3-28-2013

Pressure Induced Structural Changes and Gas Diffusion Pathways in Monomeric Fluorescent Proteins

Yuba R. Bhandari
Florida International University, bhandariyuba@gmail.com

DOI: 10.25148/etd.FI13042320

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

PRESSURE INDUCED STRUCTURAL CHANGES AND GAS DIFFUSION PATHWAYS IN MONOMERIC FLUORESCENT PROTEINS

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PHYSICS

by

Yuba R. Bhandari

2013

To: Dean Kenneth G. Furton
College of Arts and Sciences

This dissertation, written by Yuba R. Bhandari, and entitled Pressure Induced Structural Changes and Gas Diffusion Pathways in Monomeric Fluorescent Proteins, having been approved in respect to style and intellectual content, is referred to you for judgement.

We have read this dissertation and recommend that it be approved.

_____
David C. Chatfield

_____
Xuewen Wang

_____
Bernard S. Gerstman, Co-Major Professor

_____
Prem P. Chapagain, Co-Major Professor

Date of Defense: March 28, 2013

The dissertation of Yuba R. Bhandari is approved.

_____
Dean Kenneth G. Furton
College of Arts and Sciences

_____
Dean Lakshmi N. Reddi
University Graduate School

Florida International University, 2013

DEDICATION

I dedicate this dissertation to my beloved parents.

ACKNOWLEDGMENTS

I would like to thank Prof. David Chatfield and Prof. Xuewen Wang for their helpful suggestions and comments. I would like to offer special thanks to my co-major advisor Prof. Bernard Gerstman for his time and care for designing and refining my dissertation, and guiding me through my research. I would like to express great appreciation to my co-major advisor Prof. Prem Chapagain for his invaluable guidance, help and support during my entire research work and dissertation writing.

I would like to thank all members of my biophysics research group for their help and cooperation. I shall always cherish our “wonderul Friday meeting” that would keep running for hours and hours with stimulating discussions and full enthusiasm. I would also like to thank my beloved wife Bandana for her care, support and understanding during the writing of this dissertation.

ABSTRACT OF THE DISSERTATION

PRESSURE INDUCED STRUCTURAL CHANGES AND GAS DIFFUSION PATHWAYS IN MONOMERIC FLUORESCENT PROTEINS

by

Yuba R. Bhandari

Florida International University, 2013

Miami, Florida

Professor Bernard S. Gerstman, Co-Major Professor

Professor Prem P. Chapagain, Co-Major Professor

Fluorescent proteins (FPs) are extremely valuable biochemical markers which have found a wide range of applications in cellular and molecular biology research. The monomeric variants of red fluorescent proteins (RFPs), known as mFruits, have been especially valuable for *in vivo* applications in mammalian cell imaging. Fluorescent proteins consist of a chromophore caged in the beta-barrel protein scaffold. The photophysical properties of an FP is determined by its chromophore structure and its interactions with the protein barrel.

Application of hydrostatic pressure on FPs results in the modification of the chromophore environment which allows a systematic study of the role of the protein-chromophore interactions on photophysical properties of FPs. Using Molecular Dynamics (MD) computer simulations, I investigated the pressure induced structural changes in the monomeric variants mCherry, mStrawberry, and Citrine. The results explain the molecular basis for experimentally observed pressure responses among FP variants. It is found that the barrel flexibility, hydrogen bonding interactions and chromophore

planarity of the FPs can be correlated to their contrasting photophysical properties at vaious pressures.

I also investigated the oxygen diffusion pathways in mOrange and mOrange2 which exhibit marked differences in oxygen sensitivities as well as photostability. Such computational identifications of structural changes and oxygen diffusion pathways are important in guiding mutagenesis efforts to design fluorescent proteins with improved photophysical properties.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURE PAGE

ABBREVIATIONS AND ACRONYMS

FPs	Fluorescent Proteins
RFPs	Red Fluorescent Proteins
YFP	Yellow Fluorescent Protein
GFP	Green Fluorescent Protein
EGFP	Enhanced Green Fluorescent Protein
MD:	Molecular Dynamics
CPT:	Constant Pressure and Temperature
MM	Molecular Mechanics
QM	Quantum Mechanics
LJ	Lennard-Jones
PME	Particle Mesh Ewald
VDW	Van der Waals
LES	Locally Enhanced Sampling
RMSF	Root Mean Square Fluctuations
MPa	MegaPascal
RGYR	Radius of Gyration
KFP	Kindling Fluorescent Protein
FRET	Forster Resonance Energy Transfer

1. INTRODUCTION

1.1 WHAT IS FLUORESCENCE

Fluorescence is a phenomenon in which a molecule absorbs radiation, jumps to an excited state and then releases the radiation in the form of light of specific wavelength while going back to the ground state. Usually the wavelength of emitted light is longer (less energetic) than the wavelength of absorbed light. The timescale range for fluorescence process is 1 ns to 100 ns. The molecule is mostly singlet (electronic spins paired) in the ground state. After absorbing light, it transits to an excited singlet state. Once in the excited state, the molecule may lose energy due to vibrational modes. Then finally it jumps back to the ground sate emitting radiation. If the excited singlet state gets converted to the excited triplet state, it may transit to the ground state emitting radiation and this process is called phosphorescence. The phenomenon of phosphorescence takes place over longer time scales ranging from 1 ms to 100 seconds. Various processes taking place during fluorescence and phosphorescence are depicted in Figure 1.1.

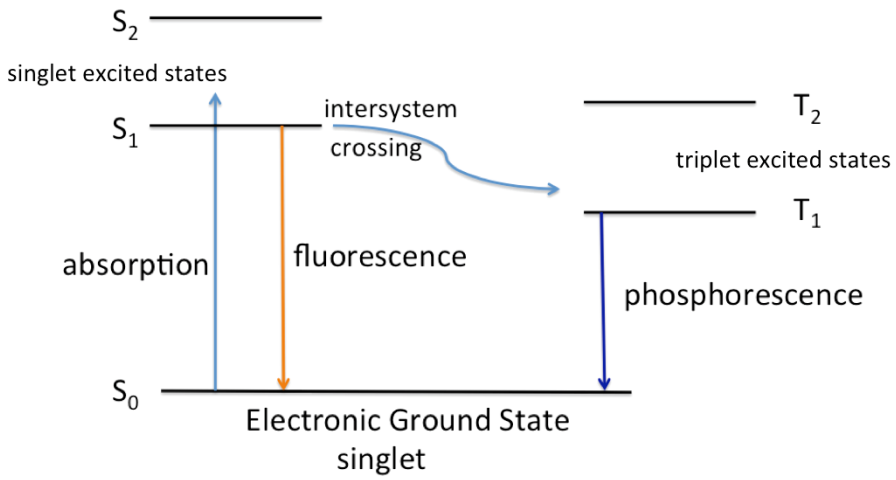


Figure 1.1 Jablonski Energy Diagram

1.2 WHY STUDY FLUORESCENT PROTEINS

Fluorescent proteins are extremely valuable biochemical markers in molecular and cell biology. The fluorescent proteins (FPs) emit in the visible range of light when they are irradiated with light of a specific shorter wavelength. The emission of light is called fluorescence. Because of the fluorescence, when fluorescent proteins are fused with other proteins of interest, it is possible to study the movement, localization and many other physiological processes of the tagged proteins *in vivo* (1-4). Fluorescent proteins are also used as genetically encoded sensors to detect protein-protein interactions on the basis of Forster resonance energy transfer (FRET) applications (5). Fluorescent proteins consist of 220-240 amino acid residues that fold into a barrel shaped beta-sheet scaffold, which encloses a tripeptide structure called the chromophore. The protein owes it fluorescence to the chromophore (2). The side chains of the residues of the beta barrel near the chromophore have an important role to play in the formation and maturation of the chromophore. The electronic molecular orbitals of the chromophore are responsible for its spectral properties, and may be modified by the surrounding amino acid side chains. Understanding the barrel-chromophore interactions and the effect on the chromophores' spectral properties in these proteins will have important implications in designing brighter and more photo-stable fluorescent protein variants that can be used as efficient biochemical markers for studying cellular processes. In order to better understand the system, we subject various fluorescent proteins to high pressures and calculate how the chromophore-barrel interactions are modified.

1.3 DEVELOPMENT OF MONOMERIC VARIANTS

Wild type fluorescent proteins have a tendency to oligomerize. Naturally occurring green fluorescent protein, derived from the jellyfish *Aequorea victoria,* forms a dimer. Similarly, the red fluorescent protein DsRed, obtained from sea coral *Discosoma sp.,* is tetrameric. Oligomeric forms of fluorescent proteins are overly bulky for the tagging purpose, and often take a long time (days) to mature and express their fluorescence. To overcome the shortcomings of the oligomers, various mutations have been carried out to generate monomeric variants. The monomeric FPs mature much faster (in hours), but may be slightly less bright and less photostable than their multimeric progenitors.

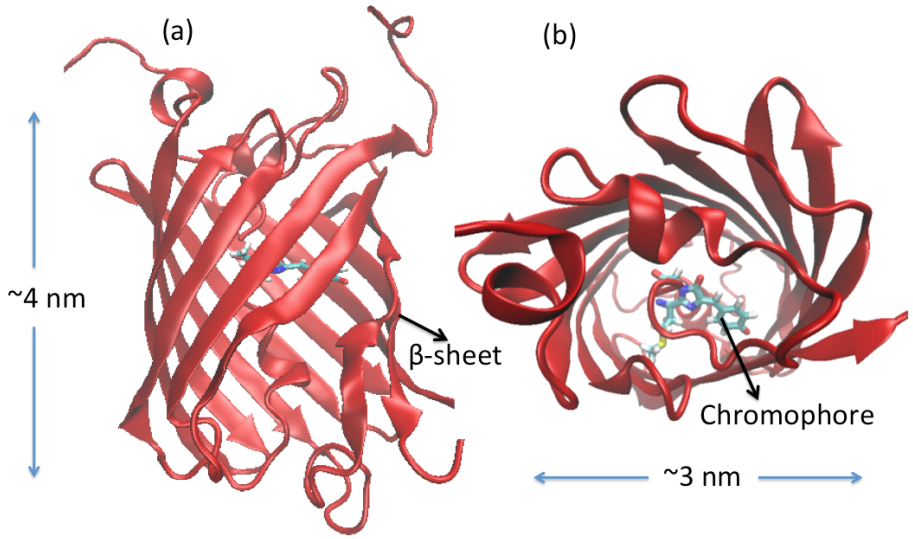


Figure 1.2 Barrel and chromophore structure of a monomeric FP. Figure 1.2(a) shows the protein barrel enclosing the chromophore. Figure 1.2(b) shows the chromophore from the top view of the protein barrel.

A typical monomeric FP consists of around 240 amino acids forming an 11 stranded beta barrel. A distorted helix runs through the axis of the cylindrical barrel. The chromophore is located at the center of the helix, which is formed by a unique post-translational modification of, usually, three amino acid residues. The barrel and chromophore structure of a monomeric FP are shown is Figure 1.2.

To be used as an efficient biochemical marker, the monomeric variant should be bright enough to produce sufficient signal above the cellular autofluorescence, and should have sufficient photostability to be imaged for the time period of the experiment. The monomeric variants investigated in my study are mOrange, mOrange2, mCherry and mStrawberry (derived from RFP) and citrine (derived from GFP).

1.4 PRESSURE EFFECTS ON PROTEINS

Subjecting a protein to higher temperatures increases the thermal fluctuations of the system and structural dynamical modes become active, which allows exploration of more protein configurations. However, there could be a large number of conformational substates unexplored by the temperature effects alone. Pressure is as equally an important thermodynamic parameter as temperature, and can be used to further unravel the protein dynamics and configurational pathways leading to folding, unfolding, or conformational changes.

Pressure experiments on proteins have shown that proteins can be squeezed (6-11). As we apply hydrostatic pressure on globular proteins, water molecules can be pushed inside or forced out. The change in the number of interior water molecules, along with the general squeezing of the protein, can change the number and arrangement of the

internal hydrogen bonds. Mozhaev et al. have reported that electrostatic interactions become much weaker at higher pressures (12). The denaturation of proteins at very high pressures of 1000-1500 MPa may be due to hydration changes occurring inside the protein cavities and changes in hydrogen bond patterns.

Pressure perturbations bring about atomic scale deformations of proteins and other biomolecules, which can affect their structural properties and biological functioning. Proteins subjected to high pressure undergo significant structural changes, varying from elastic effects that may cause spectral shifts without large conformational changes, to complete unfolding (13). For example, experimental studies on myoglobin and Ribonuclease A have shown that these proteins can be denatured when perturbed by hydrostatic pressures around 600 MPa (14, 15). Investigations of the effects on the protein structure of gradually increasing the pressure help in understanding the mechanistic details about protein functioning and interactions between proteins and solvents (7, 16, 17). Pressure effects have also been used to understand the gating phenomenon of ion channels in membrane proteins (18).

1.5 GAS DIFFUSION AND OXIDATIVE PHOTOBLEACHING

There has been a continuous search for improved fluorescent proteins to act as efficient biological markers in terms of brightness, wavelength, photostability, protein structural stability, pH and temperature stability etc. Red fluorescent proteins (RFPs) are especially preferred because the background cellular autofluorescence is low in the red region of the spectrum. Unfortunately, the wild type red fluorescent protein, DsRed is tetrameric, and is too big to be used effectively as a fusion partner (19). To create a

monomer, the tetrameric progenitor dsRed was mutated with 33 substitutions through site directed evolution to obtain the first monomeric variant, named mRFP1 (20, 21). However, obtaining the monomeric variant involves mutating key residues in the monomer that are located at the interfaces of the monomer when formed into a tetramer. These mutations may endanger the monomer stability as a whole. Structural changes resulting from these mutations may also affect the diffusion of gases and ions into the interior through gaps in the protein barrel.

Recent investigations have shown that protein flexibility plays a major role in gas access into many proteins (22-27). Conformational flexibility of the side chains of residues involved in forming transient cavities or pathways can alter the sizes of the bottlenecks for gas diffusion, as well as changing gating mechanisms at the protein surface (28-32). In FPs, in addition to affecting the structures of the chromophore and the protein barrel separately (33), the chromophore−barrel interaction can also affect the fluctuations of the chromophore and the barrel, which in turn can modify the spectral properties and lifetime of the fluorescence (34). It was shown in a recent important work on cyan fluorescent protein that the reduction in the flexibility of a beta strand in the barrel had led to a dramatic improvement in fluorescence quantum yield (35). The halide sensitivity of yellow fluorescent proteins (YFP) was eliminated by the mutation Q69M that occupies the halide-binding cavity in the vicinity of the chromophore (36). In an important work (37), Roy et al. investigated the diffusion pathways of molecular oxygen in the phototoxic KillerRed protein. In this protein, reactive oxygen is generated from molecular oxygen that diffuses into the interior of the protein. They were able to identify pores and channels that allow the oxygen to escape through the protein barrel to the bulk

solvent. This study also suggested that the ease of molecular oxygen diffusion through a channel is the cause of the high susceptibility for photobleaching (37).

Better shielding of the chromophore from the environment, as well as reducing the access of molecular oxygen to the chromophore, has been shown to significantly increase the photostability of both GFPs and RFPs (38). Two mutations (Q64H and F99Y) carried out on mRFP1, resulted in an 11 fold increase in its photostability. Similar mutations were performed on mOrange to produce mOrange2, which provided enhanced photostability and made an excellent fusion partner with other proteins (39).

There are many different ways in which amino acid mutations can affect the fluorescent properties. Though the changes in the electronic orbitals are important for spectral changes, we are especially interested in the structural effects that change the oxygen permeability, and structural flexibility that may affect the chromophore's fluorescent properties. A static image from x-ray crystallography is not enough to investigate the dynamical properties resulting from the gas diffusion and barrel fluctuations. This calls for the need of studying the various transient interactions at the molecular level *in silico*. I have performed molecular dynamics simulations with explicit oxygen to investigate the oxygen diffusions and structural fluctuations in mOrange and mOrange2.

2. METHODS

We investigate the effects of pressure on the interactions between the chromophore and the protein barrel in fluorescent proteins by performing molecular dynamics calculations. These calculations are aimed at determining structural changes, as well as changes in the bonding between the chromophore and the protein barrel. Though we do not perform quantum mechanical calculations to determine changes in the electronic molecular orbitals of the chromophore, we will see below that some aspects of the chromophore's electronic configuration are incorporated in the calculation when we discuss the parameterization of the chromophore later in this chapter.

We first discuss two major issues for determining the accuracy of an MD simulation: the mathematical scheme used to perform the numerical integration for a finite Δt, and the force fields that are incorporated to represent the forces experienced by the particles.

2.1 MOLECULAR MECHANICS: FORCE FIELD MODELS

The most detailed calculations on a molecular system employ quantum mechanical methods for calculating electronic molecular orbital properties and bonds between atoms. These calculations are computationally expensive and not practical for macromolecules containing hundreds of electrons. In this situation, a widely used approach with smaller computational cost is to choose empirical potential energy functions. The potential energy function can then be parameterized so that the energy it produces agrees with experimental data, or with the results of *ab initio* quantum mechanical calculations if available (40). Most of the MD simulation packages use

potential energy functions between atoms that have the approximate form of a harmonic oscillator potential. Use of simple models of force fields can represent the interactions in a biomolecular system from contributions such as bond stretching, angle fluctuations, and rotations about bonds, and form the basis of molecular mechanics calculations.

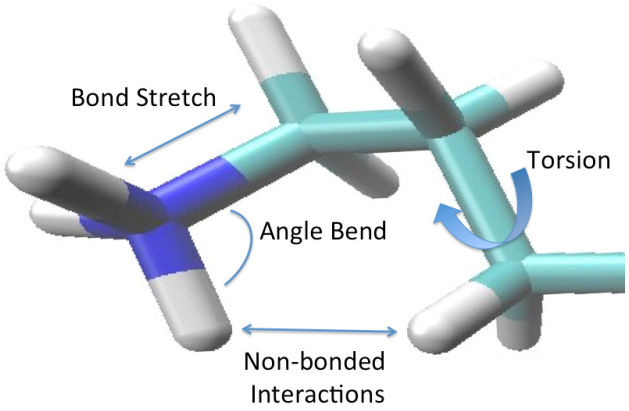


Figure 2.1 Bonded and non-bonded interactions

CHARMM (Chemistry at Harvard Macromolecular Mechanics) is a versatile molecular dynamics simulation package that not only performs energy minimization, molecular dynamics, trajectory analysis, but also enables full parameterization of the novel molecules through comparison with experimental and quantum calculations. We use CHARMM (41) version 35b2, and NAMD (42) version 2.9 with the CHARMM27 force fields for our molecular dynamics simulations. CHARMM27 force field parameters are developed to study proteins, nucleic acids, lipids, carbohydrates etc.

Molecular mechanics calculations compute the energy of a system based on its conformation. The interactions between various atoms in a molecule are classified into bonded interactions and non-bonded interactions. The sum of individual bonded and non-

bonded interactions contribute to the energy of the specific conformational state. The potential energy terms include energetic penalties when bonds and angles deviate away from their preferred arrangement. Since a system with many atoms has many interactions, it is unlikely that any configuration will allow all the interactions to be minimized simultaneously. This is an example of "frustration". However, the configuration that best minimizes the total energy of the system is the equilibrium configuration. The various potential energy terms for the nearest and non-nearest neighbor interactions are described in detail below.

2.1.1 BONDED INTERACTIONS

Bond Stretching

A chemical bond can be modeled as a spring with equilibrium bond length r_0 and stretching force constant k_s. The energy required to elongate or compress the spring can be expressed using Hooke's law.

$$U(r) = \frac{k_s}{2}(r - r_0)^2 \tag{2.1}$$

A realistic bond potential is not perfectly harmonic. Morse proposed a realistic functional form used to model bond stretching:

$$U(r) = D_e\{1 - \exp[-a(r - r_0)]\}^2 \tag{2.2}$$

In equation 2.2, D_e is the depth of the potential energy minimum at $r = r_o$ compared to the potential at $r = \infty$, and $a = w\sqrt{(\mu/2D_e)}$, where μ is the reduced mass and ω is the frequency of the bond vibration. The Morse potential is represented in Figure

2.2. Although the Morse potential is more realistic than the harmonic potential, it requires more parameters and is computationally expensive (43).

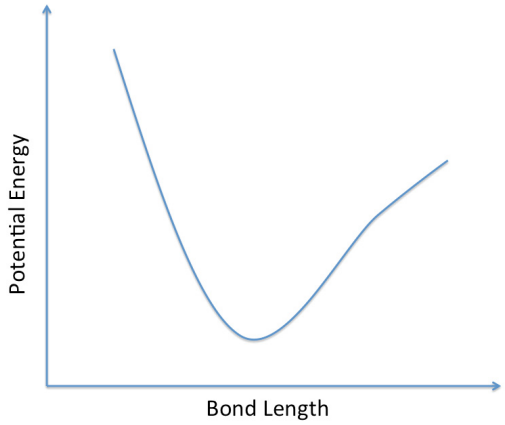


Figure 2.2 Potential energy versus separation for Morse Potential

Angle Bending

The angle bending energy is the energy required to bend a bond from its equilibrium angle, θ_0. The deviations of angles from their equilibrium values are modeled using the harmonic potential:

$$U(\theta)= \frac{k}{2}\left(\theta - \theta_0\right)^2 \qquad (2.3)$$

In equation 2.3, k is the force constant. A diagram for this potential is displayed in Fig. 2.3. The energy penalties for angle bending are less severe than those for bond stretching, and are reflected by appropriate small values for k.

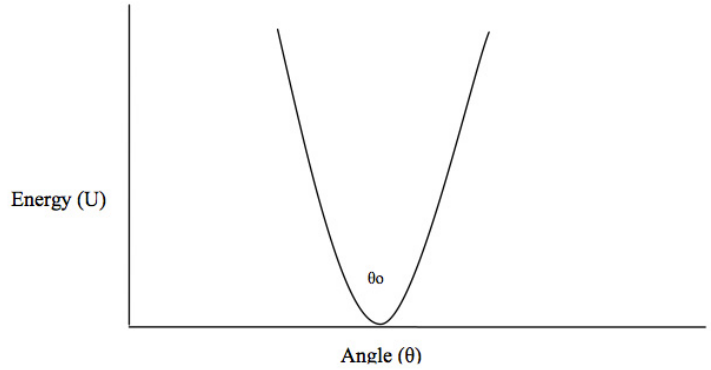


Figure 2.3 Potential energy function for bond angle bending.

Torsions

Rotation about a chemical bond causes variation in structure and relative energy. If four atoms A-B-C-D form a bonded quartet, the angle between the planes ABC and BCD is referred to as the torsion angle. A complex interplay between the torsional and non-bonded interactions result in diverse conformational states. The variations in torsional potential energy for various torsional angles of ethane are shown in Figure 2.4 (44). The torsion potential is expressed as a summation of a cosine series:

$$U(w) = \sum_{n=0}^{N} \frac{k}{2}[1 + \cos(nw - \gamma)] \quad (2.4)$$

In equation 2.4, w is the torsion angle, k is the barrier height, γ is the phase factor for minimum potential energy, and n denotes multiplicity.

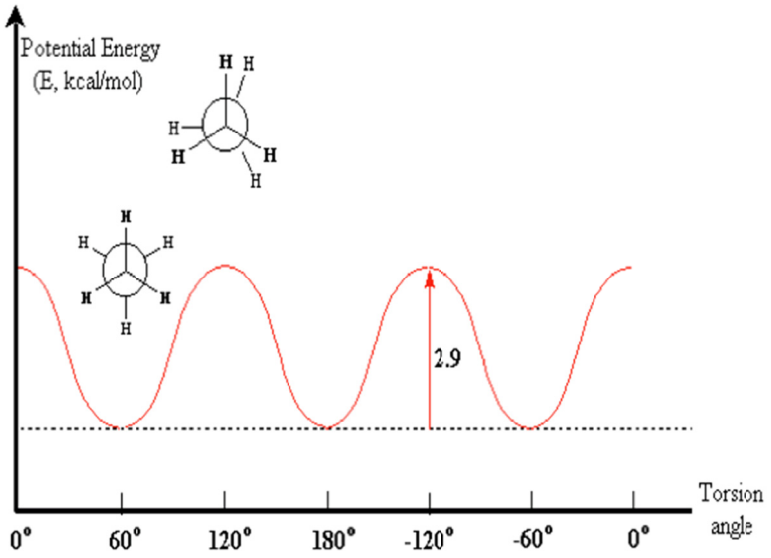


Figure 2.4 Potential energy of ethane as a function of torsion angle. Energy is maximum at eclipsed position.

Improper torsion terms are used to enforce the planar structure of a molecule when needed. Usually, improper torsion angles are defined when atoms are not bonded in a sequence. Improper torsion energy, also called the out-of-plane energy function, can be used to achieve a particular geometry.

2.1.2 NON-BONDED INTERACTIONS

The non-bonded interactions mainly consist of the electrostatic and van der Waals interactions. To compute the non-bonded interactions in a system of N atoms, we get $N^2/2$ distinct pairs. Hence, evaluation of LJ and electrostatic interactions is the computationally most expensive of all other interactions. The non-bonded interactions are modeled as an inverse power of the distance. Two types of non-bonded interactions are explained below in detail.

Electrostatic Interactions

These interactions include the attractive and repulsive forces between charged atoms and molecules. A customary approach for calculating electrostatic interactions is by assigning partial atomic charges to each atom and applying Coulomb's law:

$$U(r) = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (2.5)$$

Here, N_A and N_B are the number of point charges in the two molecules of the system, A and B. The permanent partial atomic charges are parameters that are determined by making an initial guess and then performing calculations in an attempt to reproduce the results of experiments or QM calculations. The discrepancy in the calculations is used to guide the modification of the assigned partial charges. The process is repeated until the assigned charges produce calculated results that are in good agreement with experimental results. The definition of "good" is relative and limits can be set by the user based upon considerations such as computational speed. For periodic systems, the electrostatic interactions are calculated efficiently for small as well as large distances by using the Particle-Mesh-Ewald (PME) method (45).

Lennard-Jones Interactions

The Lenard-Jones potential involves long range dispersion interactions (van der Waals attractive forces) and the short range repulsive interactions between non-bonded inert atoms. The dispersion interactions are due to the inducement of dipole moments by one atom in another, and vice versa. The induced dipole moments form the basis for van der Waals attraction (46). If the atoms come too close, the energy contribution increases

exponentially. This effect can be explained by the Pauli Exclusion Principle, which prohibits any two electrons in the system from having the same set of quantum numbers.

