Force fields simply relates structures to energies.

**Ideal properties:** fast and accurate.

**Holy grail:** If a force field is both fast and accurate, we should be able to interpret and validate current experiments as well as make prediction for future experiments.

---

**Coarse grain**

![Coarse grain diagram](image)

1 particles

---

**United atom**

![United atom diagram](image)

3 particles

---

**All atom**

![All atom diagram](image)

9 particles

---

**Example:** polypropylene polymer simulation containing $10^4$ monomer units assuming $O(N^2)$ scaling.

- 10,000 particles
  \[
  \frac{(1 \times 10^4)^2}{(1 \times 10^4)^2} = 1
  \]
  \~10^8 \text{ steps}

- 30,000 particles
  \[
  \frac{(3 \times 10^4)^2}{(1 \times 10^4)^2} = 9
  \]
  \~10^7 \text{ steps}

- 90,000 particles
  \[
  \frac{(9 \times 10^4)^2}{(1 \times 10^4)^2} = 81
  \]
  \~10^6 \text{ steps (simulate time in ns)}
Why do we use forcefields in CP2K?

- Equilibriate structures for AIMD
- QM/MM methods
- Very useful for testing / learning - FORCE_EVAL and MOTION completely decoupled
- ...
All atom type force fields

Many force fields in use today but most have common functional forms. In particular, we will focus on force fields adapted for biological systems here.

\[ E_{\text{total}} = E_{\text{nonbonded}} + E_{\text{bonded}} \]

\[ E_{\text{nonbonded}} = E_{\text{electrostatics}} + E_{\text{vdW}} \]

\[ E_{\text{bonded}} = E_{\text{bond}} + E_{\text{bend}} + E_{\text{dihedral}} + E_{\text{cross}} \]

**AMBER, CHARMM, GROMOS, OPLS, CFF, UFF, and MMFF** are a few common force fields in use today to model biopolymers. With the exception of CFF and MMFF on the list, the others are considered Class I force fields with limited transferability.

**Polarizable force fields...**

Currently, CP2K only have support for Class I type force fields such as AMBER, CHARMM, GROMOS, and OPLS.
Nonbonded Pair potential

There are many functional forms for the nonbonded pair potential either in use or proposed literature. Each functional form is ideally suited for different purposes. CP2K have too many explicitly coded up in addition to the standard LJ. You can find the list at:

\[
V(r_{ij}) = \varepsilon_{ij} \left( \left( \frac{R_{\text{min}}}{r_{ij}} \right)^{12} - 2 \left( \frac{R_{\text{min}}}{r_{ij}} \right)^{6} \right)
\]

Since there are so many different functional form, one interesting feature within CP2K is the ability to use any arbitrary pairwise potential. The input section is located at:

```plaintext
%FORCE_EVAL%MM%FORCEFIELD%NONBONDED%GENPOT
ATOMS 0 0
FUNCTION a1/(r**a2) - a3*EXP(-a4*(r-a5)**2)-a6*EXP(-a7*(r-a8)**2)
PARAMETERS a1 a2 a3 a4 a5 a6 a7 a8
VARIABLES r
```

\[
V_{oo}(r_{ij}) = \frac{a_1}{r_{ij}^2} - a_3 e^{-a_4 (r_{ij}-a_5)^2} - a_6 e^{-a_7 (r_{ij}-a_8)^2}
\]

\[
V(r_{ij}) = \frac{q_i q_j}{4r_{ij}}
\]

Note charges are non integer partial charges. Many ways to efficiently sum up the interaction of the ions to its periodic images such as Ewald etc.
Bonds and Bends

\[ V(r_{ij}) = K(r_{ij} - r_{eq})^2 \]

\[ V(\theta_{ijk}) = K(\theta_{ijk} - \theta_{eq})^2 \]

**C - C**
\[
K = 600 \frac{kcal}{mol} \frac{1}{\AA^2} \\
r_{eq} = 1.335 \AA
\]

**CA - CA - CA**
\[
K = 40 \frac{kcal}{mol} \frac{1}{rad^2} \\
\theta_{eq} = 120^\circ
\]
Urey-Bradley and Dihedral

\[ V_{UB}(r_{ik}) = K_{UB}(r_{ik} - r_{eq})^2 \]

\[ V(\theta) = K[1 + \cos(n \varphi - \gamma)] \]

CTL2 – CTL2 - CTL2 - CLT2

\[ n = 2, K = 0.10 \frac{\text{kcal}}{\text{mol}}, \gamma = 180^\circ \]

Seen in CHARMM.
Improper (Out of plane motion)

\[ V(\theta_{ijkl}) = K(\varphi_{ijkl} - \varphi_{eq})^2 \]

CPB – CPA – NPH – CPA

\[ K = 20.8 \frac{kcal}{mol \ rad^2} \]

