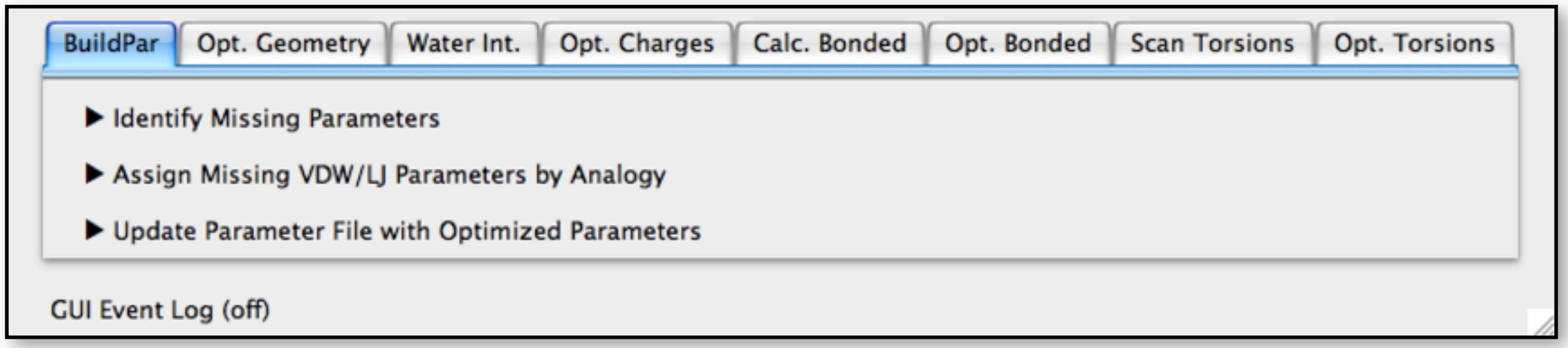
Parameterizing Small Molecules Using:

The Force Field Toolkit (ffTK)



Christopher G. Mayne, Emad Tajkhorshid
Beckman Institute for Advanced Science and Technology
University of Illinois, Urbana-Champaign

Klaus Schulten
University of Illinois, Urbana-Champaign

James C. Gumbart, Anna Pavlova
Georgia Institute of Technology

NIH Hands On Workshop | Atlanta | November 5th, 2014

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

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

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

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

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

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

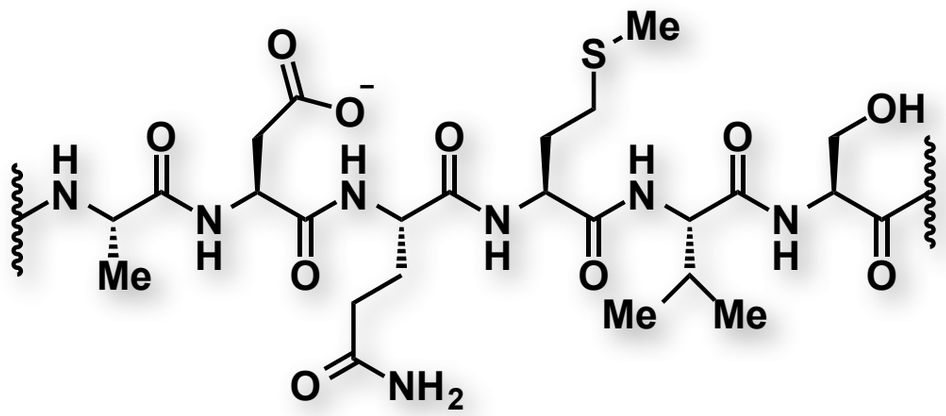
$$\sum_{dihedrals} k_i^{dihedral}[1 + \cos(n_i\phi_i + \delta_i)] +$$

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

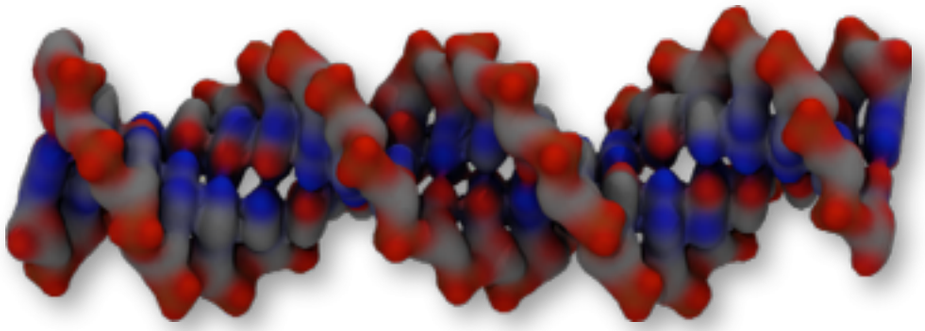
Parameter Transferability In Biopolymers

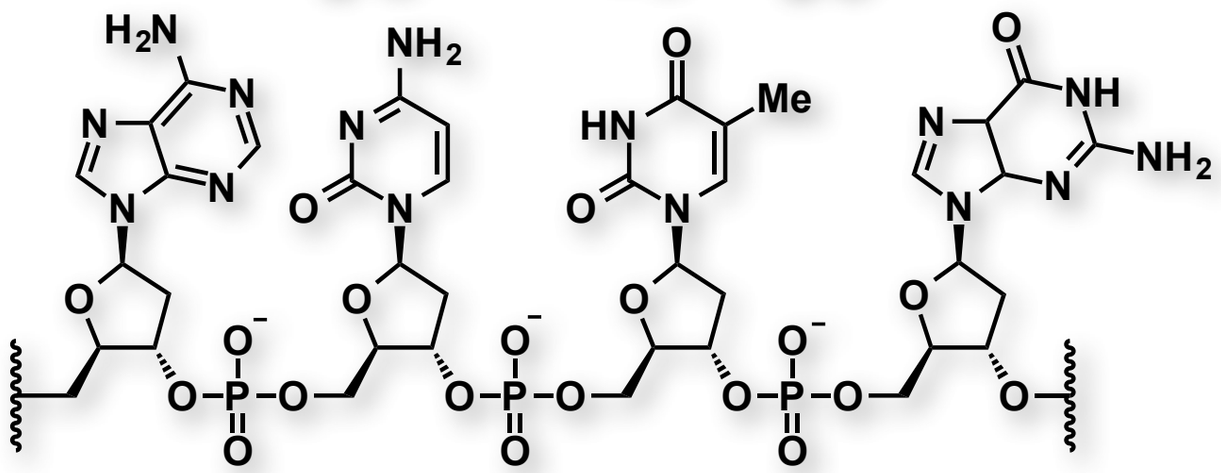
Parameter set describes molecular behavior in varied chemical (connectivity) and spatial (conformation) contexts

Peptides and Proteins

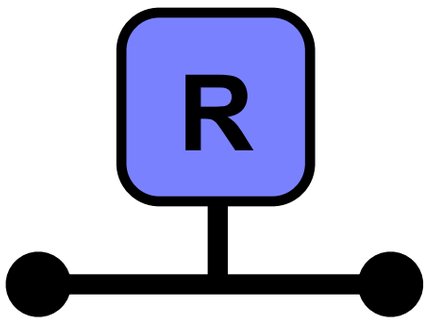


Nucleic Acids





Key Features:

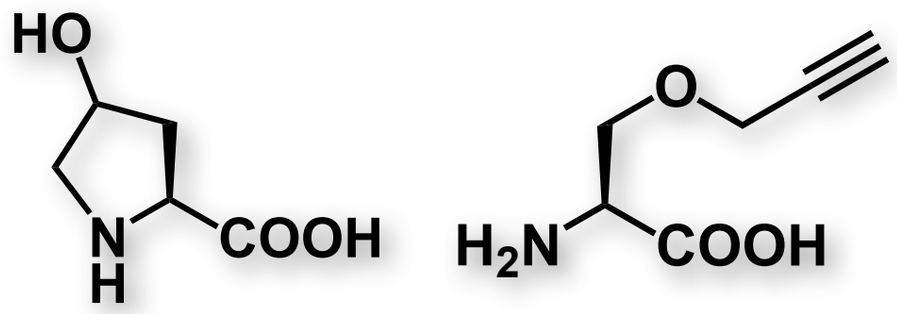


limited set of isolated building blocks

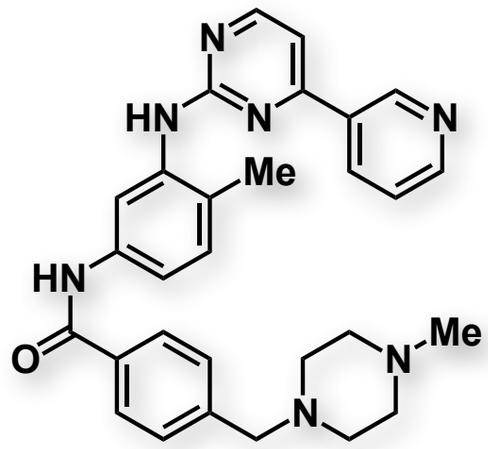
repetitive backbone unit

Parameterization as an Impasse

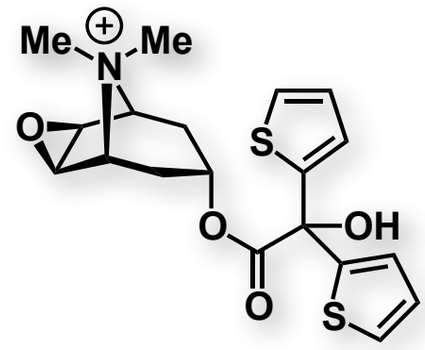
non-standard or engineered amino acids



small molecule ligands

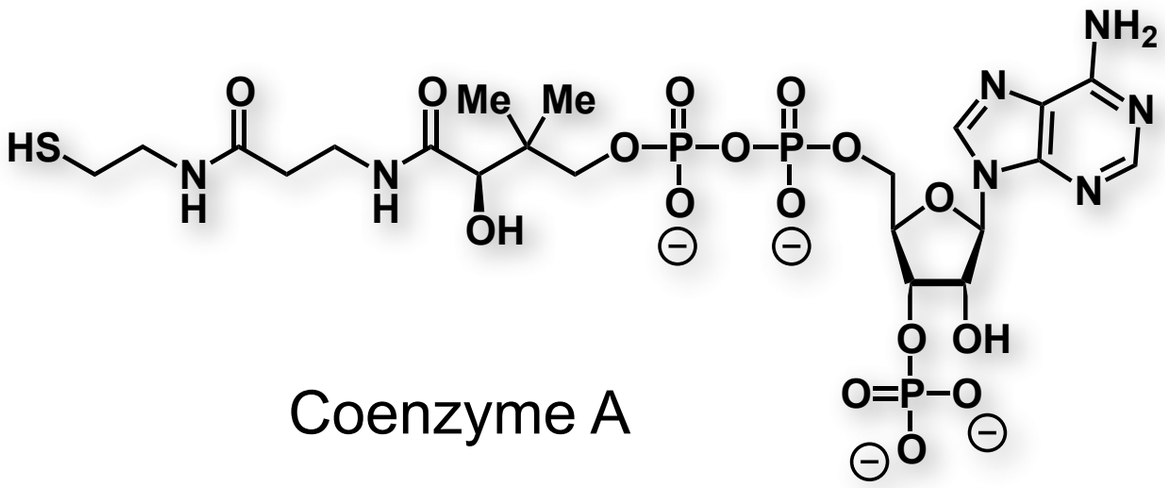


Imatinib (Gleevec)



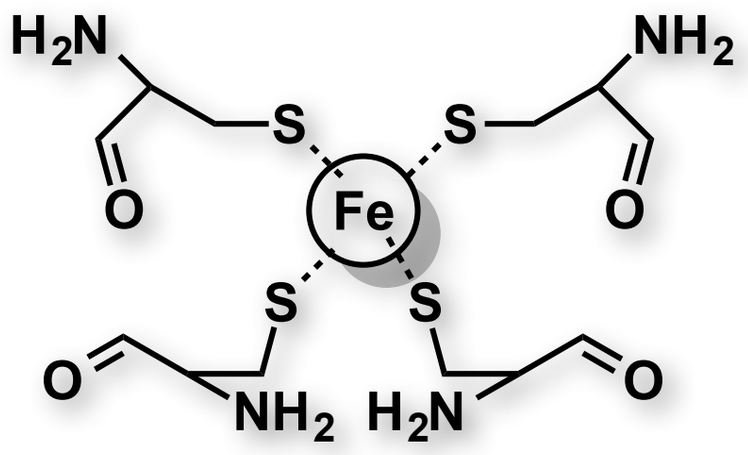
Tiotropium (Spiriva)

cofactors

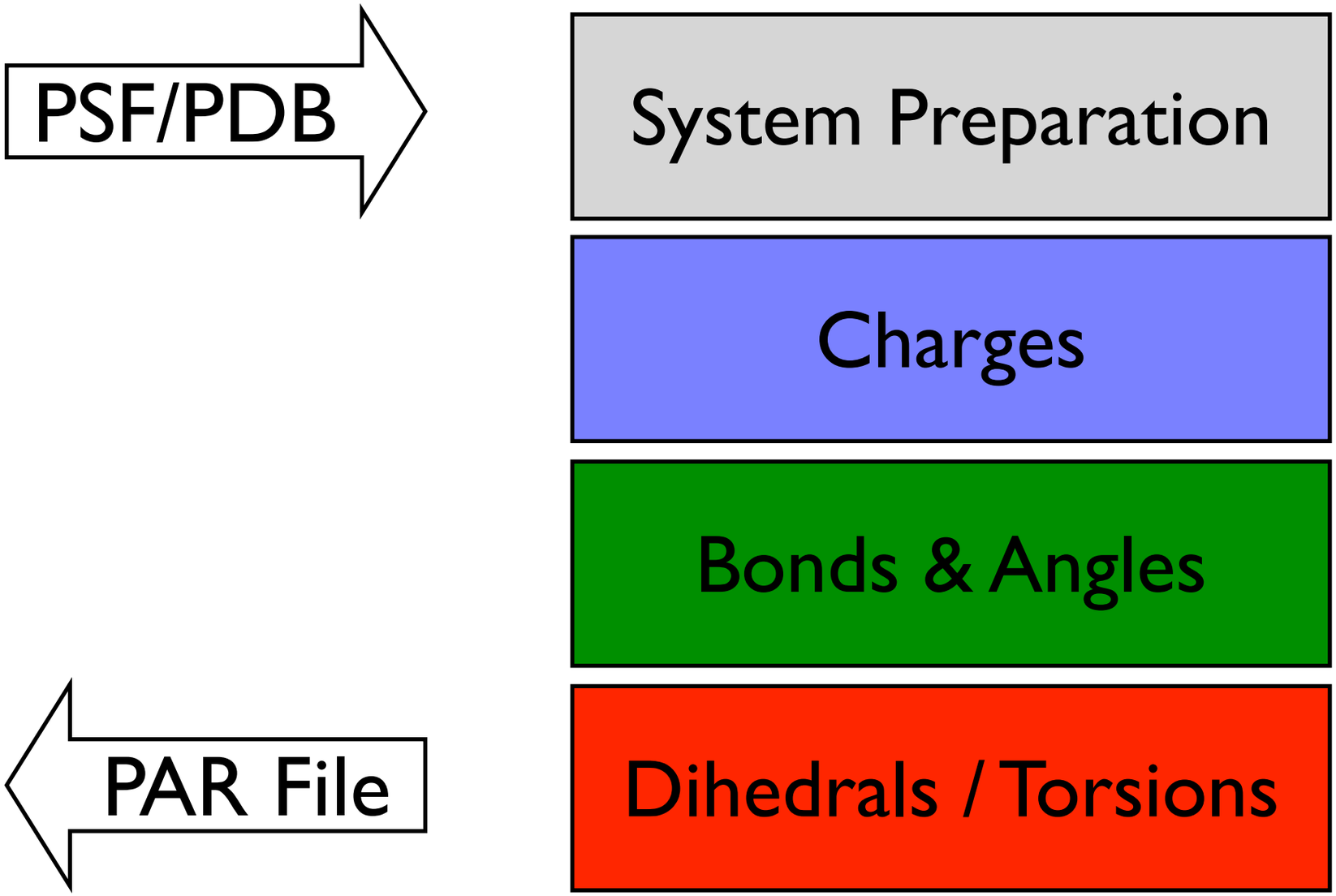


Coenzyme A

metal centers



General Parameterization Workflow



CGenFF Parameterization Workflow

PSF/PDB

Find Missing Parameters

Geometry Optimization (QM)

Water Interaction En. (QM)

Charge Optimization

Hessian Calculation (QM)

Bond & Angle Optimization

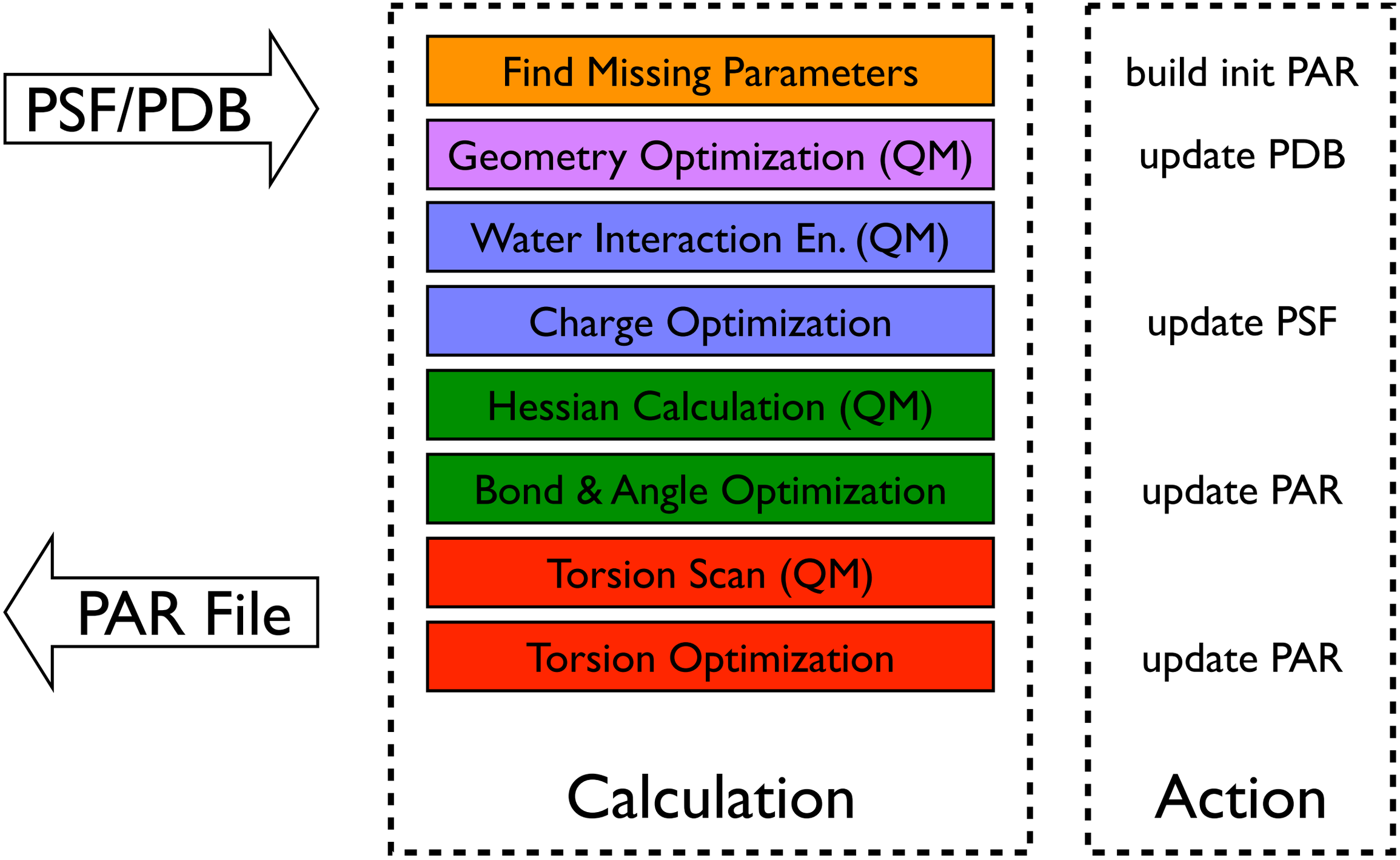
Torsion Scan (QM)

Torsion Optimization

PAR File

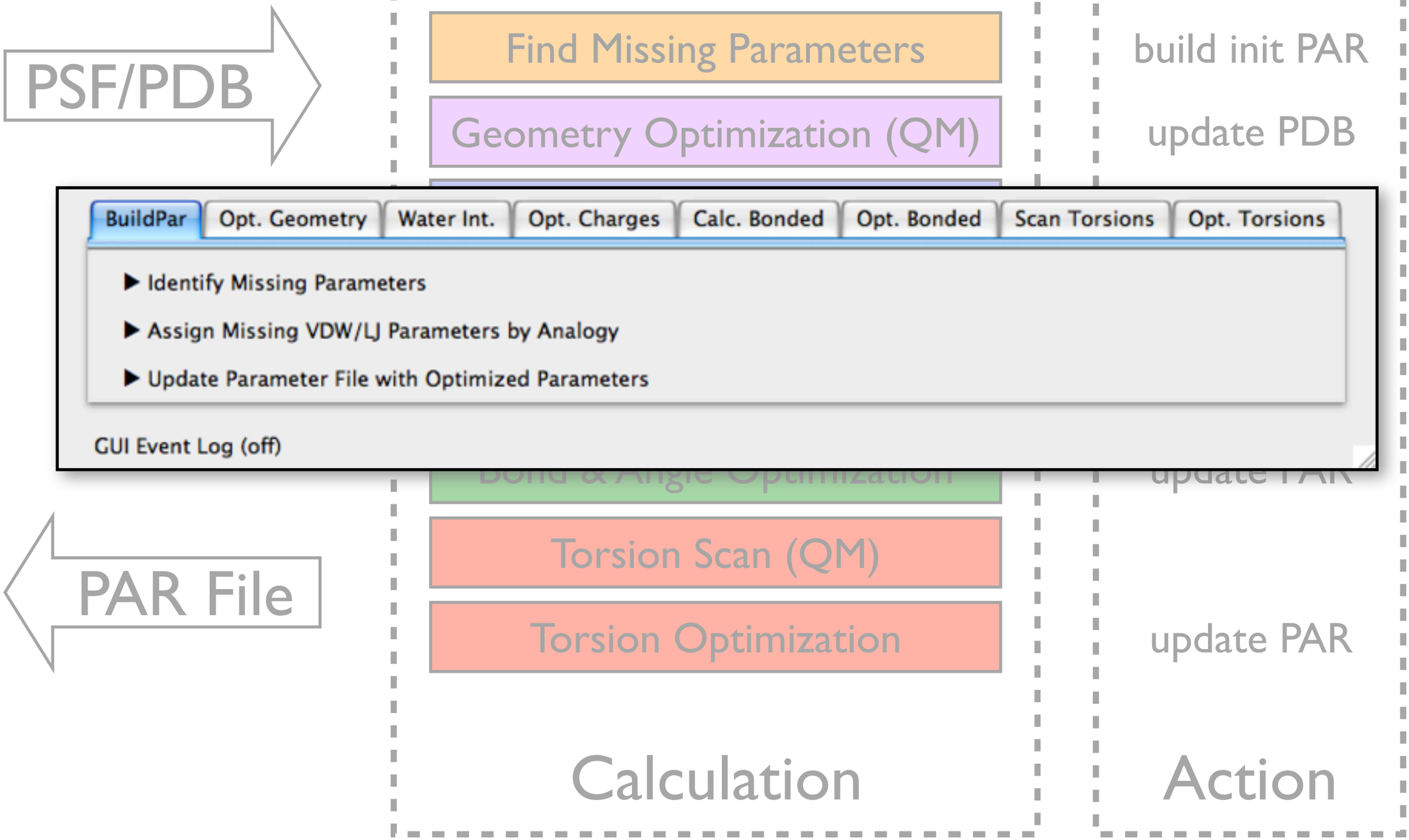
K. Vanommeslaeghe *et al.*, *J. Comput. Chem.* **2010**, 31, 671-690.

CGenFF Parameterization Workflow



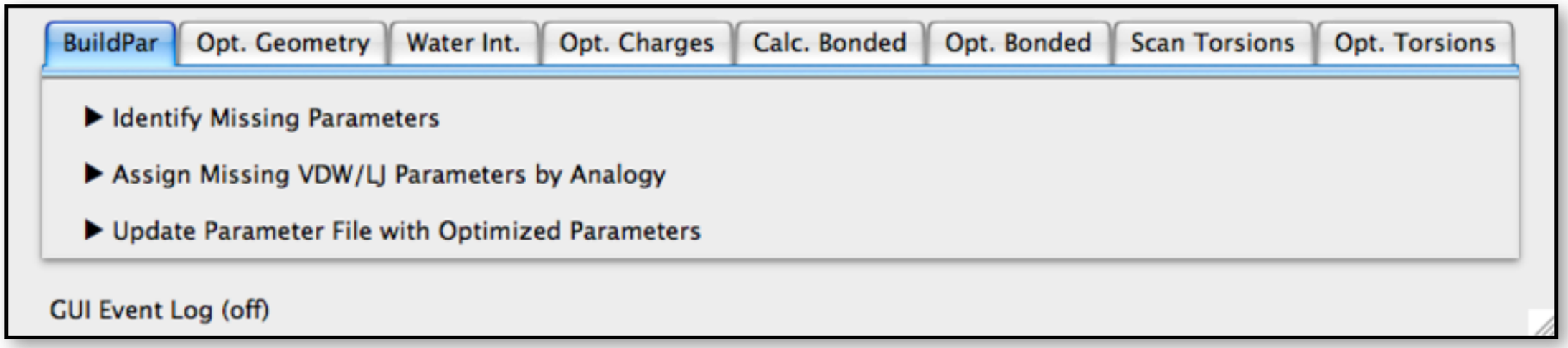
K. Vanommeslaeghe *et al.*, *J. Comput. Chem.* **2010**, 31, 671-690.

