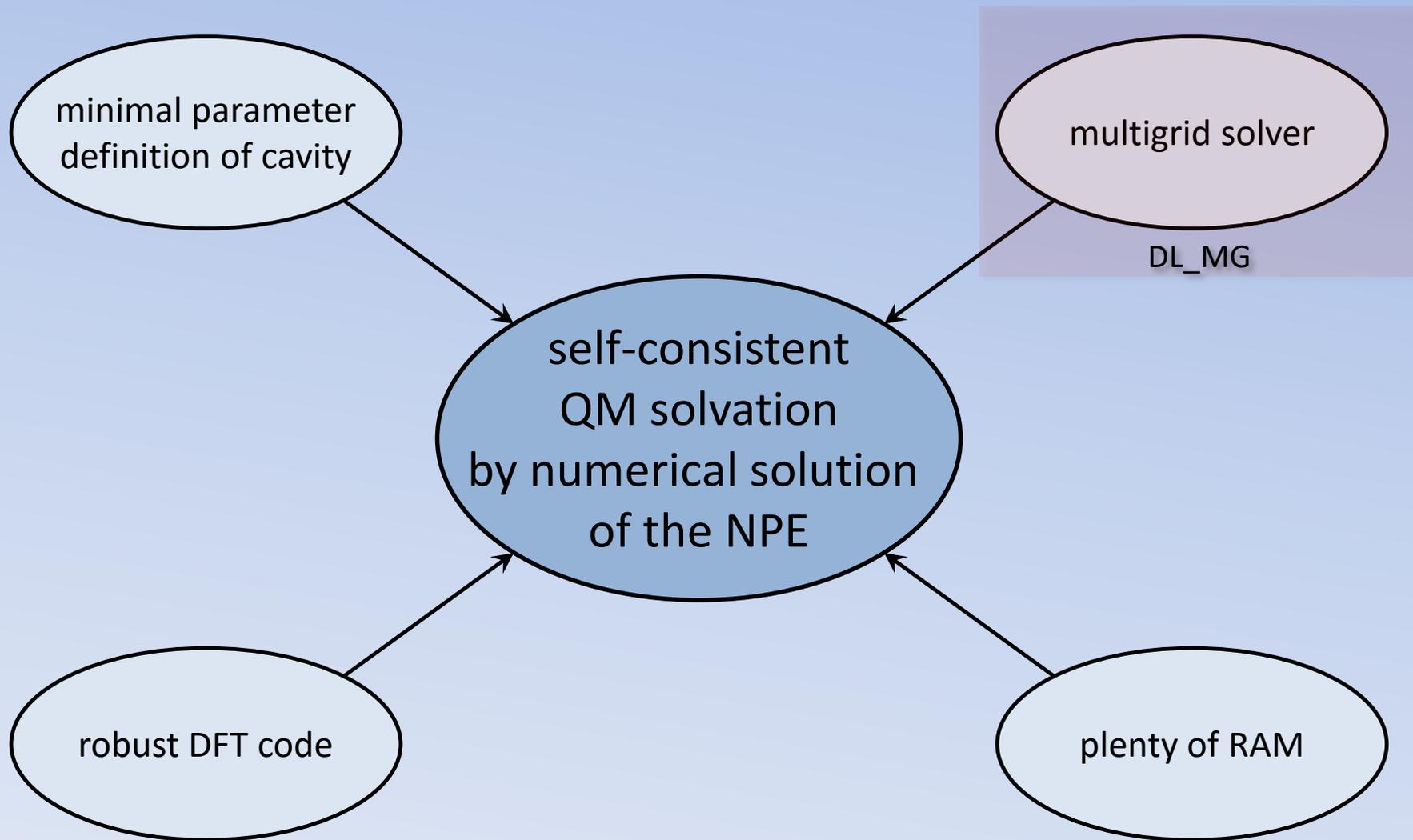
# Components of minimal parameter QM solvation



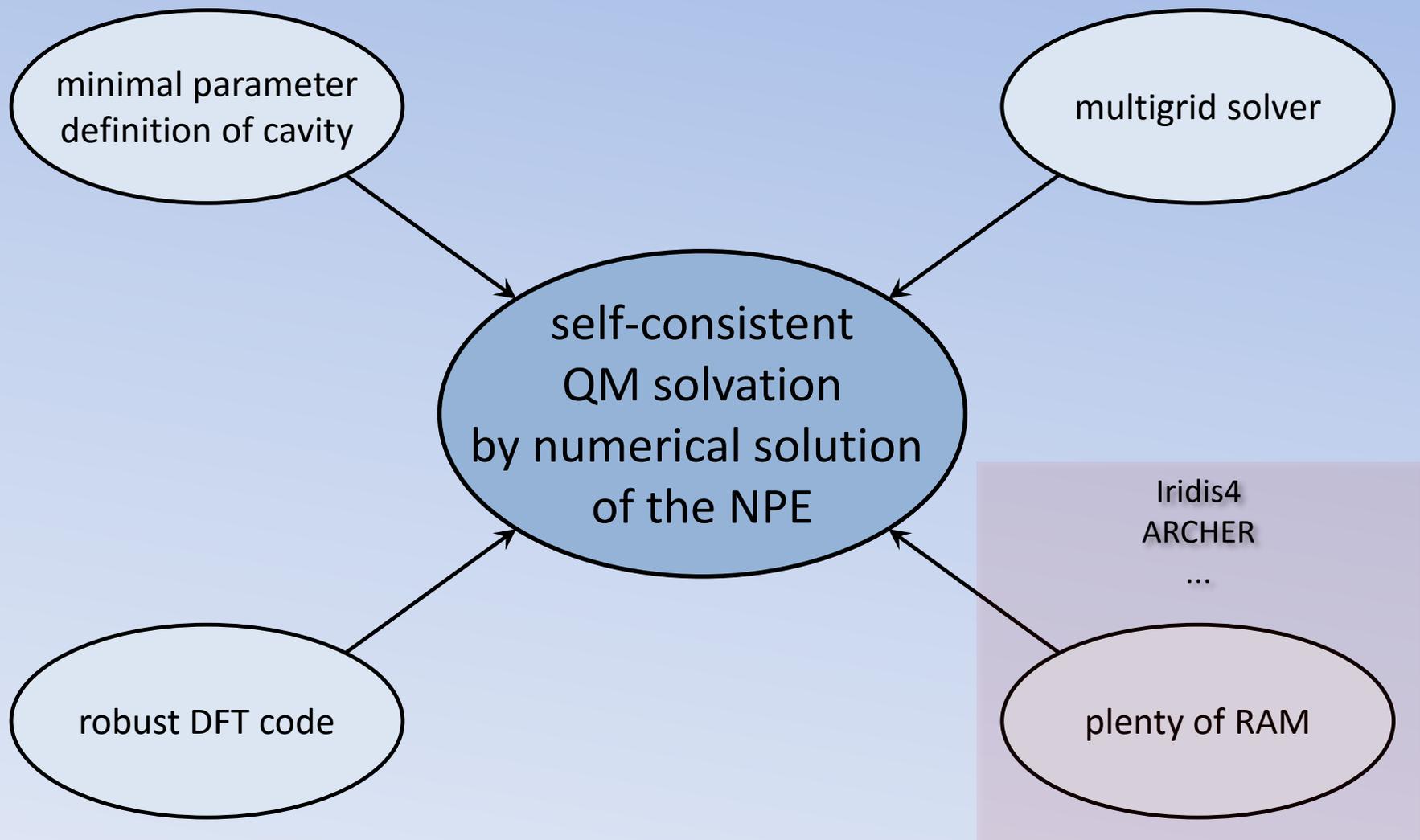
# Components of minimal parameter QM solvation



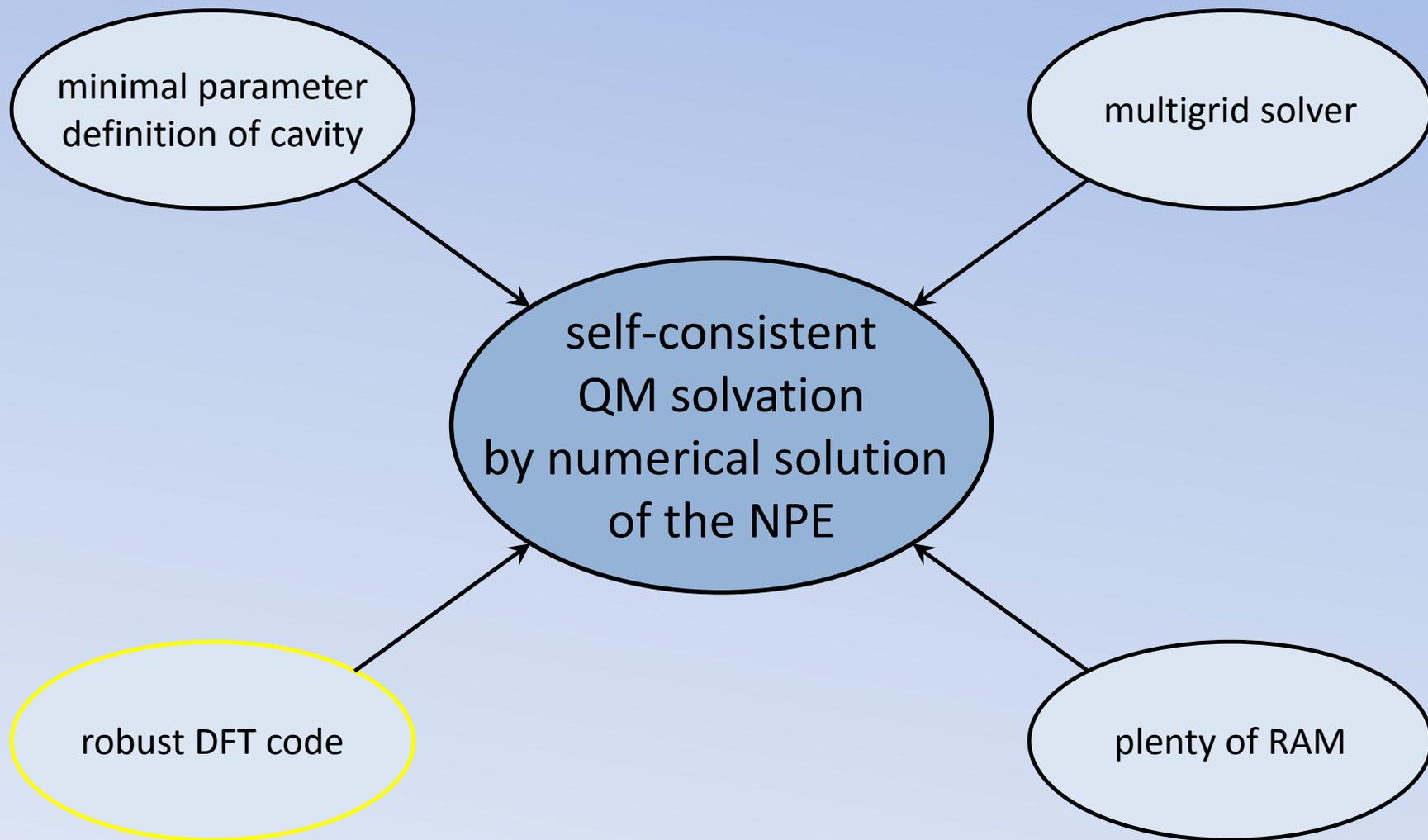
# Components of minimal parameter QM solvation