By definition, from the force field parameter and documentation, the first atom listed is the central atom. Therefore just like a normal torsion, the angle of interest is between the plane defined of particle \( ijk \) to that of the plane defined by particle \( jkl \). This is only used for special situations.
\[
V_{\text{total}} = \sum_{\text{bonds}} K_b (b - b_{eq})^2 + \sum_{\text{angle}} K_\theta (\theta - \theta_{eq})^2 + \sum_{\text{dihedrals}} \frac{V_n}{2} \left[ 1 + \cos(n\phi - \gamma) \right] + \sum_{i<j} \left[ \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} - \frac{q_i q_j}{\varepsilon R_{ij}} \right] \\
\gamma = 0^\circ \text{ or } 180^\circ
\]

There are different versions of the AMBER force fields. Select them carefully. The dielectric constant \( \varepsilon \) can have different values depending on what the system is composed of and what is simulated. Pay attention to the 1-4 scalings.
CHARMM

\[ V_{\text{total}} = \sum_{\text{bonds}} K_b (b - b_{eq})^2 + \sum_{\text{UB}} K_{UB} (S - S_{eq})^2 + \sum_{\text{angle}} K_\theta (\theta - \theta_{eq})^2 + \sum_{\text{dihedrals}} K_\phi [1 + \cos(n\phi - \gamma)] + \sum_{\text{improvers}} K_{imp} (\varphi - \varphi_{eq})^2 + \]

\[ \sum_{i<j} \varepsilon \left( \frac{R_{\text{min}_{ij}}}{r_{ij}} \right)^{12} - \left( \frac{R_{\text{min}_{ij}}}{r_{ij}} \right)^6 + \frac{q_i q_j}{r_{ij}} \]

\[ \varepsilon_{ij} = \sqrt{\varepsilon_{ii} \varepsilon_{jj}} \quad R_{\text{min}_{ij}} = \frac{1}{2} (R_{\text{min}_i} + R_{\text{min}_j}) \]

There are different version of the CHARMM force field. CHARMM22 protein, CHARMM27 nucleic acids and lipids. We do not currently have CMAP support.

Note: Specifically parameterized to be used with TIP3P. The use of “other water models would be less appropriate.”
\[ V_{total} = \sum_{\text{bonds}} K_r (r - r_{eq})^2 + \]
\[ \sum_{\text{angles}} K_\theta (\theta - \theta_{eq})^2 + \]
\[ \sum_{\text{dihedrals}} \frac{V_1}{2} [1 + \cos(\phi - \gamma_1)] + \]
\[ \frac{V_2}{2} [1 + \cos(2\phi - \gamma_2)] + \]
\[ \frac{V_3}{2} [1 + \cos(3\phi - \gamma_3)] + \]
\[ \sum_{i<j} \left[ 4\varepsilon_{ij} \left( \frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^6}{r_{ij}^6} \right) - \frac{q_i q_j e^2}{r_{ij}} \right] f_{ij} \]

\( \gamma_1 = \gamma_2 = \gamma_3 = 0^\circ \quad \sigma_{ij} = \sqrt{\sigma_{ii} \sigma_{jj}} \quad f_{ij} = 0.5 \text{ for } 1\text{-}4 \text{ interactions} \]
\( \varepsilon_{ij} = \sqrt{\varepsilon_{ii} \varepsilon_{jj}} \quad f_{ij} = 1.0 \text{ for everything else} \)
An Example of Class II force field: MMFF

\[ V_{\text{total}} = \sum_{\text{bonds}} K_{\text{bond}} (r - r_{\text{eq}})^2 \left(1 + cs(r - r_{\text{eq}}) + \frac{7}{12} cs^2 (r - r_{\text{eq}})^2\right) + \]

\[ \sum_{\text{angle}} K_\theta (\theta - \theta_{\text{eq}})^2 \left(1 + cb(\theta - \theta_{\text{eq}})\right) + \]

\[ \sum_{\text{angle,linear}} K_{\text{al}} (1 + \cos(\theta)) + \]

\[ \sum_{\text{stretch,bend}} \left(K_{ij} (r_{ij} - r_{\text{eq}}) + K_{kj} (r_{kj} - r_{\text{eq}})\right) (\theta - \theta_{\text{eq}}) + \]

\[ \sum_{\text{outofplane}} K_{\text{OOP}} (\chi)^2 + \]

\[ \sum_{\text{dihedrals}} \frac{V_1}{2} [1 + \cos(\phi)] + \frac{V_2}{2} [1 + \cos(2\phi)] + \frac{V_3}{2} [1 + \cos(3\phi)] + \]

\[ \sum_{i < j} \left[ \varepsilon_{ij} \left(\frac{1.07 \sigma}{r_{ij} + 0.07 \sigma}\right)^7 \left(\frac{1.12 \sigma^7}{r_{ij}^7 + 0.07 \sigma^7} - 2\right) - \frac{q_i q_j}{D(r_{ij} + \delta)} \right] \]
\[ E = \sum_i E_i = \frac{1}{2} \sum_{i,j} V_{ij} \]

\[ V_{ij} = f_C (r_{ij}) \left[ a_{ij} f_R (r_{ij}) + b_{ij} f_A (r_{ij}) \right] \]

\[ f_R (r) = A \exp(-\lambda_1 r) \]

\[ f_A (r) = -B \exp(-\lambda_2 r) \]

