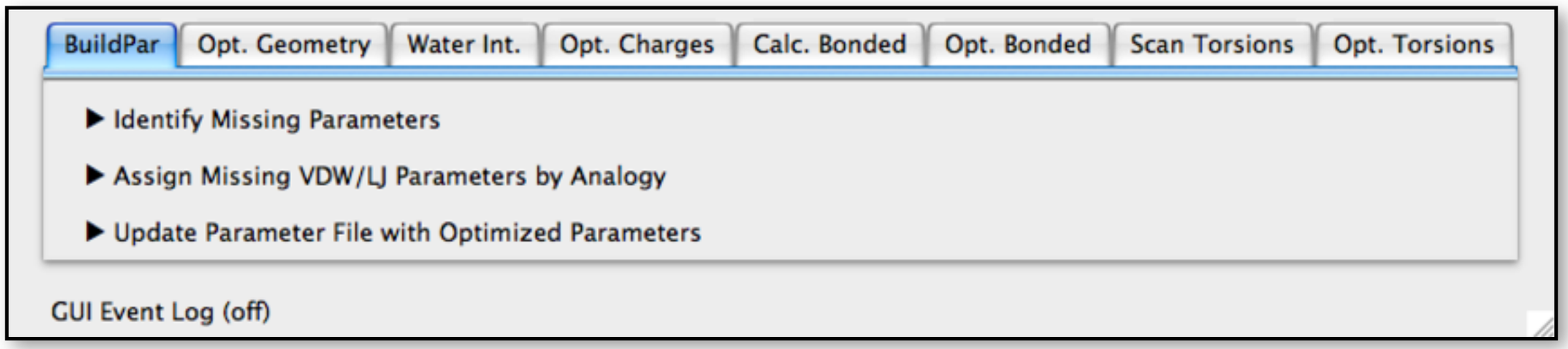
Parameterizing Small Molecules Using:

The Force Field Toolkit (*ff*TK)



Christopher G. Mayne, Emad Tajkhorshid

NIH Center for Macromolecular Modeling and Bioinformatics
Beckman Institute for Advanced Science and Technology
University of Illinois, Urbana-Champaign

James C. Gumbart

Department of Physics
Georgia Institute of Technology

NIH Hands On Workshop | TCBG | April 20th, 2017

MD Simulations of Biological Systems

Molecular Mechanics Force Fields

$$U = \underbrace{U_{bonds} + U_{angles} + U_{dihedrals}}_{\text{bonded}} + \underbrace{U_{vdW} + U_{coulombic}}_{\text{non-bonded}}$$

The CHARMM Force Field

$$U = \sum_{bonds} k_i^{bond}(r_i - r_0)^2 + \sum_{angles} k_i^{angle}(\theta_i - \theta_0)^2 +$$

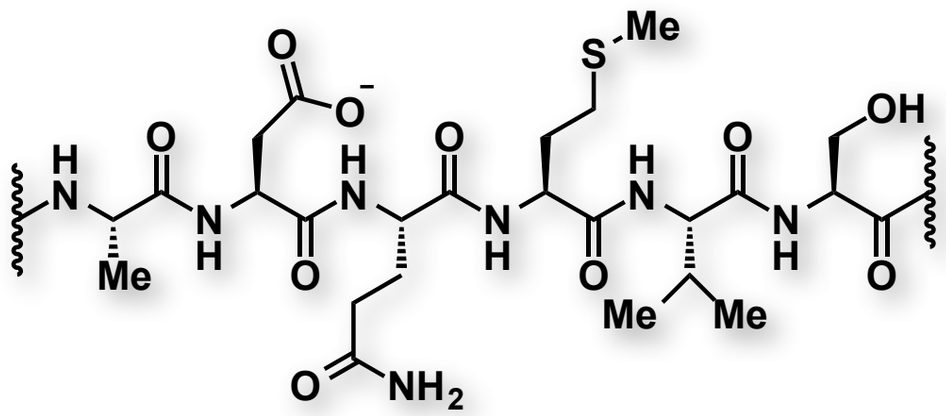
$$\sum_{dihedrals} k_i^{dihedral}[1 + \cos(n_i\phi_i + \delta_i)] +$$

$$\sum_i \sum_{j \neq i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_i \sum_{j \neq i} \frac{q_i q_j}{r_{ij}}$$

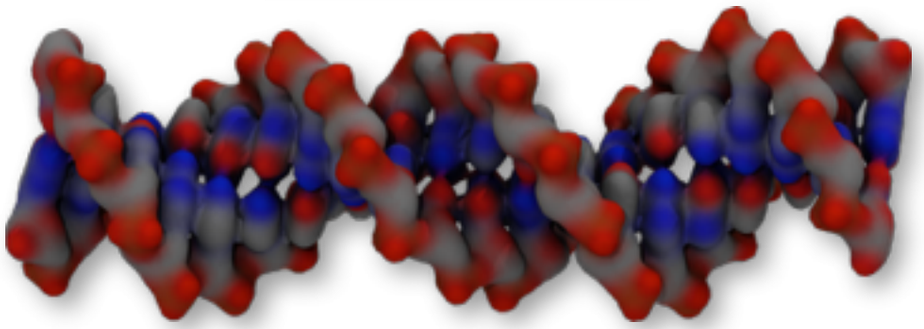
Parameter Transferability In Biopolymers

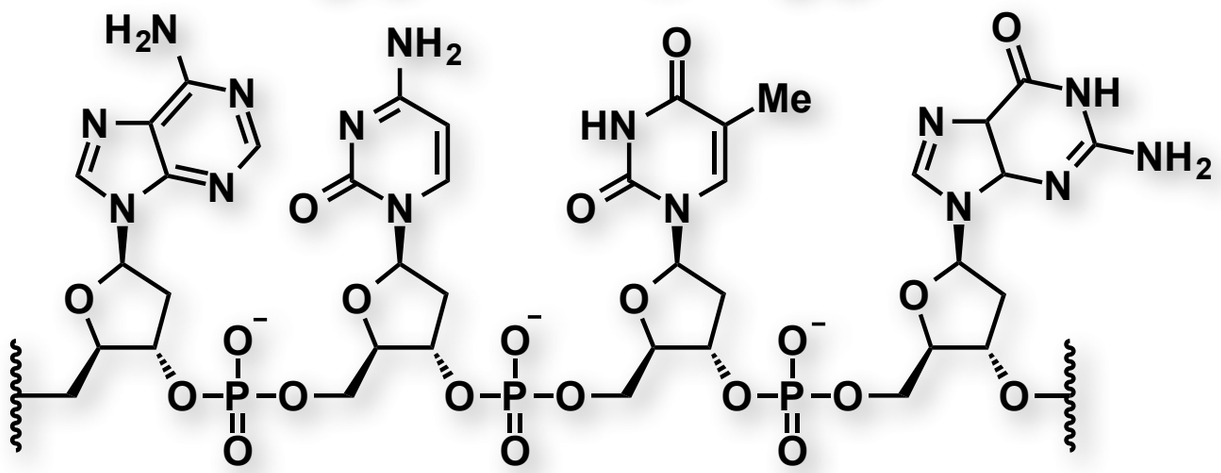
Parameter set describes molecular behavior in varied chemical (connectivity) and spatial (conformation) contexts

Peptides and Proteins

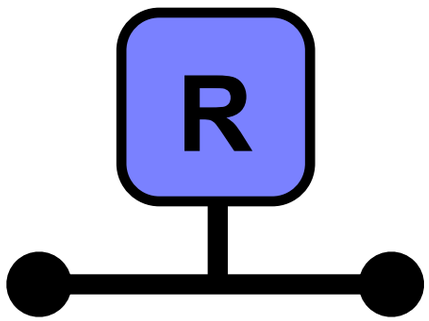


Nucleic Acids





Key Features:

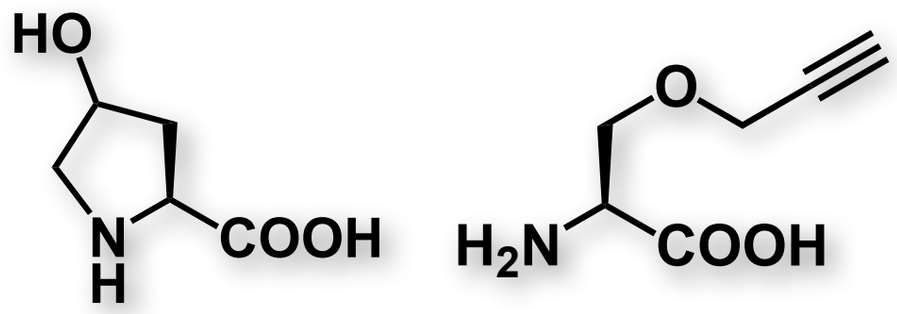


limited set of isolated building blocks

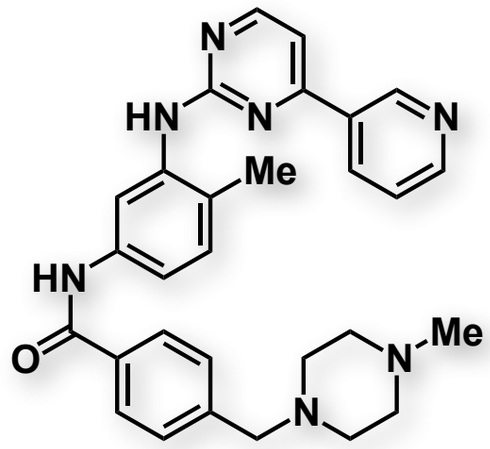
repetitive backbone unit

Parameterization as an Impasse

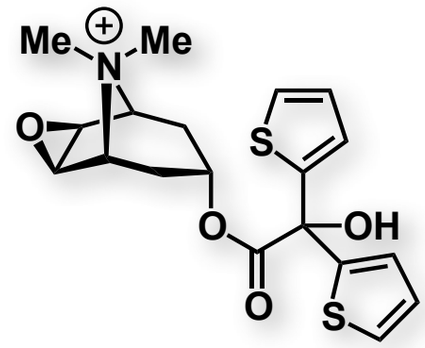
non-standard or engineered amino acids



small molecule ligands

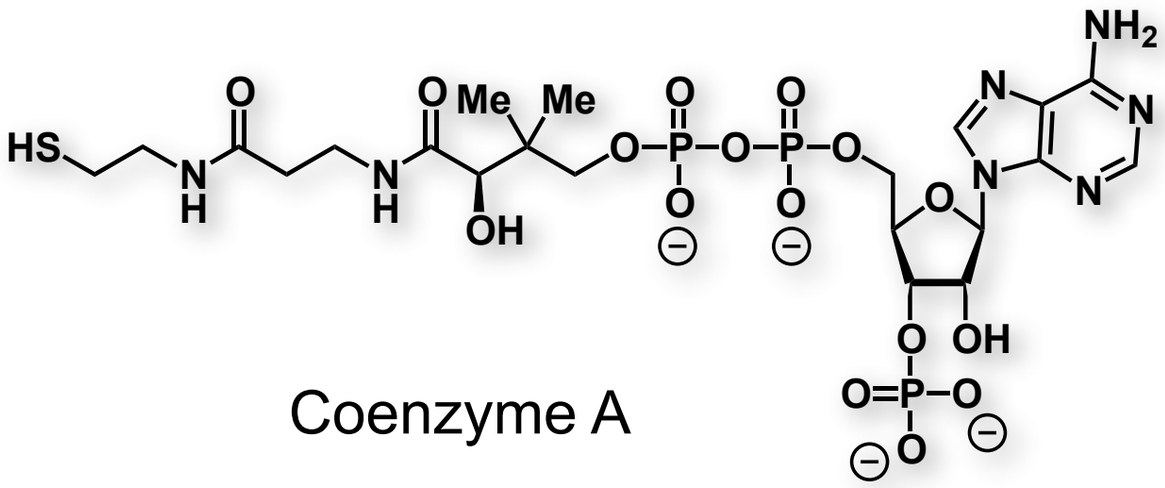


Imatinib (Gleevec)



