UNIVERSITY OF
LINCOLN

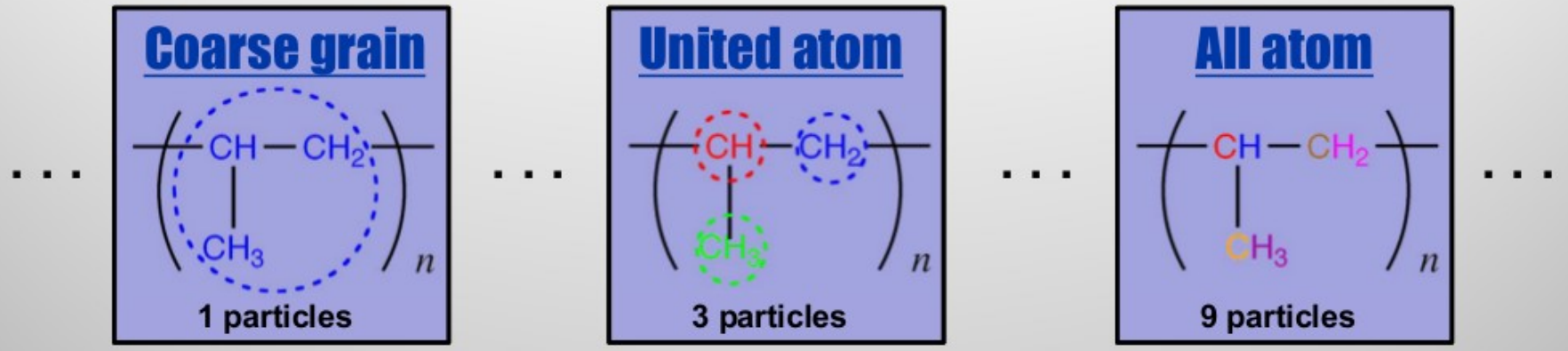
Force Fields in CP2K

Matt Watkins
mwatkins@lincoln.ac.uk

UNIVERSITY OF
LINCOLN

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.



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

10,000 particles

$$\frac{\left(1\times10^4\right)^2}{\left(1\times10^4\right)^2}=1$$

~10^8 steps

30,000 particles

$$\frac{\left(3\times10^4\right)^2}{\left(1\times10^4\right)^2}=9$$

~10^7 steps

90,000 particles

$$\frac{\left(9\times10^4\right)^2}{\left(1\times10^4\right)^2}=81$$

~10^6 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.

$$E_{total} = E_{nonbonded} + E_{bonded}$$

$$E_{nonbonded} = E_{electrostatics} + E_{vdW}$$

$$E_{bonded} = E_{bond} + E_{bend} + E_{dihedral} + E_{cross}$$

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…

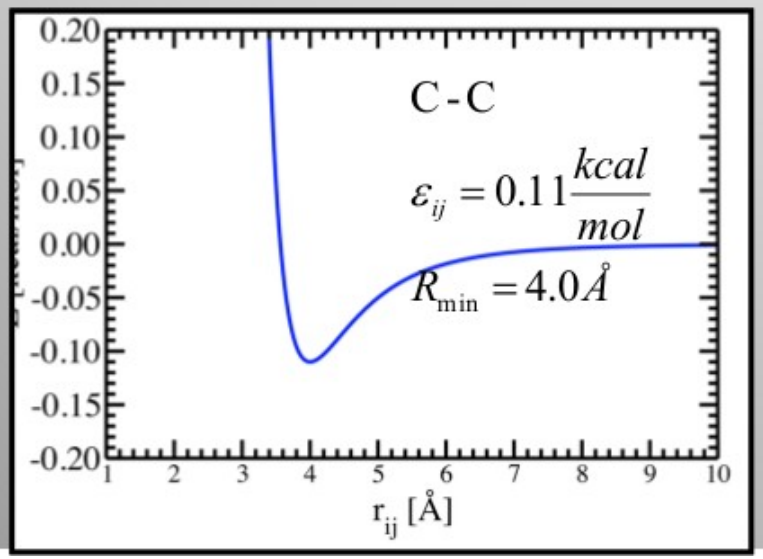
Currently, CP2K only have support for Class I type force fields such as AMBER, CHARMM, GROMOS, and OPLS.

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 :

%FORCE_EVAL%MM%FORCEFIELD%NONBONDED

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



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:

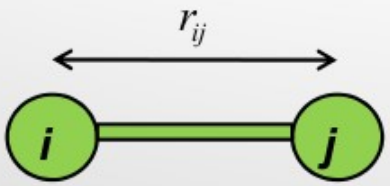
```
%FORCE_EVAL%MM%FORCEFIELD%NONBONDED%GENPOT
   ATOMS O O
   FUNCTION a1/(r**a2) -  a3*EXP(-a4*(r-a5)**2)-a6*EXP(-a7*(r-a8)**2)
   PARAMETERS   a1 a2 a3 a4 a5 a6 a7 a8
   VARIABLES    r
```

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

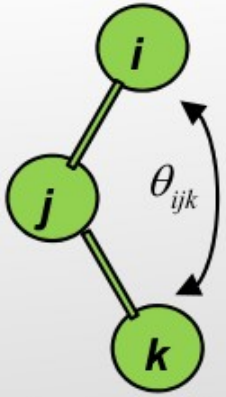
$$V(r_{ij}) = \frac{q_i q_j}{\varepsilon r_{ij}}$$

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.