\[ f_C (r) = \begin{cases} 1, & r < R - D \\ \frac{1}{2} - \frac{1}{2} \sin \left[ \frac{\pi (r - R)}{2D} \right], & R - D < r < R + D \\ 0, & r > R + D \end{cases} \]

\[ b_{ij} = (1 + \beta^n \zeta_{ij}^n)^{1/2n} \]

\[ \zeta_{ij} = \sum_{k(i,j)} f_C (r_{ik}) g(\theta_{ijk}) \exp[\lambda_3^3 (r_{ij} - r_{ik})^3] \]

\[ g(\theta) = 1 + \frac{c^2}{d^2} - \frac{c^2}{d^2 + (h + \cos \theta)^2} \]

\[ a_{ij} = (1 + \alpha^n \eta_{ij}^n)^{1/2n} \]

\[ \eta_{ij} = \sum_{k(i,j)} f_C (r_{ik}) \exp[\lambda_3^3 (r_{ij} - r_{ik})^3] \]
Parameterization of force fields

Where do all the force field parameters come from?? Many sources are used to help parameterized the force fields: X-ray, ab initio calculations (HF/6-31G(d), Cambridge Crystal Data Bank, IR, Raman, or thermodynamic properties.

CHARMM: Iterative optimization of the intermolecular and intramolecular parameters until self consistency was reached.
Intermolecular: Optimization of atomic charge follow by LJ paramters.
Intramolecular: Optimization of bond, bend dihedral and UB parameters.

“More meaningful parameter values, which have a wider range of applicability, were obtained manually with “reasonable” parameter ranges for the optimization in the iterative refinement procedure described above.”
Topology and Parameter file

RESI DMPA -1.00 ! Dimethylphosphate

GROUP

ATOM P1 PL 1.50 !
ATOM O3 O2L -0.78 !
ATOM O4 O2L -0.78 !    H11
ATOM O1 OSL -0.57 !    |
ATOM O2 OSL -0.57 !    H13- C1-H12
ATOM C1 CTL3 -0.17 !    (-) O3 O1
ATOM H11 HAL3 0.09 !    /  /
ATOM H12 HAL3 0.09 !    P1 (+)
ATOM H13 HAL3 0.09 !    /  /
ATOM C2 CTL3 -0.17 !    /  /
ATOM H21 HAL3 0.09 !    H23-C2-H22
ATOM H22 HAL3 0.09 !    |
ATOM H23 HAL3 0.09 !    H21

BOND P1 O1 P1 O2 P1 O3 P1 O4 O1 C1 O2 C2
BOND C1 H11 C1 H12 C1 H13 C2 H21 C2 H22 C2 H23

BONDS

OSL PL 270.0 1.60

ANGLES

OSL PL OSL 80.0 104.3

DIHEDRALS

OSL PL OSL CTL3 1.20 1 180.00
OSL PL OSL CTL3 0.10 2 180.00
OSL PL OSL CTL3 0.10 3 180.00
Comparison to experiments for N-methylacetamide: 1) Charges 2) Vibrations

There are a few model compounds that force field developers like to use. Each model compound usually contains a motif that is often found repeatedly in biological systems. One of the compounds that was extensively tested and was used for parameterization NMA. The interest in NMA is because it’s a small fragment that can be used to represent peptide bond linkage.

![Chemical structure of N-methylacetamide](image)

**Table 1**

<table>
<thead>
<tr>
<th>Atom</th>
<th>CHARMM</th>
<th>AMBER</th>
<th>OPLS-AA</th>
<th>GROMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.51</td>
<td>0.5869</td>
<td>0.50</td>
<td>0.38</td>
</tr>
<tr>
<td>O</td>
<td>-0.51</td>
<td>-0.5911</td>
<td>-0.50</td>
<td>-0.38</td>
</tr>
<tr>
<td>N</td>
<td>-0.47</td>
<td>-0.4192</td>
<td>-0.50</td>
<td>-0.28</td>
</tr>
<tr>
<td>H</td>
<td>0.31</td>
<td>0.2823</td>
<td>0.30</td>
<td>0.28</td>
</tr>
<tr>
<td>C₄</td>
<td>0.27</td>
<td>-0.0411</td>
<td>-0.18</td>
<td>0.00</td>
</tr>
<tr>
<td>H₂₅</td>
<td>0.09</td>
<td>0.0173</td>
<td>0.06</td>
<td>N/A</td>
</tr>
<tr>
<td>C₆</td>
<td>-0.11</td>
<td>-0.2078</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>H₆₂</td>
<td>0.09</td>
<td>0.1127</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Table 3:** Vibrational Data for N-Methylacetamide

