The Effects of Consensus Mutations on Yeast Enzyme Triosephosphate Isomerase

By Miriam Thomas

Advisor: Dr. Thomas J. Magliery

Senior Honors Thesis

The Ohio State University

Spring 2008

Abstract

Most natural proteins are only marginally stable, yet many therapeutic and industrial applications require stabilized variants. It has been shown that mutations to the most common residue in a multiple sequence alignment of protein families results in stabilization about half the time. The degree of conservation of these residues can be determined by measuring the difference between the amino acid distribution at a particular site and its neutral reference state. In this project, we examine if the degree of conservation for a position is related to its ability to affect the stability of the protein. Here, we are interested in studying the additive effects of mutating six positions in yeast triosephosphate isomerase (TIM) to consensus residues. To dissect the effects of these mutations, we are also constructing the six individual mutants which contain one highly conserved residue each. These variants will be studied for both structure and function. With this knowledge we will be able to reliably design more stable and functional proteins from genomic sequence data. This will have numerous benefits for research and industrial purposes, and will have potential health benefits in designing therapeutic proteins.

Introduction

Many proteins found in nature possess a marginally stable form, capable of surviving only under very specific conditions. This behavior is problematic for industrial and therapeutic purposes which require more stable variants. Various methods have been introduced to fix this problem, but finding one that is both accurate and efficient is challenging. In the past, rational design principles applied in protein engineering created success in stabilizing proteins. Thermostability pertains to the folding and unfolding properties of a protein. On theory thought to create stable proteins was through the analysis of the protein's crystal structure. From this information, a sequence was created that was believed to yield greater thermostability. An increase in stability among some proteins did occur, though it was not common[1]. Due to the lack of complete knowledge on protein folding and structure; the amount of rational concepts to predict stability is rather limited. It is near impossible to determine the exact folding of a protein, since various factors regarding correlated residues and its interactions with its surroundings play a significant role in its stability.

Directed evolution, another method to create stable variants, also succeeded in increasing thermostability. However, the chances of creating a more thermostable variant were rare for both random and rational methods used in this process[1]. Random mutations involved multiple rounds of cloning and screening for activity; while rational often required an in depth analysis of the protein's structure, followed by a trial and error method to find if an increase in stability was present. The difficulty with methods, such as random and rational design, is that it is time consuming, making it less efficient in creating stable proteins.

Creating mutations based off of statistical analysis is another method in engineering proteins. The benefit of using this method is that it is very efficient in deciding which residues

should be mutated, as opposed to guesswork. It has been shown that mutations to the most common residue in a multiple sequence alignment of protein families results in stabilization about half the time[1]**.** One study using this approach, found that the consensus fungal phytase unfolded 15-22 °C higher than the wildtype phytase[2]. Although only 13 sequences were aligned to create the consensus sequence, the study showed that the thermostability of a protein can be improved. Consensus design was further examined in ankyrin repeat proteins, whose function is to mediate protein-protein interactions by binding to various target polypeptides. 8,000 sequences were collected to make this consensus sequence. The thermodynamic stability of the consensus protein was clearly higher than the natural ankyrin proteins. In addition, it could tolerate variable surface residues, making it a more effective binding protein than the wildtype[7]. *Main et al. (2003)* concluded after investigating the consensus sequences in his study of repeat proteins that the increase in stability is due to a combination of factors such as the removal of unfavorable interactions and the absence of irregular insertions and deletions[10]. It illustrates the evolutionary concept of preference in selection for a particular function.

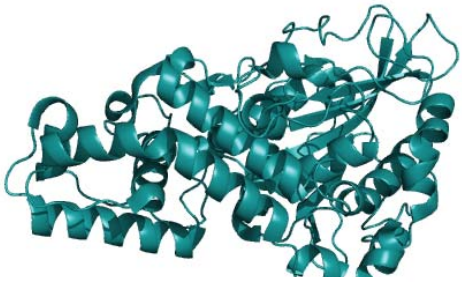


Figure I: Fungal Phytase

As mentioned before, consensus mutations are only 50 % of the time more stable than the wildtype. Some mutations can decrease in activity because of correlated mutations, in that when mutating one residue it is necessary to mutate the other in order for the protein to still be active. Also, when amino acids are poorly conserved, it is difficult to determine the appropriate consensus residue at a position. In Lehmann's work with fungal phytase, 10 of the positions on

the consensus protein were ambiguously chosen[2]. Although the synthetic protein was more stable, it is not always the case in other situations.

