May 2007

The Effect of the Protein Matrix on Fluorescence and an Analysis of the Effect of Fusing Residues 62 and 63 in Kindling Proteins

Scott L. Maddalo
Connecticut College, slmad@conncoll.edu

The Effect of the Protein Matrix on Fluorescence and an Analysis of the Effect of Fusing Residues 62 and 63 in Kindling Proteins

Scott Maddalo
Connecticut College
May 2007

Advisor:
Marc Zimmer

This work is dedicated to my grandparents
who have supported me my entire life.

Table of Contents

Acknowledgments

I would first like to thank Marc Zimmer, PhD for asking me to do research with him after my freshman year. I had no expectations of majoring in Chemistry when I came to Connecticut College but you gave me a chance to shine and for that I can never thank you enough. My appreciation for research will always be traced back to you.

To my readers, Stanton Ching, PhD and Maureen Ronau, thank you for all of your suggestions on my work. You helped make this a much better piece of writing than I thought it was going to be when I started.

One team I cannot leave unmentioned is the Zimmer research team. Many people have come and gone and I would like to give notice to Nate Lemay (thesis is finally done!), Andrew Weber, Justin Rosenberg, and Curren Mbofana. All of you were great to work with and I wish you the best of luck in your future endeavors.

There are many other people I should thank but three people I want to highlight are Bob Kyne, George Arab, and Kurt Luthy. Bob, we have formed a very strong bond over the past few years, no pun intended, and I know that our friendship will be a lasting one. George, I wish you the best of luck in Rochester and hopefully I will hear about the great work you do in the future. Kurt, I have spent far too much time at the help desk working on Chemistry with you until that damn bell rings at least a hundred times before we leave but I do not regret a minute of it.

Lastly I want to thank someone who I have never gotten to know well but has a made a profound impact on my life. Timo Ovaska, PhD, your courses have given me confidence in myself that I never had before. I will always look back on your courses as my foundation as a student.

Abstract

In the ground state of the highly conjugated green fluorescent protein (GFP), the chromophore should be planar. However, numerous crystal structures of GFP and GFP-like proteins have been reported with slightly twisted chromophores. We have previously shown that the protein cavity surrounding the chromophore in wild-type GFP is not complementary with a planar chromophore. This study shows that the crystal structure of wild-type GFP is not an anomaly: most of the GFP and GFP-like proteins in the protein databank have a protein matrix that is not complementary with a planar chromophore. When the π-conjugation across the ethylenic bridge of the chromophore is removed the protein matrix will significantly twist the freely rotating chromophore from the relatively planar structures found in the crystal structures. The possible consequences of this non-planar deformation on the photophysics of GFP are discussed.

In addition to GFP there are GFP-like proteins that can be reversibly photoswitched between a fluorescent and a nonfluorescent state. Conformational searching and molecular dynamics are used to examine the various kindling proteins with a cleaved chromophore. MD simulations supported that the chromophore does indeed start to move away from its initial *trans* configuration, and actually completes a full *trans*/*cis* isomerization. The nearby His197 residue does not provide a large barrier to the chromophore rotation when it is cleaved from the protein backbone. This implies that the rest of the protein backbone did not provide an obstacle to the isomerization. Fusion of the chromophore back to the protein backbone was also done and conformational analyses showed that the proteins with fused and cleaved chromophores did now show a great deal of change the φ and τ dihedral angles most cases.

1. Introduction

1.1 GFP

Green fluorescent protein (GFP) is a protein found in the jellyfish *Aequorea victoria*. It is commonly used as an *in vivo* fluorescent marker to help facilitate biomedical studies [1].

Due to its stability, it is widely used as a biological marker. Another interesting feature is the fact that its chromophore (see Figure 1) is formed in an autocatalytic cyclization that does not require a cofactor. GFP has also gained widespread interest in biochemistry and cell biology since it can be used as a cloneable, noninvasive marker for gene expression and protein localization in intact cells and organisms. It is interesting to note that unlike other bioluminescent reporters, GFP fluoresces in the absence of any other proteins or substrates. However the fusion of GFP to a protein does not alter the function or location of that protein [2].

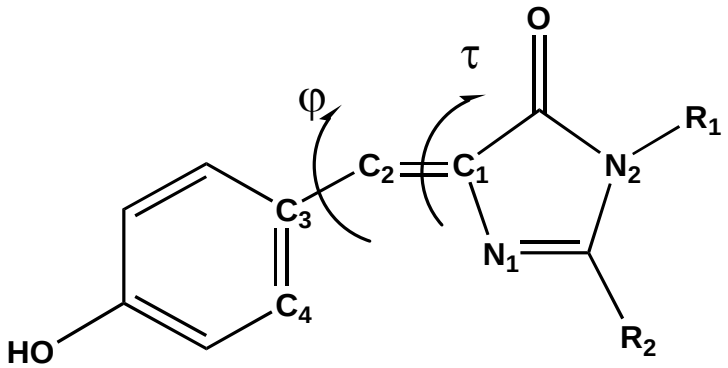

Figure 1. The τ (N_1-C_1-C_2-C_3) and φ (C_1-C_2-C_3-C_4) dihedral angles of the GFP chromophore. In the protein R_1 is Gly67 from to the end of protein and R_2 is start of protein to Ser65, and in HBDI, an often used model compound, R_1=R_2=CH_3. In τ one-bond-flips (τ-OBF) the dihedral rotation occurs around the τ torsional angle, in a φ-OBF it is around the φ dihedral angle, in a positively correlated hula-twist (+HT) the φ and τ dihedral angles concertedly rotate in the same direction (as shown above), while in a negatively correlated hula-twist (-HT) they concertedly rotate in opposite directions.

The structure of GFP can be described as a *light in a can.* The chromophore is located in the center of a can that consists of 11 β sheets. The can is a nearly perfect cylinder with a height of 42 Å and a radius of 12 Å [2].

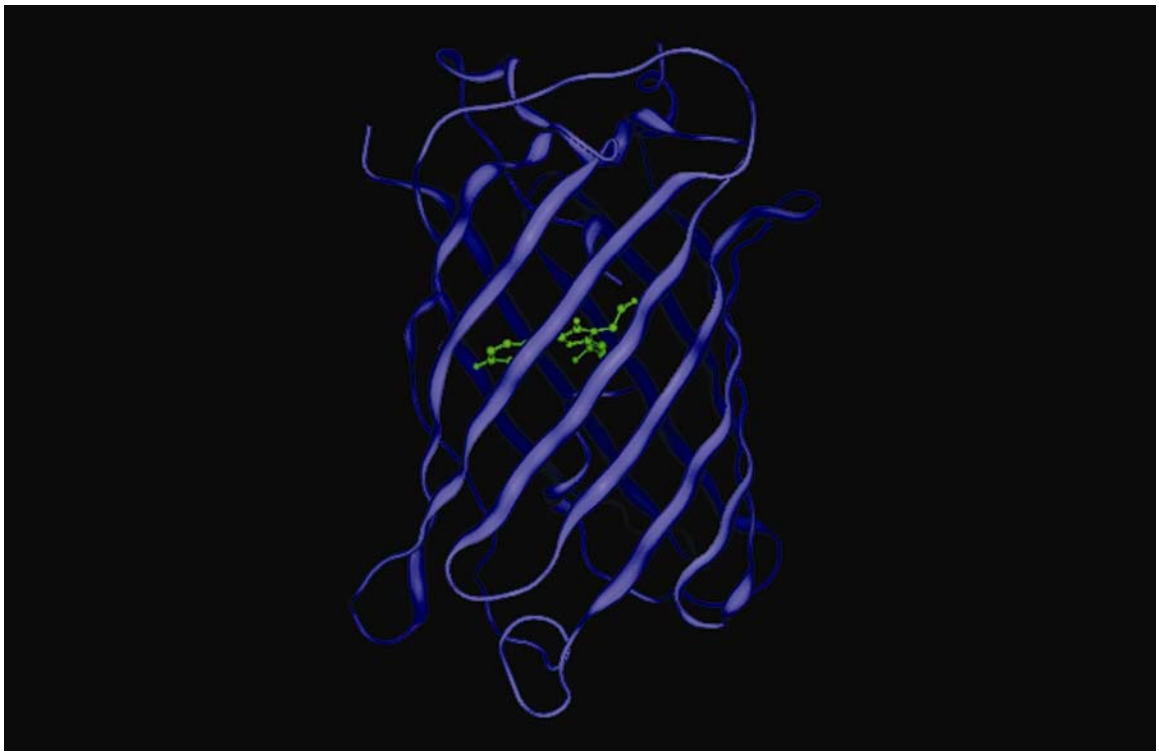

Figure 2. The basic light in a can shape is comprised of 11 beta sheets. The chromophore is located in the center of the can.

The intensive investigation of the GFP protein family began about 12 years ago. Since then many tests have been done to decipher how GFP and its many variants fluoresce. GFP is such an interesting protein because of its many potential applications in biotechnology [3], plants [4], structure and dynamics [5], reporter gene technology [6], cell biology [7], and in the drug industry [8]. Recent books about GFP have documented these uses [9, 10].

In the last ten years green fluorescent protein (GFP) has changed from a nearly unknown protein to a commonly used molecular imaging tool in biology, chemistry,

genetics and medicine. Probably the best indicator of the utility of GFP and GFP-like proteins is the fact that in 2004 about 50%, 35%, 60% and 20% of the articles in *Cell, Development, Journal of Cell Biology* and *Neuron* respectively, mentioned or used GFP-like proteins[11].

1.1.1 Photophysical Behavior of Green Fluorescent Protein

The green fluorescent protein (GFP) chromophore does not fluoresce unless there is a protein surrounding it [12]. Changing the amino acid residues around the chromophore can change the color and intensity of GFP's fluorescence [13,14]. The protein matrix of GFP has both a steric and an electronic component that influences the GFP chromophore. In this thesis the steric influence the protein matrix has on GFP chromophore will be examined along with the hydrogen bonds that the chromophore makes to the protein matrix.

GFP can absorb at two different wavelengths due to its existence in two different substates. The chromophore can exist in the neutral phenolic form (A state) or in the anionic phenolate form (B state). Wild-type GFP has a major absorption at 398nm [15] and a minor absorption at 475nm with a shoulder on the red edge [16,17]. Excitation at 398nm results in an emission maximum at 508nm, while irradiation at 475nm produces an emission with a maximum at 503nm [18]. A special mechanism has been proposed for the interconversion between the different forms of wild-type GFP [18-20]. In short, the neutral form of the chromophore can convert to the anionic species (B) by going through an intermediate state (I). In going from the neutral chromophore (species A) to the charged chromophore (B) the Tyr66 phenolic proton is shuttled through an extensive hydrogen bonding network (His148, Thr203, Ser205 and 2 waters) to the carboxylate oxygen of

Glu222.

