UNIVERSIDAD AUTÓNOMA METROPOLITANA
Unidad Iztapalapa
Casa abierta al tiempo

Departamento de

Química

Simulation methods and molecular interactions: Challenges and opportunities

José Alejandre

OUTLINE OF THE TALK

MOTIVATION

EQUATION OF MOTIONS

EVALUATION OF FORCE FIELDS

A NEW STRATEGY TO PARAMETERIZE

RESULTS FOR POLAR AND IONIC LIQUIDS

CONCLUSIONS

MOTIVATION.

DEVELOPING SOLVENTS FOR DIFFERENT APPLICATIONS.

MOST OF THE WORK IN A LAB ARE TRIAL AND ERROR

Ionic Liquids

1. SOLUBILTY OF DRUGS.
~90% OF MOLECULES WITH FARMACOLOGICAL ACTIVITY ARE LEFT IN THE LAB
BECAUSE THEIR WEAK SOLUBILITY IN WATER → **Reduced bioavailability**

2. EXTRACTION OF CONTAMINANTS IN OIL AND IN DIESEL.
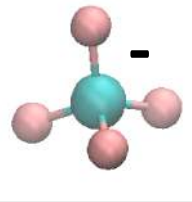OIL → AROMATIC COMPUNDS WHICH PRODUCE CANCER
DIESEL → NITROGEN COMPOUNDS THAT BLOCK THE CATALYTIC CONVERTER

3. ELECTROLYTES FOR LITHIUM ION BATTERIES
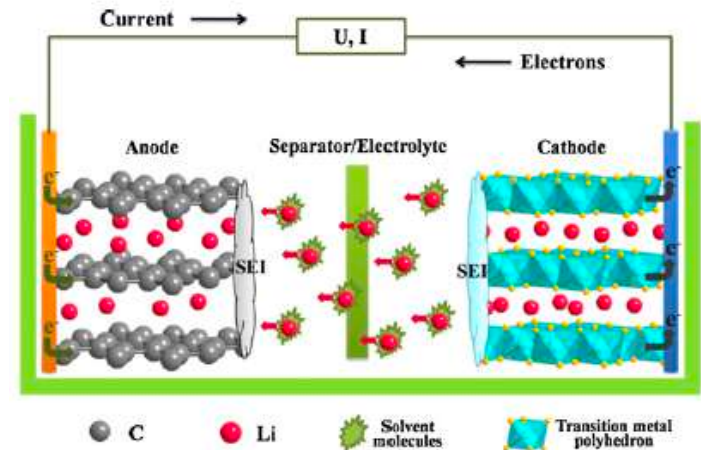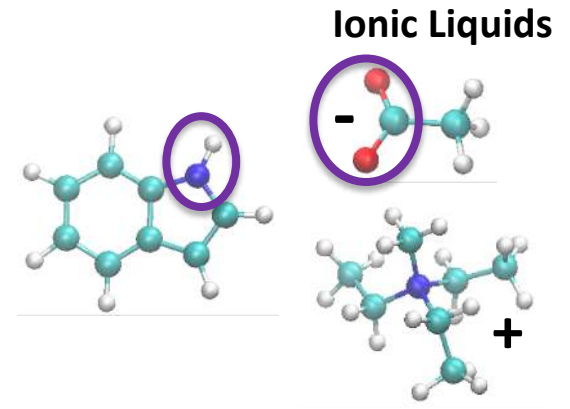SOLVENTS WITH HIGH IONIC CONDUCTIVITY
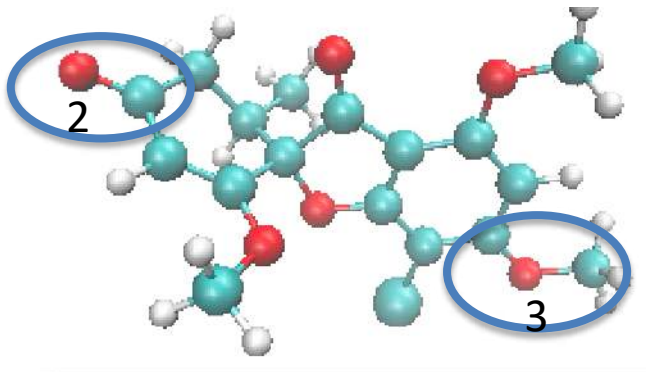AND LESS REACTIVE

Propylene Carbonate

BF4

Li

MOTIVATION.

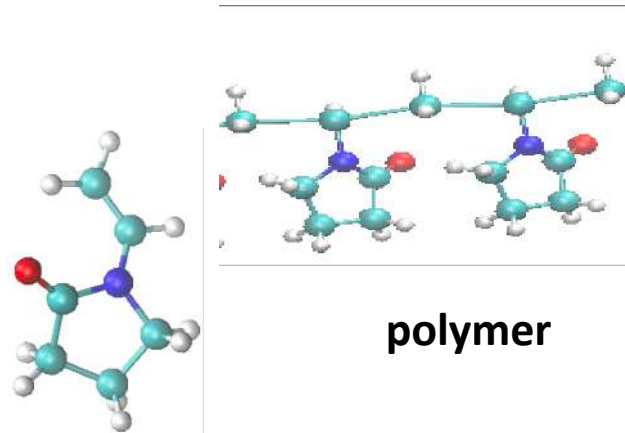## 1. SOLUBILTY OF DRUGS.

~90% OF MOLECULES WITH FARMACOLOGICAL ACTIVITY ARE LEFT IN THE LAB BECAUSE THEIR WEAK SOLUBILITY IN WATER → Reduced bioavailabilty
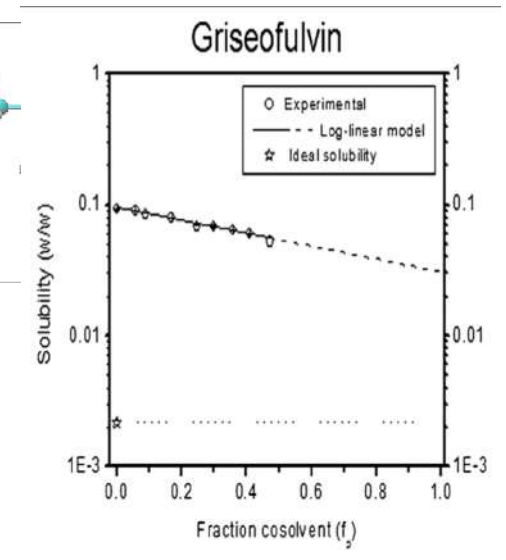
POLYMERS → Drug-Polymer interactions