In this project, we will examine if the degree of conservation for a position is related to its ability to affect the stability of the protein. The hypothesis is that more conserved positions will have more positive effects on the stability, since nature selects for function. The goal of this project is to create not only a more thermostable protein, but one which maintains wild type-like properties.

We are testing the significance of highly conserved residues by choosing the six most conserved residues of an amino acid sequence. These six residues were selected by creating a multiple sequence alignment of a protein from 644 different organisms. Unlike the fungal phytase study, this should give a more accurate sense of which residue is consensus, and who is more conserved. Highly conserved residues are defined as consensus amino acids which appear more strongly at a particular site compared to a consensus amino acid at another. The cut off for a residue considered to be "highly conserved" is determined by using the relative entropy equation.

$$\mathbf{D(p||}\boldsymbol{f}\mathbf{)= \sum p_x \ln (p_x/}\boldsymbol{f}_x\mathbf{)}$$

This equation contains variables which compare the frequency of an amino acid at a particular position coming from the multiple sequence alignment($\mathbf{p}_x$) and the frequency of that amino acid at a position coming from all the proteins existing in yeast ($\boldsymbol{f}_x$). A relative entropy value greater than 2 is considered highly conserved. This number was chosen since a relative entropy value of three contained residues already present in the sequence. The six positions mutated onto this gene are shown in Figure 1 on the following page.

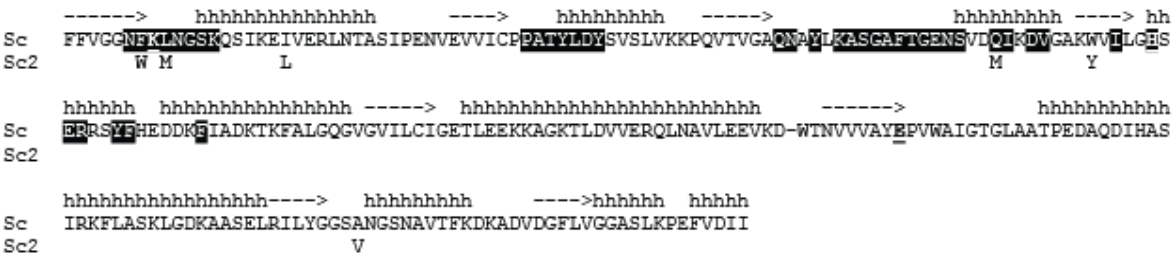


Figure 1: Six mutations made onto the wildtype yeast gene. Sc is the wildtype yeast sequence and Sc2 is the sequence with six mutations.

The protein chosen for this study was triosephosphate isomerase (TIM). It is the glycolytic enzyme used for the isomerization of dihydroxyacetone phosphate (DHAP) into glyceraldehyde 3-phosphate (G3P). Figure 2 displays the reaction TIM catalyzes in glycolysis.

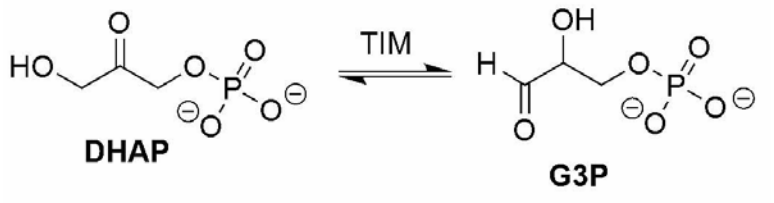


Figure 2: Isomerization of DHAP into G3P catalyzed by TIM

This enzyme was selected since it is a ubiquitous enzyme and we are aware of its structure, activity and a few of its thermodynamic properties. Also, TIM possesses an α-β barrel secondary structure, which is common in many of the proteins found in nature. This structure consists of eight alpha helices and beta strands. Yeast is selected as the model organism in the project since it is very well-studied. Much of the work done with TIM have used yeast as its model, such as the *in vitro* assay to study TIM activity[4].

The six highly conserved residues will be mapped onto the yeast triosephosphate isomerase gene and characterized (Figure 3). Six individual mutants, each containing one highly conserved residue, will be studied as well to determine the individual effect. The genes will be subcloned into a vector for *in vivo* screening via a strain of *E. coli* deficient in triosephosphate isomerase activity. Biophysical characterization is also performed to determine folding and thermal stability (Circular Dichroism spectroscopy), quaternary structure (gel filtration chromatography) and Michaelis-Menten kinetics.