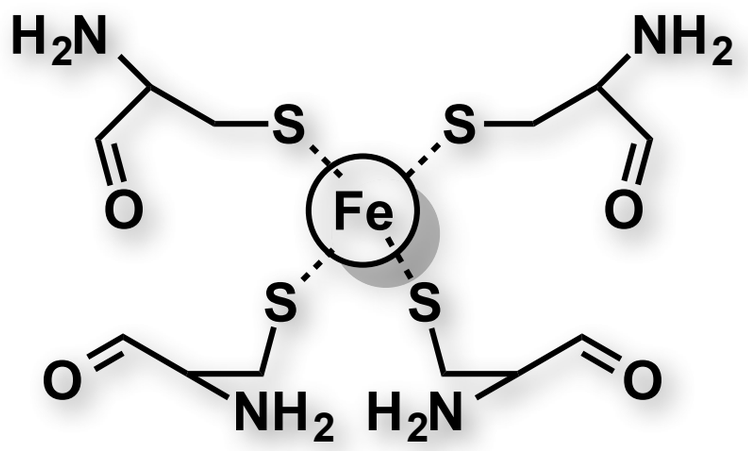
Tiotropium (Spiriva)

cofactors

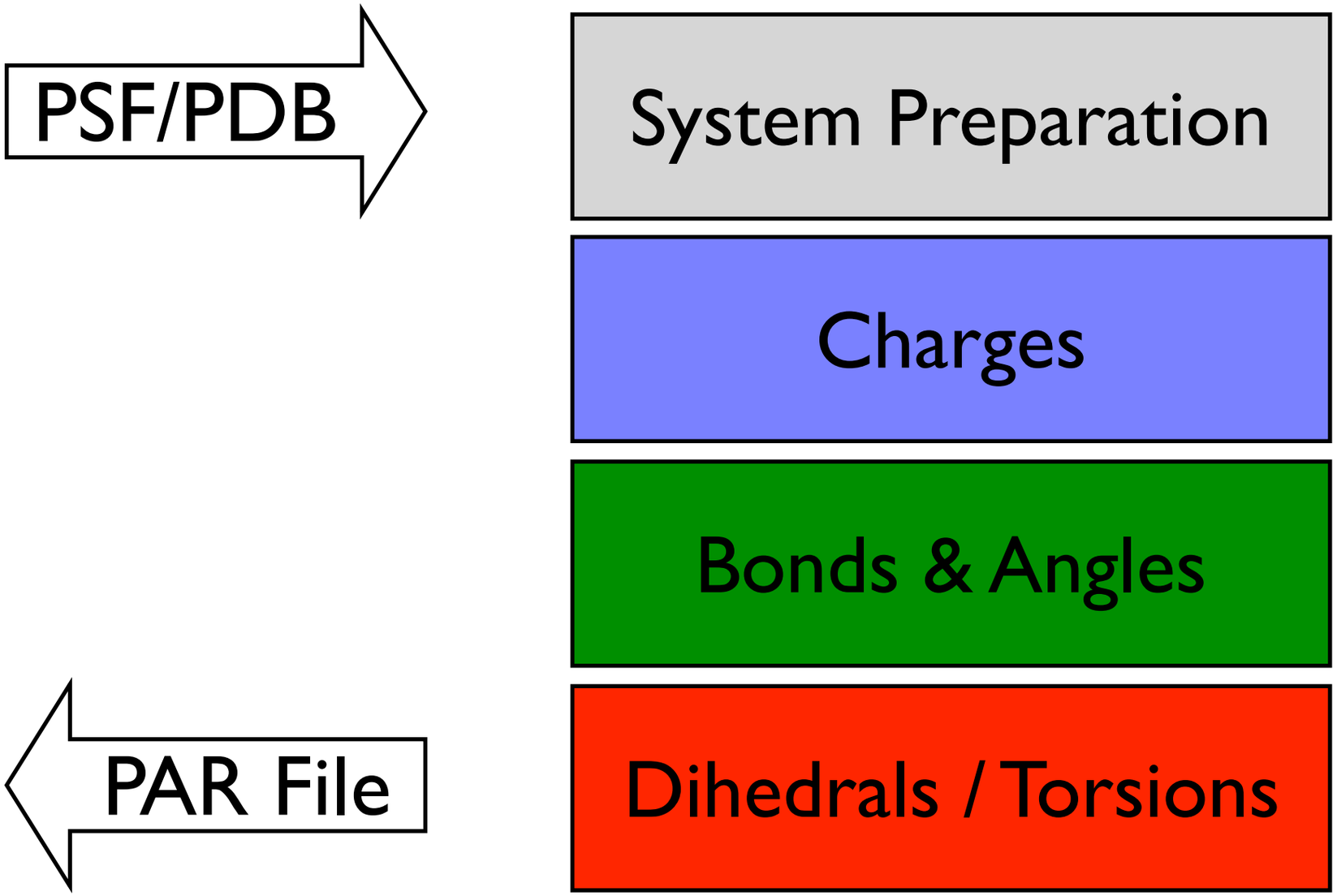


Coenzyme A

metal centers



General Parameterization Workflow



K. Vanommeslaeghe *et al.*, *J. Comput. Chem.* **2010**, 31, 671-690.

CGenFF Parameterization Workflow

PSF/PDB

Find Missing Parameters

Geometry Optimization (QM)

Water Interaction En. (QM)

Charge Optimization

Hessian Calculation (QM)

Bond & Angle Optimization

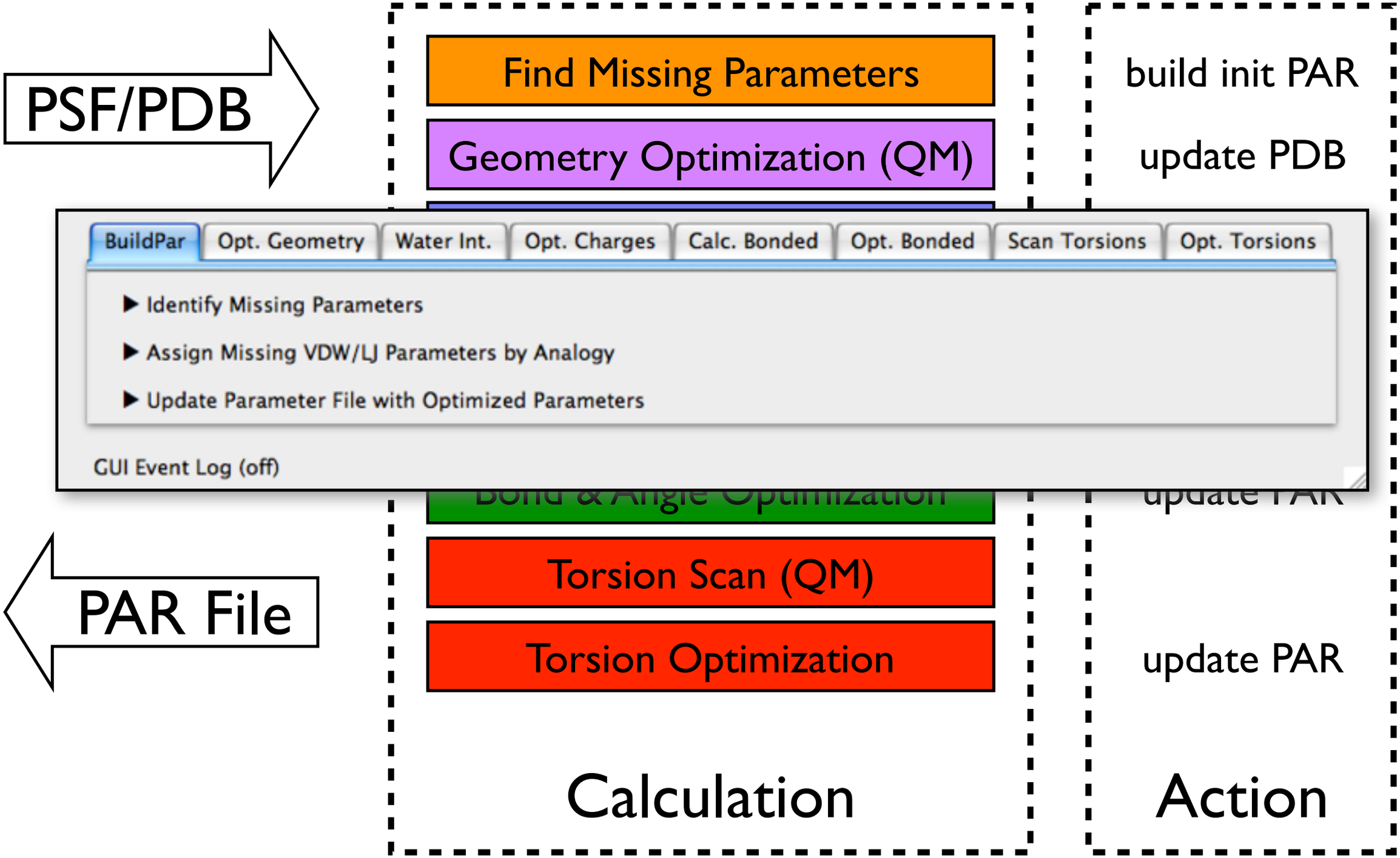
Torsion Scan (QM)

Torsion Optimization

PAR File

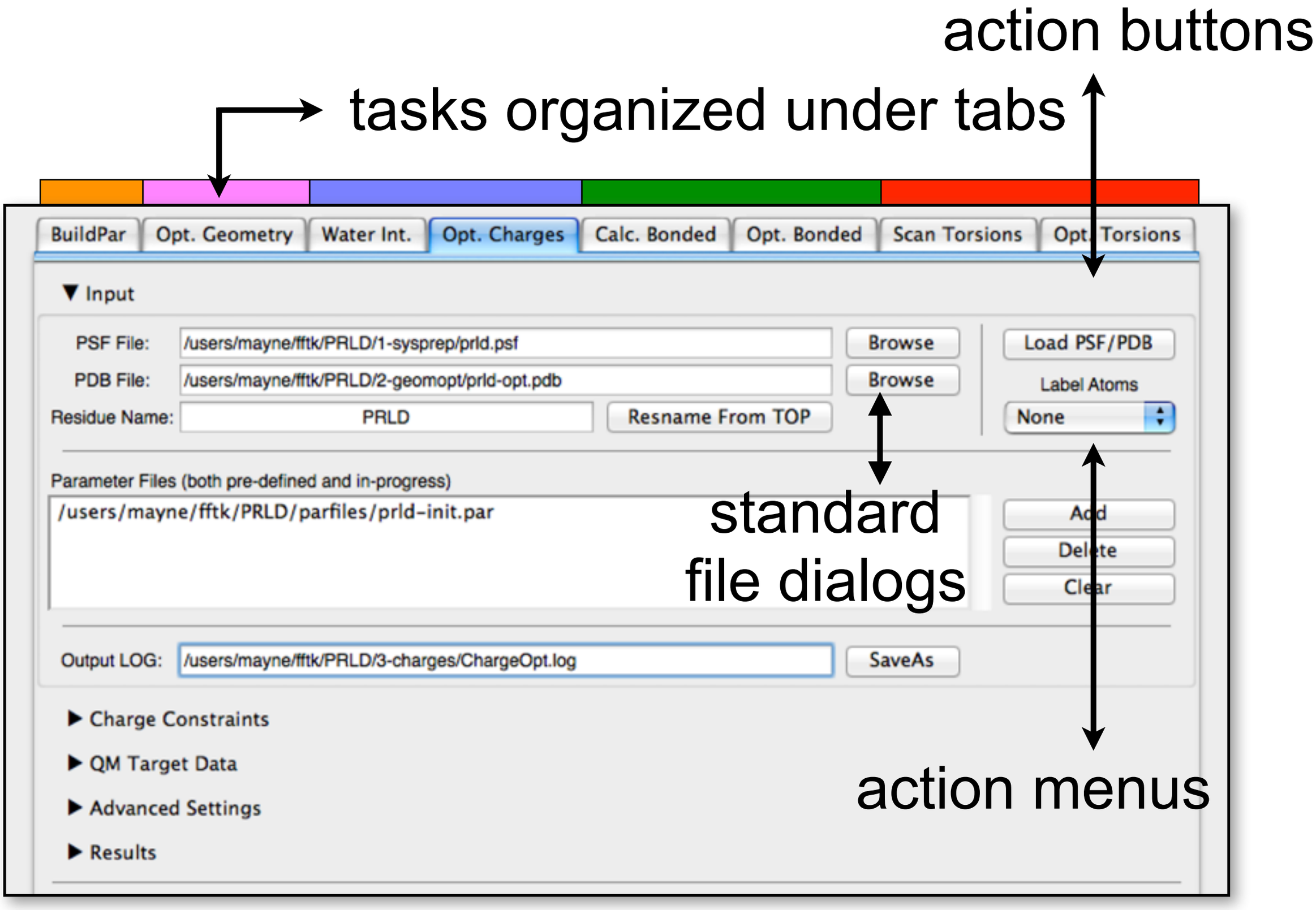
K. Vanommeslaeghe *et al.*, *J. Comput. Chem.* **2010**, 31, 671-690.

