

University of Arkansas, Fayetteville

ScholarWorks@UARK

Graduate Theses and Dissertations

7-2015

Designing FRET Assays to Study Electrostatic Interactions Pertaining to the Binding of Intrinsically Disordered Proteins

Ashley Ann Howard

University of Arkansas, Fayetteville

Follow this and additional works at: <https://scholarworks.uark.edu/etd>



Part of the [Biochemistry Commons](#), [Biophysics Commons](#), and the [Organic Chemistry Commons](#)

Citation

Howard, A. A. (2015). Designing FRET Assays to Study Electrostatic Interactions Pertaining to the Binding of Intrinsically Disordered Proteins. *Graduate Theses and Dissertations* Retrieved from <https://scholarworks.uark.edu/etd/1243>

This Dissertation is brought to you for free and open access by ScholarWorks@UARK. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of ScholarWorks@UARK. For more information, please contact uarepos@uark.edu.

Designing FRET Assays to Study Electrostatic Interactions Pertaining to the Binding of
Intrinsically Disordered Peptides

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Chemistry

by

Ashley Howard
University of Arkansas Monticello
Bachelor of Science in Chemistry, 2009

July 2015
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Dr. Colin D. Heyes
Dissertation Director

Dr. Suresh K. Thallapuranam
Committee Member

Dr. Joshua Sakon
Committee Member

Dr. Frank Millett
Committee Member

Dr. Paul Adams
Committee Member

Abstract

Fibroblast growth factor receptor plays a major role in several biological processes. Without FGFR, a human cannot live. FGFR is involved in cell differentiation and wound healing. Of course, if FGFR signaling becomes unregulated, it causes severe distress in the body. Several cancers are contributed to high signaling levels, as well as developmental conditions like rickets and Kallmann's syndrome. FGFR is thought to undergo an auto-inhibition (or self-regulatory) process in order to try to facilitate regulation. The exact method of this inhibition is currently unknown, but is proposed to involve the unstructured acid box region of FGFR. We developed a simple model system in order to further investigate current models of inhibition that FGFR may undergo. By using our model system, which contains two 15-mer homopolypeptides of polyE and polyK that mimic the acid box region and its binding site respectively, we were able to use a combination of ITC, CD, NMR, and FRET to show that one model from the literature contains flaws. We are able to characterize the binding of our polypeptide system under varying ionic conditions and pH. This model system also provides a platform to better understand general principles of charge-charge interactions in proteins, which are often characterized by FRET. One of the important findings from this study is that 15-mers of polyE and polyK bind in a parallel arrangement. One of the hurdles in applying FRET to such systems is determining the role that the attached FRET dyes play in the charge-charge interactions. Our model system allowed us to use the preference of the charged polypeptides to bind in a parallel arrangement to determine the size, charge and structural effects of the attached FRET dyes on how peptides bind under electrostatic interaction conditions and to quantify how the attached fluorescent dyes are quenched by both the charged amino acids as well as by the FRET acceptor.

Acknowledgements

There are a few people that really helped me make it through this process without going insane. First and foremost, I want to thank my advisor Colin Heyes for keeping me on track and helping me figure out what I want to do after all of this is over. Next are my committee members, thank you for being on my committee and guiding me through this journey. Thank you to my lab mates for several interesting and mostly off-topic conversations, there was never a dull moment when we were all together. The members of Dr. Kumar's lab were very helpful on several occasions, so thank you all for that. I would also like to thank Chris Mazzanti for his time helping me with everything from research questions to personal advice. I would like to give a special thanks to my husband, Michael. Without him, this would not have been possible. He pushed me when I needed it and never let me quit on myself.

Contents

I. Introduction 1

1.1 Protein Makeup and Structure.....	1
1.2 Intrinsically Disordered Proteins.....	2
1.3 Electrostatic Interactions	3
1.4 Fibroblast Growth Factor Receptor.....	4
1.5 Fibroblast Growth Factors.....	5
1.6 Fibroblast Growth Factor Receptor Signaling and Regulation	7
1.7 Charged Homopolypeptides as a Model for Charge Complementarity	10
1.8 Fluorescence and Quenching	11
1.9 Fluorescence Resonance Energy Transfer	12
1.10 Fluorescent Dyes	15
1. 11 References	16

II. Thermodynamic and Structural Characterization of Interactions between Oppositely-Charged Short-Chain Homopolypeptides 22

2.1. Abstract	22
2.2 Introduction	23
2.3. Materials and Methods.....	25
2.3.1. <i>Peptide Synthesis and Labeling</i>	25
2.3.2. <i>Circular Dichroism</i>	26
2.3.3. <i>Isothermal Titration Calorimetry</i>	26
2.3.4. <i>Nuclear Magnetic Resonance (NMR) Spectroscopy</i>	27
2.3.5. <i>Fluorescence Resonance Energy Transfer (FRET) Spectroscopy</i>	27
2.4. Results and Discussion.....	28
2.4.1 <i>Peptide Selection, Labeling and Characterization</i>	28
2.4.2. <i>Isothermal titration calorimetry (ITC) of polyE/polyK Binding</i>	31
2.4.3. <i>polyE/polyK Binding Studied by FRET</i>	33
2.4.4. <i>Far-UV CD spectroscopy of equimolar mixtures of polyE and polyK</i>	37
2.4.5. <i>TOCSY and NOESY NMR Spectroscopy</i>	40
2.4.6. <i>Combining Structural and Thermodynamic Data</i>	42
2.5. Conclusions	42

2.6. Acknowledgements	44
2.7. References	45
III. Determining the Effects of Fluorescent Labels on Electrostatic Interactions using a Polypeptide Template	50
3.1. Abstract	50
3.2. Introduction	51
3.3. Materials and Methods	53
3.3.1. <i>Peptide Labeling</i>	53
3.3.2. <i>Fluorescence Resonance Energy Transfer (FRET) Spectroscopy</i>	53
3.4. Results	54
3.5. Discussion	63
3.6. Summary and Conclusions.....	66
3.7 References	67
IV. Insights on the Quenching of Some Fluorescent Dyes Commonly Used in Biolabeling by Charged Amino Acid Residues	70
4.1. Abstract	70
4.2. Introduction	71
4.3. Materials and Methods	72
4.3.1. <i>Peptide Labeling</i>	72
4.3.2. <i>Fluorescence Spectroscopy</i>	73
4.4. Results and Discussion.....	74
4.5. Conclusions	81
4.6. References	83
V. Purification, Labeling, and FRET with FGFR and FGF	85
5.1. Abstract	85
5.2. Introduction	86
5.3. Materials and Methods	87
5.3.1. <i>Acid Box Peptide</i>	87
5.3.2. <i>PCR, Expression, Purification of D2 and D2D3 domains of FGFR2 and FGF1</i>	87
5.3.3. <i>Fluorescent Labeling of acid box, D2, D2/D3 and FGF</i>	91
5.3.5. <i>Isothermal Calorimetry</i>	92

5.3.6. <i>Fluorescence Resonance Energy Transfer (FRET) Spectroscopy</i>	92
5.4. Results	93
5.5. Discussion and Conclusion	105
VI. Conclusion	109
6.1 References	114

I. Introduction

1.1 Protein Makeup and Structure

All biological lifeforms consist of proteins¹. Even though the number, size, and shape of proteins are limitless, all are made up from the same 20 amino acids. Each amino acid has its own characteristics which contribute to the protein's structure and/or function. The various side chains on the amino acids determine how the protein will fold based upon the charge and hydrophobicity of the functional group of the amino acid². Out of the 20 total amino acids, there are 4 that have a charge at physical pH, 7.2. This small percentage of charged amino acids represents the need to regulate the total charge of a protein. Too many charges would hinder a protein from folding to a compact structure³, which many need to perform their function. The overall structure of a protein determines its function and where in the cell the protein will be located. Not all proteins fold the same way, with some proteins having a loose, flexible structure while others are rigid. It has long been thought that the structure of a protein determined its function⁴ until relatively recently, even though this is still true for most proteins. Most proteins function by binding to other proteins or small molecules inside of the cell. This binding can occur by a variety of means including, but not limited to, oppositely charged (electrostatic) interactions, hydrogen bonding and hydrophobic forces. When a ligand binds to a protein, the protein will often slightly change its structure to enable the ligand to bind more tightly or to facilitate another ligand to bind⁵. Often, the structure of the ligand will match that of the protein that it is designed to bind – the lock-and-key mechanism⁶. If binding occurs through opposite charges (electrostatic interactions), the ligand and protein will have surface-charge complementarity, which means that the protein will have a pocket of negatively charged amino acids and the ligand will contain an area of positively charged amino acids, or vice versa.

1.2 Intrinsically Disordered Proteins

Many of the molecular targets for protein binding are not shape determined, as some proteins are intrinsically disordered. This type of disorder can mean that the protein does not have specific secondary structural elements such as alpha helices or beta sheets that often contribute to protein shape. It can even mean that the protein has no shape whatsoever. Intrinsically disordered proteins (IDP's) make up roughly 33% of the known proteins today in eukaryotic systems⁷. Though once thought to have no function⁸, several very important roles have been found to be performed by IDP's⁹. There is evidence for their participation in cell signaling and transcription factors¹⁰, as well as having been connected to several important diseases such as Alzheimer's and cancer¹¹. Intrinsically disordered proteins or peptides generally contain a high number of charged amino acids¹², and most of their function arises from the polarity of charges of their amino acids as well as their sequence/density. There are different binding/folding models for IDP's that are thought to drive the binding process. One such model is the "fly-casting" mechanism. This particular mechanism uses the fact that the IDP is initially unfolded and elongated, containing a high degree of entropy. With the high density of charged residues, the protein is able to use long-range electrostatic forces to "feel" around for its binding partner. Once bound weakly through these electrostatic interactions, the protein can begin to form a more specific structure and bind more tightly to its partner¹³. This model accounts for the observation of fast binding kinetics and takes into account the extreme flexibility of unfolded proteins¹⁴. A second model recognizes that some IDP's never fold, even after binding. Termed the "fuzzy complex" mechanism, the IDP stays unfolded after binding to an ordered protein¹⁵. A typical example of an IDP binding to a structured protein is the p53 protein binding to MDM2. p53 is a mostly disordered protein that is involved in tumor suppression. MDM2 is a globular

protein that negatively regulates p53. p53 binds to MDM2 through one of the disordered regions on p53, mainly through hydrophobic and aromatic residues. Once bound, this disordered region then adopts a helical coil structure¹⁶. This protein, like many others, was initially thought to not have a function for the disordered region, since there are ordered regions that were known to be involved in DNA binding. Some proteins bind through charge-charge interactions only though. An example of this includes the Sac10b family that binds to DNA¹⁷. This family is made up of small proteins that bind to DNA with no sequence specificity. These proteins protect the DNA from digestion by Dnase I. This family of proteins is highly conserved and contains 3 pairs of positively charged residues that bind to DNA. When these amino acids are changed to uncharged residues, the protein cannot bind to DNA.

1.3 Electrostatic Interactions

One of the primary binding mechanisms of proteins is through charge-charge interactions from electrostatic forces. Coulomb's law tells us how strong electrostatic interactions are based upon distance by the formula: $F = k q_1 q_2 / r^2$ where F is the force, q_1 and q_2 are electric charges, r is distance, and k is proportionality constant. This information is very helpful when determining how neighboring charges affect each other when a protein is folding or when a ligand is trying to bind since like charges repel each other while opposite charges attract. A large percentage of proteins have charged amino acids in their binding pockets that facilitate the binding of their partner. The lock-and-key binding model¹⁸ is the most common binding model for folded proteins that have complementary shapes and charges. However, when there is no structural complementarity between partners, the charges themselves must be the leading factor for getting proteins to bind to ligands or other molecules. The electrostatic interactions between the protein and ligand are the only aspect that can contribute to binding other than hydrogen bonds which are substantially weaker. Coulomb's law tells us that opposite charges attract each other, and at

what distances these charges can be recognized by other molecules but not much information has been acquired about how the surface-charge complementarity¹⁹ and the binding affinity of protein targets relate. It is also unknown how the peptides with highly charged residues orient themselves and how do the proteins find their partner and not just bind to everything charged they come across.

1.4 Fibroblast Growth Factor Receptor

One particular interesting example for a protein with a functionally-important, as well as potentially being cancer-related, intrinsically disordered region is the fibroblast growth factor receptor (FGFR). FGFR's are known to play a role in wound healing, cell growth, and new blood vessel growth (angiogenesis) among many other biological processes²⁰. On the other hand, overexpression of FGFR's can lead to cancer with the most common types being colorectal^{21, 22}, gastric^{23, 24}, bladder^{25, 26}, and pancreatic cancers^{27, 28}. Mutations of the FGFR protein can also cause cancer by changing the activation of pathways that stimulate tumor growth²⁹. There are 4 known FGFRs, each consisting of 3 immunoglobulin-like domains (D1, D2 and D3), an acid box (AB) region linking D1 and D2, a transmembrane domain that follows D3, followed by a cytoplasmic tyrosine kinase domain³⁰. The AB region is an intrinsically disordered region, and is thus very flexible. It gets its name from the high percentage of acidic residues found there; nearly 50% of the amino acids are acidic, and are negatively charged at physiological pH³¹. The amino acid sequence of the acid box of FGFR2 is shown in figure 1.1.

VTDAISSGDEDDTDGA

Figure 1.1. The amino acid sequence of the FGFR2 acid box. The charged residues are highlighted in red. The high percentage of charged residues makes this protein region one of interest. The charged residues are potentially involved in binding to a ligand binding site on a domain in the protein.

The two main binding partners for FGFR are FGF and heparin; their binding sites are on the D2 domain with a small portion of the FGF interacting with the D3 domain. There are 2 exons that code for the last half of the D3 domain in most FGFR's that can cause two different isoforms of the receptor³². One variant is expressed in epithelial tissue while the other is found in mesenchymal tissue³³. This alternative splicing changes the specificity of the receptor to different FGF proteins that have been identified³⁴. Most published structures of FGFRs consist only of the D2 and D3 regions in a dimeric complex with its FGF and heparin partners (figure 1.2). Currently there are no crystal structures for FGFR that have been determined in the absence of ligand³⁵. Not having an uncomplexed structure available makes it very difficult to get mechanistic information about the protein.

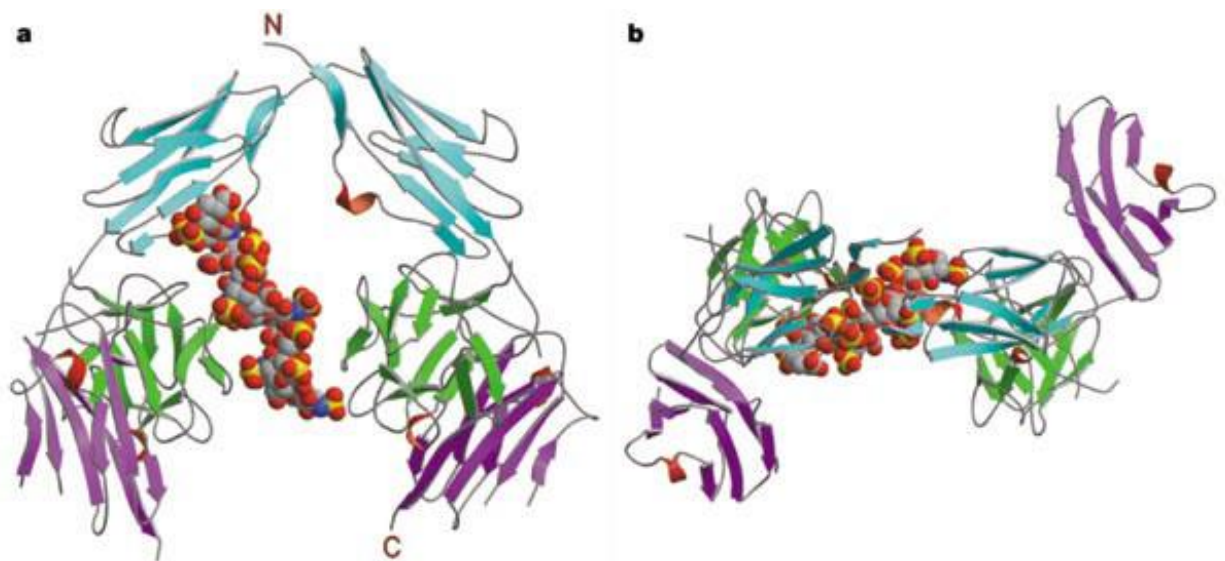


Figure 1.2. FGFR/heparin complex. A) View from the side of the complex. FGFR2 domains D2 and D3 are cyan and magenta, respectively, and FGF1 is green. The heparin molecule is shown in red. B) View from the top down³⁶.

1.5 Fibroblast Growth Factors

Fibroblast growth factors (FGF's) were the first known angiogenic factors^{37, 38}. There are 22 known human FGF's³⁹, with these proteins being found in both vertebrates and invertebrates. The FGF genes are widespread throughout the human genome and are located on several

different chromosomes. FGFs can be divided into subgroups with each subgroup having a high percentage of sequence similarity and developmental properties. The range of molecular weights for human FGFs is between 17 and 34kDa. There is a highly conserved core region that contains 28 homologous amino acids, of which 6 are identical²⁰. Ten of the core amino acids interact directly with FGFR⁴⁰. The mammalian FGFs are expressed in almost all tissues but in different patterns and at different times⁴¹. FGFs are glycoproteins that are found in the extracellular matrix and on the cell surface. In order for them to be used in signaling, they first have to be released into the extracellular matrix to be exposed to the cell surface, where they bind to heparin and the membrane-bound receptor⁴². Once the FGF is bound to the receptor, dimerization occurs. At this point it is thought that heparin sulfate glycosaminoglycan binds to a specific site on the dimer, although there is still debate on exactly when heparin binds. Some researchers propose that the heparin binds to FGF before it binds to the receptor, thereby stabilizing the protein against degradation⁴³. In any case, the formation of the dimer causes a conformational change that activates the tyrosine kinase domain which results in it becoming phosphorylated. This phosphorylated residue acts as a docking area when the FGFR interacts with other proteins, which starts a signaling cascade consisting of multiple pathways. FGF receptors are also involved in negative feedback regulations, though not much about the mechanism is known, except that the receptor is degraded. Fibroblast growth factors undergo auto-regulation to balance cell growth and death to maintain a healthy state. When the regulation breaks down, FGF's may become involved in many different diseases such as cancer^{42, 44}, rickets^{45, 46}, and Kallmann syndrome³⁹. Overexpression of some FGFR's has been shown to produce tumors but it is not clear whether it is because the higher amounts of FGF receptor binds more ligands to result in higher signaling or if the tumors respond directly to higher amounts of FGF⁴².

1.6 Fibroblast Growth Factor Receptor Signaling and Regulation

FGFR is involved in several different signaling pathways including RAS/MAPK and PI-3 kinase/Akt (figure 1.3). The signaling for this receptor starts when FGF and heparin or heparin sulfate bind to the extracellular domains, D2 and D3, on FGFR causing receptor dimerization and activation that causes the tyrosine kinase domain to auto-phosphorylate. Signaling proteins such as STAT1, Gab1, and FRS2 α are then phosphorylated so they can begin their roles in cell differentiation, cell survival, and cell proliferation²⁰. Regulation for FGFR involves pericellular matrix heparin sulfate. This small molecule controls the FGFR signaling on many levels⁴⁷ and when it's not available, the complex cannot be activated. Regulation can also come from the different isoforms of the receptor. Natural splicing can change the D3 domain but it may also remove the first domain (D1) and AB region. Both of these changes can cause the binding affinity of FGF for the receptor to change by either increasing or decreasing the affinity of the FGF for the receptor. For example, when the D1 and AB regions are removed from the receptor, binding affinity increases between FGFR and FGF but in some isoforms where the D3 domain is changed, the binding affinity of FGF1 to FGFR is greatly reduced⁴⁸. FGFR is also thought to undergo self-regulation by auto inhibition⁴⁹. There are two competing theories surrounding the auto inhibition or self-regulation of FGFR. In both cases, inhibition results in a reduction in the receptors' affinity for FGF⁴⁸. One method proposes that FGFR undergoes auto inhibition by the flexible acid box region swinging down and binding to the heparin binding site on the D2 domain (figure 1.4 A) ⁴⁹⁻⁵¹. The rationale for this hypothesis is that the acid box region is highly negatively charged while the heparin binding site in D2 is highly positively charged. Kalinina et al have recently supported this hypothesis using NMR to determine which residues are perturbed when the AB region of the protein is deleted versus when it is present⁵¹. Olsen et al. have shown using Surface Plasmon Resonance that the presence of the D1 domain decreases the binding

affinity of heparin to an FGFR⁴⁹, which was subsequently reproduced and extended by Kalinina et al⁵¹. The competing hypothesis suggests that the acid box region binds to FGF instead of FGFR, thereby blocking the FGF from binding to the FGFR's D2 domain (figure 1.4 B)⁵².

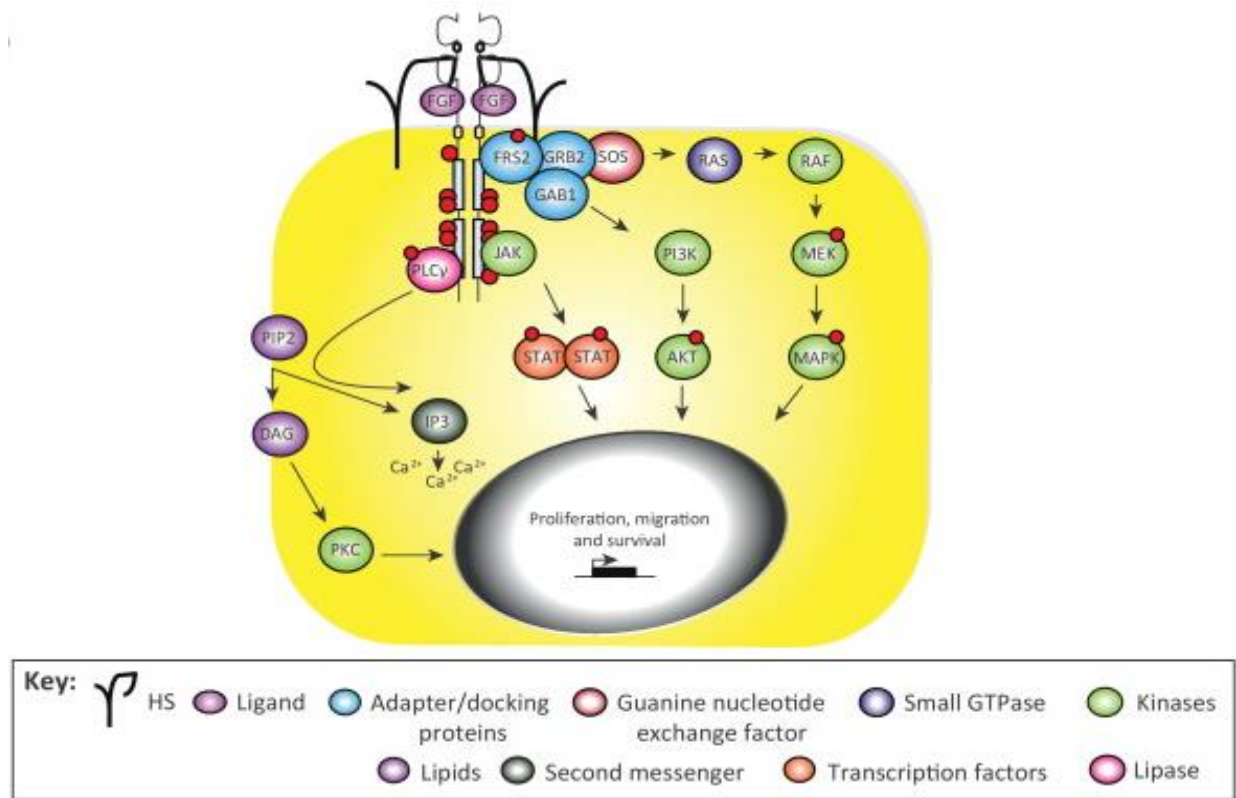


Figure 1.3. Signaling pathway for FGF. Four of the pathways that are set in motion by the phosphorylation of FGFR⁵³. The yellow box represents the inside of the cell. The top left part of the figure shows FGFR bound to heparin sulfate, FGF, as well as 3 signaling proteins: FRS2, JAK, and PLCγ. The various pathways are shown that occur once the dimerization and phosphorylation of FGFR occur.

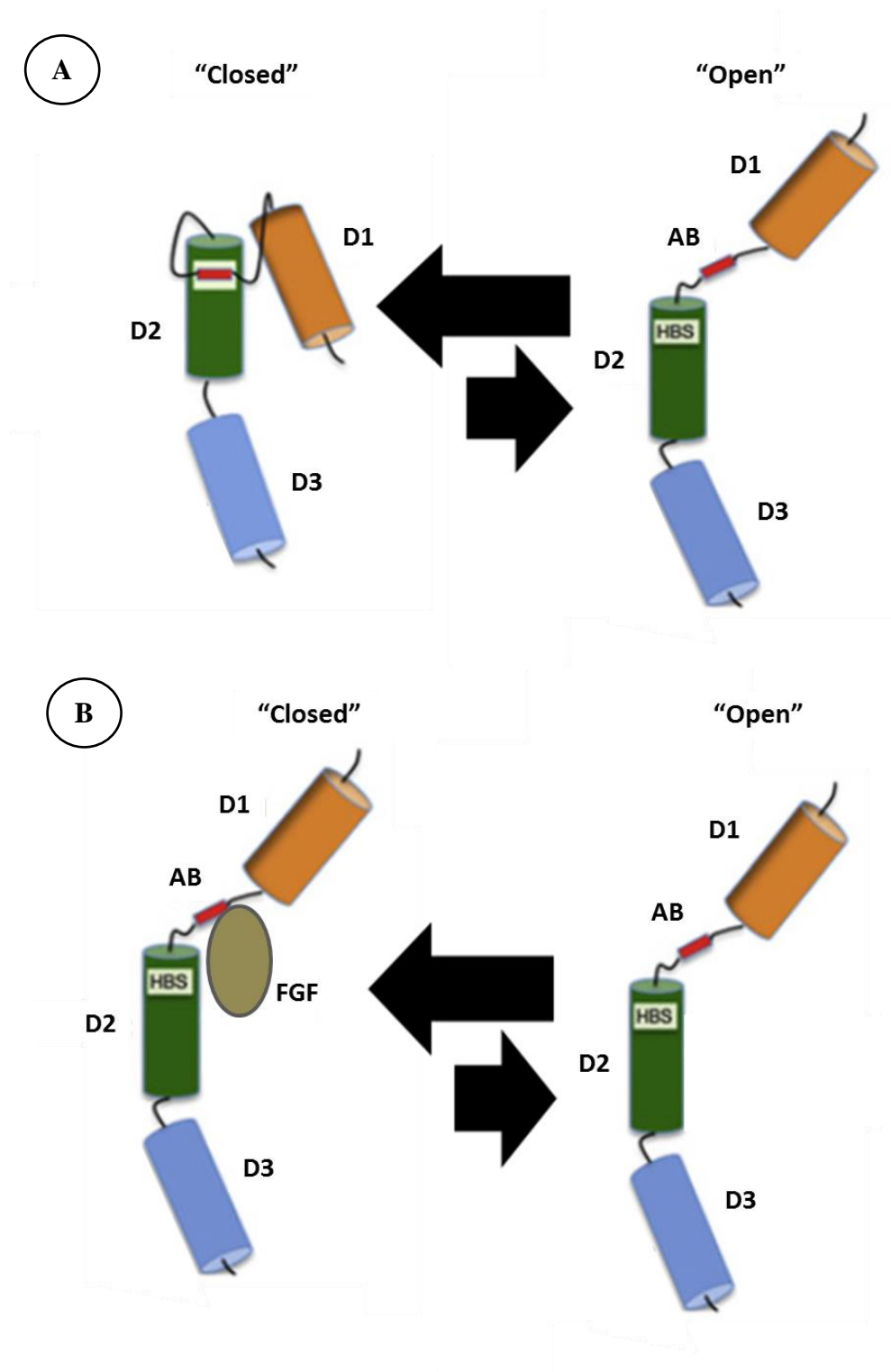


Figure 1.4. Proposed model for auto inhibition of FGFR. A) The acid box binds to the heparin binding site of the D2 domain. B) FGF binds to the acid box and blocks the FGF binding site on the D2 domain. Modified^{51, 52}.

1.7 Charged Homopolypeptides as a Model for Charge Complementarity

Homopolypeptides are constructed of a single type of amino acid repeated any number of times. Homopolypeptides were used in the late 1960's and 1970's to study the stability of side chain conformations,⁵⁴⁻⁵⁶ often by analyzing the Raman spectroscopic peaks caused by secondary structure formation^{57, 58}. At that time, it was well known that long chains of glutamic acid would form a random coil at pH 7 but could be forced into either an alpha helix by lowering the pH to 4, or a beta sheet by adding long chains of poly lysine. Poly lysine itself is a random coil at pH 7 and is alpha helical at pH 11⁵⁹. Long chains of glutamic acid with long chains of lysine were later used as films that helped to preserve the structure of DNA^{60, 61}. The size of the polypeptide chains used in previous studies have always been long (<100 amino acids). Recently shorter, more relevant homopeptides have been the focus of study. For example, by using a peptide of 34 arginine residues, the effect of the net charge per residue on the globule to coil transition was examined by replacing various arginine residues with other amino acids⁶². This study concluded that a higher average charge per residue did increase the amount of coiled structure of peptides. However, shorter, more biologically relevant versions of homopolypeptides – in terms of sizes relevant to protein regions of interest, such as the acid box of FGFR described above – have not been extensively studied. The binding between these unstructured, oppositely charged peptides arises from electrostatic interactions. Homopolypeptides constructed of glutamic acids and lysine are excellent subjects for binding interaction studies as they have charge complementarity and are largely disordered at neutral pH⁵⁵, they are also a good model for our possible acid box/heparin binding site study since the major amino acid in the acid box is glutamic acid and the possible binding site on the D2 domain consists of a positive region of residues. We will be able to get an idea if the acid box/heparin binding site theory is practical based upon the amount of binding in our model system.

1.8 Fluorescence and Quenching

Fluorescence is the emission of light from the excited state of a molecule (figure 1.5).

This emission occurs in usable timeframes and can be measured by simple machines. Quenching occurs when the intensity of fluorescence is decreased or diminished altogether⁶³. Quenching can occur through wanted as well as unwanted means. Collisional quenching occurs after collisions of the fluorophore with another molecule that causes the excited electron to return to its ground state. Static quenching can occur if the fluorophore and another molecule form a stable complex⁶⁴.

