JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Review

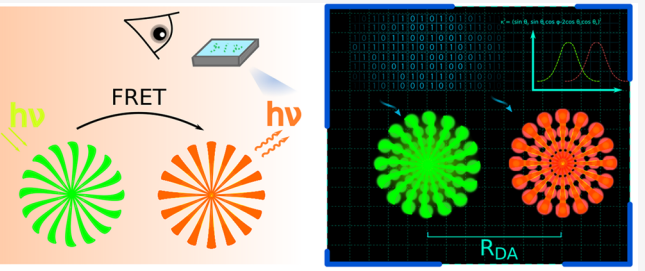
Recent Efforts and Milestones for Simulating Nucleic Acid FRET Experiments through Computational Methods

Sorin Muraru, Sebastian Muraru, Florentin Romeo Nitu, and Mariana Ionita*

Cite This: J. Chem. Inf. Model. 2022, 62, 232−239 Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Computational methods can greatly aid nucleic acid fluorescence experiments by either offering fully detailed atomic insights into the conformations and interactions present in the studied system or by providing accurate simulations of the fundamental parameters. Fluorescence-based optical biosensors show great potential for clinical diagnosis of life-altering diseases with a very high specificity. Many of the designs for such rely on the concept of Förster resonance energy transfer (FRET). Currently, the methods used experimentally make use of theoretical assumptions which fundamentally affect the results. Having a detailed atomistic overview or significant simulated parameters could improve the understanding of the calculations and provide much more accurate outcomes. However, there are many challenges that need to be addressed before standardized computational protocols can be employed. This review is meant to highlight the progress made for computational methods used to simulate FRET experiments for nucleic acid probes. Recent advances have been made in computational tools, such as force field parametrizations and improved protocols. Complementary simulations to experimental data are also comprised in the this review.



KEYWORDS: *fluorescence quenching, resonance energy transfer, optical biosensor, computational methods, molecular dynamics, Monte Carlo simulations, nucleic acid probes, molecular beacons*

INTRODUCTION

Early diagnosis of diseases may prove to be the decisive element in the battle that many patients have to endure when faced with a life-threatening condition. Biosensors are versatile detection devices that can provide useful clarity in such situations.[1] They are already functional in many fields and have established a reputation of having promising potential in a wide range of applications, from clinical diagnosis to food safety testing.[2] Specifically, optical biosensors can provide quick or real-time, simple, inexpensive, high specificity, and high sensitivity detection tests.[3] They rely on three components, namely, a recognition unit, a transducer, and a signal processor.[4] Accordingly, they are able to recognize biological elements and provide a measurable optical signal based on the detected parameters. Although there are many types of optical biosensors, two different versions are used to facilitate all the detection algorithms: label-free and label-based probe design.[5] Generally, the latter are achieved through chemical or enzymatic procedures.[6] Fluorescence-based optical biosensors have gained a lot of attention lately due to their simplicity and the technological progress in the area of new dyes and fluorescent probes.

Fluorescence-based optical biosensors rely on fluorophores that emit light following the boost to an excited energetic state. More specifically, fluorescent molecules absorb energy from a light source, leading them to an unstable excited state, from which they almost instantly return to the ground state. In this process, they may either transfer the energy to other molecules or emit light at a longer wavelengths than the one absorbed, usually both in different proportions.[7] Förster resonance energy transfer (FRET) refers to the phenomenon in which an excited donor fluorophore nonradiatively transfers energy to a ground state acceptor molecule through a dipole−dipole interaction.[8] The latter may also be a fluorophore, which ultimately emits the received energy at an even longer wavelength, or it may be a quencher, which is a nonfluorescent molecule that will emit the energy as heat. The biosensor is able to detect the intensity of the emitted light, and based on a set of different readings, certain molecular events can be quantified, which infer specific conclusions, essentially having sensed the investigated matter.

The detection of certain DNA or RNA strands has been a purpose of fluorescence-based optical sensors for a long time in

Received: August 7, 2021
Published: January 11, 2022

https://doi.org/10.1021/acs.jcim.1c00957
J. Chem. Inf. Model. 2022, 62, 232−239

the scientific community, as it can help diagnose diseases and other health conditions quickly and with high specificity. Accordingly, many probe configurations have been experimented with, and a wide collection has shown promising results.[9] Specifically, they are centered around the innate ability of the nucleic acids in DNA or RNA strands to hybridize with complementary targets. Usually, fluorophores are attached to the ends of such single strands, and an increase or decrease in the fluorescence intensity as detected by the biosensor may indicate successful hybridization, confirming the existence of the correct complementary strands. Consequently, they can provide precise diagnostics or other medical information. Lately, an increase in probe design and availability of new fluorophores has led to a rise in the interest for this field.[10]

At the moment, most such scientific investigations are performed experimentally, with computational studies carried out to a lesser extent. Lately, the use of computational methods for such experiments has grown, including Monte Carlo simulations, molecular dynamics, or even machine learning algorithms.[11] Although both areas have known an expansion lately in terms of feasibility, there are many challenges affecting both. Conveniently, one may complement the other in the course to rewarding biosensors. Explicitly, there is an obvious observation barrier at the atomic level that cannot be achieved experimentally; however, it can be simulated in silico. In turn, computational methods rely on experimental data for parametrizations and actual confirmation of their predictions.

This review aims to offer an overview on the current situation of computational methods used to simulate or complement fluorescence-based optical biosensors experiments focused on nucleic acids detection. More precisely, it highlights the challenges faced for the in silico experiments and provides a progress report on the most recent advances that have been made in the field. The required knowledge to understand the concept of FRET is provided, in addition to probe configurations and an overview of specific in silico considerations. This review is meant to help scientists interested in the opportunity of performing computational simulations for nucleic acids FRET experiments at the highest current standard.