# Components of minimal parameter QM solvation

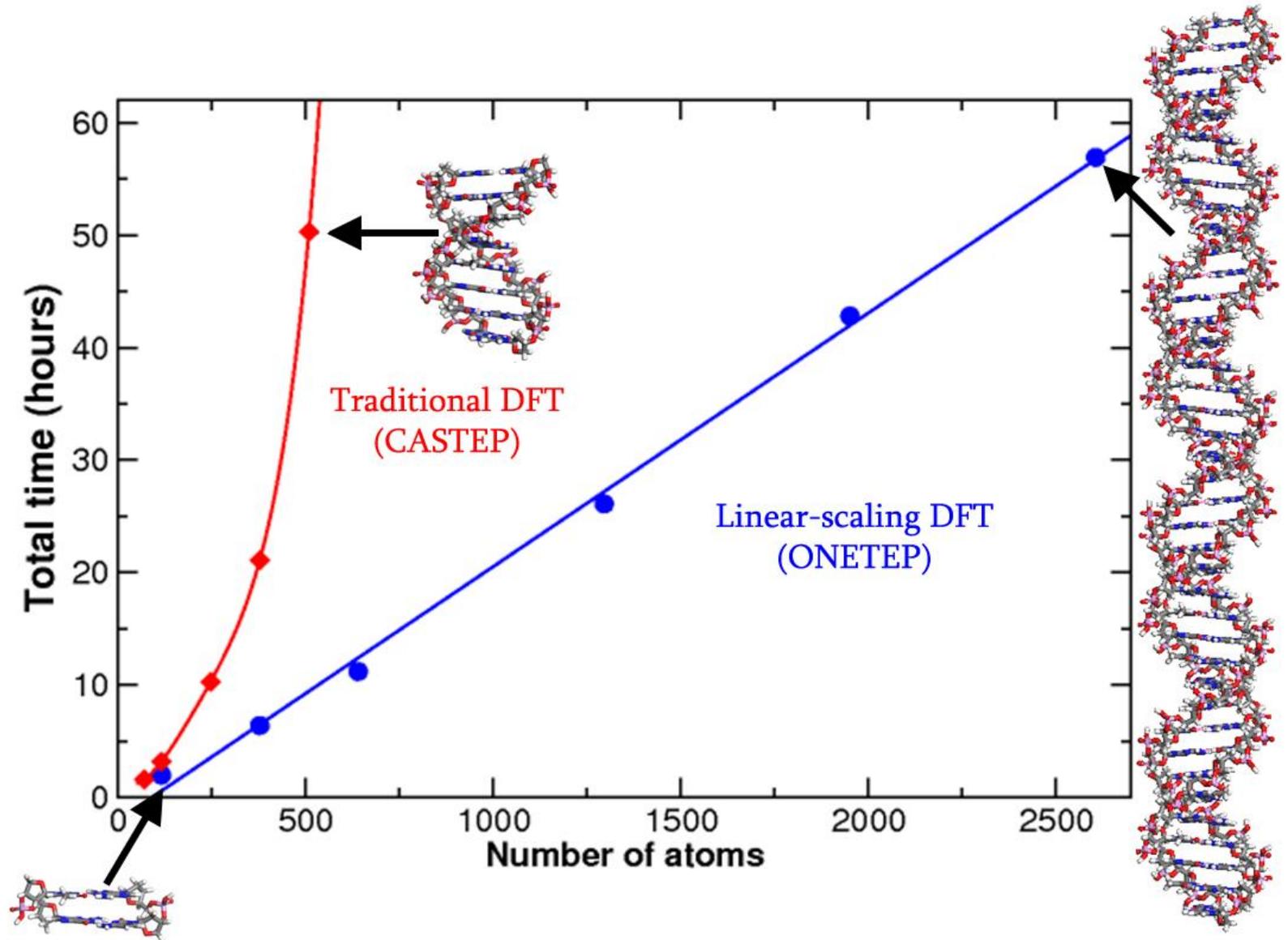


# Components of minimal parameter QM solvation

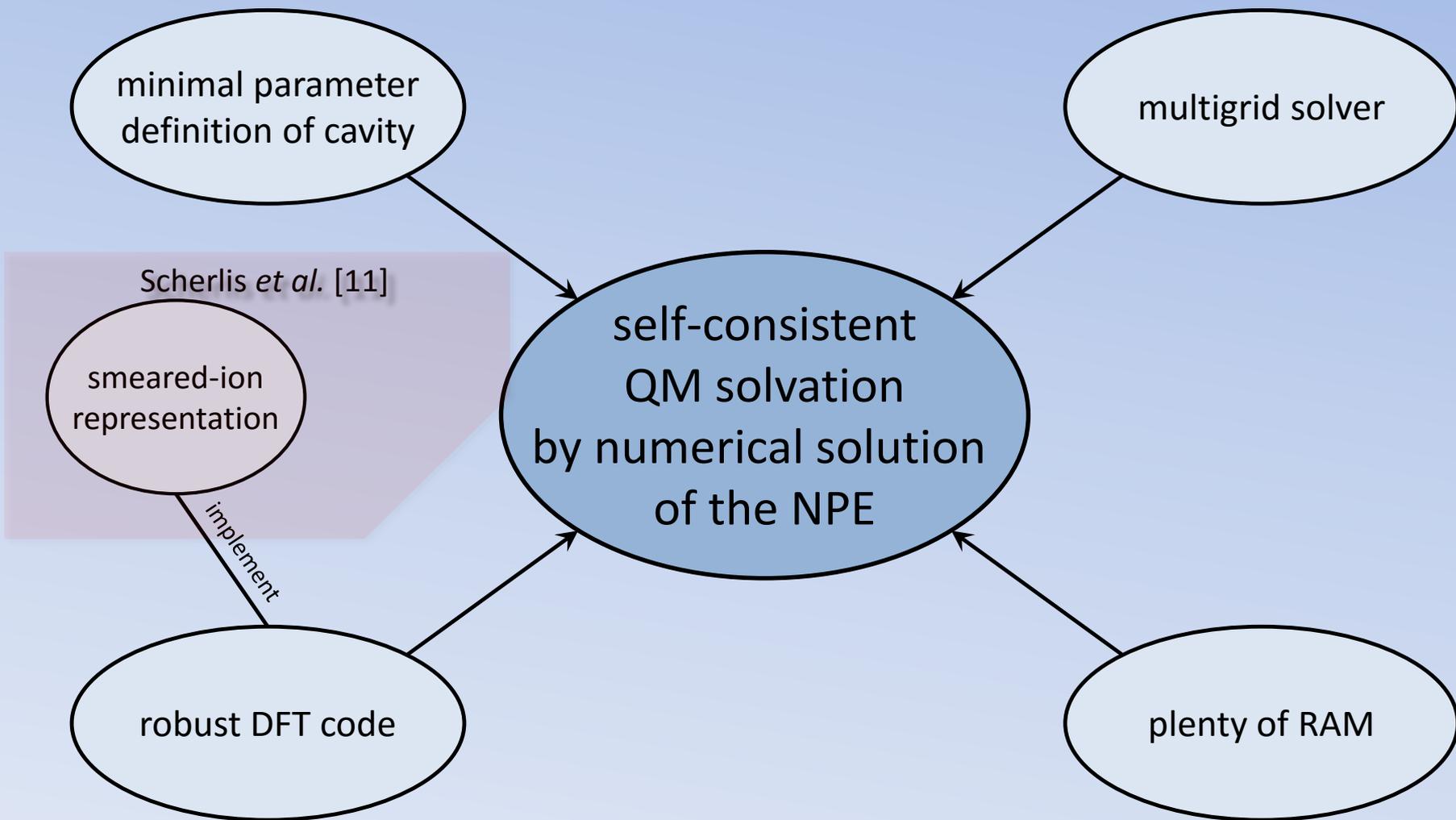




# ONETEP



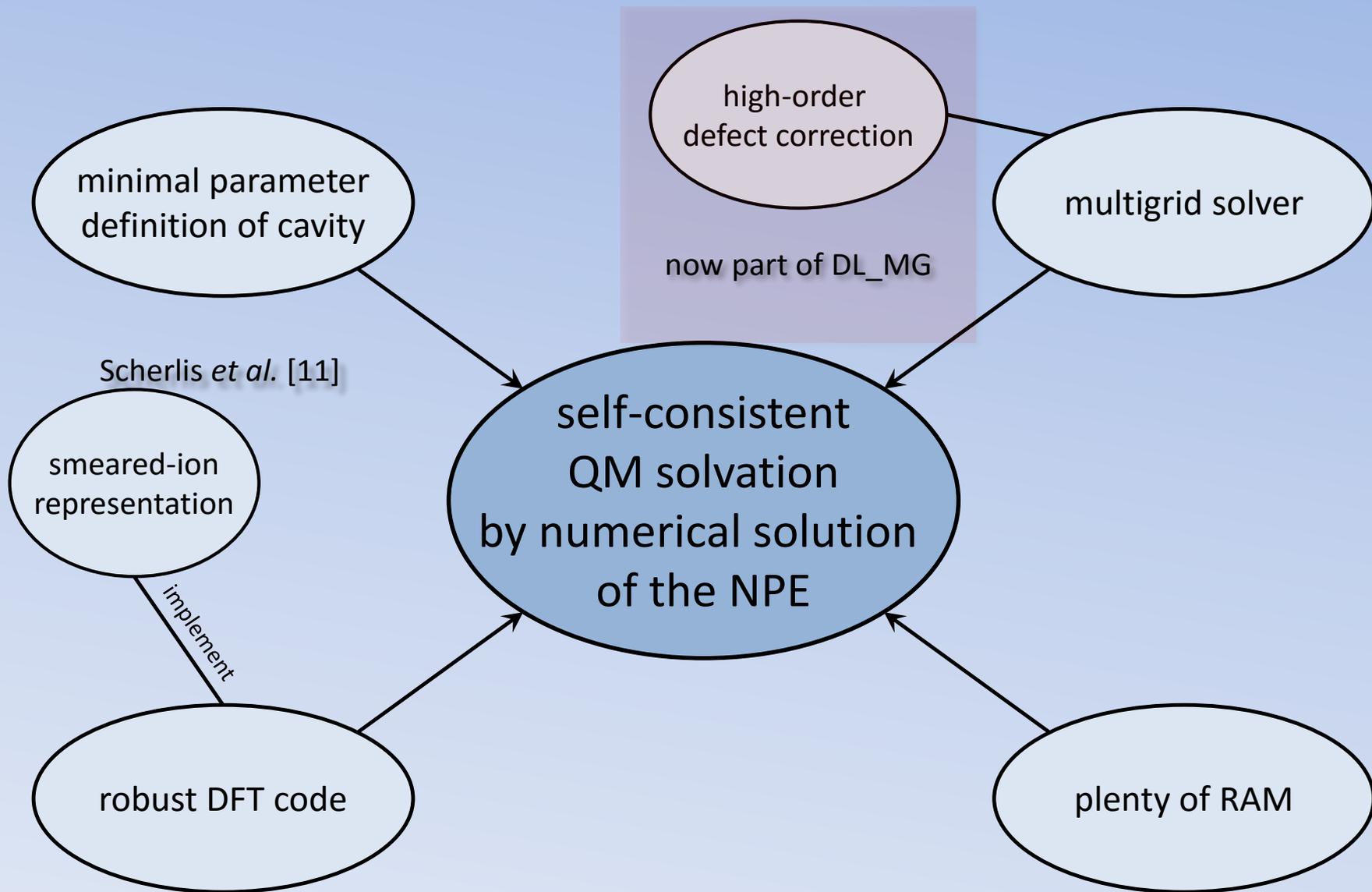
# Components of minimal parameter QM solvation



# Smearred-ion representation

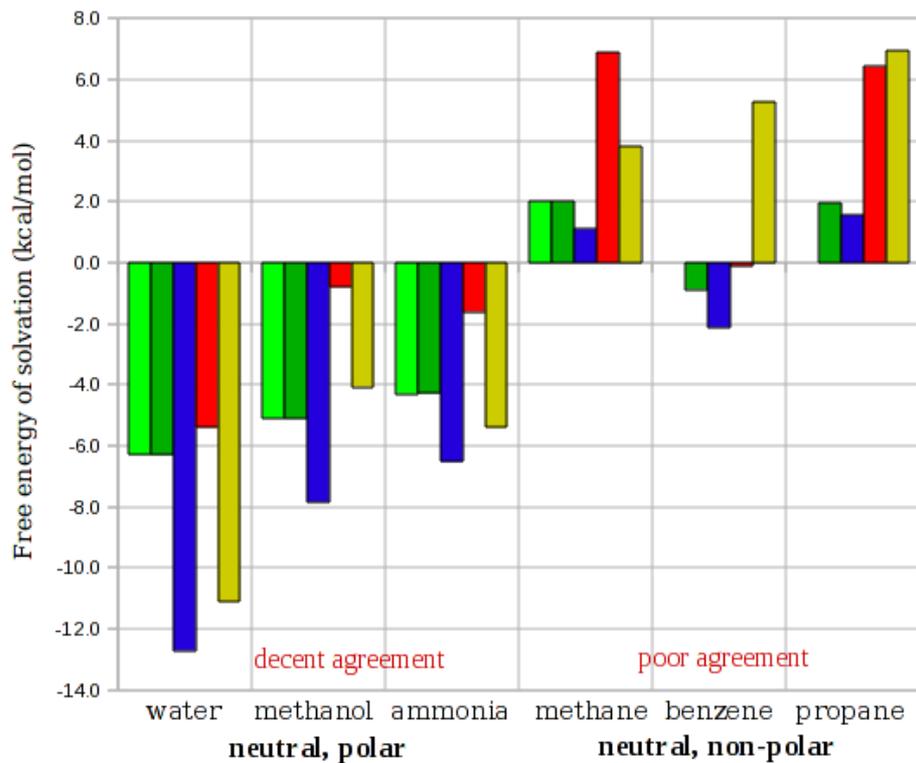
- Another numerical trick.
- From the perspective of the multigrid solver:
  - Point charges of the ionic cores are replaced by thin Gaussians. This makes them representable on a grid. Solver works with total charge density (electrons + smeared ions).
- From the perspective of DFT:
  - Point charges of the ionic cores are replaced by thin Gaussians. This makes it possible to disentangle effects of the solvent on the ionic and electronic degrees of freedom.
  - Correction terms in energy and forces are necessary to cancel out the difference.
  - Energy is re-cast into slightly different terms.
- Advantage – total charge density integrates to something close to zero – numerical errors smaller.

# Components of minimal parameter QM solvation

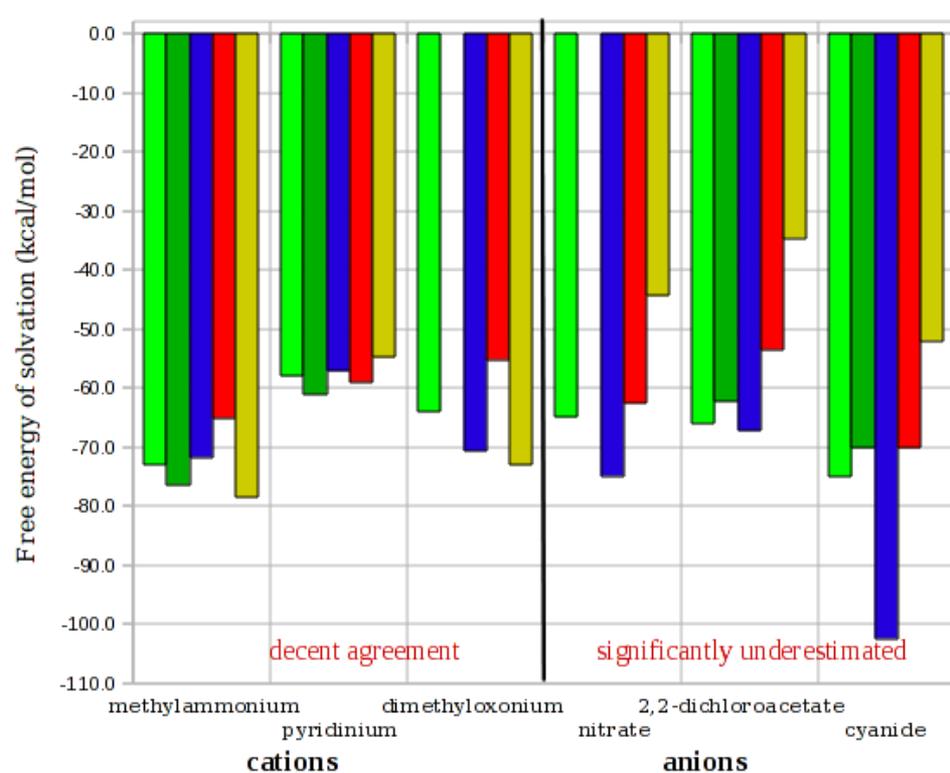


# Does it work?

Free energy of solvation of simple neutral molecules



Free energy of solvation of simple charged molecules



■ Experiment [E1]    
 ■ Experiment [E2]    
 ■ Classical (AMBER)    
 ■ Gaussian PCM  
■ ONETEP (zero BC)

[E1] Refs. 55-57 in Scherlis, Fattebert, Gygi, Cococcioni and Marzari, *J. Chem. Phys.* **124** (2006).

[E2] Marenich, Kelly, Thompson, Hawkins, Chambers, Giesen, Winget, Cramer, Truhlar, *Minnesota Solvation Database* – version 2009, University of Minnesota, Minneapolis, 2009.

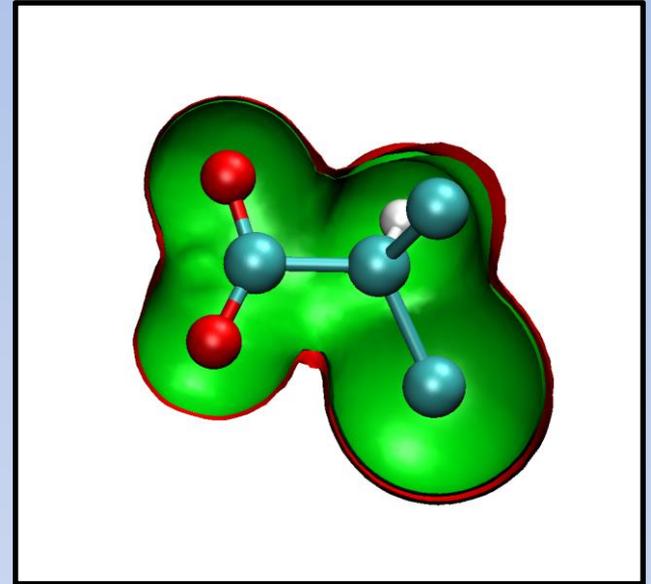
# What's wrong?

- Zero BCs are used when solving the NPE:

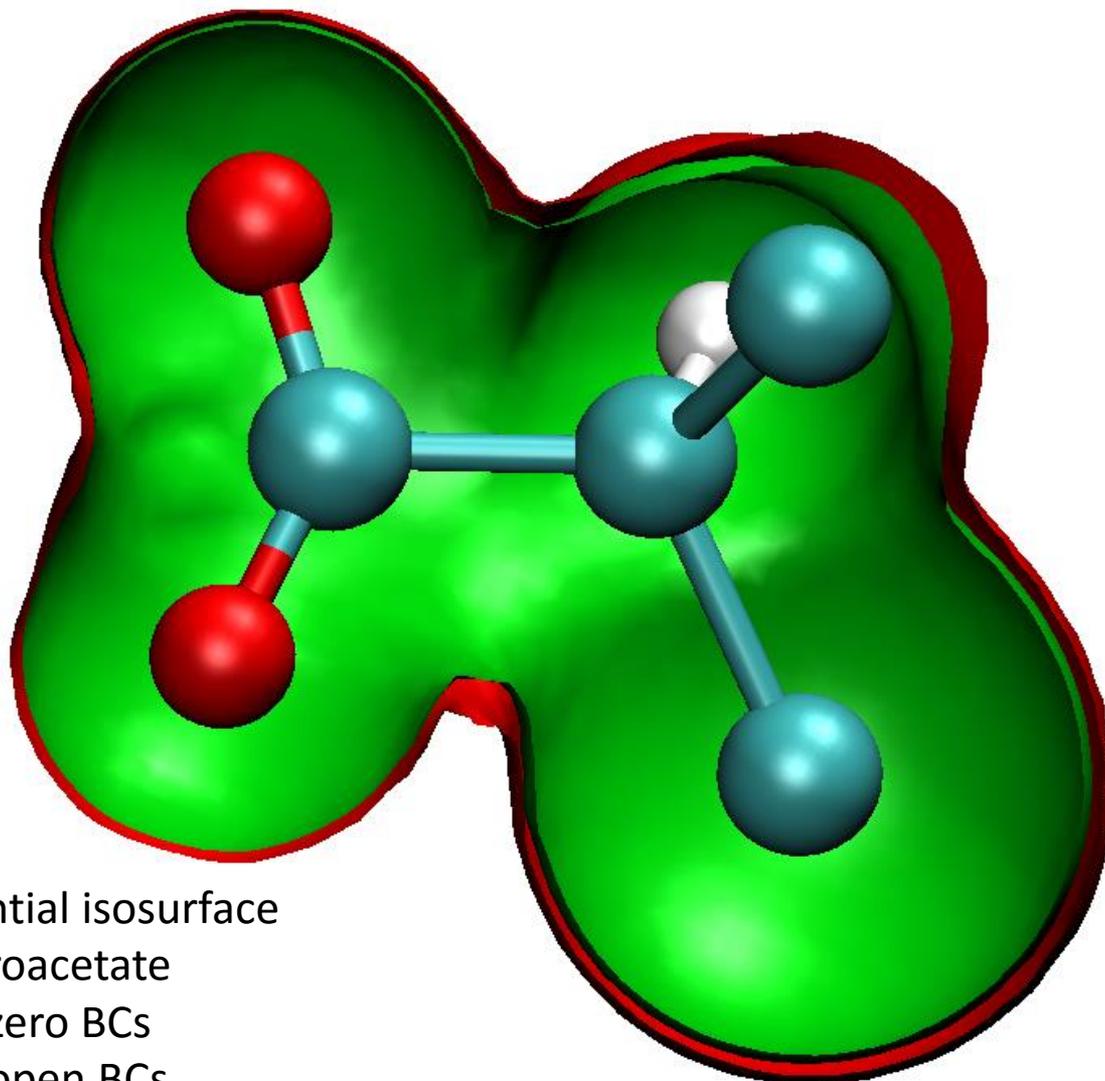
$$\nabla \varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) = -4\pi \rho(\mathbf{r})$$

$$\phi(\mathbf{r}) = 0 \text{ on } \partial\Omega$$

- ... but the potential on the boundary is clearly non-zero.
- This leads to a distortion of the obtained potential.



