DE SILVA INDRASEKARA, AGAMPODI SWARNAPALI. Green Fluorescent Protein as a Targeting Probe. (Under the direction of Stefan Franzen.)

The research presented in this thesis is focused on employing Green Fluorescent Protein (GFP) to design a cancer cell targeting probe along with its inherent ability to serve as a cell imaging agent. In this study, both chemical and genetic approaches were employed to introduce cancer cell targeting peptides to GFP.

N-cadherin targeting synthetic peptides; ADH304_1 (Ac-FHLRAHAVDINGNQVC) and ADH304_2 (Ac-FHLRAHAVDINGNQVC) was used in this study and they were chemically attached to GFP surface lysine residues using the heterobifunctional cross-linker; Sulfosuccinimidyl 4-[N-maleimidomethyl] cyclohexane-1-carboxylate (Sulfo-SMCC). GFP-peptide formulations were stable in the aqueous medium and according to the Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) results these formulations have 3-5 peptides on GFP surface. Cell delivery in A375, HeLa and MCF-7 cell lines provides evidences that GFP with ADH304_2 peptides showed enhanced cancer cell targeting over unmodified GFP.

Selected GFP loops were mutated to contain RGD and HAV amino acid sequences using site directed mutagenesis. Only the RGD mutation at 131-133 and 172-173 positions developed mature chromophore. However a significant difference in cancer cell targeting did not observe between native GFP and mutants from cell delivery study in A375 and HeLa cell lines.

Molecular dynamic analysis of GFP and the above-mentioned GFP mutants were carried out using Visual Molecular Dynamics (VMD) to study the effect of bond parameters that affects the mature chromophore formation. However, analyzing the bond distances and bond angles did not provide sufficient information to derive any correlation between the GFP loop mutations and the chromophore formation.
Green Fluorescent Protein (GFP) as a Targeting Probe

by
Agampodi Swarnapali De Silva Indrasekara

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Chemistry

Raleigh, North Carolina

2011

APPROVED BY:

_______________________________  ________________________________
Dr. Alex Deiters               Dr. Gavin Williams

______________________________
Dr. Stefan Franzen
Committee Chair
DEDICATION

This work is dedicated to my parents who always provide me with endless love, care, support and encouragement. This work is also dedicated to my husband who has been a very supportive and understanding life partner and my only sister for being such an amazing person to me and always being by my side sheltering with love.
BIOGRAPHY

Agampodi Swarnapali De Silva Indrasekara was born and raised in Sri Lanka. She graduated from University of Peradeniya in Sri Lanka with a Bachelors of Science degree in Chemistry in 2007. In 2009 she enrolled in the chemistry graduate program at North Carolina state university and worked towards her Master of Science degree in Chemistry.
ACKNOWLEDGMENTS

Dr. Stefan Franzen
Dr. Alex Deiters
Dr. Gavin Williams
Franzen group members
# TABLE OF CONTENTS

**LIST OF TABLES** ........................................................................................................................................... vii
**LIST OF FIGURES** ......................................................................................................................................... viii

**CHAPTER 1: INTRODUCCIÓN AND BACKGROUND** ................................................................. 1

1.1 Introduction .............................................................................................................................................. 2
1.2 Cancer and Cancer Cell Targeting ......................................................................................................... 2
1.3 Cellular Internalization of Targeting Peptides ..................................................................................... 3
1.4 Cancer Cell Targeting Peptides ............................................................................................................. 4
    1.4.1 N-Cadherin Targeting Peptides ....................................................................................................... 4
    1.4.2 Integrin Targeting Peptides ........................................................................................................... 5
1.5 Cancer Cell Imaging ............................................................................................................................... 6
    1.5.1 Fluorescence Spectroscopy ........................................................................................................... 6
1.6 Green Fluorescent Protein (GFP) ........................................................................................................... 7
1.7 GFP as a Targeting Probe ....................................................................................................................... 9
    1.7.1 Chemical Modifications of GFP ...................................................................................................... 9
    1.7.2 Genetic Modifications of GFP ...................................................................................................... 11
    1.7.3 Molecular Dynamic Analysis of GFP Loop Mutants .................................................................. 13

References .................................................................................................................................................. 17

**CHAPTER 2: MATERIALS AND METHODS** .................................................................................. 23

2.1 Preparation of GFP-SMCC-peptide Formulations ............................................................................... 24
2.2 GFP Mutagenesis ................................................................................................................................... 25
    2.2.1 Thermo Cycling .............................................................................................................................. 25
    2.2.2 Dpn 1 Digestion and Transformation of XL10 Gold Cells ......................................................... 26
    2.2.3 DNA Expression and Purification .................................................................................................. 26
    2.2.4 Protein Growth and Purification .................................................................................................... 26
2.3 Determination of Relative Quantum Yields ......................................................................................... 28
2.4 Cell Targeting Studies ......................................................................................................................... 29
    2.4.1 Preparation of Samples for Cell Targeting Studies .................................................................... 29
LIST OF TABLES

Table 3.1 Relative quantum yields of GFP test samples…………………………36
Table 3.2 Bond distances and bond angles of GFP loop mutants from MD trajectories…………………………………………………………44
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Solid-state structure of GFP. The chromophore is located in the center of β barrel</td>
<td>7</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Proposed mechanisms for the chromophore formation</td>
<td>8</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>Representation of lysine residues of GFP</td>
<td>10</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>Synthetic process of GFP-SMCC-peptide formulations</td>
<td>10</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>Representation of the loops of GFP subjected to mutations</td>
<td>11</td>
</tr>
<tr>
<td>Figure 1.6</td>
<td>Schematic representation of the GFP chromophore forming entity and amino acids in the close proximity</td>
<td>13</td>
</tr>
<tr>
<td>Figure 1.7</td>
<td>Interactions of GFP chromophore with amino acids in the close proximity</td>
<td>14</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Appearances of GFP-peptide formulations prepared from peptides dissolved in (a) DI Water and (b) DMSO</td>
<td>35</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>12% SDS-PAGE gel of GFP-peptide conjugates</td>
<td>35</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Targeted delivery of GFP-peptide conjugates in A375 cells</td>
<td>38</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Targeted delivery of GFP-peptide conjugates in HeLa cells</td>
<td>39</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>Targeted delivery of GFP-peptide conjugates in MCF-7 cells</td>
<td>40</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>Targeted delivery of GFP mutants in A375 cells</td>
<td>42</td>
</tr>
<tr>
<td>Figure 3.7</td>
<td>Targeted delivery of GFP mutants in HeLa cells</td>
<td>43</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction and Background
1.1 Introduction

Much research is focused on designing multifunctional nanoparticulate systems that possess cancer cell targeting, imaging and drug delivery properties. The ultimate goal of the research explained in this thesis is to design an alternative cancer cell targeting probe with imaging properties. Green fluorescent protein (GFP); a fluorescent protein isolated from jellyfish *Aequorea Victoria* which is widely used as an imaging agent is the central target for this purpose.

In this research, GFP was used to design a cancer cell targeting probe along with its inherent ability to serve as a cell imaging agent. In this study, both chemical and genetic approaches were employed to introduce cancer cell targeting peptides to GFP.

N-cadherin targeting synthetic peptides; ADH304_1 (Ac-FHLRAHAVIDINGNQVC) and ADH304_2 (Ac-FHLRAHAVIDINGNQVC) was used in this study and they were chemically attached to GFP surface lysine residues using the heterobifunctional cross-linker; Sulfosuccinimidyl 4-[N-maleimidomethyl] cyclohexane-1-carboxylate (Sulfo-SMCC).

As the genetic approach of designing GFP as a targeting probe, some selected GFP loops were mutated to contain RGD and HAV amino acid sequences using site directed mutagenesis. RGD and HAV are the recognition motifs of the integrin targeting and N-cadherin targeting peptides.

Some of the GFP loop mutants were fluorescent while some were fluorescently inactive. In order to understand the factors that affect the chromophore maturation, the Molecular Dynamic (MD) simulations were carried out and MD trajectories of those mutants were analyzed for the fluctuations of selected bond distances (BD), bond angles (BA) and dihedral angels (DA).

1.2 Cancer and Cancer Cell Targeting

Over the past few decades, a vast majority of the research in nanotechnology and biomedicine has dedicated to the development of cancer therapeutics. The ultimate goal of all these approaches is to overcome the deleterious side effects of the chemotherapeutics and radiotherapeutics and to optimize the effective destruction of the tumor cells with targeting
specificities\textsuperscript{1,2}. Although efficient carrier vectors for drug delivery have been designed, cancer therapy is still limited due to the lack of targeting specificity\textsuperscript{3}. Many efforts and approaches have been pursued in achieving the specific cancer cell targeting.

The first approach was to use antibodies for selective tumor localization and drug delivery. This has been a highly promising targeting strategy due to their ability to specifically bind to primary and metastatic cancer cells with high affinity\textsuperscript{2,4}. But the tumor specific antigens that are recognized by antibodies are not feasible for treatment by antisense, SiRNA and drug delivery\textsuperscript{3}.

