



US 20100184954A1

(19) **United States**  
(12) **Patent Application Publication**  
Glick et al.

(10) **Pub. No.: US 2010/0184954 A1**  
(43) **Pub. Date: Jul. 22, 2010**

(54) **MONOMERIC RED FLUORESCENT PROTEINS**

(75) Inventors: **Benjamin S. Glick**, Chicago, IL (US); **Brooke Bevis**, Somerville, MA (US); **Daniel E. Strongin**, Chicago, IL (US); **David Baker**, Seattle, WA (US); **Michelle Scalley-Kim**, Seattle, WA (US)

Correspondence Address:  
**MICHAEL BEST & FRIEDRICH LLP**  
100 East Wisconsin Avenue, Suite 3300  
Milwaukee, WI 53202 (US)

(73) Assignees: **THE UNIVERSITY OF CHICAGO**, Chicago, IL (US); **THE UNIVERSITY OF WASHINGTON**, Seattle, WA (US)

(21) Appl. No.: **12/689,276**

(22) Filed: **Jan. 19, 2010**

**Related U.S. Application Data**

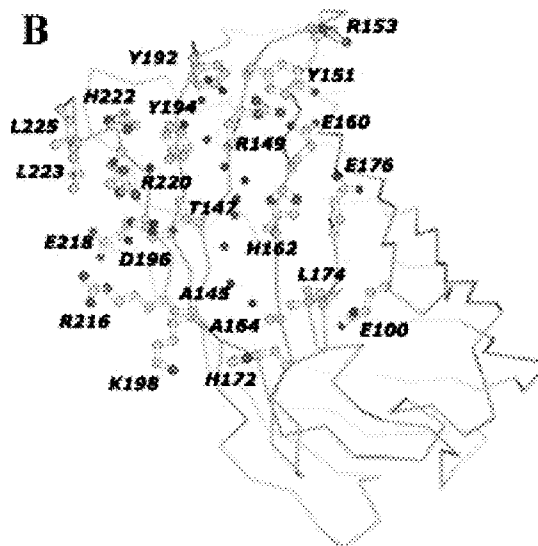
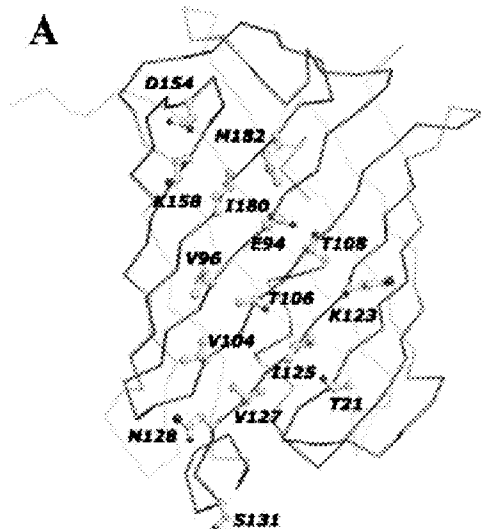
(63) Continuation of application No. 11/748,641, filed on May 15, 2007, now Pat. No. 7,671,185, which is a continuation of application No. 11/100,988, filed on Apr. 7, 2005, now Pat. No. 7,250,298.  
(60) Provisional application No. 60/560,340, filed on Apr. 7, 2004.

**Publication Classification**

(51) **Int. Cl.**  
*C07K 14/435* (2006.01)  
*C07H 21/04* (2006.01)  
(52) **U.S. Cl.** ..... **530/350; 536/23.5**

(57) **ABSTRACT**

Disclosed are sequences encoding monomeric variants of DsRed fluorescent proteins and methods of use.



