# Development and Implementation of Computational Methods for Simulation of Excitation Energy Transfer in Complex Molecular Systems with PyFREC

Yana Kosenkov<sup>1</sup> and Dmitri Kosenkov<sup>1</sup>

<sup>1</sup>Monmouth University

July 1, 2020

## Abstract

Excitation energy transfer is a ubiquitous process of fundamental importance for understanding natural phenomena, such as photosynthesis, as well as advancing technologies ranging from photovoltaics to development of photosensitizers and fluorescent labels. This work provides an overview of recent advancements in excitation energy transfer modeling with the PyFREC software package. Computational methods currently implemented in PyFREC include molecular fragmentation techniques, as well as methods for electronic coupling computations, analysis of coupled electronic excited states, and quantum dynamics simulations. Advanced functionality and possible vectors for future development of the package are also explored.

# Introduction

PyFREC has been recently used for modeling the excitation energy transfer in various molecular systems of biological and technological interest.<sup>1-4</sup> Excitation energy transfer (EET) is crucial for understanding of a wide range of phenomena including photosynthetic processes and fluorescence, as well as for development of new technologies for photovoltaics, etc.<sup>5-13</sup> A substantial progress in quantum chemical calculations based on the exciton model of excited-state properties of weakly interacting fragments (e.g. molecular aggregates or liquids) has been achieved.39 In PyFREC, excitation energy transfer modeling is based on the analysis of three-dimensional structures of donor and acceptor molecules, and properties of their electronic excited states. The approach accounts for the presence of molecular vibrations that are included in quantum dynamics simulations. This computational methodopogy employs a fragmentation technique where the properties of the donor and acceptor molecules are computed separately, and subsequently used to deduce the properties of larger donor-acceptor complexes. Such an approach helps reduce computational costs associated with modeling of complex molecular systems that contain multiple donors and acceptors, such as light-harvesting complexes.<sup>2</sup>, <sup>3</sup>

The proposed method includes the following steps. Initially, the molecular geometries of the donor and acceptor molecules are optimized, and the properties of their electronic excited states are computed with widely used electronic structure packages (e.g., GAMESS,<sup>14</sup> Gaussian,<sup>15</sup> etc.) Then, if necessary, vibrational spectra and Huang-Rhys factors are computed to account for electron-vibrational coupling. This information is further used as an input for the PyFREC software.<sup>3</sup> PyFREC performs alignment of molecular fragments (e.g., DNA bases, protein residues, photosynthetic pigments, etc.) in order to reconstruct complex molecular structures. For example, photosynthetic pigments – bilins – are treated as molecular fragments to reconstruct the structure of phycobiliprotein – a light-harvesting complex (Figure 1) – with the alignment procedure.<sup>3</sup> PyFREC enables computing electronic couplings between pairs of pigments. These electronic couplings

are used to model the exciton energy transfer in the molecular complexes.<sup>2,3</sup> This computational approach is versatile and, therefore, can be integrated with multiple density functional theory (DFT) methods for computation of electronic excited states. It also allows for integration of computed or empirical properties of the donor and acceptor molecules. For example, excitation energies and transition dipole moments used in EET modeling of energy transfer can be either computed or measured spectroscopically. There are several method choices for computing the rates of energy transfer: Förster theory based on spectral overlap of empirical donor emission and acceptor absorption spectra, as well as quantum dynamics methods. The quantum dynamics methods implemented in PyFREC are based on the quantum master equation formalism. The software has been successfully used to model electronic couplings in complexes of organic molecules,<sup>1</sup> to model EET in the Fenna-Matthews-Olson complex,<sup>2</sup>phycobiliprotein,<sup>3</sup> and halogentated bioorthogonal boron dipyrromethene photosensitizers.<sup>4</sup> The software also has options for visualization of the energy flow in quantum dynamics simulations.<sup>3</sup> In the following sections of the paper, the details of the computational method and associated software architecture features are discussed.

