##
! This code and script calculate the free energy difference and entropy difference between the target state (holo) and the reference state (apo) based on the histograms of the microscopic conformational variables in the two states of all the structural elements, like amino acid residues for a protein and the base pairs for DNA.

The microscopic conformational degrees of freedom are:

For Protein: (1) Phi, psi and the side chain dihedrals (chi1, chi2, chi3, chi4, and chi5).

For DNA: (1) the base-pair step parameters (shift, slide, rise, tilt, roll and twist)s; (2) the intra base-pair parameters (stagger, shear, stretch, buckle, propeller and opening); (3) sugar-phosphate and sugar-base torsion angles (α, β, γ, δ, ε, ζ and χ) and (4) sugar-pucker angles (ν0,ν1, ν2, ν3 and ν4).

For each microscopic conformational variable, the code and script should be run as per the following steps. For instance, if one is interested in calculating conformational thermodynamics due to the phi and psi backbone changes for a 50-residue long protein, the code and script should run twice in two separate working directories, one for phi and the other for psi. The working directory for phi will give the conformational thermodynamics values for phi of all 50 residues and similarly for psi. One can add the conformational thermodynamics values for phi and psi of each residue to get total conformational thermodynamics change for the given residue. Similarly, the other cases can be run. Python version 3 and FORTRAN codes are used here.

Follow the steps as detailed below:

Step I: The program assumes that the user knows the relevant microscopic variables and calculates them for each conformation after equilibration using standard codes, for example, GROMACS dihedral calculation module for proteins, NUPARM with BPFIND for DNA base-pair step parameters and intra base-pair parameters and so on. The data should be stored in 'apo.dat' and 'holo.dat' file respectively for the two states, written in the free format with three columns, namely, frame number, residue/base-pair/base name and the value of the microscopic variable. Angular variables should be in radian and all other variables should be made dimensionless.

Step II: The following files should be in the working sub-directory: apo.dat, holo.dat, input_parameters.dat, sort.f95, max_min.py, histogram_max_min.py, max_min_merge.sh, histogram.py, conf-thermodynamics.f90 and conftherm.sh.

Step III: Next put the input variables in input_parameters.dat as per your need and run the sort.f95 program.

Description of the input variables in 'input_parameters.dat' as needed by the sort.f95 program (the main program here refers to this program):

(1) nframe: Number of conformations you want to perform the analysis; the number must be below or equal to maxframe which is set to 100000 in parameter statement at the header of the

main program; For larger values please reset maxframe in the source code. Make sure that your computer can handle large storage place.

(2) nres: Number of residues/nucleotides you want to perform the analysis: The number must be below or equal to maxres which is set to 999 in parameter statement at the header of the main program.

(3) nwindow: Number of windows to be below maxwindow which is set to 20 in parameter statement at the header of the main program.

(4) nconf: Number of frames per window to be below maxconfpw which is set to 5000 in parameter statement at the header of the main program. Note that nconf * nwindow < maxframe.

(5) nbin: Number of bins for histogram which should be less than maxbin (set to 100) in the header of the main program .

File specification:

(1) apo.dat: Input data for microscopic variable in different conformations of reference state (apo).
(2) apo-smp.dat: Output random sample for apo consisting of nconf*nwindow values.

(3) holo.dat: Input data for the microscopic variable in different conformations of the target state (holo).

(4) holo-smp.dat: Output random sample for holo consisting of nconf*nwindow values.

The file names to be given for the windows as follows:

(1) apo_smp(i): Give file name containing random sample for the i^{th} window of the apo. The filename is (i)-apo-smp.dat.

(2) apo_maxmin(i): Give file name writing maximum and minimum values of the variables of the i^{th} window of the apo. The filename is (i)-apo-max_min.dat.

(3) apo_hist(i): Give the file name for writing histogram of the i^{th} window of the apo. The filename is (i)-apo-hist.dat.

(4) holo_smp(i): Give file name containing random sample for the i^{th} window of the holo. The filename is (i)-holo-smp.dat.

(5) holo_maxmin(i): Give file name writing maximum and minimum values of the variables of the i^{th} window of the holo. The filename is (i)-holo-max_min.dat.

(6) holo_hist(i): Give the file name for writing histogram of the i^{th} window of the holo. The filename is (i)-holo-hist.dat.