The photophysical behavior of GFP can be complicated by transitions between bright and dark fluorescent states. At the single molecule level these transitions are responsible for the reversible fast blinking and photobleaching that has been observed in single protein experiments [21-23]. The most commonly accepted models used to explain these observations are based on non-radiative relaxation pathways between the exited and ground state that involve torsional changes of the φ and τ dihedrals of the chromophore shown in Figure 1. A model for the light/dark behavior of GFP has been proposed [24]. It is based on quantum mechanical (QM) calculations of the energy barriers for the φ and τ one bond flips (OBF) and the φ/τ hula-twists (HT) that were calculated in the ground and first singlet excited states for a small non-peptide model compound. While the ground state minima of the GFP chromophore is clearly planar this is not necessarily so for the excited state. In some cases the excited state has an energy minimum with a twisted chromophore in which both rings are 90° to each other. According to the calculations the energies for the ground and excited states for the τ OBF and HT in the neutral phenolic form and the φ OBF in the zwitterionic form come very close to each other. It has been proposed that this can lead to fluorescence quenching nonadiabatic crossing (NAC), see Figure 3.

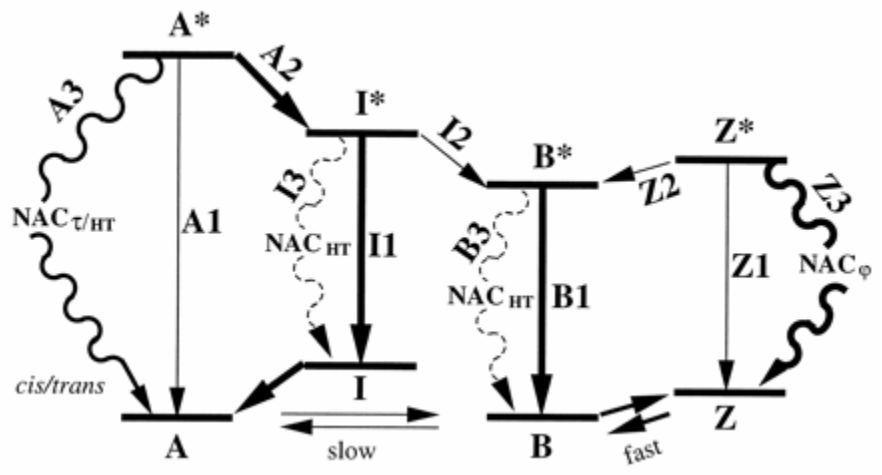

Figure 3. Model for the photophysics of GFP as inferred from our quantum chemical calculations. Excited states are labeled by asterisks. Note that barriers may exist for processes of types 2 and 3. Excitation arrows are omitted for simplicity. The relative free energies of ground state forms A, B, I, and Z depend on the protein environment and, thus, on the specific mutant [24].

The phenol and imidazolinone rings of the chromophore can twist relative to one another before undergoing fluorescence quenching NAC and relaxing back to the A state. It is also possible that the two rings systems could continue their rotation past the perpendicular form that undergoes NAC, and form the *trans* isomer. Recently a nonfluorescent dark state, state C, has been observed that is distinct from states A and B and absorbs at higher energies [25]. The chromophore in the excited state can freely twist from the planar configuration to a conformation where the two ring planes are perpendicular to each other. The NAC occurs at this point and the excited state can relax to the ground state without any fluorescence.

1.1.2 Role of the Protein Matrix

Model compounds of the chromophore do not fluoresce in solution. This is presumably due to the lack of constraints imposed by the protein. The excited state of the model compounds may freely rotate around their φ and τ dihedral angle, which allows NAC to occur, resulting in fluorescence quenching, see Figure 4A.

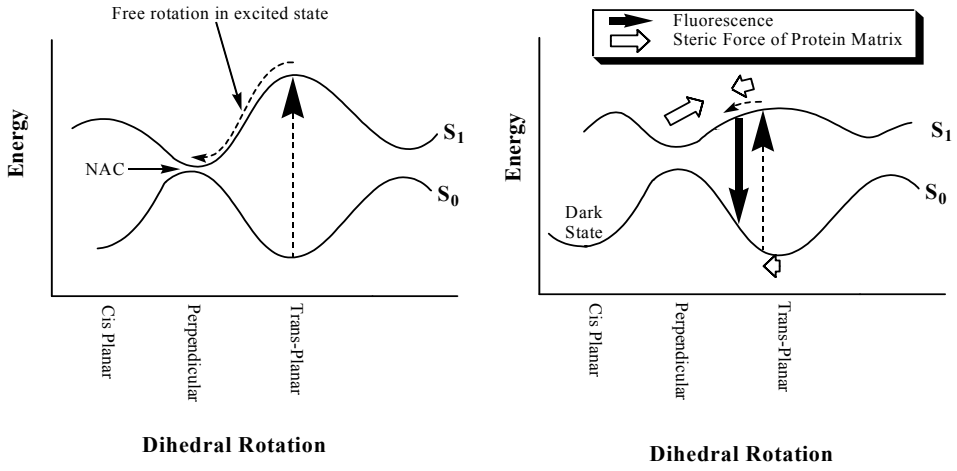

Figure 4A. Model compounds of the GFP chromophore in the ground state (S_o) can be excited to the first singlet state (S_1) in which a HT or OBF can freely occur. Upon reaching the perpendicularly twisted conformation fluorescence quenching NAC occurs.

Figure 4B. In the ground state (S_o) the residues surrounding the GFP chromophore exert a twisting force on the chromophore. Upon excitation the conjugation across the ethylenic bridge of the chromophore is reduced and it will twist, however the protein matrix prevents the chromophore from reaching the perpendicularly twisted conformation and fluorescence quenching NAC is prevented.

In the ground state, the highly conjugated GFP chromophore should be planar; however numerous crystal structures of GFP and GFP-like proteins have been reported with slightly twisted chromophores. The amino acid residues surrounding the chromophore are therefore not complementary with a planar chromophore and they exert a steric strain on the chromophore. This deviation from planarity should have an effect on the fluorescence of the chromophore.

In the excited state, the protein matrix presumably prevents the chromophore from rotating to the perpendicularly twisted conformation that has been postulated to be the conformation that leads to fluorescence quenching NAC. However, the protein also

exerts a steric force on the chromophore, twisting it away from planarity. The interplay between these forces and electronic structure of the excited chromophore will determine the excited state conformation of the fluorescing chromophore, see Figure 4B.

Based on calculations of a photoactive yellow protein (PYP), it has been concluded that this protein also prevents the chromophore from adopting a completely planar structure [26]. Yamada et al. proposed that the efficiency of photoisomerization in PYP is due to the asymmetric protein-chromophore interaction that serves as the initial accelerant for the light induced photocycle. This is very similar to our GFP findings.

One of the roles of the protein matrix in GFP is therefore to prevent the chromophore from adopting a twisted excited state conformation that can undergo fluorescence quenching NAC, see Figure 4B. It also limits *cis-trans* isomerization to the dark *trans* conformation.

Recently we have shown that wild-type GFP is not an anomaly, most of the GFP and GFP-like proteins in the PDB have a protein matrix that is not complementary with a planar chromophore [27]. When the π conjugation across the ethylenic bridge of the chromophore is removed, the protein matrix will significantly twist the freely rotating chromophore from the planar structures found in the crystal structure. These calculations were done by minimizing, with freely rotating φ and τ dihedral angles, the crystal structure of 39 GFP analogs and mutants found in the PDB. We found the energy minimum conformation of a freely rotating chromophore in the protein matrix of the GFP mutant or GFP-like protein. The energy minimizations did not provide any information about the range of low energy conformations available to a freely rotating chromophore. In order to get this information, molecular dynamics simulations of all GFP-mutants and

GFP-like proteins in the PDB were done. By running molecular dynamics simulations, with freely rotating φ and τ dihedrals, we were able to determine the range of conformations available to chromophores with complete rotational freedom. Eliminating the chromophore from the structure and then finding the volume of the cavity left behind will allow us to find out just how much space is available to the chromophore to rotate.

1.1.3 Kindling Proteins

GFP-like proteins can be divided into two groups. They can be fluorescent or non-fluorescent proteins. The non-fluorescent proteins are in the so-called "chromo" state. The "chromo" state indicates that the protein has a high extinction coefficient but a low quantum yield, whereas in the fluorescent state the protein in characterized by a high quantum yield [28]. Most interesting of these is asCP, a unique non-fluorescent GFP-like protein discovered in the sea anemone *Anemonia sulcata* [29]. Initially non-fluorescent, asCP can be made to fluoresce (kindled) by intense green light irradiation. After kindling, the protein relaxes back to its non-fluorescent state, or it can be quenched instantly by short blue light irradiation. Using mutagenesis directed at specific sites, asCP mutants were created that were always fluorescent. This same process process was also used to create some mutants that were non-fluorescent and could not be kindled. On the basis of those findings, GFP-like molecules with the chromophores in the *cis* conformation were proposed to be fluorescent, while those with the *trans* comformation were non-fluorescent (in the dark C-state). In asCP, the initial state has a *trans* chromophore – it is non-fluorescent. However, upon kindling the chromophore adopts the *cis* conformation and becomes fluorescent. Merely a mutation of residue 165 changes

the "chromo" wild-type asCP to a fluorescent form that does not have its kindling/quenching properties [30].

The crystal structure of the dark state of asCP has recently been released [31]. As predicted, it is in the *trans* conformation. However the chromophore has only one covalent link to the protein. Fragmentation of the protein backbone has occurred – this has been shown to be an intrinsic step in the maturation of the asCP chromophore, see Figure 5.

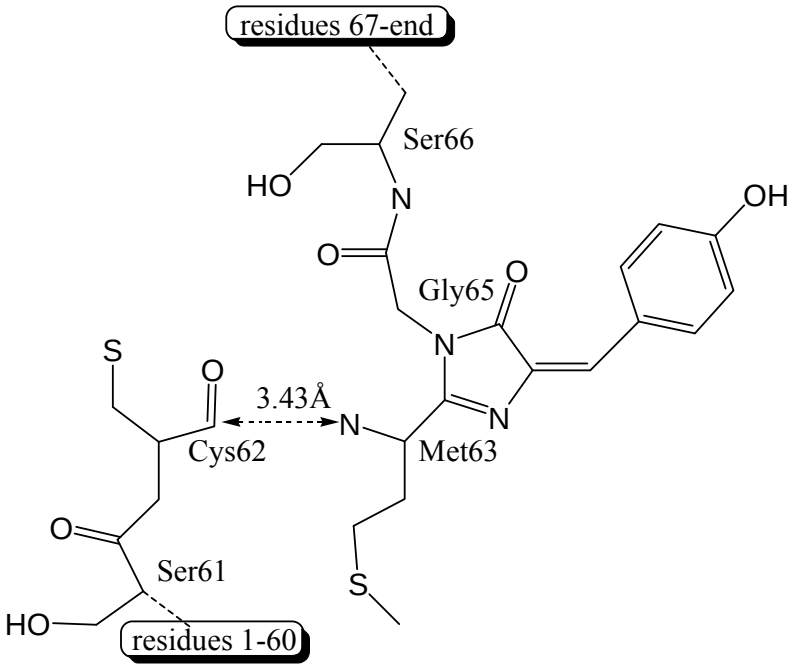

Figure 5. The chromophore of asCP. The main difference between the asCP and GFP chromophore is that polypeptide fragmentation has occurred between residues 62 and 63 in asCP. The chromophore is in a trans conformation in crystal structure of the dark state of the kindling fluorescent protein.

