

Conformation Search Using CHARMM

Description

This set of scripts performs a replica exchange molecular dynamics (a.k.a., parallel tempering) simulation of a molecule using the CHARMM molecular mechanical modeling code. The final structures of the simulations are then minimized. This provides a method to find low-energy conformations of a molecule.

Requirements

These scripts use several other molecular modeling programs in combination to generate the initial coordinates, molecular structure, topology, and parameters. For the example calculation to work, the following software must be installed:

- CHARMM (<http://www.charmm.org/>)
 - tested with version c38b2
- Antechamber (part of AmberTools, <http://ambermd.org/>)
 - tested with AmberTools14
- OpenBabel (<http://openbabel.org>)
 - only required when building molecules from SMILES strings or other formats that are not supported by antechamber
 - tested with version 2.3.2

The executables for these programs (CHARMM, antechamber, and obabel) are assumed to be in the path when the script is run.

Compiling CHARMM

This system requires that CHARMM be run in parallel using MPI. CHARMM must be compiled with MPI parallelization and the replica exchange features (repdstr). To use the semi-empirical QM module of CHARMM to represent the molecule, the SQ keyword must be used. The suggested command line for compiling CHARMM for use with these scripts is:

```
./install.com gnu large M MPIF90 SQ +REPDSTR +GENCOMM +CMPI
```

After compilation is completed, you should copy the charmm executable to the working directory or modify the conformation-search.sh script so that the fully-qualified path is included in used at both instances.

Options

The execution options for this script are specified in the file `settings.str`

- NREPLICAS
 - Specifies the number of replicas to simulation in the REMD simulation
 - This should not exceed the number of processors available to be run in parallel

- A sufficient number of replicas must be used so that neighbouring replicas are close in temperatures
 - Default value: 20
- LOWT
 - Specifies the temperature of the lowest temperature replica in Kelvin
 - If you are searching for the global potential energy minimum, this should be sufficiently low so that only lowest energy conformation is likely to be populated at the lowest temperature; however, larger temperature ranges need more replicas.
 - Default value: 50
- HIGHT
 - Specifies the temperature of the lowest temperature replica in Kelvin
 - The inclusion of high temperature replicas; however, if it is set too high (e.g. > 1000 K), the simulation can become numerically unstable and non-physical isomerizations can occur (e.g. rotations around double bonds), especially in QM simulations.
 - Default value: 500
- NSTEPS
 - Specifies the number of time steps steps to perform. Each time step is 1 fs in length ($1 \text{ fs} = 10^{-15} \text{ s}$). This should be sufficiently large to allow the simulation to find the lowest energy conformation.
 - Default value: 1000000
- XFREQ
 - The number of time steps between replica exchange attempts
 - Default value: 1000
- QM
 - Defines whether to represent the molecule with an molecular mechanical model (QM = 0) or with a semi-empirical quantum mechanical model (QM = 1)
 - Default value: 0
- QMCHARGE
 - Stores the net charge of the molecule for QM calculations
 - Default value: 0
- QMMETHOD
 - Specifies which semi-empirical QM method to use in QM calculations. The options are limited to am1, pm3, or mndo
 - Default value: pm3

Execution

The directory must contain mol.rtf, where the molecule that conformational search is being performed on is defined as "RESI MOL". The directory must also contain a file, mol.prm, that contains. It is recommended to use antechamber to generate these parameters from the Generalized Amber Force Field (GAFF), although these parameters will be ignored when a QM model is used.

```
mpirun -np 20 charmm < remd.inp > remd.out
```

It may be necessary to modify the mpirun command line to point to the host file or the correct path for the version of mpirun used in the compilation. By default, the simulation will use 20 replicas. When the simulation completes, the minimum energy structure is saved to the files min_pdb_# in PDB format. The energy of this conformation is stored in the first line of these files. The structure of the minimum temperature structure (min_pdb_0) will often be the lowest energy structure, although it is possible that one of the higher temperature replicas will hold the lowest energy structure.

When a QM method is used, it becomes possible for chemical reactions to occur in high temperature replicas, so isomerizations and fragmentations can occur. This can be prevented by adding restraints to the remd.inp script to prevent specific thermally accessible reaction mechanisms or by using a lower maximum temperature in the replica exchange.

Example

The included example builds maltose from its SMILES string, then uses CHARMM to perform a replica exchange conformation search. This example is executed by:

```
cd example  
./conformation-search.sh
```

The lowest energy conformation will be written to min_pdb_0.