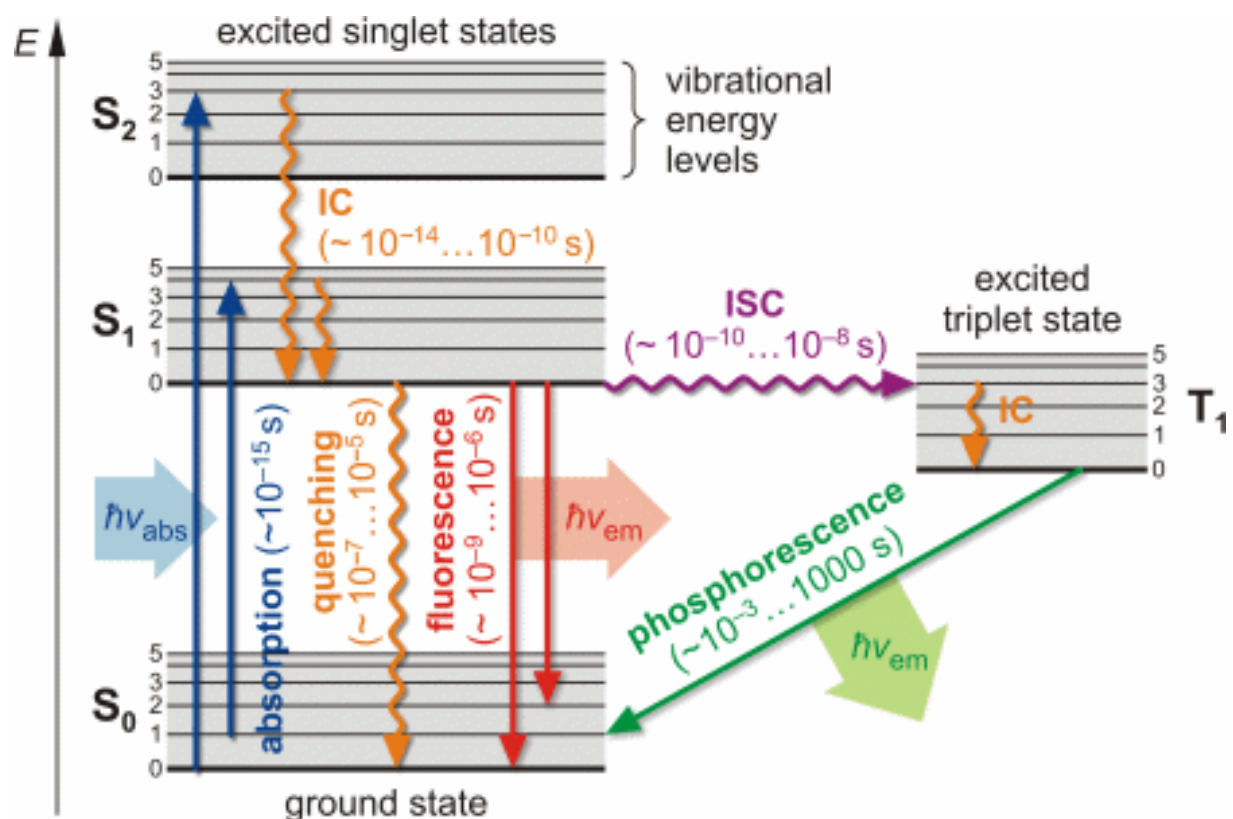


Figure 1.5. Jablonski diagram for fluorescence. When an electron in the ground state is excited, it moves into an excited singlet state (absorption). This electron can then return to the ground state through different methods, one of them being fluorescence⁶⁵.

1.9 Fluorescence Resonance Energy Transfer

One use of fluorescence is in the technique of fluorescence resonance energy transfer (FRET). In FRET, there is one fluorescent molecule (termed the donor) and another molecule (termed the acceptor) that can accept the energy of the first molecule. FRET occurs by the donor molecule being excited by an external source and becoming fluorescent. At distances unique to the donor molecule but always less than 100 Å, the donor can transfer its energy to the acceptor. This transfer of energy causes the donors' fluorescent signal to decrease. If the acceptor molecule is also fluorescent, then the acceptor will fluoresce once the energy has been transferred (figure 1.6). It should be noted that the fluorescence of the donor is not transferred to the acceptor, only the energy is transferred. The donors' fluorescence will decrease when the energy is transferred according to the amount of energy transferred. The energy transfer process is dependent upon several factors including distance between dye molecules (figure 1.7) and the overlap of the donor's emission spectrum with the acceptors absorption spectrum (figure 1.8). The amount of overlap of the two spectra will determine how much energy is transferred at a given distance. The formula for the efficiency of energy transfer (E) is $E = R_0^6 / (R_0^6 + r^6)$ where r is the distance between the donor and acceptor molecules and R_0 is the Forster distance. The Forster distance is the distance at which the energy transfer is 50%, figure 1.7. This factor is determined by the quantum yield (Q_0) of the donor in the absence of the acceptor, the orientation of the dipole (κ^2) of the molecules, as well as the refractive index (n) of the medium in which the molecules are in and the integral of the spectral overlap (J). The equation for Forster distance is $R_0^6 = (9000 Q_0 (\ln 10) \kappa^2 J) / (128 \pi^5 n^4 N_A)$ where N_A is Avogadro's number. The Forster distance is comparable to that of a biological macromolecules, since it is in the range of 20-100 Å. Because of the strong distance dependence of FRET, this technique gives researchers the ability to measure intra- and inter-molecular changes in proteins, DNA molecules, ion channels, and

many other biological molecules/systems that have structural or interactional changes. Some of the uses for FRET include determination of structural characteristics of intrinsically disordered proteins⁶⁶, monitoring kinesin motor proteins⁶⁷, and measuring the electrostatic repulsion between the domains of calmodulin⁶⁸. This is such a useful tool for analysis because it allows for varying temperatures and ionic concentrations and, importantly, does not use large amounts of sample. This technique is extremely sensitive and facilitates the use of very low concentrations of sample by employing fluorescent dyes as labels on the amino acid residues of proteins or peptides. By careful placement of the fluorescent dyes, the relative orientation of binding partners can be elucidated.

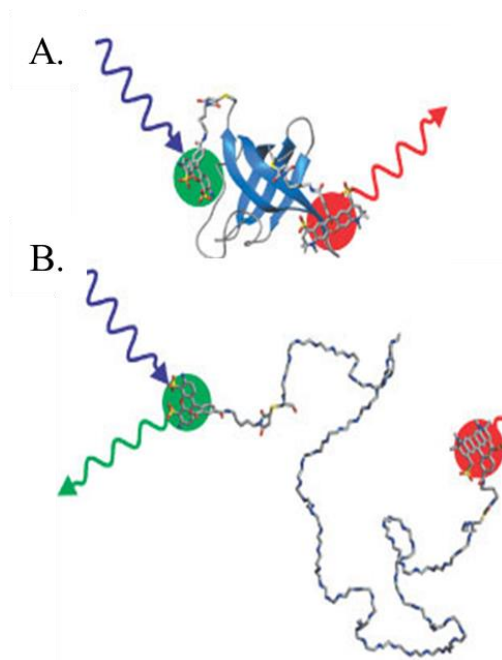


Figure 1.6. FRET occurs through energy transfer. A) The donor (green) and acceptor (red) molecules are close together so the donor's energy is transferred to the acceptor. The donor is no longer fluorescent but the acceptor is. **B)** The donor and acceptor are too far apart for energy transfer to occur. The donor is fluorescent but the acceptor is not. Adapted from⁶⁹.

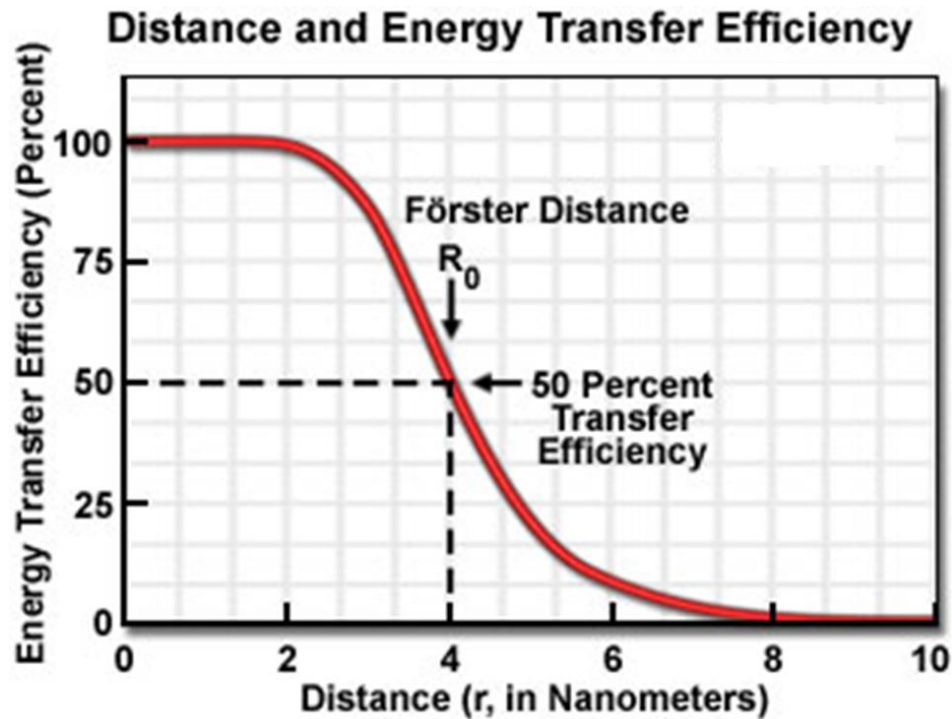


Figure 1.7. Distance dependence of energy transfer. The energy transfer efficiency varies for every donor/acceptor pair but this chart shows how the distance and the Förster distance goes together⁷⁰.

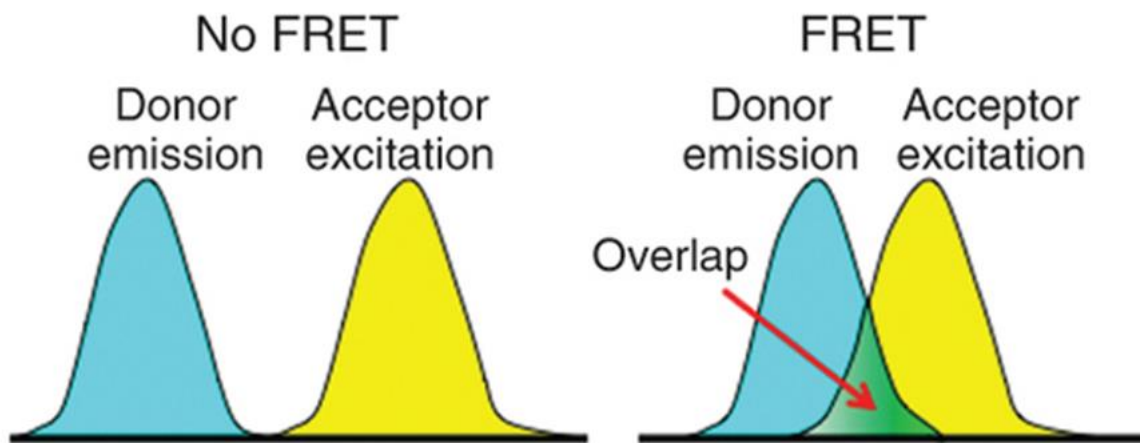


Figure 1.8. Spectral overlap for donor and acceptor needed for FRET. On the left side we can see that the spectra do not overlap and FRET does not occur as it does on the right side of the figure⁷¹.

1.10 Fluorescent Dyes

Fluorescent dyes come in a myriad of structures, sizes, and charges. There are several properties of fluorescent dyes that are important when determining which ones to use, with the most commonly considered ones being the wavelength of light that will be used and the quantum yield of emission. The environment that the dye will be used in is also paramount. There are some dyes that are designed to be used *in vivo*, namely near infrared fluorochromophores, while others are strictly for *in vitro* like rhodamines. When measuring electrostatic interactions, the charges of the dyes themselves should be taken into consideration, as well as the size of the dye, especially if they will be placed close to the binding site. One property of fluorescent dyes that is not normally but can be beneficial is fluorescence quenching. As we have seen, FRET is one method of controlled quenching that we can use for measurements. Other types of quenching like collisional and static quenching by environmental factors or other molecules in the solution are not always welcome. When quenching occurs by a method other than the one being studied then it is a hindrance. Most commercially available fluorescent dyes are stable and are thought to resist minor environmental changes. Though fluorescent dyes have been used for many years, the effects of the dyes on the measured values collected in FRET studies have not been studied explicitly. We chose a system of two highly charged homopolypeptides that bind through electrostatic interactions to measure the effects of the charges of the peptides on the fluorescence of the dyes as well as the charges of the dyes on the binding of the peptides. By using multiple dyes with various charges, we were able to determine the effects the peptide charges have on the dyes, such as decrease in quantum yield, with no other possible contributions from factors outside of our system like other proteins or molecules in solution.

1. 11 References

1. Horton, H. R., Moran, L., Scrimgeour, K. G., Perry, M., and Rawn, J. D. (2006) *Principles of Biochemistry*, Pearson Prentice Hall, New Jersey.
2. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains, *Science* *181*, 223-230.
3. Müller-Späh, S., Soranno, A., Hirschfeld, V., Hofmann, H., Rügger, S., Reymond, L., Nettels, D., and Schuler, B. (2010) Charge interactions can dominate the dimensions of intrinsically disordered proteins, *PNAS* *107*, 14609-14614.
4. Petrey, D., Fischer, M., and Honig, B. (2009) Structural relationships among proteins with different global topologies and their implications for function annotation strategies, *Proceedings of the National Academy of Sciences* *106*, 17377-17382.
5. Cox, M., and Neelson, D. (2005) Lehninger principles of biochemistry, *Biochemistry and Molecular Biology Education* *33*, 74-75.
6. Fischer, E. (1898) Meaning of stereo chemistry for the physiology. [machine translation], *Ztschr. physiol. Ch.* *26*, 60-87.
7. Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., and Jones, D. T. (2004) Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life, *Journal of Molecular Biology* *337*, 635-645.
8. Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C., and Obradovic, Z. (2001) Intrinsically disordered protein, *Journal of Molecular Graphics and Modelling* *19*, 26-59.
9. Wright, P. E., and Dyson, H. J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm, *Journal of Molecular Biology* *293*, 321-331.
10. Garza, A. S., Ahmad, N., and Kumar, R. (2009) Role of intrinsically disordered protein regions/domains in transcriptional regulation, *Life Sciences* *84*, 189-193.
11. Uversky, V. N., Oldfield, C. J., and Dunker, A. K. (2008) Intrinsically Disordered Proteins in Human Diseases: Introducing the D2 Concept, *Annual Review of Biophysics* *37*, 215-246.
12. Uversky, V., and Dunker, A. (2010) Understanding protein non-folding, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* *1804*, 1231 - 1264.

13. Shoemaker, B. A., Portman, J. J., and Wolynes, P. G. (2000) Speeding molecular recognition by using the folding funnel: The fly-casting mechanism, *Proceedings of the National Academy of Sciences* 97, 8868-8873.
14. Huang, Y., and Liu, Z. (2009) Kinetic Advantage of Intrinsically Disordered Proteins in Coupled Folding-Binding Process: A Critical Assessment of the "Fly-Casting" Mechanism, *Journal of Molecular Biology*, 1143-1159.
15. Mittag, T., Marsh, J., Grishaev, A., Orlicky, S., Lin, H., Sicheri, F., Tyers, M., and Forman-Kay, J. D. (2010) Structure/Function Implications in a Dynamic Complex of the Intrinsically Disordered Sic1 with the Cdc4 Subunit of an SCF Ubiquitin Ligase, *Structure* 18, 494-506.
16. Dawson, R., Müller, L., Dehner, A., Klein, C., Kessler, H., and Buchner, J. (2003) The N-terminal Domain of p53 is Natively Unfolded, *Journal of Molecular Biology* 332, 1131-1141.
17. Liu, Y.-F., Zhang, N., Liu, X., Wang, X., Wang, Z.-X., Chen, Y., Yao, H.-W., Ge, M., and Pan, X.-M. (2012) Molecular Mechanism Underlying the Interaction of Typical Sac10b Family Proteins with DNA, *PLoS ONE* 7, e34986.
18. Tymoczko, J. L., Berg, J. M., and Stryer, L. (2013) Kinetics and Regulation, In *Biochemistry: A short course* 2 ed., pp 100-101.
19. McCoy, A. J., Chandana Epa, V., and Colman, P. M. (1997) Electrostatic complementarity at protein/protein interfaces, *Journal of Molecular Biology* 268, 570-584.
20. Eswarakumar, V. P., Lax, I., and Schlessinger, J. (2005) Cellular signaling by fibroblast growth factor receptors, *Cytokine & Growth Factor Reviews* 16, 139-149.
21. Yao, T.-J., Zhu, J.-H., Peng, D.-F., Cui, Z., Zhang, C., and Lu, P.-h. (2015) AZD-4547 exerts potent cytostatic and cytotoxic activities against fibroblast growth factor receptor (FGFR)-expressing colorectal cancer cells, *Tumor Biol.*, Ahead of Print.
22. Abdel-Rahman, O. (2015) Targeting FGF receptors in colorectal cancer: from bench side to bed side, *Future Oncology* 11, 1373-1379.
23. Grygielewicz, P., Dymek, B., Bujak, A., Gunerka, P., Stanczak, A., Lamparska-Przybysz, M., Wieczorek, M., Dzwonek, K., and Zdzalik, D. (2014) Epithelial-mesenchymal transition confers resistance to selective FGFR inhibitors in SNU-16 gastric cancer cells, *Gastric Cancer*, Ahead of Print.
24. Xie, L., Su, X., Zhang, L., Yin, X., Tang, L., Zhang, X., Xu, Y., Gao, Z., Liu, K., Zhou, M., Gao, B., Shen, D., Zhang, L., Ji, J., Gavine, P. R., Zhang, J., Kilgour, E., Zhang, X., and Ji, Q. (2013) FGFR2 Gene Amplification in Gastric Cancer Predicts Sensitivity to the Selective FGFR Inhibitor AZD4547, *Clin. Cancer Res.* 19, 2572-2583.

25. Cheng, T., Roth, B., Choi, W., Black, P. C., Dinney, C., and McConkey, D. J. (2013) Fibroblast growth factor receptors-1 and -3 play distinct roles in the regulation of bladder cancer growth and metastasis: implications for therapeutic targeting, *PLoS One* 8, e57284.
26. di, M. E., Tomlinson, D. C., and Knowles, M. A. (2012) A Decade of FGF Receptor Research in Bladder Cancer: Past, Present, and Future Challenges, *Adv Urol* 2012, 429213.
27. Coleman, S. J., Chioni, A.-M., Ghallab, M., Anderson, R. K., Lemoine, N. R., Kocher, H. M., and Grose, R. P. (2014) Nuclear translocation of FGFR1 and FGF2 in pancreatic stellate cells facilitates pancreatic cancer cell invasion, *EMBO Mol. Med.* 6, 467-481.
28. Matsuda, Y., Yoshimura, H., Suzuki, T., Uchida, E., Naito, Z., and Ishiwata, T. (2014) Inhibition of fibroblast growth factor receptor 2 attenuates proliferation and invasion of pancreatic cancer, *Cancer Sci.* 105, 1212-1219.
29. Ho, H. K., Yeo, A. H. L., Kang, T. S., and Chua, B. T. (2014) Current strategies for inhibiting FGFR activities in clinical applications: opportunities, challenges and toxicological considerations, *Drug Discovery Today* 19, 51-62.
30. Li, S., Bock, E., and Berezin, V. (2010) Neuritogenic and Neuroprotective Properties of Peptide Agonists of the Fibroblast Growth Factor Receptor, *International Journal of Molecular Sciences* 11, 2291-2305.
31. NextProt. FGFR2, Fibroblast growth factor receptor 2 [EC 2.7.10.1].
32. Wuechner, C., Nordqvist, A.-C. S., Winterpacht, A., Zabel, B., and Schalling, M. (1996) Development expression of splicing variants of fibroblast growth factor receptor 3 (FGFR3) in mouse, *Int. J. Dev. Biol.* 40, 1185-1188.
33. McEwen, D. G., and Ornitz, D. M. (1997) Determination of fibroblast growth factor receptor expression in mouse, rat and human samples using a single primer pair, *BioTechniques* 22, 1068, 1070.
34. Givol, D. (2009) Molecular and cellular biology of FGF signalling, *Oxford monographs on medical genetics* 54, 449-460.
35. Belov, A. A., and Mohammadi, M. (2013) Molecular Mechanisms of Fibroblast Growth Factor Signaling in Physiology and Pathology, *Cold Spring Harbor Perspectives in Biology* 5.
36. Pellegrini, L., Burke, D. F., von Delft, F., Mulloy, B., and Blundell, T. L. (2000) Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin, *Nature* 407, 1029-1034.
37. Acevedo, V., Ittmann, M., and Spencer, D. (2009) Paths of FGFR-driven tumorigenesis, *Cell Cycle* 8, 580-588.

38. Armelin, H. A. (1973) Pituitary Extracts and Steroid Hormones in the Control of 3T3 Cell Growth, *Proceedings of the National Academy of Sciences* 70, 2702-2706.
39. Hu, Y., and Bouloux, P. M. (2010) Novel insights in FGFR1 regulation: lessons from Kallmann syndrome, *Trends in Endocrinology and Metabolism* 21, 385-393.
40. Plotnikov, A. N., Hubbard, S. R., Schlessinger, J., and Mohammadi, M. (2000) Crystal Structures of Two FGF-FGFR Complexes Reveal the Determinants of Ligand-Receptor Specificity, *Cell* 101, 413-424.
41. Ornitz, D., and Itoh, N. (2001) Fibroblast growth factors, *Genome Biology* 2, 1-12.
42. Turner, N. (2010) Fibroblast growth factor signalling: from development to cancer, *Nature Reviews Cancer* 10, 116-129.
43. Beenken, A., and Mohammadi, M. (2009) The FGF family: biology, pathophysiology and therapy, *Nature Reviews. Drug Discovery* 8, 235-253.
44. Wesche, J., Haglund, K., and Haugsten, E. M. (2011) Fibroblast growth factors and their receptors in cancer, *Biochem Journal*, 199-213.
45. White, K. E., Jonsson, K. B., Carn, G., Hampson, G., Spector, T. D., Mannstadt, M., Lorenz-Depiereux, B., Miyauchi, A., Yang, I. M., Ljunggren, Ö., Meitinger, T., Strom, T. M., Jüppner, H., and Econs, M. J. (2001) The Autosomal Dominant Hypophosphatemic Rickets (ADHR) Gene Is a Secreted Polypeptide Overexpressed by Tumors that Cause Phosphate Wasting, *The Journal of Clinical Endocrinology & Metabolism* 86, 497-500.
46. Fukumoto, S., and Yamashita, T. (2001) Fibroblast Growth Factor (FGF)-23 and Hypophosphatemic Rickets/Osteomalacia, *Endocrine Journal* 48, 603-610.
47. McKeehan, W. L., Wang, F., and Luo, Y. (2010) Chapter 38 - The Fibroblast Growth Factor (FGF) Signaling Complex, In *Handbook of Cell Signaling (Second Edition)* (Bradshaw, R. A., and Dennis, E. A., Eds.), pp 253-259, Academic Press, San Diego.
48. Wang, F., Kan, M., Yan, G., Xu, J., and McKeehan, W. L. (1995) Alternately Spliced NH2-terminal Immunoglobulin-like Loop I in the Ectodomain of the Fibroblast Growth Factor (FGF) Receptor 1 Lowers Affinity for both Heparin and FGF-1, *Journal of Biological Chemistry* 270, 10231-10235.
49. Olsen, S. K., Ibrahimi, O. A., Raucci, A., Zhang, F., Eliseenkova, A. V., Yayon, A., Basilico, C., Lindardt, R. J., Schlessinger, J., and Mohammadi, M. (2004) Insights into the molecular basis for fibroblast growth factor receptors auto inhibition and ligand binding promiscuity, *PNAS* 101, 935-940.
50. Schlessinger, J. (2003) Autoinhibition Control, *Science* 300, 750-752.

51. Kalinina, J., Dutta, K., Ilghari, D., Beenken, A., Goetz, R., Eliseenkova, Anna V., Cowburn, D., and Mohammadi, M. (2012) The Alternatively Spliced Acid Box Region Plays a Key Role in FGF Receptor Autoinhibition, *Structure* 20, 77-88.
52. Rutherford, L., Rajalingam, D., and Kumar, T. K. S. Understanding the molecular mechanism underlying the auto inhibition of the fibroblast growth factor signaling.
53. Carter, E. P., Fearon, A. E., and Grose, R. P. (2015) Careless talk costs lives: fibroblast growth factor receptor signalling and the consequences of pathway malfunction, *Trends in Cell Biology* 25, 221-233.
54. Townend, R., Kumosiniki, T. F., and Timasheff, S. N. (1966) The circular dichroism of the beta structure of Poly-L-Lysine, *Biochemical and Biophysical Research Communications* 23.
55. Tiffany, M. L., and Krimm, S. (1968) New chain conformations of poly(glutamic acid) and polylysine, *Biopolymers* 6, 1379-1382.
56. De Santis, P. (1968) Calculation of the most stable conformation of polypeptide side chains, *Quad. Ric. Sci. No.* 47, 65-67.
57. Fasman, G. D., Itoh, K., Liu, C. S., and Lord, R. C. (1978) Laser-excited raman spectroscopy of biomolecules. XII. Thermally induced conformational changes in poly(L-glutamic acid), *Biopolymers* 17, 1729-1746.
58. Song, S., and Asher, S. A. (1989) UV resonance Raman studies of peptide conformation in poly(L-lysine), poly(L-glutamic acid), and model complexes: the basis for protein secondary structure determinations, *Journal of the American Chemical Society* 111, 4295-4305.
59. Hammes, G. G., and Schullery, S. E. (1968) Structure of macromolecular aggregates. I. Aggregation-induced conformational changes in polypeptides, *Biochemistry* 7, 3882-3887.
60. Boulmedais, F., Schwinté, P., Gergely, C., Voegel, J. C., and Schaaf, P. (2002) Secondary Structure of Polypeptide Multilayer Films: An Example of Locally Ordered Polyelectrolyte Multilayers, *Langmuir* 18, 4523-4525.
61. Montrel, M. M., Sukhorukov, G. B., Petrov, A. I., Shabarchina, L. I., and Sukhorukov, B. I. (1997) Spectroscopic study of thin multilayer films of the complexes of nucleic acids with cationic amphiphiles and polycations: their possible use as sensor elements, *Sensors and Actuators B: Chemical* 42, 225-231.
62. Mao, A. H., Crick, S. L., Vitalis, A., Chicoine, C. L., and Pappu, R. V. (2010) Net charge per residue modulates conformational ensembles of intrinsically disordered proteins, *PNAS* 107, 8183-8188.

63. Lakowicz, J. R. (2006) *Principles of fluorescence spectroscopy*, 3rd ed., Springer, New York.
64. Jameson, D. M. (2014) *Introduction to Fluorescence*, CRC Press.
65. Pawlizak, S. (2009), University of Leipzig.
66. Haas, E. (2012) Ensemble FRET Methods in Studies of Intrinsically Disordered Proteins, In *Intrinsically Disordered Protein Analysis* (Uversky, V. N., and Dunker, A. K., Eds.), pp 467-498, Humana Press.
67. Prevo, B., and Peterman, E. J. G. (2014) Forster resonance energy transfer and kinesin motor proteins, *Chemical Society Reviews* 43, 1144-1155.
68. Hellstrand, E. a. K. S. a. S. C. F. a. S. S. a. T. E. a. K. A. a. K. B. a. L. S. a. Å. K. S. (2013) Förster resonance energy transfer studies of calmodulin produced by native protein ligation reveal inter-domain electrostatic repulsion, *FEBS Journal* 280, 2675--2687.
69. Schuler, B., Lipman, E. A., and Eaton, W. A. (2002) Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy, *Nature* 419, 743-747.
70. Herman, B., Frohlich, V., Lakowicz, J., Fellers, T., and Davidson, M. (2012) Fluorescence Resonance Energy Transfer Microscopy.
71. Broussard, J. A., Rappaz, B., Webb, D. J., and Brown, C. M. (2013) Fluorescence resonance energy transfer microscopy as demonstrated by measuring the activation of the serine/threonine kinase Akt, *Nat. Protocols* 8, 265-281.

II. Thermodynamic and Structural Characterization of Interactions between Oppositely-Charged Short-Chain Homopolypeptides

Ashley Howard, Nicole Webb, T.K.S. Kumar and Colin D. Heyes

2.1. Abstract

Interactions between charged amino acid residues have long been recognized as important in a variety of intra- and inter- protein interactions. The prevalence of disordered regions in proteins on the order of ~15 amino acids with high acidic or basic amino acid content has drawn recent attention due to a general shifting in the paradigm of the structure-function relationship in proteins and the connection of intrinsic disordered proteins (IDPs) with various diseases. In this study, we undertake a systematic characterization of the thermodynamic and structural properties underlying 15-mers of poly-K and poly-E under various conditions of pH, ionic strength and solvent hydrophobicity. We found that, after interaction, oppositely-charged homopolypeptides of this length remain as random coils at neutral pH in aqueous solutions with the interaction being endothermic and strongly entropy driven. Increasing ionic strength does decrease their affinity but does not change the sign of ΔS or ΔH . However, in more hydrophobic solvents the sign of ΔH of interaction does change to an exothermic interaction. Moreover, the interacting homopolypeptides showed secondary structure (β -sheet) in environments of much lower hydrophobicity compared to the individual non-interacting homopolypeptides. 2-D NMR showed that interaction occurred primarily via the side-groups with very little interaction between the backbone amide groups. Furthermore, we found that there is a strong preference for the polypeptide chains to bind in a parallel arrangement with the C-termini and N-termini aligned with each other. We anticipate that these results will help improve our understanding of how fundamental interactions between charged residues can result in both productive as well as harmful protein structures.

2.2 Introduction

Charged amino acids play a number of important roles in protein chemistry, including protein stability, folding, ligand binding, and protein-protein interactions and are central to function in intrinsically disordered proteins (IDP's) ¹⁻⁷. IDPs have not only been linked to important biological processes ⁸⁻¹² but also been implicated in several diseases ^{13, 14}. However, much less is known about the structure-function relationship of IDPs compared to structured proteins. A key feature of IDPs are that they contain a relatively large content of charged amino acids ¹⁵, with some regions containing up to ~15 residues with almost 100% charged composition ¹⁶. Analyses of electrostatic interactions using continuum electrostatics models have highlighted the importance of electrostatic complementarity in protein, peptide and nucleotide binding ¹⁷. For example, longer arginine-rich peptides were found to bind to SiRNA more rapidly than shorter peptides ¹⁸. In general, binding of IDPs' to their molecular targets is characterized by high-specificity and low-affinity ¹⁹. However, this description has been recently questioned in the context of relating flexibility and complementarity ²⁰. In this setting, it is important to understand the relationship between electrostatic surface complementarity and binding affinity of a protein/peptide to its molecular target(s) to better understand the interplay of structural forces governing the binding specificity in IDPs.

It is also known that electrostatic interactions play an important role in the folding of proteins and peptides ^{1, 21, 22}. Computational studies on the folding of small and medium-sized polypeptides have shown that folding thermodynamics and kinetics depend strongly on the length of the peptide ²³. Furthermore, protein misfolding is becoming an important focus due to the link between several diseases with misfolded proteins. For example, it is known that tau and prion proteins are associated with various neurological disorders when they misfold ²⁴⁻²⁶.

Moreover, there is often not a clear distinction between IDPs and misfolded proteins, complicated by the fact that disorder is often a precursor to misfolding, and some of the same diseases are associated with both IDPs and misfolded, structured proteins; Alzheimer's disease being an important example^{25, 26}. Naturally, electrostatic interactions are expected to play important roles in these proteins and associated diseases. Such interactions can be long range but are distance and ionic-strength dependent, so factors such as protein and salt concentrations as well as pH are expected to play important roles, and much work is still needed to quantify them.

In the current study, two oppositely-charged homopolypeptides are used as a model for electrostatic-mediated protein binding. The homopolypeptides selected were 15-mers of poly-L-glutamic acid (polyE) and poly-L-lysine (polyK). While the biophysics of much larger versions (>80 amino acids) of these peptides have been well-studied under various conditions, shorter peptides, which are much more relevant to protein interactions, lack similar fundamental biophysical studies. Long chains of polyK exhibit random coil structures at neutral and acidic pH but are known to form α -helix structures in basic solutions, which can be transformed into β sheets by gently heating²⁷. This polypeptide also adopts a α -helical structure in polar organic solvents such as acetonitrile²⁸ and trifluoroethanol (TFE)²⁹. Long chain polyE is a random coil at neutral and basic pH, a α -helix at acidic pH and can also transition into a β sheet when heated³⁰. It is still unknown, however, how short versions of these polypeptides behave when titrated together. Understanding the relationship between interaction thermodynamics and structure that underlies specificity for complementary charged regions of proteins on the order of ~15 residues is expected to provide valuable information on the structural forces that govern electrostatic interactions in both structured and unstructured proteins. The interaction thermodynamics between these homopolypeptides were studied using isothermal titration calorimetry (ITC) and

fluorescence resonance energy transfer (FRET) and the effects of the interactions on the structural properties were studied using circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopy. FRET showed that the C terminus of polyK interacts with the C terminus of polyE in a parallel arrangement. The interaction strength was found to depend strongly on the pH and ionic strength of the aqueous solution. Interestingly, after binding, the polypeptides remained largely as a random coil in neutral pH aqueous solutions although they showed some α -helical character in acidic or basic conditions. However, they formed β sheet secondary structures when they interacted in TFE. This is in contrast to much longer versions of these peptides, which were found to form β -pleated sheets in acidic or neutral aqueous solution, did not interact in basic aqueous solutions and formed aggregates in less-polar solvents, which were suggested to be α -helical structures³¹. We were able to determine that electrostatic interactions between side-chain groups plays a major role in the binding of these highly charged homopolypeptides in aqueous solutions but in more hydrophobic environments (high TFE concentration), hydrogen bonding interactions with backbone C=O and N-H groups play a stronger role causing large secondary structural changes within the homopolypeptides.