FLUORESCENCE DETECTION

Fluorophores are molecules known to emit absorbed light at a longer wavelengths than captured.[12] This property is very useful in fluorescence microscopy for a variety of applications. Essentially, they undergo a process of excitation in which an electron is moved in to an orbital further apart from the nucleus than the one it was initially in, leading to a quick return to its ground state following the release of energy usually by emitting light.[13] Certain fluorophores are more useful than others, depending on their absorption and emission spectra, according to the application they are selected for. Optical biosensors first illuminate these molecules and then detect the emitted light and provide readings of its intensity. On the basis of this, specific information about the studied probes can be inferred, such as successful nucleic acid hybridization.

MECHANISM OF FLUORESCENCE ENERGY TRANSFER

Depending on the chosen fluorophores, there are more mechanisms that can lead to fluorescence energy transfer. Usually, most studies involve a pair of two fluorophores, a fluorophore and a quencher, or sometimes more of either of these molecules. Experimental studies rely on measuring the intensity of emitted light of their analyzed compositions at different stages of the investigations. A change in the intensity readings provides the information necessary to infer the hypothesized events. Consequently, either of two effects is possible: an increase in intensity or a decrease, the latter of which is called quenching. Although there are many mechanisms that can result in fluorescence quenching, this review focuses on contact and energy transfer quenching, namely, FRET.

FRET relies on the energetic interaction between two molecules, a donor fluorophore and an acceptor, which can in contrast be a nonfluorescent molecule. Once the light is absorbed by the former of the two, it reaches an excited state which affects the energetic state of an electron in the acceptor. The acceptor reaches an excited state and in turn emits light of a certain wavelength given its return to the ground state, or it loses the energy as heat in case it is not a fluorophore molecule.[14] The efficiency of this phenomenon is highly dependent on the distance between the two, with the widely accepted range being between 1 and 10 nm.[15] Likewise, it is imperatively important that the emission spectrum of the donor overlaps with the absorption spectrum of the acceptor. In terms of time, this type of events happen very fast, on the order of nanoseconds.

MATHEMATICS BEHIND FRET

The most important parameter for calculating the FRET efficiency is the distance between the donor and the acceptor molecules. The efficiency E can be calculated as follows

$$E = \frac{1}{1 + \left(\frac{R_{DA}}{R_0}\right)^6} \tag{1}$$

where R_{DA} is the donor−acceptor distance, and R_0 refers to the distance at which the energy transfer between them would have a 50% efficiency. The latter can also be calculated using the formula

$$R_0^{\,6} = \frac{9 \ln 10 \kappa^2 Q_D J(\lambda)}{128 \pi^5 n^4 N_A} \tag{2}$$

where Q_D is the fluorescence quantum yield of the donor in the absence of the acceptor, $J(\lambda)$ the integration of the spectral overlap of donor's emission and acceptor's absorption, n the refractive index of the medium, and N_A Avogadro's number, while the orientational factor κ^2 can be calculated as

$$\kappa^2 = (\sin\theta_D \sin\theta_A \cos\varphi - 2\cos\theta_D \cos\theta_A)^2 \tag{3}$$

where θ_D and θ_A are the angles between the connecting vector of the centers of the donor and acceptor molecules, and φ denotes the total angle between them. Although the value for κ^2 can be anywhere between 0 and 4, it is conventionally acceptable to use a 2/3 value as found through an averaging process over all the aforementioned angles.

However, when a higher accuracy is required, it is important to surpass the limitations imposed by the averaged distances considered in experimental FRET measurements. Makarov and Plaxco[16] propose a more reliant mathematical model that can better describe the FRET efficiency according to fluorophores dynamics

https://doi.org/10.1021/acs.jcim.1c00957
J. Chem. Inf. Model. 2022, 62, 232−239

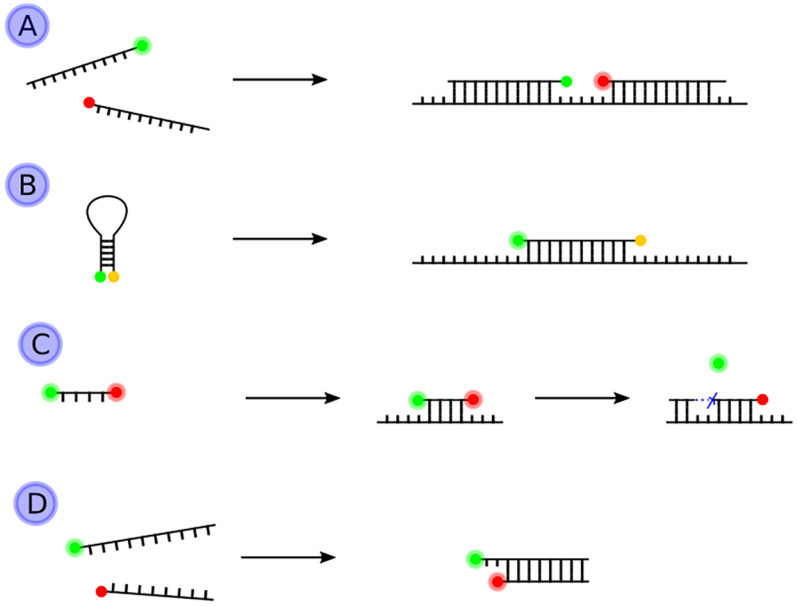


Figure 1. Most commonly used nucleic acids probe configurations, with (A) adjacent probes, (B) molecular beacons, (C) TaqMan probes, and (D) strand-displacement probes. The donor fluorophore is shown in green, the acceptor in red and the quencher in yellow.

$$\langle E \rangle = 1 - k_D \int_0^\infty I(t)\mathrm{d}t \tag{4}$$

where k_D^{-1} refers to the experimentally determined fluorescence lifetime of the donor, $I(t)$ the fluorescence intensity decay of the donor, and $\langle E \rangle$ the average FRET efficiency.

Additionally, the fluorescence intensity decay can be calculated as follows

$$I(t) = \left\langle e^{-\int_0^t [k_D + k_T(R(t'))]\mathrm{d}t'} \right\rangle \tag{5}$$

where $k_T(R)$ refers to the energy transfer rate specific to a certain distance, which varies in time.