A description of the procedure for structural alignment of molecular fragments is provided below. An analysis of electronic excited states through identification of resonances between uncoupled excited states of molecular fragments used for initial assessment of Förster resonance energy transfer (FRET) modeling<sup>1,2</sup> is also provided. Modeling of spectral overlaps of empirical emission and absorption spectra in accordance with the Förster theory<sup>16,17</sup> and a procedure for electronic coupling calculations that includes analysis of mutual orientations of transition dipole moments of fragments are described. Analysis of coupled electronic excited states with the variation method is discussed. Once a computational model of coupled states is obtained, quantum dynamics methods can be used as implemented in PyFREC. This quantum dynamics method accounts for the impact of molecular vibrations on the energy transfer via electronic-vibrational coupling and Huang-Rhys factors.<sup>18</sup> The software architecture and elements of the user input, as well as available interfaces to electronic structure packages and formats of structural information (e.g., PDB databank files)<sup>19</sup> are briefly described. Finally, the vectors for future development of computational methods and software that includes visualization, PDB database scanning, and network analysis of electronic couplings are briefly discussed.





# Methods

This section provides an overview of computational methods implemented in the PyFREC software. A detailed description of the methods is provided elsewhere.<sup>1-3</sup>

# Alignment of Molecular Fragments

An analysis of the user-provided molecular structure of the fragment (e.g., the photosensitizer molecule shown in Figure 2)<sup>4</sup> starts from the alignment procedure.

The software allows a user to specify a desired fragmentation scheme. For example, in Figure 2, the fragmentation scheme is as follows: methyl-tetrzine (m Tz) is the EET acceptor, while halogentated bioorthogonal boron dipyrromethene (BODIPY-2I) is the EET donor. Those are selected as separate fragments, accordingly.<sup>4</sup>

PyFREC utilizes the *Procrustes* analysis to perform alignment.<sup>1, 2, 20</sup> In this procedure, a translation vector  $\mathbf{a}$  and a rotation matrix  $\mathbf{R}$  are found so that the root-mean-square deviation between transformed Cartesian

coordinates of the molecular fragments  $\mathbf{F}'$  (Figure 2B,C) and corresponding coordinates of the molecular system S (Figure 2A) is minimized:

$$\|(\mathbf{FR} + \mathbf{a}) - \mathbf{S}\| \longrightarrow min(1)$$

where F are initial coordinates of the molecular fragment. Then the singular value decomposition of the matrix  $\frac{\mathbf{S}^{t}\mathbf{F}}{\|\mathbf{S}\|\|\mathbf{F}\|}$  of normalized coordinates of the fragment and corresponding atoms of the molecular system is used to find the rotation matrix R :

$$\frac{\mathbf{S}^{\mathbf{t}}\mathbf{F}}{\|\mathbf{S}\|\|\mathbf{F}\|} = \mathrm{Ub}\mathbf{V}^{\mathbf{t}}(2)$$

where U and V are unitary matrices, b is a diagonal matrix. Then the rotation matrix is:

$$\mathbf{R} = \mathbf{V}\mathbf{U}^{\mathbf{t}}$$
 (3)

The translation vector is obtained as the difference between the centroids of the molecular system and rotated fragment:

 $\mathbf{a} = |\mathbf{S}| - |\mathbf{F}| \,\mathbf{R}(4)$ 

The PyFREC software automatically computes metrics for analysis and quality assessment of the molecular fragment alignment. Two sections of output are provided. The "transformation section" provides the scaling factor, translation vector, rotation matrix, and axis–angle representation of the rotation.

PyFREC provides the sum and root-mean-square deviation between the transformed atoms of the molecular fragment and those of the molecular system in the output.



Figure 2. PyFREC performs automatic alignment of molecular fragments. Example: (A) the photosen-

sitizer *m* Tz-2I-BODIPY is split into: (B) methyl-tetrazine (*m* Tz, EET acceptor) and (C) halogentated bioorthogonal boron dipyrromethene (BODIPY-2I, EET donor).<sup>4</sup> Hydrogen atoms are not shown for clarity.

Output is generated for all atoms or a selected group of atoms used for alignment of the molecular fragment. Another example involves the alignment procedure performed on a bacteriochlorophyll molecule, where the alignment has been performed on a rigid porphyrin molecy while the flexible tail of the bacteriochlorophyll molecule was not aligned.<sup>2</sup>

The order of atoms in the reference structure and in the molecular fragment has to be the same. It is achieved by an automatic generation of molecular fragments based on the reference structures with the auxiliary script.