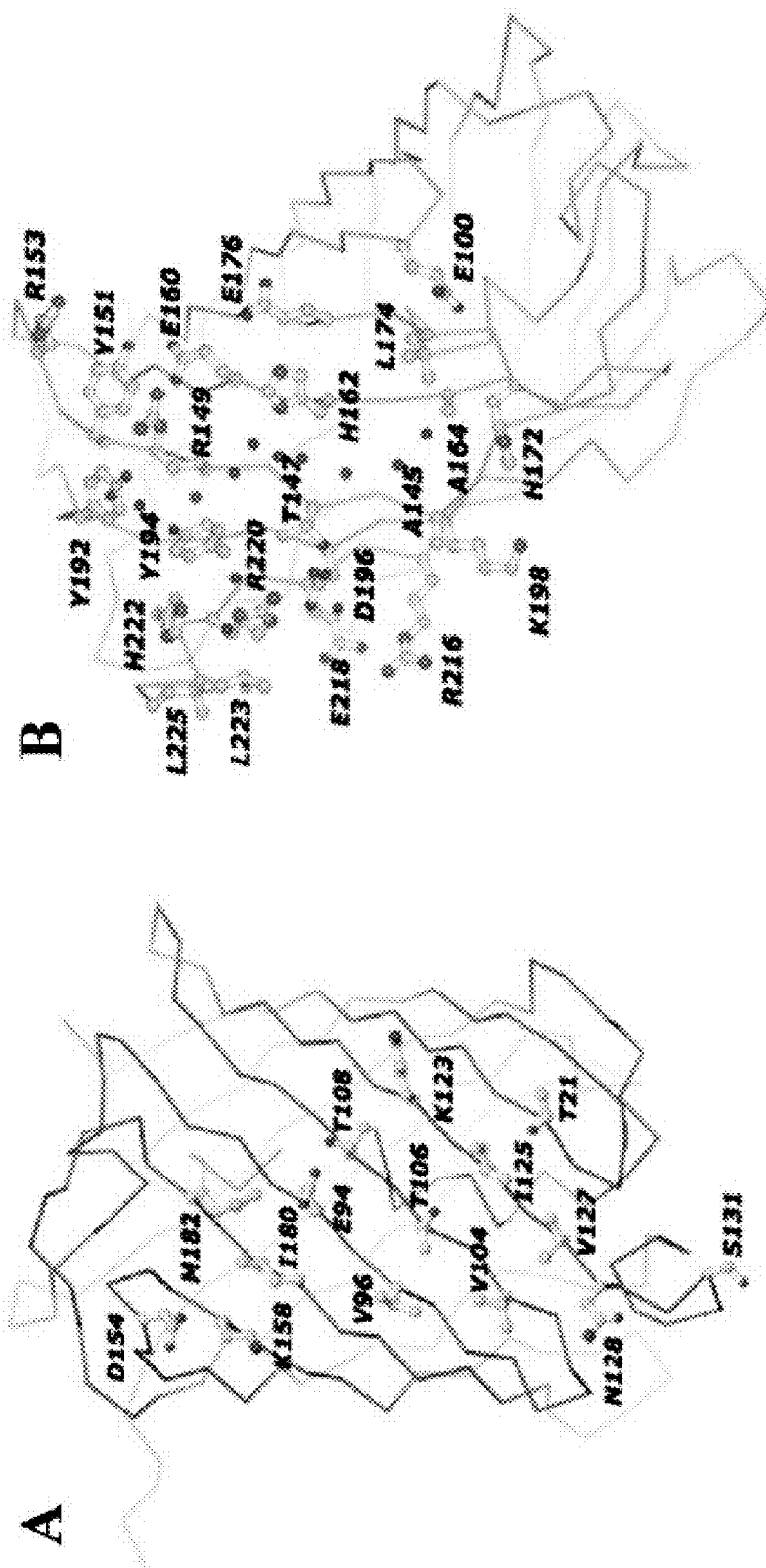


FIG. 1

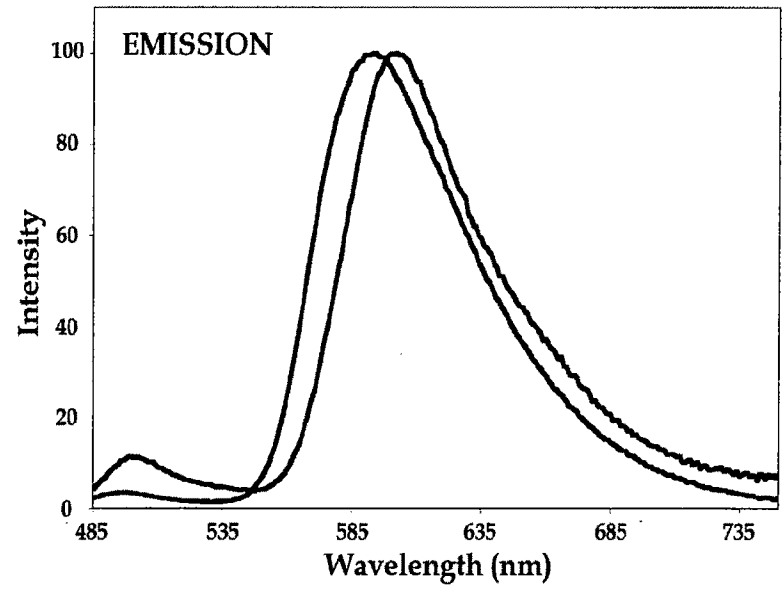
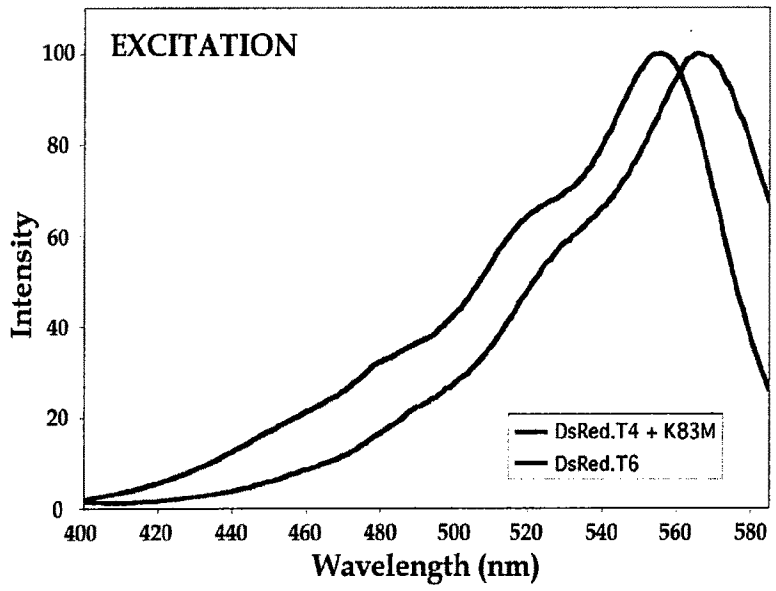


FIG 2

## MONOMERIC RED FLUORESCENT PROTEINS

### CROSS-REFERENCE TO RELATED APPLICATIONS

**[0001]** This application is a continuation of U.S. application Ser. No. 11/100,988, filed Apr. 7, 2005, which claims priority to U.S. Provisional Application No. 60/560,340, filed Apr. 7, 2004, each of which is incorporated by reference in its entirety.

### STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH

**[0002]** This invention was made with government support under MCB-9875939, awarded by the National Science Foundation, and under RPG-00-245-01-CSM, awarded by the American Cancer Society. The government has certain rights in the invention.

### INTRODUCTION

**[0003]** Fluorescent proteins such as green fluorescent protein (GFP) are invaluable tools used in pure and applied research. Although fluorescent proteins have been widely available for a relatively short time, they have had an important impact on biomedical research, contributing to our understanding of basic cellular and developmental processes that underlie health and disease. GFP and its relatives are widely used for medically-oriented research. For example, GFP has been used to analyze bacterial gene expression during infection, to visualize tumor cell behavior during metastasis, and to monitor GFP fusion proteins in gene therapy studies. Fluorescent proteins are also useful in high-throughput screens for drug discovery.

**[0004]** A red fluorescent protein produced by the coral *Discosoma* and designated DsRed (wild-type DsRed) is potentially useful as a fluorescent reporter protein or as a fusion tag. A red fluorescent protein is particularly attractive because of its suitability for use in conjunction with fluorescent proteins having different fluorescent properties, such as GFP. However, wild-type DsRed suffers from certain drawbacks.

**[0005]** First, the maturation process that yields the red fluorophore is slow, with a half-time of ~12 h at 37° C. Second, wild-type DsRed occurs as a stable tetramer of four very similar polypeptides, which makes its use as a fluorescent reporter in a fusion protein problematic. For example, tetramerization of the DsRed fusion protein may interfere with or perturb the function or localization of the protein. In addition, DsRed tetramers undergo higher-order aggregation. Fusion of DsRed to membrane proteins or to oligomeric proteins often produces large aggregates. Efforts to develop variants of DsRed having a reduced tendency to form tetramers have met with limited success, in that the variants suffer from disadvantages such as undesirable shifts in fluorescence or more rapid photobleaching.

**[0006]** There is, therefore, ongoing interest in developing new fluorescent protein labels with improved characteristics as experimental and clinical tools.

### SUMMARY OF THE INVENTION

**[0007]** The present invention provides polynucleotide encoding a variant polypeptide of wild-type DsRed or the rapidly maturing DsRed.T4, a tetrameric variant that has the substitutions described herein below. The variant polypeptide

has reduced oligomerization relative wild-type DsRed and has a fluorescence spectra similar to the fluorescence spectra of the wild-type DsRed, the variant polypeptide comprising the amino acid substitutions K83M, K163H, and Y193H and further comprises at least of one amino acid substitutions E26Y, K92T, V96S, T106E, T108Q, I125K, S131A, I180V, and M182K.