CGenFF Parameterization Workflow

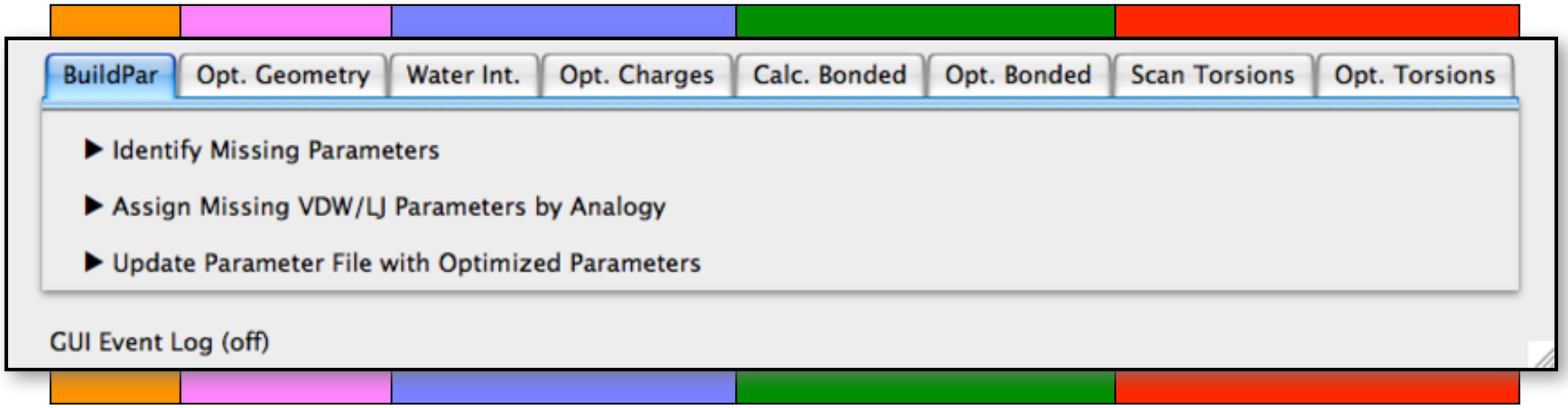


C.G. Mayne *et al.*, *J. Comput. Chem.* **2013**, 34, 2757-2270.

ffTK Interface

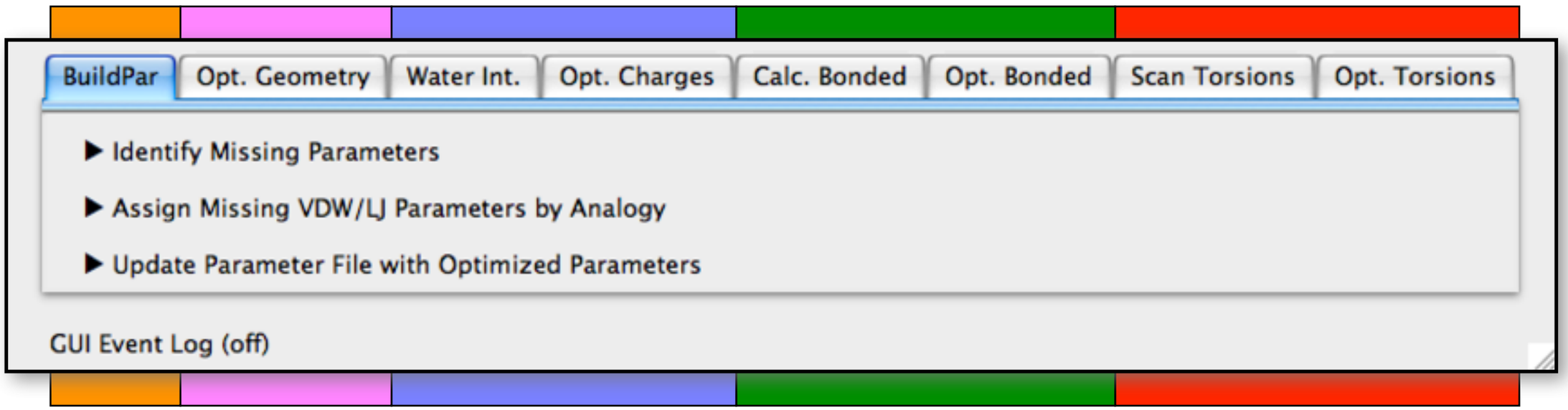


ffTK Interface

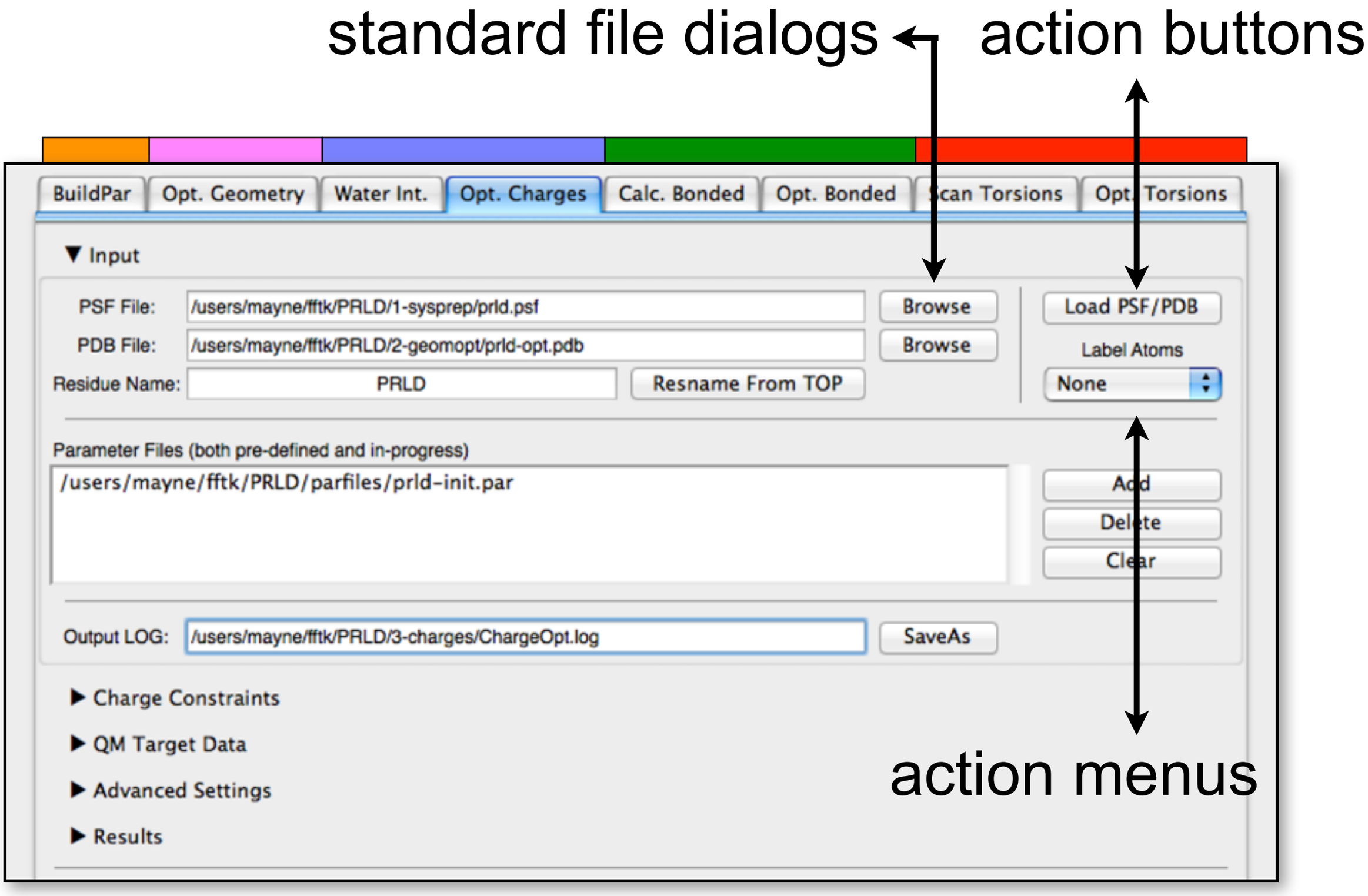


ffTK Interface

tasks organized under tabs



ffTK Interface



Functionality Provided by ffTK

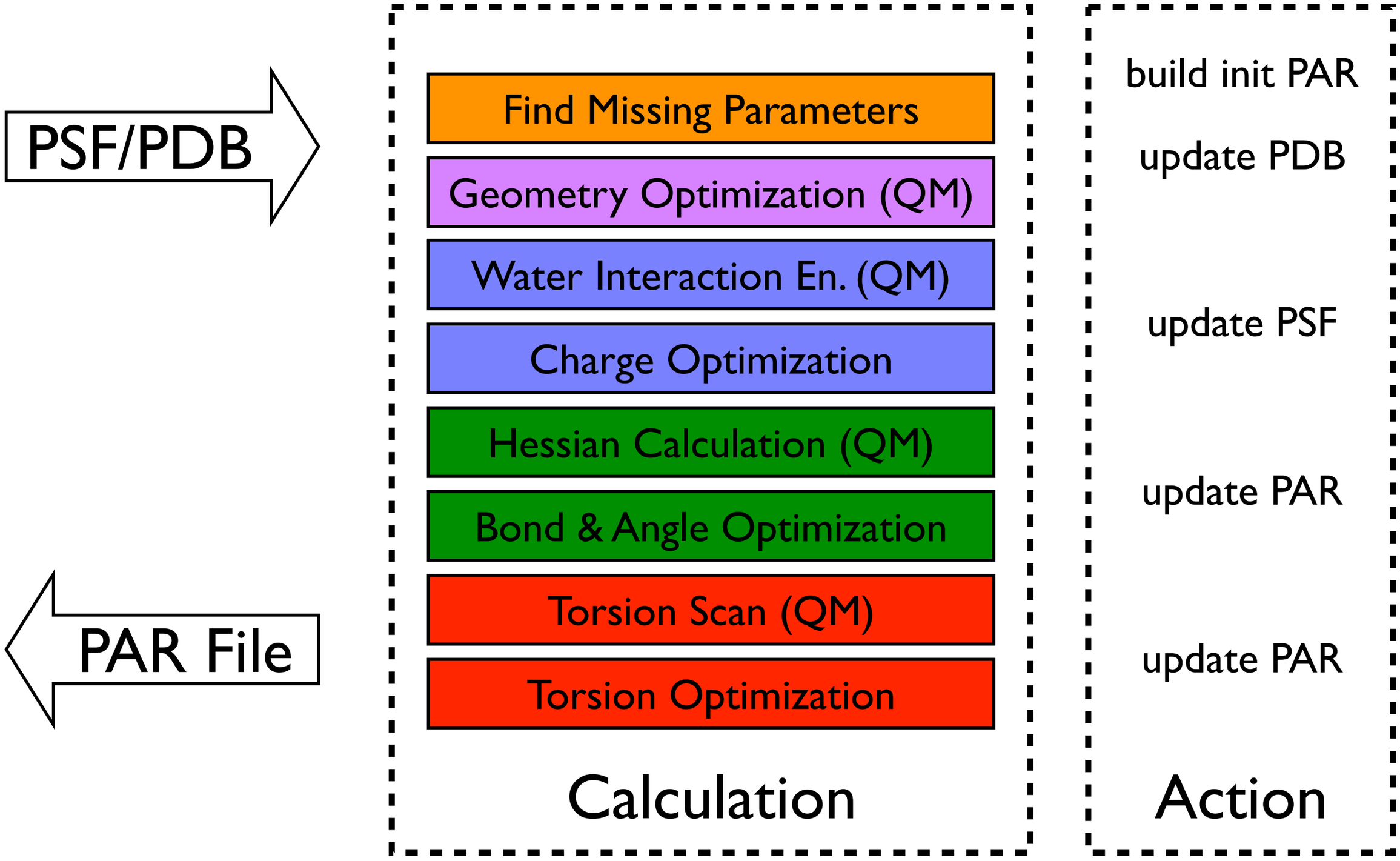
Core Functions

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

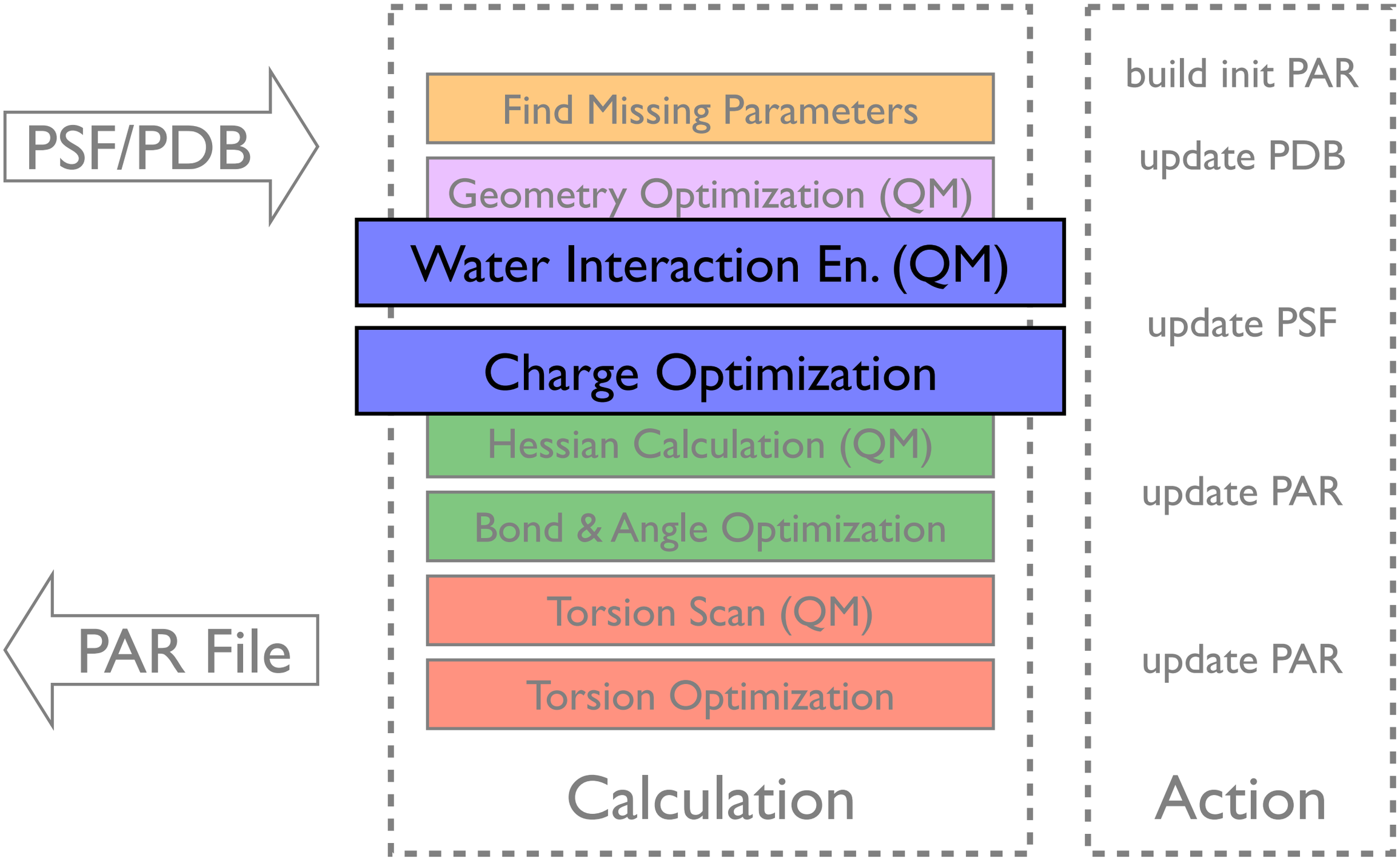
Support Functions

- 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

ffTK Exemplified by Charge Optimization



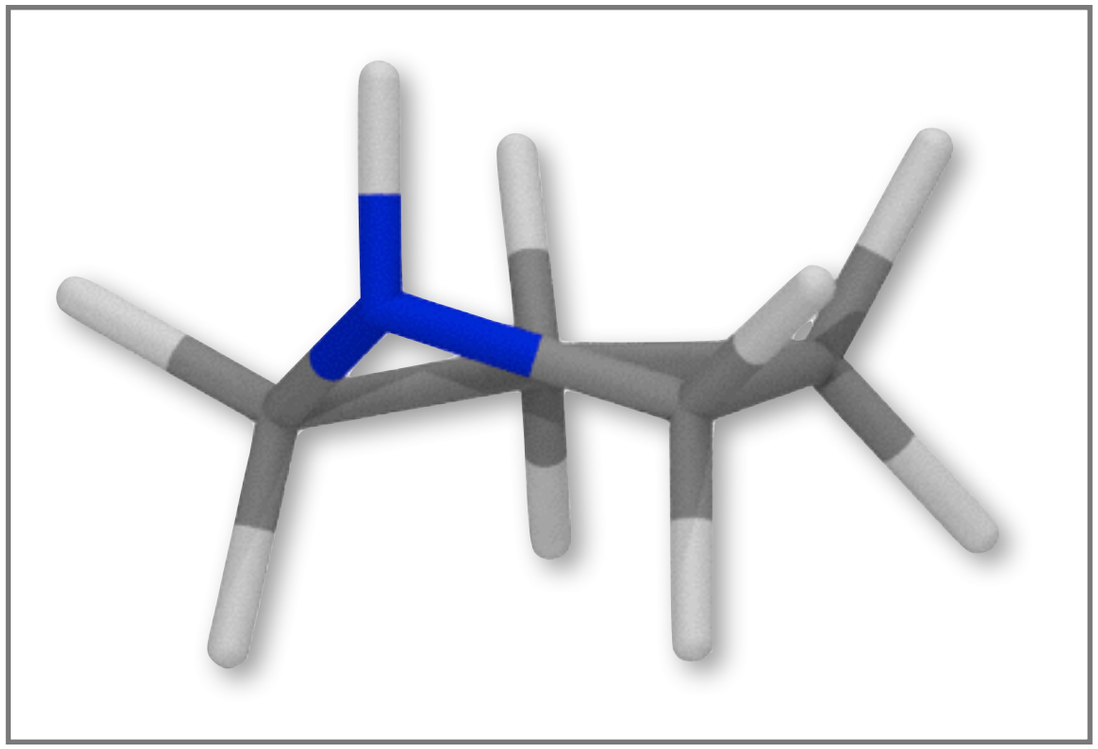
ffTK Exemplified by Charge Optimization



Generating Charge Optimization Target Data

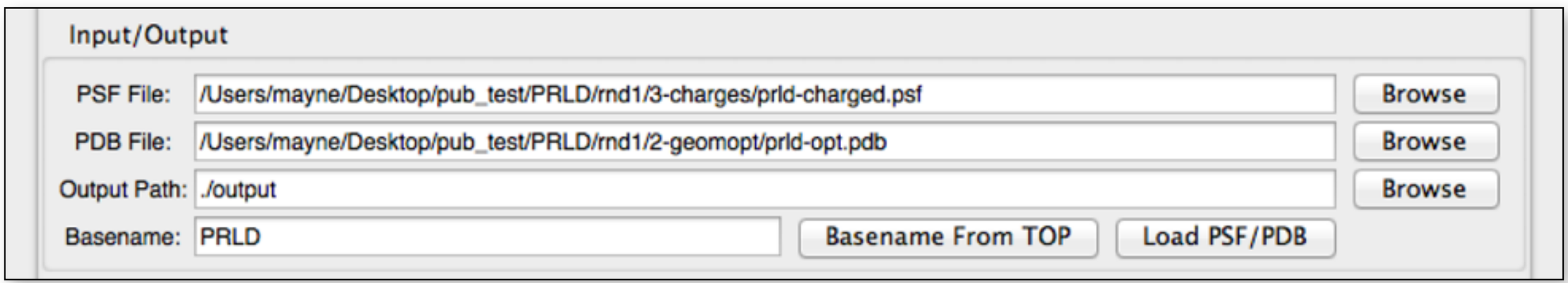
Load QM optimized geometry | Auto-detect interaction sites | Genera

pyrrolidine



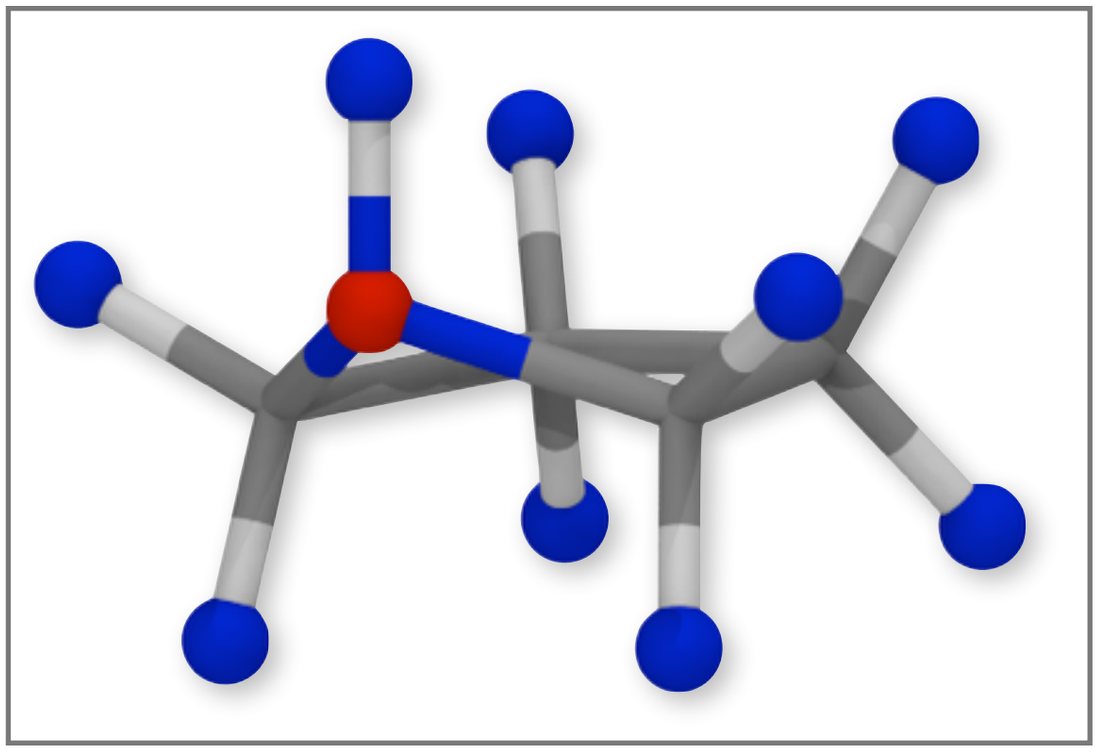
VMD main window

ffTK GUI



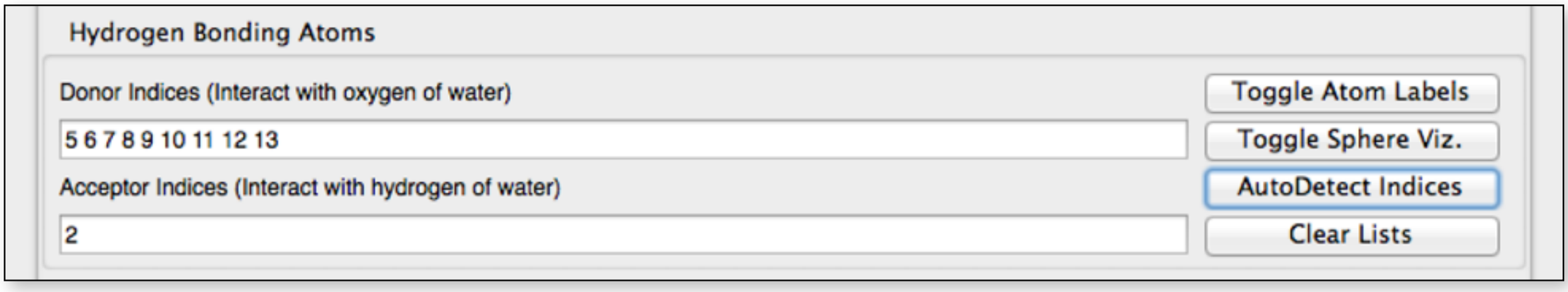
Generating Charge Optimization Target Data

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



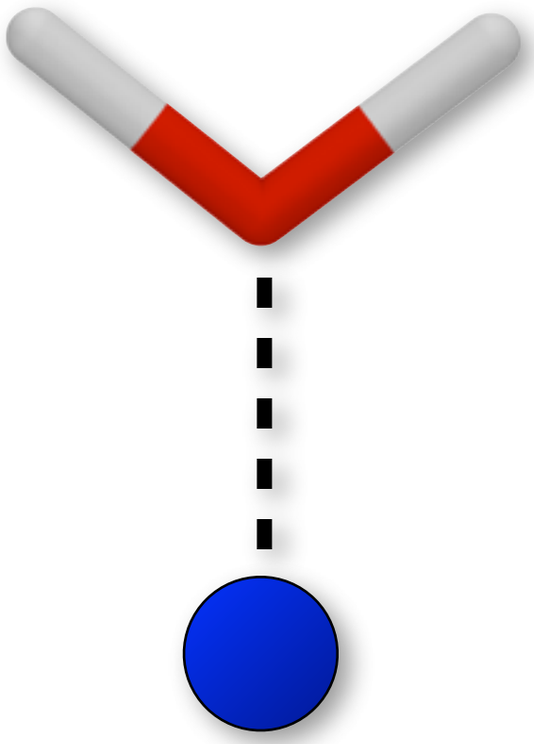
VMD main window

ffTK GUI

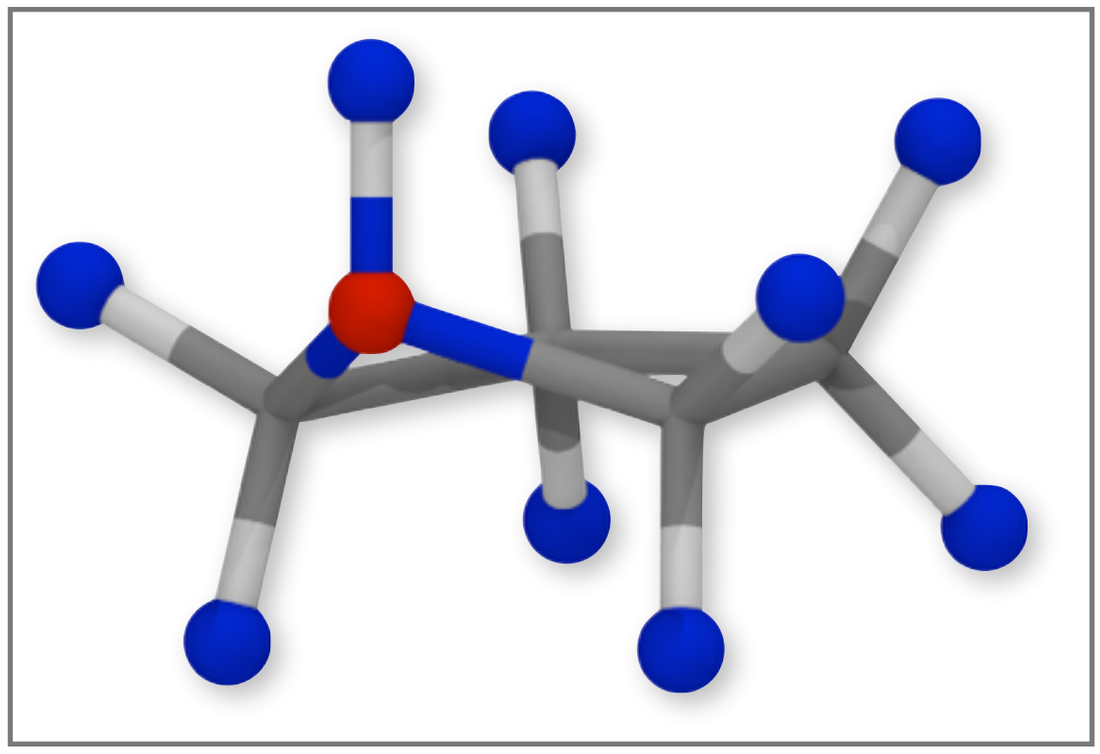


Generating Charge Optimization Target Data

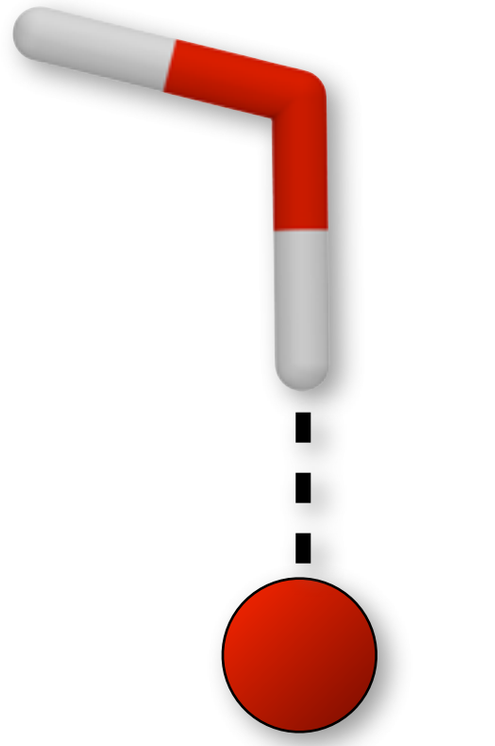
geometry | Auto-detect interaction sites | Generate Gaussian Input Files | Run



