

**Disclaimer**—This paper partially fulfills a writing requirement for first year (freshman) engineering students at the University of Pittsburgh Swanson School of Engineering. *This paper is a student, not a professional, paper.* This paper is based on publicly available information and may not provide complete analyses of all relevant data. If this paper is used for any purpose other than these authors' partial fulfillment of a writing requirement for first year (freshman) engineering students at the University of Pittsburgh Swanson School of Engineering, the user does so at his or her own risk.

## RATIONAL DESIGN PROTEIN ENGINEERING THROUGH CROWDSOURCING

Tylar Farmer, tsf13@pitt.edu, 2:00 Bursic, Patrick Bohse, pab120@pitt.edu, 10:00 Mahboobin

**Abstract**—Two popular methods exist to engineer a protein: directed evolution and rational design. Directed evolution utilizes a controlled environment to create proteins through induced mutations and selection, while rational design makes desired changes to a protein by directly manipulating its amino acids. Directed evolution is currently more commonly used, since rational design relies on structural knowledge of the protein of interest, which is often unavailable. Crowdsourcing manpower and computational power to improve protein depictions allows rational design to be more easily used to perform the manipulation of proteins. Two free programs, "Folding@home" and "Foldit", allow anyone with a computer and internet access to contribute to protein engineering. Folding@home relies on one's computational power, while Foldit relies on user intuition to improve protein models. Rational design has allowed protein engineers to create artificial proteins that can be applied to the treatment of illnesses, research of enzyme activity in a living system, genetic engineering, and biological warfare. Starting with an overview of protein engineering, this paper discusses the methods of rational design and directed evolutions and goes on to explain how computer based programs can help in the advancement of rational design as a protein engineering method.

Furthermore, this paper discusses the application of computer based programs in medicine and genetic engineering, its sustainability, and presents some ethical issues that may arise from using such technology. The paper concludes with an analysis of whether or not computer based programs for protein engineering is worth the investment.

**Key Words**—Crowdsourcing, Directed Evolution, Genetic Engineering, Foldit, Folding@Home, Protein Engineering, Rational Design

### AN OVERVIEW OF PROTEIN ENGINEERING

Protein engineering is a novel field focused on the synthesis and arrangement of amino acid chains to develop proteins that are used to benefit a living system, such as a human. Amino acids are the building block molecules of

proteins and amino acid chains are folded to create the structure of a protein molecule.

Before discussing the function of proteins, one must have an understanding of protein structure, as with most macromolecules, structure determines function. There are four stages of protein folding that are of importance in protein engineering: the primary structure, secondary structure, tertiary structure, and quaternary structure. The primary structure refers to just the sequence of the amino acids before any folding takes place and is held together via covalent bonds. There are 20 essential amino acids from which the protein sequence is created. The secondary structure refers to the local structure of the protein backbone formed by hydrogen bonds, which can be described in two ways: alpha helix, which is a spiral of amino acids, and beta pleated sheets, which are two or more parallel or anti-parallel amino acid chains. The tertiary structure refers to the folds in the amino acid chain that gives the protein its three-dimensional structure, and consists of hydrogen bonds, Van Der Waals forces, ionic interactions, salt bridges, and hydrophobic interactions. The tertiary structure is formed once the protein has reached a position in which it is most stable, or at its lowest energy state. The quaternary structure forms when the protein combines with another protein in its tertiary phase. The most important structure in reference to protein engineering is the tertiary structure, as knowledge of this structure is critical to understanding how proteins fold and how and why they interact during their quaternary phase. The protein goes through these four stages of folding in a matter of milliseconds upon being created [1][2]. These stages can be seen in Figure 1 below.

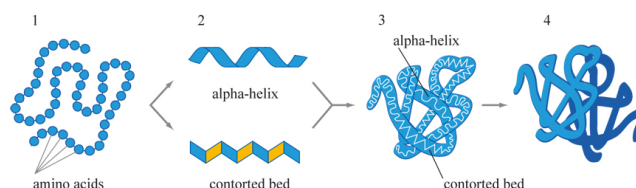


FIGURE 1 [3]  
Four Stages of Protein Folding

Step 1 shows the sequence of amino acids. Step 2 shows the amino acids folding into an alpha-helix or beta sheet (also

called a contorted bed.) Step 3 shows the rest of the amino acids folding with alpha-helices and beta sheets scattered throughout the protein. Step 4 shows two tertiary proteins coming together to form the quaternary structure.

With a fundamental understanding of protein structure, protein function can now be discussed. Proteins catalyze most of the reactions that take place within all living systems and are responsible for actions such as digestion, thinking, moving, and fighting diseases. The young nature of protein engineering means it has not yet proven what exactly it is capable of accomplishing, but already, engineered proteins, such as the Zinc Finger Nuclease, have been created in order to combat HIV [4].

