Long-Range Energy Transfer in Protein Megamolecules

Elijah L. Taylor,[†,⊥] Kevin J. Metcalf,[‡,⊥] Benedetta Carlotti,[†,§] Cheng-Tsung Lai,[‖] Justin A. Modica,[‖] George C. Schatz,[*,‖] Milan Mrksich,[*,‡,‖] and Theodore Goodson, III[*,†]

[†]Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109, United States

[‡]Department of Biomedical Engineering, and [‖]Department of Chemistry, Northwestern University, Evanston, Illinois 60208, United States

[§]Department of Chemistry, Biology and Biotechnology, University of Perugia, Perugia 06123, Italy

S *Supporting Information*

ABSTRACT: In this investigation, we report evidence for energy transfer in new protein-based megamolecules with tunable distances between donor and acceptor fluorescent proteins. The megamolecules used in this work are monodisperse oligomers, with molecular weights of ∼100−300 kDa and lengths of ∼5−20 nm, and are precisely defined structures of fusion protein building blocks and covalent cross-linkers. Such structures are promising because the study of energy transfer in protein complexes is usually difficult in this long length regime due to synthetic limitations. We incorporated fluorescent proteins into the megamolecule structure and varied the separation distance between donor and acceptor by changing the length of the cross-linker in dimer conjugates and inserting nonfluorescent spacer proteins to create oligomers. Two-photon absorption measurements demonstrated strong coupling between donor and acceptor dipoles in the megamolecules. For the dimer systems, no effect of the cross-linker length on energy transfer efficiency was observed with the steady-state fluorescence investigation. However, for the same dimer conjugates, energy transfer rates decreased upon increasing cross-linker length, as evaluated by fluorescence up-conversion. Molecular dynamics simulations were used to rationalize the results, providing quantitative agreement between measured and calculated energy transfer lengths for steady-state results, and showing that the differences between the time-resolved and steady-state measurements arise from the long time scale for large-scale fluctuations in the megamolecule structure. Our results show an increase in energy transfer length with increasing megamolecule size. This is evidence for long-range energy transfer in large protein megamolecules.
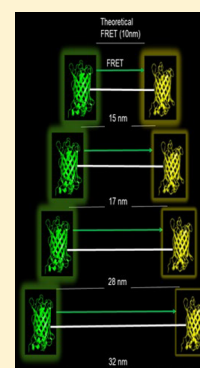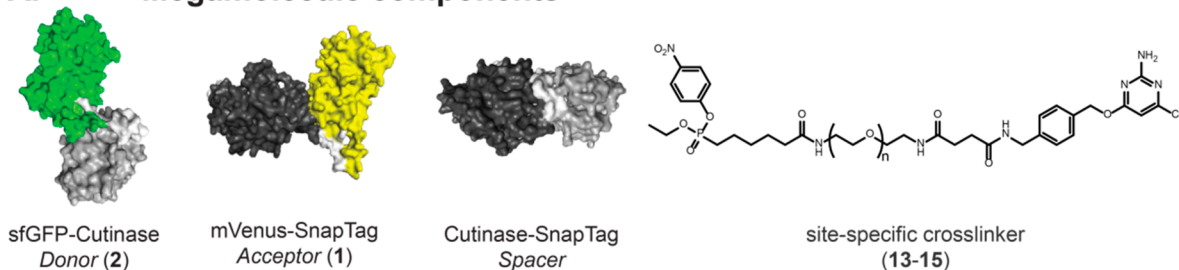
1. INTRODUCTION

The model of Förster resonance energy transfer (FRET)[1] describes an energy transfer pathway that occurs between two chromophores with overlapping emission and excitation energies[2] and that are separated by 1−10 nm.[3] Because of its sensitivity, FRET has proven useful in applications to measure protein−protein interactions.[2] The discovery of Green Fluorescent Protein (GFP) and its various fluorescent analogues[4] has revolutionized the study of FRET in studying and assaying protein−protein interactions.[3,5] First, many fluorescent proteins exhibit properties that are easily probed spectroscopically, including long lifetimes and high fluorescence quantum yields.[6,7] Second, they can be fused to essentially any biological macromolecule without losing their fluorescence, and usually without inhibiting the function of their fusion partner.[8,9] These properties, along with their overlapping excitation/emission spectra, make fluorescent proteins a useful biochemical tool for researchers to study energy transfer in protein assemblies.[10] However, studying the distance-dependence of energy transfer in protein assemblies is difficult due to synthetic limitations. We address this limitation by using a new synthetic strategy to prepare protein assemblies that allow for broad control over the donor−acceptor distances.

Energy transfer between biological macromolecules is studied by tuning interdomain distances of donor and acceptor fluorescent domains. For example, dendrimers,[11,12] nucleic acids,[13−17] and fusion proteins[18] have been used as scaffolds to vary chromophore distance. Protein macromolecules can also be used for this purpose by employing various bioconjugation strategies, although these strategies can be inefficient and often give heterogeneous conjugation products, making homogeneous oligomers difficult to produce.[19] Recently, we described the synthesis of precisely defined megamolecules by linking protein building blocks through a covalent cross-linker.[20,21] Both the protein and the cross-linker molecules are designed to allow the assembly of large (10−100 nm and 100−500 kDa), uniform covalent structures with precise connectivity. We hypothesized that the distance between the chromophore domains could be controlled with these megamolecules to probe energy transfer of protein assemblies in the 5−15 nm length regimes. It should be noted that, unlike nucleic acids, which have negatively charged backbones and a stacked π-helix, the investigated megamolecule structures do not have conducting backbones; this is also true for the covalent cross-
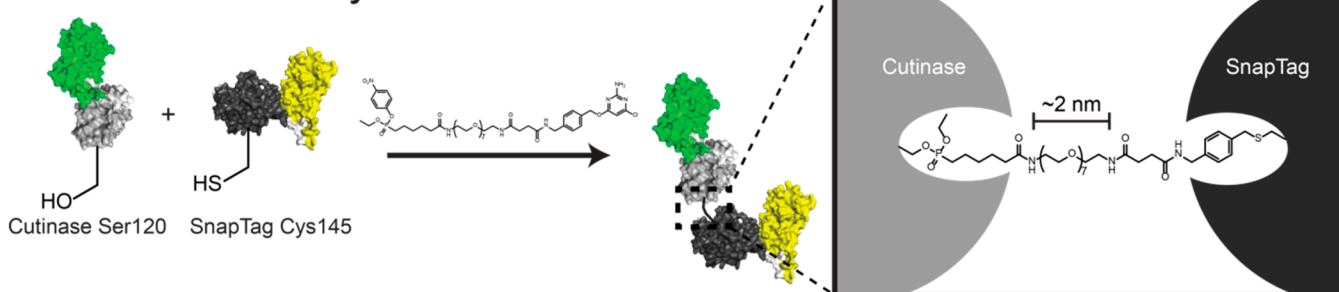
A. Megamolecule components

sfGFP-Cutinase
Donor (**2**)

mVenus-SnapTag
Acceptor (**1**)

Cutinase-SnapTag
Spacer

site-specific crosslinker
(**13-15**)

B. Generalized synthetic route

HO
Cutinase Ser120

HS
SnapTag Cys145

Cutinase

SnapTag

~2 nm

C. Variables tested

length of crosslinker

n = 7 (**4**), 9 (**5**), 11 (**6**)

number of spacer proteins

m = 0 (**4**), 1 (**7**), 2 (**8**), 3 (**9**), 4 (**10**)
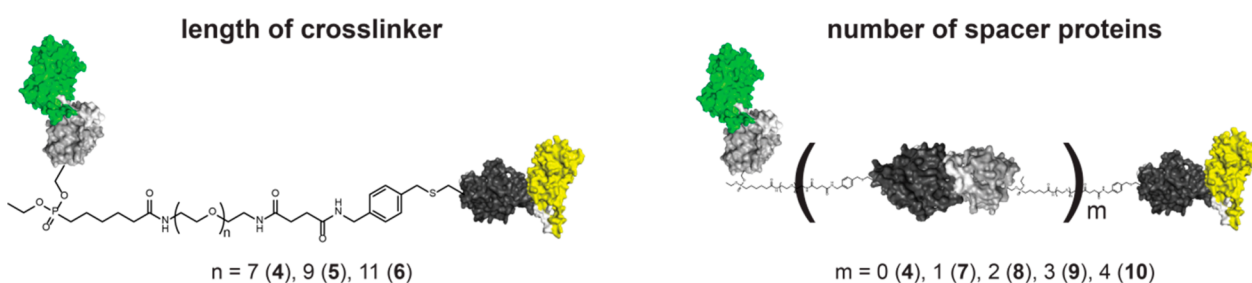
Figure 1. Experimental design. Megamolecules are numbered in bold. (A) Protein building blocks are connected using heterobifunctional site-specific cross-linkers. (B) Synthetic route scheme for synthesizing dimer FRET megamolecules. A representation of the cross-linker connecting the structural proteins, cutinase and SnapTag, and the maximal distance that could separate these two protein domains after cross-linking with the $n = 7$ cross-linker. (C) The distance between donor and acceptor was varied by two variables, cross-linker length (left) and number of spacer proteins (right).

linker. Thus, the megamolecules have the benefit that they rigorously exclude confounding modes of energy transfer, such as charge transfer and π-conjugation. As a result, only energy transfer through dipole–dipole coupling of the donor and acceptor chromophores is observed in megamolecules.

In this investigation, we characterized energy transfer within a series of megamolecules wherein the distance between donor and acceptor fluorescent proteins was varied by changing two variables: the length of the chemical cross-linker having different numbers of ethylene glycol units, and the number of nonfluorescent spacer proteins. A key to this study is the molecular-level control of megamolecule structure, where both the building blocks and the connectivity are varied systematically. Energy transfer rates measured using steady state, nonlinear, and time-resolved fluorescence methods were used to determine the distances between the donor and acceptor proteins, providing insight into the solution-phase conformations of the megamolecules, which were not previously known.[22] In addition, the energy transfer distances were found to be in agreement with the results of megamolecule structures that were determined using all-atom molecular dynamics calculations, with differences between the different time-scale measurements connected to the long time scale

(many ns) for large-scale structural changes in the megamolecules.

