Structure-guided rational design of red fluorescent proteins: towards designer genetically-encoded fluorophores
Matthew G Eason, Adam M Damry and Roberto A Chica

Red fluorescent proteins (RFPs) have become an integral part of modern biological research due to their longer excitation and emission wavelengths. Protein engineering efforts have improved many key properties of RFPs for their practical use in imaging. Even so, continued engineering is required to overcome the shortcomings of the red chromophore and create RFPs with photophysical properties rivalling those of their optimized green and yellow counterparts. Here, we highlight recent examples of structure-guided rational design of RFPs to improve brightness, monomerization, maturation, and photostability, and discuss possible pathways for the future engineering of designer RFPs tailored to specific applications.

Address
Department of Chemistry and Biomolecular Sciences, University of Ottawa, 10 Marie-Curie, Ottawa, Ontario, K1N 6N5, Canada

Corresponding author: Chica, Roberto A (rchica@uottawa.ca)

Introduction
Red fluorescent proteins (RFPs) derived from corals and sea anemones are genetically encoded fluorophores that have become an integral part of modern biological research. These proteins are extensively used to track proteins within cells [1,2], detect gene expression in vivo [3], act as acceptor chromophores in Förster resonance energy transfer [4], and transduce signal in biosensors [5,6]. Unlike other fluorescent proteins (FPs) from the GFP superfamily, RFPs contain a p-hydroxybenzylideneimidazolinone chromophore that is conjugated to an acylimine group produced by oxidation of the backbone Cα=N bond in the N-terminal chromophore-forming residue [7,8]. The presence of this group extends π-conjugation in the chromophore, enabling it to absorb and emit light at longer wavelengths. This property of RFPs makes them particularly useful for whole-animal imaging, as longer-wavelength light is less scattered by tissue and causes lower phototoxicity, enabling deeper imaging and longer acquisition times. However, RFPs have not yet reached the same level of near-optimal imaging performance as green and yellowFPs [9], leading us to focus on their improvement via structure-guided rational design in this review.

Over the years, RFPs have been extensively engineered for desired spectral properties using a combined approach of rational design followed by multiple rounds of directed evolution. To date, over 100 RFPs have been developed [10] from natural precursors such as DsRed [11], eqFP611 [12], and eqFP578 [13], including far-red FPs displaying excitation and emission wavelengths of up to 611 and 675 nm, respectively [14,15]. Unfortunately, RFP brightness tends to decrease as emission wavelength increases [16], suggesting a physical limit in the continued development of far-red FPs unless a key innovation is made to overcome this limitation. In addition, all known RFPs mature more slowly than their optimized green and yellow FP counterparts [17**,18,19], and no single RFP possesses photophysical properties that are ideally suited for all potential imaging applications [20]. Thus, continued engineering is essential to overcome the inherent limitations in brightness and maturation caused by the structural features of their acylimine-conjugated chromophore and to tailor RFPs for specific imaging applications.

In the past few years, the wealth of available high-resolution crystal structures of RFPs coupled to an ever-increasing understanding of the structural determinants of their function have opened the door to the rational design of improved RFPs without the need to rely on random mutagenesis in a directed evolution approach. Rational design can help improve desired RFP properties that are difficult to screen for in a high-throughput fashion (e.g., maturation rate) while also reducing the time and effort required to identify improved variants, as smaller numbers of mutants are typically evaluated compared to directed evolution approaches. Of all the properties of RFPs, excitation and emission wavelengths have arguably been those that have been most frequently engineered by both random [21,22] and rational [23] approaches. Because of the extensive effort dedicated to the development of far-red FPs, the structural determinants of
fluorescence red-shifts in RFPs are fairly well understood, and colour-tuning strategies have been recently reviewed [24,25]. Here, we instead review recent examples of the successful use of structure-guided rational design to engineer the brightness, oligomerization, maturation, and photostability of various RFPs, omitting studies where beneficial mutations found by directed evolution that improve these properties in other FPs were simply transplanted onto a desired scaffold. The selected examples demonstrate how the abundance of structural and functional knowledge acquired through the years can be exploited to tailor desired characteristics of RFPs in our quest to create ‘designer’ fluorophores that are optimal for the needs of a particular application.