CGenFF Parameterization Workflow



C.G. Mayne *et al.*, *J. Comput. Chem.* **2013**, 34, 2757-2270.

*ff*TK Interface



Functionality Provided by *ff*TK

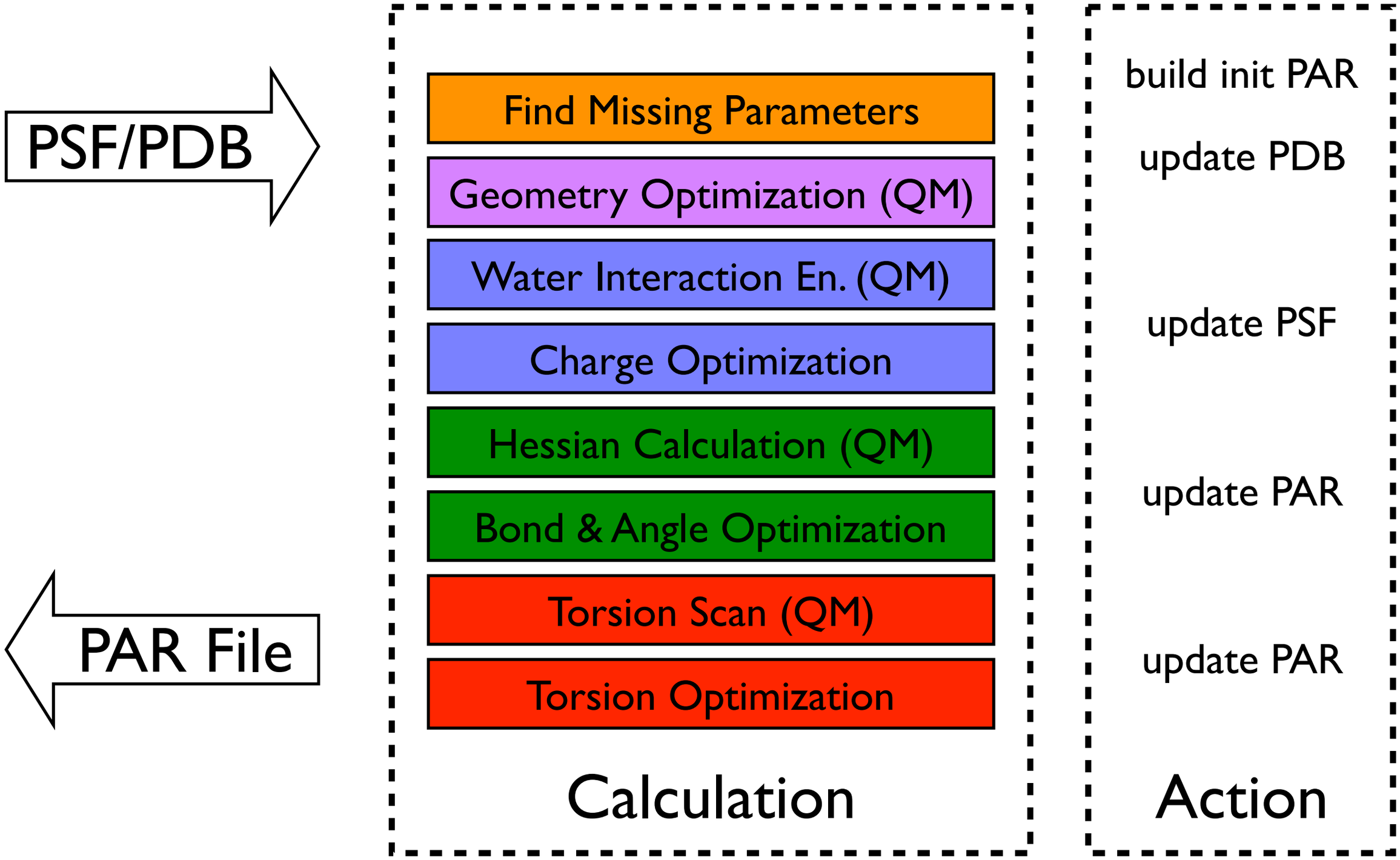
Core Functions

- Setup & Perform Multi-dimensional Optimizations
- Abstraction of Gaussian I/O (QM)
- Assess Performance of Parameters by Visualizing Optimization Data

Support Functions

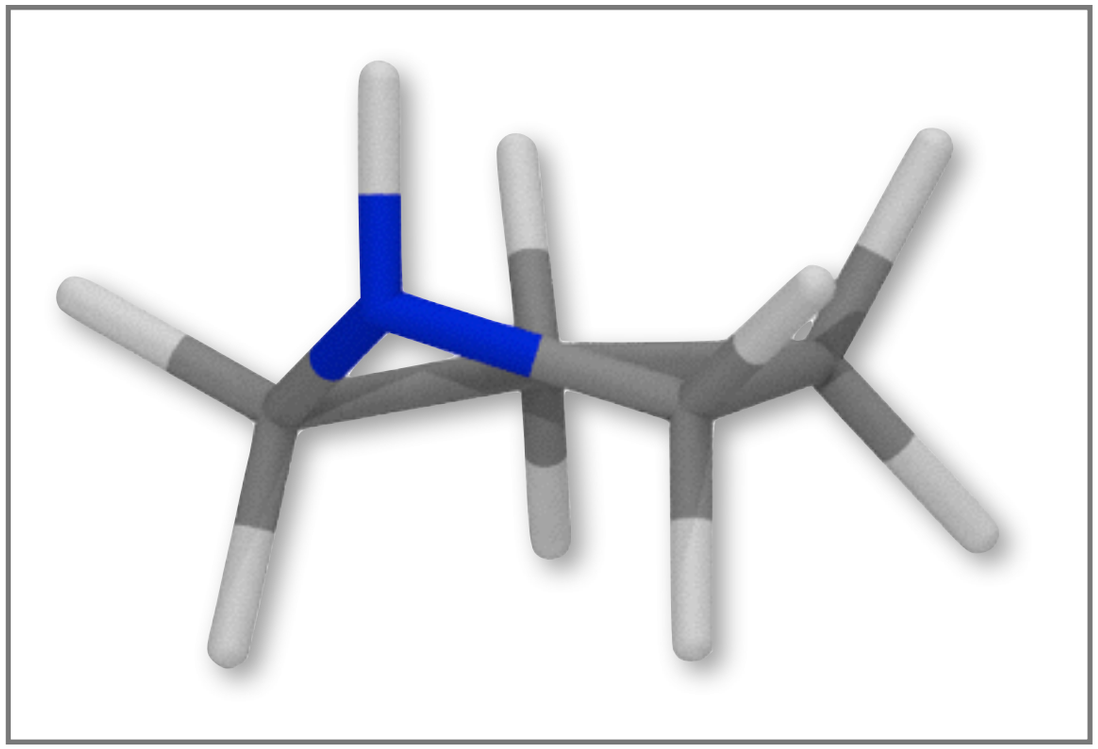
- Auto-detect Water Interaction Sites
- Auto-detect Charge Groups
- Auto-detect Non-redundant Torsions
- Build & Update Parameter Files
- Browse Existing Parameter Sets
- Write Updated Charges to PSF
- Reset Opt. Input from Output
- Visualize Target Data in VMD
- Create Graphic Objects in VMD
- Label Atoms in VMD
- Read Input Parameters from File
- Read/Write Data From Opt. Logs
- Export Plot Data to File
- Monitor Optimization Progress

*ff*TK Exemplified by Charge Optimization



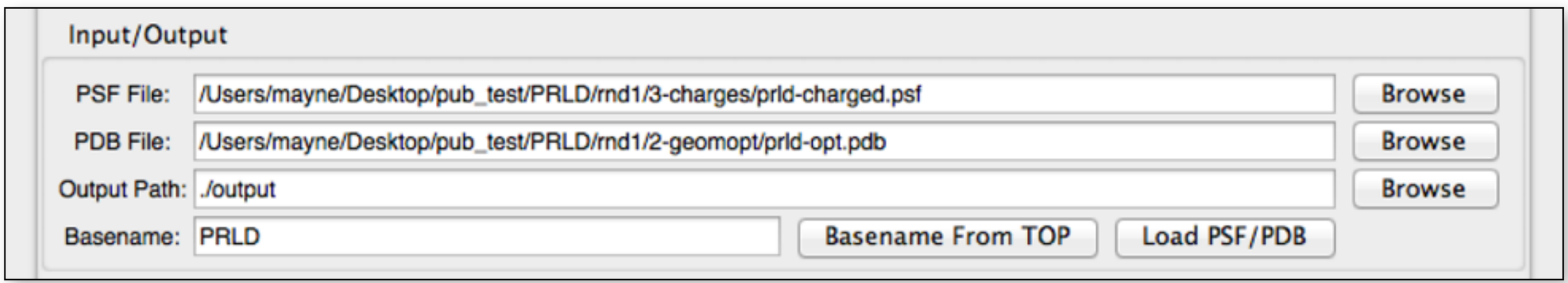
Generating Charge Optimization Target Data