<table>
<thead>
<tr>
<th>Mode</th>
<th>Experimental/frequency</th>
<th>Assignment</th>
<th>CHARMM</th>
<th>Frequency</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>VLF</td>
<td>64</td>
<td>rCC(3)</td>
<td>64</td>
<td>rCC(3)</td>
</tr>
<tr>
<td>2</td>
<td>VLF</td>
<td>89</td>
<td>rNC(3)</td>
<td>89</td>
<td>rNC(3)</td>
</tr>
<tr>
<td>3</td>
<td>171°</td>
<td>200</td>
<td>rC5-N7</td>
<td>200</td>
<td>rC5-N7</td>
</tr>
<tr>
<td>4</td>
<td>270°</td>
<td>271</td>
<td>αNCH(3)</td>
<td>271</td>
<td>αNCH(3)</td>
</tr>
<tr>
<td>5</td>
<td>361°</td>
<td>431</td>
<td>αNCH(2)</td>
<td>431</td>
<td>αNCH(2)</td>
</tr>
<tr>
<td>6</td>
<td>431°</td>
<td>579</td>
<td>rC5=O</td>
<td>579</td>
<td>rC5=O</td>
</tr>
<tr>
<td>7</td>
<td>628°</td>
<td>652</td>
<td>rC5=C4</td>
<td>652</td>
<td>rC5=C4</td>
</tr>
<tr>
<td>8</td>
<td>718°</td>
<td>776</td>
<td>rC5=C5</td>
<td>776</td>
<td>rC5=C5</td>
</tr>
<tr>
<td>9</td>
<td>812°</td>
<td>797</td>
<td>rC5=C5</td>
<td>797</td>
<td>rC5=C5</td>
</tr>
<tr>
<td>10</td>
<td>973°</td>
<td>949</td>
<td>rC3(3)</td>
<td>949</td>
<td>rC3(3)</td>
</tr>
<tr>
<td>11</td>
<td>1042°</td>
<td>996</td>
<td>rC3(4)</td>
<td>996</td>
<td>rC3(4)</td>
</tr>
<tr>
<td>12</td>
<td>1092°</td>
<td>1056</td>
<td>rC3(3)</td>
<td>1056</td>
<td>rC3(3)</td>
</tr>
<tr>
<td>13</td>
<td>1176°</td>
<td>1087</td>
<td>rC3(2)</td>
<td>1087</td>
<td>rC3(2)</td>
</tr>
<tr>
<td>14</td>
<td>1263°</td>
<td>1095</td>
<td>rC3(6)</td>
<td>1095</td>
<td>rC3(6)</td>
</tr>
<tr>
<td>15</td>
<td>1279°</td>
<td>1267</td>
<td>rC3(9)</td>
<td>1267</td>
<td>rC3(9)</td>
</tr>
<tr>
<td>16</td>
<td>1374°</td>
<td>1384</td>
<td>rC3(9)</td>
<td>1384</td>
<td>rC3(9)</td>
</tr>
<tr>
<td>17</td>
<td>1410°</td>
<td>1413</td>
<td>rC3(9)</td>
<td>1413</td>
<td>rC3(9)</td>
</tr>
<tr>
<td>18</td>
<td>1430°</td>
<td>1416</td>
<td>rC3(8)</td>
<td>1416</td>
<td>rC3(8)</td>
</tr>
<tr>
<td>19</td>
<td>1430°</td>
<td>1418</td>
<td>rC3(9)</td>
<td>1418</td>
<td>rC3(9)</td>
</tr>
<tr>
<td>20</td>
<td>1430°</td>
<td>1426</td>
<td>rC3(8)</td>
<td>1426</td>
<td>rC3(8)</td>
</tr>
<tr>
<td>21</td>
<td>1430°</td>
<td>1481</td>
<td>rC3(9)</td>
<td>1481</td>
<td>rC3(9)</td>
</tr>
<tr>
<td>22</td>
<td>1486°</td>
<td>1587</td>
<td>rC3(9)</td>
<td>1587</td>
<td>rC3(9)</td>
</tr>
<tr>
<td>23</td>
<td>1723°</td>
<td>1683</td>
<td>rC3(9)</td>
<td>1683</td>
<td>rC3(9)</td>
</tr>
<tr>
<td>24</td>
<td>2193°</td>
<td>2114</td>
<td>rC3(9)</td>
<td>2114</td>
<td>rC3(9)</td>
</tr>
<tr>
<td>25</td>
<td>2540°</td>
<td>2531</td>
<td>rC3(9)</td>
<td>2531</td>
<td>rC3(9)</td>
</tr>
<tr>
<td>26</td>
<td>2540°</td>
<td>2534</td>
<td>rC3(9)</td>
<td>2534</td>
<td>rC3(9)</td>
</tr>
<tr>
<td>27</td>
<td>2540°</td>
<td>2540</td>
<td>rC3(9)</td>
<td>2540</td>
<td>rC3(9)</td>
</tr>
<tr>
<td>28</td>
<td>2540°</td>
<td>2540</td>
<td>rC3(9)</td>
<td>2540</td>
<td>rC3(9)</td>
</tr>
<tr>
<td>29</td>
<td>2540°</td>
<td>2540</td>
<td>rC3(9)</td>
<td>2540</td>
<td>rC3(9)</td>
</tr>
<tr>
<td>30</td>
<td>2540°</td>
<td>2540</td>
<td>rC3(9)</td>
<td>2540</td>
<td>rC3(9)</td>
</tr>
</tbody>
</table>