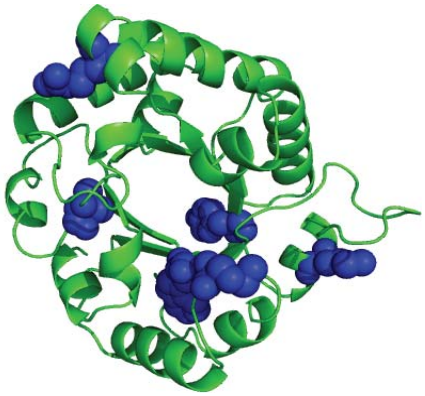


Figure 3: Sc2TIM consisting of eight alpha and eight beta strands. The blue spheres represent the six mutated residues.

Chapter 1: Analysis of Sc2TIM

I. Materials and Methods

Statistical Analysis of Relative Entropy

Because Sc2 is an unnatural protein, it was engineered by finding its DNA sequence from back-translating the amino acid sequence of the protein. The program used for backtranslating the amino acid sequence to DNA was Entelechon. After determining the DNA sequence, overlapping oligos were designed to reassemble the gene [3]. A total of ten oligos, approximately 100 nucleotides each were ordered from Sigma Genosys. Using Phusion polymerase (New England Biolabs), the oligos were assembled into a full length gene by PCR. Figure 4 shows sequences of all 10 oligos ordered for the synthesis of Sc2.

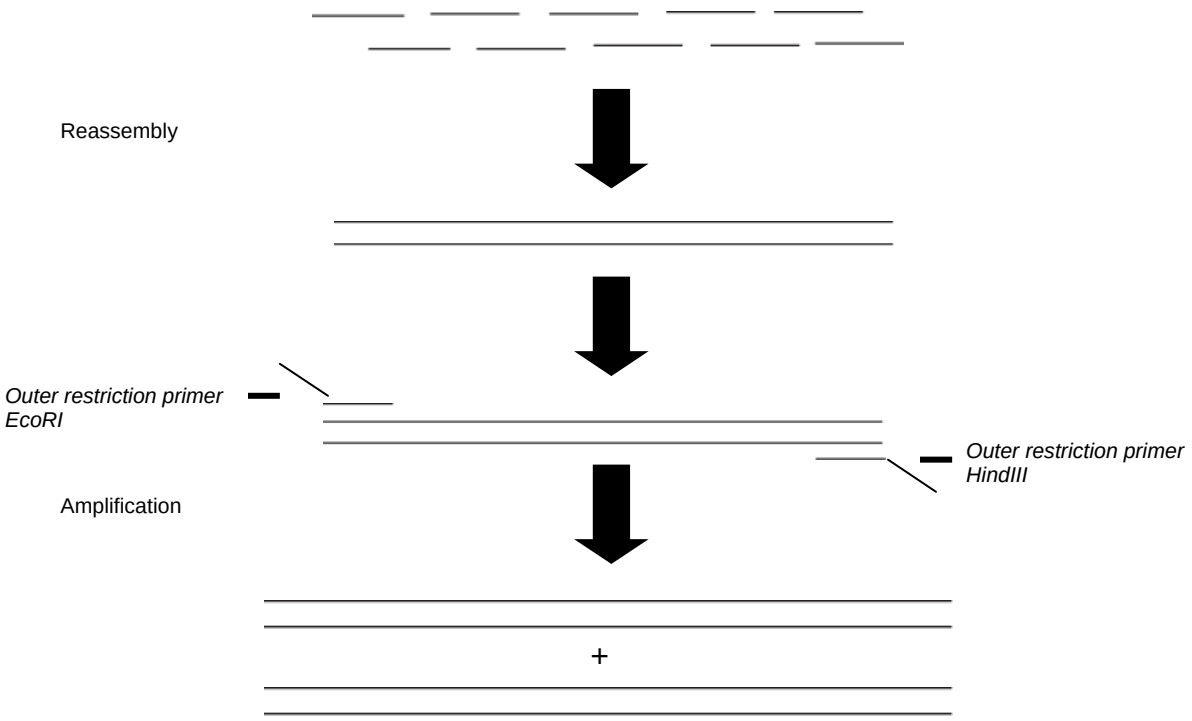
Sc2_1	GCGCGTACGTTCTTTGTTGGTGGTAACTGGAAAATGAACGGTTCTAAACAGTCTATCAAGGAAATCGTTGAACGTCTCAACACTGCGTC
Sc2_2	CTTAACCAGAGAAACAGAGTAGTCCAGGTAGGTCGCCGGAGGGCAGATAACAACTTCAACGTTTTCCGGGATAGACGCAGTGTTGAGACGTTC
Sc2_3	TGGACTACTCTGTTTCTCTGGTTAAGAAACCGCAGGTGACCGTTGGCGCCCAGAACGCGTACCTGAAGGCATCTGGTGCGTTCACCGGCGAG
Sc2_4	GAAGTAAGAACGACGTTCAGAGTGACCGAGGATAACGTATTTCGCACCAACGTCTTTGATCATATCGACAGAGTTCTCGCCGGTGAACGCAC
Sc2_5	TCACTCTGAACGTCGTTCTTACTTCCACGAAGACGACAAGTTTATCGCGGACAAAACCAAATTCGCTCTGGGTCAGGGTGTTGGCGTGATTC
Sc2_6	GTTGAGCTGACGTTCGACCACATCCAGAGTTTTACCCGCTTTACGCTCTTCCAGGGTTTCACCGATGCACAGAATCACGCCAACACCCTGAC
Sc2_7	GGTCGAACGTCAGCTCAACGCTGTTCTGGAGGAAGTTAAAGACTGGACCAATGTTGTCGTTGCGTACGAACCGGTGTGGGCGATTGGTACTG
Sc2_8	CCCAGTTTAGAAGCCAGGAATTTACGAATAGAGGCGTGGATGTCCTGCGCGTCTTCCGGGGTTGCCGCCAGACCAGTACCAATCGCCCACAC
Sc2_9	AATTCCTGGCTTCTAAACTGGGTGACAAAGCGGCGTCTGAACTGCGTATCCTCTACGGTGGTTCTGTTAACGGTAGCAACGCGGTAACCTTC
Sc2_10	GATGATGTCCACAAATTCCGGCTTCAGAGACGCACCGCCAACCAGGAAACCGTCAACGTCCGCCTTATCTTTGAAGGTTACCGCGTTGCTAC

