PyFREC v 2.7 User Manual

Usage:

pyfrec.py -f input_file.ini

or

pyfrec.py -f .\input_file.ini 1> .\output_file.log 2>&1

Parameters:

input_file.ini - main input file

output_file.log - main output file

Description

PyFREC is a command-line tool and takes all input in a form of text files and produces its output to the standard outputs. Thus, it may be used in conjunction with other command-line tools. The Windows batch script files (!run.bat) are available in each directory with samples.

Description of the main input file (*input_file.ini***)**

The file contains all global parameters of the simulation and contained references to additional input files. Each line staring with the pound sign "#" is a remark line and ignored by the PyFREC. The main input file contains the following sections:

Section: [METHODS]

This section contains global simulation parameters. The following parameters are currently available:

JOB = MAKE_FRAGMENTS - Generate molecular fragments out of given PDB file (see below)

JOB = SURVEY - Run the simulation of electronic couplings and Forster energy transfer rates

RATES = OVERLAP – Compute energy transfer rates based on overlap of donor emission (fluorescence) and donor absorption spectra

Section: [MOLECULAR_SYSTEM]

This section contains information about the molecular system (e.g. donor-acceptor pair) that will be simulated. The section contains the following parameters:

GEOMETRY_FILE = ./input_structure.pdb - The PDB (Protein Data Bank, sometimes also uses ENT file extension) a common format to store molecular structure of biological molecules. PyFREC analyzes only lines starting from "ATOM" or "HETATM" keywords. All other lines are ignored.

For example, an acceptable minimal description of an atom in a PDB file is:

"HETATM 2 C BDP A 1 1.216 -0.002 0.003 C"

Atom #2 (Carbon) that belongs to the residue BDP #1 whichbelongs to the Chain A. Cartesian coordinates (x, y, and z) of the atom are: 1.216 Å, -0.002 Å, and 0.003 Å. See also examples of acceptable PDB files in the samples directory of the PyFREC

The electrostatic screening model. The electronic coupling may be multiplied by the electrostatic screening factors (see details in Ref. X):

EL_SCR_MODEL = 0 No electrostatic screening is used (default)

EL_SCR_MODEL = 1 Uniform screening factor. Additional keyword **EL_SCR_FACTOR** specifies the value of the screening factor (default EL_SCR_FACTOR = 0.8).

EL_SCR_MODEL = 2 Exponential screening function. Additional keyword specifies parameters of the exponential screening function: **EL_SCR_EXP_FUNC = 2.68, 0.27, 0.54** (default values are: pre-exponential factor A=2.68, attenuation factor beta=0.27 $Å^{-1}$, asymptotic value s_0=0.54)

N = 1.333 - Refractive index of the solution may be specified (e.g. 1.333 for water). The default value of the refractive index is 1.0. However, if the **EL_SCR_MODEL = 1** and the **EL_SCR_FACTOR** is not specified then the **EL_SCR_FACTOR** will be set to: **$1/N^2$**. For example:

```
EL_SCR_MODEL = 1

#Refractive Index

#Acetonitrille

N = 1.3441
```

The electrostatic screening factor will be automatically set to the screening factor: $1/N^2 \approx 0.554$

Strickler-Berg (SB) calculation of the donor's lifetime is controlled with the **SB** keyword:

SB = 0 Do not run calculation, empirical values are provided in the input (Default)

SB = 1 Evaluate integrals of the donor emission spectrum

SB = 2 Estimate integrals based on the maximum of the donor emission spectrum

KAPPA_SQ - Specified the numerical value of the κ^2 orientation factor to compute homotransfer properties of the molecular fragments without specifying their molecular geometries. Commonly used value is $\kappa^2 = 2/3 = 0.6666$

For example:

```
KAPPA_SQ = 0.6666
```

Sections: [FRAGMENT…] - There are one or more sections named: [FRAGMENT1], [FRAGMENT2], etc. that contain definitions of molecular fragments (e.g. FRET donor or acceptor groups) used in simulations. Each fragment has the following parameters:

NAME – is a 3-letter residue name of the fragment that matches its name in the PDB file.

ID – is a residue ID number that matches its PDB residue ID

ATOMLIST – is comma-separated list of atomic IDs from the PDB file that belong to the fragment. Put 0 to indicate that all atoms of the given residue belong to the fragment.

ATOMTRANS – is comma-separated list of atomic IDs from the PDB file that belong to the fragment and will be used to perform molecular alignment of the fragment (see details in Ref. XX). It should be noted that PyFREC treats molecular fragment as rigid bodies. Thus, it is enough to use minimally 3 atoms (that are not on the same line) to determine the position and orientation of the molecular fragment. However, it is recommended to use at least 5-6 atoms to align molecular fragments.

EXSTATE_FILE – is an excited state parameter file (see a description below) that contains the information about excited states of the given molecular fragment (e.g. FRET donor or acceptor groups).

EXSTATE - number that identifies the excited state (e.g. 1st excited state) that will be used in the FRET modeling.