Cell targeting by small peptides emerged as a better strategy than targeting by large protein and antibodies due to their enhanced tissue penetration, ease of synthesis, availability for conjugations with cancer therapeutic agents and their selective binding to cell surface receptors specific for tumors and internalization by cancer cells\textsuperscript{1,3,5}. The peptides used in these approaches have been derived from the natural peptides and proteins that target cell surface receptors such as growth factors, hormones and cytokines\textsuperscript{6-8}. Additionally, the functional elements of viral coat proteins, antibodies, and affibodies have also studied and used to designed synthetic cancer cell targeting peptide fragments\textsuperscript{9}. Differences in the metabolic properties between healthy and cancerous cell have also been used in designing targeting agents\textsuperscript{10,11}. For example, folic acid has been used as a targeting agent due to the fact that α-hydrofolate binding protein αhFBP is over-expressed in many cancers\textsuperscript{12}.

1.3 Cellular Internalization of Targeting Peptides

Clathrin coated pit endocytosis, caveolar endocytosis, phagocytosis and macropinocytosis are the main cellular uptake routes for targeting peptides\textsuperscript{10}. Uptake of large molecules up to 10µm such as growth factors takes place via macropinocytosis. This internalization mechanism involves the activation of the signaling pathways that trigger actin-mediated membrane ruffling and blebbing followed by the formation of macropinosomes at the plasma membrane, which allows the internalization of targeting moieties\textsuperscript{13}.
Phagocytosis occurs when the targeting vectors are accompanied with antibodies and targeting peptides. This occurs mainly in cells that belong to the reticuloendothelial system. For example, it has shown that the particles coated with integrin-binding fibronectin peptide internalized through phagocytic pathway\textsuperscript{14}. Metabolic and growth factor receptors undergo receptor-mediated endocytosis for internalization, which involves the clathrin-coated pit formation followed by the endosomes formation. Unless the presence of a peptide that allows the endosomal escape it will destroy due to the recycling of endosomes by lysozyme action.

1.4 Cancer Cell Targeting Peptides

Much research has been conducted on targeted delivery of chemotherapeutic drugs specifically to the site of tumor avoiding the damage to healthy cells. There are several types of peptides that use for cancer cell targeting such as cell penetrating peptides, endosomolytic peptides, peptides that targets growth and regulatory receptors, cell adhesion proteins and metabolic products\textsuperscript{15-18}. In this study N-cadherin targeting peptides and RGD motif that belong to cell adhesion peptide category was of interest.

1.4.1 N-Cadherin Targeting Peptides

Cadherins belong to a super family of transmembrane proteins that plays a key role in maintaining the tissue integrity by forming intercellular adhesive complexes such as adheren junctions\textsuperscript{19-21}. E-cadherin can be predominantly observed in epithelial cells and forms the intercellular junction through homophilic protein interactions. The Cadherins contains evolutionary conserved Histidine-Alanine-Valine (HAV) domain. The interaction between HAV domains of two homophilic Cadherin molecules on the cell’s surface facilitates the intercellular adhesion\textsuperscript{19-21}. During a tumor progression normal epithelium cells transformed to poorly differentiated tumor cells that resemble mesenchymal cells, which exhibit high motility and invasiveness\textsuperscript{19-22}. With this epithelial-mesenchymal transition, expression of N-cadherins (neuronal) becomes prominent while losing the E-cadherin expression. Over expression of N-cadherin has been observed in solid tumors such as breast, prostate and bladder cancer. Interaction of N-cadherin with the fibroblast growth factor receptors induces
promotes the cell invasion\textsuperscript{22}. The peptides have designed to contain the HAVD sequence, which mimics the natural N-cadherin and thereby making the peptide more effective as a N-cadherin inhibitor\textsuperscript{21}. The interaction of HAV motif of these peptides and N-cadherin results the inhibition of cadherin dependent cascades and also increases the porosity of the intercellular junctions of the cells and thereby facilitating the entry to the cells\textsuperscript{23}. The peptides have designed to contain the HAV sequence, which mimics the natural N-cadherin and thereby making the peptide more effective as a N-cadherin inhibitor\textsuperscript{19}. ADH-1 (Ac-CVAHC), ADH303 (Ac-FHLRAHAVDINGNQVK) and ADH304 (Ac-FHLRAHAVDINGNQVC) are examples for N-cadherin targeting peptides\textsuperscript{24}. In this study Ac-FHLRAHAVDINGNQVC and Ac-FHLRAHAVDINGQVC peptides have been utilized and they are abbreviated as ADH304 \textsubscript{1} and ADH304 \textsubscript{2} respectively.

1.4.2 Integrin Targeting Peptides

Integrins are a class of heterodimeric transmembrane receptor proteins, which play an important role in cell adhesion, proliferation, differentiation, migration as well as tissue morphogenesis\textsuperscript{25, 26}. They facilitate the interaction between the extracellular matrix and internal cellular structure\textsuperscript{25, 26}. These heterodimeric proteins consist of $\alpha$ and $\beta$ subunits and the combination of $\alpha$ and $\beta$ subunits provides the ligand binding specificity to integrin proteins. It was also discovered that Arginine-Glycine-Asparagine (RGD) is the smallest active unit / primary recognition site in extracellular matrix proteins which acts as the ligands for integrins\textsuperscript{27}. However the flanking amino acid sequence of RGD motif provides this ligand the integrin binding specificity\textsuperscript{28}. It has been observed that the composition of different types of integrins on cell surfaces changes with the tumor development. Changes in the integrin composition mediates the metathesis and invasiveness of tumor cells and researches have been employing these altered integrin sequences for the development of diagnosis and specific targeting of cancers\textsuperscript{26, 29}. In particular, up regulation of $\alpha 5 \beta 1$ and $\alpha v \beta 3$ integrins is observed on epithelial cells while loss of $\alpha 2 \beta 1$ integrin is significant during the progression of aggressive tumors\textsuperscript{26}. Ligands that contains RGD motif and show selectively affinity towards to $\alpha 5 \beta 1$ and $\alpha v \beta 3$ integrins have been identified by phage display studies\textsuperscript{30}. The
information from those ligands have been used in designing RGD containing peptides and RGD mimetic and the literature provides sufficient amount of evidences for the use of short peptides containing RGD for cancer cell imaging and targeting. It has also been shown that the cyclic RGD peptides are much more potent than short linear peptides.

1.5 Cancer Cell Imaging

Biological imaging techniques provides valuable cellular and sub cellular information that can be employed in diagnostic purposes and development in medical field. Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) are the most widely used imaging techniques for clinical purposes. Additionally, fluorescence spectroscopic techniques based on probes that are designed with small fluorescent molecules such as Fluorescein Isothiocyanate, fluorescence proteins such as Green Fluorescent Protein (GFP) and Surface Enhanced Raman Scattering techniques have also drawn much attraction as imaging techniques. All these techniques have both pros and cons. However, with the discovery of GFP and different varieties of fluorescent proteins, fluorescent proteins based imaging techniques have drawn more attraction. Fluorescent proteins are widely used as genetic tools for cell, organelle and protein labeling as well as organism imaging. Moreover GFP and related proteins have been extensively used to study interactions, as sensors and drug screening. The more interest toward fluorescent proteins over other techniques is due to their biocompatibility, high sensitivity and relative high photostability.

1.5.1 Fluorescence Spectroscopy

Fluorescence is an optical phenomenon that is exhibited by molecules with a fluorophore. Fluorophore is a functional component in a molecule that absorbs energy at a specific wavelength and re-emits the energy at a different but specific wavelength. Electrons of a molecule at rest reside in the singlet ground state. When a fluorophore is excited with light of appropriate energy, electrons in the singlet ground state are promoted to a singlet excited state. This excitation process occurs rapidly in the order of 1fs. Upon excitation to a higher singlet excited state, it relaxes through internal conversion to higher
vibrational levels of the first excited states and then to its lowest vibrational level via vibrational relaxation. From there it relaxes back to the singlet ground state by the emission of energy as radiation. This process is called “Fluorescence”.

1.6 Green Fluorescent Protein (GFP)

In 1961, the green fluorescent protein (GFP) was discovered during the isolation of the bioluminescent protein, aequorin from jellyfish *Aequorea Victoria* and it resulted a remarkable improvement in the fields of cell biology, molecular biology and medicine.\textsuperscript{39-41} GFP is composed of 238 amino acids and has folded up to a 11-stranded β-barrel structure. These 11 β-sheets forms the wall of the barrel structure, which is 24 Å in diameter and 42 Å in height and a α-helix, goes diagonally through the can.\textsuperscript{39-41} The three consecutive amino acids in the central helix, Serine 65 (S65), Tyrosine 66 (Y66) and Glysine 67 (G67) undergo autocatalytic reactions to form the GFP chromophore.\textsuperscript{40,42}

The chromophore, para-hydroxybenzylidene-imidazolinone is well protected from the quenching by oxygen, bulk solvent due to its location in the middle of the β barrel and the shielding by an extensive network of hydrogen bonds around the surface of the β-barrel (Figure 1.1).