**Improving RFP properties by rational design**

**Brightness**

Brightness is the product of the extinction coefficient and quantum yield, which are the efficiency of the chromophore at absorbing or emitting light, respectively. Although extinction coefficient values for most monomeric RFPs are in the same range as those of other members of the GFP superfamily (40,000–100,000 M$^{-1}$ cm$^{-1}$), quantum yield values are significantly lower (e.g., <0.4 for most monomeric RFPs versus >0.6 for green and yellow FPs) [16,26]. A notable exception is the mScarlet monomeric RFP (quantum yield = 0.7) that was recently engineered by several rounds of targeted and random mutagenesis on a synthetic gene template containing many mutations previously shown to increase brightness and facilitate monomerization in other FPs [27]. While it is difficult to improve the extinction coefficient of RFPs without modifying the chemical structure of the chromophore, improvements to quantum yield can theoretically be achieved through chromophore rigidification via tighter packing interactions, which can reduce non-radiative decay by decreasing the probability of excited-state vibrational relaxation.

This idea was exploited by Chu et al. in their engineering of the far-red FP mNeptune2.5 [17]. Using a semirational approach consisting of several rounds of combinatorial saturation mutagenesis of core residues surrounding the chromophore, they identified six mutations that increased quantum yield of precursor protein mNeptune by 0.05 (22%) while also resulting in an 8-nm emission wavelength hypsochromic shift (Table 1). In a follow-up study, Chu and coworkers [28] used a similar semirational approach to increase the quantum yield of mRuby2. Over six rounds of mutagenesis, they obtained mRuby3, a variant containing 21 mutations that collectively increase quantum yield by 0.07 (18%). Similar to mNeptune2.5, mRuby3 displays an 8-nm emission wavelength hypsochromic shift relative to its parent RFP (Table 1). While quantum yield was successfully increased in both of these RFPs, the presence of tighter packing interactions causing chromophore rigidification was not validated. Thus, it remains uncertain whether the increased quantum yields of mNeptune2.5 and mRuby3 are caused by chromophore rigidification or by the increased energy difference between ground and excited singlet states, which would decrease the probability of internal conversion and therefore non-radiative decay.

Recently, a strictly rational approach was developed by Pandelieva et al. [29] to increase the quantum yield of mRojoA, a dim and red-shifted mutant of the monomeric RFP mCherry containing a π-stacked Tyr residue directly beneath the chromophore [23]. In this study, an aromatic amino acid was introduced at position 63 above the mRojoA chromophore in order to sandwich it in a triple-decker motif of aromatic rings (Figure 1a), an interaction that was hypothesized to rigidify the chromophore. To
Figure 1

Structure-guided rational design of RFPs. In all cases, the chromophore and important residues are shown as sticks. (a) Crystal structures of mRojoA (PDB ID: 3NEZ [23]) and mRojo-VYG (PDB ID: 5H89 [29*]) are shown in blue. The alpha carbon of Gly143 in mRojo-VYG is shown as a sphere. Space-filling model in grey shows tighter packing interactions around the mRojo-VYG chromophore caused by the P63Y substitution. The T16V, W143G, and L163V mutations of mRojo-VYG also contribute to increased quantum yield. (b) The crystal structure of DaRed (PDB ID: 1GGX [2]) and a computer-generated model of mLib77 (in green) show mutated interface positions. Additional subunits of DaRed are shown in grey. (c) Crystal structures of mPlum (PDB ID: 2QLG [56]), mPlum-E16P (PDB ID: 4H3L [46*]), mPlumAYC (PDB ID: 4H3N [46*]), and a computer-generated model of mPlumAYC-E16P are shown in yellow. The hydrogen bond between the Glu16 side chain and the chromophore is shown as a dashed line. (d) The crystal structure of mCherry (PDB ID: 2H5Q [67]) and a computer-generated model of Kriek are shown in red. The β-strand 7 and 10 interface is highlighted with an arrow. All computational models were prepared using the PHOENIX computational protein design software, as described in [23].

