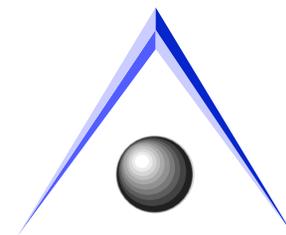
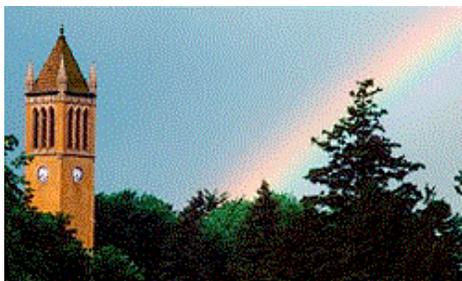


# CLUSTER-BASED APPROACHES TO SOLVATION

Mark S. Gordon  
Iowa State University  
Ames Laboratory



# OUTLINE

- Solvation Methods
  - Explicit vs. implicit methods
- Explicit Methods
  - TIP3P, TIP4P
  - SPC, SPC/E
- EFP Method for Solvation
  - Summary of EFP1 method for water
  - Sample input files
  - Monte Carlo method
  - Example applications
- Generalized EFP Method (EFP2)

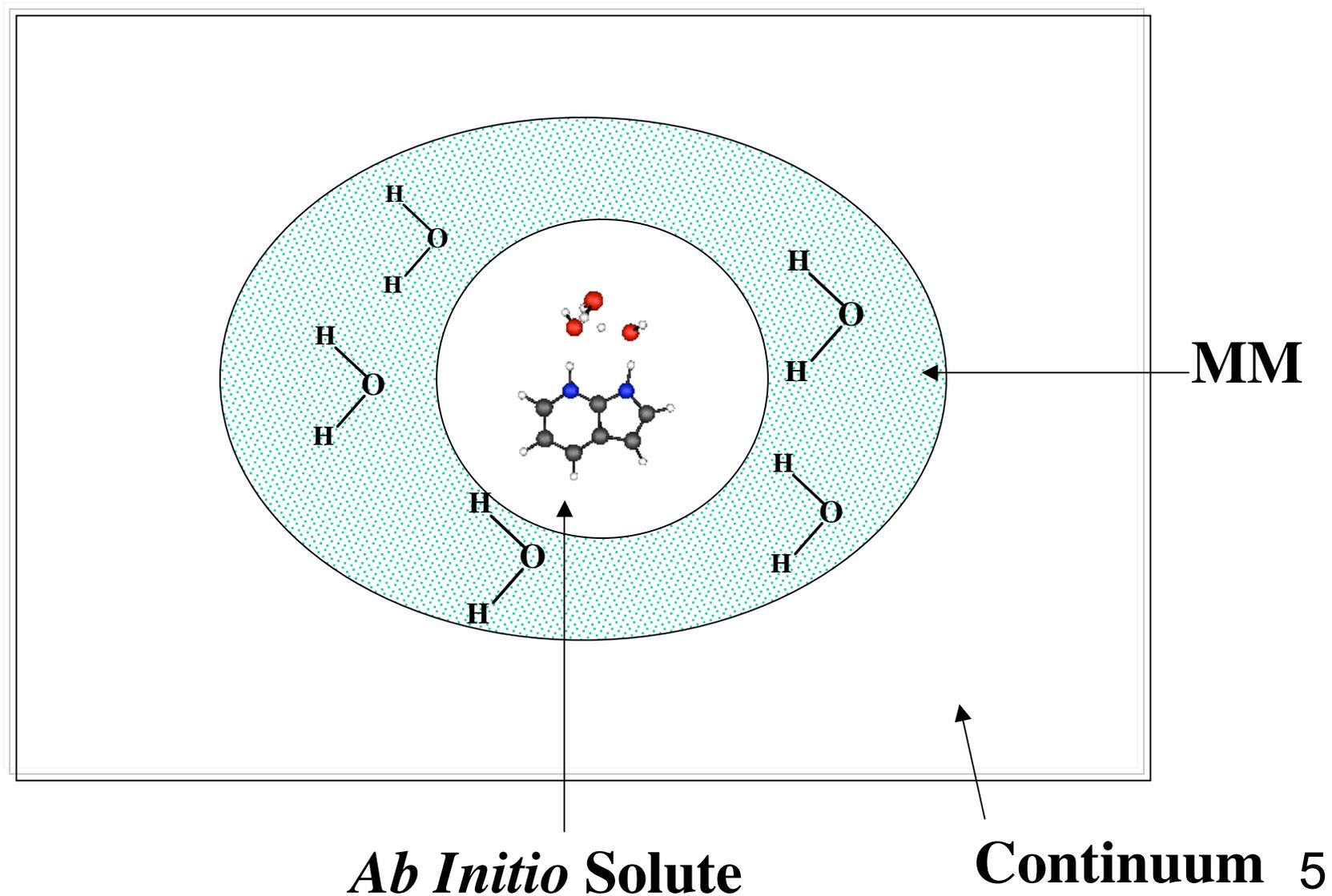
# SOLVATION MODELS: CONTINUUM

- Based on Onsager reaction field model
- Computationally efficient
- Sensitive to cavity size and shape
- Do not account for explicit solute-solvent interactions