For example:

[FRAGMENT1]

#BODIPY

NAME = BDP

ID = 1

ATOMLIST = 1, 2, 4, 13, 20

ATOMTRANS = 1, 2, 4, 13, 20

EXSTATE_FILE = ./BDP_2b_4b_ac.txt

EXSTATE = 1

In the example above a PyFREC fragment #1 is defined with the residue id #1 and name BDP in the PDB file. Atoms with PDB IDs: 1, 2, 4, 13, and 20 are used to define and transform the fragment. Excited state information for the fragment is provided in BDP_2b_4b_ac.txt file. The 1st excited state will be used for the simulations.

Excited State Parameter File

The excited state file name has to be specified with the **EXSTATE_FILE** keyword for each unique molecular fragment defined in the **[FRAGMENT…]** section in the main input file. The excited state file contains several sections described below.

Section **$GEOMETRY** specifies Cartesian coordinates (in Å) of atoms the number and order of atoms has to exactly match those in the definition of the fragment specified with **ATOMLIST and ATOMTRANS.** Initial coordinates of fragments can be automatically generated with the **JOB = MAKE_FRAGMENTS**. These coordinated are usually obtained from the output of the electronic structure package (e.g. Gaussian) calculation of the excited states. The main purpose of these coordinates is to define the coordinate system in which the transition dipole vector is defined (see **$EXCITED_STATES** section below). Sample $GEOMETRY section is shown below:

$GEOMETRY

B 2.410545 0.000031 0.008008

C -0.596064 -0.000024 0.001390

N 1.486893 -1.241495 -0.000228

F 3.228983 -0.000280 -1.127243

C -2.084312 -0.000034 0.001327

$END

$EXCITED_STATES section contains the information about excited states of the molecular fragment and usually filled out based on calculations with the general-purpose electronic structure packages. Each line of the section contains the following information:

1. Excited state number
2. Absorption maximum, cm-1 (optional if tabulated absorption spectrum is provided)*

3-5. Transition Dipole Moments x,y, and z components in atomic units (a.u.)

6. Emission maximum, cm-1 (optional if tabulated emission spectrum is provided)*
7. Molar absorption (extinction) coefficient at the absorption maximum in M-1 cm-1*
8. Quantum yield fluorescence of the fragment (donor), unitless*
9. Lifetime of the fragment (donor) excited state in the absence of the acceptor in seconds*

Parameters marked with the asterisk (*) should be set to zero if unknown.

Please note that absorption and emission maxima and their broadening are used only for Gaussian overlaps.

Sample $EXCITED_STATES section is shown below:

$EXCITED_STATES

1 23095.69 0.0 -0.0001 3.0700 0.0001 0.0 0.0 56000.0 0.6 4.3e-9

$END

The section contains description of a single (1st) excited state with the excited state (maximum of absorption) energy: 23095.69 cm^{-1} and 0 cm^{-1} FWHM (spectral broadening of absorption is unknown), transition dipole moment x, y, and z components: -0.0001, 3.0700, 0.0001 in a. u. Emission maximum and broadening: 0.0 cm^{-1} and 0.0 cm^{-1} are also unknown. The molar absorption (extinction): 56000.0 M^{-1} cm^{-1}. Quantum yield of fluorescence: 0.6; Lifetime of the fragment (donor) excited state in the absence of the acceptor: $4.3x10^{-9}$ s. Usually the excited state (absorption) energy and components of the transition dipole moment are computed with the electronic structure methods. While maximum of fluorescence, spectral broadenings, molar absorption, and quantum yield are obtained experimentally. Fluorescence lifetime may be specified is known or computed with PyFREC with the **SB** keyword (see above).

It is recommended instead of specifying spectral broadenings of absorption and emission spectra as

FWHM values above provide tabulated spectra.

$ABS_SPEC section contains tabulated absorption spectrum on the FRET acceptor

$EMS_SPEC section contains tabulated emission (fluorescence) spectrum on the FRET donor

For example:

$ABS_SPEC 1 485 605

./tetrazine_abs.txt

$END

Tabulated absorption spectrum for the excited state 1 from 485 nm through 605 nm is stored in ./tetrazine_abs.txt file

$EMS_SPEC 1 506 830

./BODIPY_ems.txt

$END

$CENTER section defines (x, y, and z) coordinates in Å of the center of the molecular fragment. The donor-acceptor distance in FRET simulations is defined as distance between centers of the fragments. If coordinates in **$GEOMETRY** section are obtained form electronic structure calculations in Gaussian then origin of the coordinates (0,0,0) corresponds to the origin of the standard orientation – i.e. center of the electric charge. Alternatively, any point can be specified as a center of the fragment. It is not uncommon to define a center of the ring (e.g. phenyl group) or position of a central metal ion (e.g. Mg^{2+} in Chlorophyll) as a center of the fragment. By default, the center is defined as the origin of the coordinate system:

$CENTER

0.0 0.0 0.0

$END