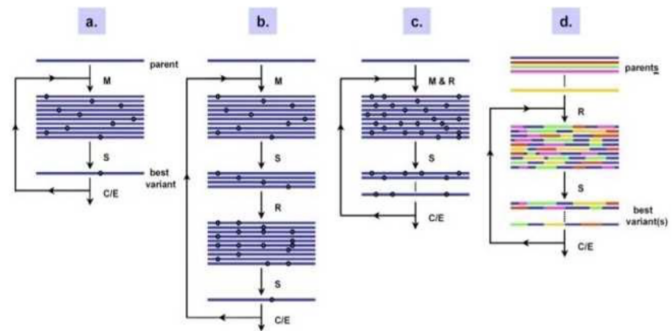
### Paper Overview

This paper seeks to analyze the importance of rational design, discuss how rational design can be made feasible via crowdsourcing protein models, show applications in the fields of medicine and genetic engineering, confront ethical issues regarding protein engineering in general, and determine if rational design is worth investing in. Rational design will be introduced and contrasted with another protein engineering method called directed evolution. Rational design's contribution towards the sustainability of human beings will also be discussed. Ultimately, rational design will be declared a better process for our applications than directed evolution. Our paper will then seek to amend rational design's main fault: its reliance on accurate protein models. Our paper will address this issue by introducing two programs called Folding@Home and Foldit that will provide more accurate protein models for rational design to use. Our paper will compare these two programs, explaining how Folding@Home uses a computer's idle processing power while Foldit uses human intuition, to show that these programs allow rational design to be a feasible process. In order to supplement our discussion of whether rational design is worth an investment, our paper will analyze various applications of protein engineering. We'll first discuss the medical applications, citing an engineered protein that improved metabolic rates in a living system. We'll then talk about genetic engineering through the use of Zinc Finger Nucleases (ZFNs), which are a type of engineered protein that binds to DNA. Next, our paper will address the various ethical concerns that come with engineering something that is normally created via natural biological processes. The main points of contention are modifications of a human's genome, modifications to agriculture, and applications to biological warfare. Sustainability in reference to agriculture will be touched upon, as well as sustainability in reference to pharmaceutical drugs. To end our paper, we'll summarize the rational design process, mentioning how one can crowdsource improved protein models to make the process feasible. Our paper will also reiterate the sustainability contributions of both rational design and protein engineering. We'll synopsise our applications and ethical concerns to

reach a conclusion about whether this technology is worth investing in.

### Analysis of Directed Evolution

Two processes are currently being utilized for the synthesis of proteins: directed evolution and rational design. Directed evolution, the more commonly used of the two, involves embedding a gene that codes for a protein that performs a task similar to what the target protein should do into a host's genome. Afterwards, a mutation would be induced on that gene and DNA recombination would occur, further increasing genetic variation. The host would then be allowed to multiply. Elements of the host's environment would then be manipulated in order to isolate a variant of the host that contains a gene that codes for a protein with the desired traits. For example, if one wanted to create a protein that digests fructose instead of sucrose, he or she would insert the gene known as SI, which codes for the sucrase-isomaltase enzyme that digests sucrose, into the well-documented host, *E. coli*, and induce mutations on the gene via Sodium Azide, a chemical mutagen. One would then allow *E. coli* to go through many rounds of replication in order to get a large population, and afterwards, one would introduce fructose into the environment and observe which variant of the host shows the greatest fit. Once the experimenter isolates the fit host, he or she would extract the host cell's genome and use restriction enzymes to splice out the gene originally inserted into the host, giving him or her a gene that codes for an enzyme that digests fructose [5]. The process of directed evolution can be seen in figure 2.



**FIGURE 2 [6]**

### Four Possible Directed Evolution Processes

Beginning with one or more parent genes, a mutagen (M) is applied and the host is allowed to replicate, giving variants of the same gene. A screening (S) is applied to isolate the genes that potentially code for the desired protein. Recombination (R) may follow, increasing genetic variation, followed by another screening. Depending on whether the experimenters have determined that they have isolated the desired variant of the parent gene, they may end the process and harvest the gene (E) or run the process again (C). The lines each

represent a different variation of the parent gene, with the black dots showing sites of mutation.

### Analysis of Rational Design

The less utilized method of protein engineering is known as rational design. Rational design involves directly modifying an existing protein (or creating a protein from scratch) in order to create a protein that performs the desired task. To use a similar example as above, if one wanted to create a protein that could digest fructose, one would load up a model of the sucrase-isomaltase enzyme into a simulation software and directly apply changes to the enzyme until the simulation indicates that the enzyme can bind to and break down fructose. Once the experimenter has the desired protein's model, he or she would create an in vitro gene that codes for the protein using a process known as Oligonucleotide synthesis with annealing. [7]. The experimenter would then implant this gene into an E. coli cell, allow the protein to be synthesized by the host, then harvest the protein from the host cell. During this process, it is crucial that the model the experimenter originally started with has a high degree of precision, as small differences between the actual protein and the simulated model can completely change the way a protein folds after inducing mutations [8]. Often, multiple rounds of the rational design process need to occur, but with each round, information about the structure of the protein and the way it functions is gained. This process can be seen in figure 3.