# Importance of BCs



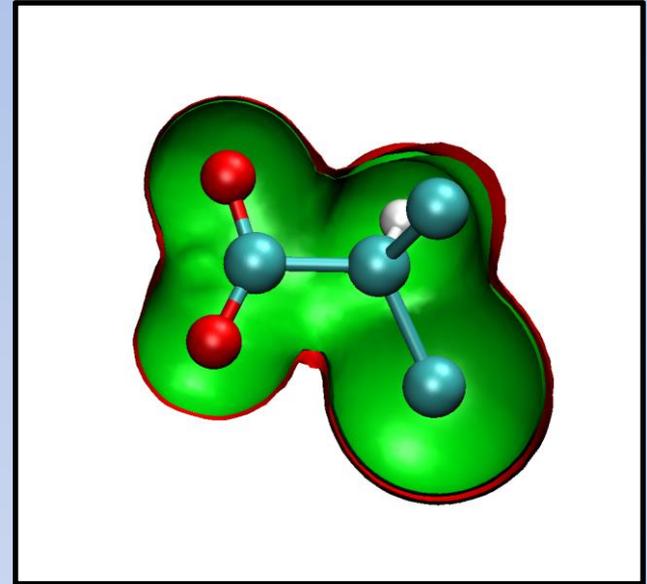
# "Proper" boundary conditions

- Use open BCs when solving the NPE:

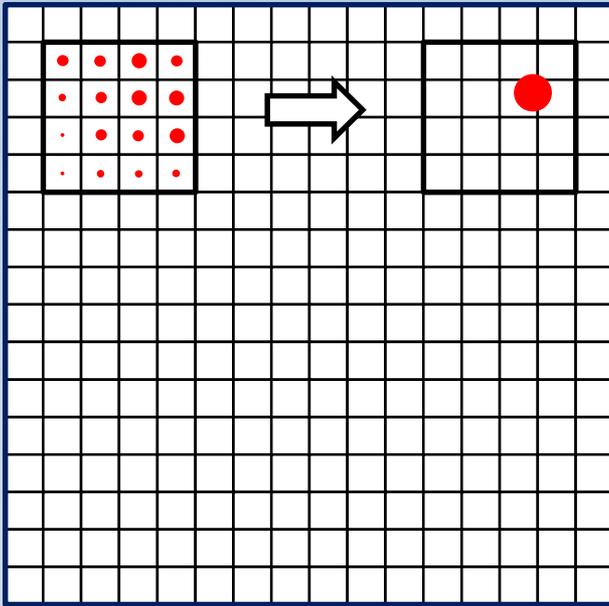
$$\nabla \varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) = -4\pi \rho(\mathbf{r})$$

$$\phi(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \text{ on } \partial\Omega$$

- ... but make sure this does not become the bottleneck.
- ... and make sure that this is consistent with the calculation of other energy terms (ion-electron, ion-ion) – these have to be recast into open BCs as well.



# "Proper" boundary conditions



$$\phi(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \text{ on } \partial\Omega$$

- Coarse-grain the charge on the fine grid over cubic blocks defined by the user (say, 4x4x4).

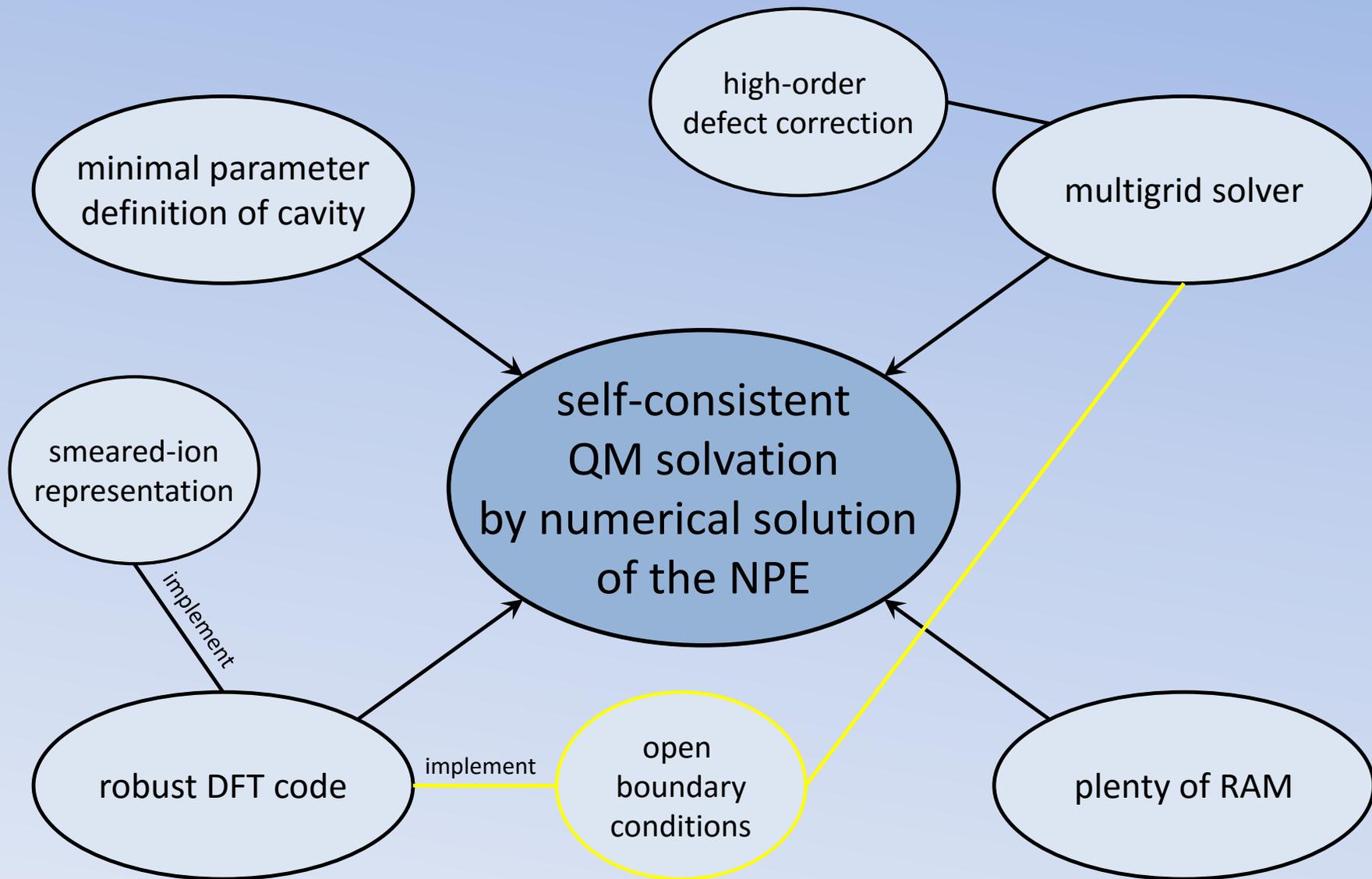
`is_bc_coarseness k`

- Evaluate only for every  $n$ -th (say  $n=3$ ) point on the boundary, use bilinear interpolation in between.

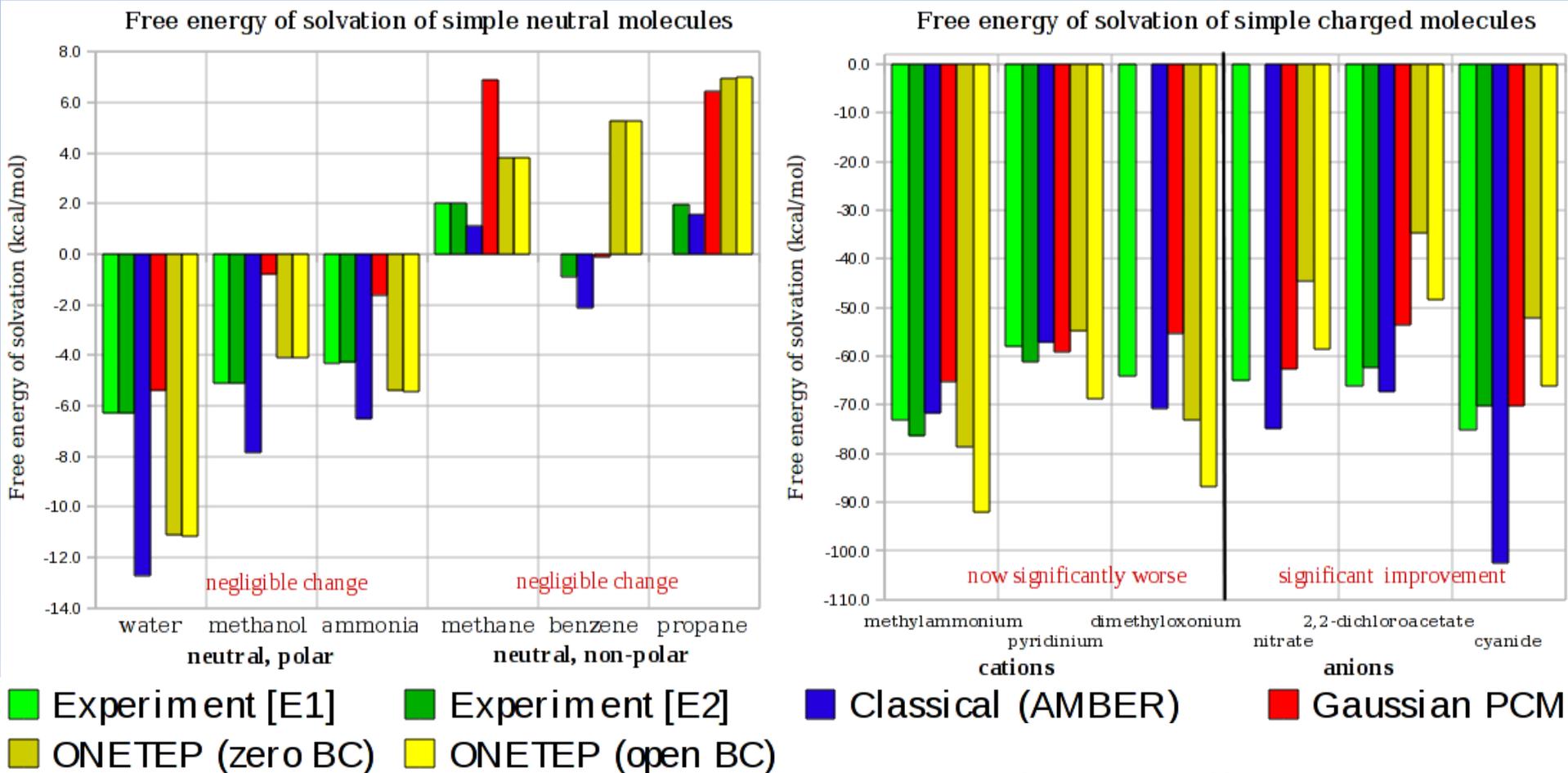
`is_bc_surface_coarseness n`

- Still  $O(L^5)$  scaling, but prefactor is reduced by 1-2 orders of magnitude (by a factor of  $k^3 n^2$ ).

# Components of minimal parameter QM solvation



# Does it work now?



[E1] Refs. 55-57 in Scherlis, Fattebert, Gygi, Cococcioni and Marzari, *J. Chem. Phys.* **124** (2006).

[E2] Marenich, Kelly, Thompson, Hawkins, Chambers, Giesen, Winget, Cramer, Truhlar, *Minnesota Solvation Database* – version 2009, University of Minnesota, Minneapolis, 2009.

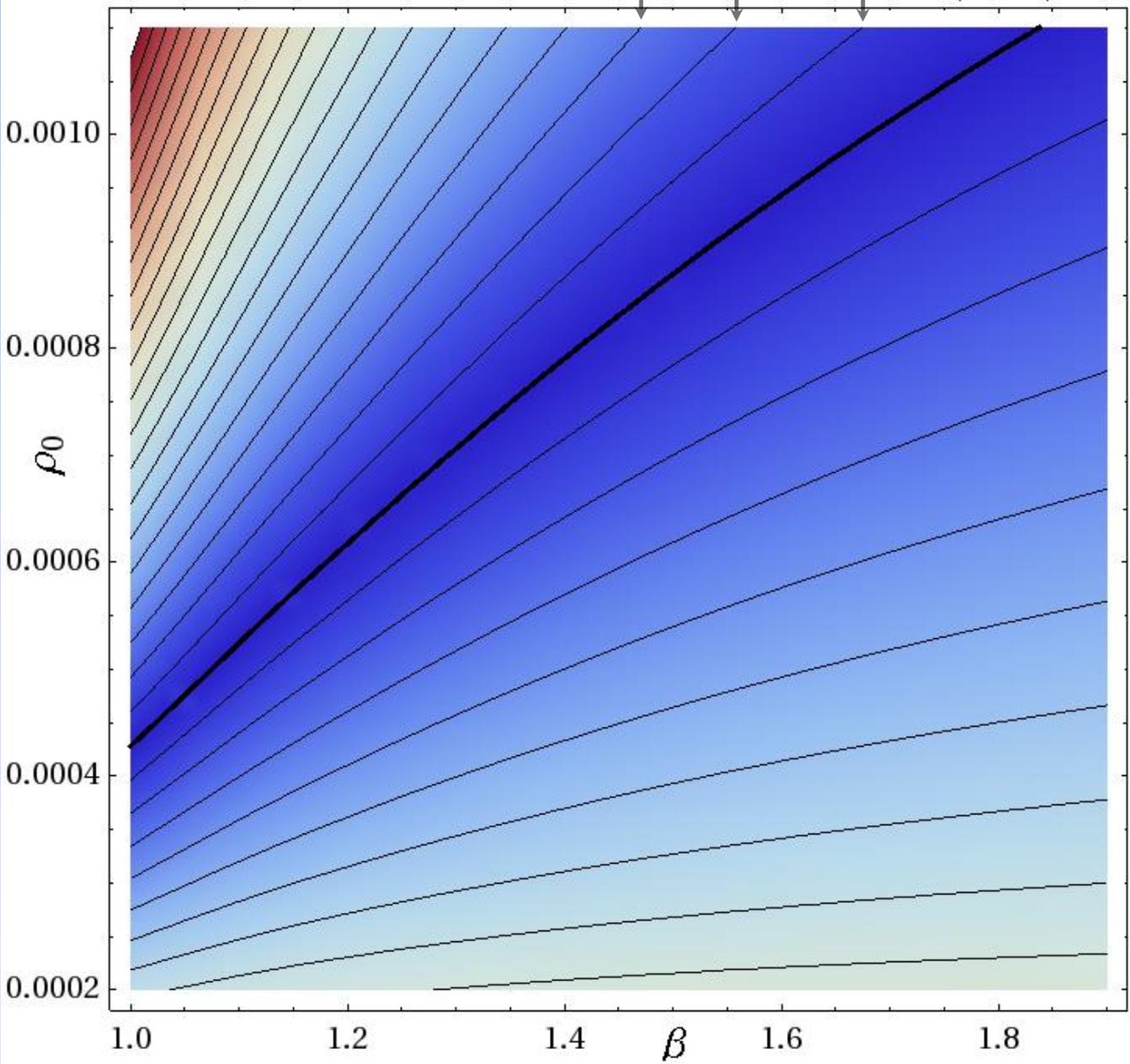
Error in obtained  $\Delta G$   
for a neutral molecule (NH<sub>3</sub>)

**overestimation**

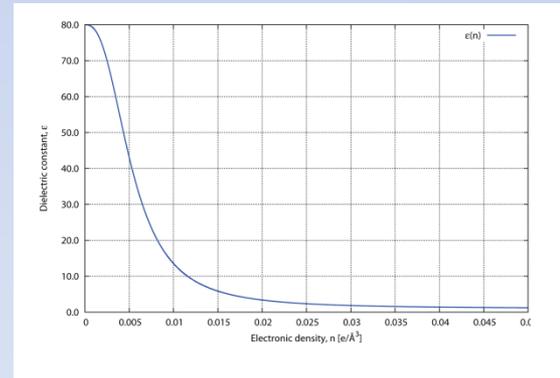
by 3 kcal/mol    by 2 kcal/mol    by 1 kcal/mol

**underestimation**

← by 1 kcal/mol  
← by 2 kcal/mol  
← by 3 kcal/mol  
← by 4 kcal/mol  
← by 5 kcal/mol