Concerning comparison with experimental data, Hoefling et al.[17] have found that the orientational factor κ^2 tends to deviate from its widely accepted value when dyes are attached to biomolecules, due to sterical hindrance and electrostatic interactions. Given the possibility of using molecular dynamics and Monte Carlo simulations, they suggest using the aforementioned eqs 1, 2, and 3 and a hybrid computational method to calculate FRET efficiencies, with parameters obtained through computational methods. The used procedure relied on molecular dynamics scenarios simulating the time-based evolution of the biomolecules, with the excitation of the donor fluorophore being chosen to have happened at a random time. Furthermore, a Monte Carlo simulating procedure was used to pinpoint a random de-excitation time, evaluated at each time step of the molecular dynamics simulation. Having obtained a time-based evolution of the excitation and de-excitation of the fluorophores, it was possible then to calculate the FRET transfer rate coefficient and equivalently the efficiency. This protocol can be equally applied to scenarios involving strictly nuclear acids.

CONTACT QUENCHING

This type of quenching refers to the case in which a fluorescent molecule forms hydrogen bonds with another fluorophore or quencher at contact range. The latter refers to cases in which the distance between the two is less than 1 nm, with the photoinduced electron transfer (PET) effect taking place.[18,19] In such a situation, a new compound is formed by the two, which in turn affects the absorption spectrum of the structure. Usually, the quenching ratios achieved in contact quenching are very high, and given any absorbed energy, virtually no light is emitted. Many fluorescent probes rely on this type of quenching, either at an initial stage or later in the experiment.

PROBE CONFIGURATIONS

Probes that are part of experiments based on fluorescence quenching events due to hybridization are known to fit into four major categories. In no particular order, they are either adjacent probes, molecular beacons, TaqMan probes, or strand-displacement probes. Although there exist many variations,[20] they all are centered around the same rationale, namely, that the nucleotide strand acting as the probe is complementary, at least in part, to a target nucleotide strand.

Specifically, two different single strands of nucleic acids labeled with the donor and acceptor molecules, at the 3′ end for the former and, respectively, at the 5′ end for the latter, will undergo energy transfer quenching once they bind to a target nucleic acid strand, with complementary bases for each of the probes.[21,22] Expectedly, the distance in between the two probe strands is meant to be configured in such a way that the efficiency of the energy transfer is relevant for the study. In a preliminary fluorescence intensity reading, the probes are expected to be freely floating in the solution. Given that quenching takes place at a subsequent reading, this can be linked to the hybridization of the probes. These types of probes are called adjacent probes. There are many variations to this probe configuration, including having a third additional fluorophore on the end of one of the probe strands.

Molecular beacons rely on a different rationale for the change in fluorescence intensity. Precisely, these probes are made out of one single strand of nucleic acid that has its starting and ending nucleotides complementary, which bind together and form a stem. The donor and acceptor molecules are linked at either of the ends, leading to contact quenching.

https://doi.org/10.1021/acs.jcim.1c00957
J. Chem. Inf. Model. 2022, 62, 232−239

This further results in a longer loop formed by its middle section, which is meant to be complementary to a target sequence. The hybridization of these two is more stable than the hybridization of the stem of the molecular beacon and proceeds to separate the donor and acceptor. In turn, this means that quenching is no longer taking place, and an increase in fluorescence intensity should be read, confirming the hybridization. Molecular beacons were first introduced by Tiagy and Kramer[23] and are one of the most popular probe configurations. Accordingly, they also have a very large number of variations and designs, with molecular beacons that may have multiple fluorophores attached to one side of the stem, one side of the stem shorter than the other, and fluorophores attached to the loop section of the probe or toward the middle part of the stem.

TaqMan probes rely on the activity of Taq DNA polymerase to cleave terminal nucleotides on its 5′ end.[24] The fluorophores are attached to its either end, with the acceptor on the 3′ end and the donor on the other. Given an intact probe, fluorescence is quenched. However, given hybridization to a target strand, the donor molecule is cleaved away, and fluorescence is restored.

Two complementary strands of equal or different lengths can also be used as hybridization probes. They are called strand-displacement probes.[25] Having the donor molecule at the end of one of them and an acceptor molecule at the corresponding end of the other would lead to either contact quenching or FRET following hybridization. These probes have been extensively studied, especially for the efficiency of FRET related to the difference in strand lengths.

Figure 1 provides a visual representation of each of the four types of probe configurations, in their most widely used forms.

CHOICE OF FLUORESCENT MOLECULES AND QUENCHERS

Depending on the desired outcome of the experiment, it is highly important to choose an optimal combination of fluorophores or quenchers. This field has been extensively studied, with the database of available such molecules significantly expanding over the years. One of the key criterions for choosing a correct pair of fluorophores or quenchers is the overlapping of their emission and absorption spectra. The most compatible pairs, called complementary pairs, have been well documented, and it has been found that nucleic acids themselves may act as quenchers to various degrees.[26]

COMPUTATIONAL SIMULATIONS OF NUCLEIC ACIDS FRET EXPERIMENTS

Currently, the field of nucleic acid fluorescence studies is dominated by experimental investigations. This is because there are a few computational challenges that have not yet been resolved. The most important of these revolve around the fact that the computational power required to simulate a full-scale nucleic acids FRET experiment is beyond the resources technologically available at the moment. Additionally, the parametrizations for the available force fields are not fully developed for the complexity of the fluorophore molecules. Recent progress and useful workarounds have appeared; however, no standard procedure has yet been developed. Consequently, the in silico studies make use of a multitude of mathematical or computational algorithms, approximations, simplifications, and combinations of methods.[27] Generally, methods such as molecular dynamics, Monte Carlo simulations, or ab initio calculations such as density functional theory tend to be the most used or any hybrid approach between them. Sponer et al.[28] provided a recent overview of simulating RNA systems. They deem that given powerful computational resources, it is currently possible to fully explore the conformational space and expect convergence for only a nucleoside of RNA or a dinucleotide at most. However, there are enhanced sampling methods that are capable of reducing the workload and allow for tetranucleotides or tetraloops to be well simulated. Despite these, it is noted that biologically relevant time scales are not within our grasp for the current computational systems. There are many approaches that aim to aid molecular dynamics simulations of nucleic acids, such as Markov state models, methods based on annealing (e.g., parallel tempering) or importance sampling. The aforementioned two main challenges are explained in more detail in the following sections.