2. EXPERIMENTAL SECTION

2.1. DNA Cloning of Fusion Proteins. Expression vectors were created using the pET28b(+) expression plasmid. All cloning reactions were performed using the Golden Gate method[23] with *Bsa*I and T4 ligase (NEB). All cloning was performed in the DH5α cell line (NEB). A list of the plasmids used in this study is given in Table S6. Linear double strand inserts for Golden Gate cloning reactions were prepared by the polymerase chain reaction (PCR) using Q5 polymerase according to vendor's instructions (NEB). A list of the primers used to make these plasmids is given in Table S7. All primers were purchased from IDT. All DNA purification was performed using Qiagen kits according to vendor's protocol. New cloned plasmids were sequence-verified using Sanger sequencing (ACGT, Inc.).

2.2. Protein Expression and Purification. Proteins were expressed in *Escherichia coli*. For cutinase-containing proteins, the Shuffle T7 express cell line (NEB) was used to ensure proper folding and formation of the two disulfide bonds in cutinase. All other proteins were expressed in the BL21 (DE3) cell line (NEB). Cultures of the production cell lines were inoculated from frozen stocks and grown in 5 mL of lysogeny broth-Lennox, supplemented with 50 μg/mL kanamycin, for ~16 h at 30 °C and 240 rpm in an incubator

shaker (Innova). Cultures were diluted 1:200 into 500 mL of fresh 2xYT media (Dot Scientific), supplemented with 50 μg/mL kanamycin and 0.005% w/w antifoam-204 (Sigma), and grown at 30 °C and 240 rpm. After reaching an OD_{600} between 0.6 and 1.0, isopropyl β-D-1-thiogalactopyranoside was added to the culture to a final concentration of 500 μM. Cells were then shifted to an incubator shaker held at 20 °C and 240 rpm and grown for ~16 h. Cells were then centrifuged, and pellets were stored at −20 °C. Frozen cell pellets were thawed, resuspended, and sonicated to disrupt cell membranes. Cell lysate was clarified by centrifugation, and soluble cell lysate was incubated with Ni-NTA agarose beads (Qiagen). Beads were washed, and captured protein was eluted. Eluted protein was then precipitated with ammonium sulfate to a final concentration of 50% saturation and incubated at 4 °C for >1 h. Precipitated protein was pelleted and resuspended in 1x PBS. Sample was then injected onto a Superdex 200 pg size exclusion chromatography column (GE), and peak fractions were collected. The molar extinction coefficient at 280 nm was predicted using the ProtParam tool (https://web.expasy.org/protparam/). Purified protein was formulated to 50 μM in 1x PBS, flash-frozen in liquid nitrogen, and stored at −80 °C for future use.

2.3. Cross-Linker Synthesis. Cross-linkers were synthesized by elaborating a oligo(ethylene glycol) core molecule. Amino-PEG-amine molecules with a varying number of ethylene glycol repeats were purchased from BroadPharm. Cross-linker cores were elaborated with two ligands, phosphonate and benzyl chloropyrimidine units, via amide coupling of the core amine to the carboxylate on the ligand in one pot. The synthesis of the two ligand groups is described in greater detail in the Supporting Information (pp S34−S41). All chemicals involved in the synthesis of target compounds were reagent grade unless stated otherwise. 4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methyl-morpholinium chloride (DMTMM), N-methylmorpholine (NMM), and anhydrous dimethyl sulfoxide (DMSO) were purchased from Sigma-Aldrich. Amino-PEG$_n$-amines (n = 7, 9, 11) were purchased from Broadpharm. The heterobifunctional linkers were then purified by HPLC and verified by MALDI-TOF MS and ESI-TOF MS. Cross-linker was stored frozen in pure DMSO at a concentration of 10 mM.

2.4. Megamolecule Synthesis and Purification. Megamole-cules are synthesized as previously described.[20] All reactions were performed at a concentration of 1−100 μM protein in PBS with an equal molar ratio of all reactive substrates for 15 min at room temperature. Protein concentration did not affect coupling yield. For the dimer structures, a one-step synthesis was used, where sfGFP fusion, cross-linker, and mVenus fusions were incubated together (Figure 1B). For higher order oligomeric structures, a multistep convergent synthesis was used (Figure S1). First, fluorescent protein fusions were incubated with cross-linker. Excess linker was removed by serial dilution and concentration in 10 kDa molecular weight cutoff spin concentrators (Millipore). Next, cutinase-SnapTag protein was added to form the dimer intermediate. Dimer intermediate was then purified by size exclusion chromatography. This process of linker conjugation, cross-linking with cutinase-SnapTag, and purification by SEC was repeated to form the trimer intermediate. To create the trimer FRET megamolecule, an mVenus intermediate dimer was incubated with cross-linker and sfGFP−cutinase fusion and purified by SEC. To create the tetramer FRET megamolecule, an mVenus intermediate dimer was incubated with cross-linker and sfGFP−cutinase intermediate dimer and purified by SEC. To create the pentamer FRET megamolecule, an mVenus intermediate trimer was incubated with cross-linker and sfGFP intermediate dimer and purified by SEC. To create the hexamer FRET megamolecule, an mVenus intermediate trimer was incubated with cross-linker and sfGFP intermediate trimer and purified by SEC. Megamolecules were purified by size exclusion chromatography using a Superdex 200 column (GE) on an AKTA pure FPLC unit (GE) using PBS + 0.02 w/w% NaN$_3$ as the mobile phase. Peak fractions were pooled, and purity was assessed by SDS-PAGE.

To determine the partition coefficient of megamolecules, purified samples were analyzed by size-exclusion chromatography using a

Superdex 200 Increase column (GE) on an AKTA pure FPLC unit (GE) using PBS + 0.02 w/w% NaN$_3$ as the mobile phase. Peak elution volumes, V_e, were determined for three technical replicate injections. A protein standard mixture of four proteins ranging from 15 to 600 kDa (Sigma) was analyzed by the same method. The void volume, V_0, was determined by injection of 2 MDa dextran blue (Sigma), and the column volume, V_t, was determined by injection of 2% acetone in water. The partition coefficient, K_{av}, is calculated using eq 1. Data were fit using a power law model.

$$K_{av} = \frac{V_e - V_0}{V_t - V_0} \tag{1}$$

2.5. Dynamic Light Scattering Analysis (DLS). Megamolecule samples were diluted to 1 μM in PBS and passed through a 0.2 μm filter. Dynamic light scattering data were collected on a Zetasizer Nano ZS at 25 °C. Samples were equilibrated in the instrument for 5 min before sample acquisition. Each replicate entailed averaging over 10 acquisitions that lasted 10 s each. Four technical replicates were performed, and the mean and standard deviation were reported. The mean of the most intense peak from the intensity particle size distribution represents the megamolecule species and was reported.

2.6. SDS-PAGE Analysis. Proteins and megamolecules were analyzed for size and purity by SDS-PAGE. 1−5 μg of sample was loaded on a 4−15% Tris-glycine precast gel (Biorad) and separated for 30 min at 200 V, and stained with Coomassie-R-250.

2.7. Protein Mass Spectrometry. Megamolecule samples were prepared for LC−MS analysis by dilution to 1 μM in water. LC−MS analysis was performed on an Agilent 1200 series HPLC connected to an Agilent 6210A time-of-flight (TOF) mass spectrometer. A 10 μL injection of each sample was captured on a C18 trap column (Waters) and eluted using a gradient from 5% to 95% acetonitrile and 0.1% formic acid in water with a flow rate of 0.25 mL/min. Data were analyzed with Agilent MassHunter Qualitative Analysis B.04.00, and spectra were deconvoluted using a maximum entropy deconvolution calculation.

2.8. Steady-State Absorption and Emission Measurements. All of the measurements were performed at room temperature. Concentrations ranging from 2.0−3.5 μM were used for the spectroscopic investigation. Absorption spectra were measured using an Agilent 8432 UV−visible absorption spectrophotometer. The emission spectrum measurements were performed with a Varian Cary Eclipse fluorescence spectrophotometer. The fluorescence quantum yields of the samples were calculated using a known procedure,[24] and Coumarin 153 in ethanol (ϕ_F = 0.501) was used as the standard.[25] Fluorescence quantum yields values were obtained according to the following equation:

$$\phi_X = \phi_{ST}\left(\frac{Grad_X}{Grad_{ST}}\right)\left(\frac{n_X^2}{n_{ST}^2}\right)$$

where the subscripts ST and X denote standard and sample, respectively; ϕ is the fluorescence quantum yield; Grad is the gradient from the plot of integrated fluorescence versus absorbance; and n is the refractive index of the solvent.

2.9. Two-Photon Excited Fluorescence Measurements. The method for two-photon experiments has been previously described.[26,27] Two-photon excited fluorescence measurements were performed using a mode-locked Ti:sapphire laser, which is tunable from 700 to 900 nm delivering 110 fs output pulses at a repetition rate of 80 MHz. Emission scans were performed at 820 nm excitation while scanning the emission in the 400−800 nm range at ~50 mW, but the exact emission detection wavelength during the power dependence scan was selected by the emission wavelength that produced the highest number of counts. Excitation power from the laser was varied using a variable neutral density filter. Two-photon power-dependent fluorescence intensity was utilized to determine the two-photon absorption cross section through a comparative method.[28] Coumarin 153 in ethanol was used as the standard (cross section 99.45 GM at 820 nm).[24] Absorption spectra were taken

before and after each experiment to ensure that there was no appreciable photodegradation due to laser irradiation.

2.10. Time-Resolved Fluorescence Measurements. The time-correlated single photon counting (TCSPC) technique, which has been described previously, was used to study the long decay component of the investigated samples.[29] The laser used for the TCSPC measurement was a Kapteyn-Murnane (KM) mode-locked Ti:sapphire laser. The output beam wavelength from the KM laser was 800 nm, with a pulse duration of ca. 30 fs. The output beam was frequency-doubled using a nonlinear β-barium borate crystal to obtain a 400 nm beam. A polarizer was used to vary the power of the 400 nm beam that excites the sample. Focus on the sample cell (quartz cuvette, 0.4 cm path length) was ensured using a lens of focal length 11.5 cm. Collection of fluorescence was carried out in a direction perpendicular to the incident beam into a monochromator, and the output from the monochromator was coupled to a photomultiplier tube, which converted the photons into counts. The instrument response function (IRF) has been determined by measuring the scattering signal of a silica gel water dispersion and was found to have a full width at half-maximum of 1 ns. Deconvolution of the IRF was operated by FluoFit software during fitting of the single photon counting results.