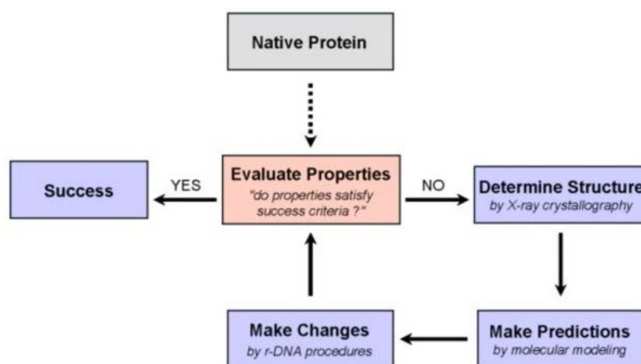


FIGURE 3 [6]

Shown Above is the Rational Design Cycle.

If the designed protein does not satisfy the demand, information regarding the protein's structure and function is gained via X-ray crystallography.

### Contrasting Directed Evolution and Rational Design

The primary benefit directed evolution has over rational design is its lack of need for accurate protein models, as directed evolution simply seeks a protein that does the desired function without ever considering its structure. Protein crystallography is typically utilized to obtain

information about a protein's structure. However, this information is often imprecise due to the vibratory nature of proteins [9]. Although the Research Collaboratory for Structural Bioinformatics' Protein Data Base provides us with about 114,500 models, the estimates taken for the distances between arms of the protein and the coordinates of atoms within the peptides make rational design an unfeasible process; if a protein is inaccurately folded, a proteopathic disease, meaning a disease caused by misfolded proteins, can occur such as Alzheimer's, Cataracts, or Sickle Cell Disease, thus the estimates given do not provide the level of precision necessary to synthesize proteins without the risk of a proteopathic disease [10][11].

Scientists and engineers have been working to improve upon these protein models and reduce their range of error so they could be used for rational design, but this task has been shown to be laborious and time-consuming due to the number and complexity of calculations needed to be done [4].

Despite directed evolution's feasibility, it's a process requiring many hours of research into both the way a protein functions and the environment the protein needs in order to mutate during the accelerated evolutionary process.

Most problematic of all, as a result of randomness of mutations and variability of DNA recombination, the chances of a gene being altered enough to the point where the desired traits are obtained are difficult to calculate; there is no guarantee one will create the desired protein using this process. Thus, rational design, being a direct alteration of a protein, is more practical for creating novel proteins since there is no reliance on random chance presented by natural selection. [12]. If protein models were more precise, rational design would allow scientists and engineers to easily create proteins that could be administered quickly to patients with certain protein deficiencies, promoting our use of biopharmaceuticals and leading to more protein-based cures [13].

The use of biopharmaceuticals, and by association, rational design, contributes to the sustainability of human beings. In this context, sustainability refers to the improvement of human fitness in our environment. As mentioned above, patients lacking a certain protein can be supplemented with a rationally designed alternative capable of performing the same function as the absent protein. This is especially beneficial if the absent protein is not easily created in vivo. For example, if a patient suffers from lactose intolerance, and is thus lacking the lactase enzyme that digests lactose, he or she can ingest a supplement containing the a rationally designed replacement protein that mimics lactases structure and be able to now ingest products containing lactose, such as milk and yogurt [14]. By increasing the amount of foodstuffs available to the patient, his or her fitness in our environment has increased. Although this process can be done via directed evolution, it is not practical as the estimated time required to create the needed protein is difficult to predict and unreliable. Thus rational design's biopharmaceutical application

supports the sustainability of human life by creating protein supplements to allow human's to thrive in our environment.

In addition to rational design's effect on sustainability in regards to human beings, protein engineering's biopharmaceutical application contributes to the sustainability of pharmaceutical drugs. In this context, sustainability is referring to the ability of pharmaceutical drugs to continue being produced at the same or higher rate using fewer resources. In the past, medicines were derived created using plant-based ingredients. Currently, however, medications are created in vitro using chemical processes that mimic their natural counterpart [15]. Creating medicine in vitro has been far more beneficial to the creation of pharmaceutical drugs as compared to their plant-based alternative; however, a laboratory responsible for creating these medicines must maintain a large supply of chemicals needed during the creation process. In contrast, a biopharmaceutical drug would only need a culture medium (typically made from agar, which comes from algae), a host cell (such as *E. coli*), and the appropriate gene in order to create the needed protein. Due to the rapid reproduction rate of bacteria, it would not take very long, nor many resources to obtain the appropriate amount of desired protein for medication. Thus, protein engineering would support the sustainability of pharmaceutical drugs, as the gene needed to create the medicinal protein returns a high yield of said protein [16].