Load QM optimized geometry | Auto-detect interaction sites | Genera



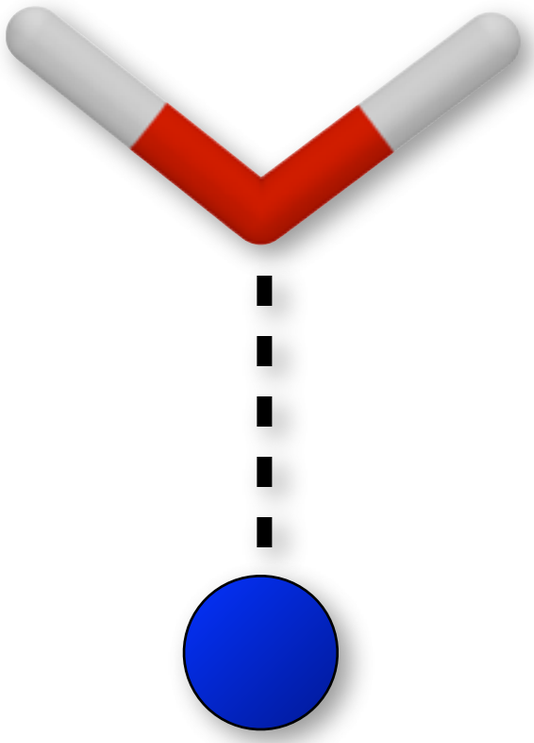
VMD main window

*ff*TK GUI

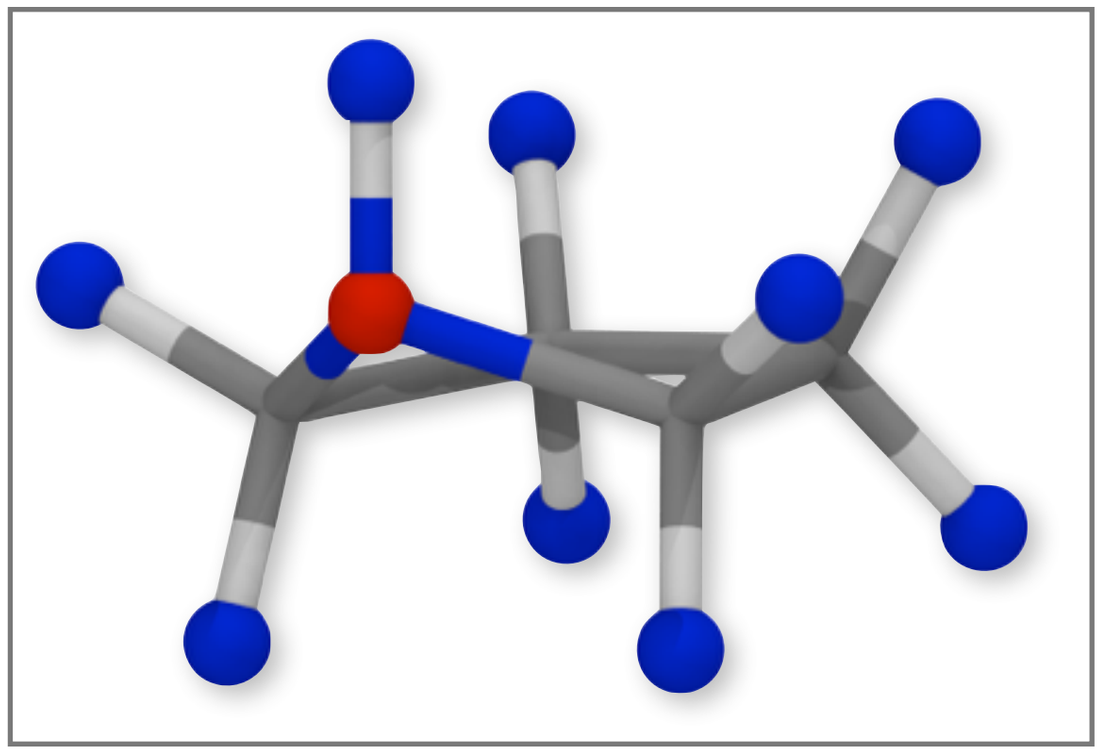


Generating Charge Optimization Target Data

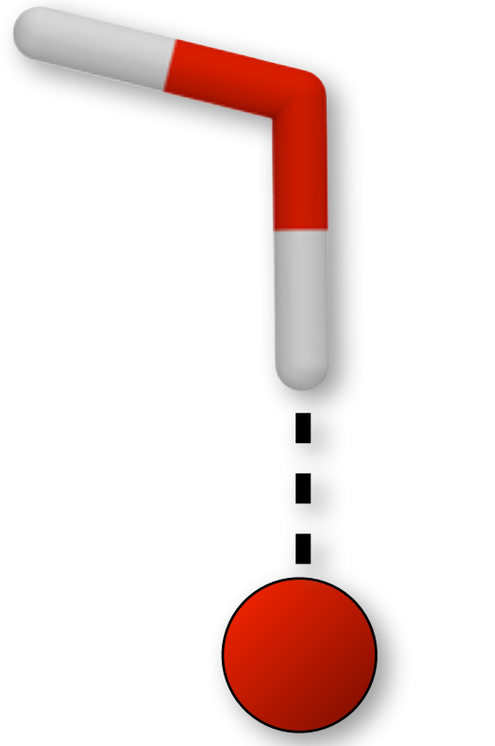
geometry | Auto-detect interaction sites | Generate Gaussian Input Files | Run



Donor

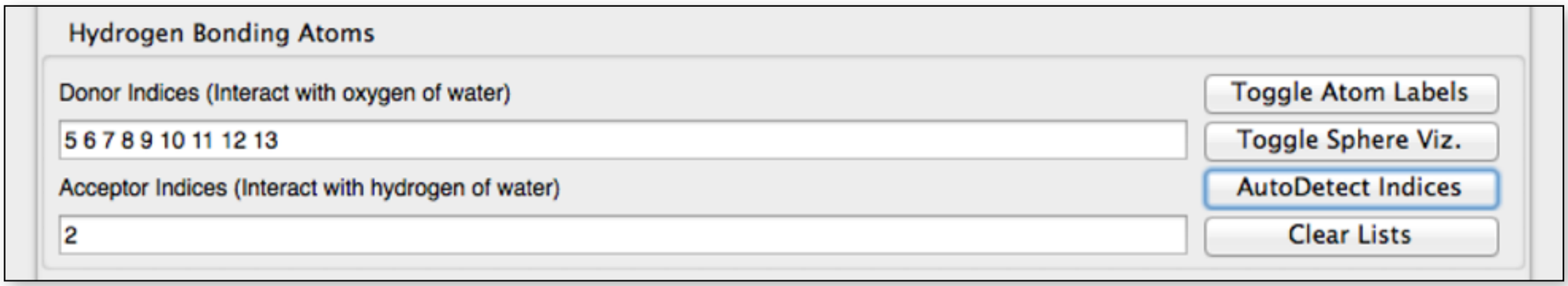


VMD main window



Acceptor

*ff*TK GUI

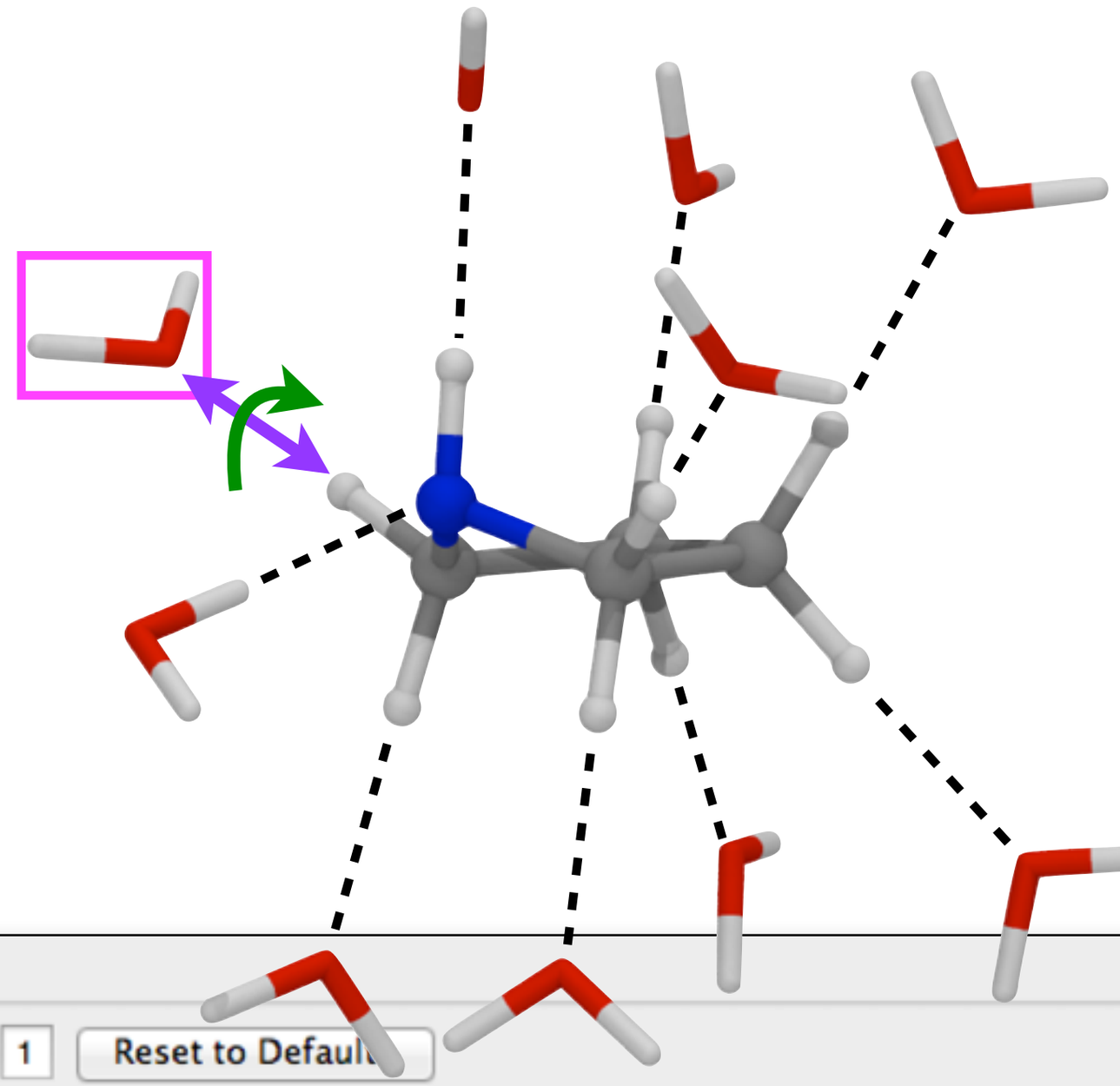


Generating Charge Optimization Target Data

on sites | Generate Gaussian Input Files | Run QM | Inspect water optimization

Compute water position

Optimize
distance & rotation



ffTK GUI

Gaussian Settings

Processors: 1 Memory (GB): 1 Charge: 0 Multiplicity: 1 Reset to Default

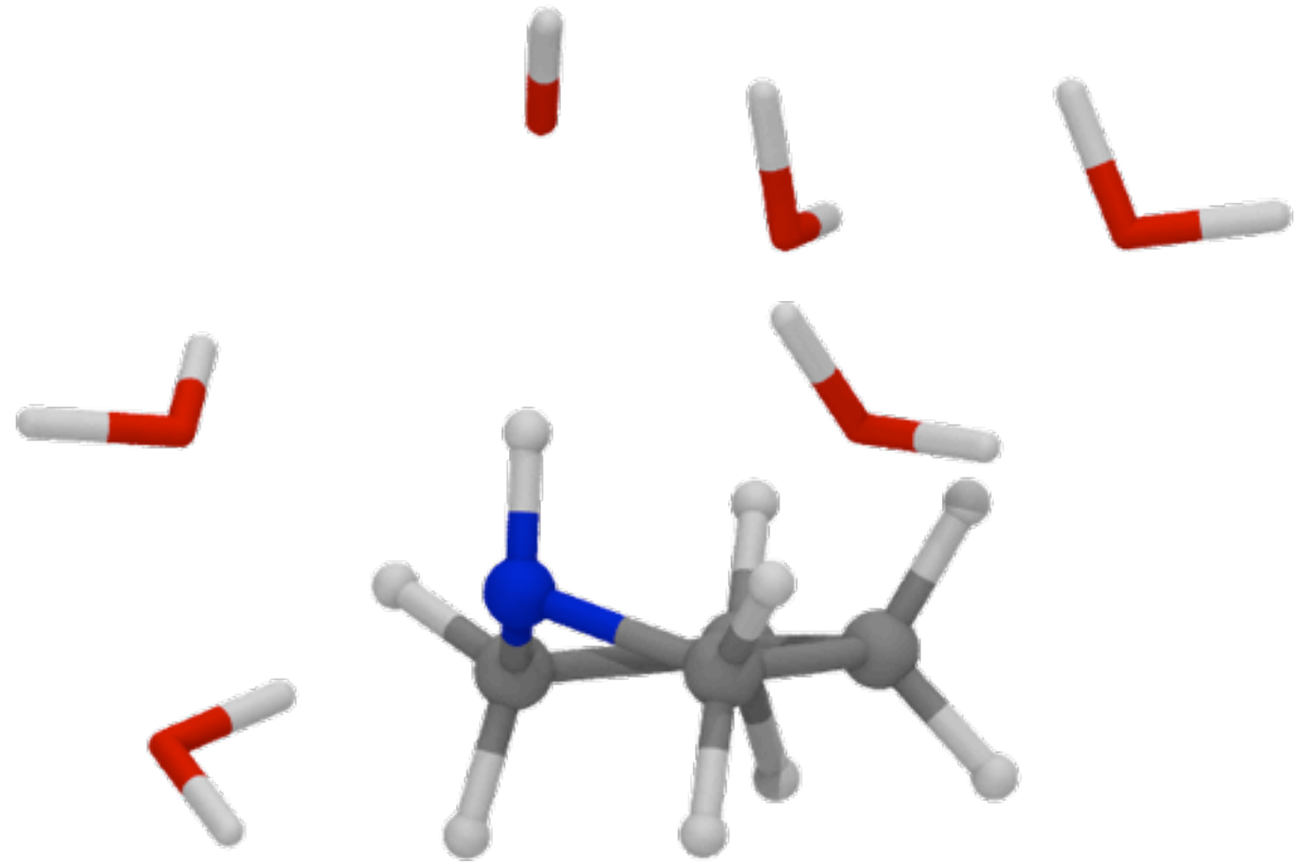
Route: # HF/6-31G* Opt=(Z-matrix,MaxCycles=100) Geom=PrintInputOrient