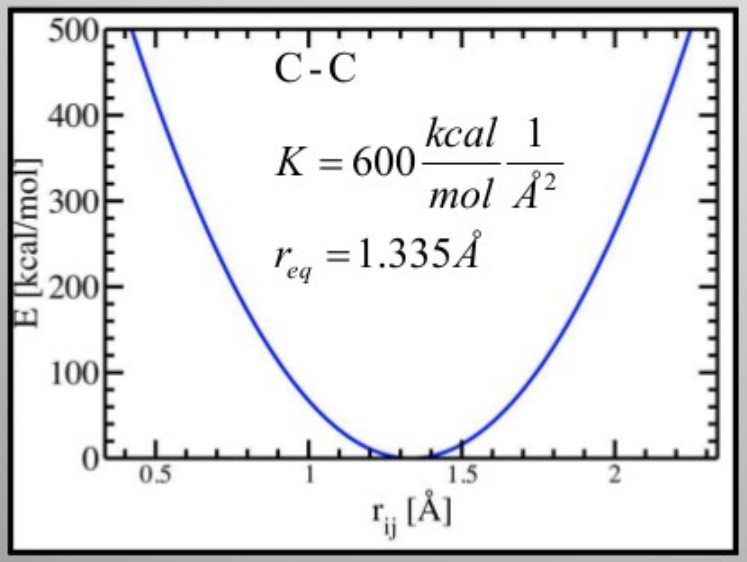
Bonds and Bends

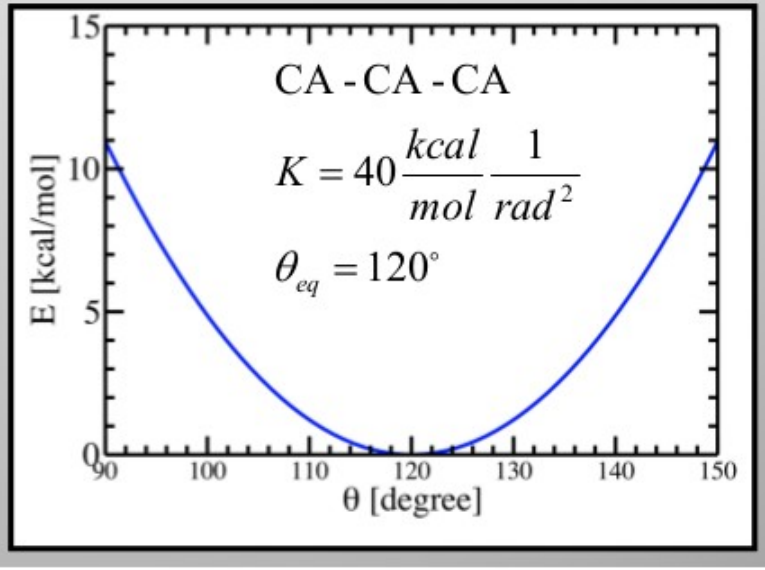


$$V(r_{ij}) = K(r_{ij} - r_{eq})^2$$

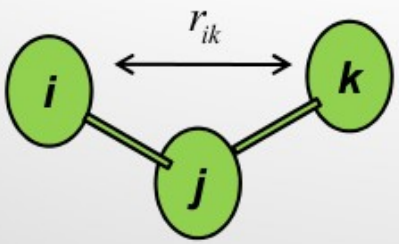


$$V(\theta_{ijk}) = K(\theta_{ijk} - \theta_{eq})^2$$

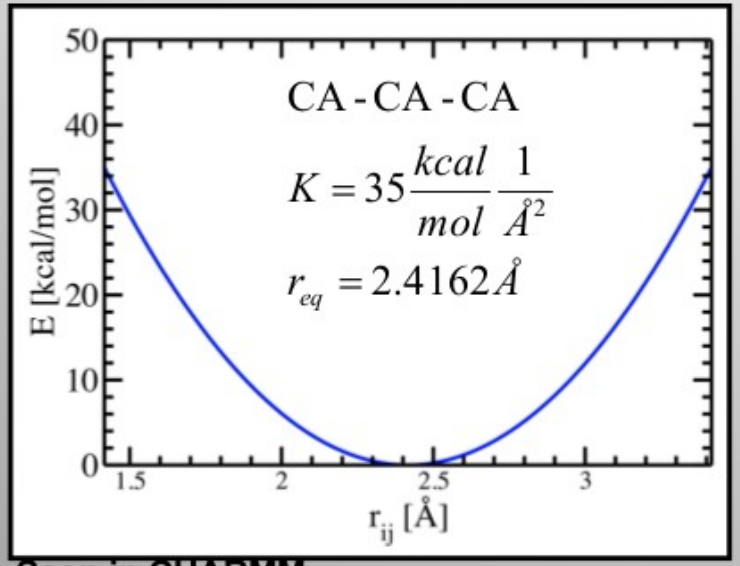




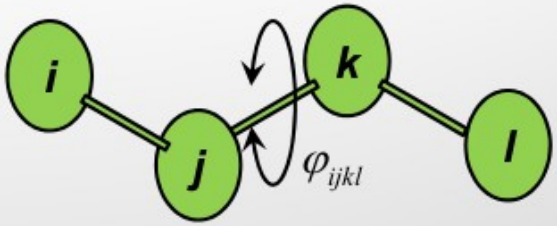
Urey-Bradley and Dihedral



$$V_{UB}(r_{ik}) = K_{UB}(r_{ik} - r_{eq})^2$$



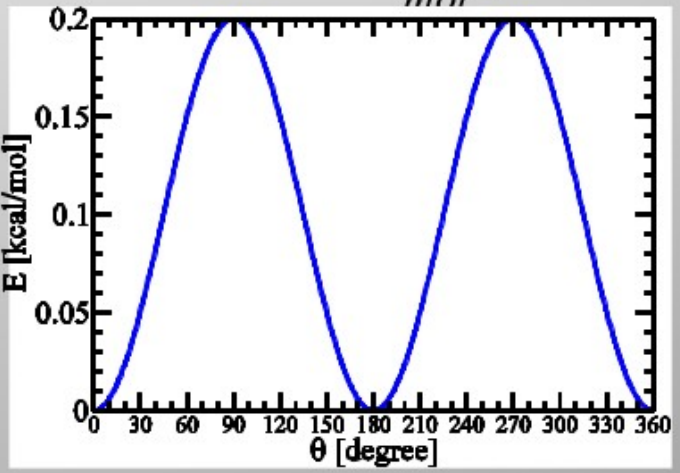
Seen in CHARMM.



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

CTL2 – CTL2 - CTL2 - CLT2

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

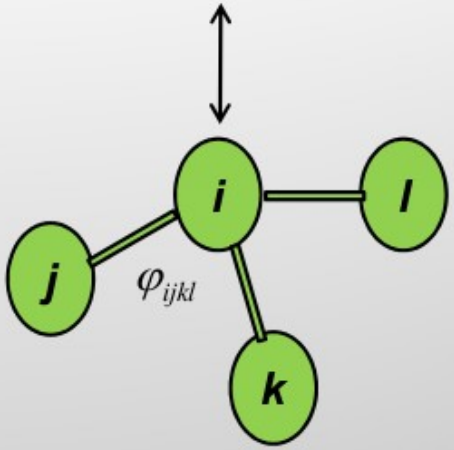


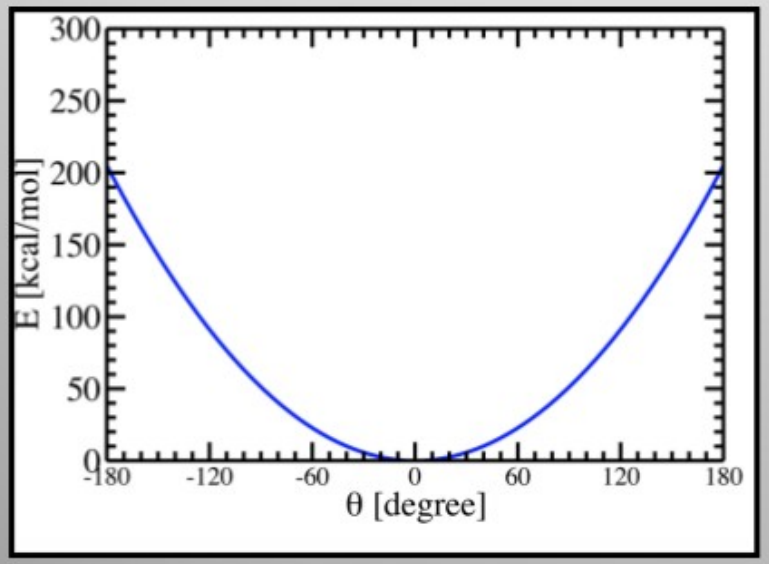
Improper(Out of plane motion)

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

CPB – CPA – NPH – CPA

$$K = 20.8 \frac{kcal}{mol} \frac{1}{rad^2}$$





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.