## **CROWDSOURCING OPTIONS TO ASSIST IN IMPROVING PROTEIN MODELS**

Although rational design is currently unfeasible, improvements in protein modelling are being made at a quicker rate than ever before. A large cause of this is the availability of two free programs capable of allowing users with no background in science or engineering to contribute to improving protein models to the point where they can be used for rational design: Folding@Home and Foldit

### **Folding@Home**

Stanford's Folding@Home runs while the user is not using his or her computer, and creates a network of processing power harvested from all computer's running the program. This network, equivalent to a 500,000 core processor, is then used to perform complex calculations that simulate a more accurate predication of a protein's folding motif. The simulation not only provides an improved model, but shows, down to increments of millisecond, every formation the protein goes through to get from its primary structure to its tertiary structure. Thus, Folding@Home provides a full account of a protein's structure from synthesis to completion.

In one of Folding@Home's 129 publications, Dr. Lillian Chong of the University of Pittsburgh discussed how Folding@Home assisted in creating a new method to identify

amino acids critical to a protein's structure. This research expedited our understanding of protein folding motifs, allowing researchers, engineers, and scientists to dedicate more time to applying protein engineering to other fields, such as cancer and viral treatment [17].

In terms of cancer treatment, Folding@Home was used to model an altered form of Interleukin-2, a protein that helps immune system T cells attack tumors, which demonstrated fewer side effects on the living system and was 300 times more effective when compared to the unaltered form. This made Interleukin-2 viable for cancer treatment, as well as provided insight to protein binding mechanisms [18].

### **Foldit**

The University of Washington's Foldit relies on man's intuition and ability to solve problems in order to improve protein models, rather than a computer's processing power. Players are presented with a protein in a 3D environment from the RCSB's Protein Data Bank, and are allowed to move and rotate the arms of the protein. The objective of the player is to achieve the highest score by creating the most stable conformation of the protein within given parameters. These parameters typically include Beta Pleated Sheets aligning a certain way and the amount of turns present on an Alpha Helix. The score is determined by an algorithm that calculates a protein's stability; the higher the score, the more stable the protein. Once a protein's high-scores plateau, the model is then updated in the PDB with the more accurate and stable model [19].

Foldit's principal claim to fame is improving the model of the Mason-Pfizer Monkey Virus retroviral protease, which assists in the onset and spread of AIDS. Players managed to improve the protease's model in only ten days, as opposed to the 15 years it took scientists. This information is being used to create drugs to combat AIDS by blocking the protease's ability to bind to the virus. The model of this molecule can be seen in Figure 4. Foldit has also facilitated scientific breakthroughs in remodeling biomolecular Diels-Alderase, an enzyme that catalyzes the reaction between conjugated diene and an alkene used in renewable fuels, to increase its binding ability with substrates and lowering its activation energy, thus creating a more effective enzyme [20][21]. Foldit players have also managed to create a remodel of Kumamolisin, an enzyme used in gluten digestion, to have high activity in more acidic environment. This improved model led to the creation of KumaMax, a supplement that allows those suffering from Celiac disease to digest gluten [22][23].



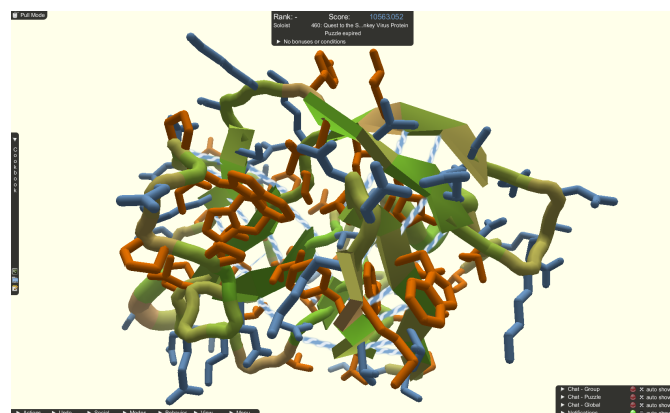


FIGURE 4 [19]

### The Mason-Pfizer Monkey Virus Retroviral Protease as Viewed in Foldit

#### Outlook for Folding@Home and Foldit

These two programs have been successful in providing improved protein models primarily due to how accessible the programs are to the average computer user, as well as the ever-improving processing capabilities of computers. The accessibility of these programs and the processing capabilities of computers are expected to continue growing well into the future, thus the capability of these programs will also continue growing [24][25]. However, the amount of active users for Folding@home (information regarding Foldit's active users over time could not be found) has dwindled since 2012. The amount of processing power continues to increase, but may level out, or even fall, as our rate of increased processing power slows. It is doubtful these programs will get to a point where they are no longer able to supply protein model improvements, thus the outlook for folding@home and Foldit looks promising [26][27].

## APPLICATIONS OF IMPROVED PROTEIN MODELS

Improved protein models have already helped in the discovery of new applications of protein engineering. These following examples demonstrate how important the integration of protein engineering can be to medicine, genetic engineering, and agriculture.

#### Treatment of HIV

Until recently, a cure for HIV was unthinkable. Most researchers thought time spent on prevention and treatment of the virus was more useful than looking for a cure. Currently, a cure for AIDS seems less urgent since it is no longer the epidemic it once was. For many people now, a diagnosis is more of a chronic treatable disease than a death sentence. But, the drugs that treat the symptoms of this disease are not

available to everyone, and once one stops taking them, the symptoms are almost certain to return. Most researchers agree that a daily dosage of drugs for the rest of one's life is not a sustainable solution for the tens of millions of infected people [4].

