NAME

OpenFECalculateRelativeBindingFreeEnergySepTop.py - Calculate relative binding free energy

SYNOPSIS

OpenFECalculateRelativeBindingFreeEnergySepTop.py [--executeDAGParams <Name,Value,..>] [--loggingLevel <Info, Warning or Error>] [--missingChargeMode <Calculate or Stop>] [--moleculePairs <MolName1,MolName2,..>] [--outfilePrefix <text>] [--overwrite] [--rbfeParams <Name,Value,...>] [--resultFileParams <Name,Value,..>] [--smallMolFileParams <Name,Value,...>] [--solventParams <Name,Value,...>] [-w <dir>] -i <infile> -s <smallmolfile> -o <outifiledir>

OpenFECalculateRelativeBindingFreeEnergySepTop.py -l | --list

OpenFECalculateRelativeBindingFreeEnergySepTop.py -h | --help | -e | --examples

DESCRIPTION

Calculate Relative Binding Free Energy (RBFE) for a pair of molecules in a small molecule input file using a Separated Topologies (SepTop) [Ref 183] approach.

You may calculate RBFEs for any arbitrary pairs of molecules. The SepTop protocol calculates RBFE by performing two ABFE calculations simultaneously in opposite directions. No atom mapping is required between pairs of molecules to create hybrid topology. Consequently, the SepTop based RBFEs calculations may be used for scaffold hopping.

A brief description of the SepTop protocol, taken from OpenFE documenation is provided below:

```
The relative binding free energy is obtained through a thermodynamic
cycle. The molecules are transformed into each other both in solvent,
giving ddG (solvent), and in the complex, giving ddG (complex), which
allows the calculation of the relative binding free energy, ddG. Each
molecule is represented with its own set of coordinates, meaning that
the interactions of all atoms of one molecule are turned off while
simultaneously turning on the interactions of all atoms of the other
molecule. Therefore, restraints are required: Molecules are restrained
to the protein in the complex states using orientational (Boresch-style)
restraints; in the solvent states molecules are restrained to remain
apart from each other using a single harmonic distance restraint between
the molecules.


The coulombic interactions of the molecule are fully turned off
(annihilated), while the Lennard-Jones interactions are decoupled,
meaning the intermolecular interactions are turned off, while keeping
the intramolecular Lennard-Jones interactions.
```

The input file must contain a macromolecule already prepared for simulation. The preparation of the macromolecule for a simulation generally involves the following tasks: identification and replacement of non-standard residues; addition of missing residues; addition of missing heavy atoms; addition of missing hydrogens.

In addition, the small molecule input file must contain molecules already prepared for simulation. It must contain appropriate 3D coordinates relative to the macromolecule along with no missing hydrogens.

```
Protocol repeats, 3


Time step size: 4.0 femtosecond


Complex equilibration phase:


Max minimization steps: 5,000
NVT equilibration length: 0.1 nanosecond
NPT equilibration length: 0.1 nanosecond
NPT length: 2.0 nanosecond


Complex production phase:


Max minimization steps: 5,000
NPT equilibration length: 1.0 nanosecond
```

```
    NPT length: 10.0 nanosecond

    Solvent equilibration phase:

    Max minimization steps: 5,000
    NVT equilibration length: 0.1 nanosecond
    NPT equilibration length: 0.1 nanosecond
    NPT length: 2.0 nanosecond

    Solvent production phase:

    Max minimization steps: 5,000
    NPT equilibration length: 1.0 nanosecond
    NPT length: nanosecond 10.0 nanosecond
```

Each complex and solvent simulation, by default, may run for 13.2 and 13.2 nanosecond respectively, for a total of 26.4 nanoseconds. The total MD simulation time for correspond to 79.2 nanosecond to repeat the protocol 3 times for the complex and solvent simulations.

The supported macromolecule input file formats are: PDB (.pdb) and CIF (.cif)

The supported small molecule input file format are : SD (.sdf, .sd)

Possible outfile prefix:

```
    <OutfilePrefix> or <SmallMolFileRoot>
```

Possible output directories:

```
    <OutfileDir>

    <OutfileDir>/Transformations
    <OutfileDir>/Results
```

Possible output files and directories under <OutfileDir>:

```
    <OutfilePrefix>_RBFE_Results.<csv or tsv>

    ... ... ...

    Transformations/<MolAName>_To_<MolBName>_Complex_Solvent.json
    ... ... ...

    Results/<MolAName>_To_<MolBName>_Complex_Solvent_Results.json
    Results/shared_SepTopComplexRunUnit-*/
    Results/shared_SepTopComplexSetupUnit-*/
    Results/shared_SepTopSolventRunUnit-*/
    Results/shared_SepTopSolventSetupUnit-*/
    ... ... ...
```

OPTIONS

-e, --examples

Print examples.