# SOLVATION MODELS: DISCRETE

- Solvent molecules described using potentials
  - Empirical potentials: lots of fitted parameters
  - *Ab initio* potentials: computationally expensive
- Include explicit solute-solvent interactions
- Configurational sampling necessary: computationally demanding

# Multi-Layered Approach to Solvation



# DISCRETE/EXPLICIT MODELS

- TIP3P, TIP4P

- Jorgensen *etal*, JCP, 79, 926 (1983)

- Basically simple Lennard-Jones model

$$\epsilon_{mn} = \sum_i^m \sum_j^n [q_i q_j / r_{ij}] + A/R^{12} - C/R^6$$

- $\epsilon_{mn}$  = interaction energy between waters m,n

- $q_O < 0 = -2q_H$ , A, C fitted to bulk properties

- Rigid molecules

- TIP4P adds a 4th center inside O toward Hs

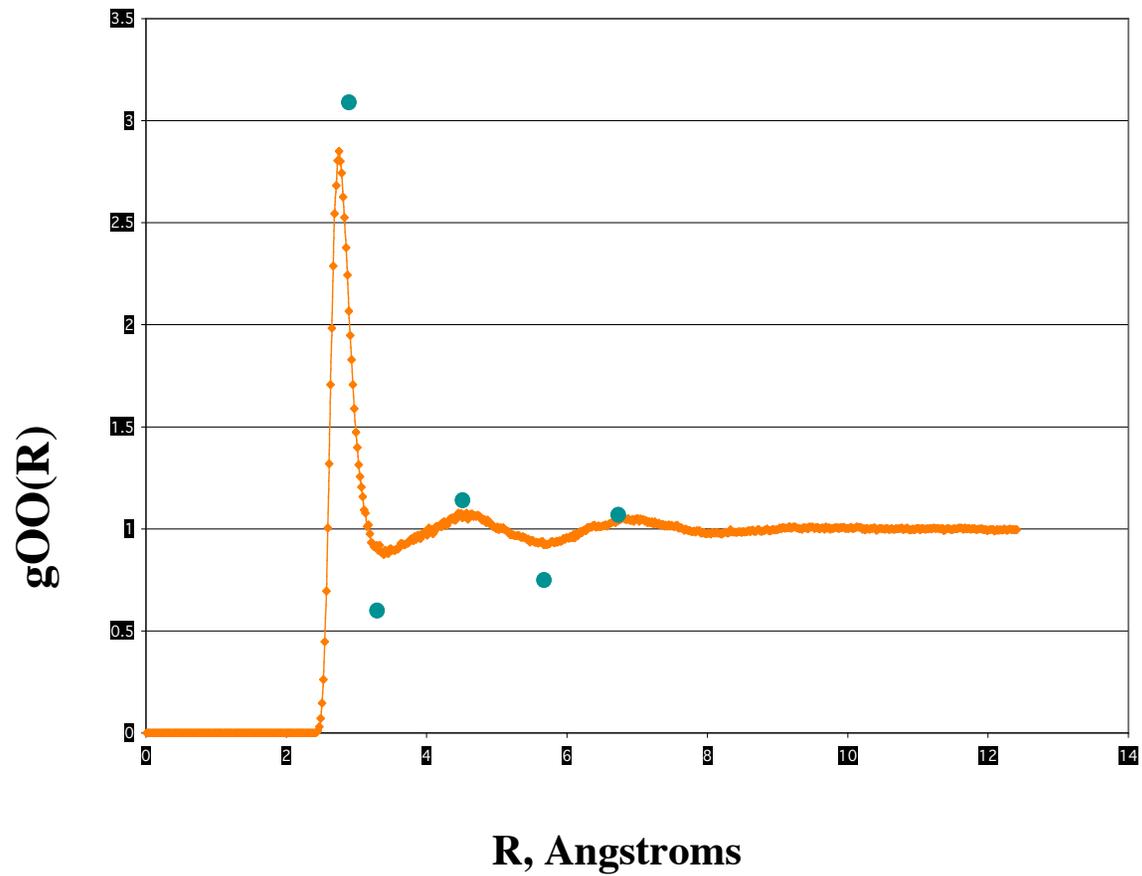
- Very popular model; doesn't get OO radial distribution function right: misses 2nd peak

# DISCRETE/EXPLICIT MODELS

- SPC (Simple Point Charge), SPC/E
  - Berendsen *etal*, JPC, 91, 6269 (1987)
  - SPC:

$$\epsilon_{mn} = \sum_i^m \sum_j^n [q_i q_j / r_{ij}] + A/R^{12} - C/R^6$$

- SPC same as TIP, but fitted to MD simulations of density and vaporization energy
- Rigid molecules
- Gets OO radial distribution function right
- SPC/E adds polarizability/induction term