Identification of Resonances of Excited States and Spectral Overlaps

The next step in user input processing is the analysis of excited states. PyFREC reads files with the excited states of each fragment and transforms the transition dipole vectors using translation vectors and rotation matrices described above. As each fragment may contain multiple electronic excited states and the molecular systems may contain multiple pigments (e.g., seven or eight bacteriochlorophyll molecules in the Fenna-Matthews-Olson complex),<sup>21, 22</sup> PyFREC has a special job type "survey" that surveys (screens) all electronic excited states of all fragments provided in the input in order to identify resonance states. The default resonance condition for states D and A with excited state energies  $\nu_D$  and  $\nu_A$ , respectively, is

 $|\nu_D - \nu_A| \le \omega_r \ (5)$ 

where  $\omega_r$  – is the resonance threshold with the default value of 1000 cm<sup>-1</sup> that can be changed by the user. As multiple factors determine broadening of spectral lines, the resonance condition above is used only for inspection of potential resonances. Alternatively, the resonance condition can be determined based on the threshold value of spectral overlap (see below). In order to compute excitation energy transfer rates (e.g., with the Förster theory, see below) the spectral overlap ( $J_{\text{DA}}$ ) is computed.<sup>8, 17, 23</sup>

$$J_{\rm DA} = \frac{1}{N_f N_a} \int_0^\infty f_D(\tilde{\nu}) a_A(\tilde{\nu}) \tilde{\nu}^{-4} d\tilde{\nu}(6)$$

where  $f_D(\tilde{\nu})$  and  $a_A(\tilde{\nu})$  are the area-normalized fluorescence and absorption line shapes, respectively, and  $N_f = \int_0^\infty f_D(\tilde{\nu}) \ \tilde{\nu}^{-3} d\tilde{\nu}$  and  $N_a = \int_0^\infty a_A(\tilde{\nu}) \tilde{\nu}^{-1} d\tilde{\nu}$  are the normalization factors. In PyFREC, Gaussian line shapes are used by default.

In PyFREC, the calculation of spectral overlaps is based on the Gaussian lineshapes approximation by default. The user provides positions and widths of absorption and emission (fluoresce) spectra of a part of the input. Properties of the excited states are either computed with general purpose electronic structure packages (e.g., Gaussian16) or from empirically based on spectroscopic observations.

### **Electronic Couplings**

Once electronic excited states of the fragments are selected, the electronic couplings are computed. The model is based on point-dipole approximation with a linear electrostatic screening factor (s):

$$V_{\rm ij} = \rm s \ V_{\rm ij}^0 \ (7)$$

Alternatively, exponentially attenuated transition dipole moments can be used in PyFREC:<sup>2, 3</sup>

$$V_{ij} = V_{ij}^0 (Aexp(-\beta P) + s_0)(8)$$

where  $V_{ij}^0$  is the point-dipole electronic coupling, and A,  $\beta$ , and  $s_0$  are parameters provided in the input.<sup>24, 25</sup> As the anisotropy of the protein environment the screening may depend on the orientation of the fragments<sup>38</sup> the proposed approximation has to be used with caution. The Förster coupling<sup>16, 17</sup> can be split into distanceand orientation-dependent parts:<sup>1-3</sup>

$$V_{ij}^{0} = \frac{1}{R^{3}} |\mu_{i}| |\mu_{j}| K_{ij}(9)$$

where R is the distance between centroids of fragments,  $|\mu_i|$ ,  $|\mu_j|$  are magnitudes of transition dipole moments, and  $-1 \le K_{ij}/2 \le 1$  is the orientation factor which depends only on mutual orientations of transition dipole moments. The centroids of fragments are defined as centroids of the electric charge of the ground state electronic density (origin in the standard orientation of Gaussian software).<sup>15</sup>



Figure 3. The methyl-tetrazine moiety is rotated by  $\sim 90^{\circ}$  to the plane perpendicular to the phenyl group.

However, if desired, the user may specify Cartesian coordinates of any point in space that is a relevant center of the fragments (e.g., the center of the porphyrin ring of a bacteriochlorophyll fragment). The orientation-dependent part can be used to analyze the quality of dipole alignment of molecular fragments.