![Figure 1.1: Solid-state structure of GFP. The chromophore is located in the center of β barrel](image)

There are three main proposed mechanisms for the chromophore formation depending on the order of oxidation, cyclization and dehydration reactions that leads to the mature
chromophore that emits glowing green light\textsuperscript{43}. Figure 1.2 shows the main proposed mechanisms for GFP chromophore formation and researches have adopted once of these mechanisms for further studies based on their interests.

**Figure 1.2: Proposed mechanisms for the chromophore formation\textsuperscript{43}**

GFP has used extensively used in cell biology as a biological marker\textsuperscript{44}, fusion tag reporter gene\textsuperscript{45} as well as an indicator for calcium sensitivity\textsuperscript{46}, protease action\textsuperscript{47} etc. As mentioned earlier, the wide applications of GFP in the development of biology and medical research area is mainly due to its autocatalytic chromophore formation ability, photostability and biocompatibility.

Researchers have mutated the native GFP to tune the excitation maximum, brightness, and photo stability as well as to obtain different colored fluorescent protein. Among them the synthesis of Enhanced Green Fluorescent Protein (EGFP) is remarkable in its applications in the field of microscopic cell imaging. Though the native GFP possess higher fluorescent and relative photo stability, it has limitations in cell imaging due to its excitation maximum 395 nm which lies in the ultraviolet region which demands modifications in the microscopic techniques and causes damage to cells during imaging. But the mutation in the chromophore region Serine 65 to Tyrosine 65, which yields EGFP, overcomes this limitation by shifting the excitation maximum to 488nm in the visible
region. Additionally EGFP is the brightest and photostable GFP available and due to all the features it has been widely used in cellular imaging applications.

In this study, we use monomeric EGFP, which has the excitation maximum 490 nm, emission maximum 510 nm, molar extinction coefficient 55000 and quantum yield of 0.60. For the ease for use, term GFP will be used instead EGFP in the rest of the context.

1.7 GFP as a Targeting Probe

The main goal of this study is to employ GFP to design a dual purpose probe which posses both cancer cell targeting potentials along with its inherent ability to serve as a cell imaging agent. In this study, both chemical and genetic approaches were employed to introduce cancer cell targeting peptides to GFP.

1.7.1 Chemical Modifications of GFP

As shown below in Figure 1.3, GFP has total of 19 Lysine residues and 18 of them are located on the outer surface and available for chemical reactions. We have employed the most widely used heterobifunctional cross-linker: Sultosuccinimidyl 4-[N-maleimidomethyl] cyclohexane-1-carboxylate (Sulfo-SMCC) through which lysine residues of GFP can be linked to Cysteine-bearing cell targeting peptide moieties. Heterobifunctional reagents contain two different reactive groups that can be used to link proteins and other macromolecules by a stepwise reaction. In SMCC, the NHS ester end can react with primary amine group, in this case amino group of Lysine of GFP to form a stable amide bond. The maleamide end of SMCC is specifically coupled to sulfhydryls when the pH of the reaction medium is in the range of 6.5-7.5. pH of the reaction medium plays a critical role in the reaction between the maleimide and sulfhydryl group. The maleimide group undergoes hydrolysis below pH 6.5 and become unreactive toward sulfhydryl group. Thus controlling pH of the reaction medium of GFP, Sulfo-SMCC and peptide is critical to achieve successful conjugation reaction. The targeting peptides of interest are designed in such a way that they contain cysteine residues at the end facilitating its sulfhydryl group to react with maleamide group of SMCC and there by introducing cell targeting peptide on GFP surface (Figure 1.4).
Figure 1.3: Representation of lysine residues of GFP

Figure 1.4: Synthetic process of GFP-SMCC-peptide formulations
1.7.2 Genetic Modifications of GFP

It has been experimentally proven that both cyclic HAV peptides and cyclic RGD peptides show greater interaction with their recognition moieties on cancer cell membrane and thereby more potent than their linear counterparts\textsuperscript{31}. Here we employ the loop structure of GFP, which mimics the cyclic presentation and mutate some of the GFP loops to have 1 HAV or RGD sequence per loop. If these mutations do not interfere with the chromophore maturation of GFP, it should be possible to create a GFP mutant protein labeled with a cancer cell targeting sequence. Such a protein would possess both fluorescent as well as cell targeting abilities. By making genetic modifications, any difficulties associated with the chemical modifications can be avoided and GFP can associate with a defined number of targeting moieties. Mutated loops and their nomenclature are shown in Figure 1.5. More details on mutated region are also given below.

\textbf{Figure 1.5:} Representation of the loops of GFP subjected to mutations
1. Loop H

GFP: 130- FKE_DGNLGHKL-141

Mutants: FHAVGNILGHKL
FRGDGNILGHKL

2. Loop J

GFP: 171- IEDGSVQLAD-180

Mutants: IHAVSVQLAD
IRGDSVQLAD

3. Loop M

GFP: 208- SKDPNEKRDH-217

Mutants: SKDPHAVRDH
SKDPRGDRDH

GFP loop mutation is achieved by site directed mutagenesis. The mutant strand synthesis reaction is employs *Pfu Turbo* DNA polymerase and thermal cycling. As the first step of the thermal cycling, the double stranded DNA vector that contains the insert of interest; in this case Pet16b vector with EGFP insert, is denatured and then annealed with the oligonucleotide primers containing the desired mutation. The nonstrand-displacing action of *Pfu Turbo* DNA polymerase can be used both for the replication and extension of DNA strands with high fidelity leading to incorporation of the mutagenic primers resulting in nicked circular strands. Mutated DNA strands are separated from parental DNA strands by *Dpn I* digestion. Parental DNA strands are methylated while newly synthesized mutated DNA strands are non-methylated. Dpn1 endonuclease targets the 5’-Gm6 ATC-3’ of methylated parent DNA strands, digest them leaving the newly synthesized mutant DNA strands. Transformation of the nicked mutant DNA strands in *E. coli* which repairs the nicks
followed by protein growth will result the mutant GFP bearing the cancer cell targeting sequences.

1.7.3 Molecular Dynamic Analysis of GFP Loop Mutants

It has already been shown experimentally that the mutations in the chromophore region and some of the amino acids in the close proximity affect the rate of one or all steps of the chromophore formation and thereby the maturation of chromophore\textsuperscript{43}. The influence of R96, E222, and Y62 on chromophore synthesis and structural stability has been extensively studied\textsuperscript{41-43, 50-54}. The spatial arrangement of these amino acids around the chromophore is shown in figure 1.6.

![Figure 1.6: Schematic representation of the GFP chromophore forming entity and amino acids in the close proximity](image)

Studies have also shown that S65, Y66 and G67 sequence is present in some non-fluorescent proteins and they undergo neither cyclization nor tyrosine oxidation\textsuperscript{50}. This suggests that the presence of the tripeptide sequence is by itself not sufficient for chromophore formation. The interactions of these amino acids with the local environment
are equally important and play a major role in rate of chromophore formation, but not necessarily directly involve in chromophore biosynthesis. Instead the amino acids located in the close proximity play catalytic, steric or electrostatic roles in protein folding and conformational stability of GFP.

Interaction of E94, R96, H148, T203 and E222 amino acids with chromophore has been extensively discussed in the literature that the current hypothesis states that basic residues such as E94, R96, and H148 stabilize the chromophore via charge delocalization\(^50\). Hydrogen bonding between T203 and Y66 stabilizes the phenolate form of the chromophore while E222 stabilizes the hydroxyl form of the chromophore through electrostatic interaction\(^50\). Figure 1.7 shows the interactions of the side chains of the surrounding amino acids to the chromophore region.

![Figure 1.7: Interactions of GFP chromophore with amino acid in the close proximity](image)

Among them the effect of R96 and E222 on the rate of maturation and the stability of GFP chromophore has been extensively studied. Reid \textit{et.al} have clearly shown that the process of chromophore formation consists of three main kinetic steps in the order of fairly slow protein folding, ring cyclization occurs at a moderate rate and the oxidation of the cyclized chromophore, which is the final step\(^54\). The rate of each of these steps determines the rate of maturation, which could take from minutes to months.
It has also been experimentally evidenced that the hydrogen bonding between the guanidinium group of Arg96 and carbonyl oxygen from Tyr66 in the chromophore enhances the acidity of the G67 nitrogen, which in turn facilitates the cyclization reaction. Additionally the electrostatic interaction of Arg96 with carbonyl oxygen from Tyr66 induces the favorable structural rearrangement important in aligning the molecular orbitals for ring cyclization. Further it acts as a base in the dehydration step of the maturation.