2.3. Materials and Methods

2.3.1. *Peptide Synthesis and Labeling*

The far UV CD, NMR, and ITC experiments were carried out using 15-amino acid long homopolypeptides (polyE and polyK) synthesized on a Perkins Elmer/ Applied Biosystems 433A synthesizer using “FastMoc” chemistry³². The molecular mass of the homopolypeptides were verified by ESI mass spectrometry. For fluorescence-based experiments, peptides with the sequence E15C, E6CE9 and K14CK were ordered from Genscript to enable site-specific labeling using maleimide chemistry³³. E15C or E6CE9 was labeled with the donor dye Alexa Fluor 488 (labeled polyE) and K14CK was labeled with the acceptor dye Atto 633 (labeled polyK).

Labeling was carried out using a 5:1 ratio of dye to peptide in 10 mM phosphate buffer at pH 7.2. The mixture of peptide and dye, protected from light, were allowed to react overnight. The unreacted free dye was separated out by dialysis using 2 kDa molecular weight cut-off dialysis tubing (Spectrum Labs). The samples were dialyzed against 10 mM phosphate buffer for 24-48 hours with the external buffer being refreshed every 2-6 hours. The labeling efficiency was verified by MALDI mass spectrometry (Bruker Ultraflex II TOF/TOF time-of-flight mass spectrometer equipped with a MALDI ion source (Bruker Daltonik GmbH, Bremen, Germany)).

2.3.2. *Circular Dichroism*

Far UV circular dichroism experiments were performed using a Jasco J720 spectropolarimeter. Homopolypeptide (polyE and polyK) solutions were prepared in deionized water. The final pH of the homopolypeptide solutions were carefully adjusted either using 1 M HCl or 1 M NaOH. A 1:1 equimolar mixture of the homopolypeptides was prepared by mixing 100 μ L of 1 mM polyK, at room temperature, with 100 μ L of polyE. The desired pH of the mixture was adjusted using HCl and NaOH, the pH meter was equipped with a microprobe. Titration experiments with 2, 2, 2, trifluoroethanol (TFE) were carried out by mixing a 1:1 mixture of polyK and polyE with appropriate volumes of clean TFE. The resultant mixture was readjusted to the desired pH. All far UV (190 nm – 250 nm) CD spectra were acquired, at room temperature, using a 1 mm path length quartz cell. CD spectra were collected using a band width of 1 nm and a scan speed of 50 nm/second and appropriate blank corrections were made in all spectra.

2.3.3. *Isothermal Titration Calorimetry*

All isothermal titration calorimetry experiments were performed at room temperature using VP-ITC (Micoal Inc.). 1 mM solution of polyK, placed in the reaction cell, was titrated incrementally (in 6 μ L aliquots) with 10 mM solution of polyE (placed in the injecting syringe)

with a 12 second interval between injections. The heats of the reaction per injection (microcalories/second) were determined by integration of the peak areas using the Origin software (version 7.0). The thermodynamic and binding stoichiometry values, characterizing the interaction(s), were calculated by fitting the raw data using the fitting models provided in the software. The best-fit of the data was judged based on the χ^2 values obtained on the individual fits.

2.3.4. Nuclear Magnetic Resonance (NMR) Spectroscopy

NMR data were acquired at 25 °C on a Bruker AVANCE 700 MHz spectrometer equipped with a 5-mm triple resonance cryoprobe. 1 mM homopolypeptide (polyK and polyE) solution was prepared in 5% v/v D₂O + 95% v/v H₂O. The final pHs of the solution(s) were adjusted with either 1 M HCl or 1 M NaOH. Two-dimensional proton TOCSY and NOESY data were acquired with 2048 data points in the F2 dimension and 512 increments in the F1 dimension. The spectral width was set to 12 ppm. TOCSY and NOESY spectra were acquired with mixing times of 60 ms and 150 ms. NOE-based distance restraints were derived from 2D ¹H NOESY data obtained using different (200, 250, 300, and 350 ms) times. NMR spectra were processed using XWIN-NMR and Sparky software³⁴.

2.3.5. Fluorescence Resonance Energy Transfer (FRET) Spectroscopy

FRET measurements were made on a PTI Quantamaster 40. The concentration of labeled polyE was set at 5.0×10^{-8} M. The concentration of labeled polyK varied from 5.0×10^{-10} to 1.5×10^{-6} M. The sample was excited at 470 nm and emission data were collected from 480 nm to 750 nm. The transfer efficiency was calculated by using the relative fluorescence intensity of the donor at 516 nm in the presence (F_{DA}) and absence (F_D) of the acceptor³⁵ using: $E = 1 - (F_{DA}/F_D)$. The K_D value was calculated by fitting to the Hill equation in Origin³⁶.

2.4. Results and Discussion

2.4.1 *Peptide Selection, Labeling and Characterization.*

Cysteine residues were introduced in polyE either as the C-terminal most amino acid (E15C) or in the middle of the peptide at position 7 (E6CE9) to examine the binding efficiency and also probe the relative orientation of the interacting homopolypeptides. However, introduction of a cysteine residue in polyK as the C-terminal most amino acid proved to be technically challenging and therefore the cysteine was introduced as the penultimate residue from the C-terminal end of polyK (K14CK).

The molecular mass and the location of the cysteine group were confirmed by MS/MS data (figure 2.1) on the homopolypeptides. PolyE (E15C and E6CE9) and polyK (K14CK) were tagged with fluorescent dyes Alexa 488 and Atto 633, respectively. MALDI mass analysis of the purified fluorescent labeled homopolypeptides showed the expected molecular mass and verified highly efficient labeling (figure 2.2).

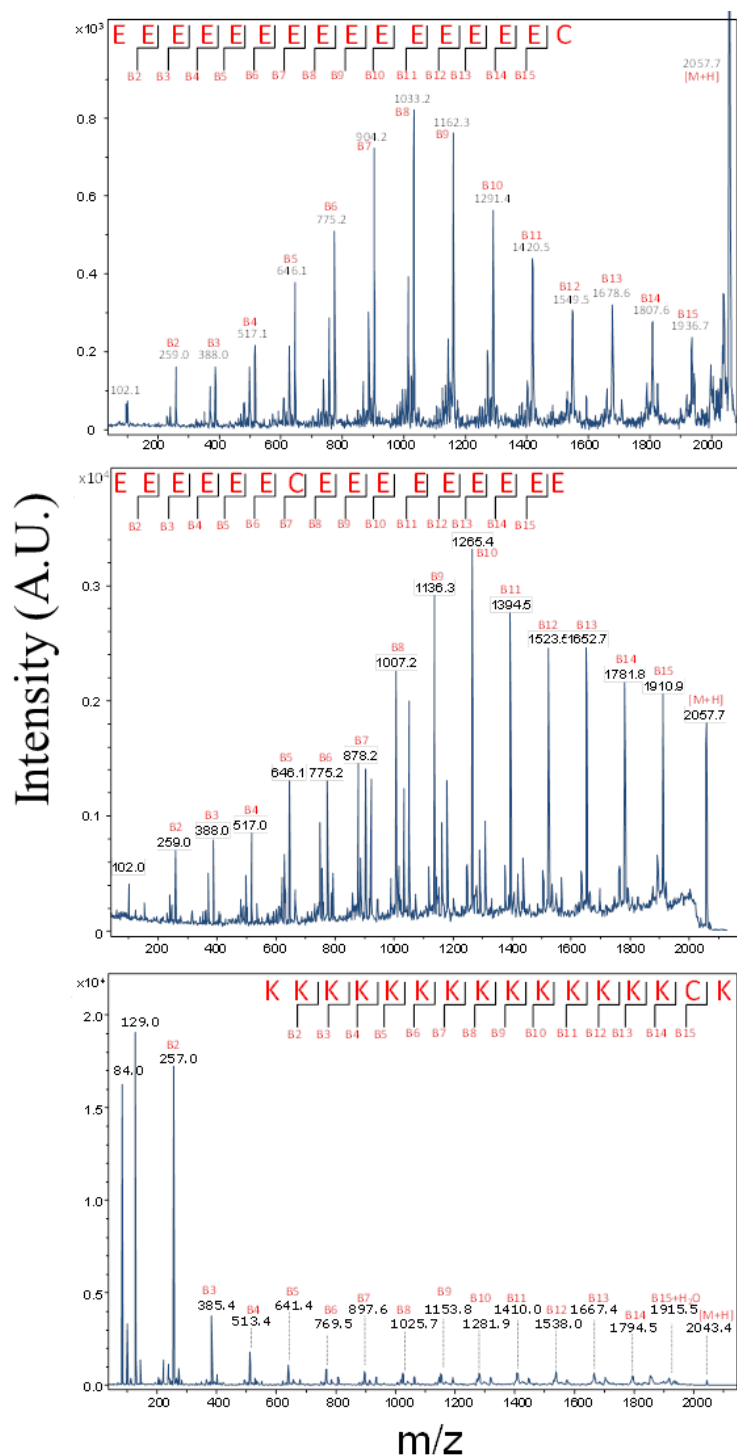


Figure 2.1. MS/MS of E15C, E6CE9 and K14CK. Top panel is for E15C, middle is E6CE9, bottom is K14CK. These results show the various fragments (labeled B2 through B15) that identifies the cysteine location via the change in mass upon removing each amino acid.

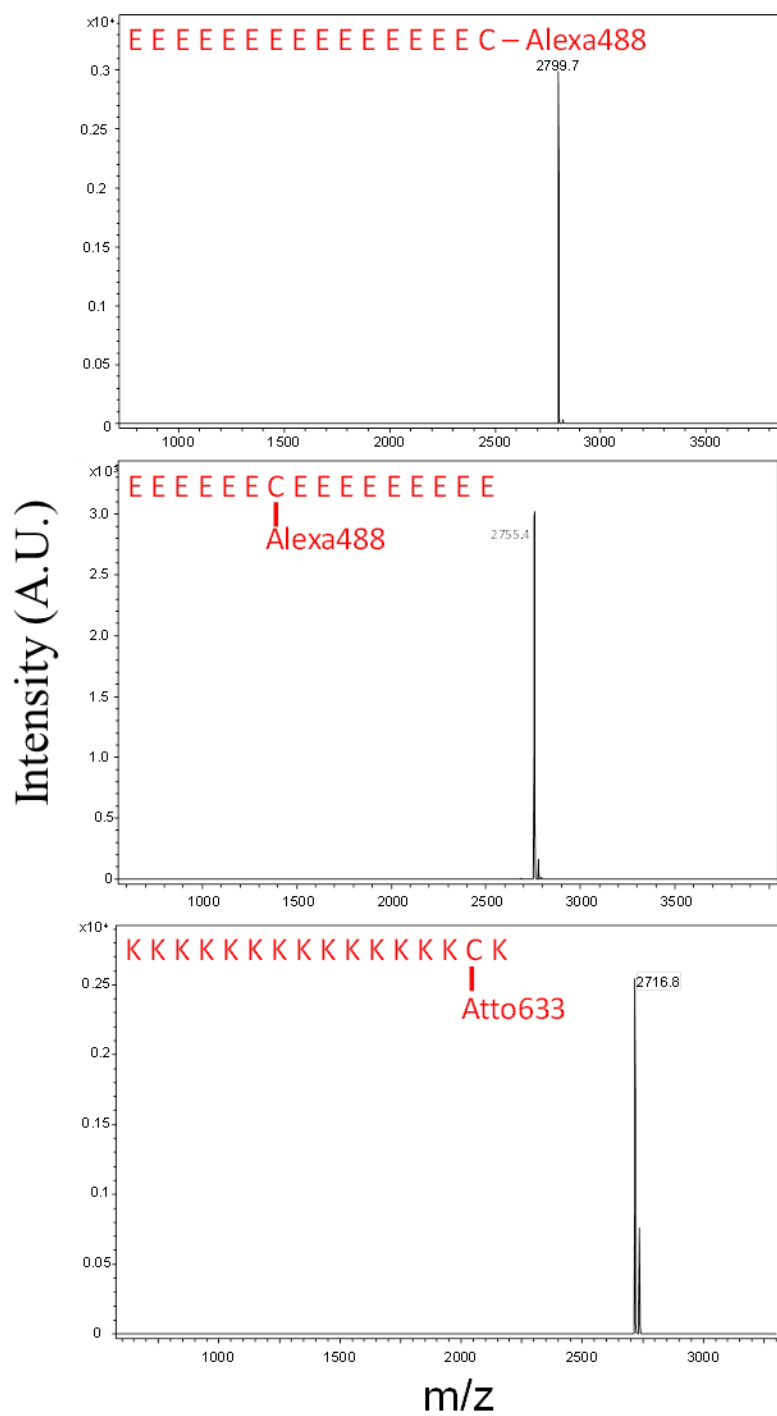


Figure 2.2. MALDI-TOF MS of labeled peptides after purification. MALDI shows a single peak highlighting complete labeling in a 1:1 dye: polypeptide ratio.

2.4.2. Isothermal titration calorimetry (ITC) of polyE/polyK Binding

ITC is a powerful technique for the measurement of binding affinity^{37, 38}, stoichiometry^{39, 40}, and the enthalpy and entropy changes^{41, 42} governing interactions. The interaction between polyE and polyK under different conditions of pH (pH – 2.0, 7.0 and 13.0) and salt concentrations (100 mM, 250 mM, and 500 mM NaCl) was examined using ITC (figure 2.3).

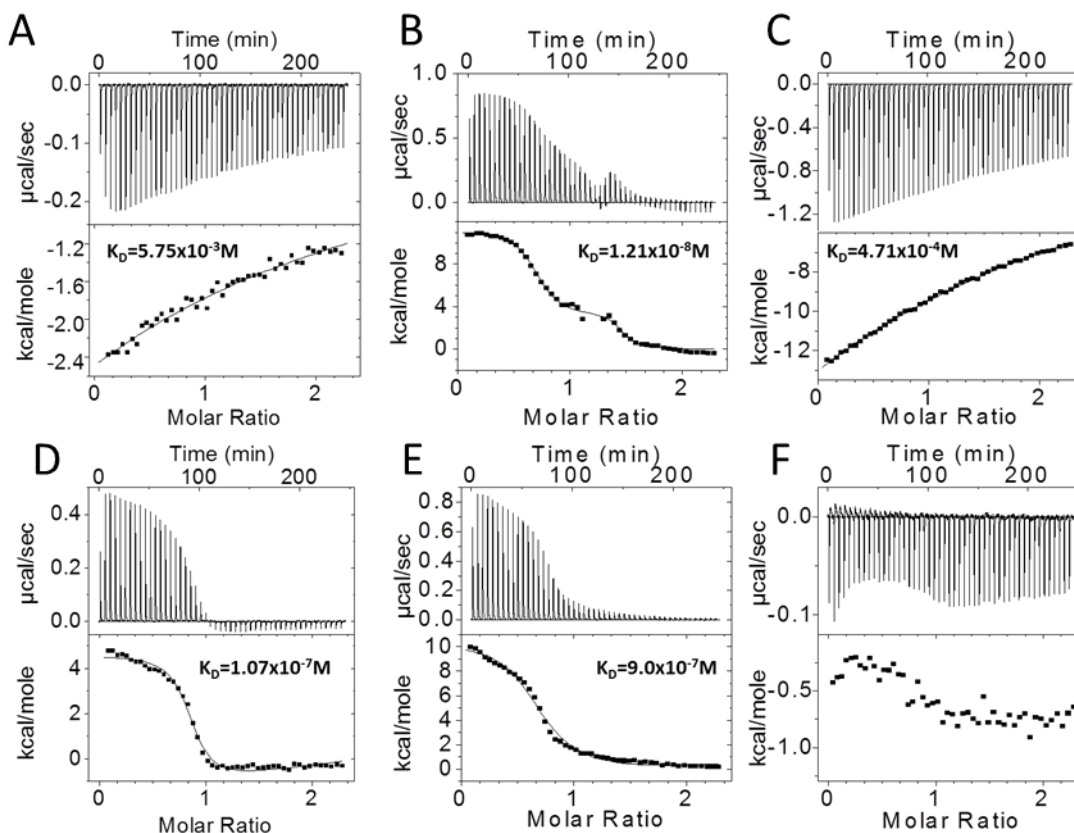


Figure 2.3. ITC data for polyE/polyK binding in A) pH2, 0 mM NaCl; B) pH7, 0 mM NaCl; C) pH13, 0 mM NaCl; D) pH7, 100 mM NaCl; E) pH7, 250 mM NaCl; F) pH7, 500 mM NaCl. From N. Webb, *Conformational Studies of Homopolyptides*.

Both polyK and polyE are charged at pH 7.0. The binding between the oppositely-charged homopolyptides, at pH 7.0, is endothermic and proceeds with an overall increase in

entropy. The binding isotherm is biphasic and best fits to a two-site binding model. The binding occurs with an average binding constant ($K_{D(app)}$) of 1.21×10^{-8} M (Table-2.1). The two observed phases plausibly represent two discrete steps involved in the binding between the two oppositely charged homopolypeptides. This suggests complex electrostatic interactions in which the charged homopolypeptides interact differently at high concentrations than at low concentrations in low ionic strength conditions, possibly via some cooperative aggregation process. The average binding affinity between polyE and polyK at the extremes of pH (pH 2.0 and 13.0), wherein one of the polypeptides is not charged, is reduced by 4-5 orders of magnitude to 5.75×10^{-3} M and 4.71×10^{-4} M, respectively. For binding interactions between the homopolypeptides at pH 2.0 and pH 13.0, the affinities are much weaker and so saturation was not observed, leading to accurate values for the enthalpy and entropy of these interactions not being able to be quantified. However, the sign of the enthalpy change is clear from the direction of change of the peaks during titration. Interestingly, unlike at pH 7, binding interactions between the homopolypeptides at pH 2.0 and pH 13.0 are exothermic.

Table 2.1. Thermodynamic parameters describing binding of polyE to polyK under various conditions of pH and NaCl concentration as measured by ITC. From N. Webb, *Conformational Studies of Homopolypeptides*.

Conditions	K_D (M)	ΔH (kJmol ⁻¹)	ΔS (JK ⁻¹ mol ⁻¹)
pH 2, 0 mM NaCl	5.75×10^{-3}	N/D (-ve)	N/D
pH 13, 0 mM NaCl	4.71×10^{-4}	N/D (-ve)	N/D
pH 7, 0 mM NaCl	1.21×10^{-8}	7.2	50.8
pH 7, 100 mM NaCl	1.07×10^{-7}	4.5	47
pH 7, 250 mM NaCl	9.0×10^{-7}	10.1	61.1

ITC data were acquired at pH 7.0, as a function of ionic strength (0 mM to 500 mM NaCl) to determine the nature of forces which are in play in the interaction between polyE and polyK (figure 2.3 B,D,E,F). The binding affinity of the charged homopolypeptides was very strong in the absence of salt ($K_D = 1.21 \times 10^{-8}$ M). A progressive loss in binding affinity was observed as the NaCl concentration was increased from 100 mM to 250 mM (figure 2.3 B,D,E,F). The binding between the charged homopolypeptides was almost completely abolished at 500 mM. It should be noted that the ΔH and ΔS values are all positive at pH 7 and they are seen to increase with increasing NaCl concentration.

In general, the decrease in binding affinity upon increasing the ionic strength suggests that the interaction between the charged homopolypeptides, at neutral pH, is predominantly electrostatic. The observed lack of binding at 500 mM NaCl, even though at 250 mM the binding is still relatively strong coupled with the weak but clear binding at pH 2 and pH 13 suggests that the binding mechanism is complex. While the ΔH changes to negative as the ionic state of either polypeptide is changed through pH, these ΔH values remain positive as you increase the NaCl concentration at pH 7. All of this information together highlights this mechanistic complexity and suggests that structural changes upon binding are different under the different pH and NaCl conditions.

2.4.3. *polyE/polyK Binding Studied by FRET*

Fluorescence techniques have extremely high sensitivity, allowing for significantly lower concentrations than are used in ITC. Moreover, due to the strong distance dependence of FRET, careful placement of the dyes can provide information on the relative orientation of the interacting homopolypeptides. By labeling polyE with the donor dye (Alexa 488), we were able to measure the decrease in its signal when the acceptor-labeled polyK (Atto 633) was titrated

into the solution. We used the same NaCl concentrations as for ITC, 0 mM to 250 mM NaCl. Figure 2.4 shows the FRET data for the donor-labeled E6CE9 (figure 2.4A) and E15C (figure 2.4B) samples upon adding acceptor-labeled K14CK polypeptide. When the donor dye was placed close to the center of the polyE chain (E6CE9), the FRET efficiency increased upon adding polyK to a maximum value of ~0.4. At pH 7, in the absence of salt, the K_D was found to be 1.65×10^{-8} M, table 2.2. With addition of NaCl, the K_D decreased to 5.37×10^{-8} M (100 mM), and then further to 2.12×10^{-7} M (250 mM NaCl). The good agreement of the FRET data with the ITC data is shown in figure 2.5, with only a slight variance in the samples that used salt. One possible reason for this difference may be the concentration difference of polyE used in the experiments (50 nM for FRET vs 500 μ M for ITC) since in the FRET experiments the concentration of the peptides approaches the K_D value. Another possible reason is that for ITC the data were not accurately fit by a single-site binding model, suggesting more complex binding mechanisms could be present at the higher concentrations used in ITC. The data were fit using both the Morrison equation and Hill equation in order to determine the binding coefficients. Inspection of the fit curves and the χ^2 values showed that the Hill equation fit the data much better, although the K_D values obtained using both methods were very similar (figure 2.4). It was observed that the Hill coefficient was higher for the lower NaCl concentrations ($n = 3$ for 0 mM and $n = 1.4$ for 250 mM NaCl), indicating that the binding cooperativity increased as the ionic strength of the solution decreased. The most likely reason for this observation is that, at low ionic strength and high polyK/polyE ratio, the lack of charge screening between the homopolypeptides causes some degree of aggregation leading to multiple polyK's binding to a single polyE, in agreement with ITC observations. However, the maximum FRET value observed in figure 2.4A

is the same 40% value independent of ionic strength, suggesting that, if aggregation does occur, the additional dyes are outside the maximum FRET distance range.

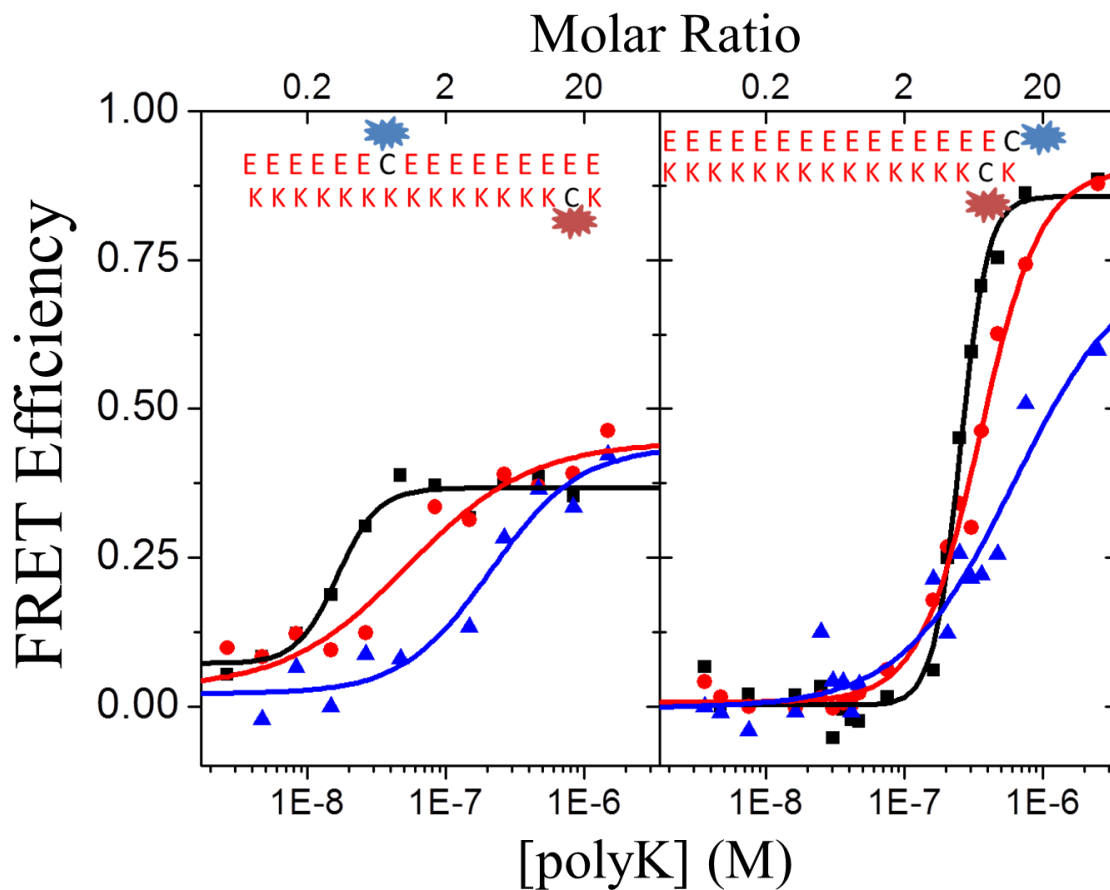


Figure 2.4. FRET data fitted using Hill equation. Labeled polyK/labeled polyE binding at pH 7 with NaCl concentrations of 0mM (black), 100 mM (red) and 250 mM (blue) with different donor dye placement positions on the polyE.

Table 2.2. FRET conditions and binding constants for polyE/polyK.

Conditions	Sample	K_D in M
0 mM NaCl, pH 7	E6CE9/ K14CK	$1.65 \times 10^{-8} \pm 1.7 \times 10^{-9}$
100 mM NaCl, pH 7	E6CE9/ K14CK	$5.37 \times 10^{-8} \pm 2.4 \times 10^{-8}$
250 mM NaCl, pH 7	E6CE9/ K14CK	$2.12 \times 10^{-7} \pm 8.3 \times 10^{-8}$
0 mM NaCl, pH 7	E15C/ K14CK	$2.53 \times 10^{-7} \pm 7.6 \times 10^{-9}$
100 mM NaCl, pH 7	E15C/ K14CK	$3.50 \times 10^{-7} \pm 2.2 \times 10^{-8}$
250 mM NaCl, pH 7	E15C/ K14CK	$6.18 \times 10^{-7} \pm 3.74 \times 10^{-7}$

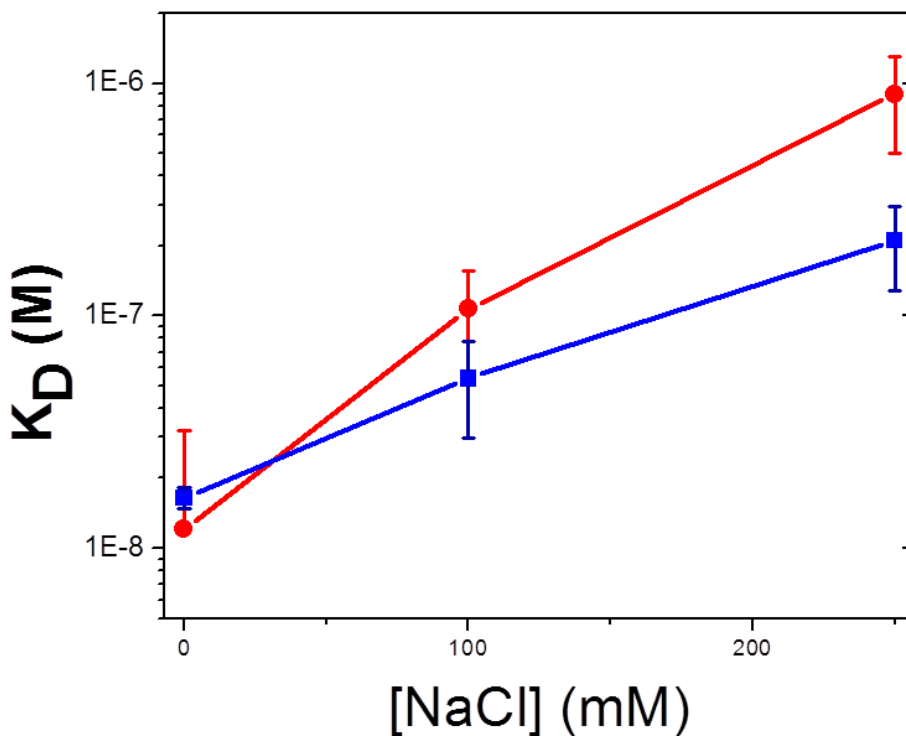


Figure 2.5. Comparison of K_D measured using FRET (blue) and ITC (red). FRET data comes from the dye in the middle position on the polyE peptide. Error bars are from the error given in fitting.

When the dyes were both placed on the C-terminus of the charged homopolypeptides (figure 2.4B), two effects were discovered. First, the maximum FRET efficiency that was observed was much higher than when the donor dye was close to the center of the polyE chain (~0.9 vs. ~0.4). This indicates that the peptides are bound in a parallel arrangement bringing the donor and acceptor dyes into very close proximity. An antiparallel disposition of the homopolypeptides would have resulted in a lower FRET value than observed in figure 2.4A. Second, the K_D values were reduced for the E15C samples compared to the E6CE9 samples at each given concentration of NaCl indicating that the binding was weaker when the dyes were forced to be close to each other. This also indicates that the preference for a parallel assembly must be so strong that the steric hindrance from the dyes was not able to force the peptides to bind in an anti-parallel arrangement.

2.4.4. *Far-UV CD spectroscopy of equimolar mixtures of polyE and polyK*

Far-UV CD spectroscopy is a powerful method to determine secondary structural information. However, the far-UV CD spectrum is an ensemble-averaged signal from all conformations in the sample and requires careful deconvolution to quantitatively estimate the constituent secondary structural elements⁴³. Nevertheless, changes in the average structure can be assessed by measuring the spectral changes at certain wavelengths (or ratios of wavelengths) upon binding or changing environmental parameters. The far-UV CD spectra of an equimolar mixture of polyE and polyK were acquired at pH values ranging from 1-14 (figure 2.6A, B). The CD spectra showed significant backbone disorder under all pH conditions. The 222/208 nm ratio is higher for α -helices than random coils and a plot of mean residue ellipticity ratio (222 nm/208 nm) *versus* pH (figure 2.6B) for the bound polyE/polyK did show some α -helical content at extreme pH values (pH 1 and pH 14). Studies of long-chain polyK homopolypeptides³¹ shows that, at high pH, they show α -helical secondary structure, while polyE is expected to be

disordered. At low pH, the converse is true; polyE forms a α -helix⁴⁴ and polyK is disordered. However, we find that, for 15-mer homopolypeptides, they are largely unstructured at all pH's, although there is a minor contribution from α -helix structure at extreme pH, as evidenced from figure 2.6.

