Quantum Mechanics/ Molecular Mechanics Methods

Jan H. Jensen

Department of Chemistry

University of Iowa

QM: numerical solution of the Schrödinger equation

$$H\Psi = E\Psi \quad \Rightarrow \quad E = \langle \Psi | H | \Psi \rangle$$

$$\sum_{i}^{L_e} \left(-\tfrac{1}{2}\nabla_i^2 - \sum_{B}^{L_N} \frac{Z_B}{r_{iB}} + \sum_{i>j}^{L_e} \frac{1}{r_{ij}} \right) \Psi\left(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_{L_e}; \{\mathbf{R}_{L_N}\}\right) = E\Psi\left(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_{L_e}; \{\mathbf{R}_{L_N}\}\right)$$

MO approximation

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_L) \approx \hat{A}\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)....\phi_L(\mathbf{r}_L)$$
$$= -\Psi(\mathbf{r}_2, \mathbf{r}_1, .., \mathbf{r}_L)$$

LCAO approximation

$$\phi_i(\mathbf{r}) \approx \sum_{\mu}^{K} C_{\mu i}\chi_\mu(\mathbf{r})$$

Energy Expression in the MO approximation
Restricted Hartree-Fock (RHF)

$$E = \sum_i^L \left\langle \phi_i(\mathbf{r}) \middle| -\tfrac{1}{2}\nabla^2 \middle| \phi_i(\mathbf{r}) \right\rangle + \sum_i^L \left\langle \phi_i(\mathbf{r}) \middle| -\sum_B^{L_N} \frac{Z_B}{r_{iB}} \middle| \phi_i(\mathbf{r}) \right\rangle$$

$$+ \sum_i^L \sum_j^L \left\langle \phi_i(\mathbf{r}_1)\phi_i(\mathbf{r}_1) \middle| r_{12}^{-1} \middle| \phi_j(\mathbf{r}_2)\phi_j(\mathbf{r}_2) \right\rangle - \sum_i^L \sum_j^L \left\langle \phi_i(\mathbf{r}_1)\phi_j(\mathbf{r}_1) \middle| r_{12}^{-1} \middle| \phi_i(\mathbf{r}_2)\phi_j(\mathbf{r}_2) \right\rangle$$

Energy determined by variational principle

$$E \geq E^{exact} \quad \Rightarrow \quad \frac{\partial E}{\partial C_{\mu i}} = 0$$

Computational Expense

$$\chi_{l,m.n}(\mathbf{r})\chi_{l',m'.n'}(\mathbf{r}) = x^{l+l'}y^{m+m'}z^{n+n'}\tilde{K}\exp\left(-\gamma\left|\mathbf{r}-\mathbf{R}_P\right|^2\right)$$

$$\sum_i^L\sum_j^L\left\langle\phi_i(\mathbf{r}_1)\phi_i(\mathbf{r}_1)\left|r_{12}^{-1}\right|\phi_j(\mathbf{r}_2)\phi_j(\mathbf{r}_2)\right\rangle = \sum_\mu^K\sum_\nu^K\sum_\lambda^K\sum_\sigma^K\sum_i^L\sum_j^L C_{\mu i}C_{\nu i}C_{\lambda j}C_{\sigma j}\left\langle\chi_\mu(\mathbf{r}_1)\chi_\nu(\mathbf{r}_1)\left|r_{12}^{-1}\right|\chi_\lambda(\mathbf{r}_2)\chi_\sigma(\mathbf{r}_2)\right\rangle$$

$$= \sum_\mu^K\sum_\nu^K\sum_\lambda^K\sum_\sigma^K P_{\mu\nu}P_{\lambda\sigma}\left\langle\chi_\mu(\mathbf{r}_1)\chi_\nu(\mathbf{r}_1)\left|r_{12}^{-1}\right|\chi_\lambda(\mathbf{r}_2)\chi_\sigma(\mathbf{r}_2)\right\rangle$$

K^4 integrals! 30-50 atoms routine

$$\langle\mu\nu|\lambda\sigma\rangle \le \sqrt{\langle\mu\nu|\mu\nu\rangle\langle\lambda\sigma|\lambda\sigma\rangle}$$ $K^{2.5\text{-}4}$integrals!

