Association of Fluorescent Protein Pairs and Its Significant Impact on Fluorescence and Energy Transfer


Fluorescent proteins (FPs) are commonly used in pairs to monitor dynamic biomolecular events through changes in proximity via distance dependent processes such as Förster resonance energy transfer (FRET). The impact of FP association is assessed by predicting dimerization sites in silico and stabilizing the dimers by bio-orthogonal covalent linkages. In each tested case dimerization changes inherent fluorescence, including FRET. GFP homodimers demonstrate synergistic behavior with the dimer being brighter than the sum of the monomers. The homodimer structure reveals the chromophores are close with favorable transition dipole alignments and a highly solvated interface. Heterodimerization (GFP with Venus) results in a complex with ≈87% FRET efficiency, significantly below the 99.7% efficiency predicted. A similar efficiency is observed when the wild-type FPs are fused to a naturally occurring protein–protein interface system. GFP complexation with mCherry results in loss of mCherry fluorescence. Thus, simple assumptions used when monitoring interactions between proteins via FP FRET may not always hold true, especially under conditions whereby the protein–protein interactions promote FP interaction.

1. Introduction

1.1. Introduction

Fluorescent proteins (FPs) have revolutionized biology through their use as genetically encoded imaging tags and biosensors.\(^{[1–3]}\) The subsequent engineering of a small subset of natural FPs,\(^{[1]}\) especially green fluorescent protein (GFP) from *Aequorea victoria*\(^{[4]}\) and DsRed from coral\(^{[3]}\) have expanded their use by changing their spectral (e.g., \(\lambda_{\text{max}}, \lambda_{\text{EM}}\), quantum yield, brightness) and structural (e.g., quaternary structure, stability, folding kinetics, chromophore maturation kinetics) properties. Many fluorescent proteins, especially those that emit in the red region, naturally exist as oligomers\(^{[5]}\) or have a tendency to oligomerize.\(^{[6]}\) There has been a great deal of protein engineering effort to generate functional monomeric forms but many commonly used FPs have a capacity to dimerize.\(^{[6,7]}\) Dimerization can be compounded by local high concentrations brought about by interactions between the fusion partner proteins that are the prime focus of such studies. Yet there is little information concerning how oligomerization influences inherent function; oligomerization potential and functional impact is especially important for studies involving FP pairs, such as Förster resonance energy transfer (FRET).

FRET is one of the most important applications of FP pairs as it can be used to monitor dynamic biological events such as...
protein–protein interactions.[8,9] FRET is largely a passive process that relies on two FPs with mutually compatible spectral properties (acceptor FP absorbance overlapping with donor FP emission wavelength) being in close proximity but not physically interacting; changes in distance between the two FPs changes efficiency of FRET between the donor and acceptor. Despite FRET being a mainstay of biomolecular interaction analysis, there are several assumptions required such as freely rotating FPs that do not interact or align in any significant manner. Here, we show that both these factors may not always hold true.

As well as absolute distance between the FPs, the angular vector between the chromophore dipoles is critical; this is \( \kappa^2 \) value in Equation (1).

\[
R_0 = 0.211 \sqrt{k^2 n^4 Q_0 J(\lambda)}
\]

where \( R_0 \) is the Förster radius, \( \kappa^2 \) is the dipole orientation factor, \( n \) is the solvent refractive index, \( Q_0 \) is the quantum yield of the donor, and \( J(\lambda) \) is the overlap integral between the donor emission and acceptor molar absorbance. \( R_0 \) is used as a constant to relate energy transfer efficiency to distance between individual components via Equation (2).

\[
r = R_0 \sqrt{(1 - E) / E}
\]

where \( r \) is the distance between two FRET chromophores and \( E \) is the observed FRET efficiency. Critically \( \kappa^2 \) is arbitrarily set to 0.667 to reflect two randomly orientated chromophores as defined distance and dipole alignment, which may in turn influence inherent fluorescence. It has previously been thought that by using FPs from different organism classes with low sequence identities (e.g., GFP with RFPs) should prevent dimerization.

We and others have previously shown that FP association can be promoted through either connecting FPs with linker sequences/protein domains, or by forming oligomers from individual monomers. In relation to the current work, we have shown that potential naturally occurring FP dimer interfaces as in turn stabilized through an SPAAC link. Regions that do not naturally associate do not promote covalent crosslinking via SPAAC. Thus, our approach stabilizes naturally feasible protein interactions.

We have previously reported the construction of FP dimers stabilized by click chemistry through the covalent coupling of genetically encoded ring-strained cyclooctyne derivative of the pyrroline (SCO-K) and \( p \)-azido-L-phenylalanine (azF)[11] (Figure 1a). It should be noted that we do not attempt to change FP dimer interface as a whole nor link them in a tandem arrangement using a spacer sequence as has been done in other approaches[12–16] but model potential naturally occurring interface sites, which are in turn stabilized through an SPAAC link. Regions that do not naturally associate do not promote covalent crosslinking via SPAAC.[11] Thus, our approach stabilizes naturally feasible protein interactions.

