**CaMBDa** Handbook

**CaMBDa** (**Ca**lculate **M**etric: **BDa**mage) is a program written in Python which calculates ***B*Damage** values for the atoms in biomolecular structures, when atomic information is input in PDB format.

NOTE: An historic version of the program written in MATLAB is also available from <https://github.com/td93/B_Damage>

**Please cite:** M. Gerstel, C. M. Deane and E.F. Garman. (2015). J. Synchrotron Radiation. **22**, 201-212 <http://dx.doi.org/doi:10.1107/S1600577515002131>

**Introduction**

For molecular structures determined using X-ray crystallography each atom in the structure is assigned an ***atomic B factor*** value. This value effectively represents our level of uncertainty about the allocated position of that atom. The most mobile atoms will have the highest B factor values. However an atom's mobility could be due to different factors: increased thermal motion due to absorbed energy from the incident X-ray photons (a sign of *radiation damage*), or it could simply be due to the atom being in a highly flexible region of the protein and it is not surrounded by other 'stationary' atoms. ***B*Damage** is a metric that attempts to deconvolute these factors to give an "effective" B factor value that is just a measure of the damage.

**B\_Damage** for a given atom is the ratio of the atom's B factor and the average B factor of atoms within a *similar packing density*. Atoms are grouped into **Similar packing density** environments, meaning that atoms in that group have a similar surrounding atomic environment which can be defined in several ways. The default here is that atoms with similar packing densities have a similar number of atoms within a given radius of the atom.

**Usage**

*Required Inputs:*

**CaMBDa** takes a minimum of one input, which directs the program to where the atomic information can be found. There are two options here, if the required input is a structure publicly available on the PDB (<http://www.rcsb.org>) then the four digit PDB accession code, enclosed in inverted commas, is sufficient to run the program. Provided the machine on which the program is running has access to an internet connection, the program will automatically retrieve the required data from the PDB, and download a local copy of the PDB file.

*Optional Inputs:*

**CaMBDa** can in addition take a number of optional inputs to vary the output of the programme. Unless you have a specific reason for changing these outputs, leaving them as the default is usually sufficient.

**Description of Errors output by Python code:**

1. Supplied filepath to PDB is neither a .pdb or .txt file nor a PDB accession code
2. Supplied filepath does not exist

HTTP error: Likely that PDB code does not exist

1. Failed to download and save PDB - cause unknown (check internet connection)
2. Failed to copy PDB to a local version. Check that supplied PDB is not in use by another program
3. Failed to generate Unit Cell PDB file. It is likely that PDBCUR failed to run to completion, or experienced an error