

University of Arkansas, Fayetteville

ScholarWorks@UARK

---

Chemistry & Biochemistry Undergraduate  
Honors Theses

Chemistry & Biochemistry

---

5-2016

## Engineering a Mutation in the Heparin Binding Pocket of the Human Fibroblast Growth Factor

Roshni Patel

*University of Arkansas, Fayetteville*

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



Part of the [Amino Acids, Peptides, and Proteins Commons](#), [Biochemistry Commons](#), [Biotechnology Commons](#), [Carbohydrates Commons](#), [Cell Biology Commons](#), [Molecular Biology Commons](#), [Molecular Genetics Commons](#), [Other Biochemistry](#), [Biophysics](#), and [Structural Biology Commons](#), [Other Microbiology Commons](#), [Pharmaceutics and Drug Design Commons](#), and the [Structural Biology Commons](#)

---

### Citation

Patel, R. (2016). Engineering a Mutation in the Heparin Binding Pocket of the Human Fibroblast Growth Factor. *Chemistry & Biochemistry Undergraduate Honors Theses* Retrieved from <https://scholarworks.uark.edu/chbcuht/14>

This Thesis is brought to you for free and open access by the Chemistry & Biochemistry at ScholarWorks@UARK. It has been accepted for inclusion in Chemistry & Biochemistry Undergraduate Honors Theses by an authorized administrator of ScholarWorks@UARK. For more information, please contact [scholar@uark.edu](mailto:scholar@uark.edu).

# Engineering a Mutation in the Heparin Binding Pocket of the Human Fibroblast Growth Factor

An Honors Thesis by Roshni Patel

March 28, 2016

## Table of Contents

<b>Acknowledgements.....</b>	<b>3</b>
<b>Abstract.....</b>	<b>4</b>
<b>Introduction.....</b>	<b>5</b>
<b>Experimental Methods and Materials.....</b>	<b>16</b>
<i>Bacterial Transformation of T137E FGF1.....</i>	<i>16</i>
<i>Small scale expression (SSE) of T137E- FGF1.....</i>	<i>16</i>
<i>Large scale expression (LSE) of T137E-FGF1.....</i>	<i>17</i>
<i>Purification of T137E-FGF1.....</i>	<i>18</i>
<i>Intrinsic Fluorescence Spectroscopy of T137E FGF.....</i>	<i>19</i>
<i>8-Anilinonaphthalene-1-sulfonic acid (ANS) Binding Assay.....</i>	<i>19</i>
<i>Far UV - Circular Dichroism (CD) Spectroscopy.....</i>	<i>20</i>
<i>Molecular Dynamics Simulation of T137E-FGF1.....</i>	<i>21</i>
<b>Results and Discussion.....</b>	<b>22</b>
<i>Expression and Purification of WTFGF1.....</i>	<i>22</i>
<i>Expression and Purification of T137E-FGF1.....</i>	<i>23</i>
<i>Biophysical Characterization of T137E-FGF1.....</i>	<i>27</i>
<b>Conclusions.....</b>	<b>33</b>
<b>References.....</b>	<b>35</b>

## **Acknowledgements**

I would like to thank Dr. T.K.S. Kumar, Dr. Srinivas Jayanthi, and Julie Eberle for being outstanding mentors during my time in the Kumar Lab. I would also like to thank the rest of the members of the Kumar Lab for their constant support and encouragement. This work is supported by NIH(P30 GM 103450), NSF grant IOS-0842937, DOE grant DE-02-01ER15161, the Arkansas Bioscience Institute grants to TSKS, and University of Arkansas Fulbright Honors College Research Grant. I would also like to thank the Honors College for allowing me to present this research project at the Southeast/Southwest American Chemical Society (ACS) Regional meeting in Memphis, TN November 7, 2015.

## **Abstract**

Fibroblast growth factors (FGFs) are family of proteins that belong to a group of growth factors that are found in mammals and play an important role in angiogenesis, differentiation, organogenesis, and tissue repair. In summary, their main functionality is involved in cell division and proliferation. Because FGFs plays such a vital role in cell proliferation, they are mainly involved in the process of wound healing and injuries. FGF binds to its ligand, heparin—a heavily sulfated glycosaminoglycan. The binding of heparin to FGF occurs through electrostatic interactions, specifically between the negatively charged sulfate groups on heparin and positively charged residues such as arginine and lysine in the heparin binding pocket of FGF.

FGF1, a prototype of the FGF family, has many potential applications since it is heavily involved in wound healing, however, FGF1 does not remain active for very long when it is not bound to heparin. With this in mind, this research project focuses on increasing the half-life of FGF1 while maintaining its stability. To achieve this objective, residue threonine at position 137, which is located near the heparin binding pocket was mutated to glutamic acid. Preliminary biophysical characterization of the mutant FGF1 protein has been discussed in this dissertation.

## **Introduction:**

The human acidic fibroblast growth factor (hFGF1) is a heparin binding protein and is a part of a larger protein family of 23 members (1, 2). Fibroblast growth factors (FGF) play an exceptional role in various fundamental cellular processes in humans. Many of these cellular processes are activated when FGF binds to heparin and upon interaction with its cell surface receptor. The resulting cell signaling pathways initiate processes such as embryogenesis, cell proliferation and differentiation, angiogenesis, and wound healing (2, 4, 5). Because of FGF's extensive involvement in important human cellular processes, FGF has significant therapeutic potential, specifically for wound healing. As soon as an injury occurs, the tissue repair process begins by releasing growth factors, which initiate cell proliferation in the wounded area (6). However, despite FGFs broad involvement in critical cellular processes, it is well known for its instability toward thermal and proteolytic degradation (2). As a consequence, further research towards engineering a more robust FGF1 can pioneer new discoveries that can potentially improve FGF1s ability to initiate wound healing processes. Many pharmaceutical companies have even recognized FGF's application as a bio-therapeutic drug candidate for wound healing, which has been of great interest for many researchers in this field (12).

The 23 members of the FGF family can be found in different organisms from nematodes to humans (2). These 23 FGFs have been grouped into 6 subfamilies according to their homology and phylogeny. It is important to note however, that FGF11-14 do not activate FGF receptors and are therefore not assigned to a subfamily (2, 7). Additionally, FGF15 has been shown to be the mouse orthologue of the human FGF19 (7). Taking FGF11-14 homologs and

FGF15 mouse orthologue out of consideration, there are a total of 18 mammalian fibroblast growth factors. The 6 FGF subfamilies are FGF1 (FGF1 and FGF2), FGF4 (FGF4, FGF5, FGF6), FGF7 (FGF3, FGF7, FGF10, FGF22), FGF8 (FGF8, FGF17, FGF18), FGF9 (FGF9, FGF16, FGF20), and FGF19 (FGF19, FGF21, FGF23) (7, 10). The subfamilies are shown in the figure below, with the seventh subfamily being the FGF11-14 homologs (8,10).

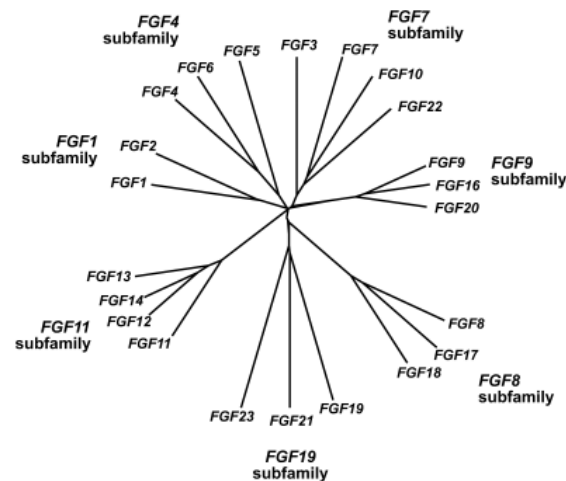


Figure 1: FGF subfamilies. FGF homologs are represented in the FGF11 subfamily (8).

The FGF subfamilies were categorized based on their various functional properties, sequence similarities, and modes of signaling (2). The first five of the subfamilies are involved in paracrine signaling, which act locally, and the last subfamily (FGF19, 21, 23) is involved in endocrine signaling, which signal across longer distances (7,9,10).

All of the human FGF genes are expressed at unique sites and many are even differentially expressed in tissue. FGF 3,4,8,15,17,19 are expressed during embryonic development only, whereas FGF 1,2,5,6,7,9-14,16,18,20-23 are expressed in embryonic and adult tissues (2). Gene locations for all 22 FGF genes are known with the exception of FGF16.

There is also a variety in gene location for the human FGF genes because they are created through translocation and chromosomal duplication during evolution (2). The evolutionary history of FGF shows that human FGF and other species share about 90% of the same amino acid sequence, except for FGF15 (the mouse orthologue of human FGF19) (2). In general, all FGFs consist of 150-300 amino acids with a molecular weight ranging between 16-34 kDa (2, 10). Crystal structures of at least one FGF member from each subfamily have been found, and these structures have generated further insight into the FGF protein core (10).

This dissertation will specifically focus on FGF1, known as human acidic FGF. The crystal structure of FGF1 has shown that the protein core consists of twelve antiparallel  $\beta$  strands within a conserved  $\beta$  trefoil fold (2,10). Six  $\beta$  strand pairs within the  $\beta$  trefoil fold create a  $\beta$ -barrel at the base of three of the  $\beta$  strand pairs, and five of the pairs are designed as a hairpin structure (3, 13). The figure below shows the structure of a biologically active human FGF1 bound to heparin (13).

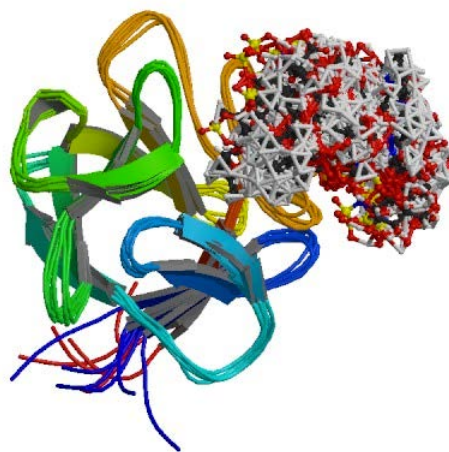


Figure 2: Human FGF1 complexed to heparin oligosaccharide (13).



In order for FGF1 to be biologically active, it must bind to heparin or heparan sulfate proteoglycans, which stabilizes FGF1 from thermal denaturation and proteolytic degradation (2). Heparan sulfate (HS) is a linear polysaccharide molecule that is a part of heparan sulfate proteoglycans (HSPG), which belongs to a family of glycosaminoglycans (14, 15). Heparan sulfate glycosaminoglycans (HS-GAG) are attached to a protein core, which make up proteoglycans, and the proteoglycans are extruded from the cell surface into the extracellular space (16). This creates a coat of HS-GAG on the cell surface, which serves as a binding region for growth factors, enzymes, and other signaling molecules in order to activate biological processes (16). HS-GAGs play a crucial role in numerous biological processes such as cell growth and development, angiogenesis, anticoagulation, and viral invasion (16, 17).