2. Results and Discussion

2.1. The Effect of sfGFP Association on Function

We have previously reported the construction of FP dimers stabilized by click chemistry through the covalent coupling of genetically encoded ring-strained cyclooctyne derivative of the pyrroline (SCO-K) and \( p \)-azido-L-phenylalanine (azF)[11] (Figure 1a). It should be noted that we do not attempt to change FP dimer interface as a whole nor link them in a tandem arrangement using a spacer sequence as has been done in other approaches[12–16] but model potential naturally occurring interface sites, which are in turn stabilized through an SPAAC link. Regions that do not naturally associate do not promote covalent crosslinking via SPAAC. Thus, our approach stabilizes naturally feasible protein interactions.

Residue Q204 is a surface exposed residue that lies close to the sfGFP chromophore (CRO; Figure 1b), with the backbone amine group making an indirect H-bond with CRO via a conserved structured water molecule, W1. In silico, molecular docking[11] revealed that Q204 consistently resided at possible dimer interfaces and is close to a region known to be involved in FP dimerization.[7] The SCO-K (sfGFP204SCO) and azF (sfGFP204azF) containing monomers were subsequently proved to dimerize, generating the dimer termed sfGFP204a2.[11] The sfGFP 204-linked dimer displayed enhanced fluorescence compared to the monomers. Dimeric sfGFP204a2 displayed positive functional synergy in which the brightness of the complex was more than the sum of the individual monomers (Figure 1c; Table S1, Supporting Information).[11] Indeed, sfGFP204a2 has a greater brightness on a per CRO basis (56 800 M\(^{-1}\) cm\(^{-1}\)) compared to the original sfGFP (36 750 M\(^{-1}\) cm\(^{-1}\)) and EGFP (34 650 M\(^{-1}\) cm\(^{-1}\)).[19] This change in fluorescence behavior is further confirmed at the single molecule level where the dimeric fluorophore is much more resistant to photobleaching, displaying longer on times compared to sfGFP\(^{WT}\) (average 0.87 s GFP204a2 compared to 0.65 s for GFP\(^{WT}\)). (Figure 1d with additional traces in Figure S1, Supporting Information; see Worthy et al. for sfGFP\(^{WT}\) single molecule analysis[11]).

Single molecule fluorescence traces of the dimer are more complex and dynamic compared to sfGFP\(^{WT}\) with a range of dynamically fluctuating fluorescent intensities observed, not well described by discrete states (Figure 1d; Figure S1, Supporting Information), which could indicate cooperative interaction between the individual monomer units. Such communication between conjoined CROs is further exemplified in the counter-intuitive ensemble fluorescence behavior of equivalently linked fluorescent protein heterodimers (vide infra). In the case of sfGFP204a2, the single molecule fluorescence time course traces are not consistent with expectations of two independent, co-localized fluorophores, with an absence of two-step photobleaching observed. Intensity state histograms compiled from single molecule traces reveal a single dominant intensity peak observed at a value similar to monomeric sfGFP\(^{WT}\) (Figure 1e). This is in contrast to the expected bimodal distribution of two independent fluorophores.
Figure 1. Click-based protein dimerization via residue 204. a) Covalent crosslink via genetically encoded p-azido-L-phenylalanine (azF) in one monomer and strained-cyclooctyne pyrrolysine (SCO-K) placed in the second monomer. Shown are the two different final regio-isomers: b) Relative positioning of residue Q204 with respect to chromophore (CRO) and local water molecules (red spheres W1 and W2). c) Steady state bulk absorbance (full line) and fluorescence emission (dashed line) of sfGFP204x2 (blue), sfGFP204azF (green), sfGFP204SCO (red) and the addition of the two monomer spectra (black). The data has been reported previously[11] and shown here for context. See Table S1, Supporting Information, for further spectral properties. d) Representative single molecule traces for sfGFP204x2 (blue) and sfGFPWT (green) measured by TIRF microscopy. Further example of single molecule traces can be found in Figure S1, Supporting Information, for sfGFP204x2. Equivalent single molecule analysis of sfGFPWT is described previously by Worthy et al.[11] e) A single molecule fluorescence intensity histogram for sfGFP204x2 consisting of 179 trajectories (2602 spots). The histogram data fit to a single log normal distribution centered around 100 counts.

If the two molecules in the dimer are acting largely independently of each other, a bimodal distribution would be expected arising from the combination of additive intensities of the two fluorophore states at any one time (i.e., ON/ON [2x intensity], OFF/ON and ON/OFF [1x intensity]). Consequently, the observed data is consistent with only 1 CRO in the dimer being fluorescent at any given time, and suggestive of a possible dependent activity relationship. The observations from single molecule and ensemble data suggest that the change in absorbance coefficient, increased resistance to photobleaching, and change in photo-dynamics observed as intensity fluctuations at the single molecule level, collectively give rise to an overall increase in ensemble fluorescence.