Curve: SPC/E  
Points: exptl

# General Effective Fragment Potential

§ **Discrete** solvation method

§ Fragment potential is one electron contribution to the *ab initio* Hamiltonian

§ Potentials

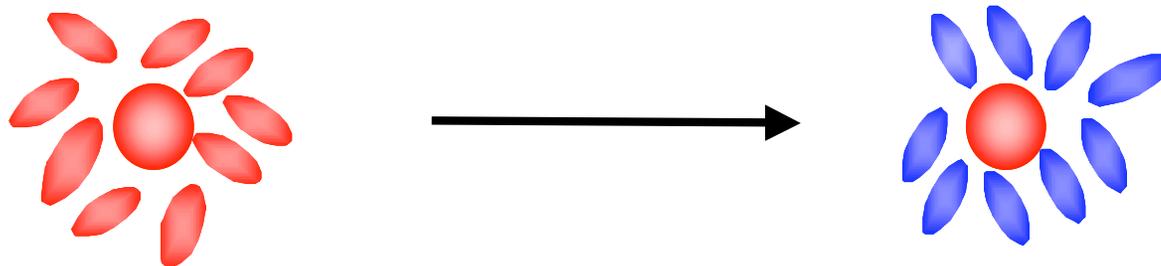
are obtained by separate *ab initio* calculations  
depend on properties of **isolated** molecules  
can be systematically improved

# Effective Fragment Potential

System is divided into

an *ab initio* region for the “solute” and  
a *fragment* region for the solvent molecules.

$$E = E_{ab\ initio} + E_{interaction}$$



## General EFP Method

In the most general implementation EFP should include all relevant energy contributions:

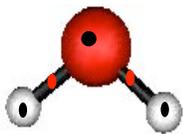
$$\begin{aligned} E_{\text{interaction}} &= E_{\text{coulomb}} + \\ &E_{\text{polarization}} + \\ &E_{\text{exch. rep. / charge transfer}} + \\ &E_{\text{dispersion}} + \\ &E_{\text{higher order terms}} \end{aligned}$$

# Hartree Fock based EFP

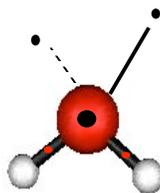
Interaction energy consists of : *electrostatic, polarization and exchange repulsion/charge transfer term*

$$E_{\text{interaction}} = E_{\text{coulomb}} + E_{\text{polarization}} + E_{\text{exchange repulsion/charge transfer}}$$

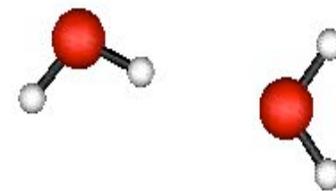
$$E_{\text{interaction}} = \sum_{k=1}^K V_k^{\text{Elec}}(\mu, s) + \sum_{l=1}^L V_l^{\text{Pol}}(\mu, s) + \sum_{m=1}^M V_m^{\text{Rep}}(\mu, s)$$



Distributed  
Multipolar expansion



LMO polarizability  
expansion



Fit to Functional  
Form

# Coulomb/Electrostatic Term

$$V_k^{Elec}(\mu, s) = \frac{q_k(\mu)q_s}{r_{sk}} - \sum_a^{x,y,z} \mu_a^k(\mu)F_a(\mathbf{r}_{sk}) - \frac{1}{3} \sum_{a,b}^{x,y,z} \Theta_{ab}^k(\mu)F_{ab}(\mathbf{r}_{sk}) - \frac{1}{15} \sum_{a,b,c}^{x,y,z} \Omega_{abc}^k(\mu)F_{abc}(\mathbf{r}_{sk})$$

where  $q$ ,  $\mu$ ,  $\Theta$  and  $\Omega$  are the charge, dipole, quadrupole and octopole, respectively

$F_a$ ,  $F_{ab}$  and  $F_{abc}$  are the solute electric field, field gradient and field hessian

This DMA is a classical point multipole approach: what if fragments approach each other or solute too closely?:

Classical treatment is then too repulsive: multiply entire expression by damping term to account for overlapping charge densities

$$V_k^{Elec}(\mu, s) \rightarrow (1 - \beta_k(\mu)e^{-\alpha_k(\mu)r_{sk}^2})V_k^{Elec}(\mu, s)$$

All calculations done once for a given solvent molecule; e.g., water

# Polarization/Induction Term

$$V_l^{Pol}(\mu) = - \sum_{a,b}^{x,y,z} F_a(\mathbf{r}_l) \alpha_{ab}^l(\mu) \langle F_b(\mathbf{r}_l) \rangle$$

$$\alpha_{xy}^l = \lim_{F_y \rightarrow 0} \frac{\mu_x^l(F_y) - \mu_x^l(0)}{F_y}$$

Polarizability is treated as a distributed tensor with each tensor located at the centroid of a localized orbital

For water, there are five such LMOs, two O-H bond orbitals, two lone pairs, one inner shell.

F is the field due to the *ab initio* part of the system

$\alpha_{xy}^l$  is a component of the polarizability of the fragment molecule in the *l*th localized orbital

Iterated to self-consistency

All calculations done once for a given solvent molecule; e.g., water

# Exchange Repulsion/Charge Transfer Terms

Fragment-*ab initio* interaction:

$$V_m^{\text{Re } p}(\mu, s) = \sum_j^J \beta_{m,j}(\mu) e^{-\alpha_{m,j}(\mu) r_{m,s}^2}$$

Gaussian functions centered at atom centers and center of mass: J=2, M=4 for water.