The femtosecond time-resolved fluorescence experiments were performed using an ultrafast fluorescence up-conversion setup that had previously been described.[30−32] A mode-locked Ti:sapphire femtosecond laser (Spectra Physics Tsunami) was used to generate 80 fs pulses at 800 nm with a repetition rate of 82 MHz. This mode-locked laser was pumped by a 532 nm continuous light output from another laser (Spectra Physics Millennia), which has a gain medium of neodymium-doped yttrium vanadate (Nd:YVO4). A 400 nm excitation pulse was generated by a second harmonic β-barium borate crystal, and the residual 800 nm beam was made to pass through a computer-controlled motorized optical delay line. The polarization of the excitation beam was controlled by a Berek compensator. The power of the excitation beam varied between 33 and 36 mW. The fluorescence emitted by the sample was up-converted by a nonlinear crystal of β-barium borate by using the residual 800 nm beam, which had been delayed by the optical delay line with a gate step of 6.25 fs. By this procedure, the fluorescence can be measured temporally. The monochromator is used to select the wavelength of the up-converted beam of interest, and the selected beam is detected by a photomultiplier tube (R152P, Hamamatsu, Hamamatsu City, Japan). The photomultiplier tube converts the detected beam into photon counts, which can be read from a computer. Coumarin 30 was used for calibrating the laser. The sigma parameter from Gaussian fit of the instrument response function (IRF) (measured from Raman signal of water) was found to be 110 fs for the fluorescence up-conversion. Lifetimes of fluorescence decays were obtained by fitting the fluorescence decay profile to the most accurate fit. Multiexponential decay function fits in OriginPro 9.1 were necessary for the data analysis.

2.11. Molecular Dynamics Simulations. The initial structures of fusion proteins were predicted using the Phyre2 prediction model[33] and were built using Robetta.[34] The protein sequences are exactly the same as in the experimental setup. Implicit solvent AMBER GB8 force fields[35] were used for the simulations. For the linker molecules, the partial charges were calculated using GAMESS[36] with the HF/6-31G* basis set, followed by the RESP[37] fitting procedure. Bond, angle, torsion, and Lennard-Jones parameters for the linker molecules were taken from the General Amber Force Field.[38] Each system was equilibrated as follows. First, the system was minimized with 1000 steps of steepest descent. The system was then gradually heated from 100 to 300 K in 200 ps using Langevin dynamics with a collision frequency of 1 ps⁻¹. A 10 kcal mol⁻¹ Å⁻² Cartesian restraint was applied to the protein and linker during the heating. The restraints on the protein and linker then were gradually removed starting with a 10 kcal mol⁻¹ Å⁻² Cartesian restraint for 200 ps, followed by 1 kcal mol⁻¹ Å⁻² Cartesian restraint for 200 ps, and last 1 kcal mol⁻¹ Å⁻² Cartesian restraint on the protein backbone for 200 ps. After equilibration, a 200 ns production run at 300 K was performed. All MD simulations were

performed using GPU-implemented pmemd in Amber 16.[39] Probability plots excluded the first 100 ns of the simulation to account for equilibration.

3. RESULTS

3.1. Synthesis and Characterization of Megamolecule Structures. We synthesized linear protein scaffolds of varying size to study the effect of distance on energy transfer between two fluorescent protein domains. Megamolecules were assembled by the sequential combination of modular, orthogonal building blocks of two classes, fusion proteins and site-specific cross-linkers (Figure 1A). The fusion proteins had either cutinase or SnapTag connected to either sfGFP or mVenus fluorescent proteins. Cutinase and SnapTag are enzymes that each react with an irreversible inhibitor to form stable covalent adducts with a nucleophilic residue in the active site. The reactions are chemoselective, in that each enzyme forms a covalent adduct exclusively with its cognate ligand. The nucleophilic residue Ser120 of cutinase forms a covalent adduct by displacement of the 4-nitrophenyl group in a phosphonate moiety to form a stable phosphonate ester adduct. The nucleophilic residue Cys145 of SnapTag forms a covalent adduct by displacement of a chloride in a benzyl chloropyrimidine moiety to form a stable thioether adduct. Reaction of sfGFP−cutinase and mVenus−SnapTag with a heterobifunctional cross-linker yielded a covalent linear heterodimeric structure where the two fusion proteins are cross-linked through the nucleophilic residues of cutinase and SnapTag. We used as the donor sfGFP and as the acceptor mVenus (Figure 1B). The distance between these chromophores was predicted to be approximately 5 nm. We tested three cross-linkers having oligo(ethylene glycol) spacers from 7−11 monomers (EG7, EG9, and EG11), which were expected to maximally vary the cross-linker length by ∼1 nm[40] (Figure 1C, left). We also tested oligomers where nonfluorescent spacer protein(s) were placed between the terminal donor and acceptor domains and were expected to increase the megamolecule length by ∼5 nm per spacer protein. In this study, oligomers from dimer to hexamer were generated and tested (Figure 1C, right; see Figure S1 for the synthetic route of oligomers).

A list of the seven megamolecules we prepared is presented in Table 1, together with the labels used for each of them throughout the Article (see also the description of each megamolecule structure in Figure 1). These molecules were

Table 1. List of Fluorescent Protein Megamolecules Tested

megamolecule	name	hypothesized end-to-end length (nm)[a]
1	mVenus−SnapTag	3
2	sfGFP−cutinase	3
3	equimolar mixture of **1** and **2**	N/A
4	dimer; EG7 linker	8
5	dimer; EG9 linker	8.5
6	dimer; EG11 linker	9
7	trimer; EG7 linker	13
8	tetramer; EG7 linker	18
9	pentamer; EG7 linker	23
10	hexamer; EG7 linker	28

[a]See the Supporting Information (p S15) for an explanation of the hypothesized distances.

purified by size-exclusion chromatography (Figure S3), and analyzed by SDS-PAGE (Figure 2A) and mass spectrometry
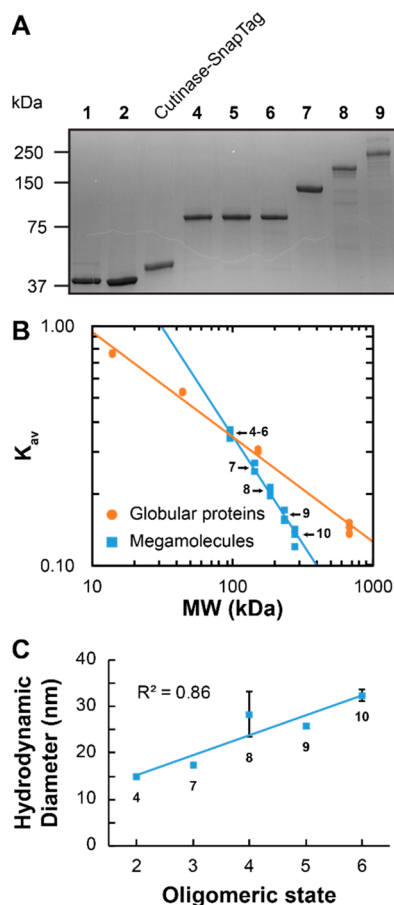

Figure 2. Characterization of FRET megamolecules. (A) SDS-PAGE analysis of megamolecule structures. (B) Plot of partition coefficient (K_{av}) for megamolecules and globular protein standards, as determined by size-exclusion chromatography. Slope for globular proteins = -0.435 ± 0.008; slope for megamolecules = -0.92 ± 0.03. Three technical replicates are plotted for each sample. (C) Plot of the intensity hydrodynamic diameter of megamolecules as determined by dynamic light scattering. The mean and standard deviation for four technical replicates are plotted. For both data plots, megamolecule data points are labeled with their corresponding number in bold.

(Figure S4). All species were produced with high purity and were confirmed to have the correct molecular weight to high precision (Table S1). The structure of the linear mega-molecules is compact and nonglobular.[20] We confirmed that fluorescent megamolecules had a similar structure by size-exclusion chromatography. The hydrodynamic radius of megamolecules increased with each successive addition of the spacer protein (**4, 7–10**), as measured by the partition coefficient (Figure 2B). In addition, the slope of this fitted line showed that the megamolecules are anisotropic and adopt a nonglobular structure (-0.92 ± 0.02 for megamolecules versus -0.435 ± 0.008 for globular proteins). The dimer mega-molecules (**4–6**) had indistinguishable hydrodynamic radius, suggesting that the dimer megamolecules may have similar structures in solution. These results were confirmed by dynamic light scattering, which also showed an increasing relationship between oligomeric state and megamolecule size (Figures 2C and S17). The intensity mean hydrodynamic

diameter increased linearly from the dimer (**4**) at 15.0 ± 0.4 nm, to the hexamer (**10**) at 32.4 ± 1.3 nm (Table S5).