2.2. Structural Basis for Association-Based Effects

The structure of sfGFP204x2 (structural statistics in Table S2, Supporting Information) reveals that each monomer unit is similar to the original starting sfGFP (Figure S2a,b, Supporting Information), with the chromophore retaining planarity in each monomer unit of the dimer. The sfGFP204x2 dimer forms a quasi-symmetrical off-set “side-by-side” monomer arrangement (Figure 2a), which is promoted by formation of a syn 1,5 triazole link that generates a reverse turn structure (Figure 2b). The two CROs point toward each other in an antiparallel arrangement 22 Å apart with a 5Å offset (Figure 2c). It is closest to the third ranked in silico model predicted previously[11] (Figure S2c, Supporting Information). Each monomer is offset by 70° with the C-termini close in space (Figure 2b). As the N- and C-termini are close to each other at the same end of the β-barrel, the proximity and orientation of the two termini in the dimer may well promote such an interaction in a fusion protein construct.

The two monomer units associate to form an extensive (∼900 Å2) and intimate interface. The main elements that comprise a natural protein–protein interface, namely hydrophobic interactions and H-bonding are observed (Figure 3). The H-bond network at the interface is not symmetrical but the hydrophobic interactions show a significant degree of symmetry (Figure 3a). The interlocking hydrophobic interface is comprised of Phe223, Val206, Leu221 from both chains (Figure 3b). These residues are surface exposed in sfGFP and form a naturally occurring hydrophobic patch[7] that can facilitate and stabilize the dimer on
Figure 2. Structure of sfGFP<sup>204x2</sup> (PDB 5N13). a) Arrangement of the azF (green) and SCO (blue) containing monomers. b) The syn conformation of the triazole linkage with the electron density map (2Fo-Fc, 1.0 sigma) shown. c) Distances and offset of the two CROs (shown as sticks).

click crosslinking (Figure 3e), or for that matter potentially other FPs. Indeed, mutation of Val206 to a charged residue is known to reduce dimerization tendency of <i>A. victoria</i> derived GFPS.<sup>[7]</sup>

The new triazole crosslink is integrated within the structure being semi-buried at the dimer interface and lying above the plane of the main hydrophobic interface patch (Figure 3c,d). The azF component is buried while one face of the SCO moiety is partially accessible to the solvent. Phe223 from both monomers forms the base of the triazole reverse turn (Figure 3c,d) while Arg73, Ser202, and Thr225 residues make putative polar interactions with oxygen and nitrogen atoms in the SCO-azide link. A more extended network linking the two chromophores is proposed in Figure S3, Supporting Information.

At the interface are two cavities filled with ordered water molecules (Figure 3f,g). The water molecules are arranged around an area where the chromophore protrudes toward the surface. A partially buried water molecule (W1; Figure 1c; Figure S4, Supporting Information) is commonly observed associated with the chromophore via a H-bond with the phenol hydroxyl group and the backbone of residue 204; this water is associated with 1 to 2 additional surface water molecules (grey spheres, Figure 3f,g) as observed for monomeric sfGFP<sup>WT</sup> (Figure S4, Supporting Information). In the sfGFP<sup>204x2</sup> dimer, these waters lie within the cavity together with several additional tightly packed water molecules.

Generation of the dimer is likely to impact on both protein and solvent dynamics, which can in turn be expected to affect functional properties such as molar absorbance and brightness.<sup>[20]</sup> Analysis of B-factors suggest that regions containing key functional residues such as S205 and E222 become less dynamic on dimerization (Figure S5, Supporting Information). E222 also adopts a different side-chain configuration in the dimers compared sfGFP<sup>WT</sup>. H148, another residue critical to function, is known to be dynamic, existing in “open” and “closed” conformation.<sup>[17,21]</sup> Only in the closed conformation can H148 help contribute to chromophore deprotonation, which in turn promotes absorbance at ≈490 nm. H148 is also located close to the dimer interface and occupies the closed configuration (Figure S5, Supporting Information). As well as changes to protein dynamics, the burial of waters at the interface is also likely to be important for enhanced function. The roles of the additional waters associated with W1 in terms of their impact on the structure–function relationship is not fully known but it has been postulated that they contribute to charge transfer and modulation of the protonated state of the CRO.<sup>[22]</sup> In solution, it is likely that the additional water molecules associated with W1 are in free exchange with the solvent when sfGFP is monomeric; exchange with bulk solvent is likely to be reduced in the dimeric sfGFP<sup>204x2</sup> so persist in a defined arrangement for longer. Thus, dimerization is likely to lead to a more rigid structure around the chromophore with more persistent bond networks and less conformational flux, with these stabilized interactions potentially contributing to the positive synergistic effect on brightness.
2.3. Heterodimers and Functional Communication by Energy Transfer