Fitted to a set of dimer calculations: several values of R(O-O) for each orientation. Total of 192 points on dimer surface. Fitted to minimize:

$$\Delta = \sum_p^P w_p \left[ \left\langle \Psi \left| \sum_m^M V_m^{\text{Re } p} \right| \Psi \right\rangle_p - E_{\text{rem}}^{(ab)}(p) \right]^2$$

$w_p$  is a weighting factor, usually = 1.0;  $\Psi$  = *ab initio* wavefunction

# Exchange Repulsion/Charge Transfer Terms

Fragment-fragment interaction: similar, except a single exponential function is fitted to the 192 points.

Analytic gradients (first derivatives) have been derived and coded for all terms. So, one can perform geometry optimizations, transition state searches, dynamics

# HIGHER-LEVELS OF EFP

- DFT-based EFP (Ivana Adamovic)
  - Same general approach
  - Based on B3LYP
  - Adds some level of correlation
- MP2-based EFP (Jie Song)
  - Same general approach
  - Separate fit for exchange repulsion, remainder
  - More effective correlation, especially at long range

# Water hexamer isomers

Binding Energy	DFT B3LYP	MP2	<b>CCSD (T)</b>	HF
prism	62.37	58.25	<b>55.10</b>	42.86
cage	61.84	57.52	<b>54.30</b>	42.49
book	61.34	56.49	<b>53.10</b>	42.44
cyclic	60.57	55.75	<b>52.20</b>	43.10
boat	59.13	54.29	<b>50.80</b>	42.12

- basis set: DH(d,p)

-units: kcal/mol

# Water hexamer isomers

Binding Energy	DFT B3LYP	EFP1/DFT	MP2	<b>CCSD (T)</b>	HF	EFP1/ HF
prism	62.37	61.08	58.25	<b>55.10</b>	42.86	42.42
cage	61.84	61.53	57.52	<b>54.30</b>	42.49	41.90
book	61.34	61.79	56.49	<b>53.10</b>	42.44	41.45
cyclic	60.57	60.65	55.75	<b>52.20</b>	43.10	41.14
boat	59.13	59.37	54.29	<b>50.80</b>	42.12	40.09

- basis set: DH(d,p)

-units: kcal/mol

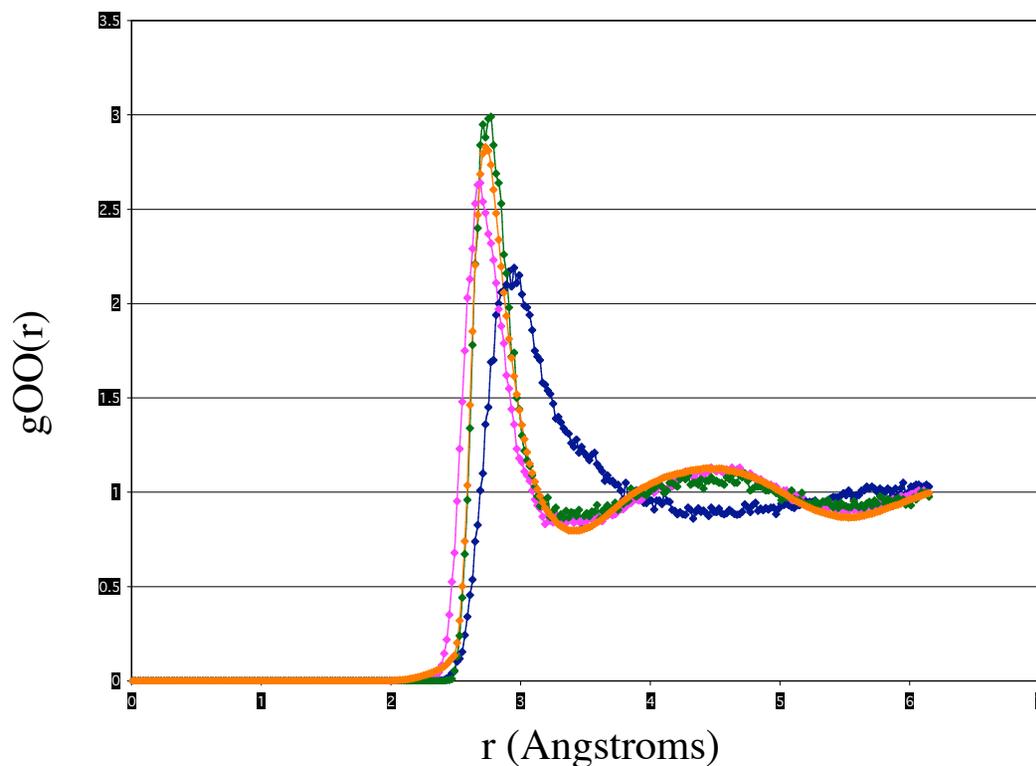
# Water hexamer isomers

Binding Energy	DFT B3LYP	EFP1/DFT	MP2	EFP1/MP2	<b>CCSD (T)</b>	HF	EFP1/HF
prism	62.37	61.08	58.25	58.26	<b>55.10</b>	42.86	42.42
cage	61.84	61.53	57.52	57.64	<b>54.30</b>	42.49	41.90
book	61.34	61.79	56.49	56.67	<b>53.10</b>	42.44	41.45
cyclic	60.57	60.65	55.75	55.24	<b>52.20</b>	43.10	41.14
boat	59.13	59.37	54.29	53.92	<b>50.80</b>	42.12	40.09

