Force fields and molecular modelling

Carine Clavaguéra

Laboratoire de Chimie Physique, CNRS & Université Paris Sud, Université Paris Saclay

carine.clavaguera@u-psud.fr
Outline

- Class I and class II force fields
- Class III force fields
- Solvent modelling
- Applications
- Reactive force fields
Multiscale modelling
History

- 1970: Proof of concept: Allinger
- 1980: Basis of force fields for classical molecular dynamics: Scheraga, Karplus, Kollman, Allinger
- 1990: Commercial distribution of program packages
- Since 2000
  - More and more publications about classical molecular dynamics simulations
  - Wide availability of codes
  - New generation force fields
$r^N : 3N$ coordinates \{x, y, z\} for N particles (atoms)

Force field = mathematical equations to compute the potential energy + parameters

\[
U(r^N) = U_s + U_b + U_\phi + U_{vdw} + U_{elec} + U_{pol}
\]

- **intramolecular interactions**
- **intermolecular interactions**

1-4 long-range interaction terms = interactions between 2 rigid spheres
Class I force fields

- CHARMM, AMBER, OPLS, GROMOS...

  - Intramolecular interactions: harmonic terms only
  - Intermolecular interactions: Lennard-Jones 6-12 and Coulomb term based on atomic charges

\[
V(r) = \sum_{s} k_s (r - r_o)^2 + \sum_{b} k_b (\theta - \theta_o)^2 + \sum_{\text{torsion}} V_n [1 + \cos((n \phi) - \delta)]
\]

\[
+ \sum_{i < j} \frac{q_i q_j}{\varepsilon r_{ij}} + 4 \varepsilon \left[ \left( \frac{r_0}{r} \right)^{12} - \left( \frac{r_0}{r} \right)^6 \right]
\]
Harmonic approximation

- Taylor series expansion around $b_o$
  \[ U_{bond} = \sum_b K_b (b - b_0)^2 \]

<table>
<thead>
<tr>
<th>Chemical type</th>
<th>$K_{bond}$ (kcal/mol/Å²)</th>
<th>$b_0$ (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C–C</td>
<td>100</td>
<td>1.5</td>
</tr>
<tr>
<td>C=C</td>
<td>200</td>
<td>1.3</td>
</tr>
<tr>
<td>C≡C</td>
<td>400</td>
<td>1.2</td>
</tr>
</tbody>
</table>

- Taylor series expansion around $\theta_o$
  \[ U_{angle} = \sum_a K_a (\theta - \theta_o)^2 \]
Torsion

- May be treated by direct 1-4 interaction terms
- Much more efficient to use a periodic form

\[ U_{\text{torsion}} = \sum_{\text{torsion}} V_{n}[1 + \cos((n\phi) - \delta)] \]
Improper torsion

- To describe out-of-plane movements
- Mandatory to control planar structures
- Example: peptidic bonds, benzene

\[ U_{\text{improper}} = \sum_{\phi} K_{\phi} (\phi - \phi_o)^2 \]
van der Waals interactions

London dispersion forces in $1/R^6$ and repulsive wall at short distances

Lennard-Jones 6-12 potential

$$E_{LJ}(R) = 4\varepsilon \left[ \left( \frac{R_o}{R} \right)^{12} - \left( \frac{R_o}{R} \right)^{6} \right]$$

Diatomic parameters

$$R_{oAB} = \frac{1}{2} \left( R_{oA} + R_{oB} \right)$$

$$\varepsilon_{AB} = \sqrt{\varepsilon^A + \varepsilon^B}$$
Electrostatic interaction

Interactions between punctual atomic charges

$$E_{elec} = \sum_{i<j} \frac{q_i q_j}{\varepsilon r_{ij}}$$

Quantum chemical calculation of atomic charges

1- Quantum chemical calculations of molecular models.

2- Electrostatic potential on a grid of points surrounding a molecule.

3- Atomic charges derived from the electrostatic potential.
Electrostatic potential fit

- Electrostatic potential (ESP) from the wave function (QM calculations)

\[
\phi_{esp}(\mathbf{r}) = \sum_{\text{nuc}} \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}|} - \int \frac{\Psi^*(\mathbf{r'})\Psi(\mathbf{r'})}{|\mathbf{r'} - \mathbf{r}|} d\mathbf{r'}
\]

- Basic idea: fit this quantity with point charges

- Minimize error function

\[
\text{ErrF}(\mathbf{Q}) = \sum_{\text{points}} \left( \phi_{esp}(\mathbf{r}) - \sum_{\text{atoms}} \frac{Q_A(\mathbf{R}_a)}{|\mathbf{R}_a - \mathbf{r}|} \right)
\]

An example: Restrained Electrostatic Potential (RESP) charges → potential is fitted just outside of the vdw radius
Class I force fields

- CHARMM, AMBER, OPLS, GROMOS...