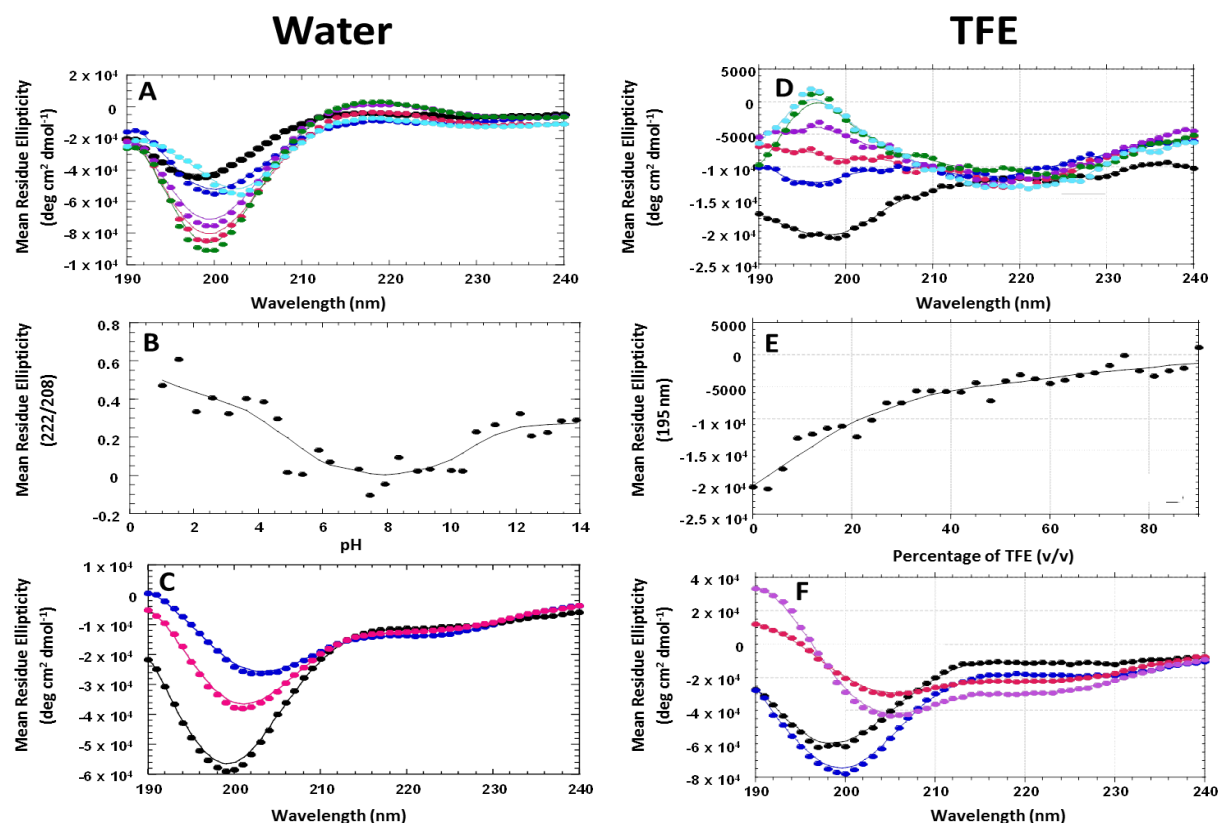


Figure 2.6. CD spectra of polypeptides in water or TFE. **A)** Far-UV CD spectra of an equimolar mixture of polyK and polyE in water at pH 2 (black), pH 4 (blue), pH 6 (pink), pH 8 (green), and pH 13 (cyan). **B)** Mean residue ellipticity ratio (222 nm/208 nm) as a function of pH. **C)** Far-UV CD spectra of the summation of the individual polyK and polyE spectra at pH 2 (blue), pH 7 (black), and pH 13 (pink). **D)** Far-UV CD spectra of an equimolar mixture of polyK and polyE at pH 7 in TFE/H₂O at 0% v/v (black), 12% v/v (blue), 30% v/v (pink), 45% v/v (purple), 75% v/v (green), and 90% v/v (cyan). **E)** Change in mean residue ellipticity at 195 nm versus concentration of TFE (v/v). **F)** Far-UV CD spectra of the summation of the individual polyK and polyE spectra at 0% v/v (black), 50% v/v (blue), 75% v/v (pink), and 90% v/v (purple). From N. Webb, *Conformational Studies of Homopolypeptides*.

••••• Black
••••• Blue
••••• Green
••••• Pink
••••• Purple
••••• Cyan

In order to follow the secondary structural changes of polyE and polyK upon binding, far-UV CD spectra of individual 15-mer polyE and polyK were separately measured at low, intermediate, and high pH and the summation of these individual spectra are shown in figure 2.6C. Comparing figure 2.6A and 2.6C for CD, the ellipticity minima at 208 nm and 222 nm are slightly more defined in the summation of individual polyE and polyK spectra at extreme pH (figure 2.6C) than that observed for an equimolar mixture of the homopolypeptides at the same pH (figure 2.6A). These results suggest that, upon interaction of the homopolypeptides, the already low propensity for α -helix formation at low and high pH is reduced even further.

Electrostatic interactions are strengthened in solvents with low dielectric constants such as 2,2,2-trifluoroethanol (TFE). In this context, the effect of dielectric of the medium was examined by treating an equimolar mixture of the polyE/polyK at pH 7.0 with different percentage concentrations of TFE v/v. Interestingly, although TFE is primarily considered to be an α -helicogenic solvent, addition of moderate TFE concentrations (~20% v/v) caused a transition of the bound polyE/polyK from primarily random coil to increased β -sheet structure (figure 2.6D). This is exemplified by the presence of a weak negative band centered near 216 nm and a strong positive band at 195 nm⁴⁵. To quantify the amount of β -sheet present in the equimolar polyE/polyK mixture the change in mean residue ellipticity at 195 nm was plotted as a function of TFE concentration (figure 2.6E).

As mentioned above, both polyE and polyK exist as random coils at pH 7. The spectra resulting from the summation of the individual spectra for polyE and polyK in increasing concentrations of TFE at pH 7 is shown in figure 2.6F. As expected, the spectra obtained by summation of the individual spectra for polyE and polyK reflect largely random coil structures below 75% v/v TFE concentrations and α -helical conformations above 75% v/v TFE. However,

when polyE and polyK were bound together, the CD spectra showed β -sheet rather than α -helical structure. Interestingly, much lower concentrations of TFE were required to induce a random coil to structured (β -sheet) transition in the bound homopolypeptides compared to the individual isolated homopolypeptides. A significant change in the mean residue ellipticity at 190 nm of the far-UV CD spectra of the equimolar polyE/polyK mixture was found between 0-20% v/v TFE. This much lower hydrophobicity threshold for interacting short-chains of polyE/polyK to form secondary structure may have significant implications in the binding-folding hypothesis of IDPs⁴⁶.

2.4.5. *TOCSY and NOESY NMR Spectroscopy*

Total Correlation Spectroscopy (TOCSY) is a nuclear magnetic resonance (NMR) spectroscopy technique used to correlate all the protons of a spin system. The resulting spectrum shows a specific pattern of crosspeaks for each amino acid. Nuclear Overhauser Enhancement Spectroscopy (NOESY) produces a crosspeak for each pair of nuclei that are coupled through space. The intensity of the NOESY crosspeaks depends on the inverse of the distance between the nuclei to the sixth power ($1/r^6$). The combination of these two techniques provides atomic-level information on the nature of interactions between the charged homopolypeptides.

TOCSY and NOESY spectra were first acquired for the isolated homopolypeptides to unambiguously assign all the resonances. TOCSY and NOESY spectra for the polyE and polyK individually are shown in work done with our collaborator⁴⁷, together with predicted chemical shifts. The TOCSY and NOESY spectra of the polyE/polyK mixture for two regions are shown in figures 2.7A and B. Three potential NOEs were observed. Two NOEs are indicated by the number “1” in figure 2.7A. Since the intensity of the NOE crosspeaks increases further downfield in the NOESY spectrum, the NOEs can be attributed to interactions between the

backbone N-H groups of the two charged homopolypeptides. For the NOE peak labeled as “2” in figure 2.7B, the intensity of the peak(s) increases further up field in the NOESY spectrum suggesting that the NOE represents the interaction between the H^ε of polyK and H^γ of polyE. This observation of NOEs shows that the polyE/polyK mixture primarily interacts via the electrostatic interactions between the oppositely charged side-chain groups of the two homopolypeptides, with only weak contributions from backbone H-bond interactions. These NMR results support and help explain the observations from the CD spectra that polyE and polyK have very little secondary structure upon their interaction.

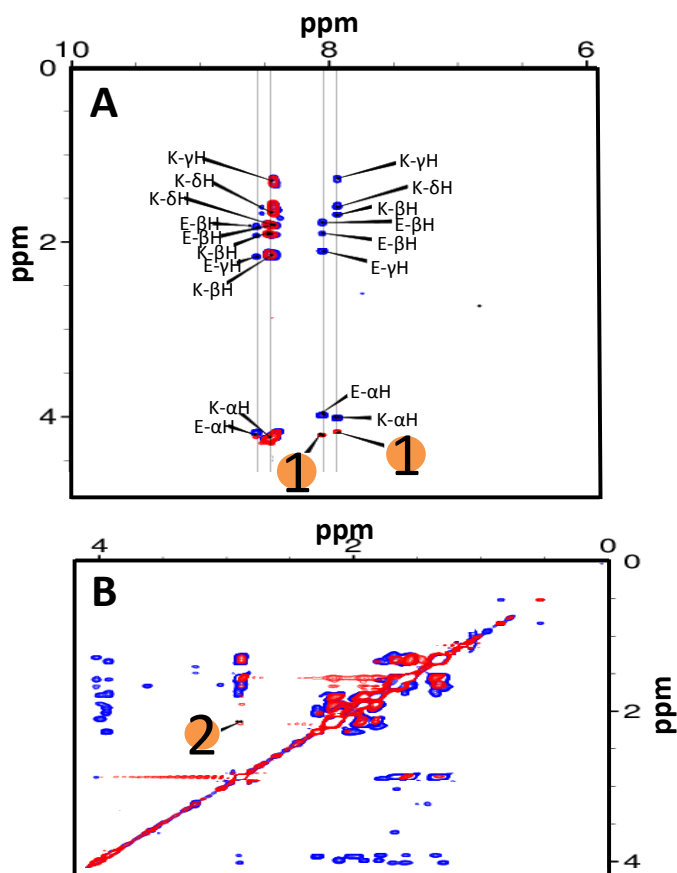


Figure 2.7. TOCSY (blue) and NOESY (red) NMR spectra of polyE/polyK mixture at pH 7. The light gray lines in A are for visual reference. From N. Webb, *Conformational Studies of Homopolypeptides*.

2.4.6. Combining Structural and Thermodynamic Data

The CD and NMR data provide some structural basis for the thermodynamic observation that the ΔH and ΔS are both positive upon binding at pH7 independent of the salt concentration, and that the binding is exothermic at the extremes of pH. The bond energy between the charged groups and the multiple solvent molecules or ions is relatively strong and this bond energy is not completely balanced by the newly-formed salt bridges between the K and E groups. However, the entropy strongly increases when solvent/ions that interact with the charged groups of polyK and polyE are released into the bulk providing the major driving force for the interaction. Clearly, the entropy decrease upon bringing together the two relatively small homopolypeptide chains together is more than made up for by this release of solvent/ions into the bulk. Naturally, as ionic strength increases, this argument holds since the only effect of the additional ions is to screen the interactions without affecting the overall structure. It seems that more solvent/ions are released when the chains bind in a parallel arrangement suggesting that the geometry of the chains play a role in the solvent/ion interactions. At the extremes of pH, where one of the homopolypeptides is uncharged, solvent molecules or ions interact less strongly with the polyK (at pH13) or polyE (at pH2). Bond energy is released by the ionization of the uncharged residues to form the salt bridge upon interaction. Although it was not quantified, one may expect that the entropy decrease from the interaction of the two chains is not as strongly balanced by the release of less solvent/ions into the bulk, leading to a more enthalpic-driven process.

2.5. Conclusions

We have presented a combined thermodynamic and structural study to better understand how oppositely-charged amino acids in short chain polypeptides interact. This interaction is measured when both polypeptides are charged or when one or both of them is/are made to be uncharged by

changing pH or solvent hydrophobicity. The main conclusions drawn from this combined study are as follows:

- 1) When the ionization state of 15 amino acid homopolypeptides are both charged they are random coils and remain as random coils after their high-affinity interaction. The interaction is strongly entropy driven by releasing solvent/ions into the bulk.
- 2) Upon changing the ionization state of 15-mers of polyE or polyK by pH, they maintain a great deal of random coil behavior both before and after interacting. There is a *slightly* increased propensity for individual polyE in low pH aqueous solvents and for individual poly-K in high pH aqueous solvents to form α -helical structures but this is reduced once they interact.
- 3) There is an increased propensity for individual 15-mers of polyE and polyK to form α -helical secondary structure in hydrophobic solvents such as TFE, but requires high concentrations (>75% v/v) to do so. After interaction, polyE and polyK form β -sheets under mildly hydrophobic conditions (<20% v/v TFE).
- 4) Although they remain as random coils after binding, there is a strong preference for the homopolypeptide chains to bind in a parallel arrangement in aqueous solvents with the C-termini and N-termini aligned with each other.

The conclusions from this study helps expand our understanding of fundamental interactions between charged residues in highly acidic or basic regions of proteins on the scale of ~15 amino acid, which play important roles in regulation of protein interactions as well as underlie structure-function relationships in IDPs. Eventually, such knowledge may help us to determine the conditions that can result in both productive as well as harmful protein structures involving highly charged regions of proteins.

2.6. Acknowledgements

This work was supported by the National Institutes of Health (NIH) through the COBRE program (P30 GM103450), and the Arkansas Biosciences Institute.

2.7. References

1. Sinha, N., and Smith-Gill, S. (2002) Electrostatics in protein binding and function, *Curr. Protein Pept. Sci.* 3, 601-614.
2. Davis, S. J., Davies, E. A., Tucknott, M. G., Jones, E. Y., and Van, D. M. P. A. (1998) The role of charged residues mediating low affinity protein-protein recognition at the cell surface by CD2, *Proc. Natl. Acad. Sci. U. S. A.* 95, 5490-5494.
3. Misra, V. K., Hecht, J. L., Yang, A.-S., and Honig, B. (1998) Electrostatic contributions to the binding free energy of the λ cI repressor to DNA, *Biophys. J.* 75, 2262-2273.
4. Vijayakumar, M., Wong, K.-Y., Schreiber, G., Fersht, A. R., Szabo, A., and Zhou, H.-X. (1998) Electrostatic enhancement of diffusion-controlled protein-protein association: comparison of theory and experiment on barnase and barstar, *J. Mol. Biol.* 278, 1015-1024.
5. Xiao, L., and Honig, B. (1999) Electrostatic Contributions to the Stability of Hyperthermophilic Proteins, *J. Mol. Biol.* 289, 1435-1444.
6. Hu, Z., Ma, B., Wolfson, H., and Nussinov, R. (2000) Conservation of polar residues as hot spots at protein interfaces, *Proteins: Struct., Funct., Genet.* 39, 331-342.
7. Sheinerman, F. B., Norel, R., and Honig, B. (2000) Electrostatic aspects of protein-protein interactions, *Curr. Opin. Struct. Biol.* 10, 153-159.
8. Dunker, K., and Obradovic, Z. (2001) The protein trinity: Linking function and disorder, *Nature* 19, 805-806.
9. Dunker, A. K. a. B. C. J. a. L. J. D. a. I. L. M. a. O. Z. (2002) Intrinsic Disorder and Protein Function†, *Biochemistry* 41, 6573-6582.
10. Dunker, K., and Brown, C. J. (2002) Identification and functions of usefully disordered proteins, *Adv. Protein Chem.* 62, 25 - 49.
11. Tompa, P. a. C. P. (2004) The role of structural disorder in the function of RNA and protein chaperones, *FASEB J.* 18, 1169-1175.
12. Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions, *Nat. Rev. Mol. Cell Biol.* 6, 197-208.
13. Das, S. a. M. D. (2011) Intrinsically unstructured proteins and neurodegenerative diseases: Conformational promiscuity at its best, *IUBMB Life* 63, 478--488.
14. Dosztányi, Z., Mészáros, B., and Simon, I. (2010) Bioinformatical approaches to characterize intrinsically disordered/unstructured proteins, *Briefings Bioinf.* 11, 225-243.

15. Uversky, V., and Dunker, A. (2010) Understanding protein non-folding, *Biochim. Biophys. Acta, Proteins Proteomics* 1804, 1231 - 1264.
16. Gast, K., Damaschun, H., Eckert, K., Schulze-Forster, K., Maurer, H. R., Mueller-Frohne, M., Zirwer, D., Czarnecki, J., and Damaschun, G. (1995) Prothymosin .alpha.: A Biologically Active Protein with Random Coil Conformation, *Biochemistry* 34, 13211-13218.
17. McCoy, A. J., Chandana Epa, V., and Colman, P. M. (1997) Electrostatic complementarity at protein/protein interfaces, *J. Mol. Biol.* 268, 570-584.
18. Kim, M., Kim, H. R., Chae, S. Y., Larson, R. G., Lee, H., and Park, J. C. (2013) Effect of Arginine-Rich Peptide Length on the Structure and Binding Strength of siRNA–Peptide Complexes, *J. Phys. Chem. B* 117, 6917-6926.
19. Uversky, V. N. (2013) Intrinsic Disorder-based Protein Interactions and their modulators, *Curr. Pharm. Des.* 19, 4191-4213.
20. Huang, Y., and Liu, Z. (2013) Do Intrinsically Disordered Proteins Possess High Specificity in Protein–Protein Interactions?, *Chem. Euro. J.* 19, 4462-4467.
21. Kumar, S., and Nussinov, R. (2002) Close-Range electrostatic Interactions in Proteins, *ChemBioChem* 3, 604-617.
22. Dill, K. A. (1990) Dominant forces in protein folding, *Biochemistry* 29, 7133-7155.
23. Duara, X. (2005) Molecular dynamics simulation of peptide folding, *Theor. Chem. Acc.* 116, 297-306.
24. Nath, A., Sammalkorpi, M., DeWitt, David C., Trexler, Adam J., Elbaum-Garfinkle, S., O'Hern, Corey S., and Rhoades, E. (2012) The Conformational Ensembles of Alpha-Synuclein and Tau: Combining Single-Molecule FRET and Simulations, *Biophys. J.* 103, 1940-1949.
25. Savelieff, M. G., Lee, S., Liu, Y., and Lim, M. H. (2013) Untangling amyloid- β , tau, and metals in alzheimer's Disease, *ACS Chem. Biol.* 8, 856-865.
26. Norrby, E. (2011) Prions and protein-folding diseases, *J. Intern. Med.* 270, 1-14.
27. Townend, R., Kumosiniki, T. F., and Timasheff, S. N. (1966) The circular dichroism of the beta structure of Poly-L-Lysine, *Biochem. Biophys. Res. Commun.* 23.
28. Arunkumar, A. I., Kumar, T. K. S., Sivaraman, T., and Yu, C. (1997) Acetonitrile-induced conformational transitions in Poly-L-Lysine, *Int. J. Biol. Macromol.* 21, 299-305.
29. Mirtič, A., and Grdadolnik, J. (2013) The structure of poly-l-lysine in different solvents, *Biophys. Chem.* 175–176, 47-53.

30. Itoh, K., Foxman, B. M., and Fasman, G. D. (1976) The two Beta forms of poly (L-glutamic acid), *Biopolymers* 15, 419-455.
31. Hammes, G. G., and Schullery, S. E. (1968) Structure of macromolecular aggregates. I. Aggregation-induced conformational changes in polypeptides, *Biochemistry* 7, 3882-3887.
32. Fields, C. G., Lloyd, D. H., Macdonald, R. L., Otteson, K. M., and Noble, R. L. (1991) HBTU activation for automated Fmoc solid-phase peptide synthesis, *Pept. Res.* 4, 95-101.
33. Hermanson, G. (2013) *Bioconjugate Techniques*, 3rd ed., Academic Press.
34. Goddard, T. D., and Kneller, D. G. Sparky, University of California, San Francisco.
35. Lakowicz, J. R. (2006) *Principles of fluorescence spectroscopy*, 3rd ed., Springer, New York.
36. Weiss, J. N. (1997) The Hill equation revisited: uses and misuses, *FASEB J.* 11, 835-841.
37. Lemmon, M. A., and Ladbury, J. E. (1994) Thermodynamic Studies of Tyrosyl-Phosphopeptide Binding to the SH2 Domain of p56lck, *Biochemistry* 33, 5070-5076.
38. Gorshkova, I., Moore, J. L., McKenney, K. H., and Schwarz, F. P. (1995) Thermodynamics of Cyclic Nucleotide Binding to the cAMP Receptor Protein and Its T127L Mutant, *J. Biol. Chem.* 270, 21679-21683.
39. Keown, M. B., Henry, A. J., Ghirlando, R., Sutton, B. J., and Gould, H. J. (1998) Thermodynamics of the Interaction of Human Immunoglobulin E with Its High-Affinity Receptor FcεRI, *Biochemistry* 37, 8863-8869.
40. Vippagunta, S. R., Dorn, A., Matile, H., Bhattacharjee, A. K., Karle, J. M., Ellis, W. Y., Ridley, R. G., and Vennerstrom, J. L. (1999) Structural Specificity of Chloroquine-Hematin Binding Related to Inhibition of Hematin Polymerization and Parasite Growth, *J. Med. Chem.* 42, 4630-4639.
41. Hu, C. Q., and Sturtevant, J. M. (1992) Thermodynamics of binding of mononucleotides to ribonuclease T1, *J. Phys. Chem.* 96, 4052-4056.
42. Turner, D. C., Straume, M., Kasimova, M. R., and Gaber, B. P. (1995) Thermodynamics of Interaction of the Fusion-Inhibiting Peptide Z-D-Phe-L-Phe-Gly with Dioleoylphosphatidylcholine Vesicles: Direct Calorimetric Determination, *Biochemistry* 34, 9517-9525.
43. Sreerama, N., Venyaminov, S. Y., and Woody, R. W. (1999) Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy, *Protein Sci.* 8, 370-380.

44. Hayashi, H., and Wada, A. (1981) High-resolution study of the helix coil transition of poly(L-glutamic acid): Spectroscopic data correlation and the discovery of a new transition, *Biopolymers* 20, 695-705.
45. Pelton, J. T., and McLean, L. R. (2000) Spectroscopic methods for analysis of protein secondary structure, *Anal. Biochem.* 277, 167-176.
46. Wright, P. E., and Dyson, H. J. (2009) Linking folding and binding, *Curr. Opin. Struct. Biol.* 19, 31-38.
47. Webb, N. (2011) Conformational Studies of Homopolypeptides, University of Arkansas.

Thermodynamic and Structural Characterization of Interactions between Oppositely-Charged Short-Chain Homopolypeptides. Ashley Howard, Nicole Webb, T.K.S. Kumar and Colin D. Heyes. **TO BE SUBMITTED**

Figures 2.3, 2.6, 2.7 and Table 2.1 are modified versions of figures from Nicole Webb's dissertation.

III. Determining the Effects of Fluorescent Labels on Electrostatic Interactions using a Polypeptide Template

3.1. Abstract

The uses of fluorescent dyes are widespread in biolabeling experiments. There are many different types of fluorescent dyes available consisting of different structural classes, sizes and charges. We have investigated these varying effects of on the apparent binding affinity associated with charge-charge interactions. We have designed a system consisting of two homopolypeptides that bind primarily through electrostatic interactions with a preferred orientation that allows us to place the dyes so that they come very close to each other during binding. By attaching fluorescent dyes to these peptides, we are able to compare the binding affinity for various dye pairs through FRET to quantify which of the various dye properties are the most important to consider. Through this approach, we were able to show that the charge of the dye can be as important as the size or structure of the dye, even though this property is not usually considered as strongly by researchers during dye selection. By using dyes too close together, you can decrease the apparent binding affinity as well as the positive cooperativity of your binding partners.

3.2. Introduction

Fluorescent dyes are used in a wide range of biological and biophysical applications. One of the most common uses involve biolabeling studies to track protein movement¹⁻³, visualize cell compartments^{4, 5}, study protein folding⁶⁻¹⁰, and measure binding between ligands and proteins¹¹. The most common structural motifs of fluorescent dyes used in biolabeling studies are coumarins, cyanines, xanthenes (fluorescein, eosin, and rhodamine), BODIPYs, oxazines and, more recently, carbopyronin, with each of the structural types having specific properties that determine how they can be used. Absorption and emission wavelength, molar absorptivity, quantum yield and photostability are the most important characteristics to determine suitability for a specific application. For example, it is known that solvent viscosity affects the quantum yield of cyanine dyes¹² and this property has been employed as a method to measure binding between biomolecules at very close distances¹³⁻¹⁵. For FRET studies, the emission wavelength and the quantum yield of the donor, together with the absorption wavelength and molar absorptivity of the acceptor, determine the sensitive distance range probed during biomolecular interaction studies. However, other dye properties such as size, molecular structure and charge could also be important in probing such interactions, although they are usually less carefully considered during selection. In order for accurate conclusions to be drawn from experimental data, all such contributions from the fluorescent dyes should be examined and quantified.

The use of fluorescent dyes attached to proteins allows one to measure binding coefficients between protein and ligand with high sensitivity. They are currently used to study DNA-ligand binding¹⁶, RNA targeting in cells¹⁷, and even heterogeneity of binding sites in stressed monoclonal antibody formulation¹⁸, among many others¹⁹. Currently in the literature, there is a very limited amount of papers that address the importance of dye charge and size

properties on experimental results. It has recently been shown that dye charge²⁰ and certain functional groups on a dye²¹ can have an effect on measured diffusion coefficients within a charged environment but little data exists as to how the charge affects biological binding, although some researchers have noted that using certain dyes can show different binding conformations depending on the protein/dye interaction under study²². Dyes can also cause steric hindrance when proteins are folding and force the protein to fold to an un-native conformation²³. It has also been shown that dyes that are in close proximity can quench each other, and mislead researchers with faulty data²⁴. In this study, we are particularly interested in quantifying the effect of the dye as it pertains to measuring the binding coefficient for interactions between charged residues, since such residues are important for long-range and short-range interactions in both structured proteins and intrinsically disordered proteins (IDPs).

In order to determine specific effects related to the structure, size, and charge of a fluorescent dye on charge-charge interactions, we have developed a system to change each of these parameters in a well-controlled biomolecular template. By using two polypeptides that bind in a parallel arrangement regardless of dye position (chapter 2), we place the dyes so that they are forced to come into close proximity of each other. One peptide contains 15 lysines (PolyK) that bind to a partner with 15 glutamic acids (PolyE). This system is well-suited for measuring the effects of different dyes on binding because these two peptides interact primarily through electrostatic interactions (chapter 2). By employing FRET, we compare the effects of various dye pairs on the binding coefficients for this peptide template system by changing the dyes on each of the peptides. We have determined that the choice of dye has a strong effect on the measured dissociation constant, K_D . If the bulky carbopyronin-based acceptor, Atto 633, which carries a +1 charge, is used, the measured affinity is significantly lower than the actual affinity (independent

of whether a +1 or -1 charged donor is used) but if the small, neutral BODIPY is used as the donor, the bulkiness and charge effect of the Atto 633 is negated. On the other hand, if the large but zwitterionic xanthene-based Atto590 is used as the acceptor, the measured K_D value is only slightly higher than the actual K_D value, and is independent of the size or charge of the donor. The results of this study shows that both the size and charge of the dyes must be considered when making decisions about which dye pair to use for quantitative FRET studies in situations when the dyes will be close to each other.

3.3. Materials and Methods

3.3.1. Peptide Labeling

Peptides with the sequence E15C and K14CK were ordered from Genscript to enable site-specific labeling using maleimide chemistry²⁵. E15C was labeled with the donor dye (termed PolyE) and K14CK was labeled with the acceptor dye (termed PolyK). Labeling was carried out using a 5:1 ratio of dye to peptide in 10mM phosphate buffer at pH 7.2. The mixture of peptide and dye, protected from light, were allowed to react overnight. The unreacted free dye was separated out by dialysis using 2 kDa cut-off dialysis tubing (Spectrum Labs). The samples were dialyzed against 10mM phosphate buffer for 24-48 hours with the external buffer being refreshed every 2-6 hours. The labeling efficiency was verified by MALDI mass spectrometry (Bruker Ultraflex II TOF/TOF time-of-flight mass spectrometer equipped with a MALDI ion source)(Bruker Daltonik GmbH, Bremen, Germany).

3.3.2. Fluorescence Resonance Energy Transfer (FRET) Spectroscopy

FRET measurements were performed on a PTI Quantamaster 40. The concentration of labeled PolyE was set at 5.0×10^{-8} M. The concentration of labeled PolyK varied from 5×10^{-10} to 1.5×10^{-6} M. The sample was excited at either 470 nm (Alexa 488 and Bodipy) or 515 nm (Cy3) and emission spectra were collected from 480 nm to 750 nm (Alexa 488 and Bodipy) or 525 nm

to 750 nm (Cy3). The energy transfer efficiency was calculated at each concentration of acceptor-labeled peptide added by using the relative fluorescence intensity of the donor peak in the emission spectrum in the presence (F_{DA}) and absence (F_D) of the acceptor using equation 1²⁶

$$E = 1 - (F_{DA}/F_D). \quad (1)$$

The K_D value was calculated by fitting the FRET data to the Hill equation (equation (2)) in Origin 8.1²⁷.

$$E = (E_{\max} * [\text{PolyK}]^n) / (K_D^n + [\text{PolyK}]^n) \quad (2)$$

3.4. Results

Figure 3.1 shows a schematic of the two labeled polypeptide chains as they bind. We showed in Chapter 2 that a parallel orientation is the favored interaction for the homopolypeptides between the charged side chains of glutamate and lysine. In this orientation, the donor and acceptor dyes are forced into close proximity of each other. The general structures of the dyes used in this study are shown in figure 3.2 and the source of each dye is given in table 3.1. The cyanine dyes are large with two fused ring structures capping a linear conjugated chain, leading to a moderate amount of structural flexibility. The xanthene dyes contain four 6-membered rings in a T-shape, together with several side groups on the rings and are much more rigid. The bodipy dye is much smaller than the others, containing a fused 6-membered and two 5-membered rings with no side groups.

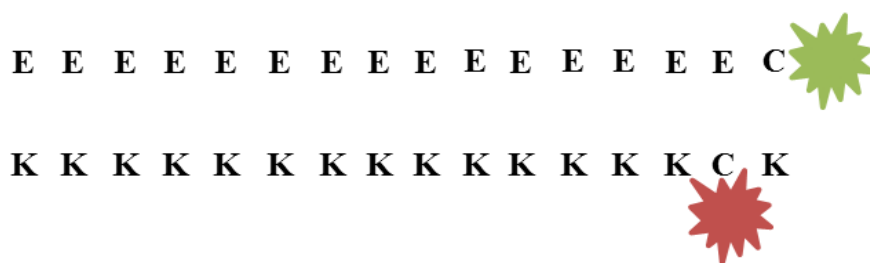


Figure 3.1. Schematic of homopolypeptides binding. The Poly E and Poly K peptides align parallel to each other so that the effect of the close proximity of the dyes can be evaluated.

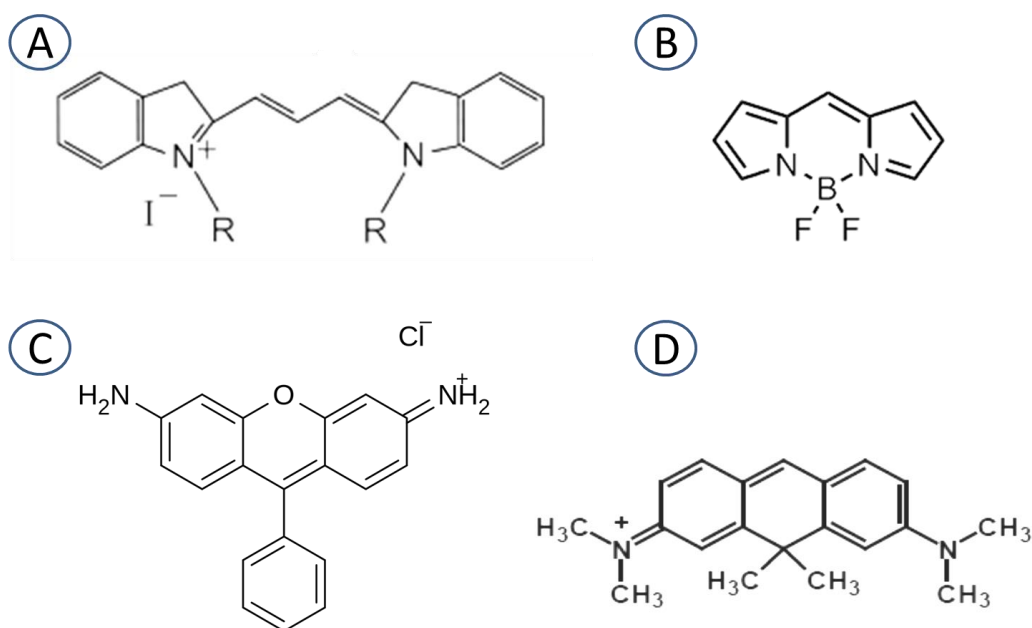


Figure 3.2. General structures of the dyes used. A) Cy3 is bulky but somewhat flexible. B) BODIPY is rigid but very small. C) The xanthene dyes (Alexa Fluor 488 and Atto 590) are very bulky and rigid. D) Carbopyronin dyes are also bulky and rigid (Atto 633).