GRISEORFULVIN
Skin problems

Monomer

polymer

Cocrystals: Hydrogen bond formation
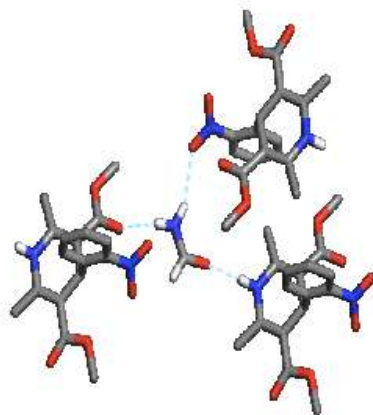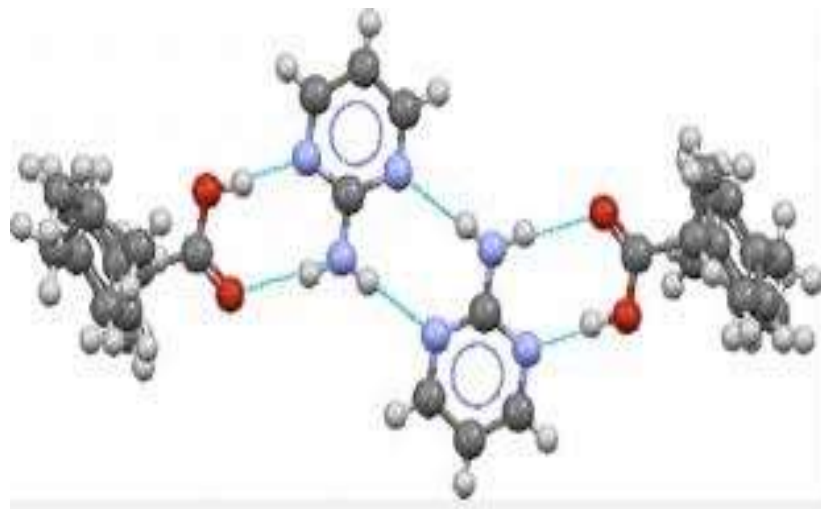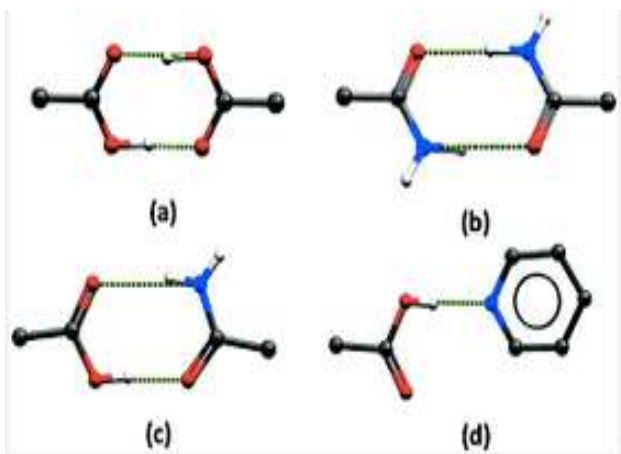
NIFEDIPINE → NOT SOLUBLE IN WATER
FORMAMIDE → SOLUBLE IN WATER

COCRYSTAL
NIFEDIPINE-FORMAMIDE

Cocrystal Itra L Malic Acid   Itraconazole
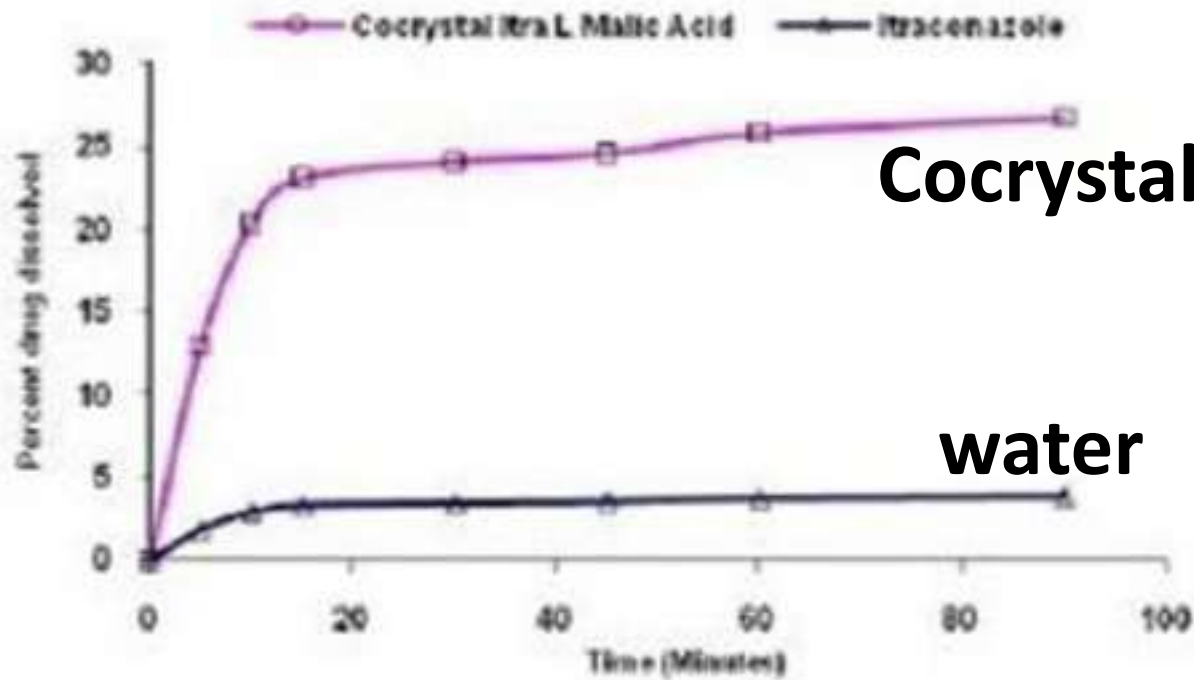
Cocrystal+water

water

Dissolution Profiles in 0.1 N HCl

Extraction with polar and ionic liquids

Pinobkansina

OH 2
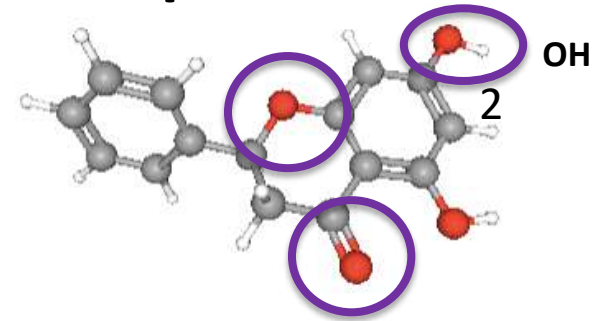
Eficiencia de extracción de los flavonoides más abundantes en Propóleos de Álamo

Legend:
- Pinobanksina (blue)
- Pinocembrina (green)
- Acetato de Pinobanksina (red)

y-axis: mg/g de propóleos

x-axis (Extractos): [Bmim][BF4], [Bmim][Br], [Omim][Br], (Omim-OH][Cl], [Umim][Br], [Umim-OH][Br], H2O, MeOH

Data values:
- [Bmim][BF4]: 5.8 ± 0.1, 6.9 ± 0.2, 3.2 ± 0.2
- [Bmim][Br]: 8.1 ± 0.1, 3.7 ± 0.0, 0.6 ± 0.1
- [Omim][Br]: 12.1 ± 0.5, 34.7 ± 1.2, 17.8 ± 0.4
- (Omim-OH][Cl]: 13.2 ± 1.8, 8.9 ± 0.5, 1.9 ± 0.2
- [Umim][Br]: 2.4 ± 1.4, 8.7 ± 4.1, 3.5 ± 1.5
- [Umim-OH][Br]: 10.8 ± 1.3, 30.5 ± 4.7, 15.9 ± 2.6
- H2O: 2.8 ± 0.4, 1.0, 0.3 ± 0.1
- MeOH: 12.6 ± 0.2, 40.9 ± 0.4, 21.8 ± 0.2

Experiments: Pale-Ezquivel , Tesis de Maestría, Universisdad Veracruzana, 2018.
Molecular Dynamics: Tayde Cadillo, Jahaziel Domínguez y Myrna Matus, UV.

