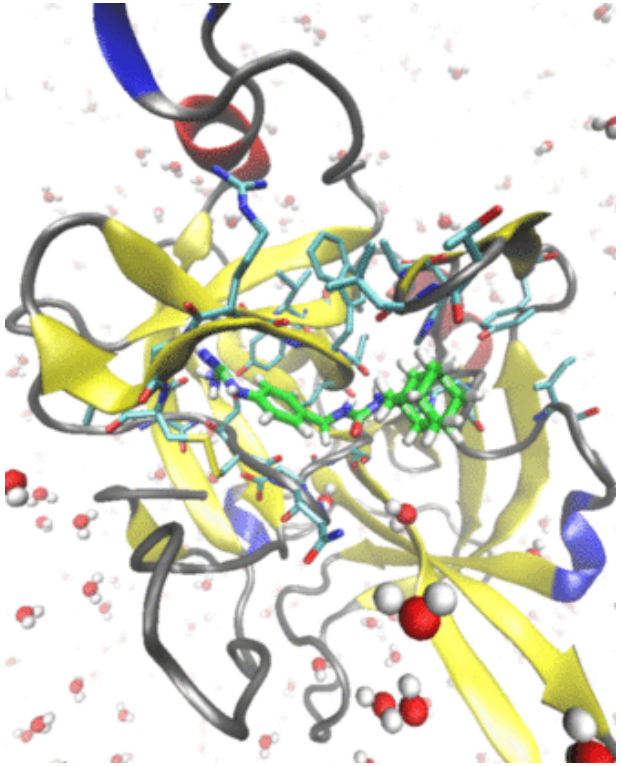
Automation of AMOEBA Polarizable Force Field for Small Molecules – Poltype 2

Brandon Walker
Advisor: Dr. Pengyu Ren
Dept of Biomedical Engineering
University of Texas at Austin

Small molecule significance

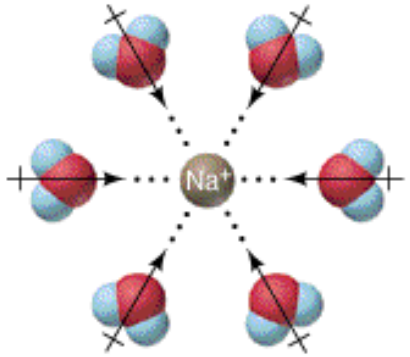


- Typical systems of interest:
  - Protein-ligand
  - DNA/RNA-ligand
- Typical applications:
  - Predicting binding affinity
  - Inhibitor design – affinity optimization

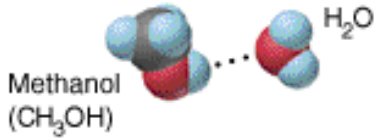
Molecular dynamics. Florent Barbault. (2010, September 10). Retrieved November 8, 2021, from https://florentbarbault.wordpress.com/research/mova-2/.

Small molecule significance

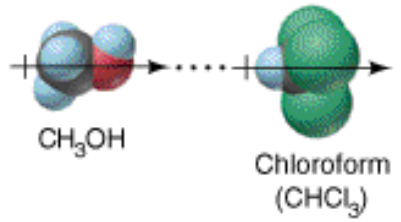
Need to model intermolecular interactions accurately to predict binding affinity !



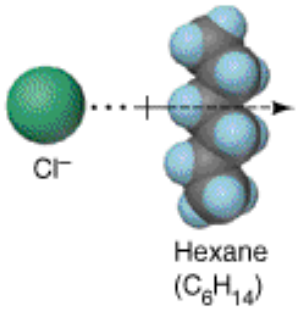
Ion-dipole



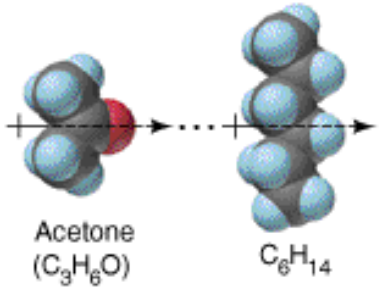
H bond



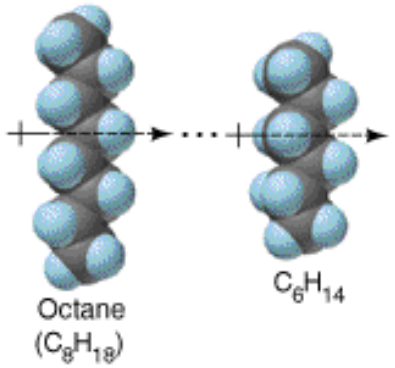
Dipole-dipole



Ion-induced dipole



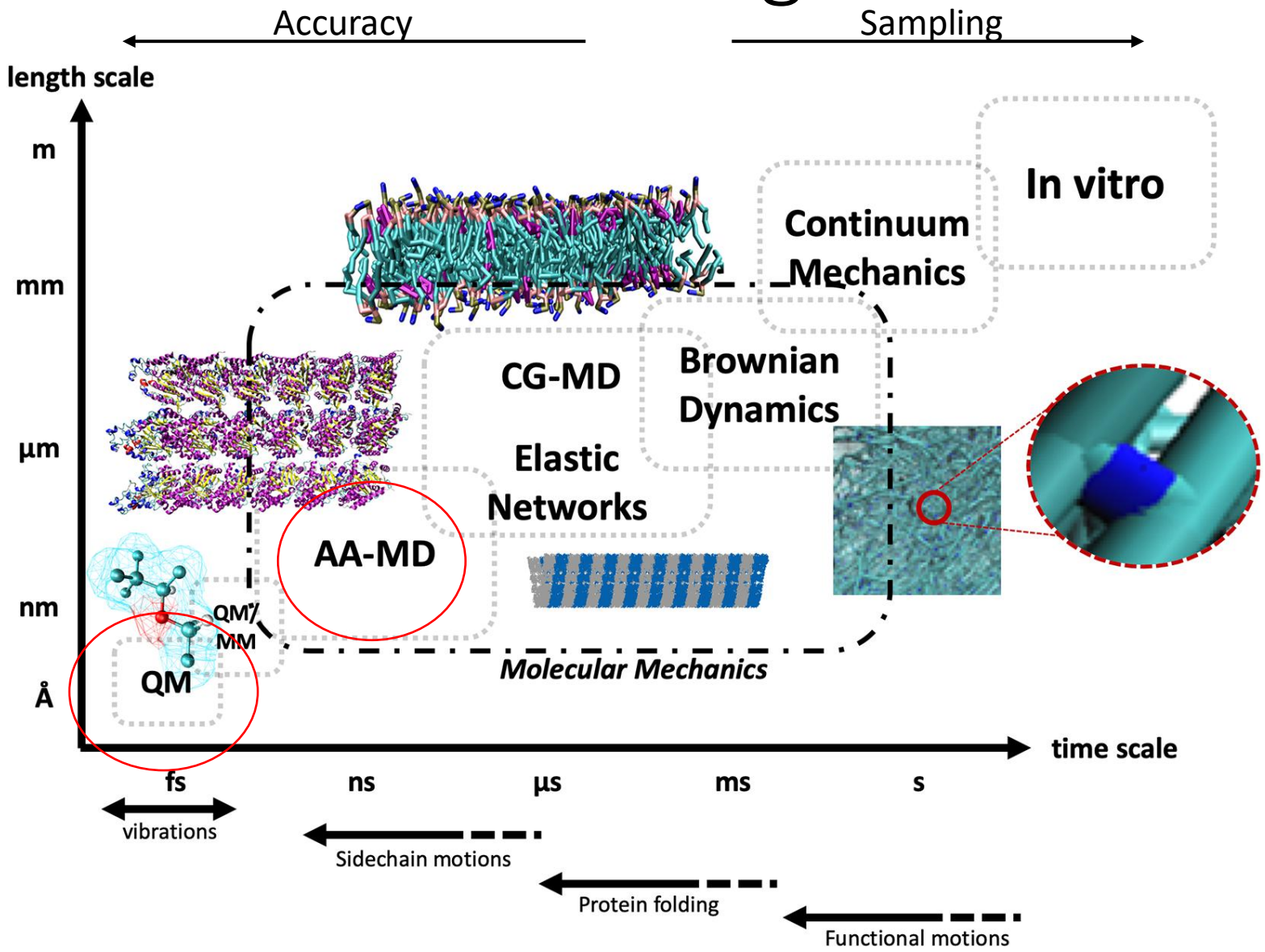
Dipole-induced dipole



Dispersion

Intermolecular Forces (IMFs) http://hdchem.weebly.com/intermolecular-forces-imfs.html (accessed Sep 16, 2020).

Small molecule significance

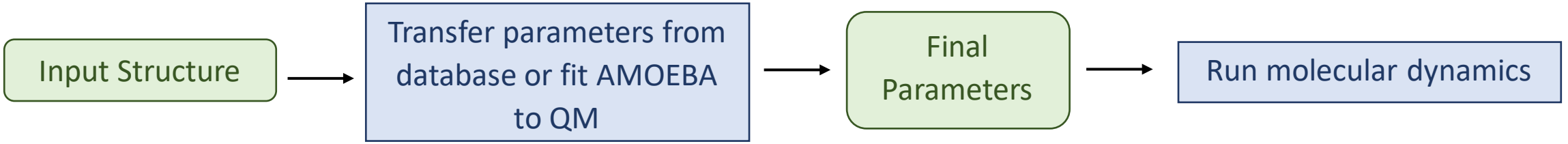


Gaetani, R., Zizzi, E. A., Deriu, M. A., Morbiducci, U., Pesce, M., & Messina, E. (2020). When stiffness matters: Mechanosensing in heart development and disease. *Frontiers in Cell and Developmental Biology*, *8*. https://doi.org/10.3389/fcell.2020.00334

Successful applications of Poltype

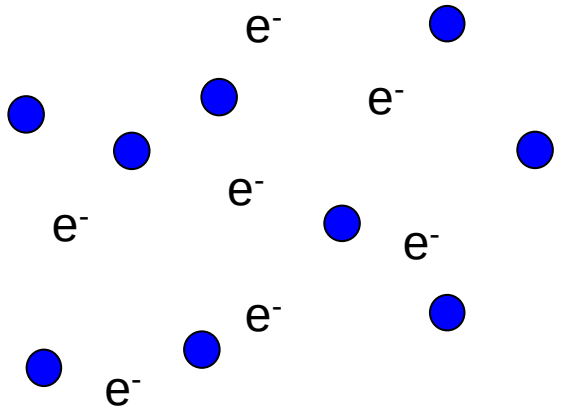
1. Protein-Ligand Studies
   - MELK cell cycle regulation – cancer target[1]
     - **Inhibitor design** – affinity prediction,
   - ALDOLASE A glycolysis regulation – cancer target[2]
     - **Inhibitor design** – affinity prediction
2. DNA/RNA – Ligand Studies
   - Modified Oligonucleotides – miRNA target[3],
   - Mango-II RNA Aptamer – biosensor[4]

Poltype 2 Overview



QM Background

$$H\psi = E\psi$$



$$H = \underbrace{\frac{1}{2}\sum_{j=1}^{N}\sum_{i=1}^{N}\frac{Z_i Z_j e^2}{|R_i - R_j|} - \sum_{j=1}^{N}\sum_{i=1}^{n}\frac{Z_j e^2}{|r_i - R_j|}}_{V_{ext}(r_i)} + \underbrace{\sum_{j=1}^{n}\left(-\frac{\hbar^2}{2m}\right)\nabla_{r_j}^2}_{T} + \underbrace{\frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{e^2}{|r_i - r_j|}}_{V_{int}}$$

QM Background

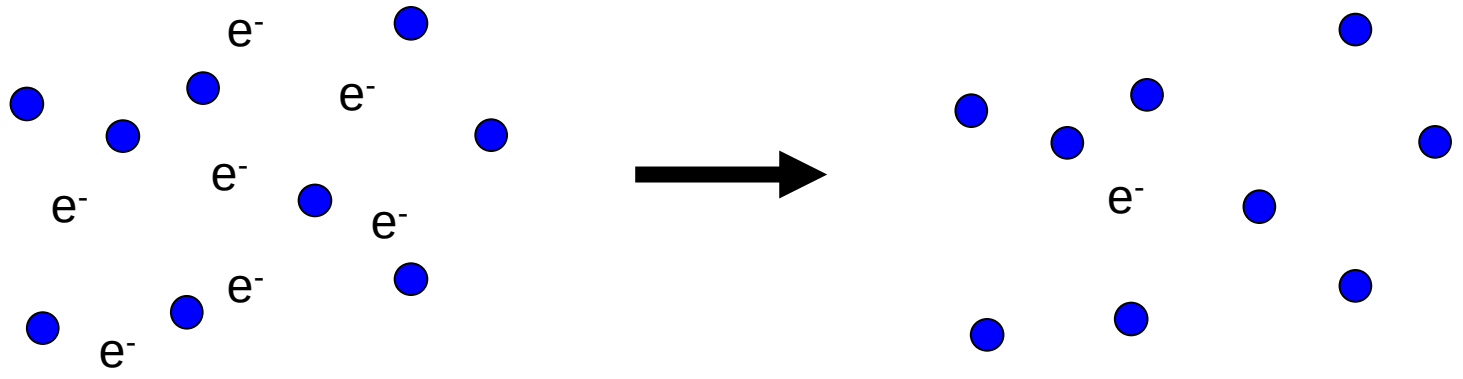
$$E_\Phi = \frac{\int \Phi H \Phi^* d\tau}{\int \Phi \Phi^* d\tau} \geq E_{exact}$$

$$H_{exact} = H_0^{HF} + \lambda V$$

Perturbation

	Density functional theory	Variational wave functions (HF,CI, CC)	Perturbative wave functions (MP2)
Basic object	Density	Many-body wave function	Many-body wave function
Basic equation	E=F[p]	Hψ=Eψ	Hψ=Eψ
Efficiency strategy	Density is only a 3D function	Efficient parameterization of the wave function	Efficient parameterization of the wave function
Accuracy limitation	Don't know the functional	Run out of computer time to add parameters	Run out of computer time to add parameters
Argument for accuracy	Argue functional is accurate, compare to experiment	Variational theorem: lower energy is closer	Perturbation theory: Truncated series expansion of correction terms

QM Background



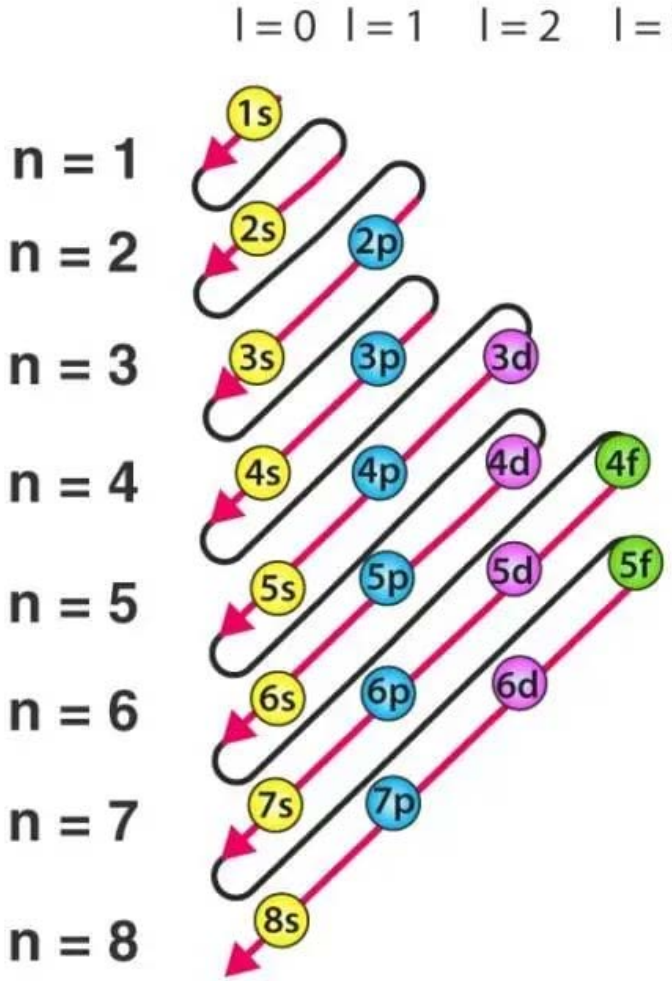
$$H = \underbrace{\frac{1}{2}\sum_{j=1}^{N}\sum_{i=1}^{N}\frac{Z_i Z_j e^2}{|R_i - R_j|} - \sum_{j=1}^{N}\sum_{i=1}^{n}\frac{Z_j e^2}{|r_i - R_j|}}_{V_{ext}(r_i)} + \underbrace{\sum_{j=1}^{n}\left(-\frac{\hbar^2}{2m}\right)\nabla_{r_j}^2}_{T} + \underbrace{\frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{e^2}{|r_i - r_j|}}_{V_{int}}$$

