 # XYZ to PDB Converter

Overview

This Python script, **XYZ to PDB Converter**, is designed to convert molecular geometries from an `.xyz` file into `.pdb` format using a base `.pdb` file. It creates multiple `.pdb` files, each corresponding to a geometry from the `.xyz` input. This tool is particularly useful for tasks such as reparameterizations in molecular dynamics simulations (e.g., for AMBER).

Features

- Reads an `.xyz` file containing multiple molecular geometries.

- Uses a base `.pdb` file to ensure consistent atom ordering and nomenclature.

- Outputs multiple `.pdb` files, each containing one geometry from the `.xyz` file.

- Handles mismatches between atom counts in the `.xyz` and `.pdb` files with error reporting.

Requirements

- Python 3.x

- A valid `.xyz` file containing molecular geometries.

- A base `.pdb` file with the same atom ordering as the `.xyz` file.

Usage

Run the script from the command line with the following syntax:

python multixyz_to_pdb.py input.xyz base.pdb output_prefix

Arguments:

1. `input.xyz`: Path to the input `.xyz` file containing molecular geometries.

2. `base.pdb`: Path to the base `.pdb` file.

3. `output_prefix`: Prefix for the output `.pdb` files.

Example:

python multixyz_to_pdb.py molecule_geometries.xyz template.pdb output_molecule

This will generate files like `output_molecule_1.pdb`, `output_molecule_2.pdb`, etc., for each geometry in `molecule_geometries.xyz`.

Input Files

1. **XYZ File:**

- Contains multiple molecular geometries in XYZ format.

- Each geometry starts with the number of atoms, followed by a comment line, and then atomic coordinates.

2. **Base PDB File:**

- A valid PDB structure with ATOM/HETATM records.

- Must have the same number of atoms as each geometry in the XYZ file.

Output

The script generates one `.pdb` file per geometry in the input `.xyz`. Each output file contains:

- The atomic coordinates from the corresponding XYZ geometry.

- The atom ordering and other metadata from the base PDB file.

Output files are named using the specified prefix followed by an index (e.g., `output_prefix_1.pdb`, `output_prefix_2.pdb`, etc.).

Error Handling

- The script checks for mismatches between atom counts in the XYZ and PDB files and skips invalid frames with a warning.

- If input files are missing or invalid, it exits with an error message.

Applications

This tool is ideal for:

1. Converting large sets of molecular geometries from quantum chemistry scans into a consistent PDB format.

2. Preparing input files for reparameterization workflows in molecular dynamics simulations (e.g., AMBER).

3. Ensuring consistency in atom ordering and nomenclature across multiple molecular structures.

Author

Richard Lopez Corbalan

GitHub: github.com/richardloopez

Citation

If you use this code, please cite:

Lopez-Corbalan, R.