**[0008]** In another aspect, the invention provides isolated polynucleotides encoding a variant polypeptide of the rapidly DsRed.T4 and exhibiting reduced oligomerization relative to DsRed.T4 and detectable red fluorescence, the variant polypeptide comprising at least one of amino substitutions K83M or K83L; K163Q, K163M, or K163H; and Y193H, and further comprising at least three amino acid substitutions selected from E26Y, K92T, V96S, T106E, T108Q, I125K, S131A, I180V, and M182K; and further comprising at least three amino acid substitutions selected from the group consisting of R149K, R153Q, H162S, L174T, E176D, Y192N, R216H, H222S, L223G, and F224S.

**[0009]** In other aspects, the invention provides genetic constructs comprising the polynucleotides, vectors comprising the constructs, cells comprising the constructs, variant polypeptides encoded by the polynucleotides, and methods of obtaining expression of the polynucleotides.

### BRIEF DESCRIPTION OF THE DRAWINGS

**[0010]** FIG. 1 depicts the hydrophobic interface (FIG. 1A) and the polar interface (1B) involved in tetramerization of DsRed, the latter of which also includes the hydrophobic tail (residues 222-225).

**[0011]** FIG. 2 compares the excitation and emission spectra for a derivative of tetrameric variant DsRed.T4 having a K83M substitution, and a derivative of DsRed.T4 having K83M, K163H, and Y193H substitutions.

### DETAILED DESCRIPTION

**[0012]** The present invention provides polypeptide variants of wild-type *Discosoma* sp. red fluorescent protein (DsRed). Wild-type DsRed forms oligomers (e.g., dimers or tetramers) under physiologic conditions, which, in the native polypeptide, appears to play a role in its fluorescence activity. A coding sequence for wild-type *Discosoma* sp. red fluorescent protein (DsRed) is shown in SEQ ID NO:1, and the amino acid sequence of DsRed is shown in SEQ ID NO:2. The polypeptide variants of the present invention have a reduced tendency to oligomerize relative to wild-type DsRed. This reduced tendency to oligomerize may be assessed by any suitable means, whether in vivo or in vitro, as described below.

**[0013]** Polypeptide variants were genetically engineered, as described in detail below, by altering a sequence encoding a rapidly maturing tetrameric variant of wild-type DsRed, designated DsRed.T4 (SEQ ID NO:3). DsRed.T4 is described in further detail in co-pending U.S. patent application Ser. No. 10/844,064, which is incorporated by reference in its entirety. Relative to the wild-type DsRed of SEQ ID NO:2, DsRed.T4 contains the following substitutions:

P(-4)L	H41T
R2A	N42Q
K5E	V44A

-continued

N6D T21S	A145P T217A
-------------	----------------

**[0014]** Relative to either wild-type DsRed or DsRedT4, the polypeptides of the present invention exhibit a reduced tendency to oligomerize or form tetramers, and exist primarily as monomers under physiologic conditions. The polypeptide variants exhibit detectable red fluorescence. By “detectable red fluorescence” it is meant that the fluorescence overlaps the emission spectra of wild-type DsRed is distinguishable over background. Preferably, the emission spectra is similar to that of wild-type DsRed.

**[0015]** To develop monomeric DsRed variants, a polynucleotide sequence encoding DsRedT4 was modified using the following general approach. As one skilled in the art will appreciate, one could also begin with a polynucleotide encoding another DsRed, for example, wild-type DsRed.

**[0016]** Mutations were introduced into a polynucleotide sequence encoding DsRedT4 to disrupt the tetramerization interfaces that form between DsRed polypeptides. Amino acids suspected of contributing to oligomerization were replaced by amino acids that are likely to reduce oligomerization.

**[0017]** Crystal structures of DsRed reveal residues that may be involved in tetramerization. The region that we have designated the “hydrophobic interface” primarily engages in hydrophobic interactions that exclude the solvent (FIG. 1A). The region that we have designated the “polar interface” primarily engages in polar interactions, although this interface also includes some hydrophobic residues (Leu-174 and the C-terminal tail) that interact with the opposing subunit (FIG. 1B).

**[0018]** Two considerations guided selection of residues to be mutated and the specific substitutions to be tested. First, certain residues are more likely than others to make a significant energetic contribution to oligomerization. For example, a statistical study revealed that isoleucine side chains are often important for protein-protein interactions (Bogan and Thorn, 1998). In designing the variants of the present invention, it was assumed that surface-exposed hydrophobic residues are likely to promote oligomerization, and therefore, such residues were considered prime candidates for mutagenesis. Thus, it is likely that generation of a stable monomer requires modification of residues Ile-180 and Ile-125 (FIG. 1A). Second, the DsRed protein will tolerate substitutions more readily at some positions than at others. Many mutations at the tetramerization interfaces will have secondary effects on the folding and/or maturation of the protein. To make educated guesses about which residues can be mutated and which alternative residues can be substituted, we used sequence alignments between DsRed, GFP, and the other known fluorescent proteins (Matz et al., 1999; Wall et al., 2000).

**[0019]** In some polypeptide variants, mutations in two of the tetramerization interfaces resulted in a loss of tetramerization, which can be measured *in vitro* or *in vivo* as described in the Examples, and loss of fluorescence. It was, therefore, speculated that tetramerization may be important for fluorescence in wild-type DsRed, possibly because the autocatalytic

formation of the chromophore may stabilize the folded structure of DsRed or because oligomerization may increase the quantum yield.

**[0020]** To restore fluorescence to an oligomerization-disrupted or monomeric DsRed variants, site-directed mutagenesis was performed to introduce specific changes that were expected to stabilize the immature and/or mature forms of the protein. Two general classes of mutations were introduced in an attempt to restore fluorescence after disrupting the tetramer: (1) mutations that may stabilize folding intermediates to permit the protein to mature even in the absence of oligomerization; and (2) mutations expected to rigidify mature DsRed to enhance quantum yield.

**[0021]** In a second strategy to restore fluorescence to an oligomerization-disrupted or monomeric DsRed variant, a wide range of mutations were introduced into either the entire DsRed coding sequence or into a selected segment of the coding sequence. The variants thus generated were screened for improved fluorescence.