The cleavage of the Cys62-chromophore bond (asCP numbering) provides the chromophore freedom of movement not observed in GFP and other GFP-like proteins. In this work, computational methods were used to examine whether the fragmentation may

lower activation barriers for *cis*/*trans* conformational transitions and may therefore be responsible for asCP kindling abilities. One of the permanently fluorescent asCP mutants was computationally examined to establish whether this mutant is preferentially found in the *cis* configuration. These results along with the others will be used to computationally design mutants of fluorescent GFP that have a protein matrix complementary with a *trans* conformation for the chromophore.

1.2 Methods

1.2.1 Molecular Modeling

Molecular modeling is a collective term that refers to theoretical methods and computational techniques used to model or mimic the behavior of molecules. The techniques are used in the fields of computational chemistry, computational biology and materials science for studying molecular systems ranging from small chemical systems to large biological molecules and material assemblies. The common feature of molecular modeling techniques is the description of the molecular systems on the atomic level; the lowest level of information is individual atoms (or a small group of atoms). The benefit of molecular modeling is that it reduces the complexity of the system, allowing many more atoms to be considered during simulations.

The function, referred to as a potential function, computes the molecular potential energy as a sum of energy terms that describe the deviation of bond lengths, bond angles and torsion angles away from equilibrium values, plus terms for non-bonded pairs of atoms describing van der Waals and electrostatic interactions.

$$E = E_{bonds} + E_{angle} + E_{dihedral} + E_{non-bonded}$$

$$E_{non-bonded} = E_{electrostatic} + E_{vanderWaals}$$

The set of parameters consisting of equilibrium bond lengths, bond angles, partial charge values, force constants and van der Waals parameters are collectively known as a force field. Different implementations of molecular mechanics use slightly different mathematical expressions, and therefore, different constants for the potential function. We will use a technique known as energy minimization to find zero gradient positions for all atoms, in other words, a local energy minimum. Lower energy states are more stable and are commonly investigated because of their role in chemical and biological processes.

1.2.2 Molecular Dynamics

A molecular dynamics simulation, on the other hand, computes the behavior of a system as a function of time. It involves solving Newton's laws of motion, principally the second law, F = ma. Integration of Newton's laws of motion, using different integration algorithms, leads to atomic trajectories in space and time. The energy minimization technique is useful for obtaining a static picture for comparison of states of similar systems, while molecular dynamics provides information about the dynamic processes with the intrinsic inclusion of temperature effects.

In this thesis, molecular dynamics simulation will be performed on every structure with freely rotating dihedral angles. These searches will show the total movement that is available to the chromophore for each structure.

1.2.3 Conformational Searches

In addition to molecular dynamics, conformational searches will be done on all the variants of the asCP crystal structure. Conformational energy searching is used to find all of the energetically preferred conformations of a molecule, which is mathematically equivalent to locating all of the minima of its energy function.

An energy surface resembles a mountain range, complete with peaks (energy barriers), valleys (energy minima), and passes (saddle points). Molecular mechanics energy approximations are most valid in and around the minima. The lowest energy minima tend to be the most populated (as per the Boltzmann distribution). However, since molecular energies are not accurately computed by molecular mechanics or semiempirical quantum mechanics, and the molecular environment in the model may not be exact, it is best to consider energy minima within a certain range of the lowest energy conformation. In this work, a large-scale low-mode setting will be used in the conformational searches.

2. Materials and Methods

Thirty-nine crystal structures of GFP, GFP-mutants, and GFP analogs in the protein databank (PDB) were downloaded and minimized with MacroModel version 8.1. Calculations were carried out using AMBER* with freely rotating τ and φ dihedral angles ($V_1 = V_2 = V_3 = 0.000$). A "hot" sphere with a radius of 8Å from residues 65-67, with a secondary constrained sphere extending a further 3Å, was used in all minimizations. The convergence criterion for all the minimizations was 0.05kJ/Å-mol. Each crystal structure was minimized for 5000 iterations and the values of the τ and φ dihedral angles of the non-minimized and the fully minimized structures were recorded.

Computational methods were then used to help determine how much space was available inside of the beta can. The chromophore was removed from the thirty-nine crystal structures used in Table 1 by cleaving residue 66 from residues 65 and 67. A castP calculation using a probe radius of 2.1 Å was then taken for each protein to find the volume of the space that the chromophore occupied (Table 2) [32]. These calculated volumes were then compared to the original dihedral calculations to find a correlation between the amount of space available to the chromophore and the dihedral movement of the chromophore.

The number of hydrogen bonds and the residues from the protein matrix that formed the hydrogen bonds to the chromophore for each of the thirty-nine crystal structures from the PDB were also found using MacroModel version 8.1. The hydrogen bonds from the hydroxyl group from the phenol ring, the carbonyl from the five-membered ring, and the nitrogen of the five-membered ring all from residue 66 were considered.

To examine the kindling proteins, large-scale low-mode conformational searches for 5000 structures were done on crystal structures 2A52, 2A53, 2A54, and 2A56 with and without residues 62 and 63 fused together using the AMBER* force field. The "hot radius" contained residues 62, 63, and 65 and the atoms within a radius of 7.00 Å. The first shell extending out from there had a radius of 2 Å and a force constant of 100 (kJ/Å). The second shell was extended another 2 Å with a force constant of 200 (kJ/Å). Initial energy minimizations were done for 500 iterations with a convergence criterion of 0.5kJ/Å-mol.

3. Results and Discussion

3.1 Cis/Trans Isomerization in GFP

3.1.1 Freely Rotating Minimization

We have previously shown that the protein cavity surrounding the chromophore in wild-type GFP is not complementary with a planar chromophore, and that a hula-twist motion could be expected if the τ and φ dihedral angles of the chromophore could freely rotate [33]. To establish whether or not this observation could only be applied to wild-type GFP or was a more common characteristic of all GFP-like molecules, all the structures of GFP, GFP-mutants and GFP-like molecules in the protein databank [34] were examined. Energy minimization calculations were carried out with a freely rotating chromophore within the protein matrix to find the conformations the protein attempts to impose on the chromophore when it has no barriers to rotation of its τ and φ dihedral angles. This can provide an indication of the conformations the chromophore can adopt within the protein assuming that the first excited state results in a loss of conjugation between the phenol and imidazolinone rings, which in turn allows for rotating τ and φ dihedral angles. The τ and φ dihedral angles of these lowest energy minimum conformations from minimizing of the freely rotating chromophore within the protein matrix are listed in Table 1. They determine whether the protein environment around the chromophore restricts chromophore twisting by imposing steric barriers to rotation of the τ and φ dihedral angles, or by forming hydrogen bonds to the chromophore. While the quantum mechanical calculations have shown the ground state minima is planar, this is not necessarily so for the excited state; in fact in some cases the excited state has an energy minimum with a perpendicularly twisted chromophore [24].

Table 1 *– τ and φ dihedral angles of solid and freely rotating minimized crystal structures GFP, GFP-mutants, and GFP-like molecules*

Protein	PDB-ID	Mutations	Name	τ (crystal)	φ (crystal)	τ (mini)	φ (mini)
GFP-mut	1B9C	F99S M153T V163A	Cycle-3, GFPuv	-2.4	3.2	22.4	-34.8
GFP-mut	1BFP	Y66H/Y145F	BlueFP	2.4	-2.1	4.6	14.1
GFP-mut	1C4F	S65T	EGFP	-0.4	0.1	-0.8	4.1
GFP-mut	1CV7			-34.6	36.3	20.9	-40.7
GFP-mut	1EMA	S65T	EGFP	17.6	-13	20.9	-38.8
GFP-mut	1EMC	F64L/I167T/K238N		-0.3	-0.8	-2.9	-48.5
GFP-mut	1EME	F64L/I167T/K238N		-0.6	-5.3	26.1	-33.4
GFP-mut	1EMF	F64L/Y66H/V163A		0.3	-1.2	2	12.3
GFP-mut	1EMG	S65T	EGFP	4.0	-4.2	11.4	-17.1
GFP-mut	1EMK	F64L/S65T/I167T/K238N		-1.0	-1.9	-1.5	-31.4
GFP-mut	1EML	F64L/I167T/K238N		-0.4	-4.3	7.3	-28.2
GFP-mut	1EMM	F64L/K238N		-0.2	0.6	0.8	-31
GFP-mut	1F0B	S65G, V68L, S72A, Q80R, T203Y, H148Q	Yellow mut	-13.3	16.7	0.2	-14.7
GFP-mut	1F09	S65G, V68L, S72A, Q80R, T203Y, H148Q		5.1	-2.9	-17	8.1
GFP	1GFL		GFP	0.5	0.0	16.6	-45.3
GFP-mut	1H6R	C48V, S65A, V68L, S72A, N149C, M153V, S202C, T203Y, D234H		7.3	4.2	7.1	-18.3
GFP-photo-prod	1HCJ		cleaved - Glu222	3.0	-3.4	22.1	-43.1
GFP-mut	1HUY	S65G, V68L, Q69M, S72A,	Citrine	1.7	0.6	30.1	-37.5
GFP-mut	1JBY	S65T,Q80R,H148G,T203C	low pH struc.	0.1	-1.9	26.5	-39.2
GFP-mut	1JBZ	S65T,Q80R,H148G,T203C	high pH struc	1.9	-1.4	7.9	4.7