--executeDAGParams <Name,Value,..> [default: auto]

A comma delimited list of parameter name and value pairs for executing protocol DAGs (Directed Acyclic Graph) to run RBFE calculations.

The supported parameter names along with their default values are are shown below:

```
    keepShared, yes  [ Possible values: yes or no ]
    keepScratch, no  [ Possible values: yes or no ]
    nRetries, 2  [ Possible values: >= 0. A value of 0 implies only
        1 try. ]
```

A brief description of parameters is provided below:

```
keepShared: Keep shared directories after the execution of DAG.
keepScratch: Keep scratch directories after the execution of DAG.
nRetries: Number of times to attempt the execution.
```

-h, --help

Print this help message.

-i, --infile <infile>

Input file name containing a macromolecule.

-l, --list

List default RBFE protocol settings provided by OpenFE module RelativeHybridTopologyProtocol.

--loggingLevel <Info, Warning or Error> [default: Error]

Logging level to configure the 'root logger' via logging.basicConfig() function. The default logging level is changed from 'logging.INFO' to 'logging.ERROR'. Otherwise, OpenFE and its associated modules may generate a lot of informational messages.

--missingChargeMode <Calculate or Stop> [default: Stop]

Calculate missing partial charges for molecules before running RBFE calculations or terminate the execution of the script. The missing partial charges will be automatically calculated by OpenFE module RelativeHybridTopologyProtocol during the calculation of RBFE. You may control the calculation of partial charges by specifying values for partialCharge* parameters using '--rbfeParams' option.

-m, --moleculePairs <MolName1,MolName2,..> [default: auto]

A comma delimited list of molecule name pairs for calculating RBFEs. Default: the names of the first and second molecule in small molecule input file.

-o, --outfileDir <outfiledir>

Output directory.

--outfilePrefix <text> [default: auto]

Prefix for generating output files under output directory.

--overwrite

Overwrite existing files.

--resultFileParams <Name,Value,..> [default: auto]

A comma delimited list of parameter name and value pairs for writing calculated RBFEs values to a results file.

The supported parameter names along with their default values are are shown below:

```
precision, 4  [ Possible values: > 0 ]
delimiter, comma  [ Possible values: comma or tab ]
```

-r, --rbfeParams <Name,Value,...> [default: auto]

A comma delimited list of parameter name and value pairs for RBFE separated topology protocol settings employed during the calculation of RBFEs.

The default values are automatically updated to match settings provided by OpenFE module SepTopProtocol.

You must specify valid OpenFE values for these parameters. An extensive validation is not performed.