E222 acts as a base facilitating the proton transfer in the maturation process and deprotonation of G67 amide. Additionally the hydrogen bonding of E222 to S65 and Y66 stabilizes the intermediates involved in the maturation process. Experimental evidences found in the literature suggesting that T62 carbonyl oxygen play a role of a base and facilitate the dehydration reaction of the chromophore maturation process.

Bond distances correlations have been established between the critical amino acids in the close proximity and chromophore region to the fluorescence activity of GFP and the rate of chromophore maturation.

If the driving forces for the chromophore maturation and stability and the involvement of the proximal amino acids in these processes can be clearly determined, it would be helpful in making fluorescent active mutant GFP for many applications in biology such as cell localization.

Our interest in computational study of the driving forces for mature GFP chromophore formation arose due to the necessity of establishing systematic approach for efficient and effective GFP mutagenesis to obtain fluorescent active GFP loop mutants that can be used in the cancer cell targeting and cell localization research.

The mutagenesis and protein purification process are time and resource consuming. If the possibility of a given mutant to be fluorescent active can be evaluated by computational analysis before the experimental synthesis of mutants, it would be a more efficient and systematic approach in developing GFP mutants to serve as a targeting probe. The basis of this molecular dynamic simulation study of the GFP mutants is to compare and contrast the bond distances and bond angle fluctuations of mutants to wild type GFP and thereby establishing correlations to predict the fluorescent nature of mutants. For that purpose, the
Molecular Dynamic (MD) simulations were carried out and MD trajectories of those mutants were analyzed for the fluctuations of selected bond distances (BD), bond angles (BA) and dihedral angels (DA)
References


Chapter 2: Materials and Methods
2.1 Preparation of GFP-SMCC-Peptide Formulations

Two different ADH304 peptides (GenScript) were used for this study and they were abbreviated as below for the ease of use; ADH 304_1 peptide (Ac-FHLRAHAVDINGNQVC-NH₂) and ADH304_2 (Ac-FHLRAHAVDINGQVC-NH₂). 10 µM, pH 7.2 GFP desalted with NAP-25 column (GE healthcare) using 50mM phosphate buffer (chemicals from Fisher), pH 7.2 was used for this study. GFP was prepared by following the same protein growth and purification protocol described under GFP loop mutagenesis.

First, 1 mg of Sulfo-SMCC (Thermo Scientific) solubilized in 100 µL of DI H₂O by sonicating for 5 min was mixed with 900 µL of 10mM GFP and the pH of the mixture was determined. The mixture was then incubated at room temperature for 30 minutes. Following this reaction, the excess Sulfo-SMCC was cleaned up by size exclusion column chromatography with NAP-25 columns pre-equilibrated with 30mL of 50mM sodium phosphate buffer at pH 7.2. The sample was eluted with the same buffer and the fractions at the highest GFP concentrations were combined. 1mg of the peptide of interest was dissolved in 100µL DMSO (MP Biomedicals) and kept in 37°C water bath for 30 min. 1mL of GFP-SMCC mixture was then combined with 25 µL of peptide. The pH of the resulting solution was determined and the samples were then incubated overnight at room temperature. 1 mL of the GFP-SMCC mixture was kept as the control sample. Following this reaction, the samples were purified over NAP-25 columns pre equilibrated with 30mL of 50mM sodium phosphate buffer at pH 7.2 by removing the excess peptides. The sample was eluted with 50mM sodium phosphate buffer at pH 7.2 and nanodrop UV –Visible spectroscopy at 490 nm, determined the GFP concentration of purified sample.

The above-mentioned GFP-peptide preparation method was followed with different amounts of the peptide per reaction (250µg, 750µg and 1000µg of peptide) to determine the optimum peptide concentration.
In order to determine the best solvent to dissolve peptide, the above-mentioned GFP-peptide preparation method was carried out with peptide dissolved in DI water and DMSO separately with both 250µg and 750µg of peptide.

GFP, GFP-SMCC and all the GFP-peptide formulations were analyzed by SDS polyacrylamide gel electrophoresis using 12% SDS pre-cast gel (Biorad) at 150 V to estimate the number of peptide attach to GFP. The gel was stained using bio-safe Coomassie (Biorad) and destained using DI water.

2.2 GFP Mutagenesis

The N-terminal 6xHis GFP PET16b vector was obtained from Tim Sit, Department of Plant Pathology, NCSU. All the primers were ordered from Eurofins MWG Operons (primer designing was done by Dustin Lockney). The GFP loop H (131 K 132 E 133D), loop J (172E 132D 133G) and loop M (212N 213E 214K) were mutated to RGD and HAV according to Strategen Quick-change Site-Directed Mutagenesis and all the reagents for mutant strand synthesis reaction (Thermal Cycling) was provided with it.

2.2.1 Thermo Cycling

For each mutagenesis reaction, 1µL 6xHis GFP PET16b vector (25ng/µL), 1µL dNTP(10mM), 1µL 3’ primer (10pmol/µL), 1µL 5’ primer (10pmol/µL), 1µL PfuTurbo DNA polymerase (2.5 U/µL), 5 µL of 10x Pfu reaction buffer and 41 µL nuclease free water was mixed together in a PCR tube. Mutagenesis was performed at different annealing temperatures, 52°C, 55°C, 57°C, 60°C, 63°C and 66°C and the cycling parameters for the Site-Directed Mutagenesis is as below. Denaturation of DNA was performed at 94 °C for 1 minute for 1 cycle and 30 seconds for 30 cycles. Annealing was performed at previously mentioned temperatures for 1 minute for 30 cycles and the extension was at 72 °C for 10 minutes for 30 cycles. Mutagenesis samples annealed at 63°C were used for the further experiments described below.
2.2.2 Dpn I Digestion and Transformation of XL10 Gold Cells

Following the thermo cycler step, the amplified DNA was subjected to Dpn I digestion. 1µL of 10 U/µL Dpn I restriction enzyme (Promega) was added to the amplified DNA, mixed the reaction mixture thoroughly and incubated at 37°C for 1 hr.

2µL of β-Mercaptoethanol was added to 50µL of XL-10 gold ultra competent cells (Agilant Tech) and incubated on ice for 10 minutes. Then 1µL of the Dpn I treated mutant DNA was added to it and incubated on ice for 30 minutes. After that the reaction mixture was heat shocked at 42°C for 30 s and incubated on ice for 2 minutes. Next 450 µL of NZY+ broth was added and the mixtures were incubated on 37°C for 1 hr. After incubation, 125µL of the transformation mixture was spread on LB agar plates containing ampicillin and the remaining 375µL on a second plate and they were incubated at 37°C overnight.

2.2.3 DNA Expression and Purification

Following the transformation, 5mL of LB broth (Fisher) with 5 µL of ampicillin was inoculated with a single colony from the LB agar plate and it was incubated at 37°C overnight shaking at 300rpm. This 5mL growth was carried out for 4 different colonies for each mutant. GFP mutant DNA was extracted from the 5 mL cell culture using the E.Z.N.A plasmid miniprep kit I. The concentration and purity of the extracted mutants DNA was recorded using the Nanodrop UV-visible absorption spectrometer and stored at -20 °C. The purified mutant plasmids were sequenced from genewiz Inc before using them for protein growth.

2.2.4 Protein Growth and Purification

BL21 cells thawed in ice for 15 minutes were transformed with 10ng of mutant GFP DNA and it was incubated on ice for 15 minutes followed by heat shock treatment at 42 °C for 2 minutes. 400 µL of NZYM+ was added to the above mixture and incubated at 37°C for 1 hour. 125µL of the cell culture was spread on a LB-ampicillin agar plates and the rest on another LB-ampicillin agar plates and they were incubated overnight at 37°C.
A single colony was picked and 50 mL of LB broth with 5µL of ampicillin was inoculated with it following the incubation overnight at 37°C with shaking at 300 rpm. On the following morning, 6L of 2XYT growth medium with 100µL/mL ampicillin was inoculated with 1mL of overnight inoculum per 1L of growth media and it was shaken for 7-8 hrs at 37°C at 250 rpm. After 8 hrs, the heater was turned off and allowed to shake at room temperature overnight. In the morning, cells were pelleted using the Piramoon, F10S rotor at 7000 rpm for 20 minutes. The supernatant was decanted; the cell pellet was removed from the bottle and deposited into a plastic 200 mL beaker. The weight of the cell pellet was recorded. The collected cell pellet was resuspended in lysis buffer (20mM imidazole, 50mM phosphate buffer, 100mM KCl, pH=7.5) at 2mL/g of cell pellet weight) followed by the addition of lysozyme from chicken egg white to a final concentration of 1mg/mL and DNase 1 and RNase 1 to a final concentration of 5µg/mL. This reaction mixture was stirred for 1 hr at room temperature and stored at -20°C freezer overnight. In the morning, the frozen cell slurry was thawed on bench top and once thawed it was stirred for approximately 1 hr at room temperature. The viscosity of the cell slurry was checked and the cell slurry was sonicated on ice for 15 minutes. After that the cell slurry was pelleted by centrifugation at 17000rpm for 30 minutes at 4 °C using KOMP KA 21.50 rotor and the supernatant was collected. Purification of 6X His tagged GFP mutant proteins from the supernatant was performed using Ni-NTA agarose column according to the Qiagen manual.