3.2. Steady-State Absorption. UV–visible absorption spectra of all samples revealed characteristic protein absorption at 280 nm due to aromatic amino acid residues in the megamolecules (Figure 3). The mVenus chromophore in **1** had an absorption maximum at 515 nm, consistent with previous reports.[41] The sfGFP chromophore in **2** had
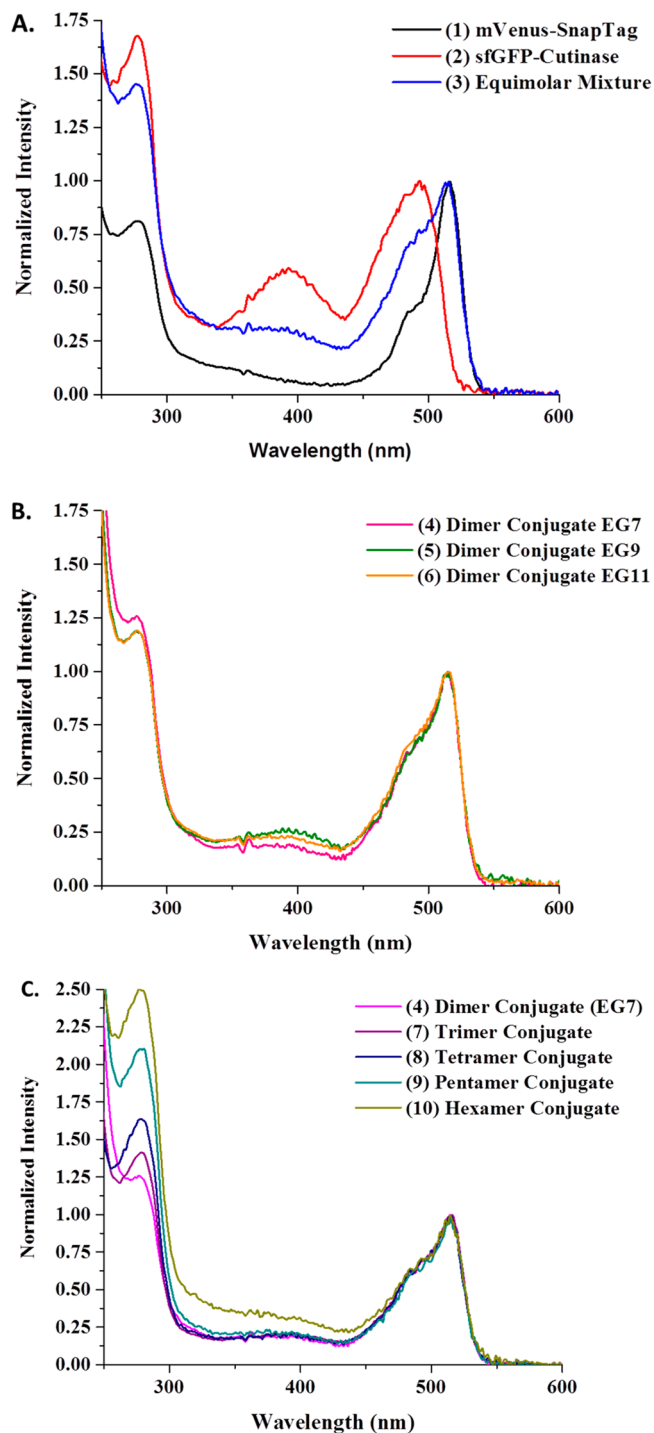
Figure 3. UV–visible absorption spectra of all samples: (A) Monomeric samples and the equimolar protein mixture; (B) each dimer conjugate with varying cross-linker lengths; and (C) each oligomeric state of the megamolecules from dimer to hexamer.

absorption maxima at 400 and 496 nm, due to the protonated and deprotonated forms of the chromophore.[9,38] Samples 3−10 showed absorption maxima at 280, 400, and 515 nm. The absorption maxima at 400 and 515 nm are contributions from the sfGFP−cutinase and mVenus−SnapTag chromophores, respectively. The spectra for the equimolar mixture and all of the linked conjugates overlapped well with one another in the visible range, suggesting that the fluorescent protein chromophores absorb light similarly in all megamolecules. However, the peak at 280 nm increased relative to the 515 nm peak with increasing oligomeric state. This is expected as increasing the oligomeric state will increase the number of aromatic residues, but will not increase the number of fluorescent protein chromophores. The UV−visible absorption spectra for the EG7 linker (13) and the cutinase-SnapTag spacer protein did not show significant absorption in the visible range and likely do not contribute to energy transfer (Figure S5).

We calculated molar extinction coefficients for each sample at 400 nm by using the extinction coefficient at 280 nm predicted from the primary sequence (see Table S2). All samples containing the sfGFP−cutinase protein fusion exhibited similar molar extinction coefficients, indicating that sfGFP absorption was not affected by its incorporation into a megamolecule.

3.3. Steady-State Fluorescence. Emission spectra were collected using 400 nm excitation for all samples to ensure selective excitation of the donor. Visible excitation at other wavelengths toward the red increases the probability of direct acceptor excitation. Choosing excitation at 400 nm allows one to maximize donor excitation while minimizing acceptor excitation, as the spectrum shows a maximum for the donor and a trough for the acceptor (see Figure 3A). The normalized spectra are shown in Figure S6. The emission spectrum for 1 revealed a peak at 527 nm, whereas 2−10 showed peaks at 510 nm. For systems in which donor and acceptor were both present in equimolar amounts, a shoulder in the red part of the spectrum at ca. 525 nm appeared, suggesting emission by the mVenus domain. In 3, where there was no linkage between donor and acceptor, a small increase in fluorescence at approximately 540 nm relative to 2 was observed. Practically only donor emission is revealed for the equimolar mixture, indicating low direct acceptor excitation. Also, this result demonstrates that no significant energy transfer takes place for the equimolar mixture. In dimeric megamolecules (4−6), the ratio of the height of the 510 nm peak to the intense 525 nm shoulder was approximately 1 to 0.92. In 7, the trimeric megamolecule with a protein spacer, the ratio of the heights between the main peak at 510 nm and the shoulder at 525 nm was approximately 1 to 0.69, indicating less emission from the mVenus domain in the trimer than in the dimers. The shoulder at ca. 525 nm for the larger oligomer megamolecules (8−10) was lower than that for the trimer, indicating that the larger megamolecules had lower levels of emission by the mVenus domain.

To measure fluorescence quantum yields, all samples were excited at 400 nm. The calculated value of 0.65 for 2 agreed with the literature value.[9] The calculated value for 1 (0.17) was much lower than expected (0.57),[42] likely due to excitation in a region with low absorption by the fluorophore. Quenching of sfGFP fluorescence was observed in the equimolar mixture (3) and the EG7 dimer (4), such that the fluorescence quantum yield was 0.20 for both samples. The decreased quantum yield

of the equimolar mixture is attributed to collisional quenching of sfGFP−cutinase by other mVenus−SnapTag molecules in solution. The decreased quantum yield of the EG7 dimer suggests that energy is transferred from the sfGFP domain to the mVenus domain within the megamolecule. The dimeric megamolecules with longer cross-linkers (5 and 6) and the trimeric megamolecule (7) had quantum yields of 0.078−0.079, evidence of greater quenching than in the dimeric EG7 megamolecule. The quantum yields were higher in the larger oligomeric megamolecules (8−10), suggesting less quenching of sfGFP fluorescence due to less efficient energy transfer to the mVenus domain. The fluorescence quantum yields for these three samples followed a decreasing trend with increasing oligomeric size, which also may be due to increased nonradiative deactivation in the more flexible megamolecules.

The distance between sfGFP and mVenus was calculated using Förster theory. Fluorescence emission spectra were fit using a linear combination of Gaussian distributions (see pp S20−S23 in the Supporting Information for a detailed description of the fitting procedure), and the relative contributions of sfGFP and mVenus to the emission spectrum of each megamolecule were calculated. The ratio of these contributions yielded the efficiency of energy transfer, E_{FRET} (Table 2). The energy transfer efficiency, ∼30%, was not

**Table 2. Mean Values of E_{FRET} for Megamolecules[a]**

no.	sample	E_{FRET}	r (nm)
3	equimolar mixture	0.048 ± 0.003	9.51 ± 0.12
4	dimer, EG7	0.341 ± 0.003	6.44 ± 0.01
5	dimer, EG9	0.329 ± 0.003	6.50 ± 0.01
6	dimer, EG11	0.342 ± 0.003	6.43 ± 0.11
7	trimer, EG7	0.180 ± 0.004	7.43 ± 0.03
8	tetramer, EG7	0.142 ± 0.008	7.79 ± 0.09
9	pentamer, EG7	0.113 ± 0.005	8.13 ± 0.06
10	hexamer, EG7	0.158 ± 0.009	7.63 ± 0.08

[a]Uncertainty is represented as the standard error of fit. A minimum of two replicate experiments were performed for each sample.

affected by cross-linker length (4−6). In addition, the energy transfer efficiency was inversely proportional to the oligomeric state in the regime of dimer to pentamer, decreasing from 34.1 ± 0.3% (4), 18.0 ± 0.4% (7), 14.2 ± 0.8% (8), and 11.3 ± 0.5% (9). This result is consistent with the insertion of spacer proteins increasing the distance between donor and acceptor. However, the energy transfer efficiency for the hexamer (10) was 15.8 ± 0.9%, a slight increase from the pentamer. For all oligomeric megamolecules, the energy transfer efficiency was greater than the noncovalent equimolar mixture of donor and acceptor (3; 4.8 ± 0.3%). We calculated the Förster radius, R_0, for the sfGFP−mVenus pair to be 5.77 nm using eq 2 (Figure S7),[43] which agrees well with the reported literature value of 5.64 ± 0.11 nm for green-yellow fluorescent protein pairs.[3]

$$R_0 = 0.2108\left[\kappa^2 \Phi_D n^{-4} \int_0^\infty I_D(\lambda)\varepsilon_A(\lambda)\lambda^4 \, d\lambda\right]^{1/6} \quad (2)$$