\[
V(r) = \sum_{s} k_{s}(r - r_{o})^{2} + \sum_{b} k_{b}(\theta - \theta_{o})^{2} + \sum_{\text{torsion}} V_{n}[1 + \cos((n\phi) - \delta)] \\
+ \sum_{i < j} \frac{q_{i} q_{j}}{\varepsilon r_{ij}} + 4\varepsilon \left[ \left( \frac{r_{0}}{r} \right)^{12} - \left( \frac{r_{0}}{r} \right)^{6} \right]
\]

<table>
<thead>
<tr>
<th>Force field</th>
<th>bond and angle</th>
<th>vdw</th>
<th>Elec</th>
<th>Cross</th>
<th>Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMBER</td>
<td>P2</td>
<td>12-6 and 12-10</td>
<td>charge</td>
<td>none</td>
<td>proteins, nucleic acids</td>
</tr>
<tr>
<td>CHARMM</td>
<td>P2</td>
<td>12-6</td>
<td>charge</td>
<td>none</td>
<td>proteins</td>
</tr>
<tr>
<td>OPLS</td>
<td>P2</td>
<td>12-6</td>
<td>charge</td>
<td>none</td>
<td>proteins, nucleic acids</td>
</tr>
<tr>
<td>GROMOS</td>
<td>P2</td>
<td>12-6</td>
<td>charge</td>
<td>none</td>
<td>proteins, nucleic acids</td>
</tr>
</tbody>
</table>
Class I force fields: differences

- **AMBER (Assisted Model Building with Energy Refinement)**
  - Few atoms available, some calculations are impossible
  - Experimental data + quantum chemical calculations
  - Possible explicit terms for hydrogen bonds and treatment of lone pairs

- **CHARMM (Chemistry at HARvard Molecular Mechanics)**
  - Slightly better for simulations in solution
  - Charges parameterized from soluted-solvent interaction energies
  - No lone pair, no hydrogen bond term

- **OPLS (Optimized Potentials for Liquid Simulations)**
  - Derived from AMBER (intramolecular)
  - Non-bonded terms optimized for small molecule solvation
  - No lone pair, no hydrogen bond term
How to obtain intramolecular parameters?

- Quantum chemistry and experimental data
  - Force constants in bonded terms: vibrational frequencies, conformational energies
  - Geometries: computations or experiments (ex. X-ray)

Example: potential energy surfaces of molecules with several rotations
Class II force fields

- **MMFF94, MM3, UFF**
  - Anharmonic terms (Morse, higher orders,...)
  - Coupling terms (bonds/angles, ...)
  - Alternative forms for non-bonded interactions

- **MM2/MM3**
  - Successful for organic molecules and hydrocarbons

- **MMFF (Merck Molecular Force Field)**
  - Based on MM3, parameters extracted on ab initio data only
  - Good structures for organic molecules
  - Problem: condensed phase properties

- **Improvements**
  - Conformational energies
  - Vibrational spectra
Morse potential → Taylor series expansion around $r_o$

Beyond harmonic approximation:

$$U_{bond} = \sum K_2(b - b_0)^2 + K_3(b - b_0)^3 + K_4(b - b_0)^4 + \ldots$$
Bond angle

Harmonic approximation:

\[ U_{\text{angle}} = \sum_a K_a (\theta - \theta_o)^2 \]

Taylor series expansion around \( \theta_o \)

MM3: beyond a quadratic expression
\( \theta^3 \) term mandatory for deformation larger than 10-15°.

\[
U_{\text{angle}} = \sum_a K_a (\theta - \theta_o)^2 \left[ 1 - 0.014(\theta - \theta_o) + 5.6\times10^{-5}(\theta - \theta_o)^2 \\
- 7.0\times10^{-7}(\theta - \theta_o)^3 + 2.2\times10^{-8}(\theta - \theta_o)^4 \right]
\]
Couplings

Coupling terms between at least 2 springs

- **Stretch-bend coupling**

  ![Diagram of stretch-bend coupling](image)

  \[ E_{sb} = \sum_{l,l'} K_{l\theta} \left[ (l - l_0) + (l - l'_0) \right] (\theta - \theta_0) \]

- Other couplings: stretch-torsion, angle-torsion, ...
Alternative to Lennard-Jones potential: Buckingham potential

\[ E_{vdw}(R) = \varepsilon \left[ e^{-\frac{aR_o}{R}} - \left( \frac{R_o}{R} \right)^6 \right] \]

- Used in the MM2/MM3 force fields
- However, LJ potential is usually preferred in biological force fields
  
  \[ \rightarrow \text{Computational cost of the exponential form in comparison with } R^{-12} \]