Algorithm with constant energy: N,V,E (Isolated system)

$$\frac{\partial \mathbf{v}}{\partial t} = \frac{\mathbf{F}}{m}; \qquad \frac{\partial \mathbf{r}}{\partial t} = \mathbf{v}$$

$$\mathbf{F}_i(r) = -\nabla_{\mathbf{r}_i} U(r)$$

$\mathbf{r}(0), \mathbf{v}(0)$ y U(r) $\rightarrow$ Obtener fuerzas $\mathbf{F}_i(0)$

$$\mathbf{v}_i\left(\frac{\Delta t}{2}\right) \leftarrow \mathbf{v}_i(0) + \frac{\Delta t}{2m_i}\mathbf{F}_i(0)$$

$$\mathbf{r}_i(\Delta t) \leftarrow \mathbf{r}_i(0) + \frac{\mathbf{v}_i(\Delta t/2)\Delta t}{m_i}$$

Obtener fuerzas $\mathbf{F}_i(\Delta t) \rightarrow U$

$$\mathbf{v}_i(\Delta t) \leftarrow \mathbf{v}_i(\Delta t/2) + \frac{\Delta t}{2m_i}\mathbf{F}_i(\Delta t) \rightarrow K$$

Constant of motion

$$E = K + U$$

How the force is calculated?

Quantum methods➔ Ab initio Molecular Dynamics
Car-Parrinello → Computationally expensive → hundred of molecules

$$\mathbf{F}_i = -\Delta_{\mathbf{r}_i} U(\mathbf{R}_i)$$

Position of nucleous and electrons

Effective potential➔ Classical Molecular dynamics
Several potential parameters are frozen → Fast methods
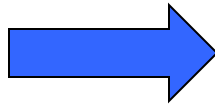**→thousands and million of atoms**

$$\mathbf{F}_{ij} = -\frac{dU(r_{ij})}{dr_{ij}}\frac{\mathbf{r}_{ij}}{r_{ij}}$$

Position of nucleous

Nosé-Hoover Chains: extended variable

System

$$\dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i}$$

$$\dot{\mathbf{p}}_i = \mathbf{F}_i - \frac{p_{\eta_1}}{Q_1}\mathbf{p}_i$$

$$\dot{\eta}_k = \frac{p_{\eta_k}}{Q_k} \qquad k = 1,\ldots,M$$

Thermostat
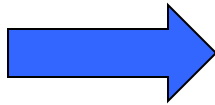
$$\dot{p}_{\eta_k} = G_k - \frac{p_{\eta_{k+1}}}{Q_{k+1}} p_{\eta_k}$$

T External temperature

$$\dot{p}_{\eta_M} = G_M$$

Thermostat forces

$$G_1 = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{m_i} - dNk_BT \qquad G_k = \frac{p_{\eta_{k-1}}^2}{Q_{k-1}} - k_BT$$

Thermostat y Barostat

System ⟶

$$\dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i} + \frac{p_\varepsilon}{W}\mathbf{r}_i$$ **Barostat**

$$\dot{\mathbf{p}}_i = \mathbf{F}_i - \left(1 + \frac{1}{N}\right)\frac{p_\varepsilon}{W}\mathbf{p}_i - \frac{p_{\eta_1}}{Q_1}\mathbf{p}_i$$ **Thermostat**

$$\dot{V} = \frac{dV}{W}p_\varepsilon$$

$$\dot{p}_\varepsilon = dV\left(P_{\text{int}} - P_{\text{ext}}\right) + \frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{p}_i}{m_i} - \frac{p_{\xi_1}}{Q_1'}p_\varepsilon \quad \Bigg\} \quad (*)$$

Thermostat

$$\dot{\eta}_k = \frac{p_{\eta_k}}{Q_k} \qquad k = 1,\dots,M$$

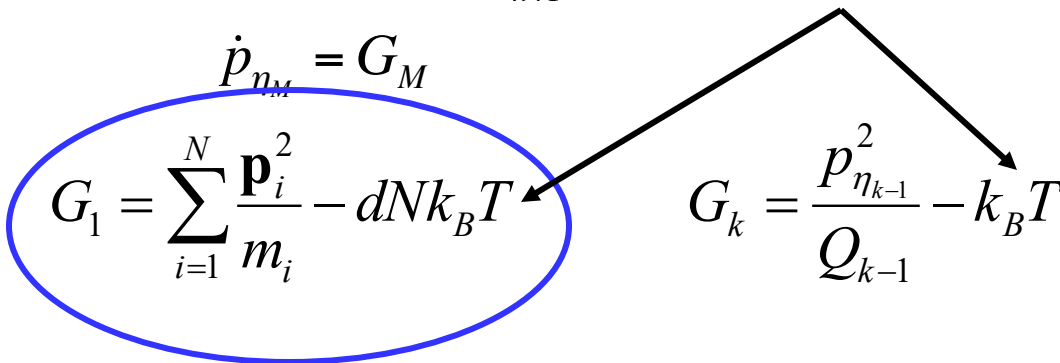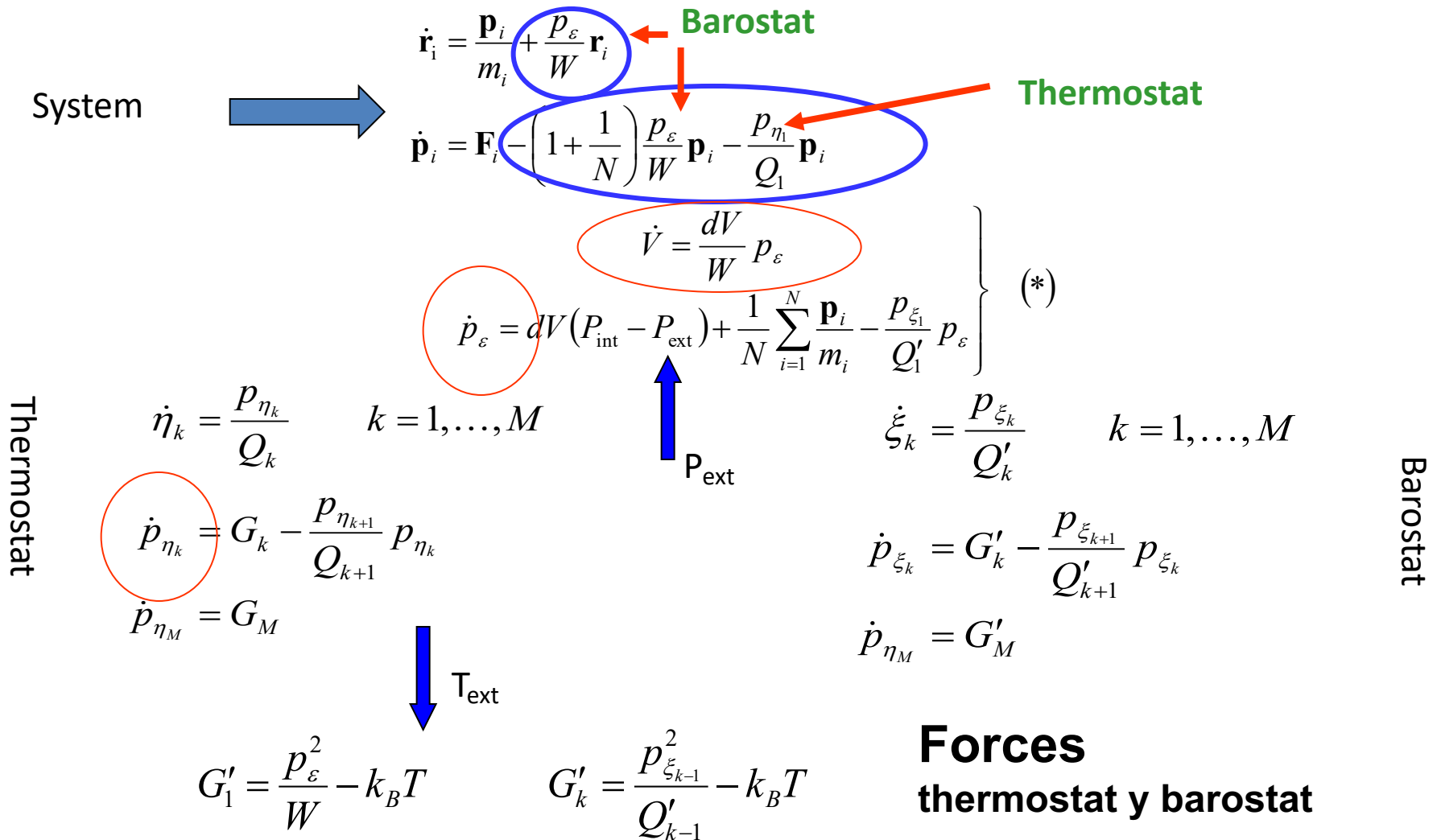$$\dot{p}_{\eta_k} = G_k - \frac{p_{\eta_{k+1}}}{Q_{k+1}}p_{\eta_k}$$