AMBER

J Comput Chem 24: 1999-2012, 2003.

$$V_{total} = \sum_{bonds} K_b \left(b - b_{eq}\right)^2 +$$

$$\sum_{angle} K_\theta \left(\theta - \theta_{eq}\right)^2 +$$

$$\sum_{dihedrals} \frac{V_n}{2}\left[1 + \cos\left(n\phi - \gamma\right)\right] +$$

$$\sum_{i<j}\left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} - \frac{q_i q_j}{\varepsilon R_{ij}}\right]$$

$$\gamma = 0^\circ \text{ or } 180^\circ$$

A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations

YONG DUAN,[1] CHUN WU,[1] SHIBASISH CHOWDHURY,[1] MATHEW C. LEE,[1] GUOMING XIONG,[1] WEI ZHANG,[1] RONG YANG,[1] PIOTR CIEPLAK,[2,3] RAY LUO,[2] TAISUNG LEE,[2,3] JAMES CALDWELL,[2] JUNMEI WANG,[2] PETER KOLLMAN[2,†]

[1]*Department of Chemistry and Biochemistry, University of Delaware, Newark, Delaware 19716*

[2]*Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94143*

[3]*Accelrys Inc., 9685 Scranton Rd, San Diego, California 92121*

Received 7 April 2003; Accepted 1 July 2003

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

CHARMM

J Phys Chem B **1998**, *102*, 3586-3616.

$$V_{total} = \sum_{bonds} K_b\left(b - b_{eq}\right)^2 + \sum_{UB} K_{UB}\left(S - S_{eq}\right)^2 + \sum_{angle} K_\theta\left(\theta - \theta_{eq}\right)^2 + \sum_{dihedrals} K_\phi\left[1 + \cos\left(n\phi - \gamma\right)\right] + \sum_{impropers} K_{imp}\left(\varphi - \varphi_{eq}\right)^2 + \sum_{i<j} \varepsilon\left[\left(\frac{R_{\min_{ij}}}{r_{ij}}\right)^{12} - \left(\frac{R_{\min_{ij}}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{r_{ij}}$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \qquad R_{\min ij} = \tfrac{1}{2}\left(R_{\min_i} + R_{\min_j}\right)$$

All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins†

A. D. MacKerell, Jr.,*,‡,§ D. Bashford,‡,⊥ M. Bellott,‡,⊥ R. L. Dunbrack, Jr.,‡,⊥ J. D. Evanseck,‡,⊥ M. J. Field,‡,⊥ S. Fischer,‡,⊥ J. Gao,‡,⊥ H. Guo,‡,⊥ S. Ha,‡,⊥ D. Joseph-McCarthy,‡,⊥ L. Kuchnir,‡,⊥ K. Kuczera,‡,⊥ F. T. K. Lau,‡,⊥ C. Mattos,‡ S. Michnick,‡,⊥ T. Ngo,‡,⊥ D. T. Nguyen,‡,⊥ B. Prodhom,‡,⊥ W. E. Reiher, III,‡,⊥ B. Roux,‡,⊥ M. Schlenkrich,‡,⊥ J. C. Smith,‡,⊥ R. Stote,‡,⊥ J. Straub,‡,⊥ M. Watanabe,‡,⊥ J. Wiórkiewicz-Kuczera,‡,⊥ D. Yin,§ and M. Karplus*,‡,∥

Department of Chemistry & Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, Department of Pharmaceutical Sciences, University of Maryland, School of Pharmacy, Baltimore, Maryland 21201, and Laboratoire de Chimie Biophysique, ISIS, Institut Le Bel, Université Louis Pasteur, 67000 Strasbourg, France

Received: September 22, 1997; In Final Form: February 6, 1998

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."

OPLS

J Am Chem Soc **1988**, *110*, 1657-1666.
J Am Chem Soc **1996**, *118*, 11225-11236.

$$V_{total} = \sum_{bonds} K_r\left(r - r_{eq}\right)^2 + \sum_{angle} K_\theta\left(\theta - \theta_{eq}\right)^2 + \sum_{dihedrals} \frac{V_1}{2}\left[1 + \cos\left(\phi - \gamma_1\right)\right] + \frac{V_2}{2}\left[1 + \cos\left(2\phi - \gamma_2\right)\right] + \frac{V_3}{2}\left[1 + \cos\left(3\phi - \gamma_3\right)\right] + \sum_{i<j}\left[4\varepsilon_{ij}\left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}}\right) - \frac{q_i q_j e^2}{r_{ij}}\right] f_{ij}$$

The OPLS Potential Functions for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin

William L. Jorgensen* and Julian Tirado-Rives

Contribution from the Department of Chemistry, Purdue University, West Lafayette, Indiana 47907. *Received January 26, 1987*

Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids

William L. Jorgensen,* David S. Maxwell, and Julian Tirado-Rives

Contribution from the Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107

Received June 27, 1996. Revised Manuscript Received September 5, 1996®

$$\gamma_1 = \gamma_2 = \gamma_3 = 0^\circ \qquad \sigma_{ij} = \sqrt{\sigma_{ii}\sigma_{jj}} \qquad f_{ij} = 0.5 \text{ for 1-4 interactions}$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \qquad f_{ij} = 1.0 \text{ for everything else}$$

An Example of Class II force field: MMFF

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

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

$$\sum_{angle,linear} K_{al}\left(1 + \cos\left(\theta\right)\right) +$$

$$\sum_{stretch,bend}\left(K_{ijk}\left(r_{ij} - r_{eq}\right) + K_{kji}\left(r_{kj} - r_{eq}\right)\right)\left(\theta - \theta_{eq}\right) +$$

$$\sum_{outofplane} K_{OOP}\left(\chi\right)^2 +$$

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

$$\sum_{i<j}\left[\varepsilon_{ij}\left(\frac{1.07\sigma}{r_{ij} + 0.07\sigma}\right)^7\left(\frac{1.12\sigma^7}{r_{ij}^7 + 0.07\sigma^7} - 2\right) - \frac{q_i q_j}{D\left(r_{ij} + \delta\right)}\right]$$

Tersoff

Phys Rev B **1988**, *38*, 9902-9905.