Donor

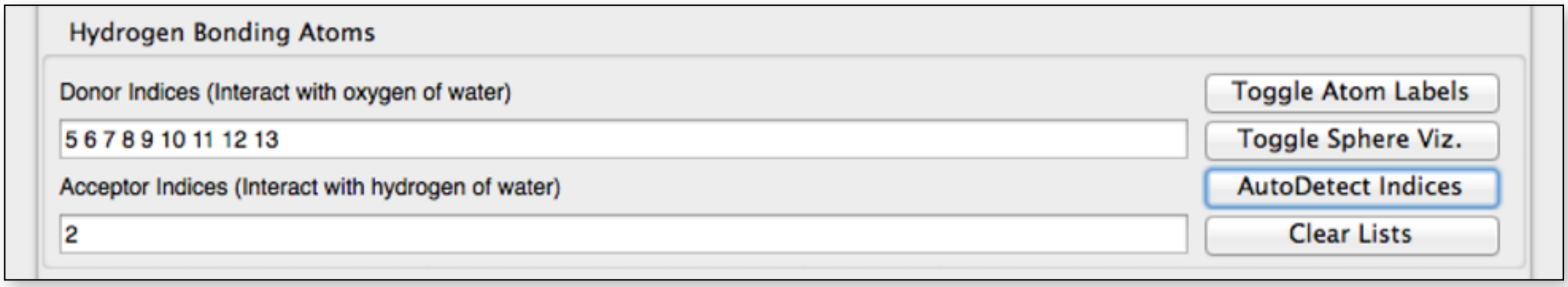


VMD main window



Acceptor

ffTK GUI

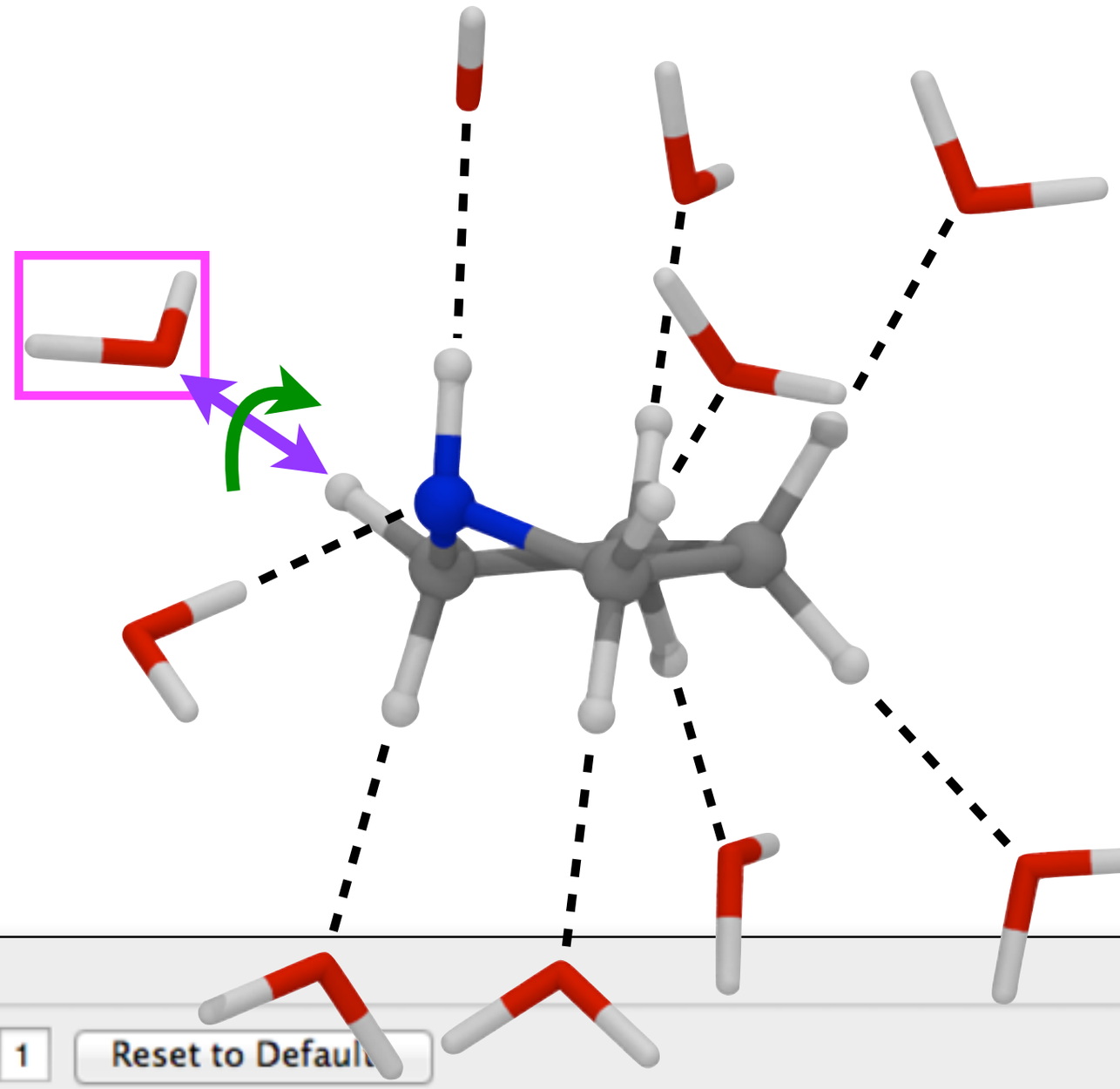


Generating Charge Optimization Target Data

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

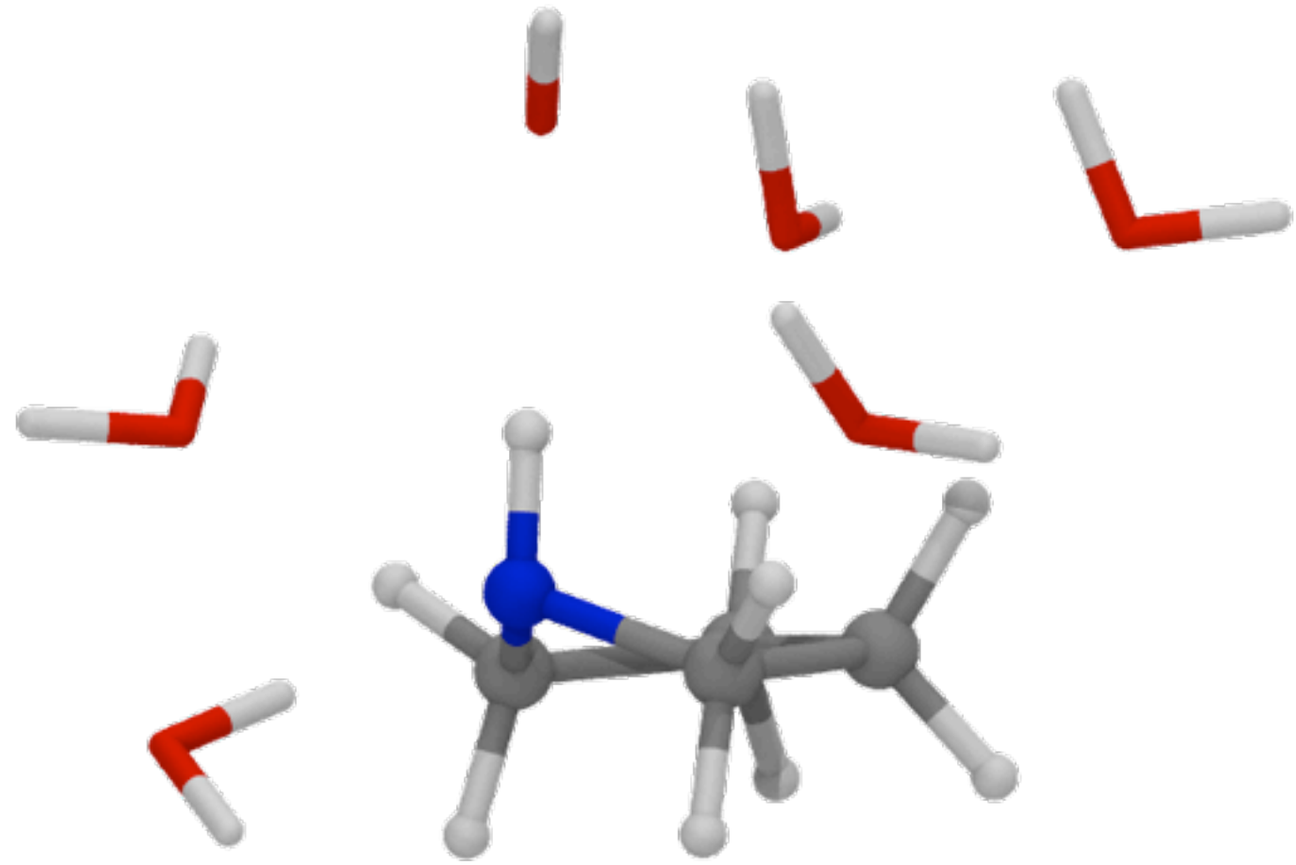
Route: # HF/6-31G* Opt=(Z-matrix,MaxCycles=100) Geom=PrintInputOrient

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

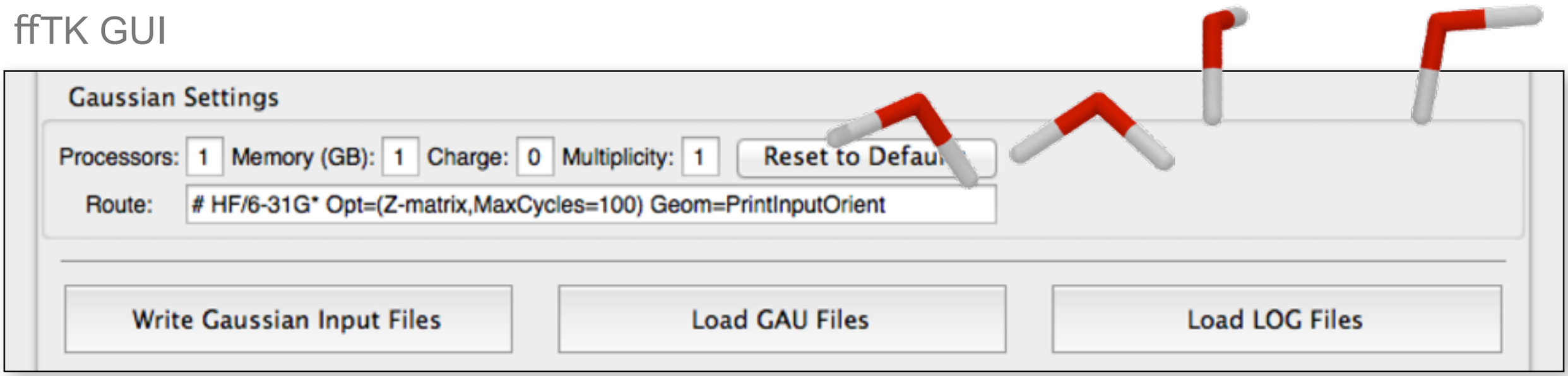
Generating Charge Optimization Target Data

| Run QM | Inspect water optimization

Visually assess QM-optimized water position(s)



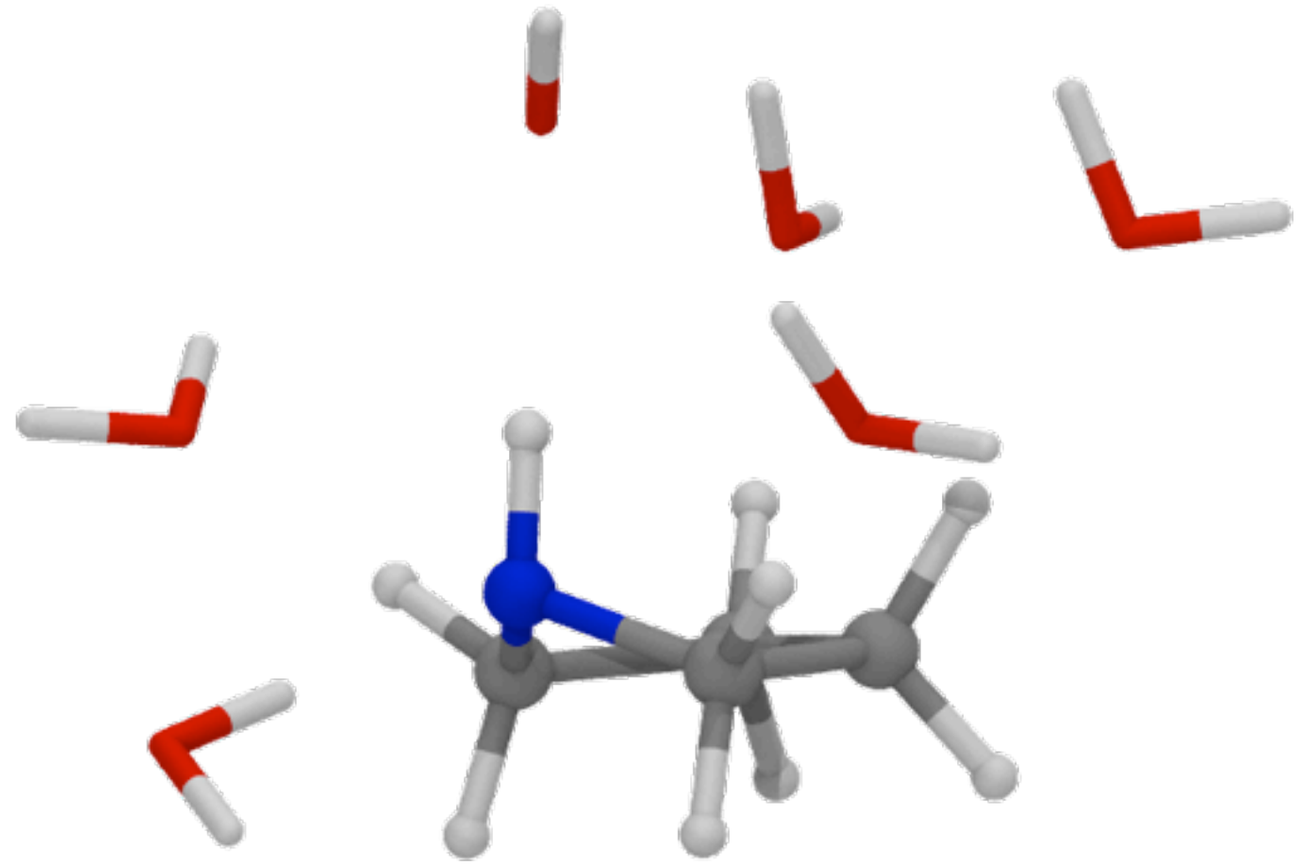
ffTK GUI



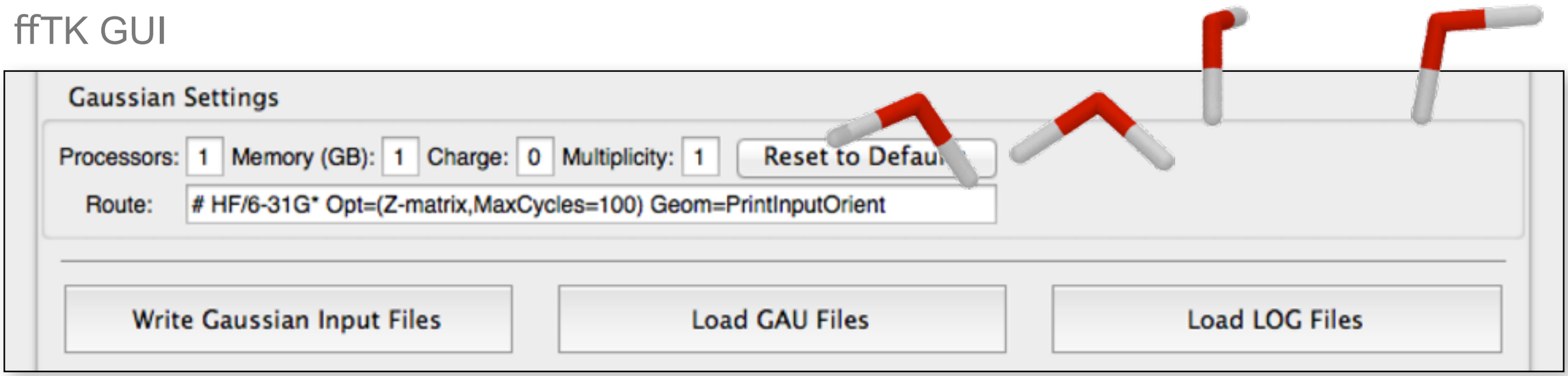
Generating Charge Optimization Target Data

| Run QM | Inspect water optimization

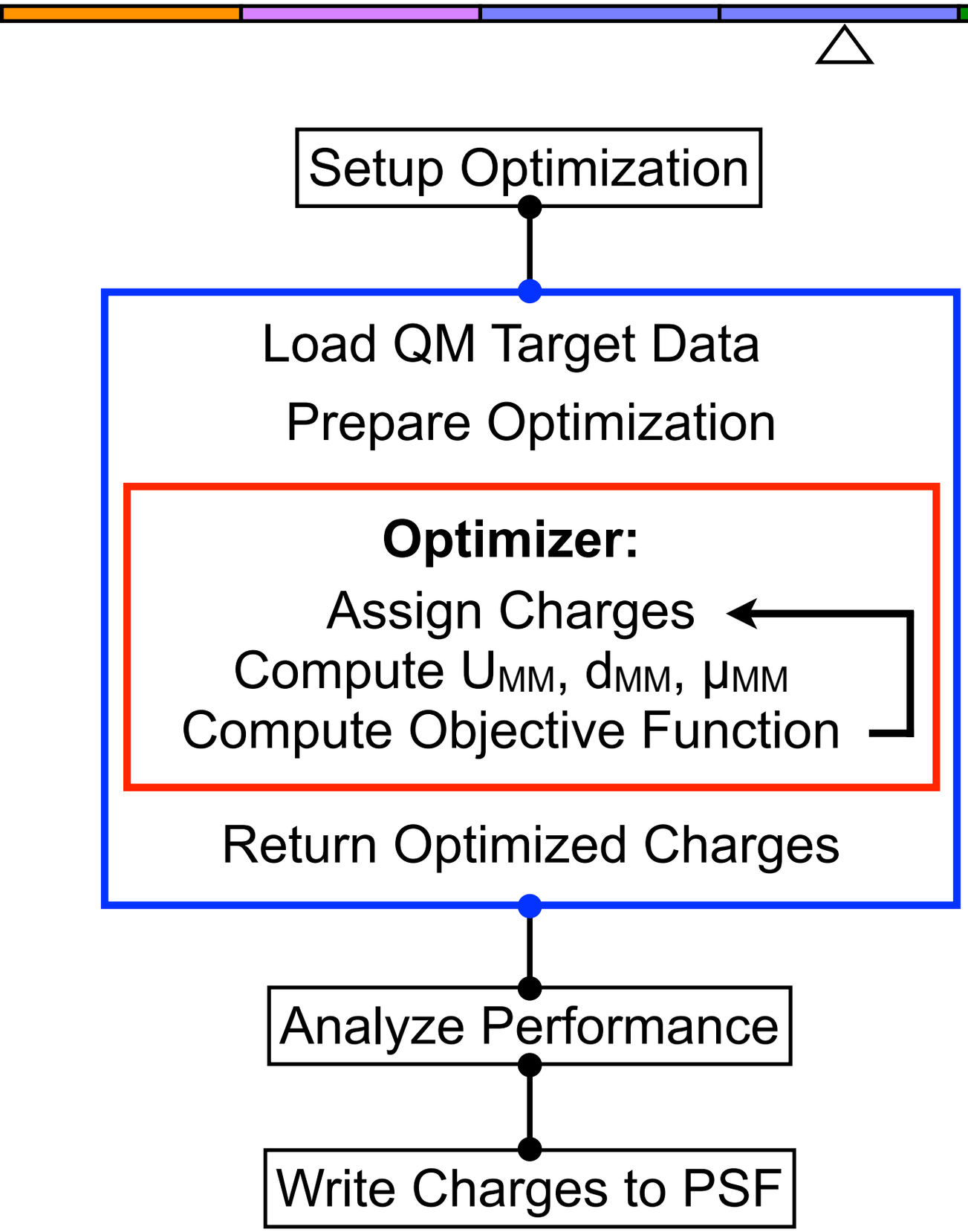
Visually assess QM-optimized water position(s)



ffTK GUI



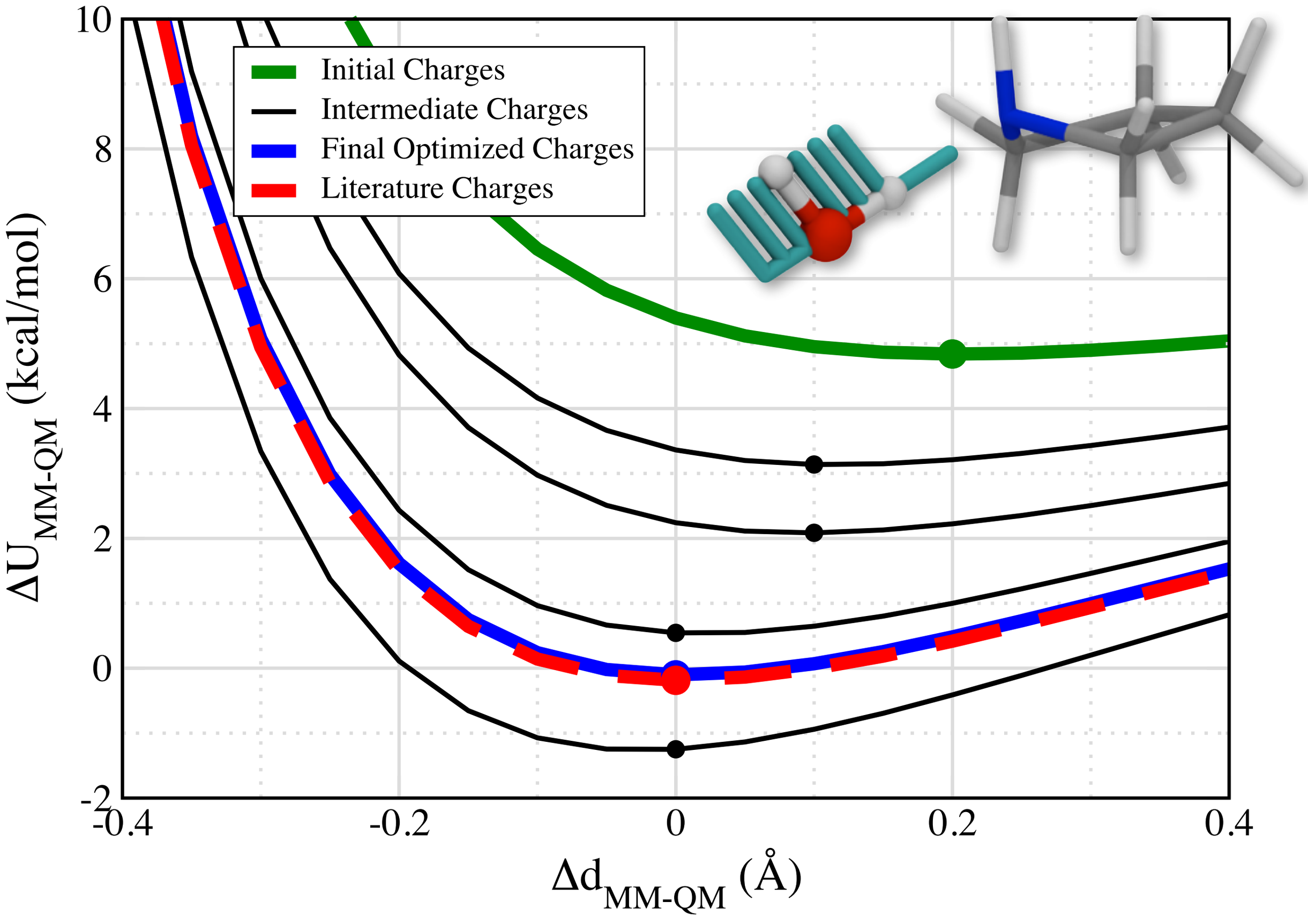
Charge Optimization



Objective Function

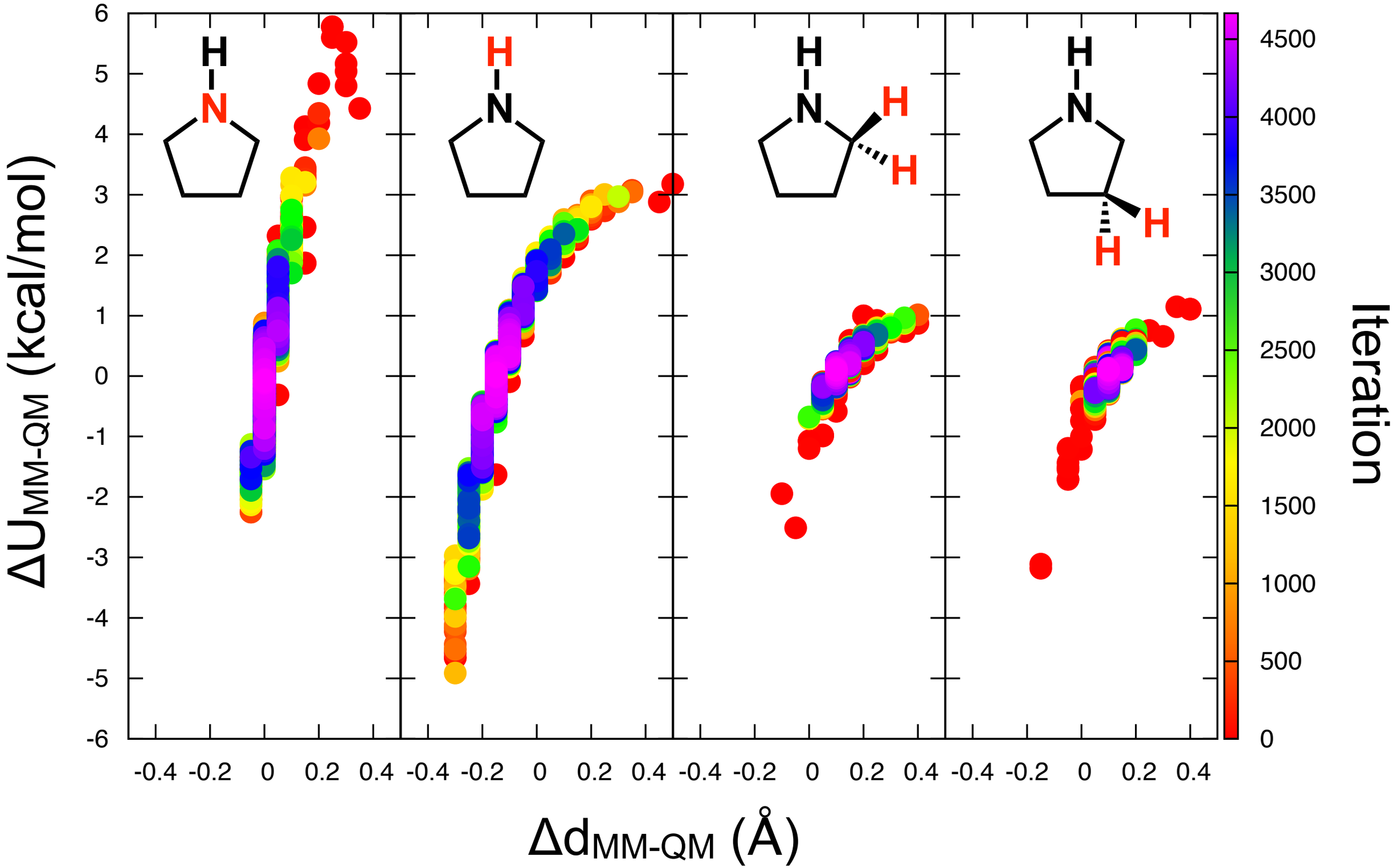
$$\sum_{\text{wat. int.}} f(U_{MM}\text{-}U_{QM}) + \sum_{\text{wat. int.}} f(d_{MM}\text{-}d_{QM}) + f(\mu_{MM}\text{-}\mu_{QM})$$