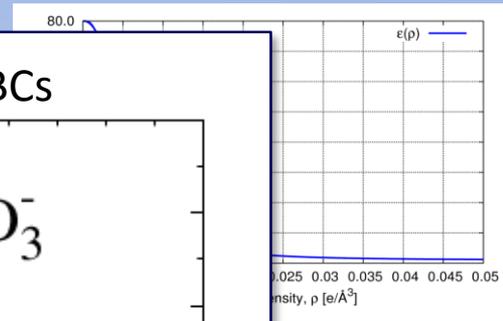
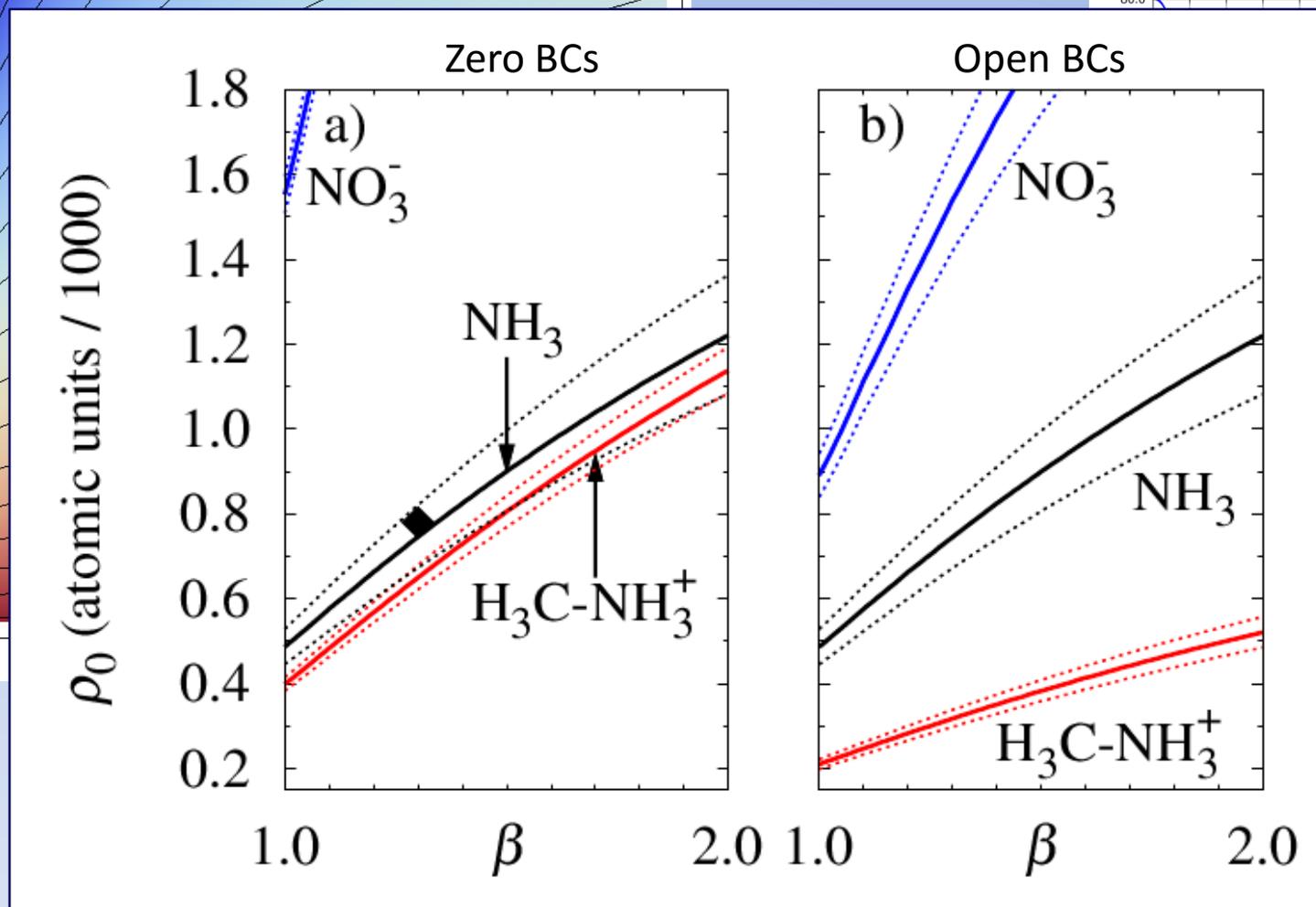
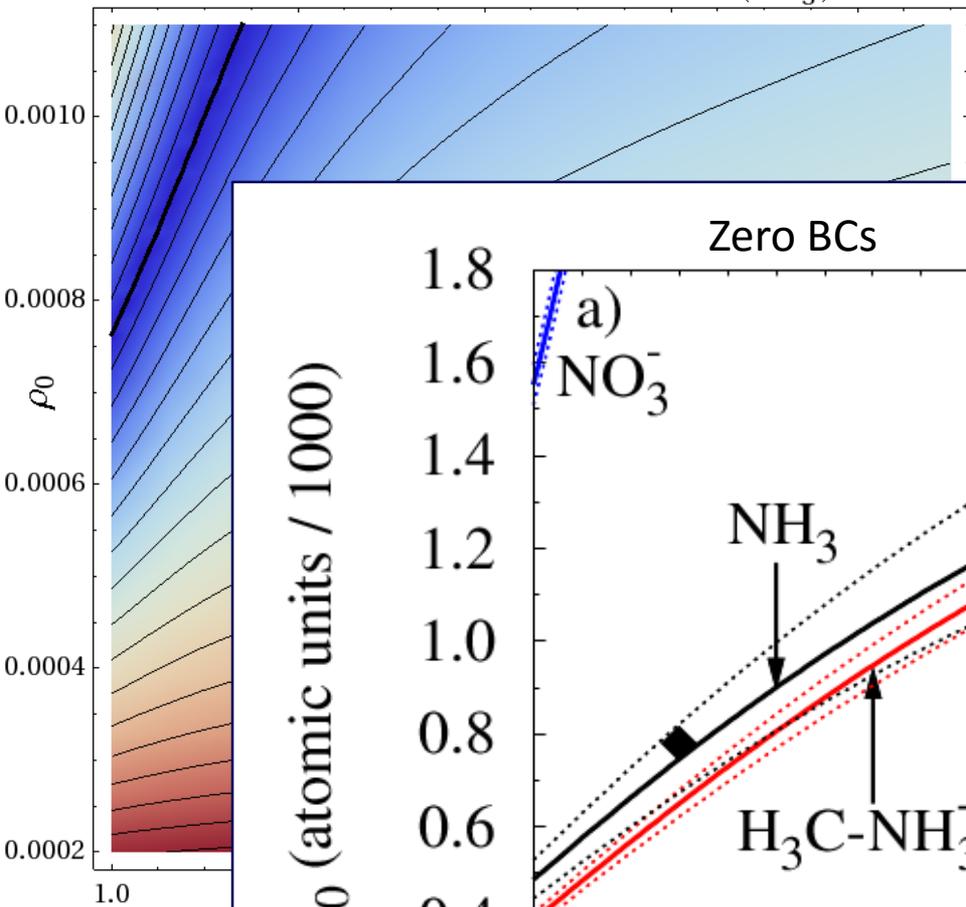


$$\epsilon(n(\mathbf{r})) = 1 + \frac{\epsilon_\infty - 1}{2} \left( 1 + \frac{1 - (n(\mathbf{r})/\rho_0)^{2\beta}}{1 + (n(\mathbf{r})/\rho_0)^{2\beta}} \right)$$

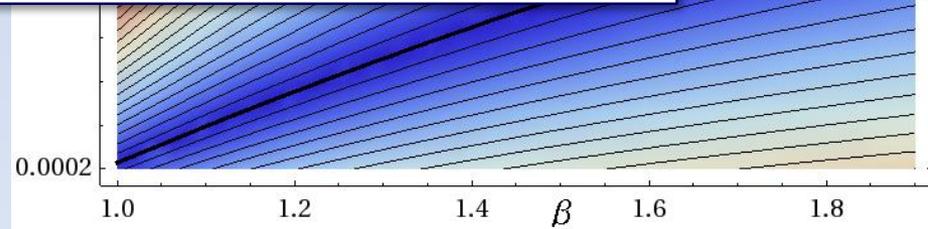


Error in obtained  $\Delta G$  for an anion ( $\text{NO}_3^-$ )

$$\epsilon(n(\mathbf{r})) = 1 + \frac{\epsilon_\infty - 1}{2} \left( 1 + \frac{1 - (n(\mathbf{r})/\rho_0)^{2\beta}}{1 + (n(\mathbf{r})/\rho_0)^{2\beta}} \right)$$

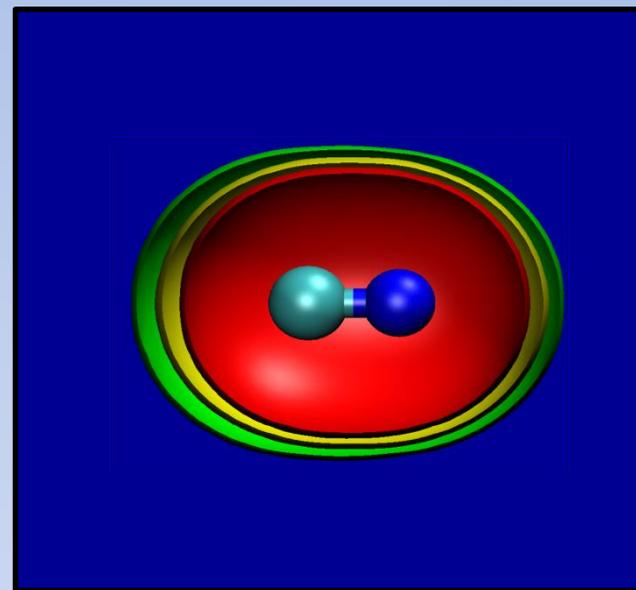


( $\text{CH}_3\text{NH}_3^+$ )



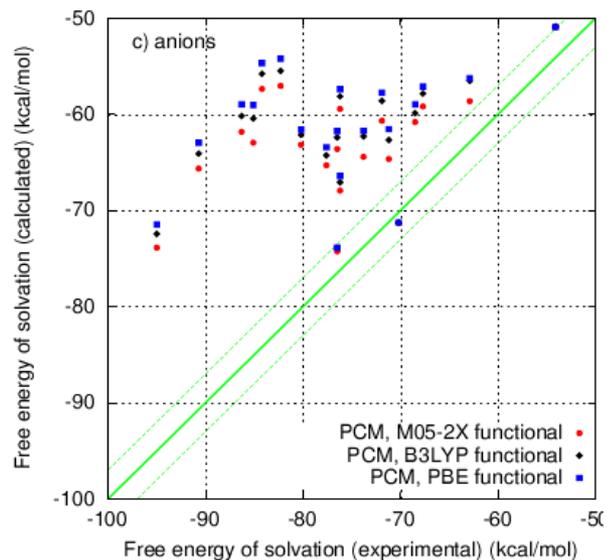
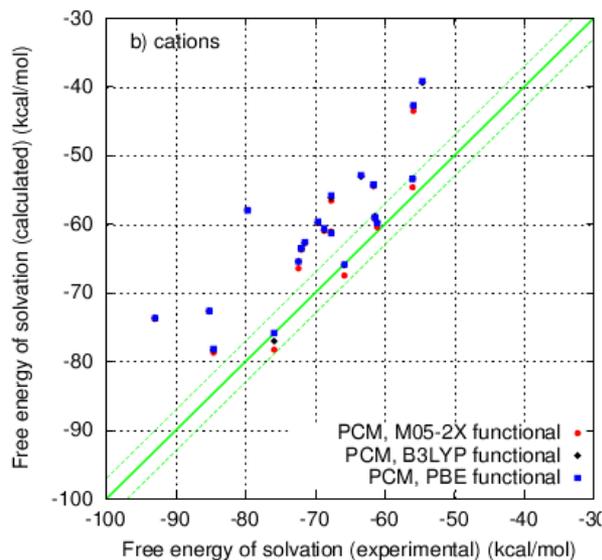
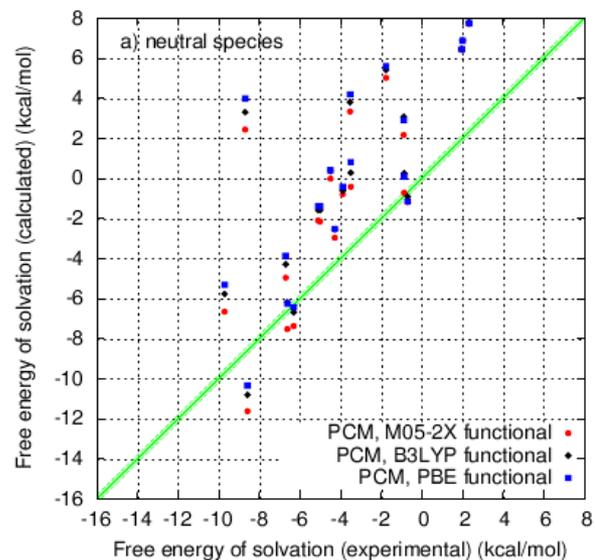
# Why are anions underestimated and cations overestimated?

- For any species  $X$ , this model predicts a larger magnitude of the solvation effect for  $X^+$  than for  $X^-$ , because the cavity for  $X^-$  is bigger.
- In reality the effect is opposite! [8]
- Poor performance of PBE functional – self-interaction error leads to excessive delocalization.



Cavities corresponding to  $\text{CN}^-$  (green) and two "imaginary" molecules:  $\text{CN}$  (yellow) and  $\text{CN}^+$  (red). The smaller the cavity, the larger the magnitude of the solvation effect.

# A better xc functional helps... a little



Approach	neutral species			cations			anions		
	rms error	max error	$r$	rms error	max error	$r$	rms error	max error	$r$
PCM, M05-2X functional	4.4	11.1	0.79	10.2	21.7	0.81	15.7	26.8	0.46
PCM, B3LYP functional	4.7	12.0	0.78	10.4	21.8	0.83	17.0	28.4	0.41
PCM, PBE functional	4.9	12.7	0.75	10.5	21.7	0.83	17.8	29.5	0.36

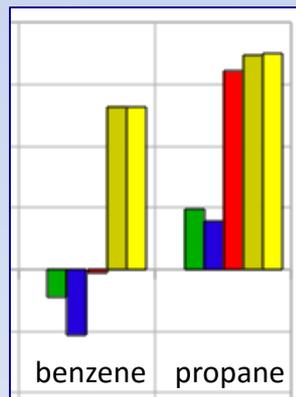
# Dispersion is missing!

only the total can  
be obtained  
experimentally

$$\rightarrow \Delta G_{sol} = \Delta G_{pol} + \Delta G_{npol}$$

- Dispersion is particularly important for large, neutral molecules, where the polar term is small and is easily dwarfed by the nonpolar term.

- The neglect of dispersion might explain the embarrassing failure for benzene and propane.



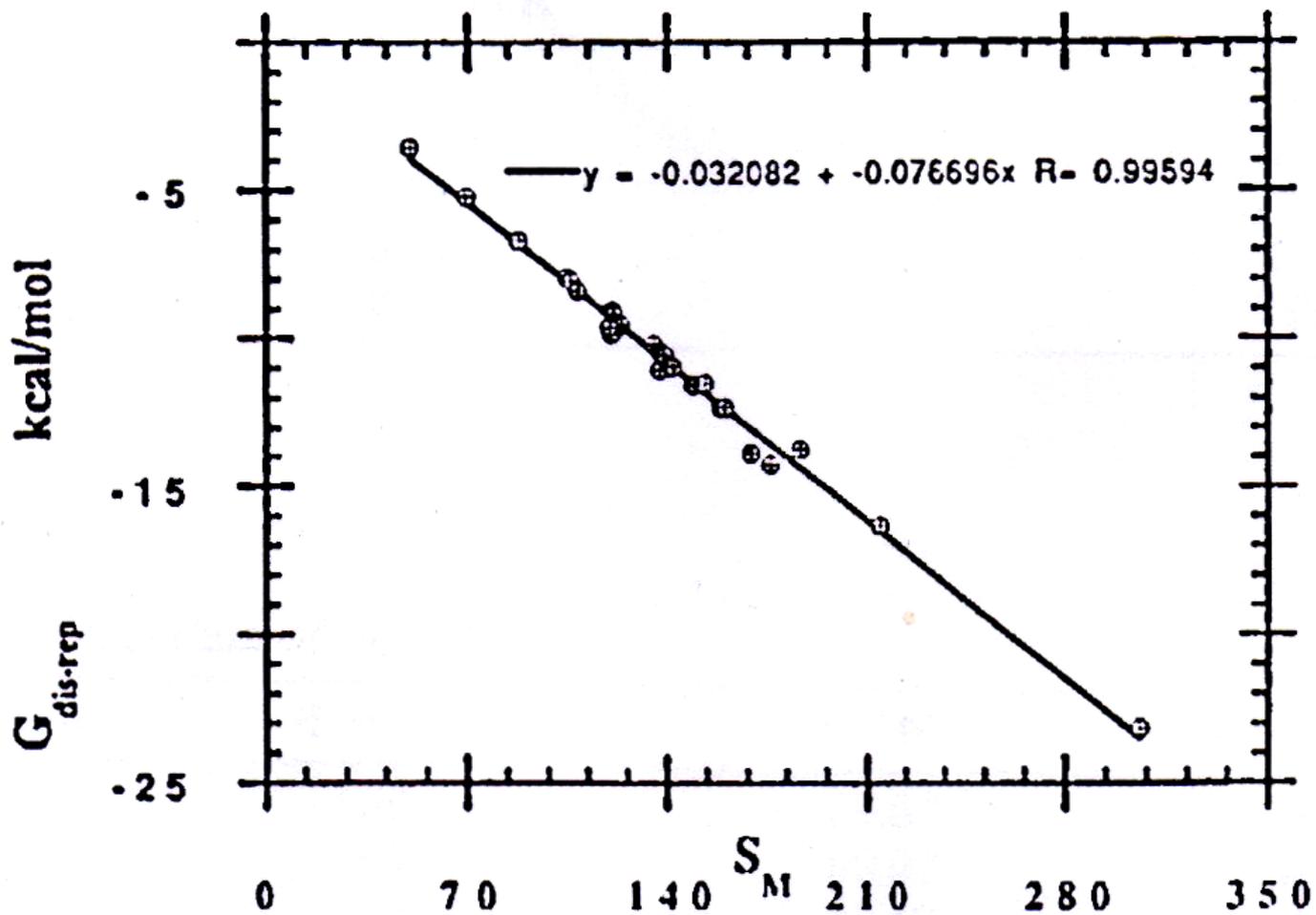
- The **nonpolar term** accounts for
  - the entropic cost of forming a cavity within the solvent (**cavitation energy**),
  - for the van der Waals interaction of the solute with the solvent [4] (**dispersion-repulsion energy**).