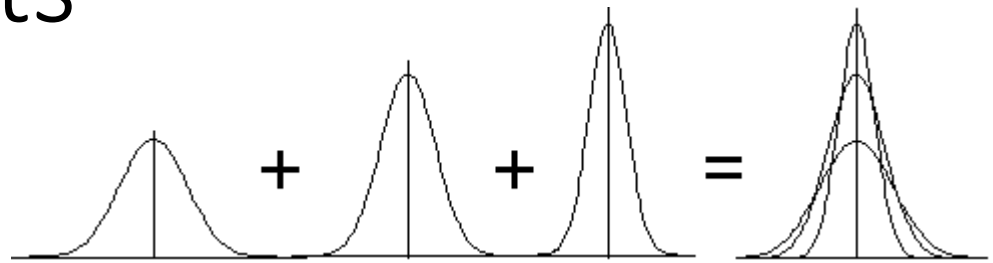
interacting particles in a real external potential

$$H = \left(-\frac{\hbar^2}{2m}\right)\nabla_r^2 + V_{eff}(r)$$

Kohn-Sham system (DFT): a set of non-interacting electrons (with the same density as the interacting system) in some effective potential

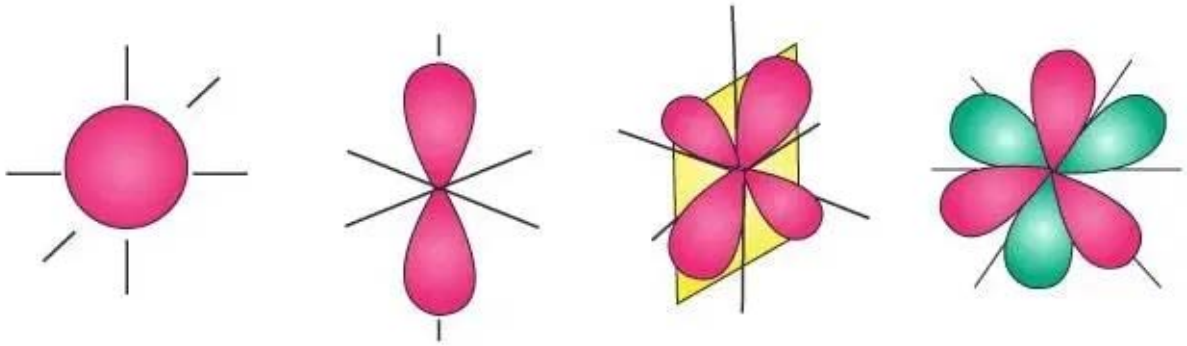
QM Basis Sets





$$\Phi(\mathbf{r}) = R_l(r)Y_{lm}(\theta,\phi)$$

$$R_l(r) = r^l \sum_{p=1}^{P} c_p B(l,\alpha_p)\exp(-\alpha_p r^2)$$



l = 0
s-orbital

l = 1
p-orbital
(1 of 3)

l = 2
d-orbital
(1 of 5)

l = 3
f-orbital
(1 of 7)

QM Basis Sets

3-21G* - Polarization functions on heavy atoms
3-21G** - Polarization functions on heavy atoms and hydrogen
3-21+G - Diffuse functions on heavy atoms
3-21++G - Diffuse functions on heavy atoms and hydrogen
3-21+G* - Polarization *and* diffuse functions on heavy atoms
3-21+G** - Polarization functions on heavy atoms and hydrogen, as well as diffuse functions on heavy atoms

X-Yzg
X= the number of primitive Gaussians comprising each core atomic orbital basis function
Y = number of primitive Gaussians for first valence orbital basis function
Z = number of primitive Gaussians for second valence orbital basis function

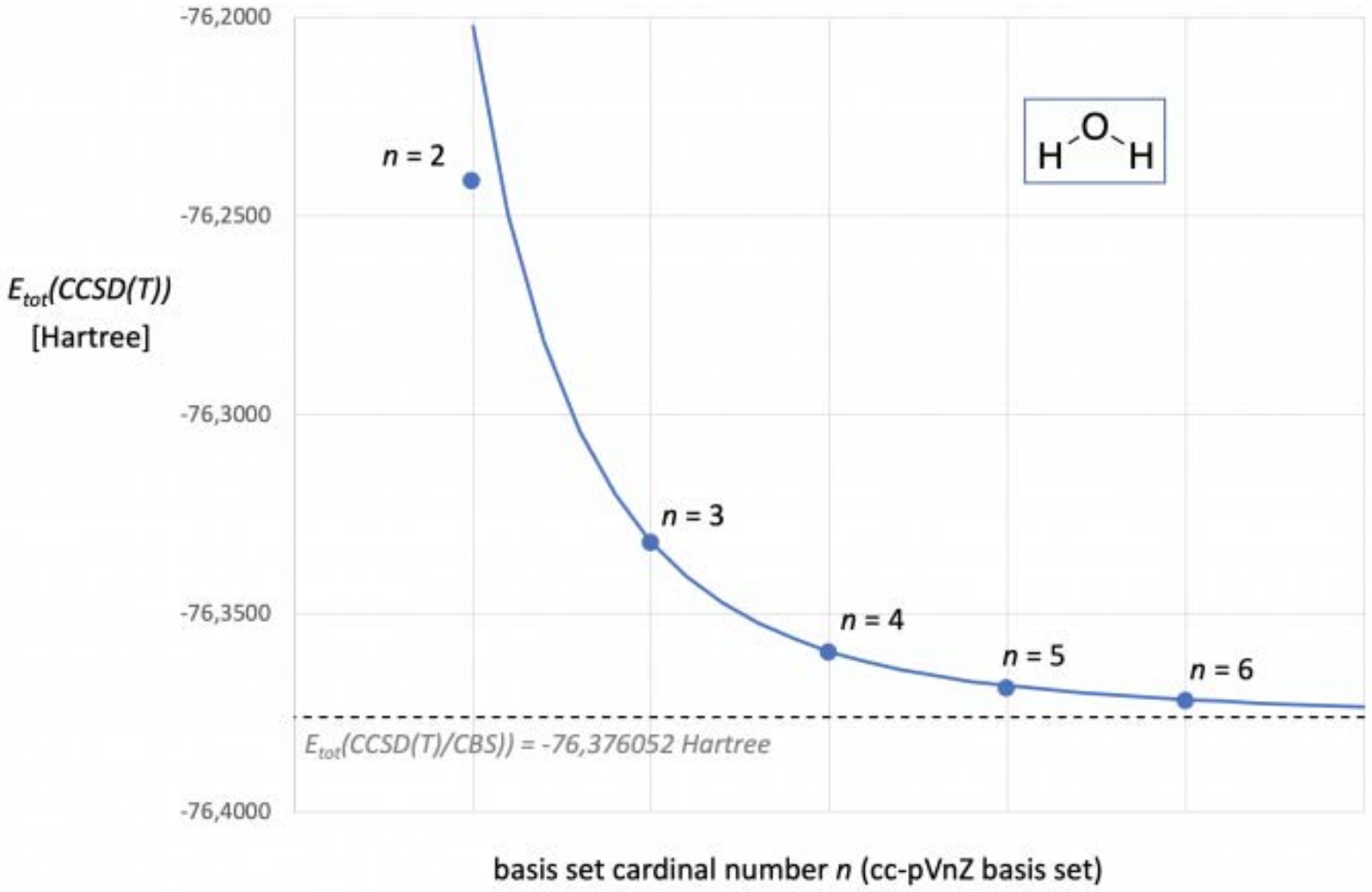
QM Basis Sets

Designed for converging Post-Hartree–Fock calculations to CBS limit

	H-He
cc-pVDZ	[2*s*1*p*] → 5 func.
cc-pVTZ	[3*s*2*p*1*d*] → 14 func.
cc-pVQZ	[4*s*3*p*2*d*1*f*] → 30 func.
aug-cc-pVDZ	[3*s*2*p*] → 9 func.
aug-cc-pVTZ	[4*s*3*p*2*d*] → 23 func.
aug-cc-pVQZ	[5*s*4*p*3*d*2*f*] → 46 func.

aug= augmented, extra functions
cc-p = correlation consistent polarized
V = valence only basis set
Z=Zeta
D = Double of minimum basis functions
T = Triple of minimum basis functions
Q = Quadruple of minimum basis functions

CBS Limit

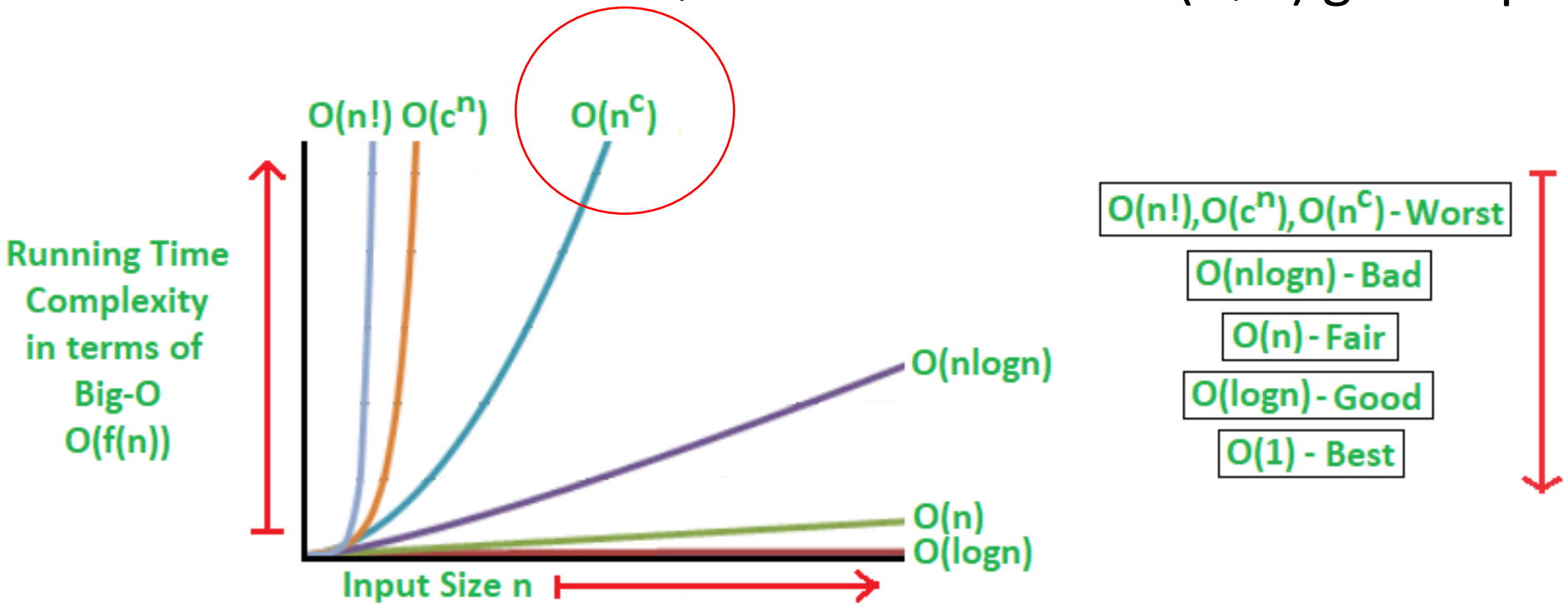


$$E_{SCF}(n) = E_{SCF}(CBS) + A \exp(-z\sqrt{n}) \qquad (1)$$

$$E_{SCF}(CBS) = \frac{E_{SCF}(n)\exp(-z\sqrt{m}) - E_{SCF}(m)\exp(-z\sqrt{n})}{\exp(-z\sqrt{m}) - \exp(-z\sqrt{n})} \qquad (2)$$

Small molecule significance

Quantum mechanics (QM) gets expensive ☹



Number of electrons/size of system

Using molecular mechanics (MM) is exponentially cheaper ☺

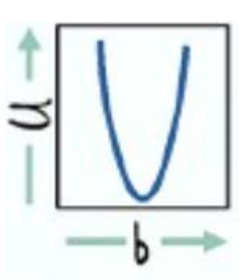
SoumyadeepDebnath The Closer You Look The Less You See, SoumyadeepDebnath, & The Closer You Look The Less You See. (2018, June 08). Analysis of Algorithms: Big-O analysis. Retrieved November 17, 2020, from https://www.geeksforgeeks.org/analysis-algorithms-big-o-analysis/

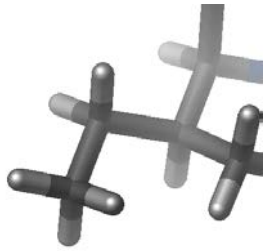
Atomic Multipole Optimized Energetics for Biomolecular Applications (AMOEBA) model

$$U_{AMOEBA} = U_{covalent} + U_{non-covalent}$$

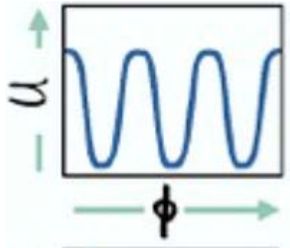
$$U_{covalent} = U_{bond} + U_{angle} + U_{torsion} + U_{oop} + U_{stretch-bend}$$

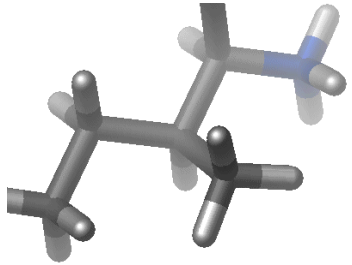
$$U_{non-covalent} = U_{perm-ele} + U_{ind-ele} + U_{vdw}$$



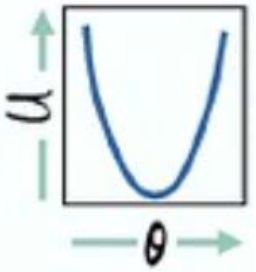


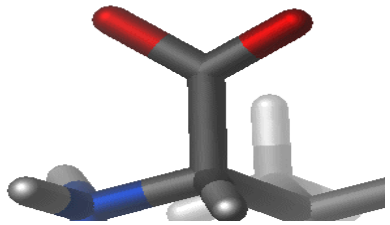
$$U_{bond} = \sum_{bonds} k_r(\mathrm{b} - b_{eq})^2 + \ldots$$



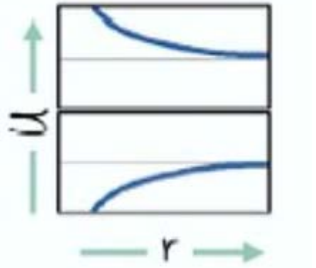


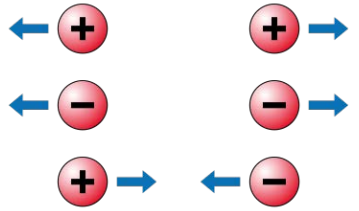
$$U_{torsion} = \sum_{torsions} \sum_{n} \frac{V_n}{2}[1 + cos(n\varphi - \varphi_n)]$$



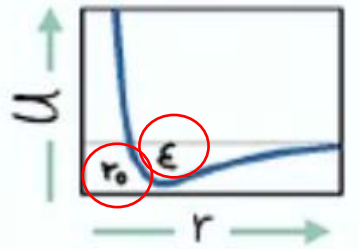


$$U_{angle} = \sum_{angles} k_a(\theta - \theta_{eq})^2 + \cdots)$$





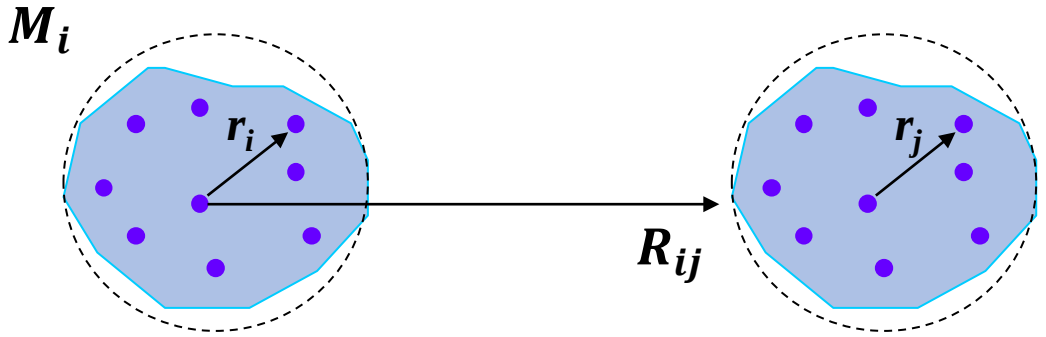
$$U_{perm-ele}$$



$$U_{vdw}$$

Introduction to Molecular Modeling Techniques. (n.d.). Retrieved from http://bionmr.unl.edu/courses/chem991a_-_protein_nmr/lectures/Intro-Min-Dyn.ppt
Levitt, M. (2001). The birth of computational structural biology. *Nature Structural Biology*, *8*(5), 392–393. https://doi.org/10.1038/87545
Sling. (n.d.). *Visualizing van der waals forces on make a GIF*. MakeAGif. Retrieved November 11, 2021, from https://makeagif.com/gif/visualizing-van-der-waals-forces-vgQnui.