Assessing MM Water-Interaction Profiles



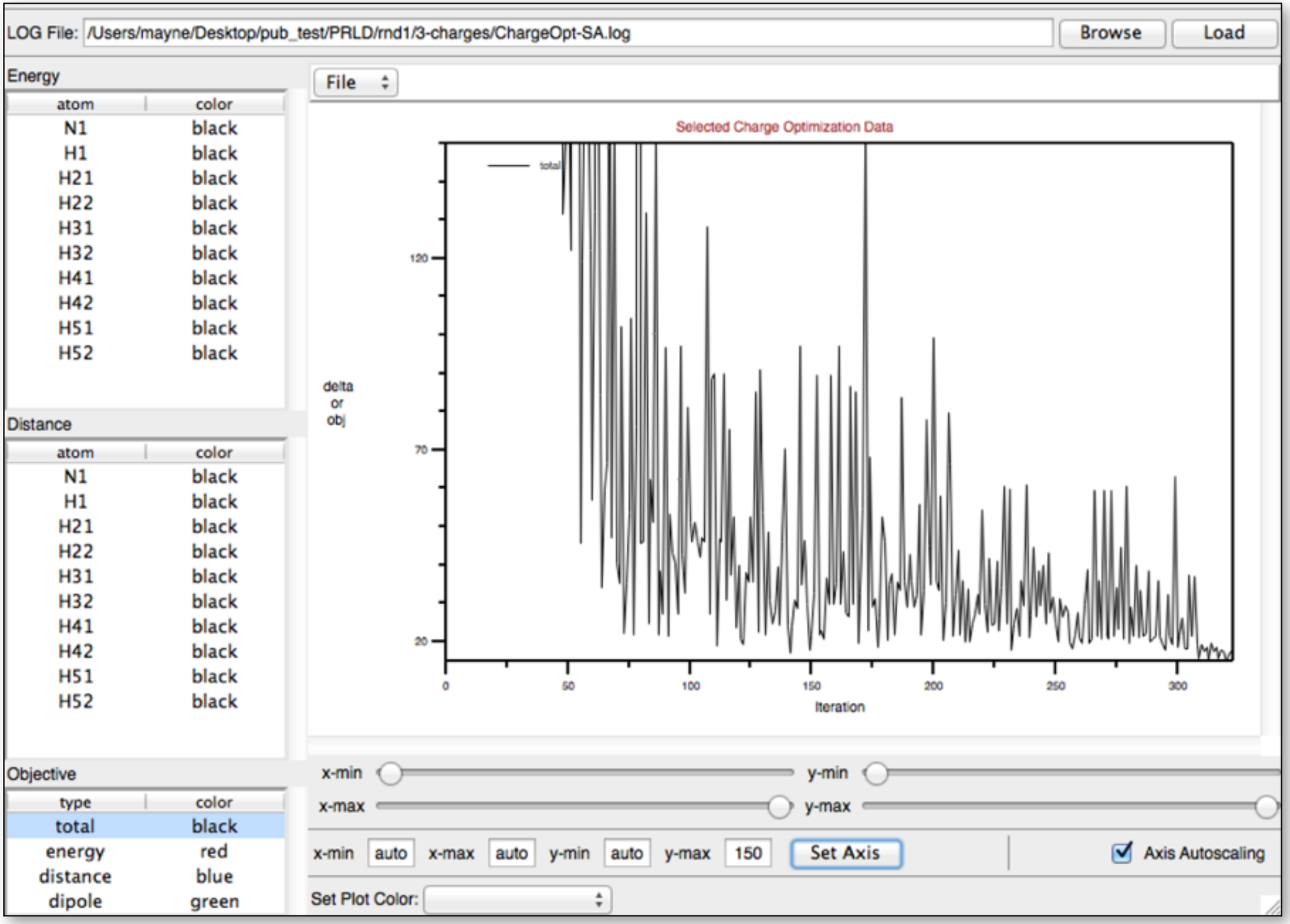
Sampling MM Water-Interaction Profiles

Mode: Simulated Annealing

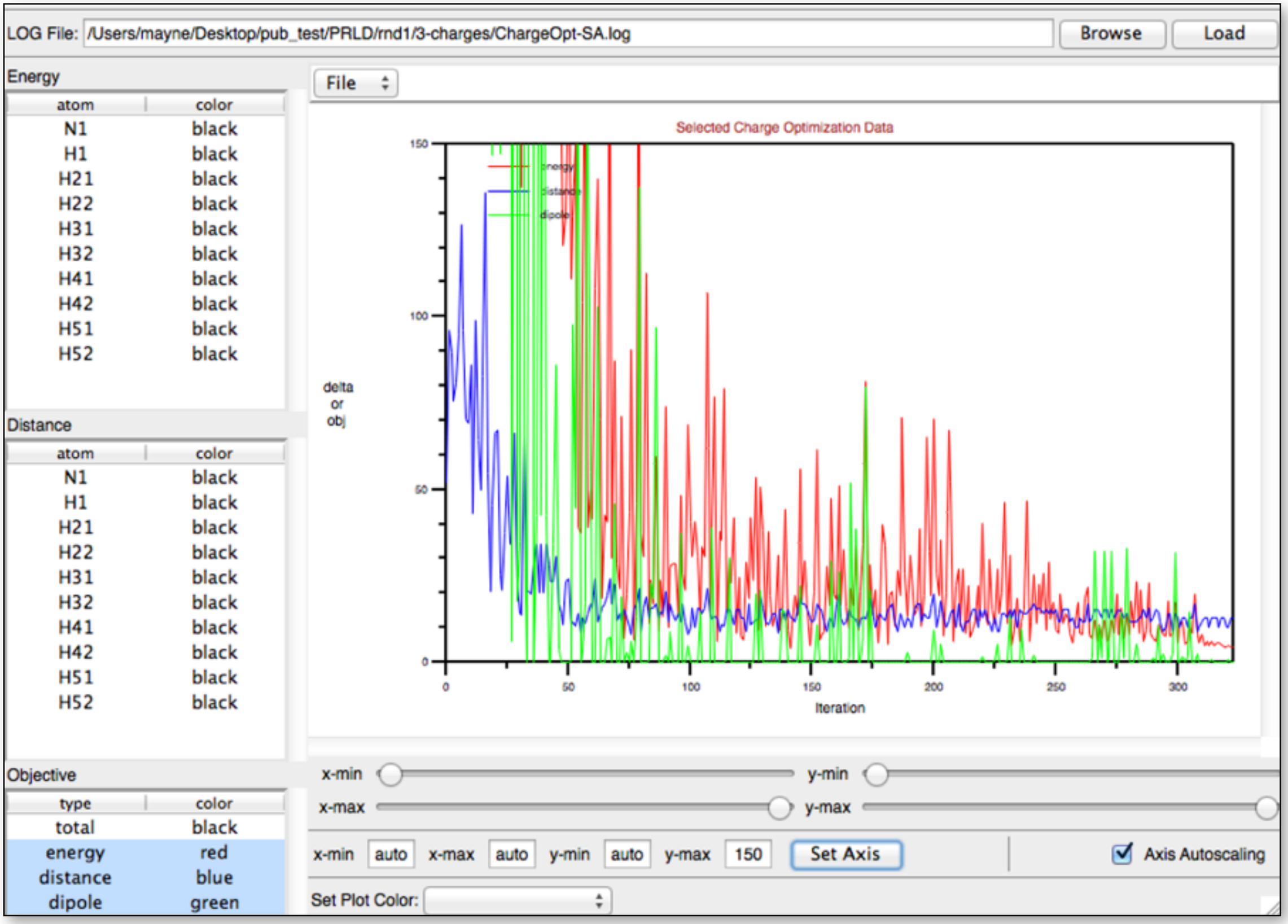


Plotting Charge Optimization Data

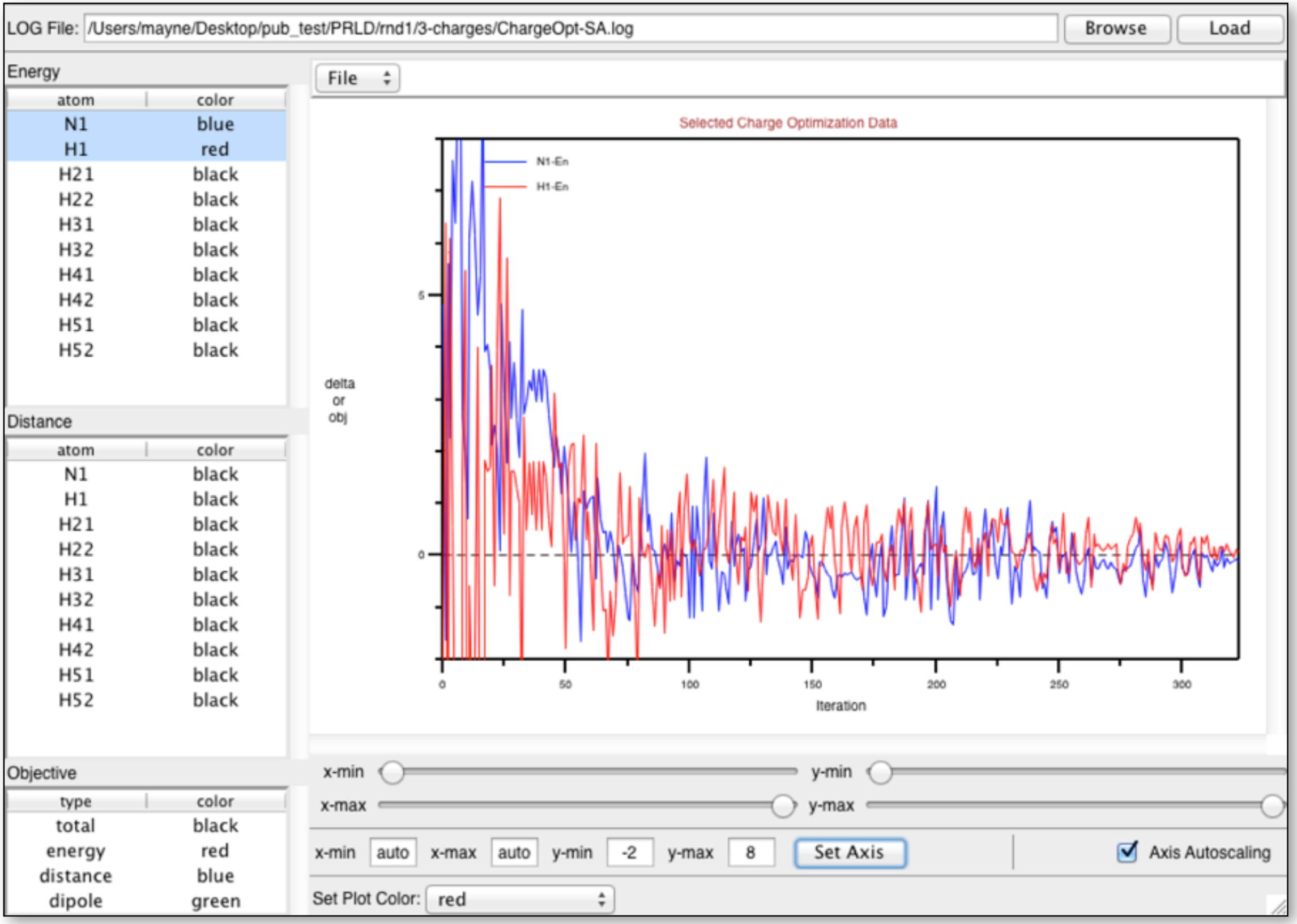
Plotting Charge Optimization Data



Plotting Charge Optimization Data



Plotting Charge Optimization Data



Plotting Charge Optimization Data

LOG File: /Users/mayne/Desktop/pub_test/PRLD/rnd1/3-charges/ChargeOpt-SA.log

File

export data

Energy

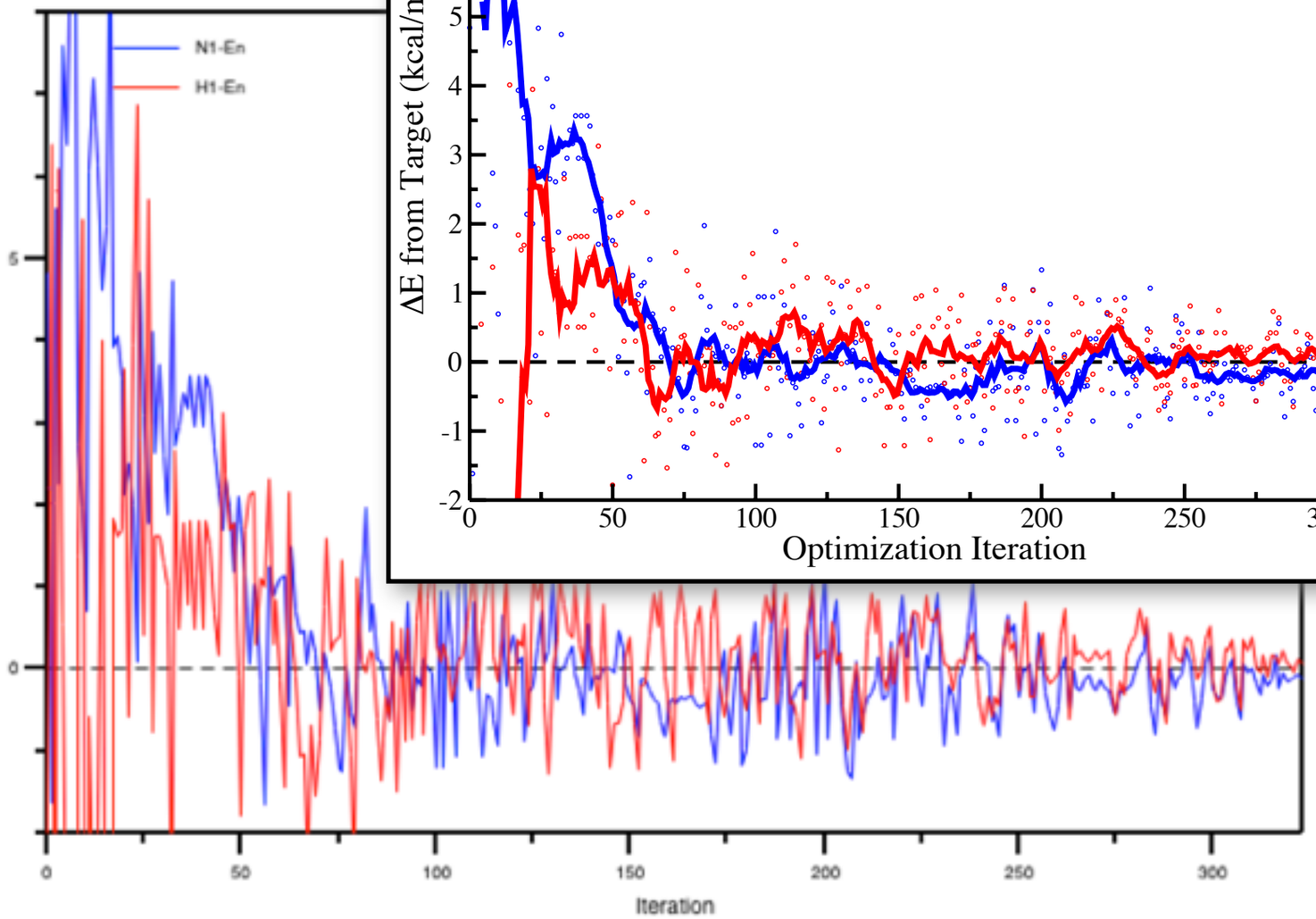
atom	color
N1	blue
H1	red
H21	black
H22	black
H31	black
H32	black
H41	black
H42	black
H51	black
H52	black

Distance

atom	color
N1	black
H1	black
H21	black
H22	black
H31	black
H32	black
H41	black
H42	black
H51	black
H52	black

Objective

type	color
total	black
energy	red
distance	blue
dipole	green



x-min y-min

x-max y-max

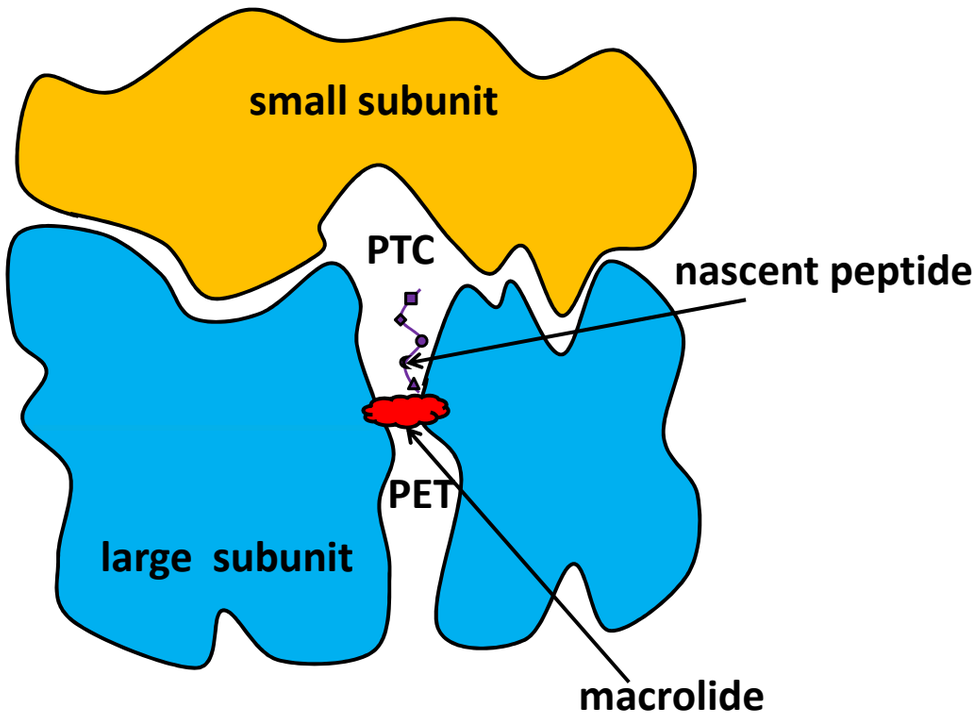
x-min auto x-max auto y-min -2 y-max 8 Set Axis