accommodate the aromatic substitutions, a small combinatorial library was designed where the bulky Trp143 and Leu163 residues surrounding position 63 were mutated to small amino acids. The best mutant identified (Table 1, mRojo-VYG) displayed a quantum yield increase of 0.03 (150%), and was isolated following the screening of 48 mutants, a library size that is several orders of magnitude smaller than those previously used to achieve comparable quantum yield increases in other RFPs [30,31]. As was the case for mNeptune2.5 and mRuby3, the observed quantum yield increase in mRojo-VYG was accompanied by a large hypochromatic shift in emission wavelength (Table 1). When the Tyr63 residue of mRojo-VYG was reverted to the native Pro, the quantum yield decreased by 0.03 without affecting the emission wavelength, suggesting that it is possible to improve quantum yield while maintaining the red-shifted emission wavelength. To confirm that the triple-decker motif of aromatic rings did improve quantum yield through increasing chromophore rigidity, crystal structures of
bright and dim variants were solved. It was found that the average chromophore B-factors of dim variants were significantly higher than the average B-factors of the rest of the protein, unlike in the brighter mRojo-VYGV variant, providing evidence that the chromophore of the brighter variant was indeed more rigid than that of the dim variants.

In the future, the development of brighter RFPs by rational design should focus on mutation of residues that are not directly interacting with the delocalized electron cloud of the chromophore so as to avoid perturbing energy levels of electronic states. One such approach is the mutation of first-shell residues that are orthogonal to the chromophore π system. The utility of this approach was demonstrated by Goedhart et al. [32] who increased the quantum yield of the cyan FP mTurquoise by 0.09 (11%) without shifting its emission wavelength. This result was achieved via the introduction of a single mutation that improved packing by creating several additional van der Waals interactions with the chromophore. An alternate approach is the introduction of distal mutations that would decrease dynamics of the chromophore pocket without altering its electrostatic environment. Alford et al. [33] showed that binding of another subunit to form an RFP heterodimer leads to a 0.05 (185%) quantum yield increase in a dim monomeric RFP without shifting its emission wavelength. The authors proposed that the quantum yield increase upon heterodimer formation results from a change to the chromophore pocket that decreases its conformational flexibility and stabilizes it in the brighter coplanar conformation. We postulate that a similar effect could be achieved in a monomer by stabilizing the entire RFP β-barrel via distal mutations.

**Oligomerization**

The effective use of RFPs as fusion tags to study in vivo processes requires that the FP itself does not influence the localization or interactions of the protein of interest. Because all natural RFPs are oligomeric, an initial step in RFP engineering is monomerization to avoid undesired aggregation of tagged proteins. Campbell et al. [34] put forth an algorithm for the rational design of monomeric RFPs whereby subunit interfaces are rationally disrupted through the introduction of repulsive charges, in the form of Arg and Lys mutations at buried interface positions. However, RFP monomerization tends to abolish fluorescence, which must be restored by additional engineering to stabilize the RFP core [13,34].

Recently, Wannier et al. [35] developed a computational approach to predict mutations enabling RFP monomerization. To avoid the usual loss of fluorescence that occurs due to monomerization, they first stabilized DsRed by introducing the core mutations found in its monomeric variant mCherry [30], then used computational protein design (CPD) [36–38] to predict combinations of surface mutations that would be likely to disrupt the oligomeric interfaces without destabilizing the RFP β-barrel. A total of 17 buried interface positions were designed (Figure 1b) by allowing 13 amino acid types at each position. Screening of a library comprising the top 96 variants, each containing 13–16 surface mutations, revealed that 97% of these were fluorescent and monomeric, as evidenced by homo-FRET experiments. A representative mutant, mLib77, contains 14 surface mutations and displays a brightness similar to that of mCherry (Table 1).