**[0022]** The effect of various mutations has, in some cases, been demonstrated, and in other cases, has been presumed or inferred. The mutations can be divided into the following six categories of effects:

1. Mutations that presumably stabilize immature folding intermediates. In tetrameric DsRed variants, K83 and K163 face the interior of the protein. The presence of these highly polar internal residues is likely to destabilize folding intermediates. Therefore, K83M and K163H substitutions were introduced to reduce the polarity at these positions. The K163H substitution is relatively conservative because histidine is still somewhat polar. The K83M substitution introduces a hydrophobic residue, and red-shifts the fluorescence spectra. However, it was discovered by random mutagenesis that a Y193H substitution reverses the spectral changes caused by K83M, and also enhances brightness. Residue 193 is close to residue 83 of the folded protein, so the Y193H substitution may compensate for the reduced polarity caused by K83M. Variants containing the trio of substitutions K83M, K163H, Y193H seem to be stabilized, as indicated by their ability to tolerate interface mutations that are not tolerated by DsRed.T4.

2. Mutations that disrupt the hydrophobic tetramerization interface. This category comprises E26Y, K92T, V96S, T106E, T108Q, I125K, S131A, I180V, and M182K, each of which is predicted to disrupt intersubunit interactions at the hydrophobic interface.

3. Mutations that disrupt the polar tetramerization interface. This category comprises R149K, R153Q, H162S, L174T, E176D, Y192N, R216H, H222S, L223G, F224S, and L225Q, each of which is predicted to disrupt intersubunit interactions at the polar interface. The hydrophobic C-terminal tail of DsRed is considered to be part of the polar interface.

4. Mutations that enhance monomer fluorescence, presumably by stabilizing and/or rigidifying the protein. This category comprises V71A, C117T, V175C, S179T, S203N, and G219A. Note that C117T also eliminates a surface cysteine residue that might otherwise be oxidized within the secretory pathway.

5. Mutations that improve expression in *E. coli*. When a variant such as DsRed.T4 is produced in *E. coli* using its own start codon, the protein levels are much lower than when the same protein is produced using an N-terminal hexahistidine tag, which suggests that the 5' end of the gene is important for expression in bacteria. The coding sequences were altered

such that the putative translation products contained substitutions in residues 2-4 and screened for strong expression in *E. coli*. A polypeptide having the substitutions A2D, S3N, and S4T was found to be strongly expressed. These N-terminal mutations probably enhance translation in *E. coli*, an effect that may not occur with expression in eukaryotic cells. In addition, the N-terminal mutations might have some beneficial effect on the maturation and/or stability of DsRed.

6. Mutations that reduce the general "stickiness" of the protein. We introduced a number of surface mutations that lie outside of the tetramerization interfaces that should reduce the tendency of the protein to adhere to other macromolecules. Specifically, we targeted basic residues (especially arginines) and large, hydrophobic residues. The substitutions in this category include R13Q, R36K, K47Q, M141A, and I210V. We have shown that these substitutions do not significantly reduce fluorescence, and expect that variants containing one or more of these substitutions may exhibit reduced aggregation with other macromolecules. Reduced aggregation may be evaluated by comparing the solubility of DsRed variants to that of wild-type DsRed.

**[0023]** An example of a DsRed monomeric variant containing numerous mutations and identified as DsRed.M1 (SEQ ID NO:5) is described below. Of course, useful monomeric variants of the presently claimed invention having fewer mutations than those contained in the DsRed.M1 may be developed using the guidance and teaching herein of this disclosure.

**[0024]** DsRed.M1 was further modified by introducing D6N, a reversion to the native sequence, in order to reduce the number of acidic residues at the N-terminus. The additional substitutions K121H, K168E, D169G, D115G, and G116N were made to improve brightness or fluorescence.

**[0025]** Once a DsRed monomeric variant having desirable characteristics is identified, one may, of course, use any polynucleotide sequence encoding the variant to express the variant. For example, the polynucleotide encoding the variant may be modified for optimal expression in a particular organism in view of the preferred codon usage of that organism. The polynucleotide may be operably linked to an inducible or constitutive promoter functional in the intended cell or organism. The polynucleotide may be linked in-frame to a second polynucleotide sequence encoding a polypeptide of interest to form a sequence encoding a fusion protein in which the polypeptide of interest is labeled with the DsRed monomeric variant at its N- or C-terminus.

## EXAMPLES

### Selection of Amino Acid Substitutions to Reduce Tetramerization

**[0026]** Site-directed mutagenesis using standard methods well-known to one of ordinary skill in the art was used to disrupt the two tetramerization interfaces of DsRed. Various substitutions at each position were made, and fluorescence was evaluated as described below.

### The Hydrophobic Interface

**[0027]** Extensive mutagenesis of the hydrophobic interface was performed, beginning with the DsRed.T4 variant. DsRed.T4 contains a T215 substitution. In addition, Met-182, Ile-180, Val-96 and Ile-125 were substituted with more polar residues, and a hydrogen-bonding residue was removed through an S131A substitution. The DsRed.T4 variant having

these substitutions was designated DsRed.D1. This variant is probably dimeric. An additional V104A substitution is tolerated in the DsRed.D1 background. It is of note that, in the related fluorescent protein HcRed, a single Leu-to-His mutation in the putative hydrophobic interface at the position corresponding to Ile-125 of DsRed was reportedly sufficient to convert HcRed into a dimer (Gurskaya et al., 2001). Bacterial colonies producing DsRed.D1 are somewhat less fluorescent than those producing DsRed.T4, but the signal with DsRed.D1 is still strong.

### The Polar Interface

**[0028]** Initial attempts to mutagenize the polar interface yielded reduced the fluorescence. For example, Leu-174 forms hydrophobic interactions with the opposing subunit, but all of the Leu-174 substitutions that we initially tried, including a conservative change to Val, virtually abolished fluorescence. It was thus concluded that Leu-174 is required for fluorescence and should not be mutated. Similar results were obtained with His-162, which associates with its counterpart on the opposing subunit in an unusual stacking interaction, and with His-222, which inserts into a groove in the opposing subunit. Conservative mutations of these His residues to Ser or Asn severely diminished the fluorescence. Indeed, mutagenesis of most the key residues at the polar interface were found to impair fluorescence.

### Assessing Oligomeric State of DsRed Variants

**[0029]** The oligomeric state of a DsRed variant may be assessed by non-denaturing SDS-PAGE or size exclusion chromatography.

**[0030]** For non-denaturing SDS-PAGE, one  $\mu\text{g}$  of each purified DsRed variant or wild-type DsRed was mixed with SDS-containing sample buffer on ice and immediately electrophoresed at 4° C. in a 10% polyacrylamide gel, followed by staining with Coomassie Blue. Additional aliquots of wild-type DsRed and DsRed variants were denatured by boiling prior to electrophoresis.

**[0031]** Gel filtration chromatography may be used to indicate whether a fluorescent protein exists predominantly as a monomer, dimer or tetramer (Gurskaya et al., 2001). Optionally, gel filtration will be conducted using the Pharmacia FPLC system. An extension of this method known as small-zone size-exclusion gel filtration chromatography could be used to measure association constants (Raffen and Stevens, 1999).