<table>
<thead>
<tr>
<th>Force field</th>
<th>bond and angle</th>
<th>vdw</th>
<th>Elec</th>
<th>Cross</th>
<th>Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM2</td>
<td>P3</td>
<td>Exp-6</td>
<td>dipole</td>
<td>sb</td>
<td>general</td>
</tr>
<tr>
<td>MM3</td>
<td>P4</td>
<td>Exp-6</td>
<td>dipole</td>
<td>sb, bb, st</td>
<td>general</td>
</tr>
<tr>
<td>MMFF</td>
<td>P4</td>
<td>14-7</td>
<td>charge</td>
<td>sb</td>
<td>general</td>
</tr>
<tr>
<td>UFF</td>
<td>P2 or Morse</td>
<td>12-6</td>
<td>charge</td>
<td>none</td>
<td>all elements</td>
</tr>
</tbody>
</table>
Figure 8 Calculated conformational energy differences between axial and equatorial methyl-cyclohexane in kcal/mol. The dashed line shows the experimental value.
Comparison of conformational energies for organic molecules

Figure 9  Calculated energy differences in kcal/mol between axial and equatorial amino-cyclohexane. The dashed line indicates the experimental value.
Comparison of conformational energies for organic molecules

[Graph showing a comparison of mean absolute errors for different force fields across various categories of compounds.]
Biological force fields

Hydration Free Energy of Amino Acid Side Chain Analogs: Differences from Experiment

Ala Val Leu Ile Ser Thr Phe Tyr Cys Met Asn Gln Trp Hid Hie

Limitations

- "Additive force fields"
  - No full physical meaning of the individual terms of the potential energy
  - No inclusion of the electric field modulations
  - No change of the charge distribution induced by the environment
  - No many-body effects

- Transferability
  - Better accuracy for the same class of compounds used for parameterization
  - No transferability between force fields

- Properties from electronic structure unavailable (electric conductivity, optical and magnetic properties)
  - Impossible to have breaking or formation of chemical bonds
Next Generation Force Fields

To overcome the limitation of additive empirical force fields

- Polarizable force fields: 3 models based on
  - Induced dipoles
  - Drude model
  - Fluctuating charges

X-Pol, SIBFA, AMOEBA, NEMO, evolution of AMBER, CHARMM

- Able to reproduce physical terms derived from QM: total electrostatic energy, charge transfer, ...

- Optimized for hybrid methods (QM/MM)

- To account for the electronegativity and for hyperconjugation

- Reactive force fields
Accurate description of electrostatic effects

\[ H_2O \quad H_2O + q \]


Multipole expansion

- Dipole-Dipole interaction: \( (\mu_1 \mu_2)/R^3 \)
- Dipole-Quadrupole interaction: \( (\mu_1 Q_2)/R^4 \)
- Quadrupole-Quadrupole interaction: \( (Q_1 Q_2)/R^5 \)

Errors (V) for electrostatic potential on a surface around N-methyl propanamide

Point charges vs. multipole expansion

Stone Science, 2008
Multipolar distributions

- Distributed multipole analysis (DMA): multipole moments extracted from the quantum wave function
- Multipoles are fitted to reproduce the electrostatic potential


Acrolein: molecular electrostatic potential
- QM (reference)
- DMA
- Fitted multipoles

Difference between ab initio ESP and
- DMA
- Fitted multipoles

Solvent induced dipoles

\[ \Delta U = -9.9 \text{ kcal/mol} \]

\[ \Delta U_{\text{pol}} = 2.5 \text{ kcal/mol} \]

- The matter is polarized proportionally to the strength of an applied external electric field.
- The constant of proportionality \( \alpha \) is called the polarizability with \( \mu_{\text{induit}} = \alpha E \)
Induced dipole polarization model

- Induced dipole on each site $i$
  \[ \mu_{i,\alpha}^{\text{ind}} = \alpha_i E_{i,\alpha} \quad (\alpha \in \{x, y, z\}) \quad \alpha_i : \text{atomic polarizability}, \ E_{i,\alpha} : \text{electric field} \]

- Mutual polarization by self-consistent iteration (all atoms included) : sum of the fields generated by both permanent multipoles and induced dipoles
  \[ \mu_{i,\alpha}^{\text{ind}} = \alpha_i \sum_{\{j\}} T_{\alpha}^{ij} M_j + \alpha_i \sum_{\{j'\}} T_{\alpha,j',\beta}^{i,j} \mu_{j',\beta}^{\text{ind}} \quad T : \text{interaction matrix} \]
  \[ \text{induced dipole on site } i \text{ by the permanent multipoles of the other molecules} \quad \text{interaction dipole on site } i \text{ by the other induced dipole} \]

Implemented in the several force fields : AMOEBA, SIBFA, NEMO, etc.
Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins

Comparison of Ramachandran potential of mean force maps for GPGG peptide

- Proline-2 : (a) AMOEBA (b) PDB
- Glycine-3 : (c) AMOEBA (d) PDB

Superimposition of the final structures from AMOEBA simulations (green) and the experimental X-ray crystal structures (gray)

(a) Crambin  
(b) Trp cage  
(c) villin headpiece  
(d) ubiquitin  
(e) GB3 domain  
(f) RD1 antifreeze protein  
(g) SUMO-2 domain  
(h) BPTI  
(i) FK binding protein  
(j) lysozyme

Overall average RMSD (ten simulated protein structures) = 1.33 Å  
Seven of them are close to 1.0 Å