$$\dot{p}_{\eta_M} = G_M$$

P$_{\text{ext}}$

T$_{\text{ext}}$

Barostat

$$\dot{\xi}_k = \frac{p_{\xi_k}}{Q_k'} \qquad k = 1,\dots,M$$

$$\dot{p}_{\xi_k} = G_k' - \frac{p_{\xi_{k+1}}}{Q_{k+1}'}p_{\xi_k}$$

$$\dot{p}_{\eta_M} = G_M'$$

$$G_1' = \frac{p_\varepsilon^2}{W} - k_B T \qquad G_k' = \frac{p_{\xi_{k-1}}^2}{Q_{k-1}'} - k_B T$$

Forces
thermostat y barostat

MOLECULAR SIMULATION PROGRAM

INPUT: $\{N, \mathbf{r}_i(0), \mathbf{v}_i(0), U(r), T_{ext}, P_{ext}\}$

TRAJECTORY OF PARTICLES

MC $\quad \mathbf{r}_i(new) = \mathbf{r}_i(old) \pm \xi \, \Delta r \qquad \dfrac{P_n}{P_o} \, \alpha \quad \exp\left[\left(\dfrac{U_n - U_o}{kT_{ext}}\right) + P_{ext}(V_n - V_o) - \left(\dfrac{N}{kT_{ext}}\right)\ln\left(\dfrac{V_n}{V_o}\right)\right]$

DM $\quad \mathbf{r}_i(\Delta t) = \mathbf{r}_i(0) + \mathbf{v}_i(0) + 0.5\,\mathbf{F}_i(0)\,\Delta t^2 / m_i \qquad\qquad \mathbf{F}_i = -\nabla_{\mathbf{r}_i} U(r)$

OUTPUT:

Termodinamics: density,
Chemical potential, surface tension,
Dinamicss: difussion, viscosity
Structure: g(r), density profile

+ Hundred of equations
+ thousands of lines in a program

GROMACS

MTTK pressure control algorithms

As mentioned in the previous section, one weakness of leap-frog integration is in constant pressure simulations, since the pressure requires a calculation of both the virial and the kinetic energy at the full time step; for leap-frog, this information is not available until *after* the full timestep. Velocity Verlet does allow the calculation, at the cost of an extra round of global communication, and can compute, mod any integration errors, the true NPT ensemble.

The full equations, combining both pressure coupling and temperature coupling, are taken from Martyna *et al.* [31] and Tuckerman [36] and are referred to here as MTTK equations (Martyna-Tuckerman-Tobias-Klein).

With constraints, the equations becomes significantly more complicated(!), in that each of these equations need to be solved iteratively for the constraint forces. The discussion of the details of the iteration is beyond the scope of this manual; readers are encouraged to see the implementation described in Ref. [37].

[36] Tuckerman, M. E., Alejandre, J., López-Rendón, R., Jochim, A. L., Martyna, G. J.   A Liouville-operator derived measure-preserving integrator for molecular dynamics simulations in the isothermal-isobaric ensemble. J. Phys. A. 59:5629--5651, 2006.

[37] Yu, T.-Q., Alejandre, J., Lopez-Rendon, R., Martyna, G. J., Tuckerman, M. E.   Measure-preserving integrators for molecular dynamics in the isothermal-isobaric ensemble derived from the liouville operator. Chem. Phys. 370:294--305, 2010.

EVALUATION OF FORCE FIELDS

Bond distance

Torsions

Intermolecular interactions

Bond angle

Intermolecular interactions
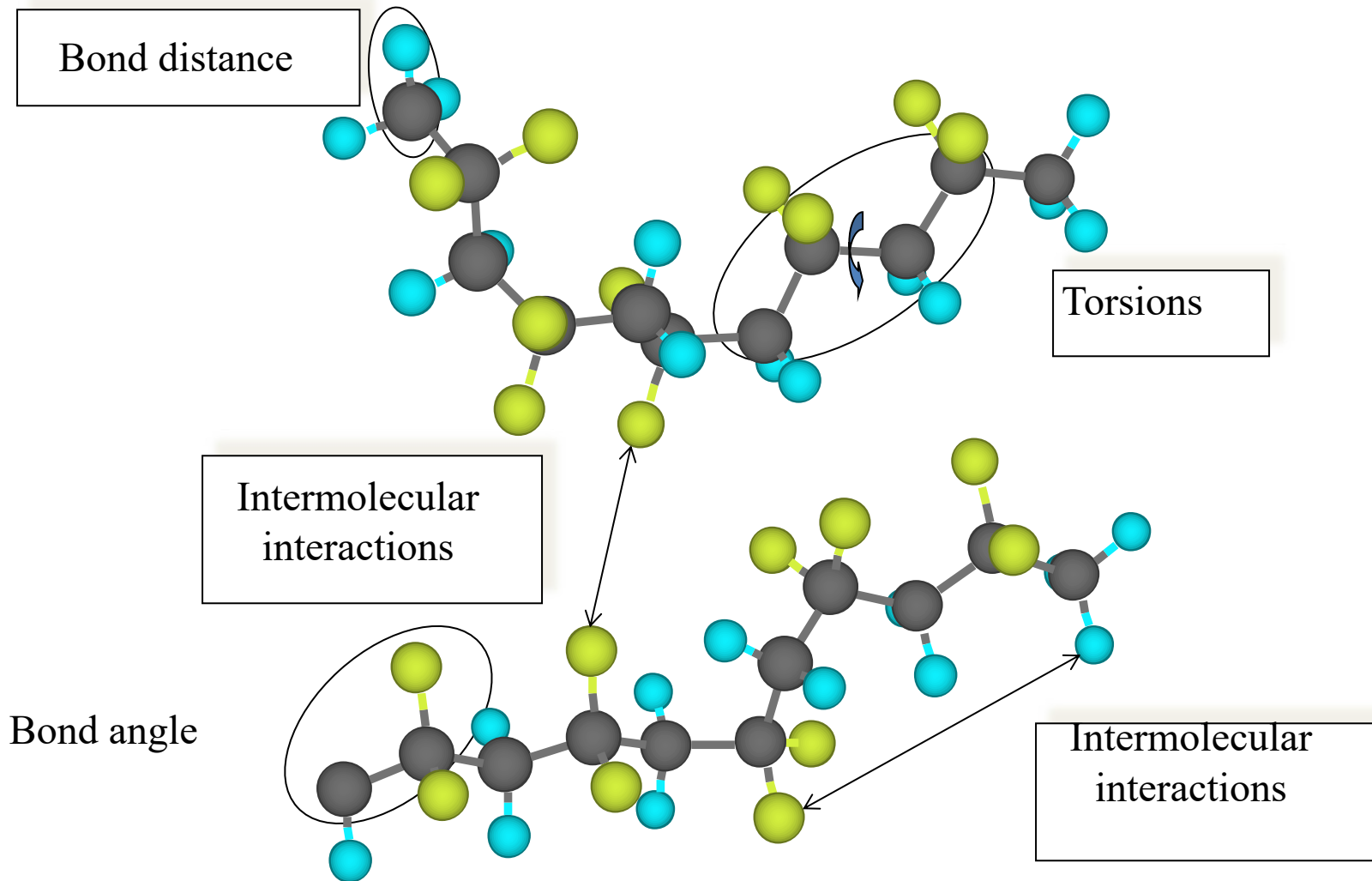
NON-POLARIZABLE INTERMOLECULAR POTENTIAL