$$E = \sum_i E_i = \tfrac{1}{2}\sum_{i \neq j} V_{ij}$$

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

$$f_R(r) = A\exp(-\lambda_1 r)$$

$$f_A(r) = -B\exp(-\lambda_2 r)$$

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

$$b_{ij} = \left(1 + \beta^n \zeta_{ij}^n\right)^{-1/2n}$$

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

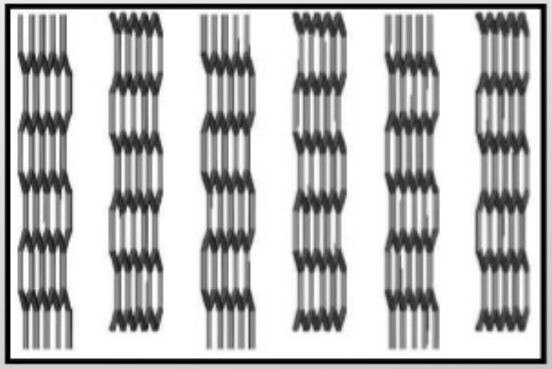
$$g(\theta) = 1 + \frac{c^2}{d^2} - \frac{c^2}{d^2 + (h + \cos\theta)^2}$$

$$a_{ij} = \left(1 + \alpha^n \eta_{ij}^n\right)^{-1/2n}$$

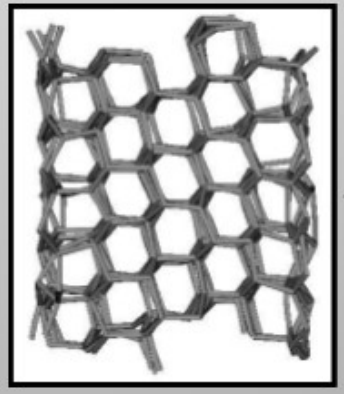
$$\eta_{ij} = \sum_{k(\neq i,j)} f_C(r_{ik}) \exp\left[\lambda_3^3 (r_{ij} - r_{ik})^3\right]$$

Empirical interatomic potential for silicon with improved elastic properties

J. Tersoff
IBM Research Division, Thomas J. Watson Research Center, Yorktown Heights, New York 10598
(Received 21 May 1988)



Shock Compression

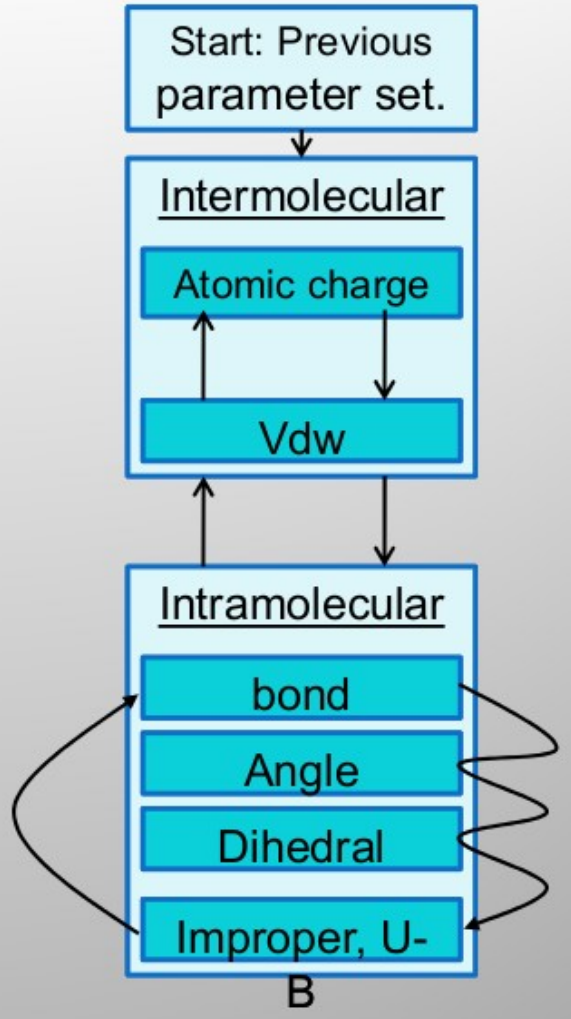


Parameterization of force fields

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 !          /  \
                        !     (-) O4   O2
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 compounds usually contains a motif that is often found repeatedly in biological system. One of the compound 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.

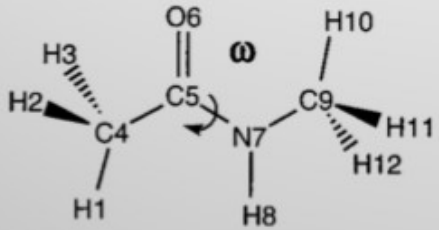


Table 1
Point charges assigned to each atom in NMA

	CHARMM	AMBER	OPLS-AA	GROMOS
C	0.51	0.5869	0.50	0.38
O	−0.51	−0.5911	−0.50	−0.38
N	−0.47	−0.4192	−0.50	−0.28
H	0.31	0.2823	0.30	0.28
C_L	−0.27	−0.0411	−0.18	0.00
H_L	0.09	0.0173	0.06	N/A
C_R	−0.11	−0.2078	0.02	0.00
H_R	0.09	0.1127	0.06	N/A

Li, Yu, Zhuang, and Mukamel Chem Phys Lett 452 (2008) 78-83.

Mackerell et al J Phys Chem B **1998**, *102*, 3586-3616.

TABLE 3: Vibrational Data for *N*-Methylacetamide[a]

	experimental/ab initio[b]		CHARMM	
mode	frequency	assignment	frequency	assignment
1	VLF		64	τCCH3(101)
2	VLF		89	τNCH3(1001)
3	171[c]	ωN7H[d]	200	τC5−N7(107)
				βCNC(62)
				βCCN(25)
				βCCN(50)
				βC5═O(50)
				νC5−C4(29)
				ωC5═O(67)
				ωN7H(30)
				νC5−N7(34)
				νC5═O(20)
				ωN7H(66)
				τCH3(15)
				τCH3(36)
				νN7−C9(34)
				τCH3(47)
				νN7−C9(26)
				τCH3(83)
				τCH3 (72)
				τCH3(67)
				ωC5═O(17)
				βN7H(44)
				νC5−C4(24)
				δCH3s(94)
				δCH3as(89)
				δCH3as(88)
				δCH3as(91)
				δCH3as(87)
				τCH3(15)
				δCH3s(50)
				βN7H(21)
				δCH3s(39)
				βN7H(20)
				νN7−C9(17)
				νC5═O(66)
				νCH3s(100)
				νCH3as(100)
				νCH3as(100)
				νCH3s(100)
				νCH3as(100)
				νCH3as(100)

TABLE 3: Vibrational Data for *N*-Methylacetamide[a]