The procedure of core stabilization followed by computational monomerization developed by Wannier et al. results in a high rate of functional monomers, even though these variants contain a large number of mutations. This approach could also be used to solubilize aggregation-prone RFPs [39], and as was done by Close et al. who disrupted inter-subunit contacts observed in the crystal lattice to solubilize an aggregation-prone variant of GFP [40]. However, the approach developed by Wannier et al. requires that core-stabilizing mutations be previously identified. In the absence of this knowledge, structure-guided rational design to increase quantum yield via core packing as described in the previous section can be used, as was done by Zhang et al. who introduced two known brightness-enhancing mutations near the chromophore of the green-to-red photoactivatable FP mEos2 to recover the brightness lost following monomerization [41]. Alternatively, CPD could be used to simultaneously predict mutations to disrupt the interfaces and to stabilize the core, since this computational procedure is designed to identify core-packing and stabilizing mutations [42,43].

**Maturation**

In order to become fluorescent, a folded FP undergoes a maturation process [44] beginning with the cyclization of the chromogenic tripeptide (Xaa-Tyr-Gly), followed by oxidation into a colourless intermediate. This intermediate is a branch point, as it can either be dehydrated into a green chromophore species, or dehydroxylated, followed by a second oxidation reaction to form a blue chromophore species. In both cases, this step is irreversible, however the blue species can then be further dehydrated to form the final red chromophore. Thus, the red and green chromophores are end-points that cannot interconvert [44], with some RFPs exhibiting inefficient maturation that results in mixed populations of molecules containing either green or red chromophores [22,45]. The dehydration step to form the red chromophore is the rate-limiting step of the maturation process [44]. Currently, the fastest-maturing RFPs have maturation half-times above 20 min [17] while maturation half-times of less than 10 min have been reported for green and yellow FPs [18,19]. Thus, there is still a need to develop RFPs that mature faster and more efficiently for in vivo imaging of protein translation in real-time and tracking of proteins with short lifetimes.
In an example of structure-guided rational design of RFP chromophore maturation, Moore et al. [46**] introduced red chromophore formation into mPlumAYC, a yellow-emitting, maturation-deficient mutant of mPlum containing the T195A, I197Y, and A217C mutations. They hypothesized that the Glu16 residue of mPlumAYC inhibits the second oxidation reaction to generate the blue intermediate by forming an H-bond with the acylimine oxygen of the chromophore (Figure 1c) that slows down the trans-cis peptide bond isomerization necessary for oxidation to occur. By replacing Glu16 with a variety of small non-polar amino acids, Moore et al. were able to restore oxidation to the blue species and thereby red chromophore maturation, with optimal results obtained from Pro and Ala substitutions. However, the red/green absorbance ratio of the mPlumAYC-E16P mutant (Table 1) remained low (15%), suggesting that the dehydration step leading to the green chromophore remained significantly faster than the dehydroxylolation and second oxidation steps leading to the blue intermediate. Interestingly, when the E16P mutation was introduced into mPlum, it increased its rate of red maturation approximately two-fold (Table 1) while slowing down green maturation, leading to a four-fold increase in maturation efficiency (red/green absorbance ratios of 370% and 90% for mPlum-E16P and mPlum, respectively). Kinetic experiments demonstrated that maturation efficiency was improved due to increased flux through the blue chromophore intermediate over the green, likely due to acceleration of the second oxidation step.

While the example described above demonstrates that structure-guided rational design can be used to improve RFP maturation efficiency, the maturation half-time for the red chromophore remained slower than for the green chromophore [46**]. Since both green and red chromophores are formed via a similar dehydration step, it is unlikely that acceleration of the dehydration step alone will result in RFPs that mature rapidly and exclusively to the red chromophore. Thus, the design of efficient-maturing and fast-maturing RFPs will require the introduction of mutations that accelerate the dehydroxylation and second oxidation steps more than they accelerate the dehydration step. Because the identities of the amino acids catalyzing these steps are unclear, structure-guided rational design of maturation remains difficult. We expect that in-depth mechanistic investigation of the structural determinants of chromophore maturation coupled to analysis of trapped intermediate structures [47] will be necessary for further designs.

