Supporting Information

Chica et al. 10.1073/pnas.1013910107

SI Text

SI Results and Discussion. ***Absorption spectroscopy.*** The absorption spectrum of mRojoB is unique among the most red-shifted mutants in that it has an additional peak centered at around 510 nm (Fig. S1*C*). This peak has been suggested to correspond to the deprotonated green form of the chromophore (1). The green form is equivalent to the *p*-hydroxybenzylideneimidazolinone group shown in Fig. 1 without the acylimine substitution on the five-membered ring. The deprotonated green form has been postulated to be a dead-end product created during chromophore maturation (2). The 510 nm peak is absent in the excitation spectrum of mRojoB monitored at 650 nm, indicating that the species associated with this peak does not contribute to far-red fluorescence. This result further validates that the 510 nm peak does not correspond to the acylimine substituted red form of the chromophore. The 510 nm peak is also found in the absorption spectrum of mRFP1 (3) from which mCherry is derived, as well as in the spectrum of mCherry point mutant Q163M (Fig. S1*E*). The 510 nm species could be caused by the combination of residues M163 and I161, which are found in mRojoB, mRFP1, and mCherry mutant Q163M, but are absent in mCherry, mRojoA, and mRouge.

mRouge has an additional absorbance peak at ∼390 nm (Fig. S1*D*) which could correspond to the protonated green form of the chromophore. This peak is also absent in the excitation spectrum monitored at 650 nm. The protonated green chromophore could be favored in this protein due to the S146C mutation, because the substitution of a Cys at this position would significantly increase the pKa of the phenol group of the chromophore, given the low H-bonding propensity of thiols. In line with this supposition, it is significant to note that the measured pKa of the red chromophore in mRouge is one pH unit higher than that of mRojoB (Table 2 and Fig. S6). Because the microenvironments of red and green chromophore species will be identical in these cases, it is safe to assume that the pKa of the green chromophore in mRouge will also be higher than that of mRojoB. This rationale implies that the assignment of the ∼390 nm peak for mRouge to the protonated green chromophore is correct.

Crystal structure of mRojoA. The crystal structure of mRojoA (PDB accession code 3NEZ) was solved at 1.70 Å resolution by molecular replacement using the structure of mCherry [PDB code: 2H5Q (4)] (Table S2). mRojoA crystallized in space group $P12_11$ with unit cell dimensions of $a = 61.2$ Å, $b = 97.4$ Å, and $c = 84.2$ Å. The asymmetric unit consisted of four molecules organized into two pairs of dimers that slant toward each other to form an A-frame-like tent structure. The tetrameric association of protomers in mRojoA is significantly different than the tight tetramer seen in DsRed and other class Anthozoa FPs (5–6). Superimposition of the backbone atoms of individual molecules of mRojoA with mCherry yielded an average rmsd of 0.21 ± 0.01 Å.

Treatment of the mRojoA chromophore was left until the end of refinement. Weak electron density for the phenolate group of the chromophore necessitated building the chromophore one atom at a time between refinement cycles. Building of the chromophore revealed a mixture of mature and immature species. The mature species is the acylimine-containing red chromophore, which is characterized by a *cis* peptide bond between F65 and M66. The immature species corresponds to the green chromophore, which is characterized by a *trans* peptide bond between F65 and M66. Occupancy refinement of the red and green forms was performed as described for DsRed (7). Each associated dimer in the structure was found to have one chromophore with ∼70%/30% red/green split and one chromophore with ∼50%/50% red/green split. The presence of the green chromophore indicates incomplete maturation. No evidence for the *trans* configuration of the chromophore phenolate ring was observed during refinement.

The H-bonding network around the chromophore of mRojoA is illustrated in Fig. S3*A*. The conserved H-bond between the chromophore imidazolinone O2 atom and catalytic R95 is present. The other catalytic residue, E215, makes a close H-bond with the chromophore imidazolinone N2 atom through $O_{\epsilon 1}$ with a similar configuration to what is seen in mCherry (4). The $O_{\epsilon 2}$ atom of E215 shows an H-bond to a conserved crystallographic water above the imidazolinone ring of the chromophore. The H-bond between S146 and the chromophore phenolate oxygen atom present in mCherry is also present in mRojoA. This H-bond has been suggested to stabilize the deprotonated phenolate form of the chromophore (8). Additionally, a water-mediated H-bond between residue T16 and the acylimine oxygen atom is observed (Fig. 2*A*). The interoxygen distances between the acylimine and water, and between water and T16 in monomer A were 2.5 Å and 2.9 Å, respectively. This water-mediated H-bond differs from the one observed in the crystal structure of Neptune (9). The H-bond in Neptune involves residue S28, analogous to G31 in mRojoA, and is located on the opposite side of the acylimine oxygen.

The structure of mRojoA shows the same type of chromophore π-stacking interaction observed in GFP-derived YFPs (10–12). In mRojoA, π-stacking occurs with Y197. The centroid to centroid distance from the chromophore phenolate to the Y197 side chain in mRojoA is 3.9 Å, and the angle between the normals of these planes is ∼8°. These values are comparable to those found in the crystal structure of the yellow fluorescent protein citrine [PDB code: 3DQO (10)] with a centroid to centroid distance of 3.6 Å and an angle of 6°.

Crystal structure of mRouge. The crystal structure of mRouge (PDB accession code 3NED) was solved at 0.95 Å resolution using direct methods (Table S2). mRouge crystallized in space group $P12_11$ with unit cell dimensions of $a = 48.9$ Å, $b = 42.9$ Å, and $c = 61.3$ Å. Unlike mRojoA, the structure of mRouge contains only one protein molecule per asymmetric unit. Superimposition of the protein with mCherry gives a backbone-atom rmsd of 0.10 Å.