- Difficult to describe rigorously.
- A widely used approach is to represent it as a linear function of the molecular surface area

$$[7]: \Delta G_{npol} = \gamma A_{SA}.$$

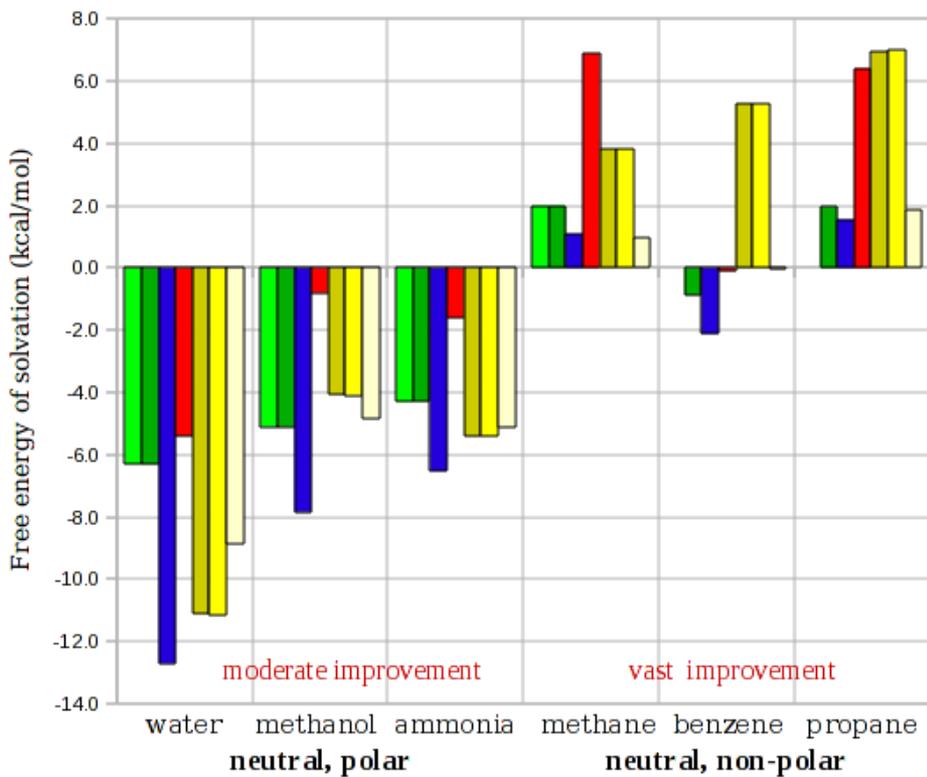
# A crude way to include dispersion-repulsion

[17] Floris, Tomasi and Ahuir, *J. Comp. Chem.* **12** (1991).

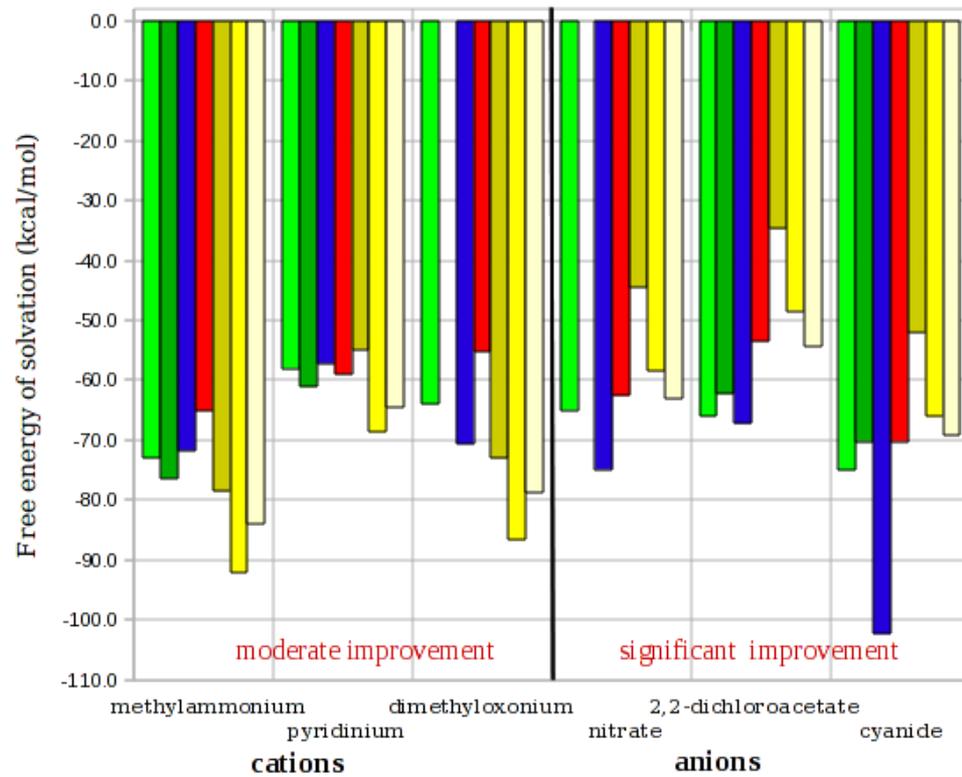


**Figure 1.** Plot of  $G_{\text{dis-rep}}$  ( $M$  in  $L$ ) against  $S_M$ , area of the van der Waals molecular surface.

Free energy of solvation of simple neutral molecules



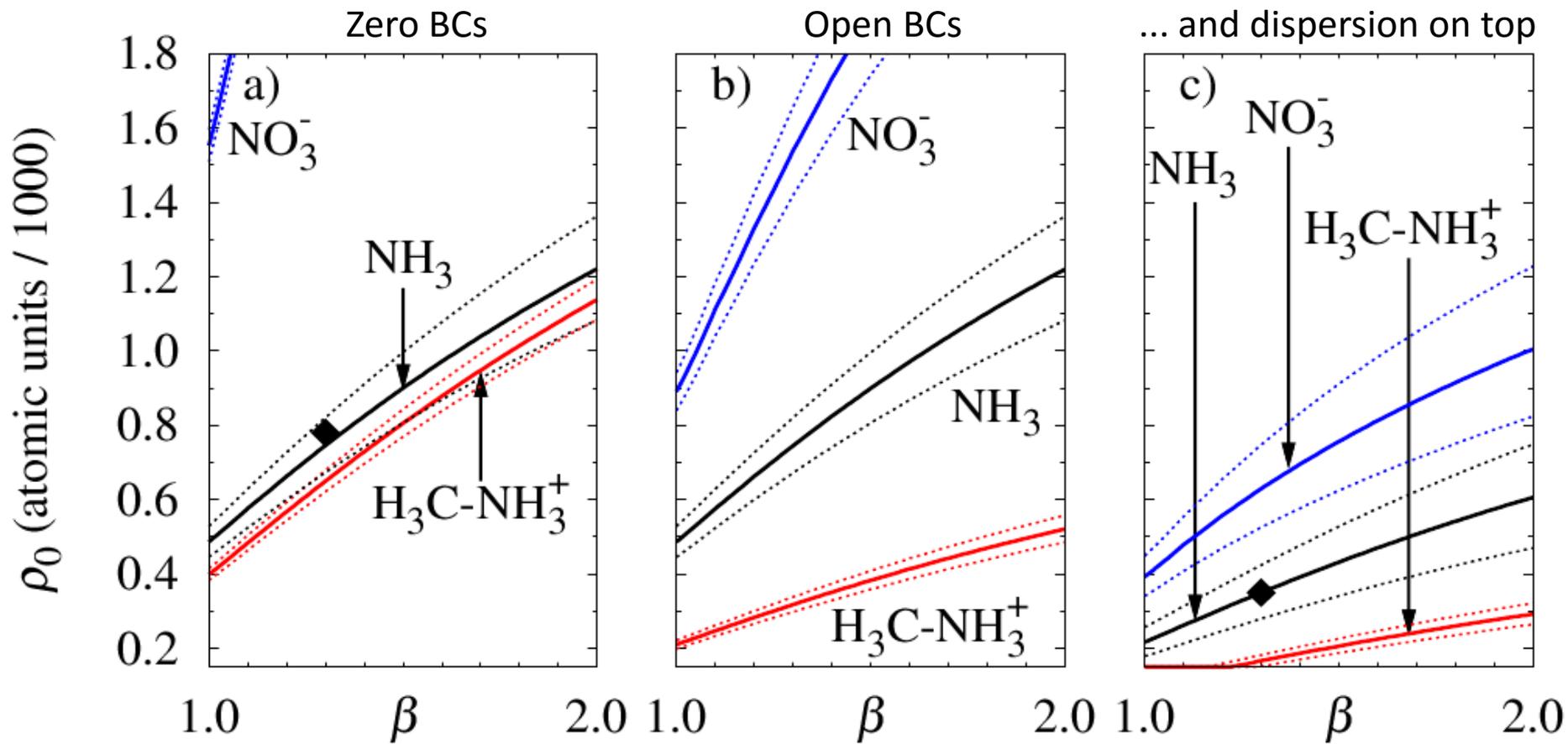
Free energy of solvation of simple charged molecules



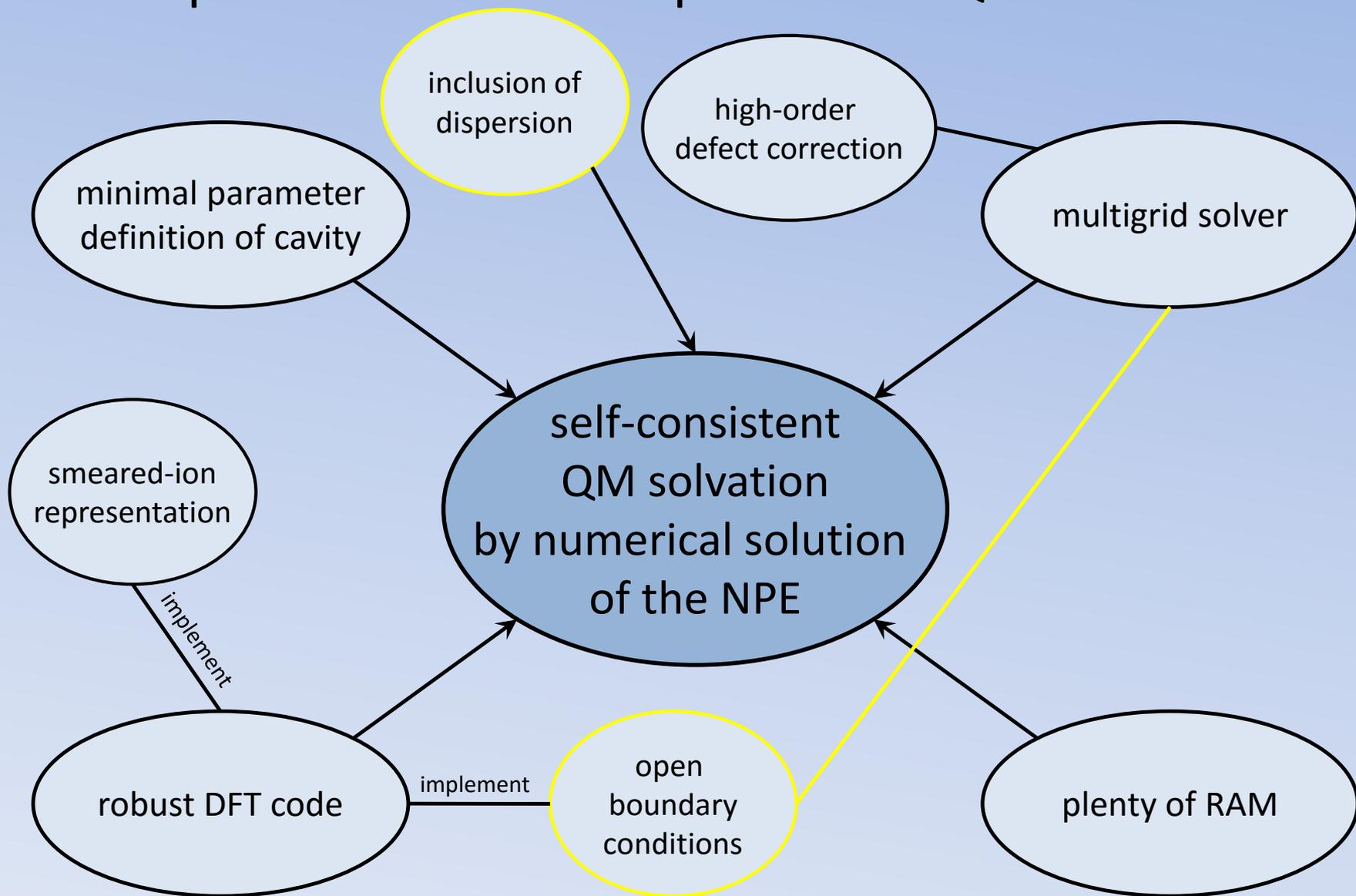
■ Experiment [E1]   
 ■ Experiment [E2]   
 ■ Classical (AMBER)   
 ■ Gaussian PCM  
■ ONETEP (zero BC)   
 ■ ONETEP (open BC)   
 ■ ONETEP (with dis-rep)

[E1] Refs. 55-57 in Scherlis, Fattebert, Gygi, Cococcioni and Marzari, *J. Chem. Phys.* **124** (2006).

[E2] Marenich, Kelly, Thompson, Hawkins, Chambers, Giesen, Winget, Cramer, Truhlar, *Minnesota Solvation Database* – version 2009, University of Minnesota, Minneapolis, 2009.