The software can be used as a tool for time-efficient screening calculations. However, the user has to be aware of the screened point-dipole approximation limitations to ensure that the given method is valid for systems of their choice. For systems where non-dipole contributions are significant, the proposed approximation may be insufficient. Alternatively, other methods beyond those implemented in PyFREC should be used instead.<sup>6</sup>, <sup>24</sup>, <sup>25</sup>

Förster Energy Transfer Rates

As follows from the Forster theory<sup>16, 17</sup> electronic couplings  $(V_{\text{DA}})$  and spectral overlaps  $(J_{\text{DA}})$  are used to calculate the resonance excitation energy transfer rate  $(k_{\text{DA}})$ :<sup>2, 6</sup>

$$k_{\rm DA} = \frac{2\pi}{\kappa} V_{\rm DA}^2 J_{\rm DA}(10)$$

The simplicity and convenience of this approach have made it a popular choice in cases where empirical emission and absorption spectra of donors and acceptors are available. Unfortunately, this expression does not provide details of the dynamics of coherent energy transfer. Therefore, quantum dynamics methodologies should be used instead. An example of application of the Förster approach is computation of EET rates in the photosensitizer m Tz-2I-BODIPY (Figure 2) that was proposed for conditional singlet oxygen generation in cells.<sup>4</sup>

In order to illustrate the effect of mutual donor-acceptor orientation change on the electronic coupling and EET rate, two geometries of the the photosensitizer are considered (Figure 3).

Table 1. Elements of PyFREC input	: components $(\mu_x, \mu_y, \mu_z)$	and magnitudes $ \mu $ c	of transition dipole moments	of lowest size
-----------------------------------	--------------------------------------	--------------------------	------------------------------	----------------

Fragment				
BODIPY-2I				
mTz				

Table 1 depicts elements of user input: transition dipole moments of the molecular fragments used by PyFREC for calculation of electric couplings with Eq. 9. Calculation of the electrostatic screening factors<sup>1-3, 6, 23, 24</sup> for this calculation is approximated as:  $s=1/n^2=0.56$  where n=1.34 is the refractive index

of acetonitrile.<sup>26</sup> The spectral overlap  $(J=3.14\times10^{-4} \text{ cm})$  is based on the empirical spectra.<sup>4</sup>



Figure 4. Example: the variation method is used to determine energies of coupled excited states ( $\varepsilon_{1-4}$  and  $\varepsilon^*_{1-4}$ ) of the Fenna-Matthews-Olson light-harvesting complex (shown in shown in the inset). The method helps to determine major contributions of uncoupled localized excitations of bacteriochlorophyll molecules (BCL 401-407) to the coupled states. Reproduced from Ref. 2.

Table 2.	Table 2.	Table 2.	Table 2.	Table 2.	Table 2.
Interfragment	Interfragment	Interfragment	Interfragment	Interfragment	Interfragment
distances $(R)$ ,					
magnitudes of					
the orientation					
factors	factors	factors	factors	factors	factors
$ \Theta  =  \mathbf{K} /2,$					
magnitudes of					
unscreened	unscreened	unscreened	unscreened	unscreened	unscreened
$( V^0 )$ and					
screened	screened	screened	screened	screened	screened
electronic	electronic	electronic	electronic	electronic	electronic
couplings $( V )$					
and EET rates					
(k) between					
fragments of					
mTz-21-	<i>m</i> Tz-21-				
BODIPY are					
shown.	shown.	shown.	shown.	shown.	shown.
Fragment	$R,  {A}$	$ \Theta ,~\%$	$ V^0 , \ cm^{-1}$	V , cm-1	k, ps <sup>-1</sup>
Orientation					
Planar <sup>[Ref. 4]</sup>	6.83	48.97	106.60	59.44	1.31
Perpendicular <sup>[a]</sup>	6.95	0.06	5.81	3.24	$3.90 \mathrm{x} 10^{-3}$
<sup>[a]</sup> See Figure 3					

Results from Table 2 illustrate that while interfragment distance has not significantly changed (6.89-6.95 Å), the orientation factor has decreased from ~49% to 0.06 % which led to significant reduction of the screened electronic coupling to 3.24 cm<sup>-1</sup> which, in turn, reduced the EET rate by a factor of ~ 336 to  $3.9 \times 10^{-3} \text{ps}^{-1}$ . Thus, effects of mutual ordinations of donor and acceptor fragments can be explored with PyFREC using the methods described above.