**[0032]** Oligomerization states and association constants of our DsRed variants may be evaluated using analytical ultracentrifugation (Laue and Stafford, 1999). This approach was used by Baird et al. (2000) for their initial demonstration that DsRed is a tetramer. A Beckman XL-A analytical ultracentrifuge will be used. Velocity sedimentation will be used to ascertain which DsRed species are present (monomers, dimers and/or tetramers), and then equilibrium sedimentation will be used to measure the oligomer association constants (Laue and Stafford, 1999).

**[0033]** Analytical ultracentrifugation will be carried out in collaboration with Borries Demeler (University of Texas Health Sciences Center at San Antonio), who has extensive experience with analytical ultracentrifugation (eg., Demeler and Saber, 1998) and is the author of the state-of-the-art UltraScan II software ([www.ultrascan.uthscsa.edu](http://www.ultrascan.uthscsa.edu)).

An *in vivo* assay for DsRed oligomerization may be used in a genetic screen. For example, a yeast two-hybrid system such as that reported to have been used to verify the oligomerization of wild-type DsRed (Baird et al., 2000) may be used to assess the tendency of DsRed variants to oligomerize.

**[0034]** Another approach used to monitor DsRed oligomerization *in vivo* was to fuse GFP or DsRed to the protein Gos1p in *S. cerevisiae*. Gos1p is a membrane protein anchored to the cytoplasmic face of yeast Golgi cisternae. The GFP-Gos1p fusion protein gives a fluorescence pattern that is typical for the Golgi in *S. cerevisiae*, whereas a wild-type DsRed Gos1p fusion generates large red blobs in the cells, presumably because multiple Golgi cisternae become crosslinked via DsRed tetramerization. Thus, visualizing DsRed-Gos1p fusions provides an *in vivo* assay for DsRed oligomerization.

**[0035]** In another screening assay, randomly mutagenized DsRed proteins will be fused to the C-terminus of glutathione S-transferase (GST; Smith and Johnson, 1988). Because GST is a dimer (McTigue et al., 1995), the fusion of GST to an oligomeric DsRed variant will generate crosslinked aggregates that will be insoluble upon gentle detergent lysis of the cells. By contrast, the fusion of GST to a monomeric DsRed variant will generate a soluble protein. The concept of using this method to evaluate the oligomeric state of DsRed variants was tested using an expression and detergent lysis protocol similar to that used to evaluate solubility or aggregation of DsRed, fluorescent DsRed.T1, a tetrameric variant, was efficiently extracted from the bacterial cells whereas the fluorescent GST-DsRed.T1 was quantitatively retained in the pellet. It is expected that this assay will provide a sensitive screen for a monomeric DsRed.

#### Optimizing Spectral Properties

**[0036]** In an earlier stage of this project, we attempted to create a bright, red-shifted DsRed variant. We began with DsRed.T4 and introduced a K83M substitution, which had been shown to red-shift the fluorescence spectra of wild-type DsRed (Baird et al., 2000). In the DsRed.T4 background, K83M red-shifted the spectra (FIG. 2), but substantially reduced the intrinsic brightness of the protein. Screening of randomly mutagenized variants was undertaken to identify variants having restored fluorescence. Surprisingly, an L174Q mutation was found to increase brightness. Subsequent tests confirmed that in the context of K83M, other substitutions at the polar interface, including H162S and H222S, preserve or restore colony fluorescence.

**[0037]** Without being limited as to theory, we speculate that the K83M mutation allows DsRed to tolerate changes at the polar interface by stabilizing the immature form of the DsRed, thereby allowing for correct folding. The cores of most proteins are hydrophobic, but in the case of DsRed, K83 is one of several charged residues that face the interior of the protein (Wall et al., 2000; Yarbrough et al., 2001). It may be that immature DsRed is stabilized by tetramerization, particularly by interactions at the polar interface, and that the K83M substitution renders the immature protein sufficiently stable to fold even in the absence of interactions at the polar interface. Based on this hypothesis, we predict that in the absence of K83M, mutations such as L174Q will decrease the yield of mature DsRed but will not reduce the intrinsic fluorescence of the mature protein. Meanwhile, regardless of the reason, K83M has enabled us to mutagenize the polar interface.

**[0038]** To address the reduced fluorescence found with variants having a K83M substitution, two rounds of random mutagenesis and screening were undertaken. Two additional substitutions (K163H and Y193H) that significantly increase the brightness of the purified protein were identified. These new mutations are distinct from L174Q, and they alter residues that face the interior of the protein. The variant of DsRed.T4 having the K83M, K163H, and Y193H substitutions was designated DsRed.T6. As judged by colony fluorescence, DsRed.T6 is comparable to DsRed.T4 in brightness. Interestingly, the spectral red-shifting observed with K83M is largely reversed by the two additional substitutions in DsRed.T6 (FIG. 2). As can be seen from FIG. 2, the K83M substitution red-shifts the excitation and emission peaks by ~20 nm. The two additional substitutions present in DsRed.T6 reverse most of this red-shifting, and also suppress the green emission. We have also introduced K83M, K163H, and Y193H substitutions into the DsRed.D1 background and the resulting variant, designated DsRed.D3, has strong fluorescence and tolerates mutations at the polar interface. We plan to start with DsRed.D3 and mutate most or all of the residues that contribute to the polar interface.

**[0039]** Preferably, the monomeric DsRed variant of the present invention retains the spectral properties of the tetramer, namely, bright red fluorescence with minimal green emission. However, fluorescent bacterial colonies expressing polypeptides containing substituted amino acids for those of the wild-type DsRed that ordinarily form the hydrophobic tetramerization interface appear somewhat dimmer than colonies of bacteria expressing tetrameric variants. Two kinds of changes may account for decreased fluorescence of the bacterial colonies expressing mutant DsRed. First, a mutation may reduce the intrinsic brightness of DsRed by lowering the extinction coefficient and/or the quantum yield. Second, a mutation might slow DsRed maturation and/or reduce the percentage of the DsRed molecules that eventually become fluorescent. To counteract such effects, random mutagenesis strategy will be used to identify brighter variants of the monomers.

#### Minimizing the Green Emission of Monomeric DsRed

**[0040]** A predicted side effect of disrupting the DsRed tetramer will be a loss of FRET and a consequent increase in the green emission. To alleviate this problem, mutations that increase the ratio of red to green molecules in mature DsRed may be introduced. Substitutions of K83M, K163H, and Y193H in DsRed.T6 correlate with brighter red fluorescence and reduced green emission, relative to that of the DsRed.T4+K83M variant.