Axis Autoscaling

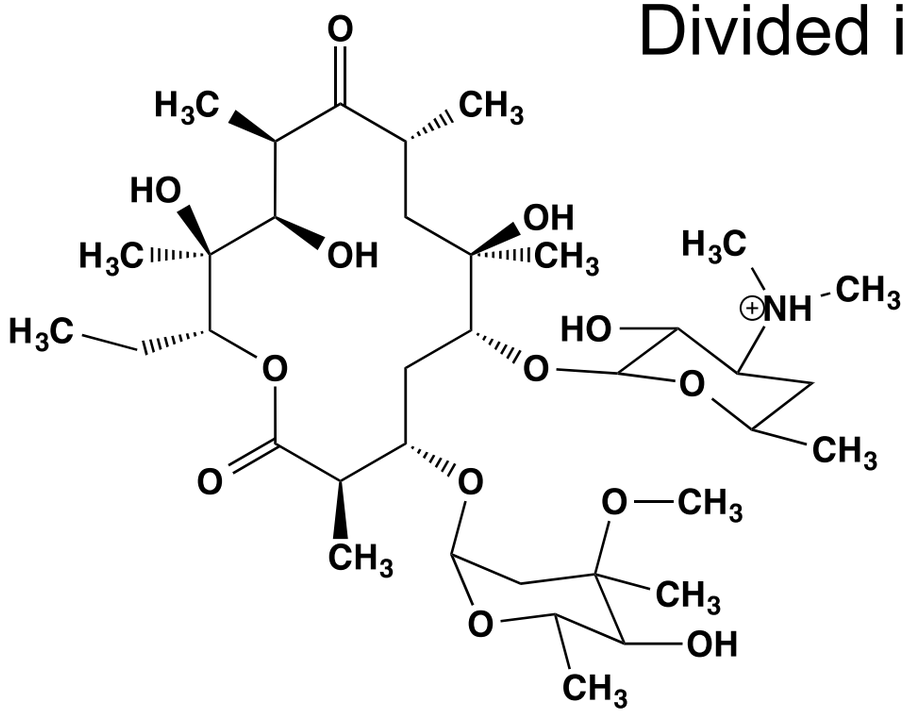
Set Plot Color: red

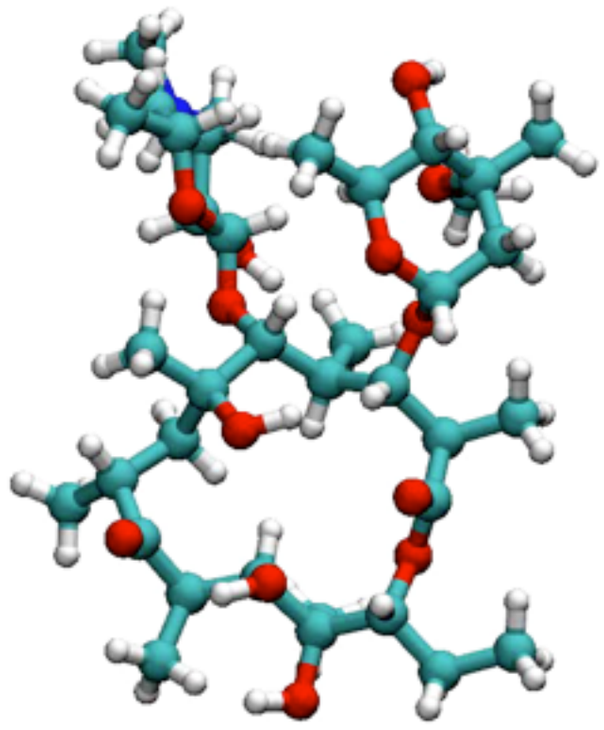
Parameterization of Erythromycin

macrolides - antibiotics that prevent peptide synthesis in bacterial ribosomes



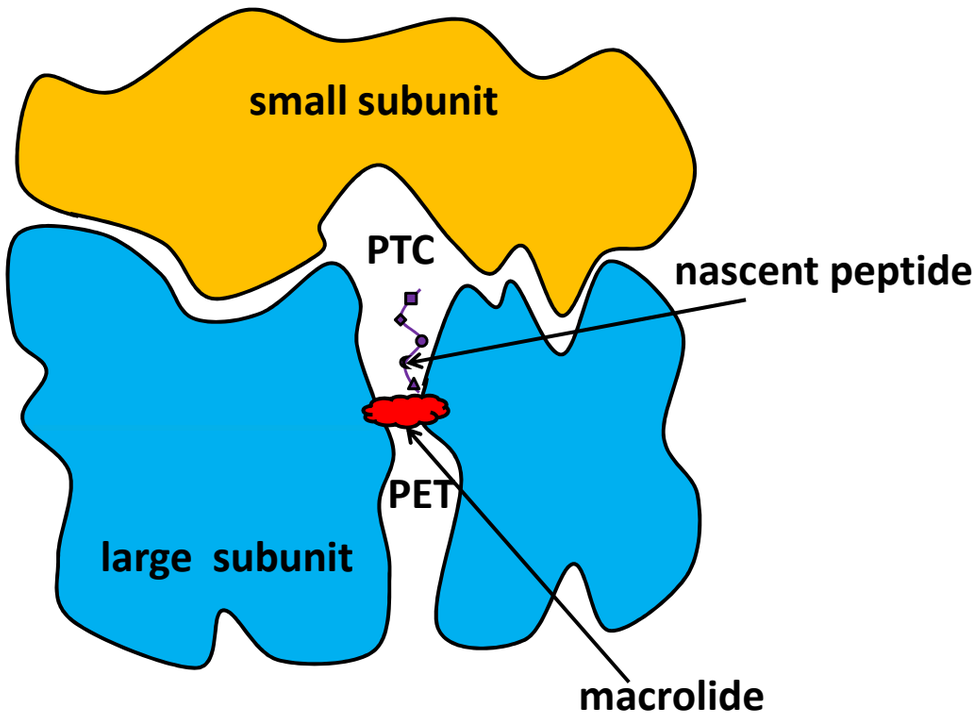
Divided into four parts



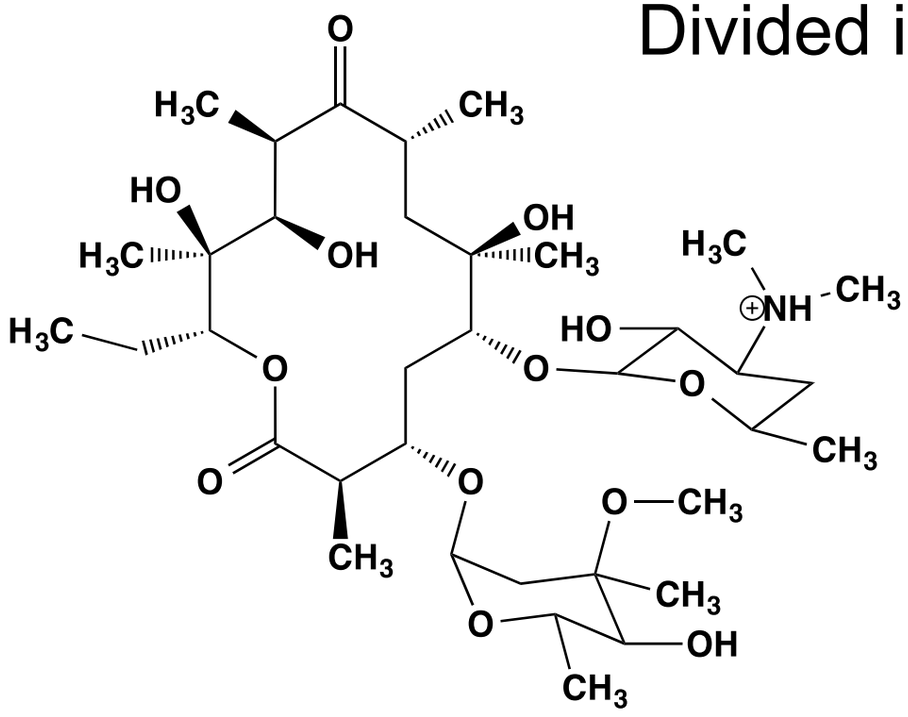


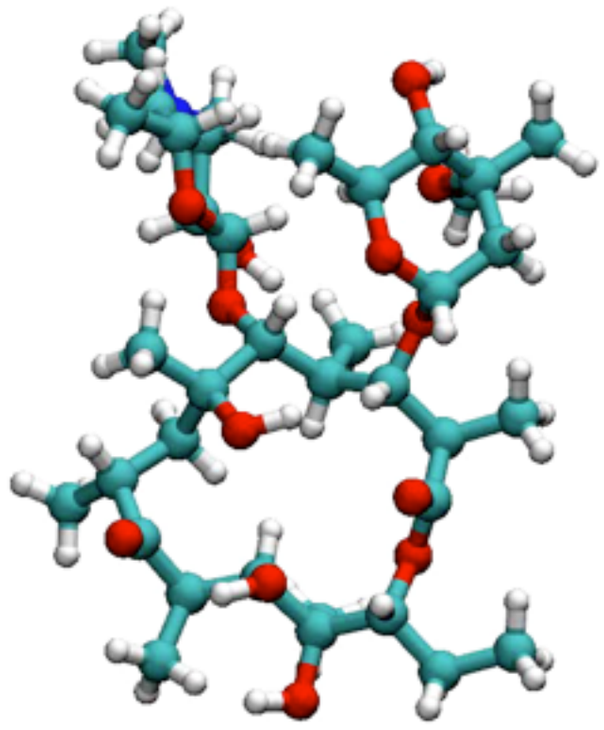
Parameterization of Erythromycin

macrolides - antibiotics that prevent peptide synthesis in bacterial ribosomes



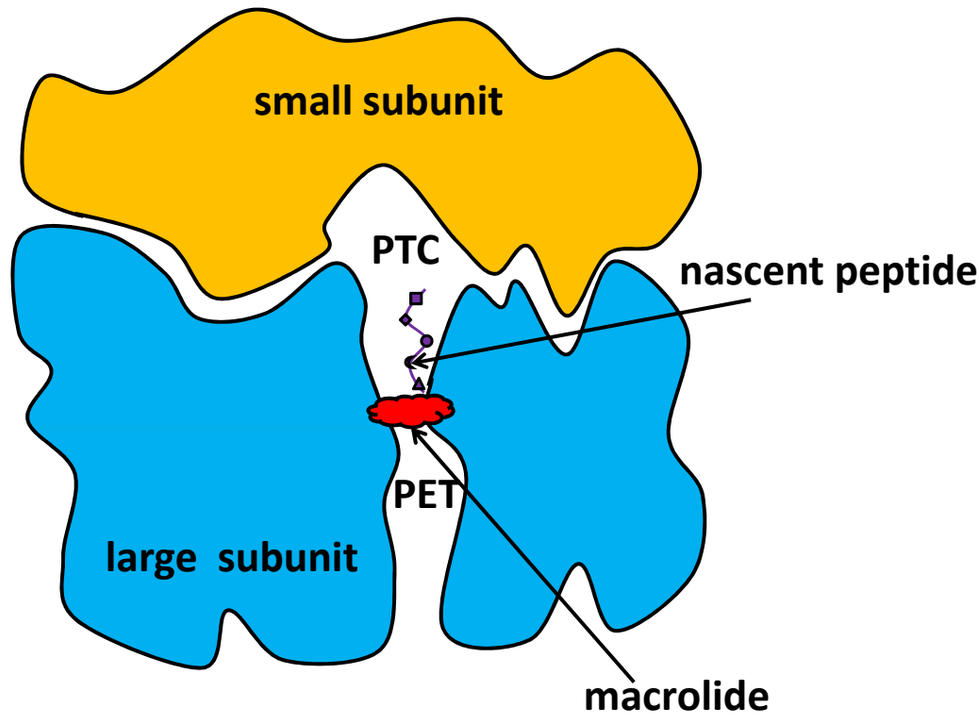
Divided into four parts



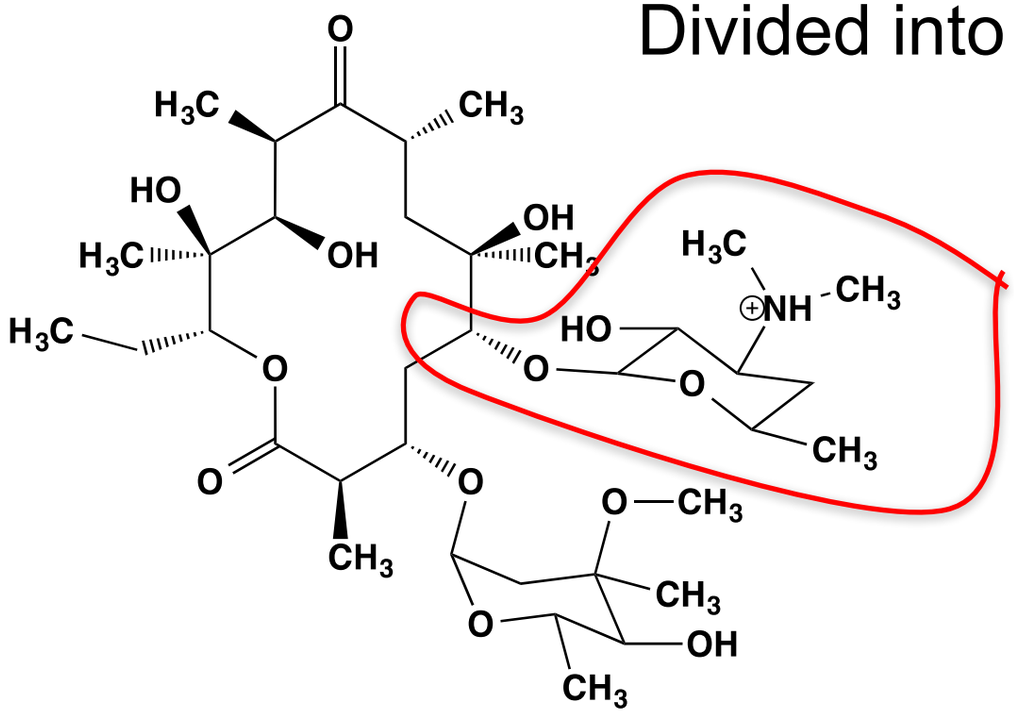


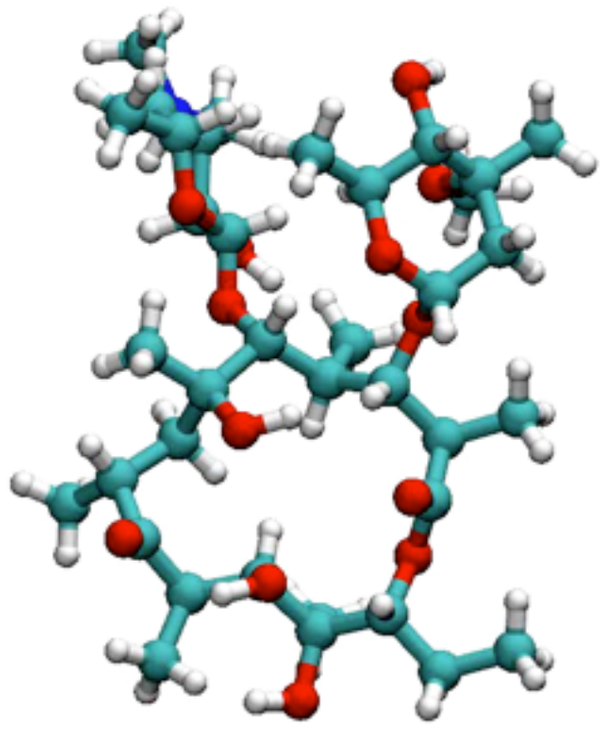
Parameterization of Erythromycin

macrolides - antibiotics that prevent peptide synthesis in bacterial ribosomes



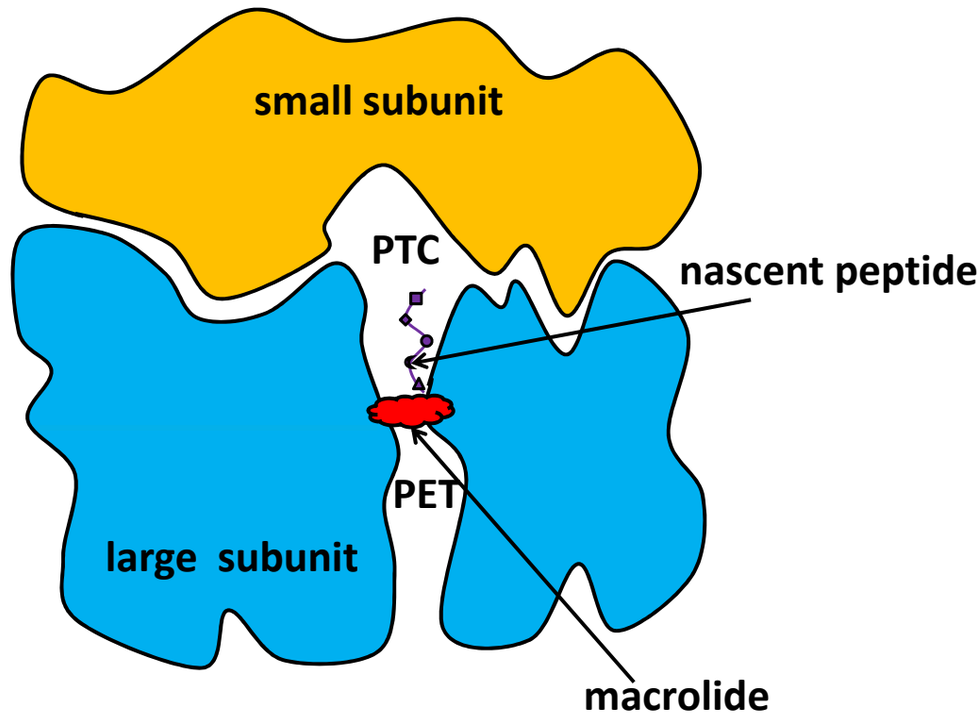
Divided into four parts



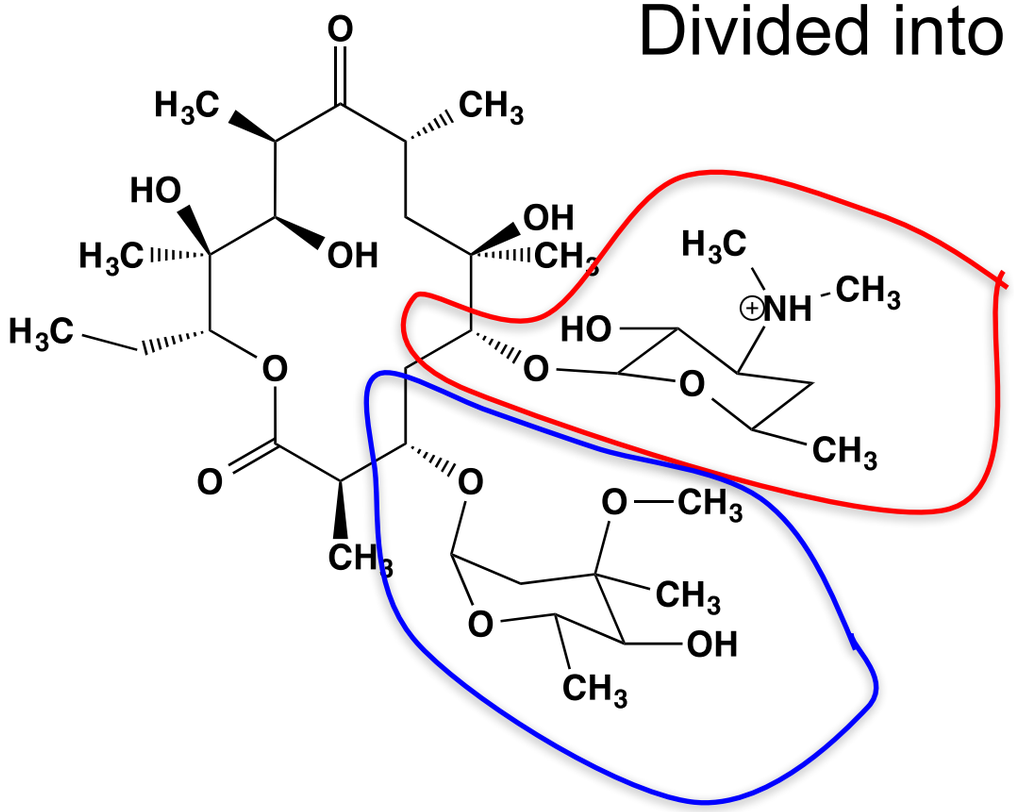


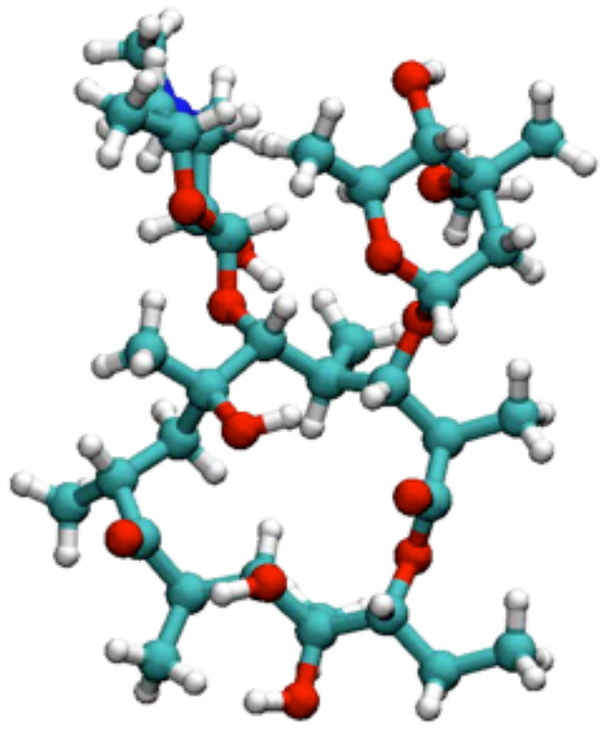
Parameterization of Erythromycin

macrolides - antibiotics that prevent peptide synthesis in bacterial ribosomes



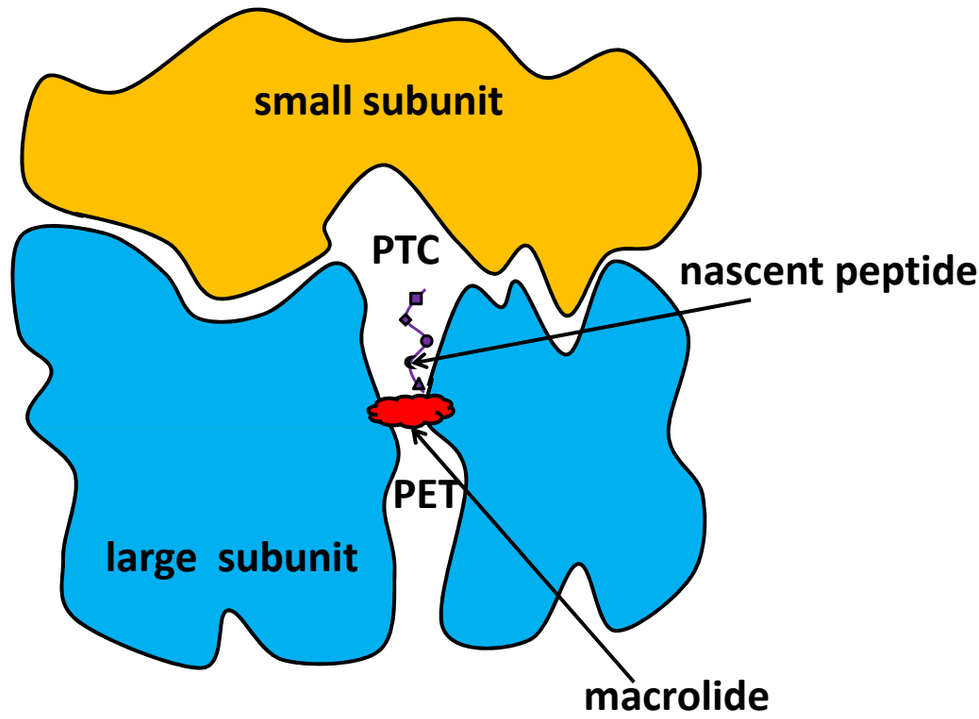
Divided into four parts



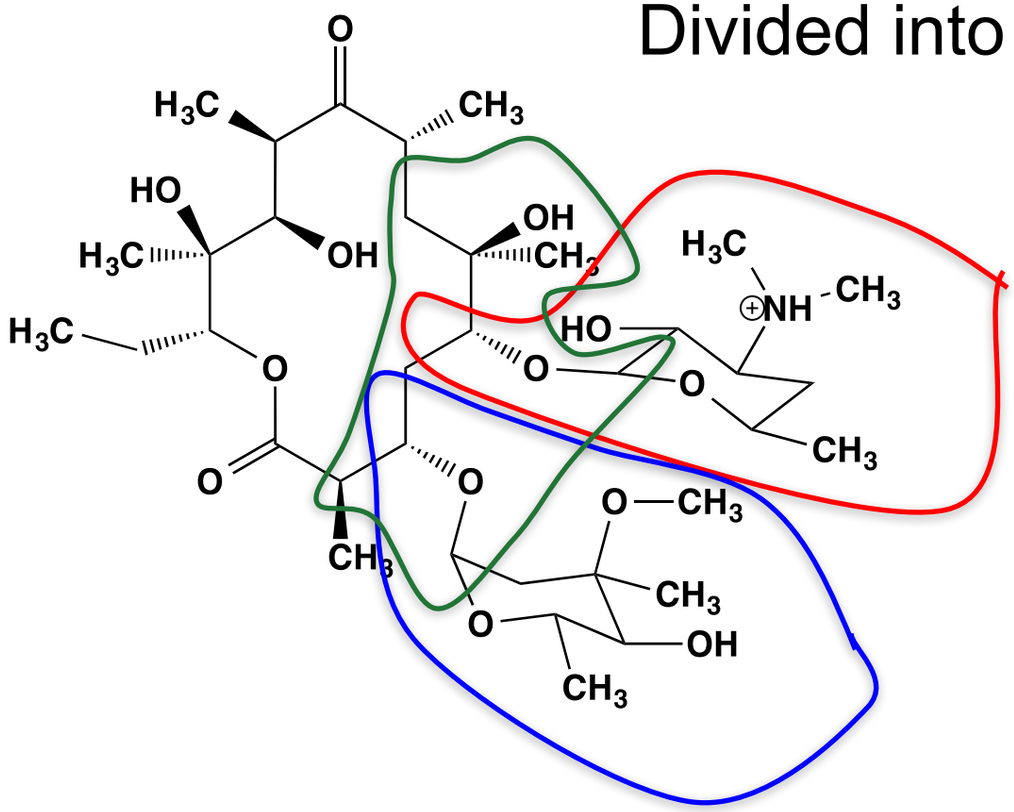


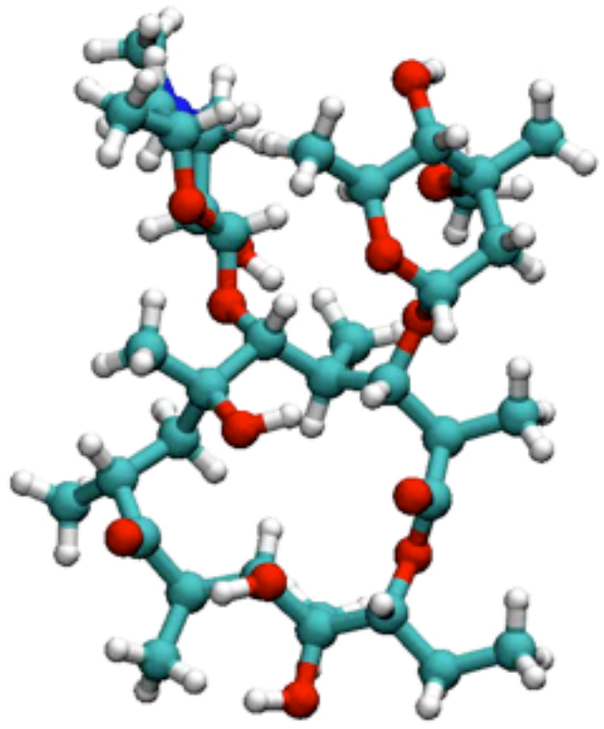
Parameterization of Erythromycin

macrolides - antibiotics that prevent peptide synthesis in bacterial ribosomes



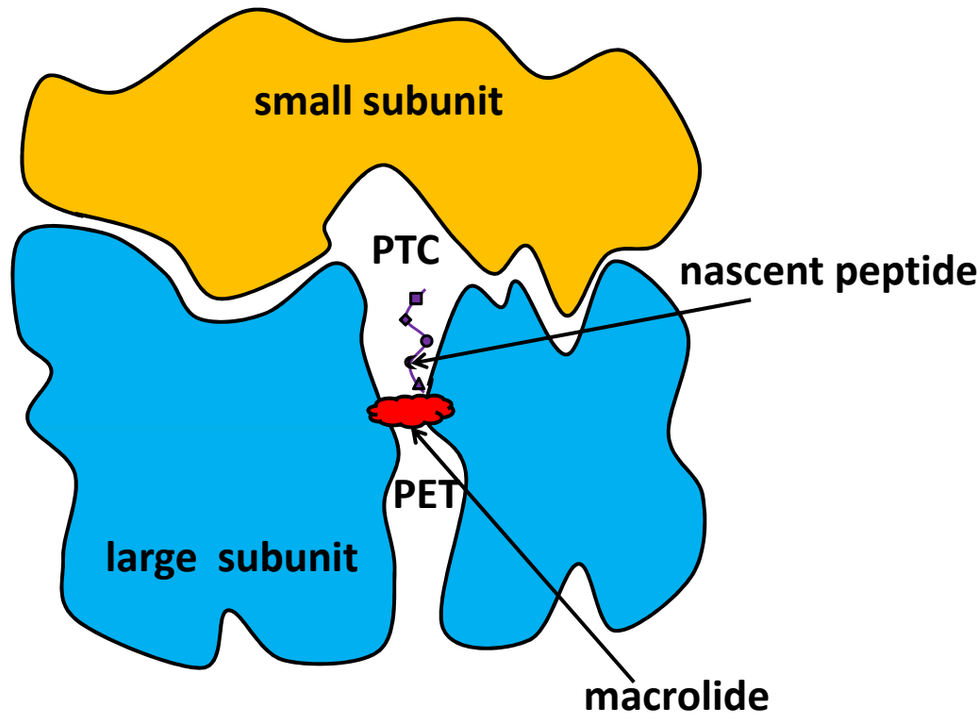
Divided into four parts



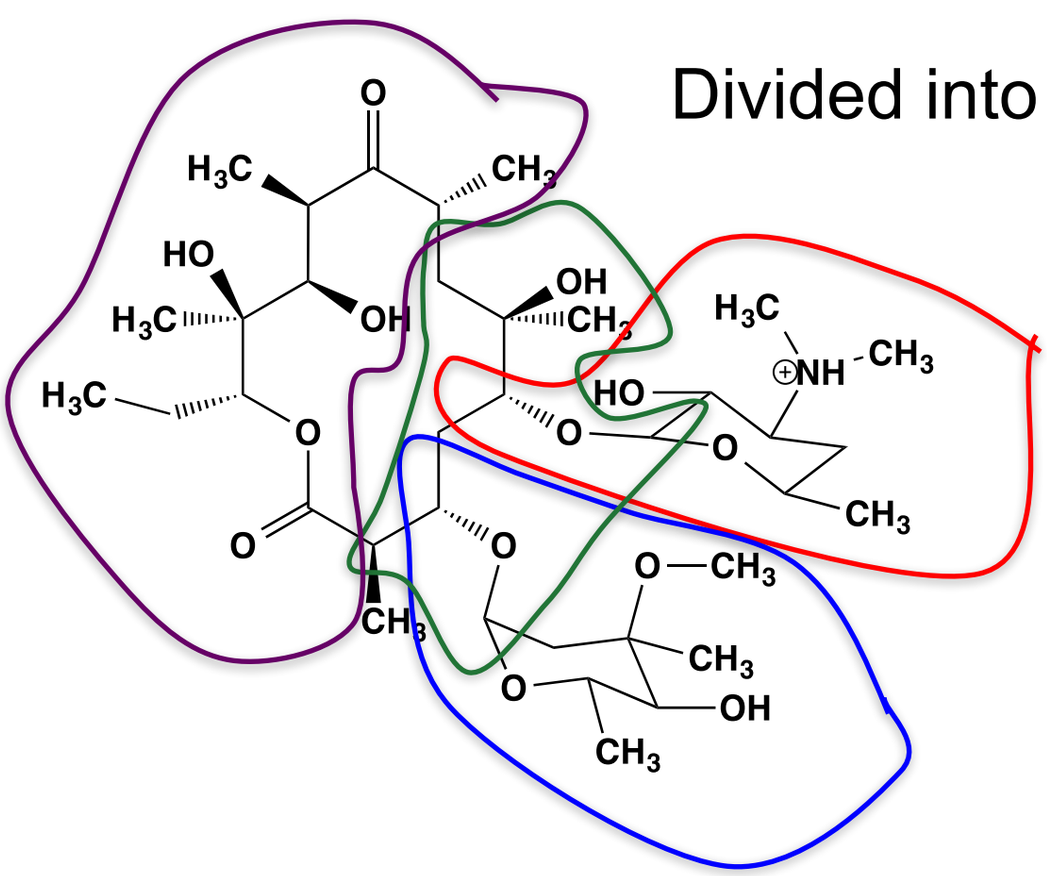


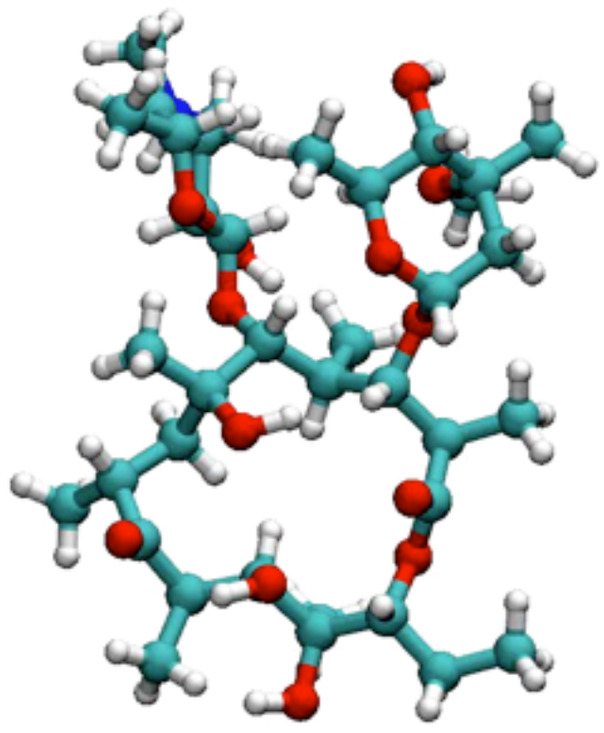
Parameterization of Erythromycin

macrolides - antibiotics that prevent peptide synthesis in bacterial ribosomes



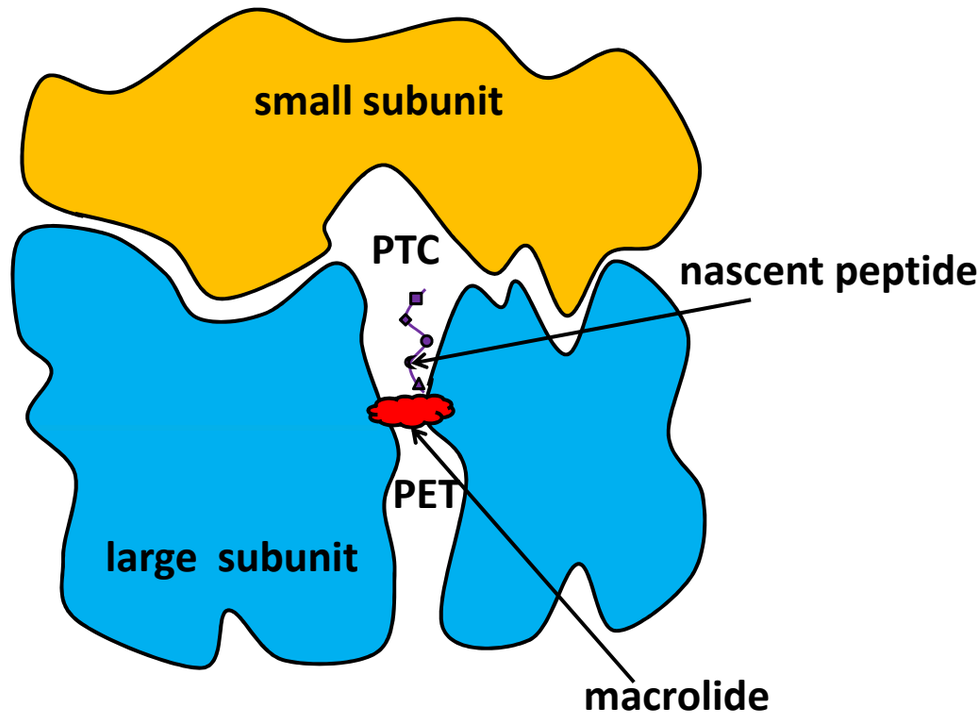
Divided into four parts



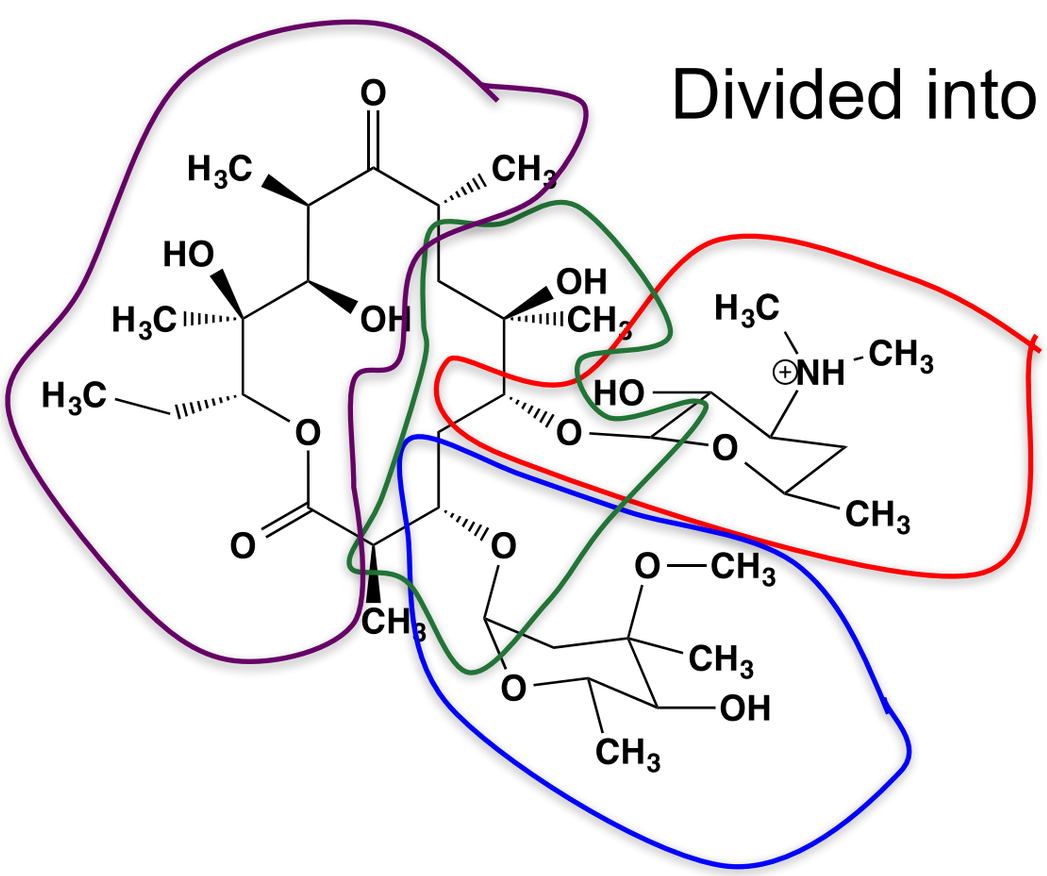


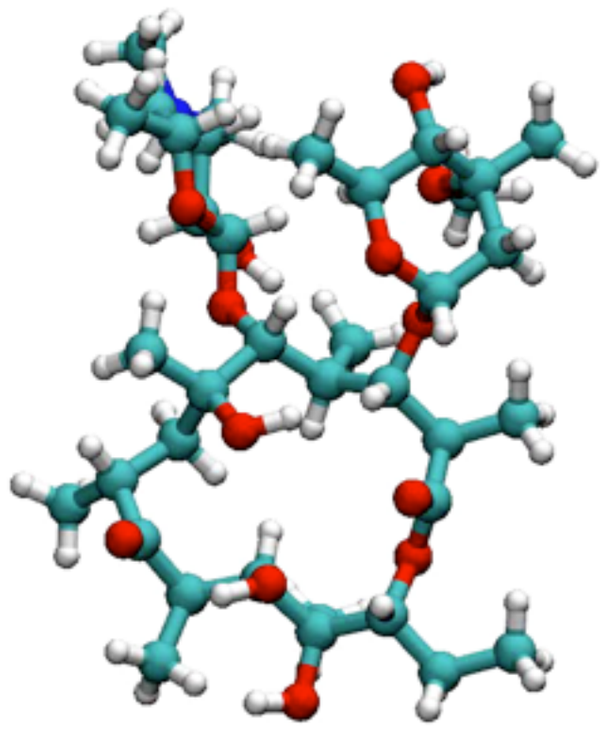
Parameterization of Erythromycin

macrolides - antibiotics that prevent peptide synthesis in bacterial ribosomes



Divided into four parts



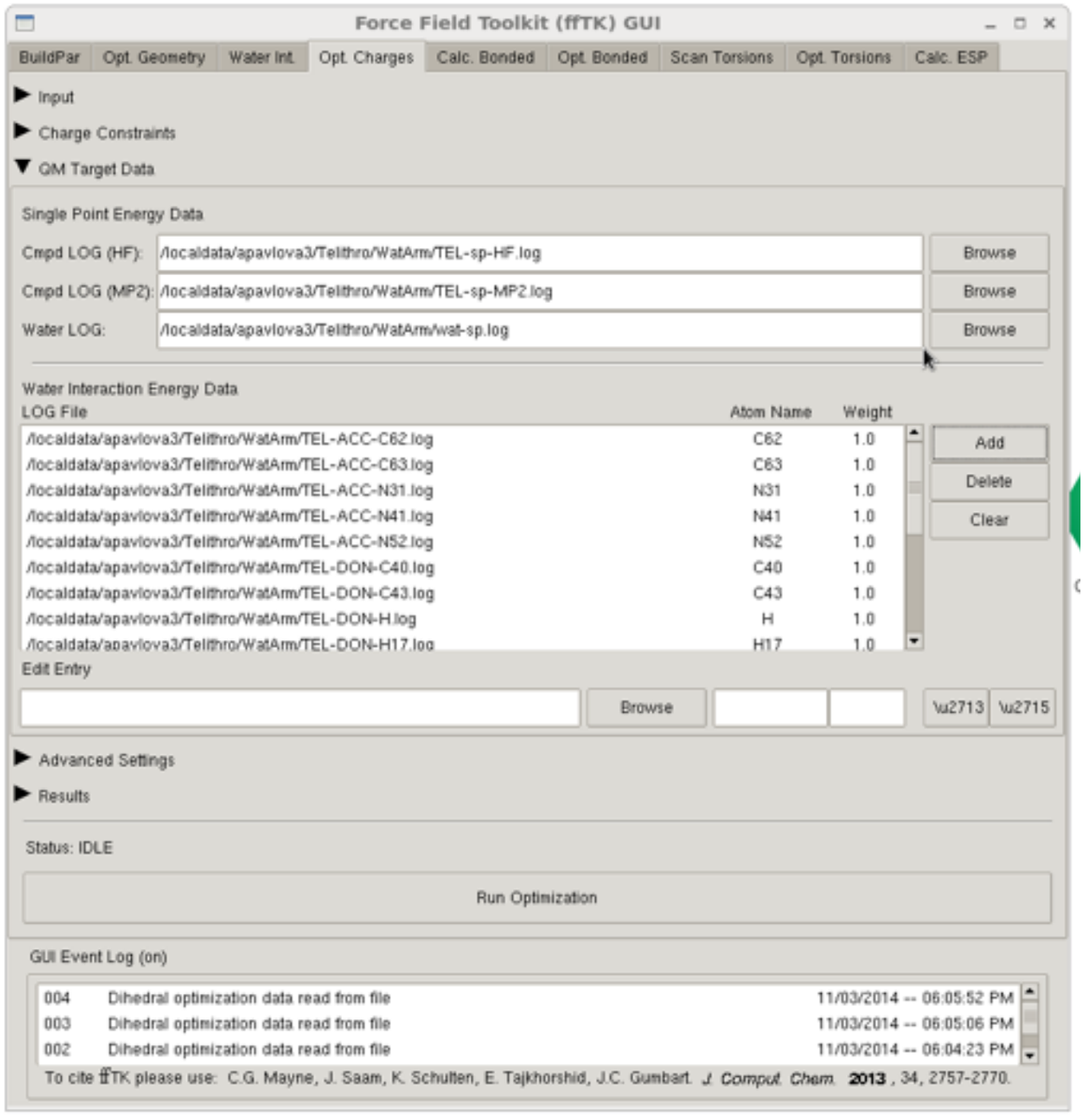


Fitting of Partial Charges

CHARMM partial charges are derived from interactions with water
Calculated for each atom that is fitted