GFP-mut	1JC1	C48S,F64L,S65T,Q80R,S147C,Q204C		-0.8	-0.5	2.9	-44.2
GFP-mut	1MYW	F46L, F64L, S65G, V68L, S72A, M153T, V163A, S175G, T203Y	Venus	0.8	-0.4	15.5	-15.4
GFP-mut	1OXD	unnatural aa	Gold	0.5	-1.7	4.5	-35
GFP-mut	1OXE	unnatural aa	Gold	0.8	-1.8	17.4	16.7
GFP-mut	1Q4A	S65T	EGFP, pH=8.5	1.7	-2.6	1.8	-48.6
GFP-mut	1Q4B	S65T	EGFP, pH=5.5	5.2	-0.9	29	-33.6
GFP-mut	1Q4C	S65T, T203C	pH=8.5	2.3	-1.2	25.8	-28.4
GFP-mut	1Q4D	S65T, T203C	pH=5.5	3.2	-2.5	28.2	-41.7
GFP-mut	1Q4E	S65T, Y145C	pH=8.5	2.2	-3.7	17.4	16.7
GFP-mut	1QYF	R96A		-11.1	5.9	11.0	-25.2
GFP-mut	2EMD	F64L, Y66H		1.7	0.0	1.3	8.9
GFP-mut	2EMN	F64L, Y66H		-0.4	-5.3	0.1	-1.6
GFP-mut	2EMO	F64L, Y66H, V163A		2.2	8.5	-1.2	-45
GFP-mut	2YFP	S65G,V68L,S72A,T203Y,H148G	YFP	-4.0	2.4	5.3	-23.2
Discosoma	1G7K		DsRed	0.0	3.5	3.1	-26.2
Discosoma	1GGX		DsRed	2.0	2.4	-25.2	14.7
	1MOV			169.2	42.4	-179.1	58.6
	1UIS			176.2	7.1	148.7	67.1
	1MOU			169.5	43.1	63.0	-49.3

The τ and φ dihedral angles of the GFP, GFP-mutant, and GFP-like protein crystal structures in the PDB, as well as the τ and φ dihedral angles of the lowest-energy minimum conformations with freely rotating τ and φ dihedral angles are listed in Table 1. Figure 6 shows the τ and φ dihedral angles of the chromophore in the crystal structures and in the

calculated freely rotating chromophores (1MOV, 1UIS, and 1MOU are not presented in Figure 6 because they have a *trans*-chromophore) and summarizes Table 1. The figure clearly shows that wild-type GFP is not an anomaly: most of the GFP and GFP-like proteins in the PDB have a protein matrix that is not complementary with a planar chromophore. When the π-conjugation across the ethylenic bridge of the chromophore is removed, the protein matrix will significantly twist the free rotating chromophore from the relatively planar structures found in the crystal structures.

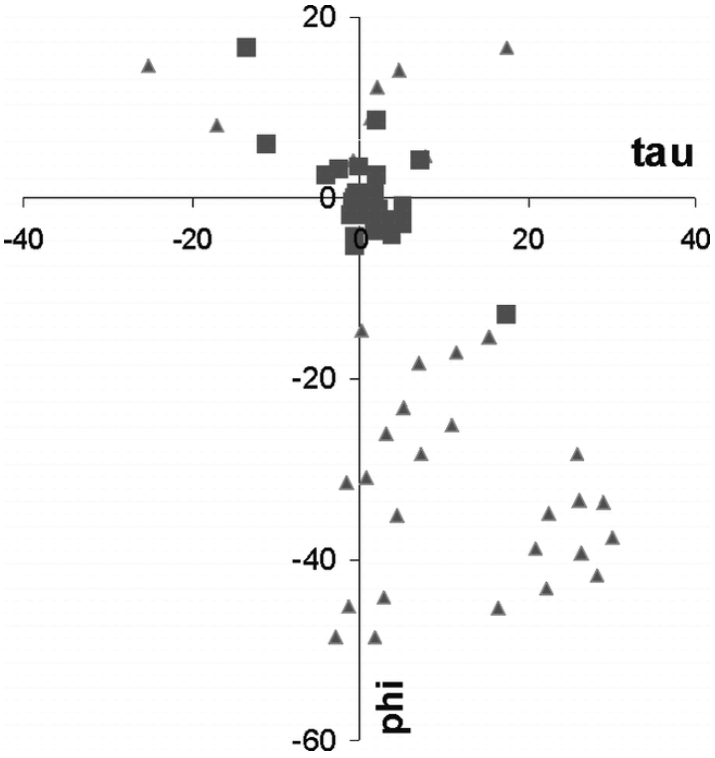

Figure 6 - τ and φ dihedral angles of all GFP and GFP-like molecules in the PDB (trans-chromophores are omitted for clarity). In the solid state (▪) the chromophores are mainly planar. However when they are allowed to freely rotate (▲), the chromophores twist in response to the steric effects of the protein matrix that surrounds the chromophore. Upon removing the rotational constraints of the τ and φ dihedral angles, most τ dihedrals underwent a clockwise (+) twist and an anticlockwise φ twist, resulting in a negatively correlated HT motion from planarity.

The cavity formed by the folding of protein around the chromophore is not symmetric: the direction (and, therefore, sign) of the τ and φ twists are important. The figure shows that, in most cases, a freely rotating chromophore will undergo a negatively correlated HT motion or a partial φ OBF. Only two chromophores undergo a significant counterclockwise (–) τ rotation of the chromophoric phenol (1F09 and 1GGX). Because they both also undergo a positive φ rotation, this is also a –HT. Two freely rotating chromophores – 1Q4E and 1JBZ – respond to the protein matrix by undergoing a positively correlated HT motion.

The protein matrix seems to have a large impact on the chromophore conformation. The unminimized crystal structures were associated with the planar conformation and the minimized structures did not maintain the planar chromophore conformation. The non-planar conformation found in the minimized structures shows that the surrounding protein matrix has an affect on how the chromophore twists. It also shows that the chromophore does not adopt the perpendicular conformation that leads to the NAC described in Figure 4a. The protein matrix forces the chromophore to rotate around its dihedral angles, but stops it before allowing the two rings of the chromophore to take on the perpendicular conformation to each other. This explains why the fluorescence quenching NAC does not occur (see Figure 4b).

3.1.2 Chromophore Cavity Volume

Since the crystal structures of the thirty-nine GFP, GFP-mutants, and GFP-like molecules did not display planar chromophores when subjected to molecular dynamics simulations with a freely rotating chromophore, it seemed to be a logical first step to examine the amount of space available to the chromophore for rotation. From there a correlation

might be drawn to understand why the chromophore moves away from planarity.

To further examine the space available to the chromophore in the protein matrix, the chromophore was eliminated from each of the structures. By doing so, a cavity was created where the chromophore once stood. The new structures with the cavities were then put through castP calculations to find the volume of all the cavities in each of the structures. A probe radius of 2.1 Å was used to help find the volume available to the chromophore (see Table 2) and the volumes were then compared to the torsion values of φ and τ from the minimization calculations shown in Table 1.

Table 2 *- Protein IDs and the Volume of space available to the Chromophore*

PDB-ID	Volume (Å^3)
1B9C	225.22
1BFP	260.52
1C4F	300.11
1CV7	504.26
1EMA	43.93
1EMC	358.21
1EME	306.41
1EMF	344.56
1EMG	359.92
1EMK	395.25
1EML	369.30
1EMM	342.45
1F0B	104.43
1F09	126.58
1G7K	128.41
1GFL	173.86
1GGX	313.65
1H6R	233.00
1HCJ	381.00
1HUY	213.63
1JBY	323.18
1JBZ	298.97
1JC1	282.98
1MOU	209.11
1MOV	223.84
1MYW	186.76
1OXD	411.84

1OXE	409.62
1Q4A	305.02
1Q4B	298.92
1Q4C	306.88
1Q4D	293.92
1Q4E	375.79
1QYF	324.87
1UIS	409.00
2EMD	244.09
2EMN	307.10
2EMO	360.87
2YFP	122.90

These volumes were compared to the φ and τ dihedral angles from the solid crystal structures as well as the freely rotating crystal structures.

Most of the proteins had cavity volumes between the ranges of 200 – 400 Å^3. However there were some that deviated from this characteristic. Most notably, the amount of free space in 1CV7 and 1EMA had that largest disparity from the average range. 1CV7 had the largest volume available to the chromophore while 1EMA had the smallest. The φ and τ dihedral angles from the freely rotating minimized crystal structure of 1CV7 showed the greatest overall change from the φ and τ dihedral angles from the solid crystal structures ($\Delta\varphi = 77.00°$ and $\Delta\tau = 55.50°$). The proteins with *trans* chromophores are discussed later. The freely-rotating minimized structure of 1EMA did not show a great degree of variance from its solid crystal structure ($\Delta\varphi = 25.80°$ and $\Delta\tau = 3.30°$) in comparison to the other proteins analyzed. The diminutive volume available to the chromophore to twist constrains the ability of the chromophore to rotate a great degree, which would account for the small difference in the φ and τ dihedral angles. It is interesting to note that 1CV7, 1EMA, and 1F0B also do not display a nearly planar chromophore in the solid crystal state. 1F0B contains the second smallest cavity volume available to the chromophore. Each of these proteins has either a very big cavity volume or a very small one. In the case of 1CV7, when the chromophore is

allowed to move about, it does not adopt a planar conformation. For 1EMA and 1F0B, the small size of their cavities forces the chromophore to take on a non-planar conformation because that is the only way it will fit into their cavities. Proteins 1OXD and 1OXE also contained fairly large chromophore cavities while 1F09, 1GFL, 1MYW, 1G7K, and 2YFP had smaller than average chromophore cavities. Overall the greatest degree of change was in the φ dihedral angle.

The proteins containing *trans* chromophores (1MOV, 1UIS, and 1MOU) were examined separately from the proteins with *cis* chromophores because their average change in φ and τ dihedral angles was large in comparison to those for the *cis* chromophores. 1UIS does contain a very large chromophore cavity while 1MOU and 1MOV are actually in the lower end of the cavity range. The large size of the 1UIS cavity probably accounts for the smaller difference between the φ and τ dihedral angles of the freely rotating minimized crystal structure. This could suggest that when given enough space and allowed to rotate freely, the chromophore will attempt to adopt a conformation closer to *cis* than to *trans*.

3.1.3 Freely Rotating Molecular Dynamics Simulation

To further examine how the chromophore behaves over the course of time molecular dynamics (MD) simulations with freely rotating τ and φ dihedral angles (V_1 = V_2 = V_3 = 0.000) were carried out. Instead of just looking at the static states from the minimizations, the MD simulations provided a good means to observing the dynamic properties of the proteins in real time. From the freely rotating molecular dynamic calculations on the thirty-nine structures, the average φ and τ torsions were calculated and the maximum and minimum φ and τ torsions were found for each protein, see Table 3.