**[0041]** In addition to mutations identified in these random screens, directed mutations aimed at weakening the tetramerization may fortuitously reduce the percentage of green molecules in mature DsRed. For instance, the A145P substitution that we incorporated into DsRed.T3 and DsRed.T4 to lower the green emission was originally generated during our attempts to mutagenize the polar interface. More recently, we found that an H222S substitution at the polar interface decreases the green emission.

#### Screening for Reduced Blue Excitation

**[0042]** Monomeric variants of interest having significant green emission, may be further mutagenized and screened for reduced green emission or reduced blue excitation. A 488-nm

laser is used to excite fluorescence in bacteria containing mutant DsRed proteins, and the bacterial cells are then sorted by flow cytometry to identify clones with reduced green emission. This approach would be more difficult with a monomeric DsRed, because undesirable mutations that caused protein aggregation would suppress the green emission due to FRET. Alternatively, mutants may be screened for reduced blue excitation. When total fluorescence is measured, the signal obtained by excitation with blue light should correlate with the percentage of the DsRed molecules having a green fluorophore, regardless of whether the emission spectrum has been modified by FRET.

**[0043]** The assay for reduced blue excitation is based on our standard slide projector method, except that the plates will be photographed with a digital camera. We will photograph each plate under two conditions. First, the total fluorescence after excitation with blue light will be recorded by illuminating through a 485±11 nm bandpass filter and capturing the emission signal through a Kodak Wratten filter #12, which passes wavelengths above 520 nm (Cronin and Hampton, 1999). The image of the plate will be colorized green using Adobe Photoshop. Second, the total fluorescence after excitation with yellow light will be recorded by illuminating through a 520±20 nm bandpass filter and capturing the emission signal through a Kodak Wratten filter #22, which passes wavelengths above 550 nm. This image of the plate will be colorized red. The red image will be due almost exclusively to DsRed molecules with the red fluorophore, whereas the green image will include a strong contribution from DsRed molecules with the green fluorophore. Upon merger of the red and green images, most of the colonies will appear yellow, but colonies having an increased ratio of red to green molecules will appear orange.

#### Evaluating Solubilities

**[0044]** One approach to evaluating solubilities of the fluorescent proteins is as follows. *E. coli* cells carrying inducible expression vectors encoding hexahistidine-tagged wild-type DsRed or DsRed variants were grown to an optical density (OD) ( $\lambda=600$  nm) of 0.5, induced for 7 h, lysed with B-PER II detergent (Pierce) and centrifuged for 20 min at 27,000×g. Equivalent amounts of the pellet and supernatant fractions were subjected to SDS PAGE followed by immunoblotting with an anti-hexahistidine antibody. The percentage of each protein in the supernatant fraction was then quantified for each protein and the percentage of protein molecules extracted (i.e., solubilized) was determined. Typically, only about 25% of the wild-type DsRed molecules are solubilized. Suitably, at least 30% of DsRed variant molecules are solubilized. Preferably, at least 50% of DsRed variant molecules are solubilized. More preferably, at least 70% of DsRed variant molecules are solubilized.

**[0045]** Another approach to evaluate variants for reduced aggregation is by nondenaturing SDS-PAGE. One  $\mu$ g of each purified DsRed variant was mixed with SDS-containing sample buffer on ice and immediately electrophoresed at 4° C. in a 10% polyacrylamide gel, followed by staining with Coomassie Blue. Additional aliquots of wild-type DsRed and DsRed variants were denatured by boiling prior to electrophoresis. Migration of proteins as a diffuse band that may reflect the formation of higher-order oligomers, whereas formation of a sharp band or bands of the appropriate size suggests reduce aggregation.

#### Creation and Characterization of a Red Fluorescent DsRed Monomer

**[0046]** A coding sequence and the amino acid sequence of one mutant according to the present invention, designated

DsRed.M1, is shown as SEQ ID NO:4 and SEQ ID NO: 5, respectively. DsRed.M1 contains the following 37 substitutions relative to DsRed.T4:

A2D*	V96S	K163H	I210V
S3N	T106E	L174T	R216H
S4T	T108Q	V175C	G219A
R13Q	C117T	E176D	H222S
E26Y	I125K	S179T	L223G
R36K	S131A	I180V	F224S
K47Q	M141A	M182K	L225Q
V71A	R149K	Y192N	
K83M	R153Q	Y193H	
K92T	H162S	S203N	

\*A2D is a second mutation of residue 2, which was an arginine in wild-type DsRed.

**[0047]** Oligomeric State: DsRed.M1 behaves functionally as a monomer. The protein appears to be monomeric as judged by nondenaturing SDS-PAGE or size exclusion chromatography.

**[0048]** Brightness: The DsRed.M1 mature protein exhibits detectable red fluorescence, but is less bright than tetrameric variants, such as DsRed.T4. DsRed.M1 is also less bright than a monomeric red fluorescent DsRed protein designated mRFP1, which was generated by Tsien et al. (US 2003/0170911 A1)

**[0049]** Spectral Properties: In contrast to mRFP1, which has a red-shift, DsRed.M1 has excitation and emission spectra similar to those of the tetrameric DsRed variants. Surprisingly, DsRed.M1 has negligible green emission. The lack of green emission is fortunate, and rather unexpected, given that the tetrameric DsRed variants exhibit green emission that is largely suppressed by intersubunit resonance energy transfer.

**[0050]** Photostability: mRFP 1 photobleaches much faster than the tetrameric DsRed variants, an effect that may correlate with the altered spectral properties of mRFP1. It is expected that DsRed.M1 will be more photostable than mRFP1 because the spectral properties of DsRed.M1 are similar to those of the tetrameric variants.

**[0051]** Maturation: DsRed.M1 appears to retain the rapid maturation of the parental DsRed.T4, which is considerably faster than that wild-type of DsRed.

**[0052]** Each publication cited in herein or in the appendix is incorporated by reference in its entirety. Also incorporated by reference in its entirety is WO 03/054158A2.

#### LITERATURE CITED

- [0053]** Baird, G. S., D. A. Zacharias, and R. Y. Tsien. 2000. Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proc. Natl. Acad. Sci. USA.* 97:11984-11989.
- [0054]** Bogan, A. A., and K. S. Thorn. 1998. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* 280:1-9.
- [0055]** Cronin, S., and R. Hampton. 1999. A genetics-friendly GFP assay. *Trends Cell Biol.* 9:36.
- [0056]** Demeler, B., and H. Saber. 1998. Determination of molecular parameters by fitting sedimentation data to finite-element solutions of the Lamm equation. *Biophys. J.* 74:444-454.
- [0057]** Gurskaya, N. G., A. F. Fradkov, A. Terskikh, M. V. Matz, Y. A. Labas, V. I. Martynov, Y. G. Yanushevich, K. A. Lukyanov, and S. A. Lukyanov. 2001. GFP-like chromoproteins as a source of far-red fluorescent proteins. *FEBS Lett.* 507:16-20.