$$U(r) = \frac{q_1 q_2}{r} + \frac{q_1 \mathbf{p}_2 \cdot \mathbf{r}_2}{r^3} + \frac{q_2 \mathbf{p}_1 \cdot \mathbf{r}_1}{r^3} + \frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{r^5} - 4\epsilon \left(\frac{\sigma}{r}\right)^6 + 4\varepsilon \left(\frac{\sigma}{r}\right)^{12}$$

charge-charge

Charge-induced dipole

induced dipole-induced dipole

Induced dipole-permanent dipole

250 kJ/mol

15 kJ/mol

1-2 kJ/mol

Computationally expensive

1-2 kJ/mol

Induced dipole moment

Electric field

$$\mathbf{p}_i = \mathbf{p}_{perm} + \mathbf{p}_{ind} \qquad \mathbf{p}_{ind} = \alpha \mathbf{E}_i$$

AMOEBA (ponder), MCDHO (Saint Martin, et al)

FORCE FIELD PARAMETERS

Electronic structure on isolated molecules:
Geometries, rotational barriers and atomic charges(for gases)

Lennard-Jones parameters
MC and MD simulations

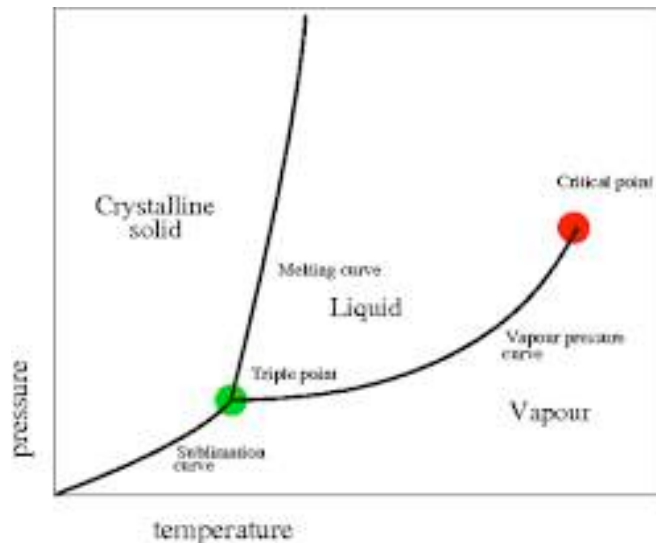**Experimental target properties for pure components**
OPLS/AA, GAFF and CHARMM
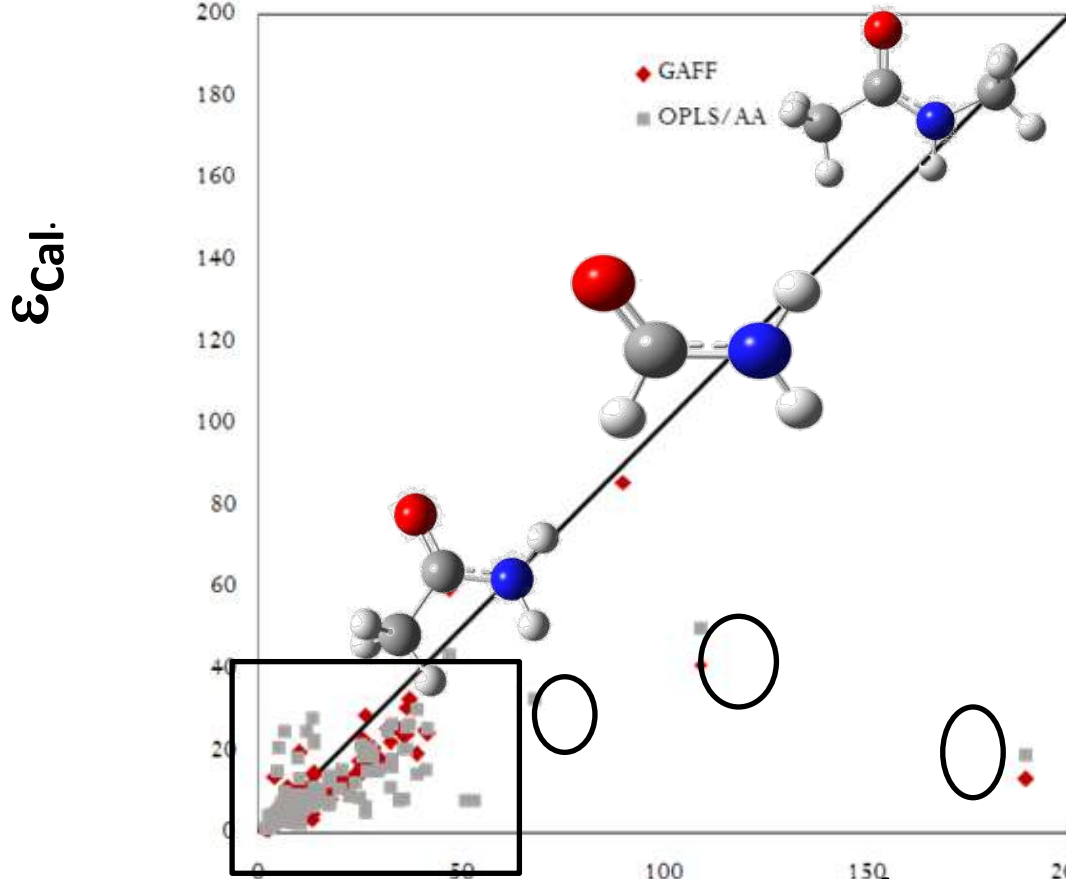**liquid density and heat of vaporization**