mode	experimental/ab initio[b] frequency	assignment	CHARMM frequency	assignment
1	VLF		64	τCCH3(101)
2	VLF		89	τNCH3(1001)
3	171[c]	ωN7H[d] τC5−N7	200	τC5−N7(107)
4	279	βCNC βCCN	271	βCNC(62) βCCN(25)
5	391	τC5−N7 ωN7H[d]	431	βCCN(50)
6	431	βCCN βC5═O	579	βC5═O(50) νC5−C4(29)
7	628	βC5═O νC5−C4	652	ωC5═O(67) ωN7H(30)
8	718[c]	βC5═O τCH3	776	νC5−N7(34) νC5═O(20)
9	812	νC5−N7 τCH3 νC5−C4	797	ωN7H(66) τCH3(15)
10	973	τCH3 νN7−C9 νC5−C4	949	τCH3(36) νN7−C9(34)
11	1042	τCH3 βC5═O	996	τCH3(47) νN7−C9(26)
12	1092	νN7−C9 τCH3	1056	τCH3(83)
13	1176	τCH3	1087	τCH3 (72)
14	1263	νN7−C9 βC5═O βN7H	1093	τCH3(67) ωC5═O(17)
15	1279[c]	τCH3	1267	βN7H(44) νC5−C4(24)
16	1374	δCH3s	1384	δCH3s(94)
17	1410	δCH3s	1413	δCH3as(89)
18	1430	δCH3as	1416	δCH3as(88)
19	1430	δCH3as	1418	δCH3as(91)
20	1430	δCH3as	1426	δCH3as(87) τCH3(15)
21	1430	δCH3as	1481	δCH3s(50) βN7H(21)
22	1494	βN7H βN7−C9	1587	δCH3s(39) βN7H(20) νN7−C9(17)
23	1723	νC5═O	1683	νC5═O(66)
24	2830	νCH3s	2852	νCH3s(100)
25	2830	νCH3s	2914	νCH3as(100)
26	2940	νCH3as	2915	νCH3as(100)
27	2940	νCH3as	2917	νCH3s(100)
28	2940	νCH3as	2975	νCH3as(100)
29	2940	νCH3as	2975	νCH3as(100)
30	3495	νN7H	3326	νN7H(99)

… 3) Geometries 4) Protein crystal structures

Mackerell et al J Phys Chem B **1998**, *102*, 3586-3616.

TABLE 1: Geometric Data on *N*-Methylacetamide[a]

		experimental			MP2/6-31G(d)[b]		
	CHARMM	gas[c]	crystal[d]	survey[e]	gas	$3H_2O$	H_2O, 2FM
			Trans Bonds				
C4−C5	1.481	1.520(5)	1.515(3)	1.52(1)	1.514	1.510	1.512
C5−N7	1.339	1.386(4)	1.325(3)	1.33(1)	1.365	1.339	1.337
N7−C9	1.444	1.469(6)	1.454(3)	1.45(2)	1.448	1.454	1.454
C5=O6	1.223	1.225(3)	1.246(2)	1.23(1)	1.232	1.255	1.254
N7−H8	0.993				1.010	1.018	1.017
			Angles				
C4−C5−N7	116.4	114.1(15)	116.3(6)	116(2)	115.3	117.1	116.6
O6=C5−N7	122.6	121.8(4)	121.7(6)	123(1)	123.1	122.1	122.6
C4−C5=O6	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[a]

Pure Solvent			
calculated		experimental	
ΔH_{vap}	mol vol.	ΔH_{vap}	mol vol.
13.85 ± 0.02	133.7 ± 0.2	14.2	135.9
Aqueous Solvent[b,c]			
ΔH_{solv}	mol vol.	ΔH_{solv}	mol vol.
−18.8(−19.4)	75(65)	−19.2	122.7

TABLE 21: Overall Protein Crystal Simulation Results[a]

property	exptl	crystal	vacuum
	Crambin		
internal pressure	1	1254 ± 1659	
temp	room	304 ± 7	286 ± 7
total energy		−810.2 ± 0.4	328.4 ± 0.1
rms difference			
backbone[d]		0.63	1.70
side chain[d]		0.94	2.16
non-hydrogen[d]		0.76	1.91
radius of gyration			
backbone	9.594	9.564	9.469
non-hydrogen	9.667	9.644	9.513
rms fluctuations			
C_α	0.46	0.32	0.50
backbone	0.47	0.34	0.51
side chain	0.55	0.45	0.68
non-hydrogen	0.50	0.39	0.58
	BPTI		
internal pressure	1	−2010 ± 1362	
temp	room	287 ± 6	295 ± 8
total energy		−2221.4 ± 0.1	−502.0 ± 0.2
rms difference			
C_α[c]		0.86	2.63
backbone[d]		0.82	2.58
side chain[d]		1.09	3.73
non-hydrogen[d]		0.96	3.19
radius of gyration			
backbone	10.607	10.838	10.348
non-hydrogen	10.944	11.222	10.562
rms fluctuations			
C_α	0.71	0.37	0.46
backbone	0.70	0.39	0.47
side chain	0.80	0.53	0.62
non-hydrogen	0.75	0.46	0.54
	MBCO		
internal pressure	1	−357 ± 828	
temp	260	268 ± 4	297 ± 5
total energy		−5331.7 ± 0.4	−173.4 ± 0.4
rms difference			
C_α[c]			1.98
backbone[d]		0.72	1.97
side chain[d]		1.16	2.59
non-hydrogen[d]		0.97	2.30
radius of gyration			
backbone	15.052	15.242	15.178
non-hydrogen[b]	15.047	15.279	15.139
rms fluctuations			
C_α	0.56	0.37	0.48
backbone	0.55	0.39	0.49
side chain	0.62	0.54	0.63
non-hydrogen	0.59	0.46	0.56

TABLE 21: Overall Protein Crystal Simulation Results[a]

property	exptl	crystal	vacuum
	Crambin		
internal pressure	1	1254 ± 1659	
temp	room	304 ± 7	286 ± 7
total energy		−810.2 ± 0.4	328.4 ± 0.1
rms difference			
backbone[d]		0.63	1.70
side chain[d]		0.94	2.16
non-hydrogen[d]		0.76	1.91
radius of gyration			
backbone	9.594	9.564	9.469
non-hydrogen	9.667	9.644	9.513
rms fluctuations			
C_α	0.46	0.32	0.50
backbone	0.47	0.34	0.51
side chain	0.55	0.45	0.68
non-hydrogen	0.50	0.39	0.58
	BPTI		
internal pressure	1	−2010 ± 1362	
temp	room	287 ± 6	295 ± 8
total energy		−2221.4 ± 0.1	−502.0 ± 0.2
rms difference			
C_α[c]		0.86	2.63
backbone[d]		0.82	2.58
side chain[d]		1.09	3.73
non-hydrogen[d]		0.96	3.19
radius of gyration			
backbone	10.607	10.838	10.348
non-hydrogen	10.944	11.222	10.562
rms fluctuations			
C_α	0.71	0.37	0.46
backbone	0.70	0.39	0.47
side chain	0.80	0.53	0.62
non-hydrogen	0.75	0.46	0.54
	MBCO		
internal pressure	1	−357 ± 828	
temp	260	268 ± 4	297 ± 5
total energy		−5331.7 ± 0.4	−173.4 ± 0.4
rms difference			
C_α[c]			1.98
backbone[d]		0.72	1.97
side chain[d]		1.16	2.59
non-hydrogen[d]		0.97	2.30
radius of gyration			
backbone	15.052	15.242	15.178
non-hydrogen[b]	15.047	15.279	15.139
rms fluctuations			
C_α	0.56	0.37	0.48
backbone	0.55	0.39	0.49
side chain	0.62	0.54	0.63
non-hydrogen	0.59	0.46	0.56

Comparisons between different force fields using NMA

Li, Yu, Zhuang, and Mukamel Chem Phys Lett 452 (2008) 78-83.