The Drude oscillator model

Classical Drude Oscillator

- Virtual sites carrying a partial electric charge $q_D$ (the Drude charge)
- Attached to individual atoms via a harmonic spring with a force constant $k_D$

\[ q_D = \sqrt{k_D \cdot \alpha} \]

- Response to the local electrostatic field $E$ via an induced dipole $\mu$

\[ \mu = \frac{q_D^2 E}{k_D} \]

- Changes in the electrostatic term in force fields and in the potential energy
  - $U_{elec}$ represents all Coulombic electrostatic interactions (atom-atom, atom-Drude, and Drude-Drude)
  - Adding a $U_{Drude}$ term that represents the atom-Drude harmonic bonds

A popular implementation: polarizable CHARMM force field

A polarizable force field for monosaccharides

Application of the Drude model derived from CHARMM

- Drude particles (blue, ‘D’) attached to non-H atoms
- Oxygen lone pairs (green, ‘LP’) connected with bond, angle and torsion terms
- Hydrogen non polarizable

- Polarization necessary to accurately model the cooperative hydrogen bonding effects
- Parameter optimization on quantum chemistry calculations in gas phase
- Transferability of the model to other monosaccharides

A polarizable force field for monosaccharides

Drude vs. CHARMM comparison

43 conformers: dipole moments and relative conformational energies

Validation of the model on condensed phase properties

- Subtil soluted-water and water-water equilibrium
- Role of concentration and temperature
- Error on densities $\approx 1\%$
- Agreement of the polarizable FF with NMR experiments
- Population of conformers with one exocyclique torsion for D-glucose and D-galactose

The fluctuating charge model

Polarization is introduced via the "movement" of the charges

- Expand energy at 2nd order as a function of the number of electrons $N$

$$E = E_0 + \Delta N \frac{\partial E}{\partial N} + \frac{1}{2} \Delta N \frac{\partial^2 E}{\partial N^2} \left( \Delta N \right)^2$$

- 2 parameters per chemical element: electronegativity ($\chi$) and hardness ($\eta$)

- Sum over sites (atomic centers)

$$E_{elec} = E_0 + \sum_A \chi_A Q_A + \frac{1}{2} \sum_{AB} \eta_{AB} Q_A Q_B$$

- Minimizing the electrostatic energy
  - Explicitly account for polarization and charge transfer
  - Main application: cooperative polarization in water

Basin-hopping: $C_{60} - (H_2O)_n$ clusters

Interaction energies $\Delta E$

\[ C_{60} + nH_2O \rightarrow C_{60} - (H_2O)_n \]

Potential energy of $C_{60} - (H_2O)_n$ with 2 contributions

- $H_2O - H_2O$ interaction: TIP4P model, class I and rigid, developed for bulk water
- $C_{60} - H_2O$ interaction: repulsion-dispersion (Lennard-Jones) + polarization

For $n = 1$, $\Delta E_a = \Delta E_b$ (1/3 from polarization)

Basin-hopping: global minima for $C_{60} - (H_2O)_n$ clusters

Method proposed by D. J. Wales

→ Potential energy surface converted into the set of basins of attraction of all the local minima

Basin-hopping: global minima for $C_{60} - (H_2O)_n$ clusters

Role of the $C_{60} - H_2O$ polarizable potential

Hydrophobic water-fullerene interaction highlighted
Summary: electrostatic effects

Class III force fields
→ To overcome the limitation of additive empirical force fields
→ To reproduce accurately electrostatic and polarization contributions

<table>
<thead>
<tr>
<th>Model</th>
<th>Polarization</th>
<th>Charge Transfer</th>
<th>Comput. Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pairwise fixed charges</td>
<td>implicit</td>
<td>implicit</td>
<td></td>
</tr>
<tr>
<td>Induced dipoles</td>
<td>✓</td>
<td>implicit</td>
<td></td>
</tr>
<tr>
<td>Drude oscillator</td>
<td>✓</td>
<td>implicit</td>
<td></td>
</tr>
<tr>
<td>Fluctuating charges</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

Lack of reactivity but important for chemists!
Solvation can have a large effect on structure and energetics.

Phosphoramidic acid is a stable "anionic zwitterion" at neutral pH.

Predicted Gas Phase Structure
P–N bond energy = 1.7 kcal/mole

Predicted Aqueous Phase Structure
P–N bond energy = 29.2 kcal/mole

Gas phase calculations find that the PMA is essentially dissociated.
Inclusion of solvation effects yields a stable P–N bond with a bond length close to the crystal structure (1.78 Angstroms).
Implicit solvent vs. explicit solvent

**Explicit**: each solvent molecule is considered

**Implicit**: the solvent is treated as a continuum description
Need to define a cavity that contains the solute
Implicit vs. explicit

- Computational efficiency for alanine dipeptide

<table>
<thead>
<tr>
<th>Solvent</th>
<th>$\Delta G$ (kcal/mol)</th>
<th>CPU Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicit</td>
<td>-13.40</td>
<td>$3.4\times10^5$</td>
</tr>
<tr>
<td>Implicit</td>
<td>-13.38</td>
<td>$1.0\times10^{-1}$</td>
</tr>
</tbody>
</table>

- Are implicit solvents sufficient?