Write Gaussian Input Files | Load GAU Files | Load LOG Files

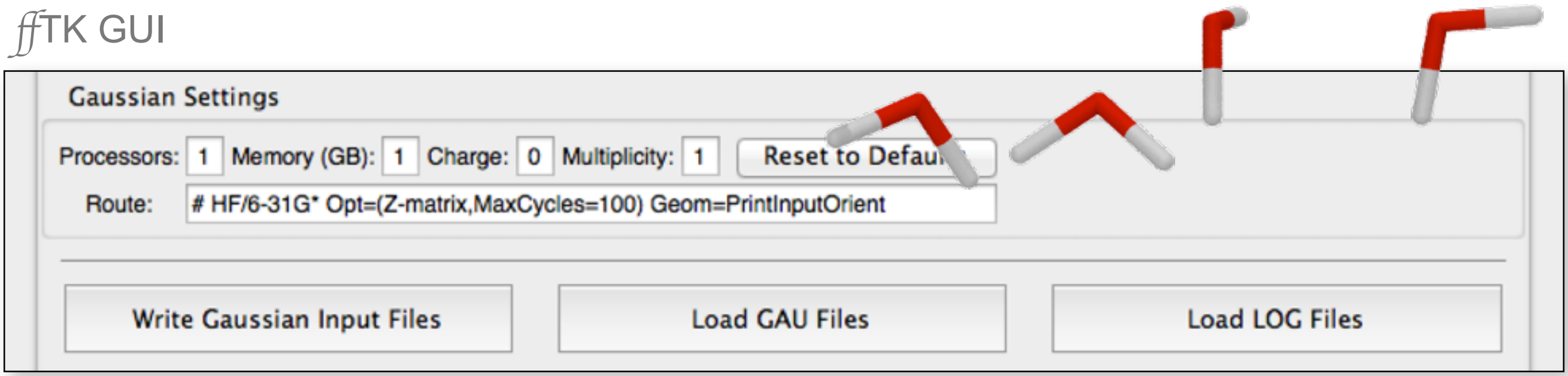
Generating Charge Optimization Target Data

| Run QM | Inspect water optimization

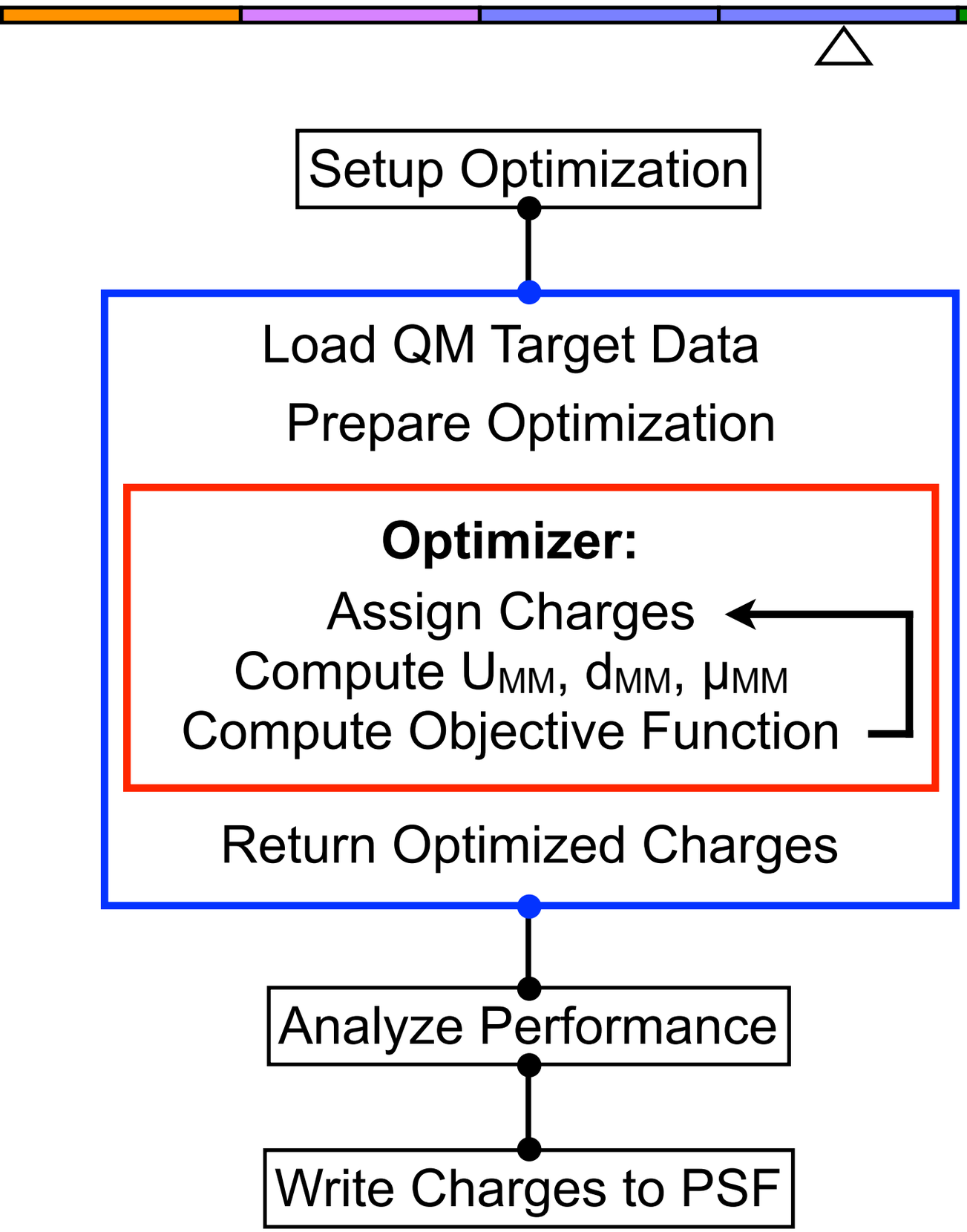
Visually assess QM-optimized water position(s)



