Rational Design Protein Engineering Through Crowdsourcing

Background
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 natural selection, while rational design makes desired changes to a protein by directly manipulating its amino acid sequence. Directed evolution is currently more commonly used, since rational design relies on precise structural knowledge of the protein of interest, which is often unavailable. Two free programs, “Folding@home and “Foldit”, are making rational design more feasible by providing more precise protein models. 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 the sustainability of human beings, and agriculture.

Tylar Farmer
Patrick Bohse

University of Pittsburgh
Swanson School of Engineering

Directed Evolution

Directed evolution utilizes a controlled environment to create proteins through induced mutations and selection. If one wanted to obtain a protein that digests fructose using this method, he or she would implant a gene for digesting sucrose into a host cell, induce mutations on it, modify its environment by adding fructose, and then isolate the most fit host and extract the protein from that host. This process of directed evolution can be seen in figure 1

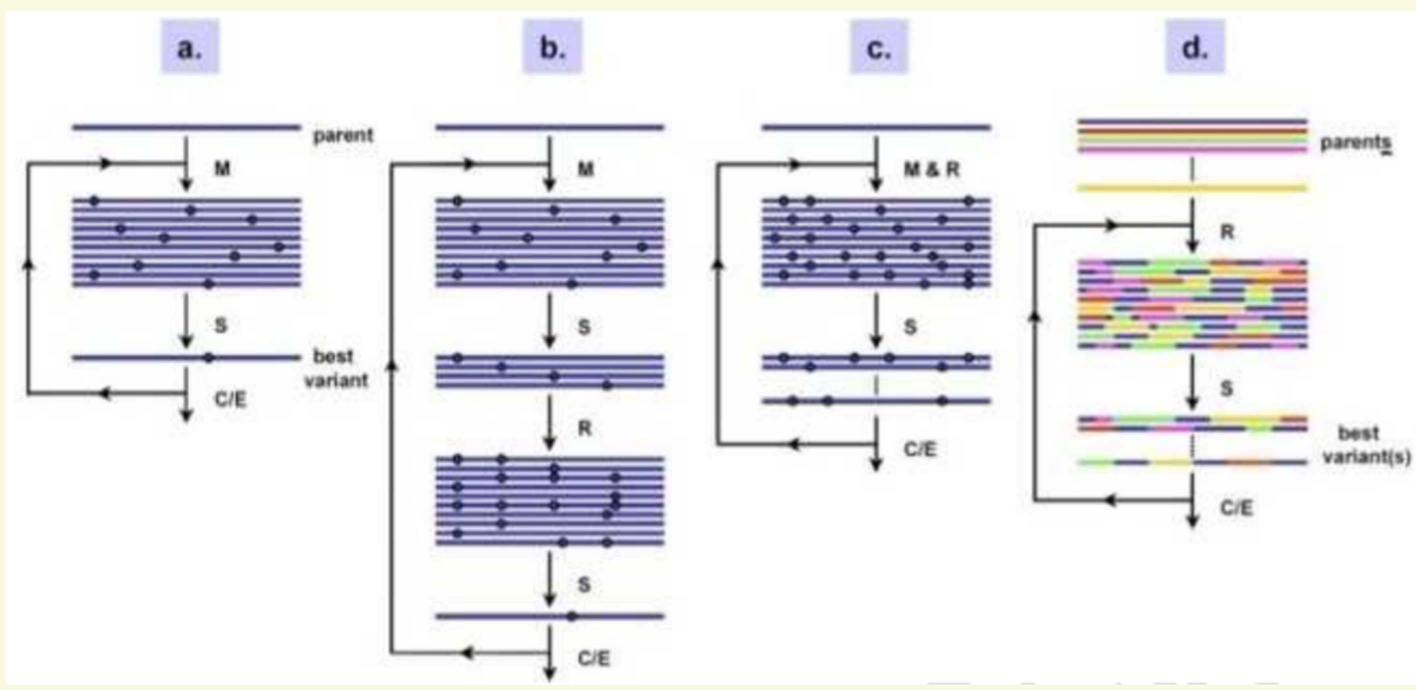

Figure 1 - Beginning with one or more parent genes, a mutagen (M) is applied and the host is allowed to replicate. A screening (S) is applied to isolate the genes that potentially code for the desired protein. Recombination (R) may follow, increasing genetic variation, followed by another screening. Experimenters will then 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.

Comparing Rational Design and Directed Evolution		
Method	**Rational Design**	**Directed Evolution**
Protein Structure Reliance	High	None
Reliance on Random Chance	Little	High
Estimated time for Product Yield	Predictable, relatively short	Unpredictable, relatively long

Rational Design

Rational design involves directly modifying an existing protein in order to create a protein that performs a desired task. If one wanted to obtain a protein that digests fructose using this method, he or she would load a model of an enzyme that digests sucrose into a simulation software and alter its amino acid sequence until the simulation indicates it can bind and break down fructose. The experimenter would then create this gene and implant it into a host and harvest the protein that the host secretes. This process can be seen in figure 2.