As a linear and heterogeneous polysaccharide, HS-GAG is formed of alternating units of D-glucosamine (GlcN), and an iduronic acid, D-glucuronic (GlcA), to form a linear chain that can widely vary in length (16, 18). Figure 3 shows a disaccharide structure of both HS and heparin with GlcN and GlcA (22). These disaccharide units can be highly sulfonated throughout the chain, giving the molecule an overall negative charge (16, 18, 19). It is also important to note that heparin forms an alpha helical secondary structure (27). Heparin is more heavily sulfated in comparison to HS (16, 19). Due to the high affinity of FGF1 towards heparin, this molecule is utilized as a tool to purify FGF1 through affinity chromatography (21).

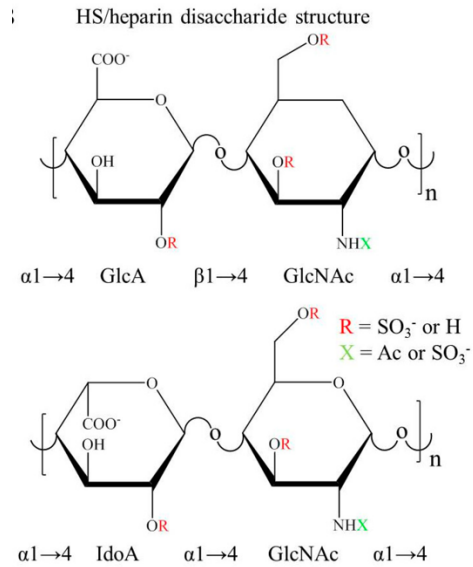


Figure 3: Comparison of the heparin (disaccharide units)

repeating unit of heparin sulfate HS and (22).

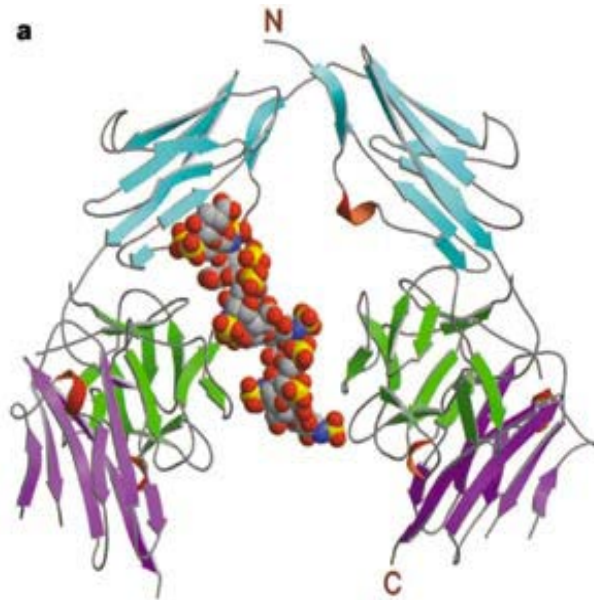


Figure 4: Binding interface between heparin (in the space-filling model) FGF1 (in the dimeric form) of the FGF and cell surface FGF receptor (23).

FGF1-heparin interaction is a critical interaction preceding FGF1 receptor binding. In order for FGF1 to mediate its cellular response, it is necessary for it to bind to one of four receptor tyrosine kinases (RTK), also known as the FGF receptors (FGFRs) (24). Heparin binding

to FGF1 both stabilizes and protects the protein from proteolytic degradation so that it may then bind its FGFR (11, 24). There are four FGFRs, all of which are located on the cell surface and contain an extracellular, transmembrane, and intracellular domain (3). The cytoplasmic domain contains the protein tyrosine kinase core and extra regulatory sequences (1, 9, 24).

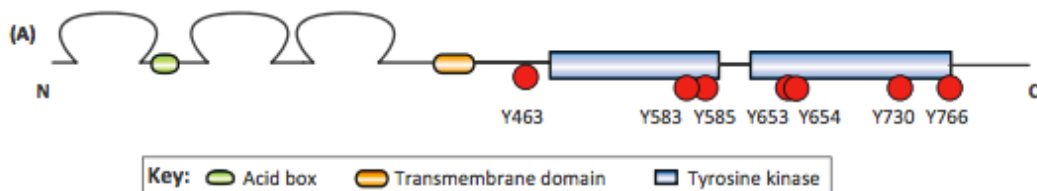


Figure 5: Depiction of the structure of FGF receptor showing the three main components: extracellular ligand-binding domain, the transmembrane domain, and the tyrosine kinase core (9).

The extracellular ligand-binding domain contains three immunoglobulin (Ig)-like domains referred to as D1-D3. In between D1 and D2, there are seven to eight acidic residues that make up a region known as the acid box. The D2-D3 region of the receptor allows for ligand binding, whereas D1 and the acid box play roles in auto inhibition (7, 24).

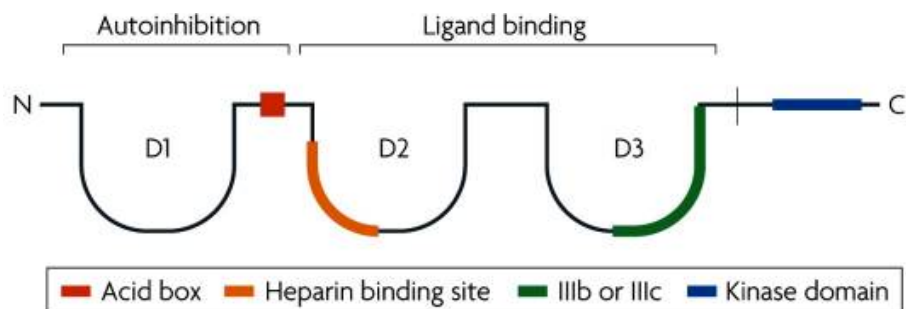


Figure 6: Structural modules (D1-D3) present in the extracellular domain of the FGF receptor. (7).

The three immunoglobulin domains of RTKs are alternatively spliced to create several isoforms of FGFRs (7, 24). Alternative splicing occurs particularly in FGFR1-3 genes coding for the D3 segment, in which exons 8 and 9 form two additional forms of FGFR (11,24). FGFR4 does not

partake in alternative splicing. These additional FGFR isoforms exhibit unique specificity for FGF ligand binding (11, 24). Figure 7 shows how the alternative splicing process occurs. Exon 7 encodes for the N-terminal of D3, which is also termed 'a.' Exons 8 and 9 encode for the C-terminal of D3, and they are referred to as 'b' and 'c' forms (24).

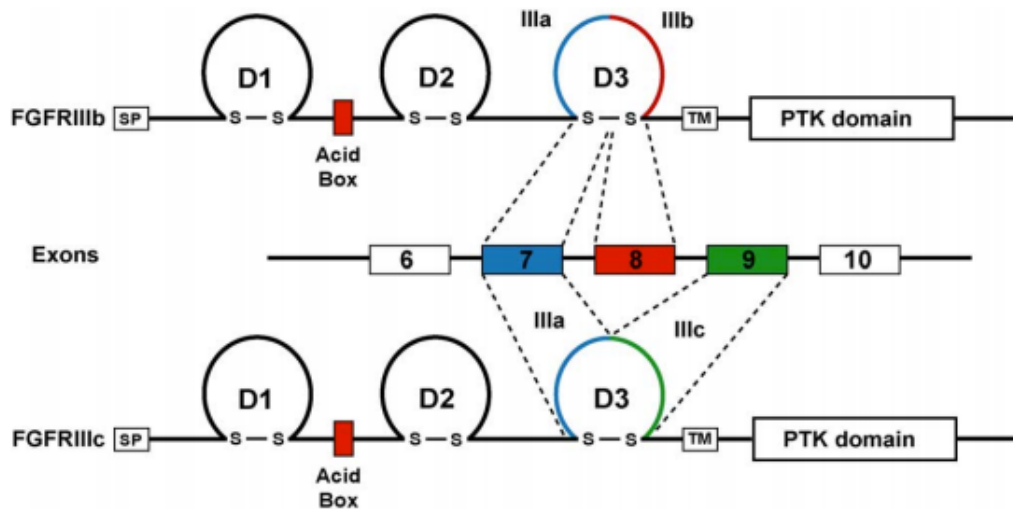


Figure 7: Alternative splicing sites in the D3 domain of FGFR to form additional isoforms of FGFR. (24).

Since FGFR1-3 are all involved in alternative splicing and each FGFR forms two additional isoforms termed as 'b' and 'c', there would be a total of seven FGFR isoforms, including FGFR4 (24). All 23 FGFs have varying ligand-binding affinity towards specific FGFR isoforms. The table below shows ligand specificities for each FGFR isoform. One significant characteristic of FGF1 that further lends towards its therapeutic potential is its ability to bind to all FGFR isoforms.

**Table 1: Ligand binding specificities of the different FGFR isoforms**

FGFR Isoform	FGF Ligand Specificity
FGFR1b	FGF1, 2, 3, 10
FGFR1c	FGF1, 2, 4, 5, 6
FGFR2b	FGF1, 3, 7, 10, 22
FGFR2c	FGF1, 2, 4, 6, 9, 17, 18
FGFR3b	FGF1, 9
FGFR3c	FGF1, 2, 4, 8, 9, 17, 18, 23
FGFR4	FGF1, 2, 4, 6, 8, 9, 16, 17, 18, 19

Once the FGF1-heparin complex binds to the D2-D3 region of the FGFR, the receptor dimerizes. FGFR dimerization initiates autophosphorylation of seven tyrosines in the protein tyrosine kinase (PTK) domain of the FGFR, thus activating various signaling pathways (7, 9, 25). In addition to the phosphorylation of the tyrosines, there are several regulatory sequences extending at the cytoplasmic PTK domain that allow for docking proteins to bind (24). In particular, the juxtamembrane region, which is longer than many RTKs, functions as a binding region for the phosphotyrosine binding (PTB) domain of FGFR substrate 2, also referred to as FRS2 (7, 9, 24). FRS2 serves as a family of docking proteins and includes FRS2 $\alpha$  and FRS2 $\beta$ . When FRS2 is phosphorylated, two signaling pathways are activated, mitogen activated protein kinase (MAPK) and phosphoinositide 3-kinase-Akt (AKT) (7, 9, 24). Phospholipase C (PLC)  $\gamma$  1, is also a docking protein which activates a specific signaling pathway (9, 26). Furthermore, one

particular intracellular tyrosine, Y766, when phosphorylated near the C-terminal, creates a binding region specific for the SH2 domain of PLC  $\gamma$  (9, 24). Once phosphorylated, activated PLC  $\gamma$  stimulates hydrolysis of phosphatidylinositol generating diacylglycerol (DAG) and Inositol triphosphate (IP3) as secondary messengers (24). All together, the four key signaling pathways are, mitogen activated protein kinase (MAPK), phosphoinositide 3-kinase (PI3K/AKT), phospholipase C  $\gamma$  (PLC  $\gamma$ ), and Janus kinase (JAK STAT) (9).