Step IV

Run sort.f95 program. This will generate the files: apo_smp.dat, holo_smp.dat, listhistprotmaxmin, listhistprot, histparameter, parameter_maxmin, num_window, listconformational_td and resname.dat.

Step V
Go to the script file conftherm.sh. Make the following changes:

(1) Change imax where the variable stands for total number of structural elements+1: for instance 141 for a 140 amino acid long protein, 21 for a 20 base pair long DNA polynucleotide etc.

(2) In the split command 'split -l 100 --numeric-suffixes=1 --suffix-length=3 --additional-suffix="-*.dat" *-smp.dat "" --verbose', the number 100 to be replaced by nwindow*nconf.

(3) In the split command 'split -l 50 --numeric-suffixes=1 --suffix-length=2 --additional-suffix="-apo-smp.dat" ${i}-apo.dat "" --verbose', the number 50 to be replaced by nconf.

The split program will generate imax number of subdirectories to work with the conformational thermodynamics program (conf-thermodynamics.f90) for each window in the respective subdirectories. Window wise data is given in the output file conf-therm_window.dat and the window averaged data and error in the output file conf-therm_av.dat. All the subdirectories are deleted upon execution of the conformational thermodynamics program.

The averaged data and the error for all the residues are given in the working directory in the output file conf-therm.dat in the unit of kJ/mol. The created subdirectories are removed.
##
Tutorial:

(1) We add two test files (apo.dat and holo.dat) for the analysis. Two different files correspond to the phi dihderal angle of two different conformations of 140 residue-long α-synuclein protein, without ZnO-nanoparticle (apo) and with ZnO-nanoparticle (holo). We calculate the conformational thermodynamics for phi of each residue in the holo state with respect to the apo state using two windows, each having 50 randomly chosen conformations from equilibrium values listed in apo.dat and holo.dat. Using the MD trajectories, phi values are generated and stored in the apo and holo states in the files apo.dat and holo.dat respectively.

(2) Set the appropriate input data in "input_parameter.dat". The calculations are done with nframe = 1000, nres = 140, nwindow = 2, nconf = 50 and nbin = 10.

(3) Run "sort.f95" with the following commands on the screen:

(i) gfortran sort.f95 -o a.out

(ii) ./a.out

(4) Finally run the script "conftherm.sh" with the following commands on the screen:

(i) chmod + conftherm.sh

(ii) ./conftherm.sh

(5) In the working directory, the averaged data and the error of conformational free energy and entropy for all the residues are given in the output file "conf-therm.dat" in the unit of kJ/mol.

##

! Jaydeb Chakrabarti, SNBNCBS provided the fundamental concepts behind of this program and wrote part of the coding (jaydebchakrabarti@gmail.com).

! Codes were written by Abhik Ghosh Moulick (abhik.ghoshmoulick@gmail.com) and Kanika Kole (kanikakole0094@gmail.com) and reviewed by Jaydeb Chakrabarti.

##

If you use the code please cite the followings:

1. Amit Das, J. Chakrabarti and Mahua Ghosh. "Conformational contribution to thermodynamics of binding in protein-peptide complexes through microscopic simulation." Biophysical journal 104, 6, **2013**, 1274-1284.

2. Amit Das, J. Chakrabarti and Mahua Ghosh. "Conformational thermodynamics of metal-ion binding to a protein." Chemical Physics Letters 581, **2013**, 91-95.

3. Amit Das, J. Chakrabarti and Mahua Ghosh. "Thermodynamics of interfacial changes in a protein–protein complex." Molecular Biosystems 10, **2014**, 437-445.

4. Samapan Sikdar, J. Chakrabarti and Mahua Ghosh. "A microscopic insight from conformational thermodynamics to functional ligand binding in proteins." Molecular Biosystems 10, 12, **2014**, 3280-3289.

5. Abhik Ghosh Moulick and J. Chakrabarti, "Fluctuation-dominated ligand binding in molten globule protein." Journal of Chemical Information and Modeling, **2023**, 5583-5591.

6. Kanika Kole, Aayatti Mallick Gupta and J. Chakrabarti. "Conformational stability and order of Hoogsteen base pair induced by protein binding." Biophysical Chemistry 301, **2023**, 107079.