- [0058] Laue, T. M., and W.F. Stafford, 3rd. 1999. Modern applications of analytical ultracentrifugation. *Annu. Rev. Biophys. Biomol. Struct.* 28:75-100.
- [0059] Matz, M. V., A. F. Fradkov, Y. A. Labas, A. P. Savitsky, A. G. Zaraisky, M. L. Markelov, and S. A. Lukyanov. 1999. Fluorescent proteins from nonbioluminescent *Anthozoa* species. *Nat. Biotechnol.* 17:969-973.
- [0060] McTigue, M. A., D. R. Williams, and J. A. Tainer. 1995. Crystal structures of a schistosomal drug and vaccine target: glutathione S-transferase from *Schistosoma japonica* and its complex with the leading antischistosomal drug praziquantel. *J. Mol. Biol.* 246:21-27.
- [0061] Pierce, D. W., and R. D. Vale. 1999. Single-molecule fluorescence detection of green fluorescence protein and application to single-protein dynamics. In *Green Fluorescent Proteins (Methods in Cell Biology, Vol. 58)*. K. F. Sullivan, and S. A. Kay, editors. Academic Press, San Diego. 49-73.
- [0062] Raffin, R., and F. J. Stevens. 1999. Small zone, high-speed gel filtration chromatography to detect protein aggregation associated with light chain pathologies. *Methods Enzymol.* 309:318-332.
- [0063] Smith, D. B., and K. S. Johnson. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene.* 67:31-40.
- [0064] Tsien, R. Y. 1998. The green fluorescent protein. *Annu. Rev. Biochem.* 67:509-544.
- [0065] Wall, M. A., M. Socolich, and R. Ranganathan. 2000. The structural basis for red fluorescence in the tetrameric GFP homolog DsRed. *Nat. Struct. Biol.* 7:1133-1138.
- [0066] Yarbrough, D., R. M. Wachter, K. Kallio, M. V. Matz, and S. J. Remington. 2001. Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-Å resolution. *Proc. Natl. Acad. Sci. USA.* 98:462-467.

## SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 5

<210> SEQ ID NO 1

<211> LENGTH: 678

<212> TYPE: DNA

<213> ORGANISM: Unknown

<220> FEATURE:

<223> OTHER INFORMATION: wild-type red fluorescent protein based on *Discosoma* species

<400> SEQUENCE: 1

```

atgagggtcctt ccaagaatgt tatcaaggag ttcattgaggt ttaaggttcg catggaagga      60
acgggtcaatg ggcacgagtt tgaatatagaa ggcgaaggag agggggaggcc atacgaaggc      120
cacaataccg taaagcttaa ggtaaccaag gggggacctt tgccatttgc ttgggatatt      180
ttgtcaccac aatttcagta tggaagcaag gtatatgtca agcaccttgc cgacatacca      240
gactataaaa agctgtcatt tcctgaagga tttaaatggg aaagggtcat gaactttgaa      300
gacggtggcg tcgttactgt aaccaggat tccagtttgc aggatggctg tttcatctac      360
aaggccaagt tcattggcgt gaactttcct tccgatggac ctgttatgca aaagaagaca      420
atgggctggg aagccagcac tgagcgtttg taccctctgt atggcgtgtt gaaaggagag      480
attcataagg ctctgaagct gaaagacggt ggtcattacc tagttgaatt caaaagtatt      540
tacatggcaa agaagcctgt gcagctacca gggactact atgttgactc caaactggat      600
ataacaagcc acaacgaaga ctatacaatc gttgagcagt atgaaagaac cgagggacgc      660
caccatctgt tcctttaa                                     678

```

<210> SEQ ID NO 2

<211> LENGTH: 225

<212> TYPE: PRT

<213> ORGANISM: Unknown

<220> FEATURE:

<223> OTHER INFORMATION: wild-type red fluorescent protein based on *Discosoma* species

<400> SEQUENCE: 2

```

Met Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val
1           5           10           15

```

-continued

---

```

Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu
      20                               25                               30
Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys Val
      35                               40                               45
Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln
      50                               55                               60
Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro
      65                               70                               75                               80
Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val
      85                               90                               95
Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser
      100                              105                              110
Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn
      115                              120                              125
Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu
      130                              135                              140
Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu
      145                              150                              155                              160
Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu
      165                              170                              175
Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr
      180                              185                              190
Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr
      195                              200                              205
Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe
      210                              215                              220
Leu
225

```

```

<210> SEQ ID NO 3
<211> LENGTH: 225
<212> TYPE: PRT
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: rapidly maturing DsRed.T4 red fluorescent
protein based on Discosoma species

```

```

<400> SEQUENCE: 3

```

```

Met Ala Ser Ser Glu Asp Val Ile Lys Glu Phe Met Arg Phe Lys Val
1      5      10      15
Arg Met Glu Gly Ser Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu
20     25     30
Gly Glu Gly Arg Pro Tyr Glu Gly Thr Gln Thr Ala Lys Leu Lys Val
35     40     45
Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln
50     55     60
Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro
65     70     75     80
Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val
85     90     95
Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser
100    105    110
Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn
115    120    125

```

-continued

```

Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu
 130                               135                               140

Pro Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu
145                               150                               155                               160

Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu
 165                               170                               175

Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr
 180                               185                               190

Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr
 195                               200                               205

Thr Ile Val Glu Gln Tyr Glu Arg Ala Glu Gly Arg His His Leu Phe
 210                               215                               220

Leu
225
    
```

```

<210> SEQ ID NO 4
<211> LENGTH: 678
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: mutant red fluorescent protein DsRed.M1 based
      on Discosoma species
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(678)
    
```