*ff*TK GUI



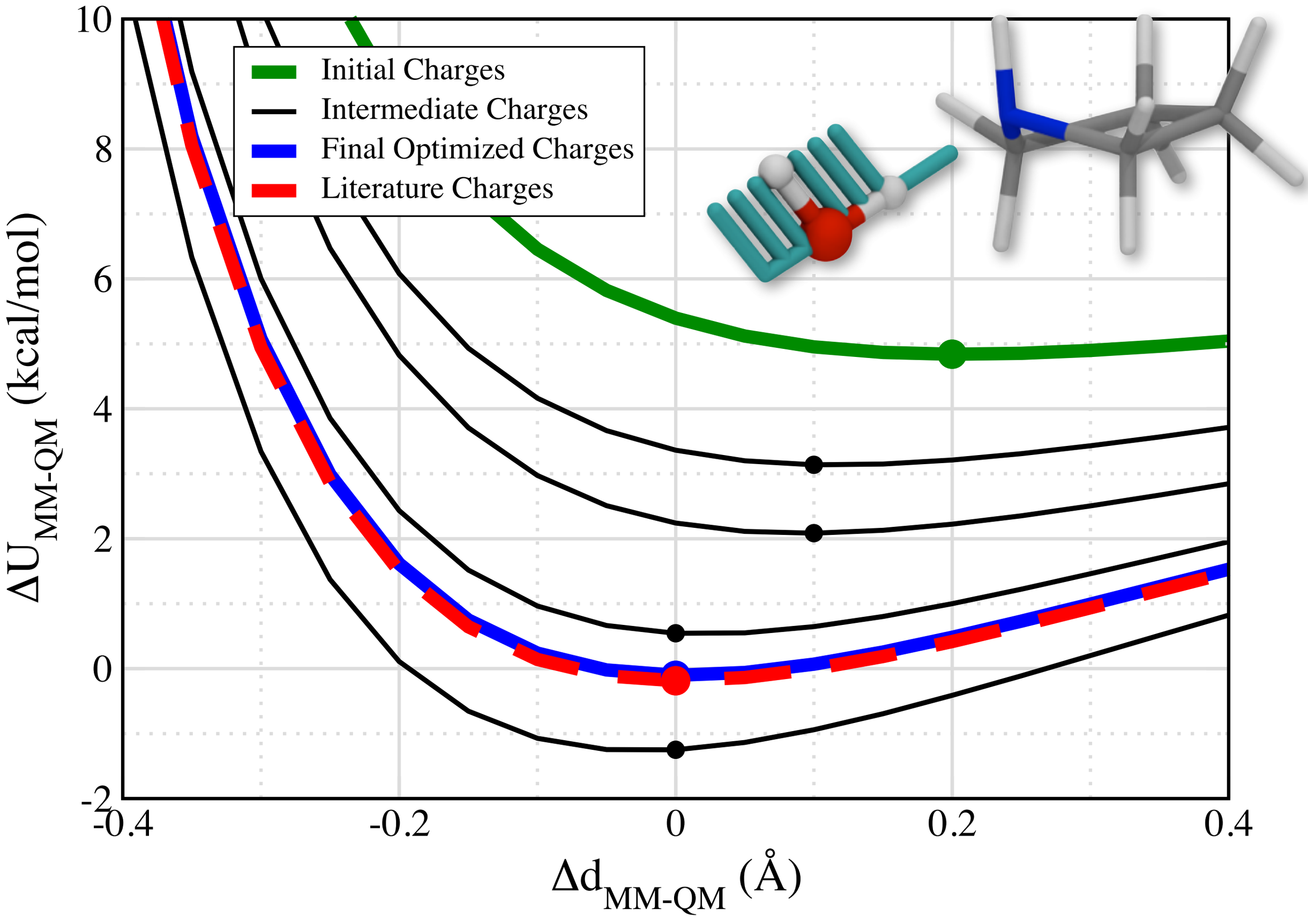
Charge Optimization



Objective Function

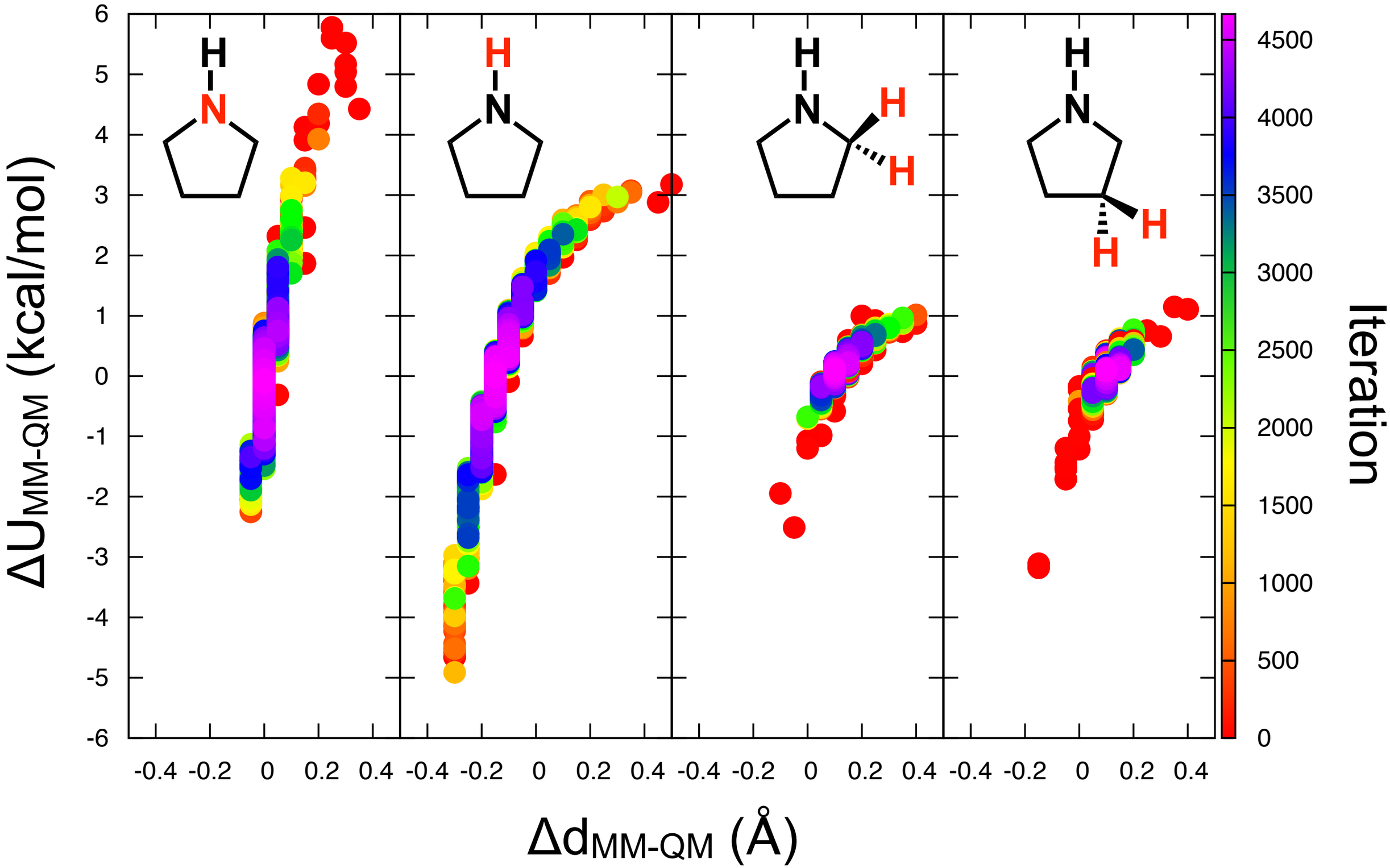
$$\sum_{\text{wat. int.}} f(U_{MM}\text{-}U_{QM}) + \sum_{\text{wat. int.}} f(d_{MM}\text{-}d_{QM}) + f(\mu_{MM}\text{-}\mu_{QM})$$