Figure 4: Sc2 Oligomers. Even numbered oligos are forward (5' to 3') and odd number labeled oligos are reverse (3' to 5').

Reassembly [3]

The reassembly reaction, ran by PCR, consisted of oligonucleotides that contained an 18 base pair overlap with each other. This was then followed by a PCR amplification reaction using DeepVent polymerase. The outer (restriction) primers used to attach onto the gene were based off of the EcoRI and HindIII restriction enzyme cutting site sequences.

Figure: Synthesis of Sc2 gene



PCR conditions for reassembly consisted of 95 ˚C for 30 seconds, 60 ˚C for 45 seconds and 72˚C for 45 seconds. There were 25 total cycles. For amplification, the conditions were 95 ˚C for 2 minutes, 95 ˚C for 30 seconds, 58.0 ˚C for 45 seconds, 72 ˚C for 45 seconds. These conditions were done for 24 cycles. A final extension time was done for 4 minutes, making 25 cycles total.

Phenol chloroform-chloroform isoamyl cleanup/ethanol precipitating:

Phenol chloroform-chlorophorm isoamyl alcohol cleanup and ethanol precipitating was done to purify the DNA . Based on the total volume of the PCR reaction, the same amount of phenol

chloroform was added into the tube with the reaction. The tube was vortexed, centrifuged for 2 minutes and the bottom organic layer was extracted. The same amount of chloroform isoamyl was then added into the tube with the reaction, vortexed for one minute and then centrifuged for two minutes. The bottom organic layer was extracted again. 1/10 the volume of NaOAc and 2.5 times the volume of absolute ethanol was added into the PCR reaction. The tube was store at -80 ºC overnight. The next day the tube spun down in the 4 ºC centrifuge at 13.2 rpm for 30 minutes. The ethanol was poured out, leaving a pellet. The same volume used to add absolute ethanol is used for the addition of 70 % ethanol to resuspend the pellet.. The mixture is spun down at 4 ºC for 30 minutes again, decanted and left to air dry.

Then a restriction digest was done using the enzymes EcoRI and HindIII to cleave the DNA. The gene was inserted into the vector, XaI Pinpoint (which also was digested by the same enzymes). The reaction was set in the 37 ˚C incubator for 3.5 hours and heat killed at 65 ˚C for 20 minutes.

Gel Purification

Restriction digested sample is loaded onto an expanded lane of 1% agarose gel and run at 175V. Gel left to run until the DNA band was sandwiched between glass fiber paper and a dialysis membrane. The sample is eluted off sandwich into spin column. The total volume of the solution is measured and 1/10 the reaction volume is added of 3M NaOAc, following the addition of 2.5 times the volume of absolute ethanol. The tube was placed in the -80 ºC freezer overnight and then spun down at room temperature for 1 hr next morning. The ethanol was decanted and replaced with 2.5 times the same volume of 70 % ethanol. The sample was spun down for 30 minutes, supernatant discarded, and allowed to air dry.