AMOEBA model – advanced electrostatics



$$U_{perm-ele} = \sum_{unique\ atom\ pairs} (M_i)^T T_{ij} M_j \qquad T_{ij} = I_{ij} \frac{1}{R_{ij}}$$

Popular models use only point charges for electrostatics

q_i $\quad d_{i\alpha}$ $\quad Q_{i\alpha\beta}$

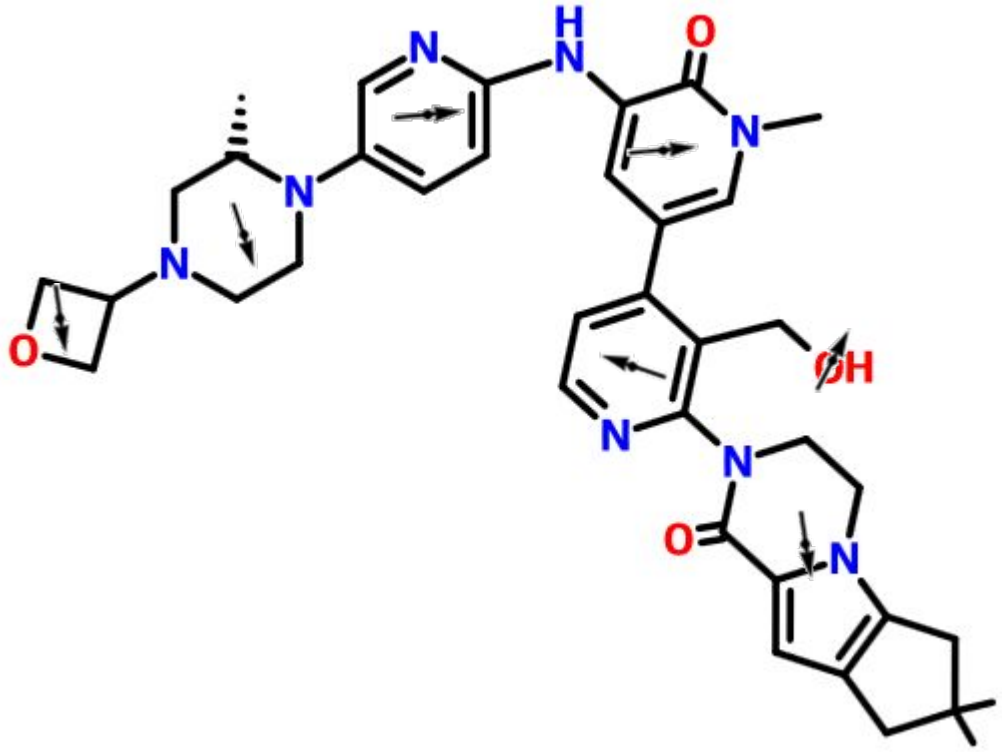
$$M_i = [q_i, d_{ix}, d_{iy}, d_{iz}, Q_{ixx}, Q_{ixy}, Q_{ixz}, Q_{iyx}, Q_{iyy}, Q_{iyz}, Q_{izx}, Q_{izy}, Q_{izz}]$$

$$I = \begin{bmatrix} 1 & \frac{\partial}{\partial x_j} & \frac{\partial}{\partial y_j} & \frac{\partial}{\partial z_j} \\ \frac{\partial}{\partial x_i} & \frac{\partial^2}{\partial x_i \partial x_j} & \frac{\partial^2}{\partial x_i \partial y_j} & \frac{\partial^2}{\partial x_i \partial z_j} \\ \frac{\partial}{\partial y_i} & \frac{\partial^2}{\partial y_i \partial x_j} & \frac{\partial^2}{\partial y_i \partial y_j} & \frac{\partial^2}{\partial y_i \partial z_j} \\ \frac{\partial}{\partial z_i} & \frac{\partial^2}{\partial z_i \partial x_j} & \frac{\partial^2}{\partial z_i \partial y_j} & \frac{\partial^2}{\partial z_i \partial z_j} \end{bmatrix}$$

Highly sensitive to changes in chemical environment, always derive parameters from ab initio data for new molecule

Chapter 3: Legendre Polynomials. Physics. (n.d.). Retrieved November 10, 2021, from https://www.physics.uoguelph.ca/chapter-3-legendre-polynomials.

AMOEBA model - polarization



$$U_{ind-ele} = \frac{-1}{2}\sum_i \mu_i^{ind} . E_i$$

$$\mu_i^{ind} = \alpha_i \left(\sum_{j,j\neq i} T_{ij}^{Thole} M_j + T_{ij}^{Thole} \mu_j^{ind} \right)$$

Transfer polarizability parameter from database

Popular models also tend to ignore polarization effects but in reality, charge distribution changes with dielectric environment changes

Array of Magnetic Dipoles. (n.d.). Retrieved November 17, 2020, from https://commons.wikimedia.org/wiki/File:Array_of_magnetic_dipoles.gif
Crasto, A. (2018, July 26). Drug Approvals International. Retrieved November 17, 2020, from http://drugapprovalsint.com/gdc-0853-fenebrutinib/
Shi, Y., Wu, C., Ponder, J. W., & Ren, P. (2010). Multipole electrostatics in hydration free energy calculations. *Journal of Computational Chemistry*, *32*(5), 967–977. https://doi.org/10.1002/jcc.21681

AMOEBA model – Thole damping

$$\mu_{mol} = \alpha_{mol} E$$

$$\alpha\ par = \frac{\alpha_A + \alpha_B + \frac{4\alpha_A\alpha_B}{r^3}}{1 - \frac{4\alpha_A\alpha_B}{r^6}}$$

$$\alpha\ perp = \frac{\alpha_A + \alpha_B - \frac{2\alpha_A\alpha_B}{r^3}}{1 - \frac{4\alpha_A\alpha_B}{r^6}}$$

$$r = 4\alpha_A\alpha_B{}^{1/6}$$

Singularity!!

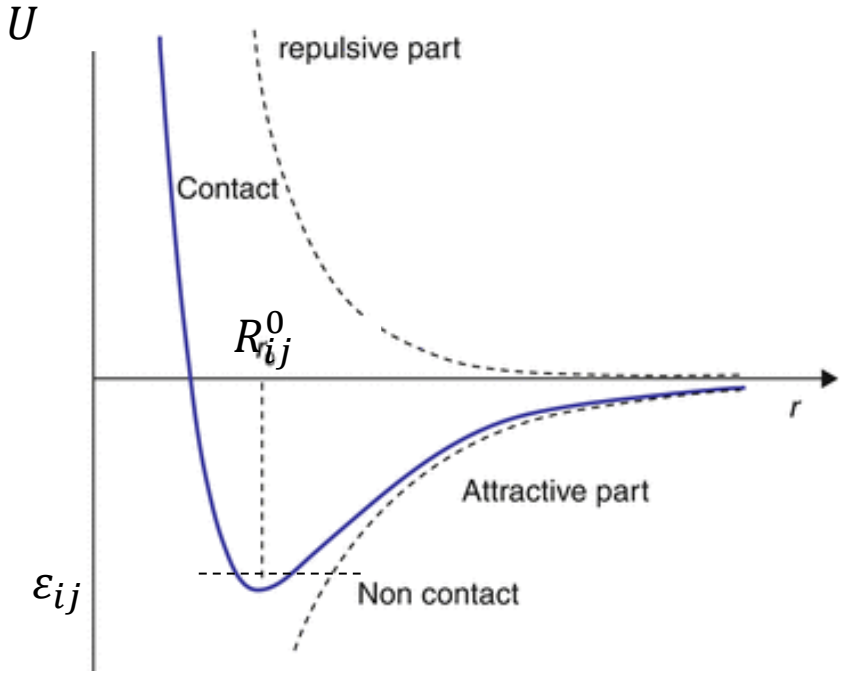
Need to “smear” and damp the interaction near these distances

$$\rho(R_{ij}) = \frac{3a}{4\pi} e^{-au\left(R_{ij}{}^3\right)}$$

$$u = \frac{R_{ij}}{(\alpha_i\alpha_j)^{1/6}}$$

$$T_{ij}^{Thole} = I_{ij}\rho(R_{ij})$$

AMOEBA van der Waals



Poltype can't parameterize only from QM, need experimental liquid property data!!

Transfer “best match” from database.

$$U_{vdw} = \sum_i \sum_{j,j\neq i} \varepsilon_{ij} \left(\frac{1.07}{\rho_{ij} + 0.07}\right)^7 \left(\frac{1.12}{\rho_{ij}^7 + 0.12} - 2\right)$$

$$\rho_{i,j} = \frac{R_{ij}}{R_{ij}^0} \qquad R_{ij}^0 = \frac{\left(R_{ii}^0\right)^3 + \left(R_{jj}^0\right)^3}{\left(R_{ii}^0\right)^2 + \left(R_{jj}^0\right)^2} \qquad \varepsilon_{ij} = \frac{4\varepsilon_{ii}\varepsilon_{jj}}{\left(\varepsilon_{ii}\right)^{1/2} + \left(\varepsilon_{jj}\right)^{(1/2)}}$$

Bruylants, Gilles. “Van Der Waals Forces.” *Encyclopedia of Astrobiology*, 2015, pp. 2583–2585, 10.1007/978-3-662-44185-5_1647. Accessed 5 May 2022.

Small molecule parameterization significance

1. Space of possible small molecules! (10^{30},10^{60})
   - Lots of chemical environments (many different possible model parameters)
2. Expensive QM computations
3. Manual parameterization requires ***expertise*** and is ***prone to human error***
4. Most people that want to use small molecule AMOEBA simulations are not AMOEBA parameterization experts
5. Thus, there is a ***critical need*** for robust automated parameterization

Poltype 2 Features

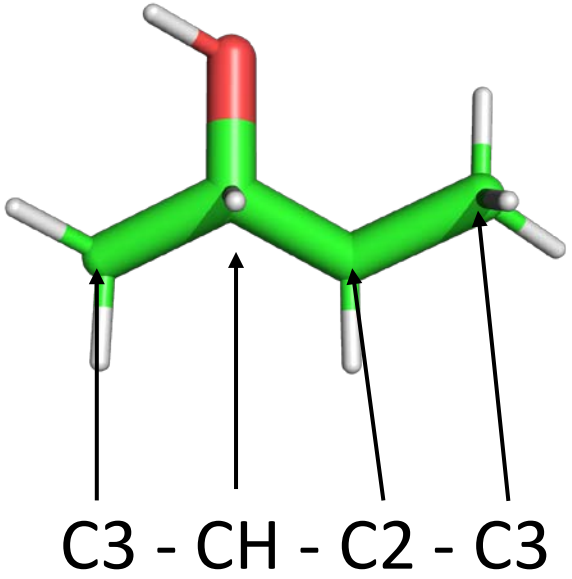
- Automated total charge assignment
- Zwitterion detection
- Ligand ionization state/tautomer enumeration
- Smart memory resource defaults for QM jobs
- Molecule fragmenter to speed up QM
- Parallelized job submission for QM jobs, fragment jobs
- Psi4/Gaussian quantum packages
- QM job error handling
- QM dimer generation for vdW fitting
- Torsion-Torsion coupling
- Missing protein residue/loop modelling with Modeller. pKa, protonation state assignment via pdb2pqr
- Tinker box set up
- BAR Free energy estimation file setup

Same day/next day delivery for customer support…

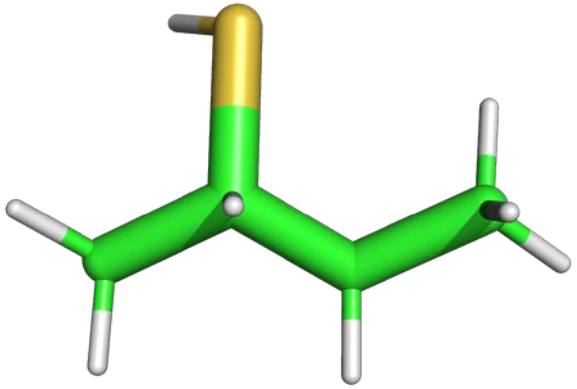
Atom Typing – Database Transfer

C3=sp^3 carbon with 3 hydrogens
CH=sp^3 carbon with 1 hydrogen
C2=sp^3 carbon with 2 hydrogens

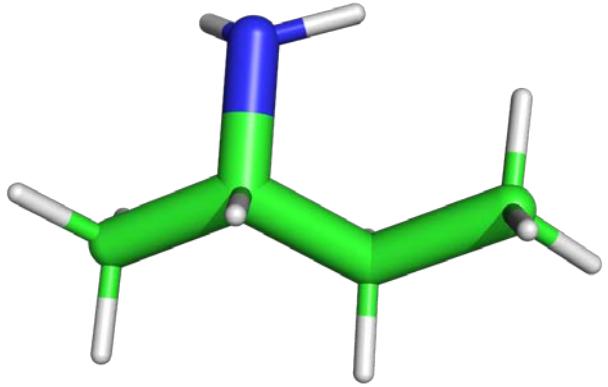
[CH3][CH]([OH])[CH2][CH3]



[CH3][CH]([SH])[CH2][CH3]



[CH3][CH]([NH2])[CH2][CH3]



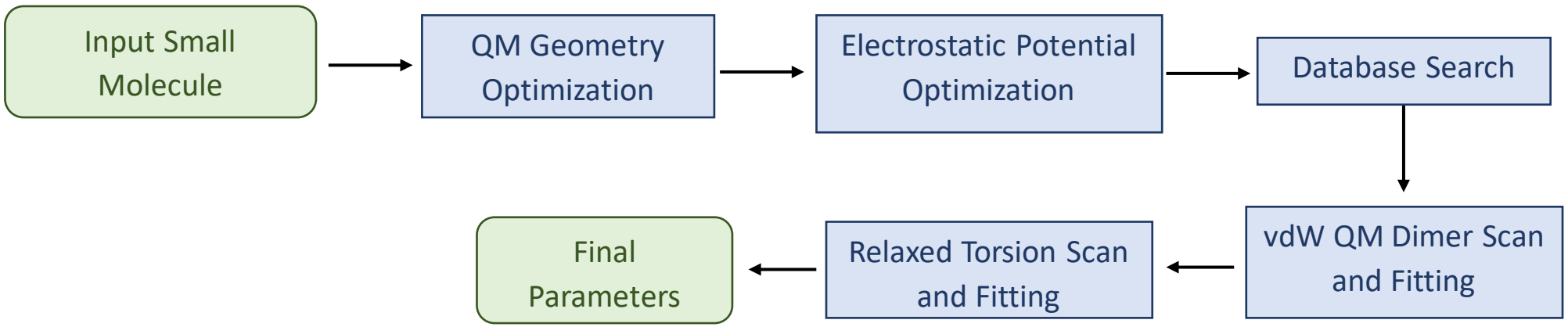
Database:

C3 - CH - C2 - C3 : $\boldsymbol{V_1, V_2, V_3}$

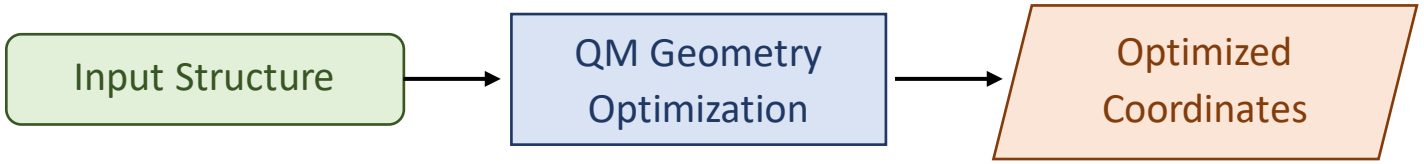
….

….