COMPUTATIONAL RESOURCES CHALLENGE

The simulation of a full-scale FRET experiment involving nucleic acids is expected to include some specific events in addition to the preliminary steps of energetic optimization, such as the fluorescence emission, probe and target hybridization, or the actual resonance energy transfer. Despite hardware and software improvements which are expected to increase the accuracy and viability of computational predictions, there are still many obstacles before such simulations can be universally performed. Conformational changes can easily be beyond the scope of the currently available computational resources based on today's technology, as it may require simulations to be done from microseconds to milliseconds and may even go as far as minutes. This is very distant from what is possible today, and considering multiple simulations would be needed for averaging purposes, it is a notably serious challenge.

Several workarounds have been proposed in order to reduce the computational power requirements. For example, most computational studies are restrained to simulating single molecule FRET (smFRET) scenarios, during which a much more limited number of particles is simulated. Additionally, the realm of nucleic acid hybridization is one of the central challenges that requires adjustments. This is mainly due to the fact that this is a very lengthy phenomenon in comparison to the simulation capabilities of today's computational methods. The available computational power is far too limited to employ such simulations as a standard procedure. Accordingly, many in silico experiments bypass this event and perform the computations for already hybridized strands. These techniques focus the attention on the donor and acceptor fluorophores interactions.

PARAMETERIZATION CHALLENGE PROGRESS

Usually, parametrizations for fluorophores are rather complex. Two of the most well-known force field families, namely, AMBER and CHARMM, cannot fully engulf the complexity of fluorophores and nor many other nucleic acid effects that cannot easily be accounted for. However, constant progress is being made and more so recently. One of the main differences between the two force fields refers to the way the electrostatic term is accounted for, the former of the force fields using

https://doi.org/10.1021/acs.jcim.1c00957
J. Chem. Inf. Model. 2022, 62, 232−239

potentials directly derived from quantum mechanical computations. In contrast, force fields such as CHARMM use an indirect approach through which the parameters are chosen so that they can reproduce the results of quantum mechanical calculations.[29] Whether developing new force fields or improving existing ones, the preferred method to validate it is considered to be the comparison with experimental data, as opposed to the otherwise still reliable quantum mechanical computations results. There have been major improvements lately regarding the AMBER and CHARMM force fields fluorophores. These force fields rely on fixed point charge approximation, which is an inconvenience when parametrizing dyes due to their anisotropic charge concentrations. In fact, they have highly polarized π-electron density, which must be well accounted for. Specifically, Graen, Hoefling, and Grubmüller[30] have developed the AMBER-DYES modular force field as part of the larger AMBER force field. They have addressed many complex parameters of the challenging fluorescent dyes, which until now could have been rather inaccurately simulated with extensive approximations within the usual major force fields, such as AMBER, OPLS-AA, CHARMM, or GROMOS. More precisely, they parametrized a total of 22 commonly used dyes from the Alexa, Atto, and Lumiprobe Cy families. The level of theory used was HF/6-31G* for bonded parameters and B3LYP/6-31G*/TIP3P for the nonbonded ones. Their results were in good agreement with quantum mechanical computations. Their work used quantum theoretical calculations to extract the mathematical values for variables in the bonded and nonbonded parameters, with particular attention on the π-electron polarizability and charge fluctuations. Although this is a significant advance for the opportunity to use the computational method of molecular dynamics to simulate FRET experiments, one of its main drawbacks is that it only works with the GROMACS molecular dynamics simulation software.

Schepers and Gohlke[31] have successfully addressed the incompatibility of the AMBER-DYES force field module with other computational simulations packages by developing AMBER-DYES in the AMBER adaptation force field module. Most notably, this is now available to be used with the AMBER MD engine and many of the other simulation softwares which are compatible to use similar input files. This adaptation was created manually, and when compared to the original AMBER-DYES module, achieved only a less than 0.01% relative error in terms of obtained energy components for each dye present. Additionally, they have also modified bonding parameters, such as adjusting for the correct representation of the Michael addition product for linker structures and expanded global charge parameters. Additionally, they fixed the issue present in the product of cysteine and maleimide in every linker ending in a cysteine. It is expected that an extension of the parametrized fluorescent labels can easily be achieved following the detailed presented workflows.

Shaw et al.[32] have developed a similar module based on the CHARMM force field for an improved parametrization of fluoresecent dyes. More precisely, the CHARMM-DYES module has been specifically focused on the Alexa, Atto, and Cy fluorescent dyes families. They used quantum mechanical calculations for the bonded, nonbonded, and charge parameters describing dye physical interactions and deemed this module as more prudent in its results that the AMBER-DYES force field, due to the tendency toward anisotropy given its approach for modeling point charges. They fit the charge parameters until there was less than 0.1% difference to quantum mechanical calculations. Their results show good agreement with the latter and are deemed much better than results obtained through the CGenFF force field, which is found to overestimate most parameters.

All these force field improvements show promising progress for the in silico FRET simulations. This is a step up from the otherwise used less exact coarse-grained force fields that make extensive use of approximations, simplifications, and averaged variables.[33] The latter type of force field relies on the approach of grouping several atoms that constitute a functional entity (e.g., one DNA base altogether, one sugar), be it a standalone chemical species, macromolecule, or even more complex components, and assigning approximate global parameters.