**Photostability**

As FPs undergo multiple cycles of excitation and emission during imaging, photobleaching of the chromophore occurs, where the chromophore is irreversibly converted into a dark state (Figure 2), leading to diminished fluorescence and limiting acquisition times [48]. While the exact mechanism of photobleaching remains uncertain [49], it is known to be an oxygen-dependent process [50,51]. In a recent example of structure-guided rational design of photostability, Dean et al. [52**] hypothesized that they could improve the photostability of mCherry by impeding oxygen access to the chromophore. To identify how oxygen enters the mCherry β-barrel, the authors performed molecular dynamics (MD) simulations that revealed disruptions in its H-bond network between β-strands 7 and 10 (Figure 1d) that create oxygen access routes to the chromophore [53]. They then prepared a combinatorial library that included mutations to positions identified by MD to decrease interstrand dynamics.

---

**Figure 2**

Maturation and photostability are inversely linked. In all cases, the chromophore is shown as sticks. During maturation, an uncyclized chromogenic tripeptide (model prepared from PDB ID: 1QXT [47]) undergoes a series of chemical reactions including two oxidation steps to yield the fluorescent, mature red chromophore (PDB ID: 2WIO [55]). The mature red chromophore can be further oxidized under light irradiation to form a non-fluorescent photobleached chromophore (PDB ID: 2WIS [55]).
(residues 143 and 163) and to positions 16, 66, and 161, which are known to affect maturation and/or brightness [30, 46]. Combinatorial mutagenesis of these sites yielded a library of 144,000 mutants that were screened using a microfluidic assay [54]. This led to the identification of Kriek, a triple mutant (W143I, I161M, and Q163V) displaying four-fold increased photostability (Table 1). MD simulations using a computational model of Kriek demonstrated that the interface between β-strands 7 and 10 had indeed been stabilized in its closed state, which should restrict oxygen access to the chromophore and thus augment photostability.

In the past, several RFPs displaying enhanced photostability have been obtained by directed evolution, either by actively screening for this property [49] or by obtaining improvements fortuitously while screening for another property [28*]. In contrast, the development of Kriek demonstrates that photostability can be enhanced by structure-guided rational design, although it required the high-throughput screening of a large combinatorial mutant library. To achieve similar enhancements to photostability using smaller libraries, a better understanding of the photobleaching mechanism is required. Possible first steps in probing this mechanism would be to solve additional structures of photobleached RFPs [55, 56] and to perform comprehensive saturation mutagenesis [57] to identify the structural determinants that impact photostability, which are still poorly understood.

**Future directions**

The examples described above demonstrate how structure-guided rational design can be used to improve many of the key characteristics of RFPs needed for their practical use in imaging. However, it is unlikely that a single RFP will ever possess optimal properties to satisfy all imaging requirements simultaneously, as many of these properties are inversely linked. For example, quantum yield is associated with chromophore rigidity, yet maturation requires sufficient flexibility to undergo the cyclization and oxidation steps. Likewise, maturation requires oxygen to access the chromophore (Figure 2), while photostability is improved by exclusion of oxygen from the chromophore environment. Thus, researchers will need to continue optimizing RFP characteristics to meet the requirements of their specific application. In this optic, structure-guided rational design can help to tailor RFP properties by focusing on positions where improvements have a high likelihood of occurring (Figure 3). However, it will be important to account for mutational epistasis when performing rational design on new RFPs, as it can limit the effectiveness of simply transplanting single mutations or motifs onto different protein scaffolds [27*, 58]. One way to help circumvent mutational epistasis is to use computationally designed combinatorial libraries, which have been shown to enable the identification of synergistic mutations that would have been difficult to predict *a priori* or to obtain through random mutagenesis [25].