Structure of DNA

Explicit solvent

Experimental

Implicit solvent
Implicit vs. explicit

Computational efficiency for alanine dipeptide

<table>
<thead>
<tr>
<th>Solvent</th>
<th>$\Delta G$ (kcal/mol)</th>
<th>CPU Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicit</td>
<td>-13.40</td>
<td>3.4x10$^5$</td>
</tr>
<tr>
<td>Implicit</td>
<td>-13.38</td>
<td>1.0x10$^{-1}$</td>
</tr>
</tbody>
</table>

Are implicit solvents sufficient?

![Diagram showing computational efficiency for alanine dipeptide](image)
Periodic boundary conditions

- Explicit solvation schemes
- The fully solvated central cell is simulated, in the environment produced by the repetition of this cell in all directions just like in a crystal
  - Infinite system with small number of particles
- Introducing a periodicity level frequently absent in reality
- Applications: liquids, solids

![Diagram of different shapes](image)
Periodic boundary conditions

- When an atom moves off the cell, it reappears on the other side (number of atoms in the cell conserved)

- Ideally, every atom should interact with every other atom

- Problem: the number of non-bonded interactions grows as $N^2$ where $N$ is the number of particles → limitation of the system size
Cutoff Methods

- Some interatomic forces decrease strongly with distance (van der Waals, covalent interactions etc.) : short-range interactions
  - Ignore atoms at large distances without too much loss of accuracy
  - Introduce a cutoff radius for pairwise interactions and calculate the potential only for those within the cutoff sphere
- Computing time dependence reduced to $N$

- Electrostatic interaction is very strong and falls very slowly : long-range interactions
  - The use of a cutoff is problematic and should be avoided
  - Several algorithms exist (Ewald summation, Particle Mesh Ewald, ....) to treat the long-range interactions without cutoffs
Setting up a simulation

- Optimize the solute geometry in vacuum (no more constraint)
- Add counter-ions if necessary
- Choose the good cell shape (cubic, hexagonal prism, etc.) depending on the shape of the system
- Add solvent molecules around the solute (water for example)
- Delete the solvent molecules which are too close to the solute
- Create the periodic boundary conditions
- "Equilibration" of the global system at constant pressure and temperature
Water as a solvent

Water models with 3 to 5 interaction sites

(a) 3 sites : TIP3P  (b) 4 sites : TIP4P  (c) 5 sites : TIP5P
The most used: 3 interaction sites such as SPC and TIP3P.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Number of point charges</th>
<th>$\mu_{\text{gaz}}$ (D)</th>
<th>$\mu_{\text{liq.}}$ (D)</th>
<th>$\epsilon_0$</th>
<th>Exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berendsen</td>
<td>3</td>
<td>2.27</td>
<td>2.27</td>
<td>65</td>
<td>1.85</td>
</tr>
<tr>
<td>Jorgensen</td>
<td>3</td>
<td>2.35</td>
<td>2.35</td>
<td>82</td>
<td>2.85</td>
</tr>
<tr>
<td>Jorgensen</td>
<td>3</td>
<td>2.18</td>
<td>2.18</td>
<td>53</td>
<td>78.4</td>
</tr>
</tbody>
</table>

(at 25 °C)
Experimental and computed thermodynamic data

25°C, 1 atm

<table>
<thead>
<tr>
<th></th>
<th>SPC</th>
<th>TIP3P</th>
<th>TIP4P</th>
<th>Exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$ (kg.dm$^{-3}$)</td>
<td>0.971</td>
<td>0.982</td>
<td>0.999</td>
<td>0.997</td>
</tr>
<tr>
<td>$E$ (kcal.mol$^{-1}$)</td>
<td>-10.18</td>
<td>-9.86</td>
<td>-10.07</td>
<td>-9.92</td>
</tr>
<tr>
<td>$\Delta H_{vap.}$ (kcal.mol$^{-1}$)</td>
<td>10.77</td>
<td>10.45</td>
<td>10.66</td>
<td>10.51</td>
</tr>
<tr>
<td>$C_p$ (cal.mol$^{-1}$.deg$^{-1}$)</td>
<td>23.4</td>
<td>16.8</td>
<td>19.3</td>
<td>17.99</td>
</tr>
</tbody>
</table>
In the solid phase, water exhibits one of the most complex phase diagrams, having 13 different (known) solid structures.

Can simple models describe the phase diagram?