When it comes to the reliability of either AMBER or CHARMM force fields for the accurate description of nucleic acids, both of them show promising results.[34] The former family's accuracy is centered around the symmetry constraints restrained electrostatic potential. In turn, this followed several improvements due to the unreliability for longer time simulations (e.g., tens of nanoseconds), which ultimately lead to worthy enhancements for A-DNA, B-DNA, and Z-DNA conformations. However, it is recommended to combine the AMBER force fields with compatible water models for the simulation of specific nucleic acids scenarios, due to issues still appearing for less common conformations. The CHARMM family similarly was well suited for short time simulations and has known many adjustments over time. However, its main drawbacks now consist of the known aberrant base-pair openings and RNA helix fraying for longer time simulations. Studies present in the literature discuss the suitability of either AMBER or CHARMM force fields for specific nucleic acids simulations, with the information depending on the intended scenarios to be investigated.[35]

Polarizable force fields represent more novel improvements brought to the level of theory that describes the electrostatic interactions and are potentially more accurate for nucleic acids than AMBER and CHARMM force field families. However, given their superior level of complexity, one major drawback is the decrease of simulation speed, along with the incomplete experimental validation of the obtainable results. Usually, a validation of new additions brought to force fields meant to describe nucleic acids makes use of so-called unusual structures, such as four-way junctions, G-quadruplexes, i-motifs, hairpins, pseudoknots, and others, which form in both DNA and RNA.[34] Specifically for the simulation of the RNA scenarios, Tan et al.[36] have developed an accurate force field, which has been validated within FRET scenarios, achieving very rigorous and timely results.

COMPUTATIONAL METHODS FOR FRET STUDIES

Computational studies of nucleic acid FRET aim to predict outcomes of simulations meant to aid or ultimately design fluorescence biosensors. The main advantage of performing in silico experiments is the ability to analyze events at the atomic level in high detail. This means that theoretical approximations used in experimental studies can be confronted, such as the averaging of the κ^2 orientation factor to 2/3.[37] One usual outcome of such simulations is the accurate measurement of the distance between the donor and acceptor molecules based on the observed FRET efficiency. Additionally, the whole setup can be analyzed at each time frame or valid snapshot, and molecule configurations, rotational motions, and kinetics can

https://doi.org/10.1021/acs.jcim.1c00957
J. Chem. Inf. Model. 2022, 62, 232−239

be investigated with a much greater accuracy. Given that many obstacles remain for simulating a complete FRET experiment, the recent computational studies relied on customized computational approaches, which are presented next.

Grotz et al.[37] provide an overview of the molecular dynamics simulation method of nucleic acids within FRET experiments. They show that given certain parametrizations it is possible to simulate smFRET experiments with ssRNA or ssDNA with reasonable accuracy, and there exist clear deviations from experimental data that must still be addressed. Specifically, the most important problem highlighted refers to the stacking of the dyes onto the nucleic acids leading to a too stable conformation. To demonstrate this, they simulated Alexa Fluor 594 and Alexa Fluor 488, along with the nucleic acids. The orientation factor κ^2 was calculated in a time-dependent manner and found to be marginally smaller or larger than the 2/3 widely accepted value. Given the complexity and increasing number of atoms arising from the use of fluorophore molecules, a method of dye mapping was used instead, which substituted these molecules with specific dye conformers present in a constructed library. The water models TIP4P-D and TIP4P were found to be best for addressing, at least in part, the stacking problem. Although found to underestimate FRET efficiencies, the most viable force fields at present for simulating nucleic acids in FRET experiments are deemed to be parmBSC1 and DESRES.

Ingargiola, Weiss, and Lerner[38] have proposed an adjusted variant of a Monte Carlo protocol that can simulate and fit data from smFRET experiments in order to compute distance information between the fluorophores. Specifically, they use a Monte Carlo diffusion-enhanced photon-interference (MC-DEPI) approach along with a machine learning algorithm and study the results of smFRET scenarios based on two dsDNAs labeled with Atto 550 and Atto 647N, once separated by seven nucleotides and later by 17. Several different modules are integrated, such as photophysics, dynamics, intralifetime diffusion, and intertimestamp diffusion modules, all of which account for different parameters and use techniques such as the Ornstein−Uhlenbeck process or the Gillespie algorithm. A cost function describes the closeness of the computational parameters to the experimental ones and is optimized through a Bayesian global optimization approach. The whole protocol is iterated enough times so that the simulated parameters match the experimental ones. The time correlation had been computed using mathematical rotational dynamics. This proposed variant of the Monte Carlo computational method achieved good results, with the simulated distances tending to incur a 10% difference than those obtained through a FRET-restrained positioning and screening tool, which however achieved a much lower standard deviation.

Schrangl et al.[39] used Monte Carlo calculations to simulate smFRET scenarios in order to statistically validate experimental data. The main advantage of their approach is that it is independent of any generated histograms. The validity of the data can be directly inferred from the obtained *p*-values, which shows how probable it is for the simulated parameters to be consistent with the experimental ones. They test their method in an experiment based on the Holliday junction two-level FRET system. Expectedly, the employed types of computational methods cannot provide time-related parameters and make use of an assumptions-based mathematical method approximating and simplifying the energy transfer lifetime while knowing the efficiency. Specifically, Alexa Fluor 555 and Alexa Fluor 647 were the used fluorophores, along with some DNA oligonucleotides. This method is proposed to work as complementary to experimental testing, as it can help reduce experimental artifacts and statistical outliers.

Chen et al.[40] have proposed a protocol of using two different computational methods to reduce errors that appear in long smFRET simulations. Although they have not explicitly applied it to nucleic acid FRET experiments, it is equally fit to be used for such. More explicitly, molecular dynamics simulations that extend to the millisecond level are known to exhibit very large errors, such as more than 40%. This technique can be used specifically for scenarios aiming to investigate the FRET phenomenon on nucleic acid probes. Precisely, following a smFRET simulation either through molecular dynamics or Monte Carlo computations, they have shown that an additional postFRET Monte Carlo simulation can reduce the so-called "camera blurring effect", which usually happens in long coarse-grained simulations due to false transitions. The method uses an evaluation function that is optimized through a local search algorithm, which in its current form may take up to a day to converge for a five-state or six-state system. This has been shown to significantly reduce false transition errors by identifying feasible conformations of the system following the FRET phenomenon.