A functional cure, which does not eliminate the virus from the body, but allows the patient to live a healthy life without taking antiviral drugs, has been successful for several patients. Because of these successes, this treatment is currently being investigated in hopes it will work on a larger scale.

Timothy Brown, a patient with both leukemia and HIV underwent two bone marrow transplants to treat his leukemia. The donor was one of few people who was resistant to the HIV virus because his immune cells lack the CCR5 protein which is located on the surface of immune cells that the virus takes advantage of to infect the cell. Four years later, Mr. Brown shows no symptoms of HIV due to his immune system basically being replaced by one resistant to the HIV virus [4].

Though this may seem like a plausible cure, in reality it is not. Bone marrow transplants are physically taxing, expensive, and risky. In addition, having to find a matching donor that is also lacking the CCR5 protein, yields a cure that is almost nonexistent.

Researchers have turned to genetic engineering, a very close relative to protein engineering to develop a treatment. Genetic engineering is the modification of a sequences of DNA base pairs to produce a different gene. Since the order of DNA base pairs directly effects the protein that the gene produces, modifying the gene will modify the corresponding protein. These researchers have attempted to modify a patient's own immune cells to be resistant to infection by extracting some of the patient's white blood cells and treating them with a gene therapy. This would cause the cells to produce proteins known as Zinc Finger Nucleases that could be used to inhibit production of the CCR5 protein by disrupting the gene that encodes it. These cells were put back into the patient's body, and the patient was told to stop taking his antiviral drugs. Initially, the amount of HIV in his blood spiked, but after 12 weeks, those levels fell drastically, soon reaching an undetectable level, and the immune cell counts increased rapidly [4]. Though this treatment has only been successful with one patient, it shows serious promise for the future of genetic and protein engineering as it applies to medicine and cures for diseases.

As mentioned above, another scientific breakthrough in the area of AIDS is the direct application of the program Foldit. Due to the ability for lay people to donate the processing power of their computers and the manpower needed to run Foldit, a 15 year old AIDS problem was solved by a group of gamers with little to no background in biochemistry. They were able to decode the Mason-Phizer monkey virus (M-PMV) retroviral protease, which is found in the monkey version of AIDS and plays a major role in the multiplication of the virus. The computer model discovered

on Foldit was compared to x-ray crystallography of the actual protein, and it was determined that the structure was correct. This newfound knowledge of protein structure will allow rational design protein engineering to be conducted in hopes of altering the protein so it will no longer aid in multiplying the virus. This may also lead to the development of new drugs that could potentially help block the replication of the virus in humans, though the main goal is to find a cure [20].

### **Aid in Metabolism**

Many of the proteins produced by our body are enzymes, which speed up the chemical reactions that occur throughout our body. These enzymes create more efficient biochemical reactions, which lead to a faster metabolism.

Glutathione transferase, an enzyme best known for detoxifying cellular environments, has recently been the focus of some studies on rational design. Two forms of the glutathione transferase protein, A1-1 and A4-4, catalyze substrates in two different ways. Through rational design, a few mutations were introduced to the A1-1 near the active site, the area on the protein where catalysis occurs, and parts of the A4-4 enzyme were added to the A1-1 form by recombination. The result was a surprising 3000-fold increase in enzymatic activity, thus creating a more effective catalyst [28].

Another form of the glutathione transferase enzyme, A2-2, was rationally designed by inserting parts of a steroid producing enzyme and mutating five amino acids that come in contact with the substrate. The results of this experiment showed an increase in catalytic efficiency of steroid hormone biosynthesis by more than two orders of magnitude [28].

These results of successful rationally designed proteins make this route to protein engineering more promising than directed evolution. The above example shows a direct application to cellular metabolism. By increasing cellular metabolism, our bodies lose less energy during chemical reactions, thus making us more efficient humans. Not only does an increased cellular metabolism apply to our bodies, but it can be applied to the creation of drugs and medicines that rely on cellular processes or proteins extracted from cells.

### **Genetic Engineering via Zinc Finger Nucleases**

Often, genetic engineering is used as a pathway to protein engineering. This is logical because genes store the information needed to create proteins. If one is to alter the genetics of a cell, they will almost surely change the proteins that are being made by the mutated genes.

Gene targeting is a way to repair or inactivate a gene, as seen in the CCR5 gene in the HIV example, by introducing a double-stranded break in the DNA. From there a repair template is placed within the break, and recombination occurs to join the new and old parts of the DNA. ZFNs are a way to target a specific part of a selected gene in basically any cell

type. Because they are extremely customizable, ZFNs are often used to induce the double-stranded breaks needed to repair or inactivate genes [29]. But the question remains on how custom ZFN's can be developed.

The odds of a ZFN being viable increase tremendously when its three-dimensional structure is known. Programs like Foldit and Folding@home can help solve for this needed information. Once the model and make-up of the ZFN is identified, it can then be created through rational design [30]. These powerful molecules are therefore products of and gateways to rational design protein engineering.