The LJ interaction is modeled as a sum of attractive and repulsive terms as:

$$U(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{2.6}$$

The LJ potential contains two adjustable parameters: the collision diameter (σ) and the potential well depth ε. Equation 2.6 can be plotted against the distance of separation as shown in Figure 2.5.

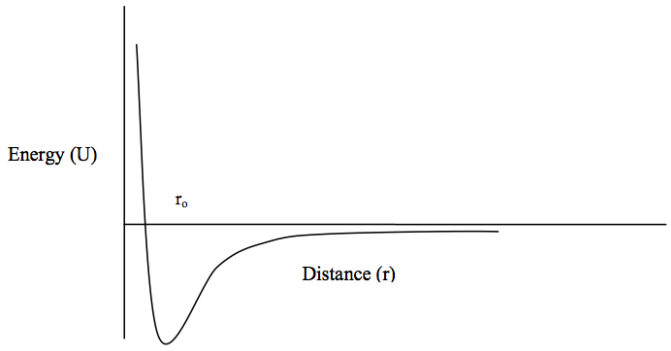


Figure 2.5 Lennard-Jones potential

Since the number of non-bonded interactions scales as the square of the number of particles of the system, calculating every non-bonded interaction between every atom imposes a huge computational cost. A commonly applied solution is to impose a cutoff distance on the non-bonded force calculations. Care must be taken because this could have serious drawbacks while simulating a system with charged particles, since the electrostatic interaction has infinite range.

To reduce the computational cost, non-bonded interactions are truncated at a certain cutoff distance, and beyond the cutoff distance no interaction is calculated. The cutoff criterion may introduce a discontinuity, as the energy may abruptly drop to zero. This can lead to unreliable results or early termination of MD simulations. For example, two atoms may be separated by a distance that is close to the cutoff. In one time frame of the MD simulation, the separation distance may be less than the cutoff and in the next time frame it may be slightly more than the cutoff. Although the separation distance may fluctuate very slightly, the interaction energy changes abruptly.

Special functions (shifting, switching etc.) are used to smoothly vary the energy at the cutoff distance (47), so that energy function smoothly goes to zero at the cutoff distance. The switching function in CHARMM introduces three different cutoff distances: ctonnb, ctofnb, and ctnb. The switching function begins to affect the LJ potential starting at the smallest of the three, ctonnb. The switching function brings the interaction smoothly to zero at ctofnb. The value of ctnb refers to the distance for generating the non-bonded pair list, and is updated periodically. The distance ctnb is slightly greater than ctofnb. The various cutoff distances are representated in Figure 2.6. The non-bonded interactions between 1-2 and 1-3 neighbors are always excluded, as they are the nearest neighbors bonded chemically.

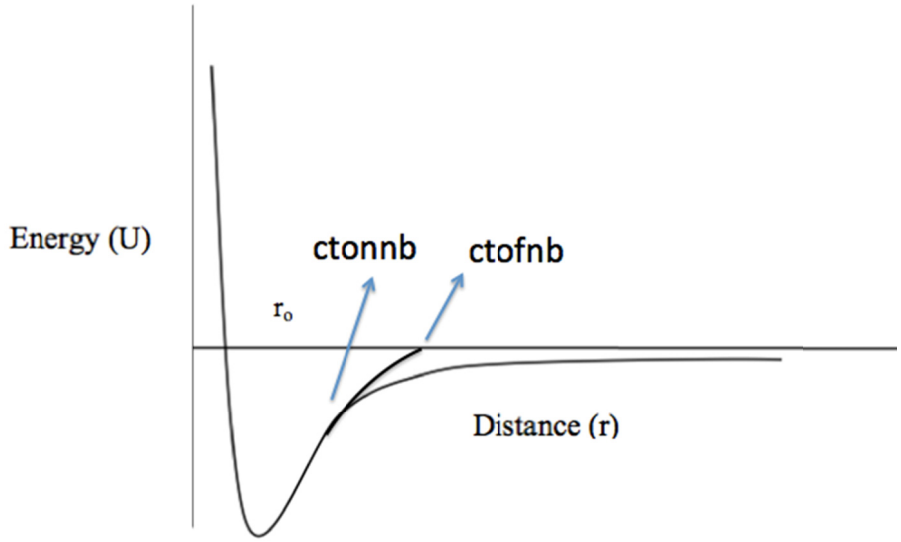


Figure 2.6 Various cutoff distances for LJ potential.

2.2 TIME INTEGRATION

Molecular dynamics uses Newton's laws of motion to calculate the position of a particle as a function of time. Basically, **F**=m**a** is used to calculate the acceleration **a** of each particle at a given time The acceleration of a particle, along with its momenta are then numerically integrated to determine the new position of a particle after a time increment Δt. Some of the commonly used numerical integration algorithms are explained below.

2.2.1 VERLET ALGORITHM

The Taylor series expansion for the position of the particle in small time step Δt can be written as:

$$r(t+\Delta t) = r(t) + \frac{dr}{dt}\Delta t + \frac{1}{2}\frac{d^2r}{dt^2}(\Delta t)^2 + \cdots$$

The same expression can be written for a reverse time step Δt as:

$$r(t - \Delta t) = r(t) - \frac{dr}{dt}\Delta t + \frac{1}{2}\frac{d^2r}{dt^2}(\Delta t)^2$$

Adding the above two equations, we get

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + a(t)(\Delta t)^2 \qquad (2.7)$$

$$a(t) = \left(\frac{d^2r}{dt^2}\right)_t$$

Equation (2.7) gives the position of a particle after a time interval Δt, provided that the position and acceleration at the current time, and the position at the previous time t-Δt is known. The Verlet algorithm does not depend on velocity, dr/dt, to calculate the new position of the particle. The acceleration at the current time can be calculated from the force acting on particle at the current time.

$$a = \frac{F}{m}$$

$$F = -\frac{dU}{dr}$$

U is the potential energy function that is incorporated in the MD program to represent the interactions among the particles.

2.2.2 LEAP-FROG ALGORITHM

Other integration schemes involve continuous updating of velocity as well as position. These integration schemes are employed when knowledge of the velocity is necessary, such as when the temperature of the system is controlled. A specific temperature means that the average kinetic energy of the particles must have a specific value. In order to assure this, periodically the MD program examines the kinetic energy and decides if an adjustment is necessary to return the system to the correct temperature. This requires knowledge of the speeds of all the particles. A useful integration scheme for updating the position of the particles that also allows velocity scaling is the Leap-frog algorithm, which is a slightly modified form of the Verlet algorithm.

The method is called Leap Frog because some quantities are calculated at integer values of Δt, whereas other quantities are calculated at Δt/2. The position of a particle can be expressed through a Taylor series expansion with a time step of Δt/2 as follows:

$$r\left(t+\frac{\Delta t}{2}+\frac{\Delta t}{2}\right)=r\left(t+\frac{\Delta t}{2}\right)+\frac{dr(t+\frac{\Delta t}{2})}{dt}\frac{\Delta t}{2}+\frac{1}{2}\frac{d^2r(t+\frac{\Delta t}{2})}{dt^2}(\frac{\Delta t}{2})^2+\cdots$$

$$r\left(t+\frac{\Delta t}{2}-\frac{\Delta t}{2}\right)=r\left(t+\frac{\Delta t}{2}\right)-\frac{dr(t+\frac{\Delta t}{2})}{dt}\frac{\Delta t}{2}+\frac{1}{2}\frac{d^2r(t+\frac{\Delta t}{2})}{dt^2}(\frac{\Delta t}{2})^2+\cdots$$

Subtracting we get

$$r(t+\Delta t)=r(t)+v(t+\tfrac{\Delta t}{2})\,\Delta t \qquad (2.8)$$

The definition for the velocity is

$$v\left(t+\frac{\Delta t}{2}\right)=\frac{dr(t+\frac{\Delta t}{2})}{dt}$$

To express this in terms of a=F/m , let us expand the velocity in the time step Δt/2

$$v\left(t+\frac{\Delta t}{2}\right)=v(t)+\frac{dv(t)}{dt}\frac{\Delta t}{2}+\frac{1}{2}\frac{d^2v(t)}{dt^2}(\frac{\Delta t}{2})^2+\cdots$$

$$v\left(t-\frac{\Delta t}{2}\right)=v(t)-\frac{dv(t)}{dt}\frac{\Delta t}{2}+\frac{1}{2}\frac{d^2v(t)}{dt^2}(\frac{\Delta t}{2})^2+\cdots$$

Subtracting we get

$$v\left(t+\tfrac{\Delta t}{2}\right)=v\left(t-\tfrac{\Delta t}{2}\right)+a(t)\Delta t \qquad (2.9)$$

$$a(t)=\frac{dv(t)}{dt}$$

Equation (2.9) is used to calculate the velocity at $(t+\frac{\Delta t}{2})$; this velocity is used to find the position at $(t+\Delta t)$ using equation (2.8). Consequently, the velocity leaps over position first, and then the position leaps over the velocity. The velocities at time t can be approximated by the relationship

$$v(t)=\frac{1}{2}\left[v\left(t+\frac{\Delta t}{2}\right)+v\left(t-\frac{\Delta t}{2}\right)\right]$$

2.2.3 TIME INCREMENTS, Δt

In molecular dynamics, the motions of atoms occur on different time scales. Characteristic times for vibrational motions involving strongly bonded atoms are high frequency motions, while non-bonded interactions are low frequency motions. For example, the highest frequency of motion in biomolecular system is the vibrational motion of a light atom chemical bond to a heavy atom and hydrogen (time period ~ 10^{-14}s). A very small time interval ~ 1 fs is required to capture this motion. Unfortunately, this requires many time steps to investigate protein motions that occur on 100 ns time scales. One way to get around this is by constraining the bond between a heavy atom and hydrogen to be a fixed length with no vibrations. This process is facilitated by the use of the SHAKE command in CHARMM. This allows the use of a slightly longer time interval (~ 2 fs).

2.3 CONSTANT PRESSURE AND TEMPERATURE SIMULATIONS

Molecular dynamics simulations perform calculations on particles in a microscopic system, but under imposed macroscopic conditions involving temperature, pressure, and volume. To be able to perform constant temperature simulations, the parameters for energy fluctuations have to be incorporated. On the other hand, to simulate isobaric conditions, volume has to be introduced as a dynamical variable. Constant temperature and pressure simulations have been made possible by the pioneering work of incorporating energy and volume fluctuations in a dynamical system, such that the trajectory average of any property is equal to its NPT ensemble average (48-51). One way of applying NPT dynamics is through the modifications of the system's

Hamiltonian (48). The terms in the Lagrangian are redefined to include the volume V, and a piston mass M, and a stochastic collision term. The piston can compress the system to vary the pressure, and the collision term can vary the energy of the system. A second approach to implementing an NPT ensemble is through the weak coupling to an external bath (49). Stochastic and friction terms are introduced in the equations of motion, resulting in a Langevin equation.

Isobaric-isothermal simulation is enabled in CHARMM through the use of command "Dynamics CPT". Constant pressure simulation is made possible only through the use of periodic boundary conditions. The size of the system is calculated and updated each step to allow the volume fluctuations for maintaining constant pressure. Two keywords 'pmass' and 'tmass' have to be specified. The term 'pmass' refers to the mass of the pressure piston, and setting any component of 'pmass' to zero will fix the size of the cell. Mass of the temperature piston "tmass" is kept ten times the mass of the pressure piston. Simulations at higher pressure cause volume compression, eventually increasing the density of the system.

2.4 PARAMETERIZATION OF THE CHROMOPHORE

The chromophore of florescent protein is formed by the auto-oxidation and cyclization of three to four amino acid residues. As a chemically modified entity, the force field parameters developed for regular amino acid residues do not accurately represent the chromophore. This calls for the need of chromophore parameterization. Fortunately, the chromophore of a green fluorescent protein has been parameterized by Reuter et al. (52) for use in CHARMM simulations. The chromophore of GFP is very

similar to the chromophore of mCherrry, mStrawberry and citrine. Hence the force field parameters for the chromophore of mCherry, mStrawberry and citrine were adopted from the anionic GFP chromophore parameters, with an addition of parameters for acylimine nitrogen from the CHARMM27 force fields.

The formation of the mOrange chromophore involves two cyclizations: one leading to the formation of imidazolinone ring, and the other forming the novel five-member oxazole ring. The parameters for the imidazolinone ring are obtained from GFP chromophore, while those for the oxazole like ring are obtained from the CGenFF parameters (53) of oxazole and similar five member heterocyclic rings. The CHARMM General Force Field (CGenFF) presents the force fields for a wide range of chemical groups, including various heterocyclic scaffolds for use in CHARMM simulations. The parameterization procedure is based on analogy to the bond, angle and dihedral types as a first approximation. The chromophore of mOrange2 is identical to that of mOrange, and the mutations are only performed on the barrel.

2.5 IDENTIFYING PROTEIN CAVITIES: CAVER CALCULATIONS

A protein barrel includes small cavities and tunnels inside, which provide the microenvironment for molecular diffusion and protein ligand interactions. The shape and structure of protein clefts and tunnels is transient, and depends on the protein barrel fluctuations and molecular interactions occurring inside the cavity. Caver (54) is a software package developed to identify such protein tunnels and cavities. The software searches for channels by exploring the interior of the protein. The protein atoms are modeled with hard sphere atoms with appropriate van der Waals radii, and represented on

a three dimensional grid space. The site and probe radius for starting the channel calculation has to be specified in the input file. Caver groups nearby tunnels into a single tunnel cluster. Performing caver calculations on a large number of frames may be computationally very expensive. Instead, frames from a long trajectory can be sampled at a specific time interval to carry out Caver calculations. The probe radius chosen in my study is 0.8 Å and the chromophore of each FP is set as the starting point for channel calculations.

2.6 LOCALLY ENHANCED SAMPLING

Locally enhanced sampling is a computational method of increasing the speed for exploring the behavior of the system. For example, we are interested in locating all the pathways for oxygen entry into the barrel. One way to determine the pathways is to place one oxygen molecule at a location outside the barrel and carry out an MD simulation to see where it enters. We would then repeat the MD simulation with another oxygen molecule initially placed in a different location outside the barrel. Each oxygen location would require the performance of its own time consuming MD simulation. Instead, the exploration of multiple pathways, i.e the sampling of the system can be done during a single MD simulation. This is done by placing multiple, non-interacting copies of the oxygen molecule, known as enhanced atoms (55, 56), at many different initial locations simultaneously. The copies of the oxygen do not interact with each other, and can occupy the same space as well. The “multiply” command in NAMD facilitates the generation of multiple copies. The beta column in the pdb file represents the presence of enhanced

atoms. A beta value of 0 represents a normal atom, and a non-zero beta value represents the enhanced atom.

I have used 15 copies of the oxygen molecules in my LES molecular dynamics simulation for identifying the oxygen diffusion pathways inside the protein barrel.

2.7 OTHER COMPUTATIONAL DETAILS

2.7.1 SIMULATION CONDITIONS FOR mCHERRY, mSTRAWBERRY AND CITRINE

Initial protein structures were obtained from the protein data bank (mCherry: pdb code 2H5Q, mStrawberry: pdb code 2H5P, and citrine: 1HUY). The parameters for the chromophores of all three FPs were obtained from the work on the anionic chromophore of the green fluorescent proteins (52). Glu215 was protonated in both mCherry and mStrawberry as observed by Shu et. al. (34). The MMTSB toolset (57) was used to set up the system. The initial structures of the FPs were separately solvated using TIP3 water molecules in rhombic dodecahedron boxes. The minimum distance of a box edge from the protein was set to be 10 Å. In addition to the water molecules already present in the crystal structures, mCherry was solvated with 9641 water molecules, mStrawberry with 8132 water molecules, and citrine with 8463 water molecules.

Energy was minimized using steepest descent followed by the adopted basis Newton-Raphson (ABNR) method. The energy-minimized systems were then neutralized by adding sodium and chloride ions, distributed throughout the volume. The systems were then heated with a linear gradient of 10 K/ps from 100 K to 300 K. At 300 K, the systems were equilibrated for 1ns with a 2 fs integration time step in the isobaric-

isothermal (constant pressure-temperature, CPT) ensemble. Finally, production runs of 20-30 ns were performed under the isobaric-isothermal ensemble at 300 K. At high pressures 750 MPa and 1000 MPa, 15 ns simulations were performed for mCherry and mStrawberry. To investigate similar pressures as done experimentally, our molecular dynamics simulations were performed at five different pressures: 0.1 MPa, 250 MPa, 500 MPa, 750 MPa and 1000 MPa for mCherry and mStrawberry, and at six different pressures for citrine: 0.1 MPa, 50MPa, 100 MPa, 250 MPa, 500 MPa and 750 MPa. To allow sufficient time for the pressure equilibration of the solvated systems, only the last 10 ns trajectory from the simulations was used for the analysis.

A 1ns CHARMM simulation takes about 2 days. The total computational time for performing 20 ns simulation of mCherry and mStrawberry at 5 different pressures is ~10000 processor hours. The total computational time for performing 30 ns simulation of citrine at six different pressures is ~9000 processor hours.

2.7.2 SIMULATION CONDITIONS FOR MORANGE AND MORANGE2

The initial X-ray crystallographic structure of mOrange (pdb code 2H5O) was obtained from the Protein Data Bank. The more photostable mOrange2 was obtained through four mutations Q64H/F99Y/E160K/G196D on the mOrange protein barrel using MODELER (58). The VMD package was used to setup the system for simulations. The initial structures of mOrange and mOrange2 with crystallographic water molecules and one molecular oxygen each were solvated by using a box cutoff of 10 Å in VMD. The solvated systems were then electrically neutralized using the VMD autoionize plugin, and the oxygen molecule was multiplied to 15 copies in each system to enable LES sampling.

The final systems contained a total of 28,892 atoms in mOrange and 29,195 atoms in mOrange2. All water molecules overlapping with the protein were removed. The particle mesh Ewald method was used to treat long-range interactions with a 12 Å non-bonded cutoff. Energy minimization was performed using the conjugate gradient and line search algorithm. Each system was then heated with a linear gradient of 20 K/ps from 20 to 300 K. At 300 K, each system was minimized for 300 steps and equilibrated for 5 ps with a 2 fs integration time step in the NVT (constant number, volume, and temperature) ensemble. Langevin dynamics was used to maintain the temperature at 300 K. For each system, an 80 ns production run was performed with NVT dynamics and 2 fs time step. The first 10 ns trajectory was not used for analysis to allow sufficient time for equilibration of each system.

The total computational time for carrying out 80 ns LES simulation of mOrange and mOrange2 is ~4000 processor hours.

3. STRUCTURAL FEATURES OF THE PROTEIN BARREL THAT ALLOW OXYGEN ACCESS AT AMBIENT PRESSURE

In this chapter I provide the results of my calculations performed at ambient pressure. For all of the proteins that I investigated, I performed computations to uncover channels that allow molecular oxygen in the solvent to penetrate the protein barrel and access the chromophore. mOrange and mOrange2 differ in only a few mutations, but have very different oxygen sensitivities. I performed additional calculations to elucidate the structural features that are responsible for the different oxygen sensitivities of mOrange and mOrange2.

3.1 PROTEIN BARREL AND CHROMOPHORE OF mORANGE, mORANGE2, mCHERRY, mSTRAWBERRY, AND CITRINE

The monomeric variants mOrange, mCherry and mStrawberry were derived from the wild type tetrameric progenitor DsRed. The naturally occurring tetramer has a very high fluorescence quantum yield of 0.79 and is very bright. The 33 mutations carried out on DsRed resulted in the first generation monomeric RFP known as mRFP1. Development of the monomeric RFP overcame the problems of tetramerization and the slow maturation of the parent protein DsRed. However, mRFP1 exhibited a very small quantum yield of 0.25 and a brightness of 21% of that of DsRed, possibly due to the compromised barrel structure caused by the mutations introduced to break the tetramer interactions at the interface between the monomers. The immediate challenge was to develop monomeric RFP variants with higher quantum yield and photostability. The following mutations were carried out on mRFP1 to obtain mOrange: Q66M, T147S, M182K, M66T, T195V, T41F and L83F. Although these mutations were able to greatly

enhance the brightness of mOrange, four additional mutations (Q64H/F99Y/E160K/G196D) in the protein barrel were needed to achieve a highly photostable variant, mOrange2. The chromophores of mOrange and mOrange2 are identical, and are composed of three connected rings formed by the oxidation and cyclization among four amino acids: Phe, Thr, Tyr and Gly. mCherry and mStrawberry were also developed from mRFP1. The chromophore of mCherry is made up of Met, Tyr and Gly, and that of mStrawberry is made up of Thr, Tyr and Gly. The mutations carried out to develop all the RFP variants are depicted in Figure 3.1. The chromophores of various FPs investigated in my study are shown in Figure 3.2.

The yellow fluorescent proteins are obtained through the mutation T203Y in green fluorescent proteins (59). Citrine is a variant of YFPs derived from the GFP with mutations S65G/V68L/Q69M/S72A/T203Y (36). The mutation T203Y is responsible for red shift, and the mutation Q69M fills a large halide binding cavity, thus decreasing the halide sensitivity. The chromophore of citrine is formed by the cyclization of Gly-Tyr-Gly, resulting in the formation of a heterocyclic imidazolinone ring (60). The β7 – β10 gap in citrine is slightly smaller than that of mCherry (61).

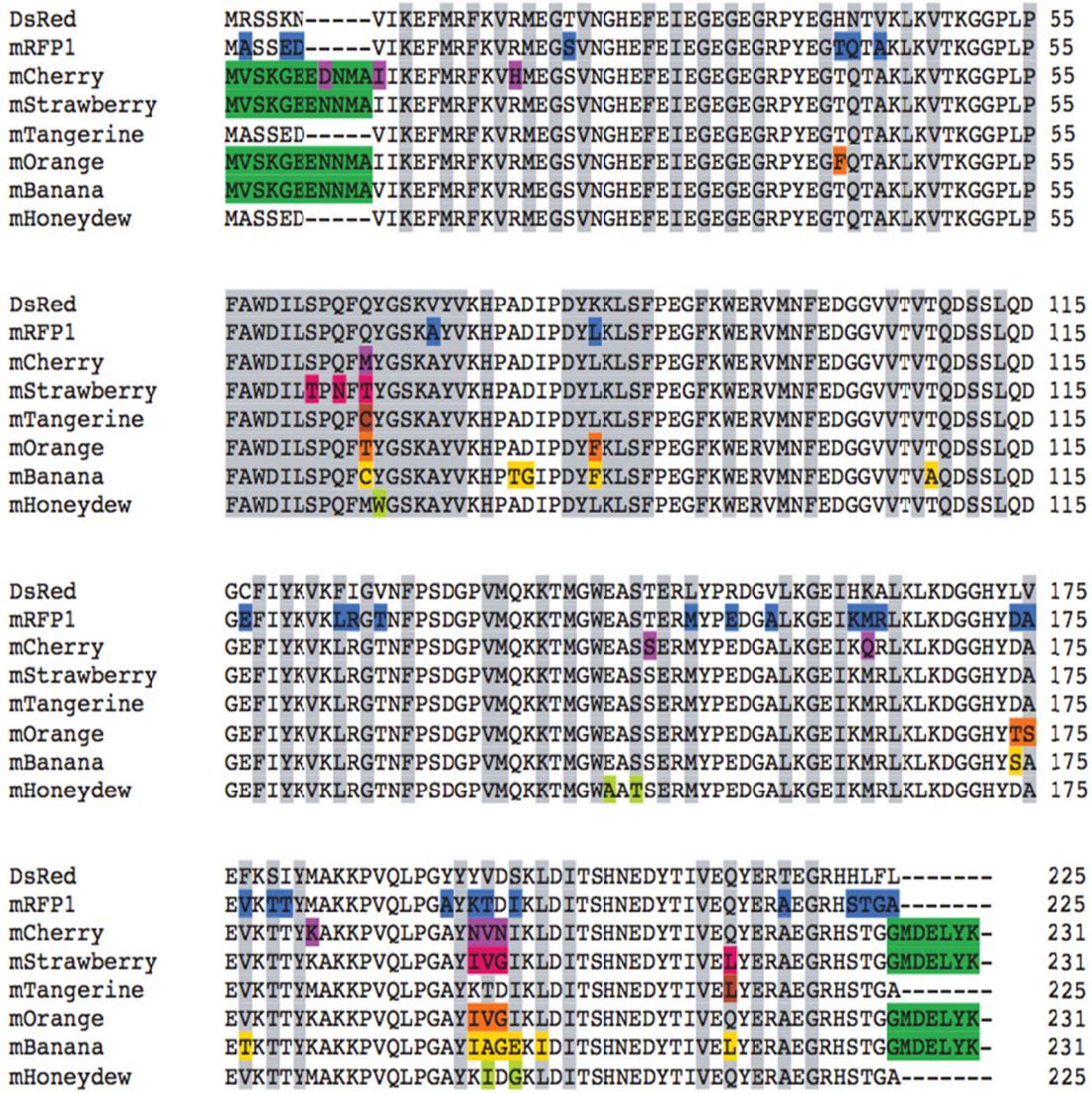


Figure 3.1 Sequence alignment of mRFP variants with the wild type DsRed and mRFP1. mRFP1 mutations are shown in blue, and critical mutations in mCherry, mStrawberry, mTangerine, mOrange, mBanana and mHoneydew are shown in colors corresponding to the color of each variant. Only the GFP type termini of mRFP variants are shown in green. Figure 3.1 is obtained from reference (62).