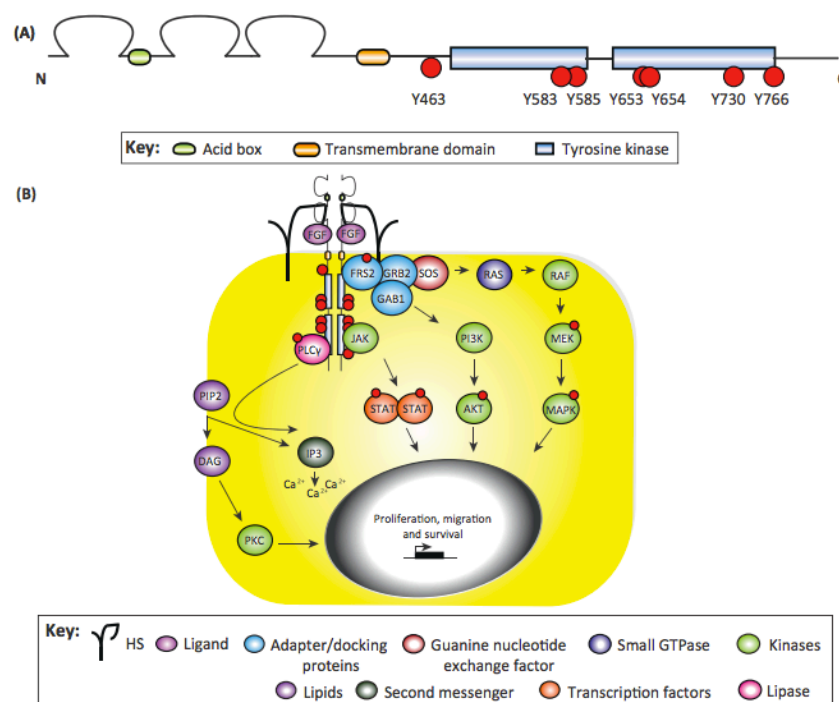


Figure 8: Panel-A, FGF receptor showing D1-D3, the transmembrane domain, and the catalytic domain with the protein tyrosine kinase core. Panel- B, Dimerization of FGFR resulting in activation of the following four main signaling pathways. MAPK, PI3K, PLC  $\gamma$ , and JAK STAT (9).

As mentioned, FGF1 has the ability to bind to all seven isoforms. This characteristic makes it of great interest in this research lab as well as in other clinical studies. FGF1 is made up of 154 amino acids; however, a truncated form of FGF1 consisting of 140 amino acids is used in Kumar's lab in order to eliminate degradation products (often produced from the first 14 amino

terminal residues) (20). Although this truncated form is shorter, it is still demonstrated to have full functionality. As mentioned earlier, FGF1 on its own is rather unstable. Under physiological conditions, nearly 50% of FGF1 is unfolded, which makes the protein more susceptible to proteases and difficult to use for medical applications such as wound healing agent, or composition in cosmetic application(s) etc. (20). Previous mutagenic research in the heparin-binding region of FGF1 has generated more robust forms of FGF1 with increased half-life and stability. Residues in the heparin binding region of FGF1 are largely positively charged arginines and lysines (27). Thus the interaction between heparin and FGF1 is mainly governed by electrostatic forces. Figure 9 below shows the heparin-binding region of FGF1. Specifically, heparin-binding residues are N32, K126, K127, K132, K142, R133, and R136.

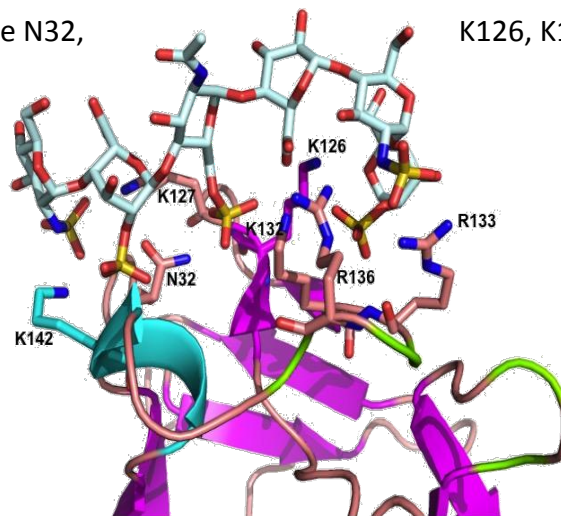


Figure 9: Heparin-binding region of FGF1 bound to heparin. The heparin binding residues are also labeled on this figure.

Previous research in the Kumar lab generated an R136E-FGF1 mutant where arginine at position 136 was replaced with a glutamic acid. Characterization of this R136E-FGF1 mutant

revealed decreased affinity for heparin, increased thermal stability and drastically increased cell proliferation activity. One of the neighboring residues of R136 is threonine at position 137, shown in figure 10. T137 will be the focus of this research project. Threonine, a polar amino acid, will be mutated to glutamic acid, a negatively charged amino acid with a longer side chain. Introduction of a glutamic acid at this position is predicted to generate repulsive electrostatic interactions that may reduce the affinity of FGF1 towards heparin molecule.

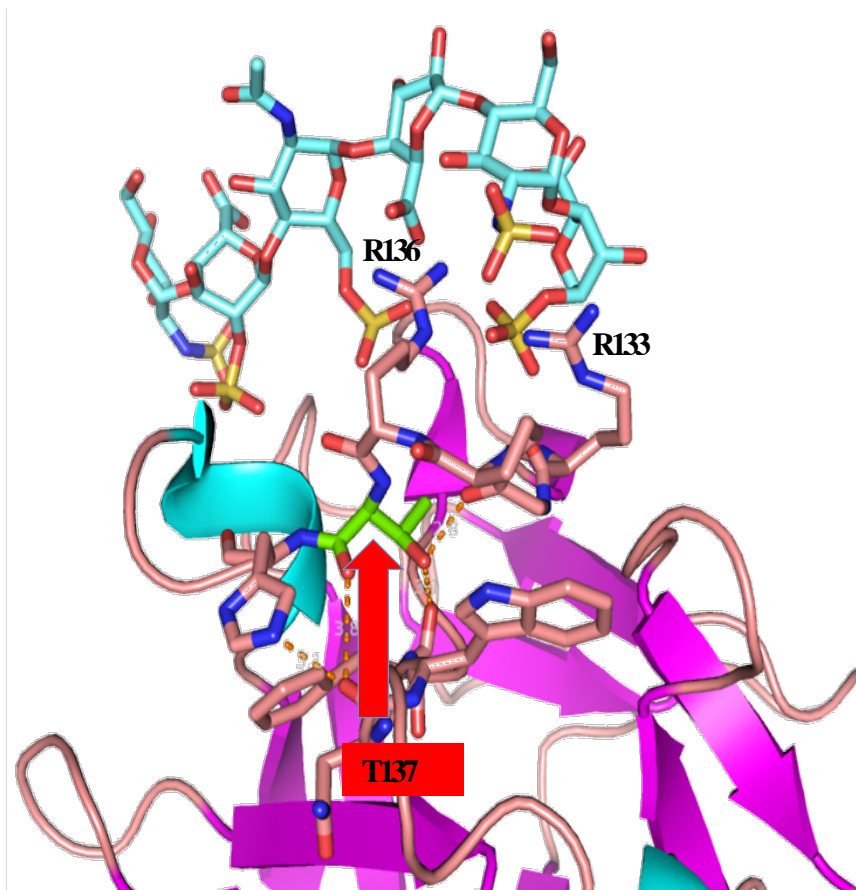


Figure 10: R136 and its neighboring residue T137, which has been mutated to E137.



## Experimental Methods and Materials

### Bacterial Transformation of T137E-FGF1:

Mutant T137E- FGF1 plasmid DNA was generated using Quick Change site directed mutagenesis kit as per the manufacturer instructions. Two types of expression hosts BL-21 (DE3) and BL-21 star *Escherichia coli* (*E. coli*) competent cells were transformed with T137E plasmid. Transformed BL-21(DE3) and star cells were selected on agar plates containing 100µg/ml of ampicillin (AMP) as a selection marker. Transformation was completed using the heat shock technique according to protocol well established in the Kumar lab.

### Small Scale Expression (SSE) of T137E-FGF1:

SSE of T137E-FGF1 protein was performed to verify the presence of the mutant protein in the cells and the ability of these cells to produce the desired protein. This is accomplished by growing cells overnight in lysogeny broth (LB) at 37°C and inoculating fresh LB the next morning with the overnight culture. After approximately 2 hours, the optical density (OD) of the culture was checked using UV-visible spectroscopy at a wavelength of 600 nm. *E. coli* cells replicate approximately every 20 minutes; therefore growth curve hits an exponential rate in the log phase at an OD between 0.4-0.6.

When the culture had reached the desired OD, it is induced with Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG), which activates the lac operon in order to induce rapid protein expression. Four hours later, cells were harvested. The cell culture is then lysed using ultrasonication, which releases cytosol content into solution. The cell debris is separated from solution—which contains our protein of interest—using high-speed centrifugation at 19000x rpm for 30 minutes. Samples of the culture are taken pre-induction, post induction, and from the insoluble pellet and supernatant following centrifugation. These samples were resolved on sodium dodecyl sulfate poly acrylamide gel electrophoresis (SDS-PAGE) to monitor protein expression.

#### Large Scale Expression (LSE) of T137E-FGF1:

LSE was performed to generate large quantities of T137E-FGF1 so that the protein may be purified for characterization studies. This expression follows the same protocol as SSE except with larger quantities of autoclaved LB media. 1 mL of glycerol stock of T137E-FGF1 and AMP (100 $\mu$ g/ml) was added to a starter flask and was incubated over night at 37°C with 250rpm. The next day, fresh, sterile LB was inoculated with the overnight culture and AMP and incubated at 37°C. At OD the culture was induced with IPTG and harvested 4 hours later using high-speed centrifugation. After centrifugation, the clear broth was discarded and the bacterial cell pellet containing the protein of interest was stored at -20°C.