Table 1
Point charges assigned to each atom in NMA

	CHARMM	AMBER	OPLS-AA	GROMOS
C	0.51	0.5869	0.50	0.38
O	−0.51	−0.5911	−0.50	−0.38
N	−0.47	−0.4192	−0.50	−0.28
H	0.31	0.2823	0.30	0.28
C_L	−0.27	−0.0411	−0.18	0.00
H_L	0.09	0.0173	0.06	N/A
C_R	−0.11	−0.2078	0.02	0.00
H_R	0.09	0.1127	0.06	N/A

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 2
Gas phase geometry

	EXP [30]	B3LYP	CHARMM	AMBER	OPLS-AA	GROMOS
d_{CO}	1.224	1.221	1.223 (1.230)	1.227 (1.229)	1.228 (1.229)	1.231 (1.230)
d_{CN}	1.386	1.367	1.339 (1.345)	1.339 (1.335)	1.338 (1.335)	1.325 (1.330)
d_{NH}	–	1.007	0.993 (0.997)	1.008 (1.010)	1.010 (1.010)	0.992 (1.000)
d_{C_LC}	1.520	1.519	1.481 (1.490)	1.524 (1.522)	1.528 (1.522)	1.535 (1.530)
d_{NC_R}	1.468	1.454	1.444 (1.430)	1.464 (1.449)	1.456 (1.449)	1.474 (1.470)
∠OCN	121.8	122.9	122.6 (122.5)	122.6 (122.9)	123.0 (122.9)	125.9 (124.0)
∠CNH	110.0	118.4	119.8 (123.0)	117.7 (120.0)	118.4 (119.8)	121.6 (123.0)
$\angle NCC_L$	114.1	115.5	116.4 (116.5)	116.8 (116.6)	117.0 (116.6)	114.3 (115.0)
$\angle CNC_R$	119.6	123.0	121.7 (120.0)	124.0 (121.9)	123.3 (121.9)	119.5 (117.0)
D_{OCNH}	–	180.0	180.0 (180.0)	180.0 (180.0)	180.0 (180.0)	180.0 (180.0)

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

Martin Fluid Phase Equilibria 248 (2006) 50-55.

Table 4
Liquid densities at 101.25 kPa for butanamide (393 K), propanamide (383 K), ethanethiol (298 K), and methylpropylsulfide (298 K)

Force field	Butanamide (g/ml)	Propanamide (g/ml)	Ethanethiol (g/ml)	Methylpropylsulfide (g/ml)
AMBER	0.904_4	0.943_7	0.814_8	0.815_6
CHARMM	0.857_{14}	0.896_5	0.864_2	0.837_3
GROMOS	0.924_4	0.947_3	0.812_2	0.811_1
OPLS	0.9068	0.941_4	0.857_2	0.826_6
TraPPE	N/A	N/A	0.808_2	N/A
UFF	0.719_7	0.678_7	0.824_4	0.855_3
Experimental	0.885	0.926	0.833	0.838

The subscripts show the standard deviation in the last digit.

Table 5
Observed frequency (expressed as a percentage) with which the force fields reproduce experimental data within one standard deviation

Force field	Phase	1%	2%	5%	10%	50%
AMBER	Liquid	36	57	89	96	100
	Vapor	42	42	46	46	79
CHARMM	Liquid	68	82	96	100	100
	Vapor	38	38	42	58	96
COMPASS	Liquid	13	21	71	92	100
	Vapor	8	8	8	13	67
GROMOS	Liquid	32	50	79	82	93
	Vapor	17	21	46	50	67
OPLS	Liquid	46	57	82	100	100
	Vapor	8	13	13	25	54
TraPPE	Liquid	88	88	100	100	100
	Vapor	17	17	21	42	96
UFF	Liquid	4	7	7	21	96
	Vapor	0	0	0	0	4

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

Price and Brooks J Comput Chem 23: 1045-1057, 2002.

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

		Total Energy (kcal/mol)	T (°C)	C_α RMSD (Å) to Experiment	C_α-RMSD (Å) to Average	r_{gyr} (Å)	r_{gyr} (Å) Experiment	SASA (Å^2)	SASA (Å^2) Experiment[a]
Calbindin									
	AMBER	−31085 (39)	298 (2)	3.02 (0.20)	1.05 (0.16)	11.70 (0.07)	11.42[b]	5225 (70)	4763[b]
	CHARMM	−31406 (68)	299 (2)	2.76 (0.13)	1.01 (0.18)	11.88 (0.08)		5253 (99)	4761[b]
	OPLS	−33104 (76)	299 (2)	2.63 (0.17)	0.98 (0.21)	11.70 (0.07)		5017 (97)	4778[b]
IL4									
	AMBER	−68836 (26)	300 (1)	1.14 (0.12)	0.76 (0.09)	14.76 (0.07)	14.55	7763 (168)	7084
	CHARMM 1	−70287 (124)	299 (2)	1.59 (0.27)	0.97 (0.17)	15.00 (0.13)		8050 (223)	7027
	CHARMM 2	−70019 (60)	300 (2)	1.36 (0.13)	0.89 (0.17)	14.82 (0.06)		7814 (118)	7027
	OPLS	−70670 (43)	300 (1)	1.37 (0.13)	0.75 (0.13)	14.79 (0.06)		7661 (104)	7077
GPIIA									
	AMBER	−47239 (32)	300 (2)	0.94 (0.10)	0.62 (0.07)	14.71 (0.05)	14.61	7724 (103)	7503
	CHARMM	−47130 (100)	300 (2)	1.20 (0.15)	0.74 (0.11)	14.78 (0.06)		7799 (96)	7458
	OPLS	−50356 (55)	300 (2)	1.25 (0.26)	0.84 (0.12)	14.81 (0.08)		7879 (153)	7496

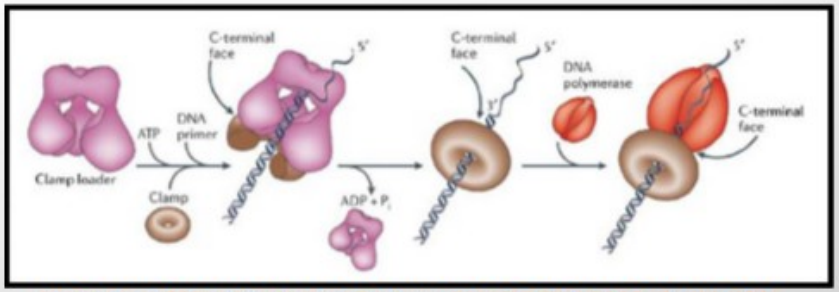
Comparison of AMBER, CHARMM, and OPLS using 3 proteins. Each simulation are 2 ns in length.

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.