### **Sustainability in Agriculture**

In the context of agriculture, sustainability refers to increased output and decreased cost of growing food throughout the world. Because of the ever-increasing global population and the debilitating effects of climate change, if we keep producing food at the rates we are now, we may not have enough food to feed the world [32].

To counteract the threat of scarcity, companies must invest in research on protein and genetic engineering in plants and foodstuffs. Much has already been done in the area of creating chemical pesticides and herbicides to spray on crops, as well as genetically modifying crops to be resistant to these chemicals. Genetically modified crops that are themselves resistant to insects and weeds have also been developed [32]. To truly avoid a worldwide famine, much more must be done.

The development of crops that grow and survive in conditions that are not optimal must be considered. For example, plants that are resistant to frost will help many farmers when there is unexpected cold weather, and will allow more crops to grow in colder conditions. Crops that can better extract water and nutrients from the ground will be extremely beneficial in areas of drought and marginal soil [32].

Though some genetically modified crops have already been developed by altering genes, and therefore the proteins they make, the danger of an impending famine has not been defeated. In order to support our global population, investments in more sustainable farming practices and more efficient plants and crops should be considered.

## **ETHICS IN PROTEIN ENGINEERING**

Whenever a new technology arises, it must be analyzed in terms of its ethicality. There are always many ethical dilemmas and philosophical questions that appear with the creation of a new technology. This is ever so prevalent in the field of biotechnology, since there are immediate health risks.

### **Altering the Human Genome**

Each person carries a copy of their own unique DNA, a sequence of guanine, thymine, adenine, and cytosine nucleic acids, in each of their cells [31]. No two people, besides

identical twins, have the exact same DNA. In a sense, your DNA determines who you are.

The positive application of human genome manipulation was discussed above with the use of ZFNs. It was shown how this treatment could prevent the reproduction of the HIV virus; inhibiting the production of the CCR5 protein by inactivating the gene that produces that protein. When considering the clinical application of genome manipulation ethical discussions such as the following are introduced: Whether human genome should be altered? Under what circumstances should genome manipulation occur? And Can genome manipulation be used for cosmetic purposes?

Some, mostly opponents of this technology, agree that one's DNA makes them who they are, and by altering it they will be changed. Advocates argue that with a genome made up of billions of base pairs, manipulating a few of them doesn't change a thing besides make one's health better and life easier [32].

Because of the philosophical nature of these questions, there is no concrete yes or no answer. There are too many details to generalize the entire technology as ethical or unethical. Situations must be analyzed on a case to case basis.

### **Agriculture**

Although protein engineering, and by association genetic engineering, is usually linked with healthcare and medicine, its effects on agriculture and the economies that are built upon agriculture cannot be ignored.

Genetic modification is sometimes associated with plants and foods, and there has been a longtime debate on whether or not these foods are safe for human consumption. Though a major concern, this is not the biggest issue. The economic and environmental implications are more troublesome.

As mentioned above, supporters of biotechnology argue that in order to avoid worldwide food shortages that are a side effect of extreme population growth, we must invest in genetically modified crops that will produce enough food to meet the growing demand. On the other side, many fear that the biotechnology industry only has their own interest in mind when developing these crops [32].

There is no argument that developing crops that can survive in less-than-ideal conditions is extremely beneficial. But instead of doing research on plants with frost resistance and better water and nutrient uptake from neglected soil, big companies develop transgenic plants resistant to pesticides and herbicides. It is no coincidence that these are the same companies that produce the chemicals they are making these plants resistant to. Ignoring the needs of small scale farms and farmers in developing countries, the biotechnology companies sell their products to the large scale farming operations [32].

These corporate practices may lead to a global economy of big business, with no chance for small scale farms to succeed. Currently, small farms and farmers in developing countries produce 15-20% of the world's food. These

genetically modified plants pose a huge threat to these people who earn their living on farming [32].

Biotechnology companies argue that these modified plants are environmentally friendly, because they result in higher crop yields with less use of chemicals. Adversaries claim that the insects and weeds that the plants are resistant to will themselves quickly become resistant to the pesticides and herbicides that the plants are resistant to as a result of natural selection. Therefore, new genetically modified crops and chemicals to kill the insects and weeds must be developed. This leads to a never ending cycle that only benefits the biotechnology companies and large scale farms and destroys the environment [32].

### **Biological Warfare**

As biotechnology becomes ever more advanced, the threat of it as a war tactic becomes more prevalent. Though the use of engineered proteins as a mechanism of war is extremely unethical, the use of them defensively seems plausible. Through rational design, proteins that deactivate pathogenic biological agents or treat diseases could be produced [33].

On the opposite side, if this technology falls into the wrong hands, it could be disastrous. Highly specific molecules that can do great damage to masses of people could possibly be developed.

The above ethical dilemmas are all complex, thus we cannot simply condemn or approve of the use of protein engineering without acknowledging that if the technology is not used, we will be giving up great opportunities, and if it is used, it comes with significant risks.