3) Geometries 4) Protein crystal structures


**TABLE 1: Geometric Data on N-Methylacetamide**

<table>
<thead>
<tr>
<th></th>
<th>experimental</th>
<th>MP2/6-31G(d)</th>
<th>H$_2$O.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CHARMM</td>
<td>gas</td>
<td>crystal</td>
</tr>
<tr>
<td>Trans Bonds</td>
<td>C4−C5</td>
<td>1.481</td>
<td>1.520(5)</td>
</tr>
<tr>
<td></td>
<td>C5−N7</td>
<td>1.339</td>
<td>1.386(4)</td>
</tr>
<tr>
<td></td>
<td>N7−C9</td>
<td>1.444</td>
<td>1.469(6)</td>
</tr>
<tr>
<td></td>
<td>C5=N6</td>
<td>1.223</td>
<td>1.225(3)</td>
</tr>
<tr>
<td></td>
<td>N7−H8</td>
<td>0.993</td>
<td>1.010</td>
</tr>
</tbody>
</table>

| Angles | C4−C5−N7 | 116.4 | 114.1(15) | 116.3(6) | 116(2) | 115.3 | 117.1 | 116.6 |
|        | O6=N5−C7 | 122.6 | 121.8(4) | 121.7(6) | 123(1) | 123.1 | 122.1 | 122.6 |
|        | C4−C5=N6 | 121.0 | 124.1 | 121.9(6) | 121(4) | 121.6 | 120.9 | 120.9 |
|        | C5−N7−C9 | 121.7 | 119.7(8) | 121.3(6) | 122(1) | 122.1 | 121.1 | 121.3 |
|        | C5−N7−H8 | 119.8 | 110.0(50) | 118.9 | 119.9 | 119.5 |

**TABLE 12: Condensed-Phase Calculated and Experimental Data for N-Methylacetamide**

<table>
<thead>
<tr>
<th>Pure Solvent</th>
<th>calculated</th>
<th>experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta H_{\text{vap}}$</td>
<td>mol vol.</td>
<td>$\Delta H_{\text{vap}}$</td>
</tr>
<tr>
<td>13.85 ± 0.02</td>
<td>133.7 ± 0.2</td>
<td>14.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aqueous Solvent$^{a,5}$</th>
<th>$\Delta H_{\text{vap}}$</th>
<th>$\Delta H_{\text{vap}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>mol vol.</td>
<td>mol vol.</td>
<td></td>
</tr>
<tr>
<td>0.88(-19.4)</td>
<td>75(65)</td>
<td>19.2</td>
</tr>
</tbody>
</table>

**TABLE 21: Overall Protein Crystal Simulation Results$^a$**

<table>
<thead>
<tr>
<th>property</th>
<th>expt</th>
<th>crystal</th>
<th>vacuum</th>
</tr>
</thead>
<tbody>
<tr>
<td>internal pressure</td>
<td>1</td>
<td>1254 ± 1659</td>
<td>286 ± 7</td>
</tr>
<tr>
<td>temp</td>
<td>room</td>
<td>304 ± 7</td>
<td>328.4 ± 0.1</td>
</tr>
<tr>
<td>total energy</td>
<td>$-810.2 ± 0.4$</td>
<td>$-910.9 ± 0.4$</td>
<td></td>
</tr>
<tr>
<td>rms difference</td>
<td>backbone$^a$</td>
<td>0.63</td>
<td>1.70</td>
</tr>
<tr>
<td></td>
<td>non-hydrogen$^a$</td>
<td>1.01</td>
<td>2.16</td>
</tr>
</tbody>
</table>

**TABLE 21: Overall Protein Crystal Simulation Results$^a$**

<table>
<thead>
<tr>
<th>property</th>
<th>Crm</th>
<th>BPTI</th>
<th>MBCO</th>
</tr>
</thead>
<tbody>
<tr>
<td>internal pressure</td>
<td>1</td>
<td>201 ± 1626</td>
<td>298 ± 8</td>
</tr>
<tr>
<td>temp</td>
<td>room</td>
<td>201 ± 4</td>
<td>302.9 ± 0.2</td>
</tr>
<tr>
<td>total energy</td>
<td>$-2221.4 ± 0.1$</td>
<td>$-203.4 ± 0.2$</td>
<td></td>
</tr>
<tr>
<td>rms difference</td>
<td>backbone$^a$</td>
<td>0.82</td>
<td>2.58</td>
</tr>
<tr>
<td></td>
<td>non-hydrogen$^a$</td>
<td>0.96</td>
<td>3.19</td>
</tr>
<tr>
<td></td>
<td>side chain$^a$</td>
<td>1.99</td>
<td>3.73</td>
</tr>
<tr>
<td></td>
<td>non-hydrogen$^a$</td>
<td>0.86</td>
<td>2.63</td>
</tr>
<tr>
<td>radius of gyration</td>
<td>backbone</td>
<td>10.807</td>
<td>10.838</td>
</tr>
<tr>
<td></td>
<td>non-hydrogen</td>
<td>10.944</td>
<td>11.222</td>
</tr>
<tr>
<td>radius of gyration</td>
<td>backbone</td>
<td>0.71</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>non-hydrogen</td>
<td>0.70</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>side chain</td>
<td>0.80</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>non-hydrogen</td>
<td>0.75</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>backbone</td>
<td>0.66</td>
<td>1.98</td>
</tr>
<tr>
<td></td>
<td>non-hydrogen</td>
<td>0.97</td>
<td>2.30</td>
</tr>
<tr>
<td></td>
<td>side chain</td>
<td>0.62</td>
<td>2.30</td>
</tr>
<tr>
<td></td>
<td>non-hydrogen</td>
<td>0.63</td>
<td>2.30</td>
</tr>
</tbody>
</table>
Comparisons between different force fields using NMA