### Purification of T137E-FGF1:

Purification of WTFGF1 and T137E-FGF1 was performed using affinity chromatography on a 20 mL heparin sepharose column. The pellet obtained from LSE was thawed and resuspended in 25 mL of 10 mM phosphate buffer containing 10 mM NaCl at pH 7.2. The cells were lysed using ultra sonication and the clear lysate was separated from cell debris by centrifugation at 19,000 rpm for 30 minutes. Immediately after the centrifugation, the supernatant was separated and loaded onto the pre-equilibrated heparin sepharose column. A smear of the pellet was collected for analysis via SDS-PAGE.

After loading, the supernatant was incubated on the column for 30 minutes to insure ligand binding. The protein was then purified of contaminant proteins by washing the column with 10 mM phosphate buffer (PB) at pH 7.2 containing increasing salt gradients. All washes were collected and samples from each wash were prepared with trichloroacetic acid (TCA) precipitation protocols for SDS-PAGE analysis. The purity of the protein was verified from the SDS-PAGE gel represented by a distinct band weighing 16kDa. The pure protein fraction collected at 1,500 mM NaCl was then buffer exchanged to 100mM NaCl and concentrated using a Millipore centrifugal concentrator with a 10kDa molecular weight cutoff. The concentration of the protein was measured using a NanoDrop at A280 nm.

Similar purification strategy was employed for BL-21 Star expression host. These cells are known to contain less proteolytic enzymes and were therefore predicted to increase

protein yield. In addition, protease inhibitors, phenylmethanesulfonyl fluoride (PMSF) and ethylenediaminetetraacetic acid (EDTA), were added to each wash buffer. Lastly, 25 mM ammonium sulfate (AS) was added to each wash buffer in order stabilize T137E-FGF1 and to avoid protein aggregation. Unfortunately analysis of protein elute via SDS-PAGE, still revealed the presence of multiple lower weight bands in addition to at 16KDa protein band. To determine whether the additional bands were degradation products or contaminants, the protein fraction was dialyzed into 100 mM NaCl/25 mM AS/10mM PB at pH 7.2. The dialyzed protein was then concentrated and characterized with biophysical techniques.

#### Intrinsic Fluorescence Spectroscopy of T137E-FGF1:

Intrinsic fluorescence spectroscopy was performed of T137E-FGF1 and WTFGF1 with and without heparin to assess any changes in structure and folding of the protein using a Hitachi F-2500 spectrofluorometer. The samples for this experiment were prepared using a protein concentration of 0.1 mg/mL in 10 mM PB which contained 100 mM NaCl, pH 7.2. All of the samples were excited at a wavelength of 280 nm, while the data was collected between 300 and 450 nm in order to analyze the fluorescence of tyrosine at 308 nm and tryptophan at 350 nm.

#### 8-Anilinonaphthalene-1-sulfonic acid (ANS) Binding Assay of T137E-FGF1:

ANS binding was performed to analyze any conformational changes in WTFGF1 & T137E-FGF1 with and without heparin using a Hitachi F-2500 spectrofluorometer. ANS is a hydrophobic probe that binds to pockets of hydrophobic residues and gives information on how well the protein is folded. The excitation wavelength was set to 380 nm, and the emission wavelength range was between 450-600 nm because ANS fluoresces with an emission maxima at around 500 nm. Samples were prepared by diluting the protein to 15  $\mu$ M, and an ANS stock solution was also prepared that allowed the addition of 1  $\mu$ L of stock ANS to the protein sample. Adding 1  $\mu$ L of the stock ANS into the protein allows the ANS concentration to increase by 20  $\mu$ M. The relative fluorescence intensity (RFI) was recorded until the ANS concentration reach 400  $\mu$ M. Data was plotted with concentration of ANS on X-axis against ANS

#### Far UV Circular Dichroism (CD) Spectroscopy of T137E-FGF1:

Far UV CD spectroscopy was utilized to assess the secondary structure of T137E-FGF1 in comparison to WTFGF1 with and without heparin using a JASCO 1500 CD spectrometer. A quartz cell of 0.2 mm path-length was loaded with 100  $\mu$ L of T137E-FGF1 protein sample. The data was scanned between a wavelength range of 190-250 nm and with 3 accumulation scans. Once collected, the data was smoothed and the buffer signal was appropriately subtracted. These steps were repeated additionally in order to collect data of the samples with heparin.

### Molecular Dynamics Simulation of T137E-FGF1:

A T137E mutant file of FGF1 was created from the WTFGF1 protein data bank structure 1RG8 (13). Molecular dynamic simulations of T137E-FGF1 were run using this base file. The protein was solvated after the structure of T137E had been minimized which left a minimum distance of 10 Å between the protein and solvent box edge. After neutralizing the system with NaCl, the effective salt concentration was set to 0.150 M. The structure was prepared for production runs by subjecting the system to numerous refinement steps in order to relax the side chains, backbone, and the molecules and ions from the surrounding water. The isothermal-isobaric (NPT) ensemble was used to equilibrate the system along with NAMD 2.9 and the CHARMM36 forcefield.

Using Langevin dynamics, the temperature was maintained at 300 K along with a damping coefficient of 1 ps<sup>-1</sup>. The Langevin piston method was used to maintain a constant pressure of 1 atm, period of 100 fs, and a decay time of 50 fs. Next, the particle mesh Ewald method with periodic boundary conditions was utilized in order to measure long-range electrostatic interactions. Once these interactions were computed, switching functions beginning at 10 Å were used to cut off electrostatic and Van der Waals interactions beyond a distance of 12 Å, and all of the hydrogen bonds were maintained rigid. The simulation time step was set to 1fs.

Minimization of the side-chains required 10,000 steps, while the backbone atoms were fixed as water and ions were not present. Following the side-chain minimization, the protein

structure was solvated. Fixing all of the protein atoms and subjecting the system to 1,500 steps of minimization and 50 ps of dynamics relaxed the water molecules around the protein. To further relax the solvent system, harmonic constraints with a force constant of 1 kcal/ (mol Å<sup>2</sup>) were utilized on the protein atoms. A gradual increase in temperature from 0 K to 300 K at a rate of 10K per 2ps was completed before running the production followed by 250 ps of dynamics. Based upon the leveling of RMSD and energy fluctuations, production run preparedness was determined.

The stability of the protein domains and protein-protein complexes was determined with 50 and 100 ns simulations. For the production runs, a 2 fs time step was used, while the remaining parameters were the same as the last 250 ps from the final equilibration run. The RMSD trajectory tool, visualized simulations, and the C-  $\alpha$  measured distances from VMD, were used to evaluate backbone RMSDs. VMD was also used to collect simulations for hydrogen bonding, solvent accessible surface area (SASA), and the gyration radius.

## Results and Discussion:

### Expression and Purification of WTFGF1

Figure 11 is the SDS PAGE analysis of the purification of WTFGF1. The expression and purification of WTFGF1 was successfully verified from lane 5 with a distinct thick band.

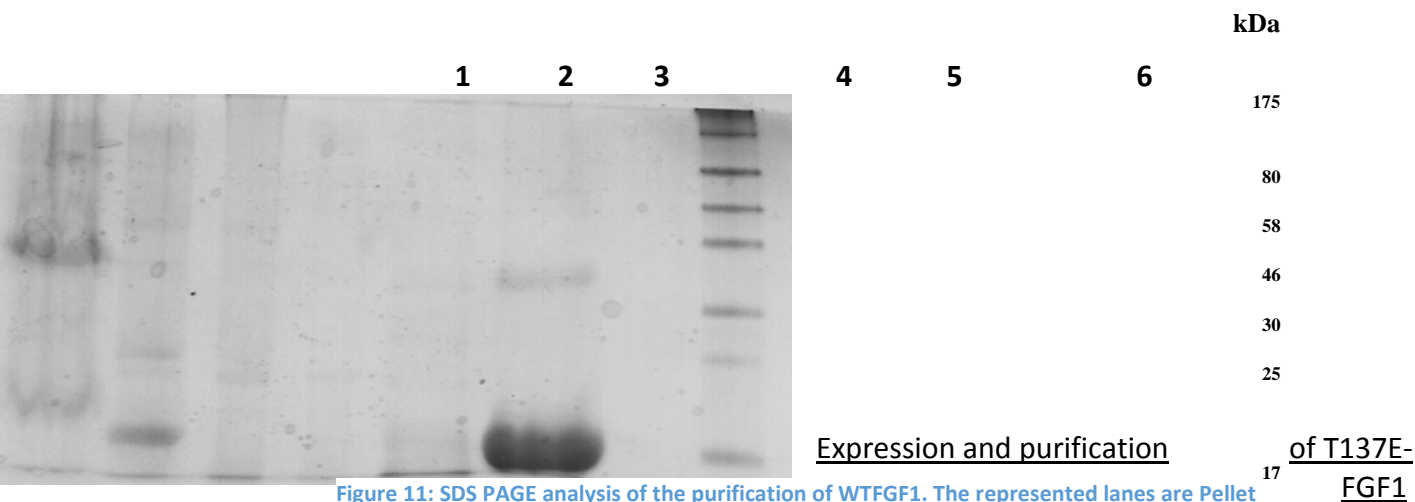


Figure 11: SDS PAGE analysis of the purification of WTFGF1. The represented lanes are Pellet (Lane-1), Supernatant (Lane-2), 100 mM NaCl (Lane-3), 800 mM NaCl (Lane-4), 1500 mM NaCl (Lane-5), Protein Marker (Lane-6). This gel shows that WTFGF1 has been successfully expressed and purified as there is a thick band in lane 5 at 1500 mM NaCl.

