MaxEnt implements a method based on the Maximum Entropy principle that can re-weight an ensemble of protein structures based on data from Residual Dipolar Couplings (RDCs). The RDCs of Intrinsically Disordered Proteins inform on its secondary structure, and our algorithm reweights the ensembles accordingly.

Why MaxEnt?

Intrinsically Disordered Proteins (IDPs) are an emerging family of proteins characterized by adopting a vast number of configurations in solution. A current challenge is to both generate and characterize the ensemble of configurations that make an IDP functional.

The amount of secondary structure on IDPs governs its binding kinetics to other proteins and small molecules and thus determine its potential as drug target. Knowing the amount of secondary structure is key to start to understand IDPs behavior.

A very suitable technique to characterize secondary structure at a residue level is the NMR technique based on Residual Dipolar Couplings (RDCs).

Maximum Entropy Prnciple

The Maximum Entropy (MaxEnt) principle derives from minimizing the information included in an ensemble to fit certain observables.

We present the application of the MaxEnt to the a posteriori re-weighting of an ensemble that has already been calculated. We also add some modifications needed to treat RDC data.

Our MaxEnt application from a primary set of RDCs data (experimental or calculated) given a second one (calculated from Molecular Dynamics or Monte Carlo simulations), reweigths the second ensembles and minimize them according to a gradient fit. These processes is repeated until the threshold is achieved. This threshold value is determined by the experimental precision.

RunPales

The present algorithm relies on the calculation of RDCs from a given set of PDBs. PALES has been thoroughly used for this purpose, and we provide a python 3 script that can call the PALES executable, generate the corresponding RDCs and read the output and covert it into a Numpy array. RunPales is nothing but an interface for call PALES program with the suitable options, and store the generated results. Is not necessary use RunPales for the correct use of MaxEnt.

Ultimately, the user is free to calculate the RDCs with other algorithms, as the input for the MaxEnt program is just the numpy array containing the RDCs for each of the structures.

In fact, nothing in the MaxEnt algorithm depends on the data being RDCs. It can be applied to any scale-invariant observable.

Additional Information

Mathemathical formulation andd further details described at the paper referred at the **citation.md** file.