For mRouge, in-depth treatment of the chromophore was again left until the end of the refinement procedure. Throughout most of the refinement only the imidazolinone heterocycle was fit due to very weak electron density corresponding to the Tyr-derived phenol ring of the chromophore (Fig. S3*B*). After refinement on the rest of the protein was completed, significant electron density in the difference map was observed within covalent bonding distance of the chromophore C1 atom. Prior studies on GFP mutants suggested that this density corresponds to hydroxylation at the C1 position (13–15). Moreover, a recent investigation of the maturation pathway for DsRed (the wild-type parent of mRouge) indicated that red chromophore maturation likely proceeds through an intermediate in which the acylimine bond preceding the chromophore is already oxidized and there is hydroxylation at the chromophore C1 position (16). Evidence for the presence of this putative intermediate in the mRouge crystal structure was strengthened by trypsin-digest mass spectrometry data, which showed a peptide fragment containing the chromo-

phore tripeptide at the appropriate mass (−4 Da from the uncyclized chromophore tripeptide) (Fig. S4*C*). Mass spectrometry analysis also demonstrated the presence of the green chromophore in mRouge (Fig. S4*B*), which had already been observed spectroscopically (Fig. S1*D*). Given the spectroscopic evidence for both red and green chromophores in mRouge along with the mass spectrometry evidence and crystallographic indication for a hydroxylated species, all three of these molecular entities were modeled into the crystal structure of mRouge. Occupancy refinement yielded 22% for the red chromophore, 45% for the green chromophore, and 33% for the hydroxylated species. No evidence for the *trans* configuration of the chromophore phenolate ring was observed during refinement in any of the difference maps.

For mRouge, many of the residues that form key interactions with the chromophore occur in multiple conformations (Fig. S3*B*). For example, C146 has two conformations occupied at 56% and 44%. In many red fluorescent proteins (RFPs) (including mCherry and mRojoA), a Ser residue at this position is H-bonded to the chromophore phenolate OH atom (4–5, 17). This interaction stabilizes the anionic chromophore, which is recognized as the dominant fluorescent species in FPs (5, 17–20). The major conformer of C146 points away from the chromophore. However, in the minor conformation, the sulfur atom points toward the chromophore in a similar conformation to what is observed for S146 in mCherry. Considering the low H-bonding strength of thiols, it is unlikely that this interaction represents a significant H-bond.

K70 has two conformations in mRouge (populated at 82% and 18%), with the major conformer pointing away from the bridging CB2 carbon of the chromophore and making an H-bond to T195. The minor conformer points toward the CB2 atom and forms a salt bridge with E148, as seen for K70 in mCherry (4).

Another position of interest with multiple conformations in mRouge is the catalytic residue E215 (populated at 75% and 25%). E215 is speculated to be deprotonated in DsRed (21), but to be protonated in mCherry because its $O_{\epsilon 1}$ atom comes within H-bonding proximity to the chromophore N2 atom in mCherry (4). The major conformation of E215 in mRouge is most similar to the conformation of that same residue in mCherry (Fig. S3*B*), with its $O_{\epsilon 1}$ oxygen forming an H-bond with the N2 atom of the chromophore. The presence of this H-bond implies that the major conformer of E215 in mRouge is protonated, analogous to what is seen in mCherry.

It is worth noting that the minor conformer of E215 in mRouge (Fig. S3*B*) is not in the same conformation as that seen for E215 in DsRed. In DsRed, E215 is oriented in such a way as to create a salt bridge with K70 across one face of the chromophore ring system (5, 7). The minor conformer of the E215 residue in mRouge attains a previously unobserved conformation for DsRed variants. In this conformation, the $O_{\epsilon 1}$ atom is within H-bonding distance of both the chromophore N2 atom and the hydroxyl group attached to the chromophore C1 atom (2.4 Å and 2.6 Å, respectively). This conformer of E215 mainly differs from the protonated, major conformer in mRouge by a ∼90° twist about the side chain χ_3 angle.

A water-mediated H-bond between residue T16 and the acylimine oxygen atom is also present in the structure of mRouge (Fig. 2*B*). This H-bond is nearly identical to the one observed in the structure of mRojoA, with interoxygen distances between the acylimine and water, and between water and T16 of 3.0 Å and 2.7 Å, respectively.

A sequence and structural motif consisting of T195-T197-N217 occurs in mRouge. This set of mutations leads to an 8 nm red shift in the emission spectra of mCherry mutants (Table 2). In mRouge, the $O_{\epsilon 2}$ oxygen of the major conformer of E215 is H-bonded to the terminal nitrogen on the side chain amide of N217 (Fig. S5). Moreover, the side chain amide oxygen from N217 forms an H-bond with the hydroxyl group of T197. Finally, the side chain of T197 forms a water-mediated H-bond with the side chain of the major conformer of T195. The means by which this set of interactions causes red-shifted emission is unclear. However, the network of H-bonds resulting from this set of mutations may explain the synergistic effects on emission wavelength encountered by including them in the mCherry scaffold.

SI Materials and Methods. ***Materials.*** All reagents used were of the highest available purity. Restriction enzymes and DNA-modifying enzymes were from New England Biolabs. Synthetic oligonucleotides were obtained from Integrated DNA Technologies, and Ni-NTA agarose resin was obtained from Qiagen. CelLytic B buffer and lysozyme were purchased from Sigma-Aldrich. All aqueous solutions were prepared using water purified with a Millipore BioCell system.

Computational design. Hydrogens were added to the crystal structure of mCherry (PDB code: 2H5Q) using Molprobity (22). Following removal of all water molecules and ions, any strain or steric clashes in the structure were removed by performing 50 steps of conjugate gradient energy minimization (23). Partial atomic charges for the chromophore were parameterized using the charges described by Sitkoff et al. (24). Computational design was performed using the PHOENIX protein design software. The energy function used was based on the DREIDING force field (23) and included a scaled van der Waals term (25), hydrogen bonding and electrostatic terms (26), and terms for implicit solvation and phi-psi propensities. Implicit solvation energies were evaluated using a model based on occluded volume (27) described below. Amino acid phi-psi propensities were derived and applied following the method of Shortle (28). Sequence optimization was carried out with FASTER; a Monte Carlo-based algorithm was then used to sample sequences around the minimum energy configurations identified by FASTER (29) and generate a list of high-scoring sequences. This design procedure was followed by a computational library design step in which combinatorial sequence libraries were defined as described by Allen et al. (30). Briefly, based on the list of scored sequences generated by computational protein design (CPD) and the list of required amino acids specified by the user, the algorithm determines the library composition that represents the best set of top-scoring sequences that can be encoded by a single degenerate codon at each position for a desired range of library sizes.