Figure

12 is an SDS PAGE analysis of a purification of T137E-FGF1 expressed in BL-21(DE3) cells (the first cell line used for the over expression of the protein). Lane 6 is the protein elute in 1500mM NaCl represented by a 16KDa band. However, there is also a band above and below the band of interest. The band above could be a possible contaminant protein, while the band below could be either a contaminant or a degradation product. The protein yield from purification was less



than 1 mg from a liter of bacterial culture which is very low compared to that of WTFGF1 approximately 60mgs (data not shown)

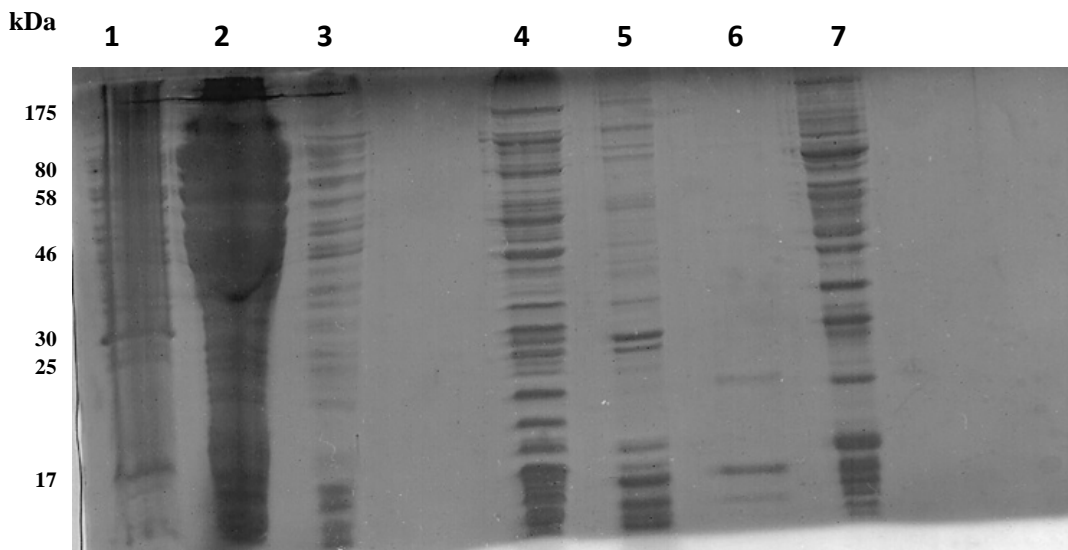


Figure 12: SDS PAGE analysis of purification of T137E-FGF1 expressed BL-21 (DE3) cells using a heparin sepharose column. The represented lanes are Protein Marker (Lane-1), Pellet (Lane-2), Supernatant (Lane-3), 500 mM NaCl/10 mM PB (Lane-4), 800 mM NaCl/10 mM PB (Lane-5), 1,500 mM NaCl/10 mM PB (Lane-6), Urea (Lane-7). In Lane-6, a band corresponding to 16 kDa size represents the presence of T137E-FGF1, however there are additional contaminant or degradation bands that are also present.

Additional purifications were also performed to collect pure protein, however, contaminant or degradation bands were still present. Due to the difficulty in collecting pure protein, another cell line, BL-21 Star, was transformed with T137E-FGF1 plasmid. This cell line was used specifically because it has fewer proteolytic enzymes. Following transformation, a small-scale expression was performed to verify the presence of T137E-FGF1. Figure 13 represents an SDS PAGE analysis of this small-scale expression. The different lanes represent snapshots of the growth phases throughout the expression. Lane 4 in particular is significant because it shows a band verifying a successful transformation. The first two lanes are T137E-FGF1 content of the cells before and after induction with IPTG. The third lane is 14kDa lysozyme

used as a marker, and the fourth and fifth lanes are the supernatant and pellet following harvest, sonication, and centrifugation of the cells.

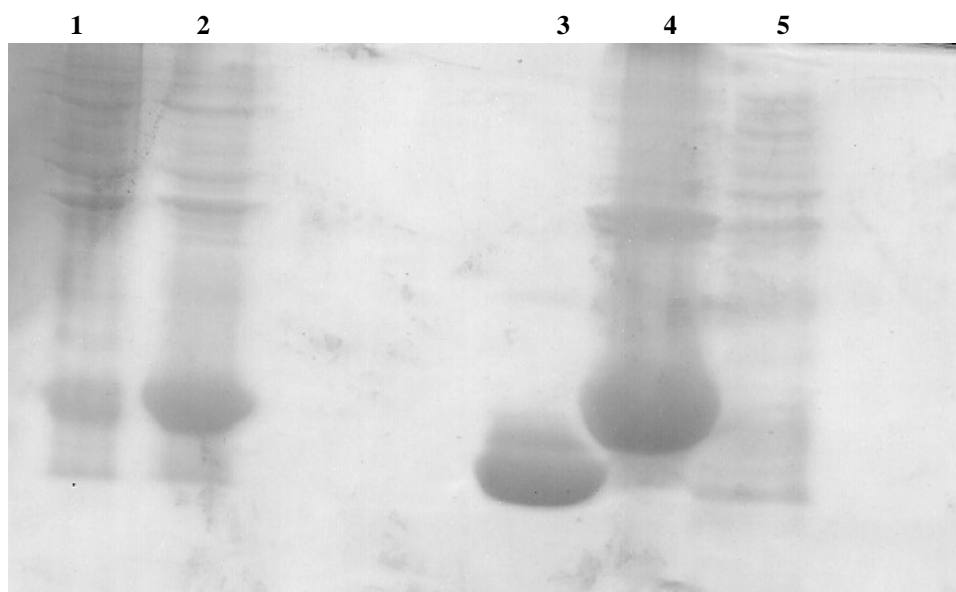


Figure 13: SDS PAGE analysis of a small-scale expression of T137E using BL-21 Star cells. The represented lanes are, Un-induced (Lane-1), Induced (Lane-2), Lysozyme (14 kDa protein molecular weight marker) (Lane-3), Supernatant (Lane-4), Pellet (Lane-5). Overall, this gel shows snapshots of the culture growth along the expression. Lane 2, the induced sample was triggered with Isopropyl B D-1-thiogalactopyranoside (IPTG), which mimics the lac operon in order to trigger protein expression at a high rate. Successful transformation and growth of T137E is verified from the band in Lane 4.

Once the transformation and growth of T137E-FGF1 was verified, purification was performed to obtain pure T137E-FGF1. The buffers used in this purification were modified in comparison to the previous purification buffers. In addition to the 10 mM phosphate and varying amounts of NaCl, 25 mM ammonium sulfate was added to the buffers as another source of salt to stabilize the protein. In addition, protease inhibitors were also added to all of the buffers to prevent any proteolytic degradation that may occur from proteolytic enzymes released into solution following sonication. Phenylmethylsulfonyl fluoride (PMSF), an inhibitor that specifically targets serine proteases (28, 29) and ethylenediaminetetraacetic acid (EDTA), a chelating agent, which can inhibit metalloid proteases were both added to the buffers (30).

Figure 14 shows the SDS PAGE analysis of the purification of T137E-FGF1 expressed in BL-21 Star, and purified with modified buffers. Lane 7 shows a band of T137E-FGF1 weighing 16kDa. However, there are additional bands below T137E-FGF1 confirming that T137E-FGF1 is not pure. The additional bands could be due to degradation or protein contaminant bands. The protein yield from this purification was approximately 0.5mgs.

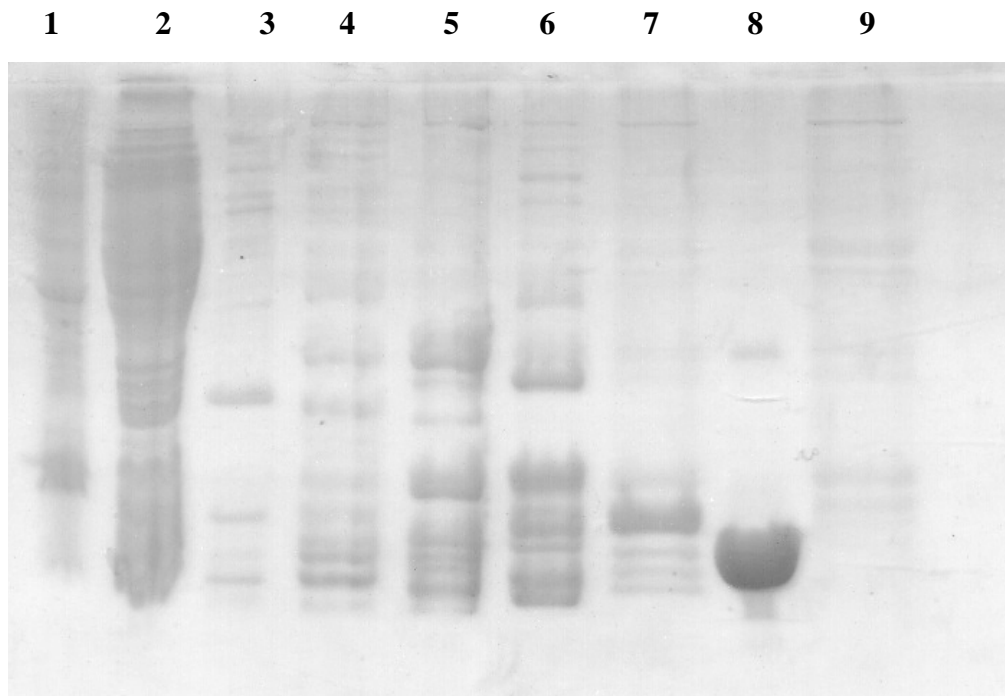


Figure 14: SDS PAGE analysis of the purification fractions of T137E-FGF1 using BL-21 Star as the cell line and modified buffers. The represented lanes are Pellet (Lane-1), Supernatant (Lane-2), 100 mM NaCl/25 mM AS/ 10 mM PB (Lane-3), 400 mM NaCl/25 mM AS/ 10 mM PB (Lane-4), 500 mM NaCl/25 mM AS/ 10 mM PB (Lane-5), 800 mM NaCl/25 mM AS/ 10 mM PB (Lane-6), 1500 mM NaCl/25 mM AS/ 10 mM PB (Lane-7), Lysozyme (14 kDa) (Lane-8), Urea (Lane-9). Lane-7 of this gel presents a band matching to 16 kDa, representing T137E-FGF1, however there are also protein contaminant or degradation bands below that even after utilizing a new less cell line with less proteolytic enzymes and buffers with protease inhibitors.

Following several rounds of purification, concentrated samples of T137E-FGF1 from both BL-21(DE3) and BL-21 Star purifications were collected. These concentrated samples were analyzed on an SDS PAGE gel along with a WTFGF1 sample, shown in figure 15. This was performed to compare the two T137E-FGF1 samples with the WTFGF. Looking at lanes 2 and 3, the T137E-FGF1 bands from BL-21(DE3) and BL-21 Star, respectively, can be seen matching up with the single WTFGF1 band in lane 1. Once again, this verifies the presence of T137E-FGF1, however, lanes 2 and 3 both show multiple bands above and below the band of interest, which represent possible contamination and degradation respectively. It is concluded that the utilization of a second cell line with ammonium sulfate and proteolytic inhibitors does not eliminate degradation or contamination bands.