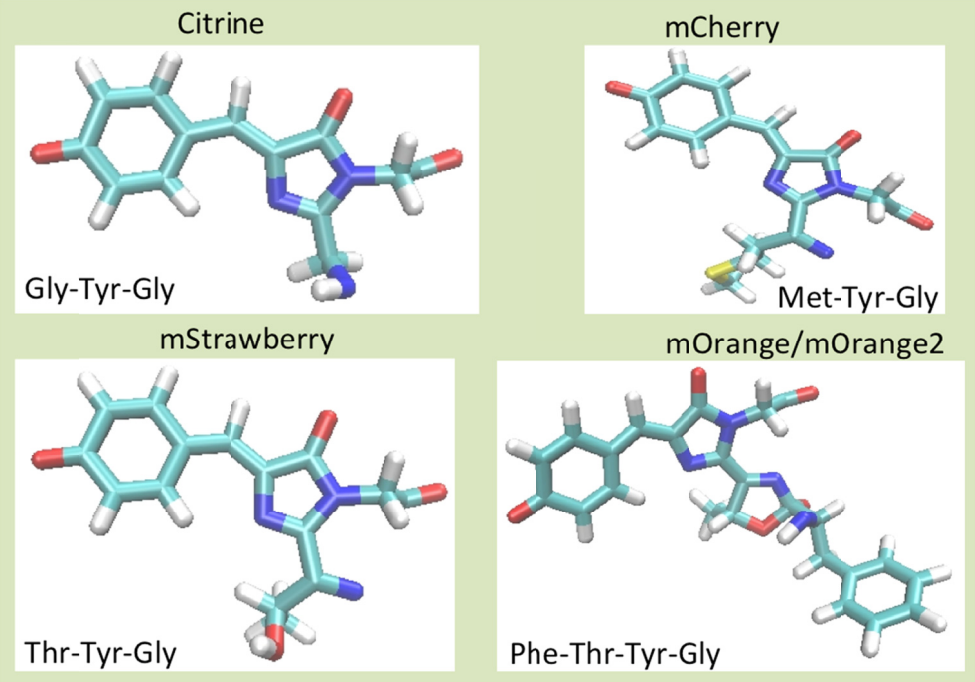


Figure 3.2 Chromophores of various fluorescent proteins

3.2 COMPUTATIONALLY DETERMINED PROTEIN CHANNELS

The protein barrel contains channels for the diffusion of water, gases and ions. The channels may be transient, opening and closing due to barrel fluctuations, and the presence of water or gases in the immediate vicinity may modify the shape and structure of the channels. The Caver software (54) was used to identify tunnels that start from the chromophore and lead to exits outside the barrel, with a probe radius of 0.8 Å. I used 200 equally spaced frames extracted from a 10 ns simulation of each FP for the Caver calculations. Table 3.1 shows the number of tunnels present in all the variants at atmospheric pressure. Nearby tunnels were grouped into single tunnel cluster.

Table 3.1 Number of tunnels in various FPs

Fluorescent Protein	**Number of tunnel clusters**	**Number of tunnels**
mOrange	35	178
mOrange2	31	95
mCherry	33	292
mStrawberry	29	191
Citrine	10	144

The insights gained from channel calculations can be useful for understanding the barrel flexibility and gas diffusion in FPs. A relatively small number of tunnel clusters is found in citrine. The greater number of tunnel clusters and significantly larger number of tunnels in mOrange provides a good explanation for the experimental observations that show that mOrange is more permeable to molecular oxygen than mOrange2. The oxygen sensitivity of mOrange and mOrange2 is investigated in detail in the next section using explicit oxygen molecular dynamics simulation. mCherry and mStrawberry also show a significantly larger number of tunnels inside the protein barrel than citrine. Figure 3.3 shows the tunnel clusters present in various FPs at atmospheric pressure. The tunnels belonging to the same cluster are represented by same color.

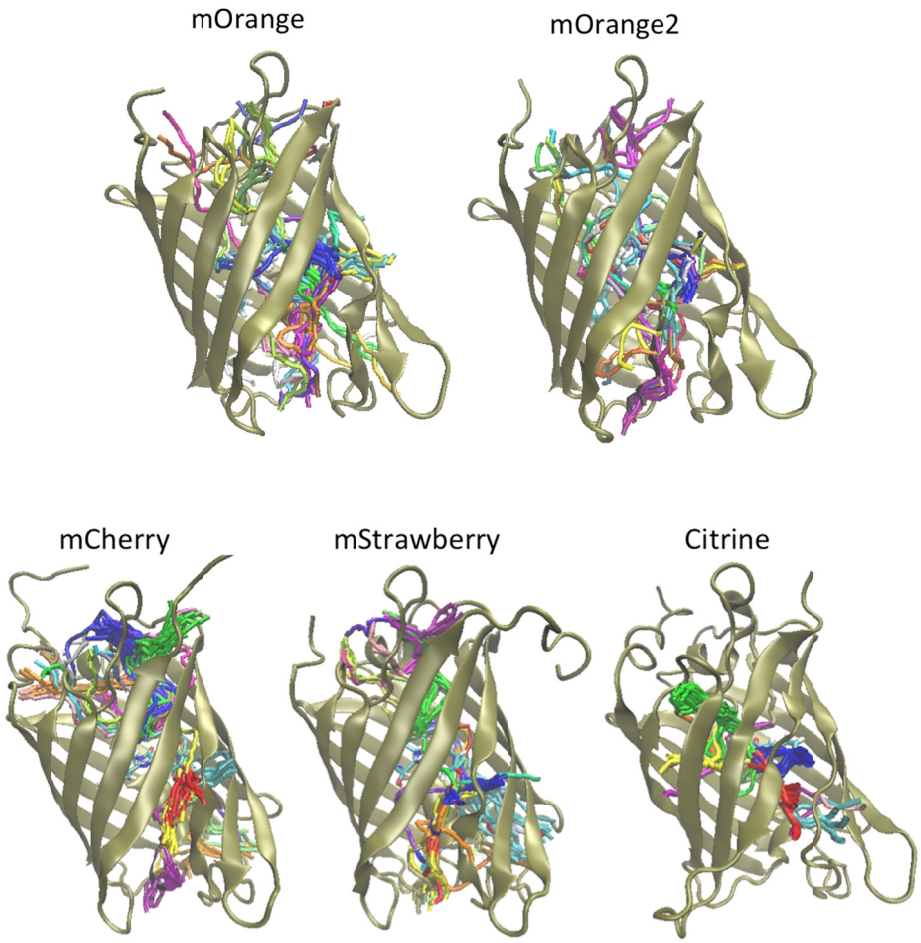


Figure 3.3 Channels present in FPs at atmospheric pressure. The tunnels belonging to the same cluster are represented by the same color.

3.3 OXYGEN DIFFUSION IN mORANGE AND mORANGE2

Mutations were performed experimentally on mOrange to produce mOrange2. These mutations were successful in significantly increasing the fluorescence lifetime. I now describe the mutations that were performed, and then show how the results of my computational investigations provide molecular level explanations for the effects of the mutations.

3.3.1 EXPERIMENTAL STUDIES ON OXYGEN SENSITIVITY

Although the mutations performed on mRFP1 greatly enhanced the fluorescent quantum yield, the photostability could not be improved for most of the variants. One possible cause of low photostability was easy oxygen access through the barrel to the chromophore which causes fluorescence quenching. Four important mutations (Q64H, F99Y, E160K and G196D) were carried out on mOrange to obtain a highly photostable variant, mOrange2. Among the mutations introduced to obtain mOrange2, a combined mutation of Q64H and F99Y imparted about a tenfold increase in photostability. The same combination of Q64H and F99Y conferred an 11-fold increase in photostability to mRFP1. Oxygen sensitivity experiments performed on both mOrange and mOrange2 showed that anoxia (low oxygen concentration) resulted in a large increase in the photobleaching half-life of mOrange, but the photobleaching half-life for mOrange2 remained unchanged at different oxygen concentrations. This implies that the mutations incorporated into mOrange2 successfully prevented the access of oxygen to the chromophore even at normal oxygen concentrations. Various photophysical properties of mRFP1, mOrange and mOrange2 are given in table 3.2 (62).

Table 3.2 Various photophysical properties of mRFP1, mOrange and mOrange2. Brightness is calculated as (extinction coefficient x quantum yield)/1,000. Arc lamp illumination causes each molecule to emit 1,000 photons/s initially. ND – not determined.

Fluorescent Protein	**Quantum Yield**	**Brightness**	**$t_{1/2}$ bleach (arc lamp) (s)**	**$t_{1/2}$ bleach (O_2 - free) (s)**
mRFP1	0.25	13	8.7	ND
mOrange	0.69	49	9.0	250
mOrange2	0.60	35	228	228

As with other mFruit variants, the barrel structural integrity of mOrange is compromised, especially in the β7 and β10 region, due to missing tetrameric interactions present in the naturally occurring DsRed. Transient thermal fluctuations (61) can allow easier oxygen access to the chromophore. This may help chromophore maturation but can cause fluorescence quenching or faster photobleaching due to oxidation. In similar work by our group, our simulation results showed that the M163Q mutation in mCherry significantly reduces molecular oxygen entry into the barrel, which might help explain the role of molecular oxygen in permanent photobleaching of FPs and improving the photostability in mCherry (63).

A floppy barrel can cause fluorescence problems in other ways. In the cyan fluorescent protein, β7 flexibility has been thought to cause collisional fluorescence quenching due to the collision of the Ile146 side chain with the chromophore. In a recent work, structure guided amino acid replacements to reduce β7 flexibility have led to a significantly brighter and highly photostable fluorescent protein mTurquoise2, with the highest quantum yield (93%) among monomeric fluorescent proteins (35).

Identification of barrels and cavities, as described in section 3.2, showed that there could be various pathways for gas diffusion and pockets that host gas molecules inside the protein barrel. I have performed molecular dynamics simulations with locally enhanced sampling of molecular oxygen which enables the investigations of gas diffusion pathways and protein barrel fluctuations on both mOrange and mOrange2 at the molecular level. Moreover, the presence of an oxygen molecule can modify the environment in a way that might allow the oxygen to enter the protein barrel. The explicit

oxygen simulations carried out in the present investigations are able to capture these possibilities.

3.3.2 MOLECULAR OXYGEN DIFFUSION PATHWAYS

For enhanced search statistics, our explicit oxygen calculation employed 15 copies of oxygen in our NAMD LES calculations. The oxygen molecules do not interact with each other, but interact with the rest of the system. With these 15 copies of non-interacting oxygen molecules in the system, we performed an 80 ns MD production run. The simulations reveal several different types of events of molecular oxygen entry into the protein barrel from various locations, which are pictured in Figure 4.2 and described below. We observed several protein pockets far from the chromophore where the molecular oxygen can enter and remain for an extended time. Some of these pockets are dead-ends with no access to the chromophore, but some of the pockets join to form pathways that ultimately lead to the chromophore.

Diffusion of Oxygen into Pockets Common to mOrange and mOrange2

There are four oxygen-hosting pockets common to both mOrange and mOrange2 as shown in Figure 3.4. Oxygen may enter through the β7 – β10 gap or through the bottom of the barrel. After 10 ns of equilibration time, the first oxygen entered into mOrange from the bottom at 3 ns, and the first oxygen entered into mOrange2 at 13.4 ns from the bottom. There are two pockets at the bottom of the barrel, which can host the oxygen entering from the bottom.

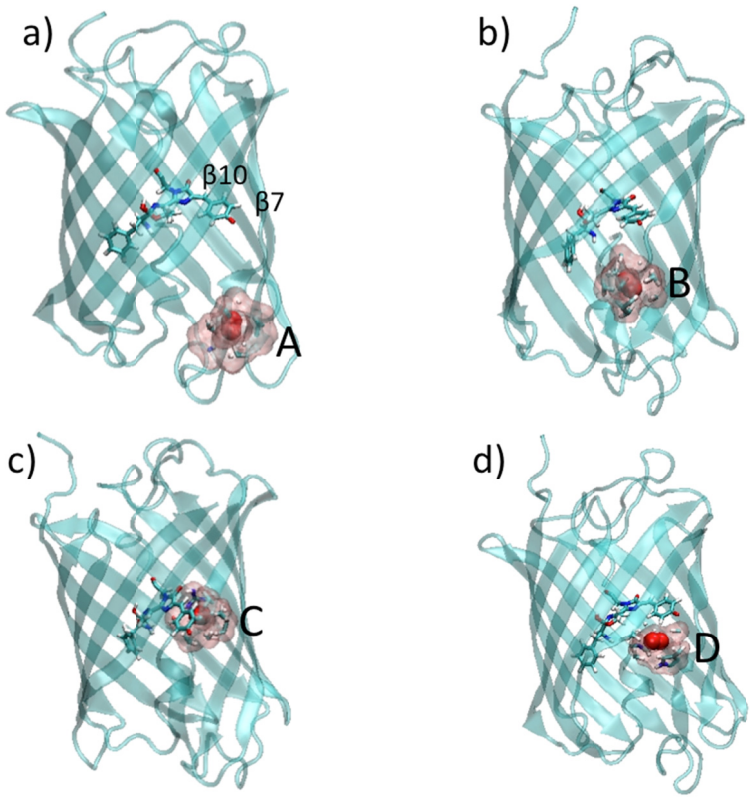


Figure 3.4 Various oxygen-hosting pockets common to mOrange and mOrange2. The largest gap between the beta strands is between β7 and β10, as shown in Figure 3.4(a).

Pocket A is at the bottom right of the protein in fig. 3.4(a), close to β7, β8 and β9, and is an entry point from the solvent. Pocket A is formed by the residues Val22, Ile60, Val105, Leu124, Phe129, Val135, Met136, Leu125, Leu167 and Tyr73. Oxygen entering in this pocket has to then pass through pocket B (fig. 3.4b) to get to the chromophore. The residues Val22, Phe27, Phe56, Leu61, Val 105, Leu124 and Ile60 form pocket B, which is located at the rear side, near the bottom of the barrel. The residues Val22, Ile60, Val105 and Leu124 form the interface between these two pockets. An oxygen molecule can move between pockets A and B multiple times. Pocket B leads to pocket C (fig. 3.4c), which is right below the chromophore. Pocket C is formed by the residues Pro63,

Arg95, Ile161, Met163, and Val177. Pocket C is close to the center of the chromophore. The fourth pocket D (fig. 3.4d) is close to pocket C, and allows entry to the interior through the β7 – β10 gap. Pocket D is formed by the residues Ser62, Pro63, Trp143, Ile161, Met163, Leu199 and Leu201. The residues Pro63, Ile161 and Met163 lie at the interface between the pockets C and D. The oxygen molecule can move between the pockets C and D.

In order to investigate oxygen diffusion into the barrel in more detail, we cut out uninteresting computational time at the beginning of the simulation during which the oxygen molecule moved around in the solvent outside the protein. In order to avoid the computational wait-time for the oxygen in the solvent to get to the gate point between β7−β10, we used 20 independent simulations, each with just one oxygen molecule placed at the β7−β10 gate. This expedited the search for pathways into the protein through the β7−β10 gate. The location of the oxygen molecule in the initial frame has to be made carefully. The oxygen molecule needs to be surrounded by a hydrophobic pocket.

A simulation was terminated if the oxygen molecule drifted out to the solvent. We considered the oxygen molecule to have completely escaped and terminated the simulation if the oxygen' s distance from the chromophore' s phenolate oxygen exceeded 15 Å. This distance ensures that a simulation will not be terminated if the oxygen has not truly drifted away from the β7−β10 gate, and an oxygen that remains in the vicinity just outside the pocket will be given time to reenter. In 1 out of 20 simulations in mOrange, the oxygen molecule escaped immediately into the solvent in less than 0.1 ns. In 8 out of the 20 simulations, the oxygen molecule entered pocket D and then exited out of the

barrel through the β7−β10 gap. In 10 out of the 20 simulations, the oxygen molecule managed to diffuse further into pocket C via pocket D. In the remaining one simulation in mOrange, the oxygen molecule escaped into pocket E, which is a unique pocket to mOrange as shown in Figure 3.7. For mOrange2, in 14 out of the 20 simulations, the oxygen molecule escaped immediately into the solvent in less than 0.2 ns. In the remaining 6 simulations, the oxygen molecule diffused into pocket C via pocket D. The fate of the oxygen molecule in the above 20 independent simulations in mOrange and mOrange2 shows that mOrange2 is less permeable to oxygen diffusion.

In another oxygen sensitivity test, an oxygen molecule is placed in pocket B in both mOrange and mOrange2, and 20 independent simulations are performed. For mOrange, in all 20 simulations, the oxygen molecule escaped out of the bottom of the barrel into the solvent in less than 6 ns. For mOrange2, in 8 out of 20 simulations, the oxygen molecule escaped from the bottom of the barrel into the solvent in less than 6 ns. In 7 out of 20 simulations in mOrange2, the oxygen molecule stayed inside the barrel for longer than 10 ns. The oxygen molecule was still inside the barrel in one simulation at the end of 30 ns. The survival curve showing the percentage of simulations in which the oxygen molecule is inside the barrel against the simulation time is given in Figure 3.5. In both mOrange and mOrange2, the oxygen molecule mostly oscillates between pockets A and B, and never reaches close to the chromophore. In 5 out of 20 runs in mOrange, the oxygen molecule exits through the bottom passing very close to the F99Y mutation within 6 ns of the simulation time. However in mOrange2, only in 1 out of 20 runs, the oxygen exits passing very close to the F99Y mutation within 6 ns of the simulation time.

Mean values were calculated through exponential fitting of the curves in Figure 3.5. The mean survival time for oxygen in mOrange is 3.2 ns and in mOrange2 is 10 ns. This shows that an oxygen molecule inside the bottom of the barrel in mOrange has higher probability of diffusing into the solvent than that in mOrange2. The same idea can be extended to oxygen diffusion from the bulk solvent into the protein barrel. This shows that mOrange2 is less permeable to molecular oxygen than mOrange.

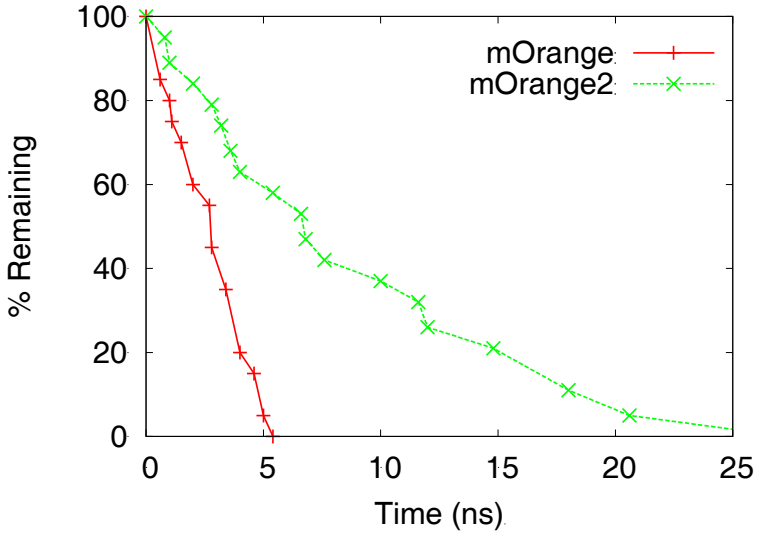


Figure 3.5 Survival curves for mOrange and mOrange2 showing the percentage of simulation runs in which the oxygen molecule is inside the barrel are plotted against the simulation time.

Trajectories that manifest the transition between the four pockets A, B, C and D are shown in Figure 3.6, which displays the distance of the oxygen molecule from a reference point on the chromophore. I chose nitrogen N3 of the imidazolinone ring as the reference point because this part of the chromophore is least flexible.

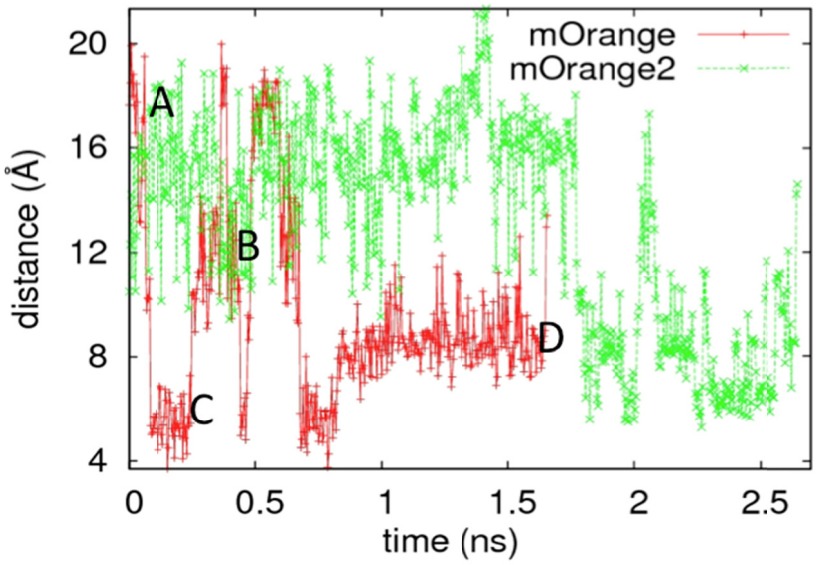


Figure 3.6 Trajectories of molecular oxygen showing its distance from the chromophore for a 2.7 ns time window. All four pockets A, B, C and D are visited in each trajectory. In the red trajectory, which represents mOrange, an oxygen molecule enters from pocket B, goes to pocket C, and then back to pocket B, then oscillates between pockets A and B. Between times 0.8 to 1.6 ns, the oxygen molecule spends a considerable amount of time in pocket D, and then exits to the solvent through the β7 – β10 gap. In the green trajectory, for mOrange2, an oxygen molecule enters through pocket A, oscillates between pockets A and B, then moves to pocket C, then oscillates between pockets C and D. The oxygen molecule finally exits through pocket D. The starting point time of 0 (ns) in the figure corresponds to 27 ns into the simulation for the red trajectory and 60 ns for the green trajectory.

In an 80 ns simulation, including the 10 ns equilibration time, a total of 13 oxygen molecule entries into the barrel were observed for mOrange, and 14 oxygen molecule entries for mOrange2. For mOrange, four entries through the β7 – β10 gap, two entries from the top, and seven entries from the bottom of the barrel were found. For mOrange2, eight entries from the β7 – β10 gap, 0 entries from the top, and six entries from the

bottom of the barrel were found. The number of oxygen entries into the barrel does not seem to explain the oxygen sensitivities. However two unique pockets were found in mOrange that were not present in mOrange2. The unique pockets could be the crucial sites for fluorescence quenching in mOrange.

Diffusion of Oxygen into Pockets Unique to mOrange

Two unique pockets for oxygen access were found in mOrange as shown in Figure 3.7. Pocket E is located to the left, on the lateral side of the chromophore. The oxygen enters into this pocket at 35 ns through the β7 – β10 gap, and stays there for about 8 ns. Various residues including Phe14, Val16, Glu32, Gln42, Ala44, Leu46, Tyr120 and Tyr214 form pocket E. The oxygen enters into this pocket from a point very close to the G196D mutation site on β10. This is one of the four mutations necessary to obtain mOrange2. The glycine at 196 seems to provide an unhindered pathway to the oxygen in mOrange. Pocket F hosts oxygen entering into the barrel from the top. Pocket F is formed by the residues Phe91, Trp93, Gln109, Asp110 and Ser111. There are two oxygen entries into this pocket at 50 ns and at 67 ns, and both of them remain in this pocket for 2 ns.

Both of these unique pockets may be important locations at which the oxygen may chemically modify the chromophore, eventually causing fluorescence quenching.

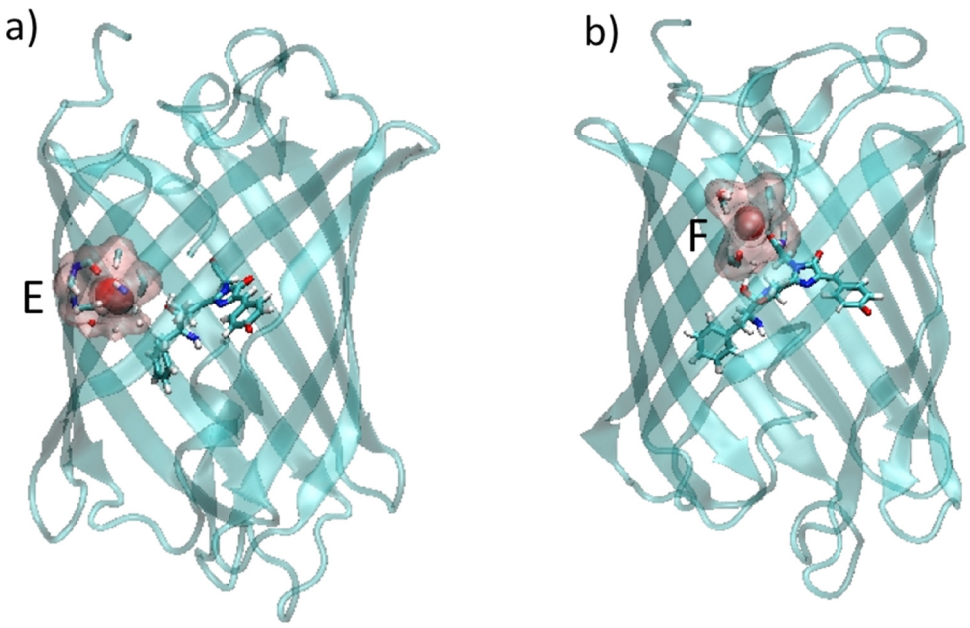


Figure 3.7 Oxygen-hosting pockets unique to mOrange.

Oxygen Diffusion Channels Connecting Multiple Pockets

A complete passage (channel) for the oxygen entering from the bottom and exiting through the β7 – β10 gap is observed for both mOrange and mOrange2. The channel passes through the four oxygen-hosting pockets A, B, C and D as shown in Figure 3.8. An oxygen molecule entering from pocket A or B, first passes through pocket C, then gets to D and then may finally exit into the solvent. The oxygen molecule can oscillate between the pockets A and B, B and C, or C and D only. This pathway is a complete channel that can bring the oxygen molecules very close to the chromophore, which may cause the decrease in the photostability of the FPs.

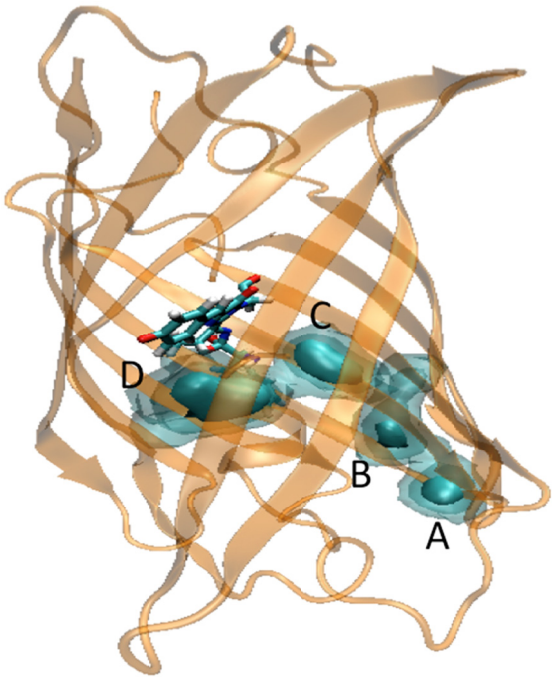


Figure 3.8 Probability density map for the molecular oxygen. The probability density map shows a clear channel for the diffusion of oxygen from the bottom of the protein to bulk solvent through the β7 – β10 gap. Three different probability density isosurfaces are depicted with isovalues 0.0241, 0.0091 and 0.0025.