Utilizes B3LYP/6-311++G** for comparison to gas phase optimized geometries. One notable exception is the C-N-H angle which simulations does not match gas electron diffraction experiments.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Point charges assigned to each atom in NMA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CHARMM</td>
</tr>
<tr>
<td>C</td>
<td>0.51</td>
</tr>
<tr>
<td>O</td>
<td>-0.51</td>
</tr>
<tr>
<td>N</td>
<td>-0.47</td>
</tr>
<tr>
<td>H</td>
<td>0.31</td>
</tr>
<tr>
<td>C_L</td>
<td>-0.27</td>
</tr>
<tr>
<td>H_L</td>
<td>0.09</td>
</tr>
<tr>
<td>C_R</td>
<td>-0.11</td>
</tr>
<tr>
<td>H_R</td>
<td>0.09</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Gas phase geometry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EXP [30]</td>
</tr>
<tr>
<td>d_{CO}</td>
<td>1.224</td>
</tr>
<tr>
<td>d_{CN}</td>
<td>1.386</td>
</tr>
<tr>
<td>d_{NH}</td>
<td>--</td>
</tr>
<tr>
<td>d_{C_C}</td>
<td>1.520</td>
</tr>
<tr>
<td>d_{NC_R}</td>
<td>1.468</td>
</tr>
<tr>
<td>(\beta)_{OCN}</td>
<td>121.8</td>
</tr>
<tr>
<td>(\beta)_{CNH}</td>
<td>110.0</td>
</tr>
<tr>
<td>(\beta)_{NCC_L}</td>
<td>114.1</td>
</tr>
<tr>
<td>(\beta)_{CNC_R}</td>
<td>119.6</td>
</tr>
<tr>
<td>D_{OCNH}</td>
<td>--</td>
</tr>
</tbody>
</table>

Distances are in Å, and angles in °. Values in bracket are equilibrium values from the force field bonded interaction parameters.
Comparisons of different force fields

Monte Carlo (NPT and GEMC) simulation of small organic molecules using different force fields. Mostly interested in liquid densities and vapor-liquid coexistence curves. Between the four force fields of interest here (AMBER, CHARMM, GROMOS, and OPLS), CHARMM might be the better one to use if interested in phase equilibria and the molecule is not a model compound used for parameterization.
Comparison to crystal structures using ns trajectories


Table 1. Averaged Overall Properties (Standard Deviations in Parenthesis).

<table>
<thead>
<tr>
<th></th>
<th>Total Energy (kcal/mol)</th>
<th>T (°C)</th>
<th>Cα RMSD (Å) to Experiment</th>
<th>Cα RMSD (Å) to Average</th>
<th>r_250 (Å) Experiment</th>
<th>r_90 (Å) Experiment</th>
<th>SASA (Å²) Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calbindin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMBER</td>
<td>31085 (39)</td>
<td>298 (2)</td>
<td>3.02 (0.20)</td>
<td>1.05 (0.16)</td>
<td>11.70 (0.07)</td>
<td>11.40^a</td>
<td>5225 (70)</td>
</tr>
<tr>
<td>CHARMM</td>
<td>31406 (68)</td>
<td>299 (2)</td>
<td>2.76 (0.13)</td>
<td>1.01 (0.18)</td>
<td>11.88 (0.06)</td>
<td>5253 (69)</td>
<td>4761^b</td>
</tr>
<tr>
<td>OPLS</td>
<td>33104 (76)</td>
<td>299 (2)</td>
<td>2.63 (0.17)</td>
<td>0.98 (0.21)</td>
<td>11.70 (0.07)</td>
<td>5017 (97)</td>
<td>4778^b</td>
</tr>
<tr>
<td>IL4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMBER</td>
<td>-68836 (26)</td>
<td>300 (1)</td>
<td>1.14 (0.12)</td>
<td>0.76 (0.09)</td>
<td>14.76 (0.07)</td>
<td>14.55</td>
<td>7763 (168)</td>
</tr>
<tr>
<td>CHARMM 1</td>
<td>70287 (124)</td>
<td>299 (2)</td>
<td>1.59 (0.72)</td>
<td>0.97 (0.17)</td>
<td>15.00 (0.13)</td>
<td>8050 (223)</td>
<td>7027</td>
</tr>
<tr>
<td>CHARMM 2</td>
<td>70019 (60)</td>
<td>300 (2)</td>
<td>1.36 (0.13)</td>
<td>0.89 (0.17)</td>
<td>14.82 (0.06)</td>
<td>7814 (118)</td>
<td>7027</td>
</tr>
<tr>
<td>OPLS</td>
<td>-70670 (43)</td>
<td>300 (1)</td>
<td>1.37 (0.13)</td>
<td>0.75 (0.13)</td>
<td>14.79 (0.06)</td>
<td>7661 (104)</td>
<td>7027</td>
</tr>
<tr>
<td>GPIIA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMBER</td>
<td>-47239 (32)</td>
<td>300 (2)</td>
<td>0.94 (0.10)</td>
<td>0.62 (0.07)</td>
<td>14.71 (0.05)</td>
<td>14.61</td>
<td>7724 (103)</td>
</tr>
<tr>
<td>CHARMM</td>
<td>-47130 (100)</td>
<td>300 (2)</td>
<td>1.20 (0.15)</td>
<td>0.74 (0.11)</td>
<td>14.78 (0.06)</td>
<td>7799 (96)</td>
<td>7458</td>
</tr>
<tr>
<td>OPLS</td>
<td>-50356 (55)</td>
<td>300 (2)</td>
<td>1.25 (0.26)</td>
<td>0.84 (0.12)</td>
<td>14.81 (0.08)</td>
<td>7879 (153)</td>
<td>7496</td>
</tr>
</tbody>
</table>