NERD, TraPPE
Liquid-vapor densities, boiling temperature and critical properties

Parameters are transferable to other molecules

DIELECTRIC CONSTANT AT 298 K AND 1 BAR
OPLSAA, CHARMM, GAFF FORCE FIELDS

$$\varepsilon = 1 + \frac{4\pi}{3kTV}\left[\langle \mathbf{M}^2 \rangle - \langle \mathbf{M} \rangle^2\right] \qquad \mathbf{M} = \sum_i q_i \mathbf{r}_i$$
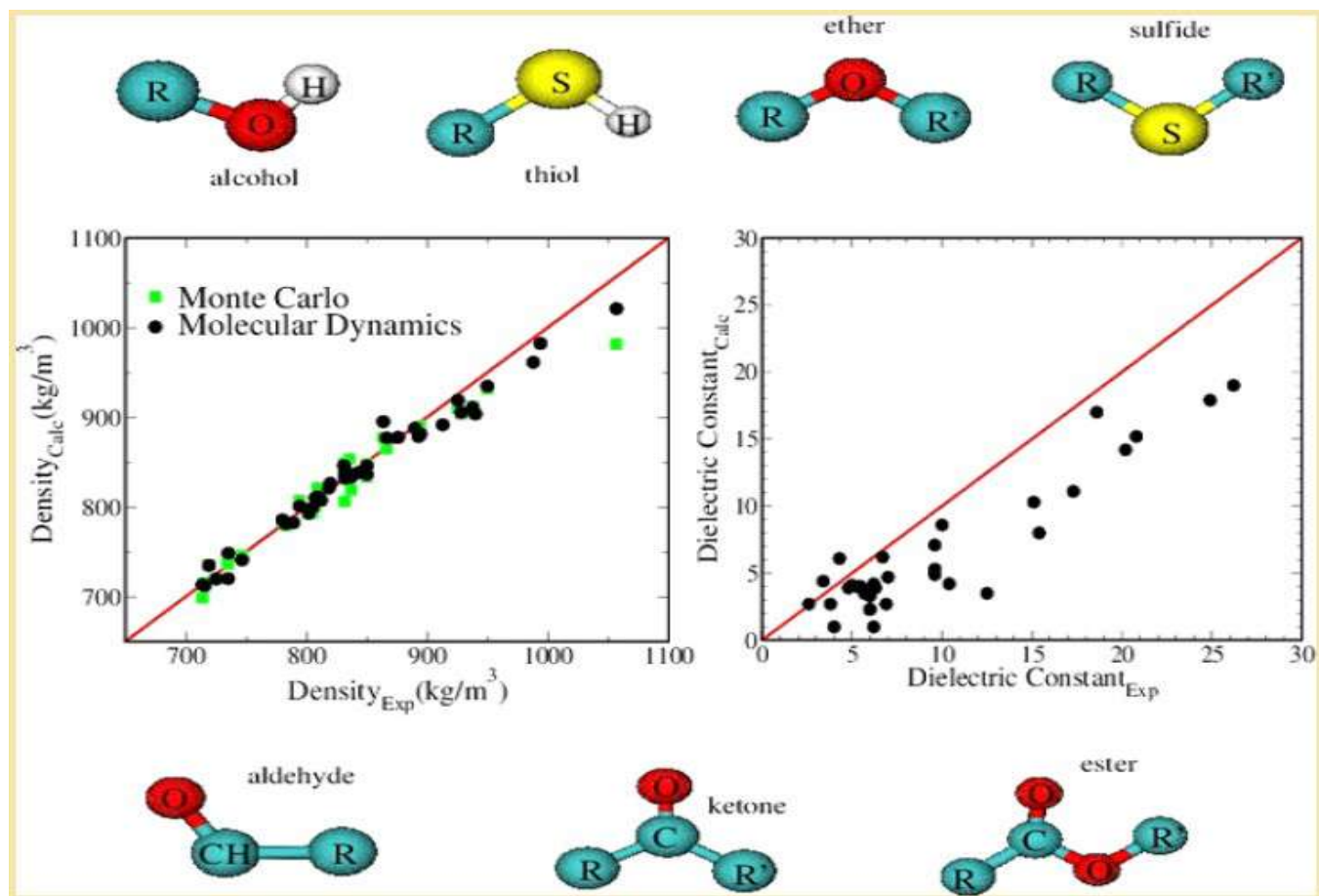

150 Polar liquids

MD values → Exp/2

Caleman, JCTC, 2012

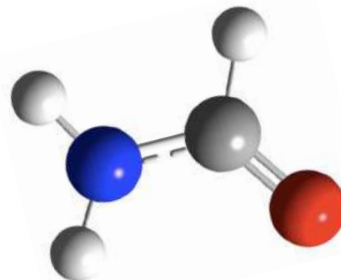# Benchmark over 41 polar liquids using the TraPPE-UA model

~40 % lower than Experiment

Intramolecular parameters taken from OPLS/AA
Charges from electronic structure of isolated molecules

CHARGES IN ATOMS

Formamide→
Isolated molecule

| Átomo	Mulliken	NBO	CHELPG	Mertz-Kollman	Scaled Hirshfeld
C	0.1203	0.5464	0.7267	0.6553	0.1884
O	-0.3798	-0.6097	-0.5835	-0.5554	-0.5171
HC	0.1035	0.1045	0.0197	0.0128	0.0835
N	-0.3919	-0.8346	-0.9407	-0.9623	-0.1442
H	0.2852	0.4000	0.4277	0.4423	0.1947
H	0.2628	0.3934	0.3895	0.4073	0.1947

Molecular dipole moment: Mulliken (5.15D) –> Hirshfeld (5.35)
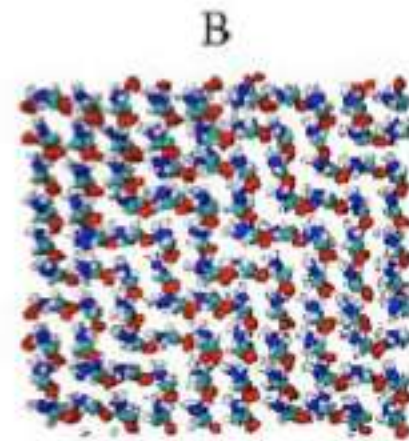
WHICH METHOD IS THE GOOD ONE?

Pérez de la Luz, Méndez-Maldonado, Núñez, Bresme, Alejandre, JCTC, 2015.
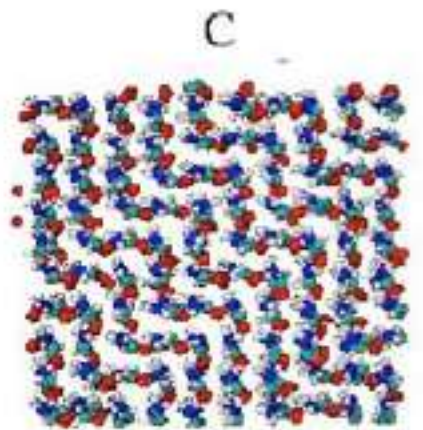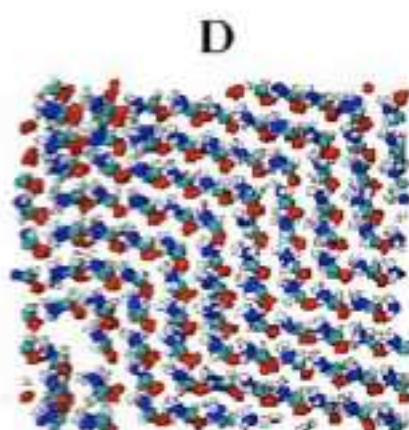
FORMAMIDE AT 298 K AND 1 BAR