4. PRESSURE EFFECTS ON mCHERRY, mSTRAWBERRY, AND CITRINE

Fluorescent proteins have significant and widespread applications in cell tagging and imaging biological cells and tissues (64). Red fluorescent proteins are preferred to green fluorescent proteins for imaging mammalian tissue because signal masking cellular autofluorescence is low in the red region of spectrum, and also because higher wavelengths can probe deeper into the tissues (19, 65, 66). We aim to explore the pressure response of two monomeric red fluorescent proteins, mCherry and mStrawberry, derived from their tetrameric progenitor DsRed. High pressure experiments performed on Blue Fluorescent Proteins show enhanced fluorescence up to a pressure of 600 Mpa (67). The restriction in chromophore vibrations at increased hydrostatic pressure contributes to increased stability of the hydrogen bonding network in the chromophore cavity and thus improves the quantum yield of these blue fluorescent proteins. In addition, chromophore planarity is believed to be an important factor in fluorescence. The chromophore of a non-fluorescent protein asFP595-A143G, also known as Kindling Fluorescent Protein (KFP), is significantly less planar (68) than the chromophore of a far red fluorescent protein eqFP611 which is highly fluorescent (69), although the chromophore is in the trans-conformation which is generally non-fluorescent in many proteins.

Recent high pressure experimental studies on the red fluorescent proteins mCherry and mStrawberry show that these proteins respond very differently to hydrostatic pressure (70), despite structural similarities in their protein barrels and in their chromophores. The difference response to pressure of these proteins are shown in Figure 4.1 (70). The fluorescence intensity of mStrawberry increases with the application of moderate pressure and decreases at higher pressure. In contrast, mCherry fluorescence

intensity shows a monotonic decrease with increasing pressure. Subsequent theoretical and computational investigations (71) showed that mStrawberry has a much floppier structure than mCherry at atmospheric pressure. Investigations were also performed involving ground state and excited state calculations on the two conformers of mStrawberry, and the results displayed a blue shift in both conformers at higher pressure. The blue shift is attributed to the change in distance between the chromophore and a neighboring residue Lys70, which is necessary for chromophore maturation in RFPs (72).

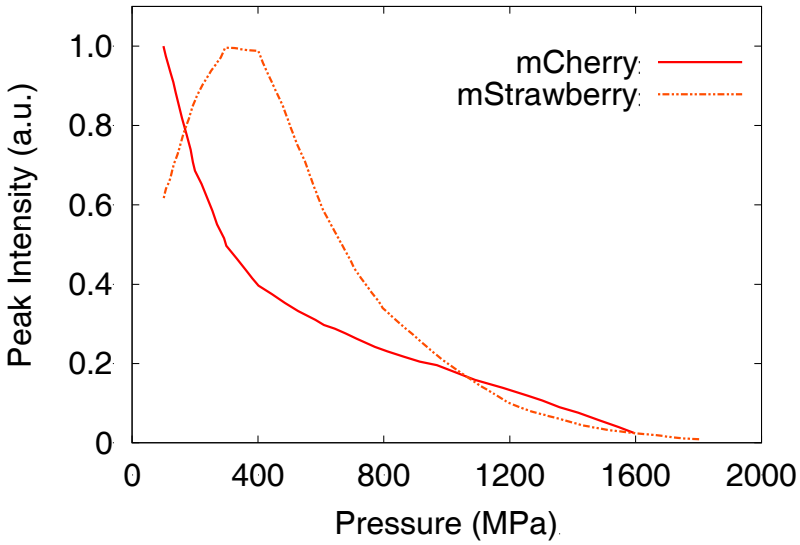


Figure 4.1 Experimental results of pressure effects on mCherry and mStrawberry. This figure is adopted from reference (70).

I have performed MD simulations and additional analyses that shed further light on the contrasting pressure response of these RFPs with respect to their structural and spectral behavior. In addition, we performed MD simulations on the pressure effects on a related monomeric yellow fluorescent protein, citrine. As in mStrawberry, the peak fluorescence intensity of citrine increases with the application of moderate (~ ambient)

pressures but decreases at higher pressures as shown in Figure 4.2 (73, 74). The effects of hydrostatic pressure on the structure of the protein barrel, the chromophore, and on the hydrogen bonding network inside the protein cavity of mCherry, mStrawberry and citrine is discussed in the following sections.

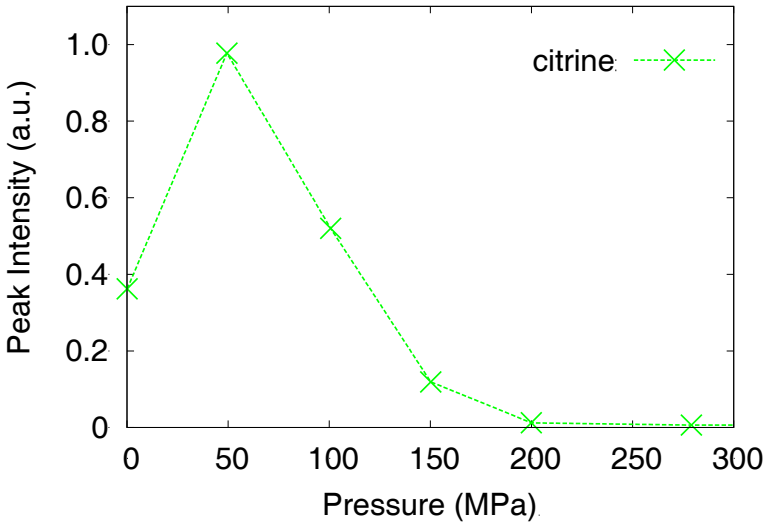


Figure 4.2 Experimental results of pressure effects on citrine. This figure is adopted from reference (74).

4.1 BARREL COMPRESSIBILITY

The compression caused by increased hydrostatic pressure was investigated by calculating the radius of gyration (rgyr) of the protein barrel at various pressures, as shown in table 4.1. Both mCherry and mStrawberry show decreases in the radius of gyration at increasing pressures. mStrawberry shows a small increase (0.01 Å) in rgyr at 500 MPa compared to that at 250 MPa. This could be because of the fluctuations in the loop regions of the fluorescent protein. The radius of gyration for citrine is slightly higher at 50 MPa than that at atmospheric pressure. The standard deviations in rgyr of citrine are observed to be higher (~ 0.058 Å) at pressures 50 MPa and 100 MPa, compared to those

(~0.040 Å) at higher pressures. In general, the changes in rgyr are on a sub-angstrom scale.

Table 4.1 Radius of gyration in angstroms at various pressures.

Pressure (MPa)	**rgyr (mCherry)**	**rgyr (mStrawberry)**	**rgyr (Citrine)**
0.1	16.95	17.07	17.09
50			17.25
100			17.08
250	16.85	16.83	17.00
500	16.65	16.84	16.89
750	16.61	16.71	16.80
1000	16.54	16.66	

The compressibility (β) is calculated by using the relation,

$$\beta = -\frac{1}{V}\left(\frac{\Delta V}{\Delta P}\right) \tag{4.1}$$

In equation 4.1, V is the volume, ΔV is the change in volume, and ΔP is the change in pressure of the fluorescent protein.

Approximating the protein barrel as a sphere with the calculated radius of gyration, the average compressibility for mCherry is 8.11 x 10^{-5} MPa^{-1}, for mStrawberry it is 9.98 x 10^{-5} MPa^{-1}, and for citrine it is 5.42 x 10^{-5} MPa^{-1}. The value of rgyr at 50 MPa is not considered for compressibility calculations in citrine, as this shows an increase in volume at higher pressure. These results show that mStrawberry is floppier than mCherry, which is consistent with the computational results of Laurent et. al. (71). Also, citrine is the least compressible. The gap between the β7 and β10 strands in citrine is

smaller than the gaps between the β7 and β10 strands in mCherry and mStrawberry. This could be a cause for the small compressibility of citrine. The compressibility values are also comparable to the experimental data obtained from the measurement of the adiabatic compressibility of globular proteins using sound velocity measurements at room temperature (75).

I analyzed the squeezing effects of increased hydrostatic pressure on the tunnel and cavities inside the protein barrel of mCherry and mStrawberry. Caver software (54) was used to identify tunnels that start from the chromophore and lead to exits outside the barrel, with a probe radius of 0.8 Å. The total number of tunnels and tunnel clusters are shown in table 4.2 for mCherry, in table 4.3 for mStrawberry, and in table 4.4 for citrine. mCherry is observed to have the most channels at pressures upto 500 MPa. This shows that mCherry has relatively more empty space inside the barrel, which also explains the higher rmsf fluctuations of mCherry as compared to mStrawberry and citrine. mStrawberry has a significantly fewer tunnels than mCherry at the pressure 250 MPa. This makes the chromophore more stable and thus able to better fluoresce at 250 MPa. However at the pressure of 500 MPa, the number of tunnels in mStrawberry is reduced to very few, and the chromophore may be distorted. The consequence is the decrease in the fluorescent intensity peak of mStrawberry at 500 MPa pressure. For citrine, it is interesting to note that the number of tunnels inside the barrel reduces significantly even at a relatively lower pressure of 100 MPa.

Tunnels inside the protein barrel are depicted for mCherry and mStrawberry at a pressure of 250 MPa, and for citrine at a pressure of 50 MPa in Figure 4.3. Similar tunnels are colored the same way. Although mCherry has significantly more tunnels than

mStrawberry at 250 MPa, the number of tunnel clusters in both RFPs is the same, and this is also represented in Figure 4.3. Citrine contains the greatest number of tunnel clusters at 50 MPa. There are more tunnel clusters leading to the bottom of the barrel in citrine than in mCherry and mStrawberry.

Table 4.2 Tunnels inside mCherry

Pressure (MPa)	Number of tunnel clusters	Number of tunnels
0.1	33	292
250	10	208
500	3	74
750	4	13
1000	1	35

Table 4.3 Tunnels inside mStrawberry

Pressure (MPa)	Number of tunnel clusters	Number of tunnels
0.1	29	191
250	10	123
500	3	7
750	1	32
1000	1	55

Table 4.4 Tunnels inside citrine

Pressure (MPa)	Number of tunnel clusters	Number of tunnels
0.1	10	144
50	25	83
100	9	22
250	3	3
500	1	1
750	1	1

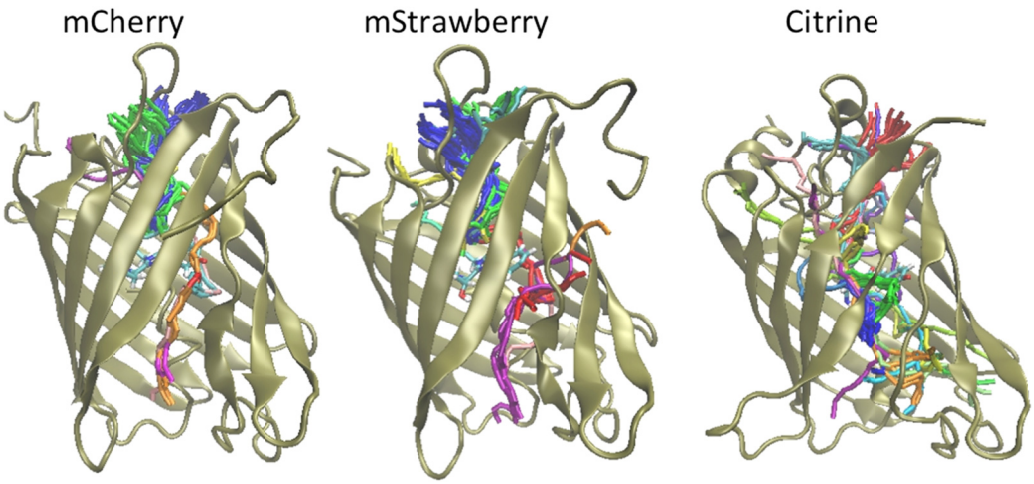


Figure 4.3 Tunnels in mCherry and mStrawberry plotted at pressures 250 MPa, and tunnels in citrine plotted at pressure 50 MPa. The fluorescence intensity of mStrawberry and citrine are maximum at 250 MPa and 50 MPa pressures respectively.

4.2 CHROMOPHORE-BARREL INTERACTION AND HYDROGEN BONDING

An intricate network of hydrogen bonds and interactions between the buried charges stabilizes the chromophore. Disruption in the hydrogen bond network can destabilize the chromophore and lead to the quenching of fluorescence. The hydrogen bonding network inside each FP is shown in Figure 4.4. The dotted lines represent the hydrogen bonds, and the two connected rings at the center represent the chromophore. The phenolate oxygen of the chromophore is hydrogen bonded to Ser146 Oγ in mCherry and mStrawberry, and to His148 Nδ in citrine. The carbonyl oxygen of the imidazolinone ring interacts with the closest arginine residue in all three FPs. In mCherry and mStrawberry, Glu215 is protonated and forms a hydrogen bond with the nitrogen in the imidazole ring and with water. However, in citrine, Glu222 is not protonated, and it forms hydrogen bonds with water only. In addition, we can also notice the position of

Tyr203 in citrine, which allows the stacking of it phenol ring against the phenol ring of the chromophore.

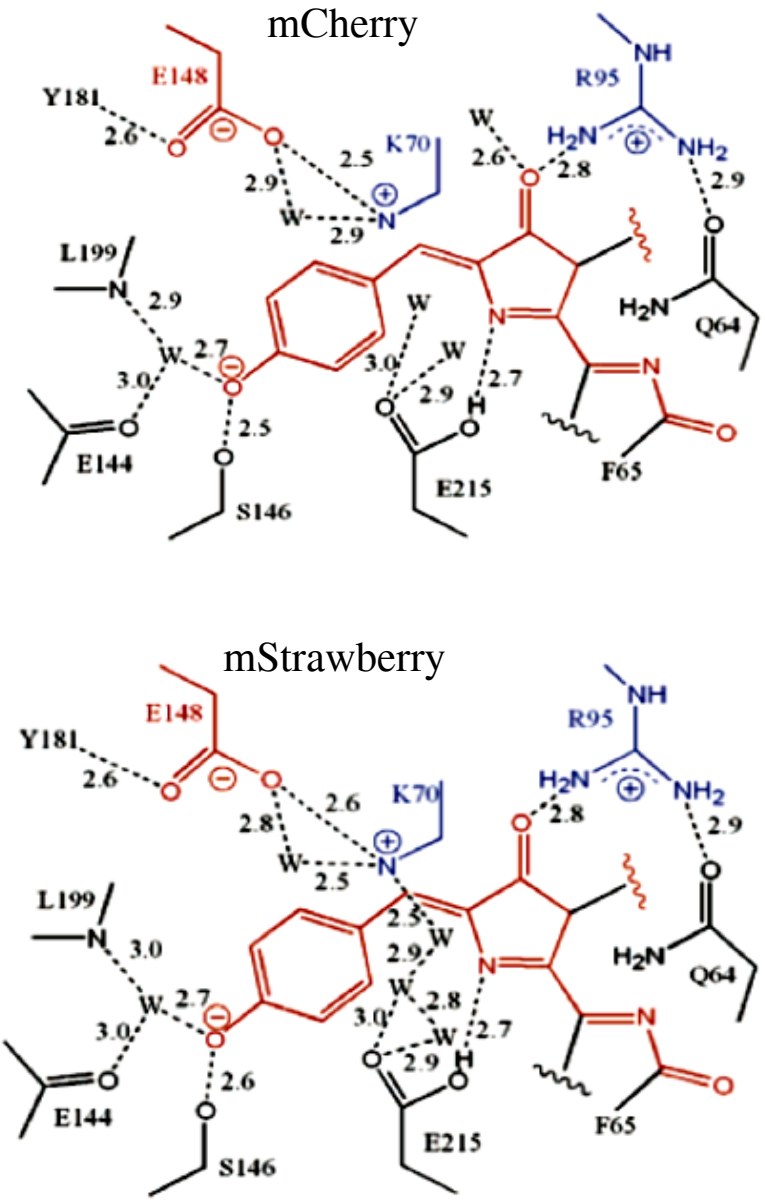


Figure 4.4(a) The hydrogen bonding network (34) in mCherry and mStrawberry. The dotted lines represent the hydrogen bonds.

Yellow Fluorescent Protein

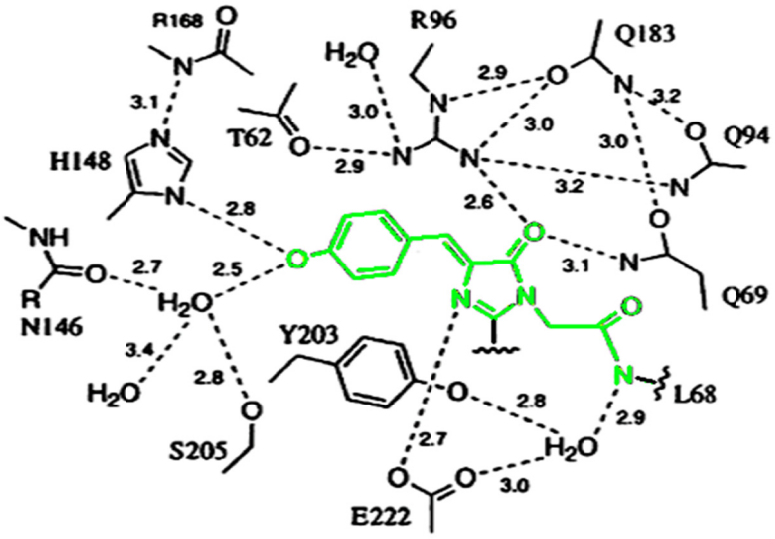


Figure 4.4(b) The hydrogen bonding network (59) in Yellow Fluorescent Protein. Citrine exhibits a hydrogen-bonding pattern very similar to YFP. The dotted lines represent the hydrogen bonds.

One end of the chromophore is connected to the central helix, and the other end (phenolic end) is stabilized by hydrogen bonding. The hydrogen bond between the chromophore tip (oxygen at the phenolic end), and the nearest residue is considered very important, as the chromophore tip acts as an important anchor point. The criteria adopted for the existence of a hydrogen bond are, the distance between the donor and acceptor atom should be less than 3 angstroms, and the hydrogen bond cutoff angle should be less than 30 degrees.

The histogram of the distance between the phenolate oxygen and Ser146 Oγ is shown in Figure 4.5(a, b) for mCherry and mStrawberry. The Ser146 adopts two different conformations – one in which it stays closer to the chromophore and forms a hydrogen bond, and the other in which it flips away from the chromophore far enough to break the

hydrogen bond. The mStrawberry chromophore is predominantly hydrogen bonded to the serine. But in mCherry, the histogram shows an appreciable hump on the right side at higher pressures representing the flipping of the serine residue away from the chromophore. Without the hydrogen bond to the serine, the chromophore is less stable. This effect may be the cause of diminished the fluorescence intensity of mCherry with increasing pressure. This mechanism also sheds light on the importance of the interaction between serine and the phenolate oxygen for forming a stable and fluorescent chromophore. However, at very high pressure (1000 MPa) the chromophore structure might be highly compressed and stiffened so that the Ser146 is in the immediate vicinity of the chromophore, allowing the formation of a hydrogen bond with the phenolate oxygen. This is depicted by the single peak of the histogram at 1000 MPa for both mCherry and mStrawberry. There are several water molecules around the chromophore that form two hydrogen bonds with the phenolate oxygen when the Ser146 Oγ is beyond the closest approach for forming hydrogen bonds.

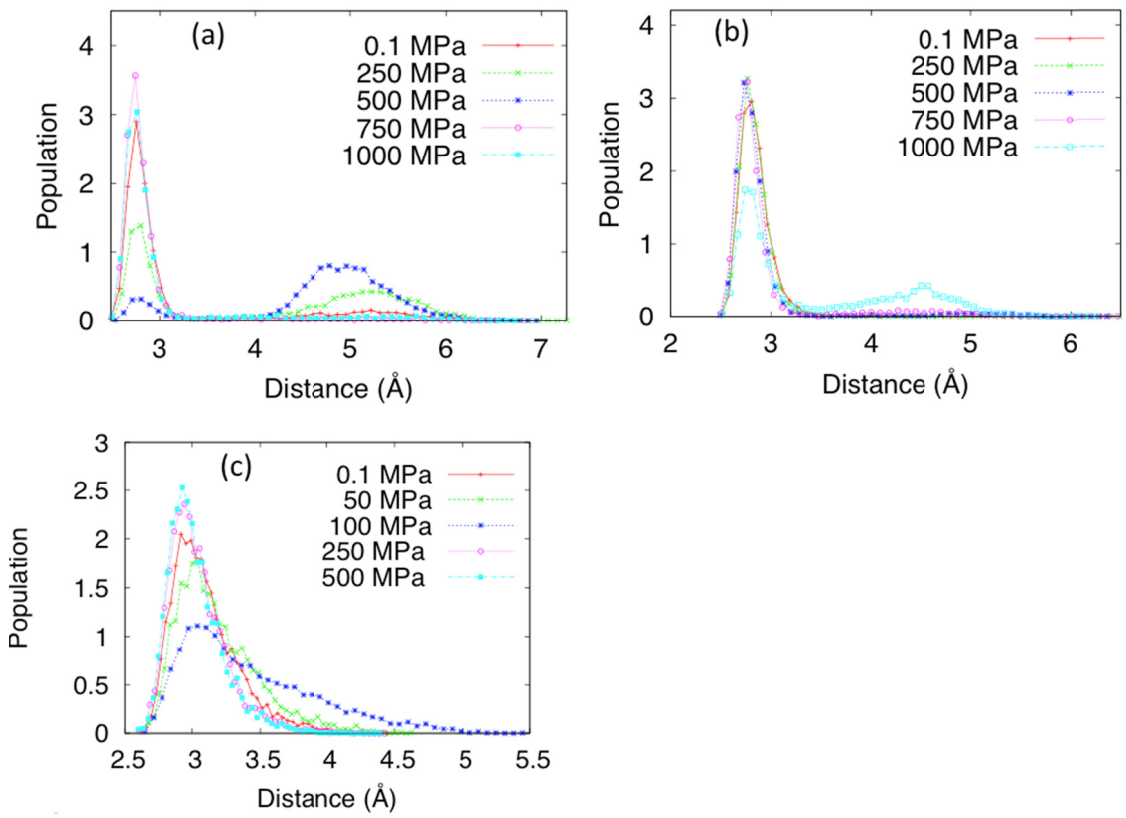


Figure 4.5. Histogram of the distance between the phenolate oxygen and Ser146 Oγ in (a) mCherry and (b) mStrawberry, (c) distance between phenolate oxygen and His148 Nδ in citrine.

In citrine, the tip of the chromophore at the phenolate terminal is stabilized by forming hydrogen bonds with His148 Nδ and/or water. The histogram of the distance between phenolate oxygen and His148 Nδ in citrine is shown in Figure 4.5(c). The peaks at pressures 0.1 MPa and 50 MPa look similar. The histogram at 100 MPa pressure is relatively more spread-out and flat, which may be the reason for the onset of a decrease in fluorescence intensity at this pressure. At pressures above 100 MPa, the barrel is highly squeezed, decreasing the distance between the phenolate oxygen and His148 Nδ in citrine.

We have also examined the percentage of hydrogen bonds formed between Ser146 Oγ and the phenolate oxygen (in mCherry and mStrawberry) as shown in Figure 4.6(a), and between His148 Nδ and the phenolate oxygen (in citrine) as shown in Figure 4.6(b). For mCherry the population of hydrogen bonds decreases significantly at 250 MPa and 500 MPa. Hence the chromophore is flexible and the fluorescent intensity decreases in mCherry. At pressures above 500 MPa the percentage of hydrogen bonds starts to increase as the barrel is highly squeezed. For mStrawberry, the percentage of H-Bonds between Ser-146 and the tip of the chromophore keeps increasing at all higher pressures. Consequently the fluorescence intensity of mStrawberry increases at 250 MPa.

For citrine, the percentage of hydrogen bonds remains similar at pressures 0.1 MPa and 50 MPa. However, the percentage of hydrogen bonds greatly decreases at 100 MPa, which explains the decrease in fluorescence intensity at this pressure. At pressures higher than 100 MPa, the protein barrel is squeezed and the percentage of hydrogen bonding increases. At even higher pressures, the pi-stacking interactions between the phenol rings of the chromophore and Tyr 203 are highly perturbed (explained in section 4.4). Hence, the stability imparted by the increment in hydrogen bonding is not helpful for increasing the fluorescent intensity at pressures above 50 MPa.

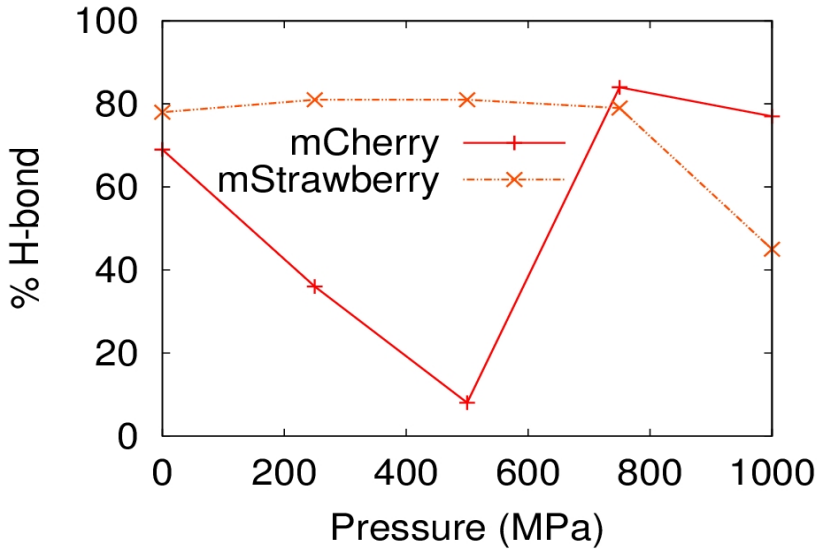


Figure 4.6(a) Percentage of hydrogen bonds formed between Ser146 Oγ and the chromophore phenolate oxygen in mCherry and mStrawberry

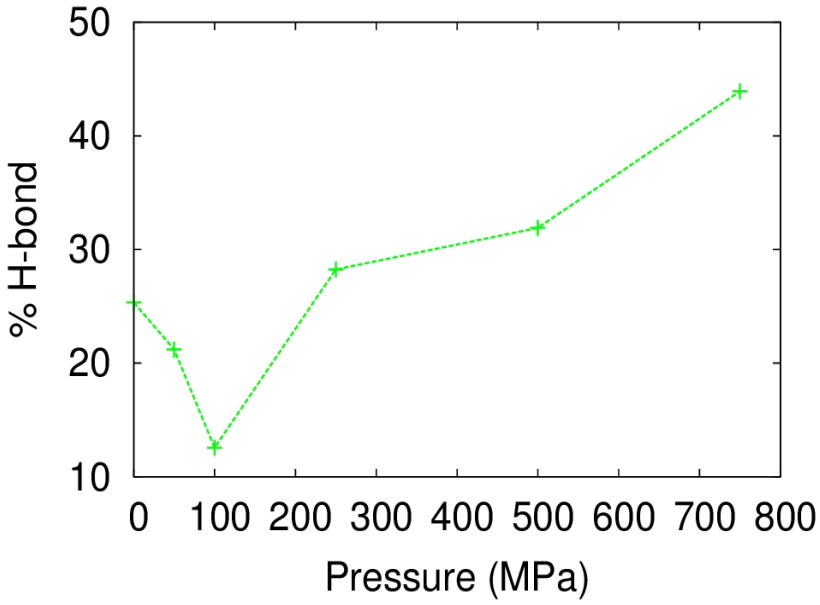


Figure 4.6(b) Percentage of hydrogen bonds formed between His148 Nδ and chromophore phenolate oxygen in citrine.