Bag et al.[41] used molecular dynamics simulations using the OPLS force field to complement experimental investigations of a FRET phenomenon taking place in a hybridized DNA duplex. The simulation parameters were coarse, as could be expected for such a large simulation. The method used in addition a Monte Carlo multiple minimum global searching protocol to search viable conformations of structures. The main purpose of using molecular dynamics was to correctly understand the molecular configuration and the configuration dynamics of the studied complex. As such, no FRET computations were calculated by the molecular dynamics engine; however, it provided insights into what is deemed a new hybridization-accompanying FRET event.

Janke et al.[42] have used molecular dynamics in combination with experimental work to compare the conformational dynamics restrictions that may appear between different families of commonly used dyes attached to single strand DNA probes. Specifically, they first simulated single stranded DNA on its own, in conditions similar to experimental ones using the corrected AMBER 99SB force field. Using AMBER-DYES, they then simulated Alexa and Cy dyes attached to single-stranded DNA and found that Cy dyes have a larger effect on conformational dynamics restrictions for the probe. Each simulation lasted for 5 μs.

Xu et al.[43] used molecular dynamics in order to gain detailed insights to the atomic configuration of experimentally studied smFRET molecular beacons, specifically the formed loops. They ran long simulations times of up to 3 μs, with a 2 fs time step using the AMBER force field for 13 different DNA sequences. The ff99 BSC1 force field was used for DNA strands, while water was modeled with TIP3P. The simulation found that, in general, tetraloops are more stable than triloops, and specifically, ACGA tetraloops were found to be more stable than CAG triloops, which is in accordance with experimental data.

Hirashima et al.[44] used time-dependent density functional theory calculations to aid in finding the theoretical FRET efficiency value for their investigated scenario. Seven strands of different DNA sequences were used: four donor strands and

https://doi.org/10.1021/acs.jcim.1c00957
J. Chem. Inf. Model. 2022, 62, 232−239

three acceptor ones. They compared this and found overall good agreement with the obtained experimental value, except a difference at a given point. The calculations were performed at the 6-31+G level in Gaussian to obtain the transition dipole moments.

Fuchtbauer et al.[45] made use of the FRETmatrix[46] tool to calculate the theoretical FRET efficiency for the investigated fluorophores pair. They utilized previous literature values for terms such as the transition dipole moments and made use of different calculation steps for either the A- or Z-form of the RNA.

Dimura et al.[47] developed a complementary software package that can be used for designing FRET experiments with biomacromolecules, including nucleic acids. Its functions include finding optimum fluorophore pairs and their conformational uncertainty, quality assessment of structural models, and model refinement for computational simulations.

A new tool for the computational generation of RNA molecular beacons has been developed by Bayer et al.[48] called PinMol. This aids with complex molecular beacon designs of up to 26 nucleotides and at least 18 nucleotides, energetically optimized through free energy optimization approaches. The entire algorithm is built in Python and provides .csv outputs with the topology of the molecular beacons, which can be easily read. This tool is specifically suitable for experiments requiring a large number of molecular beacon configurations, with up to 50 per run.

OUTLOOK

Computational methods have known significant progress lately in terms of the capability to simulate in silico FRET experiments. However, a number of important challenges remain that cannot be overcome given today's available computational power. Subsequently, resourceful workarounds have to be designed so that significant experiments can provide otherwise unavailable insights at atomic level of description. Lately, there has been an increased activity regarding fluorescent dye new force field parametrizations, ambitious computational method simulations able to question the conventional theoretical approximations made for FRET mathematical theory, and new tools able to facilitate smFRET molecular modeling scenarios. This is particularly encouraging, as in silico experiment insights can provide complementary results to this experimentally dominated field and bring close the realization of optical biosensors for clinical diagnostic of life-threatening diseases.

DATA AND SOFTWARE AVAILABILITY

The software used for creating the art in this review is Inkscape 1.0.2, which is free and open-source software licensed under the GPL.

AUTHOR INFORMATION

Corresponding Author

Mariana Ionita − *Faculty of Medical Engineering, University Politehnica of Bucharest, 011061 Bucharest, Romania; Advanced Polymer Materials Group, University Polithenica of Bucharest, 011061 Bucharest, Romania;* orcid.org/0000-0002-9563-7148; Email: mariana.ionita@polimi.it

Authors

Sorin Muraru − *Faculty of Medical Engineering, University Politehnica of Bucharest, 011061 Bucharest, Romania;* orcid.org/0000-0003-2869-2187

Sebastian Muraru − *Faculty of Medical Engineering, University Politehnica of Bucharest, 011061 Bucharest, Romania*

Florentin Romeo Nitu − *Faculty of Medical Engineering, University Politehnica of Bucharest, 011061 Bucharest, Romania*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.1c00957

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

This work was supported by a grant of the Ministry of Research and Innovation, Operational Program Competitiveness Axis 1 Section E, Program cofinanced from European Regional Development Fund "Investments for your future" under the project number 154/25.11.2016, P_37_221/2015, "A novel biosensor testing osteogenic potency; capturing best stem cell performance for regenerative medicine" (GRABTOP).

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Metkar, S. K.; Girigoswami, K. Diagnostic biosensors in medicine − A review. *Biocatal. Agric. BiotechnoL.* **2019**, *17*, 271−283.

(2) Haleem, A.; Javaid, M.; Singh, R. P.; Suman, R.; Rab, S. Biosensors applications in medical field: A brief review. *Sens. Int.* **2021**, *2*, 100100.