According to the Förster theory, the distance, r, between the donor and acceptor is a function of the energy transfer efficiency and the Förster radius. The distance between donor and acceptor was calculated using eq 3.[43] These values are presented in Table 2 for each structure.
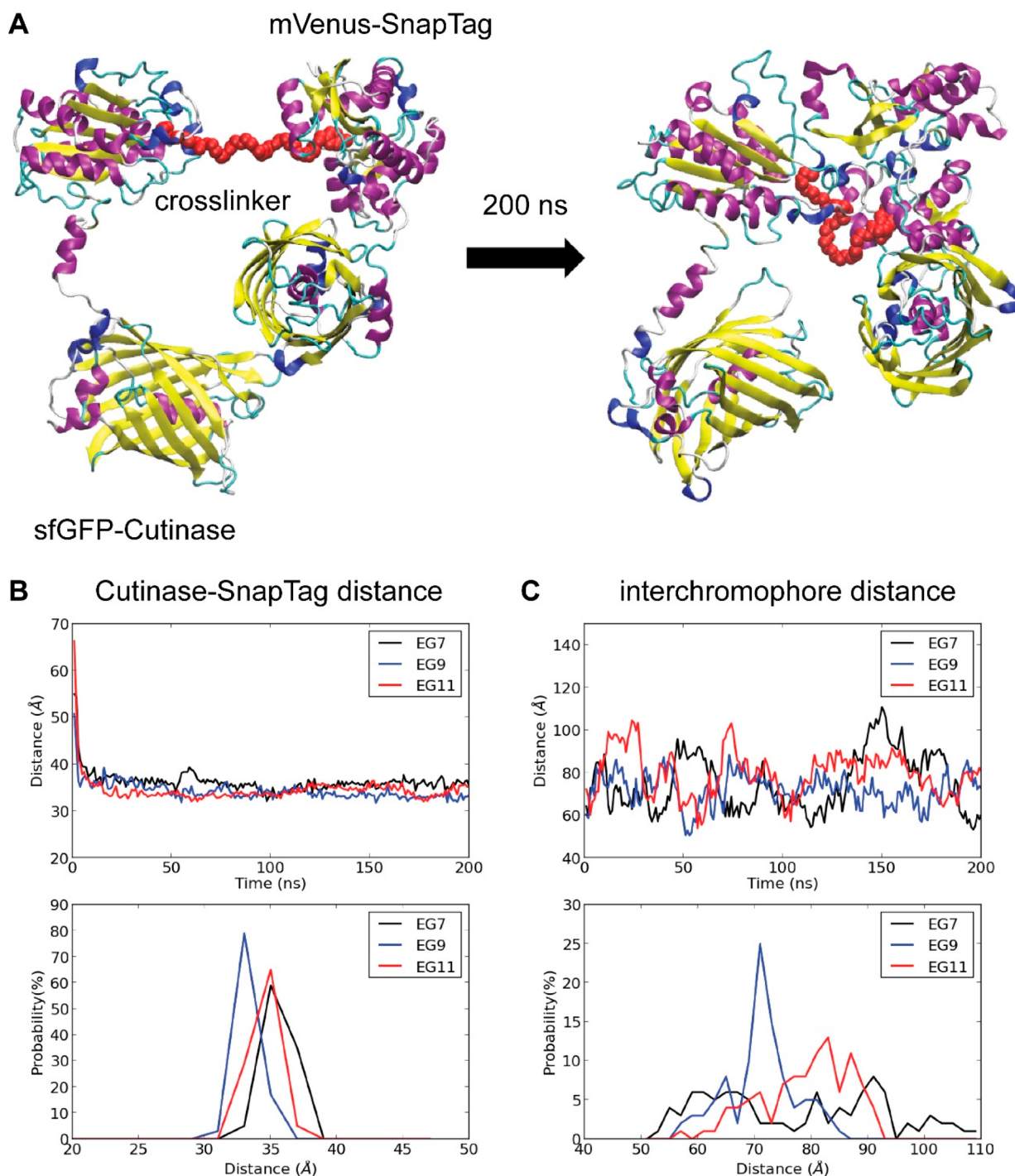
Figure 4. Molecular dynamics simulation for **4−6**. (A) A representative snapshot taken from the trajectories of **4**. (B) Distance from cutinase Ser120 to SnapTag Cys145 as a function of time. The cutinase and SnapTag proteins aggregated within 10 ns. (C) Distance between sfGFP and mVenus chromophores as a function of time.

$$r = R_0 \left(\frac{1}{E_{FRET}} - 1 \right)^{1/6}$$

(3)

To understand the lack of dependence of the cross-linker length on steady-state energy transfer efficiency, we performed all-atom molecular dynamics simulations of **4−6** to probe possible conformations. We built a molecular model from a homology model of each fusion protein (Figure S2) and varied the cross-linker length. The simulation results showed that the

cross-linker directs aggregation of sfGFP−cutinase and mVenus−SnapTag within 10 ns and forms a protein−protein contact between cutinase and SnapTag for all of the linker lengths (Figure 4A,B), such that the cutinase-SnapTag distance is 3.3−3.5 nm and independent of the linker. We also measured the distance between the sfGFP and mVenus chromophores during the simulation trajectory, and found the values fluctuate within the range of ∼6−9 nm for all dimers (Figure 4C). However, the probability distribution varies for the dimers, such that for interchromophore distances less than

6 nm, shorter linker lengths populate closer conformations. The average sfGFP−mVenus distance for **4**, **5**, and **6** was 7.75, 7.11, and 7.87 nm, respectively. To compare the simulated interchromophore distance with the experimental results, we calculated an adjusted distance (r_a) by weighting n measurements of the interchromophore distance (r) by the inverse sixth power of r (eq 4). The adjusted distance values for **4**, **5**, and **6** were 6.84, 6.92, and 7.54 nm, respectively, which is consistent with the small r dependence of the probability distributions in Figure 4C. These values are similar to the experimental values of 6.4−6.5 nm (Table 2); however, the small increases in interchromophore distance seen in the calculations, especially from **5** to **6**, are not captured in the steady-state experiments.

$$r_a = \sqrt[6]{\dfrac{1}{\dfrac{\sum_1^n r^{-6}}{n}}}$$

(4)

Several structural differences between each dimeric megamolecule were observed (Figure S12). The distance and angle fluctuations can be quite large on the 1−100 ns time scale, such that the r^{-6} weighting of the steady-state experiments misses many dynamic features of the distribution in Figure 4C. To further understand the conformations of **4−6**, we built a geometric model composed of four vertices for each protein domain and three edges for each polypeptide linker or cross-linker (Figure S12A). Again, no significant difference was observed for the length of the cross-linker edge between cutinase and SnapTag for all dimer megamolecules, which was 3.3−3.5 nm (Figure 4B). However, the length of the two polypeptide edges (between sfGFP and cutinase, and mVenus and SnapTag) was both variable and different for each dimer, ranging from 4.5−6.5 nm for the sfGFP−cutinase edge and 4.2−6.5 nm for the mVenus−SnapTag edge (Figure S12D,E). In particular, the mVenus−SnapTag distance increases systematically from **4** to **6** (Figure S12E). In addition, the angles between each neighboring edge were also both variable and different for each dimer, ranging from 50° to 150° for the angle between the sfGFP−cutinase polypeptide edge and the cross-linker edge (θ) and from 50° to 100° for the angle between the mVenus−SnapTag polypeptide edge and the cross-linker edge (φ). Finally, the angle between the two polypeptide edges (ψ) was variable and different for each dimer, ranging from −120° to 120° (Figure S12F−H). A systematic difference between the dimers is that the ψ angle switches from negative to positive (Figure S12H). Taken together, the interchromophore distances agreed well with the steady-state fluorescence data, although structures of **4−6** were not superimposable.

3.4. Two-Photon Excited Fluorescence. Two-Photon excited fluorescence (TPEF) spectra were measured by exciting all samples at 820 nm and 50 mW of power (see Figure S13). The two-photon excited fluorescence spectra of **1** and **2** revealed maxima at 527 and 510 nm, respectively. The spectrum of **3** also revealed a peak at 510 nm, although it is wider than that of **2** by approximately 150 cm^{-1} (see full width at half maximum, fwhm, in Table 3). The two-photon excited fluorescence spectra of the dimeric megamolecules (**4−6**) revealed maxima at 510 nm and a peak widening of approximately 390 cm^{-1}, relative to **2**. The spectrum of the trimeric megamolecule (**7**) showed a peak at 510 nm and a peak widening of approximately 250 cm^{-1}, relative to **2**. For all samples, the TPEF spectra of each molecule showed trends similar to the steady-state emission spectra. A key difference,

Table 3. Two-Photon Absorption Properties of Megamolecules at 820 nm Excitation

no.	sample	TPEF λ_{max}(nm)	TPEF fwhm (cm^{-1})	TPA cross-section (GM)
1	mVenus−SnapTag	527	1150	17
2	sfGFP−cutinase	504	1270	88
3	equimolar mixture	510	1420	162
4	EG7 dimer	510	1660	142
5	EG9 dimer	510	1660	437
6	EG11 dimer	510	1660	326
7	trimer	510	1520	642
8	tetramer	504	1420	172
9	pentamer	504	1420	187
10	hexamer	504	1420	113

however, was that the shoulders observed in the steady-state spectra were manifested as broadening in the TPEF spectra, due to the lower wavelength resolution (fluorescence counts are collected every 7 nm in the TPEF experiment).

In a two-photon excited-fluorescence experiment, the fluorescence intensity should have a quadratic dependence on excitation power.[44] As expected, a log−log plot of intensity (counts per second) versus beam power (mW) gave a linear fit with a positive slope of two for all samples tested (Figure S14), indicating that they all contained two-photon-absorbing fluorophores. Using this information, and the corresponding y-intercept of the plot,[45] we calculated two-photon absorption cross sections for each sample (Table 3). For **1** and **2**, the calculated cross sections were 17 and 88 GM, respectively. The calculated TPA cross sections of the dimeric and trimeric megamolecules (**4−7**) were 142, 437, 326, and 642 GM. Because the emission detected was that of the donor, these values represent enhancements of the cross section of sfGFP in the presence of mVenus in the megamolecules. For these samples (**4−7**), we see an enhancement of the TPA cross-section of sfGFP by a factor of 1.6, 4.9, 3.7, and 7.3, respectively. The TPA cross sections for the larger megamolecules (**8−10**) exhibited values of 172, 187, and 113 GM, respectively, representing enhancements in the cross-section of sfGFP by factors of 2.0, 2.1, and 1.3, respectively. Therefore, in all megamolecules, we observed larger two-photon absorption cross sections than in sfGFP−cutinase.

3.5. Time-Correlated Single Photon Counting. Time-correlated single photon counting (TCSPC) was used to study the long-lived component of the fluorescence decay kinetics of the samples. This experiment was performed under 400 nm excitation to achieve selective excitation of the donor. Fluorescence of **3−7** was detected at 527 nm, the acceptor peak emission wavelength, to observe emission from the acceptor in the samples upon donor excitation. The decays were plotted as the log of fluorescence intensity versus time, and then fitted to a linear decay to give the decay lifetime (Figure S15). These lifetime values are summarized in Table S3. Investigation of the emission wavelength effect on the lifetime obtained in the case of the trimer (see Table S4) confirmed that detecting emission at the acceptor peak is an appropriate choice. It allows one to obtain the same decay time revealed for longer emission wavelengths but with a better signal-to-noise ratio in the recorded kinetics. The fluorescence lifetime of the equimolar mixture **3** (3.36 ns) was shorter than

Table 4. Time-Resolved Fluorescence Up-Conversion Data for Megamolecules

no.	sample	λ_{em}(nm)	A_1	τ_1(ps)	A_2	τ_2(ps)	A_3	τ_3(ps)[a]
1	mVenus−SnapTag	527	0.60	9	0.28	200	0.44	3530
2	sfGFP−cutinase	510	−0.17	8	0.42	200	0.6	3060
3	equimolar mixture	510	−0.14	10	0.51	180	0.60	3360
		527	−0.20	2	0.44	210	0.57	3360
4	dimer EG7	527	−0.12	2	0.64	250	0.44	3530
5	dimer EG9	527	−0.28	3.7	0.64	360	0.40	3470
6	dimer EG11	527	−0.27	36	0.78	500	0.40	3470
7	trimer	527	−0.60	60	0.90	230	0.57	3440
8	tetramer	527	−0.82	0.70	0.25	74	0.78	3390
9	pentamer	527	−0.29	5.4	0.55	430	0.51	3340
10	hexamer	527	−0.18	3.6	0.16	140	0.81	3270

[a]During the fluorescence up-conversion data fitting, τ_3 was fixed to the value obtained by the time-correlated single photon counting experiments (see Table S3).

that of the acceptor **1** (3.53 ns) and longer than that of the donor **2** (3.06 ns), suggesting a low level of energy transfer. The fluorescence lifetime for **4** matched that of acceptor **1**, with a decay time of 3.53 ns, pointing to significant energy transfer. The fluorescence lifetimes of **5−7** were slightly shorter (3.44−3.47 ns) than that of **1**, again suggesting significant energy transfer. The fluorescence lifetimes of **8−10** were shorter (3.27−3.39 ns) and similar to that of **3**.