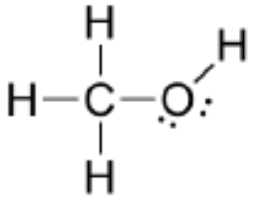
Poltype 2 Parameterization Overview



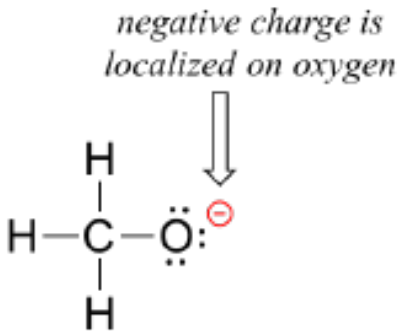
Poltype 2 Flow – Optimization



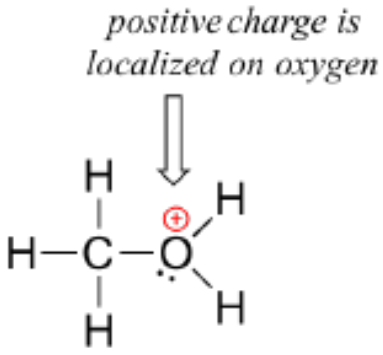
Input Structure



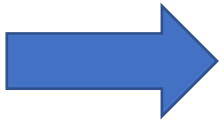
methanol



molecular anion



molecular cation



Make sure bond orders are consistent with formal charge state desired!!

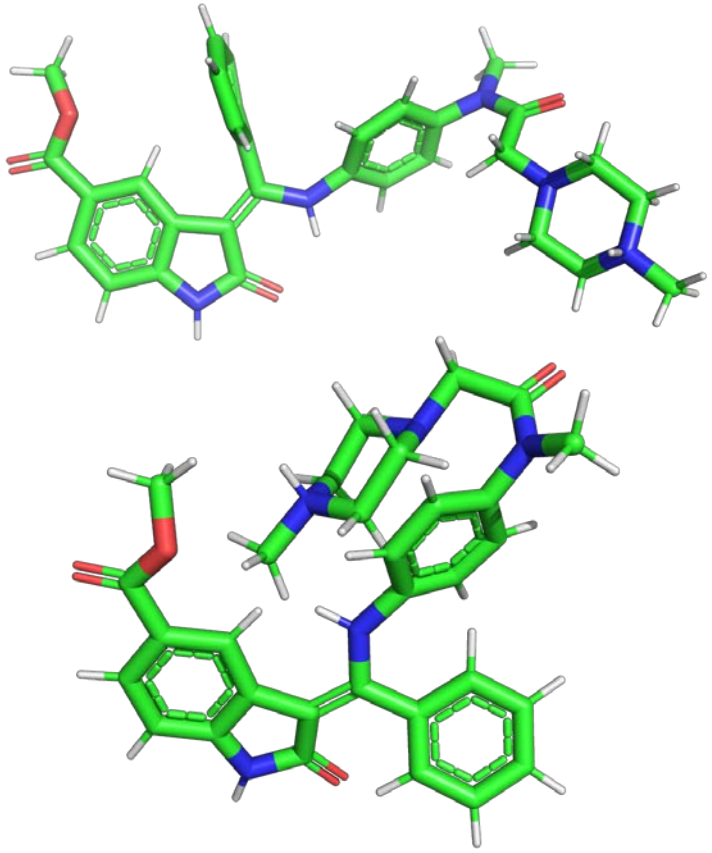
If 2D structure is given in the SDF file, poltype will generate 3D coordinates for you

Dominant protomer states at pH=7 are computed

Radical charge states require extra keyword in SDF file

Libretexts. "1.2: Drawing Organic Structures." *Chemistry LibreTexts*, Libretexts, 2 Aug. 2020, https://chem.libretexts.org/Courses/Nassau_Community_College/CHE200_-_Introduction_to_Organic_Chemistry_(Resch)/01%3A_Introduction_to_Organic_Structure_and_Bonding_I/1.02%3A_Drawing_organic_structures.

QM Geometry Optimization



QM level of theory MP2/6-31G*

Extended conformation ✓

automatically found by Poltype…

Try not to conflate overlapping densities in a folder conformation..

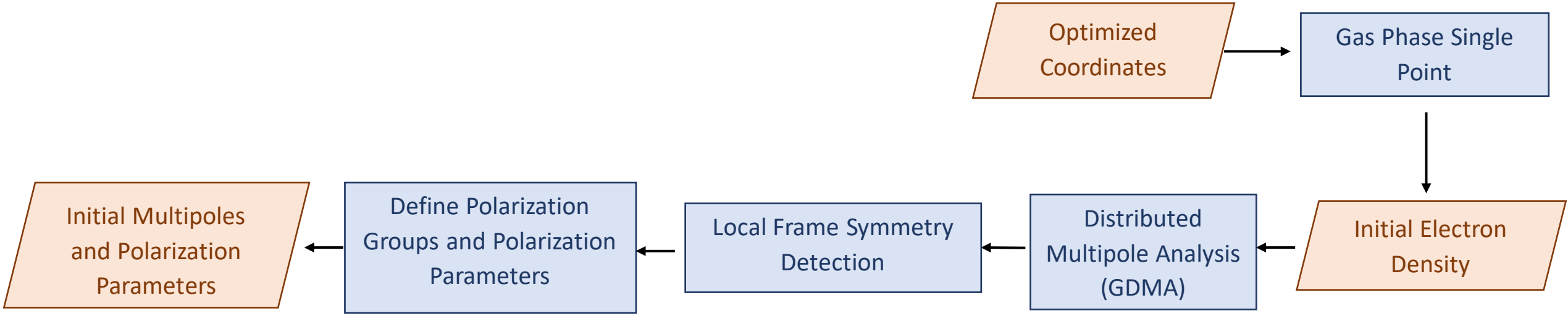
Folded conformation ✗

Rotatable bonds are restrained during initial QM geometry optimization

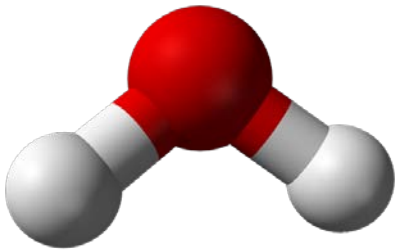
Poltype can optionally include multiple conformations for electrostatic potential fitting

The Fully-Extended Conformation in Peptides and Proteins - Wiley. https://onlinelibrary.wiley.com/doi/10.1002/bip.23100.

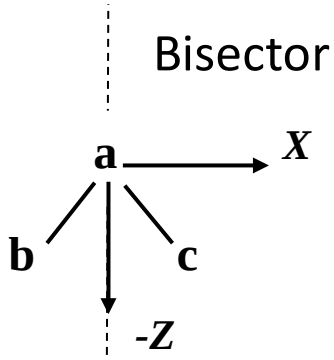
Poltype 2 Flow



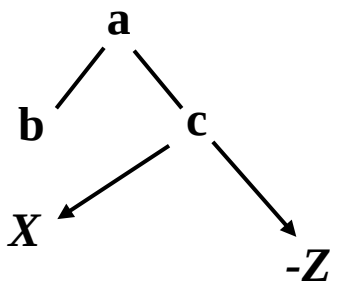
Local Frame Symmetry Detection



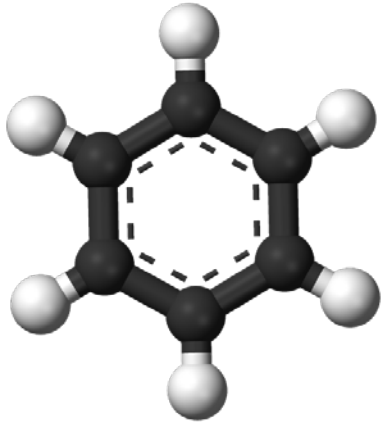
Bisector



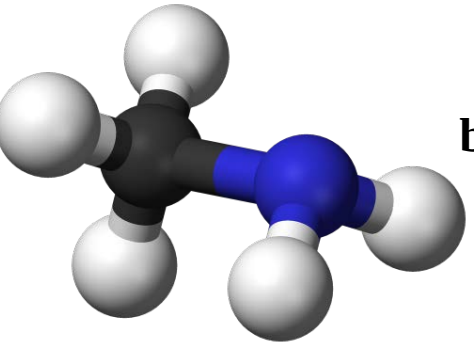
Z-then-X

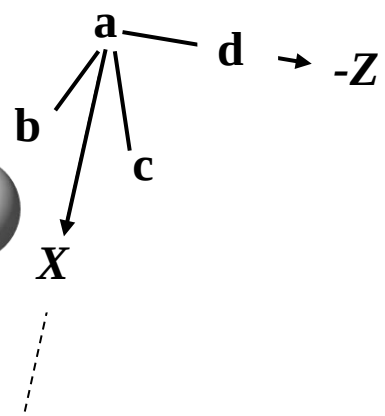


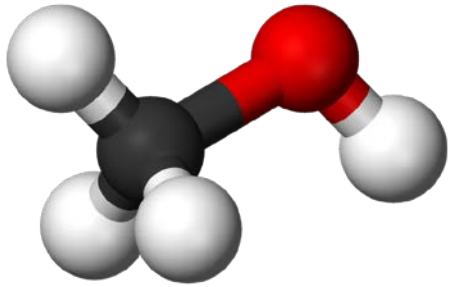
Z-Only



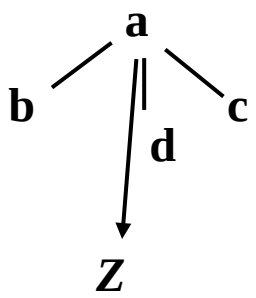
Z-then-bisector

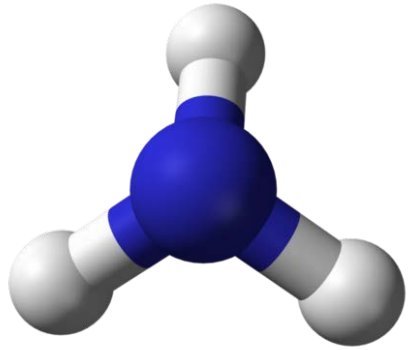






Trisector



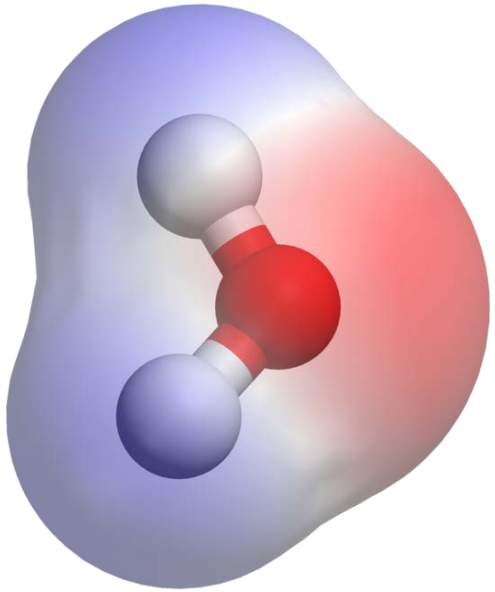


Multipoles require local frame definition! If using correct symmetry, can reduce parameters ☺

Gas Phase Single Point

Electron density grid

QM level of theory MP2/6-311G**



Initial Electron Density

Need to use electron density to obtain initial multipoles!

khaverimkhaverim 3, et al. “Easiest Way (Software) to Visualize Charge Density from an .XYZ File with Point-Charges?” *Chemistry Stack Exchange*, 1 Mar. 1965, https://chemistry.stackexchange.com/questions/76592/easiest-way-software-to-visualize-charge-density-from-an-xyz-file-with-point.

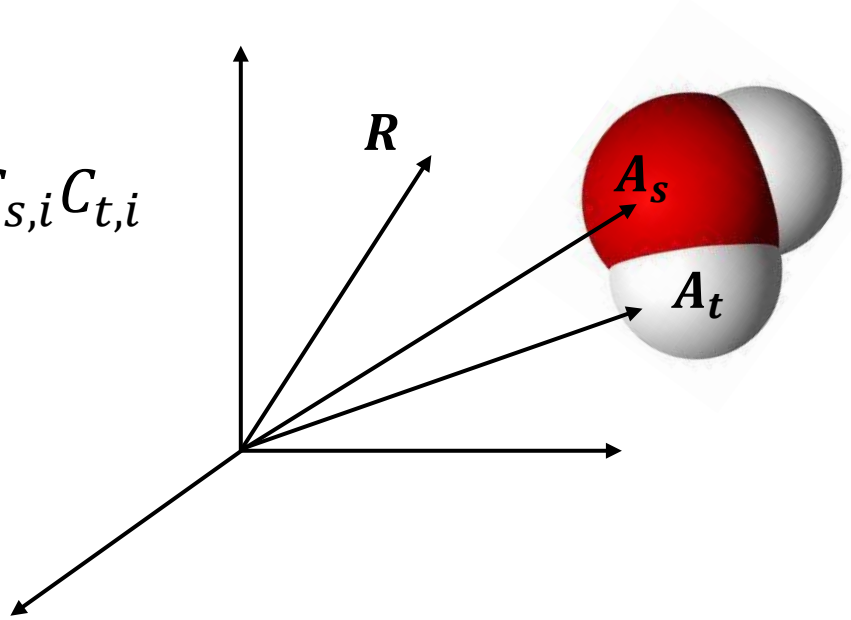
Gaussian Distributed Multipole Analysis (GDMA)

$$\phi(\boldsymbol{R}) = \sum_{s=1}^{N_{basis}} C_s \chi_{\boldsymbol{s}}(\boldsymbol{R} - \boldsymbol{A_s})$$

$$P_{s,t} = \sum_{i=1}^{N_{occupied}} C_{s,i} C_{t,i}$$

$$\chi_{\boldsymbol{s}}(\boldsymbol{R} - \boldsymbol{A_s}) = N_\chi R_{l,m}(\boldsymbol{R} - \boldsymbol{A_s}) e^{-\zeta(\boldsymbol{R}-\boldsymbol{A_s})^2}$$

Spherical harmonics



$$\rho(\boldsymbol{R}) = \sum_{s}^{N_{basis}} \sum_{t}^{N_{basis}} P_{s,t} \chi_{\boldsymbol{s}}(\boldsymbol{R} - \boldsymbol{A_s}) \chi_{\boldsymbol{t}}(\boldsymbol{R} - \boldsymbol{A_t})$$

$$Q_{l,m} = \int \rho(\boldsymbol{R}) R_{l,m}(\boldsymbol{R} - \boldsymbol{A_s})\, d^3\boldsymbol{R}$$

$$M_i = [q_i, d_{ix}, d_{iy}, d_{iz}, Q_{ixx}, Q_{ixy}, Q_{ixz}, Q_{iyx}, Q_{iyy}, Q_{iyz}, Q_{izx}, Q_{izy}, Q_{izz}]$$

Diffuse basis functions extend across many atoms causing issues with how individual atomic contributions are divided on the total grid…

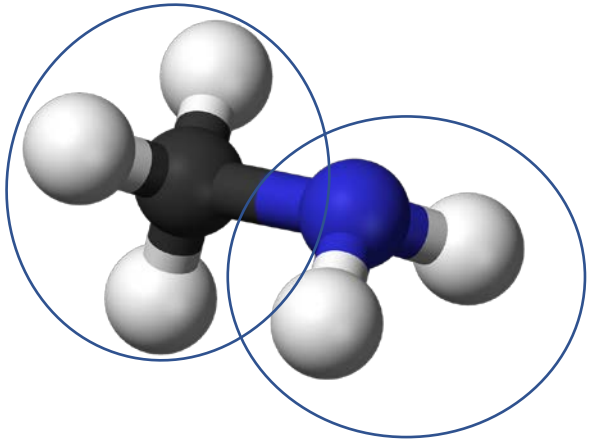
This is why we have to do electrostatic potential fitting!!