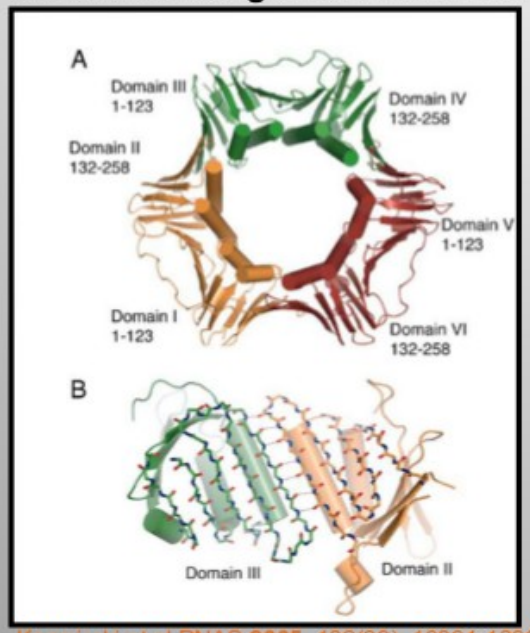
DNA Clamp: PCNA

Yao, Hurwitz, and O'Donnell J Biol Chem **2000**, 275(2), 1421-1432.

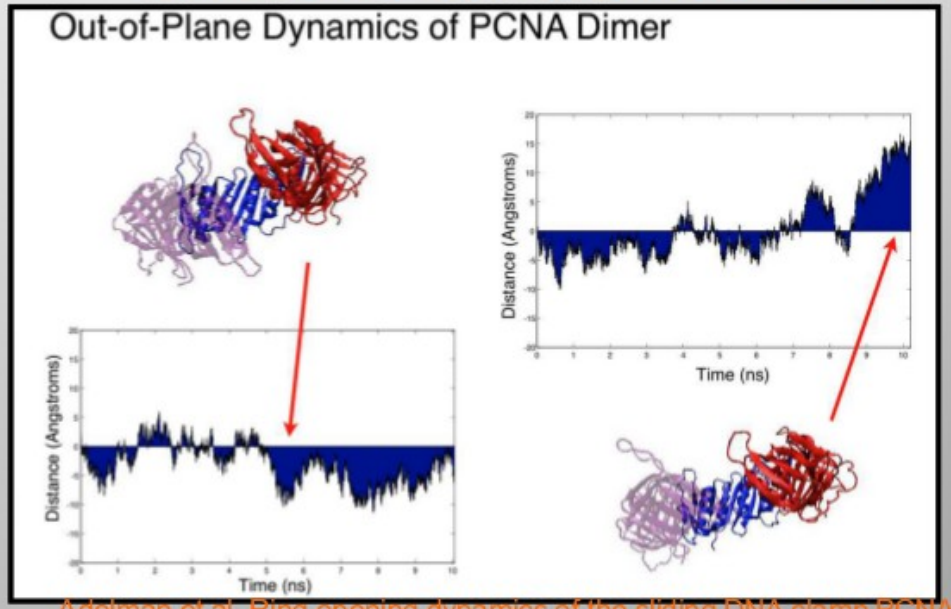
DNA polymerase, clamp, and ATP driven clamp loader together help responsible for replication of DNA. The human form of the clamp, proliferating cell nuclear antigen (PCNA), is a ring shaped protein which trimerizes to forms a six-domain ring and acts as a clamp.



Indiani and O'Donnell Nature Reviews Mol Cell Biol **2006**, 7, 751-761.



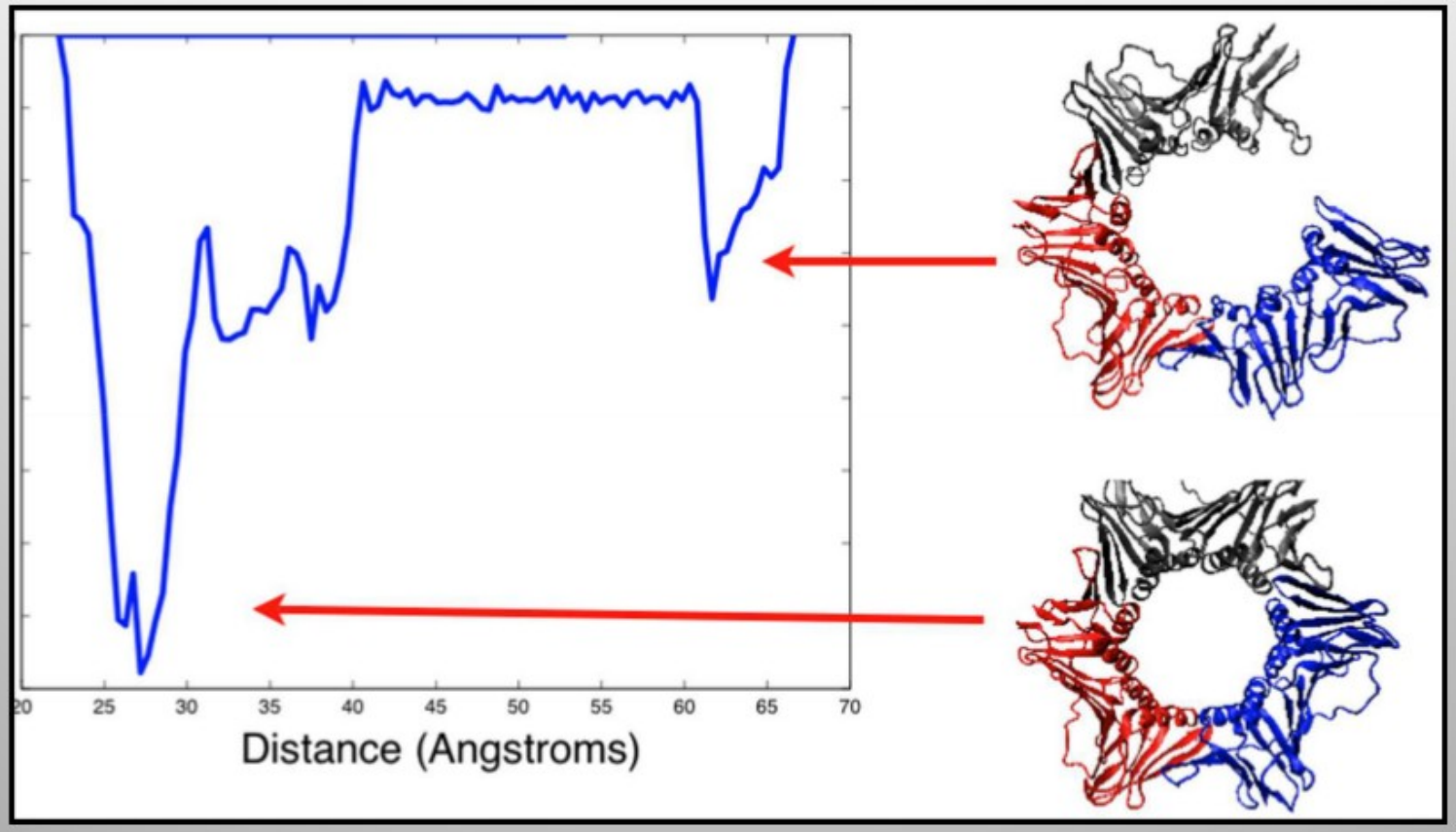
Kazmirski et al PNAS **2005**, 102(39), 13801-13806.



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

Determining the Free Energy of Ring Opening

Adelman et al. Biophysical J. 98 (2010) 3062-3069.



Force Field and Topology section

```
&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.

```
>> 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.