Ligation:

Ligation of Sc2TIM into the Xa1Pinpoint vector followed the gel purification. 1/10 reaction volume of T4 ligase buffer was added with 1 μL of insert, vector and T4 DNA ligase. A negative control of this same reaction was set up, using all the same amount of reagents, but no insert. The reaction was placed in the 16 ºC overnight.

Transformation of Sc2TIM into DH10B cells

A transformation of the plasmid containing the TIM gene into DH10B *E.coli* cells was carried out. Contents kept on ice during experiment. 30 µL DH10B electrocompetant cells and 1 µL Sc2 ligation were mixed and transferred into a cuvette. The cuvette was placed into an electroporator and pulsed. 1 mL of 2YT was added immediately. The mixture was stored in a culture tube and incubated at 37 ºC shaker-incubator for 1 hour before being plated on fresh LB-amp plates. 75 µl of the transformed cells in media was plated on LB amp media plates and were placed inside the 37 ºC incubator over night for 12 hours. Colonies were found the next morning and were picked/ inoculated in 2YT media and 1x ampicillin. The cultures were left to grow overnight in 37 ºC shaker-incubator. Approximately 16 hours later, 2 mL of the culture was taken out and placed into an 2 mL eppendorf and was spun down in a centrifuge at 13.2 rpm. The supernatant liquid was disposed and the cell pellet at the bottom of the of the tube was saved. The pellet was miniprepped using the Qiagen Miniprep Kit.

Clones were confirmed by DNA sequencing (Genewiz). When obtaining a successful sequence, protein purification is the next step. Using the miniprepped plasmid containing the Xa1 Pinpoint plasmid with Sc2TIM, external pHLIC primers were added via PCR. The same PCR conditions used to assemble Sc2TIM, was used to adhere the primers onto the plasmid. The PCR product was inserted into the pHLIC vector, using the method of ligation independent cloning. The polymerase used is T4 polymerase. This Sc2 insert combined with pHLIC vector is then transformed into DH10B cells. The colonies were miniprepped and sequenced. The plasmids from the miniprep were than transformed into BL21 (DE3) cells. The colony is then used to inoculated 25 mL media (2YT/Amp) and grown overnight.

Preparation 2YT media to grow cells for protein expression

1 L of 2YT media is prepared and autoclaved. 16 g Bacto-Tryptone, 10 g yeast extract and 5 g NaCl are weighed added into a flask which is filled to one liter of distilled water. The media is left to spin for 20 minutes before being placed in the autoclave for 20 cycles.

Protein Purification

The 1 L of 2YT/Amp is inoculated with the 25 mL overnight culture at 37 ºC. The cells are grown to an OD_{600} of approximately 0.75. 1 mL of IPTG was added to flask to create a 0.1 mM IPTG concentration. The culture is grown for 3 more hours at 37ºC in the shaker-incubator. The cells are spun down in GS3 tubes (at 6k rpm for 10 minutes). Media is then decanted and the cell pellet was stored overnight in the -80 degree Celsius freezer. The pellet is then resuspended with 25 mL of lysis buffer (50 mM Tris HCl pH8, 300 mM NaCl, 10 mM imidazole, 10 mM β-mercaptoethanol) The following reagents are then added into the tube containing the resuspended cell pellet: 75 µL 2M MgCl2, 150 µL 100 mM CaCl2, 5 µL DnaseI, 5 µL RNase A, lysozyme (used 30 mg HEW lysozyme dissolved in 1 mL lysis buffer and added entire 1 mL) and 10 % Triton X-100 (added 300 uL to make a final 0.1% concentration).
The mixture is incubated at 4 ºC for an hour. The sample is sonicated three times for 30 seconds on ice- with 2 minutes between each pulse (pulses were set at 4.0, 5.0 and 6.0). The mixture is then transferred into SS34 tubes and spin at 20 K at 4 ºC for one hour. The supernatant is decanted into a fresh 50 mL conical and add 2.0 mL of the Qiagen 50% slurry of Ni-NTA agarose. The solution is left in cold room to mix for two to three hours on the nutator mixer. The solution put through a large pre-fritted column.10 uL of flow through is saved in a 1.5 mL eppendorf tube. The mixture was washed with 10 mL of wash buffer (50 mM Tris. HCl, pH 8, 300 mM NaCl, 20 mM imidizole, 2 mM BME).10 uL of solution from wash is saved into a 1.5 mL eppendorf. 4 mL of elution buffer is then added to mixture (50 mM Tris. HCl, pH 8, 300 mM NaCl 250 mM, 2 mM BME). The gel of flow-thru, wash and elution samples with equal volume of SLB dye (12.5% SDS-PAGE, 200 V) is run. 1 mL of lysis buffer to the 4 mL elution is added to mixture, as well as DTT to make a 5 mM final concentration. One aliquot of TEV protease is also added and the solution is left to incubate overnight (>8 hours)) at room temperature without agitation. Another aliquot of TEV protease is added with 5 mM DTT and the solution is left to incubate another 3-4 hours. Two PD10 columns with 25 mL of lysis buffer each are equilibrated. 2.5 mL (half) of the cleaved eluent is added to each column and left to run through to waste. 3.5 mL of lysis buffer is added to each column and the solution is collected into a conical tube. Added 2 mL of 50 % Ni-NTA agarose slurry. The solutions are mixed in the cold room for 4 hours. They are then loaded into a small pre-fritted column and where the flow through is collected. This is where the protein is. 7 mL of the eluent is concentrated down to 1 mL by using