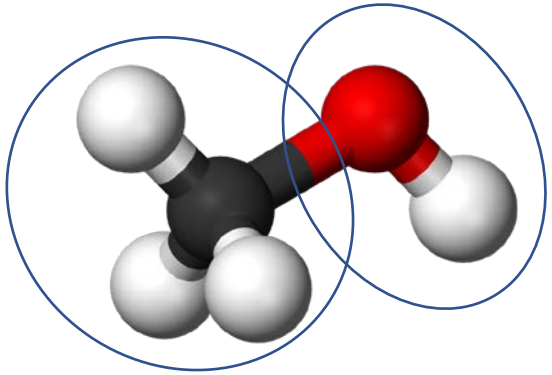
Define Polarization Groups and Polarization Parameters

Need to define polarization groups and parameters to partition the $M_i^{perm-ele}$ and $M_i^{ind-ele}$ component from total GDMA M_i

$$\mu_i^{ind} = \alpha_i \left(\sum_{j,j \neq i} T_{ij}^{Thole} M_j + T_{ij}^{Thole} \mu_j^{ind} \right)$$

Database representation ⟶ [S][C]





Initial Multipoles and Polarization Parameters

$$M_i = \left[q_i, d_{ix}, d_{iy}, d_{iz}, Q_{ixx}, Q_{ixy}, Q_{ixz}, Q_{iyx}, Q_{iyy}, Q_{iyz}, Q_{izx}, Q_{izy}, Q_{izz}\right]$$

GDMA output → $M_i = M_i^{perm-ele} + M_i^{ind-ele}$

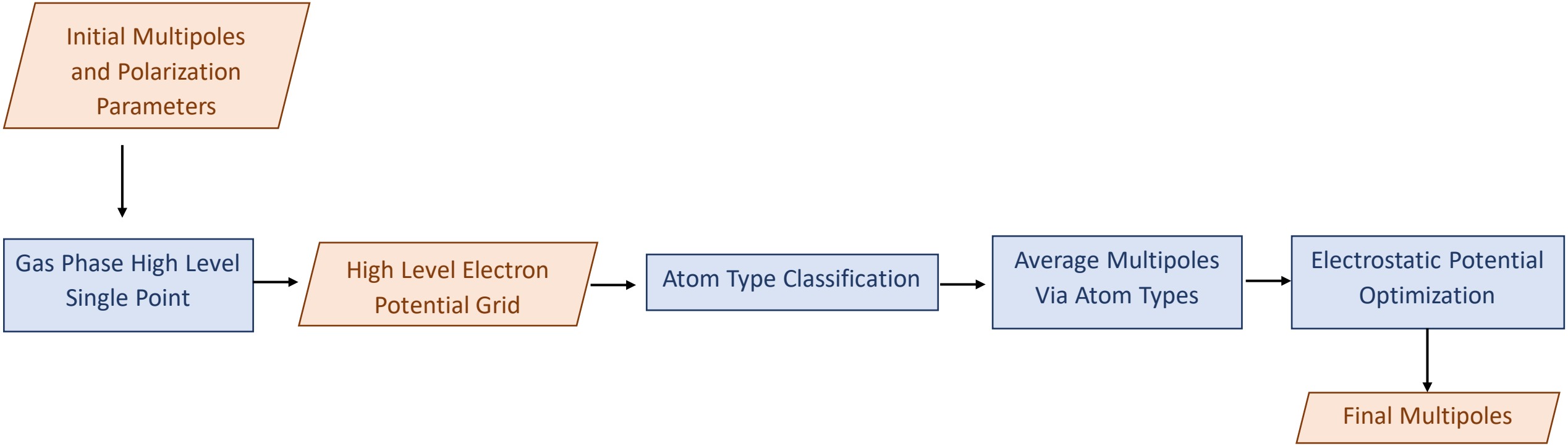
$$M_i^{ind-ele} = \alpha_i \left(\sum_{j, j \neq i} T_{ij}^{Thole} M_j^{perm-ele} + T_{ij}^{Thole} M_j^{ind-ele} \right)$$

GDMA output

$$M_i^{ind-ele} = \alpha_i \left(\sum_{j, j \neq i} T_{ij}^{Thole} (M_j - M_j^{ind-ele}) + T_{ij}^{Thole} M_j^{ind-ele} \right)$$

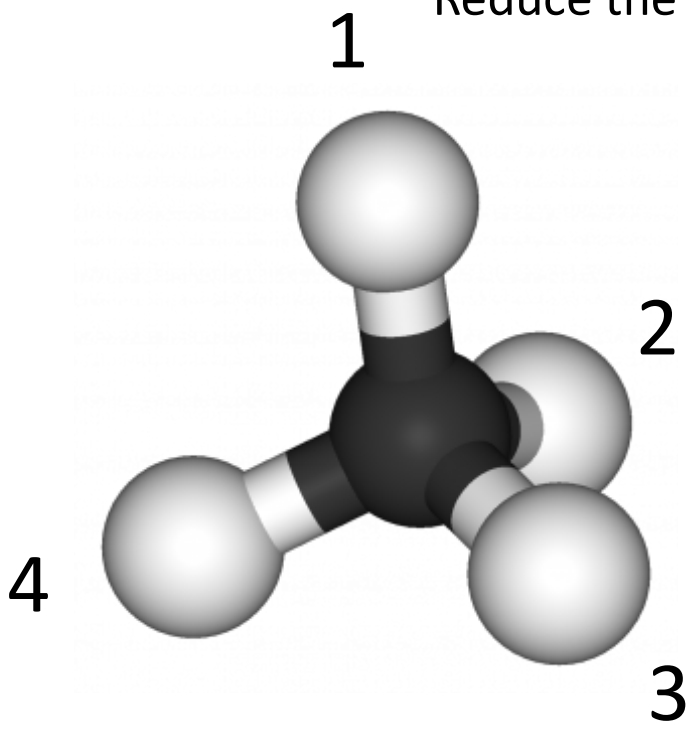
Initial multipole parameters → $M_i^{perm-ele} = M_i - M_i^{ind-ele}$

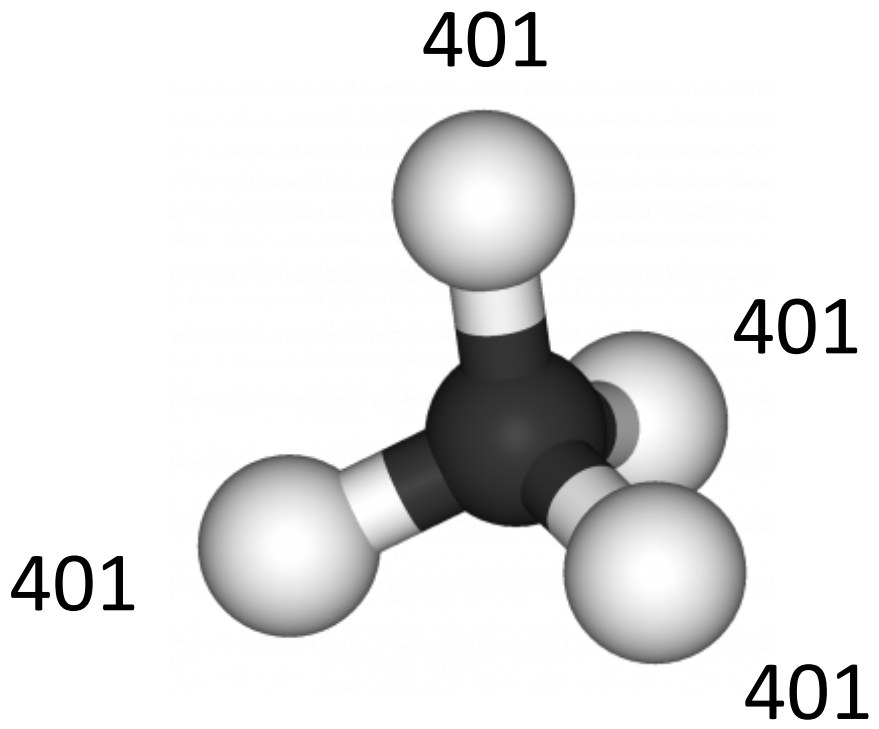
Poltype 2 Flow



Atom Type Classification

Reduce the total number of types needed to fit if possible!



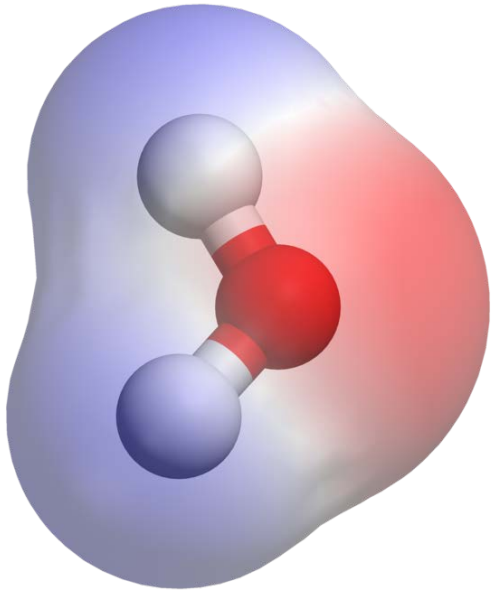


Similar electrostatic environment around atoms of the same ***type***

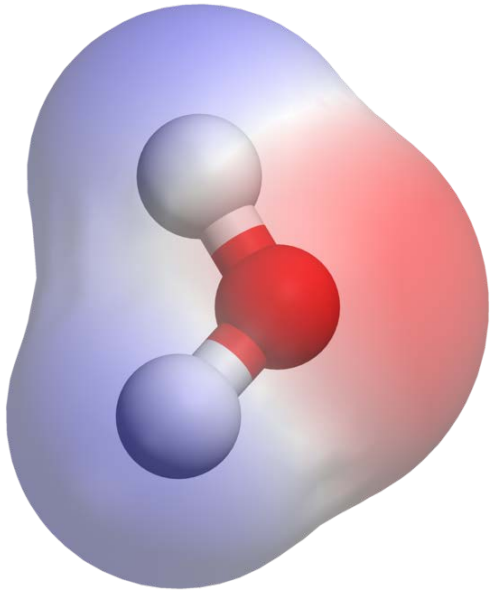
Use atomic number, graph distance, ring membership to determine symmetry

Gas Phase High Level Single Point

AMOEBA potential grid

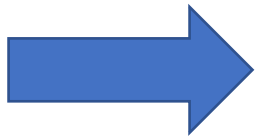


QM potential grid



QM level of theory MP2/aug-cc-pvtz

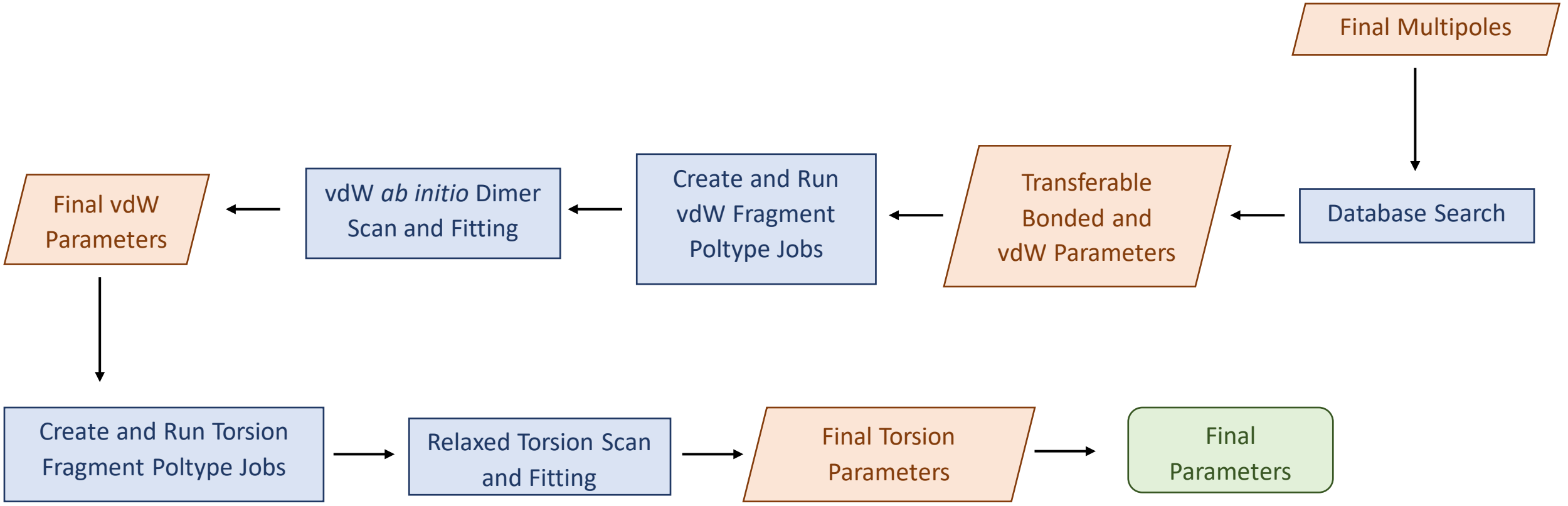
Fit dipole, quadrupole



Final Multipoles

khaverimkhaverim 3, et al. “Easiest Way (Software) to Visualize Charge Density from an .XYZ File with Point-Charges?” *Chemistry Stack Exchange*, 1 Mar. 1965, https://chemistry.stackexchange.com/questions/76592/easiest-way-software-to-visualize-charge-density-from-an-xyz-file-with-point.

Poltype 2 Flow

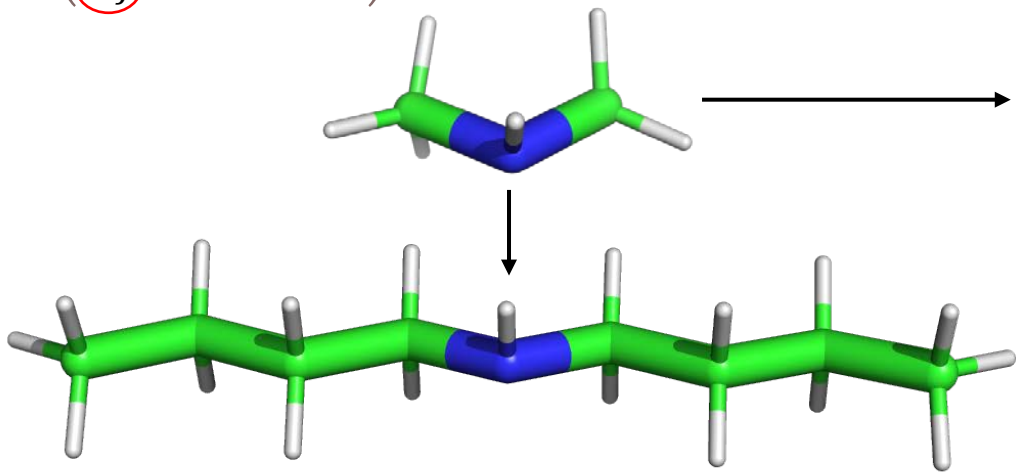


Database search

$$U_{vdw} = \sum_i \sum_{j,j\neq i} \varepsilon_{ij} \left(\frac{1.07}{\rho_{ij} + 0.07} \right)^7 \left(\frac{1.12}{\rho_{ij}^7 + 0.12} - 2 \right)$$

Database parameters

Database representation



[CH3][NH]([CH3])

ε ρ

Search valence and vdW parameters

If a “similar enough” molecule exists in a database, we don’t need to re-parameterize ☺

Similar enough = match all atom elements and neighbors + bond order to database

Database search

Poltype databases:

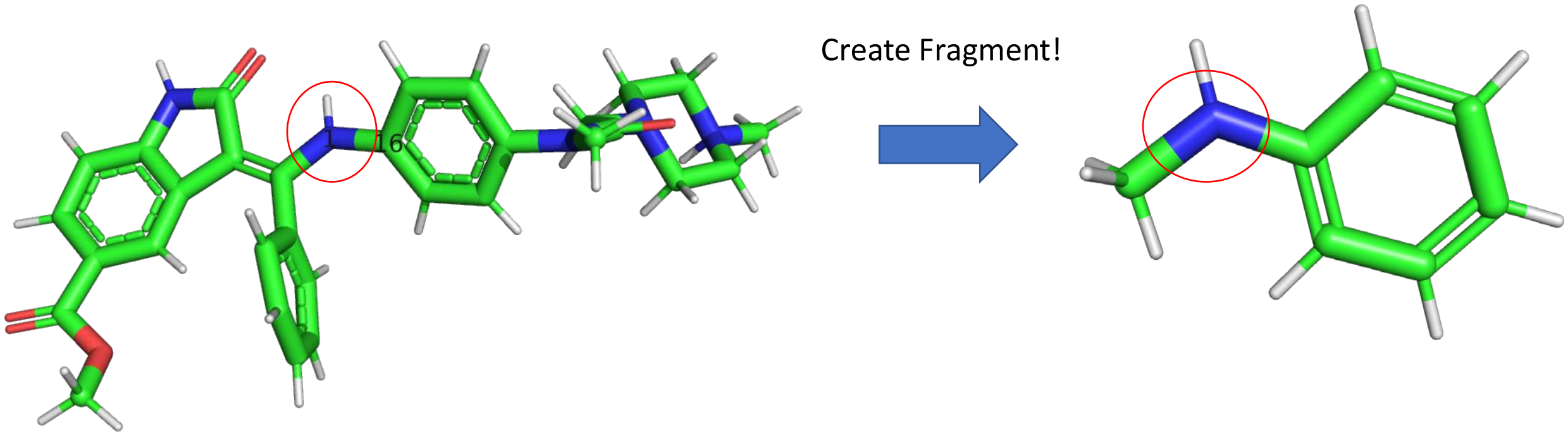
1. Amoeba21 (bond,angle,strbnd,opbend force constants)
2. Amoeba09 (bond,angle,strbnd,opbend, torsion force constants & vdw)
3. Polarizability database
4. Fragment torsion database (torsion)

