

# MCCE Tools

List of miscellaneous tools MCCE offers for your research and convenience. Most of these tools support "-h" flags for additional information and use cases.

## clear\_mcce\_folder

Deletes all MCCE outputs from the present working directory, except: run.prm, the original PDB file, prot.pdb, and any non-MCCE files.

## detect\_hbonds.py

Detect H-bonds in a PDB file, with the option to include BK (backbone) atoms. hbonds\_pdb\_collection uses this function on a collection of PDB files.

usage: detect\_hbonds.py [-h] [--include\_bk] [--no\_empty\_files] [--out\_dir OUT\_DIR] [inpdb]

## extract\_md\_frames

Extracts the trajectory's frames with the given indices into PDB files. Requires the MDAnalysis package.

## filesdiff

Obtain the column difference between two MCCE files or the differences of all files in two MCCE output folders. Use the "-threshold" flag to output absolute differences beyond a given value (0 is default).

Applicable to the following MCCE files: 'all\_pK.out', 'all\_sum\_crg.out', 'entropy.out', 'fort.38', 'head3.lst', 'pK.out', 'residues\_stats.txt', 'sum\_crg.out', 'vdw0.lst'.

## fix\_psf\_mdanalysis

Provides a reformatted PSF file if "MDAnalysis" fails to parse the given PSF. Requires the "MDAnalysis" and "parmed" packages.

# getpdb

Downloads one or more (bioassembly) PDB files from the RCSB Protein Databank. For example, to download triclinic hew lysozyme (4LZT), one could type in

usage: getpdb [RCSB protein code]

# glossary

Gives detailed information regarding the various parameters of run.prm, where MCCE looks to handle more granular customization.

You can search for specific parameters by with a given (case-sensitive) prefix string. For example, "glossary T" will return all parameters starting with T, like "TITR\_TYPE". The command "glossary --print" also prints the entire glossary.

# hbonds\_pdb\_collection

Detects Hydrogen bonds, using detect\_hbonds.py, over a collection of PDB files, in the step2\_out.pdb format. ASK HOW TO USE THIS

usage: hbonds\_pdb\_collection [-h] [-input\_dir INPUT\_DIR] [-output\_dir OUTPUT\_DIR] [--include\_bk] [--no\_empty\_files]

# mcce\_stat

Prints a table to keep track of progressing MCCE runs. Four "sentinel" files are looked for, to signify completion of each of the four basic steps of MCCE: step1\_out.pdb, step2\_out.pdb, head3.lst, and pK.out.

pK.out signifies completion of step 4, so if a book.txt exists for a protein when mcce\_stat is run, that protein will receive a "c" in book.txt to signify completion.

We recommend using mcce\_stat with p\_batch.

# ms\_hbond\_percentages.py

Creates a table displaying all Hydrogen bond connections across microstate PDBs, and their percentages. Defaults to the local directory named pdb\_output\_mc\_hbonds.

usage: ms\_hbond\_percentages.py [-h] [dir]

## ms\_top2pdb

Stands for Tautomeric Charge MicroStates. Outputs: the top N tautomeric charge microstates, along with related properties energy (E), net charge (sum\_crg), count, and occupancy (occ); a summary file identifying ionizable residues with non-canonical charge, and which residues that do not change charge over the topN set; and the top N files of each charge state in PDB and PQR format.

By default, charge microstates are retrieved at pH 7, and the number of most favorable charge microstates (N\_TOP) returned is five.

usage: ms\_top2pdb inputpdb\_filepath [-ph PH] [-n\_top N\_TOP]

## p\_batch

Starts multiple protein runs at once, using the same set of instructions, and creates a book.txt file to manage their completion status. p\_batch accepts a directory containing protein files, and (optionally) a shell script given custom instructions. If a shell script is not provided, a default one will be created, and may be edited to the user's preference. If a file named "run.prm.custom" is in the present working directory at runtime, the file will be read to override the default run.prm instructions.

p\_batch creates a run directory for each protein file, and begins running MCCE for each one. Files will be created for their respective directories as each step is completed. Use mcce\_stat to check how each run is progressing.

To stop a run in progress, delete the files or directory associated with the run.

## p\_info

Gives a high-level summary of characteristics of a PDB file, including residue, chain, and ligand counts, as well as other aspects of a PDB changed during step 1 of MCCE, including how residues are named. If step 1 has not been run on the PDB file at runtime, p\_info will automatically run step 1 before continuing as normal.

## txt\_to\_csv

A quick script that copies a given file into a .csv format. The source file does not need to be a .txt file. Recommended to use with spreadsheets.

---

Revision #4

Created 16 April 2025 18:05:41 by Jared

Updated 2 May 2025 17:19:48 by Jared