Implicit solvation energies. To account for the contributions of solvent to the free energies of folding, we applied an implicit solvation potential inspired by the methods of Dahiyat and Mayo (25) and Lazaridis and Karplus (27). The potential is intended to reward the burial of nonpolar groups, penalize the burial of polar groups, and penalize the exposure of nonpolar groups. In this scheme, atomic groups are scored based on how their volumes are occluded by other groups in the protein structure. We used a Gaussian function to compute the occlusion O_i of atom i by other atoms j:

$$O_i = \sum_j V_j e^{-\frac{R_{ij}^2}{2\lambda^2}},$$

where V_j is the volume of atom j, R_{ij} is the distance between i and j, and λ is the constant correlation length. We define the fractional exposure Θ_i for atom i as:

$$\Theta_i = \frac{O_{i,\max} - O_i}{O_{i,\max} - O_{i,\min}},$$

where the O_i minima and maxima are found from a database survey of protein crystal structures. When $O_i = O_{i,\max}$, atom i

is fully buried, and $\Theta_i = 0$; when $O_i = O_{i,\min}$, atom i is fully exposed, and $\Theta_i = 1$.

The nonpolar solvation energy of atom i is given by:

$$E_{np,i} = s_i\sigma_{np}(\kappa_{np} + 1)\Theta_i - s_i\sigma_{np},$$

where s_i is a scaling parameter specific to the amino acid type and atom type for i, σ_{np} is the nonpolar desolvation energy benefit, and κ_{np} is the nonpolar exposure scale factor. This formula allows the balance between favorable nonpolar desolvation and unfavourable nonpolar exposure to be tuned between $s_i\kappa_{np}\sigma_{np}$ energy units when i is fully exposed, and $-s_i\sigma_{np}$ energy units when fully buried. For each nonpolar atom i, s_i is the mode of exposed surface areas observed for the residue and atom type for i in truncated tripeptides with coordinates taken from proteins in a structural database previously used to generate conformer libraries (31). In this way, nonpolar atoms with greater exposed surface area in the truncated tripeptide unfolded state model receive larger magnitude energies than those with smaller unfolded state surface areas with the same fractional exposure value.

The formula for polar solvation is analogous to that for nonpolar solvation, but with opposite signs, indicating that the preference for exposure versus burial is reversed:

$$E_{p,i} = -s_i\sigma_p(\kappa_p + 1)\Theta_i + s_i\sigma_p.$$

The s_i parameters for polar groups were derived via a linear fit between the Θ_i values and Poisson-Boltzmann reaction field energies calculated for side chains in a computationally tractable subset of the structural database.

Mutagenesis. The mCherry and mRaspberry genes were PCR-amplified from plasmids mCherry-pBAD and mRaspberry-pBAD (provided by R.Y. Tsien, UCSD) and subcloned into pET11-a (Novagen) via *Nde*I/*Bam*HI. The plasmids were then transformed into *Escherichia coli* XL-1 Blue and BL21(DE3) cells. The entire *Nde*I/*Bam*HI fragments, including the whole coding region, were verified by DNA sequencing. All mutations were introduced into the mCherry gene by overlap extension mutagenesis (32) using VentR DNA polymerase. Briefly, external primers were used in combination with sets of complementary pairs of degenerate oligonucleotides containing the desired mutations in individual PCR reactions. The resulting overlapping fragments were gel-purified (Qiagen) and recombined by overlap extension PCR. The resulting amplicons were digested with *Nde*I/*Bam*HI, gel-purified, and ligated into pET11a expression vector with T4 ligase. Library composition was verified by sequencing 96 clones per library (Agencourt Biosciences).

Preparation of libraries for screening. The DNA libraries prepared as described above were transformed into chemically competent *Escherichia coli* BL21-Gold(DE3) cells (Stratagene). Colonies were picked into individual wells of Nunc V96 MicroWell polypropylene plates containing 200 μL of medium (LB with 100 μg/mL ampicillin supplemented with 10% glycerol). The plates were covered with a sterile Breathe-Easy gas permeable sealing membrane (Sigma) and incubated overnight at 37 °C with shaking. After incubation, these mother plates were used to inoculate sterile Nunc V96 MicroWell polypropylene plates ("daughter" plates) containing 300 μL of Overnight Express Instant Terrific Broth media (Novagen) supplemented with ampicillin per well. Daughter plates were sealed with breathable membranes and incubated overnight (37 °C, 250 rpm shaking). After incubation, the cells were harvested by centrifugation and the cell pellets were washed twice with PBS (pH 7.4). Washed cell pellets were then incubated at 4 °C for 72 h to allow chromophore maturation. These pellets were resuspended in PBS and transferred to a Fluotrac 96-well plate (Greiner Bio-One) for screening.

Protein expression and purification for spectral characterization. Mother plates containing the mutant libraries were used to inoculate 24-well culture plates (Whatman) containing 5 mL Overnight Express Instant TB media (Novagen) supplemented with ampicillin in each well. The 24-well plates were sealed with sterile Bugstopper venting capmats (Whatman) and incubated at 37 °C overnight with shaking. After expression, cells were harvested by centrifugation and washed twice with PBS. After maturation at 4 °C for one week, the cell pellets were resuspended in 400 μL lysis buffer [50 mM sodium phosphate buffer, pH8.0, 300 mM NaCl, 2.5 mM imidazole, 1X CelLyticB, 1 mg/mL lysozyme, and 25 U/mL benzonase nuclease (Novagen)] and incubated at 30 °C for 30 min with shaking. After centrifugation, clarified lysates were recovered and proteins were purified by affinity chromatography using His-Select plates (Sigma) according to the manufacturer's protocol.

Protein expression and purification for crystallization. Protein was expressed in 1.0 L cultures by transformation of a pET11-a vector containing the gene of interest into *E. coli* BL21-Gold(DE3) and purified by Ni-NTA affinity chromatography according to the manufacturer's protocol. Column elutions were desalted by gel filtration using a Superdex 75 10/300 GL Tricorn resin column (GE Healthcare) into a final buffer solution of 50 mM phosphate buffer, pH 7.5, and 150 mM NaCl.

Spectroscopic characterization. Proteins purified as described above were quantified using the alkali denaturation method (33). Briefly, RFPs were alkali-denatured with an equal volume of 2 M NaOH. It is known that the alkali-denatured RFP chromophore converts to a GFP-like one, with extinction coefficient 44,000 M^{-1} cm^1 at 452 nm under these conditions. Absorbance, emission, and excitation spectra were recorded in PBS with a Tecan Safire2 plate reader. Path lengths for each well were calculated ratiometrically using the difference in absorbance of PBS at 900 nm and 998 nm. Based on the absorbance spectra of native proteins and the concentration determination of alkali-denatured proteins, molar extinction coefficients were calculated. For determination of quantum yields, the integrated fluorescence intensity of mutants of interest was compared with that of equally absorbing samples of mCherry and mRaspberry (quantum yields 0.22 and 0.15, respectively) with excitation at 550 nm.

pKa measurements. pH titrations were performed using a range of buffers from pH 2 to 9. Proteins were diluted into these buffers to a concentration of 5–10 μM. Fluorescence scans were taken at each pH value using a Tecan Safire2 plate reader. The Henderson-Hasselbach equation was used to calculate the pKa for each protein (Fig. S6).