a Centriplus YM-10 tube. They are then spun down at 4 degrees Celsius at 2900 rcf. The 1 mL solution is transferred to a 1.5 mL eppendorf tube. A 1-Nap-10 column is equilibrated with 25 mL TIM storage Buffer (50 mM Sodium Phosphate, 300 mM NaCl, pH 8.0). 1 mL of concentrated protein is added to the column and allowed to flow to waste. The sample is eluted with 1.5 mL buffer and collected. To this solution, 5 mM DTT and 1 mM TCEP are added to the purified protein.

Circular Dichroism Spectroscopy

Both methods use far UV, since it is taking note of the secondary structure of the protein.

Wavelength Scan

Both Sc2TIM and ScTIM were compared. Ellipticity was monitored via AVIV Circular Dichroism spectrophotometer at 17 μM protein. Wavelength parameters were set between 195.00 nm to 300.00 nm. A sample was taken every 1.0 nm with an averaging time of 5.00 seconds. The number of scans taken was 3.

Thermal Melt

The temperature ranges from 5 degrees Celsius to 95 degrees Celsius. Data was collected every degree, with 30 second equilibrations and an averaging time of 10 seconds. Both Sc2TIM and ScTIM compared.

Gel Filtration

The sample of protein is injected into the Pharmacia FPLC and eluted from a Superdex 75 10/300 column at 0.4 mL/min with 50 mM Tris buffer, 100 mM NaCl, pH 8. The experiment is run at 4ºC. The absorbance is recorded at 280 nm. Both wildtype ScTIM alongside Sc2TIM are run in order to compare sizes since ScTIM is known to be dimeric at 4 ºC.

Michaelis-Menten Kinetics [4]

In the following order, TEA Buffer (pH 7.4), 1mM NAD+, 6.0 mM, NaAs, 5.0 mM EDTA and 0.17 mg/ mL GAPDH is added into a 1.5 mL eppendorf. Incubated tube for 2.0

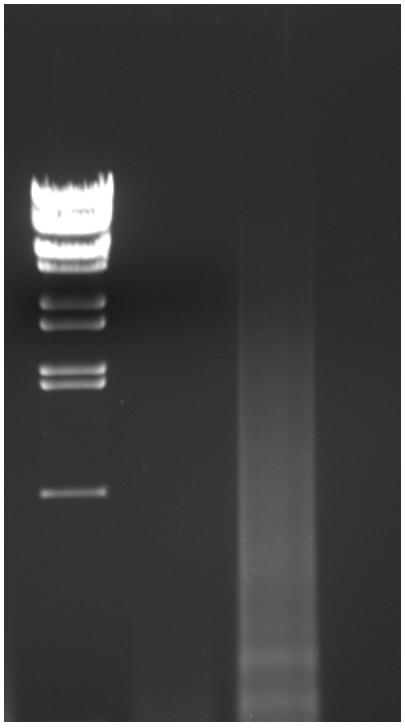
minutes in a 37 ºC water bath. Immediately added TIM (40 ng/mL) into the reaction and transferred into a cuvette and took UV absorbance. The UV spectrometer had a temperature set at 37ºC. Readings were taken every 45 seconds for 5 minutes at varying DHAP concentration. The following DHAP concentrations used to measure the rate was at 2.0 mM, 4.0 mM, 6.0 mM, 8.0 mM, 10.0 mM, 12.0 mM, and 14.0 mM. The species absorbing the UV light is NADH, which possesses an extinction coefficient of 6220 cm^{-1}. The reactions were done in triplicate to attain an accurate rate. The pH of the reaction was mainted at approximate 7.4, which is near biological conditions and is the optimim pH which yeast TIM works. The reaction conditions were based off the Plaut and Knowles paper, *pH-Dependence of the Triosephosphate Isomerase Reaction* [4]. After collecting absorbances, the data was transferred to the computer program, Kaleidoscope, which created the Michaelis-Menten curve.