The supported parameter names along with their default values are are shown below:

```
protocolRepeats, 3

Complex equil output settings:

complexEquilOutputCheckpointInterval, 1  [ Units: nanosecond ]
complexEquilOutputCheckpointStorageFilename, checkpoint.chk
complexEquilOutputEquilNPTStructure, equil_npt.pdb
complexEquilOutputEquilNVTstructure, None
complexEquilOutputForcefieldCache, db.json
complexEquilOutputLogOutput, equil_simulation.log
```

```
complexEquilOutputMinimizedStructure, minimized.pdb
complexEquilOutputIndices, all  [ Possible value: Any valid
    selection. ]
complexEquilOutputPreminimizedStructure, system.pdb
complexEquilOutputProductionTrajectoryFilename, production_equil.xtc
complexEquilOutputTrajectoryWriteInterval,  20.0  [ Units:
    picosecond ]

Complex equil simulation settings:

complexEquilSimulationEquilibrationLength, 0.1 [ Units: nanosecond ]
complexEquilSimulationEquilibrationLengthNVT, 0.1   [ Units:
    nanosecond ]
complexEquilSimulationMinimizationSteps, 5000
complexEquilSimulationProductionLength, 2.0  [ Units: nanosecond ]

Complex lambda settings:

complexLambdaElecA, 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.25 0.5 0.75
    1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0  [ Possible values: A space
    delimited list of values between 0.0 and 1.0 ]
complexLambdaElecB, 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 0.75 0.5 0.25
    0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 [ Possible values: A space
    delimited list of values between 0.0 and 1.0 ]
complexLambdaRestraintsA, 0.0 0.05 0.1 0.3 0.5 0.75 1.0 1.0 1.0
    1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 [ Possible values: A
    space delimited list of values between 0.0 and 1.0 ]
complexLambdaRestraintsB, 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0
    1.0 1.0 1.0 0.75 0.5 0.3 0.1 0.05 0.0 [ Possible values: A space
    delimited list of values between 0.0 and 1.0 ]
complexLambdaVdwA, 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
    0.143 0.286 0.429 0.572 0.715 0.857 1.0 [ Possible values: A
    delimited list of values between 0.0 and 1.0 ]
complexLambdaVdwB, 1.0 0.857 0.715 0.572 0.429 0.286 0.143 0.0
    0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 [ Possible values: A
    delimited list of values between 0.0 and 1.0 ]

Complex output settings:

complexOutputCheckpointInterval, 1.0  [ Units: nanosecond ]
complexOutputCheckpointStorageFilename, complex_checkpoint.nc
complexOutputForcefieldCache, db.json
complexOutputFilename, complex.nc
complexOutputIndices, not water  [ Possible value: Any valid
    selection. ]
complexOutputStructure, alchemical_system.pdb
complexOutputPositionsWriteFrequency, 100.0  [ Units: picosecond ]
complexOutputVelocitiesWriteFrequency, None  [ Possible
    values: > 0; Units: picosecond ]

Complex restraint settings:

complexRestraintKPhiA, 334.72  [ Units: kilojoule_per_mole/radian**2
    The default value is equivalent to 80 kcal/mol/radian**2 ]
complexRestraintKPhiB, 334.72  [ Units: kilojoule_per_mole/radian**2
    The default value is equivalent to 80 kcal/mol/radian**2 ]
complexRestraintKPhiC, 334.72  [ Units: kilojoule_per_mole/radian**2
    The default value is equivalent to 80 kcal/mol/radian**2 ]
complexRestraintKR, 4184.0  [ Units: kilojoule_per_mole/nanometer**2
     The default value is equivalent to 10 kcal/mol/angstrom**2
complexRestraintKThetaA, 334.72  [ Units:kilojoule_per_mole/radian**2
    The default value is equivalent to 80 kcal/mol/radian**2 ]
complexRestraintKThetaB, 334.72  [ Units:kilojoule_per_mole/radian**2
    The default value is equivalent to 80 kcal/mol/radian**2 ]
complexRestraintAnchorFindingStrategy, bonded  [ Possible values:
    multi-residue or bonded ]
complexRestraintDsspFilter, yes   [ Possible values: yes or no ]
complexRestraintHostMaxDistance, 1.5  [ Units: nanometer ]
complexRestraintHostMinDistance, 0.5  [ Units: nanometer ]
```

```
complexRestraintHostSelection, backbone   [ Possible value: Any valid
    selection. ]
complexRestraintRmsfCutoff, 0.1  [ Units: nanometer ]

Complex simulation settings:

complexSimulationEarlyTerminationTargetError, 0.0  [ Units:
    kilocalorie_per_mole ]
complexSimulationEquilibrationLength,  1.0  [ Units: nanosecond ]
complexSimulationMinimizationSteps, 5000
complexSimulationNReplicas, 19
complexSimulationProductionLength, 10.0  [ Units: nanosecond ]
complexSimulationRealTimeAnalysisInterval, 250.0  [ Units:
    picosecond ]
complexSimulationRealTimeAnalysisMinimumTime, 500.0  [ Units:
    picosecond ]
complexSimulationSamplerMethod, repex  [ Possible values: repex,
    sams, or independent ]
complexSimulationSamsFlatnessCriteria, logZ-flatness  [ Possible
    values: logZ-flatness, minimum-visits or histogram-flatness ]
complexSimulationSamsGamma0, 1.0
complexSimulationTimePerIteration, 2.5   [ Units: picosecond ]

Complex solvation settings:

complexSolvationBoxShape, dodecahedron  [  Possible values: cube,
    dodecahedron, or octahedron ]
complexSolvationBoxSize, None  [ Possible value: A triplet of space
    X Y Z values; Units: nanometer ]
complexSolvationSolventModel, tip3p  [ Possible values: tip3p, spce,
    tip4pew, or tip5p ]
complexSolvationSolventPadding, 1.0  [ Units: nanometer ]

Engine settings:

engineComputePlatform, CPU  [ Possible values: CPU, CUDA,
    OpenCL, or Reference ]
engineGpuDeviceIndex, None [ Possible values: 0, 0 1, etc. ]

Forcefield settings:

forcefieldConstraints, HBonds  [ Possible values: HBonds,
    AllBonds, or HAngles ]
forcefields, amber/ff14SB.xml amber/tip3p_standard.xml
    amber/tip3p_HFE_multivalent.xml amber/phosaa10.xml
    [ Possible values: A space delimited list of valid names. ]