Growth of crystals and screening. Light blue crystals of mRouge were grown in 1 μL × 1 μL hanging drops with a precipitant solution of 200 mM ammonium acetate, 100 mM Bis-Tris, pH 6.5, and 25% (w/v) polyethylene glycol 3350. Large multicrystalline chunks (0.5 mm × 2.0 mm × 0.25 mm) were prodded with a nylon loop to break off smaller shards for isolation and collection of diffraction data. All the pieces screened were crystallographically identical in terms of space group and unit cell dimensions. Dark purple crystals of mRojoA were grown in hanging drops with 1 μL protein solution and 1 μL of the same precipitant used to crystallize mRouge. These long stick-like rhomboidal crystals of mRojoA had approximate dimensions of 1.0 mm× 0.02 mm × 0.02 mm.

Data collection and processing. For mRouge, a dataset was collected locally at 2.0 Å resolution, and another dataset resolved at 0.95 Å was collected at the Stanford Synchrotron Radiation Lightsource (SSRL) beamline 12-2. The low-resolution dataset was integrated, merged, and scaled with IMOSFLM (34) and SCALA (35). For the high-resolution dataset, IPMOSFLM (34) was used for integration and ACORN was used for merging and scaling. For mRojoA, crystals were sent to the SSRL where the dataset was collected at 1.70 Å. This dataset was integrated, merged, and scaled with IMOSFLM and SCALA.

Solution and refinement of crystal structures. The 2.0 Å dataset of mRouge, and the 1.70 Å dataset of mRojoA were solved by molecular replacement using PHASERMR (36). For mRouge and mRojoA, the search model used consisted of the PDB coordinates from mCherry [2H5Q (4)] with the chromophore removed. After the higher resolution dataset of mRouge was solved by direct methods, the initial coordinates from the 2.0 Å structure were used as a starting point for further refinement of the subatomic resolution structure.

Refinement was accomplished using REFMAC5 (37–38) and PHENIX (Python-based Hierarchical ENvironment for Integrated Xtallography) (39). PHENIX was used specifically for refinement of atomic occupancies. Model building was done with COOT (40), wherein water molecules were added manually when they were within H-bonding distance of other heteroatoms (2.3–3.5 Å) and had peaks in the F_o-F_c map of greater than 3.5σ. In addition, water molecules were removed when they had equivalent isotropic B-factors greater than 60–80 $Å^2$. During generation of R-factors, 5% of data was excluded for cross-validation with an R_{free} value. Crystallographic R-factors were calculated in the standard fashion ($R = \sum |F_{obs}\text{-}F_{calc}| / \sum F_{obs}$).

In the crystal structure of mRouge, refinement was done with anisotropic atomic displacement parameters (ADPs) after the initial few cycles of refinement. In the crystal structure of mRojoA, the final refinement steps were carried out with 20 translation-libration-screw (TLS) groups per protein molecule (41). TLS groups were identified automatically by using the TLS-Motion Determination web server (TLS MD) (42). Riding hydrogens were included in the refinement of all structures for nonwater molecules, but were only retained in the final structure of mRouge due to its high resolution.

The library file for the chromophore was built based on the CH6 chromophore deposited in the Hetero-compound Information Center—Uppsala (HIC-Up) online database. Appropriate constraints and atom types were added and/or edited to account for the alternative covalencies observed for different conformations of the chromophore.

Mass spectrometry analyses. Following separation by SDS-PAGE, the ~25 kDa band from a freshly purified sample of mRouge was excised and destained. Destaining of the Coomassie dye was accomplished by a 100 μL wash of 50 mM ammonium bicarbonate followed by a 50 μL wash of a 1:1 mixture of 50 mM ammonium bicarbonate and acetonitrile; this process was repeated for a total of three times. After destaining, the gel band sample was reduced with 25 μL of 50 mM ammonium bicarbonate plus 50 μL of freshly prepared 10 mM DTT in 100 mM ammonium bicarbonate for 30 min. at 50 °C. The sample was then alkylated in the absence of light with 25 μL of 50 mM ammonium bicarbonate plus 50 μL of freshly prepared 55 mM iodoacetate in 100 mM ammonium bicarbonate for 20 min. at room temperature. Following additional washes with 100 μL of 50 mM ammonium bicarbonate and 100 μL of acetonitrile, the gel band sample was digested overnight at 37 °C with 75 μL of 50 mM ammonium bicarbonate plus 25 μL of 6 ng/μL sequencing grade porcine trypsin (Promega).

After digestion, the supernatant from the gel band sample was collected. The gel band was washed three times: once with 100 μL of 1% formic acid/2% acetonitrile in water, once with 100 μL of a 1:1 acetonitrile and water mixture, and once with 100 μL of 1% formic acid in acetonitrile. The pooled supernatant and wash solutions were then vacuum-dried overnight and resuspended in 0.1% formic acid in preparation for collection of mass spectrometry data. Samples of this nature were prepared in triplicate from the same freshly expressed and purified sample of mRouge.

These tryptic-digest samples were desalted on a 150 μm× 3 cm C18AQ precolumn (Magic 5 μm, Michrom). After desalting, separation of peptides was performed with a CapLC XE HPLC system (Waters) using a 5% to 35% acetonitrile gradient in 0.2% formic acid on a 100 μm × 15 cm column packed with the same resin as the precolumn. The flow rate during separation was 0.35 μL/ min and the HPLC column was connected directly to the mass spectrometer used for MS/MS analysis. Tandem mass spectra were acquired in data-dependent acquisition mode on a hybrid LTQ FT-ICR Ultra mass spectrometer (Thermo Fisher Scientific) with a nanoelectrospray ion source. Full scan mass spectra (400–1,600 m/z) were acquired after accumulating 500,000 ions (with a resolution of 50,000 at 400 m/z). The seven most intense ions from the full scans were trapped in the linear ion trap and fragmented by collision induced dissociation (CID) after accumulating 5,000 ions (collisional energy: 35%, isolation width: 3 Da). Ion charge state screening was employed for singly and multiply charged ions. A dynamic exclusion list was set (maximum retention time: 60 s, relative mass window: 10 ppm) and early expiration was permitted.