Only TIP4P provides a qualitatively correct phase diagram for water.
## Water dimer properties

<table>
<thead>
<tr>
<th></th>
<th>POL5</th>
<th>TIP4P/FQ</th>
<th>TIP5P</th>
<th>ab initio</th>
<th>exp</th>
</tr>
</thead>
<tbody>
<tr>
<td>U (kcal/mol)</td>
<td>-4.96</td>
<td>-4.50</td>
<td>-6.78</td>
<td>-4.96</td>
<td>-5.4 ± 0.7</td>
</tr>
<tr>
<td>r (Å)</td>
<td>2.896</td>
<td>2.924</td>
<td>2.676</td>
<td>2.896</td>
<td>2.98</td>
</tr>
<tr>
<td>θ (degrees)</td>
<td>4.7</td>
<td>0.2</td>
<td>-1.6</td>
<td>4.8</td>
<td>0 ± 6</td>
</tr>
<tr>
<td>ϕ (degrees)</td>
<td>62.6</td>
<td>27.2</td>
<td>50.2</td>
<td>57.3</td>
<td>58 ± 6</td>
</tr>
<tr>
<td>μ (D)</td>
<td>2.435</td>
<td>3.430</td>
<td>2.920</td>
<td>2.683</td>
<td>2.643</td>
</tr>
<tr>
<td>⟨μ⟩ (D)</td>
<td>2.063</td>
<td>2.055</td>
<td>2.292</td>
<td>2.100</td>
<td>?</td>
</tr>
</tbody>
</table>
### Bulk water properties

<table>
<thead>
<tr>
<th></th>
<th>POL5</th>
<th>TIP4P/FQ</th>
<th>TIP5P</th>
<th>exp</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U$ (kcal/mol)</td>
<td>-9.92±0.01</td>
<td>-9.89±0.02</td>
<td>-9.87±0.01</td>
<td>-9.92</td>
</tr>
<tr>
<td>$\rho$ (g/cm$^3$)</td>
<td>0.997±0.001</td>
<td>0.998±0.001</td>
<td>0.999±0.001</td>
<td>0.997</td>
</tr>
<tr>
<td>$\mu$ (D)</td>
<td>2.712±0.002</td>
<td>2.6</td>
<td>2.29</td>
<td>? (2.5-3.2)</td>
</tr>
<tr>
<td>$\epsilon_0$</td>
<td>98±8</td>
<td>79±8</td>
<td>82±2</td>
<td>78.3</td>
</tr>
<tr>
<td>$D$ ($10^{-9} m^2/s$)</td>
<td>1.81±0.06</td>
<td>1.9±0.1</td>
<td>2.62±0.04</td>
<td>2.3</td>
</tr>
<tr>
<td>$\tau_{NMR}$ (ps)</td>
<td>2.6±0.1</td>
<td>2.1±0.1</td>
<td>1.4±0.1</td>
<td>2.1</td>
</tr>
</tbody>
</table>
From the single water molecule to the liquid phase

<table>
<thead>
<tr>
<th></th>
<th>( \mu(D) ) (Å)</th>
<th>( \alpha_{xx} (\text{Å}^3) )</th>
<th>( \alpha_{yy} (\text{Å}^3) )</th>
<th>( \alpha_{zz} (\text{Å}^3) )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H(_2)O</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMOEBA</td>
<td>1.853</td>
<td>1.660</td>
<td>1.221</td>
<td>1.332</td>
</tr>
<tr>
<td>QM</td>
<td>1.84</td>
<td>1.47</td>
<td>1.38</td>
<td>1.42</td>
</tr>
<tr>
<td>Exp.</td>
<td>1.855</td>
<td>1.528</td>
<td>1.415</td>
<td>1.468</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>( D_\text{e} ) (kcal/mol)</th>
<th>( r_{O-O} ) (Å)</th>
<th>( \mu_{\text{tot.}} ) (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(H(_2)O)(_2)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMOEBA</td>
<td>4.96</td>
<td>2.892</td>
<td>2.54</td>
</tr>
<tr>
<td>QM</td>
<td>4.98/5.02</td>
<td>2.907/2.912</td>
<td>2.76</td>
</tr>
<tr>
<td>Exp.</td>
<td>5.44 ± 0.7</td>
<td>2.976</td>
<td>2.643</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>( \rho ) (g/cm(^3))</th>
<th>( \Delta H_v ) (kcal/mol)</th>
<th>( \epsilon_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Liquid</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMOEBA</td>
<td>1.0004 ± 0.0009</td>
<td>10.48 ± 0.08</td>
<td>81 ± 13</td>
</tr>
<tr>
<td>Exp.</td>
<td>0.9970 ± 0.7</td>
<td>10.51</td>
<td>78.3</td>
</tr>
</tbody>
</table>
Molecular dynamics

**MD flowchart**

1. Initial coordinates (X-ray, NMR)
2. Structure minimization (release strain)
3. Solvation (if explicit solvent)
4. Initial velocities assignment
5. Heating dynamics (Temp. to 300K)
6. Equilibration dynamics (control of Temp. and structure)
   - Rescale velocities
   - Temp. Ok? Structure Ok?
7. Production dynamics (NVE, NVT, NPT)
8. Analysis of trajectory
9. Calculation of macroscopic values
Trajectory analysis

- Along the trajectory, atomic coordinates and velocities can be stored for further processing.
- A number of properties, e.g. structural, can then be computed and averaged over time.