forcefieldHydrogenMass, 3.0  [ Units: amu ]
forcefieldNonbondedCutoff, 0.9   [ Units: nanometer ]
forcefieldNonbondedMethod, PME  [ Possible values: PME or
    NoCutoff ]
forcefieldRigidWater, yes  [ Possible values: yes or no ]
forcefieldSmallMoleculeForcefield, openff-2.1.1  [ Possible
    value: A valid forcefield name. ]

Integrator settings:

integratorBarostatFrequency, 25.0 * timestep  [ The specified value
    is a multiple of integratorTimestep. ]
integratorConstraintTolerance, 1e-06
integratorLangevinCollisionRate, 1.0  [ Units: 1 / picosecond ]
integratorNRestartAttempts, 20
integratorReassignVelocities, no  [ Possible values: yes or no ]
integratorRemoveCom, no  [ Possible values: yes or no ]
integratorTimestep, 4.0 [ Units: femtosecond ]

Partial charge settings:

partialChargeNaglModel, None  [ Default: Production AM1BCC model for
    NAGL; Possible value: Any valid name. ]
```

```
partialChargeNumberOfConformers, None  [ Possible value: > 0 ]
partialChargeOffToolkitBackend, AmberTools  [ Possible values:
    AmberTools or RDKit ]
partialChargeMethod, AM1BCC  [ Possble values: AM1BCC, Espaloma,
    or NAGL ]

Solvent equil output settings:

solventEquilOutputCheckpointInterval, 1.0  [ Units: nanosecond ]
solventEquilOutputCheckpointStorageFilename, checkpoint.chk
solventEquilOutputEquilNPTStructure, equil_npt.pdb
solventEquilOutputEquilNVTstructure, None
solventEquilOutputForcefieldCache, db.json
solventEquilOutputLogOutput, equil_simulation.log
solventEquilOutputMinimizedStructure, minimized.pdb
solventEquilOutputIndices, all  [  Possible value: Any valid
    selection. ]
solventEquilOutputPreminimizedStructure, system.pdb
solventEquilOutputProductionTrajectoryFilename, equil_npt.xtc
solventEquilOutputTrajectoryWriteInterval, 20.0  [ Units:
    picosecond ]

Solvent_equil_simulation_settings:

solventEquilSimulationEquilibrationLength, 0.1 [ Units: nanosecond ]
solventEquilSimulationEquilibrationLengthNVT, 0.1  [ Units:
    nanosecond ]
solventEquilSimulationMinimizationSteps, 5000
solventEquilSimulationProductionLength, 2.0  [ Units: nanosecond ]

Solvent lambda settings:

solventLambdaElecA, 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.125
    0.25 0.375 0.5 0.625 0.75 0.875 1.0 1.0 1.0 1.0 1.0 1.0 1.0
    1.0 1.0 1.0 [ Possible values: A space delimited list of values
    between 0.0 and 1.0 ]
solventLambdaElecB, 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 0.875
    0.75 0.625 0.5 0.375 0.25 0.125 0.0 0.0 0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 [ Possible values: A space delimited list of values
    between 0.0 and 1.0 ]
solventLambdaRestraintsA, 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
    0.0 [ Possible values: A space delimited list of values between
    0.0 and 1.0 ]
solventLambdaRestraintsB, 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
    0.0 [ Possible values: A space delimited list of values between
    0.0 and 1.0 ]
solventLambdaVdwA, 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0 0.0 0.15 0.23 0.3 0.4 0.52 0.64 0.76 0.88
    1.0 [ Possible values: A space delimited list of values between
    0.0 and 1.0 ]
solventLambdaVdwB, 1.0 0.85 0.77 0.7 0.6 0.48 0.36 0.24 0.12 0.0
    0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
    0.0 [ Possible values: A space delimited list of values between
    0.0 and 1.0 ]

Solvent output settings:

solventOutputCheckpointInterval, 1.0  [ Units: nanosecond ]
solventOutputCheckpointStorageFilename, solvent_checkpoint.nc
solventOutputForcefieldCache, db.json
solventOutputFilename, solvent.nc
solventOutputIndices, not water  [ Possible value: Any valid
    selection. ]
solventOutputStructure, alchemical_system.pdb
solventOutputPositionsWriteFrequency, 100.0  [ Units: picosecond ]
solventOutputVelocitiesWriteFrequency, None  [ Possible
    values: > 0; Units: picosecond ]
```

```
    Solvent restraint settings:

    solventRestraintCentralAtomsOnly, No  [ Possible values: yes or no ]
    solventRestraintSpringConstant, 1000.0 [ Units: kilojoule_per_mole /
        nanometer ** 2. The default value is equivalent to 2.40
        kilocalorie_per_mole / angstromg ** 2 ]

    Solvent simulation settings:

    solventSimulationEarlyTerminationTargetError, 0.0  [ Units:
        kilocalorie_per_mole ]
    solventSimulationEquilibrationLength, 1.0  [ Units: nanosecond ]
    solventSimulationMinimizationSteps, 5000
    solventSimulationNReplicas, 27
    solventSimulationProductionLength, 10.0  [ Units: nanosecond ]
    solventSimulationRealTimeAnalysisInterval, 250.0  [ Unit: picosecond ]
    solventSimulationRealTimeAnalysisMinimumTime, 500.0  [ Units:
        picosecond ]
    solventSimulationSamplerMethod, repex  [ Possible values: repex,
        sams, or independent ]
    solventSimulationSamsFlatnessCriteria, logZ-flatness  [ Possible
        values: logZ-flatness, minimum-visits or histogram-flatness ]
    solventSimulationSamsGamma0, 1.0
    solventSimulationTimePerIteration, 2.5  [ Units: picosecond ]

    Solvent solvation settings:

    solventSolvationBoxShape, dodecahedron  [  Possible values: cube,
        dodecahedron, or octahedron ]
    solventSolvationBoxSize, None  [ Possible value: A triplet of space
        X Y Z values; Units: nanometer ]
    solventSolvationSolventModel, tip3p  [ Possible values: tip3p, spce,
        tip4pew, or tip5p ]
    solventSolvationSolventPadding, 1.5  [ Units: nanometer ]

    Thermo settings:

    thermoPh, None  [ Possible values: > 0 ]
    thermoPressure, 1.0  [ Units: bar ]
    thermoRedoxPotential, None  [ Possible values: A valid float.
        Units: millivolts (mV) ]
    thermoTemperature, 298.15  [ Units: kelvin ]
```