<400> SEQUENCE: 4

```

atg gac aac acc gag gac gtc atc aag gag ttc atg cag ttc aag gtg      48
Met Asp Asn Thr Glu Asp Val Ile Lys Glu Phe Met Gln Phe Lys Val
 1      5      10      15

cgc atg gag ggc tcc gtg aac ggc cac tac ttc gag atc gag ggc gag      96
Arg Met Glu Gly Ser Val Asn Gly His Tyr Phe Glu Ile Glu Gly Glu
      20      25      30

ggc gag ggc aag ccc tac gag ggc acc cag acc gcc aag ctg cag gtg      144
Gly Glu Gly Lys Pro Tyr Glu Gly Thr Gln Thr Ala Lys Leu Gln Val
      35      40      45

acc aag ggc ggc ccc ctg ccc ttc gcc tgg gac atc ctg tcc ccc cag      192
Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln
 50      55      60

ttc cag tac ggc tcc aag gcc tac gtg aag cac ccc gcc gac atc ccc      240
Phe Gln Tyr Gly Ser Lys Ala Tyr Val Lys His Pro Ala Asp Ile Pro
 65      70      75      80

gac tac atg aag ctg tcc ttc ccc gag ggc ttc acc tgg gag cgc tcc      288
Asp Tyr Met Lys Leu Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ser
      85      90      95

atg aac ttc gag gac ggc ggc gtg gtg gag gtg cag cag gac tcc tcc      336
Met Asn Phe Glu Asp Gly Gly Val Val Glu Val Gln Gln Asp Ser Ser
      100      105      110

ctg cag gac ggc acc ttc atc tac aag gtg aag ttc aag ggc gtg aac      384
Leu Gln Asp Gly Thr Phe Ile Tyr Lys Val Lys Phe Lys Gly Val Asn
      115      120      125

ttc ccc gcc gac ggc ccc gta atg cag aag aag act gcc ggc tgg gag      432
Phe Pro Ala Asp Gly Pro Val Met Gln Lys Lys Thr Ala Gly Trp Glu
      130      135      140

ccc tcc acc gag aag ctg tac ccc cag gac ggc gtg ctg aag ggc gag      480
Pro Ser Thr Glu Lys Leu Tyr Pro Gln Asp Gly Val Leu Lys Gly Glu
145      150      155      160
    
```

-continued

---

```

atc tcc cac gcc ctg aag ctg aag gac ggc ggc cac tac acc tgc gac      528
Ile Ser His Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Thr Cys Asp
                165                170                175

ttc aag acc gtg tac aag gcc aag aag ccc gtg cag ctg ccc ggc aac      576
Phe Lys Thr Val Tyr Lys Ala Lys Lys Pro Val Gln Leu Pro Gly Asn
                180                185                190

cac tac gtg gac tcc aag ctg gac atc acc aac cac aac gag gac tac      624
His Tyr Val Asp Ser Lys Leu Asp Ile Thr Asn His Asn Glu Asp Tyr
                195                200                205

acc gtg gtg gag cag tac gag cac gcc gag gcc cgc cac tcc ggc tcc      672
Thr Val Val Glu Gln Tyr Glu His Ala Glu Ala Arg His Ser Gly Ser
                210                215                220

cag tag                                                                678
Gln
225

```

```

<210> SEQ ID NO 5
<211> LENGTH: 225
<212> TYPE: PRT
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

```

```

<400> SEQUENCE: 5

```

```

Met Asp Asn Thr Glu Asp Val Ile Lys Glu Phe Met Gln Phe Lys Val
1          5          10          15

Arg Met Glu Gly Ser Val Asn Gly His Tyr Phe Glu Ile Glu Gly Glu
          20          25          30

Gly Glu Gly Lys Pro Tyr Glu Gly Thr Gln Thr Ala Lys Leu Gln Val
          35          40          45

Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln
          50          55          60

Phe Gln Tyr Gly Ser Lys Ala Tyr Val Lys His Pro Ala Asp Ile Pro
          65          70          75          80

Asp Tyr Met Lys Leu Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ser
          85          90          95

Met Asn Phe Glu Asp Gly Gly Val Val Glu Val Gln Gln Asp Ser Ser
          100         105         110

Leu Gln Asp Gly Thr Phe Ile Tyr Lys Val Lys Phe Lys Gly Val Asn
          115         120         125

Phe Pro Ala Asp Gly Pro Val Met Gln Lys Lys Thr Ala Gly Trp Glu
          130         135         140

Pro Ser Thr Glu Lys Leu Tyr Pro Gln Asp Gly Val Leu Lys Gly Glu
          145         150         155         160

Ile Ser His Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Thr Cys Asp
          165         170         175

Phe Lys Thr Val Tyr Lys Ala Lys Lys Pro Val Gln Leu Pro Gly Asn
          180         185         190

His Tyr Val Asp Ser Lys Leu Asp Ile Thr Asn His Asn Glu Asp Tyr
          195         200         205

Thr Val Val Glu Gln Tyr Glu His Ala Glu Ala Arg His Ser Gly Ser
          210         215         220

Gln
225

```

---

We claim:

1. An isolated polynucleotide encoding a variant polypeptide of wild-type DsRed of SEQ ID NO:1 or a variant of the rapidly maturing *Discosoma* red fluorescent protein T4 (DsRed.T4) of SEQ ID NO:3, the variant polypeptide having reduced oligomerization relative wild-type DsRed, the variant polypeptide comprising the amino acid substitutions K83M, K163H, and Y193H, the variant polypeptide further comprising at least one amino acid substitution selected from the group consisting of E26Y, K92T, V96S, T106E, T108Q, I125K, S131A, I180V, and M182K.

2. The isolated polynucleotide of claim 1, wherein the polypeptide comprises the amino acid substitution I125K or I180V.

3. The isolated polynucleotide of claim 1, wherein the polypeptide comprises at least one amino acid substitution selected from the group consisting of R149K, R153Q, H162S, L174T, E176D, Y192N, R216H, H222S, L223G, and F224S.

4. The isolated polynucleotide of claim 1, wherein the variant polypeptide further comprises at least one amino acid substitution selected from the group consisting of V71A, C117T, S175C, S179T, S203N, and G219A.

5. The isolated polynucleotide of claim 1, wherein the variant polypeptide further comprises at least one amino acid substitution selected from the group consisting of A2D, S3N, and S4T.

6. An isolated polynucleotide encoding a variant polypeptide of the rapidly maturing *Discosoma* red fluorescent pro-

tein T4 (DsRed.T4) of SEQ ID NO:3, the variant polypeptide having reduced oligomerization relative wild-type DsRed, wherein the variant matures more rapidly than wild-type DsRed, the variant polypeptide comprising:

at least one amino acid substitution selected from the group consisting of K83M, K83L, K163Q, K163M, K163H, and Y193H;

at least three amino acid substitutions selected from the group consisting of E26Y, K92T, V96S, T106E, T108Q, I125K, S131A, I180V, and M182K;

at least three amino acid substitutions selected from the group consisting of R149K, R153Q, H162S, L174T, E176D, Y192N, R216H, H222S, L223G, and F224S; and

at least one amino acid substitution selected from the group consisting of V71A, C117T, S175C, S179T, S203N, and G219A.

7. A variant polypeptide of DsRed.T4 (SEQ ID NO:3) comprising:

at least one amino acid substitution selected from the group consisting of K83M, K83L, K163Q, K163M, K163H, and Y193H;

at least three amino acid substitutions selected from the group consisting of E26Y, K92T, V96S, T106E, T108Q, I125K, S131A, I180V, and M182K; and

at least three amino acid substitutions selected from the group consisting of R149K, R153Q, H162S, L174T, E176D, Y192N, R216H, H222S, L223G, and F224S

\* \* \* \* \*