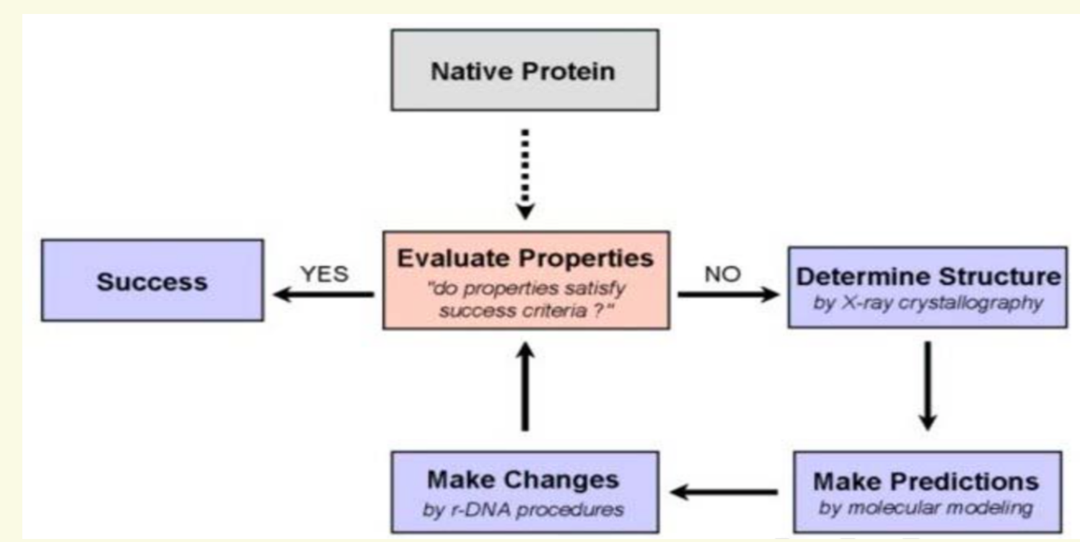

Figure 2 - Often, multiple rounds of the rational design process need to occur, but with each round, information about the structure of the protein and the way it functions is gained.

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.

Comparing Folding@Home and Foldit		
Program	**Folding@Home**	**Foldit**
Method	Computational Power	Human Intuition
Active Users	~100,000 (2004)	~59,000 (2008)
Publications	129 Publications	7 Publications

Foldit

Foldit relies on man’s ability to solve puzzles in order to improve protein models. Players are presented with a protein in a 3D environment, and are allowed to move and rotate the protein’s arms. The objective of the player is to achieve the highest score by creating the most stable conformation of the protein within given parameters; the higher the score, the more stable the protein. Figure 3 shows the game in action.

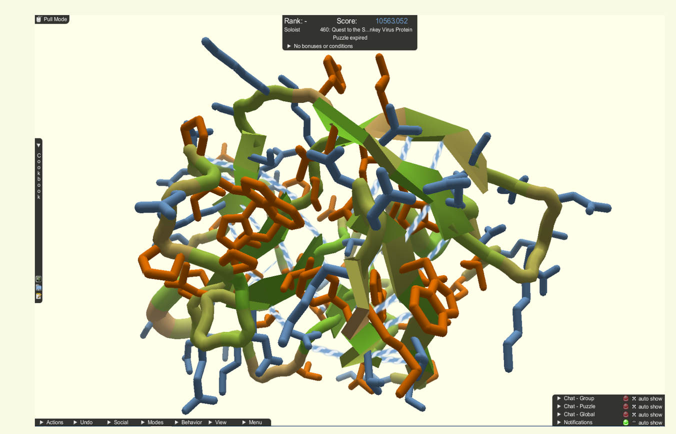

Figure 3 – The Mason-Pfizer Monkey Virus, shown here in Foldit, was improved upon by players after scientists couldn’t do it for 15 years.

Applications

Protein engineering via rational design has been used to create a custom Zinc Finger Nuclease (ZFN) with the ability to modify almost any other gene, and therefore its coded protein. As a result, the CCR5 gene, which AIDS\HIV relies on to thrive, was inhibited by a rationally designed ZFN in order to stop the onset or spread of those diseases. Additionally, through rational design, alternate forms of the enzyme glutathione transferase, known for detoxifying cellular environments, have been created, and are about 3000 times more efficient than their non-rationally-designed counterpart.

Sustainability

In relation to sustaining human life, rational design can contribute to the improvement of human fitness by the use of biopharmaceuticals. Patients lacking a certain protein can be supplemented with a rationally designed alternative capable of performing the same function as the absent protein. Additionally, rational design can be used to genetically modify crops to produce proteins that would make these crops better at obtaining water and nutrients from marginal soils, and make them resistant to insects, weeds, and frost, thus increasing our supply of food.