A brief description of parameters, taken from OpenFE documentation, is provided below:

```
    protocolRepeats: Number of completely independent repeats of the
        entire sampling process.

    Complex settings:

    Complex parameters for the system, including the solvent model and
    the solvent padding.

    Complex equil output settings:

    Parameters controlling simulation output during equilibration
    phase of complex transformation.

    complexEquilOutputCheckpointInterval: Frequency to write the
        checkpoint file.
    complexEquilOutputCheckpointStorageFilename: Checkpoint filename.
    complexEquilOutputEquilNPTStructure: NPT structure filename.
    complexEquilOutputEquilNVTstructure: NVT strucure filename.
    complexEquilOutputForcefieldCache:  Filename for caching small
        molecule residue templates.
    complexEquilOutputLogOutput: Simulation log filename.
    complexEquilOutputMinimizedStructure: Minimized structure filename.
    complexEquilOutputIndices: Selection string for selecting
        coordinates to write.
```

```
complexEquilOutputPremnimizedStructure: Initial structure filename.
complexEquilOutputProductionTrajectoryFilename: Trajectory filename.
complexEquilOutputTrajectoryWriteInterval: Frequency for writing
    velocities to trajectory file.

Complex equil simulation settings:

Parameters controlling simulation during equilibration phase of
complex transformation.

complexEquilSimulationEquilibrationLength:  Length of the NPT
    equilibration phase.
complexEquilSimulationEquilibrationLengthNVT: Length of the NVT
    equilibration phase.
complexEquilSimulationMinimizationSteps: Maximum number of
    minimization steps to perform.
complexEquilSimulationProductionLength:  Length of the NPT
    production phase.

Complex lambda settings:

Lambda protocol parameters for complex transformation.

complexLambdaElecA: List of lambda values for electrostatics. The
    values of 0 and 1 imply state A and state B respectively.
complexLambdaElecB: List of lambda values for electrostatics. The
    values of 0 and 1 imply state A and state B respectively.
complexLambdaRestraintsA: List of lambda values for restraints. The
    values of 0 and 1 imply state A and state B respectively.
complexLambdaRestraintsB: List of lambda values for restraints. The
    values of 0 and 1 imply state A and state B respectively.
complexLambdaVdwA: List of lamda values for van der Waals. The
    values of of 0 and 1 imply state A and state B respectively.
complexLambdaVdwB: List of lamda values for van der Waals. The
    values of of 0 and 1 imply state A and state B respectively.

Complex output settings:

Parameters controlling simulation output during final phase of
complex transformation.

complexOutputCheckpointInterval:  Frequency to write the checkpoint
    file.
complexOutputCheckpointStorageFilename: Checkpoint filename.
complexOutputForcefieldCache: Filename for caching small molecule
    residue templates.
complexOutputFilename: Trajectory filename.
complexOutputIndices: Selection string for selecting coordinates to
    write.
complexOutputStructure: Topology structure filename.
complexOutputPositionsWriteFrequency: Frequency for writing
    positions to trajectory file.
complexOutputVelocitiesWriteFrequency: Frequency for writing
    velocities to trajectory file.

Complex restraint settings:

Parameters to configure Boresch-style restraint between two groups
of atoms named host  (Hx) and guest (Gx).

complexRestraintKPhiA: Equilibrium force constant for the dihedral
    formed by H2-H1-H0-G0.
complexRestraintKPhiB: Equilibrium force constant for the dihedral
    formed by H1-H0-G0-G1.
complexRestraintKPhiC: Equilibrium force constant for the dihedral
    formed by H0-G0-G1-G2.
complexRestraintKR: Bond spring constant between H0 and G0.
restraintKThetaA: Spring constant for the angle formed by H1-H0-G0.
complexRestraintKThetaA: Spring constant for the angle formed by
```

```
    H1-H0-G0.
complexRestraintKThetaB:  Spring constant for the angle formed by
    H0-G0-G1.
complexRestraintAnchorFindingStrategy: Boresch atom picking strategy
    to use. bonded: pick host atoms that are bonded to each other.
    multi-residue: pick host atoms which can span multiple residues.
complexRestraintDsspFilter: Apply DSSP filter to the host atoms.
complexRestraintHostMaxDistance: Maximum distance between any
    host atom and the guest G0 atom.
complexRestraintHostMinDistance: Minimum distance between any
    host atom and the guest G0 atom
complexRestraintHostSelection: A valid selection string to
    sub-select the host atoms which will be involved in the
    restraint.
complexRestraintRmsfCutoff: Cutoff value for filtering atoms by their
    root mean square fluctuation. Atoms with values above this
    cutoff are ignored.

Complex simulation settings:

Parameters controlling simulation during final phase of complex
transformation.

complexSimulationEarlyTerminationTargetError: Target error for the