The use of different FPs with compatible spectral properties to promote FRET is essential for biomolecular analysis. The sfGFP204SCO variant can be linked to Venus (containing azF) via residue 204 to generate heterodimers. The resulting dimer, termed GFVen204, demonstrated FRET from the sfGFP component to Venus, as would be expected (Figure 4a). There is currently very little known about the relative orientation of FRET-based FP pairs with only one structure available in a biosensor configuration, which is in a single polypeptide format rather than a classical two-protein system. Given the high degree of sequence and structure similarity between sfGFP and Venus, we used the GFP204x2 structure to build models of the GFVen204 dimer so as to calculate more specific values based on the relative orientations of the two chromophores (Figure S6, Supporting Information). Using our model of GFVen204x2, together with the dipole arrangements of known transitions for both GFP and Venus[24,25] (Figure 4b), was calculated in the model to be 3.59. Using the QD and J(λ) values (Table S3, Supporting Information) together with a refractive index of 1.4 to account for a combined protein–water environment (Hellenkamp et al.[9] and Dr. Tim Craggs personal communication) we calculated R0 with the different values (Table S3, Supporting Information). The calculated R0 was ≈76 Å, which is up to 19 Å longer compared to when the arbitrary 0.667 value is used. Our calculated R0 values are consistent with those calculated using J(λ) and donor QY values available through FPbase (https://www.fpbase.org)[26] when adjusted for .

The question arises is how does our calculated R0 relate through to observed FRET efficiency. Based on the use of Equation (2) and the measured inter-chromophore distance of 29–30 Å (Figure 4b), the estimated FRET efficiency for our GFVen204 construct should be close to 100% (99.6–99.7%). However, de-convolution of the GFVen204 emission on excitation at 450 nm...
Figure 4. Heterodimer fluorescence characteristics. a) Fluorescence emission spectra of GFVen\textsuperscript{204} (black), Venus\textsuperscript{204azF} (gold), and the residual emission profile highlighting sfGFP contribution to GFVen\textsuperscript{204} spectrum (grey dashed). Excitation was at 450 nm. b) Relative positioning of the Venus (yellow) and sfGFP (green) chromophores (CROs) in the model GFVen\textsuperscript{204} structure. The dashed lines highlighted in orange represent the transition dipole moment (TDM). c) Relative distances between Venus (yellow) and sfGFP (green) based on the model of SPAAC linked dimer (top) and implied spacing calculated from observed FRET efficiency (bottom). Structures are to scale. d) FRET between sfGFP-Bcl3 and P50-Venus on excitation at 420 nm (black line), fluorescence due to P50-Venus alone (gold line), and residual sfGFP-Bcl3 fluorescence in the complex (grey dashed line). e) Modeled docking of mCherry (PDB:2H5Q) and sfGFP (PDB:2B3P) using ClusPro.\textsuperscript{[23]} Residues Q204 (sfGFP) and K198 (mCherry) that are replaced by SCO-K and azF, respectively, are highlighted. f) Fluorescence emission (on excitation at 485 nm) of sfGFP\textsuperscript{204SCO} (green solid line) and GFCh\textsuperscript{x2} dimer (dashed black line). Emission intensities are normalized to sfGFP\textsuperscript{204SCO}. Inset is the zoomed in region of the emission spectrum centered around 610 nm, the emission maximum of mCherry. The red corresponds to the subtraction of GFCh\textsuperscript{x2} from sfGFP\textsuperscript{204SCO}. g) Molar absorbance of sfGFP\textsuperscript{204SCO} (green; \(\lambda_{\text{max}}\) 488 nm), mCherry\textsuperscript{198azF} (purple; \(\lambda_{\text{max}}\) 587 nm), and GFCh\textsuperscript{x2} (black line; \(\lambda_{\text{max}}\) 488 nm and \(\lambda_{\text{max}}\) 587 nm). h) Emission profile of GFCh\textsuperscript{x2} on excitation at 485 nm (green) and 585 nm (purple). Spectral properties are detailed in Table S1, Supporting Information.

(a wavelength that will excite predominately sfGFP) reveals a significant sfGFP component (Figure 4a). Using a commonly used approach based on Equation (3), apparent relative FRET efficiency can be calculated.

\[
E_{\text{rel}} = \frac{I_A}{I_D + I_A} \quad (3)
\]

where \(E_{\text{rel}}\) is relative FRET efficiency, \(I_A\) is integrated fluorescence of the acceptor, and \(I_D\) is the integrated fluorescence of the donor. The relative FRET efficiency was determined to be 87%. While donor (sfGFP\textsuperscript{204SCO}) and acceptor (Venus\textsuperscript{204azF}) quantum yield are likely to vary between the monomers and dimer (see Table S1, Supporting Information), even taking into consideration such variation in quantum yield, FRET efficiency is the range of 85–91%. Thus, there is a clear discrepancy between the observed and theoretical FRET efficiency, as has been observed before form structure-based analysis where inter-FP interactions were observed.\textsuperscript{[12]}