Mulliken

NBO
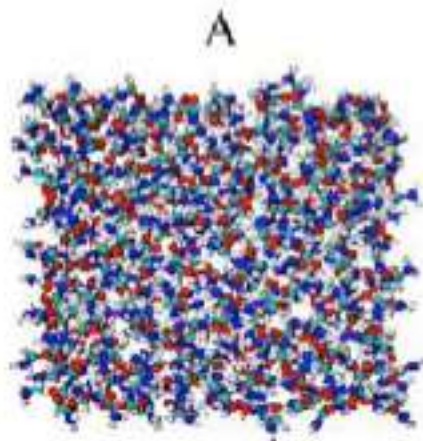
CHELPG

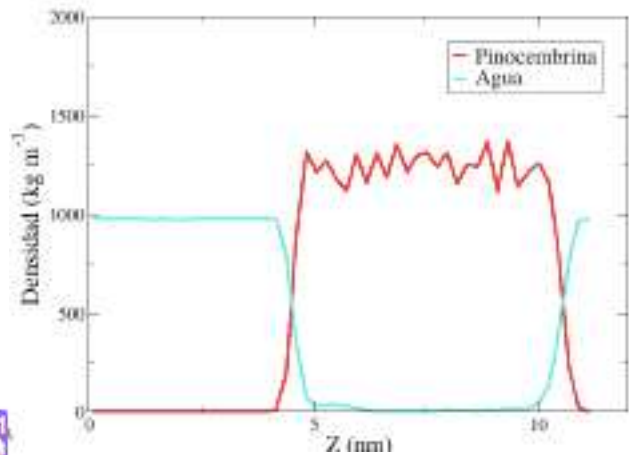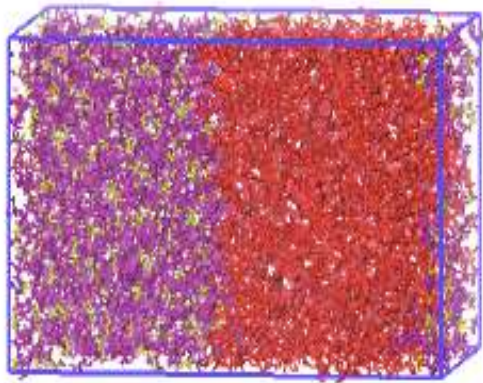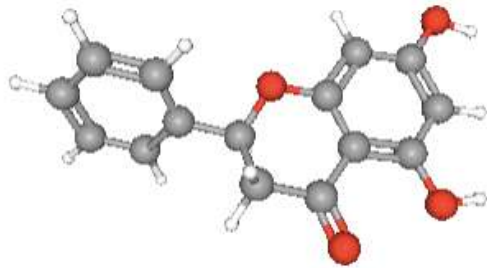**Mertz-Kollman**

Molecular dynamics: pinocembrine is not soluble in water or in ionic liquids
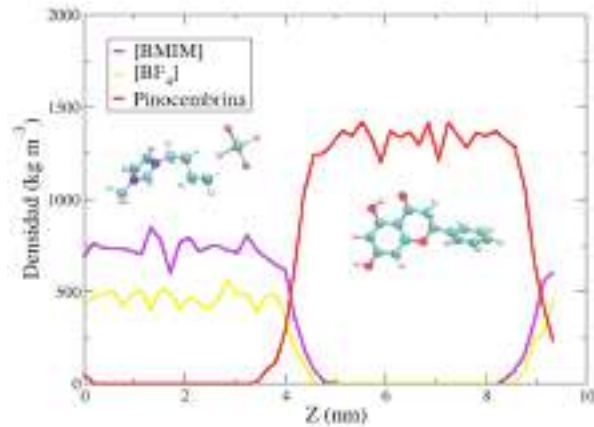
Cocrystal: Nifedipine+Pyrazine
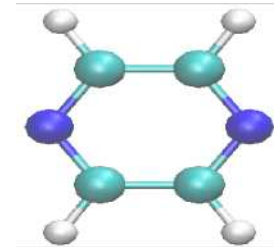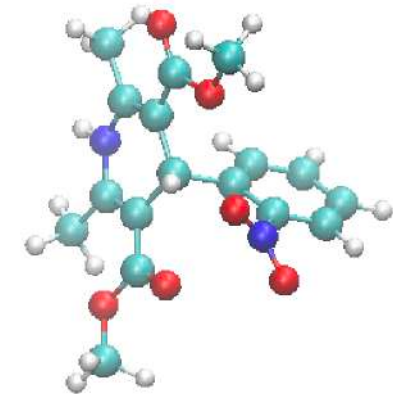
Experiments: Pirazine is soluble en water MD: It is not

A NEW STRATEGY TO PARAMETERIZE: FIRST PART

Procedure that relates experimental properties with interaction parameters

Linear scaling → Global changes

dipole moment $\rightarrow (q_i, r_i) \rightarrow$ dielectric constant

$\varepsilon_{LJ} \rightarrow$ *Surface tension*

$\sigma_{LJ} \rightarrow$ *Liquid density*

Failure: Self-diffusion coefficient lower than experimental values

Salas et al. Journal of Theory and Computational Chemistry, 2015.

SCALED CHARGE DISTRIBUTIONS FROM THE HIRSHFELD PARTITIONING
DILECTRIC CONSTANTS FROM 20 TO 180

Alexander Pérez de la Luz et. al. J. Chem. Theory Comput. 2018, 14, 11

RESULTS FOR PURE COMPONENTS
WITH HIRSHFELD CHARGE DISTRIBUTION
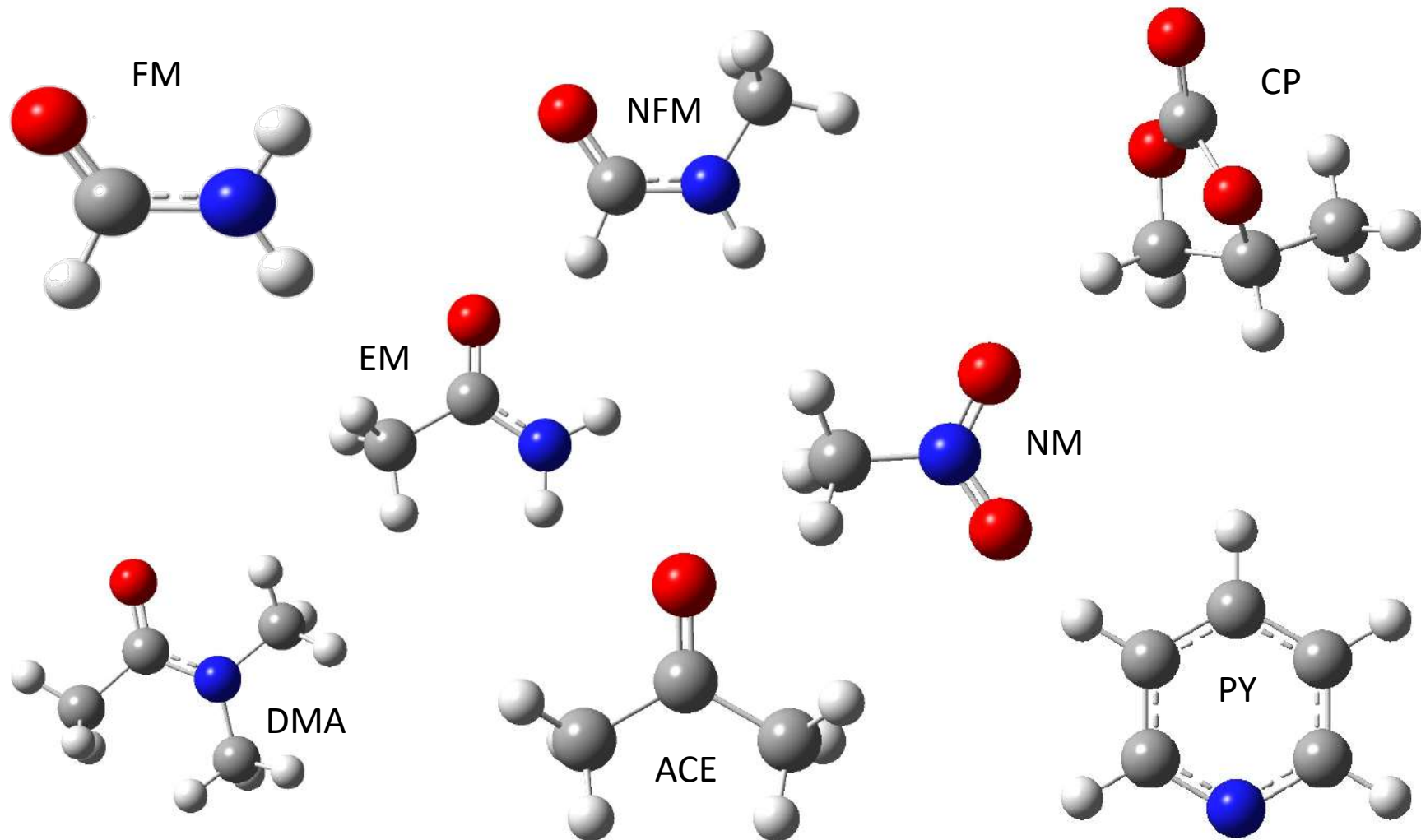
Table 1: Relative error, $|X_{exp} - X_{MD}| * 100/X_{exp}$, between calculated and experimental values for the 10 polar liquids studied in a recent work [24]. The values with green and red colors are for results closest and farthest from experimental values, repectively.