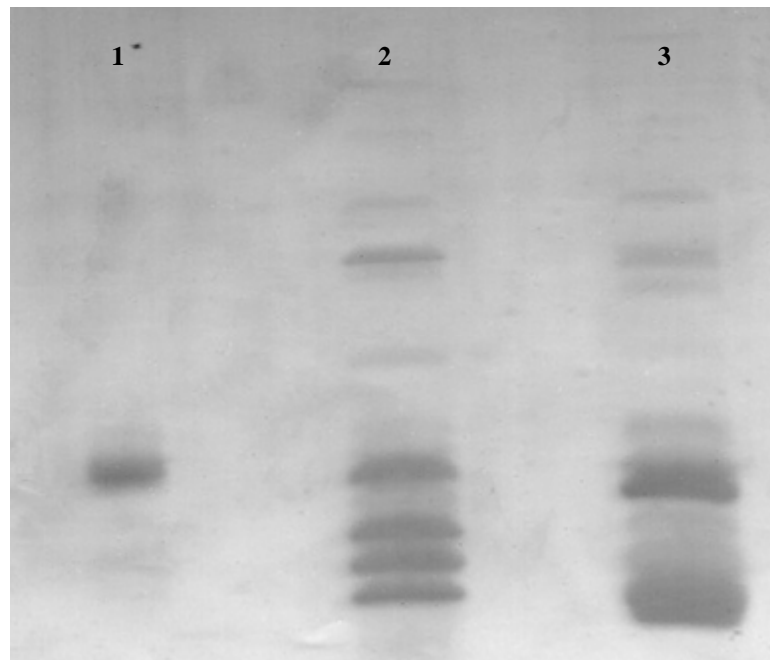


Figure 15: SDS PAGE analysis of WTFGF1 and two different samples of T137E-FGF1. The represented lanes are WTFGF1 (Lane-1), T137E (DE3 cell line) (Lane-2), T137E (Star cell line) (Lane-3). This gel shows T137E-FGF1 bands aligned with the WTFGF1 band. However, degradation or contamination bands are still present under both T137E-FGF1 bands.

## Biophysical Characterization of T137E-FGF1

### Effect of mutation on surface hydrophobicity of FGF1:

An ANS binding assay was completed to assess the structure and folding of T137E-FGF1 in the absence and presence of heparin. ANS is used as a probe to access conformational changes in protein structure upon ligand binding. The fluorescent properties of ANS differ when bound to hydrophobic residues. Therefore, the relative fluorescent intensity of ANS is representative of the amount of hydrophobic residues exposed in T137E-FGF1 in the absence and presence of heparin. Figure 16 is a graph of the ANS binding assay of T137E-FGF1 and WTGF1 with and without heparin. Comparing the mutant with the wild type, this data shows that there is not a significant difference in intensity, thus the structure and folding of the mutant does not seem to be compromised.

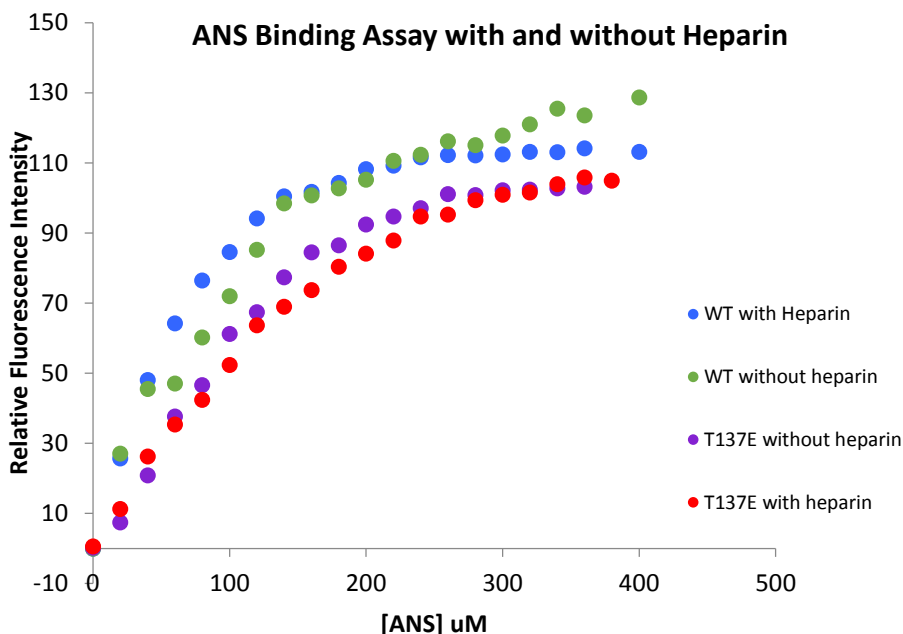


Figure 16: ANS binding to T137E-FGF1 and WTGF1 in the presence and absence of heparin. It is appears that both the mutant and wild type have similar ANS binding affinity, suggesting that the structure and folding of T137E-FGF1 has not significantly changed.

## Effect of mutation on folding of FGF1:

Intrinsic fluorescence spectroscopy was performed to analyze the changes occurring in the tertiary structure of T137E-FGF1 in comparison to WTFGF1 (figures 17 and 18). Figure 17 is the fluorescence data of the mutant from the cell line BL21 (DE3). This data shows that T137E-FGF1 has a very small peak at 308 nm, exposing tyrosine, in comparison to the WTFGF with a larger peak at 308 nm. This could be due to the small protein yield collected from the BL21 (DE3) cell line.

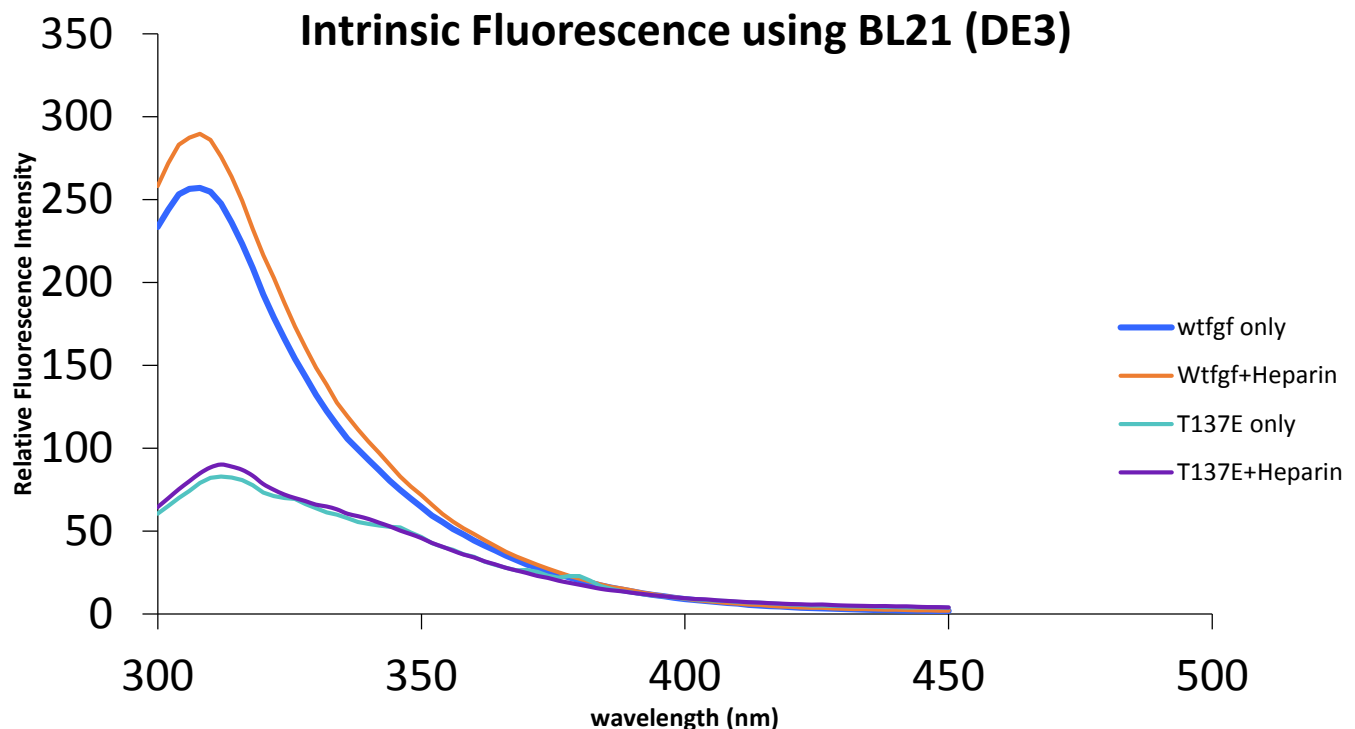


Figure 17: Intrinsic fluorescence of T137E FGF1 and WTFGF1 in the presence and absence of heparin (using BL21 (DE3) cell line).

In figure 18, the spectra of T137E-FGF1 displays a very high intensity around 350 nm compared to WTFGF1, which characteristically shows a peak only at 308nm representative of tyrosine

fluorescence. This large peak of T137E-FGF1 reveals a completely exposed tryptophan indicating that the mutant is not folded in the same manner as WTFGF1.

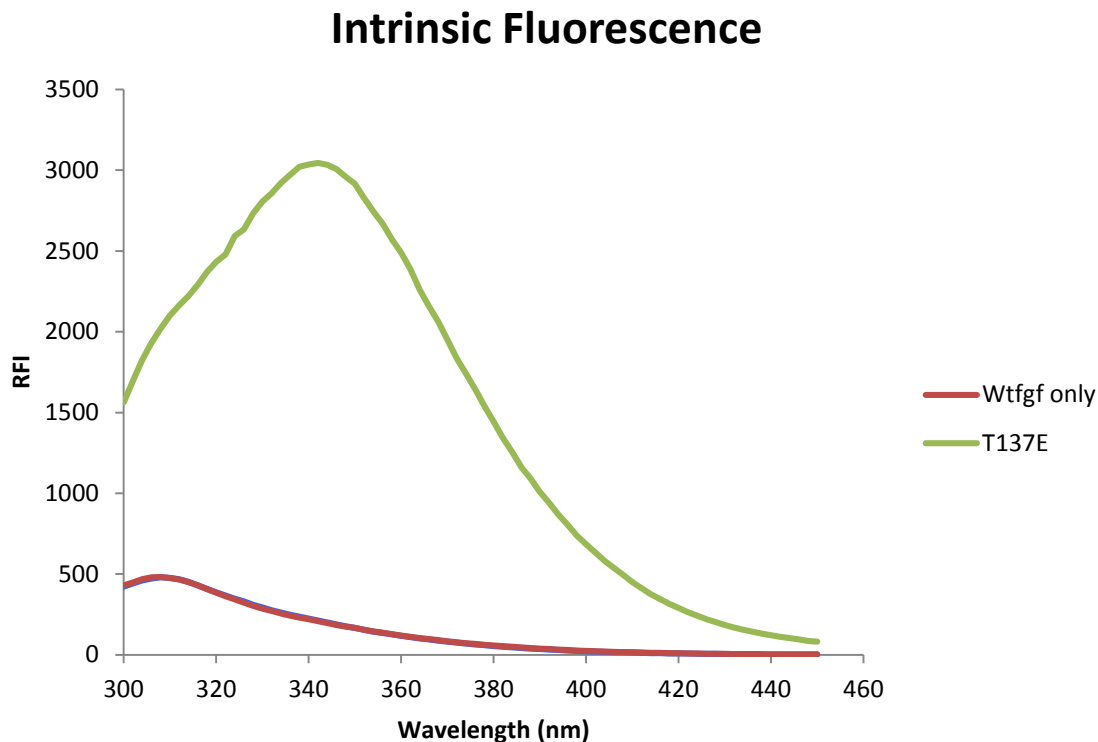


Figure 18: intrinsic fluorescence depicting the folding of T137E-FGF1 mutant shows a major peak at 350 nm, suggesting that T137E-FGF1 mutation causes the complete exposure of the lone tryptophan in the protein.

