

Obtaining a PMF from a series of EVB simulation in Amber

Kim F. Wong

Revision: April 5, 2006

1 Installation and general usage

The `AMBER_EVB_PMF.tbz2` is provided as a supporting code for generating the potential of mean force (PMF) from a set of EVB simulations within Amber 9. The user will need to modify the `Makefile` to correspond to his or her environment variables. The code was developed using `ifort` under Linux but was also tested using `pgf90`. The routines use libraries from LAPACK and BLAS so both of these need to be installed.

Unzipping and untarring the file (`tar xvfj AMBER_EVB_PMF.tbz2`), will generate the directory `AMBER_EVB_PMF`, which contains the `SRC` and the `EXAMPLES` directories. Assuming that the `Makefile` was modified appropriately, typing `make` within `SRC` will compile the executable `pmf.x`. All that is required by the program is an input file called `PMF.INPUT` as well as the biased sampling EVB output files from Amber.

The `EXAMPLES` directory contains a mapping potential example (`NRG_EMAP`) and an energy gap umbrella sampling example (`NRG_EUMB`) for the hydroxypyridine to pyridone conversion via intramolecular proton transfer (Figure 1). Within these directories are other directories containing the EVB output data from Amber as well as example input files and PBS submission scripts for using the EVB facility in Amber. Please consult the Amber manual regarding the input files. Your local systems administrator should be able to help with the submission scripts.

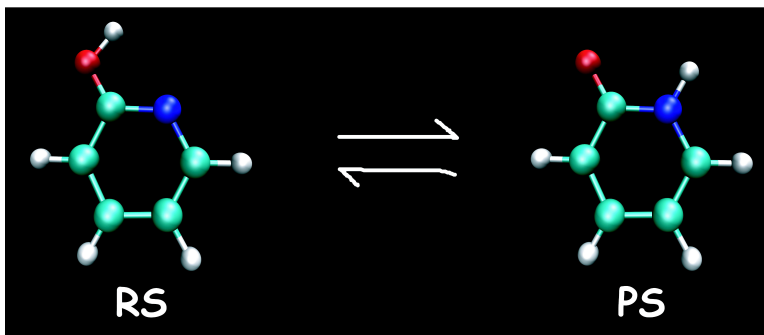


Figure 1: Hydroxypyridine to pyridone conversion via intramolecular proton transfer

Let us briefly go over the example for the mapping potential simulations. The `PMF.INPUT` file there specifies that the data contained in the files listed in the `evb_ifile(:)` variable corresponds to a set of 2-state EVB simulations with a single biasing potential applied to the dynamics. The average temperature of the canonical sampling is 300.0 K and the bin width used for constructing the average distribution of the reaction coordinate is set at 0.40 kcal/mol. Executing the PMF code (`./pmf.x`) will crank through all the data files and output a series of files that begins with `AWARSHEL_PMF_*`. Files that begin with `AWARSHEL_PMF_ELO_*` contain the piece-wise PMF for sampling on the EVB ground-state surface V_{el0} (Figure 2). Files that begin with `AWARSHEL_PMF_Vii_*` contain the PMF corresponding to molecular dynamics sampling based on the *initial* Hamiltonian, while files that begin with `AWARSHEL_PMF_Vff_*` correspond to sampling based on the *final* Hamiltonian.

Now the `PMF.INPUT` for the energy gap umbrella sampling is similar, but with two additional input variables: `wham_tol` for specifying the self-consistency criterium for the weighted histogram analysis method (WHAM) and `FE_C(:)` for specifying the initial guesses for the free energy constants. The resulting PMF is contained in the file `WHAM_PMF_UNBIASED` (Figure 3). Files that begin with `WHAM_RHO_BIASED_*` contain the average distribution of the reaction coordinate for each umbrella sampling window and those that begin with `WHAM_PMF_BIASED_*` are the corresponding PMF slices.