Generally applicable

Chemical Accuracy: Beyond the MO approximation

$$\Psi(\mathbf{r}_1,\mathbf{r}_2,..,\mathbf{r}_L) \approx \hat{A}\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)....\phi_L(\mathbf{r}_L) \equiv \Phi(\mathbf{r}_1,\mathbf{r}_2,..,\mathbf{r}_L)$$

$$\Psi(\mathbf{r}_1,\mathbf{r}_2,..,\mathbf{r}_L) \approx \sum_I^K C_I \Phi_I(\mathbf{r}_1,\mathbf{r}_2,..,\mathbf{r}_L)$$

$$\sum_i^K\sum_j^K\sum_k^K\sum_l^K \Gamma_{ijkl}\left\langle \phi_i(\mathbf{r}_1)\phi_j(\mathbf{r}_1)\left|r_{12}^{-1}\right|\phi_k(\mathbf{r}_2)\phi_l(\mathbf{r}_2)\right\rangle = \sum_\mu^K\sum_\nu^K\sum_\lambda^K\sum_\sigma^K\sum_i^K\sum_j^K\sum_k^K\sum_l^K C_{\mu i}C_{\nu i}C_{\lambda j}C_{\sigma j}\Gamma_{ijkl}\left\langle \chi_\mu(\mathbf{r}_1)\chi_\nu(\mathbf{r}_1)\left|r_{12}^{-1}\right|\chi_\lambda(\mathbf{r}_2)\chi_\sigma(\mathbf{r}_2)\right\rangle$$

K^8 multiplications!

Γ_{ijkl} determined by perturbation theory or coupled cluster theory

$K^{5\text{-}7}$ multiplications!

MP2, MP4, CCSD(T), CI

Chemical Accuracy: DFT

$$E = \sum_i^L \left\langle \phi_i(\mathbf{r}) \middle| -\tfrac{1}{2}\nabla^2 \middle| \phi_i(\mathbf{r}) \right\rangle + \sum_i^L \left\langle \phi_i(\mathbf{r}) \middle| -\sum_B^{L_N} \frac{Z_B}{r_{iB}} \middle| \phi_i(\mathbf{r}) \right\rangle$$

$$+ \sum_i^L \sum_j^L \left\langle \phi_i(\mathbf{r}_1)\phi_i(\mathbf{r}_1) \middle| r_{12}^{-1} \middle| \phi_j(\mathbf{r}_2)\phi_j(\mathbf{r}_2) \right\rangle - \sum_i^L \sum_j^L \left\langle \phi_i(\mathbf{r}_1)\phi_i(\mathbf{r}_1) \middle| \upsilon_{xc}[\rho] \middle| \phi_j(\mathbf{r}_2)\phi_j(\mathbf{r}_2) \right\rangle$$

Speed: AM1, PM3

$$\left\langle \mu_A \nu_B \middle| \lambda_C \sigma_D \right\rangle \approx \delta_{AB}\delta_{CD} \left\langle \mu_A \nu_B \middle| \lambda_C \sigma_D \right\rangle \approx \sum_{a,b} Q_a Q_b \frac{q_a q_b}{r_{ab}}$$

≥ 12 parameters per atom

Molecular Mechanics

$$E = \sum_{bonds} \tfrac{1}{2} k_r (r - r_0)^2 + \sum_{angles} \tfrac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} \tfrac{1}{2} k_\phi \cos(n\phi - \delta)$$

$$+ \sum_{m,n} \frac{q_m q_n}{r_{mn}} + \sum_{m,n} \left(\frac{A_{mn}}{r_{mn}^{12}} - \frac{B_{mn}}{r_{mn}^{6}} \right)$$

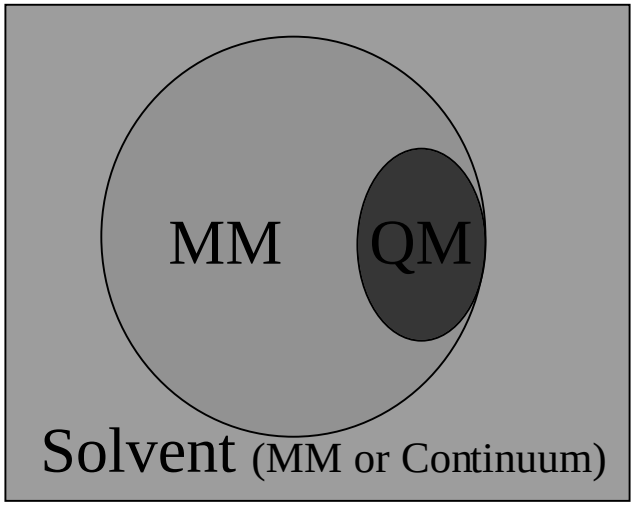
Very fast: ~10,000 atoms routine

Cannot treat bond making/breaking

Many empirical parameters

Bio force fields:
AMBER, CHARMM, GROMOS

Philosophy of QM/MM Approach



Warshell & Levitt (1977)
Bond making/breaking
in localized region

$$E = E_{QM} + E_{MM} + E_{QM\,/\,MM}$$

$$E_{QM\,/\,MM} = -\sum_{i,n}\left\langle \Psi \left| \frac{q_m}{r_{im}} \right| \Psi \right\rangle + \sum_{a,m} \frac{Z_a q_m}{r_{im}} + \sum_{a,m}\left(\frac{A_{am}}{r_{am}^{12}} - \frac{B_{am}}{r_{am}^{6}} \right)$$