#### Variation Method

Electronic coupling between excited states leads to changes in their energies and the formation of coupled states. The variation method is a simple approach that enables calculation of the energies and structures of coupled states in the basis of localized (uncoupled) states:

#### $H_{\rm el}c_l = \varepsilon_l c_l \ (10)$

where  $H_{\rm el}$  is the Hamiltonian containing excitation energies of localized states (diagonal elements) and electronic couplings (off-diagonal elements),  $\varepsilon_l$  are energies of coupled states and  $c_l$  are eigenvectors that show contributions of localized excitations to each coupled excited state.<sup>1-3</sup> The variation method is based on Frenkel (or Frenkel-Davydov) exciton model.<sup>39,40</sup>

The excitation energies of uncoupled molecular fragments (diagonal elements of the Hamiltonian) are specified in the input file. The excitation energies can be obtained from quantum chemical computations or form experimental studies (see Refs. 6, 8 for review).

An example of the coupled states in the Fenna-Matthews-Olson  $\operatorname{complex}^2$  is shown in Figure 4



Quantum Dynamics Approaches

In PyFREC, quantum dynamics simulations are implemented with the quantum master equation formalism.<sup>3, 18, 27</sup> The Hamiltonian consists of a pure electronic part  $(H_{\rm el})$  that describes coupled electronic excited states, a vibrational part  $(H_{\rm vib})$  that includes molecular vibrations of fragments, and electron-vibrational couplings  $(H_{\rm el-vib})$ :

 $H = H_{\rm el} + H_{\rm vib} + H_{\rm el-vib} \ (11)$ 

Huang-Rhys factors are used to describe electron-vibrational coupling:

$$H_{\rm el-vib} = \sum_{i=1}^{N} \sum_{k=1}^{n(i)} \hbar \omega_{\rm ik} \sqrt{S_{\rm ik}} \left( a_{\rm ik}^{\dagger} + a_{\rm ik} \right) |i\rangle \langle i| (12)$$

where  $\omega_{ik}$  is the vibrational mode, the Huang-Rhys factors  $S_{ik}$ ,  $a_{ik}^{\dagger}$  and  $a_{ik}$  are standard raising and lowering operators, N is the number of fragments, n(i) is the number of vibrational modes coupled to the electronic state i. A Lindblad-type quantum master equation is written as:

$$\frac{\mathrm{d}}{\mathrm{dt}} = -\frac{i}{\hbar} \left[ H, \rho \right] + \mathcal{L}_{\mathrm{deph}} \left( \rho \right) + \mathcal{L}_{\mathrm{vib}} \left( \rho \right) (13)$$

where  $\rho$  is the density matrix,  $\mathcal{L}_{deph}$  is the electronic dephasing operator (to describe interactions with the environment) and  $\mathcal{L}_{vib}$  is the vibrational damping operator. Sample quantum dynamics of the fully coherent regime determined by the  $\frac{i}{\hbar}[H,\rho]$  term in Eq. 13 is shown in Figure 5. The dynamics that include all terms from Eq. 13 is presented in Figure 6.

The parameters such as electronic decoherence rate and vibrational damping rates<sup>3</sup> in the Lindblad equations are entered as parameters. The user is free to choose desirable parameters based on other calculations or empirical (spectroscopic) findings as a part of the user input. The user may either choose quantum master equation or Förster theory calculations.

## Software Architecture

A general view of the PyFREC architecture is shown in Scheme 1. The software consists of several modules that provide reading and initial processing of input data: configuration manager, calculation manager, excited states reader, reader of molecular structures, and a module that performs alignment of molecular fragments.



Once the information on excited states and molecular structures is received, PyFREC identifies potential resonances between electronic excited states of fragments, calculates spectral overlaps, and computes electronic couplings. The coupling calculation module calls the module that computes electrostatic screening factors. Rates of excitation energy transfer are further computed using the Förster theory. The variation method is used to model coupled electronic excited states. A separate module performs quantum dynamics simulations that are based on the master equation formalism and account for molecular vibrations. There also

are shared auxiliary library modules that store constants, conversion factors, standard data structures (e.g., atoms, Cartesian coordinates, etc.) and routines for data exchange among other modules. While many common standard Python libraries are used by PyFREC (e.g. os, sys, re, etc.),<sup>28</sup> the computationally intensive routines (matrix algebra, integration of differential equations, etc.) rely on NumPy and SciPy libraries.<sup>29, 30</sup>