Table 3 - *Protein IDs and their Maximum, Minimum, and Average τ and φ values*

PDB-ID	Avg τ	Avg φ	Max τ	Max φ	Min τ	Min φ
1B9C	-15.88	-16.61	25.95	53.41	-63.08	-65.77
1BFP	1.80	-2.02	31.87	40.52	-53.87	-47.45
1C4F	5.65	-45.46	51.02	53.00	-28.29	-76.60
1CV7	22.97	-44.32	59.79	-3.93	-14.38	-81.29
1EMA	21.27	-54.67	81.83	-2.24	-25.45	-103.51
1EMC	-18.76	25.29	45.47	90.91	-63.03	-84.45
1EME	5.89	-9.26	50.76	51.96	-50.81	-72.87
1EMF	-1.65	-2.17	23.48	56.71	-32.78	-195.47
1EMG	9.76	-37.55	46.35	47.74	-32.12	-118.47
1EMK	-5.24	-27.47	16.14	22.34	-26.84	-66.01
1EML	-8.73	-43.88	34.01	47.16	-46.57	-104.94
1EMM	-3.75	-27.89	39.37	64.26	-61.19	-60.60
1F0B	7.22	-45.26	55.91	6.48	-40.99	-84.15
1F09	-12.82	-6.02	22.29	57.84	-46.83	-42.05
1G7K	-24.26	24.17	25.75	61.06	-63.79	-43.37
1GFL	0.68	-45.93	42.54	30.31	-38.01	-92.95
1GGX	-16.85	49.81	54.24	101.21	-74.93	-21.44
1H6R	1.98	46.62	54.49	86.06	-42.13	-71.67
1HCJ	-24.36	-39.22	42.74	148.97	-92.54	-149.59
1HUY	15.74	-26.67	58.59	88.63	-56.71	-74.23
1JBY	-4.52	-146.37	71.11	-71.40	-50.31	-264.79
1JBZ	15.02	-56.52	65.97	46.17	-20.47	-94.45
1JC1	9.36	-52.77	61.93	-8.58	-32.30	-98.60
1MOU	7.64	144.43	63.93	196.55	-28.01	65.50
1MOV	-166.86	-122.72	-144.38	-98.15	-192.20	-250.16
1MYW	24.08	-34.29	62.51	62.13	-41.91	-78.04
1OXD	-2.66	11.09	48.34	55.65	-44.30	-55.05
1OXE	16.15	-41.44	48.89	-9.67	-14.85	-65.72
1Q4A	2.02	-39.14	38.20	-2.54	-31.55	-70.48
1Q4B	18.64	-49.49	62.32	21.49	-23.60	-86.59
1Q4C	5.79	2.87	47.61	69.96	-48.02	-73.86
1Q4D	27.58	-45.27	63.49	42.87	-32.46	-82.80
1Q4E	16.15	-41.44	48.89	-9.67	-14.85	-65.72
1QYF	15.85	-53.10	57.45	6.63	-30.00	-104.62
1UIS	-156.07	-149.34	-128.50	-114.17	-190.51	-183.21
2EMD	0.12	-25.20	29.83	45.68	-30.02	-66.22
2EMN	-1.81	15.09	28.13	69.99	-34.23	-155.91
2EMO	-1.32	-9.88	30.01	58.44	-27.31	-65.77
2YFP	-6.18	39.79	49.65	91.69	-53.87	-39.22

The MD calculations produced structures for 1MOU, 1MOV, and 1UIS that had average φ and τ torsions that were much higher in magnitude than the other structures because these structures contained *trans*-chromophores. From the maximum value and

minimum values of the φ and τ dihedrals, the changes for these torsions in the proteins were quantified in Table 4.

Table 4 *– The change in the maxima and minima φ and τ torsions from the freely rotating molecular dynamics calculations*

PDB-ID	Δτ	Δφ
1B9C	89.04	119.18
1BFP	85.74	87.97
1C4F	79.31	129.6003
1CV7	74.17	77.35
1EMA	107.28	101.28
1EMB	111.11	152.46
1EMC	108.49	175.36
1EME	101.57	124.83
1EMF	56.26	107.81
1EMG	78.47	166.21
1EMK	42.98	88.35
1EML	80.57	152.10
1EMM	100.56	124.86
1F0B	96.89	90.63
1F09	69.12	99.88
1G7K	89.53	104.43
1GFL	80.56	123.26
1GGX	129.17	122.65
1H6R	96.62	157.72
1HCJ	135.29	61.44
1HUY	115.30	162.87
1JBY	121.42	166.61
1JBZ	86.44	140.62
1JC1	94.23	90.02
1MOU	91.93	131.04
1MOV	47.82	152.01
1MYW	104.43	140.17
1OXD	92.64	110.71
1OXE	63.74	56.05
1Q4A	69.75	67.94
1Q4B	85.92	108.09
1Q4C	95.63	143.82
1Q4D	95.95	125.67
1Q4E	63.74	56.05
1QYF	87.45	111.25
1UIS	62.02	69.03
1YFP	98.04	97.71
2EMD	59.84	111.91
2EMN	62.37	134.10
2EMO	57.32	124.20
2YFP	103.53	130.91

Each structure appeared to have a fairly mobile chromophore when it was allowed to rotate freely. 1CV7, which was shown to have the largest chromophore cavity volume, displays a relatively small change in its φ and τ dihedrals compared to some of the other crystal structures. This could be an indication that if the chromophore cavity is very big the chromophore will adopt a preferred conformation and not deviate that far from it even when the chromophore is allowed to rotate without its ethylenic bride conjugation. 1OXE, which also has a large chromophore cavity, had smaller changes in its φ and τ torsions, but a closely related protein 1OXD did show a larger degree of change in the two dihedrals. Some of the proteins with smaller chromophore cavities displayed larger changes in their φ and τ dihedral, which would indicate that with a smaller chromophore cavity it is most likely struggling to find a comfortable conformation taking into account the steric effects of the protein matrix. It is interesting to note that in no case does the chromophore remain at its starting position of the MD simulation. This once again reinforces the idea that the protein matrix plays a large role in chromophore conformation and therefore fluorescence by preventing the fluorescence quenching NAC.

3.1.4 Hydrogen Bonding

The hydrogen bonding of the chromophore was also examined to find a trend that correlates with the non-planar chromophore deformations observed in the MD simulations. It was hypothesized that the formation of new hydrogen bonds could be another driving force for the twisting of the chromophore away from planarity. It is possible that in addition to the steric force exerted by the protein matrix, the formation of new hydrogen bonds could also have a stabilizing effect on the chromophore by forcing it to twist, and at the same time, allowing it to make more hydrogen bonds to the protein matrix. Table 5 summarizes the

hydrogen bonds formed by the OH of the phenol ring in the chromophore as well as the carbonyl oxygen and sp^2 hybridized nitrogen of the imidazolinone ring of the chromophore to other amino acid residues of the protein matrix.

Table 5 *– The hydrogen bonding from each crystal structure of GFP, GFP-mutants, and GFP-like proteins before and after minimization*

Mutant	Carbonyl of chromophore	Phenol of chromophore	Nitrogen of carbonyl ring of chromophore
1B9C	GLN 94, ARG 96	HIS 148	
1B9C mini	GLN 94, ARG 96(2)	TYR 145, THR 203	
1BFP	GLN 94, ARG 96		IIC 66
1BFP mini	GLN 94, ARG 96		IIC 66
1C4F	GLN 94, ARG 96		
1C4F mini	ARG 96(2)	TYR 145	
1CV7	GLN 94, ARG 96		
1CV7 mini	GLN 94, ARG 96	GLU 222	
1EMA	GLN 94, ARG 96	THR 203	
1EMA mini	GLN 94	TYR 145, THR 203	
1EMC	ARG 96		
1EMC mini	GLN 94, ARG 96(2)	TYR 145, GLU 222	
1EME	GLN 94, ARG 96		
1EME mini	GLN 94, ARG 96	TYR 145	
1EMF	GLN 94, ARG 96		CSH 66
1EMF mini	GLN 94, ARG 96		
1EMG	GLN 94, ARG 96		
1EMG mini	GLN 94, ARG 96	TYR 145, HIS 148	
1EMK	GLN 94, ARG 96	THR 203	

1EMK mini	GLN 94, ARG 96(2)	TYR 145, THR 203	
1EML	GLN 94, ARG 96	THR 203	
1EML mini	GLN 94, ARG 96	TYR 145, THR 203	
1EMM	GLN 94, ARG 96	HIS 148, THR 203	
1EMM mini	GLN 94, ARG 96	HIS 148, THR 203	
1F09	GLN 94, ARG 96		
1F09 mini	ARG 96	TYR 145, GLN 148	
1F0B	ARG 96		
1F0B mini	GLN 94, ARG 96		
1G7K	ARG 95	SER 146, LYS 163	
1G7K mini	ARG 95, TYR 181	LYS 163	
1GFL	GLN 94, ARG 96	HIS 148	
1GFL mini	GLN 94, ARG 96		
1GGX	ARG 95	LYS 163	
1GGX mini	CRO 68, ARG 95	SER 146, LYS 163, SER 197	
1H6R	GLN 94, ARG 96	HIS 148	
1H6R mini	GLN 94, ARG 96(2)	TYR 145, HIS 148	
1HCJ	GLN 94, ARG 96		
1HCJ mini	ARG 96	TYR 145, THR 203	
1HUY	GLN 94, ARG 96	HIS 148	
1HUY mini	GLN 94, ARG 96	TYR 145, HIS 148	
1JBY	GLN 94, ARG 96		
1JBY mini	ARG 96	TYR 145	
1JBZ	GLN 94, ARG 96		
1JBZ mini	ARG 96		
1JC1	GLN 94, ARG 96		

1JC1 mini	GLN 94, ARG 96(2)	ASN 149, THR 203	
1MOU	ARG 95	ASN 161	
1MOU mini		HIS 146, ASN 161	
1MOV	ARG 95	ASN 161	
1MOV mini	ARG 95(2)	GLU 148	
1MYW	GLN 69, GLN 94, ARG 96		
1MYW mini	ARG 96	TYR 145	
1OXD	GLN 94, ARG 96		
1OXD mini	ARG 96	GLU 222	
1OXE	GLN 94, ARG 96		
1OXE mini	ARG 96(2)		GLU 222
1Q4A	GLN 94, ARG 96	THR 203	
1Q4A mini	CRO 66, ARG 96	TYR 145, GLU 222	
1Q4B	GLN 94, ARG 96	HIS 148	
1Q4B mini	CRO 66, ARG 96	TYR 145	
1Q4C	GLN 94, ARG 96		
1Q4C mini	GLN 94, ARG 96(2)	TYR 145	
1Q4D	GLN 94, ARG 96		
1Q4D mini	GLN 94, ARG 96	TYR 145	
1Q4E	GLN 94, ARG 96	THR 203	
1Q4E mini	ARG 96(2)		GLU 222
1QYF	GLN 94	HIS 148, THR 203	
1QYF mini	GLN 94	TYR 145, THR 203	
1UIS	ARG 92	ASN 143	
1UIS mini	ARG 92	ASN 143, SER 158	
1YFP			

1YFP mini			
2EMD	GLN 94, ARG 96		CSH 66
2EMD mini	GLN 94, ARG 96(2)		
2EMN	GLN 94, ARG 96		CSH 66
2EMN mini	GLN 94, ARG 96(2)		
2EMO	GLN 94, ARG 96		CSH 66, THR 203
2EMO mini	GLN 94, ARG 96		THR 203
2YFP	GLN 94, ARG 96		
2YFP mini	ARG 96	ASN 146	

Twenty-one structures exhibited more hydrogen bonds in the freely rotating minimized form than in the solid crystal structure. Thirteen had no net change in the number of hydrogen bonds, but the residues to which hydrogen bonds were made did change in some cases. Only five showed a net loss in hydrogen bonds from the solid to the minimized crystal structures. The phenol portion of the chromophore seems to be the most active player in seeking out new hydrogen bonds when the chromophore starts rotating. Since it is the portion of the chromophore that is not directly bound to the protein backbone it will be able to seek out other parts of the protein matrix that the five-membered ring cannot. The steric force of the protein seems to have an additional effect of allowing the chromophore to make more hydrogen bonds to the protein matrix. This would in theory stabilize the twisted conformation of the chromophore even further not taking into account any other factors that the protein could impose of the chromophore.