Practical Considerations

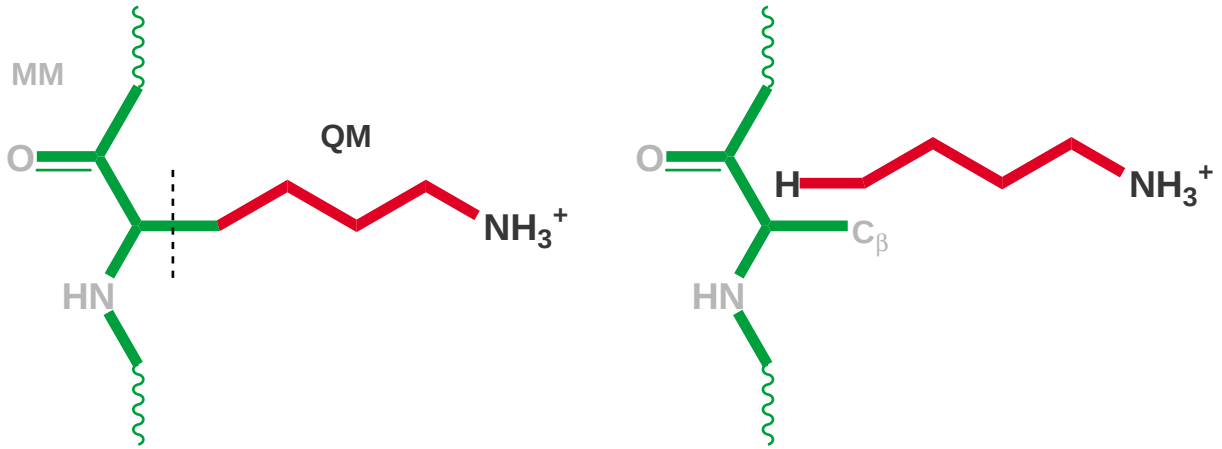
E_{QM}: HF, DFT, AM1/PM3 ?

E_{MM}: AMBER, CHARMM, GROMOS?
Are parameters available?

$E_{QM/MM}$: Boundary across covalent bond?

Singh & Kollman (1986)

E_{QM}: HF E_{MM}: AMBER $E_{QM/MM}$: Link atom



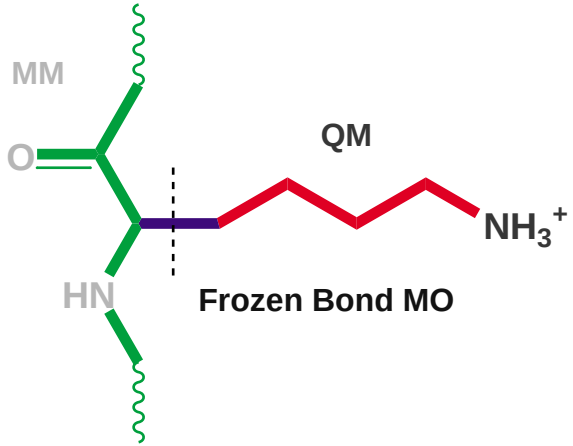
Field, Bash, Karplus (1990)

E_{QM}: AM1 E_{MM}: CHARMM $E_{QM/MM}$: Link atom

Application: Triosephosphate Isomerase 1991

Rivail & co-workers (1994)

E_{QM}: AM1 E_{MM}: AMBER $E_{QM/MM}$: **Hybrid MO**



Rivail & co-workers (1996)

E_{QM}: **HF** E_{MM}: AMBER $E_{QM/MM}$: Hybrid MO

Summary of current approaches

Karplus & co-workers

E_{QM}: DFT, HF, AM1 E_{MM}: CHARMM $E_{QM/MM}$: Link atoms
CHARMM interface with GAMESS or CADPAC

Friesner & co-workers

E_{QM}: DFT, HF E_{MM}: OPLS-AA $E_{QM/MM}$: Hybrid Orbital
Qsite (Macromodel interface with Jaguar)

Gao & co-workers

E_{QM}: AM1 E_{MM}: CHARMM $E_{QM/MM}$: Hybrid Orbital*

Yang & co-workers

E_{QM}: HF, DFT E_{MM}: CHARMM $E_{QM/MM}$: Link atom*

*Specific parameterization allows bond length change

Considerations

- Link atoms *vs.* Hybrid MOs?
  Karplus (2000): Similar for E_{QM}: AM1

- CHARMM *vs.* AMBER *vs.* GROMOS *vs.* OPLS-AA?
  Not studied in the context of QM/MM!