The percentage of hydrogen bonds formed by oxygen at the phenolate ring, and carbonyl oxygen and nitrogen at the imidazolinone ring is shown in table 4.5 for mCherry, in table 4.6 for mStrawberry, and in table 4.7 for citrine. The general trend is increase in hydrogen bonding at higher pressures. The percentage of hydrogen bonding is comparatively high for the interaction between Arg95 N and the chromophore, because it offers two donor sites for the formation a hydrogen bond. Other residues offer only a single donor site. In citrine, the population of hydrogen bonds between Arg96 N and the chromophore increases considerably at 50 MPa, while that between His148 Nδ and the chromophore remains almost unchanged as seen in table 4.7. This could be partially responsible for the maximized fluorescence peak intensity of citrine at 50 MPa.

Table 4.5 Hydrogen bond percentage of mCherry chromophore at various pressures.

H-Bond Type	**0.1 MPa**	**250 MPa**	**500 MPa**	**750 MPa**	**1000 MPa**
Ser146 Oγ - Cro	69	36	8	84	77
Arg95 N - Cro	86	89	58	88	93
Glu215 Oε - Cro	47	59	68	68	71

Table 4.6 Hydrogen bond percentage of mStrawberry chromophore at various pressures.

H-Bond Type	**0.1 MPa**	**250 MPa**	**500 MPa**	**750 MPa**	**1000 MPa**
Ser146 Oγ - Cro	78	81	81	79	45
Arg95 N - Cro	77	86	91	88	82
Glu215 Oε - Cro	25	36	40	51	50

Table 4.7 Hydrogen bond percentage of citrine chromophore at various pressures.

H-Bond Type	0.1 MPa	50 MPa	100 MPa	250 MPa	500 MPa	750 MPa
His148 Nδ - Cro	25	21	13	28	32	44
Arg96 N - Cro	65	76	62	45	57	55

4.3 CHROMOPHORE FLEXIBILITY

As with other FPs, the chromophores of both mCherry and mStrawberry are formed by cyclization and oxidation of peptide sequences; the mCherry chromophore is formed by Met-Tyr-Gly and the mStrawberry chromophore is formed by Thr-Tyr-Gly. Although Tyr and Gly are common to both tripeptides, Met in mCherry is replaced by Thr in mStrawberry. The mStrawberry chromophore constituent Thr is less bulky than Met in mCherry. Therefore, the mStrawberry chromophore has more conformational degrees of freedom at ambient pressure conditions. When a moderate pressure is applied this conformational freedom is reduced and the chromophore is more stabilized. The improved chromophore rigidity at moderate pressure (~250 MPa) results in a better fluorescence in mStrawberry. The overall effect of pressure in the conformational freedom of the chromophore is shown in Figure 4.7 in terms of root mean square fluctuations (rmsf). For mStrawberry, the chromophore rmsf decreases with increasing pressure, which explains the improved intensity at 250 MPa. But mCherry shows higher rmsf fluctuations at all pressures, which implies that the chromophore is more mobile and the excited state chromophore is more likely to follow the path of radiationless decay through vibrational relaxations. At very high pressures of 750 MPa and 1000 MPa, the chromophore structure may be distorted, resulting in the loss of fluorescent intensity.

The chromophore of the yellow fluorescent protein citrine shows a slight increase in rmsf at moderate pressures (50 MPa and 100 MPa) as shown in Figure 4.7. This seems contrary to the fact that the peak intensity of citrine is maximum at 50 MPa. However the increased interaction between the chromophore and the phenol ring of Tyr203 may be a dominant cause of the peak fluorescence intensity in citrine at 50 MPa (explained in section 4.4). At pressures higher than 100 MPa, the citrine chromophore shows a monotonic decrease in rmsf, similar to mStrawberry.

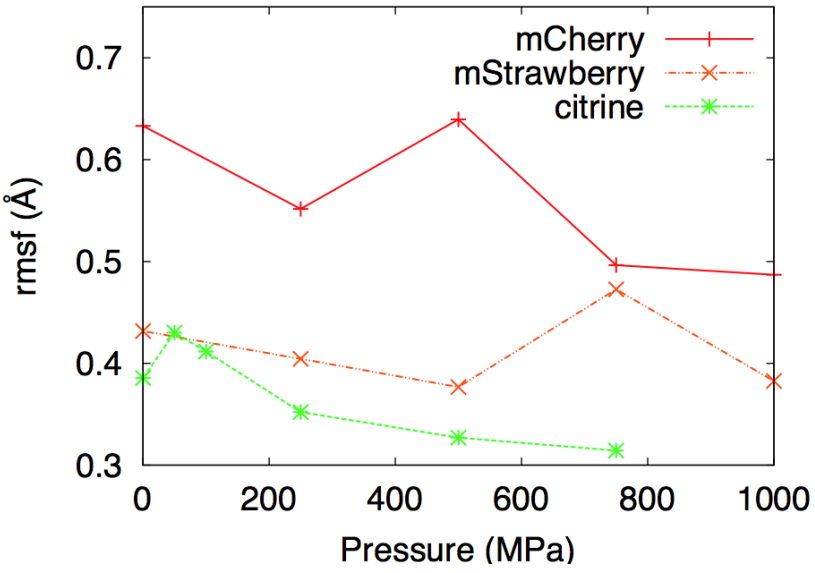


Figure 4.7 Plot of root mean square fluctuations (rmsf) of the chromophore of mCherry, mStrawberry and citrine at various pressures.

4.4 CHROMOPHORE PLANARITY AND THE FLUORESCENCE

In order to investigate the pressure effects on chromophore planarity, best-fit planes were constructed for the two rings of the chromophore using the CHARMM trajectory analysis package. To find the least square planes, I have considered atoms C1, C2, N2, N3, O2, CA1, CA2, CB2 and CA3 for imidazolinone, and atoms CB2, CG2,

CD1, CD2, CE1, CE2, CZ and OH for p-hydroxybenzyl. The atom names are exactly as given in the PDB files of the three FPs. The average dihedral angle between these two least squares planes is plotted against different pressures in Figure 4.8.

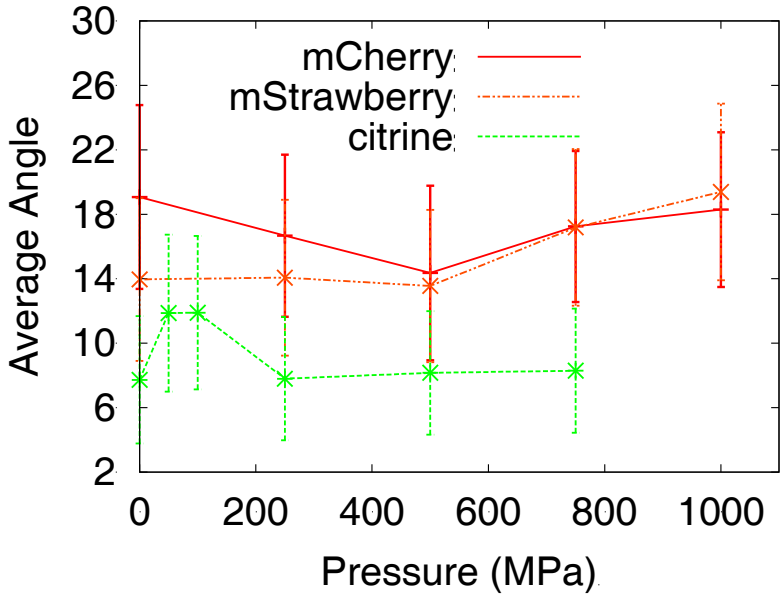


Figure 4.8 Plot of average angle between the planes of phenolate ring and imidazolinone ring of the chromophore of mCherry, mStrawberry, and citrine at various pressures.

At moderate pressures up to 500 MPa, mCherry shows an improvement in chromophore planarity while the mStrawberry chromophore maintains the same planarity. The total change in the angle between the planes of the two chromophore rings is about 6 degrees, and this effect may not be significant enough to alter the fluorescence intensity.

Interestingly, the average angle between the two chromophore rings in citrine increases at 50 MPa and 100 MPa. This may adversely affect the fluorescence intensity peak. However, the phenolate ring of the citrine chromophore interacts with the phenolate ring of neighboring residue Tyr203, and this phenomenon is not observed in mCherry and mStrawberry. In addition, the structure of citrine is more planar (less than 12 degrees) than that of mCherry or mStrawberry at all pressures. Also, at pressures above 100 MPa, the chromophore structure in citrine shows no change in planarity, compared to the monotonic decrease in chromophore planarity observed in mCherry and mStrawberry at equivalent pressures.

The YFP variant citrine stems from an important mutation T203Y in the green fluorescent protein family. The interaction between the chromophore and the phenolic ring of Tyr203, as shown in Figure 4.9(b) causes the fluorescence peak shift from green to yellow in citrine (59). Barstow and coworkers observed an initial increase of fluorescence peak intensity in citrine by applying moderate hydrostatic pressure, followed by a monotonic decrease in intensity at higher pressure. The maximum peak intensity was observed at around 50MPa of pressure. The study reported that the relative displacement between the phenolic ring of Tyr203 and the phenolic ring of the chromophore explains the observed pressure response to the peak fluorescent intensity (73, 74). I have investigated the effects of hydrostatic pressure on the dihedral angle between these phenolic rings. Best-fit planes were constructed for the chromophore phenolic ring and Tyr203 Phenolic ring, using the CHARMM trajectory analysis tool. The fitting involve atoms CB2, CG2, CD1, CD2, CE1, CE2, CZ and OH of the chromophore and CB, CG, CD1, CE1, CD2, CE2, CZ and OH of Tyr203. The average

dihedral angle between these two least squares planes is plotted against various pressures in Figure 4.9(a). As the average angle reaches a minimum at 50 MPa, this would result in favorable pi-stacking interactions between the chromophore and the phenol. At 50 MPa pressure, although the planarity of the chromophore itself decreases, the increased interaction between the chromophore and the neighboring phenol ring of Tyr203, helps increase the fluorescence intensity. Hence the fluorescence intensity peak is maximum at 50 MPa. At pressures beyond 50 MPa, the planarity between the two phenolate rings decreases and so does the peak intensity. This directly correlates with the experimentally observed changes in peak fluorescence intensity with the application of hydrostatic pressure in citrine.

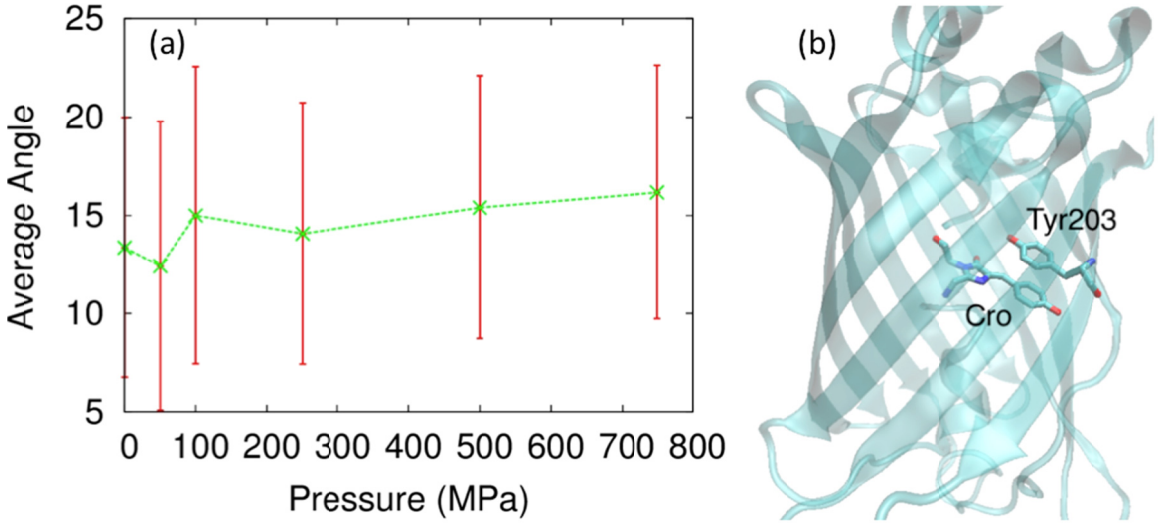


Figure 4.9 (a) Average angle between the least square panes through the phenol rings of chromophore and Tyr203 in citrine. (b) Stacking between phenolate rings of chromophore and neighboring residue Tyr203.

5. CONCLUSIONS

I investigated several structural aspects related to fluorescence in fluorescent proteins. I now summarize these results.

Access for molecular oxygen to the chromophore in fluorescent proteins is required for chromophore maturation. However, oxygen penetration through the barrel to the mature chromophore can also cause irreversible photo-bleaching and significantly reduce the photostabilty of an FP. In this work, I performed molecular dynamics simulations to investigate the diffusion of molecular oxygen into the protein barrel of the monomeric RFP variants mOrange and mOrange2. A clear channel for oxygen diffusion into the protein, common to both mOrange and mOrange2, is described. The pathway contains several oxygen hosting pockets, which are identified by the amino acid residues that form each pocket. One end of the channel is accessed from the solvent through the floppy β7−β10 gap, which leads immediately to a gateway pocket (D) that provides some access to the chromophore. Another entrance from the solvent is at the bottom of the protein barrel. Two unique oxygen-hosting pockets have been identified in mOrange, that are not present in mOrange2. In both pockets, oxygen spends a significant amount of time in the immediate vicinity of the chromophore. This explains the significantly lower photostability of mOrange as compared to mOrange2. Such computational identifications of oxygen diffusion pathways can be helpful in guiding mutagenesis efforts to design fluorescent proteins with improved photophysical properties.

A large chromophore like the tripeptide in fluorescent proteins needs to be fairly rigid to be able to fluoresce. Otherwise, the energy contained in the excited electronic orbital can dissipate through radiationless, vibrational modes. For example, the

chromophore isolated from the Aequorea green fluorescent protein was observed to fluoresce highly when present in a frozen solution, but was non-fluorescent at room temperature (76). In fluorescent proteins, rigidity of the chromophore can be attained by anchoring to the protein. The chromophore in the fluorescent protein eqFP611 interacts with the residue Phe174, but KFP lacks this interaction due to a different residue at this position and is non-fluorescent (77). Our computational results show that high pressures stabilize the mStrawberry chromophore more than the mCherry chromophore. The citrine chromophore is more stabilized at 50 MPa than at 0.1 MPa through hydrogen bonding with Arg96 N and His148 Nδ, which may be the cause of the increased fluorescence intensity in citrine at that pressure.

The chromophore interaction with neighboring residues and the movement of water molecules in the chromophore cavity seem to play a crucial role in imparting rigidity and stability to the chromophore. The Y66H/Y145F mutations of the green fluorescent protein result in the formation of blue fluorescent proteins. As a consequence of the smaller size of the histidine (H) as compared to tyrosine (Y), the chromophore in BFP cannot link to the hydrogen bonding network in the cavity (78). Mauring and his group showed that the fluorescence intensity of BFP can be increased by increasing pressure (67). This can be a result of squeezing of the cavity, which allows the formation of additional hydrogen bonds to the histidine, causing it to become more rigid. Similarly, in mStrawberry, with a small threonine residue in the cavity, the chromophore may have increased mobility at 1 atm. High pressure may squeeze the pocket enough to reduce the flexibility of the chromophore, and the increased rigidity at higher pressure improves the fluorescence intensity.

Pressure can be used as an effective tool to explore details of protein conformations. Fluorescent proteins are complex biopolymers with numerous degrees of freedom. It is difficult to elucidate the role of each bond in the chromophore cavity of these proteins, but pressure can be used to produce structural changes that allow a better understanding of the structural factors associated with their biological functions. We have investigated pressure effects on planarity and other interactions of the chromophore that contribute to its overall fluorescent behavior. Our investigations show that there is a complex interplay involving various interactions between the chromophore, the barrel, and the aqueous environment. All of these are responsible for the fluorescent behavior of the FPs. This approach of varying the pressure to determine important structure-function relationships is useful in designing and engineering novel fluorescent proteins with optimized photophysical properties.

REFERENCES

1. Griesbeck O. Fluorescent proteins as sensors for cellular functions. Current opinion in neurobiology. 2004;14(5):636-41. doi: 10.1016/j.conb.2004.08.002. PubMed PMID: 15464898.

2. Tsien RY. The green fluorescent protein. Annual review of biochemistry. 1998;67:509-44. doi: 10.1146/annurev.biochem.67.1.509. PubMed PMID: 9759496.

3. Zimmer M. Green fluorescent protein (GFP): applications, structure, and related photophysical behavior. Chemical reviews. 2002;102(3):759-81. PubMed PMID: 11890756.

4. Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. Green fluorescent protein as a marker for gene expression. Science. 1994;263(5148):802-5. PubMed PMID: 8303295.

5. Rizzo MA, Springer GH, Granada B, Piston DW. An improved cyan fluorescent protein variant useful for FRET. Nature biotechnology. 2004;22(4):445-9. doi: 10.1038/nbt945. PubMed PMID: 14990965.

6. Paci E, Marchi M. Intrinsic compressibility and volume compression in solvated proteins by molecular dynamics simulation at high pressure. Proceedings of the National Academy of Sciences of the United States of America. 1996;93(21):11609-14. PubMed PMID: 8876183; PubMed Central PMCID: PMC38105.

7. Silva JL, Foguel D, Royer CA. Pressure provides new insights into protein folding, dynamics and structure. Trends in biochemical sciences. 2001;26(10):612-8. PubMed PMID: 11590014.

8. Hillson N, Onuchic JN, Garcia AE. Pressure-induced protein-folding/unfolding kinetics. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(26):14848-53. PubMed PMID: 10611301; PubMed Central PMCID: PMC24736.

9. Boonyaratanakornkit BB, Park CB, Clark DS. Pressure effects on intra- and intermolecular interactions within proteins. Biochimica et biophysica acta. 2002;1595(1-2):235-49. PubMed PMID: 11983399.

10. Prehoda KE, Mooberry ES, Markley JL. Pressure denaturation of proteins: evaluation of compressibility effects. Biochemistry. 1998;37(17):5785-90. doi: 10.1021/bi980384u. PubMed PMID: 9558311.

11. Herberhold H, Marchal S, Lange R, Scheyhing CH, Vogel RF, Winter R. Characterization of the pressure-induced intermediate and unfolded state of red-shifted green fluorescent protein--a static and kinetic FTIR, UV/VIS and fluorescence spectroscopy study. Journal of molecular biology. 2003;330(5):1153-64. PubMed PMID: 12860135.

12. Mozhaev VV, Heremans K, Frank J, Masson P, Balny C. High pressure effects on protein structure and function. Proteins. 1996;24(1):81-91. doi: 10.1002/(SICI)1097-0134(199601)24:1<81::AID-PROT6>3.0.CO;2-R. PubMed PMID: 8628735.

13. Frauenfelder H. Proteins and pressure. J Phys Chem. 1990;94(3):1024–37. doi: 10.1021/j100366a002.

14. Panick G, Winter R. Pressure-induced unfolding/refolding of ribonuclease A: static and kinetic Fourier transform infrared spectroscopy study. Biochemistry. 2000;39(7):1862-9. PubMed PMID: 10677237.

15. Doster W, Gebhardt, R. High pressure – unfolding of myoglobin studied by dynamic neutron scattering. Chemical Physics. 2003;292:383–7. doi: 10.1016/S0301-0104(03)00064-8.

16. Heremans K. High pressure effects on proteins and other biomolecules. Annual review of biophysics and bioengineering. 1982;11:1-21. doi: 10.1146/annurev.bb.11.060182.000245. PubMed PMID: 7049058.

17. Collins MD, Kim CU, Gruner SM. High-pressure protein crystallography and NMR to explore protein conformations. Annual review of biophysics. 2011;40:81-98. doi: 10.1146/annurev-biophys-042910-155304. PubMed PMID: 21275639.

18. Macdonald AG. Experiments on ion channels at high pressure. Biochimica et biophysica acta. 2002;1595(1-2):387-9. PubMed PMID: 11983411.

19. Merzlyak EM, Goedhart J, Shcherbo D, Bulina ME, Shcheglov AS, Fradkov AF, et al. Bright monomeric red fluorescent protein with an extended fluorescence lifetime. Nature methods. 2007;4(7):555-7. doi: 10.1038/nmeth1062. PubMed PMID: 17572680.

20. Bevis BJ, Glick BS. Rapidly maturing variants of the Discosoma red fluorescent protein (DsRed). Nature biotechnology. 2002;20(1):83-7. doi: 10.1038/nbt0102-83. PubMed PMID: 11753367.

21. Campbell RE, Tour O, Palmer AE, Steinbach PA, Baird GS, Zacharias DA, et al. A monomeric red fluorescent protein. Proceedings of the National Academy of Sciences of the United States of America. 2002;99(12):7877-82. doi: 10.1073/pnas.082243699. PubMed PMID: 12060735; PubMed Central PMCID: PMC122988.

22. Amara P, Andreoletti P, Jouve HM, Field MJ. Ligand diffusion in the catalase from Proteus mirabilis: a molecular dynamics study. Protein science : a publication of the Protein Society. 2001;10(10):1927-35. doi: 10.1110/ps.14201. PubMed PMID: 11567083; PubMed Central PMCID: PMC2374231.

23. Bossa C, Anselmi M, Roccatano D, Amadei A, Vallone B, Brunori M, et al. Extended molecular dynamics simulation of the carbon monoxide migration in sperm whale myoglobin. Biophysical journal. 2004;86(6):3855-62. doi: 10.1529/biophysj.103.037432. PubMed PMID: 15189882; PubMed Central PMCID: PMC1304287.

24. Brunori M, Gibson QH. Cavities and packing defects in the structural dynamics of myoglobin. EMBO reports. 2001;2(8):674-9. doi: 10.1093/embo-reports/kve159. PubMed PMID: 11493595; PubMed Central PMCID: PMC1083996.

25. Carlson ML, Regan RM, Gibson QH. Distal cavity fluctuations in myoglobin: protein motion and ligand diffusion. Biochemistry. 1996;35(4):1125-36. doi: 10.1021/bi951767k. PubMed PMID: 8573567.

26. Gibson QH, Regan R, Elber R, Olson JS, Carver TE. Distal pocket residues affect picosecond ligand recombination in myoglobin. An experimental and molecular dynamics study of position 29 mutants. The Journal of biological chemistry. 1992;267(31):22022-34. PubMed PMID: 1429552.

27. Lamb DC, Arcovito A, Nienhaus K, Minkow O, Draghi F, Brunori M, et al. Structural dynamics of myoglobin: an infrared kinetic study of ligand migration in mutants YQR and YQRF. Biophysical chemistry. 2004;109(1):41-58. doi: 10.1016/j.bpc.2003.10.002. PubMed PMID: 15059658.

28. Cohen J, Kim K, King P, Seibert M, Schulten K. Finding gas diffusion pathways in proteins: application to O2 and H2 transport in CpI [FeFe]-hydrogenase and the role of packing defects. Structure. 2005;13(9):1321-9. doi: 10.1016/j.str.2005.05.013. PubMed PMID: 16154089.

29. Feher VA, Baldwin EP, Dahlquist FW. Access of ligands to cavities within the core of a protein is rapid. Nature structural biology. 1996;3(6):516-21. PubMed PMID: 8646537.

30. Lakowicz JR, Weber G. Quenching of protein fluorescence by oxygen. Detection of structural fluctuations in proteins on the nanosecond time scale. Biochemistry. 1973;12(21):4171-9. PubMed PMID: 4200894.

31. Nadler W, Stein DL. Biological transport processes and space dimension. Proceedings of the National Academy of Sciences of the United States of America. 1991;88(15):6750-4. PubMed PMID: 1713690; PubMed Central PMCID: PMC52166.

32. Wang PH, Blumberger J. Mechanistic insight into the blocking of CO diffusion in [NiFe]-hydrogenase mutants through multiscale simulation. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(17):6399-404. doi: 10.1073/pnas.1121176109. PubMed PMID: 22493222; PubMed Central PMCID: PMC3340082.

33. Branchini BR, Nemser, A. R., Zimmer, M. A Computational Analysis of the Unique Protein-Induced Tight Turn That Results in Posttranslational Chromophore Formation in Green Fluorescent Protein. J Am Chem Soc. 1998;120(1):1-6. doi: 10.1021/ja973019j.

34. Shu X, Shaner NC, Yarbrough CA, Tsien RY, Remington SJ. Novel chromophores and buried charges control color in mFruits. Biochemistry. 2006;45(32):9639-47. doi: 10.1021/bi060773l. PubMed PMID: 16893165.

35. Goedhart J, von Stetten D, Noirclerc-Savoye M, Lelimousin M, Joosen L, Hink MA, et al. Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. Nature communications. 2012;3:751. doi: 10.1038/ncomms1738. PubMed PMID: 22434194; PubMed Central PMCID: PMC3316892.

36. Griesbeck O, Baird GS, Campbell RE, Zacharias DA, Tsien RY. Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. The Journal of biological chemistry. 2001;276(31):29188-94. doi: 10.1074/jbc.M102815200. PubMed PMID: 11387331.

37. Roy A, Carpentier P, Bourgeois D, Field M. Diffusion pathways of oxygen species in the phototoxic fluorescent protein KillerRed. Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology. 2010;9(10):1342-50. doi: 10.1039/c0pp00141d. PubMed PMID: 20820672.

