Osteogenesis imperfecta (OI), or brittle bone disease, is an inherited genetic disorder that can occur in different stages of a person’s lifetime. It is estimated that 1 in 20,000 individuals are affected globally, making this condition rather rare \(^3\). It occurs equally in both females and males, making no distinctions among ethnic groups. OI is considered primarily an autosomal dominant disease, meaning that a parent has at least one copy of a mutated gene that gets passed down a generation \(^1\). Since it is a congenital disease, there is a 50% chance that it will be passed on to the next generation. For this particular disease, being homozygous can exist in a different gene which can be the cause of one type of OI. Of the eight recognized forms of OI, some of the most common are type I and IV \(^1\).

Many effects of OI include skeletal features such as short stature, loose joints, scoliosis, fragile bones and bone deformity. Bones also break easily with little to no trauma\(^1\). Non-skeletal symptoms affect the heart, eyes, blood, teeth, heart defects, hearing, breathing and tendons. Examinations like a physical diagnosis, skin punch biopsy, and chorionic villus sampling are commonly used to test for OI, if family history exists. As of now, there is no cure for OI; however, surgery, exercise and therapy are available to alleviate the symptoms \(^1\). Within this disease, there are mutations that are involved within the collagen structure. Some instances of the mutation do not manifest in the patient until later in life, while others who are carriers of the mutation show
symptoms prior to birth. Because the disease is inherited, it is essential to know that there are several levels of severity.

As mentioned, this is a congenital disease; an example of this mutation can be attributed to the scientific researcher's contribution at Université de Nantes, France. The researchers discussed a case in which mutations were found while testing the DNA of the fetus \(^1\). Through a model, they performed a prediction of clinical lethality with glycine mutations of the alpha chain and applied the strategy by illustrating the lethality of COL1A1 glycine mutations. The researchers used the model to illustrate that mutations cause lethality \(^1\). The illustration showed that Glycine mutations occurred in the \(\alpha\) 1 chain of type 1 collagen and lethality of Gly \(\rightarrow\) Ser and Gly-Cys mutations occurred in the 2-\(\alpha\) chains \(^1\). Bulky amino acids were replaced with glycine per their position of the N terminal of the triple helix. Ala was predicted to be non-lethal because it had a melting temperature, defined as \(T_{m}^{[+1]}\). The \([+1]\) represents the lower stability of the peptide GPOSPAGFA, than the average prediction. Alanine is non-lethal and higher compared to a lethal mutation that has a low \(T_{m}^{[+1]}\) value in comparison to the Gly \(\rightarrow\) Ser model where c terminal triplets were compared. According to the research done, the following amino acids were found to be the most unstable due to their position in the c terminal of the helix; Arg, Val, Asp, and Glu. In contrast, Gly \(\rightarrow\) Ser showed that environmental factors such as neighboring Ser substitution affected the mutation \(^1\). Residues found through the study of this model were necessary to conclude severity levels, predicting the clinical lethality from glycine mutations within the \(\alpha\) 1 chain of type I collagen.

In this study, the researchers collected collagen Gly mutations published by the OI consortium. This data was essential to the model the researchers developed through
this study. After identical mutations were found in various patients, a data set was made based on the position of the mutation and the substitution of the amino acid replacing Gly. To each mutation, a lethality was assigned and observed in the patients with OI associated with the mutation. The researchers used Modeling and Statistical Analysis to perform calculations with a program called Weka. Calculations facilitated a decision tree analysis and logistic regression to compute the differences under the null hypothesis with the use of one or more random variables consisting of two lethal (control) and one non-lethal (non-experimental).

The next step of the process included peptide synthesis. Purification was used to find peptide concentrations at an absorption amount of 275nm of the amino acid tyrosine. By using a mass spectrometer with a matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) peptides were analyzed for purity. Peptides were synthesized at Tufts University Core Facility (Boston, MA). Using reverse-phase high-pressure liquid chromatography system (RPC) (Shimadzu), a protein was purified and separated. The next method used a hydrophobic stationary phase.

Consequently, molecules that were hydrophobic in the polar mobile phase were absorbed to the hydrophobic stationary phase, and the hydrophilic molecules in the mobile phase went through the column and were the first to be eluted. Decreasing the polarity of the mobile phase using a (non-polar) solvent reduced hydrophobic interactions. Furthermore, the Matrix-Assisted Laser Desorption Ionization (MALDI-TOF) was used for the analysis of biomolecules and large molecules to confirm the results and identity of the residues.
Another approach attempted was Circular Dichroism Spectroscopy (CD); it measured the difference in absorption between left and right handed polarized light. The outcome of this approach occurs when a molecule contains one or more chromophores. The wavelength of the peptides helped determine the structure of the helix.