Table 3.1. Full dye name, manufacturer, product number, and structural motif for all dyes used.

Dye Name	Company	Product Number	Structural Motif
Alexa Fluor 488	Molecular Probes	A-10254	Xanthene
Atto 590	Atto-Tec	AD 50-41	Xanthene
Atto 633	Atto-Tec	AD 633-41	Carbopyronin
Bodipy FL	Molecular Probes	B-10250	BODIPY
Cyanine3	Lumiprobe	11080	Cyanine

Using FRET, we measured the binding affinities between the two homopolypeptides labeled with different dyes and under different conditions. We varied the NaCl concentration from 0 mM to 250 mM at pH 7.2. Figure 3.3 shows a decrease in binding affinity as the concentration of salt is increased, as expected for electrostatic interactions (chapter 2). A similar trend is seen in all dye-pairs studied. The K_D values of all of the dye pairs that we examined are collated in Table 3.2. We have chosen dyes that we can pair together to investigate structural and size effects as well as charge effects. The salt concentration, Hill coefficients and maximum FRET values (E_{\max}) are also shown in Table 3.2.

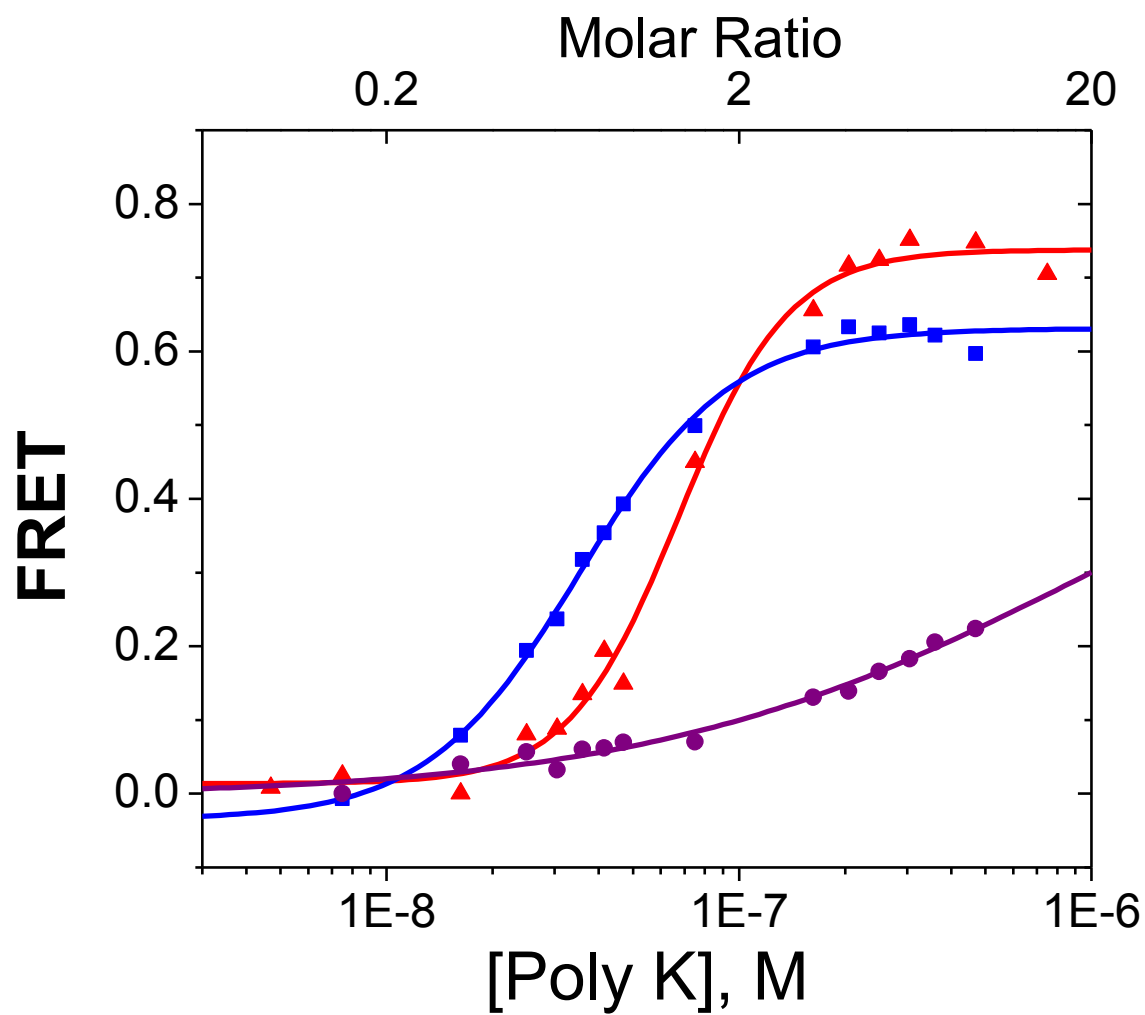


Figure 3.3. FRET curves at different NaCl concentrations: 0 mM (blue), 100 mM (red) and 250mM (purple). With increasing salt the slope of the curve decreases, consistent with the binding between the peptides being primarily electrostatic.

Table 3.2. Measured binding parameters for the different FRET dye pairs. Dyes carried either a +1 charge, -1 charge, were zwitterionic (+/-) or were neutral (0).

Dye Pair (D/A)	Donor Charge	Acceptor Charge	[NaCl]	K_D	Hill Coefficient	Maximum FRET
Alexa 488/Atto 633	-1	+1	0	253 +/- 7.2	4.4	.86
Alexa 488/Atto 633	-1	+1	100	350 +/- 22	2.0	.90
Cy3/Atto 633	+1	+1	0	348 +/- 13	3.0	.93
Cy3/Atto 633	+1	+1	100	578 +/- 24	1.3	.88
Alexa 488/Atto 590	-1	+/-	0	103 +/- 15	5.0	.80
Alexa 488/Atto 590	-1	+/-	100	48 +/- 5	2.2	.78
Cy3/Atto 590	+1	+/-	0	94 +/- 5	4.2	.78
Cy3/Atto 590	+1	+/-	100	65 +/- 4	2.6	.69
Alexa 488/ Cy3	-1	+1	0	67 +/- 3.5	2.8	.74
Alexa 488/ Cy3	-1	+1	100	35 +/- 1.4	2.0	.63
Bodipy/ Atto 633	0	+1	0	49 +/- 2.3	2.9	.79
Bodipy/ Atto 633	0	+1	100	90 +/- 5.3	3.1	.70
Bodipy/ Atto 590	0	+/-	0	85 +/- 3.1	4.0	.84
Bodipy/ Atto 590	0	+/-	100	75 +/- 3.9	4.2	.43

In figure 3.4, the differences in K_D values are plotted for each of the dye pairs at 0 mM and 100 mM NaCl. We also collected FRET data for all samples at 250 mM NaCl, but the much weaker binding between the polypeptides at this salt concentration precluded the data from being accurately fit (figure 3.3). Along the bottom axis, the data are sorted by the charge on the acceptor dye. The control K_D value was obtained by ITC on unlabeled polypeptides and found to be 12.1 nM at 0 mM NaCl and 107 nM at 100 mM NaCl (chapter 2). In the case of the acceptor dye being a zwitterion, the measured K_D values did not vary much between the differently charged donor dyes with $K_D = 103$ nM (donor with -1 charge), $K_D = 94$ nM (donor with +1

charge), and $K_D = 85$ nM (neutral donor) at 0 mM NaCl. This is only slightly larger than the control value. In contrast, when an acceptor dye with a +1 charge was used, the K_D values varied widely depending upon the charge of the donor. When the donor was neutral, the K_D value was comparable to the control with $K_D = 49$ nM at 0 mM NaCl. When the donor carried a +1 charge, $K_D = 348$ nM at 0 mM NaCl, while the donor carrying a -1 charge had $K_D = 253$ nM at 0 mM NaCl. When the concentration of NaCl was increased to 100 mM, the control K_D decreased by about 1 order of magnitude compared to that at 0 mM NaCl. At 100 mM NaCl, the zwitterionic acceptor showed K_D values that still did not depend on the charge on the donor, but did not become weaker at 100 mM NaCl, as the control did; $K_D = 65$ nM for the donor with a +1 charge and $K_D = 48$ nM for the donor with a -1 charge. However, when the +1 charged acceptor was used at 100 mM NaCl, the affinity did decrease compared to 0 mM, whether the donor carried a +1 ($K_D = 578$ nM) or a -1 charge (350 nM), although the factor of the decrease was not as much as for the control. With the +1 charged acceptor and the neutral donor $K_D = 90$ nM, effectively the same as the control.

Due to the fact that the effect of the charge on the dye showed some unexpected behavior, such as oppositely-charged dyes not always leading to an increased affinity and that increasing the NaCl concentration did not always decrease the affinity, the effect of the dyes' structure was also considered as a potential source for the effect. Depending on commercial availability, one can keep the charge on the acceptor as +1 but use different structural motifs. In the case of the -1 charged donor (the xanthene-based Alexa Fluor 488), we compared a +1 charged acceptor with either the very similarly structured carbopyronin-based structure of Atto 633 or the somewhat differently structured cyanine-based structure of Cy3. When the smaller, slightly more flexible cyanine dye was used as an acceptor the K_D value at 0 mM NaCl was 67

nM compared to 235 nM for the larger, more rigid carbopyronin dye acceptor. At 100 mM NaCl, the K_D values for the cyanine and carbopyronin acceptors were 35 nM and 350 nM, respectively, effectively the same as at 0 mM NaCl.

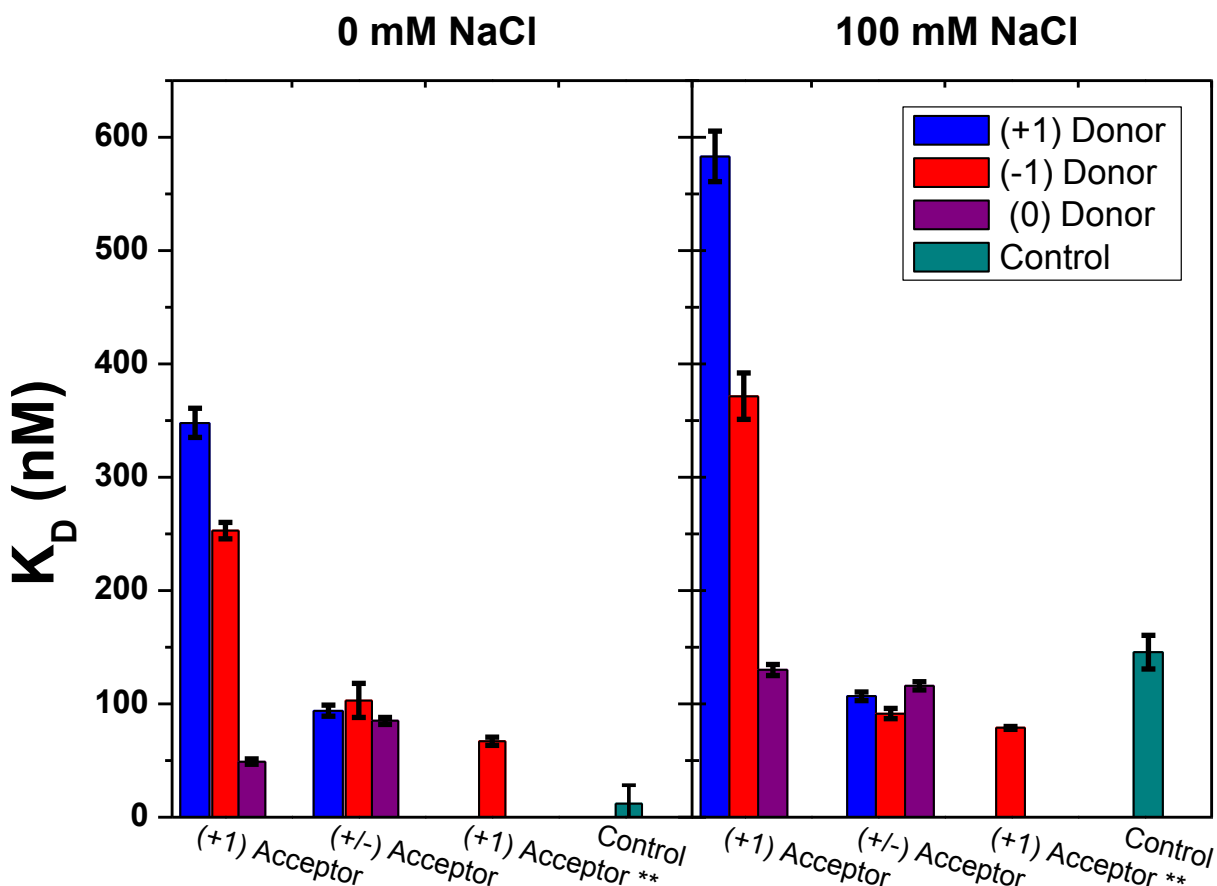


Figure 3.4. K_D values for dye pairs as a function of charge. The +1 charged acceptor was either the rigid Atto 633 or the slightly flexible Cy3 (denoted by **), whereas the +1 charged donor was only Cy3 (see table 1). The control data was obtained by ITC without dyes²⁹.

Figure 3.5 highlights the difference in the Hill coefficients based upon the charges on the dye pairs and the salt concentration. When the acceptor is zwitterionic, the Hill coefficient does not depend on the charge of the donor, the Hill coefficient stays high, between 4 and 5 for all

donors. The Hill coefficient was also similarly high for the -1/+1 D/A dye pair as well. When the acceptor has a +1 charge and the donor's charge is either neutral or +1, the Hill coefficients are a little lower at ~3. Clearly, at 0 mM NaCl, the Hill coefficient is >1 for all dye pairs. In all but one case, the Hill coefficient decreases to ~2 as the concentration of salt increases to 100 mM. The dye pair where the donor is neutral and the acceptor has a +1 charge does not decrease at 100 mM NaCl, remaining at ~3.

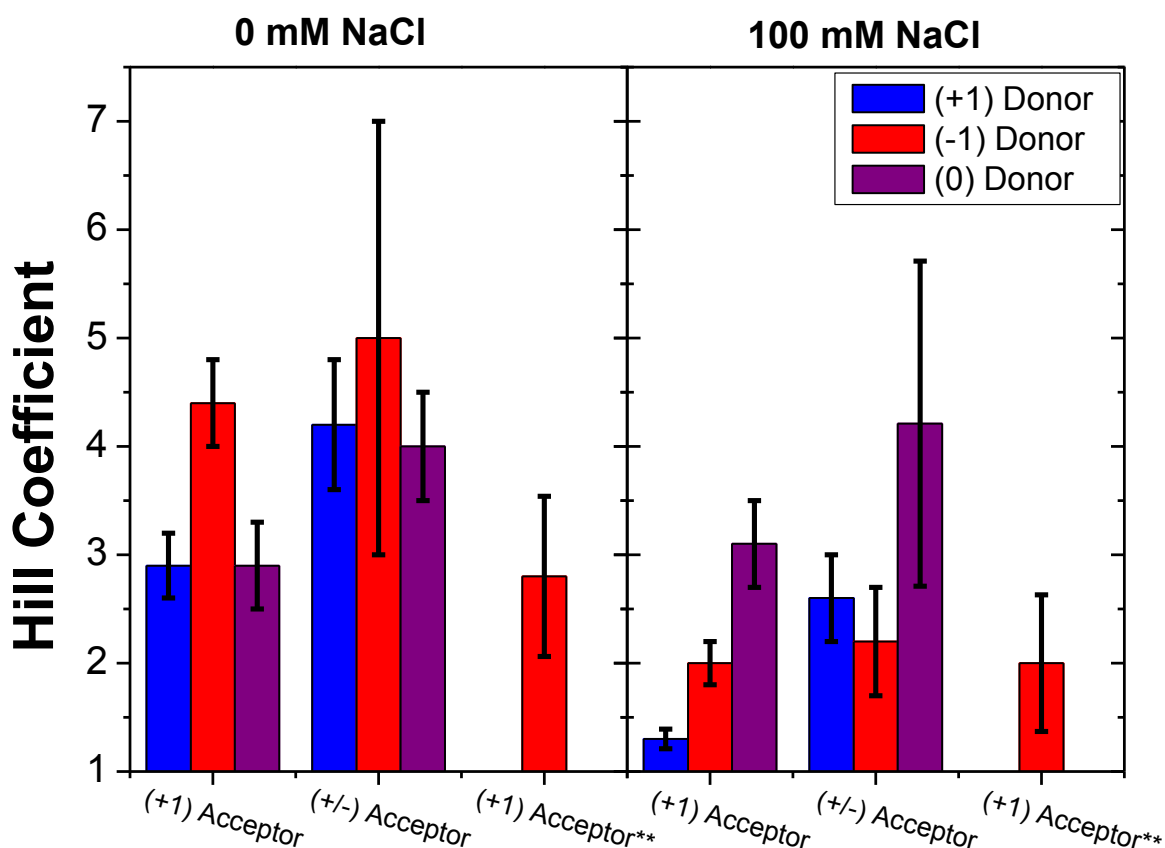


Figure 3.5. Comparison of Hill coefficient for every dye pair with varying NaCl concentration. ** denotes when the +1 charged acceptor was Cy3 rather than Atto 633.

Finally, we quantify the maximum change in FRET (E_{\max}) at the saturating concentration of acceptor-labeled PolyK for each of the dye pairs at 0 mM and 100 mM NaCl in figure 3.6. All of the E_{\max} values at 0 mM NaCl are high (above 0.7), with most between 0.8 and 0.9. Interestingly, the D/A pair consisting of +1 xanthene/+1 xanthene shows the highest E_{\max} value, while the D/A pair consisting of +1 xanthene/+1 cyanine shows the lowest E_{\max} value. All the other dye pairs show effectively the same E_{\max} value (between 0.80 and 0.85) independent of dye charge and structure. At 100 mM NaCl, the E_{\max} values are still high (above 0.6), but show slightly more variability.

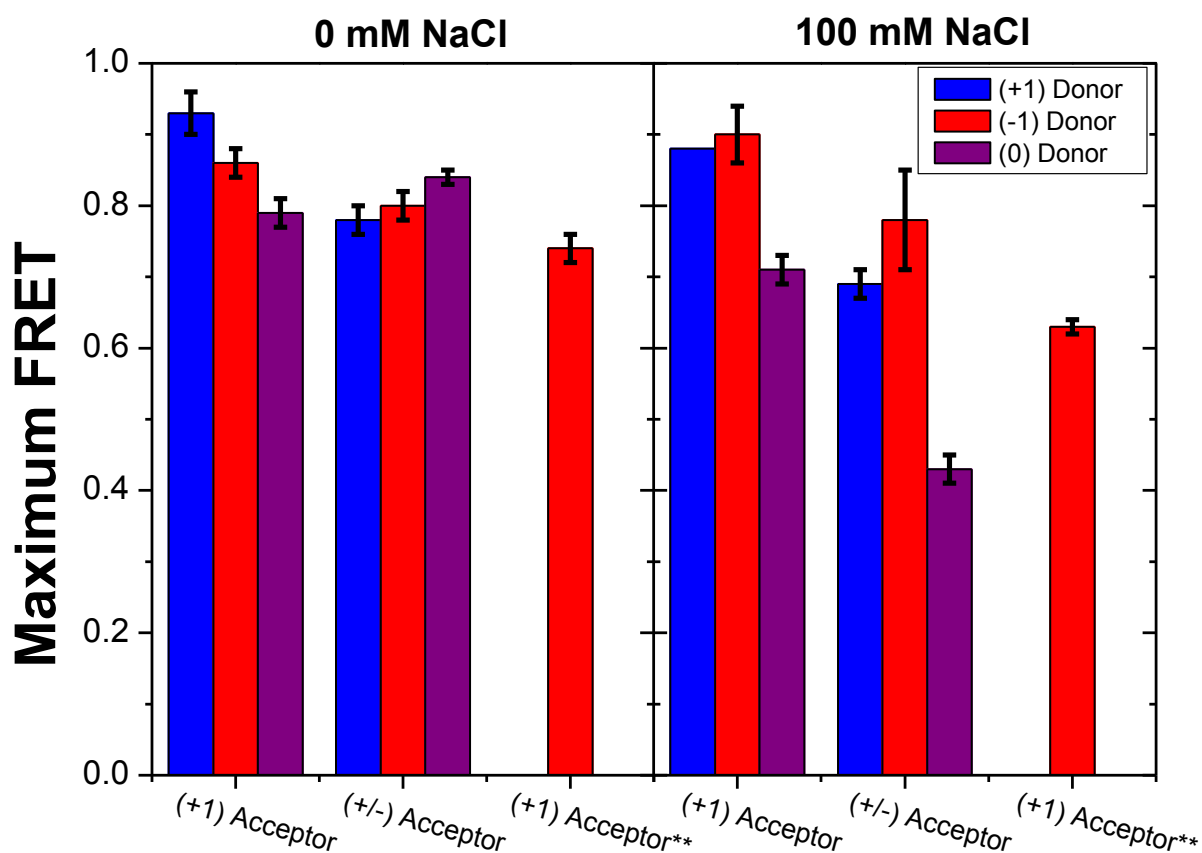


Figure 3.6. Maximum FRET value for dye pairs. ** denotes when the +1 charged acceptor was Cy3 rather than Atto 633.

3.5. Discussion

The central focus for this work was to determine if the choice of fluorescent dyes that are commonly used in FRET-based binding studies affects the results when electrostatic interactions are being probed. We purposely chose conditions in which the dyes are forced into close proximity of each other to determine the worst-case scenario. We have developed a homopolypeptide template system that forces this condition. It is clear from our results that, when the fluorescent labels come into close proximity of each other, they do have an effect on the measured K_D values, although the effects are not as clear cut as one might anticipate. They depend on the size, charge, and structure of the dye label in varying degrees, which further varies with the ionic strength of the solution in which binding is studied. The control K_D value for the binding of our peptides (PolyE + PolyK) was obtained by ITC, without using fluorescent labels, (Chapter 2) to enable absolute effects rather than relative effects to be quantified. At 0 mM NaCl, all of the K_D values that were quantified using fluorescent labels showed weaker binding than the control although some dye combinations has only a marginal effect. The decrease in binding most likely arises from steric hindrance from the dyes interfering with the peptide interaction. The degree of the steric effect depends on both the size/rigidity of the dye as well as the charge. The peptides still bind very tightly ($K_D < 100$ nM in most cases) because there are still effectively 15 charges on each peptide that attract each other. If the commonly used carbopyronin-based acceptor, Atto 633, which carries a +1 charge, is used, the measured affinity is significantly lower than the actual affinity (control value) when either the +1 charged Cy3 or the -1 charged Alexa Fluor 488 is used as the donor, with the +1/+1 dye combination being ~40% weaker than the +1/-1 dye combination, even though Cy3 is more flexible than Alexa Fluor 488. However, if the small, neutral BODIPY is used as the donor, the bulkiness and charge of the Atto 633 is negated, with the measured K_D value being almost the same as the actual K_D . On the other hand,

if the large, rigid but zwitterionic xanthene-based Atto590 is used as the acceptor, the measured K_D value is only a little higher than the actual K_D value, and is independent of the size or charge of the donor. For the situation when commercial availability allowed us to keep the charge on both the donor and acceptor the same (but complementary to each other) but change the dye structural motif, the bulkiness/rigidity effects can be studied. Comparing Alexa Fluor 488 (-1)/Atto 633(+1) vs Alexa Fluor 488 (-1)/Cy3(+1), we can see that the Alexa Fluor 488/Cy3 pair was much closer to the actual K_D value, being very similar to the cases in which the zwitterionic xanthene-based acceptor was used. Clearly the complementary dye charge is not enough to overcome the bulkiness of the carbopyronin-based Atto 633, whereas it is enough for the cyanine dye. It is possible that the flexibility of the cyanine dye allows it to be pushed into a position to allow the homopolypeptides to bind in a more natural way.

When 100 mM NaCl is added to the buffer, the actual K_D value for the homopolypeptide binding is increased, as expected for electrostatically-mediated interactions (chapter 2). However, only the situations in which the carbopyronin-based Atto 633 was used as the acceptor showed an increase in measured K_D , and these were still much higher than the actual K_D value at 100 mM – except when BODIPY was used as the donor. In this case, BODIPY again allowed us to measure the actual K_D value, as was the case at 0 mM NaCl. Clearly, BODIPY is an excellent choice of dye when dyes must be placed close to each other. Interestingly, when the zwitterionic Atto 590 or the cyanine Cy3 was used as the acceptor, no effect of increasing NaCl to 100 mM was observed. Since the K_D value for these situations in 0 mM NaCl was very similar to the actual K_D value at 100 mM, perhaps the increased salt concentration negated the remaining dye effect, although more experiments will be needed to confirm this possibility.

The Hill coefficient is a measure of positive interactions. For simple, sequential binding, the Hill coefficient (n) should be less than 2. In all cases at 0 mM NaCl, $n > 2$. This suggests that the binding of the peptides is not a simple model in the sense that 2 peptides come together and bind. What is more likely happening is small clusters of peptides are forming. The Hill coefficient values behave proportionally to the K_D values, i.e. as the K_D value shows weaker binding, the n value shows less positive cooperativity. These effects are from the size of the dye as well as the charge. When the bulky dye causes steric hindrance, it not only reduces the binding affinity but it also stops the peptides from forming the tightly bound small complexes. The charges act in the same way, the repelling charges won't allow the peptides to bind as tightly as they would otherwise.

Under the conditions of 100 mM NaCl, the n values are lower which is to be expected since the peptides interact through electrostatic interactions. In some cases, $n < 2$, mainly when the K_D value was very high. The same conditions apply here as in the 0 mM NaCl conditions, the dye is keeping the peptides from binding and forming the complexes that increase the Hill coefficient. When the BODIPY dye is used however, the n value stays very close to the same regardless of the salt concentration just like with the binding affinity. This neutral dye has the minimal effect on the binding and stacking of the peptides.

When analyzing the E_{\max} values, a clear trend becomes apparent for both the 0 mM and the 100 mM NaCl experiments. In the peptide complexes that bind less tightly and have less positive cooperativity, the maximum FRET value is higher. Since the FRET is a measure of quenching, it can be presumed that the large complexes reduce the apparent quenching, by either peptides donating electrons to increase the quantum yield of the donor or by the

closeness/number of dyes present. Further experiments are needed to accurately ascertain which phenomenon is occurring.

3.6. Summary and Conclusions

With all of the fluorescent dye choices available, it is important to consider the system in which the dyes will be used. In FRET studies, the dyes are quite often placed near a binding site in order to get high FRET values upon the binding of a ligand. We have shown that putting the dyes in an arrangement so that they come in very close contact with one another can be detrimental to the accuracy of the results. All parameters explored in this study, binding affinity and cooperativity, were affected by the nearness of the donor and acceptor dyes. In instances where the experimental design does not allow for options other than extremely close dye placements, it would be advisable to use small, neutral dyes to mitigate the effects the dyes cause, like decrease in binding and loss of positive cooperativity. At this point, the decrease in the E_{\max} value is not clear and future work will need to be performed to help answer this question.

3.7 References

1. Hildick, K. L., Gonzalez, G., Iez, I. M., Jaskolski, F., Ric, and Henley, J. M. (2012) Lateral Diffusion and Exocytosis of Membrane Proteins in Cultured Neurons Assessed using Fluorescence Recovery and Fluorescence-loss Photobleaching, *e3747*.
2. Niu, C., Anstead, J., and Verchot, J. (2012) Analysis of protein transport in the Brassica oleracea vasculature reveals protein-specific destinations, *Plant Signaling & Behavior*7, 361-374.
3. Prevo, B., and Peterman, E. J. G. (2014) Forster resonance energy transfer and kinesin motor proteins, *Chemical Society Reviews*43, 1144-1155.
4. Zhang, J., Campbell, R. E., Ting, A. Y., and Tsien, R. Y. (2002) Creating new fluorescent probes for cell biology, *Nat Rev Mol Cell Biol*3, 906-918.
5. Spring, O., and Krauß, W. (1987) In vivo fluorescence labeling of plasma membrane from plant tissue and cells, *Plant Science*48, 203-213.
6. Rhoades, E., Gussakovsky, E., and Haran, G. (2003) Watching proteins fold one molecule at a time, *Proceedings of the National Academy of Sciences*100, 3197-3202.
7. Uversky, V., and Dunker, A. (2010) Understanding protein non-folding, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*1804, 1231 - 1264.
8. Schuler, B., and Eaton, W. A. (2008) Protein folding studied by single-molecule FRET, *Current Opinion in Structural Biology*18, 16-26.
9. Schuler, B., Lipman, E. A., and Eaton, W. A. (2002) Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy, *Nature*419, 743-747.
10. Kuzmenkina, E. V., Heyes, C. D., and Nienhaus, G. U. (2005) Single-molecule Förster resonance energy transfer study of protein dynamics under denaturing conditions, *Proceedings of the National Academy of Sciences of the United States of America*102, 15471-15476.
11. Ferreon, A. C. M., Ferreon, J. C., Wright, P. E., and Deniz, A. A. (2013) Modulation of allostery by protein intrinsic disorder, *Nature (London, U. K.)*498, 390-394.
12. Aramendia, P. F., Negri, R. M., and Roman, E. S. (1994) Temperature Dependence of Fluorescence and Photoisomerization in Symmetric Carbocyanines. Influence of Medium Viscosity and Molecular Structure, *The Journal of Physical Chemistry*98, 3165-3173.
13. Hwang, H., and Myong, S. (2014) Protein induced fluorescence enhancement (PIFE) for probing protein-nucleic acid interactions, *Chemical Society Reviews*43, 1221-1229.

14. Marko, R. A., Liu, H.-W., Ablenas, C. J., Ehteshami, M., Götte, M., and Cosa, G. (2013) Binding Kinetics and Affinities of Heterodimeric versus Homodimeric HIV-1 Reverse Transcriptase on DNA–DNA Substrates at the Single-Molecule Level, *The Journal of Physical Chemistry B*117, 4560-4567.
15. Hwang, H., Kim, H., and Myong, S. (2011) Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity, *Proceedings of the National Academy of Sciences*108, 7414-7418.
16. Dupureur, C. M., Bashkin, J. K., Aston, K., Koeller, K. J., Gaston, K. R., and He, G. (2012) Fluorescence assay of polyamide–DNA interactions, *Analytical Biochemistry*423, 178-183.
17. Kim, M. Y., Kim, J., and Hah, S. S. (2012) Poly(A)-targeting molecular beacons: Fluorescence resonance energy transfer-based in vitro quantitation and time-dependent imaging in live cells, *Analytical Biochemistry*429, 92-98.
18. Hawe, A., Rispen, T., Herron, J. N., and Jiskoot, W. (2011) Probing bis-ANS binding sites of different affinity on aggregated IgG by steady-state fluorescence, time-resolved fluorescence and isothermal titration calorimetry, *Journal of Pharmaceutical Sciences*100, 1294-1305.
19. Ebrecht, R., Don Paul, C., and Wouters, F. (2014) Fluorescence lifetime imaging microscopy in the medical sciences, *Protoplasma*251, 293-305.
20. Daniels, C. R., Reznik, C., and Landes, C. F. (2010) Dye Diffusion at Surfaces: Charge Matters, *Langmuir*26, 4807-4812.
21. Ranjit, S., and Levitus, M. (2012) Probing the Interaction Between Fluorophores and DNA Nucleotides by Fluorescence Correlation Spectroscopy and Fluorescence Quenching†, *Photochemistry and Photobiology*88, 782-791.
22. DeVore, M. S., Gull, S. F., and Johnson, C. K. (2013) Reconstruction of calmodulin single-molecule FRET states, dye interactions, and CaMKII peptide binding by MultiNest and classic maximum entropy, *Chemical Physics*422, 238-245.
23. Nettels, D., Müller-Späth, S., Küster, F., Hofmann, H., Haenni, D., Rügger, S., Reymond, L., Hoffmann, A., Kubelka, J., Heinz, B., Gast, K., Best, R. B., and Schuler, B. (2009) Single-molecule spectroscopy of the temperature-induced collapse of unfolded proteins, *Proceedings of the National Academy of Sciences*106, 20740-20745.
24. Di Fiori, N., and Meller, A. (2010) The Effect of Dye-Dye Interactions on the Spatial Resolution of Single-Molecule FRET Measurements in Nucleic Acids, *Biophysical Journal*98, 2265-2272.
25. Hermanson, G. (2013) *Bioconjugate Techniques*, 3rd ed., Academic Press.