- basis set: DH(d,p)

-units: kcal/mol

# $g_{OO}(r)$ : EFP1/HF, EFP1/DFT, SPC/E 62 waters

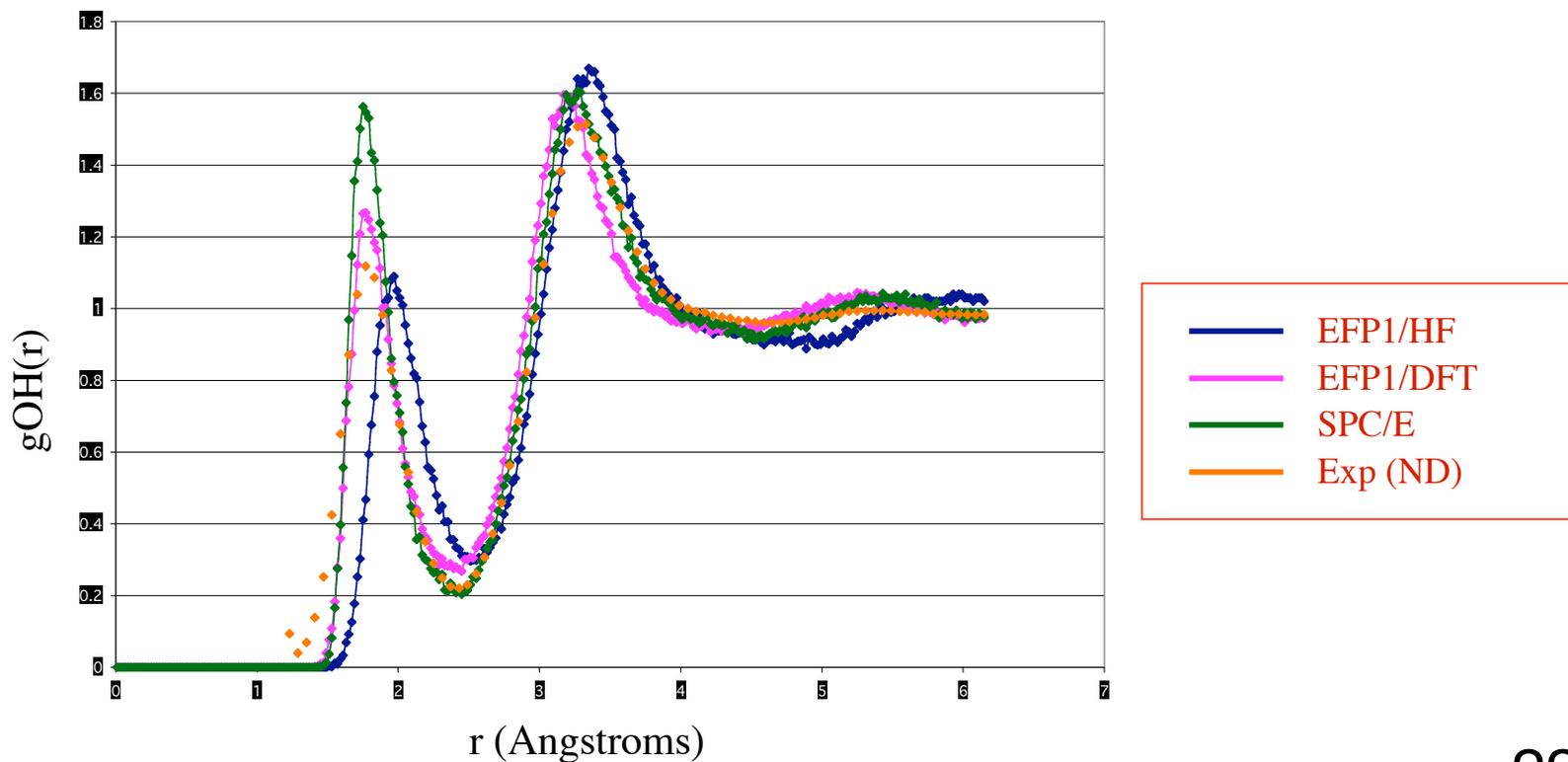


Initial structure: 62 waters,  
26 ps equilibration

Timestep size = 1 fs,  
Simulation = 5000 fs

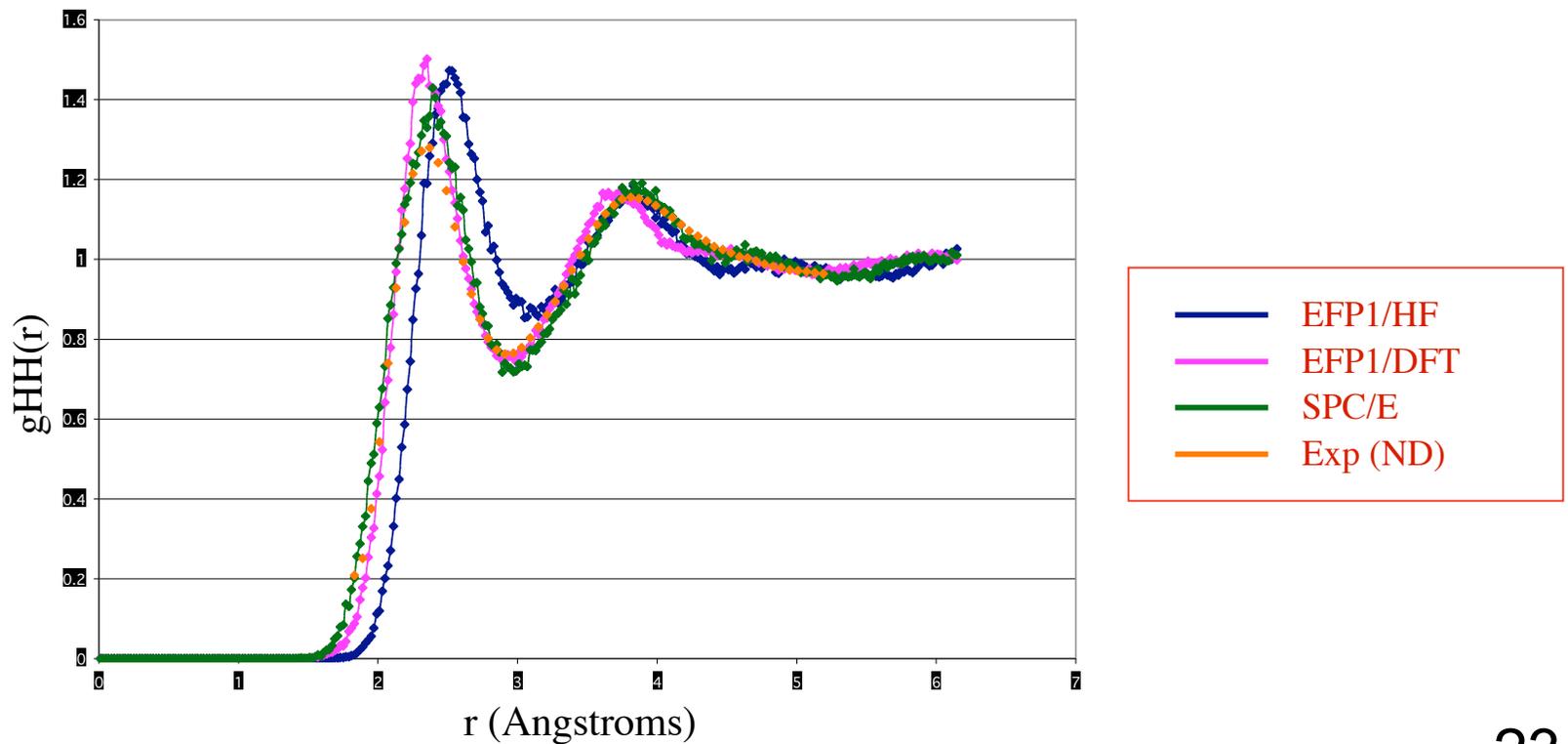
— EFP1/HF  
— EFP1/DFT  
— SPC/E  
— Exp (THG)

# gOH(r): EFP1/HF, EFP1/DFT, SPC/E 62 waters



Exp (ND): Neutron Diffraction; Soper et. al.

# $g_{HH}(r)$ : EFP1/HF, EFP1/DFT, SPC/E 62 waters



# GENERAL EFP METHOD

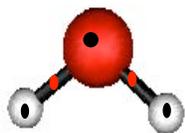
- Jan Jensen
- Mark Freitag
- Ivana Adamovic
- Mol. Phys., 89, 1313 (1996)
- JCP, 108, 4772 (1999)
- JCP, 112, 7300 (2000)

# Hartree Fock based EFP

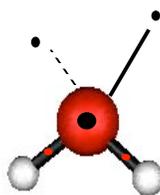
Interaction energy consists of : *electrostatic, polarization and exchange repulsion/charge transfer term*