Fitting of Partial Charges

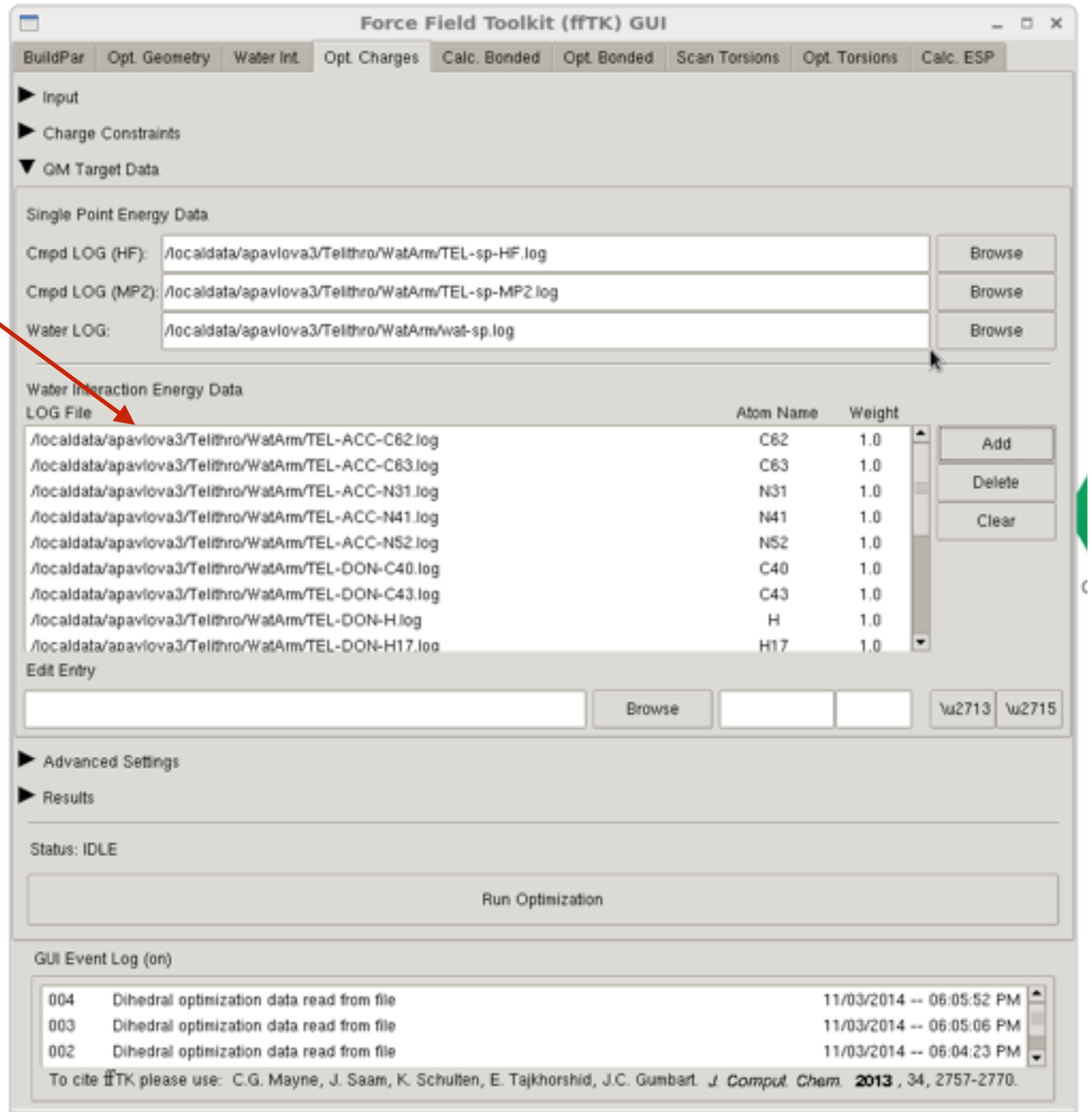
CHARMM partial charges are derived from interactions with water
Calculated for each atom that is fitted



Fitting of Partial Charges

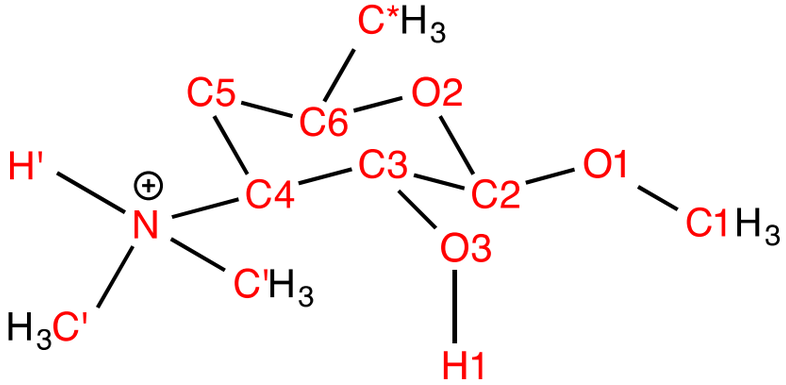
CHARMM partial charges are derived from interactions with water
Calculated for each atom that is fitted

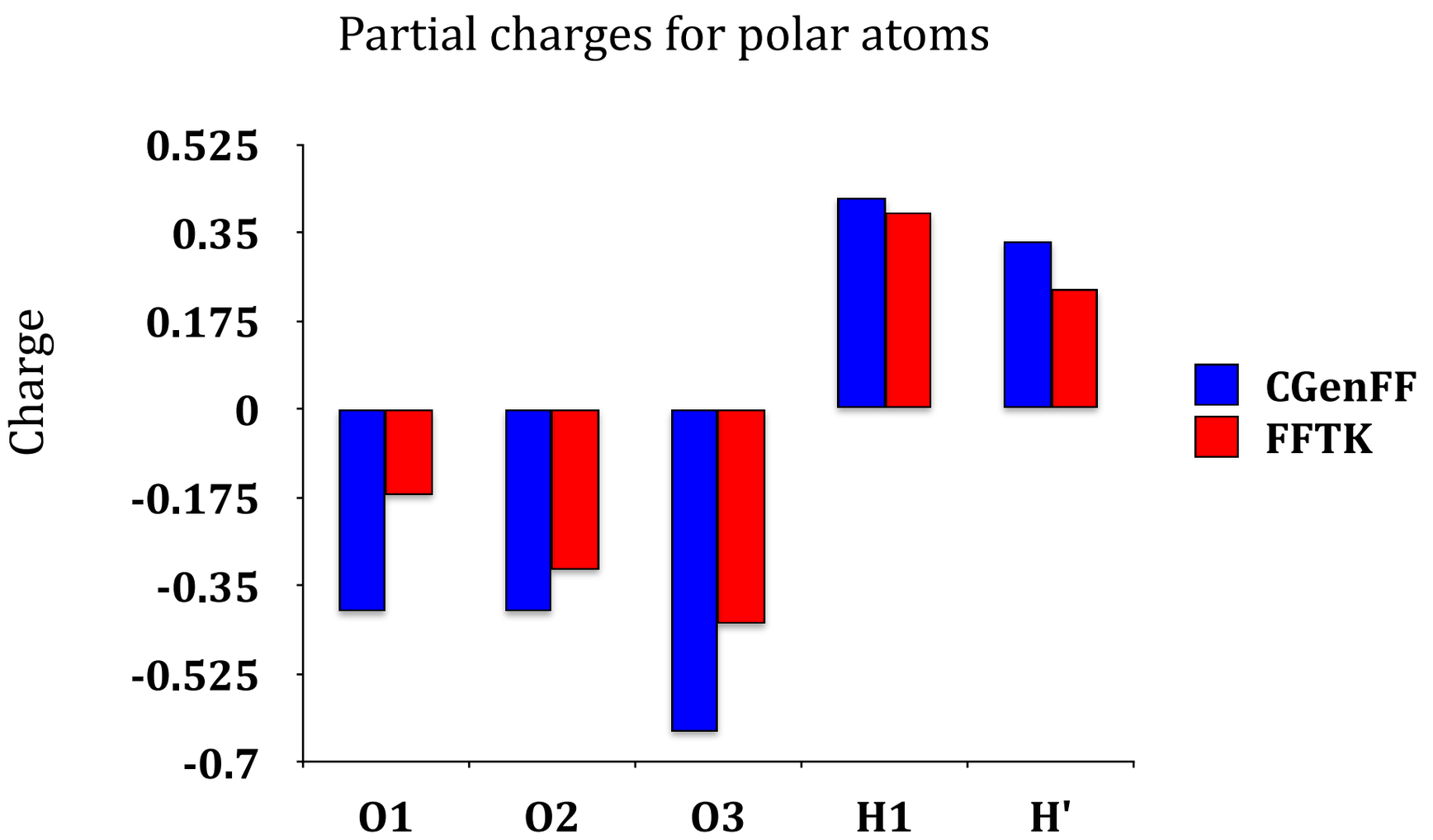
Loaded into FFTK



Accuracy of Fitting: Charges

Desosamine
charges fitted for red atoms

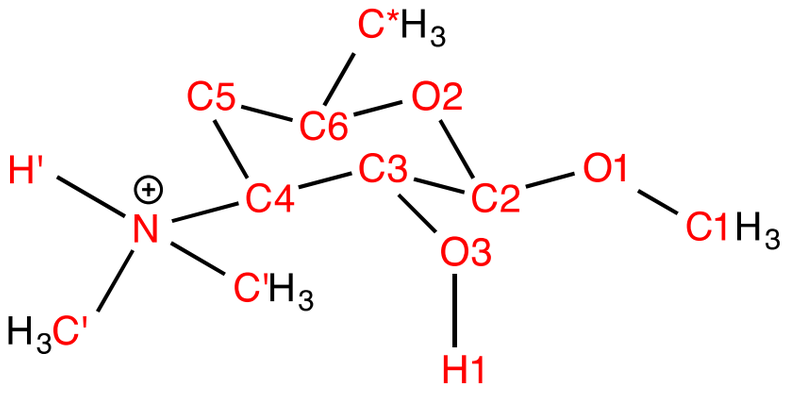




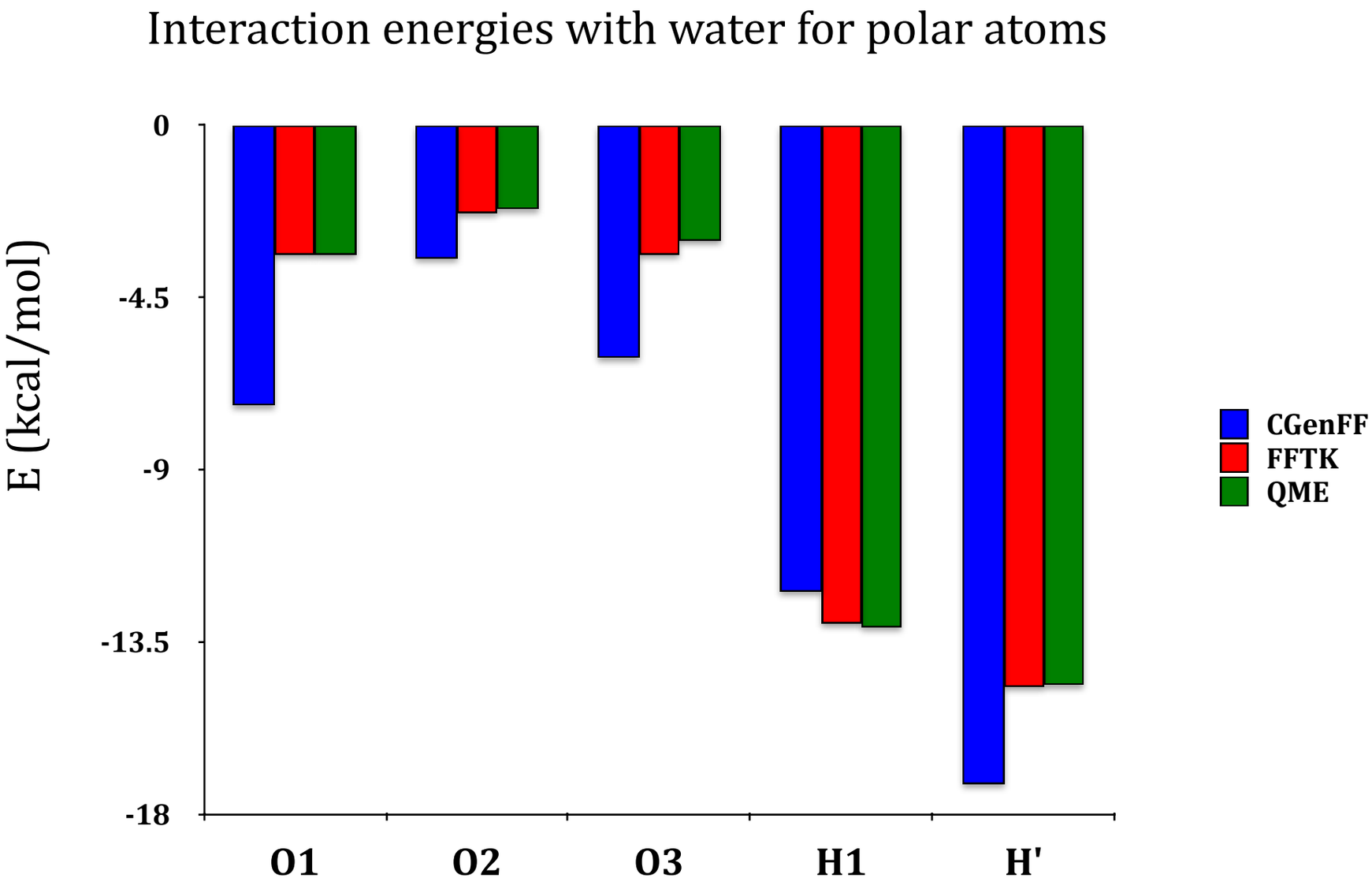
FFTK can be used to get better agreement with QM data

Accuracy of Fitting: Charges

Desosamine
charges fitted for red atoms



Interaction energies with water for polar atoms



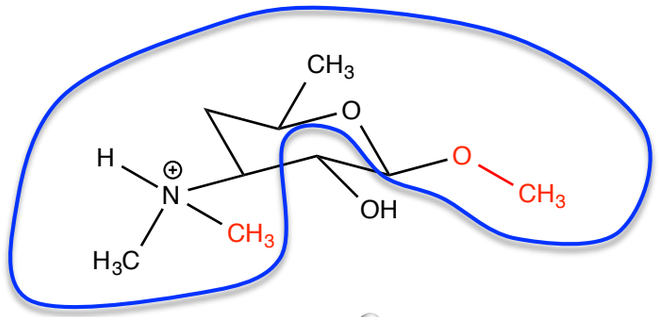
FFTK can be used to get better agreement with QM data

Fitting of Partial Charges

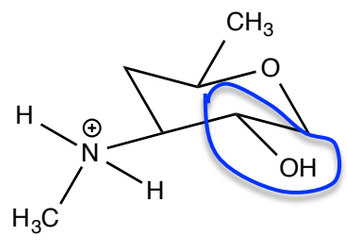
Optimization of water interactions can be difficult for sterically hindered groups

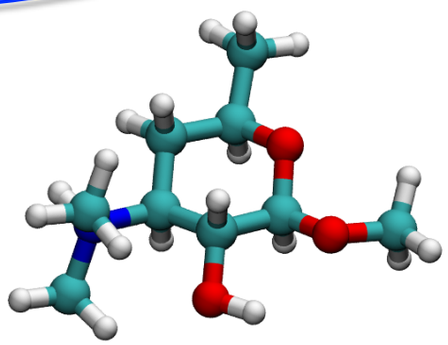
Our self consistent iterative charge optimization scheme

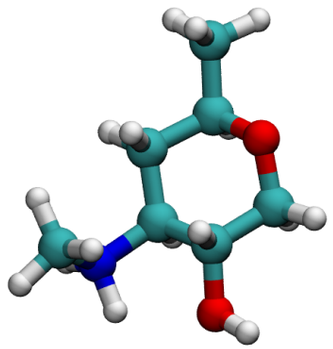
Desosamine



Simplified Desosamine





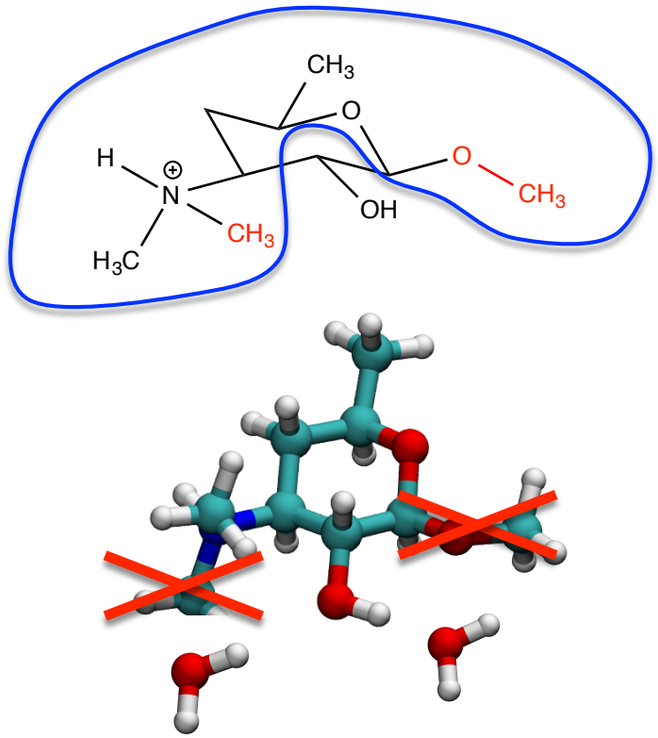


Fitting of Partial Charges

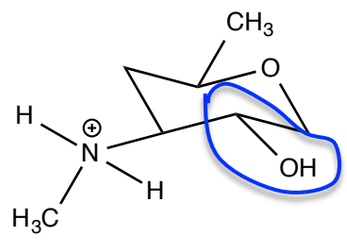
Optimization of water interactions can be difficult for sterically hindered groups

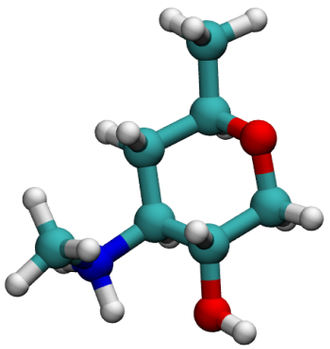
Our self consistent iterative charge optimization scheme

Desosamine



Simplified Desosamine



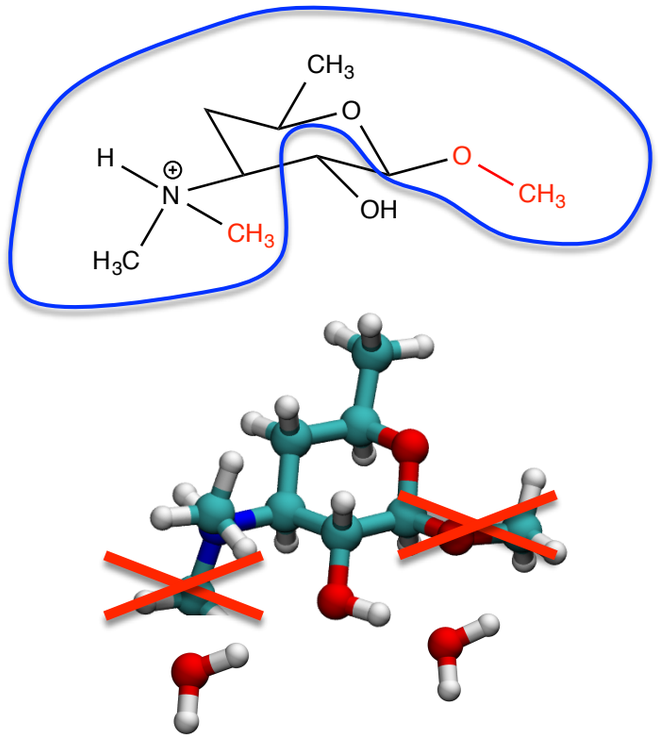


Fitting of Partial Charges

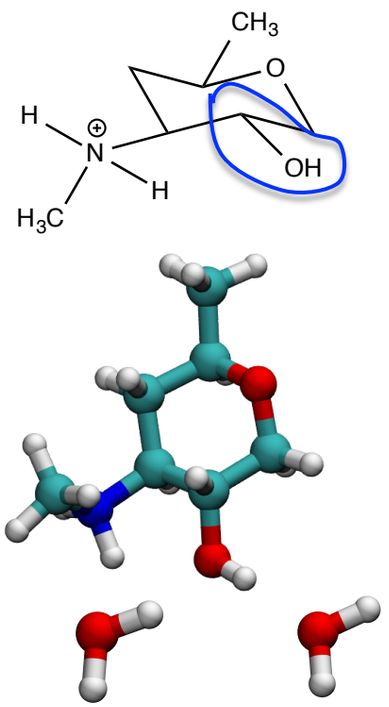
Optimization of water interactions can be difficult for sterically hindered groups

Our self consistent iterative charge optimization scheme

Desosamine



Simplified Desosamine



An Alternative to Water Interactions for Charges

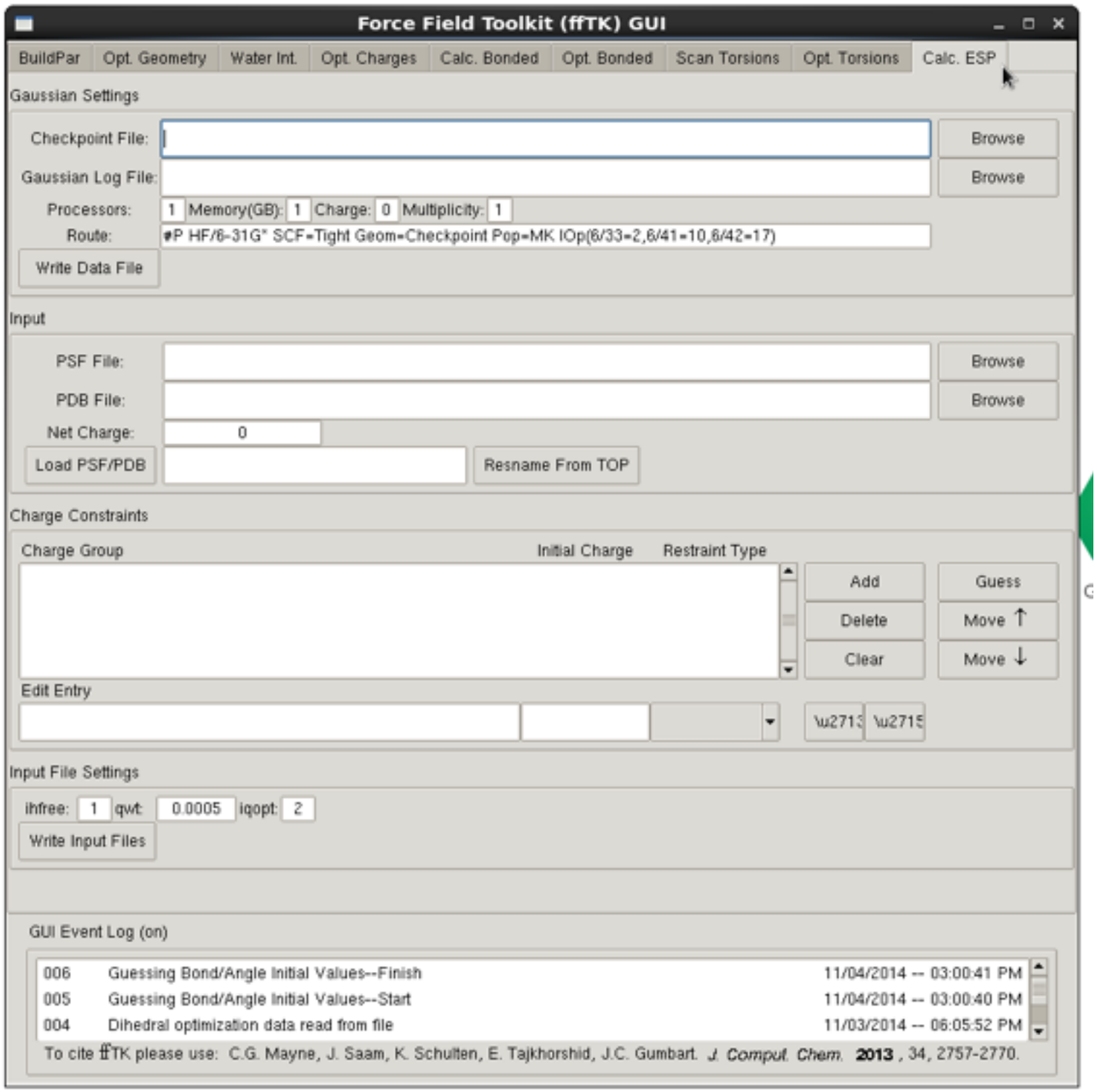
RESP fitting of charges
Based electrostatic potential