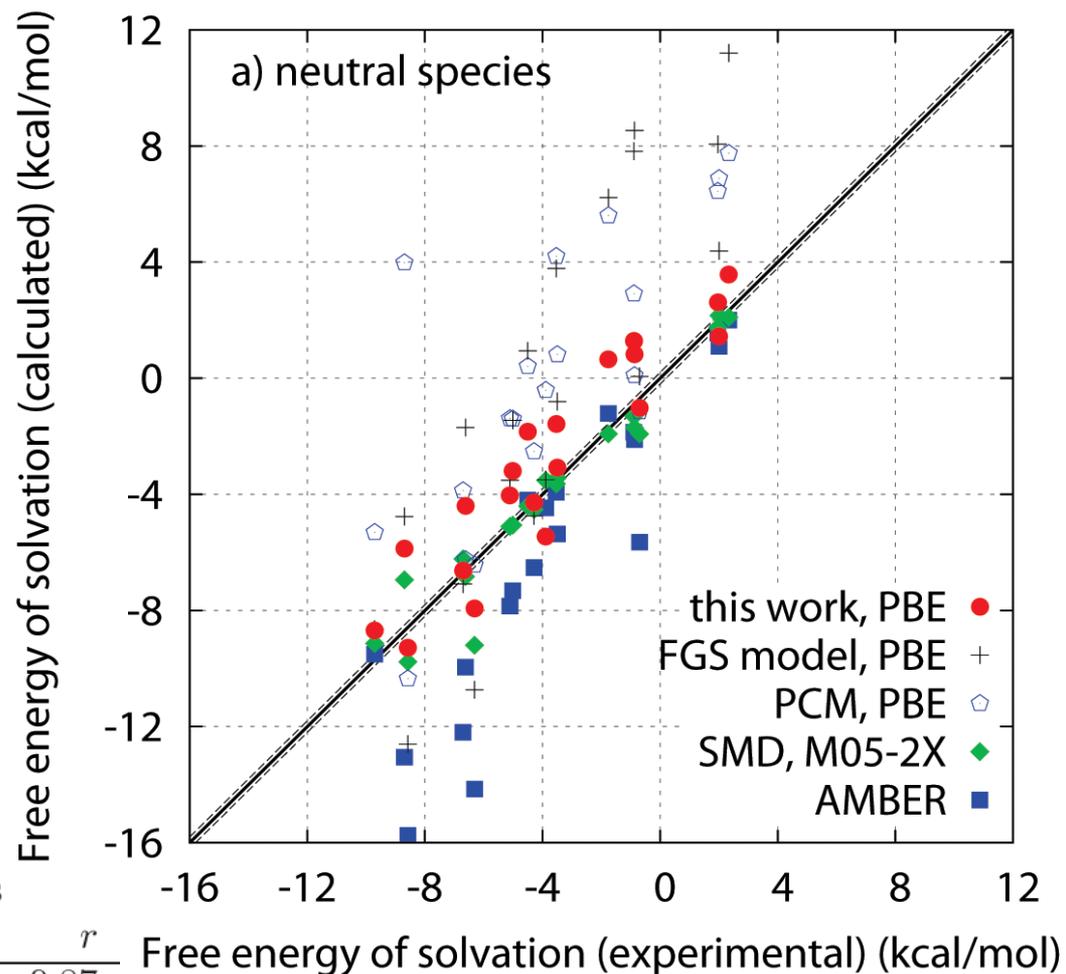


# Components of minimal parameter QM solvation



# Now it clearly works

Validation on 20 neutral molecules:  
*ammonia, water, methane, propane, n-pentane, n-octane, benzene, toluene, methanol, ethanol, 1-hexanol, phenol, diethyl ether, acetaldehyde, 2-pentanone, acetic acid, hydrogen peroxide, ethylamine, acetonitrile, acetamide, hydrogen sulfide, trimethylphosphate.*

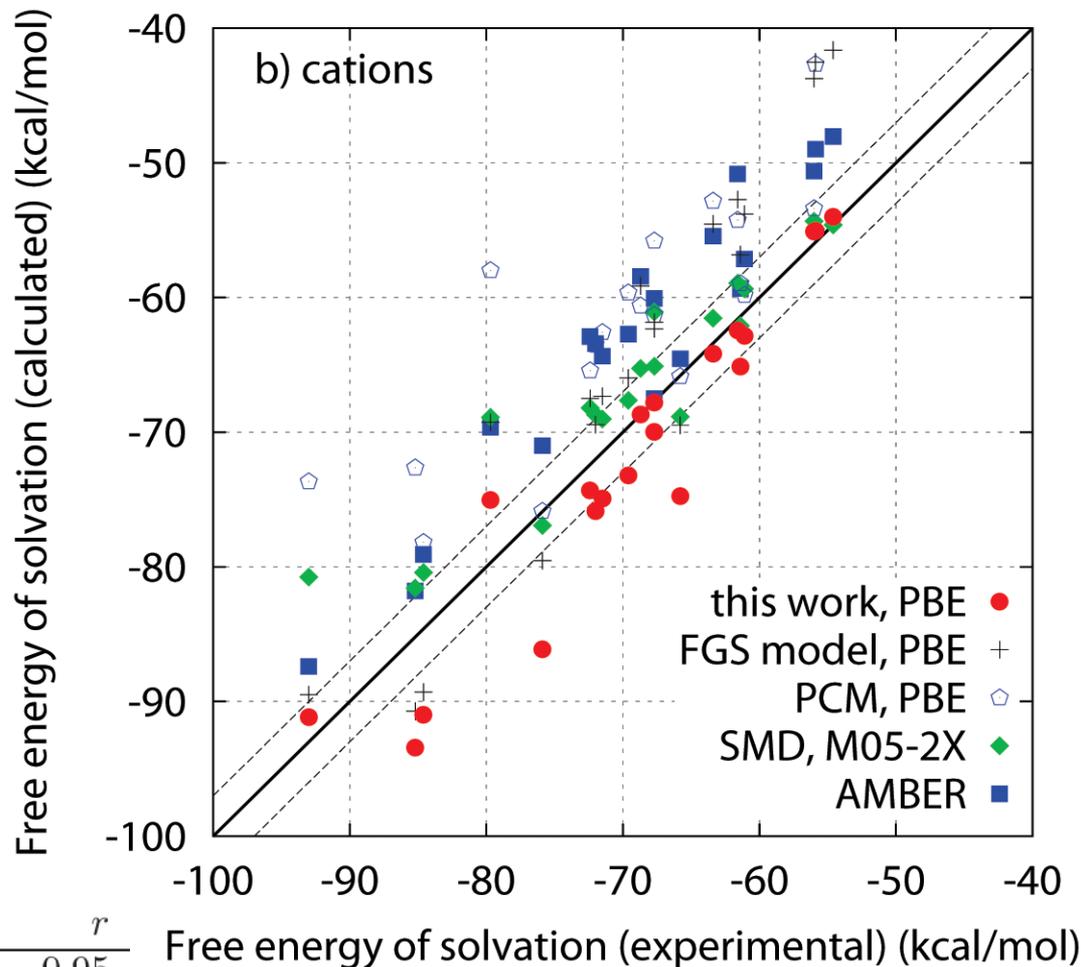


Approach	XC functional	neutral species		
		rms err.	max err.	$r$
+ FGS	PBE	5.2	9.4	0.87
● this work	PBE	1.6	2.8	0.93
PCM	B3LYP	4.7	12.0	0.78
PCM	M05-2X	4.4	11.1	0.79
◆ SMD	M05-2X	0.9	2.9	0.97
■ AMBER	(classical)	3.3	7.84	0.64

# Now it clearly works

Validation on 20 cations:

*ethanol<sup>+</sup>, dimethyl ether<sup>+</sup>, hydrazine<sup>+</sup>, methanol<sup>+</sup>, 3-aminoaniline<sup>+</sup>, 4-methyl-N,N-dimethylaniline<sup>+</sup>, allylamine<sup>+</sup>, aniline<sup>+</sup>, azetidine<sup>+</sup>, cyclohexanamine<sup>+</sup>, diallylamine<sup>+</sup>, diethylamine<sup>+</sup>, isopropylamine<sup>+</sup>, N-ethylaniline<sup>+</sup>, n-propylamine<sup>+</sup>, pyridine<sup>+</sup>, pyrrole<sup>+</sup>, quinoline<sup>+</sup>, triethylamine<sup>+</sup>, 4-nitroaniline<sup>+</sup>, dimethyl sulfoxide<sup>+</sup>, ammonia<sup>+</sup>.*

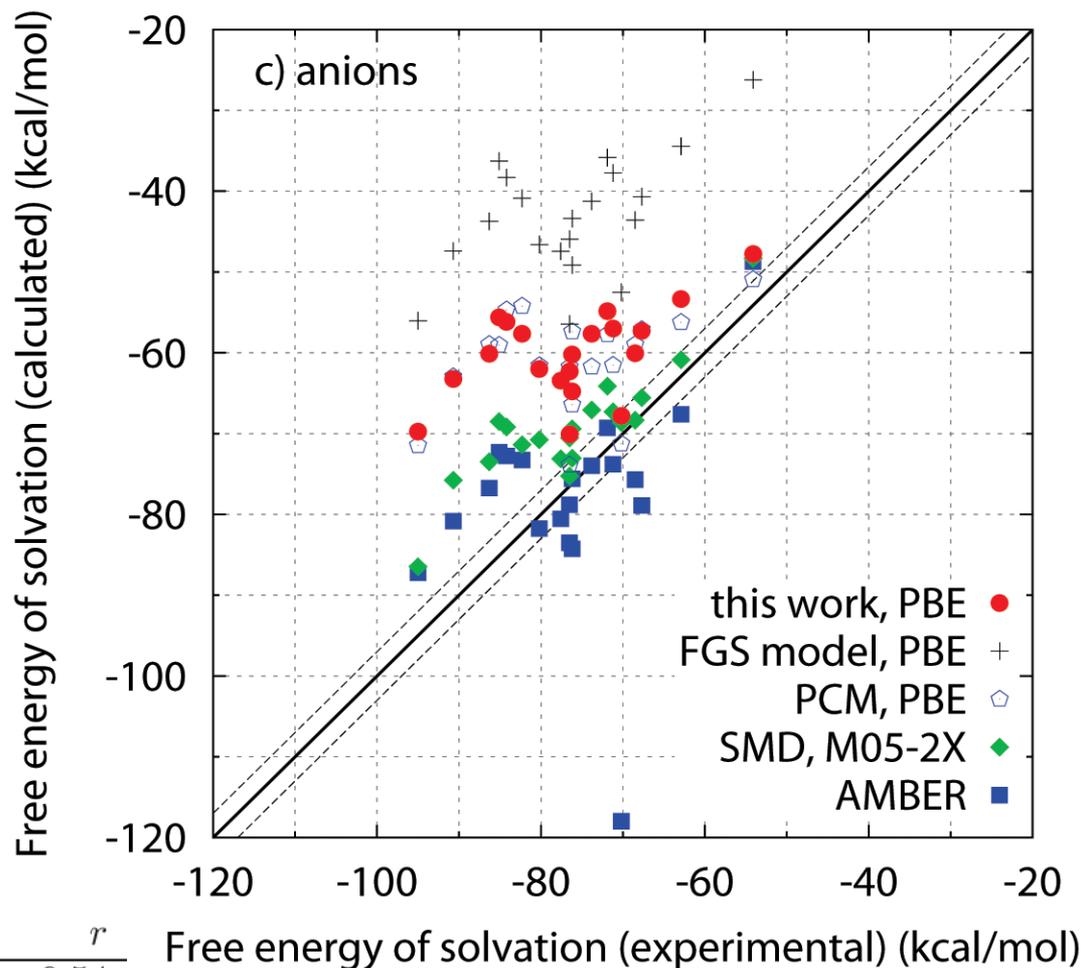


Approach	XC functional	rms err.	max err.	$r$
+ FGS	PBE	7.6	13.4	0.95
● this work	PBE	4.4	10.2	0.95
PCM	B3LYP	10.4	21.8	0.83
PCM	M05-2X	10.2	21.7	0.81
◆ SMD	M05-2X	4.6	12.2	0.95
■ AMBER	(classical)	6.9	10.8	0.96

# Now it clearly works

Validation on 20 anions:

*acetylene<sup>-</sup>, 2-butanol<sup>-</sup>, acetaldehyde<sup>-</sup>, acetic acid<sup>-</sup>, ethanol<sup>-</sup>, phenol<sup>-</sup>, hydrogen cyanide<sup>-</sup>, 3-hydroxyphenol<sup>-</sup>, acetone<sup>-</sup>, benzoic acid<sup>-</sup>, benzyl alcohol<sup>-</sup>, formica acid<sup>-</sup>, isopropanol<sup>-</sup>, methyl hydroperoxide<sup>-</sup>, pyruvic acid<sup>-</sup>, t-butanol<sup>-</sup>, aniline<sup>-</sup>, tricholormethane<sup>-</sup>, acetamide<sup>-</sup>, dimethyl sulfoxide<sup>-</sup>.*



Approach	XC functional	rms err.	max err.	$r$
+ FGS	PBE	34.2	48.8	0.54
● this work	PBE	18.1	29.5	0.53
PCM	B3LYP	17.0	28.4	0.41
PCM	M05-2X	15.7	26.8	0.46
◆ SMD	M05-2X	8.5	16.6	0.86
■ AMBER	(classical)	12.8	47.8	0.32

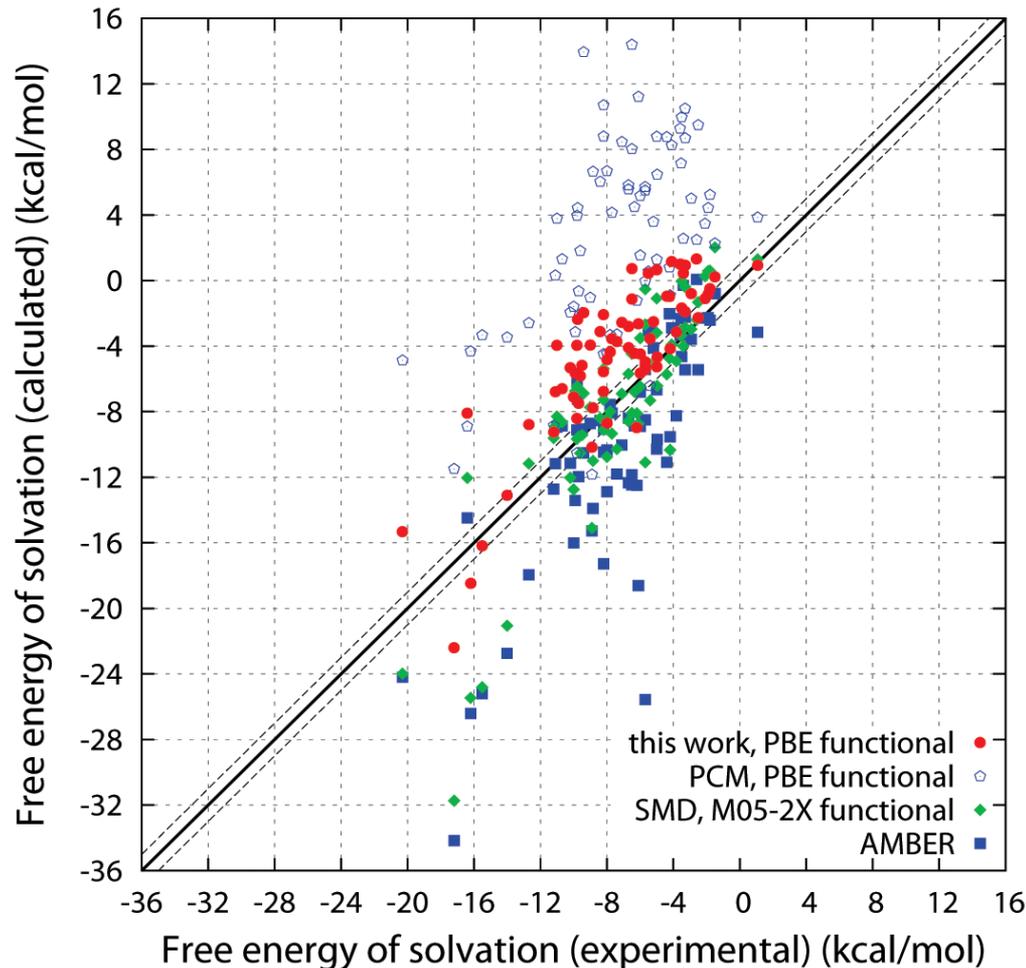
- Nicholls, Mobley, Guthrie, Chodera, Bayly, Cooper and Pande, "Predicting Small-Molecule Solvation Free Energies: An Informal **Blind Test** for Computational Chemistry", *J. Med. Chem.* **51** (2008).
- Guthrie, "A **Blind Challenge** for Computational Solvation Free Energies: Introduction and Overview", *J. Phys. Chem. B* **113** (2009)