**MD Analysis**

1) Mean Energy

\[ \langle E \rangle = \frac{1}{N} \sum_{i=1}^{N} E_i \]

2) RMS difference between two structures

\[ RMS = \left( \langle r_i^\alpha - r_i^\beta \rangle \right) = \sqrt{\frac{1}{N_i} \sum_{i} \left( r_i^\alpha - r_i^\beta \right)^2} \]

3) RMS fluctuations

\[ RMS_i^{\text{fluct}} = \sqrt{\frac{1}{N_j} \sum_{j} \left( r_i^j - r_i^{\text{average}} \right)^2} \]

Note the relation between the RMS fluctuations and the crystallographic B factors:

\[ B_i = \frac{8}{3} \pi^2 \left( RMS_i^{\text{fluct}} \right)^2 \]

4) Radius of Gyration

\[ \text{Radius of Gyration} = \sqrt{\frac{1}{N_i} \sum_{i} (r_i - r_{cm})^2} \]

where \( r_i - r_{cm} \) is the distance between atom \( i \) and the center of mass of the molecule.
The radial distribution function

- \( g(r) \): probability of finding a particle at a distance \( r \) from another particle, relative to the same probability for an ideal gas.
- \( g(r) \): can be determined experimentally, from X-ray and neutron diffraction patterns.
- \( g(r) \): can be computed from MD simulations.
- Example: solvation of the \( K^+ \) ion:
  - Simulation time: 2 ns
  - Time step: 0.1 fs
  - \( T = 298 \text{ K} \)
  - 216 water molecules d’eau or 100 formamide molecules.
Coordination number CN

CN : number of atoms or ligands

\[ CN(r) = 4\pi \rho \int_r^{r_1} g(r') r'^2 \, dr \]

Example : solvation of K\(^+\), Na\(^+\) and Cl\(^-\) ions

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Force Field</th>
<th>Potassium</th>
<th>Sodium</th>
<th>Chloride</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>AMOEBA</td>
<td>7.0</td>
<td>6.0</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td>OPLS-AA</td>
<td>6.6</td>
<td>6.5</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>CHARMM27</td>
<td>6.9</td>
<td>5.9</td>
<td>7.2</td>
</tr>
<tr>
<td>Formamide</td>
<td>AMOEBA</td>
<td>6.3</td>
<td>6.0</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td>OPLS-AA</td>
<td>6.9</td>
<td>6.0</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td>CHARMM27</td>
<td>6.6</td>
<td>5.9</td>
<td>7.5</td>
</tr>
</tbody>
</table>

*Values reported are calculated by integrating over the water–ion radial distribution functions (shown for K\(^+\) and Cl\(^-\) in Figures 3 and 4, respectively). The position of the first minimum in the radial distribution function was taken as the outer limit of the first solvation shell.
Micro-solvation of the K⁺ ion

- Cluster of 34 water molecules around the cation
- Simulation time: 10 ns
- Timestep: 1 fs
- Force field: AMOEBA polarizable force field
- \( T = 200 \text{ K} \)
Micro-solvation of the $K^+$ ion
Time Correlation Functions $C(t)$

- A(t) and B(t) : two time dependent signals
- A correlation function $C(t)$ can be defined

$$C(t) = \langle A(t_0) \cdot B(t_0 + t) \rangle = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau A(t_0)B(t_0 + t) dt_0$$

- If A and B represent the same quantity (A ≡ B), then $C(t)$ is called "autocorrelation function"
- The autocorrelation function shows how a value of A at $t = t_0 + t$ is correlated to its value at $t_0$
Example: transport coefficients

Diffusion coefficient $D$ directly related to the time integral of the velocity autocorrelation function

$$D = \frac{1}{3} \int_0^\infty \langle v_i(0) \cdot v_i(t) \rangle dt$$

Example: Diffusion coefficient for $K^+$, $Na^+$ and $Cl^-$ ions in water
Molecular Dynamics Study of the Liquid-Liquid Interface

  - (BMI$^+/Tf_2N^-$) ionic liquid/water biphasic system
  - Influence of the solvent of the Cs$^+$ extraction by the ligand "L" (calix[4]arene-crown-6)
  - Comparison with chloroform
Conditions of the simulations

- AMBER force field and corresponding code
  - TIP3P water model
  - OPLS chloroform model
  - Extraction of new parameters for IL
  - Cs\(^+\) parameters: QM micro-hydration

- Solvents
  - Adjacent cubic boxes of IL and water for the interface representation
  - System size
    - IL (219 molecules)/water (2657 molecules)
    - CLF (722 molecules)/water (2657 molecules)
  - Energy minimization of the systems mandatory!