First, Ni-NTA column (10 mL volume of resin) was washed and equilibrated with 100 mL of wash buffer (100 mM NaH₂PO₄, 100 mM NaCl, 20mM Imidazole, pH 7.0). Then 60 mL of the collected supernatant was loaded on to the column, washed with 100 mL of wash buffer and eluted with 100 mL elution buffer (100 mM NaH₂PO₄, 100 mM NaCl, 250mM Imidazole, pH 7.0). Collected fractions were analyzed by nanodrop UV-visible spectroscopy at 490nm and the fractions with highest absorbance (which have higher GFP concentration) were combined together and further purified by dialysis. Ni-NTA column purified sample was placed in molecular weight cutoff 6000-8000 tube and dialyzed against 4L dialysis buffer (50mM sodium phosphate buffer, pH 7.3) with stirring at 4°C overnight.
and it was repeated. The dialyzed GFP/ mutant GFP was filtered using 0.2µm filter and frozen at 4°C for storage.

2.3 Determination of Relative Quantum Yields

The quantum yield is an indication of the efficiency of the fluorescence process. It can be defined as the ratio of the number of fluorescence photons emitted by the fluorophore to the number of photons absorbed\(^1\). Here, the quantum yield of GFP-peptide conjugates and GFP loop mutations were calculated and compared. The gradient method\(^2\) was used to determine the relative quantum yields. In this method, the integrated fluorescence intensity is plotted against the absorbance for both standard and test solutions and the gradient from the resulting straight line is used in the following equation to calculated the relative quantum yield of the test sample;

\[
\phi_X = \phi_{ST} \frac{\text{Gradient}_X}{\text{Gradient}_{ST}} \left( \frac{n_X^2}{n_{ST}^2} \right)
\]

Where, \(X\) and \(ST\) denotes the test sample and standard solution respectively. \(\phi\) is the quantum yield and \(n\) is the refractive index of the solvent.

The fluorescence and absorbance measurements were conducted using a Perkin-Elmer Luminescence spectrometer and Hewlet Packard 8453 spectrometer equipped with the HP 845x UV-visible system, respectively, using 1.5 mL of test solution in 10 mm path length quartz fluorescence cuvette (Starna cells).

A solution of fluorescein in 0.1M NaOH was used as the fluorescence standard. 0.1M NaOH was used as the blank solution for fluorescein while 50mM phosphate buffer at pH 7.2 was used as the blank solution for all the GFP test samples. Eight test solutions for the standard and each test sample were prepared by dilution corresponding to absorbance of 0.012, 0.008, 0.006, 0.004, 0.003, 0.002, 0.0012 and 0.0006 at the excitation wavelength of 490nm.

The absorbance measurement of the stock solution for both standard and test samples were recorded at 490 nm and the absorbance of the solutions in the above dilution series were calculated using the Beer-Lambert law. The calculated absorbance values were corrected for the absorbance of the corresponding blank solution at the same wavelength and used to
determine the fluorescence quantum yield. Fluorescence measurements were recorded at \( \lambda_{\text{excitation}} = 490 \) nm, scan rate =250 nm/min and scan width =5 nm. The corrected and integrated fluorescence intensity was used to calculate the fluorescence quantum yield each of the samples according to the gradient method. The refractive index of phosphate buffer; solvent of the samples is assumed to be equal to that of water at 20 \(^{\circ}\)C, which is 1.33 and that of NaOH in which fluorescein was prepared, is 1.412.

2.4 Cell Targeting Studies

HeLa (Human cervical cancer) cells, MCF-7 (human breast adenocarcinoma) cells and A375 (human melanoma) cells were purchased from American Type Culture Collection (Rockville, MD). Minimum Essential Medium Eagle (MEME), DMEM, fetal bovine serum (FBS) were purchased from Bio-whittaker, Inc (Walkersville, MD). HeLa and MCF-7 cells were cultured in EMEM while A35 were cultured in DMEM media supplemented with 10% FBS and 1% penicillin-streptomycin (Bio-whittaker, Inc) at 37 \(^{\circ}\)C in a 95% air-5% CO2 atmosphere.

2.4.1 Preparation of Samples for Cell Targeting Studies

GFP, GFP-SMCC GFP-peptide formulations, GFP loop mutants and 50mM sodium phosphate buffer at pH 7.2 in which all the above samples have prepared used as the test samples. All the test samples were filter sterilized using 0.2\( \mu \)M syringe filter (Corning) before cell studies. The GFP concentration of the filter sterilized test samples was determined by UV –Vis spectroscopy at 490 nm and was reconstituted in appropriate amount of the 50mM sodium phosphate buffer at pH 7.2 to bring the initial GFP concentration to 2\( \mu \)M.

2.4.2 Sample Delivery in Cancer Cells

The day before the cell targeting experiment, 96-well plate (Corning) was seeded with 1.0\( \times \)10^4 cells per well in 100 \( \mu \)L of the appropriate complete growth medium with 100% Fetal Bovine serum (FBS) and streptomycin but without phenol red. Three columns were kept empty as a control.
After 24 hrs, the growth media was removed and replaced with 100 µL of fresh growth medium without Phenol red following the addition of 200 µL of test samples at 2µM concentration to the first row of wells. A serial dilution of the samples was then performed and incubated under humidified conditions with CO₂ and 37 °C. As control experiments, cells were also treated with 50mM sodium phosphate buffer at pH 7.2 and the complete cell growth medium separately and the complete growth medium without cells was analyzed as the blank. All experiments were performed in triplicates.

After 48 hrs, the fluorescence of samples under 560/580 filters was recorded using the Synergy HT BioTek plate reader with KC4 V.31 program and re-recorded after washing the cells twice with 100µL DPBS buffer (Bio-whittaker, Inc).

2.4.3 AlamarBlue Cell Viability Assay

AlamarBlue reagent is a non-toxic and rapid indicator of cellular health. Resazurin, the active component of this reagent is blue in color, non-fluorescent and cell permeable. In the presence of live cells, Resazurin is reduced to Resorufin, which is red colored fluorescent compound. The fluorescence intensity as well as the intensity of the red color increases with the amount of live cells and thereby it can be used as an accurate quantitative measurement of the proliferation of cells to establish the cytotoxicity of chemical agents.

Cells treated with GFP test samples and incubated for 48 hrs followed by washing twice with DPBS were used for this experiment. Each well was treated with 100 µL of the relevant cell growth medium without phenol red followed by the addition of 10µL of the 10 X AlamarBlue reagents (Invitrogen). Cells were incubated for 4 hours at 37 °C in a cell culture incubator and after that the absorbance of AlamarBlue at 570nm was recorded by using the plate reader.

2.5 Molecular Dynamic Analysis of GFP Loop Mutants

Wild-type GFP X-ray crystal structure (1GFL) was obtained from the Protein Data Bank and was modified to obtain immature chromophore (1GFL_broken). Minimized 1 GFL_broken structure was used to perform mutations and all computational analysis.
The GFP loop mutations were performed using Visual Molecular Dynamics (VMD) on the minimized 1GFL _broken structure followed by solvation and ionization. The MD simulations were carried out in NCSU High Performance Cluster where NAMD is employed. MD simulation was carried out at 298 K for 100ps on energy minimized structures and the trajectories were stored every 1 ps. Two additional runs were performed under same conditions for detailed analysis.

Quantitative analysis of trajectories for fluctuations of the atomic coordinates of interests was carried out in VMD. As quantitative analysis, changes in selected bond distances, bond angles and dihedral angles with time were plotted and statistical values such as mean, median, minimum and maximum were calculated.

These statistical values of mutant GFP were compared to wild- type GFP to establish any relationship between mature chromophore formation to bond distances, bond angles and dihedral angles of interest.
References:

(1) Periasamy, A.; Elangovan, M.; Elliott, E.; Brautigan, D. L. In Fluorescence Lifetime Imaging (FLIM) of Green Fluorescent Fusion Proteins in Living Cells; Green Fluorescent Protein; 2002; Vol. 183, 89-100.

Chapter 3: Results and Discussion
3.1 GFP-Peptide Conjugations

3.1.1 Optimizing the Conjugation Reaction Conditions

1mg of Sulfo-SMCC completely dissolved in 100 µL of DI water was allowed to react with 900 µL of 10 µM GFP for 30 minutes at room temperature. The resulted GFP-SMCC solution was separated from excess Sulfo- SMCC by size exclusion chromatography and 1 mL of purified GFP-SMCC was then used for conjugation reaction with ADH304 peptides. In order to determine the best solvent to dissolve ADH304 peptides and the optimum peptide concentration that will yield more stable GFP-SMCC-peptide formulations, reactions of 1 mL GFP-SMCC with 25 µL, 75 µL and 100 µL of 10 µg/µL ADH304 peptides dissolved separately in DMSO and DI water were carried out.