- Is $E_{QM/MM}$ expression sufficient?

$$E_{QM/MM} = -\sum_{i,n}\left\langle \Psi \left| \frac{q_m}{r_{im}} \right| \Psi \right\rangle + \sum_{a,m} \frac{Z_a q_m}{r_{im}} + \sum_{a,m}\left(\frac{A_{am}}{r_{am}^{12}} - \frac{B_{am}}{r_{am}^{6}} \right)$$

- Free energy: E $\rightarrow$ G = H - TS

- Conformational sampling

Is $E_{QM/MM}$ expression sufficient?

Theory of Intermolecular Interactions

***ab initio* QM** ← $\hat{H}\Psi = E\Psi$ $\Rightarrow$ $E = \left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle$

RHF/6-31G*
MP2/aug-cc-pVDZ
B3LYP/6-311+G**

Perturbation Theory
(Taylor Expansion)

$$\hat{H} = \hat{H}_A + \hat{H}_B + \lambda \hat{H}_{AB}$$

$$\Psi = \Psi^{(0)} + \lambda\Psi^{(1)} + \lambda^2\Psi^{(2)} + \mathrm{K} \qquad \text{where} \quad \Psi^{(0)} = \Psi_A\Psi_B$$

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \mathrm{K}$$

Theory of intermolecular interactions continued

$$E^{(0)} = \left\langle \Psi^{(0)} \middle| \hat{H}_A + \hat{H}_B \middle| \Psi^{(0)} \right\rangle = E_A + E_B$$

$$E^{(1)} = \left\langle \Psi^{(0)} \middle| \hat{H}_{AB} \middle| \Psi^{(0)} \right\rangle = E_{AB}^{Electrostatic}$$

$$E^{(2)} = \left\langle \Psi^{(0)} \middle| \hat{H}_{AB} \middle| \Psi^{(1)} \right\rangle = E_A^{Induction} + E_B^{Induction} + E_{AB}^{Dispersion}$$

$$\Psi^{(1)} = \Psi_A^{Induction} \Psi_B + \Psi_A \Psi_B^{Induction} + \Psi_{AB}^{Dispersion}$$

Long Range (no overlap)

$$E_{AB} \approx E_{AB}^{Electrostatic} + E_A^{Induction} + E_B^{Induction} + E_{AB}^{Dispersion}$$

$$\Psi^{(0)} = \mathrm{A}\, \Psi_A \Psi_B$$ Pauli Exclusion Principle

Short Range (overlap)

$$E_{AB} \approx E_{AB}^{Electrostatic} + E_A^{Induction} + E_B^{Induction} + E_{AB}^{Dispersion}$$
$$+ E_{AB}^{Exchange-Repulsion} + E_A^{Exchange-Induction} + E_B^{Exchange-Induction} + E_{AB}^{Exchange-Dispersion}$$

$$\hat{H}_{AB} = \sum_{i\in A}\sum_{j\in B}\frac{1}{r_{ij}} = \frac{\hat{q}_A\hat{q}_B}{R_{AB}} + \frac{\hat{q}_A\hat{\mu}_B\cdot\dot{R}_{AB}}{R_{AB}^3} + \frac{\hat{q}_A f(\hat{\Theta}_B,\dot{R}_{AB})}{R_{AB}^5} + \frac{\hat{q}_A f(\hat{\Omega}_B,\dot{R}_{AB})}{R_{AB}^7}\mathsf{K}$$
$$+ \frac{\hat{\mu}_A\hat{q}_B\cdot\vec{R}_{AB}}{R_{AB}^3} + \frac{\hat{\mu}_A f(\hat{\mu}_B,\vec{R}_{AB})}{R_{AB}^5} + \frac{\hat{\mu}_A f(\hat{\Theta}_B,\vec{R}_{AB})}{R_{AB}^7} + \mathsf{K}$$
$$+ \frac{\hat{\Theta}_A f(\hat{q}_B,\vec{R}_{AB})}{R_{AB}^5} + \frac{\hat{\Theta}_A f(\hat{\mu}_B,\vec{R}_{AB})}{R_{AB}^7} + \mathsf{K}$$
$$+ \frac{\hat{\Omega}_A f(\hat{q}_B,\vec{R}_{AB})}{R_{AB}^7} + \mathsf{K}$$