3.1.5 His148

From the crystal structures it is known that His148 forms a hydrogen bond to Tyr66 stabilizing the anionic form of the phenolate. With this knowledge, we decided to examine the χ_1 and χ_2 dihedral angles of His148 in 1GFL to see how they responded during the MD simulation of 1GFL. It was already known that the chromophore did not perform a *cis*/*trans* isomerization, and if His148 did not show much conformational change, this would support the assertion that the protein matrix plays a major role in inhibiting the fluorescence quenching NAC. The complete range of conformations that the chromophore and the nearby histidine adopted with this rotational freedom are shown in Figures 7 through 10.

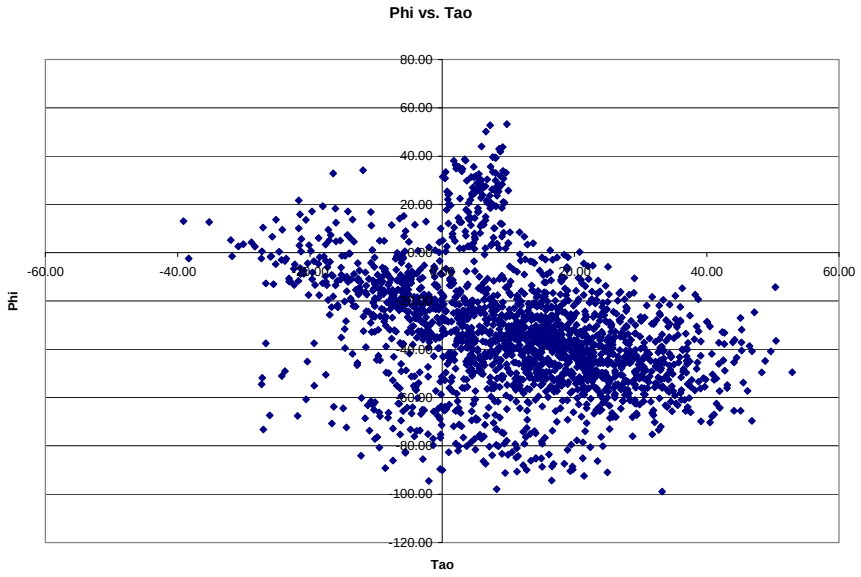

Figure 7. The φ and τ dihedral angles for 1GFL.

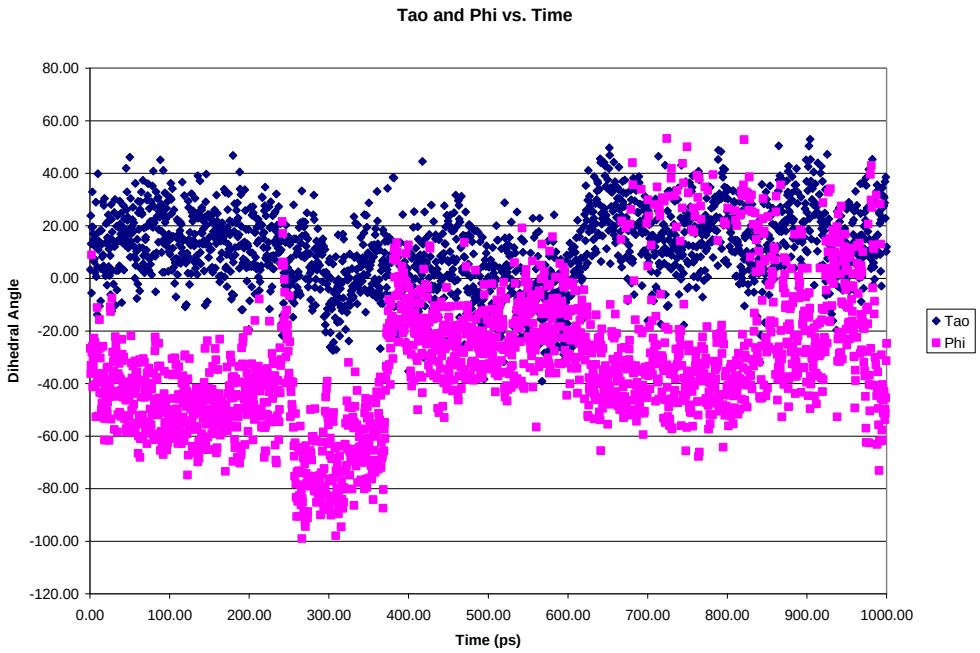

Figure 8. For 1GFL, the φ dihedral angle had a much greater conformational change than the τ dihedral.

The chromophore did not adopt a planar conformation (τ = φ = 0°). The dihedral angles were constrained by the protein matrix surrounding the chromophore, so the *cis*/*trans* isomerization did not occur, which prevented the NAC from occurring. The dihedrals for 1GFL also demonstrate that when they are allowed to rotate freely, they do not deviate much from their original conformation. Even when the chromophore is excited and allowed to rotate, it will not behave much differently than it would in the ground state.

Upon examination of the χ torsions from His148 there was not a great degree of overall rotation during the MD simulation, see Figures 9 and 10.

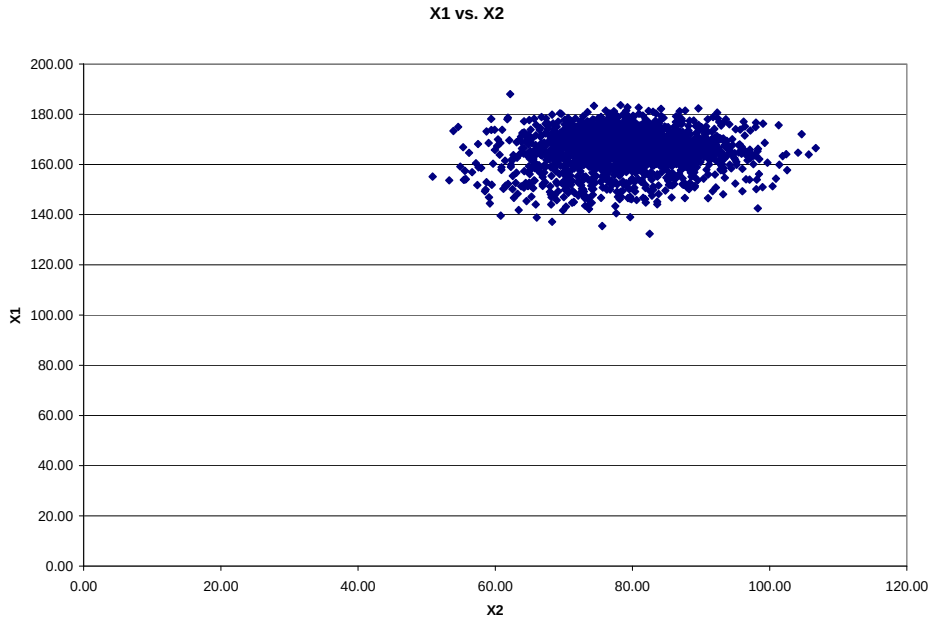

Figure 9. The χ dihedrals of His148 for 1GFL do not show a great degree of difference.

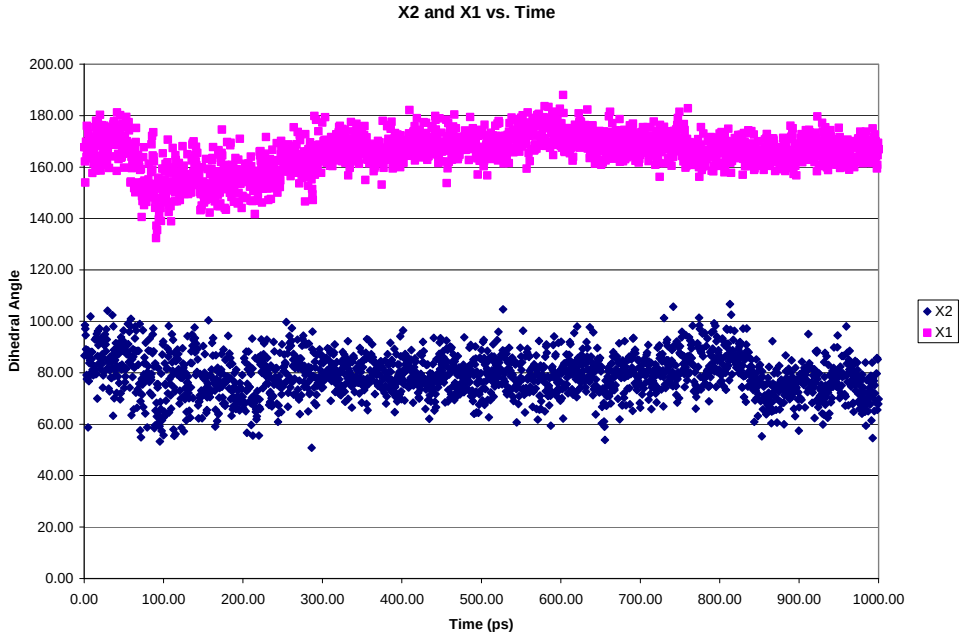

Figure 10. The χ dihedral angles of 1GFL do not take on very different conformations throughout the MD simulation. The histidine that is located close to the chromophore does not move away from the chromophore.