3.6. Femtosecond Time-Resolved Fluorescence Up-Conversion. Ultrafast fluorescence dynamics were investigated by fluorescence up-conversion. We again note that in megamolecules having both donor and acceptor, there was selective excitation of the donor, as the experiment was performed upon 400 nm excitation of the samples. The best correlation of the experimental decays was achieved with a triexponential fit, setting the long component of the decay equal to the corresponding lifetime found in TCSPC experiments. The lifetimes and the amplitudes for the three exponential components of all samples are summarized in Table 4.

3.6.1. Monomers/Unlinked Donor and Acceptor (1−3). Fluorescence of **1** and **2** was collected at their respective emission peaks and served as standards. Kinetics for the equimolar mixture (**3**) were measured at 510 and 527 nm to detect any differences in the properties of **1** and **2** in the presence of one another (Figure 5). Negative amplitude in the exponential decay function of **2** indicated a rise-time of 8 ps, due to excited-state proton transfer within the fluorophore region of the fluorescent protein,[9] and not related to energy transfer. No rise-time was found in **1**. The rise-time of **3** at 510 nm fluorescence detection was 10 ps, similar to **2**, suggesting that the dynamics of sfGFP in the presence and absence of mVenus were similar to one another. A rise-time of 2 ps was found for **3** at 527 nm detection, likely due to the influence of the intense fluorescence of the donor even at the acceptor emission wavelength.

3.6.2. Linked Conjugates (4−10). Fluorescence kinetics for megamolecules **4−10** were collected at 527 nm to determine if fluorescence of the acceptor was observed when only the donor was excited (Figure 6). In the case of the trimer, investigation at the donor emission wavelength revealed small donor quenching as expected for a low efficiency energy transfer (see Figure S16 and Table S4). The initial amplitude A_1 was negative for all samples, indicating a rise-time. Rise-times calculated from this experiment indicate energy transfer. Additionally, **4−7** showed a trend of increasing rise-time (2,
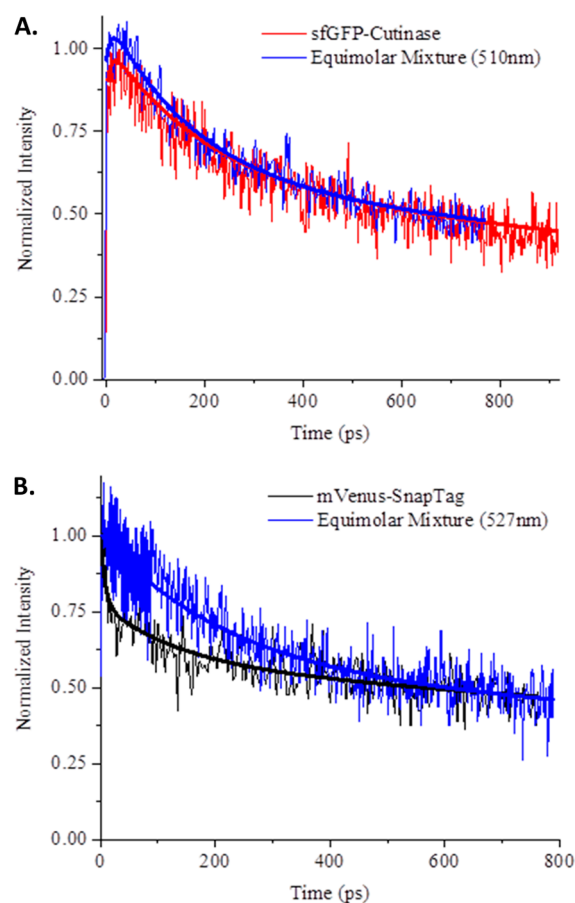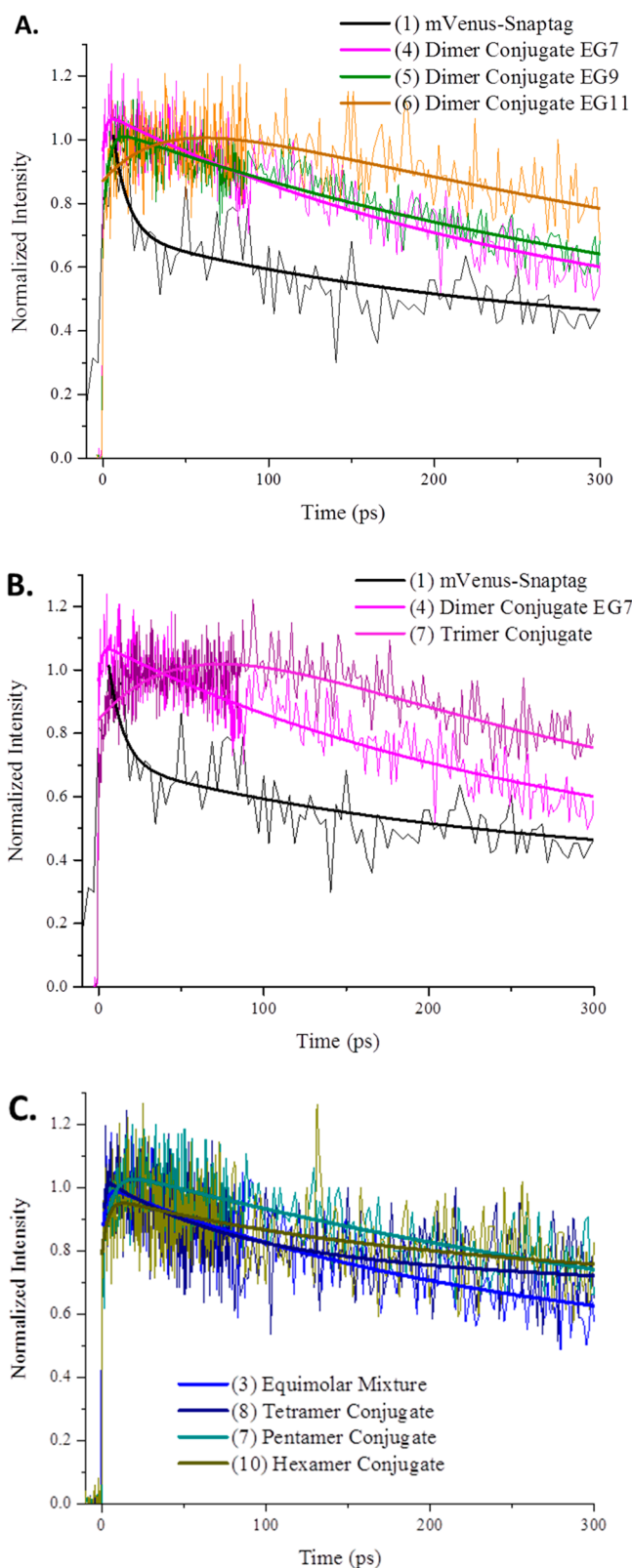

Figure 5. Time-resolved fluorescence decay curves of **3** overlaid with (A) sfGFP−cutinase and (B) mVenus−SnapTag.

3.7, 36, and 60 ps) and amplitude (−0.12, −0.28, −0.27, and −0.60) with increasing megamolecule size, with **7** showing the longest rise-time. From the rise-time values, energy transfer rates ($k_{FRET} = 1/\tau_1$) were computed, which exhibit values of 5.0, 2.7, 0.27, and 0.16 × 10^{11} s^{-1} for **4−7**, respectively. This trend of rise-times did not hold for larger oligomeric megamolecules (**8−10**), which featured shorter rise-times (0.70, 5.4, and 3.6 ps).

4. DISCUSSION

In this investigation, the first clear indication of energy transfer in these protein megamolecules came from the steady-state

Figure 6. Time-resolved fluorescence decay curves at 527 nm for (A) each of the three dimer samples as compared to **1**; (B) the EG7 dimer and trimer conjugate as compared to **1**; and (C) **8**−**10** as compared to **3**.

fluorescence spectra. Here, significant emission from the acceptor protein was observed for the megamolecules. This points to excitation energy being transferred to the acceptor

after absorption by the donor. In addition, we observed fluorescence quantum yield quenching in the megamolecules with respect to the donor protein. Another requirement of energy transfer is coupling between the dipoles of donor and acceptor groups.[46] In our two-photon absorption experiments, we verified this coupling between the fluorescent proteins in the megamolecules. The two-photon absorption cross section of a molecule is squarely proportional to its change in transition dipole moment.[42] Thus, a larger cross section implies a larger change in transition dipole moment. We would expect to see the strength of the interaction between donor and acceptor proteins weaken as the distance between them increases, and we should observe significantly smaller cross sections as we increase megamolecule size. However, this is not the case, as we see an increase in the two-photon absorption cross section in larger megamolecules up to the trimeric species. For this trend to be observed, the donor and acceptor dipoles in these megamolecules must be strongly coupled to one another.

Cooperative two-photon absorption enhancement has been repeatedly observed in multichromophoric systems with respect to the isolated chromophores. This has usually been revealed in conjugated branched or conjugated dendritic structures as a result of through-bond coherent coupling between chromophoric subunits. However, TPA enhancement has also been reported for multichromophoric compounds where two or more chromophores are linked through nonconjugated, covalent linkers (e.g., saturated bonds).[47] This has been justified considering through-space interactions between the monomeric subunits in the dimers. In our opinion, through-space interactions between sfGFP and mVENUS should be considered to rationalize the TPA cross section enhancement observed in the megamolecules with respect to the fluorescent protein monomers. In a previous work, Clark et al. discussed the case of amyloid peptides labeled with the GFP chromophore.[48] The exhibited TPA dependence on peptide aggregation was explained considering the relative orientations of the transition dipole moments. According to the model used for this work, a decrease or an increase in TPA probability with shorter separations was expected depending on parallel or linear orientation of the transition dipole moments of the peptides, respectively. In our opinion, the increase of TPA cross section upon increasing megamolecule size observed in the dimers and in the trimer here investigated can be explained considering a parallel orientation of the transition dipole moments of the donor and acceptor chromophores in these conjugates. Further insight into the dynamics of the donor−acceptor interaction was investigated by time-resolved experiments.