Comparison of AMBER, CHARMM, and OPLS using 3 proteins. Each simulation are 2 ns in length.

Conclusion that no force field showed any consistent trend in variations and is remarkably close to experimental structure, SASA, R_g, and backbone order parameters.
DNA Clamp: PCNA


DNA polymerase, clamp, and ATP driven clamp loader together help responsible for replication of DNA. The human form of the clamp, proliferating cell nuclear antigen (PCNA), is a ring shaped protein which trimerizes to form a six-domain ring and acts as a clamp.

Kazmirski et al PNAS 2005, 102(39), 13801-13806.

Adelman et al. Ring opening dynamics of the sliding DNA clamp PCNA Poster Biophysical Society Meeting.

Determining the Free Energy of Ring Opening

&FORCE_EVAL
 &MM
 &FORCEFIELD
 PARMTYPE CHM
 PARM_FILE_NAME par_all27_prot_lib.inp
 VDW_SCALE14 1.0
 EI_SCALE14 1.0
 IGNORE_MISSING_CRITICAL_PARAMS F
 &END FORCEFIELD
 &END MM
 &END FORCE_EVAL

&FORCE_EVAL
 &SUBSYS
 &TOPOLOGY
 COORD_FILE_FORMAT PDB
 COORD_FILE_NAME filename.pdb
 CONN_FILE_FORMAT PSF
 CONN_FILE_NAME filename.psf
 &END TOPOLOGY
 &END SUBSYS
 &END FORCE_EVAL
MOL_SET for Monte Carlo

&FORCE_EVAL
&MM
&FORCEFIELD
PARMTYPE CHM
PARM_FILE_NAME par_all27_prot_lib.inp
VDW_SCALE14 1.0
EI_SCALE14 1.0
IGNORE_MISSING_CRITICAL_PARAMS F
&END FORCEFIELD
&END MM
&END FORCE_EVAL

&FORCE_EVAL
&SUBSYS
&TOPOLOGY
COORD_FILE_FORMAT PDB
COORD_FILE_NAME filename.pdb
CONN_FILE_FORMAT MOL_SET
&MOL_SET
&MOLECULE
NMOL 64
CONN_FILE_FORMAT PSF
CONN_FILE_NAME 1_water.psf
&END MOLECULE
&END MOL_SET
&END TOPOLOGY
&END SUBSYS
&END FORCE_EVAL
Generating appropriate PDB and PSF files

Goal: Generate necessary PDB and PSF file for DMPA that can be used by CP2K. There are many ways to read in your systems besides using PDB and PSF files.

```bash
>> cd DMPA/PSFGEN
```

Files in the directory:
1. `command.vmd`
   Script command for PSFGEN
2. `init.pdb`
   Initial structure.
3. `psfgen.x`
   Standalone executable. Often found as plugin with VMD.
4. `topology.rtf`
   CHARMM 27 topology file.

Using the psfgen script “command.vmd”, we can generate PDB/PSF files of DMPA in the gas phase by utilizing standard distribution CHARMM topology files and a valid initial starting structure “init.pdb” by issuing the following command.

```bash
>> ./psfgen.x < command.vmd
```

Two new files generated are “dmpa_only.pdb” and “dmpa_only.psf”. (Note: The generation of these two file can be done through CP2K also via the DUMP_PDB and DUMP_PSF section.) Ultimately, we are interested in solvated species in a charge neutralized system most of the time. To generated a solvated DMPA with appropriate counter ions, you will need VMD for the next step.

Easy way to load both PDB and PSF from VMD command line is:

```bash
vmd >> mol load psf dmpa_only.psf pdb dmpa_only.pdb
```
Quick ways to add solvents and counter ions

Solvating DMPA using the “Add Solvation Box” plugin under the menu Extensions:Modeling.

Now DMPA is solvated in a 20Åx20Åx20Å box of water. Information saved under “dmpa_wat”.