This lack of change indicates that the *cis*/*trans* isomerization does not occur and the limited conformational change in His148 provides a good correlation with the fact that the *cis*/*trans* isomerization is absent. The lack of conformational change in His148 also indicates that the protein matrix plays a significant role in preventing the *cis*/*trans* isomerization of the chromophore.

3.2 Kindling Proteins

3.2.1 His197

When the chromophore is separated from the rest of the protein, it starts displaying the ability to switch between the on and off states, which is why proteins with a cleaved chromophore are considered kindling proteins. The chromophore can undergo a drastic conformational change when φ and τ are allowed to freely rotate. A molecular dynamics simulation of 2A50 was done to see how the chromophore and His197 would respond if the chromophore was allowed to rotate freely over an extended period of time.

In the case of 2A50, since the chromophore was essentially on its own in the chromophore cavity, it was expected that there would be a good amount of rotation in the φ and τ torsions. The cleaved chromophore in 2A50 has a greater ability to rotate because it has one less bond connecting it to the rest of the protein to restrict its rotation when excited, see Figure 5. The φ and τ torsions from the MD simulation were plotted out to see how they changed over time in the MD simulation (see Figures 11 and 12).

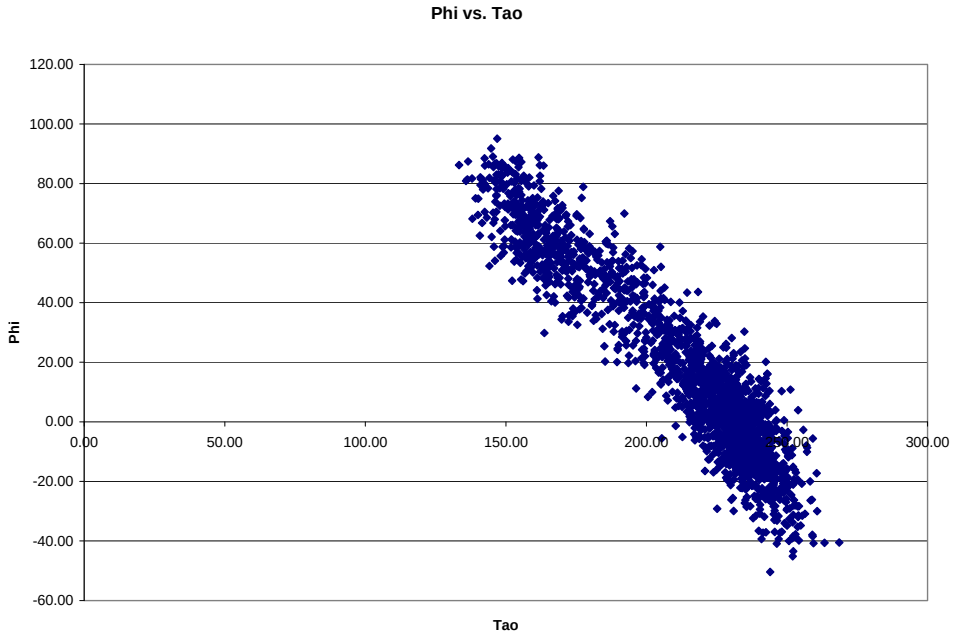

Figure 11. Both the φ and τ dihedral angles vary greatly for the various structures of 2A50 from the MD simulation. The difference in φ however is much greater than that in τ.

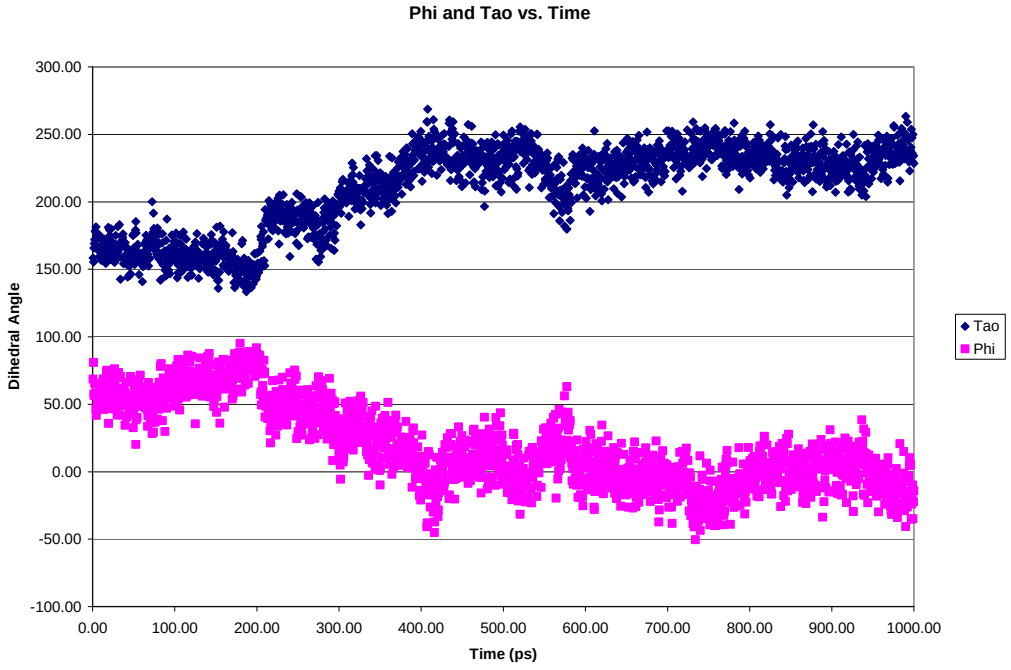

Figure 12. Throughout the MD simulation for 2A50, the φ and τ dihedrals gradually change their conformations from the beginning of the simulation to the end of the simulation.

Over the time of the MD simulation, the chromophore starts to change its conformation from a *trans* configuration, and it does complete an isomerization to the *cis* configuration. It does this by twisting both of the chromophore dihedral angles at the same time in a volume conserving HT-type motion. To examine if the nearby His197 residue had an effect on the chromophore conformation, the χ_1 and χ_2 dihedral angles were found and plotted in Figures 13 and 14 to see how it changed over the MD simulation. By monitoring His197, we were able to see whether its conformation changes were related to chromophore rotation. Since there was less restriction to the chromophore to rotate it was hypothesized that there would be a good deal of chromophore rotation.

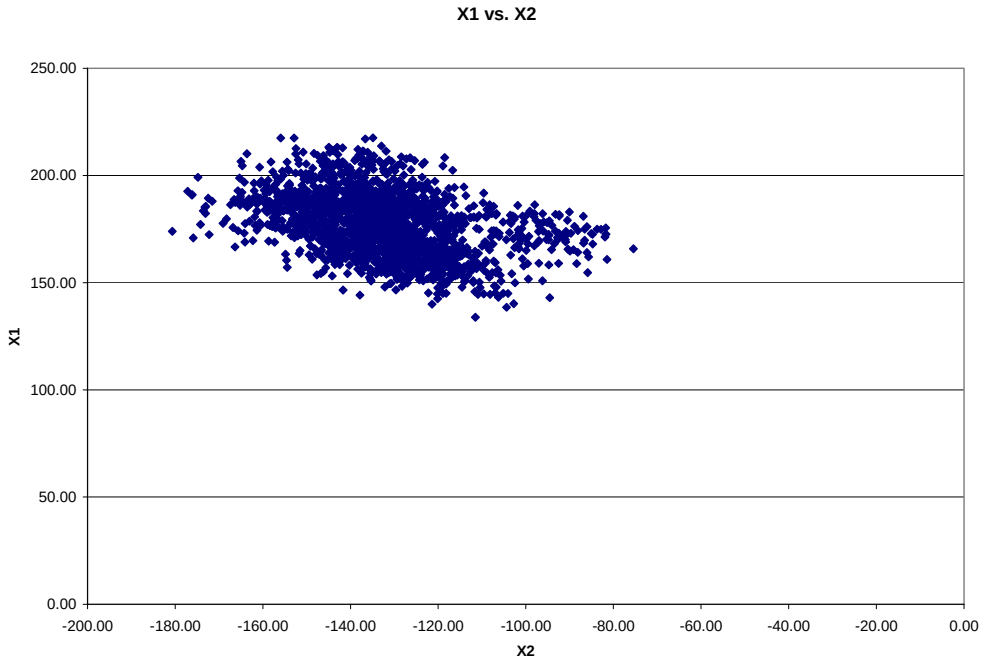

Figure 13. The χ dihedrals for 2A50 do not demonstrate a great degree of variance.

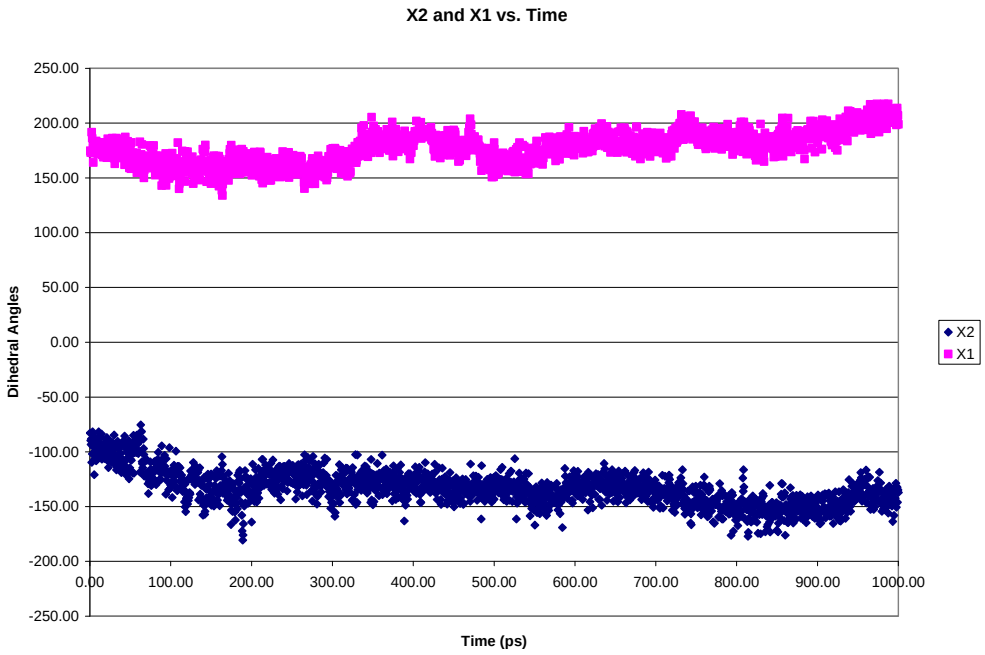

Figure 14. Throughout the MD simulation for 2A50, the χ dihedrals do not demonstrate a great degree of change in their conformations.