26. Lakowicz, J. R. (2006) *Principles of fluorescence spectroscopy*, 3rd ed., Springer, New York.
27. Weiss, J. N. (1997) The Hill equation revisited: uses and misuses, *FASEB J.* *11*, 835-841.
28. Dey, D., Saha, J., Bhattacharjee, D., and Hussain, S. A. (2014) Development of an ion-sensor using Fluorescence resonance energy transfer, *arXiv.org, e-Print Arch., Condens. Matter*, 1-29, arXiv:1408.6556v1401 [cond-mat.mtrl-sci].
29. Webb, N. (2011) Conformational Studies of Homopolypeptides, University of Arkansas.

IV. Insights on the Quenching of Some Fluorescent Dyes Commonly Used in Biolabeling by Charged Amino Acid Residues

4.1. Abstract

The use of fluorescent dyes in biolabeling studies has greatly increased in the past several years. One very popular method of measuring protein binding using fluorescent dyes is by FRET since it can be used to extract structural details due to its well-known distance dependence. We have determined that changes in the immediate environment of the dyes caused by binding of charged residues can affect the quantum yield of these dyes via quenching which will affect the extracted structural parameters. The effects of the charged environment are also determined by the charge of the dye used. If the fluorescent dyes must be used in the very near vicinity of charged residues, very careful consideration should be used in selecting the dye and the method of measurement.

4.2. Introduction

Fluorescent dyes have been used for many years for a myriad of different applications from protein labeling and assays with DNA to individual proteins measured on a single molecule basis¹⁻⁴. There is such a broad range of fluorescent dyes available and it is important to know which will best suit the needs of the system being probed. Several reviews have been published that describe various aspects of available fluorescent dyes and their uses⁵⁻⁷. Some of the most commonly used dyes include rhodamines, which have been used for labeling human cardiac tissue for use in imaging⁸, BODIPY's, which have been used to probe lipids in microalgae⁹, and coumarins, fluoresceins, and oxazines that are used for fluoride sensing in cellular systems¹⁰. All of these dye types are popular because they have high stability and long fluorescent lifetimes^{11, 12}.

Sometimes it is necessary to use fluorescent dyes in environments that are not the most ideal; e.g. near to highly charged regions of a protein or near to a binding site. Since most dyes have a charge themselves, the combination of these factors could pose a problem for obtaining accurate results. Others have started to investigate the effects of fluorescent dyes on results, from the charge of the dye interacting with glass surfaces¹³, to trying to distinguish between observable protein folding and dye photophysical induced data¹⁴. Dye effects have also been noted in studies of nucleic acids¹⁵. In microarray work, dye bias can lead to over estimations in expression levels¹⁶.

Our lab has been particularly interested in studying charge-charge interactions in biomolecules (chapters 2 and 3), and we have been developing fluorescence-based assays to study the binding affinity and orientation between oppositely-charged homopolypeptides. We have also studied the effect of putting differently sized and charged dyes very close to each other

and the interaction. The results of that study concluded major differences in binding affinities due to the dyes' interaction. In those studies, the reduction in the fluorescence intensity of a donor dye was used as a probe for interaction between oppositely-charged homopolypeptides. However, we did not expand upon which quenching mechanisms were responsible for the fluorescence decrease. The affinity can be determined without such knowledge, but in order to gain additional information from these experiments, such as interfluorophore distances that can be used to extract structural details, it is necessary to determine which mechanisms are responsible for quenching. For example, quenching by FRET to an acceptor fluorophore has a different distance dependence than by electron transfer to/from electron acceptor/donor residues. This study aims at addressing the relative contribution of FRET compared to these other quenching mechanisms. We found that the near-vicinity of charged residues can affect the quantum yield of a dye, which depends upon the charge of the dye, the charge on the residue, as well as the distance between the two. In some measurements, the quantum yield of the dye will not affect the results of the study; however when using a technique such as FRET, quantifying the intensity of the dye in the environment that it is placed is very important. When using a method such as donor-quenching FRET, the intensity of the donor's dye is used to measure the binding of the two partners, either protein-protein or protein-ligand. If the quantum yield (QY) of the dye changes at any time in the experiment, then the results can show either an increase in binding (for a decrease in QY) or a decrease in binding (increase in QY). We have developed a system that allows for direct comparisons of parameters changed in order to determine the exact effects of charges on fluorescent dye QY as well as how this changes apparent binding affinity.

4.3. Materials and Methods

4.3.1. Peptide Labeling

Peptides with the sequence E15C, E6CE9 and K14CK were ordered from Genscript to enable site-specific labeling using maleimide chemistry. E6CE9 was labeled with the donor dye Alexa Fluor 488 (labeled polyE*). The E15C and K14CK were labeled with either a donor dye or acceptor dye (labeled polyE and polyK). Labeling was carried out using a 5:1 ratio of dye to peptide in 10 mM phosphate buffer at pH 7.2. The mixture of peptide and dye, protected from light, were allowed to react overnight. The unreacted free dye was separated out by dialysis using 2 kDa molecular weight cut-off dialysis tubing (Spectrum Labs). The samples were dialyzed against 10 mM phosphate buffer for 24-48 hours with the external buffer being refreshed every 2-6 hours. The samples were further cleaned up using 3 kDa spin columns (Millipore). The labeling efficiency and separation of unreacted dye was verified by MALDI mass spectrometry (Bruker Ultraflex II TOF/TOF time-of-flight mass spectrometer equipped with a MALDI ion source) (BrukerDaltonik GmbH, Bremen, Germany).

4.3.2. Fluorescence Spectroscopy

Fluorescence measurements that were carried out in the absence of an acceptor dye were performed on a Perkin Elmer LS55. The concentration of labeled peptide was set at 1.0×10^{-6} M. The unlabeled peptide concentration was 1.67×10^{-6} M to 1.65×10^{-8} M. The fluorescence measurements that did include acceptor dyes were performed on a PTI Quantamaster 40. The concentration of labeled polyE was set at 5.0×10^{-8} M. The concentration of labeled polyK varied from 5×10^{-10} M to 1.5×10^{-6} M. For all experiments the sample was excited at a wavelength depending upon the dye used (Table 4.1). Emission data was collected 10 nm from the excitation wavelength up to 750 nm. The transfer/quenching efficiency, E, was calculated by using the relative fluorescence intensity of the donor at the maximum intensity in the presence (F_{DA}) and absence (F_D) of the oppositely charge polypeptide¹⁷ using: $E = 1 - (F_{DA}/F_D)$. The K_D value was calculated by fitting to the Hill¹⁸ equation in Origin.

Table 4.1. Excitation and detection wavelengths as well as charges of dyes used.

Donor Dye	Excitation (nm)	Detection (nm)	Charge
Alexa 488	470	480-750	-1
Bodipy	470	480-750	0 (neutral)
Cy3	515	525-750	+1
Alexa 590	540	550-750	+/- (zwitterionic)
Alexa 568	580	580-750	-1
Atto 633	610	610-750	+1

4.4. Results and Discussion

We have investigated the FRET efficiency vs quenching by highly charged polypeptides on attached fluorescent dyes. The magnitude of difference between the two mechanisms depends greatly on the charge of the peptide, the charge of the dye, and the location of the dyes on the peptide. When a fluorescent label was placed on a positively charged homopolypeptide, the binding of a highly charged negative peptide significantly quenched the dye by 35-75%, figure 4.1. We believe this quenching is caused by the electrons being transferred from the excited state of the fluorophore to the polyK peptide (figure 4.2) with an electron being transferred from the polyE peptide back to the ground state of the fluorophore. This process is much more efficient when the dye is placed on the polyK peptide vs placement on polyE. We believe this is due to the distance that the excited electron has to travel in order to be transferred. The shorter distance to the polyK means a quicker and more efficient transfer. When the dye is on the polyE, it is farther away from the electron acceptor. In the case of the donor dye being in the middle of the polyE chain, the fluorescence is quenched almost as efficiently as when the dye is attached

directly on the polyK peptide. This drastic increase is likely due to the fact that there are now multiple residues in the near vicinity of the dye to accept the excited electrons.

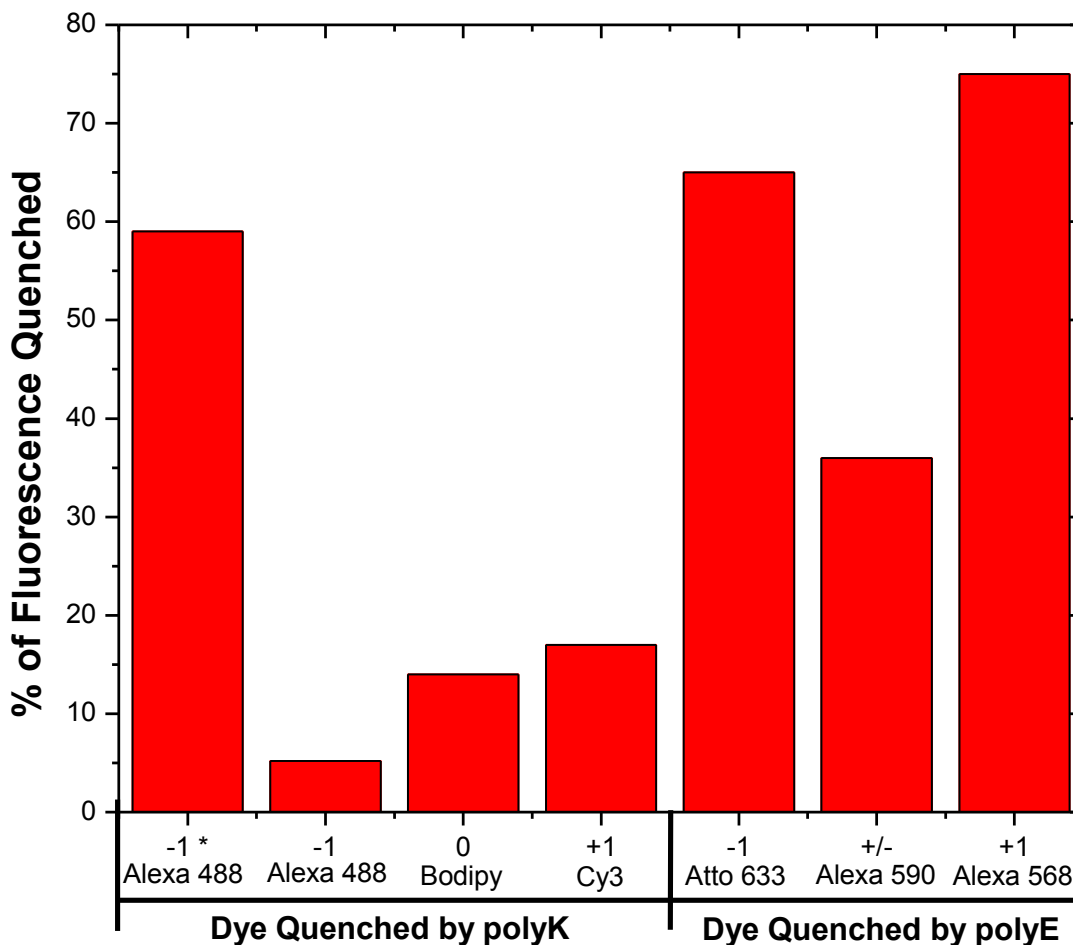


Figure 4.1. Percent of fluorescence quenched by charged peptides. On left side, the dye is attached to polyE peptide and the addition of polyK decreases the signal. On the right side, the dye is attached to the polyK peptide and the addition of polyE decreases the signal. The * represents the polyE peptide where the cysteine is in the middle of the chain instead of at the end.

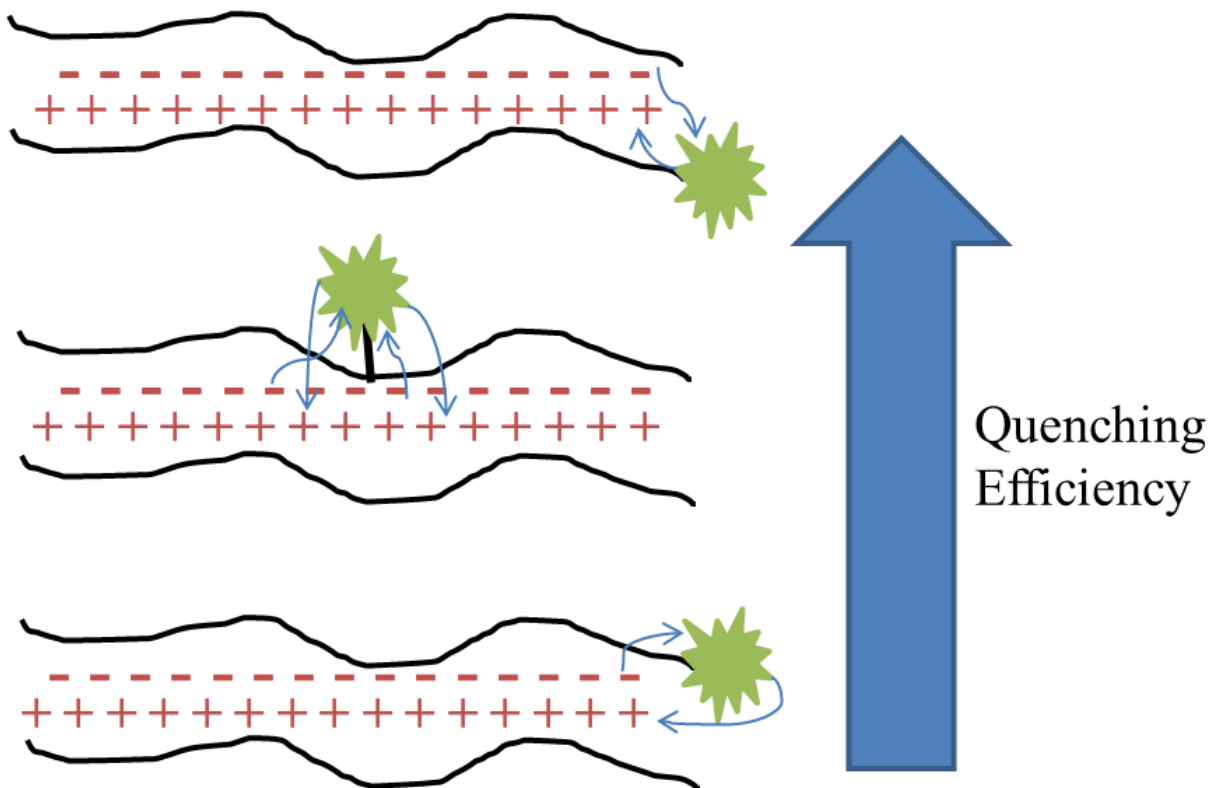


Figure 4.2. Schematic of peptide binding with relative quenching efficiency. A) The dye is attached to the polyK peptide. An electron is easily transferred from the dye to the positively charged peptide. PolyE gives up an electron in the ground state to the dye. B) The dye is attached in the middle of the polyE peptide. The bound polyK takes an electron from the dye; the polyE donates an electron back to the dye. Here the distance of travel is longer for the electron, but there are multiple pathways. C) PolyE has the dye attached at the C-terminus. The polyK takes an electron and the polyE donates one back. Here the distance is as long as in the middle example, but there is only one possible peptide to except the charge.

Conjugating the Cy3 dye to the polyE peptide increases the quantum yield (figure 4.3) of the dye by 2.4 times from 68% (free in solution) to 100%. This increase in brightness may be the reason why the polyK does not appear to quench the dye, figure 4.1. In contrast, conjugation of the dye to the polyK reduces the quantum yield from 68% to 42%. The change in quantum yields are from the electron donating (polyE) or electron accepting (polyK) properties of the peptides that the dyes are attached to. When there are an excess of electrons, the quantum yield increase

to the maximum amount while the electron acceptor reduces the quantum yield by removing the excited electrons before they have time to fully relax back to the ground state through the normal fluorescence pathway.

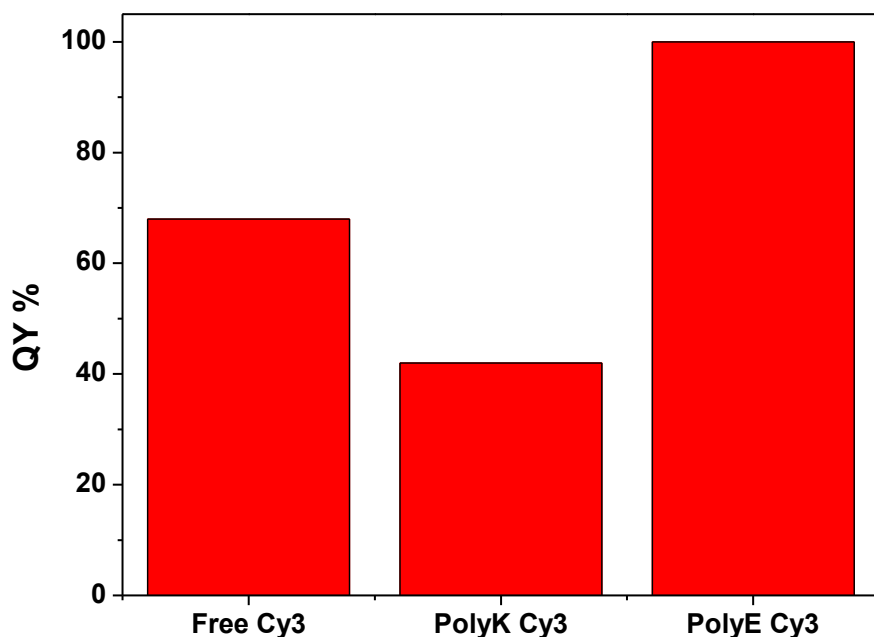


Figure 4.3. Quantum yield measurements for Cy3. The quantum yield of Cy3 was measured for dye alone, dye attached to polyK, and dye attached to polyE.

The maximum transfer/quenching efficiency was examined to determine the relationship between the energy transfer and electron transfer mechanisms, figure 4.4. In all cases studied, the energy transfer results in a higher efficiency and greater reduction in donor fluorescence than when the electron transfer quenching only is present, which is as expected when there are 2 processes occurring at the same time vs only a single quenching process. Since energy transfer (FRET) occurs at longer distances with weaker distance dependence than the electron transfer of standard quenching, it is not surprising that it gives a much larger signal change when comparing

the unlabeled polyK to labeled polyK. In the case where the donor dye was placed in the middle of the peptide instead of at the end, the energy transfer and electron transfer quenching both gave very similar results with the electron transfer only being slightly more efficient. The energy transfer in this situation is lower than when the dyes are both placed at the end because of the longer distance between the two dyes. The electron transfer is higher than the other samples where polyK is quenching the dye but lower than those where the polyE is the quencher (figure 4.2). We expect this is due to the peptide orienting itself in such a way that the dye is close to the polyK peptide but not as close as it would be if the dye was conjugated to the peptide, figure 4.4. In the bottom of figure 4.4, the energy transfer as well as electron transfer graphs are shown for the Cy3 sample. The K_D value is obtained by finding the midpoint of the curve. Both of these curves have very similar midpoints suggesting the K_D value is the same regardless of the method used to measure binding.

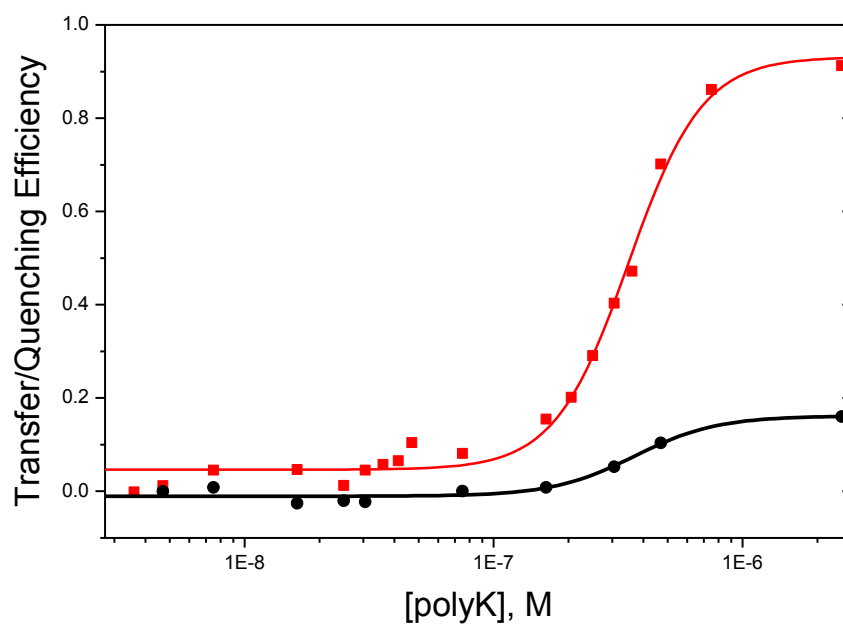
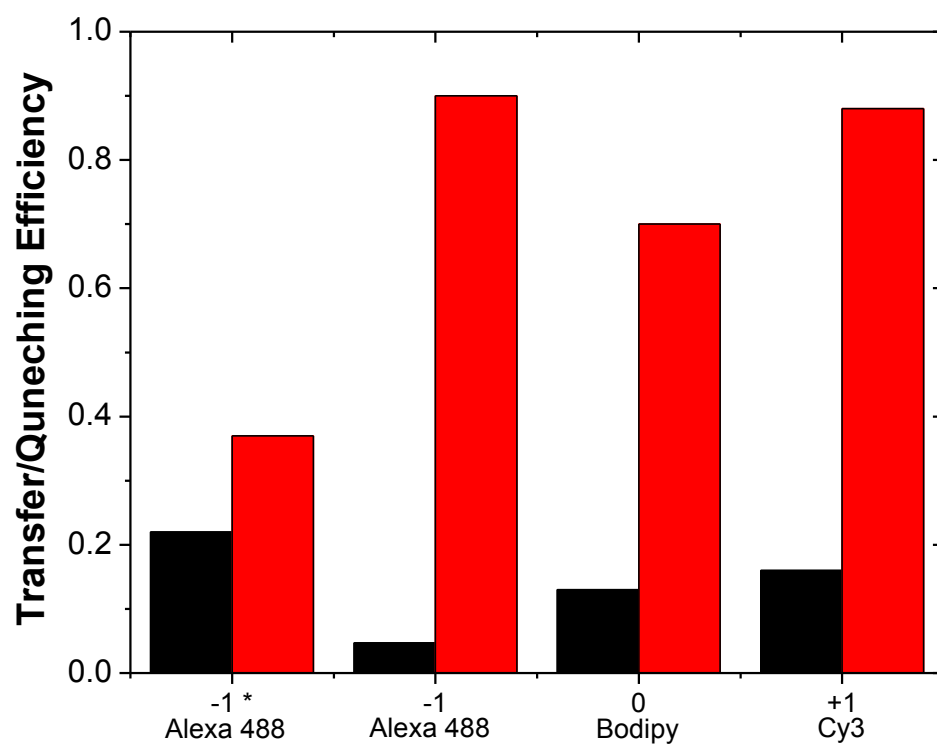


Figure 4.4. The maximum transfer/quenching efficiency is compared between labeled and unlabeled samples sorted by charges. **A)** Labeled acceptor peptide in red, unlabeled acceptor peptide in black. In all labeled cases, the acceptor is Atto 633. **B)** Transfer/Quenching curves are shown for Cy3 with labeled acceptor (red) unlabeled acceptor (black).

In figure 4.5, the measured binding constants are presented for each of the dye pairs depending upon the charge and whether or not the acceptor peptide has a dye. The K_D values for labeled vs unlabeled polyK for each pair are equal or very close to equal for all pairs except when the -1 charged dye is placed at the end of the peptide chain. When the dye was in the middle of the peptide the binding constant that was measured through energy transfer and electron transfer were both very close ($K_D = 54 \text{ nM} \pm 24 \text{ nM}$ vs $K_D = 59 \text{ nM} \pm 24 \text{ nM}$, respectively) and comparable to that of the control (measured using ITC where neither peptide is labeled, $K_D = 12 \text{ nM}$). When the +1 charged dye was used, the K_D values for both labeled and unlabeled acceptor peptide were very close to each other, $K_D = 348 \text{ nM} \pm 13 \text{ nM}$ vs $368 \text{ nM} \pm 42 \text{ nM}$, but were over an order of magnitude larger than the control, signifying a large effect from the donor dye. When the neutral dye was used, the unlabeled sample measured a slightly higher K_D value than the labeled sample $K_D = 160 \text{ nM} \pm 15 \text{ nM}$ vs $90 \text{ nM} \pm 6 \text{ nM}$, respectively. It would appear that the neutral dye is somehow reducing the binding, either sterically or through electrostatics. These results are all different from the -1 charged dye pair that show a tighter binding or lower K_D value when the acceptor dye is not present, $K_D = 67 \text{ nM} \pm 10 \text{ nM}$ vs $350 \text{ nM} \pm 22 \text{ nM}$. This increase in K_D may come from the acceptor dye as well as the donor dyes being repelled from their respective peptides and causing the peptides themselves to not be able to bind as they would if the interfering dyes were not attached.

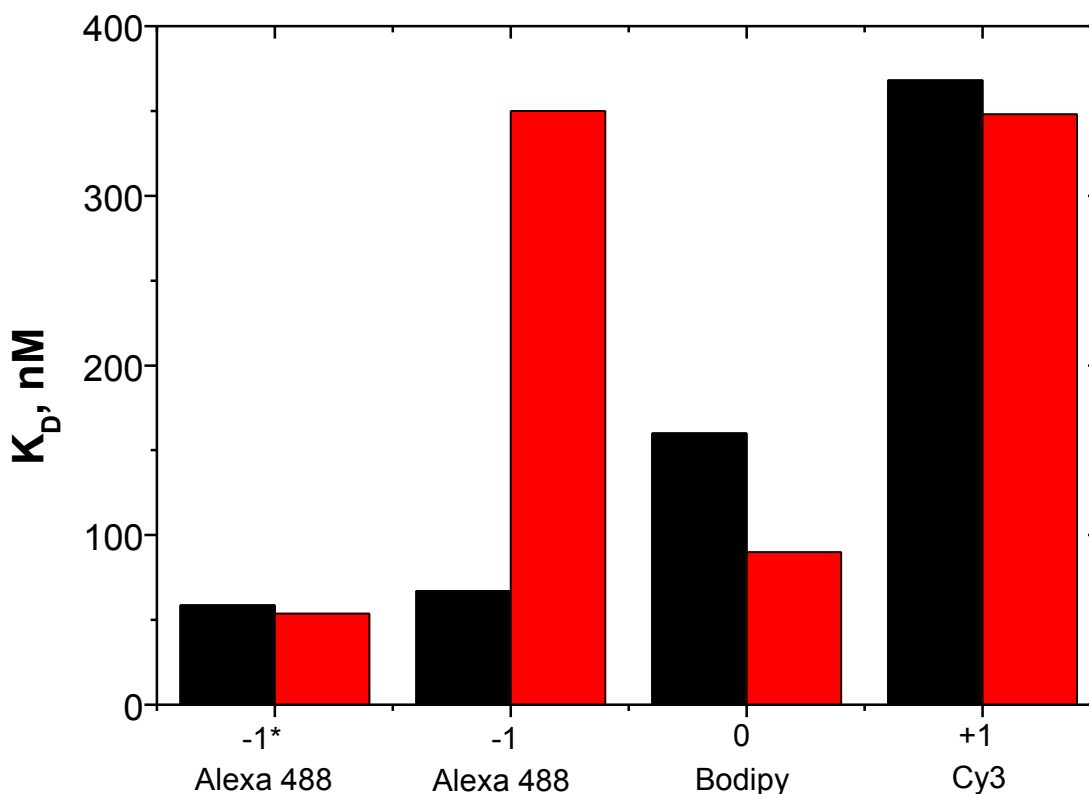


Figure 4.5. Comparison of the measured binding coefficients based upon the charge of the donor dye on polyE and whether or not the acceptor was labeled. Black: PolyK Peptide is unlabeled; Red: PolyK peptide is labeled with acceptor dye. The * represents the polyE peptide where the cysteine is in the middle of the chain instead of at the end.

4.5. Conclusions

Taking all of these results into consideration, it is clear that using fluorescent dyes around highly charged residues can severely effect results when reporting the total degree of quenching, rather than expanding on which quenching mechanism is responsible. In using FRET, it is common to use the decrease in donor fluorescence to measure binding. We determined that the effects of the charged environment on this parameter are unique to the dye as well as the environment. Each dye charge exhibited different effects in the highly charged environment, although each one did

give a similar K_D value, with the exception of the -1 charged dye. The amount of quenching through either FRET or electron transfer also varied by dye. It would appear as though the acceptor dye may not always be important for the measurement of binding affinity with fluorescence, as long as there is some sort of quencher available for the donor signal. In the case that a donor and acceptor dye are required, the donor dye should not be placed in a highly charged environment. If the options for dye placement are limited, then the best option is to use a neutral dye or to label your molecules in a manner so that the fluorescent dyes only come close enough to show high FRET but not close enough to interact. The distance needed will depend on the Forster distance for your chosen dye pair. It would also be advisable to not place your fluorescent label next to a charged amino acid since the quantum yield of the dye can be either reduced or increased based upon the charge of the amino acid.

4.6. References

1. Benitez, J. J., Keller, A. M., and Chen, P. (2010) Nanovesicle Trapping for Studying Weak Protein Interactions by Single-Molecule FRET, In *Methods in Enzymology*, pp 41-60, Elsevier Inc.
2. Hofmann, H., Nettels, D., and Schuler, B. (2013) Single-molecule spectroscopy of the unexpected collapse of an unfolded protein at low pH, *The Journal of Chemical Physics* 139, -.
3. Haas, E. (2012) Ensemble FRET Methods in Studies of Intrinsically Disordered Proteins, In *Intrinsically Disordered Protein Analysis* (Uversky, V. N., and Dunker, A. K., Eds.), pp 467-498, Humana Press.
4. Murno, J. B., Vaiana, A., Sanbonmatsu, K. Y., and Blanchard, S. C. (2008) Single Molecule Series: A New View of Protein Synthesis: Mapping the Free Energy Landscape of the Ribosome Using Single-Molecule FRET, *Biopolymers* 89, 565-577.
5. Giepmans, B. N. G., Adams, S. R., Ellisman, M. H., and Tsien, R. Y. (2006) The Fluorescent Toolbox for Assessing Protein Location and Function, *Science* 312, 217-224.
6. Ni, Y., and Wu, J. (2014) Far-red and near infrared BODIPY dyes: synthesis and applications for fluorescent pH probes and bio-imaging, *Org. Biomol. Chem.* 12, 3774-3791.
7. Yan, Q., and Bruchez, M. P. (2015) Advances in chemical labeling of proteins in living cells, *Cell Tissue Res.* 360, 179-194.
8. Crossman, D. J., Hou, Y., Jayasinghe, I., Baddeley, D., and Soeller, C. Combining confocal and single molecule localisation microscopy: A correlative approach to multi-scale tissue imaging, *Methods*.
9. Rumin, J., Bonnefond, H., Saint-Jean, B., Rouxel, C., Sciandra, A., Bernard, O., Cadoret, J.-P., and Bougaran, G. (2015) The use of fluorescent Nile red and BODIPY for lipid measurement in microalgae, *Biotechnol. Biofuels* 8, 1-32.
10. Jiao, Y., Zhu, B., Chen, J., and Duan, X. (2015) Fluorescent sensing of fluoride in cellular system, *Theranostics* 5, 173-187.
11. Ebrecht, R., Don Paul, C., and Wouters, F. (2014) Fluorescence lifetime imaging microscopy in the medical sciences, *Protoplasma* 251, 293-305.
12. Sobek, J., Aquino, C., and Schlapbach, R. (2007) Analyzing Properties of Fluorescent Dyes Used for Labeling DNA in Microarray Experiments, SigmaAldrich, BioFiles.