$$E_{\text{interaction}} = E_{\text{coulomb}} + E_{\text{polarization}} + E_{\text{exchange repulsion/charge transfer}}$$

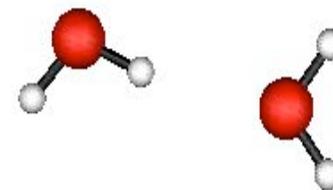
$$E_{\text{interaction}} = \sum_{k=1}^K V_k^{\text{Elec}}(\mu, s) + \sum_{l=1}^L V_l^{\text{Pol}}(\mu, s) + \sum_{m=1}^M V_m^{\text{Rep}}(\mu, s)$$



Distributed  
Multipolar expansion



LMO polarizability  
expansion

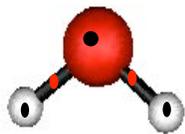


Fit to Functional  
Form

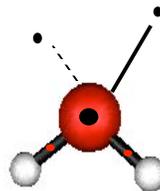
# Generalized EFP2 Method

Interaction energy consists of : *electrostatic*, *polarization* and *exchange repulsion term*

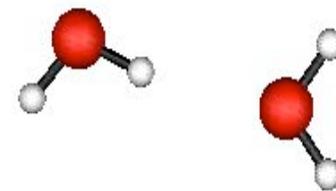
$$E_{\text{interaction}} = E_{\text{electrostatic}} + E_{\text{polarization}} + E_{\text{exchange repulsion}}$$