Furthermore, determination of $T_m$ was performed through CD by refolding the peptides. It was found that the spectra of the two peptides GPOSPAGFA and GFASPAGPO exhibited a triple helix characteristic. However, peptide GFASPAGPO requires more energy to unfold because it absorbs more heat, therefore becoming more stable because of higher temperatures. Also, CD folding experiments were done through the heating of peptides and quick water bath. Ellipticity was then observed to see if coiling was done correctly. Lastly, through Differential Scanning Calorimetry (DSC), it was determined how much energy was released and the helix’s folding stability. More fold with black $Tm$ 0 was higher, GPO more stable. The more alpha helices the more folding the higher melting temperature the more stable the peptide, and vice versa.

The authors concluded that there was no link between lethality and the set of mutations based on this model. Insufficient information has been accumulated thus far and has been used to determine all of the possible outcomes. Factors such as integrins, which are transmembrane receptors that allow for cell-to-cell communication, and proteoglycans, proteins in the connection tissue, found in the matrix may be the missing link that regulate the phenotype. To fully understand this mutation, more research will need to be done in encoding genes and how it is linked to lethality of OI mutations. Studies demonstrated that substitutions of Gly could have structural effects regardless of what amino acid replaces glycine.
Considering the information above, researchers investigated Gly and its effects within a bacterial collagen protein found in E.coli in a different study done in VU University Medical Centre in Amsterdam, the Netherlands. The bacterium was then analyzed through SDS-PAGE. The reason this organism was chosen was because the sequence, Gly-Xaa-Yaa, confirmed the triple-helical structure. It also didn't contain hydroxyproline, known to stabilize the triple-helix structure. In addition, through Trypsin digestion, VCL (vinculin) proteins in PBS were digested, and the proteins were inhibited with Trypsin, which stopped the digestion. The importance of PMSF (phenylmethane sulfonyl fluoride) was to prepare cell lysates.

Other experiments done were circular dichroism spectroscopy (CD) and thermal transitions. They determined fraction folding of the triple helix through observation of temperature fluctuation. Using differential spectrometry, data was collected, and dynamic light scattering measurements were performed to determine the size of the bacterium to conclude the size of the collagen molecule.

The effects of the Gly mutation positions are replaced within a bacterial collagen protein and are explored in the triple helix near the N-terminal trimerization domain in which symmetry begins. The Gly mutations located near the trimerization could explain the severity of OI, for example, nucleation, a phase in which the physical change of the helix begins to occur (formation from a liquid to a crystal). The affects of these mutations near the folding domain of the triple helix may relate to the mutations near the C-terminus belonging to type 1 and type III collagens, and may lead to further experimentation with the same mutations.

Similar to the study on collagen, a group of researchers from Ethics Review
Committee on Human Research of the University of Tartu focused on a phenotype of OI II and the new mutation responsible for the causation of the phenotype in the patient \(^3\). OI Type II results from a new dominant mutation in type I collagen gene or parental mosaicism. Primarily, the researchers presented the correlation between the two genes, \(COL1A1\) and \(COL1A2\), using sequencing analysis from two related Chinese patients’ DNA. The test subjects were given identification numbers for the duration of the experiment. The group collected blood samples from the proband (716), her brother (715), and parents (710, 711) \(^3\). The parents were found absent of the mutation and proved that OI has a dominant phenotype. The PCR method was used to detect exons by mapping locations and exon-intron sequences found in \(COL1A1\) and \(COL1A2\).

Comparisons of both sequences were made, by analyzing both OI and Ehlers-Danlos syndrome patients to discover a novel mutation.

Within the same study, another experiment conducted was the whole genome sequencing which determined the DNA of an organism, one genome at a time, although no direct results were found. To complete this, the experiment was performed with both the child and the parents affected using the TruSeq Exome Enrichment kit (Illumina Dye Sequencing) a technique used to determine the series of base pairs in DNA \(^3\). It is based on a reversible dye-terminators to enable the identification of single bases as they are introduced into the DNA strands.

After performing the whole-exome sequencing in case parent–child trio, it was identified that a heterozygous missense mutation within the \(COL1A1\) gene, Gly to Cys transversion in the triple helical domain of the collagen type I \(\alpha\) chain \(^3\). It has been noted that Gly missense mutations disrupting collagen mutations type I collagen could lead to
phenotypes with lethal consequences. Even though a new mutation was found, in the 
*COL1A1* gene, to be severe and life threatening, it’s still unidentifiable by a genetic 
counselor because there is insufficient research to support specific unidentified genes.