$$E^{(0)} = \left\langle \Psi^{(0)} \middle| \hat{H}_A + \hat{H}_B \middle| \Psi^{(0)} \right\rangle = E_A + E_B$$

$$E^{(1)} = \left\langle \Psi^{(0)} \middle| \hat{H}_{AB} \middle| \Psi^{(0)} \right\rangle = E_{AB}^{Electrostatic}$$

$$E^{(2)} = \left\langle \Psi^{(0)} \middle| \hat{H}_{AB} \middle| \Psi^{(1)} \right\rangle = E_A^{Induction} + E_B^{Induction} + E_{AB}^{Dispersion}$$

$$E_{AB}^{Electrostatic} \approx \frac{q_A q_B}{R_{AB}} + \frac{q_A\mu_B\cdot\dot{R}_{AB}}{R_{AB}^3} + \frac{q_A f(\Theta_B,\dot{R}_{AB})}{R_{AB}^5} + \frac{q_A f(\Omega_B,\dot{R}_{AB})}{R_{AB}^7}\mathsf{K}$$

$$E_B^{Induction} \approx \frac{q_A\mu_B^{induced}\cdot\vec{R}_{AB}}{R_{AB}^3} + \frac{q_A f(\Theta_B^{induced},\vec{R}_{AB})}{R_{AB}^5} + \frac{q_A f(\Omega_B^{induced},\vec{R}_{AB})}{R_{AB}^7}\mathsf{K}$$

$$E_{AB}^{Dispersion} \approx \frac{f[\alpha_A(i\nu),\alpha_B(i\nu),\vec{R}_{AB}]}{R_{AB}^{10}} + \mathsf{K}$$

$$\approx \frac{f[\alpha_A(i\nu),\alpha_B(i\nu)]}{R_{AB}^6} + \mathsf{K}$$

Short Range (overlap)

$$E_{AB} \approx E_{AB}^{Electrostatic} + E_{A}^{Induction} + E_{B}^{Induction} + E_{AB}^{Dispersion} + E_{AB}^{Exchange-Repulsion} + E_{A}^{Exchange-Induction} + E_{B}^{Exchange-Induction} + E_{AB}^{Exchange-Dispersion}$$

$$E_{AB}^{Electrostatic} \approx \frac{q_A q_B}{R_{AB}} + \frac{q_A \mu_B \cdot \dot{R}_{AB}}{R_{AB}^3} + \frac{q_A f(\Theta_B, \dot{R}_{AB})}{R_{AB}^5} + \frac{q_A f(\Omega_B, \dot{R}_{AB})}{R_{AB}^7} + \mathrm{K}$$

$$E_{B}^{Induction} \approx \frac{q_A \mu_B^{induced} \cdot \vec{R}_{AB}}{R_{AB}^3} + \frac{q_A f(\Theta_B^{induced}, \vec{R}_{AB})}{R_{AB}^5} + \frac{q_A f(\Omega_B^{induced}, \vec{R}_{AB})}{R_{AB}^7} + \mathrm{K}$$

$$E_{AB}^{Dispersion} \approx \frac{f[\alpha_A(iv), \alpha_B(iv), \vec{R}_{AB}]}{R_{AB}^{10}} + \mathrm{K}$$

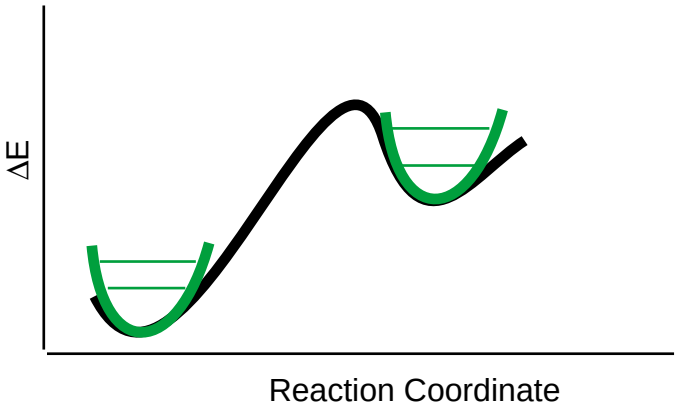
$$\approx \frac{f[\alpha_A(iv), \alpha_B(iv)]}{R_{AB}^6} + \mathrm{K}$$

$$E_{QM/MM} = -\sum_{i,n} \left\langle \Psi \left| \frac{q_m}{r_{im}} \right| \Psi \right\rangle + \sum_{a,m} \frac{Z_a q_m}{r_{im}} + \sum_{a,m} \left(\frac{A_{am}}{r_{am}^{12}} - \frac{B_{am}}{r_{am}^6} \right)$$