    real time analysis measured in kcal/mol. Once the MBAR error of
    the free energy is at or below this value, the simulation will
    be considered complete. The suggested value of 0.12 has shown to
    be effective in both hydration and binding free energy
    benchmarks.
complexSimulationEquilibrationLength: Length of the equilibration
    phase. The specified value must be divisible by
    'integratorTimestep'.
complexSimulationMinimizationSteps: Maximum number of minimization
    steps to perform.
complexSimulationNReplicas: Number of replicas to use.
complexSimulationProductionLength: Length of the production phase.
    The specified value must be divisible by 'integratorTimestep'.
complexSimulationRealTimeAnalysisMinimumTime: Time interval for
    performing analysis of the free energies. At each interval, real
    time analysis data will be written to a yaml file named
    <outputFileName>_real_time_analysis.yaml. The current error
    in the estimate will also be assessed and the simulation will
    be terminated when it drops below
    'complexSimulationEarlyTerminationTargetError'.
complexSimulationSamplerMethod: Alchemical sampling method to use:
    REPEX (Hamiltonian REPlica EXchange), SAMS (Self-Adjusted
    Mixture Sampling), or Independent (Independently sampled lambda
    windows).
complexSimulationSamsFlatnessCriteria:Method for assessing when to
    switch to asymptomatically optimal scheme for SAMS.
complexSimulationsamsGamma0: Initial weight adaptation rate for
    SAMS.
complexSimulationTimePerIteration: Simulation time between each
   MCMC move attempt

Complex solvation settings:

Solvation parameters for the system, including the solvent model and
the solvent padding.

complexSolvationBoxShape: Shape of the periodic solvent box.
complexSolvationBoxSize:  Lengths of the unit cell for a solvent box.
complexSolvationSolventModel: Forcefield water model to use during
    solvation and defining the model properties.
complexSolvationSolventPadding: Minimum distance from any solute
    bounding sphere to the edge of the box.

Engine settings:
```

```
Parameters configuring the compute platform used by the OpenMM to
perform the simulation.

engineComputePlatform: Platform to use for running OpenMM MD
    calculations.
engineGpuDeviceIndex: Space delimited list of device indices
    to use for running OpenMM MD calculations.

Forcefield settings:

forcefieldConstraints:Constraints  to use.
forcefields: List of valid forcefield paths for all components
    except small molecules.
forcefieldHydrogenMass: Mass to be repartitioned to hydrogens
    from neighboring heavy atoms.
forcefieldNonbondedCutoff: Cutoff for short range nonbonded
    interactions.
forcefieldNonbondedMethod: Method for treating nonbonded
    interactions.
forcefieldRigidWater: Use a rigid water model.
forcefieldSmallMoleculeForcefield: A valid forcefield name to use
    for small molecules.

Integrator settings:

Parameters controlling the LangevinSplittingDynamicsMove integrator
used for simulation.

integratorBarostatFrequency: Frequency at which volume scaling
    changes should be attempted.
integratorConstraintTolerance: Tolerance for constraint solver.
integratorLangevinCollisionRate: Collision frequency.
integratorNRestartAttempts: Number of attempts to restart from
    Context in case there are NaNs in the energies after
    integration.
integratorReassignVelocities: Reassign velocities  from the
    Maxwell-Boltzmann distribution at the beginning of each
    Monte Carlo move.
integratorRemoveCom: Remove the center of mass motion.
integratorTimestep: Size of the simulation timestep.

Partial charge settings:

Parameters for automatically assigning missing partial charges to
small molecules, including the partial charge method.

partialChargeNaglModel: Model to use for partial charge assignment.
    A value of None implies the use of the latest available
    production AM1BCC model.
partialChargeNumberOfConformers: Number of conformers to generate
    as part of the partial charge assignment. A value of None
    implies the use of the existing conformer.
partialChargeOffToolkitBackend: OpenFF toolkit registry backend to
    use for calculating partial charges.
partialChargeMethod: Method to use for calculating partial charges.

Solvent equil output settings:
Solvent equil simulation settings:
Solvent lambda settings:
Solvent output settings:
Solvent restraint settings:
Solvent simulation settings:
Solvent solvation settings:

The solvent settings are similar to the complex settings already
described under various sections for complex. The prefix 'solvent'
is used for the names of the pramaters instead of the prefix
'complex.'
```

```
Thermo settings:

Thermodynamic parameters, including the temperature and the pressure
of the system.

thermoPh: Simulation pH
thermoPressure: Simulation pressure.
thermoRedoxPotential:Simulation redox potential.
thermoTemperature: Simulation temperature.
```