*COL1A1* is lethal and is a silencing gene that reduces expression of mutant 
collagen found in fibroblasts in a mouse model \(^4\). For the first screening, a human 
embryonic kidney cell line was used through two types of techniques in vitro and ex vivo. 
The researchers located the SiRNAs, short interfering RNA, a common tool for inducing 
short term silencing of the gene and degradation, as well as also noted their effects on 
collagen transcripts and protein. The data supported shRNAs, a small short hairpin, an 
artificial RNA molecule, which silenced target gene expression, and suppressed the Mut 
allele. SiRNAs were chosen based on their specificity and corresponding shRNAs for the 
effects of collagen and protein. Some materials that were used in the experiment include, 
cell cultures, short hairpin RNAs and RNAi molecules; the exact use was not noted \(^4\). The 
experiments performed were: generation of Col1a1 WT or Mut-LucF-HEK cells for 
siRNA screening, siRNA transient transfection, generation of fibroblasts stabling 
expressing shRNAs, proliferation assay RNA isolation and real time allele specific PCR, 
collagen analysis, Western Blotting and statistical analysis.

Methods performed included: cell cultures designed from primary dermal 
fibroblasts that were isolated from a newborn Brlt mouse skin \(^4\). Small interfering RNA 
were designed as target Mut Col1a1 of the mouse \(^4\). Three siRNA molecules were used to 
observe allele-specific gene silencing using Mut *Coll1a1* from the Brlt mouse, a common 
model for human OI. first, a negative control, Lac Z-siRNA (targeted bacterialsidase), 
second, a positive control, LucF-siRNA(designed against the firefly luciferase), and third,
a second positive control e-Wt sRNA(Wt murine Col1a1). To design two oligonucleotides, short hairpin RNAs, such as LacZ- and Mut-shRNAs were used.

In addition to the materials used, an experiment was performed: Generation of Col1a1 WT- or Mut-LucF-HEK cells for siRNA screening were cloned in the plasmid. The siRNA transient transfection was another experiment performed, Hek-LucF-co1a1 cells were seeded along with primary Brtl fibroblasts and transfected. The application of the Luciferase activity was used to measure gene expression at a transcriptional level after the transfection was performed yielding high protein.

Another experiment showed that oligonucleotides LacZ- and Mut-shRNAs were annealed and cloned into the pSUPER vector which produced lentiviral particles and were further analyzed for expression to demonstrate how the mutated genes were transferred. The lentiviral delivered viral information from the DNA to the RNA. Cell proliferation was completed to measure the amount of cells from untransduced and transduced Brtl primary fibroblasts. Total RNA was extracted from modified and unmodified Brtl fibroblasts. The expression of the Col1a1 gene was observed by using real-time PCR to monitor the RNA target Col1a1 allele to monitor gene growth.

The synthesis of collagen for the next experiment was completed using F-Mut shRNA transduced and parental Brtl fibroblasts. Extraction and purification of the collagen was obtained from the media and separated using SDS-PAGE, finally analyzed for DyLight fluorescence. DyLight illuminated the secondary antibodies. Western blotting was then used to identify a specific amino acid to determine if the right antibody was made, this method also helped to quantify the amount of proteins. In summary, E-Mut and f-Mut siRNAs did not result in effective suppression towards luciferase activity,
in contrast, the labeled small interfering RNA, a-, b-, c- and d-Mut siRNAs, were more effective. When Brtl fibroblasts were injected with f-Mut siRNA or e WT siRNA, overall expression of RNA was reduced, moreover the F-Mut shRNA improved COL1 protein synthesis in Brtl fibroblasts. Lentiviral particles produced efficient cells, but no changes in morphology or density was shown. Western blot results to indicate the reduction in F-Mut-transduced Brtl cells in comparison of untransduced or transduced Brtl fibroblasts.

The potential of gene silencing becoming a powerful tool in monogenic diseases has great potential. The findings of this article suggested that testing of the delivery of shRNAs by the lentivirus successfully suppressed the Mut protein with about 40% decrease. An indication of this result was the improvement of normal mRNA using hammerhead ribozymes in both cell-free assays and patient fibroblasts. Within the RNA molecule, hammerhead ribozymes improved specificity without affecting efficiency. They are enzymes that serve as therapeutic agents to catalyze reversible cleavage and joining reactions at specific sites. Hammerhead mismatches were used within the binding arm. Both siRNAs, specifically, e and f were adequately successful in down regulating the Mut allele yielding a two-fold increase in the WT versus Mut allele ratio.

Moreover, the Genetics and Molecular Biology journal identified a frameshift mutation within the COL1A1 gene. The authors focused on how a missense mutation in collagen can lead to structural consequences at a strongly charged site within the supercoiled triple helix. Electrons become attracted to areas of higher density and as a result the area becomes altered chemically, leaving a residue of collagen at the location p263. Several experiments were done: such as nuclear magnetic resonance, circular dichroism, and differential scanning. They indicated that substitution of Gly to Ala
caused a loss of stability in the triple helix as well as a non-equivalence of Ala residue in the three chains. A proposal in this article has suggested that structural defects leading OI to occur due to genetic and environmental factors influencing the severity. There were several techniques involved to prove the Gly to Ala loss of stability.