The exact reason for the lower than expected FRET efficiency is not entirely clear. A simple and obvious explanation is that some free monomeric sfGFP\textsuperscript{204SCO} is present. Analysis of polyacrylamide gels and mass spectrum suggests little or no monomeric protein is present (see Worthy et al.\textsuperscript{[11]} and Figure S7, Supporting Information, for details). Are the considerable number of water molecules present at the domain interface observed for sfGFP\textsuperscript{204azF} (Figure 3f,g) playing a role in quenching? Water can quench fluorescence,\textsuperscript{[4,27]} especially if collisional events are promoted through free dynamic exchange. However, the crystal structure suggests local water molecules are likely to be less dynamic in the dimer compared to monomeric forms. Furthermore, if water quenching was significant then the residual donor sfGFP signal would not be observed. Is the arrangement of the monomers in GFVen\textsuperscript{204azF} similar to the assumed sfGFP\textsuperscript{204azF}? One possibility is that there is a mixed population of dimers; a major population in which FRET is highly efficient and a second minor conformation in which energy transfer from sfGFP to Venus is negligible. While modeling predicts GFVen\textsuperscript{204} is likely to occupy...
a conformation similar to sfGFP204 SCO, we cannot rule out a minor state in which energy transfer is compromised. For example, a major shift in transition dipole arrangement could result in \( k^2 \) being close to 0 or the chromophore conformation being altered (e.g., cis-trans isomerization). While we cannot rule out some rotation of one FP with respect to another, the triazole link should restrict such rotation and the CROs should retain a similar vector configuration in terms of the transition dipole moments. With an \( R_0 \) of 76.6 Å, the two CROs will need to be at least 50 Å apart (shown schematically in Figure 4c) for FRET efficiency to be close to that observed. Even using the arbitrary \( k^2 \) value of 0.667 produces an \( R_0 \) of 56 Å, which will require the CROs to be \( \approx 40 \) Å apart to generate the observed FRET efficiency. Given the relationship of residue 204 to the CRO (Figure 1), neither distances are feasible in a covalently linked dimer. What is clear is that bringing two different FPs in close proximity so promoting inter-FP interactions does influence apparent FRET, which results in an observed FRET efficiency that generates an overestimation of the distance between the pair. Thus, while FRET is routinely assessed using differences in donor and acceptor fluorescence intensities based on the general idea presented in Equation (3), it does bring into question whether such analysis is appropriate in all cases.

To understand the impact of the observed effects for typical in vitro and in vivo investigations into protein interactions using FRET capabe FPs fused to proteins of interest, we then assessed FRET between sfGFP and Venus in a non-covalently linked complex. The FPs were fused to two known interacting proteins: Bcl3 and P50.[28] The first construct comprised Bcl3 fused at the N-terminal to sfGFP with no artificial linker (sGFP-Bcl3; Figure S8a, Supporting Information). The second construct comprised Venus fused to the C-terminal of P50 via a largely unstructured 10 amino acid sequence (including a GSS artificial linker sequence) to allow for flexibility (Figure S8a, Supporting Information). The proteins were produced and purified as the complex, as outlined in the Supporting Methods and Figure S8b, Supporting Information. In this configuration the FPs will be positioned at the same end in the complex potentially allowing the FPs to interact in manner similar to that for GFVen204 (Figure S8c, Supporting Information). FRET was observed on excitation at 420 nm, with the major emission peak equivalent to Venus (\( \lambda_{EM} \), 529 nm; Figure 4d). The equivalent emission for P50-Venus alone was much lower. The overall FRET efficiency was 85%, similar to that observed for the covalent GFVen204 dimer. This in turn raises the question concerning which system (closely associated or free rotating) is being sampled given the similarity of FRET efficiency with the covalently linked dimer. However, even using standard \( R_0 \) of 55 Å, the two FP chromophores will be \( \approx 41 \) Å apart and thus relatively close to each other in space so increasing the likelihood of association (and the accompanying effects on fluorescence as noted above), despite the relatively long flexible sequence that comprises the P50 construct. Alternatively, if a preferential TDM alignment between the CROs occurs, then \( R_0 \) would be closer to 75 Å resulting in an interchromophore distance of \( \approx 57 \) Å, which rudimentary structural modeling suggests is a more likely distance given that the Venus element is likely to sample greater dynamical freedom (Figure S8c, Supporting Information). Thus, using fluorescence intensity changes to measure FRET with no understanding of positional relationship of FP pairs may not fully reflect the molecular events in protein complexes.

2.4. Association of Green and Red Fluorescent Proteins

We next linked sfGFP with a DsRed-derived monomeric protein, mCherry.[29] Green fluorescent proteins can be used as an FRET partner with mCherry\(^{[14,30]}\) with an estimated \( J \) coupling of \( 1.8 \times 10^{-25} \) M\(^{-1}\) cm\(^{-1}\) nm\(^{-4}\) (FPbase FRET tool; www.fpbase.org/fret/).\(^{[26]}\) The sfGFP204 SCO variant was reacted with mCherry containing azF at the structurally equivalent position, residue 198 (Figure 4e). Molecular docking suggested the two proteins can associate at the interface between residues 204 sfGFP and 198 mCherry (Figure 4e), with covalent coupling via SPAAC subsequently proved by SDS PAGE (Figure S9, Supporting Information). Incorporation of azF at residue 198 in mCherry had little effect on the spectral properties of the monomer with a similar molar absorbance and brightness to the wt mCherry (69 000 M\(^{-1}\) cm\(^{-1}\) nm\(^{-1}\) with a quantum yield of 24% compared to 72 000 M\(^{-1}\) cm\(^{-1}\) for wt mCherry with quantum yield of 22% at 587 nm; Figure 4g; Figure S9c, Table S1, Supporting Information).