For bond,angle,strbnd,opbend: amoeba21 >> amoeba09

Missing Parameters:
Torsion -> Fit to QM
Bond,angle -> Use tinker program valence to assign defaults
Strbnd -> Zero out parameters (don't use)
Opbend ->Assign default from MM3 forcefield

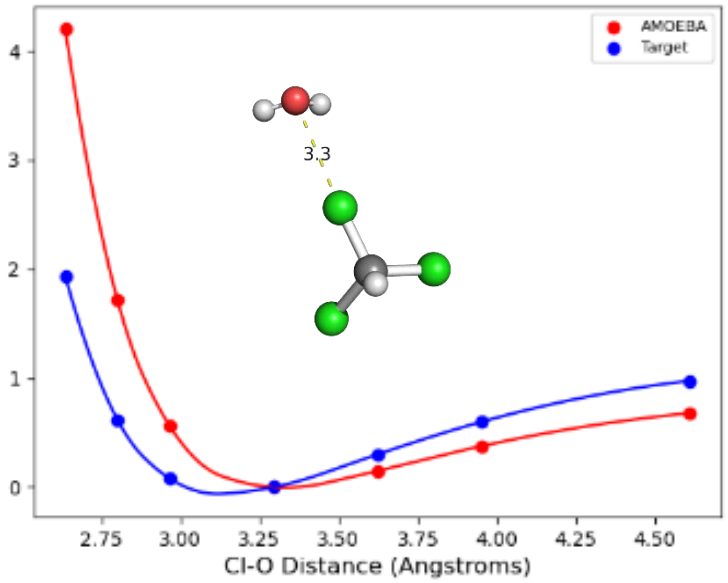
Create and Run vdW Fragment Poltype Jobs

Large floppy molecules can produce steric clashes during fitting procedure ☹



Smaller molecules = less steric clashes and huge decrease in computational cost = ☺

vdW ab initio Dimer Scan and Fitting



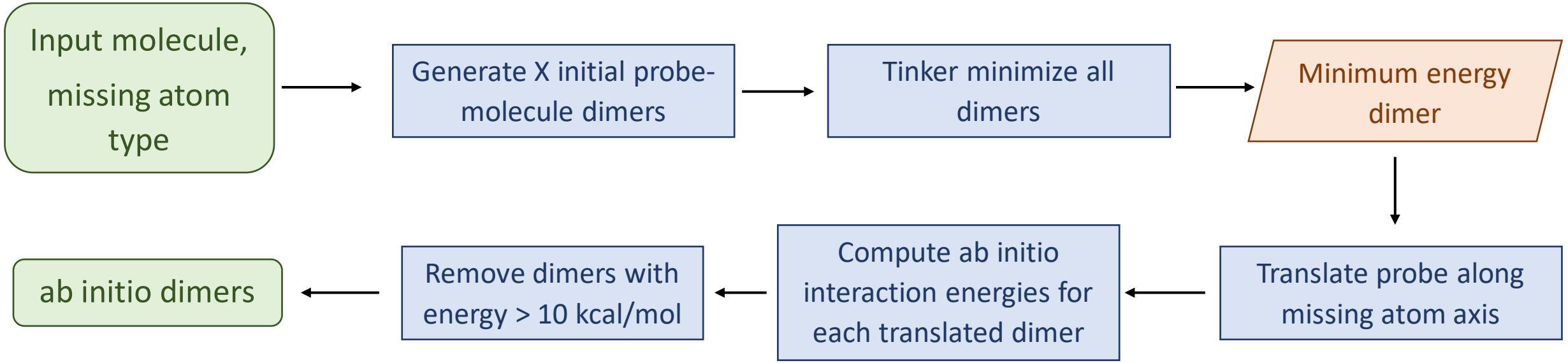
Target=MP2/aug-cc-pV[TQ]Z

↑

CBS extrapolation using triple and quadruple zeta

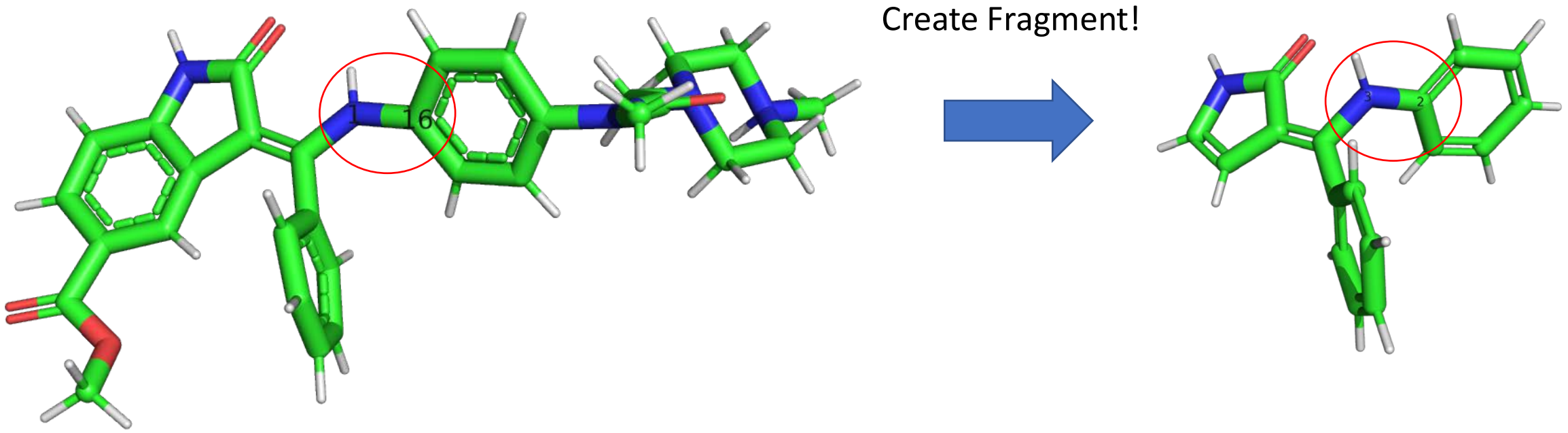
Poltype generates ab initio dimer data and does initial vdW parameter refinement to ab initio data as a starting point for fitting to liquid and ab initio targets

ab initio dimer generation



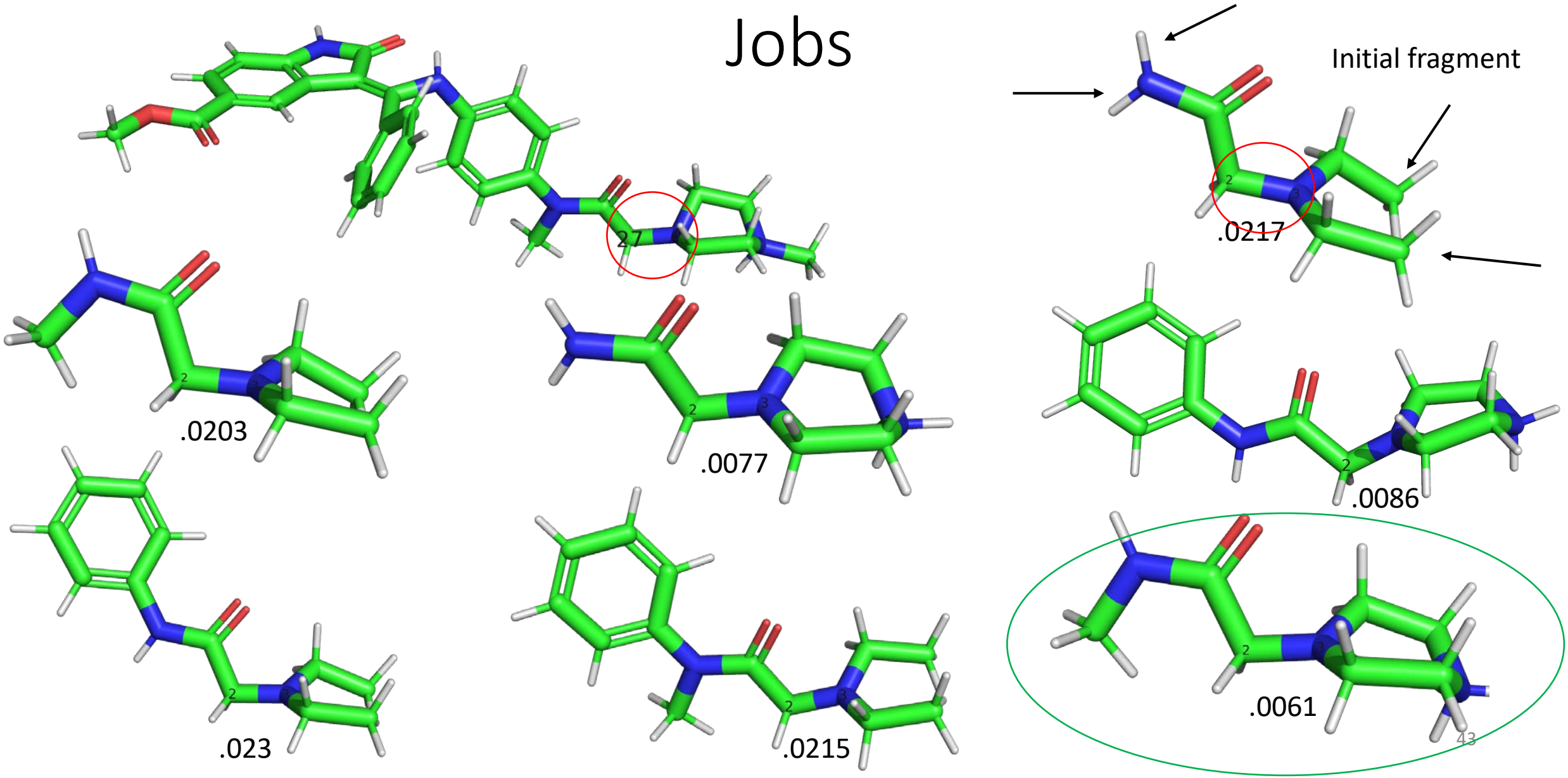
Create and Run Torsion Fragment Poltype Jobs

Large floppy molecules can produce steric clashes during fitting procedure ☹

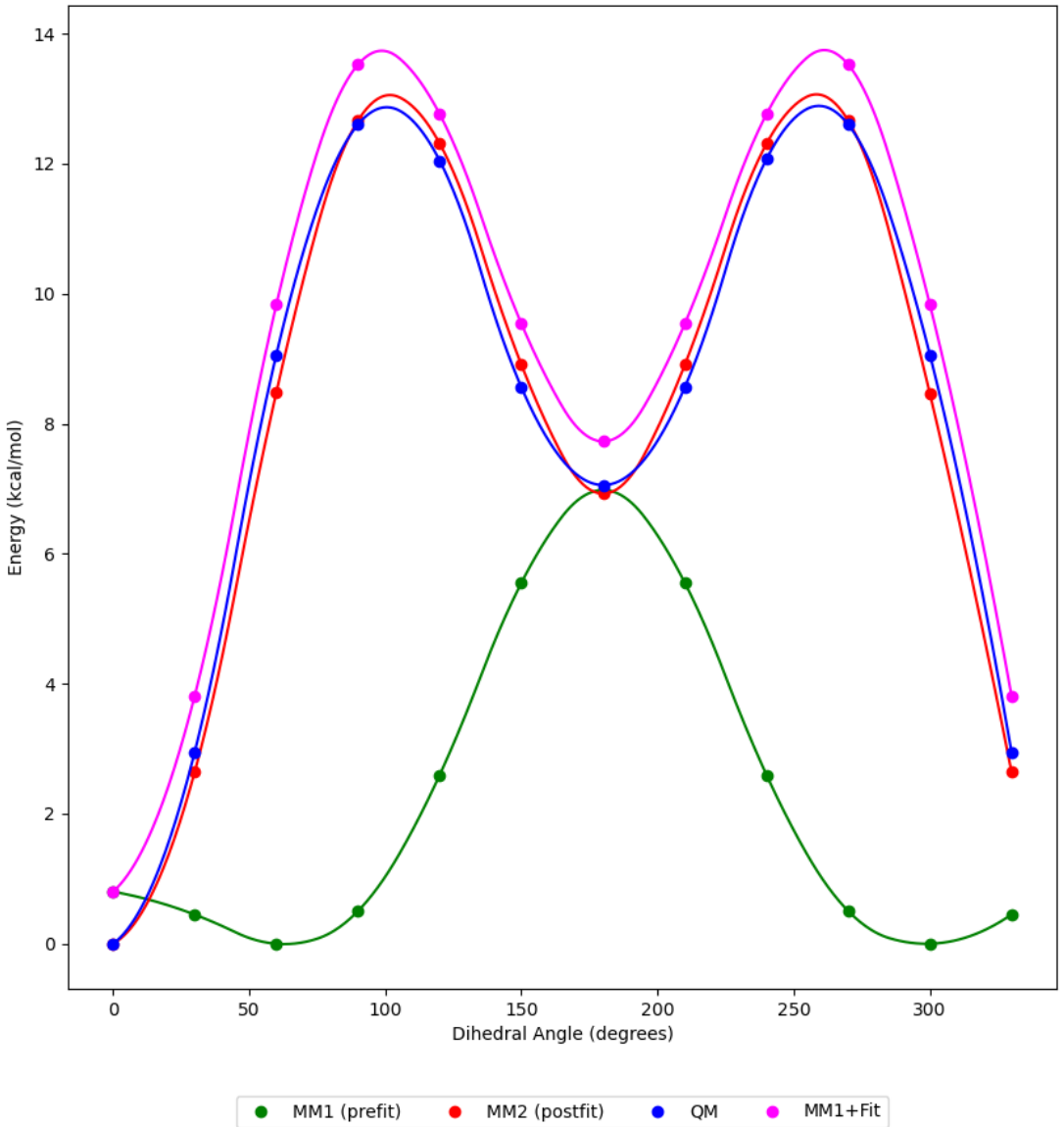


Smaller molecules = less steric clashes and huge decrease in computational cost = ☺