38. Subach FV, Verkhusha VV. Chromophore transformations in red fluorescent proteins. Chemical reviews. 2012;112(7):4308-27. doi: 10.1021/cr2001965. PubMed PMID: 22559232; PubMed Central PMCID: PMC3394910.

39. Shaner NC, Lin MZ, McKeown MR, Steinbach PA, Hazelwood KL, Davidson MW, et al. Improving the photostability of bright monomeric orange and red fluorescent proteins. Nature methods. 2008;5(6):545-51. doi: 10.1038/nmeth.1209. PubMed PMID: 18454154; PubMed Central PMCID: PMC2853173.

40. Field MJ. A practical introduction to the simulation of molecular systems. Cambridge: Cambridge University Press; 1999.

41. Brooks BR, Brooks CL, 3rd, Mackerell AD, Jr., Nilsson L, Petrella RJ, Roux B, et al. CHARMM: the biomolecular simulation program. Journal of computational

chemistry. 2009;30(10):1545-614. doi: 10.1002/jcc.21287. PubMed PMID: 19444816; PubMed Central PMCID: PMC2810661.

42. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. Journal of computational chemistry. 2005;26(16):1781-802. doi: 10.1002/jcc.20289. PubMed PMID: 16222654; PubMed Central PMCID: PMC2486339.

43. Leach AR. Molecular modelling principles and applications. 2nd ed. Great Britain: Pearson Education Limited; 2001.

44. Jones M. Organic Chemistry, 3rd ed. New York, NY 10110: W. W. Norton & Company Inc; 2005.

45. Darden T, Perera L, Li L, Pedersen L. New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. Structure. 1999;7(3):R55-60. PubMed PMID: 10368306.

46. Kittel C. Introduction to solid state physics. 8th ed. USA: John Wiley & Sons, Inc.; 2005.

47. Steinbach PJaB, B. R. New Spherical-Cutoff Methods for Long-Range Forces in Macromolecular Simulation. Journal of computational chemistry. 1994;15(7):667-83. doi: 10.1002/jcc.540150702.

48. Andersen HC. Molecular dynamics simulations at constant pressure and/or temperature J Chem Phys. 1980;72(4):10. doi: 10.1063/1.439486.

49. Berendsen HJC, Postma, J. P. M., Gunsteren, W. F. V., diNola, A. and Haak, J. R. Molecular dynamics with coupling to an external bath J Chem Phys. 1984;81(8):7. doi: 10.1063/1.448118.

50. Hoover WG. Canonical dynamics: Equilibrium phase-space distributions. Phys Rev A. 1985;31(3):1695–7. doi: 10.1103/PhysRevA.31.1695.

51. Nose S. A unified formulation of the constant temperature molecular dynamics methods J Chem Phys. 1984;81(1):9. doi: 10.1063/1.447334.

52. Reuter N, Lin, H. and Thiel, W. Green Fluorescent Proteins: Empirical Force Field for the Neutral and Deprotonated Forms of the Chromophore. Molecular Dynamics Simulations of the Wild Type and S65T Mutant. J Phys Chem B. 2002;106(24):6310–21. doi: 10.1021/jp014476w.

53. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, et al. CHARMM general force field: A force field for drug-like molecules compatible with the

CHARMM all-atom additive biological force fields. Journal of computational chemistry. 2010;31(4):671-90. doi: 10.1002/jcc.21367. PubMed PMID: 19575467; PubMed Central PMCID: PMC2888302.

54. Petrek M, Otyepka, M., Banas, P., Kosinova, P., Koca, J. and Damborsky, J. CAVER: a new tool to explore routes from protein clefts, pockets and cavities. BMC Bioinformatics. 2006;7:316:1-9. doi: 10.1186/1471-2105-7-316.

55. Roitberg AaE, R. Modeling side chains in peptides and proteins: Application of the locally enhanced sampling and the simulated annealing methods to find minimum energy conformations J Chem Phys. 1991;95(12):11. doi: 10.1063/1.461157.

56. Simmerling C, Fox, T. and Kollman, P. A. Use of Locally Enhanced Sampling in Free Energy Calculations: Testing and Application to the α → β Anomerization of Glucose. J Am Chem Soc. 1998;120(23):5771–82. doi: 10.1021/ja972457n.

57. Feig M, Karanicolas, J. and Brooks, C. . MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. J Mol Graph Model. 2004;22(5):377-95.

58. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, et al. Comparative protein structure modeling using Modeller. Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al]. 2006;Chapter 5:Unit 5 6. doi: 10.1002/0471250953.bi0506s15. PubMed PMID: 18428767.

59. Wachter RM, Elsliger MA, Kallio K, Hanson GT, Remington SJ. Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein. Structure. 1998;6(10):1267-77. PubMed PMID: 9782051.

60. Heim R, Prasher DC, Tsien RY. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. Proceedings of the National Academy of Sciences of the United States of America. 1994;91(26):12501-4. PubMed PMID: 7809066; PubMed Central PMCID: PMC45466.

61. Chapagain PP, Regmi CK, Castillo W. Fluorescent protein barrel fluctuations and oxygen diffusion pathways in mCherry. The Journal of chemical physics. 2011;135(23):235101. doi: 10.1063/1.3660197. PubMed PMID: 22191901; PubMed Central PMCID: PMC3248888.

62. Shaner NC, Campbell RE, Steinbach PA, Giepmans BN, Palmer AE, Tsien RY. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nature biotechnology. 2004;22(12):1567-72. doi: 10.1038/nbt1037. PubMed PMID: 15558047.

63. Regmi CK, Bhandari YR, Gerstman BS, Chapagain PP. Exploring the Diffusion of Molecular Oxygen in the Red Fluorescent Protein mCherry Using Explicit Oxygen Molecular Dynamics Simulations. The journal of physical chemistry B. 2013. doi: 10.1021/jp308366y. PubMed PMID: 23363049.

64. Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA. Fluorescent proteins and their applications in imaging living cells and tissues. Physiological reviews. 2010;90(3):1103-63. doi: 10.1152/physrev.00038.2009. PubMed PMID: 20664080.

65. Shcherbo D, Merzlyak EM, Chepurnykh TV, Fradkov AF, Ermakova GV, Solovieva EA, et al. Bright far-red fluorescent protein for whole-body imaging. Nature methods. 2007;4(9):741-6. doi: 10.1038/nmeth1083. PubMed PMID: 17721542.

66. Shcherbo D, Murphy CS, Ermakova GV, Solovieva EA, Chepurnykh TV, Shcheglov AS, et al. Far-red fluorescent tags for protein imaging in living tissues. The Biochemical journal. 2009;418(3):567-74. doi: 10.1042/BJ20081949. PubMed PMID: 19143658; PubMed Central PMCID: PMC2893397.

67. Mauring K, Deich J, Rosell FI, McAnaney TB, Moerner WE, Boxer SG. Enhancement of the fluorescence of the blue fluorescent proteins by high pressure or low temperature. The journal of physical chemistry B. 2005;109(26):12976-81. doi: 10.1021/jp0448595. PubMed PMID: 16852610.

68. Quillin ML, Anstrom DM, Shu X, O'Leary S, Kallio K, Chudakov DM, et al. Kindling fluorescent protein from Anemonia sulcata: dark-state structure at 1.38 A resolution. Biochemistry. 2005;44(15):5774-87. doi: 10.1021/bi047644u. PubMed PMID: 15823036.

69. Petersen J, Wilmann PG, Beddoe T, Oakley AJ, Devenish RJ, Prescott M, et al. The 2.0-A crystal structure of eqFP611, a far red fluorescent protein from the sea anemone Entacmaea quadricolor. The Journal of biological chemistry. 2003;278(45):44626-31. doi: 10.1074/jbc.M307896200. PubMed PMID: 12909624.

70. Pozzi EA, Schwall LR, Jimenez R, Weber JM. Pressure-Induced Changes in the Fluorescence Behavior of Red Fluorescent Proteins. The journal of physical chemistry B. 2012. doi: 10.1021/jp306093h. PubMed PMID: 22861177.

71. Laurent AD, Mironov VA, Chapagain PP, Nemukhin AV, Krylov AI. Exploring structural and optical properties of fluorescent proteins by squeezing: modeling high-pressure effects on the mStrawberry and mCherry red fluorescent proteins. The journal of physical chemistry B. 2012;116(41):12426-40. doi: 10.1021/jp3060944. PubMed PMID: 22988813; PubMed Central PMCID: PMC3500579.

72. Bravaya KB, Subach OM, Korovina N, Verkhusha VV, Krylov AI. Insight into the common mechanism of the chromophore formation in the red fluorescent proteins: the elusive blue intermediate revealed. Journal of the American Chemical Society.

2012;134(5):2807-14. doi: 10.1021/ja2114568. PubMed PMID: 22239269; PubMed Central PMCID: PMC3310345.

73. Barstow B, Ando N, Kim CU, Gruner SM. Alteration of citrine structure by hydrostatic pressure explains the accompanying spectral shift. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(36):13362-6. doi: 10.1073/pnas.0802252105. PubMed PMID: 18768811; PubMed Central PMCID: PMC2533195.

74. Barstow B, Ando N, Kim CU, Gruner SM. Coupling of pressure-induced structural shifts to spectral changes in a yellow fluorescent protein. Biophysical journal. 2009;97(6):1719-27. doi: 10.1016/j.bpj.2009.06.039. PubMed PMID: 19751677; PubMed Central PMCID: PMC2749779.

75. Gekko K, Hasegawa Y. Compressibility-structure relationship of globular proteins. Biochemistry. 1986;25(21):6563-71. PubMed PMID: 3790543.

76. Niwa H, Inouye S, Hirano T, Matsuno T, Kojima S, Kubota M, et al. Chemical nature of the light emitter of the Aequorea green fluorescent protein. Proceedings of the National Academy of Sciences of the United States of America. 1996;93(24):13617-22. PubMed PMID: 8942983; PubMed Central PMCID: PMC19369.

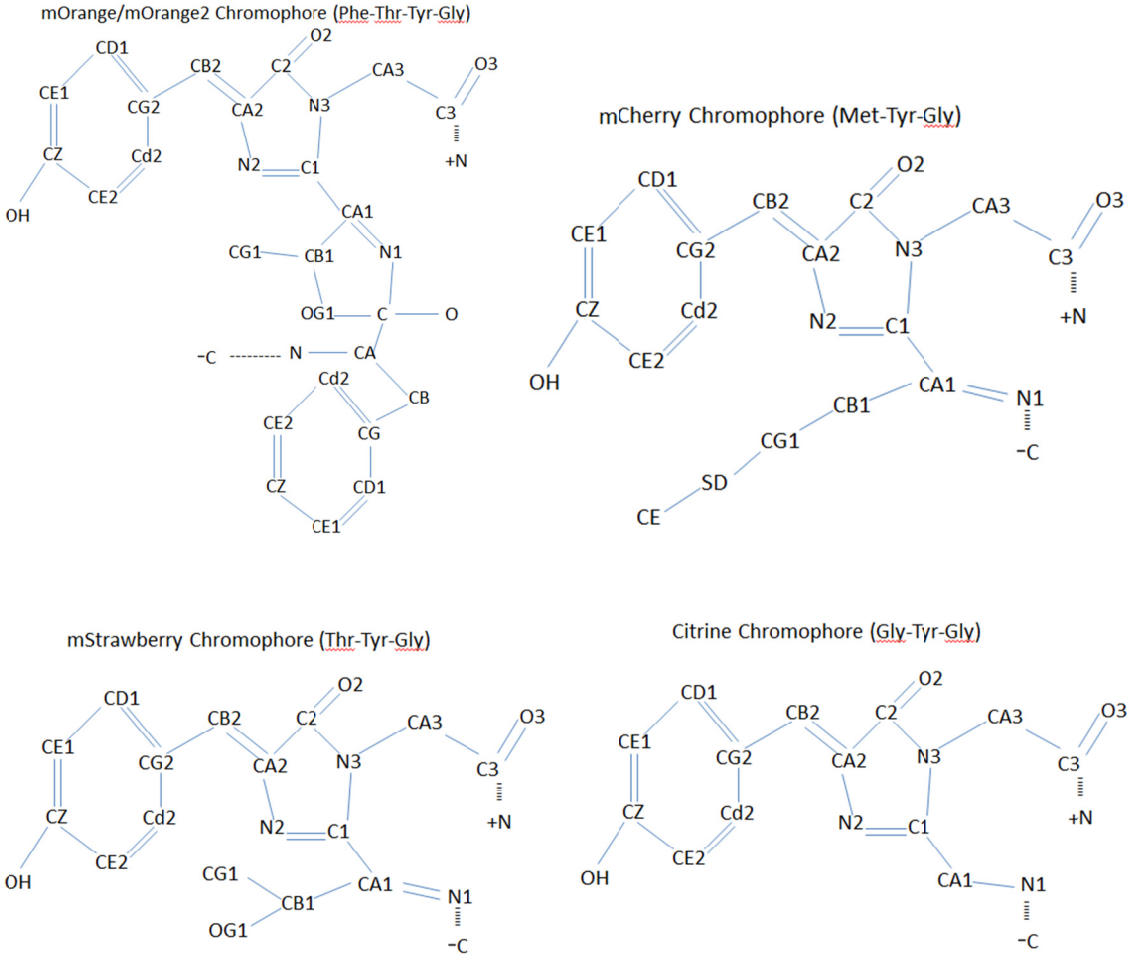
77. Wilmann PG, Petersen J, Devenish RJ, Prescott M, Rossjohn J. Variations on the GFP chromophore: A polypeptide fragmentation within the chromophore revealed in the 2.1-A crystal structure of a nonfluorescent chromoprotein from Anemonia sulcata. The Journal of biological chemistry. 2005;280(4):2401-4. doi: 10.1074/jbc.C400484200. PubMed PMID: 15542608.

78. Wachter RM, King BA, Heim R, Kallio K, Tsien RY, Boxer SG, et al. Crystal structure and photodynamic behavior of the blue emission variant Y66H/Y145F of green fluorescent protein. Biochemistry. 1997;36(32):9759-65. doi: 10.1021/bi970563w. PubMed PMID: 9245407.

79. Prescott M, Ling M, Beddoe T, Oakley AJ, Dove S, Hoegh-Guldberg O, et al. The 2.2 A crystal structure of a pocilloporin pigment reveals a nonplanar chromophore conformation. Structure. 2003;11(3):275-84. PubMed PMID: 12623015.

APPENDICES

Appendix A: Chromophore Structure of various fluorescent proteins. Hydrogen atoms are not shown for clarity.



Appendix B: Residue topology and parameter files for the chromophore of mOrange and mOrange2

Table 5.1 Residue topology file for the chromophore of mOrange and mOrange2

```
MASS    251 NRc2     14.00700 N ! neutral his unprotonated ring nitrogen
MASS    252 NRc1     14.00700 N ! neutral his protonated ring nitrogen
MASS    253 HAc1     1.00800 H ! for alkene; RHC=CR
MASS    254 HPc      1.00800 H ! aromatic H
```

```
MASS   255 Oc2      15.99900 O ! carbonyl oxygen
MASS   256 OcH      15.99900 O ! from OH1
!MASS   257 HcH     1.00800 H ! polar H  ! not needed for anionic
MASS   258 HAc      1.00800 H ! nonpolar H
MASS   259 CA1      12.01100 C ! aromatic C
MASS   260 CA2      12.01100 C ! aromatic C
MASS   261 CA3      12.01100 C ! aromatic C
MASS   262 CPc2     12.01100 C ! his CE1 carbon
MASS   263 CEc1     12.01100 C ! for alkene; RHC=CR
MASS   264 CPc1     12.01100 C ! his CG and CD2 carbons
MASS   265 CA4      12.01100 C ! aromatic C
!MASS   266 CT3c    12.01100 C ! aliphatic sp3 C for CH3
MASS   267 CAg1     12.01100 C !
MASS   268 CBg1     12.01100 C !
MASS   269 Cg4      12.01100 C !
MASS   270 OGg1     15.99900 O  ! carbonyl oxygen
MASS   271 Og4      15.99900 O ! carbonyl oxygen
MASS   272 HBg1     1.00800 H ! aromatic H
MASS   273 HOg1     1.00800 H ! aromatic H
MASS   274 Ng1      14.00700 N ! neutral his protonated ring nitrogen

DECL -CA
DECL -C
DECL -O
DECL -C3
DECL +N4
DECL +N
DECL +HN
DECL +CA

DEFA FIRS NTER LAST CTER
AUTO ANGLES DIHE

RESI CRO  -1.000
GROUP                    ! Imidazolinone ring
ATOM C1    CPc2   0.50
ATOM N2    NRc2  -0.60
ATOM N3    NRc1  -0.57
ATOM C2    CPc1   0.57
ATOM O2    Oc2   -0.57
ATOM CA2   CPc1   0.10
ATOM CB2   CEc1  -0.14
ATOM HB2   HAc1   0.21
ATOM CG2   CA1   -0.09 ! Tyr ring : charges from charmm22
ATOM CD1   CA2   -0.08
ATOM HD1   HPc    0.14
ATOM CD2   CA2   -0.08
ATOM HD2   HPc    0.14
ATOM CE1   CA3   -0.28
ATOM HE1   HPc    0.10
ATOM CE2   CA3   -0.28
ATOM HE2   HPc    0.10
ATOM CZ    CA4    0.45
ATOM OH    OcH   -0.62 ! Absence of H ---- anionic
```

```
!
!Glycine part from Charmm22
GROUP
ATOM CA3   CT2   -0.18  !     |
ATOM HA31  HB     0.09  !     |
ATOM HA32  HB     0.09  ! HA1-CA-HA2
GROUP                   !     |
ATOM C3    C      0.51  !     |
ATOM O3    O     -0.51  !     C=O
!
!Thr and Phe cyclized part part from charmm22
GROUP
ATOM N1      Ng1       -0.61 ! oxazole like ring
ATOM CA1     CAg1       0.22
ATOM CB1     CBg1       0.34 ! from compensating
ATOM CG1     CT3       -0.27
ATOM HG11    HA         0.09
ATOM HG12    HA         0.09
ATOM HG13    HA         0.09
ATOM HB1     HBg1       0.11
ATOM OG1     OGg1      -0.29 ! from oxazole
ATOM C4      Cg4        0.34 ! from compensating
ATOM O4      Og4       -0.54
ATOM HO1     HOg1       0.43
GROUP
ATOM CA4     CT1        0.07
ATOM N4      NH1       -0.47
ATOM HN4     H          0.31
ATOM HA4     HB         0.09

ATOM CB4     CT2        -0.18
ATOM HB41    HA          0.09
ATOM HB42    HA          0.09
GROUP
ATOM CG4     CA          0.00  ! benzene ring
ATOM CD4     CA         -0.115
ATOM HD4     HP         0.115
ATOM CD5     CA         -0.115
ATOM HD5     HP          0.115
ATOM CE4     CA         -0.115
ATOM HE4     HP          0.115
ATOM CE5     CA         -0.115
ATOM HE5     HP          0.115
ATOM Cz4     CA         -0.115
ATOM Hz4     HP          0.115

BOND CA1 C1 N4 -C C3 +N
BOND N2 CA2 CB2 HB2 CB2 CG2 CD1 HD1 CD1 CE1 CE1 HE1 CZ OH
BOND CZ CE2 CE2 HE2 CD2 HD2 CD2 CG2 CA2 C2
BOND N3 CA3 CA3 HA31 CA3 HA32 CA3 C3 N3 C1 N3 C2
BOND N1 C4 C4 O4 O4 HO1 C4 OG1 OG1 CB1 CB1 CG1 CB1 HB1 CB1 CA1
BOND CG1 HG11 CG1 HG12 CG1 HG13
BOND C4 CA4 CA4 HA4 CA4 N4 N4 HN4 CA4 CB4 CB4 HB41 CB4 HB42 CB4 CG4
BOND CD5 HD5 CD5 CE5 CE5 HE5 Cz4 Hz4 Cz4 CE4 CE4 HE4
BOND CD4 HD4 CD4 CG4
```

```
DOUBLE C1 N2 CA2 CB2 C2 O2 C3 O3 CD1 CG2 CD2 CE2 CZ CE1 CA1 N1 CG4 CD5
CE5 Cz4 CE4 CD4
```

Table 5.2 Parameter file for the chromophore of mOrange and mOrange2

```
*charmm parameter file of mStrawberry chromophore (phe-thr-tyr-gly)
*
!parameter file

! GFP Chromophore parameters, protonated form
!
BONDS
!
!V(bond) = Kb(b - b0)**2
!
!Kb: kcal/mole/A**2
!b0: A
!
!atom type Kb          b0

CPc2 CTc1  354.000     1.4900 !ion for RFP C1-CA1 connection
NRc1 CT2   396.000     1.4400 !ion
NRc1 CPc2  400.000     1.3900 !
NRc1 CPc1  400.000     1.4100 !
CPc1 Oc2   854.000     1.2400 !ion
NRc2 CPc2  400.000     1.3000 !
CPc1 CPc1  410.000     1.4600 !ion
NRc2 CPc1  400.000     1.4000 !
CPc1 CEc1  500.000     1.3900 !ion
HAc1 CEc1  360.500     1.1000 !
CEc1 CA1   437.000     1.4100 !ion
CA1  CA2   305.000     1.4300 !ion
HPc  CA2   340.000     1.0800 !
CA2  CA3   305.000     1.3500 !ion
HPc  CA3   340.000     1.0800 !
CA3  CA4   305.000     1.4550 !ion
OcH  CA4   842.000     1.2500 !ion
CTc1 NH1   463.000     1.3650 !RFP CA1-N1 connection
CT1  CTc1  222.500     1.5000 ! ALLOW   ALI

ANGLES
!
!V(angle) = Ktheta(Theta - Theta0)**2
!
!V(Urey-Bradley) = Kub(S - S0)**2
!
!Ktheta: kcal/mole/rad**2
!Theta0: degrees
```

```
!Kub: kcal/mole/A**2 (Urey-Bradley)
!S0: A
!
!atom types     Ktheta    Theta0   Kub     S0
!

NRc2 CPc2 NRc1  130.000   113.3000 !ion
CPc2 NRc2 CPc1  130.000   106.6000 !ion
CPc2 NRc1 CPc1  130.000   107.9000 !
NRc2 CPc1 CPc1  130.000   108.3000 !
NRc2 CPc1 CEc1   45.800   129.5000 !
NRc1 CPc1 Oc2    50.000   124.0000 !ion
NRc1 CPc1 CPc1  130.000   103.0000 !
Oc2  CPc1 CPc1   44.000   133.0000 !ion
CPc1 CPc1 CEc1   45.800   122.7000 !ion
CPc1 CEc1 CA1   130.000   133.2000 !ion
CPc1 CEc1 HAc1   42.000   112.0000 !ion
CEc1 CA1  CA2    45.800   120.0000 !ion
HAc1 CEc1 CA1    42.000   115.0000 !ion
CA1  CA2  CA3    40.000   122.0000 !ion
CA2  CA1  CA2    40.000   116.0000 !ion
CA2  CA3  CA4    40.000   122.0000 !ion
CA3  CA4  CA3    40.000   115.0000 !ion
!
HPc  CA3  CA4    30.000   120.0000 !
HPc  CA3  CA2    30.000   120.0000 !
HPc  CA2  CA3    30.000   120.0000 !
HPc  CA2  CA1    30.000   120.0000 !
!
OcH  CA4  CA3    45.200   120.0000 ! ALLOW   ARO ALC

!Link to the thr(65) fragment
NH1  CTc1 CPc2   50.000   107.0000 ! ALLOW   PEP POL ARO ALI
NRc2 CPc2 CTc1   40.000   125.0000 !
NRc1 CPc2 CTc1   40.000   121.7000 !ion
!CT2  CT1  CPc2   52.000   108.0000 ! ALLOW   ALI PEP POL ARO
CT1  CTc1 CPc2   52.000   108.0000 ! ALLOW   ALI PEP POL ARO

CTc1 NH1  C      50.000   120.0000
NH1  CTc1 CT1    70.000   113.5000
HA   CT1  CTc1   34.500    110.10   22.53   2.17900 ! ALLOW   ALI
CT3  CT1  CTc1   53.350    108.50    8.00   2.56100 ! ALLOW   ALI
OH1  CT1  CTc1   75.700   110.1000 ! ALLOW   ALI ALC ARO  (Yuba)

!Link to the gly(67) fragment
CPc2 NRc1 CT2    40.000   128.3000 !ion
CPc1 NRc1 CT2    40.000   123.8000 !ion
NRc1 CT2  C      50.000   107.0000
NRc1 CT2  HB     48.000   108.0000

!
DIHEDRALS
!
!V(dihedral) = Kchi(1 + cos(n(chi) - delta))
!
```

```
!Kchi: kcal/mole
!n: multiplicity
!delta: degrees
!
!atom types             Kchi    n   delta
!
CPc2 NRc2 CPc1 CPc1    14.0000  2   180.00 !
CPc2 NRc1 CPc1 CPc1    14.0000  2   180.00 !
NRc2 CPc2 NRc1 CPc1    14.0000  2   180.00 !
NRc2 CPc1 CPc1 NRc1     4.0000  2   180.00 !
NRc1 CPc2 NRc2 CPc1     4.0000  2   180.00 !
CA1  CA2  CA3  CA4      3.1000  2   180.00 !
CA2  CA1  CA2  CA3      3.1000  2   180.00 !
CA2  CA3  CA4  CA3      3.1000  2   180.00 !
!Oc2  CAC  CAC  CAC      3.1000  2   180.00 !
CA2  CA3  CA4  OcH      3.1000  2   180.00 !
CA1  CA2  CA3  HPc      4.2000  2   180.00 !
CA2  CA1  CA2  HPc      4.2000  2   180.00 !
CA3  CA4  CA3  HPc      4.2000  2   180.00 !
HPc  CA2  CA3  CA4      4.2000  2   180.00 !
HPc  CA2  CA3  HPc      2.4000  2   180.00 !
!HcH  Oc2  CAC  CAC      0.9900  2   180.00 !
!HcH  OcH  CA4  CA3      0.9900  2   180.00 !
HPc  CA3  CA4  OcH      4.2000  2   180.00 !
!
CPc2 NRc2 CPc1 CEc1     3.000   2   180.00 !
NRc1 CPc1 CPc1 CEc1     3.00    2   180.00 !
Oc2  CPc1 CPc1 CEc1     2.00    2   180.00 !
CEc1 CA1  CA2  HPc      4.20    2   180.00 !
CEc1 CA1  CA2  CA3      3.10    2   180.00 !

!connection CA-CB
CPc1 CPc1 CEc1 HAc1     3.9000  2   180.00 !
CPc1 CPc1 CEc1 CA1      3.9000  2   180.00 !
NRc2 CPc1 CEc1 HAc1     3.9000  2   180.00 !
NRc2 CPc1 CEc1 CA1      3.9000  2   180.00 !

!connection CB-CG2
CPc1 CEc1 CA1  CA2      2.7000  2   180.00 !
HAc1 CEc1 CA1  CA2      2.7000  2   180.00 !
!
CPc2 NRc1 CPc1 Oc2     14.0000  2   180.00 !
NRc2 CPc2 NRc1 CT2     14.0000  2   180.00 !
NRc2 CPc1 CPc1 Oc2     14.0000  2   180.00 !
CPc1 NRc1 CPc2 CTc1    14.0000  2   180.00 !
Oc2  CPc1 NRc1 CT2     14.0000  2   180.00 !
CPc1 NRc2 CPc2 CTc1    14.0000  2   180.00 !
CPc1 CPc1 NRc1 CT2     14.0000  2   180.00 !
CTc1 CPc2 NRc1 CT2     14.0000  2   180.00 !

! Linking the chromophore and the glycine(67) fragment
O    C    CT2  NRc1     0.0000  1     0.00 !
NH1  C    CT2  NRc1     0.6000  1     0.00 !
CPc2 NRc1 CT2  HB       0.0320  3     0.00 !
CPc2 NRc1 CT2  C        0.0320  3     0.00 !
```

```
CPc1 NRc1 CT2  HB       0.0320  3   180.00 !
CPc1 NRc1 CT2  C        0.0320  3   180.00 !


! Linking the chromophore and the thr(65) fragment
C    NH1  CTc1 CPc2     2.2500  2   180.00 !Taken from X-C-NC2-X
Charmm22
NRc2 CPc2 CTc1 CT1      0.1050  3   180.00 !
NRc2 CPc2 CTc1 NH1      0.1050  3   180.00 !
NRc1 CPc2 CTc1 CT1      0.1050  3     0.00 !
NRc1 CPc2 CTc1 NH1      0.1050  3     0.00 !

!connecting N1=CA1 region due to new type CTc1
O    C    NH1  CTc1     2.5000  2   180.00
CT1  C    NH1  CTc1     2.5000  2   180.00
CT1  CTc1 NH1  C        1.8000  1     0.00
CPc2 CTc1 CT1  HA       0.2000  3     0.00
CPc2 CTc1 CT1  CT3      0.2000  3     0.00
NH1  CTc1 CT1  HA       0.2000  3     0.00
NH1  CTc1 CT1  CT3      0.2000  3     0.00   ! X    CT1  CT1  X
CPC2 CTC1 CT1  OH1      0.2000  3     0.00 ! ALLOW   ALI
NH1  CTC1 CT1  OH1      0.2000  3     0.00 ! ALLOW   ALI

IMPROPER
!
!V(improper) = Kpsi(psi - psi0)**2
!
!Kpsi: kcal/mole/rad**2
!psi0: degrees
!note that the second column of numbers (0) is ignored
!
!atom types           Kpsi                   psi0
!

CPc2 NRc2 NRc1 CTc1    50.0000         0      0.0000
CPc2 NRc1 NRc2 CTc1    50.0000         0      0.0000
!
CPc1 NRc1 CPc1 Oc2     50.0000         0      0.0000
CPc1 CPc1 NRc1 Oc2     50.0000         0      0.0000
!
NRc1 CPc1 CPc2 CT2     50.0000         0      0.0000
NRc1 CPc2 CPc1 CT2     50.0000         0      0.0000
!
CPc1 NRc2 CPc1 CEc1    50.0000         0      0.0000
CPc1 CPc1 NRc2 CEc1    50.0000         0      0.0000
!
CEc1 CPc1 CA1  HAc1    30.0000         0      0.0000
CEc1 CA1  CPc1 HAc1    30.0000         0      0.0000

!
!V(Lennard-Jones) = Eps,i,j[(Rmin,i,j/ri,j)**12 - 2(Rmin,i,j/ri,j)**6]
!
!epsilon: kcal/mole, Eps,i,j = sqrt(eps,i * eps,j)
!Rmin/2: A, Rmin,i,j = Rmin/2,i + Rmin/2,j
!
```

```
!atom  ignored    epsilon      Rmin/2   ignored   eps,1-4     Rmin/2,1-4
!
!CAc     5.000000  -0.070000    1.992400 ! ALLOW   ARO

NONBONDED nbxmod  5 atom cdiel shift vatom vdistance vswitch -
cutnb 14.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5

CA1    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CA2    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CA3    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CA4    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CEc1   0.000000  -0.068000     2.090000 ! ! for propene, yin/adm jr.,
CPc1   0.000000  -0.050000     1.800000 ! ALLOW ARO
CPc2   0.000000  -0.050000     1.800000 ! ALLOW ARO
HAc    0.000000  -0.022000     1.320000 ! ALLOW PEP ALI POL SUL
HAc1   0.000000  -0.031000     1.250000 !
HPc    0.000000  -0.030000   1.358200   0.000000  -0.030000    1.358200
NRc1   0.000000  -0.200000     1.850000 ! ALLOW ARO
NRc2   0.000000  -0.200000     1.850000 ! ALLOW ARO
Oc2    0.000000  -0.120000     1.700000 ! ALLOW   PEP POL,
OcH    0.000000  -0.152100     1.770000 ! ALLOW   ALC ARO
CTc1   0.000000  -0.020000  2.275000 ! 0.000000  -0.010000     1.900000
```