(3) Chen, C.; Wang, J. Optical biosensors: an exhaustive and comprehensive overview. *Analyst* **2020**, *145*, 1605−1628.

(4) Fu, Z.; Lu, Y.-C.; Lai, J. J. Recent Advances in Biosensors for Nucleic Acid and Exosome Detection. *Chonnam Med. J.* **2019**, *55*, 86−98.

(5) Vigneshvar, S.; Sudhakumari, C. C.; Senthilkumaran, B.; Prakash, H. Recent Advances in Biosensor Technology for Potential Applications − An Overview. *Front. Bioeng. Biotechnol.* **2016**, *4*, 11.

(6) Terai, T.; Nagano, T. Small-molecule fluorophores and fluorescent probes for bioimaging. *Pflugers Arch. − Eur. J. Physiol.* **2013**, *465*, 347−359.

(7) Algar, W. R.; Hildebrandt, N.; Vogel, S. S.; Medintz, I. L. FRET as a biomolecular research tool − understanding its potential while avoiding pitfalls. *Nat. MethodS* **2019**, *16*, 815−829.

(8) Quan, K.; Yi, C.; Yang, X.; He, X.; Huang, J.; Wang, K. FRET-based nucleic acid probes: Basic designs and applications in bioimaging. *TRAC, Trends Analyt. Chem.* **2020**, *124*, 115784.

(9) Okamoto, A. Next-generation fluorescent nucleic acids probes for microscopic analysis of intracellular nucleic acids. *Appl. Microsc.* **2019**, *49*, 14.

(10) Freidus, L. G.; Pradeep, P.; Kumar, P.; Choonara, Y. E.; Pillay, V. Alternative fluorophores designed for advanced molecular imaging. *Drug Discov.* **2018**, *23*, 115−133.

(11) Cui, F.; Yue, Y.; Zhang, Y.; Zhang, Z.; Zhou, H. S. Advancing Biosensors with Machine Learning. *ACS Sens.* **2020**, *5*, 3346−3364.

(12) Pietraszewska-Bogiel, A.; Gadella, T. W. J. FRET microscopy: from principle to routine technology in cell biology. *J. Microsc.* **2011**, *241*, 111−118.

(13) Lichtman, J. W.; Conchello, J.-A. Fluorescence microscopy. *Nat. Methods* **2005**, *2*, 910−919.