Create and Run Torsion Fragment Poltype Jobs

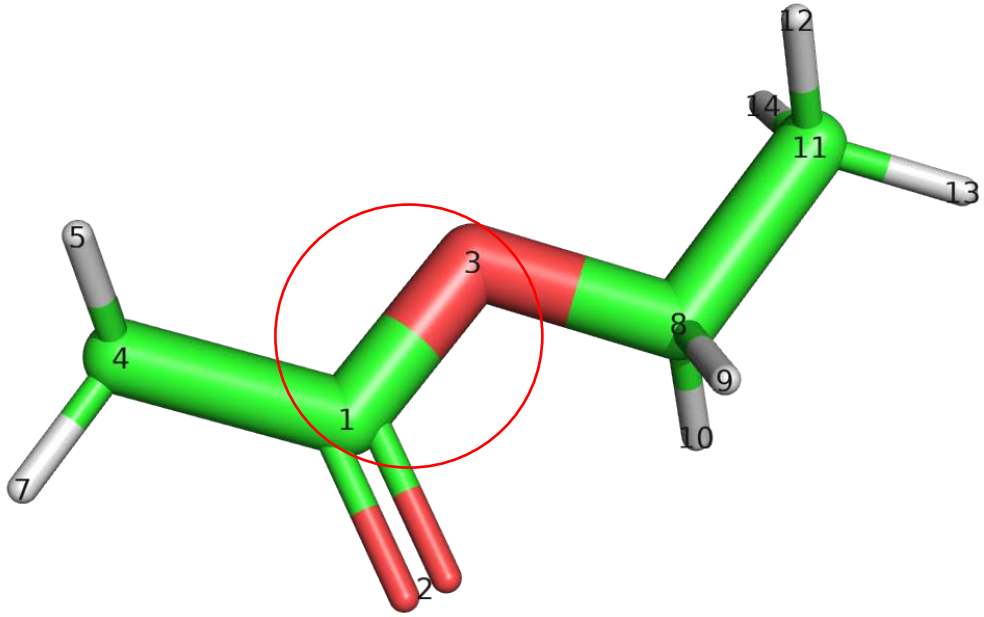


Relaxed Torsion Scan and Fitting



Opt : QM level of theory ωB97X-D/6-311G

SP : QM level of theory ωB97X-D/6-311+G*

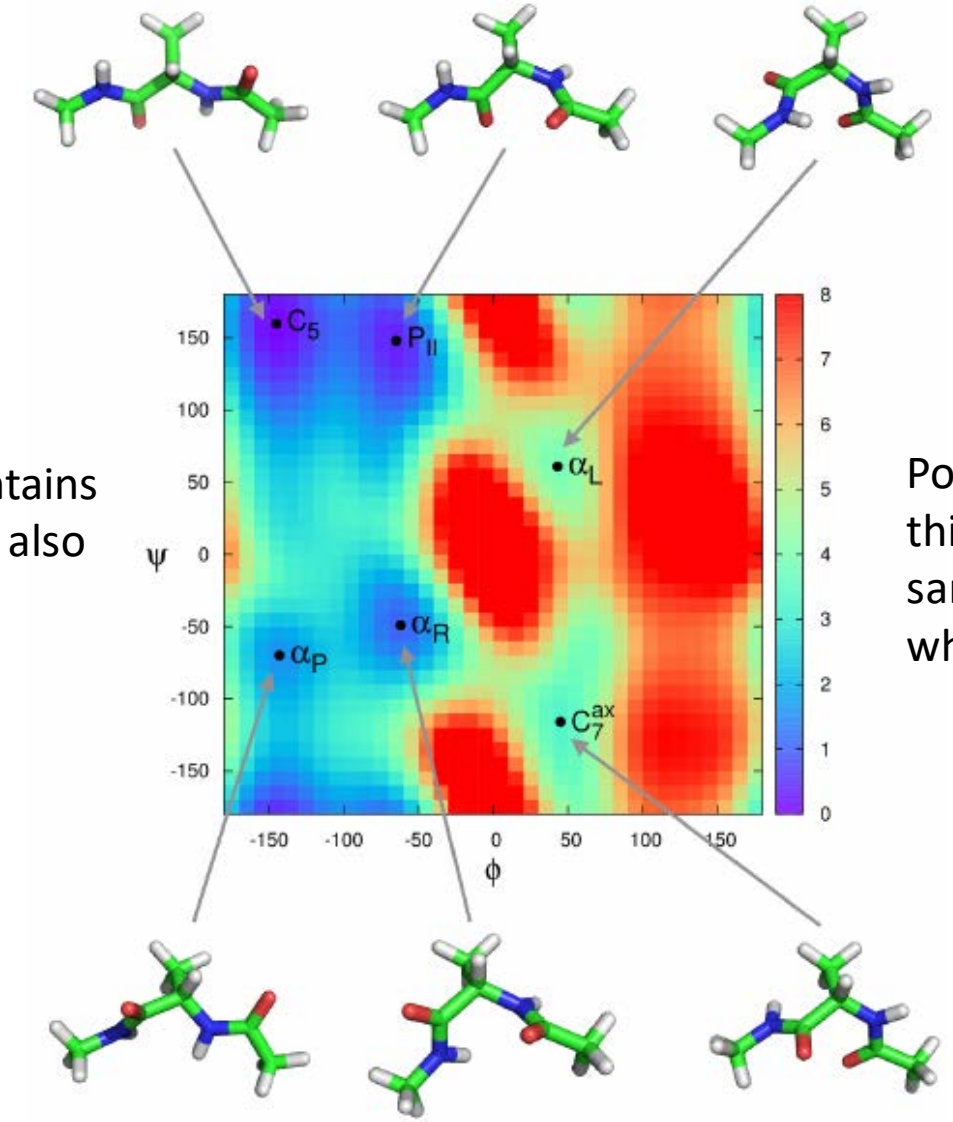


Very expensive to do many rigid scan parameterizations, especially for large molecules!

After fitting can add each fragment molecule to database ☺

Poltype 2 - torsion scanning – 2D

2D conformational energy surface contains more information, but computing it is also more expensive ☹



Poltype can sample which parts of grid it thinks are closer to QM minima and sample around those points instead of whole surface ☺

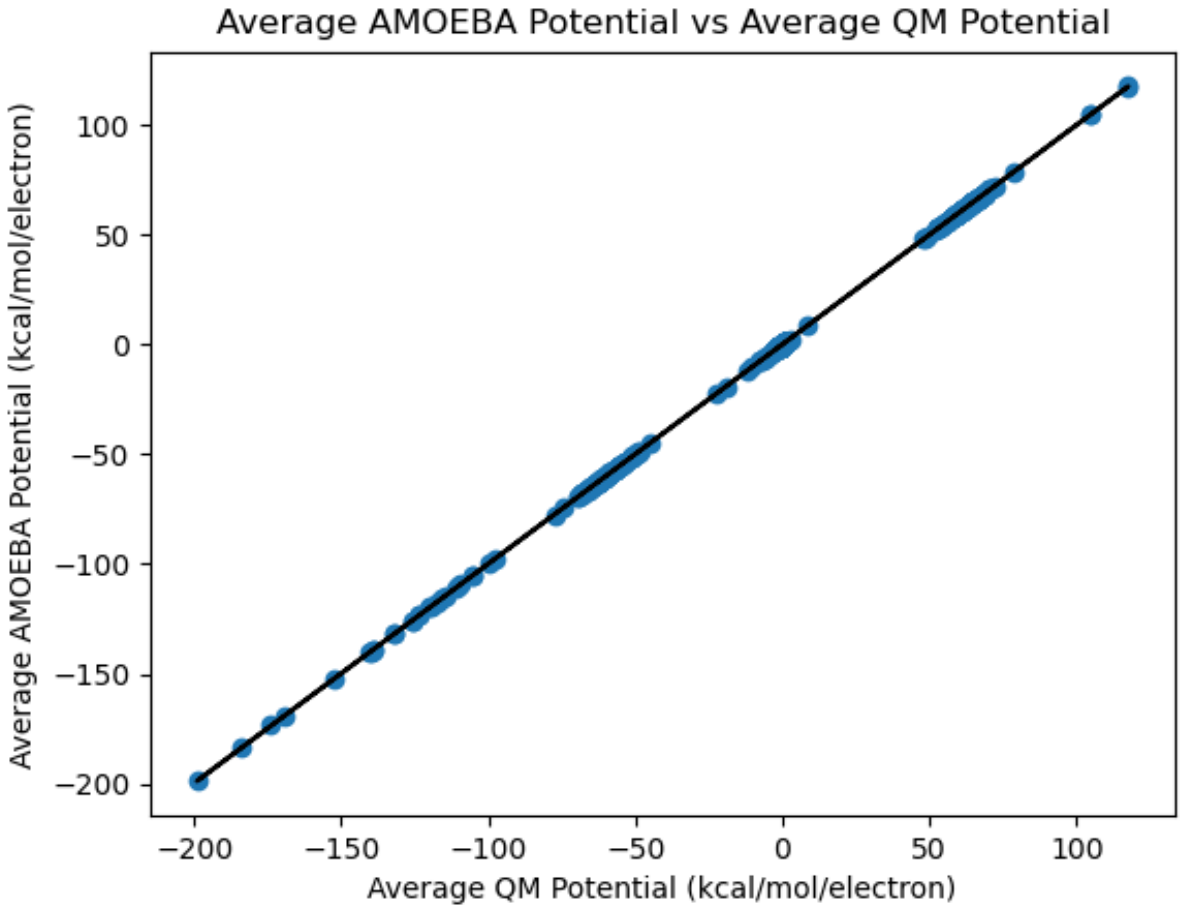
(PDF) a multi-tree approach to compute transition paths on … (n.d.). Retrieved November 11, 2021, from https://www.researchgate.net/publication/250612549_A_multi-tree_approach_to_compute_transition_paths_on_energy_landscapes.

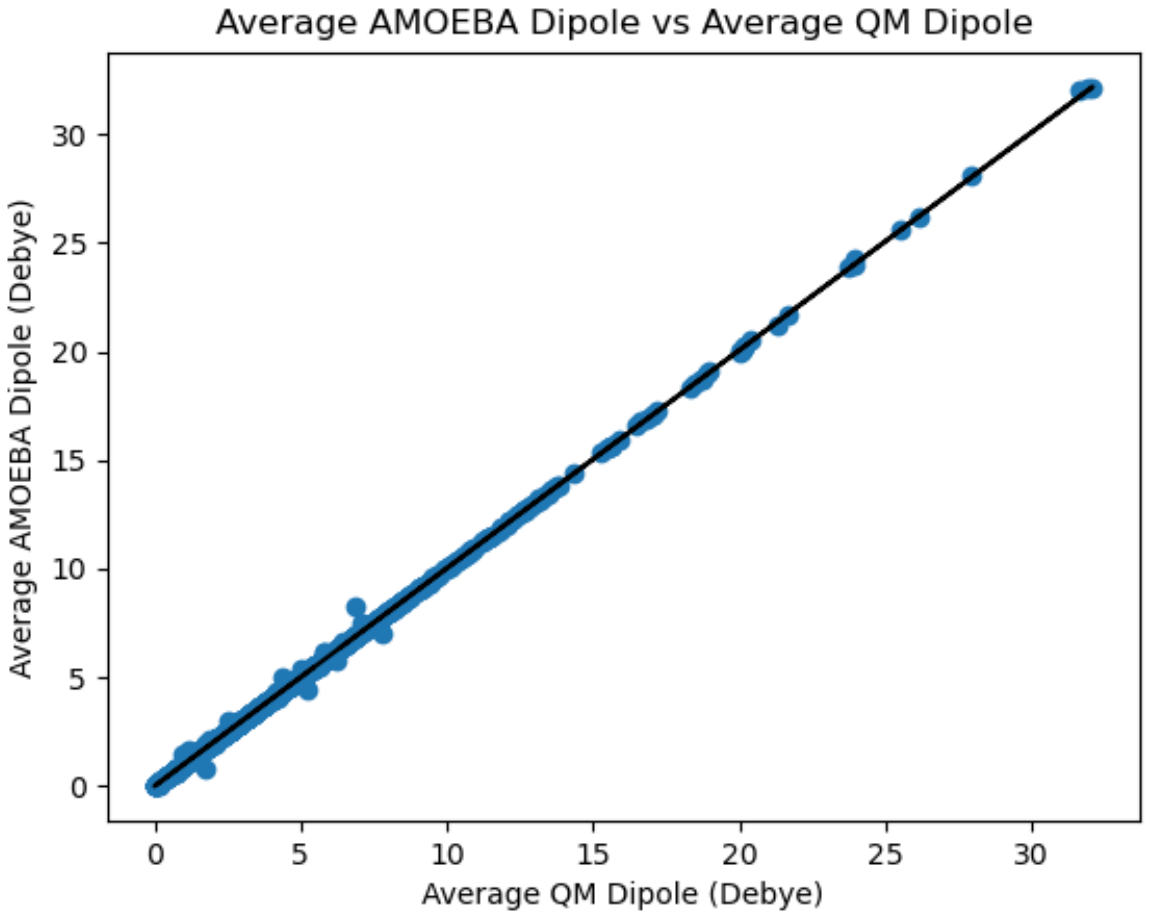
Software validation test set

- Randomly chose ~ 1000 FDA approved small molecules from DrugBank database
- 952/1000 completed Poltype jobs
- 3030 unique fragments out of 4010 total

Fitting results

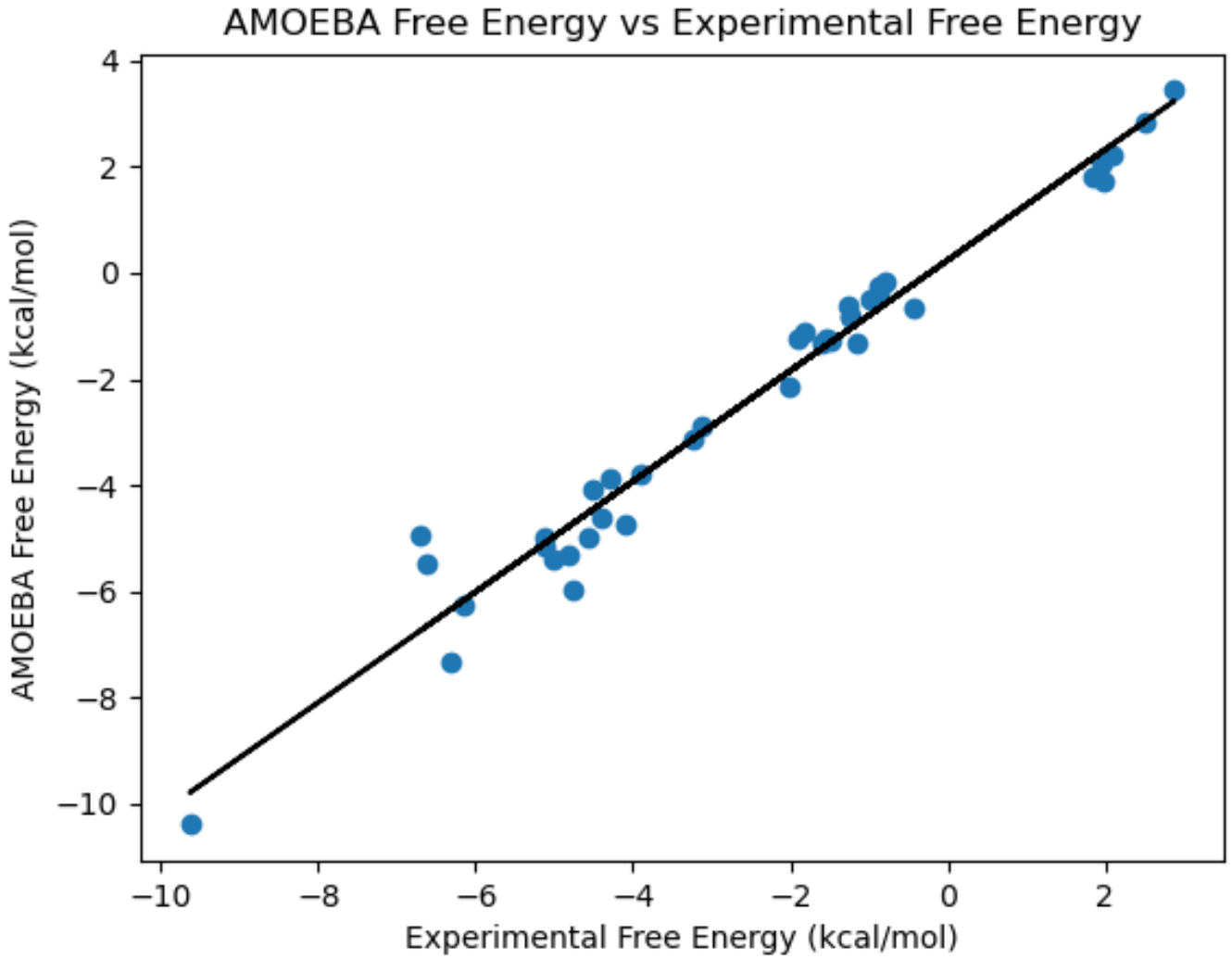
$r^2 \sim .99$





Average torsion fitting RMSE ~ .55 kcal/mol

Hydration free energy data results



Final questions?