1 mg of both ADH304_1 and ADH304_2 peptide were soluble in 100 µL of DI water and DMSO. But, for both peptides, the addition of peptide dissolved in DI water to GFP-SMCC resulted immediate precipitation in the reaction mixture at all tested peptide concentrations. However, the conjugation reactions with ADH304_1 and ADH304_2 peptides dissolved in DMSO yielded stable formulations with 25 µL of 10 µg/mL peptide after 24 hrs incubation at room temperature while 75 µL of 10 µg/mL of peptides resulted aggregation of GFP-SMCC-peptide formulations (Figure 3.1).

Thus, for optimal conjugation 25µL of 10 µg/mL of ADH304 peptides dissolved in DMSO was used for the preparation of GFP-peptide formulations.
3.1.2 SDS–PAGE Analysis of GFP–Peptide Conjugates

1. Molecular Ladder
2. GFP
3. GFP–SMCC
4. GFP–SMCC–ADH304_1
5. GFP–SMCC–ADH304_2

Figure 3.1: Appearance of GFP–peptide formulations prepared from peptides dissolved in (a) DI water and (b) DMSO

Figure 3.2: 12% SDS-PAGE gel of GFP–peptide conjugates

The shifts in the positions of the bands of GFP–peptide formulations relative to GFP indicate the successful conjugations of peptides to GFP through SMCC. The molecular weight of the ADH304_1 peptide is 1.83kDa and that of ADH304_2 is 1.72kDa. Based on the molecular weights of the peptides and the magnitude of the gel shifts, it can be estimated
that GFP-SMCC-ADH304_1 has 2-3 peptides per GFP while 3-4 peptides per GFP in GFP-SMCC-ADH304_2 conjugate. However, the data also indicate that the GFP-peptide conjugates are heterogeneous. These GFP-peptide conjugations are stable for weeks at 4°C storage condition.

3.2 Mutagenesis of GFP Loops

Only the site directed mutagenesis of GFP loops performed at 63°C annealing temperature was used for this study. DNA sequencing provided the evidence that the site directed mutagenesis of all these GFP loop mutants is successful. However, only Loop J-RGD and Loop H-RGD yielded fluorescently active mutant GFP while the rest of the mutants yielded proteins but not green fluorescent. The spectral properties such as absorbance maximum, excitation and emission wavelengths and extinction coefficients of Loop J_RGD and Loop H-RGD mutants are as same as native GFP.

3.3 Quantum Yields of GFP-Peptide Conjugates and GFP Loop Mutants

Quantum yields of GFP and the conjugates were calculated relative to fluorescein and given below in Table 3.1. Quantum yield of GFP agrees with the literature and it indicates that both loop mutations and chemical attachment of peptides to GFP has significantly changed the quantum yield. The differences in quantum yield of mutant GFP and peptide conjugated GFP may be due to the changes in the rates of radioactive decays involve in the fluorescence mechanism upon mutations and chemical modifications of GFP.

Table 3.1: Relative quantum yields of GFP test samples

<table>
<thead>
<tr>
<th>Sample</th>
<th>Relative quantum yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluoroscein (0.1M NaOH)</td>
<td>0.95(^2)</td>
</tr>
<tr>
<td>GFP</td>
<td>0.64821</td>
</tr>
<tr>
<td>GFP-SMCC</td>
<td>0.60656</td>
</tr>
<tr>
<td>GFP-SMCC-ADH304_1</td>
<td>0.7189</td>
</tr>
<tr>
<td>GFP-SMCC-ADH304_2</td>
<td>0.5924</td>
</tr>
<tr>
<td>H-RGD mutant</td>
<td>0.7136</td>
</tr>
<tr>
<td>J-RGD mutant</td>
<td>0.6163</td>
</tr>
</tbody>
</table>
3.4 Cell Targeting Studies

3.4.1 Targeted Cell Delivery of GFP-Peptide Conjugates

Targeting specificities of ADH304 peptides attached to GFP was studied in A375, HeLa and MCF-7 cell lines. Both A375 and HeLa cells exhibit over expression of N–cadherin receptors which is the target of ADH304 peptides that contains the HAV recognition motif and therefore a significantly higher uptake of GFP-ADH304 peptide conjugations is expected. On the other hands, MCF-7 under express N-cadherin receptors and thus any significant difference in terms of cellular uptake is not expected between native GFP and GFP-ADH304 peptide conjugations.

The fluorescence intensity of cells treated with GFP test samples after 48hrs incubation at 37 °C followed by washing twice with DPBS to remove residual test samples were measured. In order to compensate for variations in quantum yields, the fluorescence was normalized to corresponding quantum yield. For more accurate comparison of, fluorescence data were also normalized to cell viability after sample treatment to minimize the errors due to variations in cell density and any cytotoxicity from sample, which would make data more comparable to well-to-well, plate-to-plate and experiment-to-experiment.

The fluorescence observed from sample treated cells is due to either internalized GFP-peptide formulation and/or attach to cell membrane. Thus the differential effect in fluorescence response from cells treated with GFP test samples were used to compare the targeting abilities of GFP- peptide formulations relative to unmodified GFP.

In both HeLa and A375 cells, the cells treated with GFP-SMCC-ADH304_2 conjugate indicate significantly higher fluorescence response, which is nearly 3000 fluorescence counts compare to unmodified GFP (Figure 3.3 and 3.4). GFP by itself does not have any specific cell targeting properties. Therefore, this observation indicates a differential effect in terms of cell targeting of GFP upon introducing ADH304_2 peptide as a targeting agent. However, the cells treated with GFP –SMCC- ADH304_1 conjugate exhibit as same level of fluorescence response as unmodified GFP. This indicates the lack of targeting ability of ADH304_1 peptide in GFP-SMCC-peptide formulations and this could be due to changes
in ADH304_1 peptide that disfavor the interaction with N-cadherin receptors. Additionally, the presence of an additional asparagine molecule in ADH304_2 peptide in compare to ADH304_1 may have contributed to enhanced interaction with N-cadherin receptors.

Figure 3.3 Targeted delivery of GFP-peptide conjugates in A375 cells (N=4)
In both HeLa and A375 cell lines, GFP-SMCC, even without any cell targeting moiety showed relatively higher fluorescence response in cell uptake experiments (Figure 3.3 and 3.4). The nonspecific uptake of Sulfo-SMCC labeled targeting systems is a major disadvantage of using SMCC as a heterobifunctional linker to synthesize GFP-peptide conjugates.
conjugates. The literature provides evidences for this observation and it suggests that the unreacted sites on Sulfo-SMCC are responsible for the nonspecific cell adhesion and uptake.\(^3\)

**Figure 3.5** Targeted delivery of GFP-peptide conjugates in MCF-7 cells (N=2)
Unlike in A375 and HeLa cells, a significant difference in fluorescence response of GFP test samples cannot be observed in MCF-7 cells (Figure 3.5). The fluorescence intensities of GFP-peptide conjugates, unmodified GFP and GFP-SMCC lie within 400 fluorescence counts. The absence of specific targeting ability of ADH304 peptide labeled GFP compared to unmodified GFP in MCF-7 cells that under express N-cadherin suggests that ADH304 peptides exhibits specific targeting and involves cellular uptake in the cells with N-cadherin receptors that only present in cancer cells (Figure 3.5).

3.4.2 Targeted Cell Delivery of GFP Loop Mutants

The GFP loop mutants, namely H-RGD and J-RGD contains only one RGD motif per GFP molecule. RGD recognition motif shows selective affinity towards α5β1 and αvβ3 integrins that over expressed in cancer cells. A375 and HeLa cell lines do over express α5β1 and αvβ3 integrins.

The cell targeting results of the GFP loop mutants in A375 and HeLa cells are very similar (Figure 3.6-3.7). The normalized fluorescence responses of sample treated cells are dose dependant. However a significant difference in fluorescence response that indicates the specific cell targeting cannot be observed between unmodified GFP and GFP loop mutants. This could be due to several reasons. It might be due to changes in favorable conformations of RGD motif on GFP that reduces the proper interaction with integrins and thereby specific cell internalization. These loop mutants have only one RGD motif, which might not be sufficient for effective interaction with integrins that leads to cellular uptake. Additionally in both cell lines, cells treated with Loop J-RGD mutant showed a higher fluorescence response over Loop H-RGD mutant. This might be due to the location of RGD sequence on GFP, which allows more space for the interaction with integrins in Loop J-RGD over Loop H-RGD.
Figure 3.6 Targeted delivery of GFP mutants in A375 cells (N=4)
Figure 3.7 Targeted delivery of GFP mutants in HeLa cells (N=2)
3.5 Molecular Dynamic Analysis of GFP Loop Mutants

MD trajectories of each mutant were analyzed for the fluctuations of bond distances, bond angles and dihedral angles of interests. The minimum, maximum and average for each bond distance and bond angle fluctuation were calculated and are summarized in table 1. Only the bond distances and angles that directly involve in the chromophore maturation reactions and interactions with Arg96, Thr62 and Glu 222 have tabulated. BD, BA and DA are abbreviations for bond distance, bond angle and dihedral angle respectively.