Our femtosecond time-resolved fluorescence up-conversion data showed energy transfer between donor and acceptor fluorescent proteins across tunable distances in these large protein megamolecules. A trend of increasing rise-time upon increasing megamolecule size was observed for dimeric and trimeric species (**4**−**7**), indicating a slower energy transfer rate with increasing megamolecule size. Thus, the time-resolved behavior gives clear evidence for an increase in energy transfer length upon increasing megamolecule size and achieves long-range energy transfer in megamolecules. Interestingly, we still observed energy transfer in the larger conjugates even though the hypothesized distance between donor and acceptor in **7**−**10** was greater than 10 nm, outside of the normal FRET range. To understand the reason behind such behavior, the ultrafast

spectroscopy results should be considered in conjunction with the steady-state data.

Energy transfer efficiencies computed from the steady-state spectra were found to be quite low (≤33%), as is expected for long donor–acceptor distances. Additionally, efficiencies were insensitive to a change in the cross-linker length of the dimers (**4**–**6**). The dimers were also indistinguishable by size-exclusion chromatography (Figure 2B). However, both of these measurements sample molecular conformations at long time scales (≥μs). To understand the lack of effect of the cross-linker length on energy transfer efficiency, we performed an all-atom molecular dynamics simulation of each dimer megamolecule to understand dynamics on the ns time scale. Each trajectory resulted in the formation of a non-native protein–protein contact between the cutinase and SnapTag domains, independent of cross-linker length. The cross-linker drove this interaction by restraining the diffusion of SnapTag and cutinase. For the dimers, the Förster-adjusted interchromophore distances compared well with the experimental values of 6.4–6.5 nm. This is consistent with the bent U-shaped structure determined by molecular dynamics simulations (Figure S12), although we note that the calculations clearly show the r^{-6} weighted distance increasing in going from **4** to **6**, consistent with the time-resolved results. The trimer megamolecule (**7**) was calculated to have a donor–acceptor distance 1.0 nm greater than the dimers based on steady-state fluorescence, much shorter than hypothesized. These results imply a bent conformation for the trimer megamolecule, as would be expected from the simulations of the dimers. In addition, the tetramer, pentamer, and hexamer all exhibited similar low energy transfer efficiencies. Such behavior suggests that steady-state measurements were not sensitive to potential differences in chromophore distance for **8**–**10**. Interestingly, the hydrodynamic radius increased with oligomeric state by both size-exclusion chromatography (Figure 2B) and dynamic light scattering (Figure 2C), suggesting that the solution phase size of these molecules is increasing. However, conformations with shorter interchromophore distances are more heavily weighted in FRET measurements, and conformational fluctuations during the excited state lifetime may bias energy transfer measurements to shorter length scales. Taken together, this implies that the hydrodynamic radius of the ensemble of conformations is increasing with oligomeric state, but the conformations of each oligomer with short interchromophore distances and the probability of these states are similar. These results suggest a dynamic bent structure, with end-to-end lengths shorter than hypothesized. From the simulation data, we built a geometric model of dimer megamolecules (**4**–**6**) (Figure S12A). The shape of these structures is a relatively constrained and dynamic U-shaped structure. We modeled the four protein domains as vertices and the three linkers (polypeptide and cross-linker) as edges. The distance between cutinase and SnapTag was similar for the three structures, while all other geometric measurements showed differences between structures, as well as dynamics in the same structure. The effect of the cross-linker length on the structure had effects distal from the cross-linker; for example, the distance between the mVenus and SnapTag domains in the fusion protein showed a strong effect on cross-linker length, although these domains are connected though a polypeptide linker that did not change chemically between the dimers (Figure S12E). Each structure featured small fluctuations on the nanosecond time scale of <1 nm between adjacent protein domains (Figure

S12C–E), and <50° between linkers (Figure S12F–H), suggesting structural constraints that restrict megamolecule dynamics. Thus, it is likely that the cross-linker drives a protein–protein contact between cutinase and SnapTag that is different for each dimer, such that the fluorescent protein domains access different conformations in the different dimers, leading to differences in the interchromophore distance and the angle of the dipole moments. The MD simulations show that megamolecules exhibit constrained structural dynamics with large fluctuations on the nanosecond time scale.

The steady-state fluorescence measurements showed no difference in energy transfer efficiency between the three dimers (**4**–**6**); however, we observed an increasing time constant (τ_1) with increasing cross-linker length by time-resolved fluorescence spectroscopy, suggesting a greater interchromophore distance.[49] While these results may seem incompatible, we note that the measurements capture different phenomena. The steady-state fluorescence experiments measure the average fluorescence intensity over the nanosecond lifetime of the donor–acceptor system, while the time-resolved experiments measure energy transfer on the picosecond time scale[50,51] and which is largely dominated by distances and transition dipole directions where the energy transfer rate is high.[52,53] On the basis of the results in Figures 4 and S12, these structures with fast transfer rates are infrequently sampled. Further, we found large fluctuations in structure on the nanosecond time scale in the simulations. This means that the time-resolved measurements average over a broad distribution of donor–acceptor distances and transition moment directions that are accessible within picoseconds of the ultrafast excitation, while nanosecond fluctuations of the excited megamolecule are not significant. This should suppress energy transfer for **6** as compared to **4** in the time-resolved measurements, as **6** involves larger interchromophore distances. In contrast, the nanosecond fluctuations (that lead to short donor–acceptor distances) are likely to play a more important role in the steady-state experiments. In this case, the energy transfer efficiency is less dependent on cross-linker length, as a fluctuation that leads to a short donor–acceptor distance is more likely on the nanosecond time scale than on the picosecond time scale. In future work, it will be important to perform excited-state dynamics studies to quantify this conclusion.

5. CONCLUSIONS

This Article reports the synthesis and characterization of megamolecules that incorporate fluorescent donors and acceptors, and where the defined connectivity and large sizes enabled a structure–function study of energy transfer across long and tunable distances. Two-photon absorption measurements revealed strong dipole–dipole coupling between donor and acceptor proteins in the megamolecules. Additionally, our time-resolved results showed an increase in energy transfer length upon increasing megamolecule size. Even though, according to our synthetic strategy, the theoretical lengths are outside of the FRET range, we still observed energy transfer in the larger megamolecules. Our results suggest that the flexibility and conformational dynamics of the megamolecules create favorable orientations between donor and acceptor proteins in which energy transfer occurs. The dynamic behavior of these structures on the 1–100 ns time scale allows for long-range energy transfer in the large protein megamolecules.

■ AUTHOR INFORMATION

Corresponding Authors
*g-schatz@northwestern.edu
*milan.mrksich@northwestern.edu
*tgoodson@umich.edu

ORCID ⓘ
Benedetta Carlotti: 0000-0002-2980-2598
Cheng-Tsung Lai: 0000-0002-1192-2815
George C. Schatz: 0000-0001-5837-4740
Milan Mrksich: 0000-0002-4964-796X
Theodore Goodson, III: 0000-0003-2453-2290

Author Contributions
⊥E.L.T. and K.J.M. contributed equally.

Notes
The authors declare no competing financial interest.

■ REFERENCES

(1) Förster, T. Intermolecular Energy Migration and Fluorescence. *Ann. Phys.* **1948**, *437*, 55−75.

(2) Truong, K.; Ikura, M. The use of FRET imaging microscopy to detect protein-protein interactions and protein conformational changes *in vivo*. *Curr. Opin. Struct. Biol.* **2001**, *11*, 573−578.

(3) Patterson, G. H.; Piston, D. W.; Barisas, B. G. Forster Distances between Green Fluorescent Protein Pairs. *Anal. Biochem.* **2000**, *284*, 438−440.

(4) Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y. Improved monomeric red, orange, and yellow fluorescent proteins derived from Discosoma sp. Red fluorescent protein. *Nat. Biotechnol.* **2004**, *22* (12), 1567−1573.

(5) Bajar, B. T.; Wang, E. S.; Zhang, S.; Lin, M. Z.; Chu, J. A Guide to Fluorescent Protein Pairs. *Sensors* **2016**, *16*, 1488.

(6) Blacker, T. S.; Chen, W.; Avezov, E.; Marsh, R. J.; Duchen, M. R.; Kaminski, C. F.; Bain, A. J. Investigating State Restriction in Fluorescent Protein FRET Using Time-Resolved Fluorescence and Anisotropy. *J. Phys. Chem. C* **2017**, *121*, 1507−1514.

(7) Goedhart, J.; Van Weeren, L.; Hink, M. A.; Vischer, N. O. E.; Jalink, K.; Gadella, T. W. J., Jr. Bright cyan fluorescent protein variants identified by fluorescence lifetime screening. *Nat. Methods* **2010**, *7* (2), 137−139.

(8) Miyawaki, A.; Tsien, R. Y. Monitoring Protein Conformations and Interactions by Fluorescence Resonance Energy Transfer between Mutants of Green Fluorescent Protein. *Methods Enzymol.* **2000**, *327*, 472−500.

(9) Cotlet, M.; Goodwin, P. M.; Waldo, G. S.; Werner, J. H. A comparison of the Fluorescence Dynamics of Single Molecules of a Green Fluorescent Protein; One- versus Two-Photon Excitation. *ChemPhysChem* **2006**, *7*, 250−260.

(10) Bruns, T.; Angres, B.; Steuer, H.; Weber, P.; Wagner, M.; Schneckenburger, H. Forster Resonance energy transfer-based total internal reflection fluorescence reader for apoptosis. *J. Biomed. Opt.* **2009**, *14* (2), 012003.

(11) Wang, Y.; Ranasinghe, M. I.; Goodson, T., III Ultrafast Fluorescence Investigation of Excitation Energy Transfer in Different Dendritic Core Branched Structures. *J. Am. Chem. Soc.* **2003**, *125*, 9562−9563.

(12) Ranasinghe, M. I.; Varnavski, O.; Pawlas, J.; Hauck, S. I.; Louie, J.; Hartwig, J. F.; Goodson, T., III Femtosecond Excitation Energy Transport in Triarylamine Dendrimers. *J. Am. Chem. Soc.* **2002**, *124*, 6520−6521.