References

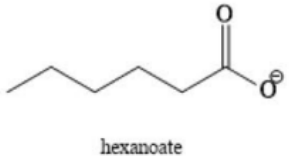
[1] Harger, M., Lee, J.-H., Walker, B., Taliaferro, J. M., Edupuganti, R., Dalby, K. N., & Ren, P. (2019). **Computational insights into the binding of IN17 inhibitors to Melk**. *Journal of Molecular Modeling*, *25*(6). https://doi.org/10.1007/s00894-019-4036-1

[2] Qi, R., Walker, B., Jing, Z., Yu, M., Stancu, G., Edupuganti, R., Dalby, K. N., & Ren, P. (2019). Computational and Experimental Studies of Inhibitor Design for Aldolase A. The journal of physical chemistry. B, 123(28), 6034–6041. https://doi.org/10.1021/acs.jpcb.9b04551

[3] Jing, Z., Qi, R., Thibonnier, M., & Ren, P. (2019). **Molecular dynamics study of the hybridization between RNA and modified oligonucleotides**. Journal of Chemical Theory and Computation, 15(11), 6422–6432. https://doi.org/10.1021/acs.jctc.9b00519

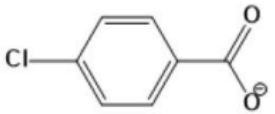
[4] Computational Study on the Binding of Mango-II RNAAptamer and Fluorogen Using the Polarizable Force Field AMOEBA

Accuracy of AMOEBA model - SAMPL7 Contest



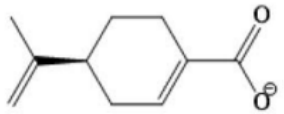
hexanoate

g1



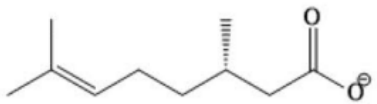
4-chlorobenzoate

g2



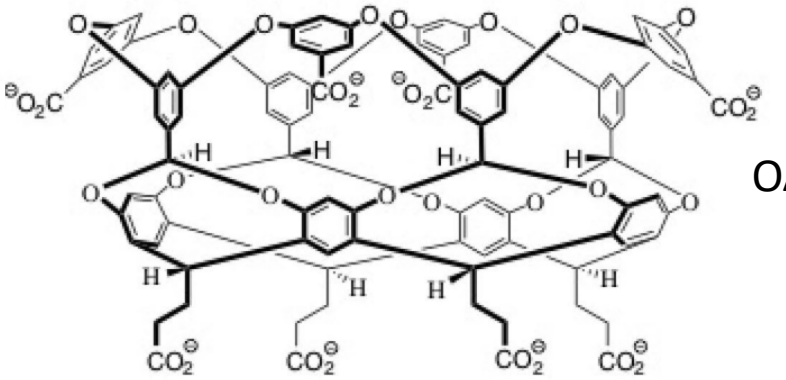
(*S*)-4-(prop-1-en-2-yl)cyclohex-1-enecarboxylate

g3

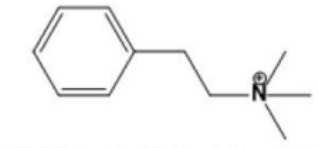


(*S*)-3,7-dimethyloct-6-enoate

g4

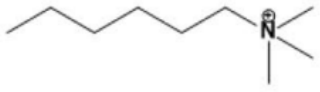


OA



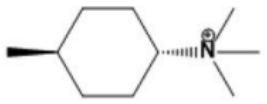
N,*N*,*N*-trimethyl-2-phenylethanaminium

g5



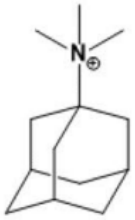
N,*N*,*N*-trimethylhexan-1-aminium

g6



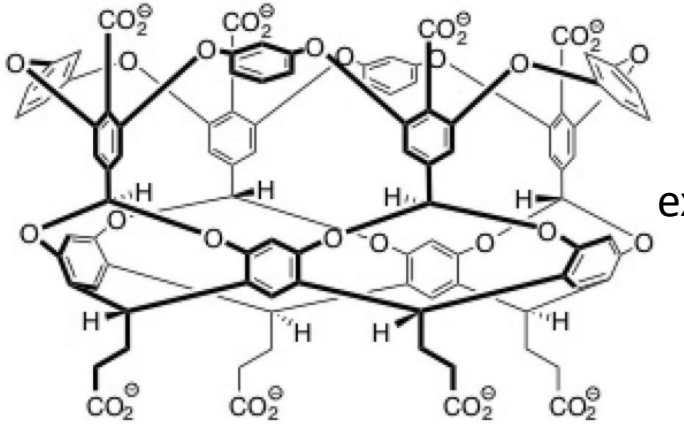
(1*r*,4*r*)-*N*,*N*,*N*,4-tetramethylcyclohexanaminium

g7



(3*s*,5*s*,7*s*)-*N*,*N*,*N*-trimethyladamantan-1-aminium

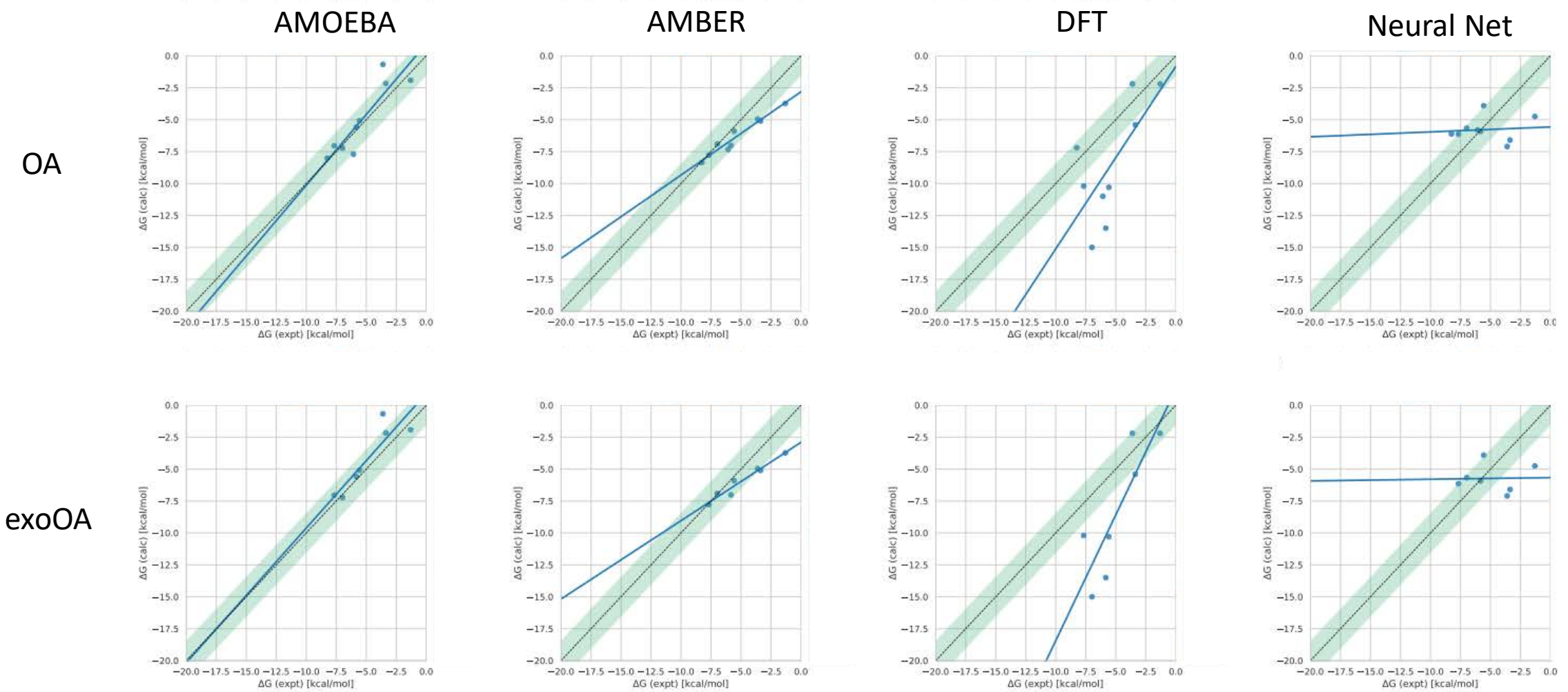
g8



exoOA

Amezcua, M.; Mobley, D. SAMPL7 Challenge Overview: Assessing the Reliability of Polarizable and Non-Polarizable Methods for Host-Guest Binding Free Energy Calculations. **2020**. https://doi.org/10.26434/chemrxiv.12768353.v1

AMOEBA SAMPL7 contest



Amezcua, M.; Mobley, D. SAMPL7 Challenge Overview: Assessing the Reliability of Polarizable and Non-Polarizable Methods for Host-Guest Binding Free Energy Calculations. **2020**. https://doi.org/10.26434/chemrxiv.12768353.v1

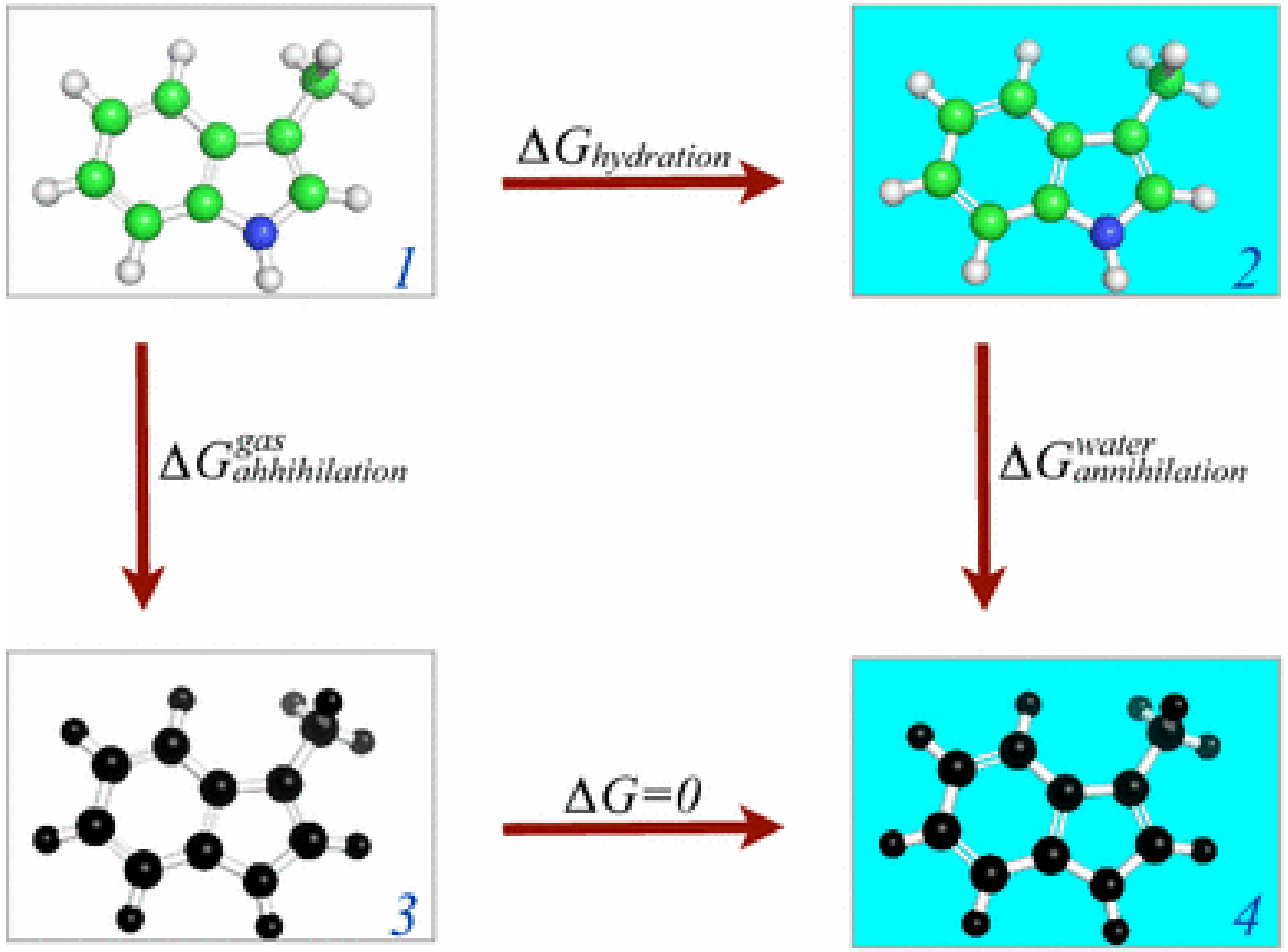
Poltype vs popular parametrization programs

Model	Program	Point charge parameterization	Multipole parameterization	1D torsion scan	Advanced torsion scan	Fragmenter
AMOEBA	Poltype 2		X	X	X	X
AMOEBA+	Poltype 2		X	X	X	X
AMBER	antechamber	X		X		
CHARMM	CGenFF	X		X		
GROMACS	STaGE	X				
OPLS	LigParGen	X				
AMBER?	OpenFF	X		X	X	X
OPLS	Schrodinger	X				

AMOEBA vs popular models

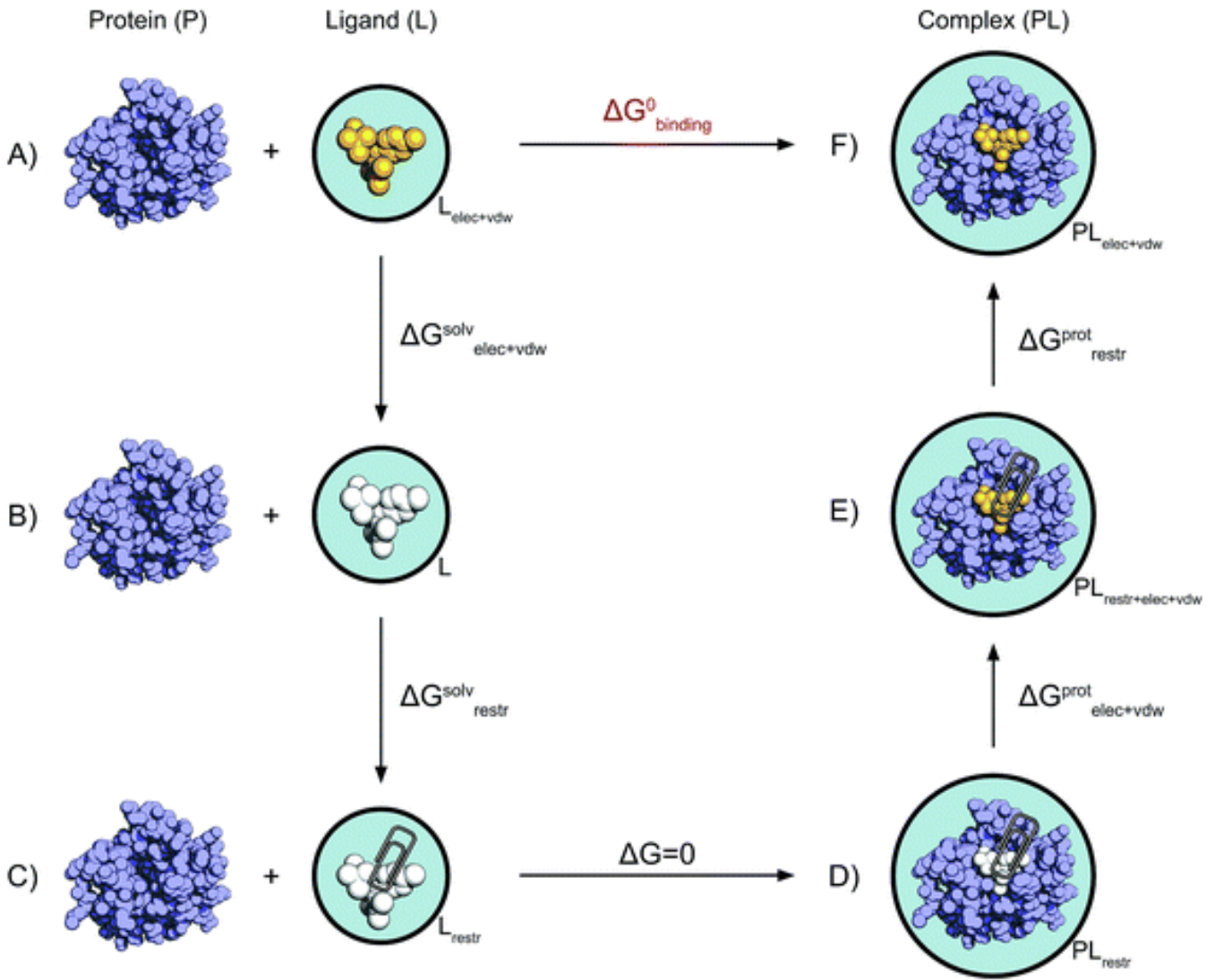
	Point Charge	Multipole Expansion	Polarization	Charge Penetration	Charge Transfer	Experimental Liquid Data/NMR/ X-ray	SAPT Data
AMOEBA		X	X			X	
AMOEBA+		X	X	X	X	X	X
AMBER	X		X			X	
CHARMM	X		X			X	
GROMACS	X		X			X	
Schrodinger	X					X	

Hydration free energy (HFE) thermodynamic cycle



Klimovich, P. V., Shirts, M. R., & Mobley, D. L. (2015, March 26). *Guidelines for the analysis of Free Energy Calculations*. Journal of Computer-Aided Molecular Design. Retrieved November 11, 2021, from https://link.springer.com/article/10.1007/s10822-015-9840-9/.

Binding free energy thermodynamic cycle



Aldeghi, M., Heifetz, A., Bodkin, M. J., Knapp, S., & Biggin, P. C. (2016). Accurate calculation of the absolute free energy of binding for drug molecules. *Chemical Science, 7*(1), 207-218. doi:10.1039/c5sc02678d