# Elements of the User Input

The philosophy of the user input assumes maximal reuse of the provided information. Therefore, the user defines properties of each unique molecular fragment only once, and then those properties are reused by PyFREC. The PyFREC is a command line tool (all input and output are text files). The user provides definitions of molecular fragments and parameters of the simulation (e.g. Förster simulations or quantum dynamics, etc.) in the main user input file (Scheme 2). The main user input file (Scheme 2) contains the following sections: "Methods", "Molecular System", followed by the definitions of molecular fragments. The section "Methods" defines the type of calculations to be requested (e.g., calculation of electronic couplings or quantum dynamics simulations). The "Molecular System" section defines general parameters of the entire system (e.g., a PDB file with the complete protein complex and type and parameters of the electrostatic screening model used in the requested calculation). The next section, "Quantum Dynamics", provides the initial conditions and numerical integration parameters of the simulation. In addition, properties of each molecular fragment (e.g. excitation energies and transition dipole moments) are provided in a separate input files for each fragment (Scheme 2). These properties of the fragment are usually computed with the general purpose electronic structure packages (e.g. Gaussian). Auxiliary scripts can be used to convert electronic structure package outputs to PyFREC input format.



Scheme 1 Architecture of PyFREC software (see explanation in text).



The definition includes: name and identification "serial" number of the fragment (numbers based on the

residue ID numbers from the PDB file), and a list of atoms from the PDB file that belong to the fragment. This enables the most flexible way to split any PDB file into several fragments by selecting specific atoms from each residue included into the simulation.

The section provides information about vibrational modes coupled to the electronic excited state. The section contains: the vibrational mode number and its wavenumber, electron-vibrational coupling, and the vibrational decay rate. These properties of electronic excited states can be obtained from quantum chemical calculations or empirically. Thus, PyFREC provides a flexible way of integrating computed and experimental data into a single simulation.

## Advanced functionality and future development

Advanced functionality features of PyFREC geared toward expert users and prospective developers. The discussed features are currently under development. The functions described below are intended as auxiliary routines to build novel units of PyFREC.

# 3D Visualization of Quantum Dynamics

PyFREC has a module for 3D visualization of populations of excited states of quantum dynamics trajectory. The PyFREC visualization module takes CUBE files<sup>31</sup> with densities of localized excited states of separate fragments and the density matrix obtained from the quantum dynamics simulation and generates JMOL script for visualization.<sup>32</sup> First JMOL loads the PDB file with coordinates of the molecular complex. Then, CUBE files for all excited states are loaded and the transparency of the excited state surface is based on populations of the excited state (unpopulated states are entirely transparent, completely populated states are totally opaque). Transparency of orbitals changes as populations of the excited states evolve during the dynamics simulations (See supporting information for a sample visualization.)

# Conversions of Quantum Dynamics Trajectory into Audio files

An alternative way to perceive quantum dynamics simulations is generation of sound-based density matrix oscillations of populations of excited states (implemented in the "qd" module). This may be of help for outreach activities and for visually impaired users. *PyAudio* <sup>33</sup> and *SciPy* <sup>34</sup>libraries are used. The user specifies a particular excited state. Then, excited state population changes along the trajectory are automatically slowed down to the audible frequency range (20- 20,000 Hz). For example, a trajectory of 50 ps is slowed down by a factor of  $2\times10^{10}$  and converted into 1 s of a sound track. Moreover, nearly resonant energy levels of pigments are audibly distinguishable as they lead to low frequency beats (See supporting information for samples.)

## Automatic Scanning of the PDB Databank

As the number of high-throughput computational methods increases, and PyFREC provides means for quick screening of excited state resonances, electronic couplings, and quantum dynamics simulations, it is convenient to have a tool for automatic extraction of structural information. The Protein Data Bank (PDB)<sup>19</sup> provides a convenient interface for such operations. Employing *urllib2*, <sup>35</sup> PyFREC automatically downloads and parses PDB files based on a user-provided PDB ID list. PyFREC then analyzes the downloaded PDB structures (e.g., identifies chlorophyll pigments inside PDB files) in order to compute electronic couplings between the selected fragments. Currently, the identification of pigments is based on chemical structure and topology of chemical bonds (e.g., the central Mg atom surrounded by nitrogen and oxygen atoms at particular distances).<sup>2</sup> In the future, machine-learning algorithms (see below) will be used for this analysis.