(14) Marras, S. A. E. Selection of Fluorophore and Quencher Pairs for Fluorescent Nucleic Acid Hybridization Probes. *Methods Mol. Biol.* **2006**, *335*, 3−16.
(15) Piston, D. W.; Kremers, G.-J. Fluorescent protein FRET: the good, the bad and the ugly. *Trend. Biochem. Sci.* **2007**, *32*, 407−414.
(16) Makarov, D. E.; Plaxco, K. W. Measuring distances within unfolded biopolymers using fluorescence resonance energy transfer: the effect of polymer chain dynamics on the observed fluorescence resonance energy transfer efficiency. *J. Chem. Phys.* **2009**, *131*, 085105.
(17) Hoefling, M.; Lima, N.; Haenni, D.; Seidel, C. A. M.; Schuler, B.; Grubmuller, H. Structural Heterogeneity and Quantitative FRET Efficiency Distributions of Polyprolines through a Hybrid Atomistic Simulation and Monte Carlo Approach. *PLoS One* **2011**, *6*, e19791.
(18) Torimura, M.; Kurata, S.; Yamada, K.; Yokomaku, T.; Kamagata, Y.; Kanagawa, T.; Kurane, R. Fluorescence-Quenching Phenomenon by Photoinduced Electron Transfer between a Fluorescent Dye and a Nucleotide Base. *Anal. Sci.* **2001**, *17*, 155−160.
(19) Kurata, S.; Kanagawa, T.; Yamada, K.; Torimura, M.; Yokomaku, T.; Kamagata, Y.; Kurane, R. Fluorescent quenching-based quantitative detection of specific DNA/RNA using a BODIPY FL-labeled probe or primer. *Nucl. Acids Res.* **2001**, *29*, 34e.
(20) Juskowiak, B. Nucleic acid-based fluorescent probes and their analytical potential. *Anal. Bioanal. Chem.* **2011**, *399*, 3157−3176.
(21) Tsuji, A.; Sato, Y.; Hirano, M.; Suga, T.; Koshimoto, H.; Taguchi, T.; Ohsuka, S. Development of a time-resolved fluorometric method for observing hybdrization in living cells using fluorescence resonance energy transfer. *Biophys. J.* **2001**, *81*, 501−515.
(22) Yang, C. J.; Martinez, K.; Lin, H.; Tan, W. Hybrid Molecular Probe for Nucleic Acid Analysis in Biological Samples. *J. Am. Chem. Soc.* **2006**, *128*, 9986−9987.
(23) Tyagi, S.; Kramer, F. R. Molecular Beacons: Probes that Fluoresce upon Hybridization. *Nat. Biotech.* **1996**, *14*, 303−308.
(24) Lyamichev, V.; Brow, M. A. D.; Dahlberg, J. E. Structure-specific endonucleolytic cleavage of nucleic acids by eubacterial DNA polymerases. *Science* **1993**, *260*, 778−783.
(25) Dai, Y.; Furst, A.; Liu, C. C. Strand Displacement Strategies for Biosensor Applications. *Trends Biotechnol.* **2019**, *37*, 1367−1382.
(26) Marras, S. A. E.; Kramers, F. R.; Tyagi, S. Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucl. Acids ReS.* **2002**, *30*, 122e.
(27) Sponer, J.; Banas, P.; Jurecka, P.; Zgarbova, M.; Kuhrova, P.; Havrila, M.; Krepl, M.; Stadlbauer, P.; Otyepka, M. Molecular Dynamics Simulations of Nucleic Acids. From Tetranucleotides to the Ribosome. *J. Phys. Chem. Lett.* **2014**, *5*, 1771−1782.
(28) Sponer, J.; Bussi, G.; Krepl, M.; Banas, P.; Bottaro, S.; Cunha, R. A.; Gil-Ley, A.; Pinamonti, G.; Poblete, S.; Jurecka, P.; Walter, N. G.; Otyepka, M. RNA Structural Dynamics As Captured by Molecular Simulations: A Comprehensive Overview. *Chem. Rev.* **2018**, *118*, 4177−4338.
(29) Yoo, J.; Aksimentiev, A. Improved Parametrization of Li^+, Na^+, K^+, and Mg^{2+} Ions for All-Atom Molecular Dynamics Simulations of Nucleic Acids Systems. *J. Phys. Chem. Lett.* **2012**, *3*, 45−50.
(30) Graen, T.; Hoefling, M.; Grubmüller, H. AMBER-DYES: Characterization of Charge Fluctuations and Force Field Parameterization of Fluorescent Dyes for Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 5505−5512.
(31) Schepers, B.; Gohlke, H. AMBER-DYES in AMBER: Implementation of fluorophore and linker parameters into AmberTools. *J. Chem. Phys.* **2020**, *152*, 221103.
(32) Shaw, R. A.; Johnston-Wood, T.; Ambrose, B.; Craggs, T. D.; Hill, J. G. CHARMM-DYES: Parameterization of Fluorescent Dyes for Use with the CHARMM Force Field. *J. Chem. Theory Comput.* **2020**, *16*, 7817−7824.
(33) Minhas, V.; Sun, T.; Mirzoev, A.; Korolev, N.; Lyubartsev, A. P.; Nordenskiold, L. Modeling DNA Flexibility: Comparison of Force Fields from Atomistic to Multiscale Levels. *J. Phys. Chem. B* **2020**, *124*, 38−49.
(34) Salsbury, A. M.; Lemkul, J. A. Recent developments in empirical atomistic force fields for nucleic acids and applications to studies of folding and dynamics. *Curr. Opin. Struct. Biol.* **2021**, *67*, 9−17.
(35) Yoo, J.; Aksimentiev, A. Improved Parameterization of Amine-Carboxylate and Amine-Phosphate Interactions for Molecular Dynamics Simulations Using the CHARMM and AMBER Force Fields. *J. Chem. Theory Comput.* **2016**, *12*, 430−443.
(36) Tan, D.; Piana, S.; Dirks, R. M.; Shaw, D. E. RNA force field with accuracy comparable to state-of-the-art protein force fields. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, E1346−E1355.
(37) Grotz, K. K.; Nueesch, M. F.; Holmstrom, E. D.; Heinz, M.; Stelzl, L. S.; Schuler, B.; Hummer, G. Dispersion Correction Alleviates Dye Stacking of Single-Stranded DNA and RNA in Simulations of Single-Molecule Fluorescence Experiments. *J. Phys. Chem. B* **2018**, *122*, 11626−11639.
(38) Ingargiola, A.; Weiss, S.; Lerner, E. Monte-Carlo Diffusion-Enhanced Photon Interference: Distance Distributions and Conformational Dynamics in single-molecule FRET. *J. Phys. Chem. B* **2018**, *122*, 11598−11615.
(39) Schrangl, L.; Gohring, J.; Schutz, G. J. Kinetic analysis of single molecule FRET transitions without trajectories. *J. Chem. Phys.* **2018**, *148*, 123328.
(40) Chen, J.; Pyle, J. R.; Piecco, K. W. S.; Kolomeisky, A. B.; Landes, C. F. A Two-Step Method for smFRET Data Analysis. *J. Phys. Chem. B* **2016**, *120*, 7128−7132.
(41) Bag, S. S.; Das, S. K.; Pradhan, M. K.; Jana, S. Hybridization accompanying FRET event in labeled natural nucleoside-unnatural nucleoside containing chimeric DNA duplexes. *J. Photochem. Photobiol.* **2016**, *162*, 669−673.
(42) Jahnke, K.; Grubmüller, H.; Igaev, M.; Gopfrich, K. Choice of fluorophore affects dynamic DNA nanostructures. *Nucl. Acids Res.* **2021**, *49*, 4186−4195.
(43) Xu, P.; Pan, F.; Roland, C.; Sagui, C.; Weninger, K. Dynamics of strand slippage in DNA hairpins formed by CAG repeats: roles of sequence parity and trinucleotide interrupts. *Nucl. Acids Res.* **2020**, *48*, 2232−2245.
(44) Hirashima, S.; Sugiyama, H.; Park, S. Construction of a FRET System in a Double-Stranded DNA Using Fluorescent Thymidine and Cytidine Analogs. *J. Phys. Chem. B* **2020**, *124*, 8794−8800.
(45) Fuchtbauer, A. F.; Wranne, M. S.; Bood, M.; Weis, E.; Pfeiffer, P.; Nilsson, J. R.; Dahlen, A.; Grotli, M.; Wilhelmsson, L. M. Interbase FRET in RNA: from A to Z. *Nucl. Acids Res.* **2019**, *47*, 9990−9997.
(46) Preus, S.; Kilsa, K.; Miannay, F. A.; Albinsson, B.; Wilhelmsson, L. M. FRETmatrix: a general methodology for the simulation and analysis of FRET in nucleic acids. *Nucl. Acids Res.* **2013**, *41*, e18.
(47) Dimura, M.; Peulen, T.-O.; Sanabria, H.; Rodnin, D.; Hemmen, K.; Hanke, C. A.; Seidel, C. A. M.; Gohlke, H. Automated and optimally FRET-assisted structural modeling. *Nat. Comm.* **2020**, *11*, 5394.
(48) Bayer, L. V.; Omar, O. S.; Bratu, D. P.; Catrina, I. E. PinMol: Python application for designing molecular beacons for live cell imaging of endogenous mRNAs. *RNA* **2019**, *25*, 305−318.

https://doi.org/10.1021/acs.jcim.1c00957
J. Chem. Inf. Model. 2022, 62, 232−239