## **SHOULD WE INVEST IN RATIONAL DESIGN?**

While directed evolution is currently more feasible when attempting to modify a protein, rational design is quicker and has no reliance on random chance, thus increasing our modified protein yield rate. Rational design's central issue is its dependence on precise protein models in order to synthesize novel proteins. Two free programs are actively working to remedy this: Folding@home and Foldit. Folding@home relies on the user's computer's idle processing power to create a processor network that works through calculations to predict how a protein will fold. Foldit takes advantage of a user's problem-solving abilities and presents the protein as a puzzle, with the highest score going to the most stable protein.

The primary implications of a higher yield rate for engineered proteins are increased use and production of biopharmaceuticals, improved human fitness in our environment, and decreased cost and higher output of crops in agriculture. Within biopharmaceuticals, engineered proteins can be used to aid chemical processes within a living

system and modify one's DNA to prevent the production of diseases, such as the multiplication of the HIV virus. ZFNs are also a product of protein engineering. These highly specific molecules have the potential to alter any sequence of DNA in almost any cell type, and therefore can be a pathway to protein engineering through genetic engineering.

As with any new technology, there are some ethical concerns. Though introducing rational design to medicine could potentially save lives, many are concerned that altering the human genome is immoral, and fear that it will not stop at immediate medical concerns, eventually being used cosmetically. There is also the dilemma between the big biotechnology companies and their beneficiaries, the large scale farming operations, and the small farms and farmers in developing countries. Some feel that genetically modified crops will help save the environment while others feel it will destroy it, while at the same time putting small farms out of business. Additionally, this technology could be used to defend against biological warfare. At the same time, if it falls into the wrong hands, it could be used offensively as a form of war.

After taking into consideration both the benefits and the risks, it seems as though this technology should be invested in. While caution should be taken and the ethicality of the technology continuously re-evaluated, research in this field has the promise to save lives and cure diseases that were once thought to be incurable. If more research is done to create crops that will thrive in malnourished conditions, we could solve the food shortage problem across the world. All of these benefits, and many more appear to outweigh the negatives. With caution in mind, rational design can solve many problems that our world faces today.

## REFERENCES

- [1] "Welcome to The Fold." (2016). Stanford University. (website). <http://folding.stanford.edu/home/>
- [2] "Protein Folding." (2011). *UC Davis*. (online article). [http://chemwiki.ucdavis.edu/Core/Biological\\_Chemistry/Proteins/Protein\\_Structure/Protein\\_Folding](http://chemwiki.ucdavis.edu/Core/Biological_Chemistry/Proteins/Protein_Structure/Protein_Folding)
- [3] "Proteins." (2011). Glossary of Nanotechnology and related Terms. (website). <http://eng.thesaurus.rusnano.com/wiki/article561>
- [4] A. Pollack. (2011). "New Hope of a Cure for H.I.V." *New York Times*. (online article). [http://www.nytimes.com/2011/11/29/health/new-hope-of-a-cure-for-hiv.html?\\_r=0](http://www.nytimes.com/2011/11/29/health/new-hope-of-a-cure-for-hiv.html?_r=0)
- [5] "SI." (2008). *Genetics Home Reference*. (online article). <https://ghr.nlm.nih.gov/gene/SI>
- [6] D.N. Rubingh, R.A. Grayling. (2000). "Protein Engineering." *Encyclopedia of Life Support Systems*. (online article). <http://www.eolss.net/sample-chapters/c17/E6-58-03-06.pdf>
- [7] "Solid-phase Oligonucleotide Synthesis." (2016). Atdbio. (website). <http://www.atdbio.com/content/17/Solid-phase-oligonucleotide-synthesis>
- [8] "Gene Clusters, rational design and metabolic engineering." (2012). *ATG:biosynthetics*. (online article). <https://www.atg-biosynthetics.com/index.php?mact=News,cntnt01,detail,0&cntnt01articleid=55&cntnt01origid=252&cntnt01returnid=25>
- [9] D. Lawson. (2016). "A Brief Introduction to Protein Crystallography by Dave Lawson." *John Innes Centre*. (online article). <https://www.jic.ac.uk/staff/david-lawson/xtallog/summary.htm>
- [10] M. De Los Rios et al. (2006). "On the precision of experimentally determined protein folding rates and phi-values." *Protein Science*. (online article). <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2249776/>
- [11] F. Chiti, C.M. Dobson. (2006). "Protein misfolding, functional amyloid, and human disease." *National Center for Biotechnology Information*. (online article). <http://www.ncbi.nlm.nih.gov/pubmed/16756495>
- [12] U.T. Bornscheuer, M. Pohl. (2001). "Improved biocatalysts by directed evolution and rational protein design." *National Center for Biotechnology Information*. (online article). <http://www.ncbi.nlm.nih.gov/pubmed/11282339>
- [13] R. Rader. (2008). "(Re)defining biopharmaceutical." *Nature Biotechnology*. (online article). [http://www.nature.com/nbt/journal/v26/n7/box/nbt0708-743\\_BX3.html](http://www.nature.com/nbt/journal/v26/n7/box/nbt0708-743_BX3.html)
- [14] "Lactase." (2015). WebMD. <http://www.webmd.com/vitamins-supplements/ingredientmono-540-lactase.aspx?activeingredientid=540>
- [15] C. Veeresham. (2012). "Natural products derived from plants as a source of drugs." (Online Journal). *Journal of Advanced Pharmaceutical Technology & Research*. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3560125/>
- [16] I. Tomlinson. (2004). "Next-generation protein drugs." (online article). *Nature Biotechnology*. <http://www.nature.com/nbt/journal/v22/n5/full/nbt0504-521.html>
- [17] L. Chong et al. (2006). "Kinetic Computational Alanine Scanning: Application to p53 Oligomerization." *Journal of Molecular Biology*. (online article). [https://www.researchgate.net/publication/7316090\\_Kinetic\\_Computational\\_Alanine\\_Scanning\\_Application\\_to\\_p53\\_Oligomerization](https://www.researchgate.net/publication/7316090_Kinetic_Computational_Alanine_Scanning_Application_to_p53_Oligomerization)
- [18] "Scientists boost potency, reduce side effects of IL-2 protein used to treat cancer." (2012). *Medical Xpress*. (online article). <http://medicalxpress.com/news/2012-03-scientists-boost-potency-side-effects.html>
- [19] "The Science behind Foldit." (2015). Fold.it. (website). <https://fold.it/portal/info/about>
- [20] D. Praetorius. (2011). "Gamers Solve 15-Year AIDS Problem in 3 Weeks." *The Huffington Post*. (online article). [http://www.huffingtonpost.com/2011/09/19/aids-protein-decoded-gamers\\_n\\_970113.html](http://www.huffingtonpost.com/2011/09/19/aids-protein-decoded-gamers_n_970113.html)