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

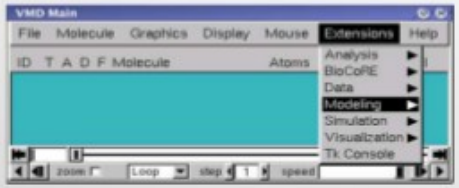
Two new files generated are "dmpa_only.pdb" and "dmpa_only.psf". (Note: The generation of these two file can be done through CP2K also via the DUMP_PDB and DUMP_PSF section.) Ultimately, we are interested in solvated species in a charge neutralized system most of the time. To generated a solvated DMPA 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:

```
vmd >> mol load psf dmpa_only.psf pdb dmpa_only.pdb
```

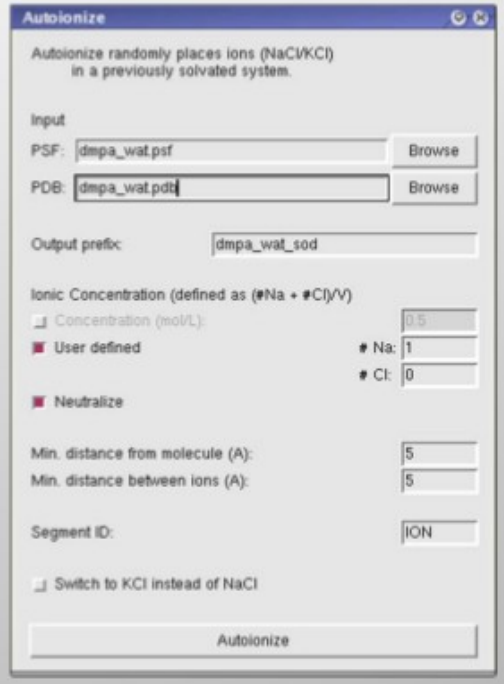
****Quick ways to add solvents and counter ions****

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 neturalize 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
   parm_file_name par_all27_prot_lipid.inp
   parmtype CHM
   &SPLINE
    EMAX_SPLINE    1.0E8
    RCUT_NB        10.0
   &END SPLINE
  &END FORCEFIELD
  &POISSON
   &EWALD
    EWALD_TYPE spme
    ALPHA .35
    GMAX 21 21 21
   &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
   COORD_FILE_NAME     dmpa_wat_sod.pdb
   CONN_FILE_NAME      dmpa_wat_sod.psf
  &END TOPOLOGY
 &END SUBSYS
&END FORCE_EVAL
```

Overcome initial close overlap problem.

$$\alpha \oplus \frac{7}{2} r_{cut}^{-1}$$

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
  &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
   &THERMOSTAT
     TYPE NOSE
     REGION MASSIVE
     &NOSE
       TIMECON   [wavenumber_t] 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
```

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
 &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.

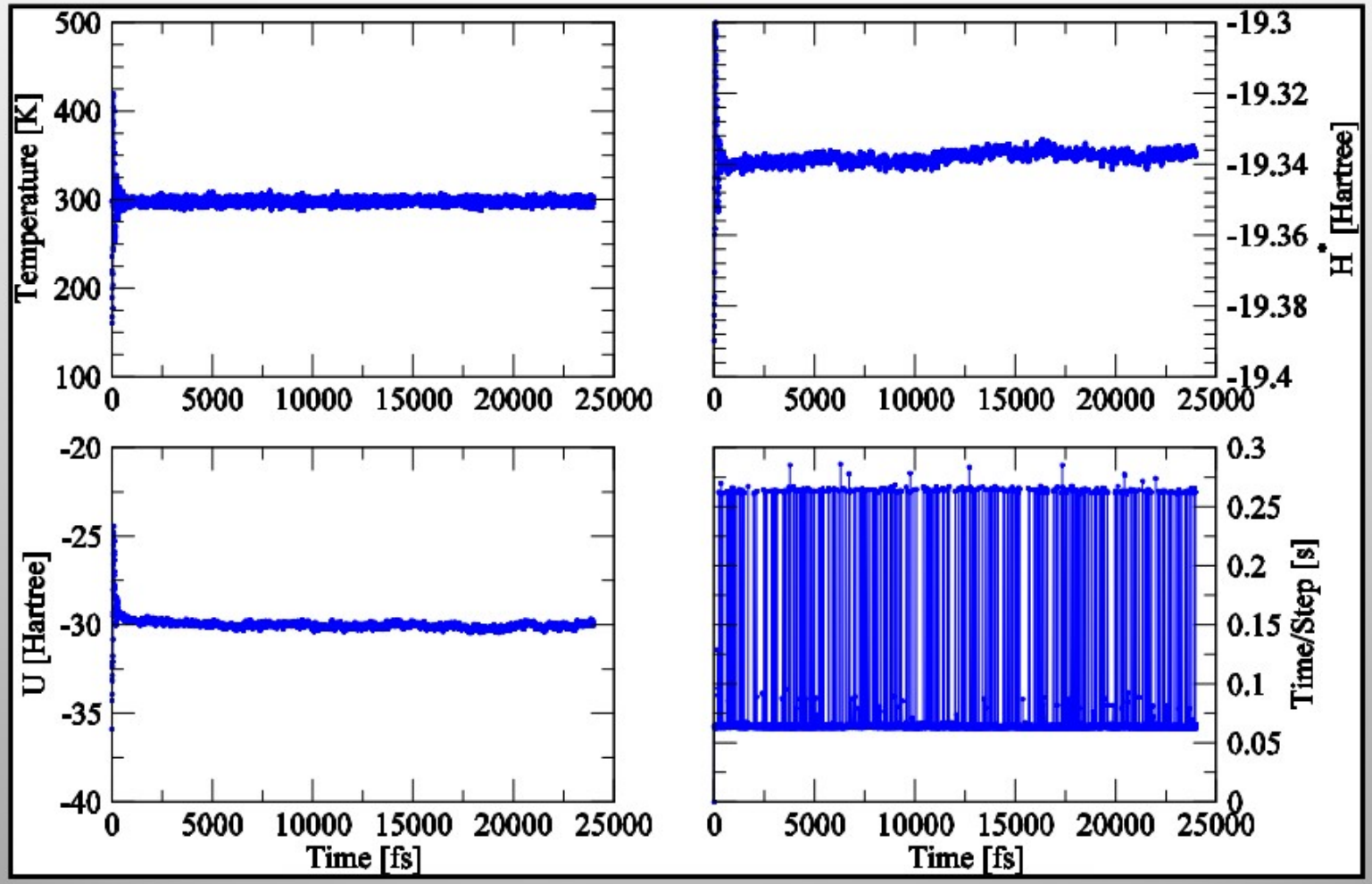
```
>> cp2k.x md.inp
```

```
&MOTION
 &MD
  ENSEMBLE  NVT
  STEPS     1000
  TIMESTEP  0.48
  TEMPERATURE 298.0
  &THERMOSTAT
   TYPE NOSE
   REGION MASSIVE
   &NOSE
    TIMECON  [wavenumber_t] 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
```

Usually a good idea.

Only want to use the minimized positions.

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.