ID in MnSol	Species	ID in MnSol	Species
2853	1,3-diacetyloxypropan-2-yl acetate	2889	die thoxy-[(2-isopropyl-6-methyl-4-pyrimidinyl)oxy]-thioxophosphorane (diazinon)
2854	m-bis(trifluoromethyl) benzene	2890	3,6-dichloro-2-methoxybenzoic acid (dicamba)
2855	N,N-dimethyl-p-methoxybenzamide	2891	N1,N1-dimethyl-2,6-dinitro-4-trifluoromethyl-m-phenylenediamine (dinitramine)
2856	N,N,4-trimethylbenzamide	2892	(RS)-2-sec-butyl-4,6-dinitrophenol (dinoseb)
2857	bis(2-chloroethyl) ether	2893	6,7,8,9,10,10-hexachloro-1,5,5a,6,9,9a-hexahydro-6,9-methano-2,4,3-benzodioxathiepine-3-oxide (endosulfan alpha)
2858	1,1-diacetoxyethane	2894	(1R,4S,4aS,5S,6S,7R,8R,8aR)-1,2,3,4,10,10-hexachloro-1,4,4a,5,6,7,8,8a-... ...octahydro-6,7-epoxy-1,4:5,8-dimethanonaphthalene (endrin)
2859	1,1-dithoxyethane	2895	O,O,O',O'-tetraethyl S,S'-methylene bis(phosphorodithioate) (ethion)
2860	diethyl propanedioate	2896	1,4,5,6,7,8,8-hexachloro-3a,4,7,7a-tetrahydro-4,7-methanoindene (heptachlor)
2861	dimethoxymethane	2897	3,5-trimethyl-2-cyclohexen-1-one (isophorone)
2862	ethylene glycol diacetate	2898	1,2,3,4,5,6-hexachlorocyclohexane (lindane)
2863	1,2-dithoxyethane	2899	2-(dimethoxyphosphinothioylthio) butanedioic acid diethyl ester (malathion)
2864	phenyl formate	2900	N-methylcarbamic acid [1-(methylthio)ethylideneamino] ester (methomyl)
2865	imidazole	2901	methyl 2-(4-methoxy-6-methyl-1,3,5-triazin-2-ylcarbamoylsulfamoyl)benzoate (metsulfuron methyl)
2866	nitroglycol	2902	4-methyl-2,6-dinitro-N,N-dipropylaniline (nitralin)
2867	1,2-dinitroxypropane	2903	nitroxyacetone
2868	butyl nitrate	2904	O,O-dimethyl-O-4-nitrophenylthiophosphate (parathion)
2869	2-butyl nitrate	2905	S-propyl butyl(ethyl)thiocarbamate (pebulate)
2870	isobutyl nitrate	2906	O,O-dimethyl S-ethylthiomethyl phosphorodithioate (phorate)
2871	ethylene glycol mononitrate	2907	N-cyclopropylmethyl-2,6-dinitro-N-propyl-4-trifluoromethylaniline (profluralin)
2872	2-chloro-N-(2,6-dimethylphenyl)-N-(methoxymethyl)acetamide (alachlor)	2908	N2,N4-diisopropyl-6-methylthio-1,3,5-triazine-2,4-diamine (prometryn)
2873	2-methyl-2-(methylthio)propanal O-(N-methylcarbamoyl)oxime (aldicarb)	2909	N-(3,4-dichlorophenyl)propanamide (propanil)
2874	2-(ethylamino)-4-isopropylamino-6-methylthio-s-triazine (ametryn)	2910	5-amino-4-chloro-2-phenyl-3-(2H)-pyridazinone (pyrazon)
2875	O,O-dimethyl S-[(4-oxo-1,2,3-benzotriazin-3(4H)-yl)methyl] dithiophosphate (azinphos methyl)	2911	6-chloro-N,N'-dimethyl-1,3,5-triazine-2,4-diamine (simazine)
2876	N-butyl-N-ethyl-2,6-dinitro-4-(trifluoromethyl)aniline (benefin)	2912	methyl 2-(4,6-dimethylpyrimidin-2-ylcarbamoylsulfamoyl)benzoate (sulfometuron-methyl)
2877	a-[(4,6-dimethoxypyrimidin-2-ylcarbamoyl)sulfamoyl]-o-toluic acid (bensulfuron)	2913	3-t-butyl-5-chloro-6-methyluracil (tebacil)
2878	5-Bromo-3-sec-butyl-6-methyl-uracil (bromacil)	2914	N2-tert-butyl-N4-ethyl-6-methylthio-1,3,5-triazine-2,4-diamine (tebutryn)
2879	3a,4,7,7a-tetrahydro-2-[(trichloromethyl)thio]-1H-indole-1,3(2H)-dione (captan)	2915	3-(4-methoxy-6-methyl-1,3,5-triazin-2-ylcarbamoylsulfamoyl)thiophene-2-carboxylic acid (thifensulfuron)
2880	1-naphthyl methylcarbamate (carbaryl)	2916	dimethyl (RS)-2,2,2-trichloro-1-hydroxyethylphosphonate (trichlorfon)
2881	2,3-dihydro-2,2-dimethyl-7-benzofuran-1-ylmethylcarbamate (carbofuran)	2917	a,a,a-trifluoro-2,6-dinitro-N,N-dipropyl-p-toluidine (trifluralin)
2882	S-4-chlorophenylthiomethyl O,O-dimethyl phosphorodithioate (carbophenothion)	2918	S-propyl dipropyl(thiocarbamate) (vermolate)
2883	octachloro-4,7-methanohydroindane (chlordane)	2919	4-amino-4'-nitroazobenzene
2884	phosphoric acid [(E)-2-chloro-1-(2,4-dichlorophenyl)vinyl] diethyl ester (chlorfenvinphos)	2920	1-amino-4-anilinoanthraquinone
2885	ethyl 2-(4-chloro-6-methoxypyrimidin-2-ylcarbamoylsulfamoyl)benzoate (chlorimuron ethyl)	2921	1,4,5,8-tetraminoanthraquinone
2886	trichloro(nitro)methane (chloropicrin)	2922	1-aminoanthraquinone
2887	O,O-dimethyl O-3,5,6-trichloro-2-pyridyl phosphorothioate (chlorpyrifos)	2923	(2-dimethylamino-5,6-dimethyl-pyrimidin-4-yl) N,N-dimethylcarbamate (pirimor, pirimicarb)
2888	S-(RS)-2-chloro-1-phthalimidomethyl O,O-dimethyl phosphorodithioate (dialifor)		

# Now it clearly works

Validation on 71 medium-size neutral molecules from

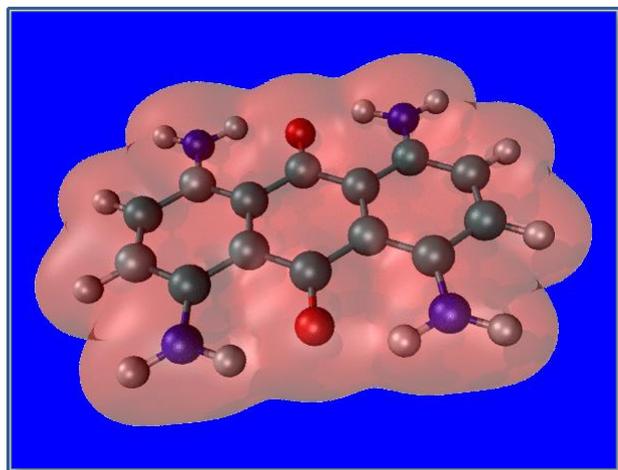
- Nicholls, Mobley, Guthrie, Chodera, Bayly, Cooper and Pande, "Predicting Small-Molecule Solvation Free Energies: An Informal **Blind Test** for Computational Chemistry", *J. Med. Chem.* **51** (2008).
- Guthrie, "A **Blind Challenge** for Computational Solvation Free Energies: Introduction and Overview", *J. Phys. Chem. B* **113** (2009).



Approach	XC functional	rms error	max error	$r$
● this work <sup>a</sup>	PBE	3.8	8.3	0.83
● this work <sup>b</sup>	PBE	4.1	9.1	0.83
◇ PCM	PBE	10.9	23.3	0.53
◆ SMD	M05-2X	3.4	14.5	0.87
■ AMBER	(classical)	5.1	19.9	0.77

# Now it clearly works

Approach	XC functional	rms error	max error	$r$
● this work <sup>a</sup>	PBE	3.8	8.3	0.83
● this work <sup>b</sup>	PBE	4.1	9.1	0.83
◇ PCM	PBE	10.9	23.3	0.53
◆ SMD	M05-2X	3.4	14.5	0.87
■ AMBER	(classical)	5.1	19.9	0.77

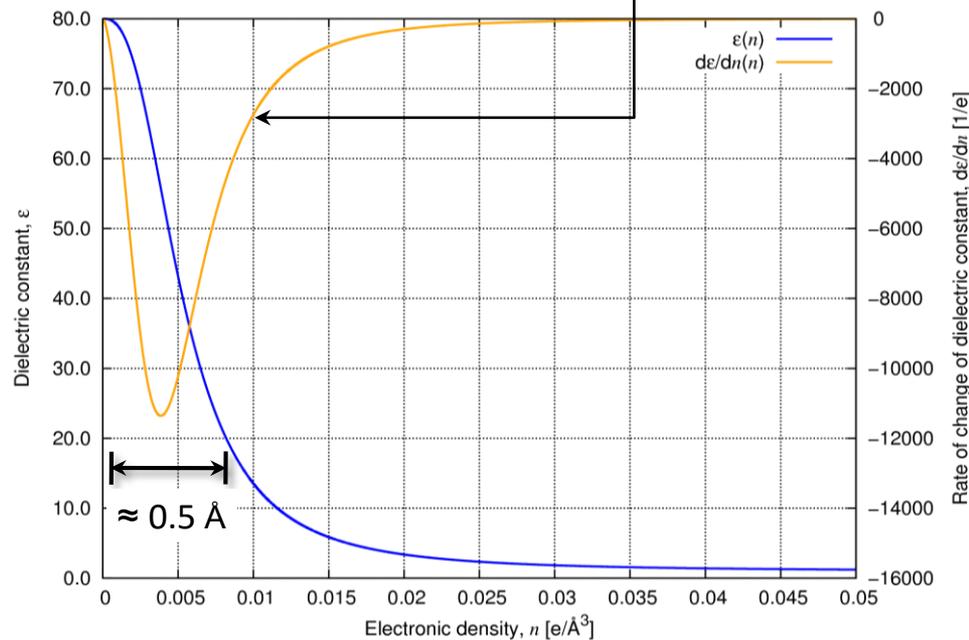


NB how the density hardly changes at all in solvent.

# Self-consistent updating of the cavity is very costly

- When the dielectric cavity reacts to changes in density, extra terms appear in the energy gradient:

$$\frac{\delta E_{es}}{\delta n}(\mathbf{r}) = \phi(\mathbf{r}) - \frac{1}{8\pi}(\nabla\phi(\mathbf{r}))^2 \frac{\delta\epsilon}{\delta n}(\mathbf{r}) + \frac{\Delta G_{cav}}{\delta n}(\mathbf{r})$$

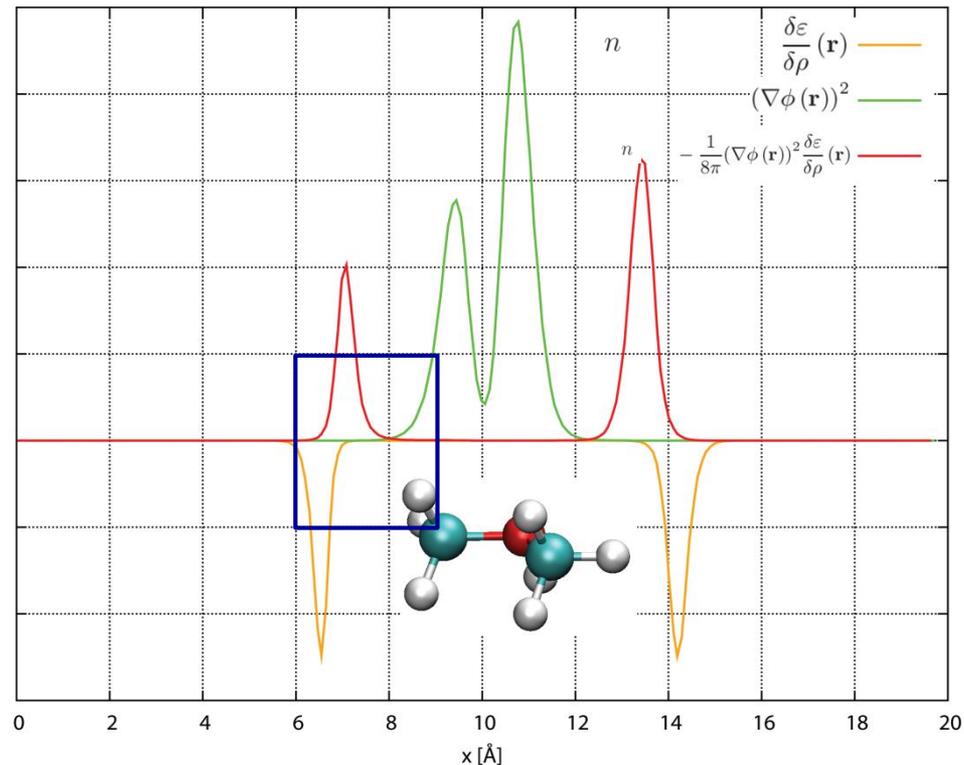
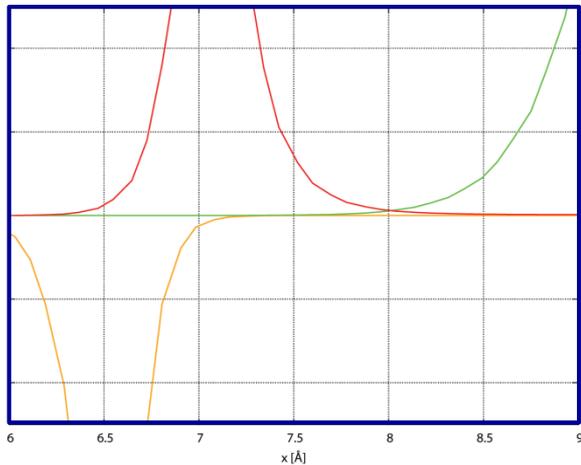


$$\epsilon(n(\mathbf{r})) = 1 + \frac{\epsilon_{\infty} - 1}{2} \left( 1 + \frac{1 - (n(\mathbf{r})/\rho_0)^{2\beta}}{1 + (n(\mathbf{r})/\rho_0)^{2\beta}} \right)$$

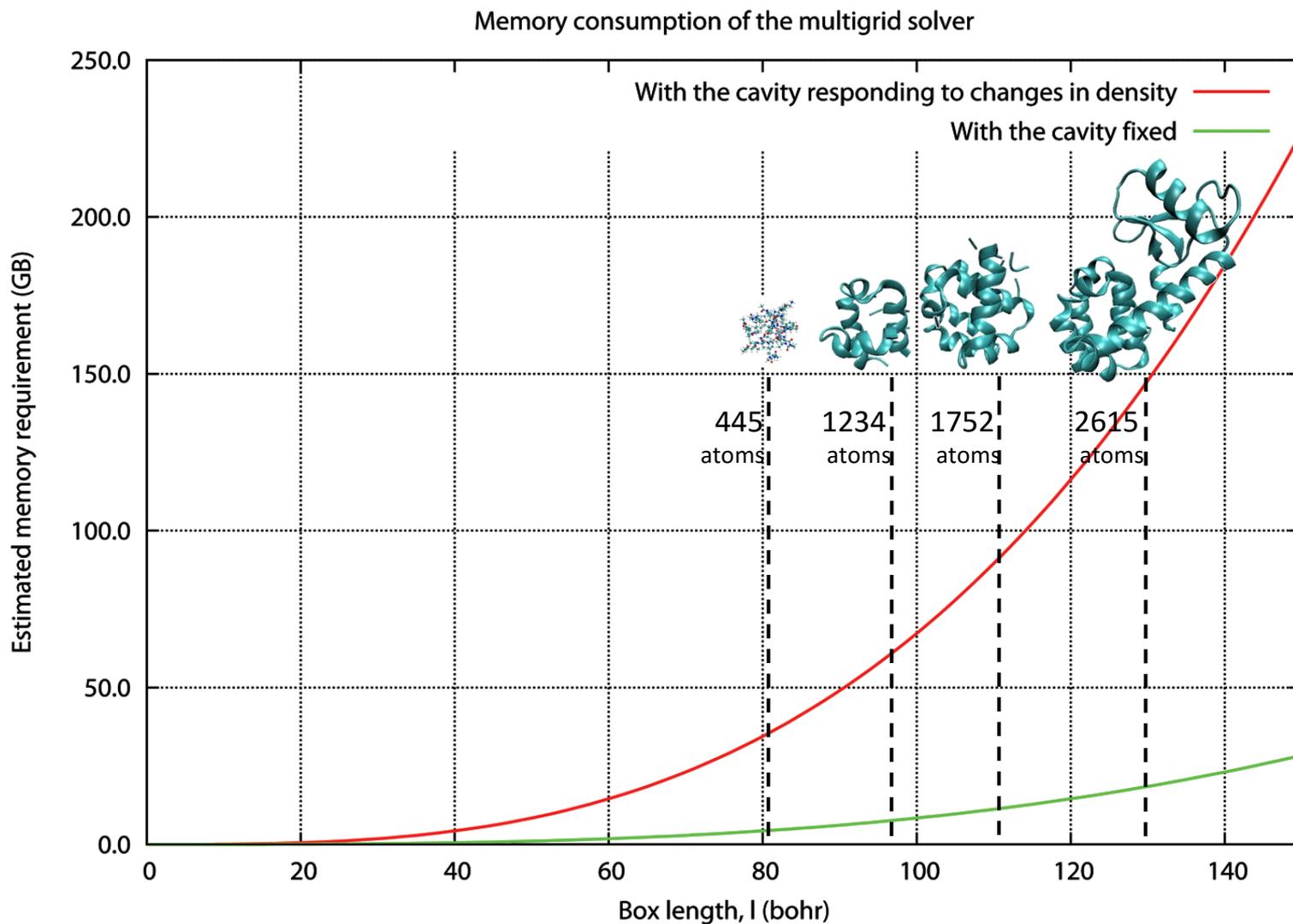
# Self-consistent updating of the cavity is very costly

- When the dielectric cavity reacts to changes in density, extra terms appear in the energy gradient:

$$\frac{\delta E_{es}}{\delta n}(\mathbf{r}) = \phi(\mathbf{r}) - \frac{1}{8\pi}(\nabla\phi(\mathbf{r}))^2 \frac{\delta\varepsilon}{\delta n}(\mathbf{r}) + \frac{\Delta G_{cav}}{\delta n}(\mathbf{r})$$