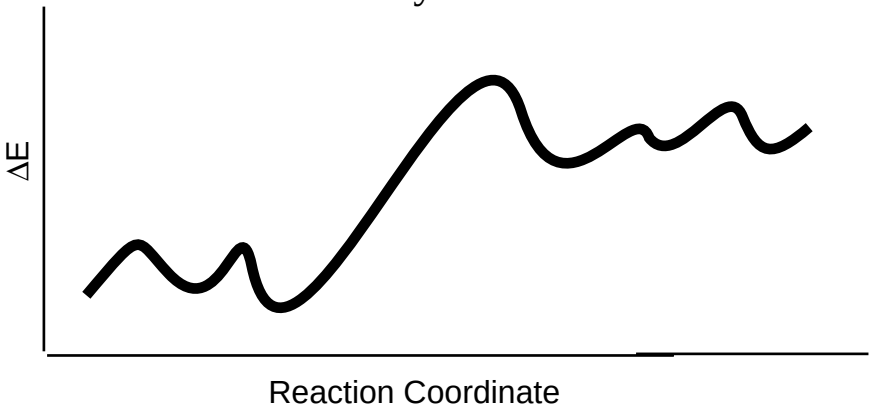
• Free energy: E → G = H - TS

• Conformational sampling

Harmonic Approximation



Molecular Dynamics or Monte Carlo



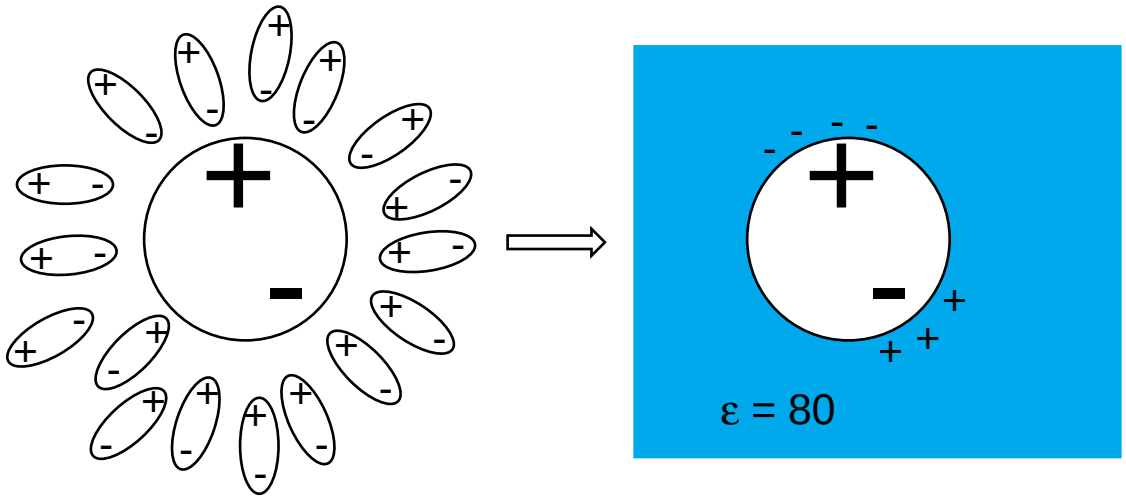
<u>Frequencies</u>
Karplus (2000)

<u>Free Energies</u>
Karplus (2001)

Practically limited to AM1
Gao & Truhlar (2000)