Liquid chromatography, a technique, commonly was used to separate molecules that are dissolved in a solvent. In order to perform this technique, peptides were purified in a high pressure C-18 column. To confirm the results and identity of the residues, a soft ionization was used, known as Matrix-Assisted Laser Desorption Ionization (MALDI), and was used for the analysis of biomolecules and large molecules.

Circular Dichroism Spectroscopy, was another approach used in the lab. Samples were prepared in a temperature-controlled environment in a phosphate-buffered solution. To determine the concentrations of the peptides, the absorbance at around 275 nM was used to quantitate the amount of protein present.

To study thermal transitions of a polymer, which define the changes in temperature, Differential Scanning Calorimetry, a thermoanalytical approach involving the application of heat had been used. Measurements of DSC were taken using a scanning calorimeter where peptides were purified in buffers and concentrations were then recorded.

Additionally, for compounds of particular interest, such as gly-pro and pro-hyp, NMR Spectroscopy was used to determine the physical and chemical properties of the molecules within those compounds. Using N-15 labeled proteins, peptides were synthesized for NMR characterization at selective positions. Selective peptides, for example T1-655 were labeled at position G7 and G16 and peptide T1-645 were selected
for G16 only. Once labeled, the peptides were prepared in PBS (phosphate buffered saline) and the pH of the solution dropped by adding HCl. 3D HNHA allowed the measurement of homonuclear coupling constants of N15 labeled proteins and induced hyperacetylation, which occurs when breaking lysine residues. The measurements of amide protons and temperature were recorded using linear aggression analysis.

Furthermore, a peptide was designed to model a missense mutation of Gly to Ala and led to a severe case of OI where the mutation site on a KGE sequence had contributed to stabilization of the triple helix. Using the Differential Scanning Calorimetry method, Gly to Ala replacement was observably more stable at the higher temperature. The application of this method determined the thermal stability of the helix by observing temperatures compared to the parent peptide. When a pH of approximately 3 has been reached, T1-655 shows a hydrogen bond formation of Gly7 and Gly16. It lowers the pH and weakens the hydrogen bonds that have been mutated within Ala, but do not affect the Gly hydrogen bonds in the N terminal because of inaccessibility. Using NMR, investigations were performed using T1-645 and its mutant and parent peptides to confirm the features of the helix that are typically found. GL7 was shown to have a monomer suggesting a partially folded conformation around the mutation and an unfolded N terminal. The mutant and parent peptide were modified to promote folding and stability at the Gly7 and Gly16 locations with N labeled residues.

Undeniably, the level of pH at the mutation is a predisposing factor that leads to disruption or pathology of OI. The KGD and KGE sequences also greatly affect the stability when the T1-645 and T1-655 peptides are isolated within the peptides’ environment. Although stability is a major contributing factor in the helix, the stability
does not predict substitutions in the amino acids, gly to ala. In fact, the authors have deduced that there hasn't been any concluding evidence found to support the severity of OI. Nevertheless, the environment of the charged site itself had no impact on the stability of the helix. The conclusion for this experiment showed that sequence analysis represented no mutations in COL1A2 gene from the proband or her mother. On the other hand, a heterozygous frameshift mutation was identified from both mother and proband but not from other family members. If translation would occur for the mother and proband, a short form of collagen α1 chain without the helical found in the wild type. This in turn would form haploinsufficiency of the collagen α1 chain.

Given these points, it is clear that more information is needed to find the cause of OI’s severity. Understanding how molecules interact and behave in different states of environment is vital to understand a disease. As mentioned before, findings by various institutes have researched clinical lethality, missense mutation, and allele deletions providing insight to this rare disease. Specifically, one article reported that gly → Ser mutations with low melting points resulted in lower stability. Another article suggested that the environment of the mutation, highly charged sites, could be a predisposing factor to the destabilization of the local sequence. Another article was able to accurately determine the suppress Mut protein using quantitation of the Mut dimeric α1-α1 collagen; because of this, they were able to observe a reduction in Mut collagen. The Mut collagen could help lower the stress caused by aberrant collagen retention. Furthermore, researchers discovered that variants in the proteins interacting with collagen may compensate for a collagen α1 or α2 variant and contribute to OI symptoms as proposed by the authors of the article. The researchers’ proposal that the novel COL1A1 variant
Torres

could lead to haploinsufficiency of collagen α1 chain and the frameshift variant could be a possible molecular basis of OI type I in this Han Chinese family \(^4\).

Because OI is a hereditary disease and has no cure, alternative forms of treatment are available such as physiotherapy, exercise, therapy, and medical treatments aimed at increasing bone mass and strength \(^8\). Surprisingly, hopes for bone marrow transplantation (BMT) are a potential future for OI. Adult bone marrow transplants in utero has shown to decrease lethality in OI \(^4\).
Literature Cited