(13) Yun, C. S.; Javier, A.; Jennings, T.; Fisher, M.; Hira, S.; Peterson, S.; Hopkins, B.; Reich, N. O.; Strouse, G. F. Nanometal Surface Energy Transfer in Optical Rulers, Breaking the FRET Barrier. *J. Am. Chem. Soc.* **2005**, *127*, 3115−3119.

(14) Cooper, J. P.; Hagerman, P. J. Analysis of Fluorescence Energy Transfer in Duplex and Branched DNA molecules. *Biochemistry* **1990**, *29*, 9261−9268.

(15) Cunningham, Paul; Khachatrian, A.; Buckhout-White, S.; Deschamps, J. R.; Goldman, E. R.; Medintz, I. L.; Melinger, J. S. Resonance Energy Transfer in DNA Duplexes Labeled with Localized Dyes. *J. Phys. Chem. B* **2014**, *118*, 14555−14565.

(16) Krainer, G.; Hartmann, A.; Schlierf, M. farFRET: Extending the Range in Single-molecules FRET Experiments beyond 10nm. *Nano Lett.* **2015**, *15*, 5826−5829.

(17) Komatsua, N.; Aoki, K.; Yamadac, M.; Yukinagac, H.; Fujitac, Y.; Kamioka, Y.; Matsuka, M. Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Mol. Biol. Cell* **2011**, *22*, 4647−4656.

(18) Mizuno, H.; Sawano, A.; Eli, P.; Hama, H.; Miyawaki, A. Red Fluorescent Protein from Discosoma as a fusion tag and a Partner for Fluorescence Resonance Energy Transfer. *Biochemistry* **2001**, *40*, 2502−2510.

(19) Stephanopoulos, N.; Francis, M. B. Choosing an Effective Protein Bioconjugation Strategy. *Nat. Chem. Biol.* **2011**, *7*, 876−884.

(20) Modica, J. A.; Skarpathiotis, S.; Mrksich, M. Modular Assembly of Protein Building Block to Create Precisely Defined Megamolecules. *ChemBioChem* **2012**, *13*, 2331−2334.

(21) Modica, J. A.; Lin, Y.; Mrksich, M. Synthesis of Cyclic Megamolecules. *J. Am. Chem. Soc.* **2018**, *140*, 6391−6399.

(22) Vogel, S. S.; Wieb van der Meer, B.; Blank, P. S. Estimating the distance separating fluorescent protein FRET pairs. *Methods* **2014**, *66*, 131−138.

(23) Engler, C.; Gruetzner, R.; Kandzia, R. Marillonnet. Golden Gate Shuffling, A one-pot DNA Shuffling Method Based on Type IIs Restriction Enzymes. *PLoS One* **2009**, *4* (5), e5553.

(24) Furgal, J. C.; Jung, J. H.; Goodson, T.; Laine, R. M. Analyzing Structure-Photophysical Property Relationships for Isolated T8, T10, and T12 Stilbenevinylsilsesquioxanes. *J. Am. Chem. Soc.* **2013**, *135*, 12259−12269.

(25) Kim, D.; Sambasivan, S.; Nam, H.; Kim, K. H.; Kin, J. Y.; Joo, T.; Lee, K.; Kim, K.; Ahn, K. H. Reaction-based two-photon probes for *in vitro* analysis and cellular imaging of monoamine oxidase activity. *Chem. Commun.* **2012**, *48*, 6833−6835.

(26) Doan, P.; Pitter, D. R. G.; Kocher, A.; Wilson, J. N.; Goodson, T., III A New Design Strategy and Diagnostic to Tailor the DNA-Binding Mechanism of Small Organic Molecules and Drugs. *ACS Chem. Biol.* **2016**, *11*, 3202−3213.

(27) Xu, C.; Webb, W. W. Measurement of two-photon excitation cross sections of molecular fluorophores with data from 690 to 1050 nm. *J. Opt. Soc. Am. B* **1996**, *13*, 481.

(28) Ho-Wu, R.; Sun, K.; Goodson, T., III Synthesis and enhanced Linear and Nonlinear Optical properties of chromophore-Au Metal Cluster Oligomers. *J. Phys. Chem. C* **2018**, *122* (4), 2315−2329.

(29) Lahankar, S. A.; West, R.; Varnavski, O.; Xie, X.; Goodson, T.; Sukhomlinova, L.; Twieg, R. Electronic interactions in a branched chromophore investigated by nonlinear optical and time-resolved spectroscopy. *J. Chem. Phys.* **2004**, *120*, 337−344.

(30) Varnavski, O.; Samuel, I. D. W.; Palsson, L. O.; Beavington, R.; Burn, P. L.; Goodson, T. Investigations of Excitation Energy Transfer and Intramolecular Interactions in a Nitrogen Corded Distrylbenzene Dendrimer System. *J. Chem. Phys.* **2002**, *116*, 8893−8903.

(31) Adegoke, O. O.; Jung, I. H.; Orr, M.; Yu, L.; Goodson, T. 3rdEffect of Acceptor Strength on Optical and Electronic Properties in Conjugated Polymers for Solar Applications. *J. Am. Chem. Soc.* **2015**, *137*, 5759.

(32) Vázquez, R. J.; Kim, H.; Kobilka, B. M.; Hale, B. J.; Jeffries-El, M.; Zimmerman, P.; Goodson, T., III Evaluating the Effects of Heteroatoms on the Photophysical Properties of Donor-Acceptor Conjugated Polymers Based on 2,6-Di(thiophen-2-yl)benzo[1,2-b:4,5-b"]difuran: Two-Photon Cross-Section and Ultrafast Time-Resolved Spectroscopy. *J. Phys. Chem. C* **2017**, *121*, 14382−14392.

(33) Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat. Protoc.* **2015**, *10*, 845−858.

(34) Raman, S.; Vernon, R.; Thompson, J.; Tyka, M.; Sadreyev, R.; Pei, J.; Kim, D.; Kellogg, E.; DiMaio, F.; Lange, O.; Kinch, L.; Sheffler, W.; Kim, B.; Das, R.; Grishin, N. V.; Baker, D. Structure prediction for CASP8 with all-atom refinement using Rosetta. *Proteins: Struct., Funct., Genet.* **2009**, *77* (59), 89−99.

(35) Nguyen, H.; Pérez, A.; Bermeo, S.; Simmerling, C. Refinement of Generalized Born Implicit Solvation Parameters for Nucleic Acids and Their Complexes with Proteins. *J. Chem. Theory Comput.* **2015**, *11* (8), 3714−3728.

(36) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., Jr. General atomic and molecular electronic structure. *J. Comput. Chem.* **1993**, *14* (11), 1347−1363.

(37) Bayly, C. I.; Ciepak, P.; Cornlee, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97* (40), 10269−10280.

(38) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general force field. *J. Comput. Chem.* **2004**, *25* (9), 1157−1174.

(39) Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Grand, S. L.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* **2012**, *8* (5), 1542−1555.

(40) Yoshida, H.; Takikawa, K.; Kaneko, I.; Matsuura, H. Conformational Analysis of Model Compounds of non-ionic Surfactants in Aqueous Solution: Raman Spectroscopy and ab initio MO Calculations. *J. Mol. Struct.: THEOCHEM* **1994**, *311*, 205−210.

(41) Rekas, A.; Alattia, J.; Nagai, T.; Miyawaki, A.; Ikura, M. Crystal structure of Venus, A yellow fluorescent protein with improved maturation and reduced environmental sensitivity. *J. Biol. Chem.* **2002**, *277*, 50573−50578.

(42) Nagai, T.; Ibata, K.; Park, E. S.; Kubota, M.; Mikoshiba, K.; Miyawaki, A. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **2002**, *20*, 87−90.

(43) Valeur, B.; Berberan-Santos, M. N.; Berberan-Santos, M. N. *Molecular Fluorescence: Principles and Applications*; John Wiley & Sons, Inc.: Weinheim, Germany, 2013.

(44) Drobizhev, M.; Makarov, N. S.; Tillo, S. E.; Hughes, T. E.; Rebane, A. Two-photon absorption properties of fluorescent proteins. *Nat. Methods* **2011**, *8* (5), 393−399.

(45) Bhaskar, A.; Ramakrishna, G.; Lu, Z.; Twieg, R.; Hales, J. M.; Hagan, D. J.; Van Stryland, E.; Goodson, T., III Investigation of Two-Photon Absorption Properties in Branched Alkene and Alkyne Chromophores. *J. Am. Chem. Soc.* **2006**, *128*, 11840−11849.

(46) Piston, D. W.; Kremers, G. Fluorescent Protein FRET: the good, the bad, and the ugly. *Trends Biochem. Sci.* **2009**, *32* (9), 407−414.

(47) Terenziani, F.; Parthasarathy, V.; Pla-Quintana, A.; Maishal, T.; Caminade, A.; Majoral, J.; Blanchard-Desce, M. *Angew. Chem.* **2009**, *121* (46), 8847−8850.

(48) Clark, T. B.; Ziólkowski, M.; Schatz, G. C.; Goodson, T., III Two-Photon and Time-Resolved Fluorescence Spectroscopy as Probes for Structural Determination in Amyloid-*β* Peptides and Aggregates. *J. Phys. Chem. B* **2014**, *118* (9), 2351−2359.

(49) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd ed.; Springer: New York, 2003.

(50) Beechem, J. M.; Brand, L. Time-Resolved Fluorescence of Proteins. *Annu. Rev. Biochem.* **1985**, *54*, 43−71.

(51) Amaro, M.; Šachl, R.; Jurkiewicz, P.; Coutinho, A.; Prieto, M.; Hof, M. Time-Resolved Fluorescence in Lipid Bilayers: Selected Applications and Advantages over Steady-State. *Biophys. J.* **2014**, *107* (12), 2751−2760.

(52) Khrenova, M.; Topol, I.; Collins, J.; Nemukhin, A. Estimating Orientation Factors in the FRET Theory of Fluorescent Proteins: The TagRFP-KFP Pair and Beyond. *Biophys. J.* **2015**, *108* (1), 126−132.

(53) Ansbacher, T.; Srivastava, H. K.; Stein, T.; Baer, R.; Merkx, M.; Shurki, A. Calculation of Transition Dipole Moment in Fluorescent Proteins − Towards Efficient Energy Transfer. *Phys. Chem. Chem. Phys.* **2012**, *14*, 4109−4117.