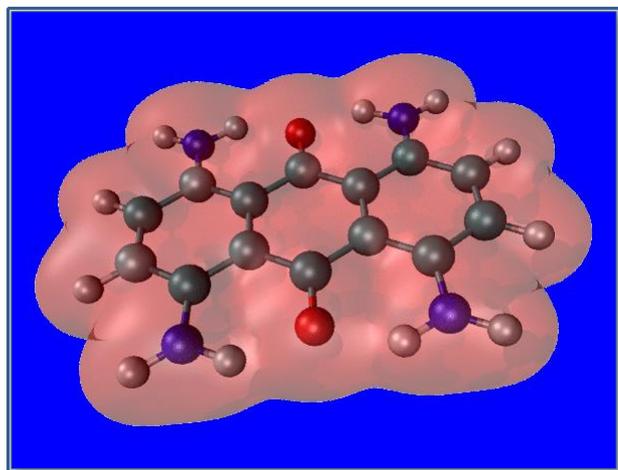


# Self-consistent updating of the cavity is very costly



# Using a fixed cavity is a decent approximation

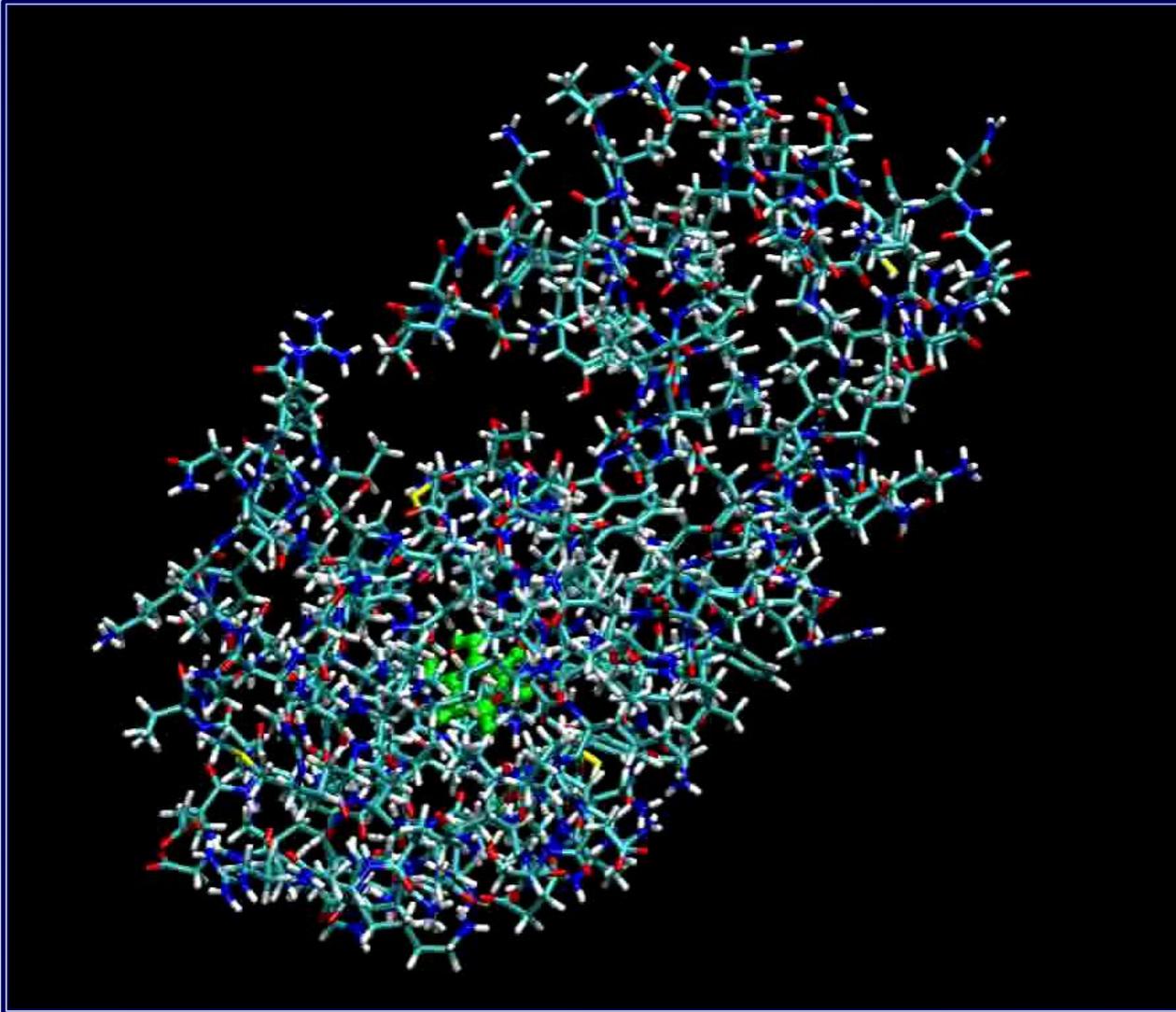
Approach	XC functional	rms error	max error	$r$
this work <sup>a</sup>	PBE	3.8	8.3	0.83
this work <sup>b</sup>	PBE	4.1	9.1	0.83
PCM	PBE	10.9	23.3	0.53
SMD	M05-2X	3.4	14.5	0.87
AMBER	(classical)	5.1	19.9	0.77



NB how the density hardly changes at all in solvent.

# L99A/M102Q mutant of T4 lysozyme

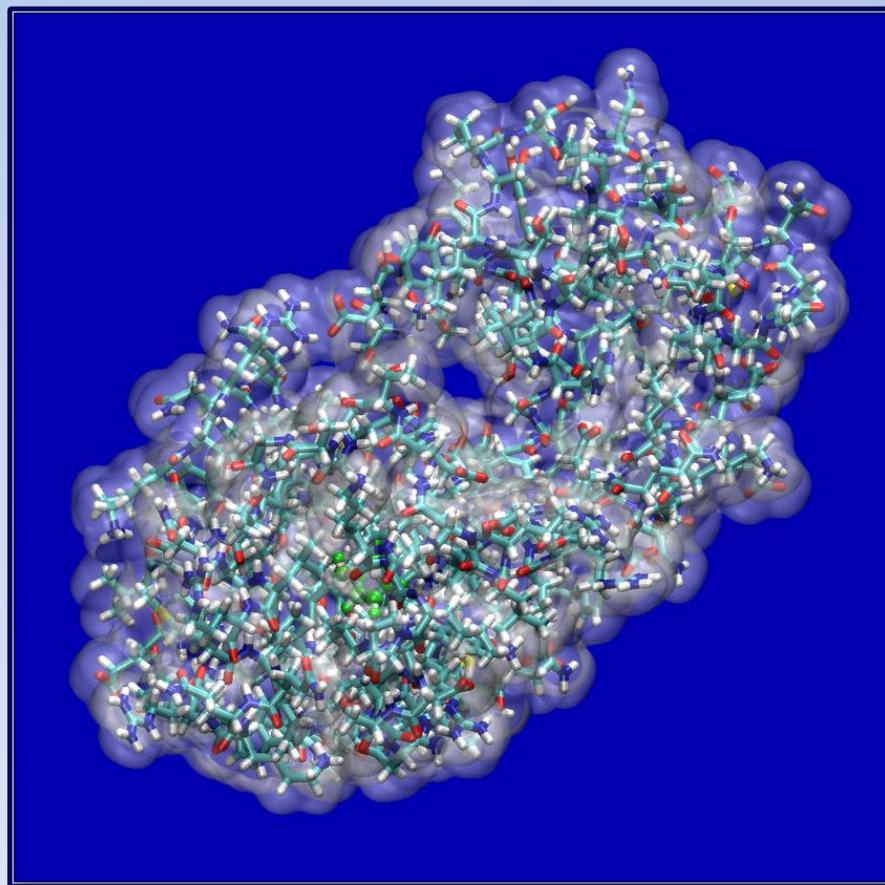
Calculation performed by S. Fox



- 2601 atoms.
- Here shown bound with a catechol ligand.
- A long (10 ns) classical simulation is first performed and a series of snapshots is taken.
- QM free energies of binding, in solution, are obtained with ONETEP for every snapshot.

# L99A/M102Q mutant of T4 lysozyme

Approach	$\Delta G_{\text{cplx}}$	$\Delta G_{\text{host}}$	$\Delta G_{\text{lig}}$	$\Delta G_{\text{d}}$	$\Delta E_{\text{gas}}$	$\Delta E_{\text{sol}}$
this work	-2423.0	-2421.3	-7.5	5.8	-28.6	-22.8
AMBER	-2428.3	-2433.0	-17.6	22.4	-27.7	-5.3
expt.	–	–	-9.3	–	–	–



# Conclusions

- Building on the isodensity model of Fattebert and Gygi, we have developed a solvation model which:
  - is *ab initio* based,
  - has predictive power that is
    - superior to classical models,
    - superior to PCM,
    - comparable to the state-of-the-art SMD model,
  - has only two parameters ( $\beta$ ,  $\rho_0$ ), whose values have been optimized and which appear to be universal (do not depend on the solute),
  - uses only two fundamental quantities ( $\epsilon$ ,  $\gamma$ ), with clear physical interpretation, to describe the solvent.

# Conclusions

- We observe that the following must be taken care of to obtain a successful QM-based solvation model:
  - high-order defect correction in the solution of NPE,
  - inclusion of dispersion-repulsion, at least crudely,
  - careful implementation of open boundary conditions.
- With our current implementation (LS-DFT ONETEP) we are able to calculate accurate free energies of solvation at the DFT level for molecules...
  - of about 500 atoms on a desktop PC workstation.
  - of about 3000 atoms on a parallel machine (40 cores or so).... within one day.

# Recommended reading

- J. Dziedzic, H. H. Helal, C.-K. Skylaris, A. A. Mostofi, and M. C. Payne, *Minimal parameter implicit solvent model for ab initio electronic structure calculations*, Europhysics Letters 95, 43001 (2011).
- J. Dziedzic, S. J. Fox, T. Fox, C. S. Tautermann, and C.-K. Skylaris, *Large-scale DFT calculations in implicit solvent - a case study on the T4 lysozyme L99A/M102Q protein*, Int. J. Quantum Chem 113, 771 (2013).
- S. Fox, J. Dziedzic, T. Fox, C.S. Tautermann, and C.-K. Skylaris, *DFT calculations on entire proteins for free energies of binding: application to a model polar binding site*, Proteins: Structure, Function and Bioinformatics 82, iss. 12 (2014).

# Recommended reading



# ONETEP

Print | [Login](#)

## Documentation for specific functionality

For general help on standard total energy and force calculations, see the [tutorials](#) and the [list of input variables](#).

[Advice on environment variables and other settings to consider when launching ONETEP](#)

Below are various documents outlining the use of specific elements of functionality:

### Ground State Calculation Setup

[Pseudoatomic Solver](#)

[Conduction NGWF Optimisation](#)

[Finite-temperature Ensemble-DFT calculations](#)

[Density kernel and Hamiltonian mixing \(kernel\\_diis\)](#)

[Empirical dispersion corrections](#)

[Van der Waals Density Functional Calculations](#)

[Realspace Local Pseudopotentials](#)

[Implicit Solvation \(older versions\) | Implicit Solvation \(newer versions\)](#)

[Implicit Solvation -- explanation of changes | Implicit Solvation examples](#)

[Energy Decomposition Analysis \(EDA\)](#)

# Recommended reading



# ONETEP

Print | [Login](#)

## Documentation for specific functionality

For general help on standard total energy and force calculations, see the [tutorials](#) and the [list of input variables](#).

[Advice on environment variables and other settings](#)

Below are various documents outlining the use of

### Ground State Calculation Setup

[Pseudoatomic Solver](#)

[Conduction NGWF Optimisation](#)

[Finite-temperature Ensemble-DFT calculations](#)

[Density kernel and Hamiltonian mixing \(kernel\\_c\)](#)

[Empirical dispersion corrections](#)

[Van der Waals Density Functional Calculations](#)

[Realspace Local Pseudopotentials](#)

[Implicit Solvation \(older versions\)](#) | [Implicit Solvation \(newer versions\)](#)

[Implicit Solvation -- explanation of changes](#)

[Energy Decomposition Analysis \(EDA\)](#)

Name	Size	Type
1_ammonia_in_water	451.5 kB	Folder
2_acetonitrile_in_water	363.8 kB	Folder
3_DMSO_anion_in_water	792.4 kB	Folder
4_aniline_cation_in_acetonitrile	479.2 kB	Folder
5_methane_in_water	480.5 kB	Folder
6_methane_in_water_FGS_model	505.8 kB	Folder
7_methane_in_water_FG_model	557.9 kB	Folder
8_azinphos_methyl_in_water_high_beta_param	692.6 kB	Folder

[Implicit Solvation examples](#)

# In practice

```
! --- Turns on solvation, model with fixed cavity ---
is_auto_solvation T
is_smeared_ion_rep T
is_dielectric_model fix_initial

! --- Physical characteristics of the solvent: ---
! --- Dielectric permittivity, surface tension. ---
is_bulk_permittivity 78.54
is_solvent_surf_tension 0.07415 N/m

! --- Include apolar terms to solvation: ---
! --- * Cavitation ---
! --- * Solute-solvent dispersion-repulsion ---
! --- is included approximately through ---
! --- rescaling of the cavitation term ---
is_include_apolar T
is_apolar_scaling_factor 0.281075

! --- Parameters of the model. These mirror the defaults ---
! --- and are expected to be transferable across solvents, ---
! --- XC functionals and pseudopotentials. I suggest these ---
! --- be included in inputs in case defaults change in the ---
! --- future. ---
is_solvation_beta 1.3
is_density_threshold 0.00035
```

# In practice

```
=====
IS: Calculation in solvent completed (NGWFs converged).
IS: Total energy in vacuum:                -11.67205886619922 Ha
IS: Total energy in solvent:              -11.67847977708559 Ha
IS: Apolar cavitation energy:              0.01364956829068 Ha
IS: Apolar solute-solvent disp-repulsion energy: -0.00981301588337 Ha
IS: Total apolar term to solvation:        0.00383655240730 Ha
IS: Total polar term to solvation:        -0.01025746329367 Ha
IS: Total solvation energy:               -0.00642091088637 Ha
=====
```

# Recent developments

- **Solvation forces** (geometry optimisation in solvent, MD in solvent) – exact only with the adaptive cavity and somewhat approximate with fixed cavity. [JD]
- **Non-zero ionic concentrations** („salty water”) – machinery in place, but model needs refinement to be viable. [JD, LA, CKS]
- **Defect-correction as part of DL\_MG.** [JCW, LA]
- **Periodic and mixed BCs in solvent** – soon. [JCW, LA, CKS]

# References

1. Fattebert and Gygi, *J. Comp. Chem.* **23** (2002).
2. Foresman, Keith, Wiberg, Snoonian and Frisch, *J. Phys. Chem.* **100** (1996).
3. Moreira, Fernandes and Ramos, *J. Mol. Struct.: THEOCHEM* **729** (2005).
4. Feig and Brooks III, *Curr. Opinion in Struct. Biol.* **14** (2004).
5. Chen, Brooks III and Khandogin, *Curr. Opinion in Struct. Biol.* **18** (2008).
6. Dong, Wagoner and Baker, *Phys. Chem. Chem. Phys.* **10** (2008).
7. Baker, *Curr. Opinion in Struct. Biol.* **15** (2005).
8. Feig, Onufriev, Lee, Wonpil, Case and Brooks III, *J. Comput. Chem* **25** (2004).
9. Miertus, Scrocco and Tomasi, *Chem. Phys.* **55** (1981).
10. Fattebert and Gygi, *International Journal of Quantum Chemistry* **93** (2003).
11. Scherlis, Fattebert, Gygi, Cococcioni and Marzari, *J. Chem. Phys.* **124** (2006).
12. Holst, "Multilevel Methods for the Poisson-Boltzmann Equation", PhD thesis (1993).
13. Trottenberg, Oosterles and Schuller, "Multigrid" (2001).
14. Schaffer, *Math. Comp.* **43** (1984).
15. Clark, Segall, Pickard, Hasnip, Probert, Refson and Payne, *Zeitschrift für Kristallographie* **220** (2005).
16. Skylaris, Haynes, Mostofi and Payne, *J. Chem. Phys.* **122** (2005).
17. Floris, Tomasi and Ahuir, *J. Comp. Chem.* **12** (1991).
18. Marenich, Kelly, Thompson, Hawkins, Chambers, Giesen, Winget, Cramer, and Truhlar, *Minnesota Solvation Database – version 2009*, University of Minnesota, Minneapolis, 2009.
19. Marenich, Cramer, and Truhlar, *J. Phys. Chem. B* **113** (2009).
20. Cramer and Truhlar in "Free Energy Calculations in Rational Drug Design" (2001).
21. Shestakov, Milovich and Noy, *Journal of Colloid and Interface Science* **247** (2002).
22. Lu, Zhang and McCammon, *J. Chem. Phys.* **122** (2005).
23. Fogolari, Brigo and Molinari, *Journal of Molecular Recognition* **15** (2002).
24. Tomasi and Persico, *Chem. Rev.* **94** (1994).
25. Swanson, Henchman and McCammon, *Biophysical Journal* **86** (2004).
26. Rocchia, Sridharan, Nicholls, Alexov, Chiaberra and Honig, *J. Comput. Chem.* **23** (2002).
27. Tan, Yang and Luo, *J. Phys. Chem. B* **110** (2006).
28. Tomasi, Mennucci and Cammi, *Chem. Rev.* **105** (2005).
29. Sanchez, Sued and Scherlis, *J. Chem. Phys.* **131** (2009).
30. Kaukonen, Soederhlejm, Heimdal and Ryde, *J. Phys. Chem. B* **112** (2008).
31. Baker, Sept, Joseph, Holst and McCammon, *Proc. Natl. Acad. Sci. USA* **98** (2001).