Effect of mutation on secondary structure of FGF1:

An additional experiment was performed to investigate the mutant's secondary structure by utilizing far-UV Circular Dichroism (CD) as a probe. Figure 19 shows the CD data of both T137E-FGF1 and WTFGF1 with and without heparin. The spectra of WTFGF1 shows a positive peak in the region spanning from 240-220nm, which is representative of the  $\beta$ -trefoil motif. However, the mutant spectra is drastically different and represents the secondary structure similar to that of an alpha helix. This important finding confirms that the bands present above and below the desired mutant protein band in the SDS PAGE gels of figure 4 are

indeed contaminant bands. Because FGF1 is composed of beta strands arranged in beta barrel, the alpha helical nature of the mutant spectra is due to other contaminant proteins present in the sample.

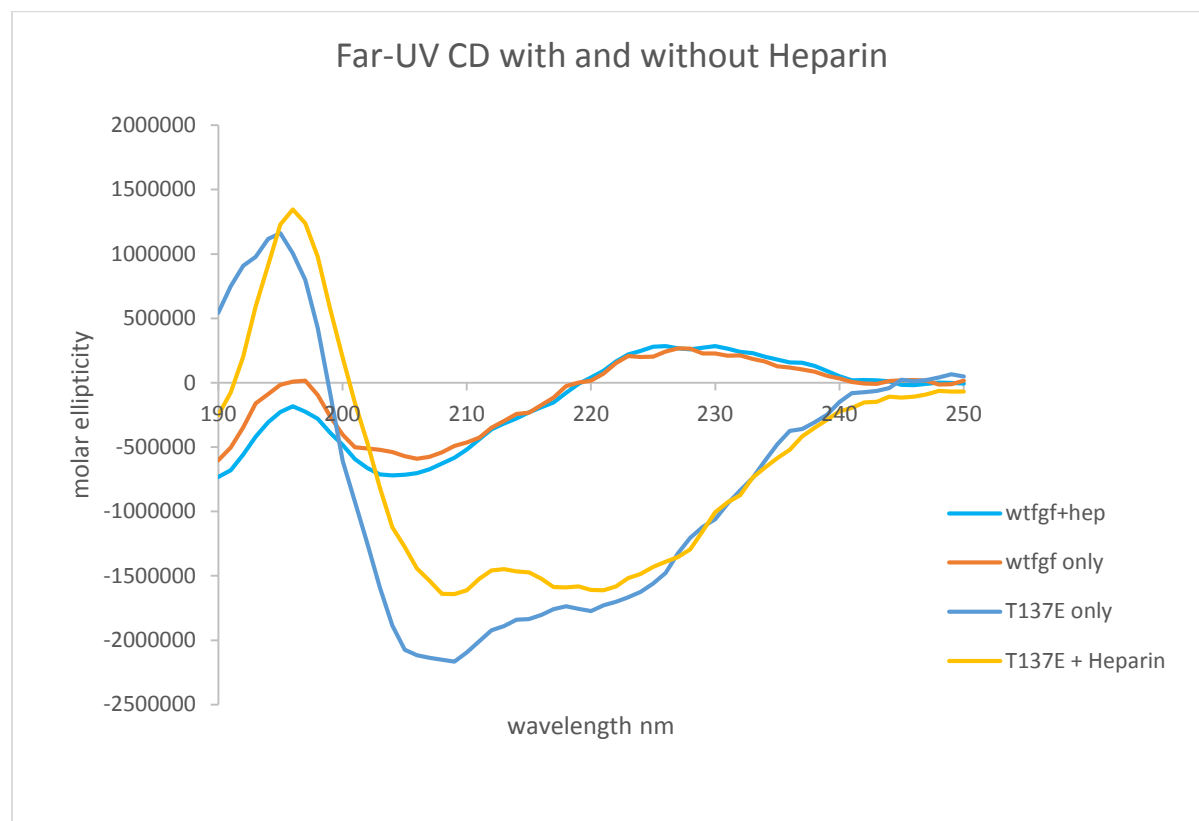


Figure 19: Far-UV CD spectra of T137E FGF1 and WT FGF1 in the presence and absence of heparin.

PyMOL, a molecular visualization program, was utilized to look further into the location and surroundings of T137. The PyMOL image of T137-FGF1, figure 20, exhibits that the functional group of threonine, a hydroxyl, is facing the hydrophobic core of the protein. Considering that threonine is a polar amino acid, this was an interesting discovery. Furthermore, it was found that the hydroxyl group of threonine was hydrogen bonded with several residues within 5 Å distance around that hydrophobic core. The following residues were a part of this hydrogen-bonding network with threonine: glycine at position 134 (G134), phenylalanine at position 122



(F122), and asparagine at position 120 (N120). Glycine is an amphipathic molecule, meaning that it contains a non-polar end as well as a polar end. Phenylalanine is a non-polar molecule, while asparagine is a polar molecule. These three residues are all within 3 Å from the threonine. The significance of this hydrogen-bonding network is that T137 is very important for stabilizing the hydrophobic core as well as the heparin-binding region. In particular, T137 is stabilizing two important loops that are in the heparin-binding pocket. Since T137 was mutated with E137, it was predicted that the heparin-binding region and the hydrophobic core would be severely disrupted.

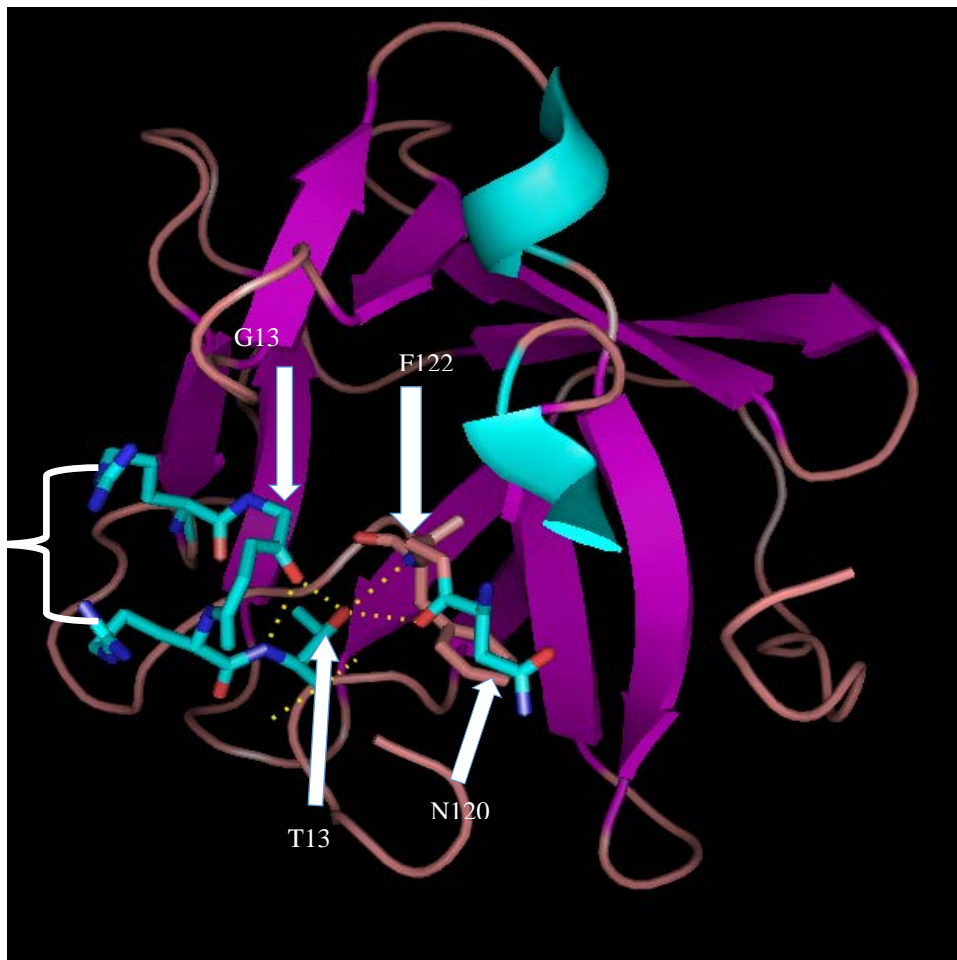


Figure 20: PyMOL image of FGF-1 showing the hydrogen-bonding network created by T137. T137 is hydrogen-bonded with G134, F144, and N120.

The following PyMOL image, figure 21, displays the mutation, E137. As predicted, this image shows loop disruption around the heparin binding region. The functional group of glutamic acid is shown on the outer surface of the protein, as it is a polar amino acid. However, as the amino acid is shifted outwards away from the hydrophobic core where the threonine was located, the hydrogen-bonding network with G134, F122, and N120, is completely disrupted. In fact, these residues are within 11 Å in distance from the glutamic acid, compared to being within 3 Å in distance from the threonine. The loop disruption caused by the mutation interrupted the heparin binding region by unfolding it.