- [21] C. Eiben, J. Siegal, et al. (2012). "Increased Diels-Alderase activity through backbone remodeling guided by Foldit players." *Nature Biotechnology*. (online article).
- [22] "Gluten Destruction: Methods." (2011). *University of Washington* (online article). <http://2011.igem.org/Team:Washington/Celiacs/Methods>. <http://www.nature.com/nbt/journal/v30/n2/full/nbt.2109.html>
- [23] B. Gray (2013). "A Pill So People With Celiac Disease Can Eat Freely?" *Health*. (online article). <http://news.health.com/2013/02/08/a-pill-so-people-with-celiac-disease-can-eat-freely/>
- [24] G. Moore. (1965). "Cramming more components onto integrated circuits." *Monolithic 3D*. (online article). [http://www.monolithic3d.com/uploads/6/0/5/5/6055488/gordon\\_moore\\_1965\\_article.pdf](http://www.monolithic3d.com/uploads/6/0/5/5/6055488/gordon_moore_1965_article.pdf)
- [25] J. Poushter. (2016). "Internet access growing worldwide but remains higher in advanced economies." *Pew Research Center*. (online article). <http://www.pewglobal.org/2016/02/22/internet-access-growing-worldwide-but-remains-higher-in-advanced-economies/>
- [26] J. Niccolai. (2015). "Intel pushes 10nm chip-making process to 2017, slowing Moore's Law." *InfoWorld*. (online article). <http://www.infoworld.com/article/2949153/hardware/intel-pushes-10nm-chipmaking-process-to-2017-slowing-moores-law.html>
- [27] "Folding Forum." (2012). Folding Forum. (website). <https://foldingforum.org/viewtopic.php?f=16&t=21317>
- [28] K. Hult, P. Berglund. (2003). "Engineered enzymes for improved organic synthesis." *ScienceDirect*. (online article). <http://www.sciencedirect.com/science/article/pii/S0958166903000958>
- [29] "Scientific Background." (2008). Zinc Finger Consortium. (website). <http://www.zincfingers.org/scientific-background.htm>
- [30] S. Durai, M. Mani, K. Kandavelou, J. Wu, M. H. Porteus, S. Chandrasegaran. (2005). "Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells." (online article). *Oxford Journals*. <http://nar.oxfordjournals.org/content/33/18/5978.full>
- [31] "Human Genome." (2015). *Science Daily*. (online article).
- [32] J.J. Gair, M.D. Mehta. (2001). "Social, political, legal and ethical areas of inquiry in biotechnology and genetic engineering." *ScienceDirect*. (online article). <http://www.sciencedirect.com/science/article/pii/S0160791X01000124>
- [33] S. Wolpert. (2008). "Chemists Create 'Designer Enzymes'." *MNT*. (online article). <http://www.medicalnewstoday.com/releases/101236.php>

## ACKNOWLEDGMENTS

Special thanks to the Milwaukee School of Engineering for hosting the protein modelling event that inspired our interest in protein engineering. We would also like to thank the Folding@home and Foldit organizations for providing the programs discussed in this paper. In addition, we would like to recognize our writing instructor, Diane Kerr, for providing comments and critical review of our previous works.