1. Tsien RY (1998) The green fluorescent protein. *Annu Rev Biochem* 67:509–544.
2. Verkhusha VV, et al. (2001) An enhanced mutant of red fluorescent protein DsRed for double labeling and developmental timer of neural fiber bundle formation. *J Biol Chem* 276:29621–29624.
3. Campbell RE, et al. (2002) A monomeric red fluorescent protein. *Proc Natl Acad Sci USA* 99:7877–7882.
4. Shu X, Shaner NC, Yarbrough CA, Tsien RY, Remington SJ (2006) Novel chromophores and buried charges control color in mFruits. *Biochemistry* 45:9639–9647.
5. Wall MA, Socolich M, Ranganathan R (2000) The structural basis for red fluorescence in the tetrameric GFP homolog DsRed. *Nat Struct Biol* 7:1133–1138.
6. Petersen J, et al. (2003) The 2.0 Å crystal structure of eqFP611, a far-red fluorescent protein from the sea anemone Entacmaea quadricolor. *Journal of Biological Chemistry* 278:44626–44631.
7. Tubbs JL, Tainer JA, Getzoff ED (2005) Crystallographic structures of Discosoma red fluorescent protein with immature and mature chromophores: linking peptide bond trans-cis isomerization and acylimine formation in chromophore maturation. *Biochemistry* 44:9833–9840.
8. Gurskaya NG, et al. (2001) GFP-like chromoproteins as a source of far-red fluorescent proteins. *FEBS Lett* 507:16–20.
9. Lin MZ, et al. (2009) Autofluorescent proteins with excitation in the optical window for intravital imaging in mammals. *Chem Biol* 16:1169–1179.
10. Barstow B, Ando N, Kim CU, Gruner SM (2008) Alteration of citrine structure by hydrostatic pressure explains the accompanying spectral shift. *Proc Natl Acad Sci USA* 105:13362–13366.
11. Wachter RM, Elsliger MA, Kallio K, Hanson GT, Remington SJ (1998) Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein. *Structure* 6:1267–1277.
12. Griesbeck O, Baird GS, Campbell RE, Zacharias DA, Tsien RY (2001) Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. *J Biol Chem* 276:29188–29194.
13. Barondeau DP, Putnam CD, Kassmann CJ, Tainer JA, Getzoff ED (2003) Mechanism and energetics of green fluorescent protein chromophore synthesis revealed by trapped intermediate structures. *Proc Natl Acad Sci USA* 100:12111–12116.
14. Barondeau DP, Kassmann CJ, Tainer JA, Getzoff ED (2005) Understanding GFP chromophore biosynthesis: Controlling backbone cyclization and modifying posttranslational chemistry. *Biochemistry* 44:1960–1970.
15. Rosenow MA, Huffman HA, Phail ME, Wachter RM (2004) The crystal structure of the Y66L variant of green fluorescent protein supports a cyclization-oxidation-dehydration mechanism for chromophore maturation. *Biochemistry* 43:4464–4472.
16. Strack RL, Strongin DE, Mets L, Glick BS, Keenan RJ (2010) Chromophore formation in DsRed occurs by a branched pathway. *J Am Chem Soc* 132:8496–8505.
17. Wilmann PG, et al. (2005) The 2.1 Å crystal structure of the far-red fluorescent protein HcRed: Inherent conformational flexibility of the chromophore. *J Mol Biol* 349:223–237.

18. Pletnev S, et al. (2008) A crystallographic study of bright far-red fluorescent protein mKate reveals pH-induced cis-trans isomerization of the chromophore. *J Biol Chem* 283:28980–28987.
19. Quillin ML, et al. (2005) Kindling fluorescent protein from Anemonia sulcata: Dark-state structure at 1.38 Å resolution. *Biochemistry* 44:5774–5787.
20. Battad J, et al. (2007) A structural basis for the pH-dependent increase in fluorescence efficiency of chromoproteins. *J Mol Biol* 368:998–1010.
21. Taguchi N, et al. (2009) Fragment molecular orbital calculations on red fluorescent proteins (DsRed and mFruits). *Journal of Physical Chemistry B* 113:1153–1161.
22. Davis IW, et al. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35:W375–383.
23. Mayo SL, Olafson BD, Goddard WA (1990) DREIDING—A generic force field for molecular simulations. *J Phys Chem* 94:8897–8909.
24. Sitkoff D, Sharp KA, Honig B (1994) Accurate calculation of hydration free energies using macroscopic solvent models. *J Phys Chem* 98:1978–1988.
25. Dahiyat BI, Mayo SL (1997) Probing the role of packing specificity in protein design. *Proc Natl Acad Sci USA* 94:10172–10177.
26. Dahiyat BI, Gordon DB, Mayo SL (1997) Automated design of the surface positions of protein helices. *Protein Sci* 6:1333–1337.
27. Lazaridis T, Karplus M (1999) Effective energy function for proteins in solution. *Proteins* 35:133–152.
28. Shortle D (2003) Propensities, probabilities, and the Boltzmann hypothesis. *Protein Sci* 12:1298–1302.
29. Allen BD, Mayo SL (2006) Dramatic performance enhancements for the FASTER optimization algorithm. *J Comput Chem* 27:1071–1075.
30. Allen BD, Nisthal A, Mayo SL (2010) Experimental library screening demonstrates the successful application of computational protein design to large structural ensembles.*Proc Natl Acad Sci USA*, in press.
31. Lassila JK, Privett HK, Allen BD, Mayo SL (2006) Combinatorial methods for small-molecule placement in computational enzyme design. *Proc Natl Acad Sci USA* 103:16710–16715.
32. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 77:51–59.
33. Gross LA, Baird GS, Hoffman RC, Baldridge KK, Tsien RY (2000) The structure of the chromophore within DsRed, a red fluorescent protein from coral. *Proc Natl Acad Sci U S A* 97:11990–11995.
34. Leslie AGW (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 and ESF-EAMCB Newsletter on Protein Crystallography*, http://www.mail-archive.com/ccp4bb@jiscmail.ac.uk/msg12057.html.
35. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallographica Section D* 62:72–82.
36. Mccoy A, et al. (2007) Phaser crystallographic software. *Journal of Applied Crystallography* 40:658–674.
37. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D* 53:240–255.
38. Vagin AA, et al. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallographica Section D* 60:2184–2195.
39. Afonine PV, Grosse-Kunstleve RW, Adams PD (2005) The Phenix refinement framework. *CCP4 Newsletter*, http://www.ccp4.ac.uk/newsletter.php#back_issues.
40. Emsley P, Cowtan K (2004) COOT: Model-building tools for molecular graphics. *Acta Crystallographica Section D* 60:2126–2132.
41. Painter J, Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS-motion. *Acta Crystallographica Section D* 62:439–450.
42. Painter J, Merritt EA (2006) TLSMD web server for the generation of multigroup TLS models. *Journal of Applied Crystallography* 39:109–111.