-s, --smallMolFile <SmallMolFile>

Input file containing small molecules.

--smallMolFileParams <Name,Value,...> [default: auto]

A comma delimited list of parameter name and value pairs for reading molecules from files. The supported parameter names for different file formats, along with their default values, are shown below:

```
SD: removeHydrogens,no,sanitize,yes,strictParsing,yes
```

--solventParams <Name,Value,...> [default: auto]

A comma delimited list of parameter name and value pairs for solvent component. You must specify valid OpenFE values. No extensive validation is performed. These parameters are used in conjunction with solvation* parameters available through '--rbfeParams' to perform solvation.

The supported parameter names along with their default values are are shown below:

```
positiveIon, Na+ [ Possible value: Li+, Na+, K+, Rb+, or Cs+ ]
negativeIon, Cl- [ Possible values: Cl-, Br-, F-, or I- ]
neutralize, yes  [ Possible values: yes or no ]
ionConcentration, 0.15  [ Units: molar ]
```

A brief description of parameters is provided below:

```
positiveIon, negativeion: Pair of ions used to neutralize and bring
    the solvent to required ionic concentration.
neutralize: Neutralize the net charge on the chemical state by the
    ions in the solvent component.
ionConcentration: Ionic concentration.
```

-w, --workingdir <dir>

Location of working directory which defaults to the current directory.

EXAMPLES

The sample protein and ligand files for tyrosine kinase 2 (Tyk2) are distributed with MayaChemTools and are available in data directory. These files have been taken from OpenFE distribution for example notebooks. The AM1BCC partial charges have been calculated for the ligands in SD file to facilitate calculations. You may review OpenFE tutorial notebooks for the expected results.