The histidine did not seem to undergo any significant conformational change in the MD simulation. It also did not remain in its starting conformation but, after a period of time it settled into a conformation. This would imply that the chromophore rotation does cause the protein matrix to change its conformation. Unlike GFP and GFP-mutants, the kindling proteins are not constrained as much by the protein matrix, and this allowed for the greater change in the chromophore dihedral angles during the MD simulation. The fragmentation of the chromophore from the rest of the protein provides freedom to the chromophore and allows it to rotate and deviate from the *trans* configuration. The steric effect of the protein matrix on the conformation of the chromophore is most likely not as strong because of this fragmentation. The MD simulations of 2A50 provided strong evidence that the chromophore did complete a *cis*/*trans* isomerization.

3.2.2 Conformational Searching

Conformational searches were done on crystal structures 2A52, 2A53, 2A54, and 2A56 to find out how the chromophore adapts to being severed from the protein backbone. The conformational searches were then done again, but this time the chromophores were artificially reattached by connecting the carbonyl from residue 62 with the free nitrogen from residue 63. This was done to determine if there was a great difference when the chromophore is restricted in its rotation. The difference between the maximum and minimum values of the φ and τ dihedral angles and the distance between residues 62 and 63 were then found from all the structures produced from the conformational searches with and without the fused chromophore, see Table 6.

Table 6 *– The average φ and τ dihedral angles, and the distance between residues 62 and 63 for the freely-rotating crystal structures with the fused and cleaved chromophore*

Protein	PDB-ID	Δτ	Δφ	Avg. Dist (Å)
asFP595 (S158V) on-state	2A52	10.70	60.13	3.35
asFP595 (S158V) on-state	2A52fused	10.55	26.88	1.34
asFP595 (A143S) off-state	2A53	12.91	9.64	3.55
asFP595 (A143S) off-state	2A53fused	14.28	15.98	1.33
asFP595 (A143S) on-state 1 min irradiation	2A54	4.93	11.06	3.57
asFP595 (A143S) on-state 1 min irradiation	2A54fused	3.47	10.50	1.34
asFP595 (A143S) on-state 5 min irradiation	2A56	8.46	21.82	3.70

asFP595 (A143S) on-state 5 min irradiation	2A56fused	6.38	14.63	1.34

The φ and τ dihedral angles do seem to show a great deal of change in conformation. All of the crystal structures without fused chromophores, however, always had a larger change in their φ and τ dihedral angles over the course of the conformational search. Most of the time the difference between the proteins with a fused and cleaved chromophore was minimal, but the φ dihedral from 2A52 did exhibit a much larger change than that for 2A52fused ($\Delta\varphi = 33.25°$). The additional torsional freedom to the chromophore in the cleaved state most likely allowed the chromophore to take on more drastic conformations. By fusing residues 62 and 63, the chromophore becomes more restricted in its rotations but not so much so that the chromophore becomes static.

Conclusion

The role of the protein matrix on GFP fluorescence is summarized in Fig. 4B. In the ground state, the highly conjugated GFP chromophore should be planar. However, numerous crystal structures of GFP and GFP-like proteins have been reported to have slightly twisted chromophores (Table 1). The amino acid residues surrounding the chromophore are not complementary with a planar chromophore and they exert a steric strain. This deviation from planarity should have an effect on the fluorescence of the chromophore. When the chromophore is computationally permitted to freely rotate it will adopt a conformation that complements the protein matrix. In most cases the freely rotating chromophore undergoes φ rotations of at least 20°, and in some cases these rotations are accompanied by an equal but opposite rotation of the τ dihedral angle (a −HT). None of the proteins examined have a cavity that only causes a rotation solely around the τ dihedral angle. Interestingly, a similar study of photoactive yellow protein (PYP) by Yamada *et al.* [26] concluded that the protein prevents the chromophore from adopting a completely planar structure. On the basis of their calculations, they proposed that the efficiency of photoisomerization in PYP is due to the asymmetric protein-chromophore interaction that can serve as the initial accelerant for the light-induced photocycle.

In the excited state the protein matrix presumably prevents the chromophore from rotating to the perpendicularly twisted conformation that has been postulated to be the conformation leading to fluorescence-quenching NAC. The protein also exerts a steric force on the chromophore, twisting it away from planarity by means of a negatively correlated HT motion. The interplay between these forces and the electronic structure of

the excited chromophore will determine the excited-state conformation of the fluorescing chromophore (Fig. 4B).

MD simulations for 2A50 provided evidence that the chromophore can undergo a *cis*/*trans* isomerication via some HT type motion. The nearby His197 residue shows a small change in its conformation over the MD simulation. This implies that His197 is not involved in a HT *cis*/*trans* isomerization.

Conformational analysis of the kindling proteins 2A52, 2A53, 2A54, and 2A56 showed that fusion of the chromophore back to the protein backbone does force a change in the chromophore conformation, but it is not large as was expected. The φ dihedral angle from 2A52 was the only torsion that showed a significant difference from its fused partner, 2A52fused. However all of the changes in the φ and τ dihedral angles for the crystal structures with cleaved chromophores were still greater than those with the fusing of residues 62 and 63.

References

(1) Shagin, D. A.; Barsova, E. V.; Yanushevich, Y. G.; Fradkov, A. F.; Lukyanov, K. A.; Labas, Y. A.; Semenova, T. N.; Ugalde, J. A.; Meyers, A.; Nunez, J. M.; Widder, E. A.; Lukyanov, S. A.; Matz, M. V. *Mol. Biol. Evol.* **2004***, 21*, 841-850.

(2) Baffour-Awuah, N. Y. A.; Zimmer, M. *Chem. Phys.* **2004***, 303*, 7-11.

(3) March, J. C.; Rao, G.; Bentley, W. E. *Appl. Microbiol. Biotechnol.* **2003***, 62*, 303-315.

(4) Ehrhardt, D. *Curr. Opin. Plant Biol.* **2003***, 6*, 622-628.

(5) Phillips, G. N. *Curr. Opin. Struct. Biol.* **1997***, 7*, 821-827.

(6) Naylor, L. H. *Biochem. Pharmacol.* **1999***, 58*, 749-757.

(7) Misteli, T.; Spector, D. L. *Nat. Biotechnol.* **1997***, 15*, 961-964.

(8) Taylor, D. L.; Woo, E. S.; Giuliano, K. A. *Curr. Opin. Biotechnol.* **2001***, 12*, 75-81.

(9) Zimmer, M. In *Glowing Genes: A Revolution in Biotechnology* Prometheus Books: Amherst, NY, 2005.

(10) Pieribone, V.; Gruber, D. F. In *Aglow In the Dark: The Revolutionary Science of Biofluorescence;* Harvard University Press: 2006.

(11) Chalfie, M.; Kain, S. R., Eds.; In *Green Fluorescent Protein: Properties, Applications, and Protocols (Second Edition);* John Wiley & Sons, Inc.: 2006.

(12) Niwa, H.; Inouye, S.; Hirano, T.; Matsuno, T.; Kojima, S.; Kubota, M.; Ohashi, M.; Tsuji, F. I. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 13617-13622.

(13) Patterson, G. H. *Nature Biotechnology* **2004**, *22*, 1524-1525.

(14) Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y. *Nature Biotechnology* **2004**, *22*, 1567-1572.

(15) Morise, H.; Shimomura, O.; Johnson, F. H.; Winant, J. *Biochemistry* **1974**, *13*, 2656-2662.

(16) Ward, W. W. B., S.H. *Biochemistry.* **1982**, *21*, 4535-4540.

(17) Creemers, T. M. H.; Lock, A. J.; Subramaniam, V.; Jovin, T. M.; Volker, S. *Nature Struct Biology* **1999**, *6*, 557-560.

(18) Heim, R.; Prasher, D. C.; Tsien, R. Y. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 12501-12504.

(19) Wachter, R. M.; King, B. A.; Heim, R.; Kallio, K.; Tsien, R. Y.; Boxer, S. G.; Remington, S. J. *Biochemistry* **1997**, *36*, 9759-9765.

(20) Chattoraj, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8362-8367.

(21) Garcia-Parajo, M. F.; Segers-Nolten, G. M. J.; Veerman, J.-A.; Greeve, J.; Hulst, N. F. v. *Proc. Nat. Acad. Sci.* **2000**, *97*, 7237-7242.

(22) Chirico, G.; Cannone, F.; Diaspro, A. *Journal of Physics D-Applied Physics* **2003**, *36*, 1682-1688.

(23) Chirico, G.; Cannone, F.; Diaspro, A.; Bologna, S.; Pellegrini, V.; Nifosi, R.; Beltram, F. *Physical Review E* **2004**, *7003*, 901-901.

(24) Weber, W.; Helms, V.; McCammon, J.; Langhoff, P. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 6177-6182.

(25) Nifosi, R.; Ferrari, A.; Arcangeli, C.; Tozzini, V.; Pellegrini, V.; Beltram, F. *Journal of Physical Chemistry B* **2003**, *107*, 1679-1684.

(26) Yamada, A., Ishikura, T. & Yamato, T. *Proteins-Structure Function and Bioinformatics* **2004**, *55*, 1063-1069.

(27) Maddalo, S.L. & Zimmer, M. *Photochemistry and Photobiology* **2006**, *82*, (2), 367-372.

(28) Chudakov, D.M., Feofanov, A.V., Mudriku, N.N., Lukyanov, S. & Lukyanoz, K.A. *Journal of Biological Chemistry* **278**, 7215-7219 (2003).

(29) Lukyanov, K.A.; Fradkov, A.F.; Gurskaya, N.G.; Matz, M.V.; Labas, Y.A.; Savitsky, A.P.; Markelov, M.L.; Zaraisky, A.G.; Zhao, X.; Fang, Y.; Tan, W.; Lukyanov, S.A. *Journal of Biological Chemistry* **2000**, *275*, 25879-25882.

(30) Chudakov, D.M.; Verkhusha, V.V.; Staroverov, D.B.; Souslova, E.A.; Lukyanov, S.; Lukyanov, K.A. *Nature Biotechnology* **2004**, *22*, 1435-1439.

(31) Wilmann, P.G., Peterson, J., Devenish, R.J., Prescott, M. & Ross, J. *Journal of Biological Chemistry* **2005**, *280*, 2401-2404.

(32) Binkowski, T.A, Naghibzadeg, S., Liang, J. *Nucl. Acid Res*. **2003**, *31*, 3352-3355.

(33) Chen M. C., C. R. Lambert, J. D. Urgitis, M. Zimmer. *Chem. Phys*. **2001**, *270*, 157–164.

(34) Berman H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne. *Nucl. Acid Res.* **2000**, *28*, 235–242.

(35) Toniolo A., S. Olsen, L. Manohar, T. J. Martinez. *Far. Disc*. **2004**, *127*, 149–163.