Distributed  
Multipolar expansion



LMO polarizability  
expansion



From first principles  
using LMO overlaps

## Summary

If all approximations presented above are combined,  $E_{exch}$  can be approximated as (provided LMOs are used)

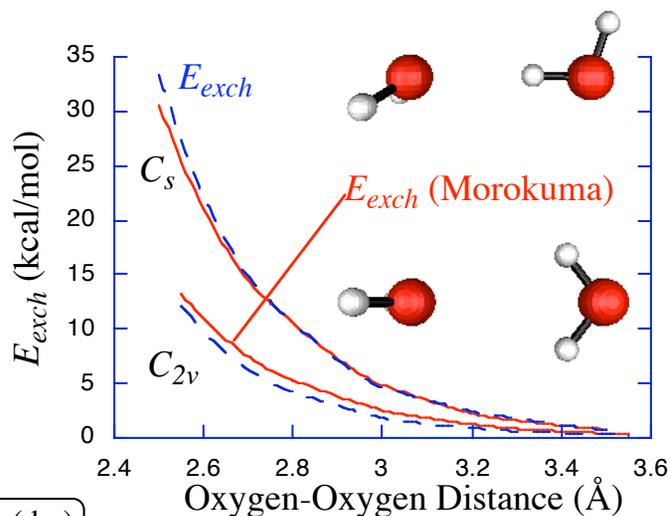
$$E_{exch} \approx -2 \sum_{i \in A} \sum_{j \in B} 2 \sqrt{\frac{-2 \ln S_{ij}}{\pi}} \frac{S_{ij}^2}{R_{ij}} - 2 \sum_{i \in A} \sum_{j \in B} S_{ij} \left[ \sum_{k \in A} F_{ik}^A S_{kj} + \sum_{l \in B} F_{jl}^B S_{li} - 2T_{ij} \right]$$

$$+ 2 \sum_{i \in A} \sum_{j \in B} S_{ij}^2 \left[ \sum_{j \in B} -Z_J R_{iJ}^{-1} + 2 \sum_{l \in B} R_{il}^{-1} + \sum_{l \in B} -Z_l R_{lj}^{-1} + 2 \sum_{k \in A} R_{kj}^{-1} - R_{ij}^{-1} \right]$$

This equation requires only the computation of intermolecular overlap and electronic kinetic energy integrals, i.e. no two-electron integrals other than those calculated once for the isolated molecules.

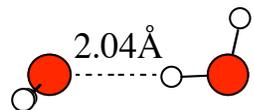
*It contains no adjustable parameters*, only fixed parameters computed for the isolated molecules, such as the LMOs in some AO basis, Fock matrices in the LMO bases, and the LMO centroids of charge

## Numerical Tests



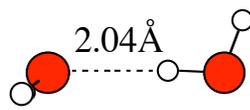
RHF/6-31++(d,p)

**EFP(nc)[6-31G(d,p)]**

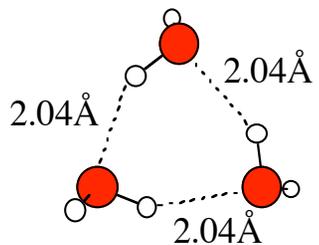


$E_{\text{Binding}}=5.6$  kcal/mol

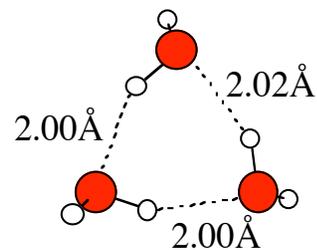
**RHF/6-31G(d,p)**



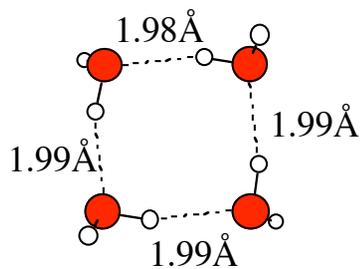
5.6 kcal/mol



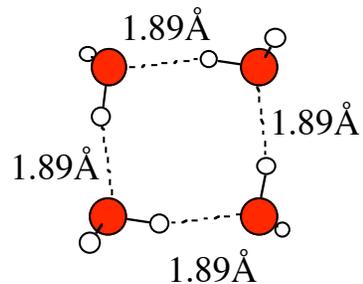
16.2 kcal/mol



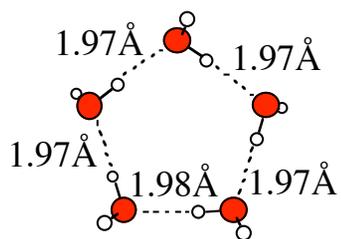
17.3 kcal/mol



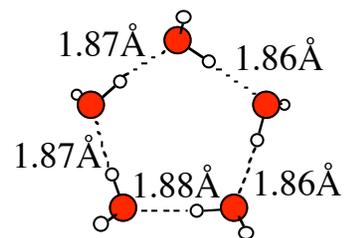
26.3 kcal/mol



29.1 kcal/mol



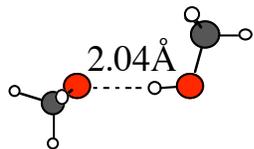
35.0 kcal/mol



37.7 kcal/mol

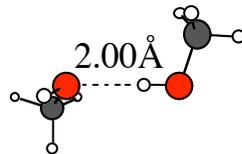
**Figure 3.** Structures and binding energies of  $(\text{H}_2\text{O})_n$ ,  $n=2-5$  at the EFP(nc)[6-31G(d,p)] and RHF/6-31G(d,p) level of theory.