**Table 3.2:** Bond distances and bond angles of GFP loop mutants from MD trajectories

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Min/°</th>
<th>Max/°</th>
<th>Average/°</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFP</td>
<td>111.28</td>
<td>129.31</td>
<td>119.96</td>
</tr>
<tr>
<td>H_hav</td>
<td>110.45</td>
<td>129.04</td>
<td>120.25</td>
</tr>
<tr>
<td>H_rgd</td>
<td>111.60</td>
<td>129.97</td>
<td>120.34</td>
</tr>
<tr>
<td>J_hav</td>
<td>109.68</td>
<td>129.65</td>
<td>120.27</td>
</tr>
<tr>
<td>J_rgd</td>
<td>110.40</td>
<td>129.35</td>
<td>119.81</td>
</tr>
<tr>
<td>M_rgd</td>
<td>11.618</td>
<td>131.66</td>
<td>120.92</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Min/Å</th>
<th>Max/Å</th>
<th>Average/Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFP</td>
<td>2.78</td>
<td>3.46</td>
<td>3.03</td>
</tr>
<tr>
<td>H_hav</td>
<td>2.87</td>
<td>3.64</td>
<td>3.11</td>
</tr>
<tr>
<td>H_rgd</td>
<td>2.84</td>
<td>3.58</td>
<td>3.17</td>
</tr>
<tr>
<td>J_hav</td>
<td>2.87</td>
<td>3.52</td>
<td>3.14</td>
</tr>
<tr>
<td>J_rgd</td>
<td>2.82</td>
<td>4.63</td>
<td>3.59</td>
</tr>
<tr>
<td>M_rgd</td>
<td>2.89</td>
<td>3.59</td>
<td>3.18</td>
</tr>
</tbody>
</table>
Table 3.2 continued

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Min/Å</th>
<th>Max/Å</th>
<th>Average/Å</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BD_Ser65C_Tyr66N</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Important in dehydration followed by oxidation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GFP</td>
<td>1.27</td>
<td>1.45</td>
<td>1.36</td>
</tr>
<tr>
<td>H_hav</td>
<td>1.25</td>
<td>1.45</td>
<td>1.35</td>
</tr>
<tr>
<td>H_rgd</td>
<td>1.24</td>
<td>1.46</td>
<td>1.36</td>
</tr>
<tr>
<td>J_hav</td>
<td>1.23</td>
<td>1.45</td>
<td>1.35</td>
</tr>
<tr>
<td>J_rgd</td>
<td>1.24</td>
<td>1.46</td>
<td>1.35</td>
</tr>
<tr>
<td>M_rgd</td>
<td>1.25</td>
<td>1.46</td>
<td>1.35</td>
</tr>
<tr>
<td><strong>BD_Tyr66O_Arg96HH21</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Important in H bonding</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GFP</td>
<td>4.6</td>
<td>7.96</td>
<td>5.97</td>
</tr>
<tr>
<td>H_hav</td>
<td>3.86</td>
<td>6.38</td>
<td>5.11</td>
</tr>
<tr>
<td>H_rgd</td>
<td>4.79</td>
<td>7.60</td>
<td>6.17</td>
</tr>
<tr>
<td>J_hav</td>
<td>3.85</td>
<td>7.71</td>
<td>5.49</td>
</tr>
<tr>
<td>J_rgd</td>
<td>2.65</td>
<td>7.07</td>
<td>4.25</td>
</tr>
<tr>
<td>M_rgd</td>
<td>4.40</td>
<td>6.59</td>
<td>5.60</td>
</tr>
<tr>
<td><strong>BD_Tyr66O_Arg96HH22</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Important in H bonding</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GFP</td>
<td>4.6</td>
<td>7.96</td>
<td>5.89</td>
</tr>
<tr>
<td>H_hav</td>
<td>3.42</td>
<td>6.38</td>
<td>4.82</td>
</tr>
<tr>
<td>H_rgd</td>
<td>5.32</td>
<td>7.68</td>
<td>6.21</td>
</tr>
<tr>
<td>J_hav</td>
<td>5.18</td>
<td>8.23</td>
<td>6.79</td>
</tr>
<tr>
<td>J_rgd</td>
<td>4.20</td>
<td>8.21</td>
<td>5.68</td>
</tr>
<tr>
<td>M_rgd</td>
<td>4.46</td>
<td>7.49</td>
<td>6.04</td>
</tr>
</tbody>
</table>
Table 3.2 continued

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Min/Å</th>
<th>Max/Å</th>
<th>Average/Å</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BD_Thr62O_Arg96HH1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Important in H bonding</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GFP</td>
<td>1.57</td>
<td>4.38</td>
<td>2.72</td>
</tr>
<tr>
<td>H_hav</td>
<td>1.46</td>
<td>2.29</td>
<td>1.71</td>
</tr>
<tr>
<td>H_rgd</td>
<td>1.91</td>
<td>6.73</td>
<td>3.28</td>
</tr>
<tr>
<td>J_hav</td>
<td>1.56</td>
<td>6.03</td>
<td>4.82</td>
</tr>
<tr>
<td>J_rgd</td>
<td>2.63</td>
<td>5.56</td>
<td>3.95</td>
</tr>
<tr>
<td>M_rgd</td>
<td>1.59</td>
<td>4.23</td>
<td>2.64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Min/Å</th>
<th>Max/Å</th>
<th>Average/Å</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BD_Ser65OH_Glu222O</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Important in H bonding</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GFP</td>
<td>1.4</td>
<td>7.96</td>
<td>3.66</td>
</tr>
<tr>
<td>H_hav</td>
<td>1.49</td>
<td>2.36</td>
<td>1.76</td>
</tr>
<tr>
<td>H_rgd</td>
<td>1.42</td>
<td>2.51</td>
<td>1.73</td>
</tr>
<tr>
<td>J_hav</td>
<td>1.48</td>
<td>3.02</td>
<td>1.81</td>
</tr>
<tr>
<td>J_rgd</td>
<td>1.45</td>
<td>4.91</td>
<td>1.89</td>
</tr>
<tr>
<td>M_rgd</td>
<td>2.41</td>
<td>3.48</td>
<td>2.69</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Min/Å</th>
<th>Max/Å</th>
<th>Average/Å</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BD_Tyr66C_Arg96N</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Important in electrostatic interactions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GFP</td>
<td>2.92</td>
<td>4.71</td>
<td>3.49</td>
</tr>
<tr>
<td>H_hav</td>
<td>3.32</td>
<td>5.52</td>
<td>4.20</td>
</tr>
<tr>
<td>H_rgd</td>
<td>4.34</td>
<td>6.46</td>
<td>5.19</td>
</tr>
<tr>
<td>J_hav</td>
<td>4.03</td>
<td>6.42</td>
<td>5.09</td>
</tr>
<tr>
<td>J_rgd</td>
<td>4.02</td>
<td>6.26</td>
<td>4.93</td>
</tr>
<tr>
<td>M_rgd</td>
<td>3.79</td>
<td>6.49</td>
<td>4.89</td>
</tr>
</tbody>
</table>
Table 3.2 continued

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Min/°</th>
<th>Max/°</th>
<th>Average/°</th>
</tr>
</thead>
<tbody>
<tr>
<td>DA_Ser65C_Ser65CO_Tyr66N_Tyr66C (ω)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GFP</td>
<td>-179.99</td>
<td>179.99</td>
<td>-110.61</td>
</tr>
<tr>
<td>H_hav</td>
<td>-179.99</td>
<td>179.99</td>
<td>89.91</td>
</tr>
<tr>
<td>H_rgd</td>
<td>-179.99</td>
<td>179.99</td>
<td>-93.12</td>
</tr>
<tr>
<td>J_hav</td>
<td>-179.99</td>
<td>179.99</td>
<td>81.03</td>
</tr>
<tr>
<td>J_rgd</td>
<td>-179.99</td>
<td>179.99</td>
<td>130.78</td>
</tr>
<tr>
<td>M_rgd</td>
<td>-179.99</td>
<td>179.99</td>
<td>-25.75</td>
</tr>
<tr>
<td>DA_Ser65C_Tyr66N_Tyr66C_Tyr66CO (ϕ)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GFP</td>
<td>30.27</td>
<td>86.66</td>
<td>60.97</td>
</tr>
<tr>
<td>H_hav</td>
<td>42.65</td>
<td>98.84</td>
<td>71.25</td>
</tr>
<tr>
<td>H_rgd</td>
<td>33.26</td>
<td>93.44</td>
<td>66.54</td>
</tr>
<tr>
<td>J_hav</td>
<td>34.83</td>
<td>93.43</td>
<td>69.16</td>
</tr>
<tr>
<td>J_rgd</td>
<td>-14.92</td>
<td>93.00</td>
<td>59.43</td>
</tr>
<tr>
<td>M_rgd</td>
<td>44.89</td>
<td>109.30</td>
<td>73.30</td>
</tr>
<tr>
<td>DA_Tyr66N_Tyr66C_Tyr66C_O_Gly67N (ψ)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GFP</td>
<td>-68.06</td>
<td>57.92</td>
<td>-56.08</td>
</tr>
<tr>
<td>H_hav</td>
<td>-84.19</td>
<td>8.71</td>
<td>-33.31</td>
</tr>
<tr>
<td>H_rgd</td>
<td>-96.89</td>
<td>0.56</td>
<td>-57.87</td>
</tr>
<tr>
<td>J_hav</td>
<td>-87.77</td>
<td>21.81</td>
<td>-32.94</td>
</tr>
<tr>
<td>J_rgd</td>
<td>-82.14</td>
<td>137.21</td>
<td>46.37</td>
</tr>
<tr>
<td>M_rgd</td>
<td>-87.95</td>
<td>16.17</td>
<td>-52.43</td>
</tr>
</tbody>
</table>
All the bond distances and bond angles that have calculated from MD trajectories are important only in evaluating the impact of mutations on ring cyclization and dehydration followed by double bond formation reactions of the chromophore biosynthetic pathway.