II. Results

A. Cloning of Sc2TIM

The 10 oligos ordered from Sigma Genosys, as mentioned before, were put together in the PCR using DeepVent polymerase. When running a 1% agarose gel, a smear is seen, meaning various sized products have been formed in the PCR. The ladder used to identify the approximate band sizes is BstEII Digested Lambda ladder. The size of the Sc2 TIM gene is approximately 744 nucleotides in length. Figure 5 shows reassembly reaction was assumed to be successful since there is product formed at the 744 base pair line.

Figure 5. PCR reassembly displays a smear of product around the 700 basepair line of ladder, signifying product is made.

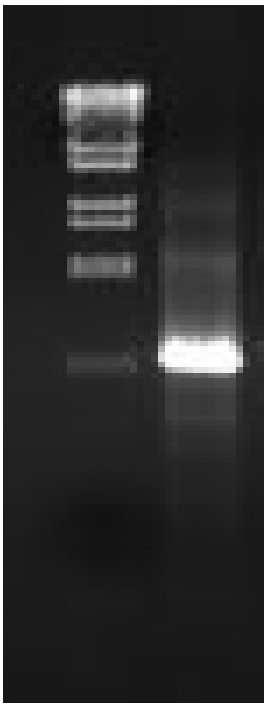


Lambda Ladder Sc2

An amplification of the reassembly was done to form the specific product; once again, using DeepVent polymerase. The PCR conditions for this reaction of 25 cycles involved 95ºC for 2

minutes, 95ºC for 30 seconds, 58ºC for 45 seconds, 72ºC for 45 seconds, this is repeated in 24 cycles before going onto the final extension step of 72ºC for 4 minutes. A sharp band appearing at the 750 base pair line signifies that the Sc2 band is formed (Figure 6).

Figure 6: PCR amplification product displays desired bright band of correct size.



Lambda Ladder Sc2

After the phenol chloroform-chlorophorm isoamyl cleanup/ethanol precipitating, both the insert and vector, XaI Pinpoint were digested with Hind III and EcoRI and gel purified. After the products were ethanol precipitated, ligated together and transformed, colonies were grown to saturation and miniprepped. These were then analytically digested using the same restriction enzymes. The top band appears near the 3300 base pair line (which is the presence of cut vector) and the bottom band appears at the 750 base pair line, meaning the Sc2 insert was successfully cut.

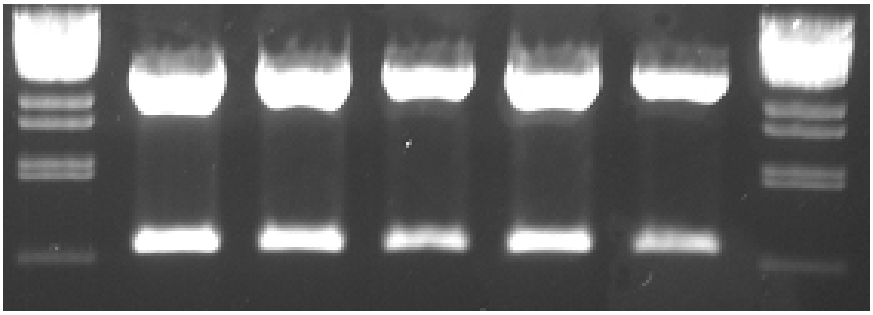


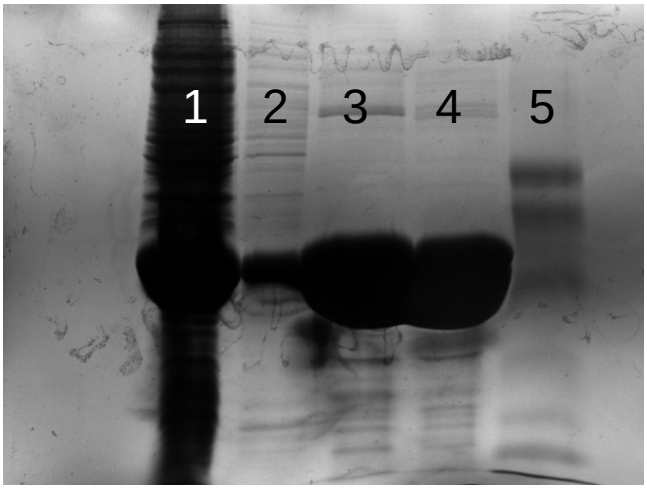
Figure 7: Lambda Ladder and five colonies containing Sc2 which were miniprepped and analytically digested