**EFP[6-31G(d,p)]**

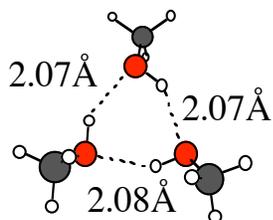


$E_{\text{Binding}}=5.2$  kcal/mol

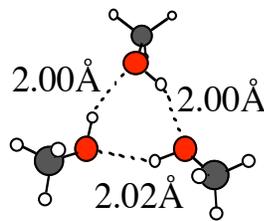
**RHF/6-31G(d,p)**



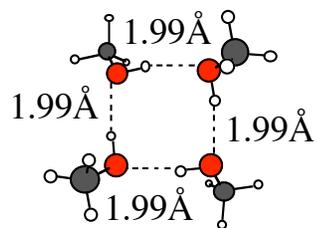
5.5 kcal/mol



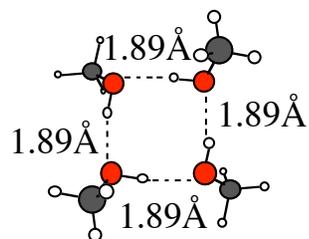
15.2 kcal/mol



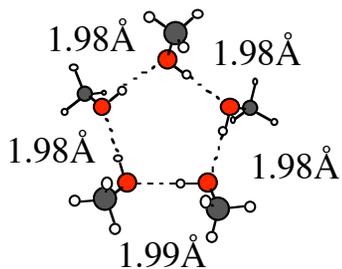
16.6 kcal/mol



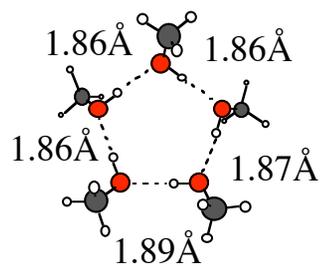
24.6 kcal/mol



27.9 kcal/mol



32.4 kcal/mol



36.5 kcal/mol

**Figure 4.** Structures and binding energies of  $(\text{MeOH})_n$ ,  $n=2-5$  at the EFP(nc)[6-31G(d,p)] and RHF/6-31G(d,p) level of theory.

# EFP Performance: Energy + Gradient Calculation

Method <sup>1</sup>	20 water molecules	62 water molecules	122 water molecules	512 water molecules
<i>Ab initio</i> <sup>2</sup>	3.19 hrs	---	---	~157 yrs <sup>3</sup>
EFP2	3.3 sec	26.1 sec	95.3 sec	26.8 min
EFP1/HF	0.2 sec	2.6 sec	5.1 sec	97.8 sec
SPC/E <sup>4</sup>	0.02 sec	0.02 sec	0.1 sec	0.7 sec

<sup>1</sup>Run on 1200 MHz Athlon/Linux machine

<sup>2</sup>*Ab initio*: DZP basis set, <sup>3</sup>Assuming N<sup>4</sup> scaling,

<sup>4</sup>SPC/E = Simple Extended Point Charge model

# CONCLUSIONS/FUTURE

- Developments in Progress
  - Analytic *ab initio*/EFP gradient for EFP2
  - Dispersion Term
    - EFP1: Fit to MP2, similar to EFP1/DFT
    - EFP2: From frequency-dependent polarizabilities
  - Generalized Charge Transfer (EFP2)
  - Molecular Dynamics Code (~done)
  - Flexible fragments

# CONTINUUM SOLVATION MODELS

- Simplest is Onsager self-consistent reaction field (SCRF) model (\$SCRF)
  - Solute (QM) dipole moment  $\mu$  polarizes “medium” (solvent) through dielectric  $\epsilon$
  - Newly polarized solvent alters solute dipole moment
  - Iterated to self-consistency
  - Accomplished by adding new term to QM Hamiltonian

# ONSAGER SCRF MODEL

- $V_{\sigma} = -r \cdot R$ 
  - $r$  = position vector
  - $R$  is proportional to molecular dipole moment
    - $R = g\mu$
    - $\mu$  = dipole vector
    - $g = 2(\epsilon - 1) / [(2\epsilon + 1)(a^3)]$
    - $a$  = radius of cavity
  - Usually assume spherical cavity
  - Radius can be pretty arbitrary
  - Adds simple  $\langle \chi_i | x | \chi_j \rangle$  integrals - very cheap

# PCM MODEL

- Polarizable Continuum Model (PCM)
  - Tomasi et al., Chem. Phys. Lett., 255, 327 (1996)
  - Van der Waals type surface cavity
  - Uses detailed knowledge of electrostatic potential
  - Cavity dispersion potential determined from surface area
  - Can include dispersion effects
  - Interfaced with EFP in GAMESS: \$PCM
  - See both input and reference sections in manual