13. Daniels, C. R., Reznik, C., and Landes, C. F. (2010) Dye Diffusion at Surfaces: Charge Matters, *Langmuir* 26, 4807-4812.
14. Chung, H. S., Louis, J. M., and Eaton, W. A. (2010) Distinguishing between Protein Dynamics and Dye Photophysics in Single-Molecule FRET Experiments, *Biophysical Journal* 98, 696-706.
15. Di Fiori, N., and Meller, A. (2010) The Effect of Dye-Dye Interactions on the Spatial Resolution of Single-Molecule FRET Measurements in Nucleic Acids, *Biophysical Journal* 98, 2265-2272.
16. Dobbin, K. K., Kawasaki, E. S., Petersen, D. W., and Simon, R. M. (2005) Characterizing dye bias in microarray experiments, *Bioinformatics* 21, 2430-2437.
17. Lakowicz, J. R. (2006) *Principles of fluorescence spectroscopy*, 3rd ed., Springer, New York.
18. Weiss, J. N. (1997) The Hill equation revisited: uses and misuses, *FASEB J.* 11, 835-841.

V. Purification, Labeling, and FRET with FGFR and FGF

5.1. Abstract

Fibroblast growth factor receptor is involved in many important biological processes across different species as well as various tissue types. When FGFR goes unregulated, the result can be deadly. Almost all major cancers have a connection to unregulated FGFR. This protein is thought to have a self-regulatory process that reduces the ability for the receptor to bind to FGF, which is necessary for signaling. There are two competing models of self-regulation. We started looking at ways to either prove or disprove the two models. Our preliminary results show that the leading model in the literature may not be correct.

5.2. Introduction

Fibroblast growth factors (FGF's) are involved in numerous biological processes including wound healing and development¹⁻³. FGF's activate the fibroblast growth factor receptor (FGFR). There are more than 20 different FGF's but only 4 FGFR's⁴. These receptors are found in several different tissue types and the amount of each depends on the tissue it is in. FGFR participates in many different signaling processes⁵. The physical makeup contains 3 immunoglobulin-like domains (D1-3) as well as a transmembrane region and cytoplasmic region⁶. All of the ligand binding occurs in the second (D2) and third (D3) domains with the D3 domain only playing a small role. There is a disordered region (D1) and a flexible region (acid box, AB). The AB connects the first and second regions (figure 5.1). This flexible region is termed the acid box because of the high percentage of acidic residues it contains. It is thought to participate in an auto-inhibition mechanism that involves the acid box and the ligand binding domain on D2^{7,8}. When FGFR goes unregulated, there are a plethora of diseases that can be obtained, including cancer⁹⁻¹¹. FGFR must bind to both FGF and heparin in order for it to signal. There are two models proposed for the self-regulation or auto-inhibition of FGFR. The first model predicts the AB will bind to the heparin binding site on the D2 domain⁷. This binding event is assumed to block FGF from binding and therefore inhibits signaling. We have devised experiments to determine if in fact the acid does bind to the D2 domain or if the other competing model is true. The second model hypothesizes that the AB binds to FGF instead⁸. In this model, when AB binds to FGF, the FGF binding site on D2 is blocked and keeps other FGF molecules from binding. Through ITC and FRET both, we were only able to measure very little to no binding between the acid box and the D2 domain of FGFR leading us to believe the secondary model of auto-inhibition is more correct than the primary model.

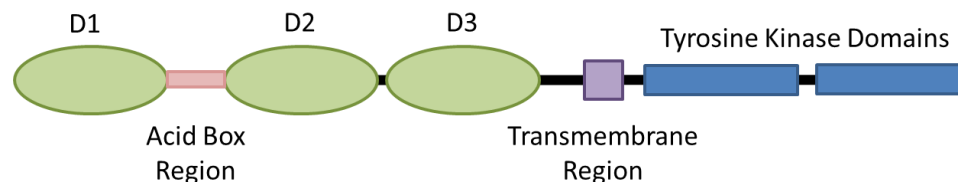


Figure 5.1. FGFR linear layout. FGFR contains 3 IG like domains (D1-3), an acid box region, a transmembrane region, and 2 tyrosine kinase domains.

5.3. Materials and Methods

5.3.1. Acid Box Peptide

The acid box peptide was ordered from Genscript with the sequence of V-T-D-A-I-S-S-G-D-D-E-D-D-T-D-G-A. This sequence aligns with residues 174 through 190 of FGFR2¹².

5.3.2. PCR, Expression, Purification of D2 and D2D3 domains of FGFR2 and FGF1

The D2, D2D3 domains of FGFR2 as well as FGF1 was mutated to include cysteines at single locations using a QuikChange XL Site Directed Mutagenesis Kit from Stratagene. A his-tag was added to the N-terminal end of the D2D3 domain using the mutagenesis kit as well. All primers used in PCR were ordered from IDT DNA Technologies. The thermocycler used for PCR was a master cycler gradient. Overexpression of the D2, D2D3 domains as well as the FGF1 protein was performed as previously reported by Hung¹³ using BL-21DE3 PLys cells from *E. coli* (Novagen). After shaking at 37°C and 160 rpms until OD at 600 nm reached 0.6, IPTG (Isopropylthio-β-D-galactoside, 0.5 mM/L) was added to induce growth. Shaking and heat stayed constant until OD at 600 nm reached 1.5-2.0. The bacteria were then centrifuged at 6000 rpms for 20 minutes. The pellet was then resuspended in a buffer containing 1xPBS (pH=7.2), 1% v/v BME (β-mercaptoethanol), and 10 mM PMSF (phenylmethanesulfonylfluoride). The cells were then lysed using a sonicator in 1 sec pulses for 55 pulses. The sample was then centrifuged again at 16,000 rpms for 20 minutes to remove cellular debris. For the D2, D2D3 proteins, the

pellet is resuspended again in buffer containing 8 M urea and centrifuged at 10,000 rpms for 20 minutes. Purification of D2, D2D3 proteins was performed using a nickel affinity column (Clontech), as described by Hung¹³. The protein was loaded onto the column after equilibration with 8 M urea. The protein was eluted with a gradient of imidazole with urea (20 mM to 500 mM imidazole, 8 M to 0 M urea). The D2, D2D3, and FGF proteins were also purified by heparin affinity chromatography (GE Healthcare). The protein was loaded onto the column using a buffer containing 10 mM phosphate, pH 7.2. The proteins were eluted using a gradient of NaCl, from .2 M to 1.5 M. Purification was verified by SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis), figures 5.2-5.8. Figure 5.2 displays the gel after purification by a heparin column. The protein of interest (D2 wild type) has a molecular weight of 14 kDa. The band on the gel appears slightly below the molecular weight marker of 17 kDa. In Figure 5.3, the D2 mutant 236C was purified also on a heparin column. The band on this gel is also slightly below the 17 kDa molecular weight marker. Figure 5.4 shows the gel after the purification of D2D3 wild type. Here the weight of the protein is 24.8 kDa. The band appears level with the 25 kDa marker. It should be noted that the molecular weight marker erroneously showed a double band for 25 kDa. Figure 5.5, 5.6, and 5.7, are all from the purification of the D2D3 protein. Each figure represents a different mutant however. For all cases, the band is around the 25 kDa marker. In figure 5.8, the gel shown is from the purification of FGF wild type. The molecular weight of this protein is 16 kDa and the band is present alongside the 17 kDa marker.

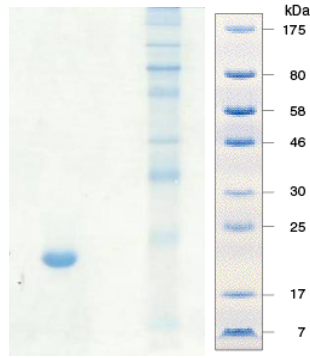


Figure 5.2. Gel of D2 wild type purification after the heparin column. The molecular weight of D2 is ~14 kDa. The band is slightly below the 17 kDa marker.

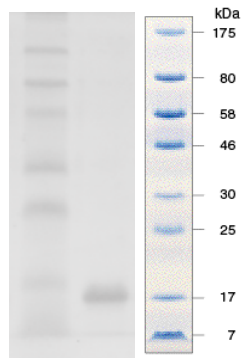


Figure 5.3. Gel of D2 236C purification after the heparin column. Our protein is slightly below the marker for the 17 kDa control, which is where it should be.

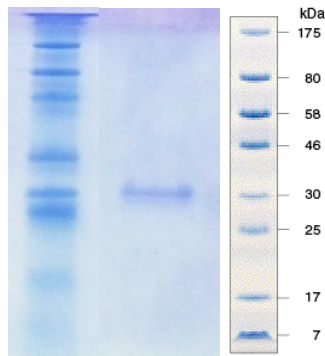


Figure 5.4. Gel of D2D3 wild type purification after the heparin column. The molecular weight of D2D3 is 24.8 kDa. The molecular weight marker that was used erroneously showed a double band at 25 kDa.

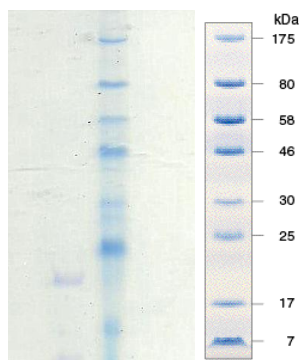


Figure 5.5. Gel of D2D3 219/236C purification after the heparin column. The protein of interest produced a band slightly below the 25 kDa molecular marker control band. Protein of interest = 24.8 kDa.

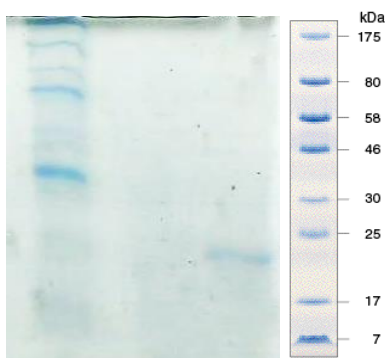


Figure 5.6. Gel of D2D3 236/281C purification after the heparin column. Protein of interest molecular weight is 24.8 kDa. Band is present below 25 kDa marker.

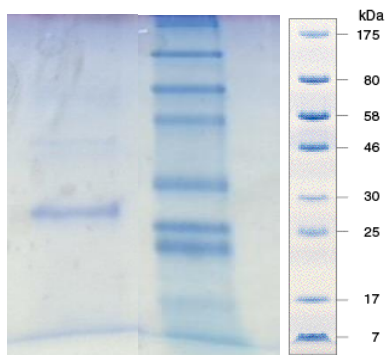


Figure 5.7. Gel of D2D3 222C purification after the heparin column. The molecular weight marker that was used erroneously showed a double band at 25 kDa. The protein of interest's band was right above the 25 kDa molecular weight marker. Molecular weight of protein = 24.8 kDa.

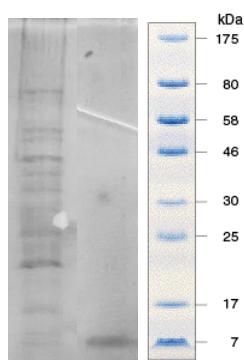


Figure 5.8. Gel of FGF wild type purification after the heparin column. The molecular weight of FGF is ~16 kDa. The band is almost even with the 17 kDa marker band.

5.3.3. Fluorescent Labeling of acid box, D2, D2/D3 and FGF

Acid box peptide was labeled with the donor dye Alexa Fluor 488. Labeling was carried out using a 5:1 ratio of dye to peptide in 10 mM phosphate buffer at pH 7.2. The mixture of peptide and dye, protected from light, were allowed to react overnight. The unreacted free dye was separated out by dialysis using 2 kDa molecular weight cut-off dialysis tubing (Spectrum Labs). The samples were dialyzed against 10 mM phosphate buffer for 24-48 hours with the external buffer being refreshed every 2-6 hours. The labeling efficiency was verified by MALDI mass spectrometry for the acid box; figure 5.9 (Bruker Ultraflex II TOF/TOF time-of-flight mass spectrometer equipped with a MALDI ion source (Bruker Daltonik GmbH, Bremen, Germany)). The proteins (D2, D2D3, FGF) were labeled using between 10:1 and 15:1 ratios of dye to protein in 10 mM phosphate buffer at pH 7.2. Tris (2-carboxyethyl) phosphine (TCEP) was added to the solution 30 minutes before the dye was introduced in order to reduce surface disulfide bonds. The mixtures were allowed to sit for 2 hours after the addition of dye, at room temperature and protected from light. The samples were then centrifuged in 3 kDa spin tubes (Millipore) to remove all unbound dye. The absorbances of the samples were taken to determine amount of labeling.

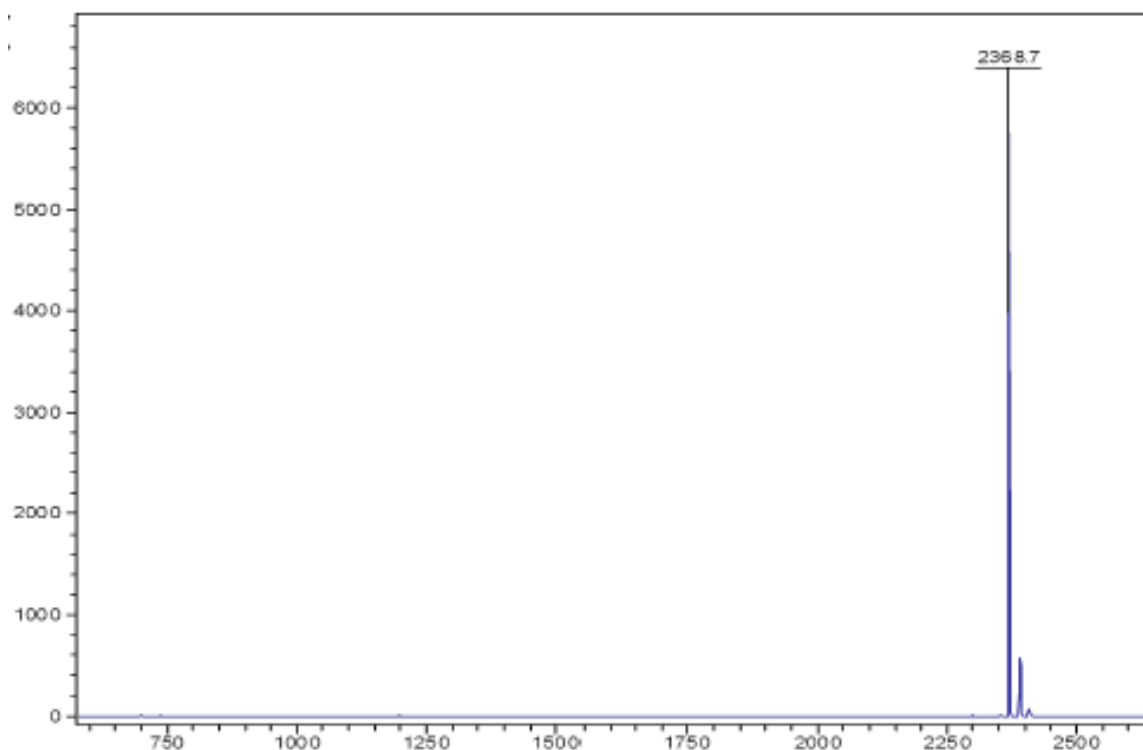


Figure 5.9. MALDI results from acid box labeling with Alexa488. The molecular weight of the acid box peptide is 1670.7 Da. The weight of the peptide + dye is 2368.7 Da. The peaks on the right side of the major peak presents the sample with sodium ions attached.

5.3.5. Isothermal Calorimetry

All ITC data was collected on a Microcal VP titration calorimeter (Northampton, MA, USA). For D2 with acid box, the D2 concentration was 53 μM , the acid box was 499 μM . The injection volume was 6 μL with 30 injections. The buffer was 10 mM Phosphate, pH=7.2. The data was fit using Origin and the χ^2 value was used to determine the best fit.

5.3.6. Fluorescence Resonance Energy Transfer (FRET) Spectroscopy

FRET measurements were made on a PTI Quantamaster 40. The concentration of labeled acid box was set at 5.0×10^{-8} M. The concentration of labeled FGF varied from 5.0×10^{-10} M to 1.5×10^{-6} M. For D2 or D2D3 with FGF, the D2 (or D2D3) sample had the same concentration as the acid box (above) and the FGF remained the same as listed above. The sample was excited

at 470 nm and emission data was collected from 480 nm to 750 nm. The transfer efficiency was calculated by using the relative fluorescence intensity of the donor at 516 nm in the presence (F_{DA}) and absence (F_D) of the acceptor¹⁴ using: $E = 1 - (F_{DA}/F_D)$. The K_D value was calculated by fitting to the Hill equation in Origin.¹⁵

5.4. Results

In order to measure the binding of the FGFR with FGF using FRET, we had to add a cysteine into the sequence of the protein to add a fluorescent label. The un-changed sequences for D2, D2D3, and FGF are shown in figures 5.10-5.12, respectively. By using site directed mutagenesis, we were able to make a mutation at amino acid 236 for D2. Some of the mutations that we tried to make but never could get to work were: 155C, 219C, and 281C. For the D2D3 construct, the mutations that worked were: 219/236C (double), 236/281 (double), 222C, 241C, and 356C. The mutations that did not work were: 265C, 267C, 308C, 309C, and 358C. For FGF, the mutations that worked were 2C and 49C. After the mutations were made, some of the samples could not be purified to a high degree. These were samples D2D3 241C and D2D3 356C. After receiving the sequence for the D2D3 protein, it was discovered that there was not a His-tag on the protein. We performed PCR in order to add a His-tag. The amino acid sequence for the D2D3 domains after the addition of the His-tag is shown in figure 5.13.

**Met NSNNKRAPYWTNTEKMEKRLHAVPAANTVKFRCPAGGNP
MPTMRWLKNGKEFKQEHRIGGYKVRNQHWSLIMESVVP SDK
GNYTCVVENEYGSINHTYHLDVVLEHHHHHHH Stop**

Figure 5.10. Primary sequence for D2. This is the wild type sequence of D2. The his-tag is at the C-terminus of the protein.

```
Met NSNNKRAPYWTNTEKMEKRLHAVPAANTVKFRCPAGGNP  
MPTMRWLKNGKEFKQEHRIGGYKVRNQHWSLIMESVVPSDK  
GNYTCVVENEYGSINHTYHLDVVERSPHRPILQAGLPANAST  
VVGGDVEFVCKVYSDAQPHIQWIKHVEKNGSKYGPDGLPYL  
KVLKHSGINSSNAEVLALFNVTEADAGEYICKVSNIYIGQANQ  
SAWLTVLPKQQAPGRE Stop
```

Figure 5.11. Primary sequence for D2D3. This is the original sequence for the D2D3 protein. Notice the lack of a his-tag. Without this tag, we could not purify the protein using affinity chromatography.

```
Met FNLPPGNYKKPKLLYCSNGGHFLRILPDGTVDGTRDRSDQ  
HIQLQLSAESVGEVYIKSTETGQYLAMDTDGLLYGSQTPNEE  
CLFLERLEENHYNTYISKKHAEKNWFVGLKKNGSCKRGPRTH  
YGQKAILFLPLPVSSD Stop
```

Figure 5.12. Primary sequence for FGF. This sequence is the wild type sequence for FGF. We were able to purify FGF by using a heparin column alone, without the need to include a nickel column.

```
Met NSNNKRAPYWTNTEKMEKRLHAVPAANTVKFRCPAGGNP  
MPTMRWLKNGKEFKQEHRIGGYKVRNQHWSLIMESVVPSDK  
GNYTCVVENEYGSINHTYHLDVVERSPHRPILQAGLPANAST  
VVGGDVEFVCKVYSDAQPHIQWIKHVEKNGSKYGPDGLPYL  
KVLKHSGINSSNAEVLALFNVTEADAGEYICKVSNIYIGQANQ  
SAWLTVLPKQQAPGREHHHHHH Stop
```

Figure 5.13. Sequence for D2D3 after addition of His-tag. This is sequence of the D2D3 after the addition of a his-tag at the C-terminus. This sequence will be referred to as the wild type D2D3.

Labeling the mutants that we made proved to be difficult. Before labeling, the absorbance of every protein was measured in order to determine the concentration, figure 5.14. After labeling, the samples were sent to the Arkansas Statewide Mass Spectrometry Facility for analysis. Figure 5.15 an example of the MALDI results after labeling. This MALDI spectrum

shows that we had protein with several variations of numbers of dyes attached. Since the maximum that we could have had with the protein folded was 1, the rest are indicative of unfolded protein, which we also saw with the absorbance measurement in figure 5.16. We successfully labeled D2 236C with Alexa488, figure 5.17, even though a small amount of the sample did appear to have 2 dyes attached. When we were trying to label the D2D3 double mutants, we would add both of the dyes in at the same time. It did not matter where either of the dyes attached, only that both dyes were on each molecule. Several times we would end up with aggregated samples, figure 5.18. Other times we would have samples that only had a single dye attached, figure 5.19. We decided to start attaching one dye then the other, but this too yielded poor results. We obtained samples that had a low labeling efficiency; figure 5.20 for example, had a 5:1 ratio of protein of dye.

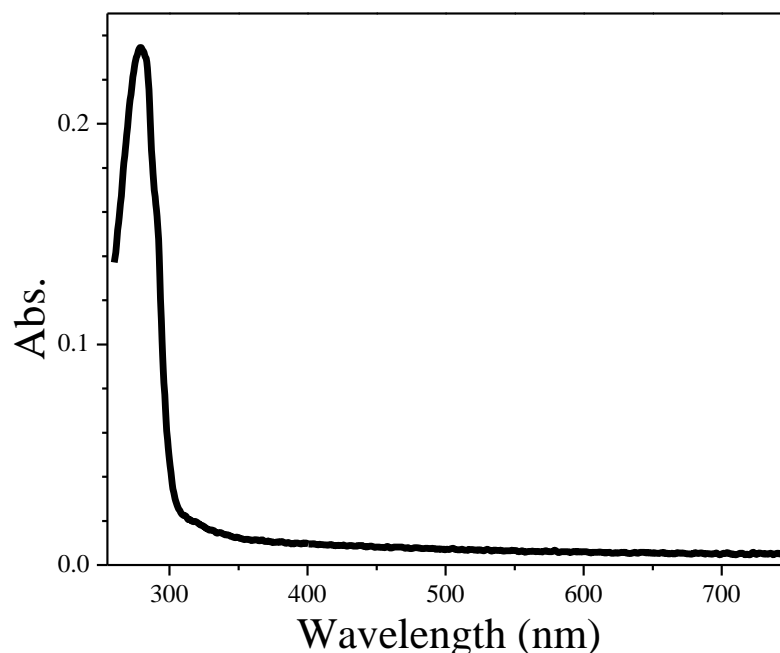


Figure 5.14. Spectra of D2 236C before labeling. The peak at 280 nm signifies the presence of protein that is not aggregated.

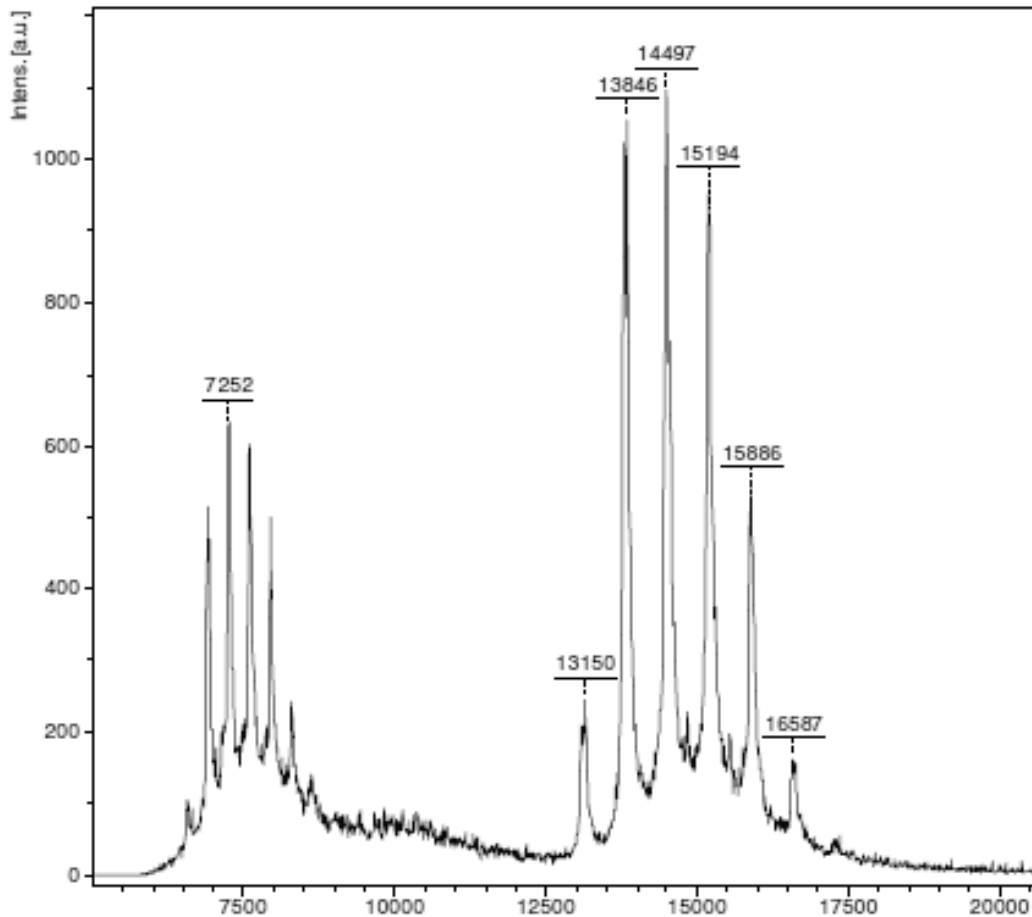


Figure 5.15. MALDI results from D2 236C with Alexa488. The peak at 13150 represents the protein. All of the peaks with a higher M/Z ratio represent the protein with dye, +2 dye, +3 dye and so forth. Here the extra dye that are attached come from the protein most likely being unfolded so the native cysteines were available for labeling.

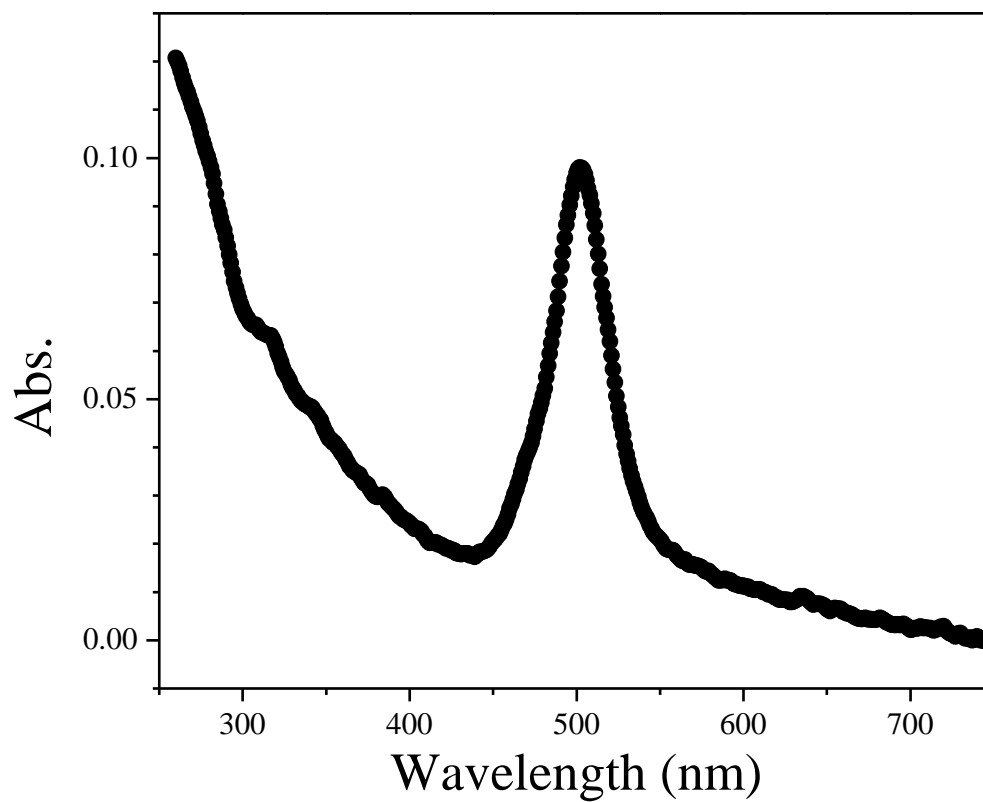


Figure 5.16. Spectra of D2 236C after labeling with Alexa488. The peak at ~500 nm is from the dye, Alexa 488. The large peak that does not come back down indicates aggregated protein.

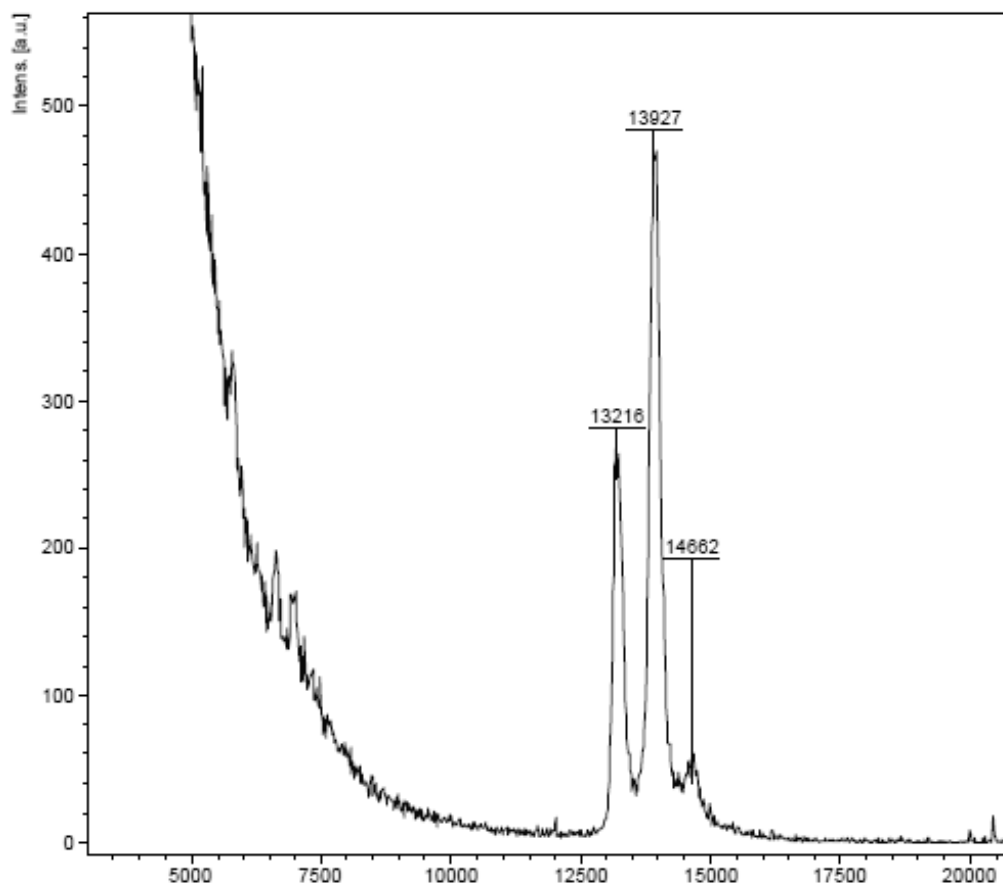


Figure 5.17. D2 236C with Alexa488. The peak at 13216 indicates the protein. The next downfield peak is for protein + dye, the last labeled peak is for protein +2 dyes. The extra dye is likely a result of unfolding of the protein, allowing the normally buried native cysteines to be available to accept the dye.