DMPA has a net -1 charge. Next, we will convert one water into Na⁺ to neutralize the system. Use the “Autoionize” plugin under the menu Extensions:Modeling.

Procedure outline here can be applied to any biological system in which case, the “init.pdb” can be obtained from the protein data bank such as www.rcsb.org.
Simulation of DMPA with FIST

Goal: Perform a stable MD simulation of DMPA with FIST. The potential chosen for this demo is the CHARMM27 force field.

>> cd DMPA/FIST

Files in the directory:
1. par_all27_prot_lipid.inp
   CHARMM 27 all atom force field
2. dmpa_wat_sod.psf/dmpa_wat_sod.pdb
   Initial structure and topology file
3. minimize.inp
   Geometry minimization to remove overlap
4. md.inp
   Molecular dynamics simulation input.

A typical &FORCE_EVAL section for FIST:

&FORCE_EVAL
METHOD FIST
&MM
&FORCEFIELD
  par_file_name par_all27_prot_lipid.inp
  parmtype CHARMM
&&SPLINE
  EMAX_SPLINE 1.0E8
  RCUT_NB 10.0
&END SPLINE
&END FORCEFIELD
&POISSON
&EWALD
  EWALD_TYPE spme
  ALPHA .35
  GMX 2121
&END EWALD
&END POISSON
&END MM
&SUBSYS
&CELL
  ABC 20.0 20.0 20.0
&END CELL
&TOPOLOGY
  CONN_FILE_FORMAT PSF
  COORD_FILE_FORMAT PDB
  CONN_FILE_NAME dmpa_wat_sod.pdb
&END TOPOLOGY
&END SUBSYS
&END FORCE_EVAL

Overcome initial close overlap problem.

\[ \alpha \Theta \frac{7}{2} r_{cut}^{-1} \]
Simulation of DMPA with FIST cont...

Generation of the initial solvation box resulted in some close overlaps for water at the edge of the simulation box. Therefore, the system must be minimized in order to prevent presence of large force in MD simulation. To do the minimization, we will use the input file “minimize.inp” with the following command.

```
>> cp2k.x minimize.inp
```

This is what a barebone representation of what a minimization input looks like in the &MOTION section.

```
&MOTION
&MD
  ENSEMBLE NVT
  STEPS 1000
  TIMESTEP 0.48
  TEMPERATURE 298.0
&MOTION
  &GEO_OPT
    OPTIMIZER CG
    MAX_ITER 50
  &END GEO_OPT
&END MOTION
```

Total minimization is not necessary and is counter productive!!

Now we can start the MD simulation with the following command. The simulation is carried out using NHC in the NVT ensemble.

```
>> cp2k.x md.inp
```

```
&MOTION
&MD
  ENSEMBLE NVT
  STEPS 1000
  TIMESTEP 0.48
  TEMPERATURE 298.0
&MOTION
  &NOSE
    REGION MASSIVE
    NOSE
    TIMECON [wavenumber_0] 1000
  &END NOSE
&MOTION
  &END THERMOSTAT
&END MD
&MOTION
  &EXT_RESTART
    RESTART_FILE_NAME dmpa_minimize.restart
    RESTART_DEFAULT F
    RESTART_POS T
  &END EXT_RESTART
```

Usually a good idea.

Only want to use the minimized positions.
Example of this procedure using VMD to perform the psf generation and solvation are on the wiki for the workshop.

Generation of the initial solvation box resulted in some close overlaps for water at the edge of the simulation box. Therefore, the system must be minimized in order to prevent presence of large force in MD simulation. To do the minimization, we will use the input file “minimize.inp” with the following command.

>> cp2k.x minimize.inp

This is what a barebone representation of what a minimization input looks like in the &MOTION section.

&MOTION
&MD
ENSEMBLE NVT
STEPS 1000
TIMESTEP 0.48
TEMPERATURE 298.0
&THERMOSTAT
TYPE NOSE
REGION MASSIVE
&NOSE
TIMECON [wavenumber_1] 1000
&END NOSE
&END THERMOSTAT
&END MD
&MOTION
&EXT_RESTART
RESTART_FILE_NAME dmpa_minimize.restart
RESTART_DEFAULT F
RESTART_POS T
&END EXT_RESTART

Now we can start the MD simulation with the following command. The simulation is carried out using NHC in the NVT ensemble.

>> cp2k.x md.inp

Usually a good idea.

Only want to use the minimized positions.

Total minimization is not necessary and is counter productive!!
FIST Run diagnostic: deca-alanine
Solid state forcefields also implemented

- Williams (Buckingham)
- Shell model
- Embedded Atom Method (EAM)
- General potentials supported
- QUIP library (http://www.libatoms.org/Home/Software) – provides extra functionality

Forcefield module is actually very flexible
Conclusions

1. Keep in mind the origin of the parameters. Is the parameters applicable to the type of system that you are interested in?
2. Classical force fields can sometimes predict properties very close to experiments but sometimes can be off.
3. The class I force fields discussed here are all very similar to each other. Therefore use the one that you’re most comfortable with.