**Figure 3**

Positions mutated in rationally designed RFPs. Sites where mutations have been introduced to improve the quantum yield, monomerization, or photostability of RFPs discussed herein [17*, 29*, 35*, 46*, 52*] are coloured red, green, or blue, respectively. Sites that contribute to optimization of quantum yield and either maturation or photostability are coloured orange or purple, respectively. The RFP structure is based on that of DsRed (PDB ID: 1ZGO [68]).
In the long run, the robust application of structure-guided rational design to improve any combination of RFP properties, such as those discussed here as well as additional ones including pKa, folding kinetics, protein stability, and phototransformation efficiency, will require increased functional and structural information. Functional information can be acquired by systematic high-throughput studies of the local fitness landscape of RFPs [59], which are being facilitated by the development of comprehensive saturation mutagenesis methods [57] and microfluidic assays enabling dynamic single cell analysis [60]. Structures of alternative RFP states that have hitherto been difficult to obtain, such as the photoexcited or oxygen-bound states, will also be necessary. These structures could be obtained by new biophysical techniques such as time-resolved serial femtosecond crystallography [61,62], or calculated using QM/MM techniques [63–65]. Ultimately, structure-guided rational design of improved RFPs will require a holistic view of FP engineering, where all relevant structural states that the RFP can adopt are considered simultaneously for optimization. We expect that the availability of alternative structural states will help the rational design of RFPs displaying any desired combination of properties, taking us closer to an age of designer genetically encoded fluorophores on demand.

Acknowledgements

We gratefully acknowledge support from the Natural Sciences and Engineering Research Council of Canada (NSERC) through Discovery Grant RGPIN-2016-04831 (R.A.C.). M.G.E. is the recipient of an Ontario Graduate Scholarship and A.M.D. is the recipient of an NSERC Postgraduate Scholarship. We dedicate this review to the memory of Prof. Roger Y. Tsien (1952–2016).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- **of outstanding interest

This study details the semi-rational design of mNeptune to create mNeptune2, mNeptune2.5, and mCardinal, which are among the brightest, fastest maturing, and most photostable far-red FPs currently in use.
23. Chica RA, Moore MM, Allen BD, Mayo SL: Generation of longer emission wavelength red fluorescent proteins using...


During this new process, Bindels et al. published their engineering of mScarlet, the first reported monomeric RFP displaying a quantum yield >0.5. In this work, a synthetic RFP gene template containing many mutations previously shown to increase brightness and facilitate mono- 
merization in other RFPs was built. The resulting dim RFP mRed7 (quantum yield < 0.01) was subjected to a combination of targeted and random mutagenesis to yield mScarlet. While mScarlet is the bright- est monomeric RFP developed to date, its emission wavelength (594 nm) is significantly blue-shifted compared to other RFPs and its maturation rate is low. Nonetheless, mScarlet represents an excellent starting point for the engineering of brighter far-red FPs.


In this study, the authors used semi-rational directed evolution to create mRuby3, a variant containing 21 substitutions relative to its parent mRuby2 that were found over 6 rounds of combinatorial saturation mutagenesis. mRuby3 is among the brightest monomeric RFPs available to date.


In this study, the authors make use of rational design to create a triple-decker motif of aromatic rings around the mRhoA chromatophore to rigidify it and increase its quantum yield. A quantum yield improvement of 0.03 was obtained after screening a library of only 48 members, and x-ray crystallography confirmed that the chromatophore was indeed rigi- 
dified.


The authors here outline a computational methodology for the monomerization of red fluorescent proteins without significant loss of fluores- 
cence. Following screening of a 96-member computationally designed library of DsRed mutants containing 13–16 surface substitutions, the authors found that 97% of these were fluorescent and monomeric.


In this first example of rational design of chromophore maturation, the authors introduced red chromophore maturation into a yellow-emitting, maturation-deficient mutant of mPlum. To do so, they mutated residue Glu16, which forms an H-bond preventing the peptide bond trans-cis isomerization necessary for the second oxidation reaction, to non-polar amino acids Ala and Pro.


Here, the authors engineered mCherry for increased photostability by removing a pathway for oxygen access to the chromatophore. Mutagen- 
esis was guided by molecular dynamics simulations, and a recently developed microfluidics-based platform was used to screen for photo-

Rational design of red fluorescent proteins

Eason, Damry and Chica