Assessing MM Water-Interaction Profiles

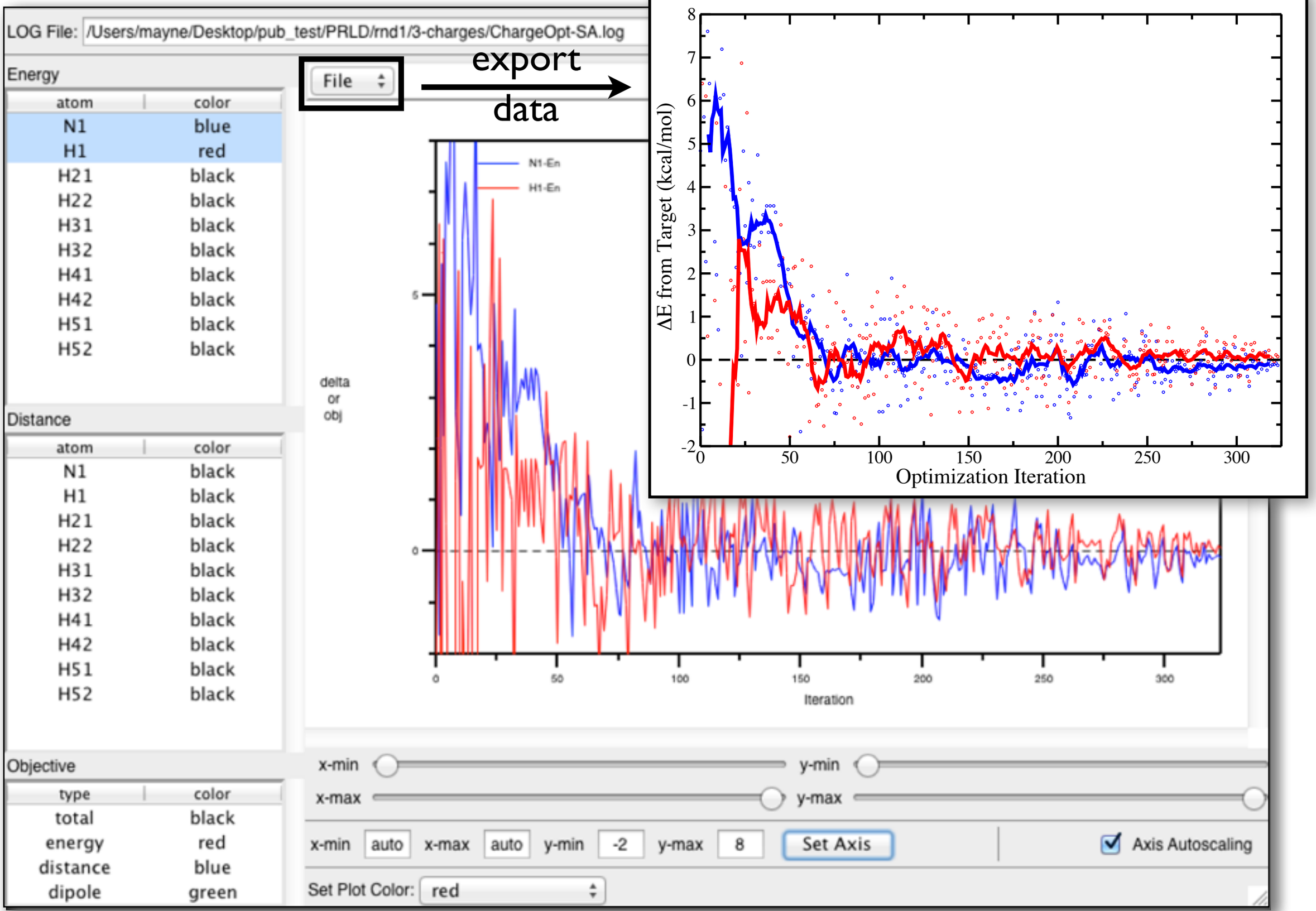


Sampling MM Water-Interaction Profiles

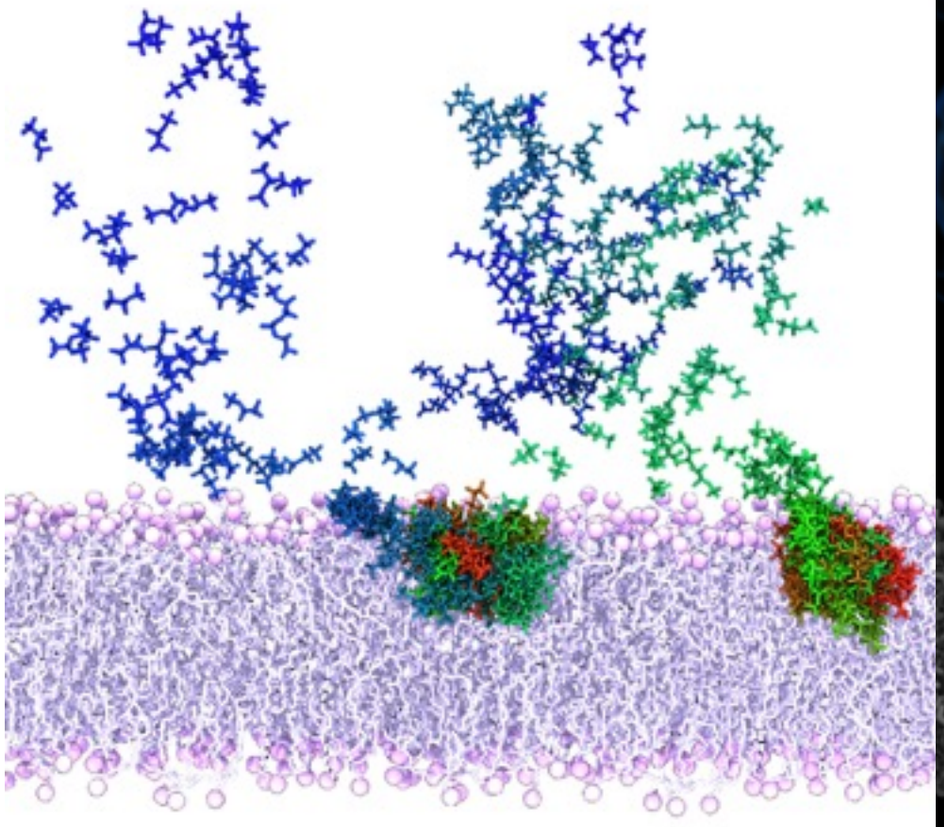
Mode: Simulated Annealing

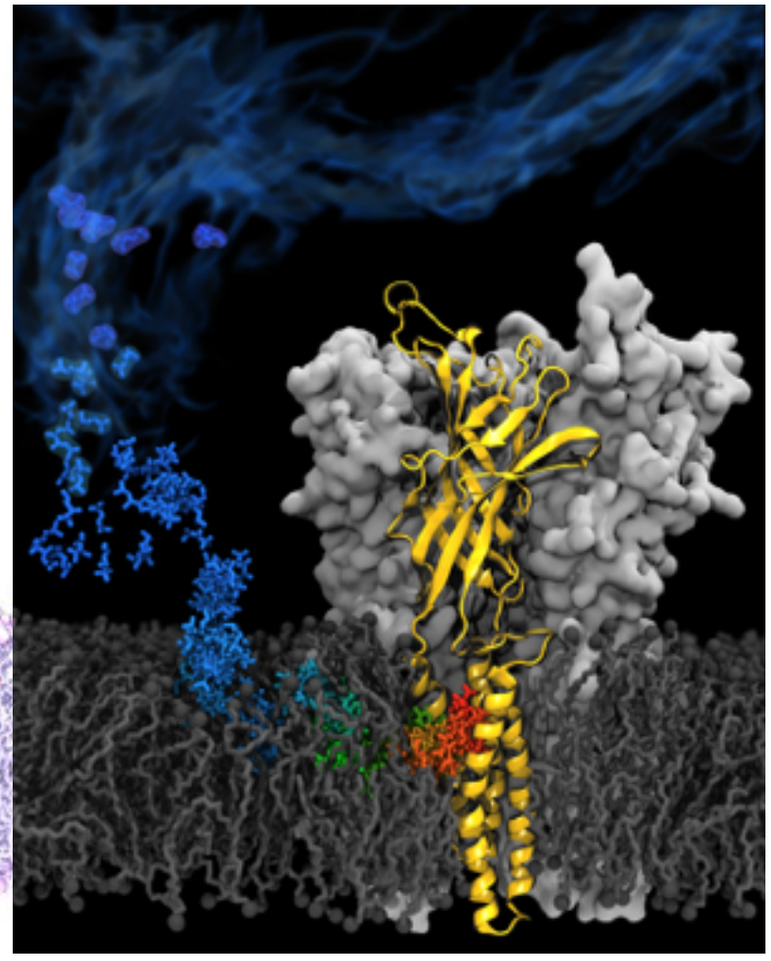


Plotting Charge Optimization Data



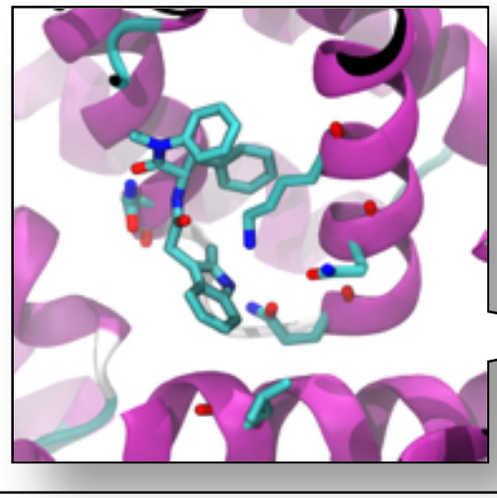
*ff*TK Enables Exciting Science

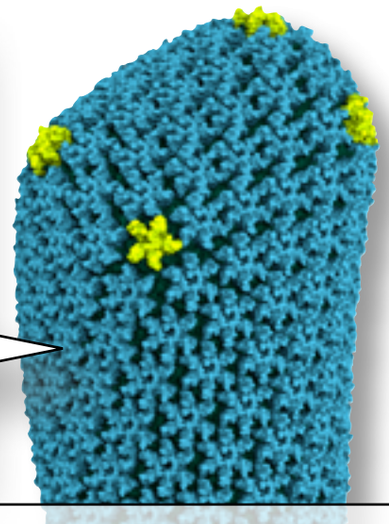




binding mechanisms of inhaled anesthetics

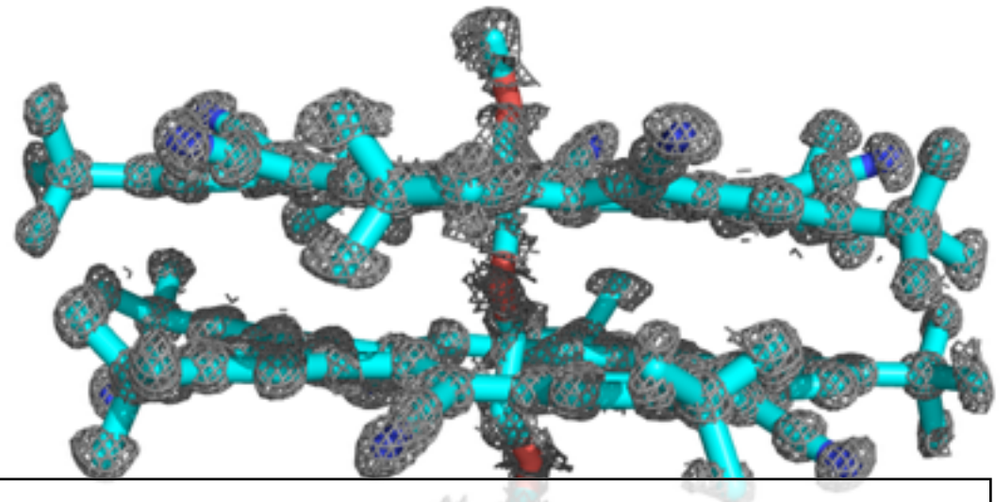
with
Mark Arcario
(Tajkhorshid)





mechanism of action for anti-retroviral drugs targeting the HIV capsid

with Juan Perilla (Schulten)



improved molecular models from low resolution x-ray data

with Abhi Singharoy (Schulten)

NEW

Starting from Somewhere: CGenFF Output

Force Field Toolkit (ffTK) GUI

BuildPar | Opt. Geometry | Water Int. | Opt. Charges | Calc. Bonded | Opt. Bonded | Scan Torsions | Opt. Torsions

▶ Identify Missing Parameters

▶ Assign Missing VDW/LJ Parameters by Analogy

▼ Prepare Parameterization from CGenFF Program Output

For information on the CGenFF Program see: http://cgenff.paramchem.org

Input/Output

Input PDB/MOL2: /Users/mayne/Research/projects_completed/cyanostar-Abhi/abhi.old/pc/abhi_00.mol2 Browse

CGenFF STR File: /Users/mayne/Research/projects_completed/cyanostar-Abhi/abhi.old/pc/abhi_00.str Browse

Output Folder: Browse

Resname: LIG Chain: L Segment: L Get From Input

Analyze Input | Write PSF/PDB | Write PAR | Clear

CGenFF Parameter Data (existing parameters found, only showing missing parameters)

BONDS

Type Def.	k	b_0	Penalty
CG1N1 CG2DC1	345.00	1.4350	140
CG2R61 CG301	230.00	1.4900	8

ANGLES

Type Def.	k	θ	kub	s	Penalty
CG2DC1 CG1N1 NG1T1	40.00	180.00			21
CG1N1 CG2DC1 CG2DC1	48.00	123.00			37.5
CG1N1 CG2DC1 CG2R61	48.00	113.00			70.4
CG2DC1 CG2DC1 CG2R61	29.00	122.00			3.5

DIHEDRALS

Type Def.	k	n	d	Penalty
CG2R61 CG2R61 CG2R61 CG301	3.1000	2	180.00	1.2
CG301 CG2R61 CG2R61 HGR61	2.4000	2	180.00	1.2
CG2R61 CG2R61 CG301 CG331	0.2300	2	180.00	8
CG2R61 CG301 CG331 HGA3	0.0400	3	0.00	8

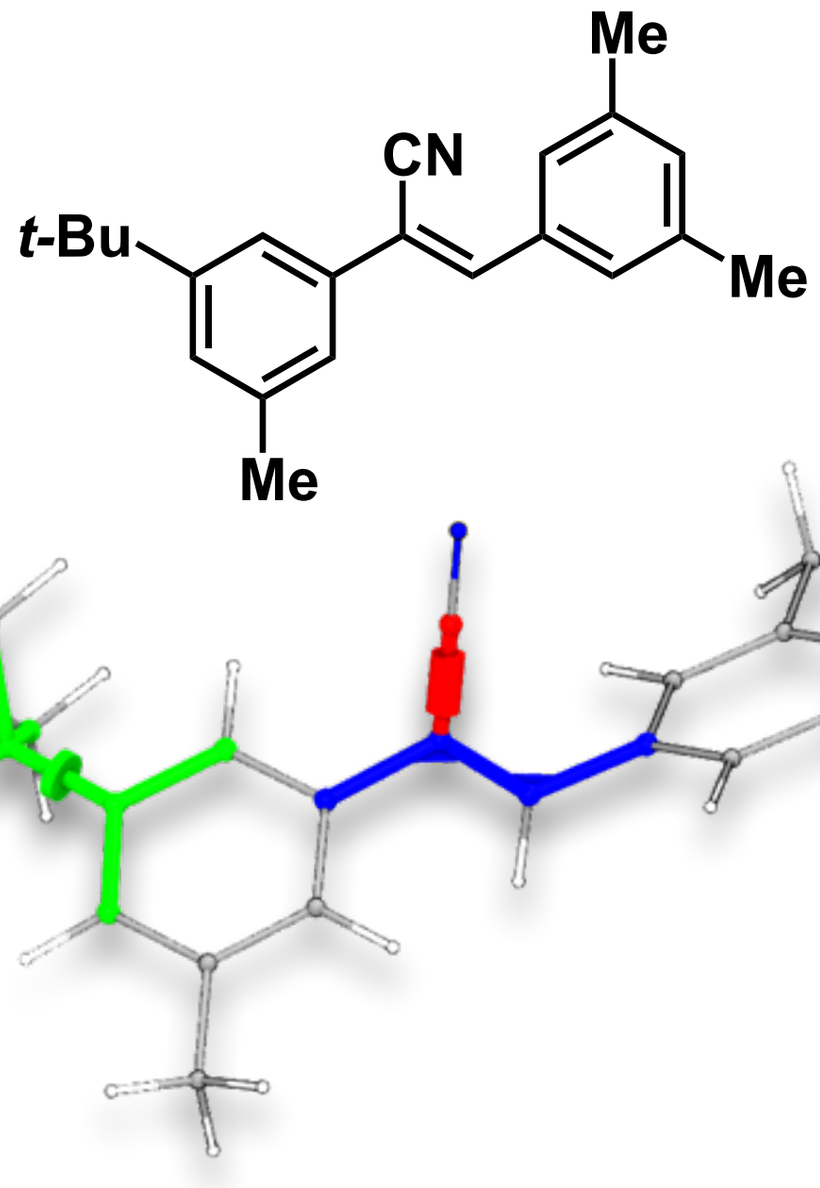
IMPROPERS

▶ Update Parameter File with Optimized Parameters

GUI Event Log (off)

I) preparation of PSF, PDB, and initial PAR files

II) use the CGenFF Program for atom typing and “first guess” at missing parameters



NEW

Analyzing QM Dihedral Target Data

Torsion Explorer

PSF File: /Users/mayne/Desktop/work_IN_PROGRESS/ffTK/fftk_demo/prld/3-charges/prld-charged.psf Browse

PDB File: /Users/mayne/Desktop/work_IN_PROGRESS/ffTK/fftk_demo/prld/2-geomopt/prld-opt.pdb Browse

Log Files:

PRLD.scan.inversion.log
PRLD.scan1.neg.log
PRLD.scan1.pos.log
PRLD.scan2.neg.log
PRLD.scan2.pos.log
PRLD.scan3.neg.log
PRLD.scan3.pos.log