The purified dimer, termed GFChx2 did not appear to display any significant FRET on excitation at 490 nm (Figure 4f). Indeed, very little observable fluorescence can be attributed to mCherry in the dimer even on excitation at 585 nm (Figure 4h), which is confirmed visually through general UV excitation (Figure S9b, Supporting Information). The mCherry associated peak at \( \approx 585 \) nm is reduced in terms of molar absorbance compared to the mCherry198 azF monomer. As with other dimeric forms (vide supra and ref.\(^{[11]}\)), the sfGFP molar absorbance increased above the simple addition of the two monomeric forms (\( \approx 16 \) 000 M\(^{-1}\) cm\(^{-1}\) taking into account the contribution from the mCherry chromophore) confirming the role of dimerization via residue 204 in enhancing sfGFP function.

Given that donor fluorescence from the sfGFP component is prevalent in GFChx2 (Figure 4f), it is unlikely that non-radiative energy release from mCherry is the main cause of fluorescence loss as quenching of sfGFP fluorescent signal would also be observed. Data suggests that it is the interaction and subsequent conformation changes on interfacing with sfGFP that is responsible for loss of mCherry associated fluorescence, through altering the intrinsic properties of the chromophore. The reduction in the molar of absorbance of the mCherry component provide some credence to this idea given that molar absorbance is generally enhanced in dimers comprised solely of A. victoria derived FPs.\(^{[11]}\) One possibility is that dimerization is shifting the mCherry chromophore to a chemically modified, non-functional conformation, akin to that observed for reduced versions of the chromophore\(^{[13]}\) or for dark state versions of photosensitive versions of the protein (e.g., PAmCherry).\(^{[32]}\) However, the presence of a clear, albeit less intense mCherry absorbance peak in GFChx2 suggest that a significant proportion of the chromophore has not been chemically altered. Thus, it is more likely that chromophore is switched and then trapped in an alternative non-fluorescent conformation (e.g., cis to trans isomerization),\(^{[33]}\) as occurs in the mCherry photoswitching variant nsCherry.\(^{[34]}\) Given the proximity of the mCherry chromophore to the dimer interface,
dimerization may alter the conformation surrounding residues so promoting the 
\textit{trans} over the \textit{cis}.\cite{11} Such phenomena would open up the opportunity for engineering FP probes capable of photoswitching upon the interaction of their fusion protein partners. Reaction of mCherry\textsubscript{198az} with the SCO-K ncAA alone does not appear to affect fluorescence suggesting the covalent linkage per se is not responsible (Figure S9c, Supporting Information). Attachment with the bulker azide containing Cy3 dye also does not result in loss of fluorescence, with FRET observed as expected (Figure S9f, Supporting Information). Thus, placing the two FPs in close proximity to promote their interaction is the likely cause of the loss in fluorescence. As donor fluorescence is still observed, in a classical ensemble FRET experiment this could be interpreted as the two target proteins not interacting when the opposite may in fact be the case. As pointed out earlier, basing FRET efficiency solely on intensity changes may thus not be appropriate in all cases.

\subsection*{2.5. Comparison with Alternative sfGFP Dimer sfGFP\textsubscript{148az2}}

The structure of another click-linked dimer joined via residue 148 (termed sfGFP\textsubscript{148az2}) has recently been reported.\cite{11} Dimerization effectively switched sfGFP\textsubscript{148az2} on, with the dimer displaying improved function compared to both monomers and the original wild type sfGFP (sfGFP\textsubscript{WT}). We used the structure of the sfGFP\textsubscript{148az2} dimer to calculate $\chi^2$ as a representative alternative CRO arrangement. This will in turn allow us to investigate how different configurations of one monomer to the other affect dipole alignments and hence FRET. Residue 204 lies close to 148 on the adjacent $\beta$-strand (Figure 5a) but they adopt very different sidechain and thus monomer arrangements in their crystal structures (Figure 5b,c). In contrast to sfGFP\textsubscript{204az2}, the triazole link in sfGFP\textsubscript{148az2} forms the extended anti form that is re-enforced with both polar and hydrophobic interactions between the monomers generating a quasi-symmetrical “head-to-tail” arrangement of the monomers. The result of such a configurational change between the two monomers units results in the relative positioning of the CROs being very different (compare Figures 5b and 5c).