Property	NEWFFP	OPLS/AA	GAFF	CGenFF	OPLS16
Dielectric Constant	3.4	52.6	37.6	42.8	28.8
Surface Tension	2.7	10.5	26.1	22.1	20.7
Density	0.29	1.8	4.1	3.2	2.6
Heat of Vaporization	4.5	3.4	16.3	9.9	5.7
Self-Diffusion	19.2	37.4	49.0	60.6	32.9
Shear Viscosity	21.8	16.6	52.2	35.2	14.0
Isothermal Compressibility	11.6	8.7	27.5	19.2	25.5
Volumetric Expansion Coefficient	15.5	6.2	8.0	4.0	13.4

Pérez de la Luz et. al. JCTC, 2018

MOLECULRA DYNAMICS: SOLUBILITY WITH EXPLICIT WATER

TRAPPE-UA

Molecule	Solubility in water (g/L)	
	Exp	Calc
1-propanol	1000.0 [63]	220.0
2-butanol	181.0 [64]	52.9
ethanethiol	15.6 [65]	0.0
methanethiol	15.4 [66]	21.4
dimethyl-ether	46.0 [66]	540.0
1,2-dimethoxyethane	1000.0 [67]	soluble
diethyl disulfide	0.3 [68]	0.0
dimethyl sulfide	22.0 [69]	0.0
2-pentanone	41.2 [70]	0.0
acetaldehyde	1000.0 [67]	103.3
methyl acetate	205.0 [70]	20.4
methyl acrylate	49.4 [71]	6.0

Wrong

Wrong

CHARMM: Ketones are soluble from C2 TO C6.

OPLS/AA: None of the ketones are solubles

TRAPPE-UA: None of the ketones are solubles

OUR PARAMETERS: Acetone, pyridine and propylene carbonate are not soluble

WHEN FORCE FIELD PARAMETERS ARE GOOD BUT FAIL TO REPRODUCE RESULTS OF BINARY MIXTURES

1. **THE MIXING RULES OF SHORT RANGE INTERACTIONS ARE WRONG**

$$\sigma_{ij} = f_{ij}\left(\frac{\sigma_{ii}+\sigma_{jj}}{2}\right) \qquad \epsilon_{ij} = f_{ij}\sqrt{\varepsilon_{ii}\sigma\varepsilon_{jj}}$$

PROBLEMS TO DEFINE THE VAlUES OF FIJ

2. **THE USE OF POLARIZABLE MODELS**

$$\frac{q_1\mathbf{p}_2 \cdot \mathbf{r}_2}{r^3} + \frac{q_2\mathbf{p}_1 \cdot \mathbf{r}_1}{r^3} + \frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{r^5}$$

COMPUTATIONALLY EXPENSIVE

PARAMETERIZATION WITH EXPLICIT WATER

FOLLOWING THE HYDROGEN BOND, SOLUBILITY OR SELF-DIFFUSION COEFFICIENT

+ MODIFYING CHARGES KEEPING CONSTANT THE MOLECULAR DIPOLE MOMENT

Atom	2-pentanone			
	3SSPP	Dist. 1	Dist. 2	Dist. 3
C1	0.1069	0.1301	0.1373	0.1099
C2	0.2522	0.1800	0.1500	0.1400
C3	0.0464	0.0564	0.0596	0.0703
C4	0.0403	0.0490	0.0518	0.0611
C5	0.0242	0.0295	0.0313	0.0367
O1	-0.4700	-0.4450	-0.4300	-0.4180
μ (D)	3.44	3.45	3.41	3.41
ε	15.6	15.4	15.0	14.1
ρ (kg/m^3)	817.0	817.2	816.0	813.1
γ (mN/m)	22.6	22.8	23.1	22.8
Solubility (g/L)	107.0	64.0	54.0	38.0
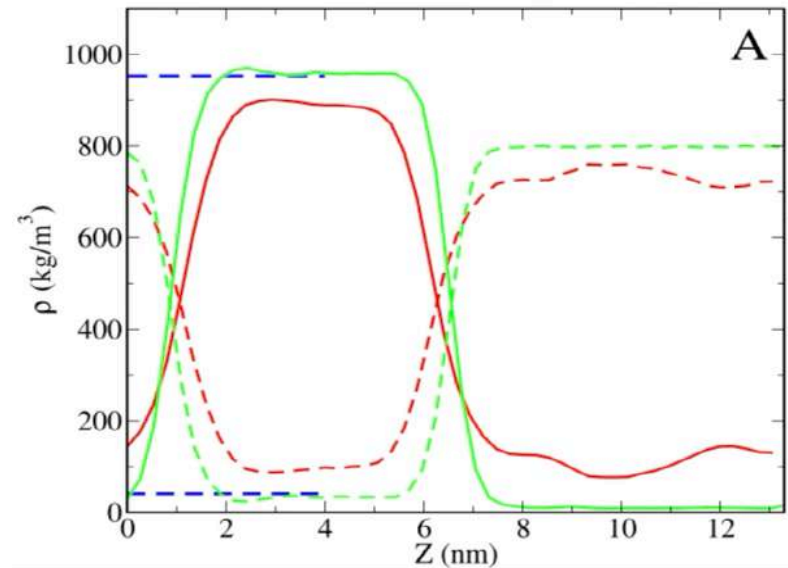
+ MODIFYING THE LENNARD JONES PARAMETERS

CONCLUSIONS

1. PARAMETERIZATION WITH EXPLICIT WATER LOOKING THE LIQUID TO LIQUID-LIQUID TRANSITION HELP TO EQUILIBRATE THE COULOMB AND LENNARD-JONES PARAMETERS

2. THE PARAMETERS CAN BE TRANSFERED SUCCESFULLY TO OTHER MOLECULES AND THERMODYNAMIC STATES

MUCHAS GRACIAS