Appendix C: Residue topology and parameter files for the chromophore of mCherry

Table 5.3 Residue topology file for the chromophore of mCherry

```
MASS   197 NRc2  14.00700 N ! neutral his unprotonated ring nitrogen
MASS   198 NRc1  14.00700 N ! neutral his protonated ring nitrogen
MASS   199 HAc1   1.00800 H ! for alkene; RHC=CR
MASS   200 HPc    1.00800 H ! aromatic H
MASS   201 Oc2   15.99900 O ! carbonyl oxygen
MASS   202 OcH   15.99900 O ! from OH1
!MASS   203 HcH    1.00800 H ! polar H
MASS   204 HAc    1.00800 H ! nonpolar H
MASS   205 CA1   12.01100 C ! aromatic C
MASS   206 CA2   12.01100 C ! aromatic C
MASS   207 CA3   12.01100 C ! aromatic C
MASS   208 CPc2  12.01100 C ! his CE1 carbon
MASS   209 CEc1  12.01100 C ! for alkene; RHC=CR
MASS   210 CPc1  12.01100 C ! his CG and CD2 carbons
MASS   211 CA4   12.01100 C ! aromatic C
!MASS   212 CT3c  12.01100 C ! aliphatic sp3 C for CH3
MASS   213 CTc1  12.01100 C ! aliphatic sp3 C for CH

DECL -CA
DECL -C
DECL -O
DECL -C3
DECL +N1
```

```
DECL +N
DECL +HN
DECL +CA

DEFA FIRS NTER LAST CTER
AUTO ANGLES DIHE

RESI CH6  -1.000
GROUP                    ! Imidazolinone ring
ATOM C1    CPc2   0.50
ATOM N2    NRc2  -0.60
ATOM N3    NRc1  -0.57
ATOM C2    CPc1   0.57
ATOM O2    Oc2   -0.57
ATOM CA2   CPc1   0.10
ATOM CB2   CEc1  -0.14
ATOM HB2   HAc1   0.21
ATOM CG2   CA1   -0.09 ! Tyr ring : charges from charmm22
ATOM CD1   CA2   -0.08
ATOM HD1   HPc    0.14
ATOM CD2   CA2   -0.08
ATOM HD2   HPc    0.14
ATOM CE1   CA3   -0.28
ATOM HE1   HPc    0.10
ATOM CE2   CA3   -0.28
ATOM HE2   HPc    0.10
ATOM CZ    CA4    0.45
ATOM OH    OcH   -0.62

!Glycine part from Charmm22
GROUP
ATOM CA3   CT2   -0.18  !     |
ATOM HA31  HB     0.09  !     |
ATOM HA32  HB     0.09  ! HA1-CA-HA2
GROUP                   !     |
ATOM C3    C      0.51  !     |
ATOM O3    O     -0.51  !    C=O
!
!met part from charmm22
GROUP
ATOM N1    NH1    -0.16
ATOM CA1   CTc1    0.16 !atom type changed
GROUP
ATOM CB1   CT2    -0.18
ATOM HB11  HA      0.09
ATOM HB12  HA      0.09
GROUP
ATOM CG1   CT2    -0.14
ATOM HG11  HA      0.09
ATOM HG12  HA      0.09
ATOM SD    S      -0.09
ATOM CE    CT3    -0.22
ATOM HE11  HA      0.09
ATOM HE12  HA      0.09
ATOM HE13  HA      0.09
```

```
BOND CA1 C1 N1 -C C3 +N
BOND N2 CA2 CB2 HB2 CB2 CG2 CD1 HD1 CD1 CE1 CE1 HE1 CZ OH
BOND CZ CE2 CE2 HE2 CD2 HD2 CD2 CG2 CA2 C2
BOND N3 CA3 CA3 HA31 CA3 HA32 CA3 C3 N3 C1 N3 C2
BOND CB1 HB11 CB1 HB12 CB1 CG1 CG1 HG11 CG1 HG12 CA1 CB1
BOND CG1 SD SD CE CE HE11 CE HE12 CE HE13
DOUBLE C1 N2 CA2 CB2 C2 O2 C3 O3 CD1 CG2 CD2 CE2 CZ CE1 CA1 N1
```

Table 5.4 Parameter file for the chromophore of mCherry

```
*charmm parameter file of mCherry chromophore (met-tyr-gly)
*
!parameter file

! GFP Chromophore parameters, protonated form
!
BONDS
!
!V(bond) = Kb(b - b0)**2
!
!Kb: kcal/mole/A**2
!b0: A
!
!atom type Kb          b0

CPc2 CTc1  354.000     1.4900 !ion for RFP C1-CA1 connection
NRc1 CT2   396.000     1.4400 !ion
NRc1 CPc2  400.000     1.3900 !
NRc1 CPc1  400.000     1.4100 !
CPc1 Oc2   854.000     1.2400 !ion
NRc2 CPc2  400.000     1.3000 !
CPc1 CPc1  410.000     1.4600 !ion
NRc2 CPc1  400.000     1.4000 !
CPc1 CEc1  500.000     1.3900 !ion
HAc1 CEc1  360.500     1.1000 !
CEc1 CA1   437.000     1.4100 !ion
CA1  CA2   305.000     1.4300 !ion
HPc  CA2   340.000     1.0800 !
CA2  CA3   305.000     1.3500 !ion
HPc  CA3   340.000     1.0800 !
CA3  CA4   305.000     1.4550 !ion
OcH  CA4   842.000     1.2500 !ion
CTc1 NH1   463.000     1.3650 !RFP CA1-N1 connection
CT2  CTc1  222.500     1.5380 !RFP CB1-CA1 connection

ANGLES
!
!V(angle) = Ktheta(Theta - Theta0)**2
!
!V(Urey-Bradley) = Kub(S - S0)**2
```

```
!
!Ktheta: kcal/mole/rad**2
!Theta0: degrees
!Kub: kcal/mole/A**2 (Urey-Bradley)
!S0: A
!
!atom types     Ktheta    Theta0   Kub     S0
!

NRc2 CPc2 NRc1  130.000   113.3000 !ion
CPc2 NRc2 CPc1  130.000   106.6000 !ion
CPc2 NRc1 CPc1  130.000   107.9000 !
NRc2 CPc1 CPc1  130.000   108.3000 !
NRc2 CPc1 CEc1   45.800   129.5000 !
NRc1 CPc1 Oc2    50.000   124.0000 !ion
NRc1 CPc1 CPc1  130.000   103.0000 !
Oc2  CPc1 CPc1   44.000   133.0000 !ion
CPc1 CPc1 CEc1   45.800   122.7000 !ion
CPc1 CEc1 CA1   130.000   133.2000 !ion
CPc1 CEc1 HAc1   42.000   112.0000 !ion
CEc1 CA1  CA2    45.800   120.0000 !ion
HAc1 CEc1 CA1    42.000   115.0000 !ion
CA1  CA2  CA3    40.000   122.0000 !ion
CA2  CA1  CA2    40.000   116.0000 !ion
CA2  CA3  CA4    40.000   122.0000 !ion
CA3  CA4  CA3    40.000   115.0000 !ion
!
HPc  CA3  CA4    30.000   120.0000 !
HPc  CA3  CA2    30.000   120.0000 !
HPc  CA2  CA3    30.000   120.0000 !
HPc  CA2  CA1    30.000   120.0000 !
!
OcH  CA4  CA3    45.200   120.0000 ! ALLOW   ARO ALC

!Link to the met(65) fragment
NH1  CTc1 CPc2   50.000   107.0000 ! ALLOW   PEP POL ARO ALI
NRc2 CPc2 CTc1   40.000   125.0000 !
NRc1 CPc2 CTc1   40.000   121.7000 !ion
CT2  CT1  CPc2   52.000   108.0000 ! ALLOW   ALI PEP POL ARO
CT2  CTc1 CPc2   52.000   108.0000 ! ALLOW   ALI PEP POL ARO
CTc1 NH1  C      50.000   120.0000
NH1  C    CTc1   80.000   116.5000
NH1  CTc1 CT2    70.000   113.5000
HA   CT2  CTc1   33.430   110.1000 22.53   2.17900
CT2  CT2  CTc1   58.350   113.50   11.16   2.56100

!Link to the gly(67) fragment
CPc2 NRc1 CT2    40.000   128.3000 !ion
CPc1 NRc1 CT2    40.000   123.8000 !ion
NRc1 CT2  C      50.000   107.0000
NRc1 CT2  HB     48.000   108.0000

!
DIHEDRALS
!
```

```
!V(dihedral) = Kchi(1 + cos(n(chi) - delta))
!
!Kchi: kcal/mole
!n: multiplicity
!delta: degrees
!
!atom types             Kchi    n   delta
!
CPc2 NRc2 CPc1 CPc1    14.0000  2   180.00 !
CPc2 NRc1 CPc1 CPc1    14.0000  2   180.00 !
NRc2 CPc2 NRc1 CPc1    14.0000  2   180.00 !
NRc2 CPc1 CPc1 NRc1     4.0000  2   180.00 !
NRc1 CPc2 NRc2 CPc1     4.0000  2   180.00 !
CA1  CA2  CA3  CA4      3.1000  2   180.00 !
CA2  CA1  CA2  CA3      3.1000  2   180.00 !
CA2  CA3  CA4  CA3      3.1000  2   180.00 !
!Oc2  CAC  CAC  CAC      3.1000  2   180.00 !
CA2  CA3  CA4  OcH      3.1000  2   180.00 !
CA1  CA2  CA3  HPc      4.2000  2   180.00 !
CA2  CA1  CA2  HPc      4.2000  2   180.00 !
CA3  CA4  CA3  HPc      4.2000  2   180.00 !
HPc  CA2  CA3  CA4      4.2000  2   180.00 !
HPc  CA2  CA3  HPc      2.4000  2   180.00 !
!HcH  Oc2  CAC  CAC      0.9900  2   180.00 !
!HcH  OcH  CA4  CA3      0.9900  2   180.00 !
HPc  CA3  CA4  OcH      4.2000  2   180.00 !
!
CPc2 NRc2 CPc1 CEc1     3.000   2   180.00 !
NRc1 CPc1 CPc1 CEc1     3.00    2   180.00 !
Oc2  CPc1 CPc1 CEc1     2.00    2   180.00 !
CEc1 CA1  CA2  HPc      4.20    2   180.00 !
CEc1 CA1  CA2  CA3      3.10    2   180.00 !
!connection CA-CB
CPc1 CPc1 CEc1 HAc1     3.9000  2   180.00 !
CPc1 CPc1 CEc1 CA1      3.9000  2   180.00 !
NRc2 CPc1 CEc1 HAc1     3.9000  2   180.00 !
NRc2 CPc1 CEc1 CA1      3.9000  2   180.00 !

!connection CB-CG2
CPc1 CEc1 CA1  CA2      2.7000  2   180.00 !
HAc1 CEc1 CA1  CA2      2.7000  2   180.00 !
!
CPc2 NRc1 CPc1 Oc2     14.0000  2   180.00 !
NRc2 CPc2 NRc1 CT2     14.0000  2   180.00 !
NRc2 CPc1 CPc1 Oc2     14.0000  2   180.00 !
CPc1 NRc1 CPc2 CTc1    14.0000  2   180.00 !
Oc2  CPc1 NRc1 CT2     14.0000  2   180.00 !
CPc1 NRc2 CPc2 CTc1    14.0000  2   180.00 !
CPc1 CPc1 NRc1 CT2     14.0000  2   180.00 !
CTc1 CPc2 NRc1 CT2     14.0000  2   180.00 !

! Linking the chromophore and the glycine(67) fragment
O    C    CT2  NRc1     0.0000  1     0.00 !
NH1  C    CT2  NRc1     0.6000  1     0.00 !
CPc2 NRc1 CT2  HB       0.0320  3     0.00 !
```

```
CPc2 NRc1 CT2  C        0.0320  3     0.00 !
CPc1 NRc1 CT2  HB       0.0320  3   180.00 !
CPc1 NRc1 CT2  C        0.0320  3   180.00 !

! Linking the chromophore and the met(65) fragment
C    NH1  CTc1 CPc2     2.2500  2   180.00 !Taken from X-C-NC2-X
Charmm22
NRc2 CPc2 CTc1 CT2      0.1050  3   180.00 !
NRc2 CPc2 CTc1 NH1      0.1050  3   180.00 !
NRc1 CPc2 CTc1 CT2      0.1050  3     0.00 !
NRc1 CPc2 CTc1 NH1      0.1050  3     0.00 !

!connecting N1=CA1 region due to new type CTc1
O    C    NH1  CTc1     2.5000  2   180.00
CT1  C    NH1  CTc1     2.5000  2   180.00
CT2  CTc1 NH1  C        1.8000  1     0.00
CPc2 CTc1 CT2  HA       0.2000  3     0.00
CPc2 CTc1 CT2  CT2      0.2000  3     0.00
NH1  CTc1 CT2  HA       0.2000  3     0.00
NH1  CTc1 CT2  CT2      0.2000  3     0.00



IMPROPER
!
!V(improper) = Kpsi(psi - psi0)**2
!
!Kpsi: kcal/mole/rad**2
!psi0: degrees
!note that the second column of numbers (0) is ignored
!
!atom types           Kpsi                   psi0
!

CPc2 NRc2 NRc1 CTc1    50.0000         0      0.0000
CPc2 NRc1 NRc2 CTc1    50.0000         0      0.0000
!
CPc1 NRc1 CPc1 Oc2     50.0000         0      0.0000
CPc1 CPc1 NRc1 Oc2     50.0000         0      0.0000
!
NRc1 CPc1 CPc2 CT2     50.0000         0      0.0000
NRc1 CPc2 CPc1 CT2     50.0000         0      0.0000
!
CPc1 NRc2 CPc1 CEc1    50.0000         0      0.0000
CPc1 CPc1 NRc2 CEc1    50.0000         0      0.0000
!
CEc1 CPc1 CA1  HAc1    30.0000         0      0.0000
CEc1 CA1  CPc1 HAc1    30.0000         0      0.0000



!
!V(Lennard-Jones) = Eps,i,j[(Rmin,i,j/ri,j)**12 - 2(Rmin,i,j/ri,j)**6]
!
!epsilon: kcal/mole, Eps,i,j = sqrt(eps,i * eps,j)
!Rmin/2: A, Rmin,i,j = Rmin/2,i + Rmin/2,j
```

```
!
!atom  ignored    epsilon      Rmin/2   ignored   eps,1-4     Rmin/2,1-4
!
!CAc     5.000000  -0.070000    1.992400 ! ALLOW   ARO

NONBONDED nbxmod  5 atom cdiel shift vatom vdistance vswitch -
cutnb 14.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5

CA1    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CA2    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CA3    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CA4    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CEc1   0.000000  -0.068000     2.090000 ! ! for propene, yin/adm jr.,
12/95
CPc1   0.000000  -0.050000     1.800000 ! ALLOW ARO
CPc2   0.000000  -0.050000     1.800000 ! ALLOW ARO
HAc    0.000000  -0.022000     1.320000 ! ALLOW PEP ALI POL SUL
HAc1   0.000000  -0.031000     1.250000 !
HPc    0.000000  -0.030000   1.358200   0.000000  -0.030000    1.358200
NRc1   0.000000  -0.200000     1.850000 ! ALLOW ARO
NRc2   0.000000  -0.200000     1.850000 ! ALLOW ARO
Oc2    0.000000  -0.120000     1.700000 ! ALLOW   PEP POL,
OcH    0.000000  -0.152100     1.770000 ! ALLOW   ALC ARO
CTc1   0.000000  -0.020000     2.275000
```

Appendix D: Residue topology and parameter files for the chromophore of mStrawberry

Table 5.5 Residue topology file for the chromophore of mStrawberry

```
MASS   197 NRc2  14.00700 N ! neutral his unprotonated ring nitrogen
MASS   198 NRc1  14.00700 N ! neutral his protonated ring nitrogen
MASS   199 HAc1   1.00800 H ! for alkene; RHC=CR
MASS   200 HPc    1.00800 H ! aromatic H
MASS   201 Oc2   15.99900 O ! carbonyl oxygen
MASS   202 OcH   15.99900 O ! from OH1
MASS   204 HAc    1.00800 H ! nonpolar H
MASS   205 CA1   12.01100 C ! aromatic C
MASS   206 CA2   12.01100 C ! aromatic C
MASS   207 CA3   12.01100 C ! aromatic C
MASS   208 CPc2  12.01100 C ! his CE1 carbon
MASS   209 CEc1  12.01100 C ! for alkene; RHC=CR
MASS   210 CPc1  12.01100 C ! his CG and CD2 carbons
MASS   211 CA4   12.01100 C ! aromatic C
MASS   212 CTc1  12.01100 C ! aliphatic sp3 C for CH

DECL -CA
DECL -C
DECL -O
DECL -C3
DECL +N1
```

```
DECL +N
DECL +HN
DECL +CA

DEFA FIRS NTER LAST CTER
AUTO ANGLES DIHE

RESI CRO  -1.000
GROUP                    ! Imidazolinone ring
ATOM C1    CPc2   0.50
ATOM N2    NRc2  -0.60
ATOM N3    NRc1  -0.57
ATOM C2    CPc1   0.57
ATOM O2    Oc2   -0.57
ATOM CA2   CPc1   0.10
ATOM CB2   CEc1  -0.14
ATOM HB2   HAc1   0.21
ATOM CG2   CA1   -0.09 ! Tyr ring : charges from charmm22
ATOM CD1   CA2   -0.08
ATOM HD1   HPc    0.14
ATOM CD2   CA2   -0.08
ATOM HD2   HPc    0.14
ATOM CE1   CA3   -0.28
ATOM HE1   HPc    0.10
ATOM CE2   CA3   -0.28
ATOM HE2   HPc    0.10
ATOM CZ    CA4    0.45
ATOM OH    OcH   -0.62

!
!Glycine part from Charmm22
GROUP
ATOM CA3   CT2   -0.18  !     |
ATOM HA31  HB     0.09  !     |
ATOM HA32  HB     0.09  ! HA1-CA-HA2
GROUP                   !     |
ATOM C3    C      0.51  !     |
ATOM O3    O     -0.51  !    C=O
!
!thr part from charmm22
GROUP
ATOM N1    NH1    -0.16
ATOM CA1   CTc1    0.16 !atom type changed
GROUP
ATOM CB1   CT1     0.14
ATOM HB1   HA      0.09
ATOM OG1   OH1    -0.66
ATOM HG1   H       0.43
GROUP
ATOM CG1   CT3    -0.27
ATOM HG11  HA      0.09
ATOM HG12  HA      0.09
ATOM HG13  HA      0.09

!!!Now needed to define bond
```

```
BOND CA1 C1 N1 -C C3 +N
BOND N2 CA2 CB2 HB2 CB2 CG2 CD1 HD1 CD1 CE1 CE1 HE1 CZ OH
BOND CZ CE2 CE2 HE2 CD2 HD2 CD2 CG2 CA2 C2
BOND N3 CA3 CA3 HA31 CA3 HA32 CA3 C3 N3 C1 N3 C2
BOND CA1 CB1 CB1 HB1 CB1 OG1 CB1 CG1 OG1 HG1
BOND CG1 HG11 CG1 HG12 CG1 HG13
DOUBLE C1 N2 CA2 CB2 C2 O2 C3 O3 CD1 CG2 CD2 CE2 CZ CE1 CA1 N1
```