As the suffix tag of the filenames implies, each file is a concatenation

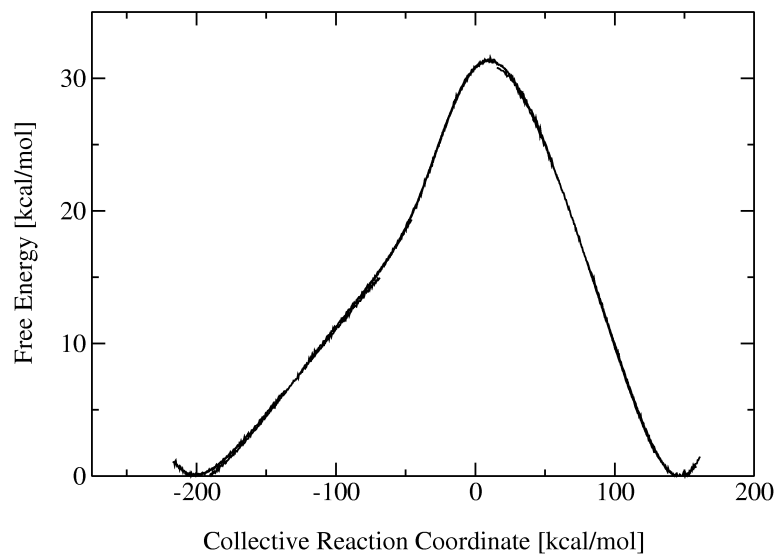


Figure 2: Potential of mean force obtained from a series of mapping potential simulations

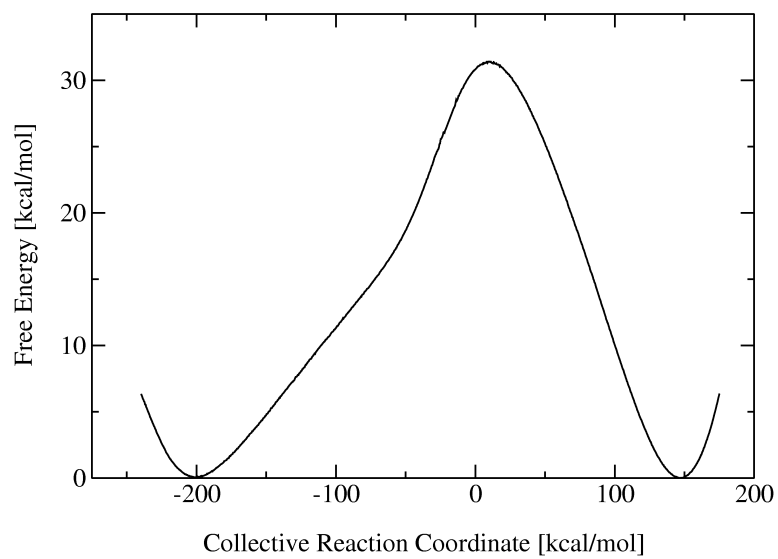


Figure 3: Potential of mean force obtained from umbrella sampling of the energy gap reaction coordinate

of multiple trajectories at that particular biased sampling window. No pre-processing of the EVB output files from Amber is necessary. Yet what prevents the user from accidentally concatenating a trajectory from the `-25.00` kcal/mol window with that from the `+25.00` kcal/mol window? The code does check that the header information from all the trajectory slices in a data file are consistent; otherwise, the program exits with an error. Setting the input variable `Lchk_data_slices` to `.false.` will disable this check but the standard user should leave this at the default setting.

2 Description of variable names in `PMF.INPUT`

The variables for the `PMF.INPUT` file are described below. Default values are enclosed in curly braces `{ ... }`.

| | | |
|------------------|-----------|--|
| nevb | [integer] | Number of EVB states {2} |
| nbias | [integer] | Number of biasing potentials included in the system Hamiltonian {1} |
| bin_width | [real] | Bin width used for constructing the reaction coordinate distribution {0.40 kcal/mol} |
| temperature | [real] | Temperature of canonical MD sampling {300.0 K} |
| nhits_tol | [integer] | Minimum number of hits/bin. This is used to determine which bin results are written to file {25}. |
| evb_ifile(:) | [string] | Array containing the names of the EVB data files from Amber. Each file name is enclosed in quotes and the array element index should be sequential |
| FE_C(:) | [real] | Array containing the initial guesses for the free energy constants associated with each biased simulation window specified in evb_ifile {0.0 kcal/mol} |
| wham_tol | [real] | WHAM self-consistency tolerance {1.0e-5} |
| reject_1stdata | [integer] | Discard this many data points from each EVB data file. Useful for discarding preliminary unequilibrated data {0} |
| Lchk_data_slices | [logical] | Check the consistency of concatenated data slices in each EVB data file {.true.} |