To calculate RBFE for a pair molecules corresponding to the fist and second molecules in a SD file, performing 3 independent repeats of the entire MD sampling process to calculate RBFE for a pair of molecules, each complex and solvent MD repeat consisting of equilibration phase (Complex: Minimization - 5,000; NVT - 0.1 ns; NPT - 0.1; NPT prod - 2.0 ns; Solvent: Minimization - 5,000; NVT - 0.1; NPT - 0.1 ns; NPT prod - 2.0 ns) and production phase (Complex: Minimization - 5,000; NPT equil - 1.0 ns; NPT prod: 10.0 ns; Solvent: Minimization - 5,000; NPT equil - 1.0 ns; NPT prod - 10 ns) using a step size of of 4 fs, writing out appropriate trajectory and PDB files for each MD repeat in Results subdirectory under output directory, type:

```
% OpenFECalculateRelativeBindingFreeEnergySepTop.py -i SampleTyk2.pdb
  -s SampleTyk2Ligands.sdf -o SampleTyk2LigandsRBFESepTop
```

To run the first example for calculating RBFE for a specific pair molecules using CUDA platform on your machine to perform MD simulations and generate various output files, type:

```
% OpenFECalculateRelativeBindingFreeEnergySepTop.py -i SampleTyk2.pdb
  -s SampleTyk2Ligands.sdf -o SampleTyk2LigandsRBFESepTop
  --moleculePairs "lig_ejm_31, lig_ejm_47"
  --rbfeParams "engineComputePlatform,CUDA"
```

To run the second example to see all warning messages produced by OpenFE modules and write various output

files, type;

```
% OpenFECalculateRelativeBindingFreeEnergySepTop.py -i SampleTyk2.pdb
  -s SampleTyk2Ligands.sdf -o SampleTyk2LigandsRBFESepTop
  --moleculePairs "lig_ejm_31, lig_ejm_47"
  --rbfeParams "engineComputePlatform,CUDA"
  --loggingLevel Warning
```

To run the first example for calculating RBFE for a specific pair molecules using CUDA platform on your machine to perform MD simulations, automatically calculate missing partial charges for molecules, and generate various output files, type:

```
% OpenFECalculateRelativeBindingFreeEnergySepTop.py -i SampleTyk2.pdb
  -s SampleTyk2Ligands.sdf -o SampleTyk2LigandsRBFESepTop
  --moleculePairs "lig_ejm_31, lig_ejm_47"
  --rbfeParams "engineComputePlatform,CUDA"
  --missingChargeMode Calculate
```

To run the second example by specifying explict values for various parametres and generate various output files, type:

```
% OpenFECalculateRelativeBindingFreeEnergySepTop.py -i SampleTyk2.pdb
  -s SampleTyk2Ligands.sdf -o SampleTyk2LigandsRBFESepTop
  --moleculePairs "lig_ejm_31, lig_ejm_47"
  --loggingLevel Warning
  --executeDAGParams "keepShared, yes, nRetries, 2"
  --missingChargeMode Stop
  --rbfeParams "engineComputePlatform,CUDA,
  protocolRepeats,3,
  engineComputePlatform,CUDA,
  forcefieldConstraints, HBonds, forcefieldHydrogenMass, 3.0,
  forcefieldNonbondedMethod, PME, integratorTimestep, 4.0,
  complexEquilSimulationMinimizationSteps, 5000,
  complexEquilSimulationEquilibrationLength, 0.1,
  complexEquilSimulationEquilibrationLengthNVT, 0.1,
  complexEquilSimulationProductionLength, 2.0,
  complexSimulationMinimizationSteps, 5000,
  complexSimulationEquilibrationLength,  1.0,
  complexSimulationProductionLength, 10.0,
  solventEquilSimulationMinimizationSteps, 5000,
  solventEquilSimulationEquilibrationLength, 0.1,
  solventEquilSimulationEquilibrationLengthNVT, 0.1,
  solventEquilSimulationProductionLength, 2.0,
  solventSimulationMinimizationSteps, 5000,
  solventSimulationEquilibrationLength, 1.0,
  solventSimulationProductionLength, 10.0"
  --solventParams "positiveIon, Na+, negativeIon, Cl-"
```

AUTHOR

Manish Sud(msud@san.rr.com)

SEE ALSO

OpenFECalculateAbsoluteBindingFreeEnergy.py,
OpenFECalculateAbsoluteHydrationFreeEnergy.py, OpenFECalculatePartialCharges.py,
OpenFECalculateRelativeHydrationFreeEnergy.py, OpenFEGenerateLigandNetwork.py

COPYRIGHT

The functionality available in this script is implemented using OpenFE, an open source molecuar for alchemical free energy calculations.

This file is part of MayaChemTools.

MayaChemTools is free software; you can redistribute it and/or modify it under the terms of the GNU Lesser General Public License as published by the Free Software Foundation; either version 3 of the License, or (at your option) any later version.