Carr-Parinello QM/MM
unpublished

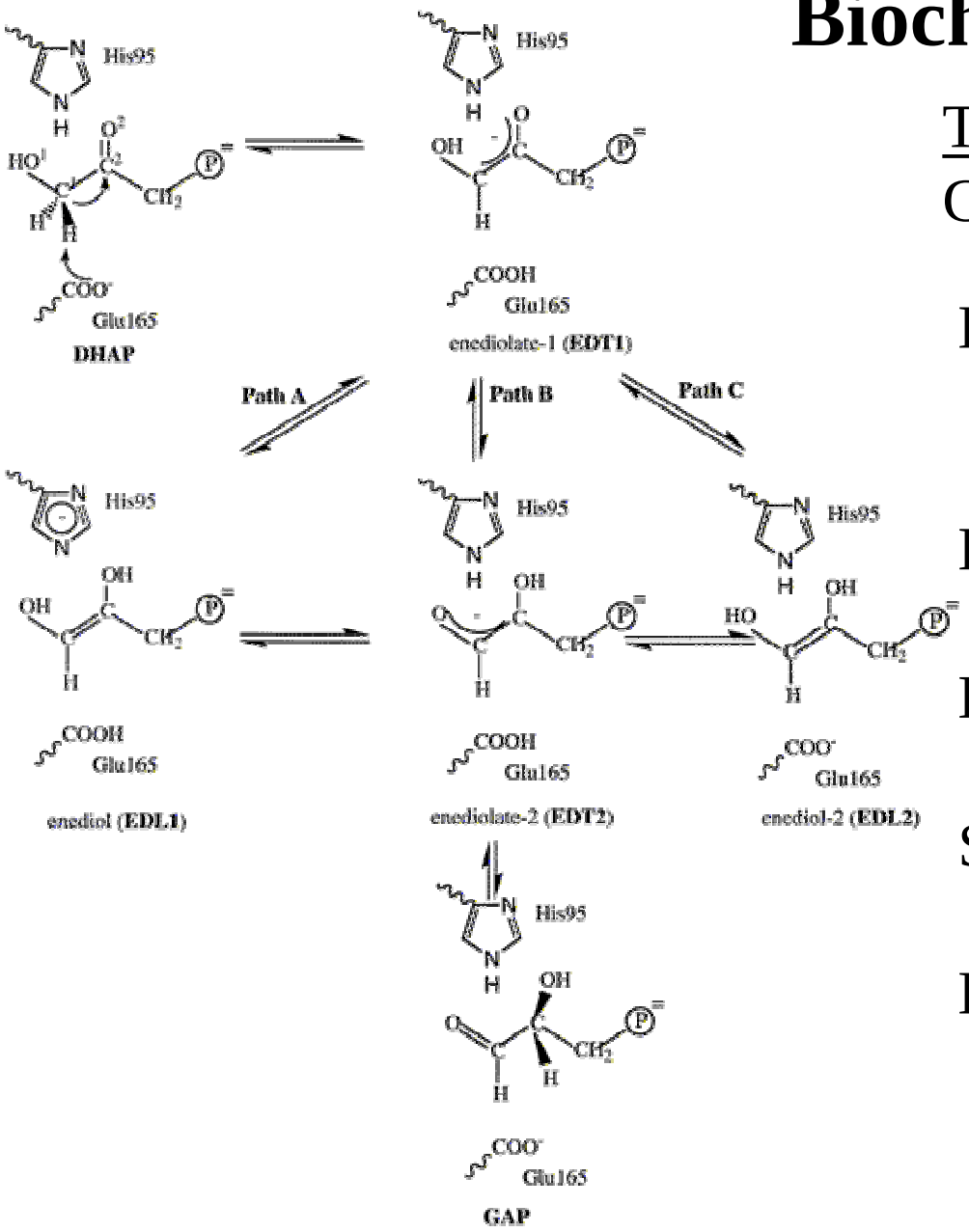
Solvation: explicit vs. continuum



$$q = -D(\varepsilon)^{-1} b(\vec{E})$$

Friesner (2000) [charge screening: Karplus (1997)]