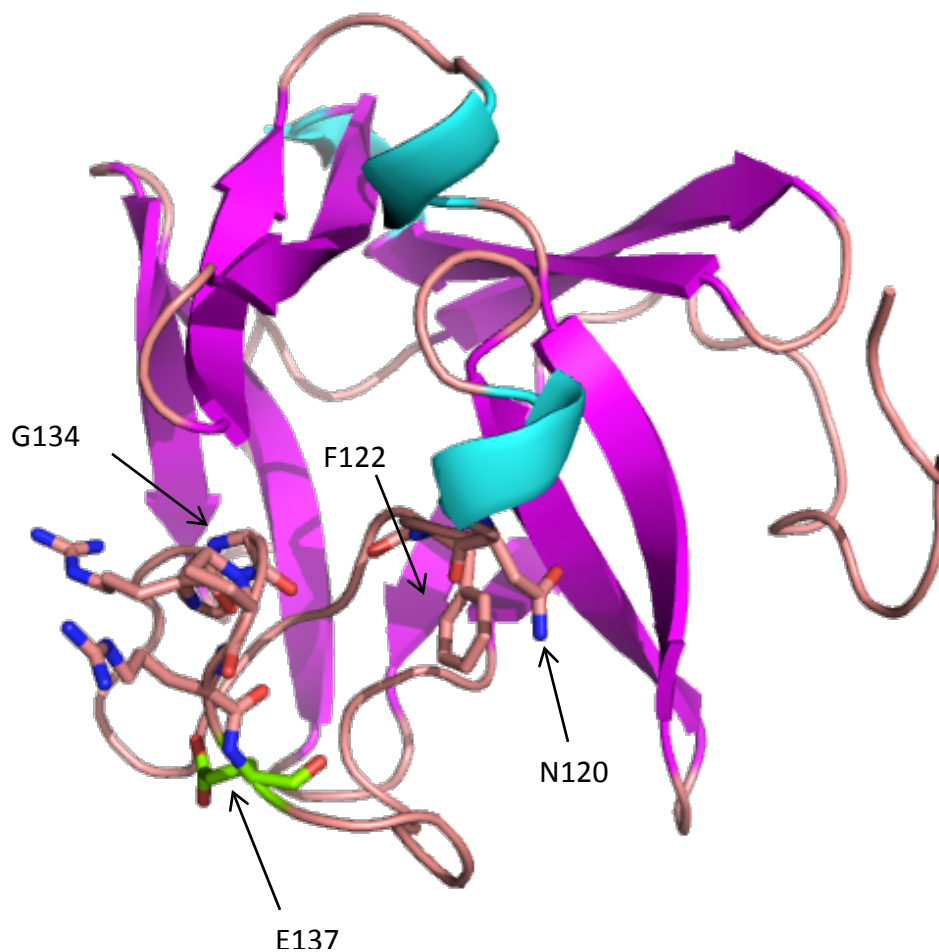


Figure 21: PyMOL image of FGF1 E137. This structure shows that the hydrogen-bonding network has been disrupted due to replacement of threonine with glutamic acid at position 137.

## Conclusions:

After several purifications of T137E-FGF1, it was concluded that homogeneous form of T137E-FGF1 is very difficult to obtain due to the inherent instability of the protein molecule generated because of the exposed loop from the disruption in hydrogen bonds. Initially, BL-21 (DE3) cell line was used to perform a bacterial transformation and move further in obtaining pure T137E-FGF1 via protein purification. The first purification of T137E-FGF1 using the BL-21 (DE3) cell line showed that there was very low expression of the protein. Comparing this to the purification of T137E-FGF1 using the BL-21 Star cell line, BL-21(DE3) showed less contamination with similar yields. This can be exemplified from the intrinsic fluorescence data of T137E-FGF1 performed using the protein obtained through expression in the BL-21(DE3) cell line. Intrinsic fluorescence was also performed using T137E-FGF1 from the BL-21 Star cell line. The data was significantly different compared to the previous cell line because it showed that T137E-FGF1 was completely unfolded due to tryptophan being entirely exposed at 350 nm. Additionally, CD data of T137E-FGF1 using the BL-21 Star cell line showed that the mutant had an alpha helical secondary structure whereas WT FGF1 is a beta barrel. From these preliminary characterization experiments, it was concluded that overexpression of the protein in BL-21(DE3) cells produces lower yields of the mutant protein but interestingly results in the presence of lesser protein contaminants. BL-21 Star, on the other hand, results in slightly higher protein yield. However, the presence of higher amounts of contaminating protein products hampered the unambiguous characterization of the mutant protein.

The core of this project is based on the mutation from threonine to glutamic acid in order to introduce a negative charge in the heparin binding region of FGF1 and analyze the

effects it can have on heparin binding to FGF1. It was very evident that introducing a glutamic acid in place of threonine caused a drastic change in the folding of the heparin binding region. The hydrogen bonds that threonine shared with several residues in the hydrophobic core of the heparin binding region of FGF1 were severely disrupted when it was replaced with glutamic acid. This disruption in the hydrogen-bonding network appears to cause the heparin binding loop to significantly unfold and be exposed to the solvent. This conclusion is also consistent with increased susceptibility of the mutant protein to proteolytic degradation resulting in the appearance of several degradation bands on the SDS PAGE.

Although pure T137E-FGF1 was not obtained, the overall findings of this project highlight the critical role of T137 in the maintenance of the structural integrity of FGF1.

## References:

1. Schlessinger, J. (2000). Cell signaling by receptor tyrosine kinases. *Cell*, 103(2), 211-225.
2. Ornitz, D. M., & Itoh, N. (2001). Fibroblast growth factors. *Genome Biol*, 2(3), 1-12.
3. Powers, C. J., McLeskey, S. W., & Wellstein, A. (2000). Fibroblast growth factors, their receptors and signaling. *Endocrine-related cancer*, 7(3), 165-197.
4. Hung, K. W., Kumar, T. K. S., Kathir, K. M., Xu, P., Ni, F., Ji, H. H., ... & Yu, C. (2005). Solution structure of the ligand binding domain of the fibroblast growth factor receptor: role of heparin in the activation of the receptor. *Biochemistry*, 44(48), 15787-15798.
5. Hung, K. W., Kumar, T. K. S., Chi, Y. H., Chiu, M., & 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(1), 253-258.
6. Werner, S., & Grose, R. (2003). Regulation of wound healing by growth factors and cytokines. *Physiological reviews*, 83(3), 835-870.
7. Beenken, A., & Mohammadi, M. (2009). The FGF family: biology, pathophysiology and therapy. *Nature reviews Drug discovery*, 8(3), 235-253.
8. Itoh, N. (2007). The Fgf families in humans, mice, and zebrafish: their evolutionary processes and roles in development, metabolism, and disease. *Biological and Pharmaceutical Bulletin*, 30(10), 1819-1825.
9. Carter, E. P., Fearon, A. E., & Grose, R. P. (2015). Careless talk costs lives: fibroblast growth factor receptor signalling and the consequences of pathway malfunction. *Trends in cell biology*, 25(4), 221-233.
10. Belov, A. A., & Mohammadi, M. (2013). Molecular mechanisms of fibroblast growth factor signaling in physiology and pathology. *Cold Spring Harbor perspectives in biology*, 5(6), a015958.
11. Zhang, X., Ibrahimi, O. A., Olsen, S. K., Umemori, H., Mohammadi, M., & Ornitz, D. M. (2006). Receptor specificity of the fibroblast growth factor family The complete mammalian fgf family. *Journal of Biological Chemistry*, 281(23), 15694-15700.
12. Li, X., Wang, C., Xiao, J., McKeenan, W. L., & Wang, F. (2016, January). Fibroblast growth factors, old kids on the new block. In *Seminars in Cell & Developmental Biology*. Academic Press.
13. Canales, A., Lozano, R., López-Méndez, B., Angulo, J., Ojeda, R., Nieto, P. M., ... & Jiménez-Barbero, J. (2006). Solution NMR structure of a human FGF-1 monomer, activated by a hexasaccharide heparin-analogue. *Febs Journal*, 273(20), 4716-4727.
14. Casu, B., Naggi, A., & Torri, G. (2015). Re-visiting the structure of heparin. *Carbohydrate research*, 403, 60-68.
15. Xu, R., Rudd, T. R., Hughes, A. J., Siligardi, G., Fernig, D. G., & Yates, E. A. (2013). Analysis of the fibroblast growth factor receptor (FGFR) signalling network with heparin as coreceptor: evidence for the expansion of the core FGFR signalling network. *Febs Journal*, 280(10), 2260-2270.

16. Sasisekharan, R., & Venkataraman, G. (2000). Heparin and heparan sulfate: biosynthesis, structure and function. *Current opinion in chemical biology*, 4(6), 626-631.
17. Perrimon, N., & Bernfield, M. (2000). Specificities of heparan sulphate proteoglycans in developmental processes. *Nature*, 404(6779), 725-728.
18. Angulo, J., Ojeda, R., de Paz, J. L., Lucas, R., Nieto, P. M., Lozano, R. M., ... & Martín-Lomas, M. (2004). The Activation of Fibroblast Growth Factors (FGFs) by Glycosaminoglycans: Influence of the Sulfation Pattern on the Biological Activity of FGF-1. *ChemBioChem*, 5(1), 55-61.
19. Pineda-Lucena, A., Jiménez, M. Á., Lozano, R. M., Nieto, J. L., Santoro, J., Rico, M., & Giménez-Gallego, G. (1996). Three-dimensional structure of acidic fibroblast growth factor in solution: effects of binding to a heparin functional analog. *Journal of molecular biology*, 264(1), 162-178.
20. Zakrzewska, M., Krowarsch, D., Wiedlocha, A., & Otlewski, J. (2004). Design of fully active FGF-1 variants with increased stability. *Protein Engineering Design and Selection*, 17(8), 603-611.
21. Sarrazin, S., Lamanna, W. C., & Esko, J. D. (2011). Heparan sulfate proteoglycans. *Cold Spring Harbor perspectives in biology*, 3(7), a004952.
22. Farrugia, B. L., Lord, M. S., Melrose, J., & Whitelock, J. M. (2015). Can we produce heparin/heparan sulfate biomimetics using "mother-nature" as the gold standard?. *Molecules*, 20(3), 4254-4276.
23. Pellegrini, L., Burke, D. F., von Delft, F., Mulloy, B., & Blundell, T. L. (2000). Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. *Nature*, 407(6807), 1029-1034.
24. Eswarakumar, V. P., Lax, I., & Schlessinger, J. (2005). Cellular signaling by fibroblast growth factor receptors. *Cytokine & growth factor reviews*, 16(2), 139-149.
25. Plotnikov, A. N., Schlessinger, J., Hubbard, S. R., & Mohammadi, M. (1999). Structural basis for FGF receptor dimerization and activation. *Cell*, 98(5), 641-650.
26. Mohammadi, M., Honegger, A. M., Rotin, D., Fischer, R., Bellot, F., Li, W., ... & Schlessinger, J. (1991). A tyrosine-phosphorylated carboxy-terminal peptide of the fibroblast growth factor receptor (Flg) is a binding site for the SH2 domain of phospholipase C-gamma 1. *Molecular and cellular biology*, 11(10), 5068-5078.
27. Wong, P., Hampton, B., Szylobryt, E., Gallagher, A. M., Jaye, M., & Burgess, W. H. (1995). Analysis of putative heparin-binding domains of fibroblast growth factor-1 using site-directed mutagenesis and peptide analogues. *Journal of Biological Chemistry*, 270(43), 25805-25811.
28. James, G. T. (1978). Inactivation of the protease inhibitor phenylmethylsulfonyl fluoride in buffers. *Analytical biochemistry*, 86(2), 574-579.
29. Sekar, V., & Hageman, J. H. (1979). Specificity of the serine protease inhibitor, phenylmethylsulfonyl fluoride. *Biochemical and biophysical research communications*, 89(2), 474-478.
30. National Center for Biotechnology Information. PubChem Compound Database; CID=6049, <https://pubchem.ncbi.nlm.nih.gov/compound/6049> (accessed Mar. 17, 2016).