Much easier to implement
Tests on compatibility with CHARMM in progress

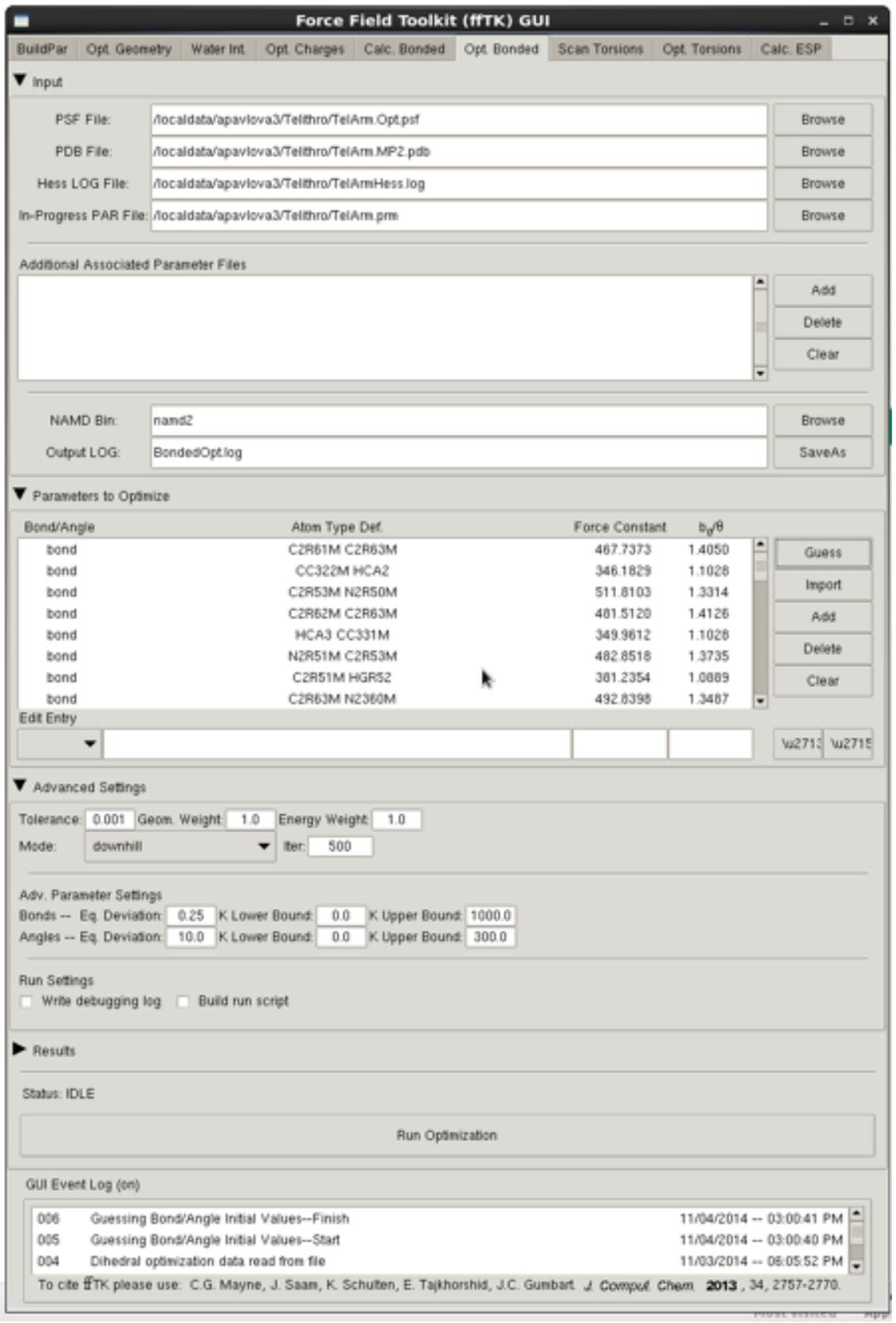
An Alternative to Water Interactions for Charges