- Other data
  - T = 300 K (Berendsen thermostat)
  - Leapfrog Algorithm, 2 fs time step
  - Long equilibration! Can be until 600 ps
  - Long simulations
    - From 20 to 60 ns for IL/water
    - From 0.4 to 2 ns for CLF/water
Simulations with the Cs$^+$/NO$_3^-$ ion pair

- Position of the interface defined by the intersection of the density curves
- 3 distinct domains including a 12 Å interface
- Starting point: 12 ion pairs in water
Simulations with the Cs\(^+\)L/NO\(_3^-\) complex

- From previous conclusion: the Cs\(^+\)/NO\(_3^-\) pair behaviour is different in IL and CLF

- A complex is made with Cs\(^+\) and the ligand L, NO\(_3^-\) as counterion
- Starting point: 3 complexes
  - 1 in water
  - 1 at the interface
  - 1 in IL
- Same simulation with chloroform
- Questions
  - Where will the complexes go?
  - Will Cs\(^+\) remain as a complex?
  - Which differences will be observed in the 3 domains?
  - What will be the differences between IL and CLF?
Results

- Hydrophobic complex
- Cs\(^{+}\)/L activity at the interface
- Ion transfer easier in IL than in CLF
Reactive force fields

How to bridge the gap between quantum chemistry and empirical force fields?
How to deal with chemical reactivity with a force field?

- **REAXFF, REBO, AIREBO**
  - To get a smooth transition from non-bonded to single, double and triple bonded systems
  - Use of bond length/bond order relationships updated along the simulation

- **BUT**
  - Very large number of parameters, expensive computational cost
  - The validation step is crucial

- **ReaxFF is 10-50 times slower than non-reactive force fields**

- Better behaviour than quantum methods
  - $N\log N$ for ReaxFF vs $>N^3$ for QM
\[ E_{\text{potential}} = (E_{\text{bond}} + E_{\text{over}} + E_{\text{under}}) + (E_{\text{angle}} + E_{\text{penalty}}) + E_{\text{torsion}} \\
+ E_{\text{lone-pair}} + E_{\text{conjugaison}} + E_{\text{H-bond}} + E_{\text{vdW}} + E_{\text{elec}} \]

All connectivity-dependent interactions are made bond-order dependent

→ Ensure that energy contributions disappear upon bond dissociation

- Over- and under-coordination of atoms must be avoided

→ Energy penalty when an atom has more bonds than its valence allows

→ e.g. Carbon no more than 4 bonds, Hydrogen no more than 1

- Electrostatic term = fluctuating charges to account for polarization effects


Electrostatic interactions

- Fluctuating charge model = electrostatic and polarization effects
- Good reproduction of Mulliken charges
- Largest part of the computational cost in the reactive force field
- Must be computed at each step of the simulation
Oxidation of aluminum nanoparticles

- ANPs: a solid propellant in the application of solid-fuel rockets
  - high combustion enthalpy
  - result of an oxidation process

- Growth of the oxide layer on ANPs: combined effects
  - inward diffusion of oxygen atoms
  - role of temperature, pressure of oxygen gas, induced electric field through the oxide layer

- ReaxFF reactive force field to assess the full dynamics of the oxidation process of ANPs

- Aluminum/oxygen interactions from an atomistic-scale view

Growth of an oxide film with temperature

Starting point 473 K, 573 K, 673 K

Oxidation of aluminum nanoparticles

- Mecanism of oxidation
  1. adsorption of the oxygen molecules → highly exothermic reaction
  2. formation of hot-spot regions on the surfaces of the ANPs
  3. void space created in these regions → reduction of reaction barrier for O$_2$ diffusion
  4. growth of the oxide layer

- Temperature and pressure of oxygen gas
  - Role on the number of void spaces
  - Density of the oxide layer

Role of the pressure of the oxygen gas

Growth of the oxide layer with temperature

- Low pressure of oxygen gas at 500 K
- High pressure of oxygen gas at 500 K

300 K 500 K 900 K
Nanotube formation with the aid of Ni-catalysis

Synthesis of chains of $C_{60}$ molecules inside single-wall carbon nanotubes

→ REAX-FF: various levels for the potential energy
   ▶ C-C
   ▶ Ni/Ni-C
   ▶ Ni cluster

- Understand the first stage of the nanotube growth
  ▶ 5 $C_{60}$ in a (10,10) nanotube
  ▶ van der Waals interactions
  ▶ $\sim 3.5$ eV / $C_{60}$

- Validation with quantum chemistry calculations: geometries, binding energies, reaction barrier height, heat capacity, etc.

- Simulations at 1800, 2000, 2400 and 2500 K

Nanotube formation with the aid of Ni-catalysis

**Without Ni catalyst, 5 C\textsubscript{60} in the nanotube**

- At 16 ps, coalescence only at \( T = 2500 \) K
- In agreement with the energy barrier for the formation of the C\textsubscript{60} dimer

**With Ni catalyst : 5 C\textsubscript{60} + 2 Ni between 2 C\textsubscript{60}**

- Lower coalescence temperature : \( T = 1800 \) K


Molecular Dynamics: survey of methods for simulating the activity of proteins, S. A. Adcock and J. A. McCammon, Chem. Rev. 2006, 105, 1589