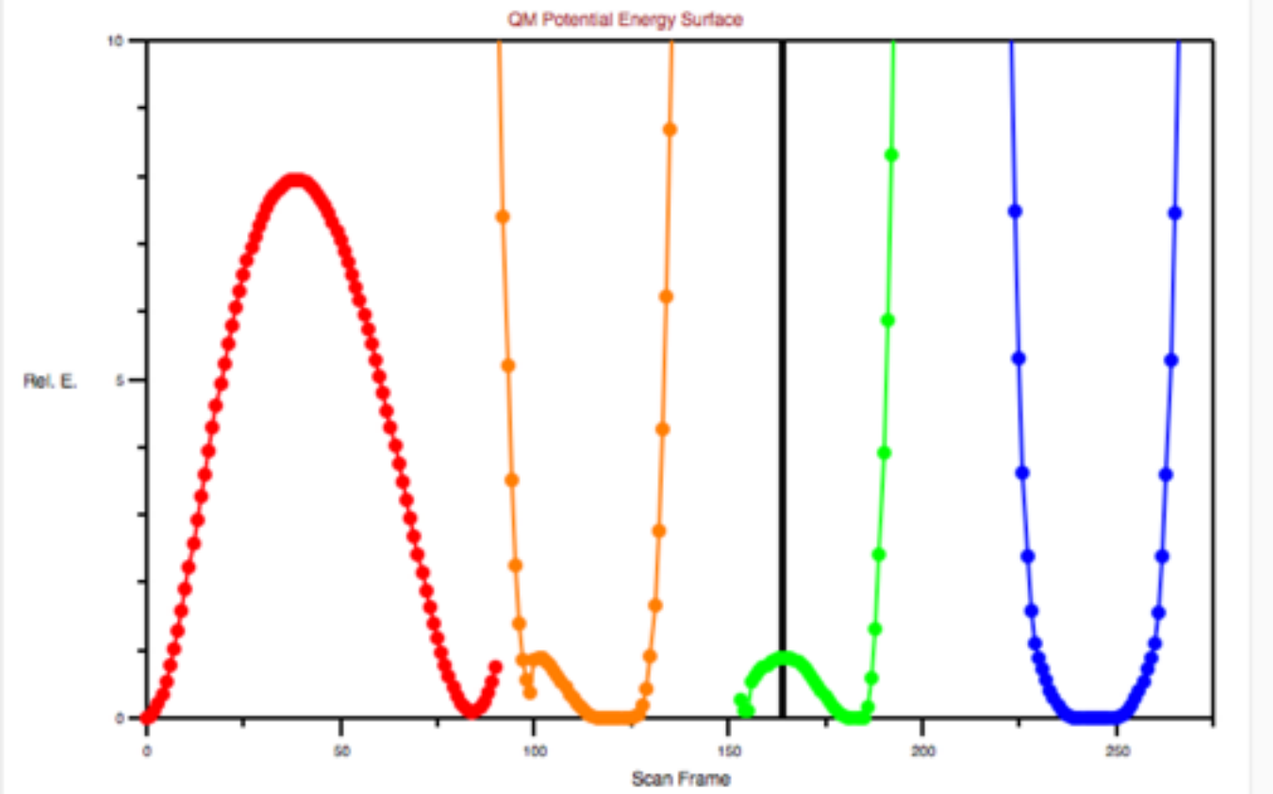
Add
Delete
Clear
Load

Dihedral Analysis

Analyze Trajectory Update Dihedral Data On Off

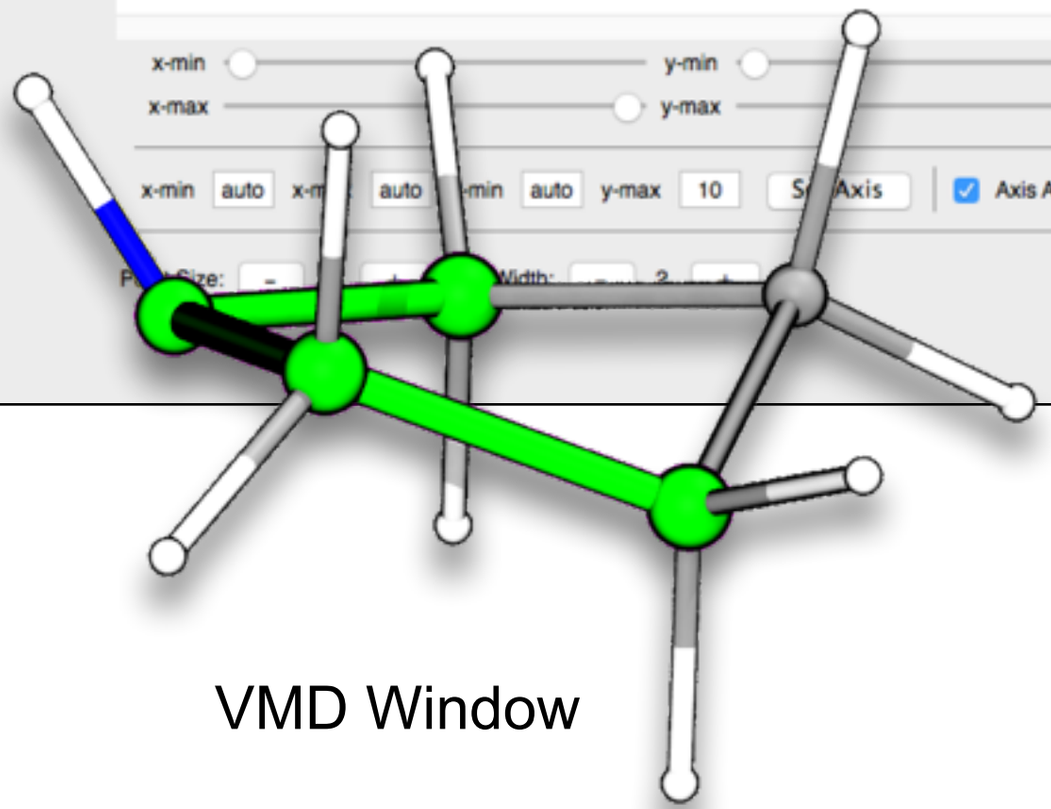
Indices	Type	θ	dθ
5 2 3 11	HGP1 NG3C51 CG3C52 HGA2	9.47	4.18
5 2 3 10	HGP1 NG3C51 CG3C52 HGA2	127.74	4.05
7 1 2 5	HGA2 CG3C52 NG3C51 HGP1	21.28	3.95
5 2 3 4	HGP1 NG3C51 CG3C52 CG3C52	113.67	3.90
0 1 2 5	CG3C52 CG3C52 NG3C51 HGP1	139.86	3.75
6 1 2 5	HGA2 CG3C52 NG3C51 HGP1	98.93	3.68
1 2 3 11	CG3C52 NG3C51 CG3C52 HGA2	130.39	3.44
1 2 3 10	CG3C52 NG3C51 CG3C52 HGA2	111.35	3.31
7 1 2 3	HGA2 CG3C52 NG3C51 CG3C52	99.72	3.21
1 2 3 4	CG3C52 NG3C51 CG3C52 CG3C52	7.24	3.16
0 1 2 3	CG3C52 CG3C52 NG3C51 CG3C52	18.86	3.00
6 1 2 3	HGA2 CG3C52 NG3C51 CG3C52	140.07	2.93
10 3 4 12	HGA2 CG3C52 CG3C52 HGA2	155.63	2.07
10 3 4 0	HGA2 CG3C52 CG3C52 CG3C52	87.91	2.06
2 3 4 12	NG3C51 CG3C52 CG3C52 HGA2	86.25	2.02
2 3 4 0	NG3C51 CG3C52 CG3C52 CG3C52	30.21	2.01
11 3 4 12	HGA2 CG3C52 CG3C52 HGA2	35.86	1.94
11 3 4 0	HGA2 CG3C52 CG3C52 CG3C52	152.32	1.93
10 3 4 13	HGA2 CG3C52 CG3C52 HGA2	34.44	1.91
2 3 4 13	NG3C51 CG3C52 CG3C52 HGA2	152.57	1.86
11 3 4 13	HGA2 CG3C52 CG3C52 HGA2	85.32	1.78

File



x-min y-min
x-max y-max
x-min auto x-max auto y-min auto y-max 10 Set Axis
Axis Autoscaling

Torsion Explorer GUI



VMD Window

Conclusions

Find Missing Parameters

Geometry Optimization (QM)

Water Interaction En. (QM)

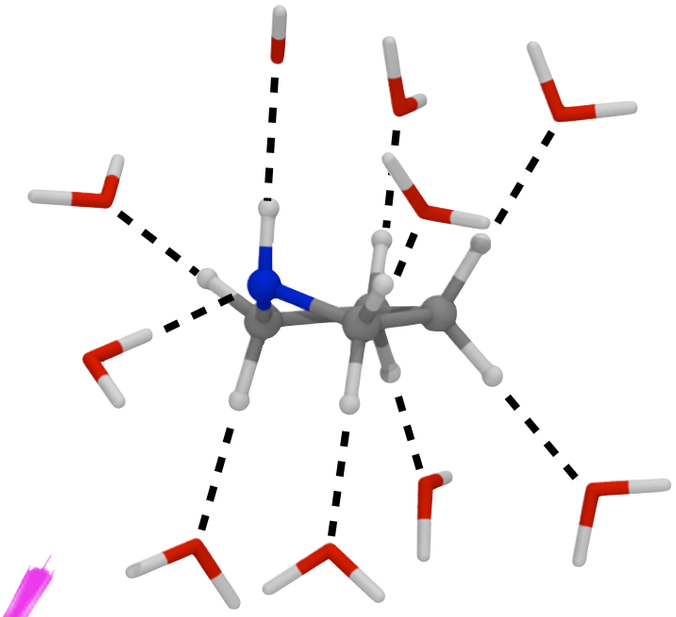
Charge Optimization

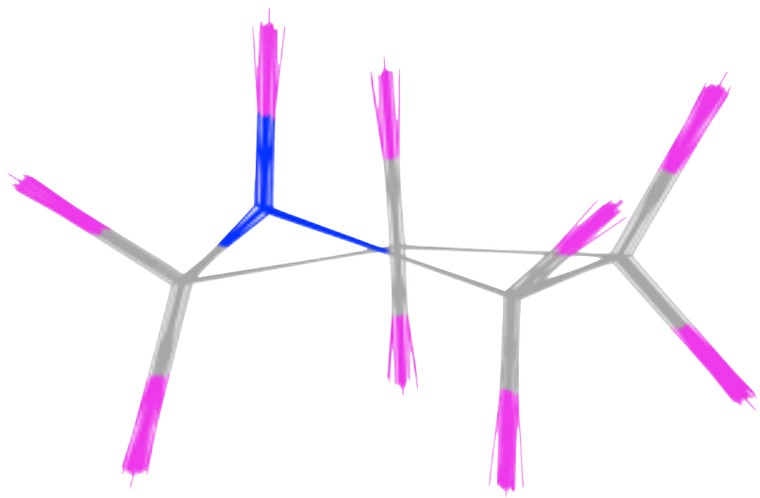
Hessian Calculation (QM)

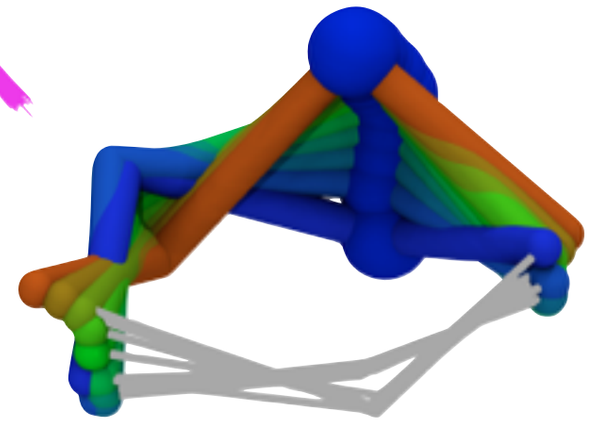
Bond & Angle Optimization

Torsion Scan (QM)

Torsion Optimization







*ff*TK:

- Simplifies the parameterization workflow
- Offers opportunity for extensive customization
- Provides analytical tools to assess parameter performance

www.ks.uiuc.edu/Research/vmd/plugins/fftk

Mayne *et al.; J. Comp. Chem.* **2013,** 34, pp. 2757-2770 (Cover Article)

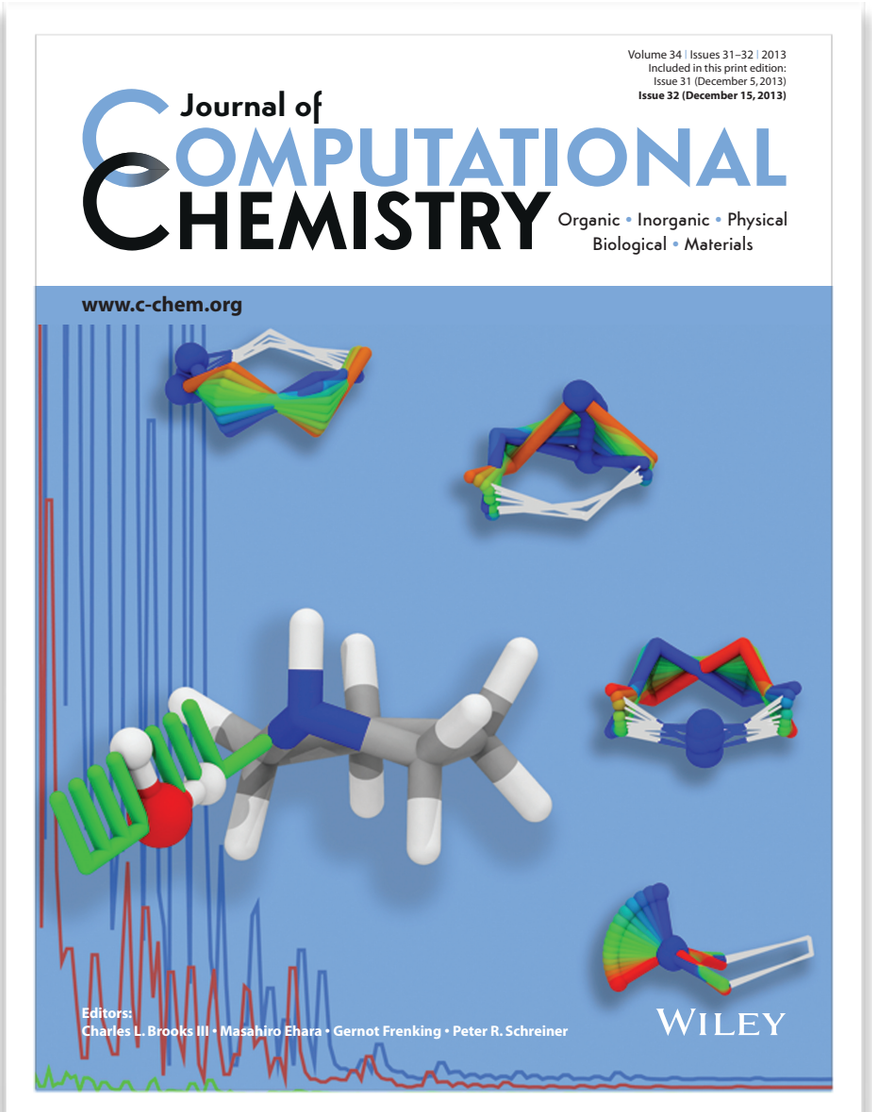
*ff*TK is available as a VMD Plugin (1.9.2 and greater)

Full Documentation and Screencast & Paper Tutorials

http://www.ks.uiuc.edu/Research/vmd/plugins/fftk
http://www.ks.uiuc.edu/Training/Tutorials/#FFTK

May the Force Field Be With You!

http://www.ks.uiuc.edu/Highlights/?section=2013&highlight=2013-09



Questions?