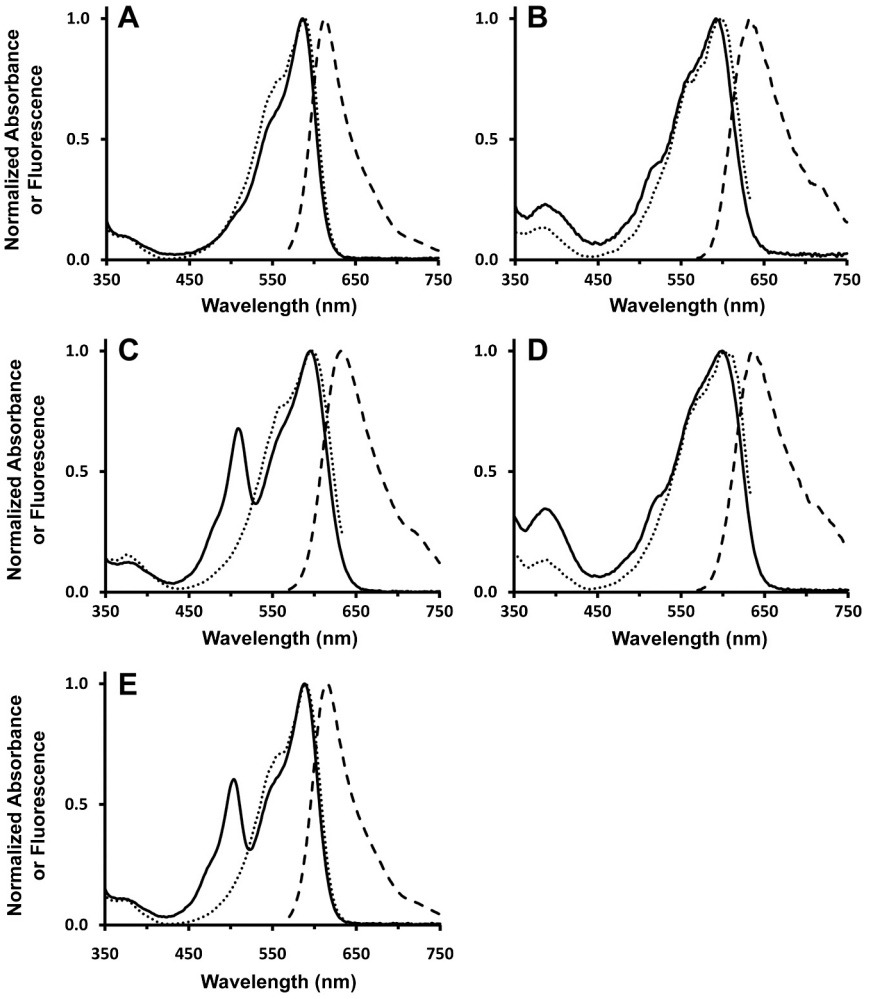

Fig. S1. Fluorescence and absorption spectra of mCherry (*A*), mRojoA (*B*), mRojoB (*C*), mRouge (*D*), and mCherry Q163M (*E*). The absorption spectra are shown with a solid line, the excitation spectra are shown with a dotted line, and the emission spectra are shown with a dashed line.

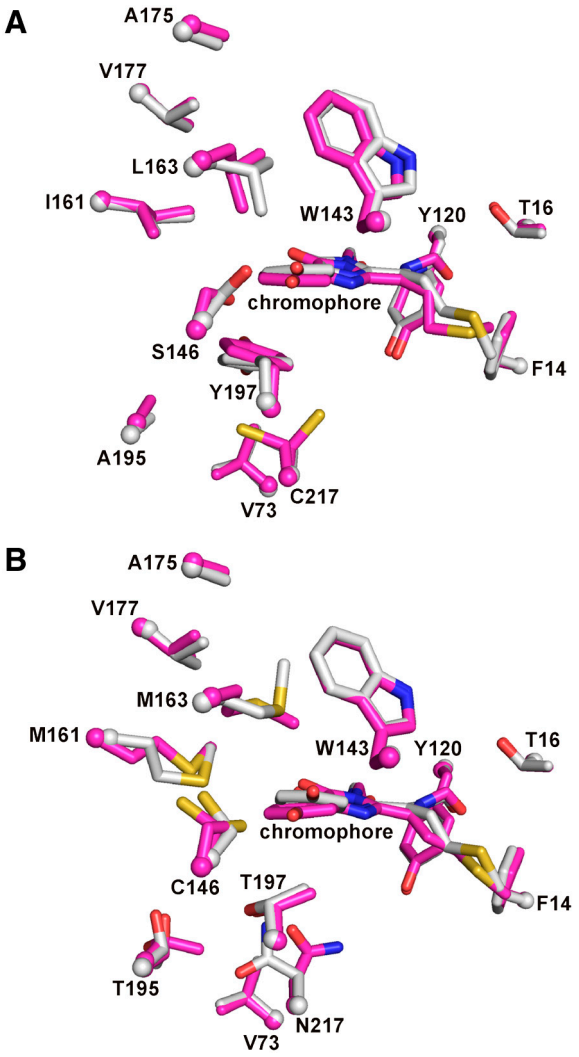

Fig. S2. Comparison of structure predicted by CPD (white) and crystal structure (magenta) for red-shifted mutants mRojoA (*A*) and mRouge (*B*). Residue C217 in mRojoA (*A*) and residues C146 and T195 in mRouge (*B*) exhibit two conformations in the crystal structure. The chromophore was not designed.