Biochemical Application



<u>Tiosephosphate Isomerase</u>

Cui and Karplus, JACS 2001

E_{QM}: B3LYP/6-31+G(d,p)
CADPAC

E_{MM}: CHARMM

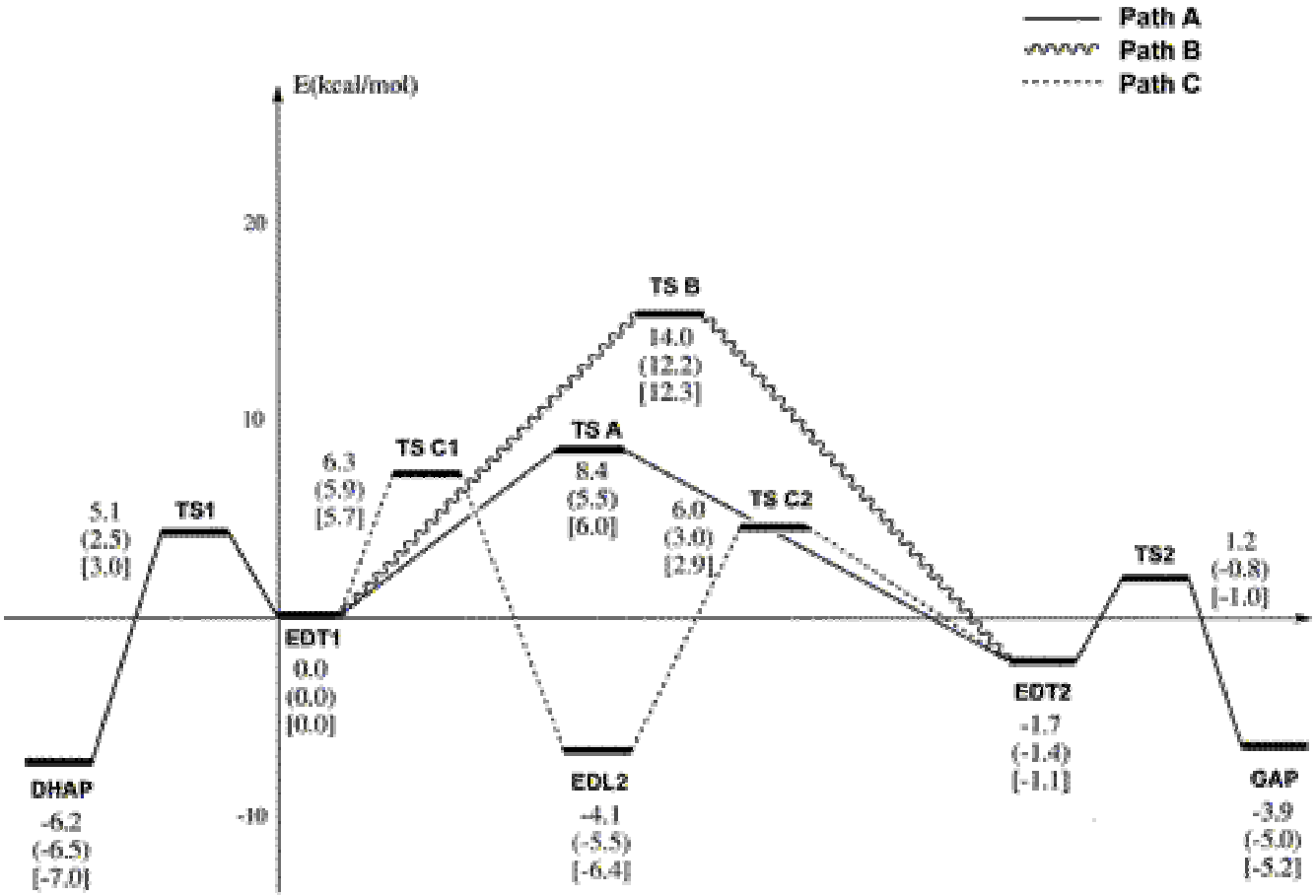
$E_{QM/MM}$: Link atom

Solvation: “Continuum”

Dynamics: Harmonic

Tiosephosphate Isomerase

Cui and Karplus, JACS 2001



Beyond Biomolecules

Force fields

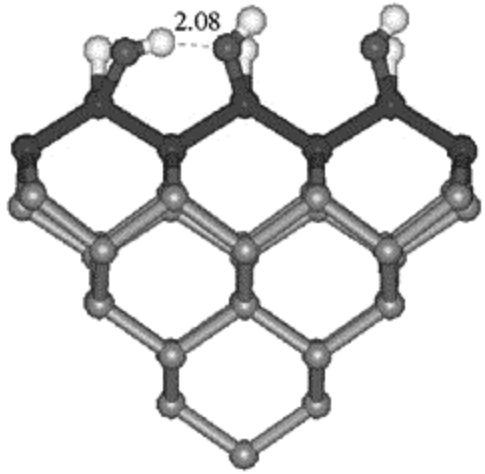
New parameters for covalent boundaries

Radicals

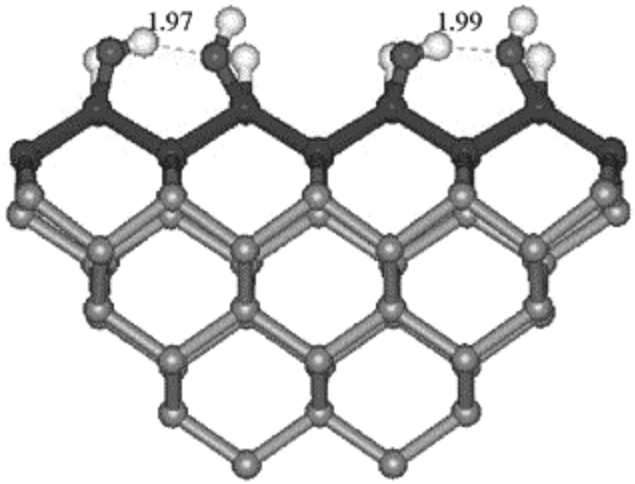
Solvation

Delocalization

Gordon & co-workers (2001)



(c) $Si_{48}H_{36}/3H_2O$



(d) $Si_{64}H_{44}/4H_2O$