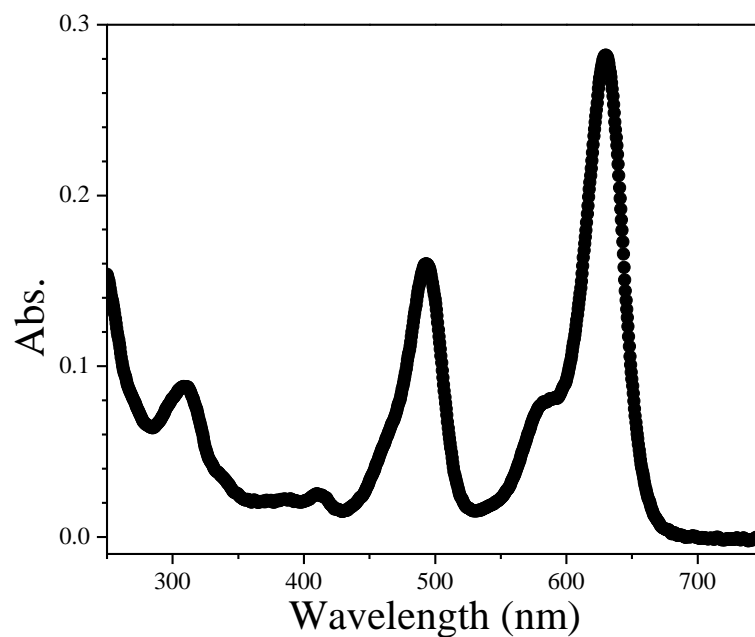


Figure 5.18. Absorbance of D2D3 219/236C. The sample contains both of the dyes, but it also appears to be aggregated by the peak that starts at 285 nm and does not come back down.

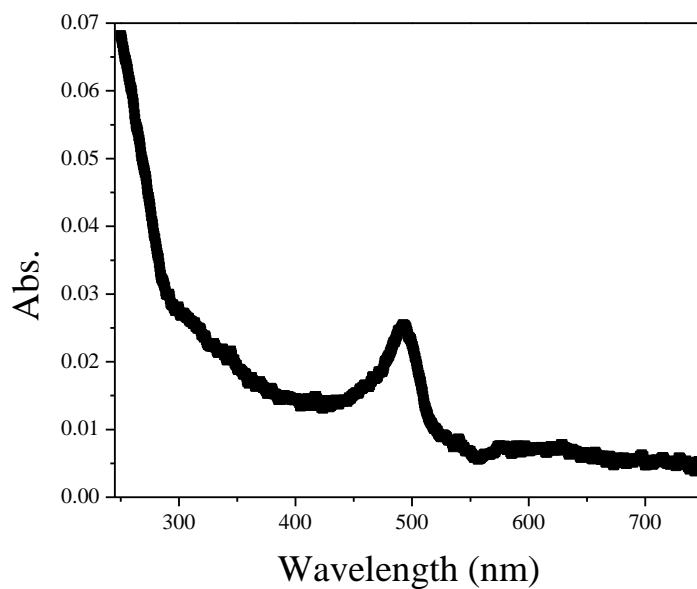


Figure 5.19. Absorbance of D2D3 236/281C. The sample contains one of the dyes, but not the other. It also appears to be aggregated by the peak that starts at 400 nm and does not come back down. The overall concentration of the dye is very low, $\sim 1 \times 10^{-7}$ M.

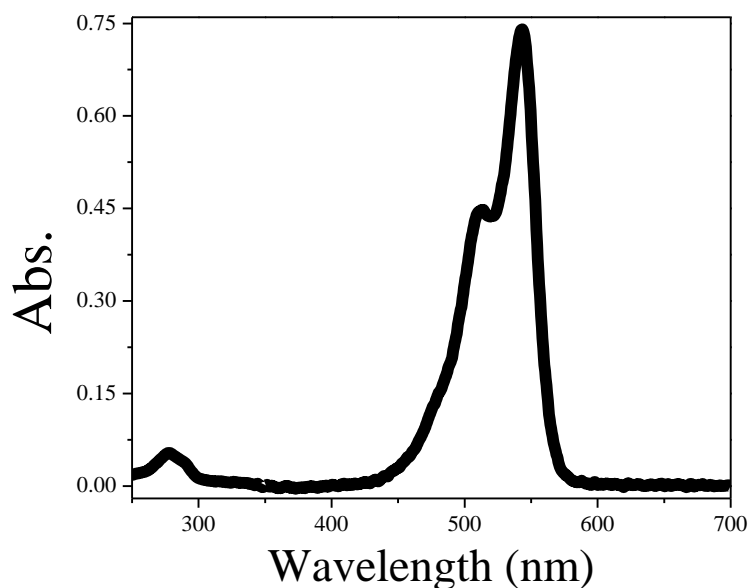


Figure 5.20. D2D3 236/281C with Cy3. The peak at ~550 nm donates the presence of Cy3 while the peak at 280 nm indicates the presence of protein.

The wild type FGF has 3 native cysteines. We determined that none of these cysteines were available for labeling by adding dye to the protein and then purifying the sample. In every case, the ratio of protein to dye was very high, anywhere from 37:1 to 20:1, figure 5.21. Because of this, we decided to introduce a cysteine at position 49. This mutant was successful and was easily purified. The labeling was also somewhat successful with a 7:1 ratio of protein to dye. This sample was not, however, refolded properly. After labeling, the sample was aggregated (figure 5.22). After the aggregates were dissolved, the protein would not bind to the D2D3 protein, figure 5.23.

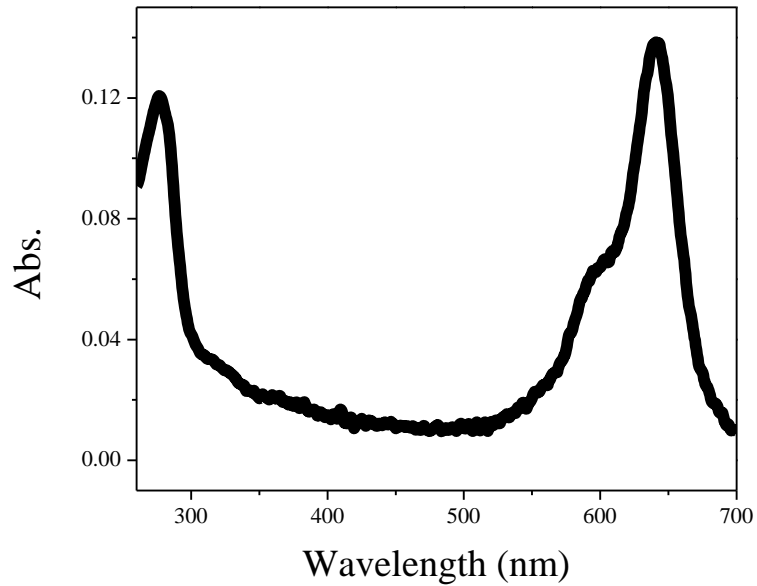


Figure 5.21. Absorbance spectrum of FGF with Cy5 dye. The peak at ~650 nm is from the Cy5 dye while the peak at 280 nm is from protein.

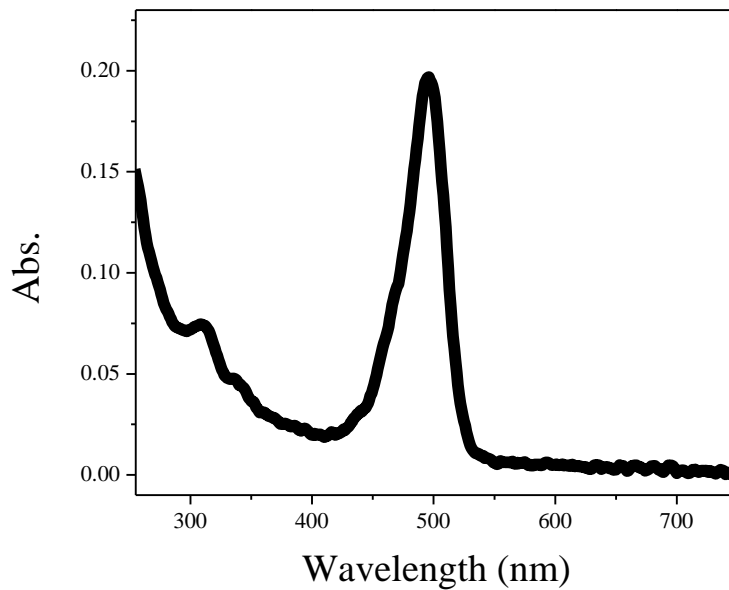


Figure 5.22. Absorbance spectrum of FGF-49C with Alexa488. The presence of Alexa 488 is denoted by the peak at 500 nm while the lack of a defined peak at 280 nm suggests aggregated protein.

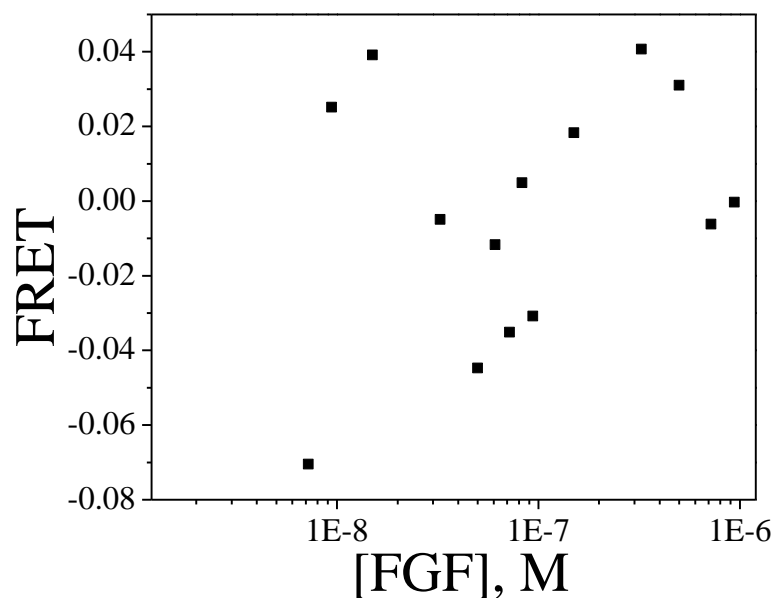


Figure 5.23. Ensemble FRET with FGF and FGFR. The scattered data points suggest no binding between the receptor and FGF.

In our ITC experiments, D2 wild type was titrated with the acid box peptide. This experiment showed very weak to no binding, figure 5.24. We performed the same experiment using FRET to determine if the two were binding on a small scale. FRET also showed no binding, figure 5.25. We also decided to try ITC with the D2 wild type with a polypeptide that was designed to mimic the acid box peptide, but only contained glutamic acid. This peptide is termed PolyE and contains 15 amino acids, E15. For this experiment, there was binding between the protein and peptide even though it was weak (37 mM), figure 5.26.

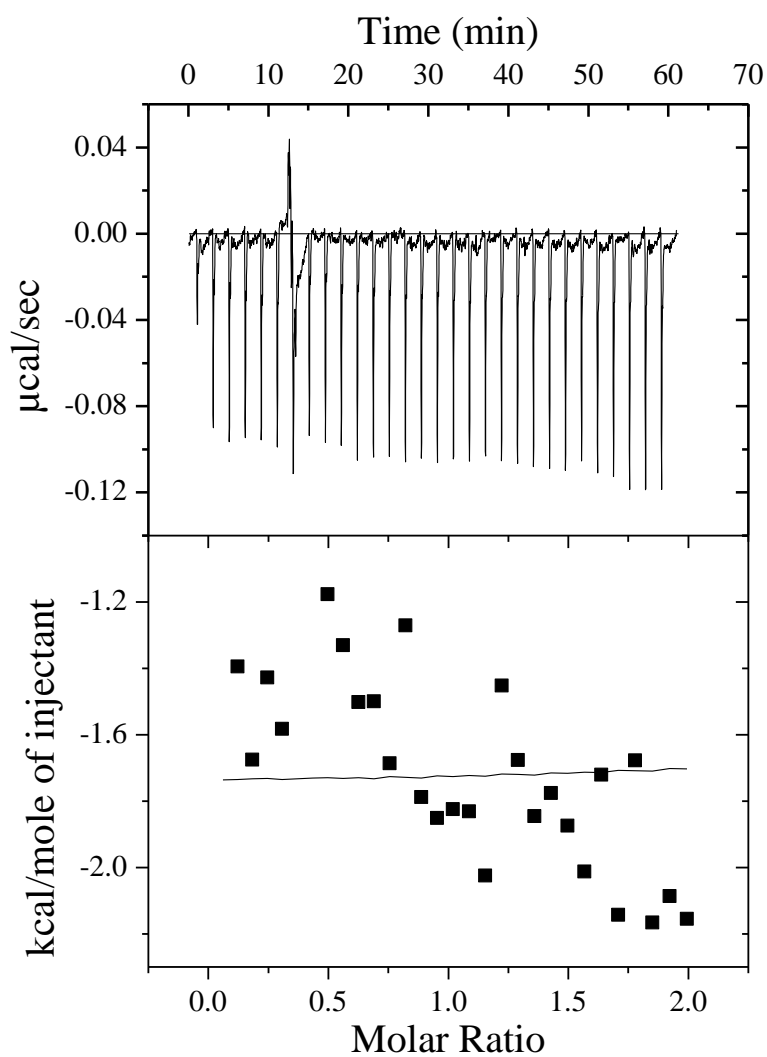


Figure 5.24. ITC of D2 wild type with acid box peptide. The ITC results indicate that only a very small amount of binding took place between the acid box peptide and the D2 domain.

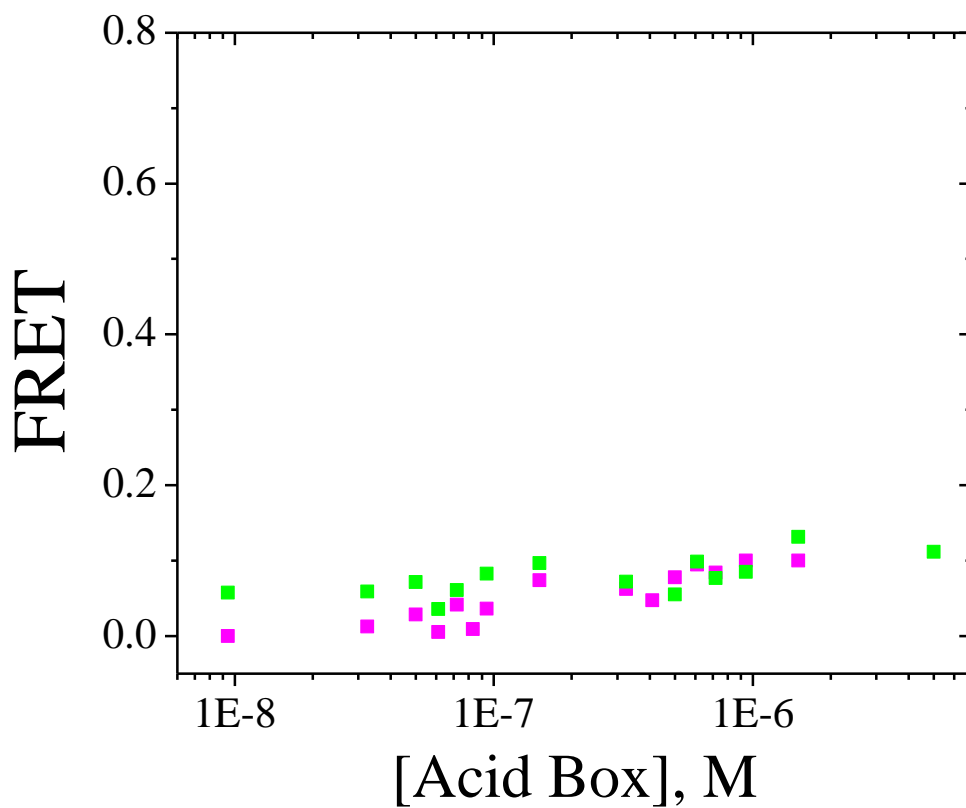


Figure 5.25. FRET of D2 236C with acid box peptide. Green and purple represent two different trials. There was extremely weak binding in both. Data could not be fit to model because of extremely weak binding between the acid box peptide and the D2 domain.

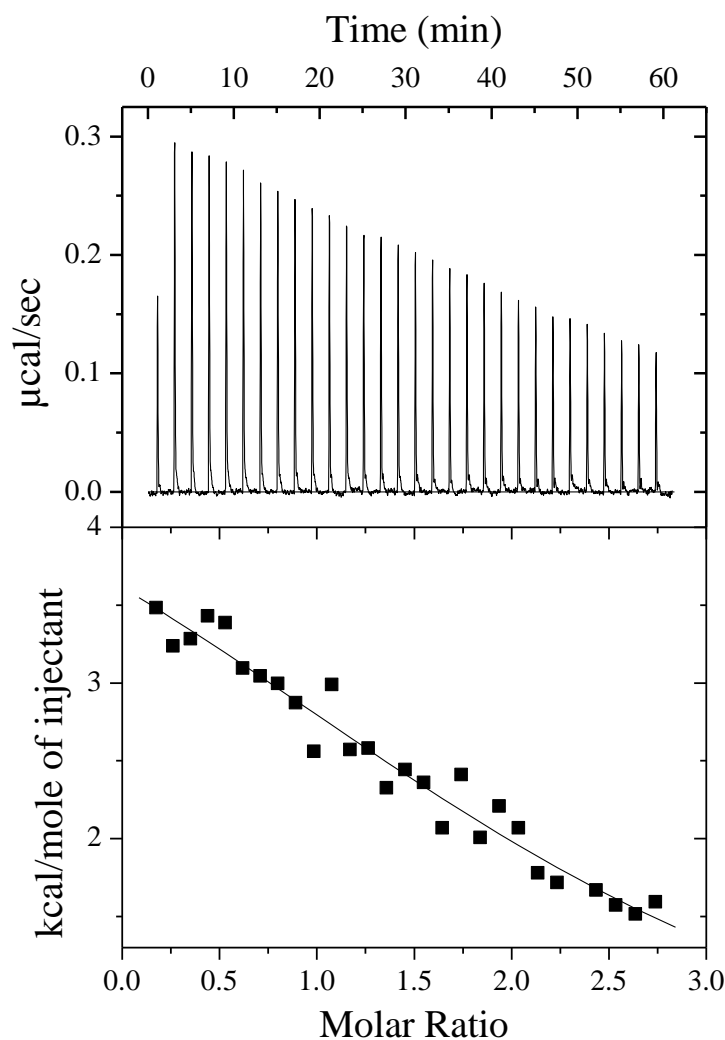


Figure 5.26. ITC of D2 wild type with PolyE peptide. ITC confirms that the polyE peptide does bind to the D2 domain, with relatively weak binding, $K_D = 37$ mM.

5.5. Discussion and Conclusion

The auto-inhibition of FGFR is thought to occur by the acid box region of the receptor binding to the heparin binding site of the receptor¹⁶. In this model, the acid box binding on the heparin binding site is supposed to block heparin and the FGF from binding to the receptor⁷. Even though this is a popular model in the literature^{7, 16}, our results from ITC (figure XX) as well as FRET (figure XX) both show little to no binding between the acid box and FGFR. Our results

are preliminary and we hope future results will help to clear up the questions left at this point.

Our results do however, correlate with Rutherford et al. who saw that the acid box has a higher affinity for FGF than for the receptor⁸.

5.6. References

1. Zhang, X., Ibrahimi, O. A., Olsen, S. K., Umemori, H., Mohammadi, M., and Ornitz, D. M. (2006) Receptor Specificity of the Fibroblast Growth Factor Family: THE COMPLETE MAMMALIAN FGF FAMILY, *Journal of Biological Chemistry* 281, 15694-15700.
2. Beenken, A., and Mohammadi, M. (2009) The FGF family: biology, pathophysiology and therapy, *Nature Reviews. Drug Discovery* 8, 235-253.
3. Givol, D. (2009) Molecular and cellular biology of FGF signalling, *Oxford monographs on medical genetics* 54, 449-460.
4. Ornitz, D., and Itoh, N. (2001) Fibroblast growth factors, *Genome Biology* 2, 1-12.
5. Eswarakumar, V. P., Lax, I., and Schlessinger, J. (2005) Cellular signaling by fibroblast growth factor receptors, *Cytokine & Growth Factor Reviews* 16, 139-149.
6. Wang, F., Kan, M., Yan, G., Xu, J., and McKeehan, W. L. (1995) Alternately Spliced NH2-terminal Immunoglobulin-like Loop I in the Ectodomain of the Fibroblast Growth Factor (FGF) Receptor 1 Lowers Affinity for both Heparin and FGF-1, *Journal of Biological Chemistry* 270, 10231-10235.
7. Kalinina, J., Dutta, K., Ilghari, D., Beenken, A., Goetz, R., Eliseenkova, Anna V., Cowburn, D., and Mohammadi, M. (2012) The Alternatively Spliced Acid Box Region Plays a Key Role in FGF Receptor Autoinhibition, *Structure* 20, 77-88.
8. Rutherford, L., Rajalingam, D., and Kumar, T. K. S. Understanding the molecular mechanism underlying the auto inhibition of the fibroblast growth factor signaling.
9. Dieci, M. V., Arnedos, M., Andre, F., and Soria, J. C. (2013) Fibroblast Growth Factor Receptor Inhibitors as a Cancer Treatment: From a Biologic Rationale to Medical Perspectives, *Cancer Discovery* 3, 264-279.
10. Belov, A. A., and Mohammadi, M. (2013) Molecular Mechanisms of Fibroblast Growth Factor Signaling in Physiology and Pathology, *Cold Spring Harbor Perspectives in Biology* 5.
11. Kelleher, F. C., O'Sullivan, H., Smyth, E., McDermott, R., and Viterbo, A. (2013) Fibroblast growth factor receptors, developmental corruption and malignant disease, *Carcinogenesis* 34, 2198-2205.
12. Uni-Prot. (<http://www.uniprot.org/uniprot/P21802#sequences>) **P21802-1**.
13. Hung, K.-W., Kumar, T. K. S., Chi, Y.-H., Chiu, I.-M., and Yu, C. (2004) Molecular cloning, overexpression, and characterization of the ligand-binding D2 domain of fibroblast growth factor receptor, *Biochemical and Biophysical Research Communications* 317, 253-258.

14. Lakowicz, J. R. (2006) *Principles of fluorescence spectroscopy*, 3rd ed., Springer, New York.
15. Weiss, J. N. (1997) The Hill equation revisited: uses and misuses, *FASEB J.* *11*, 835-841.
16. Olsen, S. K., Ibrahimi, O. A., Raucci, A., Zhang, F., Eliseenkova, A. V., Yayon, A., Basilico, C., Linhardt, R. J., Schlessinger, J., and Mohammadi, M. (2004) Insights into the molecular basis for fibroblast growth factor receptor autoinhibition and ligand-binding promiscuity, *Proceedings of the National Academy of Sciences of the United States of America* *101*, 935-940.

VI. Conclusion

This body of work was aimed at determining if the auto-inhibition mechanism for FGFR that is popular in the literature is accurate¹. This work was started by making a simplified homopolyptide that would eventually morph into the actual amino acid sequence proposed to be involved in auto inhibition by FGFR. One model peptide was made to mimic the acid box – a 15-mer of glutamic acid (polyE). In the actual protein, the binding site on FGFR was thought to be a positively charged region in the D2 domain that binds to heparin. We used a 15-mer of polylysine (polyK) to mimic this binding site. For the polypeptides that were used in FRET studies, a cysteine residue was added to enable site-specific labeling using maleimide-functionalized dyes that react specifically with the terminal –SH group on cysteine. The binding of the peptides were characterized by ITC, CD, FRET, and NMR (chapter 2) and it was discovered that they bind in a parallel arrangement almost exclusively via long-range electrostatic interactions of the side-groups with little-to-no involvement of backbone hydrogen bonds. This leads to the polypeptides to remain as random coils after interacting. A parallel arrangement and electrostatic only interactions are somewhat counterintuitive, since longer versions of these peptides have been shown to form beta-sheets after interacting. Our data clearly shows that polypeptides of this length do not form beta-sheets in aqueous solution until the solvent polarity is decreased by adding trifluoroethanol. A recent molecular dynamics study² showed that for 5-mers of polyE and polyK in dried layers, the difference in energy between parallel and anti-parallel arrangements was rather small, with the parallel arrangement resulting in more compact sheets. It is possible that, when the solvent is included and the length of the polypeptides is increased, such as in our studies, the balance of entropy and enthalpy shifts towards maintain a more entropically-favored random coil. In fact, our ITC results showed that both the DS and the DH of interaction is positive at neutral pH in both lower and higher ionic strength condition, verifying that the interaction is largely entropically driven. In the MD study² a parallel arrangement was shown to result in a larger distance between the peptides, which may allow the solvent to interact with the charged side groups enough to increase the entropy enough to balance out the enthalpy that would be gained by forming hydrogen bonds that leads to the formation of beta sheets. The question remains as to the

minimum length and charge composition of peptide that is needed to maintain this entropy balance before the formation of intra-peptide hydrogen bonds shifts the balance towards the formation of secondary structure. Indeed, answering such a question may prove to be extremely important in the area of amyloid formation that involve IDPs coming together to form ordered structures (often beta sheets) that have been implicated in several diseases such as Alzheimer's³, Parkinson's⁴, Cystic Fibrosis and the human equivalent of Bovine Spongiform Encephalopathy (BSE), Creutzfeldt-Jakob disease (CJD)⁴.

Through the course of the FRET studies with the polypeptides, we noticed that the exact quantification of our FRET data seemed to depend on the choice of dyes used. This forced us to start examining the role that fluorescent dyes played in the interaction of the charged polypeptides that we were using. In chapters 3 and 4, we were able to quantify the effects of the charges on the dyes themselves, as well as the effect of the charged amino acids on the dye fluorescence – both of which are very important for correct interpretation of FRET data. In chapter 3, we elucidated the fact that bulky, rigid dyes inhibited peptide binding and reduced the binding affinity but also decreased the cooperativity of binding, as evidenced by the change in the Hill coefficient extracted from our fits. Furthermore, dyes that had the same charges also reduced the K_D value since they repelled each other, keeping the peptides apart. The structural flexibility of the dye also plays a major role, since this allows the dye to adopt conformations that interact less with other dyes. These effects are amplified when the dyes are forced into close proximity of each other, which was allowed in our system by the strong preference of the polypeptides to bind in a parallel arrangement. One important question to answer would be what the minimum distance is that the dyes must be placed apart from each other for them to have a negligible effect on the data. It is anticipated that, at larger distances, the charge of the dyes would become more important than their bulkiness and rigidity, since charge-charge interactions are longer range than the dimensions of the dye. Answering such questions is important when designing FRET assays to look at protein interactions, especially those involving electrostatic interactions. The results of chapter 3 have

gone a long way towards this goal, since it shows which dye combinations work better than others when dyes must be placed relatively close together.

Since our peptides were so highly charged, it was also necessary to uncover the effect of these charges on the quantum yield of our dyes when attached to, or brought into close proximity to, such residues (chapter 4). How the quantum yield of the dyes is affected by these highly-charged environments are very important to accurately measure FRET efficiencies which are used to extract inter-fluorophore separation and, in turn, structural information about the labeled biomolecules. After determining that the dye quantum yield was in fact influenced by the amino acid chain it was attached to, we monitored the change in quantum yield after the peptides bind to each other. We were able to establish a general mechanism for the quenching of the fluorescence of our dyes. This mechanism involves the transfer of an electron from the dye to the polyK peptide, which depends strongly on the distance between the dye and the amino acid chain as well as the number of quenching pathways (number of lysine residues) available within the short quenching distance. Since this quenching competes with FRET, lacking such knowledge of quenching will overestimate the FRET efficiency and may lead to inaccurate conclusions to be drawn regarding the structural details uncovered by FRET. The results of chapter 4 will help to correctly interpret FRET results when dyes are placed into close proximity of charged amino acids to enable more accurate inter-fluorophore distances to be extracted.

Once we had quantified the how interacting charged residues affected each other's structure and the roles that attached dyes played on this interaction, we then moved on to studying the more complex FGFR system. After the addition of a hexahistidine tag to the obtained FGFR plasmid through PCR (chapter 5), we were able obtain purified protein. FGFR contains 2 cysteines per domain (D2 and D3), but these residues are not available for labeling since they form disulfide bridges after folding. Because of this, we decided to add a cysteine into the sequence at various points in order to have several protein labeling options. Each mutation was made for a single protein variant, some containing single cysteines to enable inter-molecular FRET between FGFR and its binding partner and some containing double

mutations to enable intra-molecular FRET to be performed on FGFR. The single mutations enabled binding between FGFR and FGF or FGFR and the acid box peptide to be measured by FRET. After initial studies with ITC and FRET, the expected binding between FGFR and the acid box could not be verified, bringing into question one of the proposed mechanisms for FGFR auto-regulation¹. Unfortunately, we have not been able to measure binding between FGFR and FGF, which is another proposed mechanism of FGFR autoregulation, due to the FGF being unstable after labeling and aggregation issues.

Since one of the proposed autoregulation mechanisms have been brought into doubt, future work should investigate alternative mechanisms. Such a project will include the following steps: 1) Producing new mutants of FGF that could potentially be more stable. After obtaining stable FGF, FRET and even single molecule FRET could be performed between FGFR and FGF. These studies should be able to determine if the acid box binds to FGF to determine if the alternative FGFR autoregulation mechanism is viable. If the binding between FGF and the acid box is present, NMR studies can also be conducted to determine where the binding is taking place on FGF. 2) The binding between FGF and the acid box should also be measured in the presence and absence of the D2D3 regions. This study will help to uncover whether the binding of FGF to the acid box is preferred over D2D3 or if this interaction is only a side effect. Just because the acid box will bind, does not mean that it is going to bind when other potential partners are available. 3) Measure the binding of peptides with all of the different FGFR acid box sequences since each receptor has a slightly different one. FGFR1 is the highest regulated with the longest acid box. The binding of these acid boxes could be compared to the regulation level for each one, after the binding mechanism is determined. 4) Determine the minimum number of charged amino acids that are required for the acid box to bind to its partner (either FGF or FGFR). It could be very helpful with future binding mechanisms if the minimal amount of charged residues is known. 5) Expand the homopolypeptide study by changing the amino acids one at a time to glycine and then measuring the binding between the polyE and polyK. Since there are many intrinsically disordered proteins in the body, these polypeptide systems could be changed to mimic some of those systems more closely. By changing

one amino acid at a time, we could determine if small changes could force folding changes, aggregation, or even binding changes that we could not foresee.

6.1 References

1. Kalinina, J., Dutta, K., Ilghari, D., Beenken, A., Goetz, R., Eliseenkova, Anna V., Cowburn, D., and Mohammadi, M. (2012) The Alternatively Spliced Acid Box Region Plays a Key Role in FGF Receptor Autoinhibition, *Structure* 20, 77-88.
2. Zhao, W., Zheng, B., and Haynie, D. T. (2006) A Molecular Dynamics Study of the Physical Basis of Stability of Polypeptide Multilayer Nanofilms, *Langmuir* 22, 6668-6675.
3. Savelieff, M. G., Lee, S., Liu, Y., and Lim, M. H. (2013) Untangling amyloid- β , tau, and metals in alzheimer's Disease, *ACS Chem. Biol.* 8, 856-865.
4. Norrby, E. (2011) Prions and protein-folding diseases, *J. Intern. Med.* 270, 1-14.