Processing of multiple molecular structures (e.g., proteins from the PDB databank) produces datasets that can be interpreted and analyzed using the network (graph) theory. For example, electronic couplings or orientation factors that characterize interactions between pigments and affect the exciton energy transfer can be rationalized in terms of network theory. PyFREC employs *NetworkX*library<sup>36</sup> to generate and analyze networks. Various properties of the network are computed, including average shortest path length, average clustering coefficient, and current-flow closeness centrality.

### Conclusions

PyFREC software provides a versatile tool for modeling excitation energy transfer in such diverse systems as light-harvesting protein complexes, fluorescents labels, and photosensitizers. The software provides alignment of molecular fragments, calculation of electronic couplings and orientation factors followed by calculations of spectral overlaps and Förster energy transfer rates. The variation method can be additionally used to analyze coupled electronic excited states. Quantum dynamics is implemented with the quantum master equation approach that provides a prediction of density matrix dynamics including populations of the electronic excited states and may be coupled to molecular vibrations. Finally, a set of additional modules provides 3D visualization, generation of audio files, PDB databank scanning, and network topology analysis functionality. Future development of PyFREC will include adding non-dipole interactions for calculations of electronic couplings in order to account for triplet excited states, since quantum dynamics of photophysical processes (fluorescence quenching, phosphorescence) proceeds with involvement of triplet states. It is also planned to implement deep learning algorithms for automation of multiple routines in PyFREC, such as recognition of molecular fragments inside proteins in PDB files, prediction of excitation energies, and electronic couplings of molecular fragments based on molecular structure (coordinates) of molecular systems that contains fragments (e.g., pigments inside a light-harvesting protein).

Molecular Education and Research Consortium in Undergraduate computational chemistRY (MERCURY)<sup>37</sup> provides dynamic and supportive environment for undergraduate students and faculty involved in our research projects.<sup>1-3</sup> The consortium promotes the development of PyFREC and helps students to gain experience in modern computational quantum chemistry.

# **Funding Information**

The project was funded in part by the Research Corporation for Science Advancement through a Cottrell Scholar Award. Acknowledgment is also made to the Donors of the American Chemical Society Petroleum Research Fund for partial support of this research though grant #58019-UR6 and National Science Foundation MRI #1662030.

## **Research Resources**

Computational resources were provided in part by the MERCURY consortium.

Keywords: Förster, FRET, Exciton, Quantum Dynamics

Additional Supporting Information may be found in the online version of this article.

### **References and Notes**

- 1. D. Kosenkov, J. Comp. Chem. 2016, 37, 1847-1854.
- Y. Kholod, M. DeFilippo, B. Reed, D. Valdez, G. Gillan, and D. Kosenkov, J. Comp. Chem. 2018, 39, 438-449.
- 3. Y. K. Kosenkov and D. Kosenkov, J. Chem. Phys. 2019, 151, 144101.
- G. Linden, L. Zhang, F. Pieck, U. Linne, D. Kosenkov, R. Tonner, and O. Vazquez, Angew. Chem. Int. Ed. 2019, 58, 12868-12873.
- R. E. Blankenship, D. M. Tiede, J. Barber, G. W. Brudvig, G. Fleming, M. Ghirardi, M. R. Gunner, W. Junge, D. M. Kramer, A. Melis, T. A. Moore, C. C. Moser, D. G. Nocera, A. J. Nozik, D. R. Ort, W. W. Parson, R. C. Prince, R. T. Sayre, *Science*, **2011**, 332, 805-809.
- 6. C. Curutchet, B. Mennucci, Chem. Rev.  $\mathbf{2017}$  , 117, 294–343.
- 7. G. D. Scholes, G. R. Fleming, A. Olaya-Castro, R. van Grondelle, Nat. Chem. 2011, 3, 763–774.
- 8. A. Chenu, G. D. Scholes, Annu. Rev. Phys. Chem. 2015, 66, 69-96.
- 9. O. G. Reid, G. Rumbles, J. Phys. Chem. C , 2016, 120, 87–97.
- G. S. Engel, T. R. Calhoun, E. L. Read, T.-K. Ahn, T. Mančal, Y.-C. Cheng, R. E. Blankenship, G. R. Fleming, *Nature*, 2007, 446, 782–786.
- 11. G. Li, R. Zhu, Y. Yang Nat. Photonics, 2012, 6, 153-161.