Using the same approach as for GFVen\textsubscript{204}, we calculated $\chi^2$ values for a model of GFVen\textsubscript{148} based on the sfGFP\textsubscript{148az2} structure. It should be noted that unlike linkage through residues 204 (or 198 in mCherry), covalent coupling via residue 148 was designed to instigate a functional change through synergistic conformation events.\cite{11} However, it does allow us to assess how changing the orientation and inter-unit interactions of one monomer to another along a quasi-similar interface region alters dipole arrangements. The calculated $\chi^2$ was 3.79, even closer to the maximal value of 4 than sfGFP\textsubscript{204az2}. While this would suggest a longer $R_o$ distance than GFVen\textsubscript{204}, the inherent function of the GFVen\textsubscript{148} dimer system makes calculating $R_o$ problematic; the donor, sfGFP\textsubscript{SCO148}, is essentially switched off in monomeric state and only becomes activated on dimerization. However, the main effect that will influence any FRET analysis is the shift in $\lambda_{\text{EM}}$, which is blue shifted by 10 nm in the GFVen\textsubscript{148} dimer compared to Venus\textsubscript{WT} (Figure 5d) when excited at a wavelength corresponding to sfGFP. If single wavelength readings are taken with 530 nm assumed to be the Venus emission maximum, fluorescence emission would be underestimated by up to 35% so impacting on perceived FRET efficiency. While residue 204 is more applicable in terms of understanding association and FRET due to the non-perturbative nature of the initial mutations, sfGFP\textsubscript{148az2} and GFVen\textsubscript{148} still act as good examples of how association is once again having a significant effect on the spectral properties. It also demonstrates that FP dimerization are not restricted to a defined interaction configuration but that different inter-FP orientations are available.

\section*{3. Conclusion}

Our ability to construct dimers of FPs coupled with structural analysis has allowed us to look at how association can influence two of their key functions: inherent electronic excitation/light emission and communication through energy transfer. With regards to the latter, we can use structural information to predict dipole alignments of two CROs, which is critical to FRET through defining $\chi^2$. In our case, the arbitrary 0.6667 for the $\chi^2$ value provides a significant underestimate of the predicted values that impacts on $R_o$. There have been several studies to date that measure FRET in constructs whereby FPs are coupled via linker sequences or whole protein domains. However, by linking two FPs together they can no longer freely interact with each other due to, for example, steric hindrance (e.g., when using linker sequences)\cite{13} or spatially forced apart (e.g., when linked to via a single whole protein),\cite{12} schematically outlined in Figure 6. Our use of bio-orthogonal chemistry allows broader sampling and stabilization

Figure 5. Comparison of sfGFP\textsubscript{204az2} structure with sfGFP\textsubscript{148az2}. a) Location of the azF moiety at residue 148 (grey) or 204 (green). Structure of b) sfGFP\textsubscript{204az2} and c) sfGFP\textsubscript{148az2}. Each structure is orientated identically with respect to the azF containing ncAA (colored green) to highlight the relative differences in monomer arrangements. d) Normalized emission on excitation at 490 nm for GFVen\textsubscript{148} (black) and wt Venus (gold).
of mutually compatible FP interfaces (Figure 6c); non-compatible FP surfaces do not form covalent bonds so the interface will not persist.\cite{11} The most powerful use of FRET is monitoring protein–protein interactions whereby the FPs are fused to separate proteins. It can be argued that most FP fusions will not associate in most FRET experiments. However, as FPs will be attached to partner protein that normally associate, and if the two FPs are in close proximity they may well align or even interact in preferential arrangements (Figure 6d), which can define dipole alignment and influence inherent FP function. Naïve docking of FPs along with empirical evidence highlights FPs tendency to oligomerize, which will be enhanced by local high concentrations. Promotion of FP interactions can then potentially impact on calculated FRET efficiencies, critical for reporting of protein–protein interactions, using commonly used fluorescence intensities approaches related to Equation (3). Closer attention to the initial construction of FP fusion protein to prevent such interactions should help address this issue (Figure 6d), which in turn will allow standard approaches to FRET analysis to be applied in a more meaningful manner. It is thus clear from our work that by placing FPs in close proximity so promoting physical interaction can result in changes in the expected fluorescence behavior.

4. Experimental Section

Protein Production: The monomeric sfGFP\textsuperscript{204x2F} and sfGFP\textsuperscript{204SCO} proteins and the sfGFP\textsuperscript{204x2} dimer were produced as described previously.\cite{11} The WT mCherry, mCherry\textsuperscript{198azF}, sfGFP-Bcl3, and P50-Venus proteins were produced as outlined in the section Methods, Supporting Information.

Protein Dimerization and Conjugation: The procedures for generating sfGFP homodimers and sfGFP-Venus heterodimers have been reported previously.\cite{11} Generation of the sfGFP-mCherry dimers was performed as follows. The sfGFP\textsuperscript{204SCO} and mCherry\textsuperscript{198azF} were mixed at an equimolar concentration (50 µm, 50 mM Tris-HCl pH 8.0) and left at room temperature for \approx 16 h. Dimers were purified by size exclusion chromatography (Superdex 200, 16/600). Protein dimerization and separation was also monitored by SDS PAGE gel. Conjugation with non-proteinaceous molecules is described in the section Methods, Supporting Information.