Table 5.6 Parameter file for the chromophore of mStrawberry

```
*charmm parameter file of mStrawberry chromophore (thr-tyr-gly)
*
!parameter file

! GFP Chromophore parameters, protonated form
!
BONDS
!
!V(bond) = Kb(b - b0)**2
!
!Kb: kcal/mole/A**2
!b0: A
!
!atom type Kb          b0

CPc2 CTc1  354.000     1.4900 !ion for RFP C1-CA1 connection
NRc1 CT2   396.000     1.4400 !ion
NRc1 CPc2  400.000     1.3900 !
NRc1 CPc1  400.000     1.4100 !
CPc1 Oc2   854.000     1.2400 !ion
NRc2 CPc2  400.000     1.3000 !
CPc1 CPc1  410.000     1.4600 !ion
NRc2 CPc1  400.000     1.4000 !
CPc1 CEc1  500.000     1.3900 !ion
HAc1 CEc1  360.500     1.1000 !
CEc1 CA1   437.000     1.4100 !ion
CA1  CA2   305.000     1.4300 !ion
HPc  CA2   340.000     1.0800 !
CA2  CA3   305.000     1.3500 !ion
HPc  CA3   340.000     1.0800 !
CA3  CA4   305.000     1.4550 !ion
OcH  CA4   842.000     1.2500 !ion
CTc1 NH1   463.000     1.3650 !RFP CA1-N1 connection
CT1  CTc1  222.500     1.5000 ! ALLOW   ALI

ANGLES
!
!V(angle) = Ktheta(Theta - Theta0)**2
!
!V(Urey-Bradley) = Kub(S - S0)**2
!
```

```
!Ktheta: kcal/mole/rad**2
!Theta0: degrees
!Kub: kcal/mole/A**2 (Urey-Bradley)
!S0: A
!
!atom types     Ktheta    Theta0   Kub     S0
!

NRc2 CPc2 NRc1  130.000   113.3000 !ion
CPc2 NRc2 CPc1  130.000   106.6000 !ion
CPc2 NRc1 CPc1  130.000   107.9000 !
NRc2 CPc1 CPc1  130.000   108.3000 !
NRc2 CPc1 CEc1   45.800   129.5000 !
NRc1 CPc1 Oc2    50.000   124.0000 !ion
NRc1 CPc1 CPc1  130.000   103.0000 !
Oc2  CPc1 CPc1   44.000   133.0000 !ion
CPc1 CPc1 CEc1   45.800   122.7000 !ion
CPc1 CEc1 CA1   130.000   133.2000 !ion
CPc1 CEc1 HAc1   42.000   112.0000 !ion
CEc1 CA1  CA2    45.800   120.0000 !ion
HAc1 CEc1 CA1    42.000   115.0000 !ion
CA1  CA2  CA3    40.000   122.0000 !ion
CA2  CA1  CA2    40.000   116.0000 !ion
CA2  CA3  CA4    40.000   122.0000 !ion
CA3  CA4  CA3    40.000   115.0000 !ion
HPc  CA3  CA4    30.000   120.0000 !
HPc  CA3  CA2    30.000   120.0000 !
HPc  CA2  CA3    30.000   120.0000 !
HPc  CA2  CA1    30.000   120.0000 !
OcH  CA4  CA3    45.200   120.0000 ! ALLOW   ARO ALC

!Link to the thr(65) fragment

NH1  CTc1 CPc2   50.000   107.0000 ! ALLOW   PEP POL ARO ALI
NRc2 CPc2 CTc1   40.000   125.0000 !
NRc1 CPc2 CTc1   40.000   121.7000 !ion
CT1  CTc1 CPc2   52.000   108.0000 ! ALLOW   ALI PEP POL ARO

CTc1 NH1  C      50.000   120.0000
NH1  CTc1 CT1    70.000   113.5000
HA   CT1  CTc1   34.500    110.10   22.53   2.17900 ! ALLOW   ALI
CT3  CT1  CTc1   53.350    108.50    8.00   2.56100 ! ALLOW   ALI
OH1  CT1  CTc1   75.700   110.1000 ! ALLOW   ALI ALC ARO  (Yuba)

!Link to the gly(67) fragment
CPc2 NRc1 CT2    40.000   128.3000 !ion
CPc1 NRc1 CT2    40.000   123.8000 !ion
NRc1 CT2  C      50.000   107.0000
NRc1 CT2  HB     48.000   108.0000
!
DIHEDRALS
!
!V(dihedral) = Kchi(1 + cos(n(chi) - delta))
!
!Kchi: kcal/mole
```

```
!n: multiplicity
!delta: degrees
!
!atom types             Kchi    n   delta
!

CPc2 NRc2 CPc1 CPc1    14.0000  2   180.00 !
CPc2 NRc1 CPc1 CPc1    14.0000  2   180.00 !
NRc2 CPc2 NRc1 CPc1    14.0000  2   180.00 !
NRc2 CPc1 CPc1 NRc1     4.0000  2   180.00 !
NRc1 CPc2 NRc2 CPc1     4.0000  2   180.00 !
CA1  CA2  CA3  CA4      3.1000  2   180.00 !
CA2  CA1  CA2  CA3      3.1000  2   180.00 !
CA2  CA3  CA4  CA3      3.1000  2   180.00 !
!Oc2  CAC  CAC  CAC      3.1000  2   180.00 !
CA2  CA3  CA4  OcH      3.1000  2   180.00 !
CA1  CA2  CA3  HPc      4.2000  2   180.00 !
CA2  CA1  CA2  HPc      4.2000  2   180.00 !
CA3  CA4  CA3  HPc      4.2000  2   180.00 !
HPc  CA2  CA3  CA4      4.2000  2   180.00 !
HPc  CA2  CA3  HPc      2.4000  2   180.00 !
!HcH  Oc2  CAC  CAC      0.9900  2   180.00 !
!HcH  OcH  CA4  CA3      0.9900  2   180.00 !
HPc  CA3  CA4  OcH      4.2000  2   180.00 !
!
CPc2 NRc2 CPc1 CEc1     3.000   2   180.00 !
NRc1 CPc1 CPc1 CEc1     3.00    2   180.00 !
Oc2  CPc1 CPc1 CEc1     2.00    2   180.00 !
CEc1 CA1  CA2  HPc      4.20    2   180.00 !
CEc1 CA1  CA2  CA3      3.10    2   180.00 !

!connection CA-CB
CPc1 CPc1 CEc1 HAc1     3.9000  2   180.00 !
CPc1 CPc1 CEc1 CA1      3.9000  2   180.00 !
NRc2 CPc1 CEc1 HAc1     3.9000  2   180.00 !
NRc2 CPc1 CEc1 CA1      3.9000  2   180.00 !

!connection CB-CG2
CPc1 CEc1 CA1  CA2      2.7000  2   180.00 !
HAc1 CEc1 CA1  CA2      2.7000  2   180.00 !
!
CPc2 NRc1 CPc1 Oc2     14.0000  2   180.00 !
NRc2 CPc2 NRc1 CT2     14.0000  2   180.00 !
NRc2 CPc1 CPc1 Oc2     14.0000  2   180.00 !
CPc1 NRc1 CPc2 CTc1    14.0000  2   180.00 !
Oc2  CPc1 NRc1 CT2     14.0000  2   180.00 !
CPc1 NRc2 CPc2 CTc1    14.0000  2   180.00 !
CPc1 CPc1 NRc1 CT2     14.0000  2   180.00 !
CTc1 CPc2 NRc1 CT2     14.0000  2   180.00 !

! Linking the chromophore and the glycine(67) fragment
O    C    CT2  NRc1     0.0000  1     0.00 !
NH1  C    CT2  NRc1     0.6000  1     0.00 !
CPc2 NRc1 CT2  HB       0.0320  3     0.00 !
CPc2 NRc1 CT2  C        0.0320  3     0.00 !
```

```
CPc1 NRc1 CT2  HB       0.0320  3   180.00 !
CPc1 NRc1 CT2  C        0.0320  3   180.00 !

! Linking the chromophore and the thr(65) fragment
C    NH1  CTc1 CPc2     2.2500  2   180.00 !Taken from X-C-NC2-X
Charmm22
NRc2 CPc2 CTc1 CT1      0.1050  3   180.00 !
NRc2 CPc2 CTc1 NH1      0.1050  3   180.00 !
NRc1 CPc2 CTc1 CT1      0.1050  3     0.00 !
NRc1 CPc2 CTc1 NH1      0.1050  3     0.00 !

!connecting N1=CA1 region due to new type CTc1
O    C    NH1  CTc1     2.5000  2   180.00
CT1  C    NH1  CTc1     2.5000  2   180.00
CT1  CTc1 NH1  C        1.8000  1     0.00
CPc2 CTc1 CT1  HA       0.2000  3     0.00
CPc2 CTc1 CT1  CT3      0.2000  3     0.00
NH1  CTc1 CT1  HA       0.2000  3     0.00
NH1  CTc1 CT1  CT3      0.2000  3     0.00   ! X    CT1  CT1  X  (Yuba)
CPC2 CTC1 CT1  OH1      0.2000  3     0.00 ! ALLOW   ALI
NH1  CTC1 CT1  OH1      0.2000  3     0.00 ! ALLOW   ALI

IMPROPER
!
!V(improper) = Kpsi(psi - psi0)**2
!
!Kpsi: kcal/mole/rad**2
!psi0: degrees
!note that the second column of numbers (0) is ignored
!
!atom types           Kpsi                   psi0
!

CPc2 NRc2 NRc1 CTc1    50.0000         0      0.0000
CPc2 NRc1 NRc2 CTc1    50.0000         0      0.0000
!
CPc1 NRc1 CPc1 Oc2     50.0000         0      0.0000
CPc1 CPc1 NRc1 Oc2     50.0000         0      0.0000
!
NRc1 CPc1 CPc2 CT2     50.0000         0      0.0000
NRc1 CPc2 CPc1 CT2     50.0000         0      0.0000
!
CPc1 NRc2 CPc1 CEc1    50.0000         0      0.0000
CPc1 CPc1 NRc2 CEc1    50.0000         0      0.0000
!
CEc1 CPc1 CA1  HAc1    30.0000         0      0.0000
CEc1 CA1  CPc1 HAc1    30.0000         0      0.0000



!
!V(Lennard-Jones) = Eps,i,j[(Rmin,i,j/ri,j)**12 - 2(Rmin,i,j/ri,j)**6]
!
!epsilon: kcal/mole, Eps,i,j = sqrt(eps,i * eps,j)
!Rmin/2: A, Rmin,i,j = Rmin/2,i + Rmin/2,j
```

```
!
!atom  ignored    epsilon      Rmin/2   ignored   eps,1-4     Rmin/2,1-4
!
!CAc     5.000000  -0.070000    1.992400 ! ALLOW   ARO

NONBONDED nbxmod  5 atom cdiel shift vatom vdistance vswitch -
cutnb 14.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5

CA1    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CA2    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CA3    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CA4    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CEc1   0.000000  -0.068000     2.090000 !
CPc1   0.000000  -0.050000     1.800000 ! ALLOW ARO
CPc2   0.000000  -0.050000     1.800000 ! ALLOW ARO
HAc    0.000000  -0.022000     1.320000 ! ALLOW PEP ALI POL SUL
HAc1   0.000000  -0.031000     1.250000 !
HPc    0.000000  -0.030000     1.358200 0.000000  -0.030000    1.358200
NRc1   0.000000  -0.200000     1.850000 ! ALLOW ARO
NRc2   0.000000  -0.200000     1.850000 ! ALLOW ARO
Oc2    0.000000  -0.120000     1.700000 ! ALLOW   PEP POL,
OcH    0.000000  -0.152100     1.770000 ! ALLOW   ALC ARO
CTc1   0.000000  -0.020000  2.275000 ! 0.000000  -0.010000     1.900000
```

Appendix E: Residue topology and parameter files for the chromophore of citrine

Table 5.7 Residue topology file for the chromophore of citrine

```
MASS   197 NRc2  14.00700 N ! neutral his unprotonated ring nitrogen
MASS   198 NRc1  14.00700 N ! neutral his protonated ring nitrogen
MASS   199 HAc1   1.00800 H ! for alkene; RHC=CR
MASS   200 HPc    1.00800 H ! aromatic H
MASS   201 Oc2   15.99900 O ! carbonyl oxygen
MASS   202 OcH   15.99900 O ! from OH1
MASS   204 HAc    1.00800 H ! nonpolar H
MASS   205 CA1   12.01100 C ! aromatic C
MASS   206 CA2   12.01100 C ! aromatic C
MASS   207 CA3   12.01100 C ! aromatic C
MASS   208 CPc2  12.01100 C ! his CE1 carbon
MASS   209 CEc1  12.01100 C ! for alkene; RHC=CR
MASS   210 CPc1  12.01100 C ! his CG and CD2 carbons
MASS   211 CA4   12.01100 C ! aromatic C
MASS   213 CTc1  12.01100 C ! aliphatic sp3 C for CH

DECL -CA
DECL -C
DECL -O
DECL -C3 !Chola
DECL +N1 !Chola
DECL +N
```

```
DECL +HN
DECL +CA

DEFA FIRS NTER LAST CTER
AUTO ANGLES DIHE

RESI CRO  -1.000
GROUP                    ! Imidazolinone ring
ATOM C1    CPc2   0.50
ATOM N2    NRc2  -0.60
ATOM N3    NRc1  -0.57
ATOM C2    CPc1   0.57
ATOM O2    Oc2   -0.57
ATOM CA2   CPc1   0.10
ATOM CB2   CEc1  -0.14
ATOM HB2   HAc1   0.21
ATOM CG2   CA1   -0.09 ! Tyr ring : charges from charmm22
ATOM CD1   CA2   -0.08
ATOM HD1   HPc    0.14
ATOM CD2   CA2   -0.08
ATOM HD2   HPc    0.14
ATOM CE1   CA3   -0.28
ATOM HE1   HPc    0.10
ATOM CE2   CA3   -0.28
ATOM HE2   HPc    0.10
ATOM CZ    CA4    0.45
ATOM OH    OcH   -0.62

!
!Glycine (67) part from Charmm22
GROUP
ATOM CA3   CT2   -0.18  !     |
ATOM HA31  HB     0.09  !     |
ATOM HA32  HB     0.09  ! HA1-CA-HA2
GROUP                   !     |
ATOM C3    C      0.51  !     |
ATOM O3    O     -0.51  !    C=O
!
!gly (65) part from charmm22
GROUP
ATOM N1    NH1    -0.47 !atom type changed
ATOM HN1   H       0.31
ATOM CA1   CTc1   -0.02
ATOM HA11  HB      0.09
ATOM HA12  HB      0.09

BOND CA1 C1 N1 -C C3 +N CA1 N1 N1 HN1 CA1 HA11 CA1 HA12
BOND N2 CA2 CB2 HB2 CB2 CG2 CD1 HD1 CD1 CE1 CE1 HE1 CZ OH
BOND CZ CE2 CE2 HE2 CD2 HD2 CD2 CG2 CA2 C2
BOND N3 CA3 CA3 HA31 CA3 HA32 CA3 C3 N3 C1 N3 C2
DOUBLE C1 N2 CA2 CB2 C2 O2 C3 O3 CD1 CG2 CD2 CE2 CZ CE1
```

Table 5.8 Parameter file for the chromophore of citrine

```
*charmm parameter file of citrine chromophore (gly-tyr-gly)
*
!parameter file

! GFP Chromophore parameters, protonated form
!
BONDS
!
!V(bond) = Kb(b - b0)**2
!
!Kb: kcal/mole/A**2
!b0: A
!
!atom type Kb          b0

CPc2 CTc1  354.000     1.4900 !ion for RFP C1-CA1 connection
NRc1 CT2   396.000     1.4400 !ion
NRc1 CPc2  400.000     1.3900 !
NRc1 CPc1  400.000     1.4100 !
CPc1 Oc2   854.000     1.2400 !ion
NRc2 CPc2  400.000     1.3000 !
CPc1 CPc1  410.000     1.4600 !ion
NRc2 CPc1  400.000     1.4000 !
CPc1 CEc1  500.000     1.3900 !ion
HAc1 CEc1  360.500     1.1000 !
CEc1 CA1   437.000     1.4100 !ion
CA1  CA2   305.000     1.4300 !ion
HPc  CA2   340.000     1.0800 !
CA2  CA3   305.000     1.3500 !ion
HPc  CA3   340.000     1.0800 !
CA3  CA4   305.000     1.4550 !ion
OcH  CA4   842.000     1.2500 !ion

NH1  CTc1  320.000     1.4300
HB   CTc1  330.000     1.0800

ANGLES
!
!V(angle) = Ktheta(Theta - Theta0)**2
!
!V(Urey-Bradley) = Kub(S - S0)**2
!
!Ktheta: kcal/mole/rad**2
!Theta0: degrees
!Kub: kcal/mole/A**2 (Urey-Bradley)
!S0: A
!
!atom types     Ktheta    Theta0   Kub     S0
!

NRc2 CPc2 NRc1  130.000   113.3000 !ion
CPc2 NRc2 CPc1  130.000   106.6000 !ion
```

```
CPc2 NRc1 CPc1  130.000   107.9000 !
NRc2 CPc1 CPc1  130.000   108.3000 !
NRc2 CPc1 CEc1   45.800   129.5000 !
NRc1 CPc1 Oc2    50.000   124.0000 !ion
NRc1 CPc1 CPc1  130.000   103.0000 !
Oc2  CPc1 CPc1   44.000   133.0000 !ion
CPc1 CPc1 CEc1   45.800   122.7000 !ion
CPc1 CEc1 CA1   130.000   133.2000 !ion
CPc1 CEc1 HAc1   42.000   112.0000 !ion
CEc1 CA1  CA2    45.800   120.0000 !ion
HAc1 CEc1 CA1    42.000   115.0000 !ion
CA1  CA2  CA3    40.000   122.0000 !ion
CA2  CA1  CA2    40.000   116.0000 !ion
CA2  CA3  CA4    40.000   122.0000 !ion
CA3  CA4  CA3    40.000   115.0000 !ion
HPc  CA3  CA4    30.000   120.0000 !
HPc  CA3  CA2    30.000   120.0000 !
HPc  CA2  CA3    30.000   120.0000 !
HPc  CA2  CA1    30.000   120.0000 !
OcH  CA4  CA3    45.200   120.0000 ! ALLOW   ARO ALC

!Link to the gly(65) fragment
NH1  CTc1 CPc2   50.000   107.0000 ! ALLOW   PEP POL ARO ALI
HB   CTc1 CPc2   50.000   109.5000 ! ALLOW   ALI PEP POL ARO
NRc2 CPc2 CTc1   40.000   125.0000 !
NRc1 CPc2 CTc1   40.000   121.7000 !ion
CTc1 NH1  C      50.000   120.0000
NH1  CTc1 HB     48.000   108.0000
H    NH1  CTc1   35.000   117.0000
HB   CTc1 HB     36.000   115.0000

!Link to the gly(67) fragment
CPc2 NRc1 CT2    40.000   128.3000 !ion
CPc1 NRc1 CT2    40.000   123.8000 !ion
NRc1 CT2  C      50.000   107.0000
NRc1 CT2  HB     48.000   108.0000

!
DIHEDRALS
!
!V(dihedral) = Kchi(1 + cos(n(chi) - delta))
!
!Kchi: kcal/mole
!n: multiplicity
!delta: degrees
!
!atom types             Kchi    n   delta
!

CPc2 NRc2 CPc1 CPc1    14.0000  2   180.00 !
CPc2 NRc1 CPc1 CPc1    14.0000  2   180.00 !
NRc2 CPc2 NRc1 CPc1    14.0000  2   180.00 !
NRc2 CPc1 CPc1 NRc1     4.0000  2   180.00 !
NRc1 CPc2 NRc2 CPc1     4.0000  2   180.00 !
CA1  CA2  CA3  CA4      3.1000  2   180.00 !
```

```
CA2  CA1  CA2  CA3      3.1000  2   180.00 !
CA2  CA3  CA4  CA3      3.1000  2   180.00 !
!Oc2  CAC  CAC  CAC      3.1000  2   180.00 !
CA2  CA3  CA4  OcH      3.1000  2   180.00 !
CA1  CA2  CA3  HPc      4.2000  2   180.00 !
CA2  CA1  CA2  HPc      4.2000  2   180.00 !
CA3  CA4  CA3  HPc      4.2000  2   180.00 !
HPc  CA2  CA3  CA4      4.2000  2   180.00 !
HPc  CA2  CA3  HPc      2.4000  2   180.00 !
HPc  CA3  CA4  OcH      4.2000  2   180.00 !
!
CPc2 NRc2 CPc1 CEc1     3.000   2   180.00 !
NRc1 CPc1 CPc1 CEc1     3.00    2   180.00 !
Oc2  CPc1 CPc1 CEc1     2.00    2   180.00 !
CEc1 CA1  CA2  HPc      4.20    2   180.00 !
CEc1 CA1  CA2  CA3      3.10    2   180.00 !

!connection CA-CB
CPc1 CPc1 CEc1 HAc1     3.9000  2   180.00 !
CPc1 CPc1 CEc1 CA1      3.9000  2   180.00 !
NRc2 CPc1 CEc1 HAc1     3.9000  2   180.00 !
NRc2 CPc1 CEc1 CA1      3.9000  2   180.00 !

!connection CB-CG2
CPc1 CEc1 CA1  CA2      2.7000  2   180.00 !
HAc1 CEc1 CA1  CA2      2.7000  2   180.00 !
!
CPc2 NRc1 CPc1 Oc2     14.0000  2   180.00 !
NRc2 CPc2 NRc1 CT2     14.0000  2   180.00 !
NRc2 CPc1 CPc1 Oc2     14.0000  2   180.00 !
CPc1 NRc1 CPc2 CTc1    14.0000  2   180.00 !
Oc2  CPc1 NRc1 CT2     14.0000  2   180.00 !
CPc1 NRc2 CPc2 CTc1    14.0000  2   180.00 !
CPc1 CPc1 NRc1 CT2     14.0000  2   180.00 !
CTc1 CPc2 NRc1 CT2     14.0000  2   180.00 !

! Linking the chromophore and the glycine(67) fragment
O    C    CT2  NRc1     0.0000  1     0.00 !
NH1  C    CT2  NRc1     0.6000  1     0.00 !
CPc2 NRc1 CT2  HB       0.0320  3     0.00 !
CPc2 NRc1 CT2  C        0.0320  3     0.00 !
CPc1 NRc1 CT2  HB       0.0320  3   180.00 !
CPc1 NRc1 CT2  C        0.0320  3   180.00 !


! Linking the chromophore and the gly(65) fragment
C    NH1  CTc1 CPc2     0.2000  1   180.00 !
NRc2 CPc2 CTc1 NH1      0.1050  3   180.00 !
NRc1 CPc2 CTc1 NH1      0.1050  3     0.00 !
!
H    NH1  CTc1 CPc2     0.0000  1     0.00 !
NRc2 CPc2 CTc1 HB       0.1050  3   180.00 !
NRc1 CPc2 CTc1 HB       0.1050  3     0.00 !
CT1  C    NH1  CTc1     1.6000  1     0.00 !  ALLOW PEP
HB   CTc1 NH1  C        0.0000  1     0.00 ! ALLOW PEP
```

```
O    C    NH1  CTc1     2.5000  2   180.00 !  ALLOW PEP
HB   CTc1 NH1  H        0.0000  1     0.00 ! ALLOW PEP



IMPROPER
!
!V(improper) = Kpsi(psi - psi0)**2
!
!Kpsi: kcal/mole/rad**2
!psi0: degrees
!note that the second column of numbers (0) is ignored
!
!atom types           Kpsi                   psi0
!

CPc2 NRc2 NRc1 CTc1   50.0000      0      0.0000
CPc2 NRc1 NRc2 CTc1   50.0000      0      0.0000
!
CPc1 NRc1 CPc1 Oc2    50.0000      0      0.0000
CPc1 CPc1 NRc1 Oc2    50.0000      0      0.0000
!
NRc1 CPc1 CPc2 CT2    50.0000      0      0.0000
NRc1 CPc2 CPc1 CT2    50.0000      0      0.0000
!
CPc1 NRc2 CPc1 CEc1   50.0000      0      0.0000
CPc1 CPc1 NRc2 CEc1   50.0000      0      0.0000
!
CEc1 CPc1 CA1  HAc1   30.0000      0      0.0000
CEc1 CA1  CPc1 HAc1   30.0000      0      0.0000



!
!V(Lennard-Jones) = Eps,i,j[(Rmin,i,j/ri,j)**12 - 2(Rmin,i,j/ri,j)**6]
!
!epsilon: kcal/mole, Eps,i,j = sqrt(eps,i * eps,j)
!Rmin/2: A, Rmin,i,j = Rmin/2,i + Rmin/2,j
!
!atom  ignored    epsilon      Rmin/2   ignored   eps,1-4     Rmin/2,1-4
!
!CAc     5.000000  -0.070000    1.992400 ! ALLOW   ARO

NONBONDED nbxmod  5 atom cdiel shift vatom vdistance vswitch -
cutnb 14.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5

CA1    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CA2    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CA3    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CA4    5.000000  -0.070000     1.992400 ! ALLOW   ARO
CEc1   0.000000  -0.068000     2.090000 ! ! for propene, yin/adm jr.,
CPc1   0.000000  -0.050000     1.800000 ! ALLOW ARO
CPc2   0.000000  -0.050000     1.800000 ! ALLOW ARO
HAc    0.000000  -0.022000     1.320000 ! ALLOW PEP ALI POL SUL
HAc1   0.000000  -0.031000     1.250000 !
HPc    0.000000  -0.030000     1.358200 0.000000  -0.030000    1.358200
```

```
NRc1    0.000000  -0.200000     1.850000 ! ALLOW ARO
NRc2    0.000000  -0.200000     1.850000 ! ALLOW ARO
Oc2     0.000000  -0.120000     1.700000 ! ALLOW   PEP POL,
OcH     0.000000  -0.152100     1.770000 ! ALLOW   ALC ARO
CTc1    0.000000  -0.020000     2.275000 !
```

VITA

YUBA R. BHANDARI

2008 – 2013	Teaching Assistant Florida International University Miami, Florida, USA
2004 – 2005	Master of Science (Physics) Tribhuvan University Kathmandu, Nepal
2001 – 2003	Bachelor of Science (Physics) Tribhuvan University Bharatpur, Nepal

LIST OF PUBLICATIONS IN REFERRED JOURNALS:

1. Chola K. Regmi, Yuba R. Bhandari, Bernard S. Gerstman and Prem P. Chapagain "Exploring the diffusion of molecular oxygen in the red fluorescent protein mCherry using explicit oxygen molecular dynamics simulations", Journal of Physical Chemistry B, 2013, 117 (8), 2247–2253

2. Yuba R. Bhandari, Prem P. Chapagain and Bernard S. Gerstman “Lattice model simulations of the effects of the position of a peptide trigger segment on helix folding and dimerization” Journal of Chemical Physics 137, 105103 (2012).

3. Ilan E. Chemmama, Adam Colt Pelea, Yuba R. Bhandari, Prem P. Chapagain and Bernard S. Gerstman “Structural propensities and entropy effects in peptide helix-coil transitions” Physical Review E 86, 031915 (2012).

4. Prem P. Chapagain, Bernard Gerstman, Yuba R. Bhandari, and Dipak Rimal, “Free-energy landscapes and thermodynamic parameters of complex molecules from nonequilibrium simulation trajectories”, Physical Review E 83, 061905 (2011).

ORAL/POSTER/PARTICIPATION IN CONFERENCES/SEMINARS:

1. American Physical Society March Meeting, Baltimore, March 2013 “Structural Transformation and Aggregation of cc-beta Peptides into Amyloid Proto-fibrils”
2. American Physical Society March Meeting, Boston, February 2012
    a) “Effects of pressure on the protein barrel and the chromophore interactions in mCherry”
    b) “Barrel fluctuation and oxygen diffusion pathways in the monomeric fluorescent proteins”
3. American Physical Society March Meeting, Dallas, March 2011 “Heterogeneous helical propensity and its effects on dimerization and the stability of a model protein dimer”
4. American Physical Society March Meeting, Portland, March 2010 “Position dependence of the trigger sequence in the folding and dimerization of a lattice model coiled-coil peptide”
5. Presentation on “Protein Folding” at Central Department of Physics, Tribhuvan University, Nepal on August 2005 organized by the Students Association of Physics.