- M. Kaltenbrunner, M. S.White, E. D. Głowacki, T. Sekitani, T. Someya, N. S. Sariciftci, S. Bauer, Nat. Commun. 2012, 770.
- 13. M. B. Schubert, J. H.Werner, Mater. Today, 2006, 9, 42–50.
- M.S. Gordon, M.W. Schmidt, In Theory and Applications of Computational Chemistry: the first forty years. C.E. Dykstra, G. Frenking, K.S. Kim, G.E. Scuseria, (Eds.), Elsevier: Amsterdam, 2005, 1167-1189.
- Gaussian 16, Revision B.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, **2016**.
- 16. T. Forster, Naturwissenschaften, 1946, 33, 166-175.
- 17. Th. Förster, *Delocalized Excitation and Excitation Transfer*, **1965**, Florida State University, Tallahassee, Florida.
- 18. E. K. Irish, R. Gómez-Bombarelli, B. W. Lovett, Phys. Rev. A ,2014, 90, 012510.
- 19. RCSB Protein Data Bank https://www.rcsb.org/ (Accesses Apr 8, 2020)
- 20. J. C. Gower, G. B. Dijksterhuis, Procrustes Problems; Oxford University Press: Oxford, 2004 .
- J. Wen, J. Harada, K. Buyle, K. Yuan, H. Tamiaki, H. Oh-oka, R.A. Loomis, R.E. Blankenship, Biochemistry, 2010, 49, 5455-63.
- 22. E. L. Read, G. S. Schlau-Cohen, G.S. Engel, J. Wen, R.E. Blankenship, G.R. Fleming, *Biophys. J.* 2008, 95, 847-856.
- 23. T. Renger, Photosynth. Res. 2009, 102, 471-485.
- 24. C. Curutchet, G. D. Scholes, B. Mennucci, R. Cammi, J. Phys. Chem. B. 2007, 111, 13253-13265.
- 25. G. D. Scholes, Annu. Rev. Phys. Chem. 2003, 54, 57-87.
- 26. O'Neil, M.J.(ed.). TheMerck Index -AnEncyclopedia ofChemicals. Drugs, andBiologicals. Cambridge, UK: Royal Society ofChemistry, 2013., 14 p.https://pubchem.ncbi.nlm.nih.gov/compound/acetonitrile#section=Spectral-Properties (Accessed) and acetonitrile#section=Spectral-Properties (Apr 8, 2020)
- 27. H-P. Breuer, F. Petruccione, The Theory of Open Quantum Systems; Oxford Univ. Press: Oxford 2002
- 28. The Python Standard Library https://docs.python.org/2.7/library/ (Accessed Apr 8, 2020)
- 29. NumPy https://numpy.org/ (Accessed Apr 8, 2020)
- 30. SciPy https://www.scipy.org/ (Accessed Apr 8, 2020)
- 31. cubegen https://gaussian.com/cubegen/ (Accessed Apr 8, 2020
- Jmol: an open-source Java viewer for chemical structures in 3D http://jmol.sourceforge.net/ (Accessed Apr 8, 2020)
- 33. PyAudio 0.2.11 https://pypi.org/project/PyAudio/ (Accessed Apr 8, 2020)
- 34. P. Virtanen, R. Gommers, Nat. Methods , 2020 , 17, 261–272.
- urllib2 extensible library for opening URL https://docs.python.org/2/library/urllib2.html#moduleurllib2 (Accessed Apr 8, 2020)
- 36. NetworkX Software for complex networks https://networkx.github.io/ (Accessed Apr 8, 2020)
- 37. Shields, G. C., Twenty Years of Exceptional Success: The Molecular Education and Research Consortium in Undergraduate computational chemistRY (MERCURY). Int. J. Quantum Chem. 2020, under review.
- 38. C.-P. Hsu. J. Chem. Phys. 2001, 114, 3065.

39. J. Frenkel, J. Phys. Rev. **1931**, 37, 17-44.

40. A. Davydov, Theory of Molecular Excitons, 1971, Springer US: New York, 1971 .