RESP fitting of charges
Based electrostatic potential

Much easier to implement
Tests on compatibility with CHARMM in progress

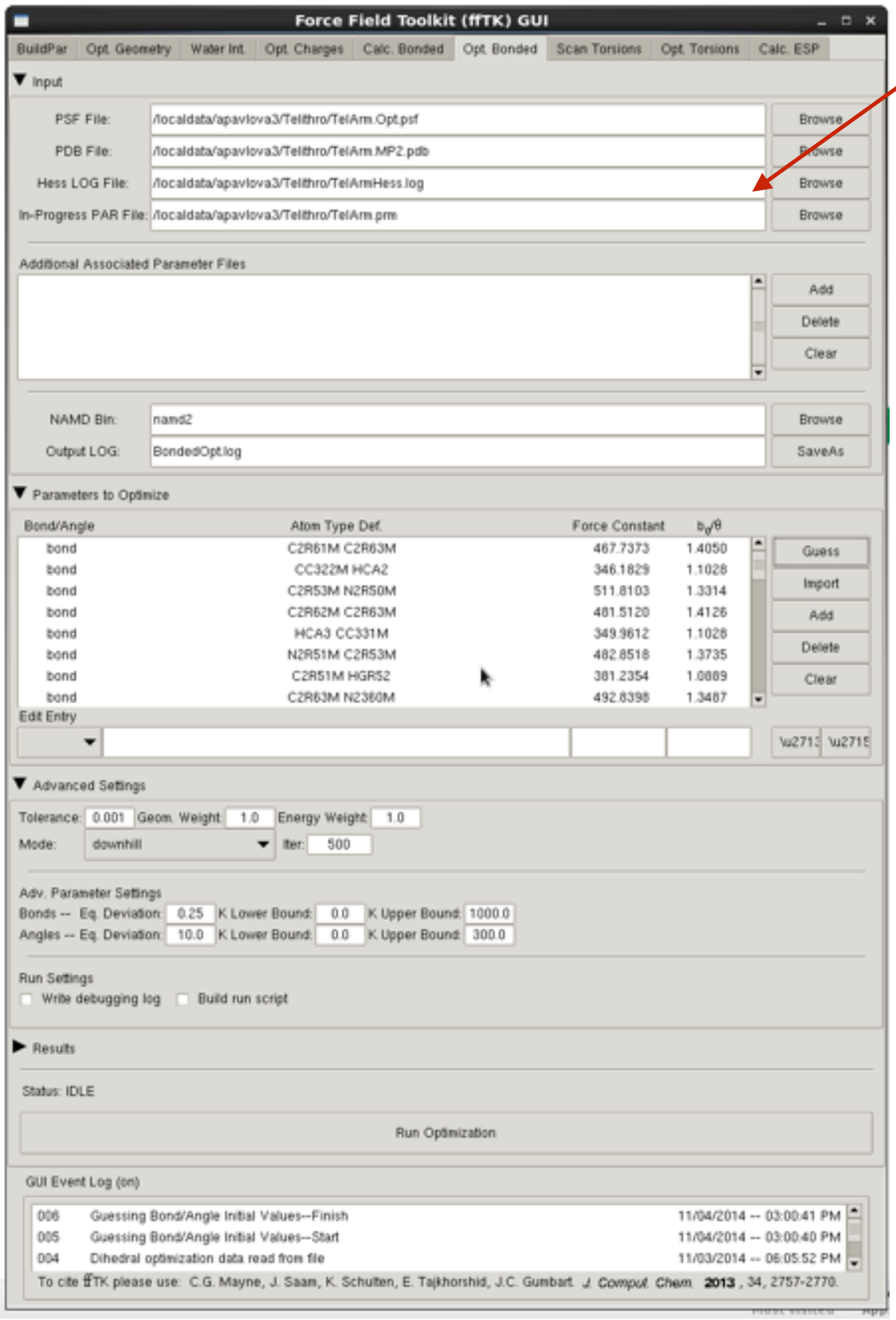


Fitting of Bonds and Angles



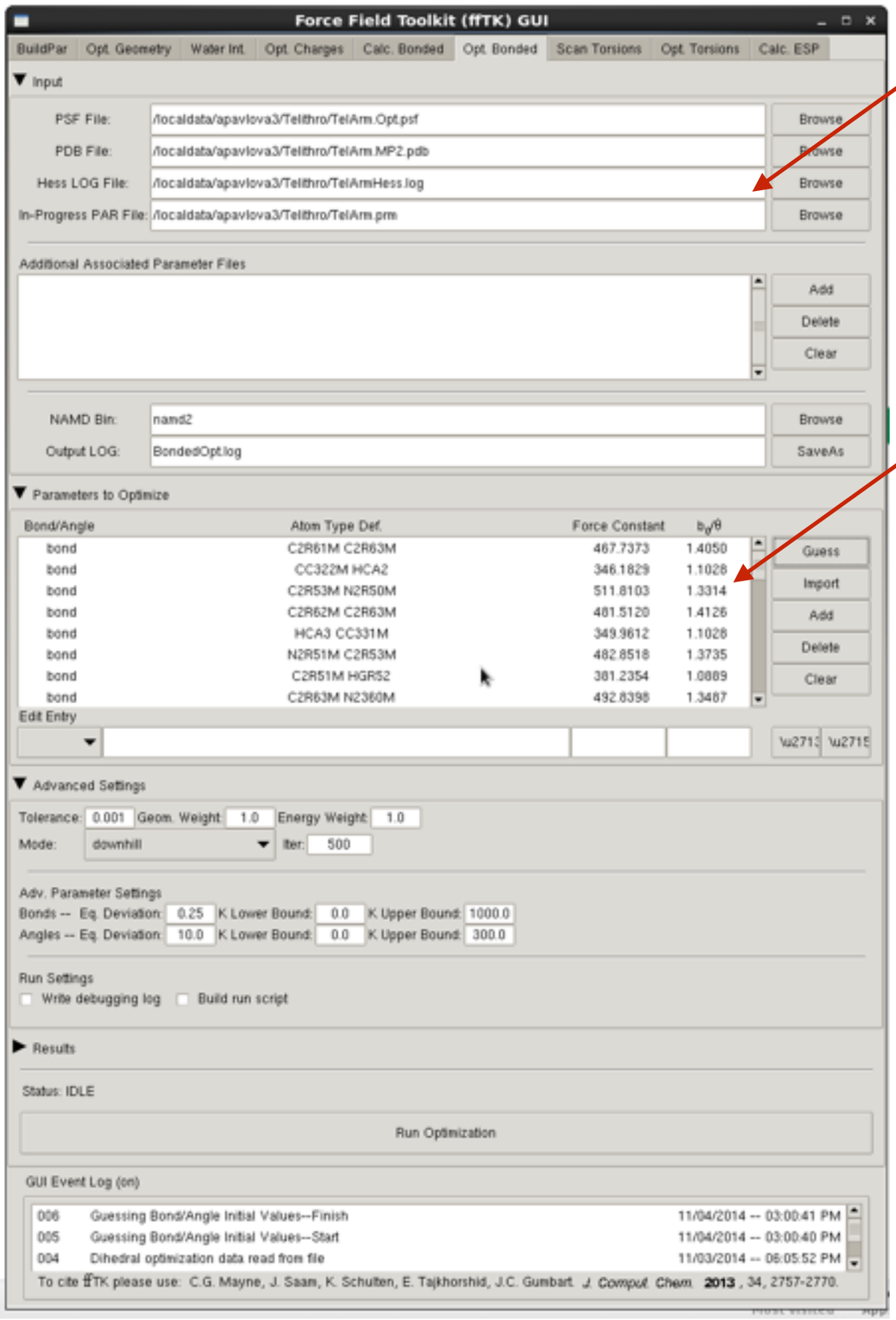
Fitting of Bonds and Angles

input Hessian file



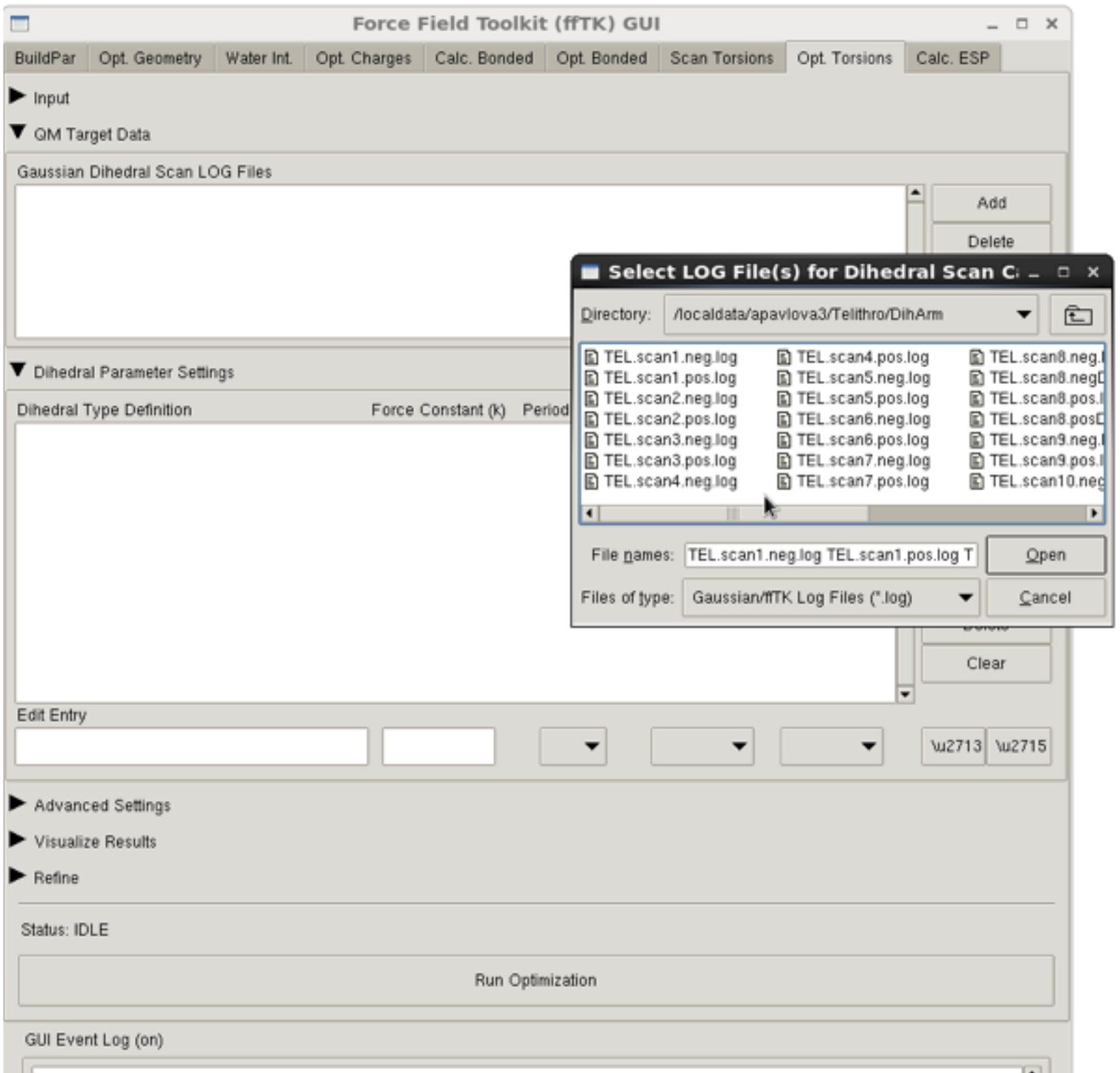
Fitting of Bonds and Angles

input Hessian file



Obtained force constants and values

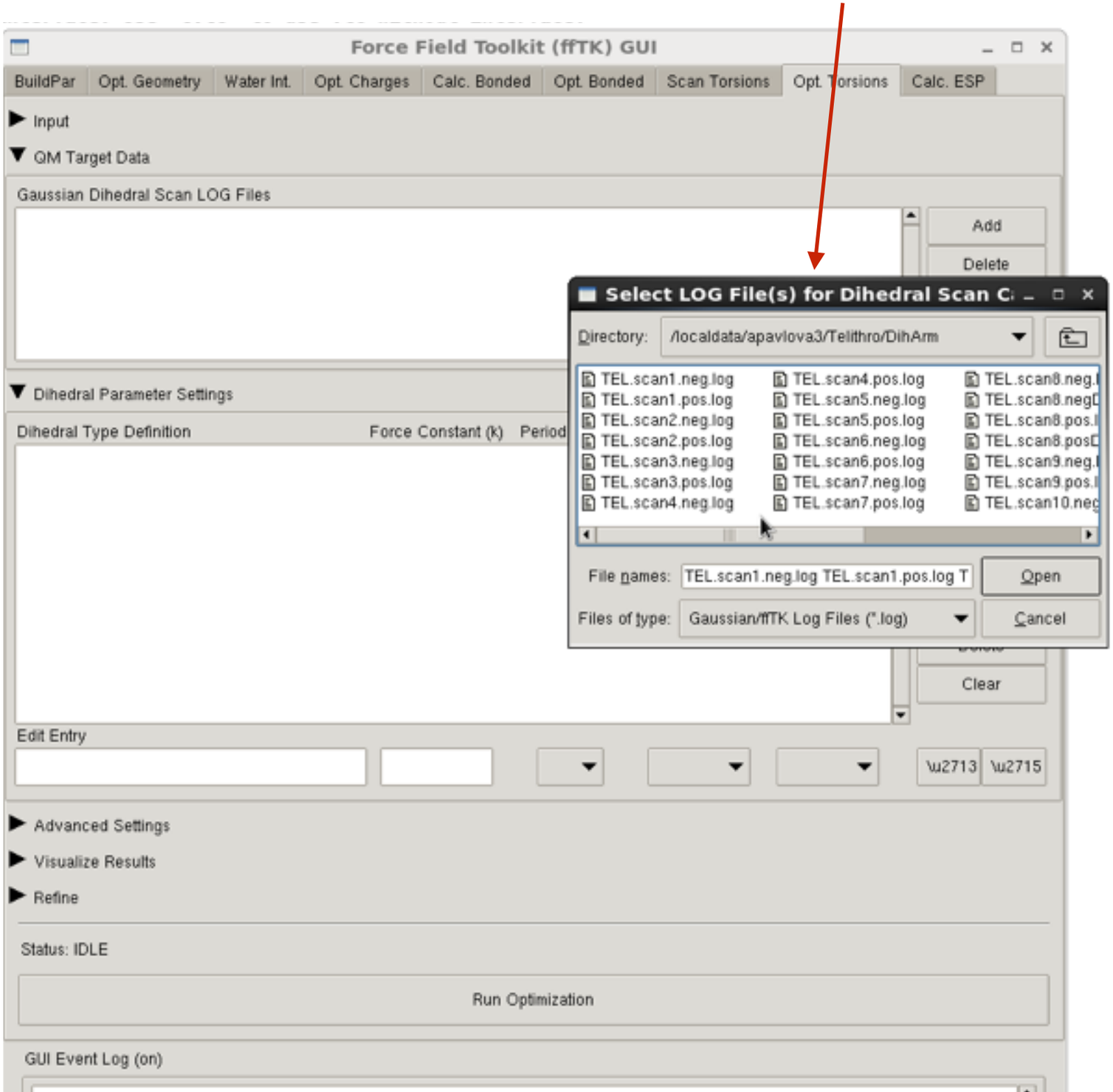
Fitting of Dihedrals



$$K_\phi(1 + \cos(n\phi - \delta))$$

Fitting of Dihedrals

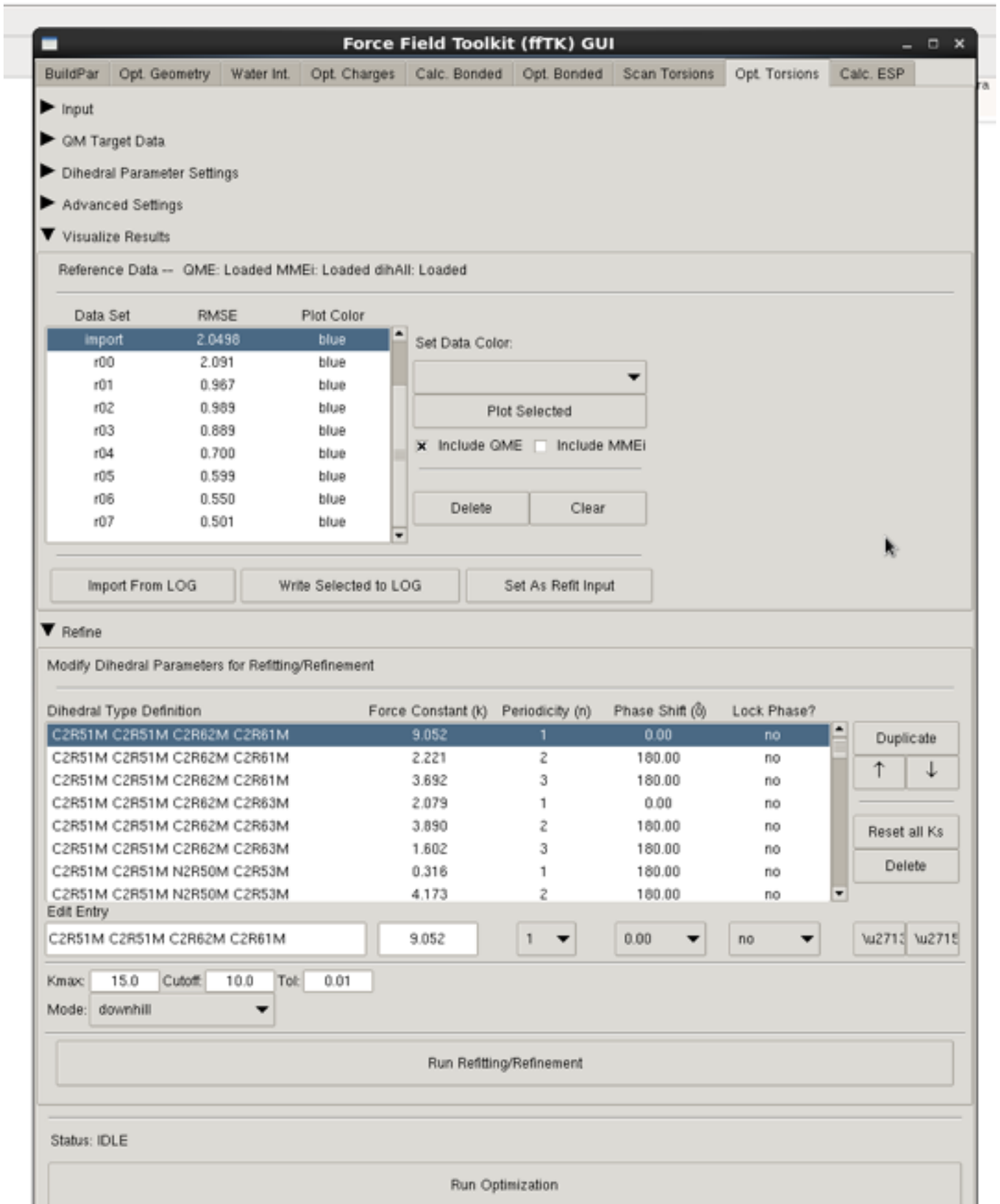
Feed in the dihedral scans



$$K_\phi(1+\cos(n\phi-\delta))$$

Fitting of Dihedrals

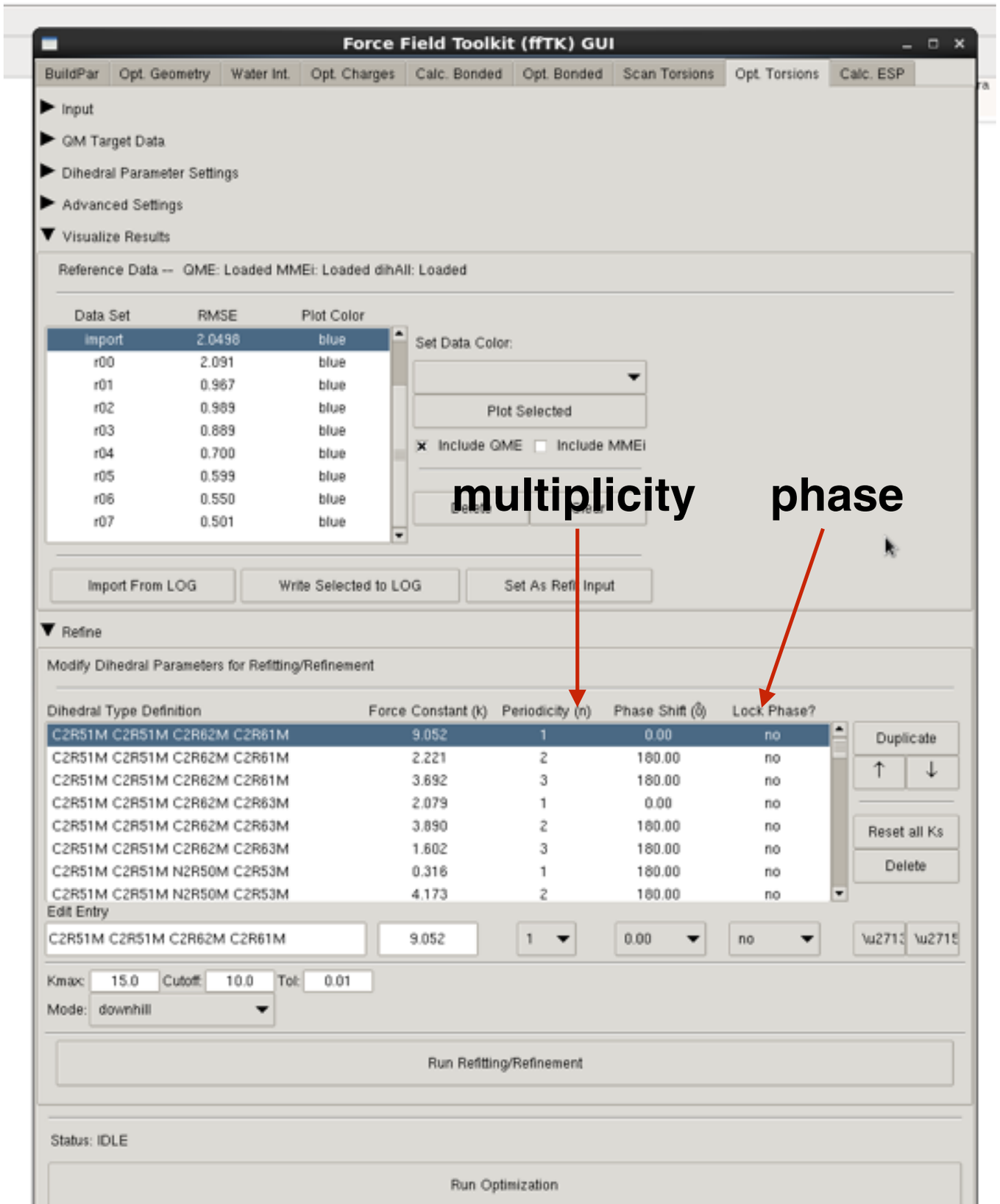
Select multiplicity and phase settings



$$K_\phi(1+\cos(n\phi-\delta))$$

Fitting of Dihedrals

Select multiplicity and phase settings



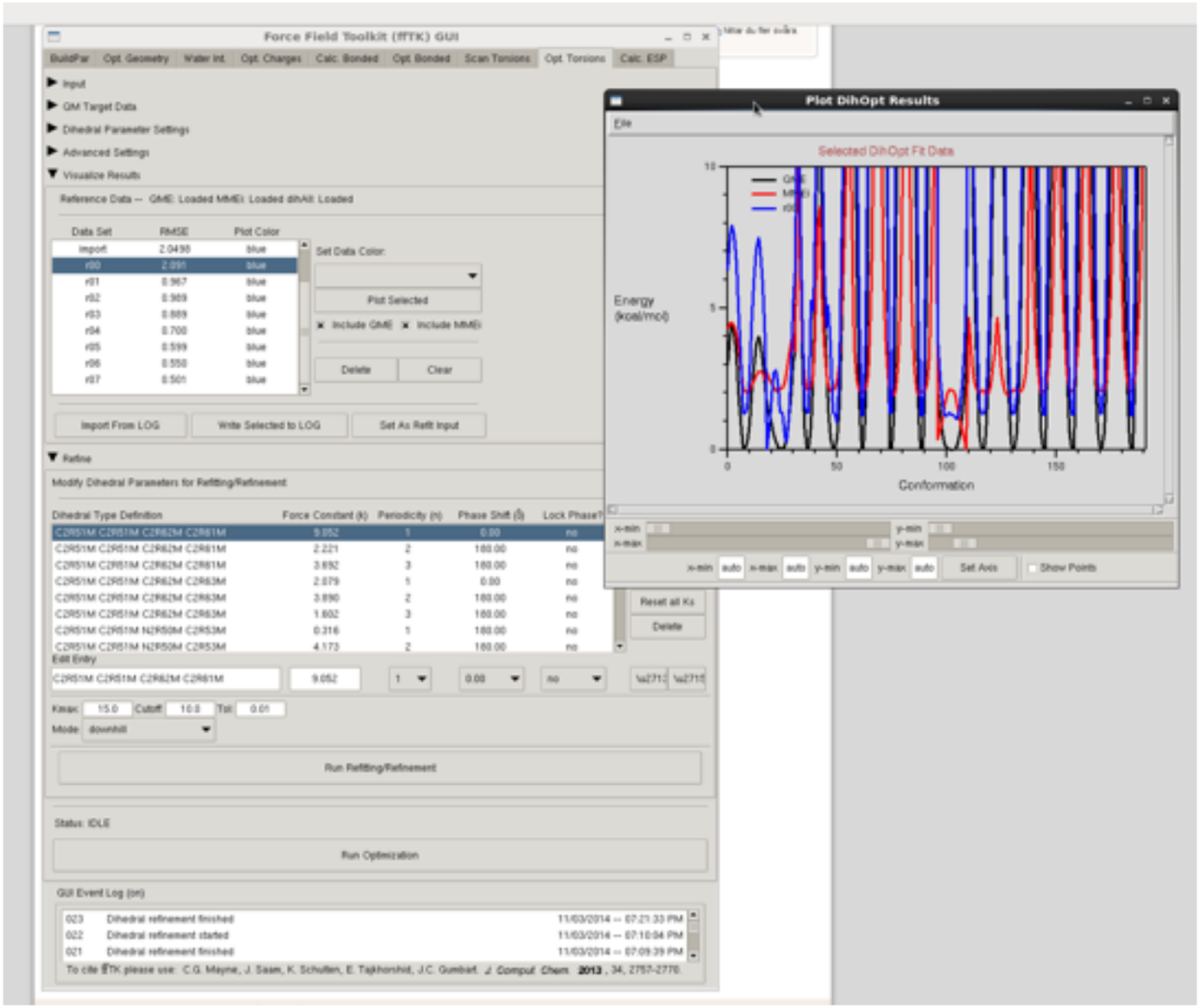
multiplicity phase

$$K_\phi(1+\cos(n\phi-\delta))$$

Fitting of Dihedrals

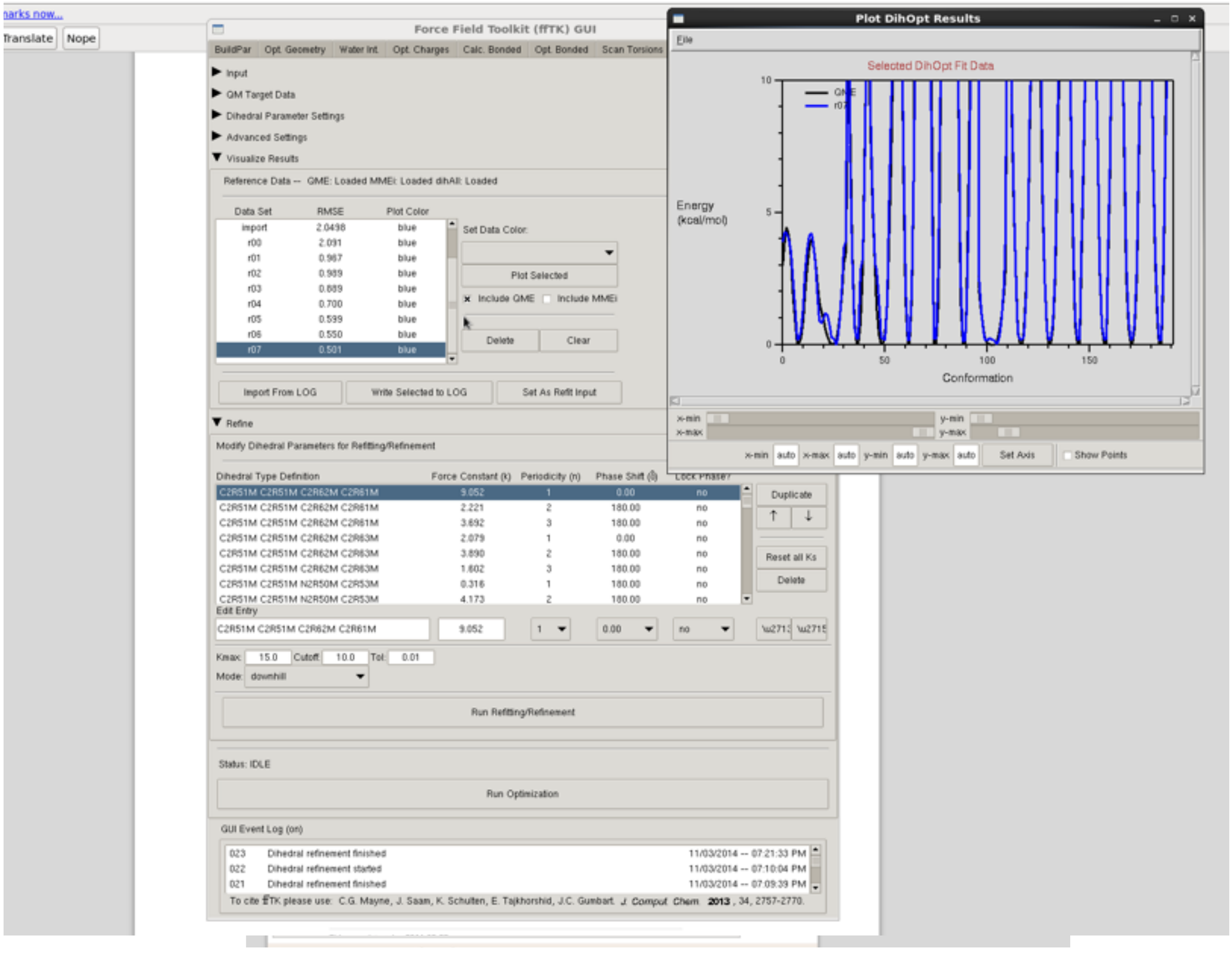
First fit is not great
but better than the starting fit

QM PES: black
First fit: blue
Starting fit: red



Fitting of Dihedrals

After 7 iterations the fit is much better



Two Approaches to Fitting the Dihedrals

Several multiplicities and free phase

Pro: very good fit of QM PES
Cons: possible incorrect behavior multiple sets of force constants

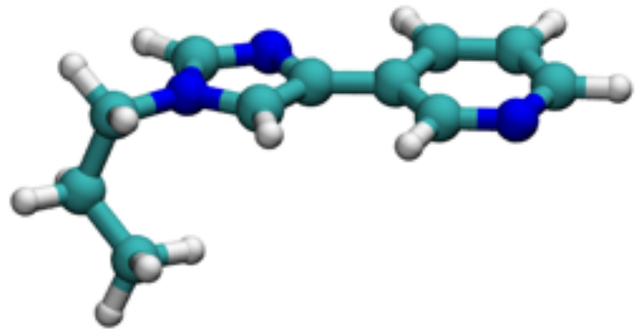
One multiplicity and locked phase

Pros: limits incorrect behavior, sets of force constants
Con: fit to QM PES not always possible

In practice a tradeoff is needed!

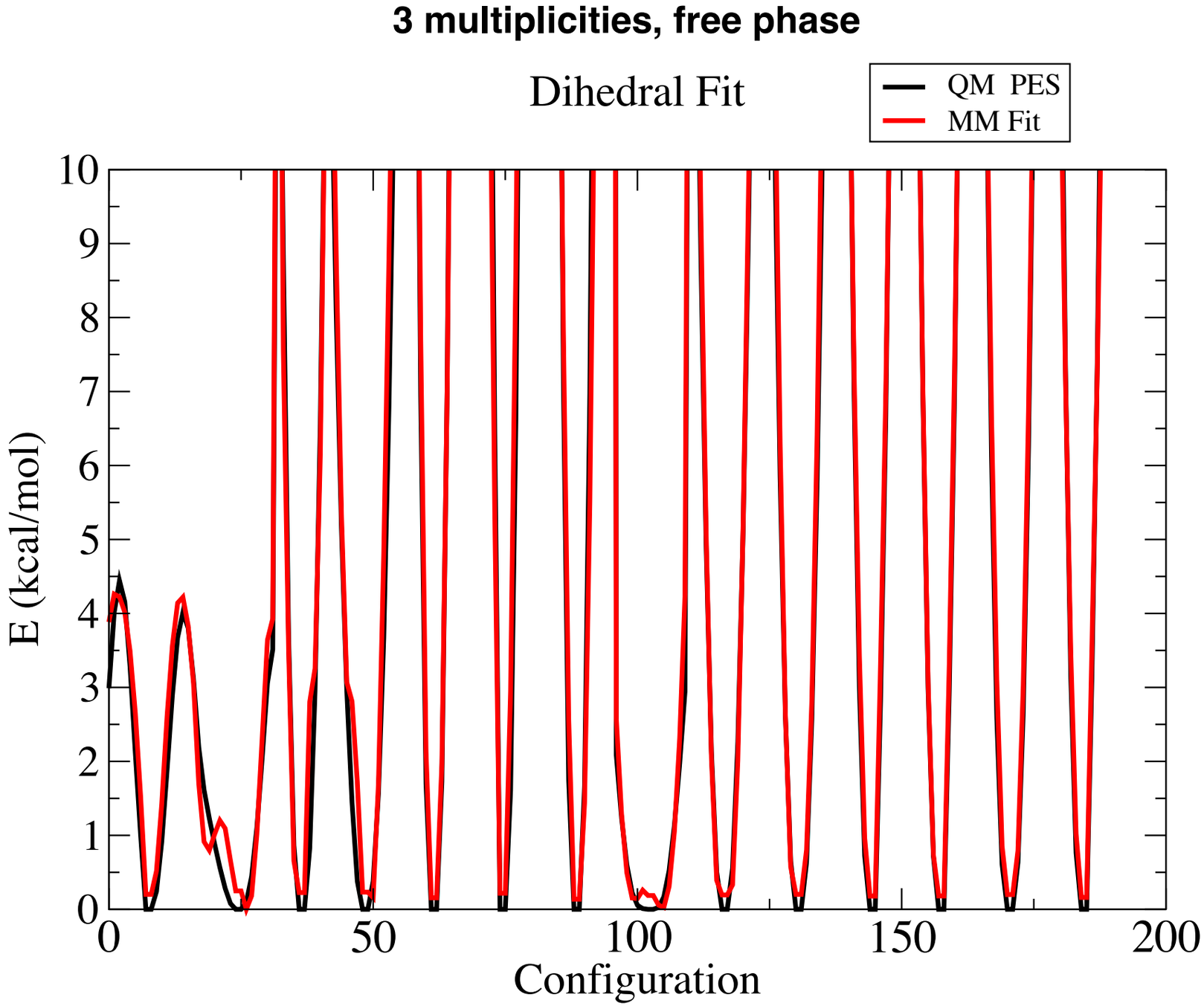
Example of Overfitting

imidazole-pyrridine moiety of antibiotic telithromycin



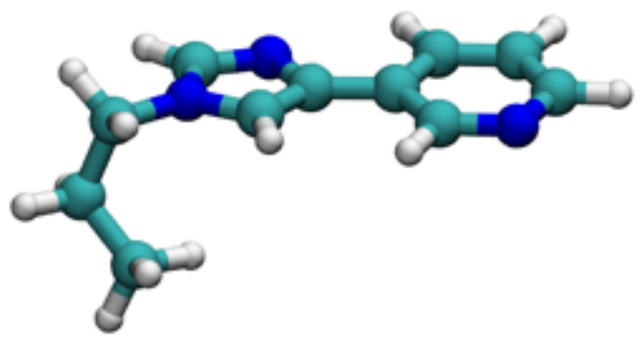
Example of Overfitting

imidazole-pyrridine moiety of antibiotic telithromycin



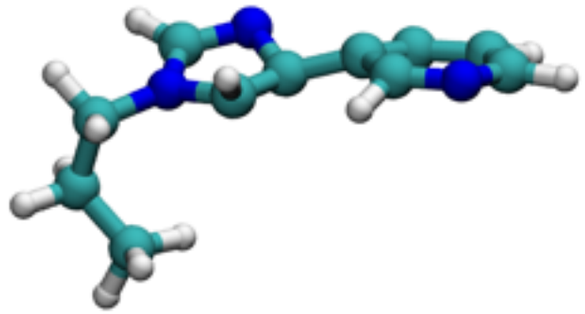
Example of Overfitting

imidazole-pyrridine moiety of antibiotic telithromycin



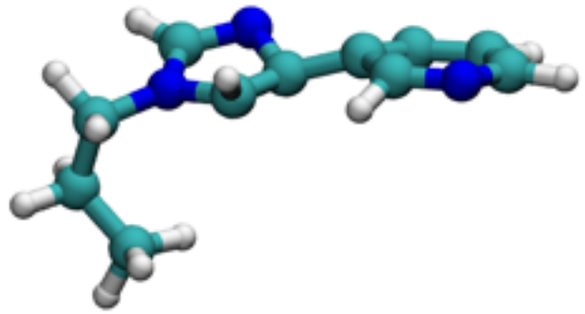
Example of Overfitting

imidazole-pyrridine moiety of antibiotic telithromycin



Example of Overfitting

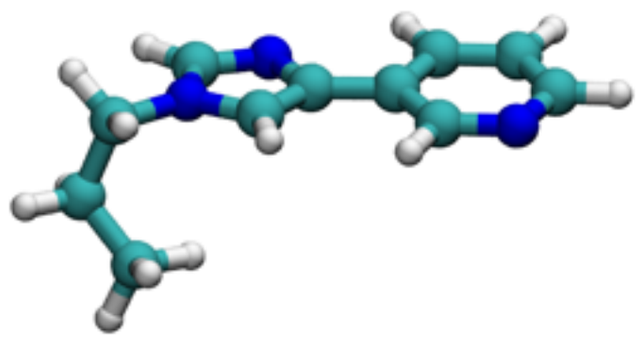
imidazole-pyrridine moiety of antibiotic telithromycin



Using too many dihedral multiplicities
can leads to distortion of a planar molecule!

Example of Overfitting

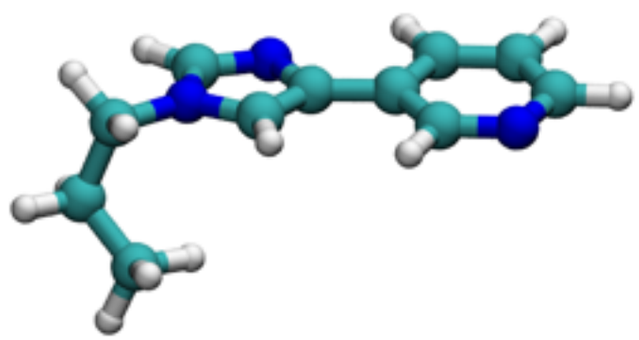
imidazole-pyrridine moiety of antibiotic telithromycin



Example of Overfitting

imidazole-pyrridine moiety of antibiotic telithromycin

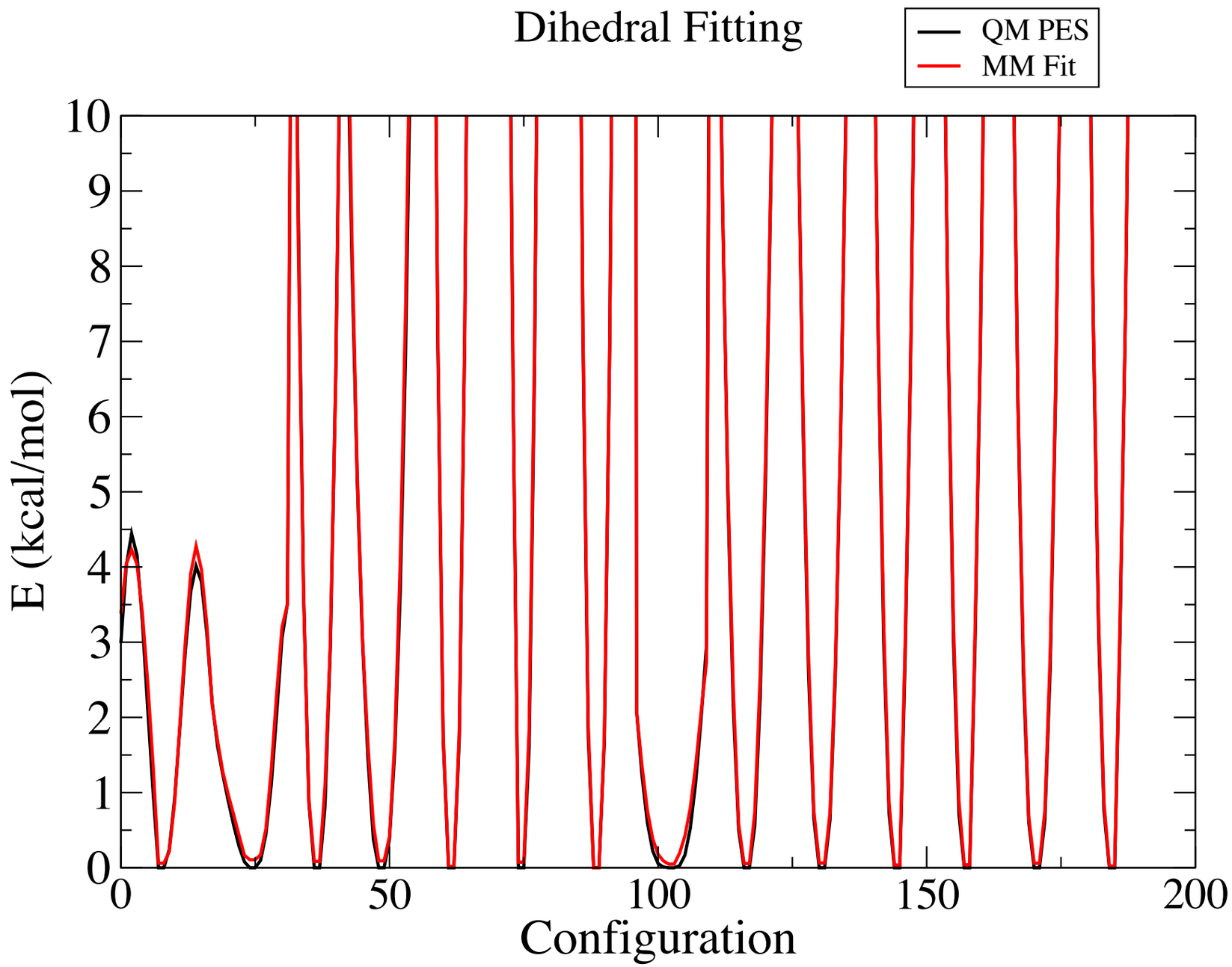
planar dihedrals have multiplicity 2 and phase 180



Example of Overfitting

imidazole-pyrridine moiety of antibiotic telithromycin

planar dihedrals have multiplicity 2 and phase 180

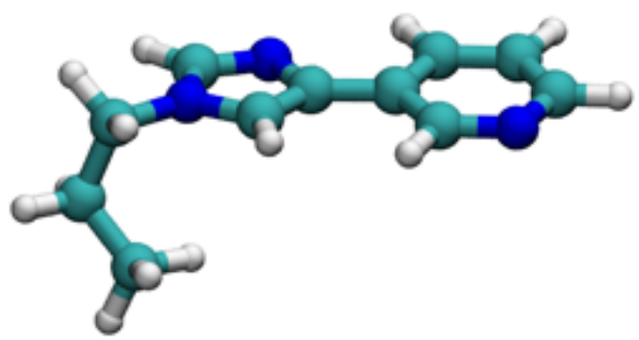


The fit looks even better!

Example of Overfitting

imidazole-pyrridine moiety of antibiotic telithromycin

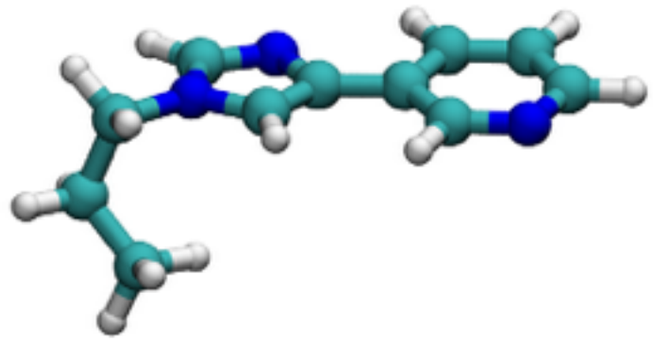
planar dihedrals have multiplicity 2 and phase 180



Example of Overfitting

imidazole-pyrridine moiety of antibiotic telithromycin

planar dihedrals have multiplicity 2 and phase 180

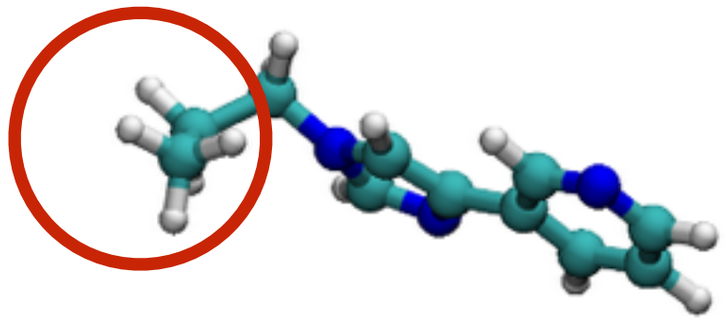


Planarity is maintained!

Example of Overfitting

imidazole-pyrridine moiety of antibiotic telithromycin

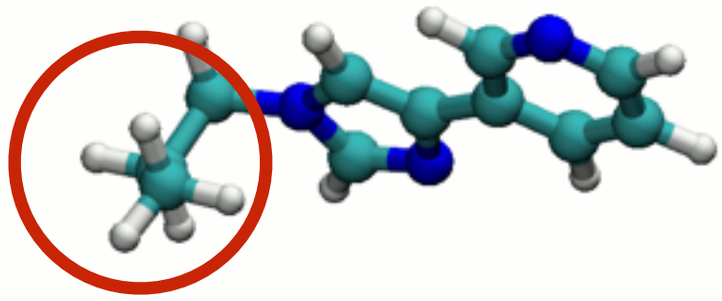
Problems persist
eclipsed conformation of the alkane part



Example of Overfitting

imidazole-pyrridine moiety of antibiotic telithromycin

Problems persist
eclipsed conformation of the alkane part



Restraining phase of CH dihedrals to 0 prevents eclipsed confirmations

How to know what terms to include?

Back to CGenFF main page

CGenFF Frequently Asked Questions

Contents

How can dihedral contributions cancel out?

Why does a wind turbine or a propeller with 2 or 3 blades not wobble?

There are a number of common cases in which dihedral contributions cancel out:

- If a dihedral involves an sp^2 center as one of the inner atoms, and both substituents on this center are identical or have identical dihedrals, all terms with an odd multiplicity will cancel out.
- If a dihedral involves an sp^3 center as one of the inner atoms, and all three substituents on this center are identical or have identical dihedrals, all terms with a multiplicity that is not a multiple of 3 (ie. all terms other than 3-fold and 6-fold) will cancel out.
- Consequence 1: if both of the above conditions are satisfied (example: methyl rotation in toluele), only the 6-fold term will *not* cancel out.
- Consequence 2: if the substituents are not identical but some of their dihedral terms have the same phase and multiplicity, there may be partial cancellation and some of the terms can be omitted. (In practice, one would typically choose to omit H-X-X-H terms because these are more prone to transferability problems.)

All these observations follow from the expression for the dihedral energy:

http://mackerell.umaryland.edu/~kenno/cgenff/faq.html

Conclusions

Find Missing Parameters

Geometry Optimization (QM)

Water Interaction En. (QM)

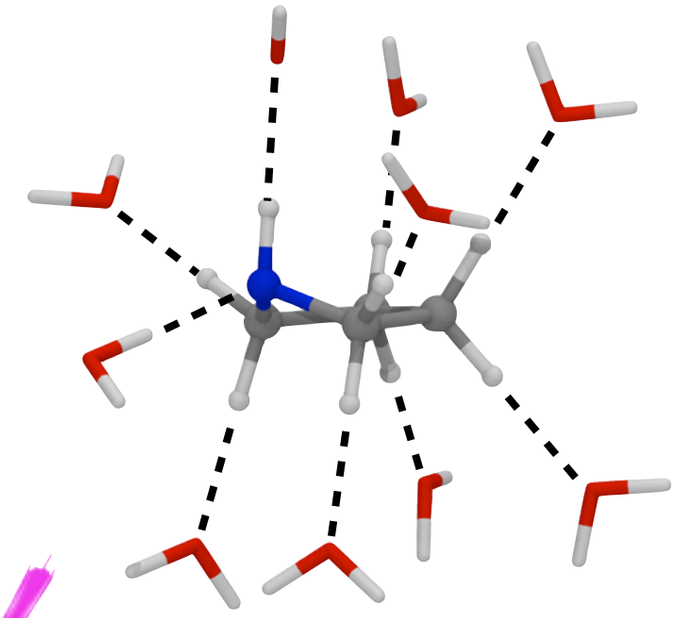
Charge Optimization

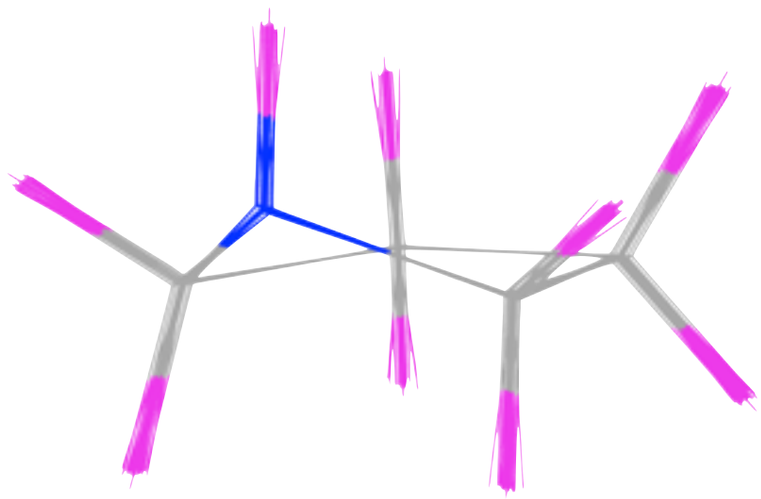
Hessian Calculation (QM)

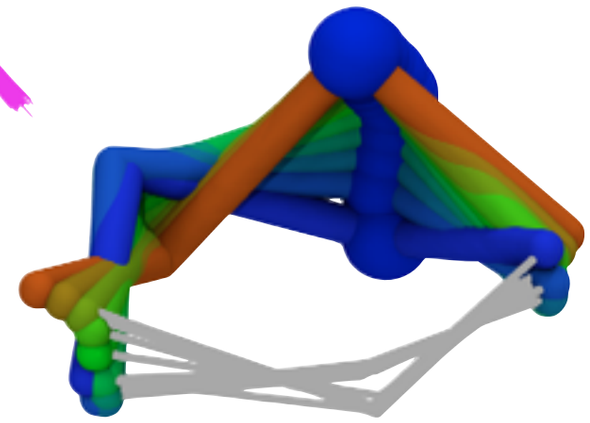
Bond & Angle Optimization

Torsion Scan (QM)

Torsion Optimization







ffTK:

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

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

Future Directions

Investigate forcefield performance

Support additional force fields (e.g. AMBER)

Improve optimization speed

Expand analysis tools

Support multiple QM packages

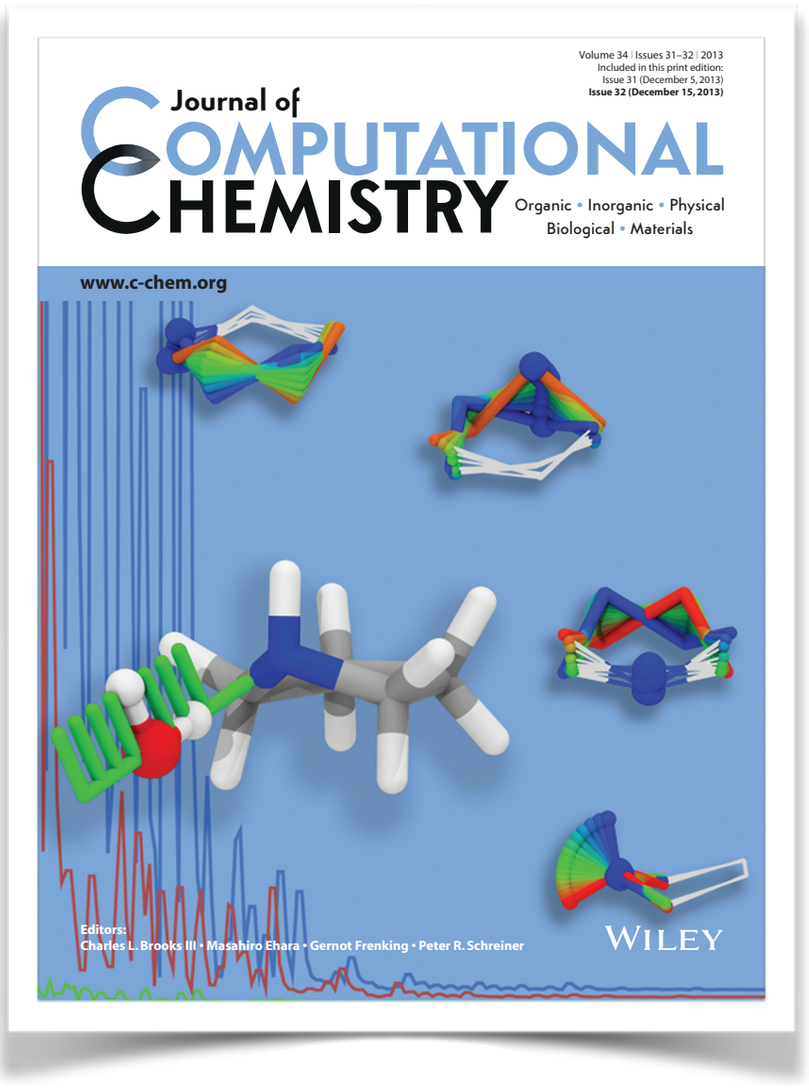
www.ks.uiuc.edu/Research/vmd/plugins/fftk

ffTK

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

ffTK is available as a VMD Plugin (1.9.1 or newer)

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



Questions?