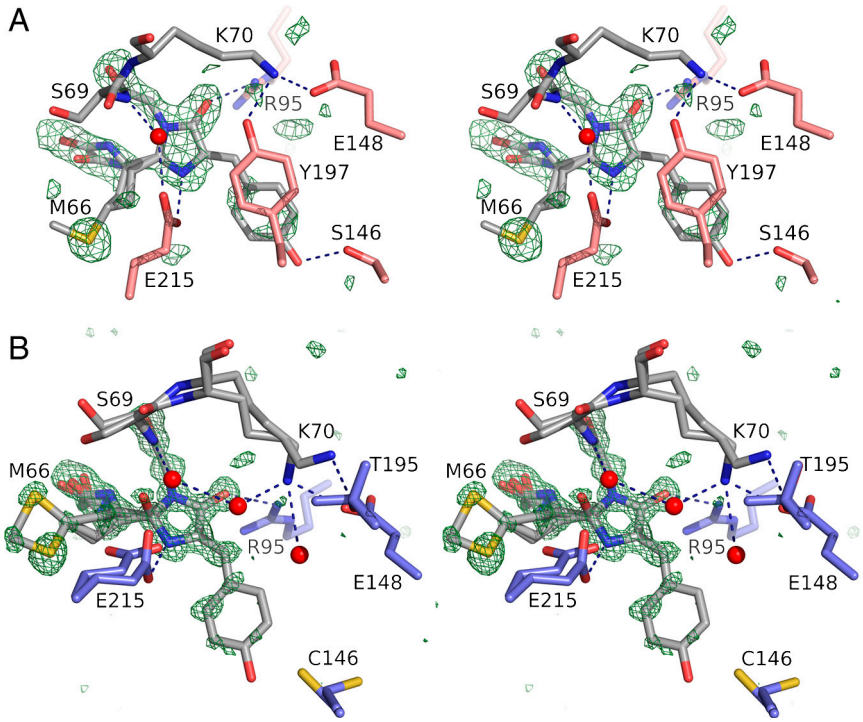

Fig. S3. Chromophore environment of mRojoA (*A*) and mRouge (*B*). H-bonds are indicated by dashed lines and crystallographic waters are indicated by red spheres. Omit maps (green) for the chromophores were contoured at 3.0σ for mRojoA and at 3.5σ for mRouge. Noncontiguous residues are colored pink (mRojoA) and purple (mRouge). The chromophore and its contiguous residues are colored gray. Note that the O_γ atom in both conformations of T195 for mRouge (*B*) is hidden behind the C_β atoms of that same residue. The refinement model for the hydroxylated conformation of mRouge did not include atoms from the phenolate ring of the chromophore due to a lack of electron density.

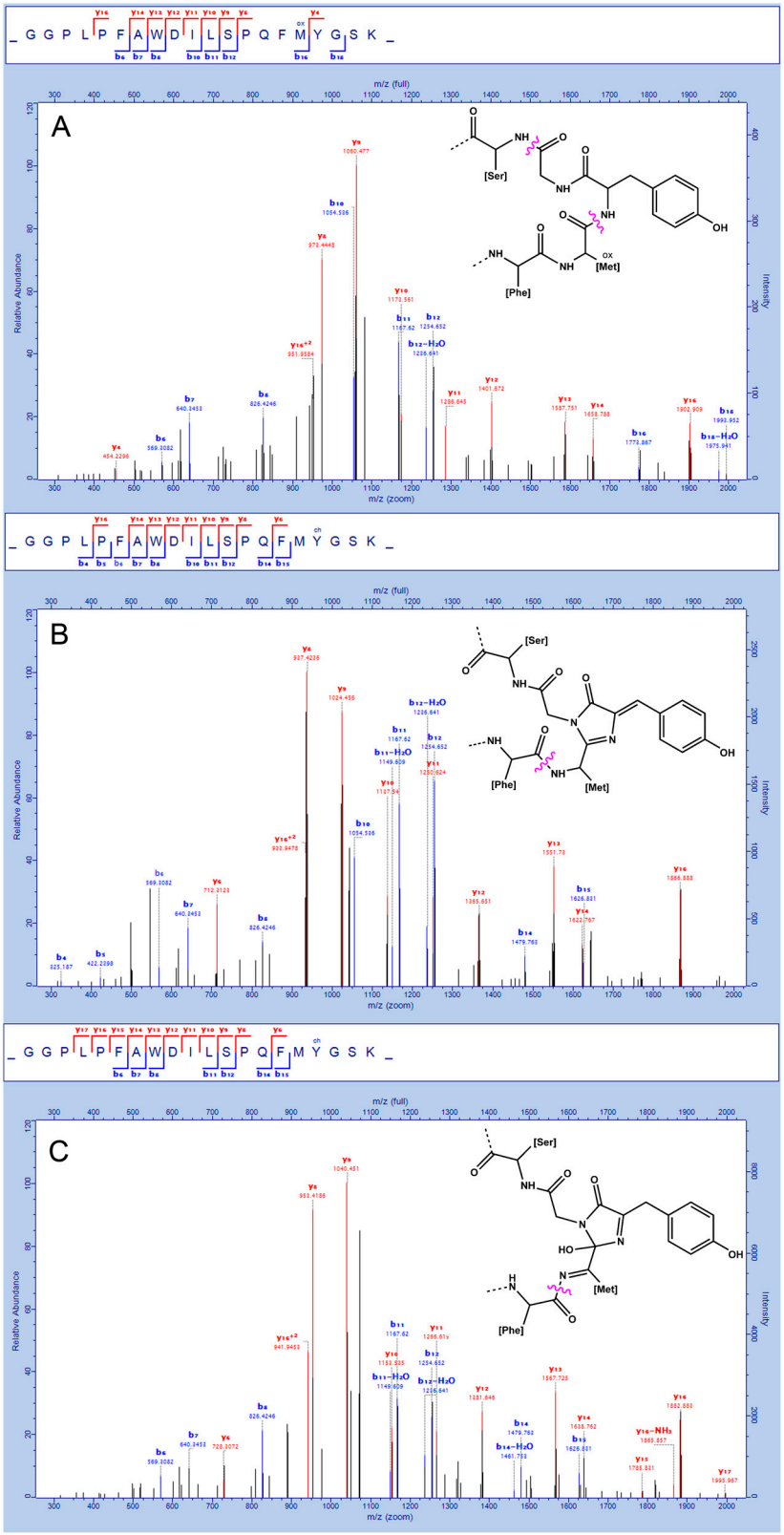

Fig. S4. Trypsin-digest mass spectrometry confirms the presence of a hydroxylated chromophore species in mRouge. Peptide fragments isolated by tandem mass spectrometry from an in gel trypsin-digest of mRouge indicate the presence of the uncyclized chromophore (*A*), the green chromophore species (*B*), and a hydroxylated species consistent with current literature on the maturation of DsRed (*C*). Note that in the uncyclized chromophore (*A*), the side chain sulfur atom of M66 was spontaneously oxidized to a sulfoxide during the process of data acquisition.