According to the literature, interaction between the chromophore-forming region and the amino acid side chains in the close proximity (Thr62, Arg96 and Glu222) via hydrogen bonding and electrostatic attraction has a significant impact on the stability of chromophore. MD trajectory analysis reveals no significant difference in BD_Ser65OH_Glu222O and BD_Tyr66O_Arg96HH21/22 between wild type GFP and none of the RGD and HAV GFP loop mutants. Based on that, it can be concluded that the hydrogen bonding interactions between Arg96 and Glu222 and the chromophore region have not been affected by the three point mutations on GFP loops. Hydrogen bonding to Arg96 and Glu222 indirectly involve in facilitating the favorable structural rearrangements for ring cyclization and dehydration reaction of the biosynthetic pathway respectively. Since these interactions have not changed due to three point GFP loop mutations, it may also possible to conclude that the rates of ring cyclization and water elimination steps have not significantly affected from above mutations and there by the chromophore formation.

However, no solid conclusions can be derived from BD_Tyr66C_Arg96N data, which represent the electrostatic catalytic function of Arg96. The above bond distance for fluorescently active Loop J-RGD and Loop H-RGD mutants as well as non-fluorescent GFP loop mutants is larger than that of GFP.

BD_Ser65C_Gly67N reveals about the feasibility of nuclciophilic attack of amide N of Gly67 to carbonyl carbon of Ser65 while BD_Ser65C_Tyr66N and BA_Ser65O_Ser65C_Tyr66N are important parameters in dehydration followed by double bond formation. Average bond distances and angles of these for all HAV and RGD mutants are very similar to that of GFP. Therefore these bond distances and angles also cannot be employed to differentiate the GFP loop mutants that are capable of fluorescence emission from non-fluorescent mutants.
The ability to form the “tight turn” conformation within the chromophore forming region of the immature GFP also plays a critical role in the chromophore maturation. This tight turn confirmation has 2.9 Å distance between i carbonyl carbon to i+2 amide nitrogen with $\phi=60\pm30^\circ$ and $\psi=30\pm15^\circ$ and in GFP, it refers to DA_Ser65C_Ser65CO_Tyr66N_Tyr66C ($\omega$), DA_Ser65CO_Tyr66N_Tyr66C_Tyr66CO ($\phi$) and DA_Tyr66N_Tyr66C_Tyr66CO_Gly67N ($\psi$) respectively. This “tight turn” is also restrict the conformational space of the chromophore-forming region that keeps the amino acids in place for autocatalytic cyclization. The values of above dihedral angles for all the mutants are within the required range for the “tight turn” formation. This also evidenced that the loop mutations do not directly affect the conformational requirements for the chromophore formation.

Even considering all these calculations for bond distances and angles, any computational prediction for the rate or probability of a mature chromophore formation in GFP loop mutations cannot be established.

As stated earlier, the calculations carried out here provide information about the first two steps of the chromophore biosynthesis. But any important information about the rate determining oxidation step can be extracted from MD trajectory analysis.

Thus, it might be the rate of oxidation step that greatly influence in chromophore maturation. It might be some other factors apart from the interactions with proximal amino acids responsible for slower rate of maturation. Evidences on GFP mutations that take much longer time for maturation compare to GFP can be found in literature. Change in the polarity/charge of the overall structure or in a given region due to the mutation may have an effect on the network of water molecules around GFP, which is believed to participate in proton exchange throughout the chromophore biosynthetic pathway, which ultimately affect the chromophore maturation process. Additionally if more parameters can be computationally evaluated which account for the rate of oxidation, more information can be gathered in predicting and establishing the conditions that leads to mature chromophore formation.
Based on the calculations and regardless of the oxidation step of the biosynthesis, it is reasonable to predict that all GFP loop mutations to yield fluorescent active mutants. A possibility of not yielding fluorescent mutants may be due to the trapping of immature GFP with unfavorable conformations or folding inside inclusion bodies\textsuperscript{12-14}. Therefore, as an experimental solution to that, necessary procedures to recover proteins from inclusion bodies can be carried out prior to regular purification protocol.

One other possibility for immature chromophore may be due to the slower rate of maturation of these GFP loop mutants. Thus as another experimental approach, the fluorescence of the purified non fluorescent GFP loop mutants can be monitored over a longer time frame to gather evidences for slower maturation.

If none of these experimental approaches provide adequate evidences for the presence and development of immature chromophore in GFP loop mutants, a strong conclusion can be made on alteration of the rate of oxidation reaction of the chromophore biosynthesis due to mutations. More computational approaches to evaluate this can be fruitful in obtaining computational evidences for the maturation of GFP chromophore.

Finally, it can be clearly concluded that any accurate prediction about the fluorescence activity of GFP loop mutants cannot be derived solely considering the bond distances, bond angles and dihedral angles fluctuation data gathered from the molecular dynamic simulation calculations.
References:

(1) Patterson GH; Knobel SM; Sharif, W. D. Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. **1997**.


Chapter 4: Conclusions
Conclusions

GFP- peptide conjugations with both ADH304_1 and ADH304_2 peptides were soluble and stable for weeks in 4 °C. SDS-PAGE provides evidences that GFP-peptide conjugates are heterogeneous and average of 2-4 peptides have chemically conjugated per GFP molecule. A major disadvantage of using SMCC as a heterobifunctional linker in preparing GFP-peptide conjugates is the SMCC cross linking which decreases SMCC available for the reaction with GFP and peptide. This results heterogeneous mixture of GFP-peptide conjugates and inability to control the number of peptides chemically linked per GFP between separate preparations.

Cell targeting studies of GFP-ADH304 peptide formulations in A375 and HeLa cell lines showed a differential effect in terms of cell targeting upon introducing ADH304_2 peptides to GFP. The presence of an additional asparagine molecule in ADH304_2 peptide may have contributed to the enhance cell targeting response compare to GFP conjugated to ADH304_1. Additionally, similar targeting response of both GFP-ADH304 formulations and GFP in N-cadherin lack MCF-7 cell line is also an important evidence for the selective targeting of GFP-ADH304 in cancer cells expressing N-cadherin.

Genetic modification of GFP was designed as an alternative approach to overcome the problems associated with the SMCC chemistry that was employed in making GFP-peptide conjugates. GFP loop mutations allow controlling the position and the number of cancer cell targeting moieties on GFP. DNA sequencing provided the evidences that HAV and RGD mutations in all the GFP loops of interest namely, loop H, loop J and loop M are successful. However only the loop J-RGD and loop H-RGD mutants were green fluorescent while the rest of the mutations are non-fluorescent. The loop H-RGD and loop J-RGD loop mutations have changed the quantum yield but they still own almost the same spectral properties (excitation and emission wavelengths) as unmodified GFP.

Cell targeting study of loop J-RGD and loop H-RGD in A375 and HeLa did not show a significant targeting ability over unmodified GFP. This could be due to the lack of targeting moieties per GFP minimizing the probability of the GFP loop mutants to interact with integrins. Improved cancer cell targeting might be observed upon introducing double or triple
RGD loop mutations per GFP and thereby increasing the total number of targeting moieties on GFP.

In order to understand the effect of GFP loop mutations in the chromophore maturation, GFP loop mutants were analyzed using VMD and NAMD. Molecular dynamic analysis of these mutants does not provide any information to significantly differentiate any bond parameters that would affect the chromophore formation and maturation.

According to the research described in this thesis, GFP-ADH304_2 peptide formulations provide promise in developing as a targeting probe and multiple loop mutations of GFP could also have a potential in designing GFP derived specific cancer cell targeting probe.