3 Example PMF.INPUT input file (mapping potential)

```
&evb nevb = 2, nbias = 1, bin_width = 0.40,
      reject_1stdata = 100,
      temperature = 300.0,
      nhits_tol = 25,
```

```
evb_ifile(1) = "evbout_L-0.05_CAT",
evb_ifile(2) = "evbout_L-0.10_CAT",
evb_ifile(3) = "evbout_L-0.15_CAT",
evb_ifile(4) = "evbout_L-0.20_CAT",
evb_ifile(5) = "evbout_L-0.25_CAT",
evb_ifile(6) = "evbout_L-0.30_CAT",
evb_ifile(7) = "evbout_L-0.35_CAT",
evb_ifile(8) = "evbout_L-0.40_CAT",
evb_ifile(9) = "evbout_L-0.45_CAT",
evb_ifile(10) = "evbout_L-0.50_CAT",
evb_ifile(11) = "evbout_L-0.55_CAT",
evb_ifile(12) = "evbout_L-0.60_CAT",
evb_ifile(13) = "evbout_L-0.65_CAT",
evb_ifile(14) = "evbout_L-0.70_CAT",
evb_ifile(15) = "evbout_L-0.75_CAT",
evb_ifile(16) = "evbout_L-0.80_CAT",
evb_ifile(17) = "evbout_L-0.85_CAT",
evb_ifile(18) = "evbout_L-0.90_CAT",
evb_ifile(19) = "evbout_L-0.95_CAT",
&end
```

4 Example PMF.INPUT input file (energy gap umbrella sampling)

```
&evb nevb = 2, nbias = 1, bin_width = 0.40,
neject_1stdata = 100,
wham_tol = 1.0e-5,
temperature = 300.0,
nhits_tol = 25,
evb_ifile(1) = "evbout_L_250.00_CAT",
evb_ifile(2) = "evbout_L_225.00_CAT",
evb_ifile(3) = "evbout_L_200.00_CAT",
evb_ifile(4) = "evbout_L_175.00_CAT",
evb_ifile(5) = "evbout_L_150.00_CAT",
evb_ifile(6) = "evbout_L_125.00_CAT",
evb_ifile(7) = "evbout_L_100.00_CAT",
evb_ifile(8) = "evbout_L_75.00_CAT",
evb_ifile(9) = "evbout_L_50.00_CAT",
```

```
evb_ifile(10) = "evbout_L_25.00_CAT",
evb_ifile(11) = "evbout_L_15.00_CAT",
evb_ifile(12) = "evbout_L_0_CAT",
evb_ifile(13) = "evbout_L_-15.00_CAT",
evb_ifile(14) = "evbout_L_-25.00_CAT",
evb_ifile(15) = "evbout_L_-50.00_CAT",
evb_ifile(16) = "evbout_L_-75.00_CAT",
evb_ifile(17) = "evbout_L_-100.00_CAT",
evb_ifile(18) = "evbout_L_-125.00_CAT",
evb_ifile(19) = "evbout_L_-150.00_CAT",
evb_ifile(20) = "evbout_L_-175.00_CAT",
evb_ifile(21) = "evbout_L_-200.00_CAT",
evb_ifile(22) = "evbout_L_-225.00_CAT",
evb_ifile(23) = "evbout_L_-250.00_CAT",
FE_C(1) = 0.0000,
FE_C(2) = -14.1673,
FE_C(3) = -24.3710,
FE_C(4) = -30.6089,
FE_C(5) = -32.9725,
FE_C(6) = -31.6637,
FE_C(7) = -27.1443,
FE_C(8) = -20.2554,
FE_C(9) = -12.3303,
FE_C(10) = -5.0491,
FE_C(11) = -2.9157,
FE_C(12) = -4.8976,
FE_C(13) = -8.9926,
FE_C(14) = -11.4053,
FE_C(15) = -16.0755,
FE_C(16) = -19.5214,
FE_C(17) = -22.7678,
FE_C(18) = -26.0634,
FE_C(19) = -29.2672,
FE_C(20) = -31.8910,
FE_C(21) = -33.0264,
FE_C(22) = -31.6773,
FE_C(23) = -27.5030,
```

&end