Steady State Absorbance and Fluorescence Analysis: Spectrophotometry and fluorescence were performed essentially as described previously for sfGFP monomers, dimers and Bcl3 fusion, Venus monomers, P50 fusions,
molecules and sfGFP–Venus hybrid dimers. Analysis of variants involving mCherry followed a similar analysis procedure, using proteins concentrations of 5 μM in 50 mM Tris-HCl pH 8.0. Absorbance spectra were recorded on a Cary Win UV, using a 300 nm min⁻¹ scan rate at 1 nm intervals. Absorbance at λmax for each variant, was used to determine the molar extinction coefficients (ε) for each variant, using the Beer–Lambert equation and measured protein concentrations. Emission spectra were collected on a Cary Varian fluorimeter at a scan rate of 60 nm min⁻¹ and 1 nm intervals. Emission and excitation slit widths were set to 10 nm and a detector voltage of low. Samples were excited at 5 nm from 460 nm to 590 nm as stated in the main text and emission was scanned from the excitation wavelength to 800 nm. J coupling constants (J[A]) were calculated using either available parameters on FPFbase,’26 via the FRET tool or calculated from experimental data using a|e software (http://www.fluortools.com/software/ae). FRET efficiency was calculated using Equation (3). To determine the contribution of donor fluorescence in dimer constructs, dimer emission spectra were deconvoluted by subtracting the monomeric acceptor emission spectra so as to calculate residual emission of the donor.

**Single Molecule Fluorescence:** Measurement and analysis of single molecule sfGFP²042 fluorescence by total internal resonance fluorescence microscopy was performed as described previously.[¹¹]

**Structure Determination of sfGFP²042:** The sfGFP²042 dimer variant was concentrated in 50 mM Tris pH 8.0 to a final concentration of 10 mg mL⁻¹, and used to set up vapor diffusion crystal trays. A JBScreen membrane (Jena Bioscience, Germany) was used initially to facilitate crystal growth, where large green crystals grew in a multitude of buffer conditions. Large green crystals grew in 20% polyethylene glycol w/v, 100 mM HEPES, which were harvested and transferred to mother liquor supplemented with 13% w/v PEG 200 as a cryo-protectant, and vitrified in liquid nitrogen. X-ray diffraction data was collected at the Diamond light source, Harwell, UK (beamline I02). Structure refinement was performed using the CCP4 program suite.[³] The structure was solved initially using the molecular replacement program Phaser[¹⁶] with wt sfGFP (PDB accession 2B3P) used as a model. Structures were manually adjusted using COOT,[¹⁷] and refined with TLS restrained refinement using REFMAC.[³] The final coordinates were deposited in the Protein Data Bank (PDB) under accession code 5N13.

**Kappa² Calculation:** The dipole orientation factor, κ², was calculated using an approach as described previously.[³⁹] The model structure of GFVen²⁰⁴ dimer was built by overlapping the WT structure of Venus (1MYW) onto the sfGFP²⁰⁴² component of sfGFP²⁰⁴² structure. While Venus and sfGFP used here had 15 amino acid differences in the core β-barrel structure only Ala206 in sfGFPSCO204 and Val206 in Venus contributed to the domain interface. Residue 204 was then replaced with azF and linked to SCO using the PyMOL bond building tool. The GFVen²⁰⁴ model overlaid with sfGFP²⁰⁴² is shown in Figure S4, Supporting Information. Using previously established approaches,[²⁴,³⁹] the model structure κ² was calculated as outlined in Equation (4) using the distance between the centers of the donor and acceptor dyes (rda), and the orientations of the transition dipole moments of the donor (d̅) and the acceptor (a̅).[¹⁹]

\[
kappa² = \left[ \begin{array}{cc} d & -3 \\ -rda & d \end{array} \right] \left( \begin{array}{c} a \\ -rda \end{array} \right) = (\cos θ₁ - 3 \cos θ₁ \cos θ₂)^²
\]

(4)

The atomic positions of CG2 and C2 of the chromophore, as shown in Figure S6b, Supporting Information, were used to define the vector for the TDM for both Venus and sfGFP.²⁴ The κ² was then used in Equation (1) together with available experimental to calculate R0 for FRET pairs.

**Supporting Information**

Supporting Information is available from the Wiley Online Library or from the author.

---

**Acknowledgements**

The authors would like to thank the staff at the Diamond Light Source (Harwell, UK) for the supply of facilities and beam time, especially Beamline I02 staff. The authors would also like to thank Edward Lemke and his group at EMBL Heidelberg for donating the pEVOL-SCO plasmid. The authors thank BBSRC (BB/H003746/1 and BB/M000249/1), EPSRC (EP/J015318/1), and Cardiff SynBio Initiative/SynBioCite for supporting this work. R.L.J. and R.A. was supported by Knowledge Economy Skills Scholarship (KESS2) Ph.D. studentships, W.D.J. by Wellcome ISSF, H.L.W. by a BBSRC-facing Cardiff University Ph.D. studentship, H.S.A. by the Higher Committee for Education Development in Iraq. The authors would like to thank the Protein Technology Hub, School of Biosciences, Cardiff University for use of facilities. They also thank Dr. Joachim Goedhart from the University of Amsterdam for his helpful and insightful comments on the manuscript. [Correction added after publication on January 06, 2020: Figure 2 was corrected because some elements were not visible in the previous version]

**Conflict of Interest**

The authors declare no conflict of interest.

**Keywords**

fluorescence, fluorescent proteins, Förster resonance energy transfer (FRET), oligomerization, protein design

Received: August 18, 2020
Published online: November 23, 2020