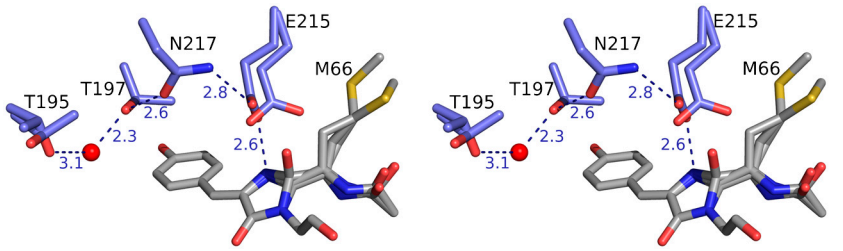

Fig. S5. H-bond network in the red-shifting TTN (T195-T197-N217) motif of mRouge. H-bonds are indicated by dashed lines. Numbers in purple indicate distances in Å.

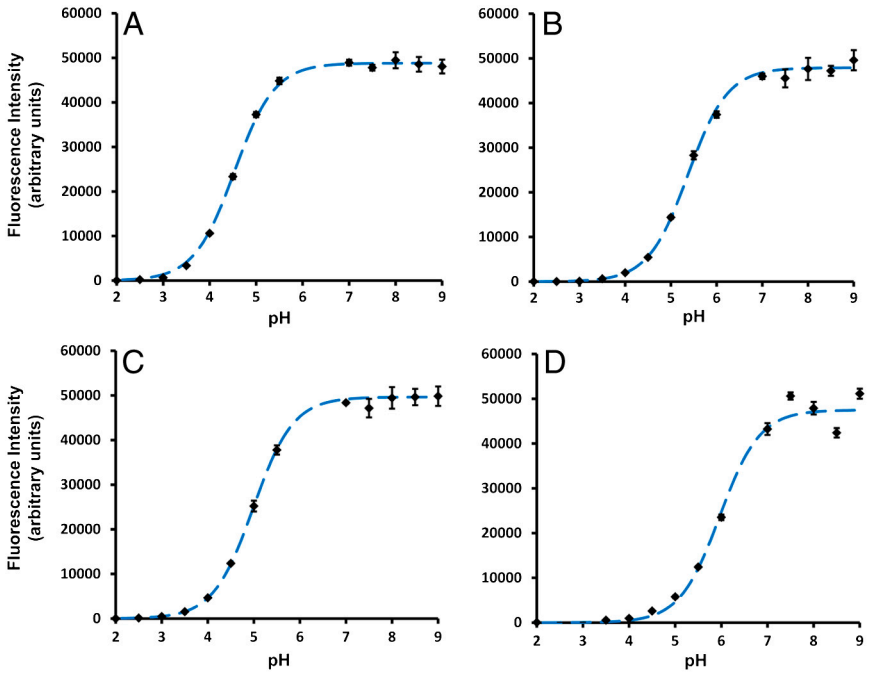

Fig. S6. pH profiles for mCherry (*A*), mRojoA (*B*), mRojoB (*C*), and mRouge (*D*).

Table S1. Sequences and emission wavelengths of the most red-shifted mutants found in each library

	Design hypothesis I			Design hypothesis II				Design hypothesis III				
Protein	14	16	120	143	146	161	163	73	195	197	217	λ_{em} (nm)
mCherry	F	V	Y	W	S	I	Q	V	V	I	A	611
Library 1												
1.13B9	-	T	-	-	-	-	-	-	-	-	-	615
1.10B1	Y	T	-	-	-	-	-	-	-	-	-	616
1.8D2	T	I	-	-	-	-	-	-	-	-	-	615
1.9H8	T	T	-	-	-	-	-	-	-	-	-	616
1.1B10	T	S	-	-	-	-	-	-	-	-	-	617
1.9F2	I	T	-	-	-	-	-	-	-	-	-	617
1.5G10	I	S	-	-	-	-	-	-	-	-	-	617
Library 2												
2.14H12	-	-	-	L	C	-	L	-	-	-	-	621
2.14C2	-	-	-	L	C	-	M	-	-	-	-	620
2.10E2	-	-	-	L	C	M	L	-	-	-	-	620
2.1H7	-	-	-	-	C	-	M	-	-	-	-	622
2.7A11	-	-	-	V	C	-	M	-	-	-	-	620
2.15E2	-	-	-	V	C	M	L	-	-	-	-	621
2.13B11	-	-	-	S	C	M	L	-	-	-	-	621
2.18G3	-	-	-	C	C	M	M	-	-	-	-	621
2.14D10	-	-	-	A	C	M	L	-	-	-	-	621
2.2E11	-	-	-	T	C	M	M	-	-	-	-	621
2.5H4	-	-	-	-	C	M	M	-	-	-	-	622
2.8F1	-	-	-	-	T	-	M	-	-	-	-	624
2.18C7	-	-	-	V	C	M	M	-	-	-	-	621
2.15H12	-	-	-	I	C	M	L	-	-	-	-	621
2.15G9	-	-	-	I	C	M	M	-	-	-	-	622
2.16H8	-	-	-	A	A	M	M	-	-	-	-	620
Library 3-1												
31.5H10	-	-	-	-	-	-	-	-	-	Y	C	621
31.7F9	-	-	-	-	-	-	-	-	-	Y	S	620
31.2G4	-	-	-	-	-	-	-	-	T	Y	C	621
31.3G3	-	-	-	-	-	-	-	I	-	Y	C	620
31.2D8	-	-	-	-	-	-	-	-	T	Y	S	620
31.4D8	-	-	-	-	-	-	-	-	A	Y	C	621
3-1.7H9	-	-	-	-	-	-	-	-	-	Y	T	620
3-1.7H1	-	-	-	-	-	-	-	I	A	Y	C	621
3-1.8D4	-	-	-	-	-	-	-	I	-	Y	T	620
Library 3-2												
3-2.8C9	-	-	-	-	-	-	-	-	T	T	N	619

Table S2. Crystallographic Data

	mRouge	mRojoA
Resolution, Å	56.7–0.95	38.0–1.70
Space group	$P12_11$	$P12_11$
Observations	568,101	234,983
Unique observations	135,356	102,737
R_{merge}, %	12.1 (49.6)	4.2 (39.3)
Completeness, %	91.9 (62.9)	95.4 (97.0)
$I/\sigma I$	9.7 (1.7)	11.9 (2.2)
R_{work}/R_{free}, %	13.3/15.6	18.2/22.0

Values in parenthesis are statistics for the highest resolution shell of data.