The analytical digest shows that the Sc2 insert successfully ligated with the digested Pinpoint vector and was cloned (figure 7). The minipreps of the samples were sent off for sequencing to Genewiz and the results showed the samples containing no unwanted mutations present.

B. Protein Purification of Sc2TIM

Sc2 was made through ligation independent cloning using T4 polymerase and single stranded DNA ends from pHLIC primers. This was cloned into the pHLIC vector and transformed. The colonies were miniprepped and sequenced to make sure it was the correct sequence. Once verified, the miniprep was then transformed into BL21 (DE3) cells and the protocol for protein purification was followed.

Figure 8. SDS PAGE gel of the purified protein



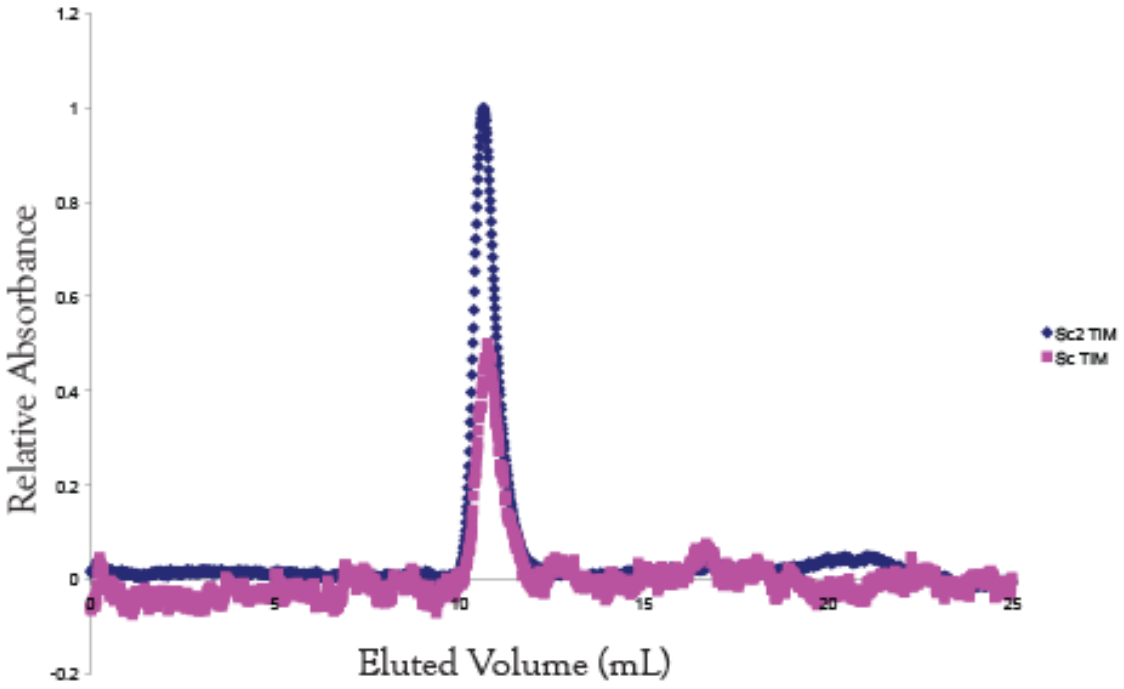
1 Flow Through of Protein
2 Wash buffer
3 Elution buffer
4 Purified protein after Tev reaction (this is done to remove the His-6-Tag site on protein, so only protein is present
5 Ladder

C. Biophysical Characterization of Sc2TIM

i. Gel Filtration (FPLC)

Gel filtration gave information about Sc2TIM's quaternary structure. Wildtype ScTIM was run along side of it in comparison. Wildtype ScTIM is a dimeric protein at 4 degrees Celsius (molecular weight is 55 kD). When Sc2TIM was put through the FPLC, it was found to elute at the same exact volume as the wildtype, meaning it is a dimeric protein as well. If it were to be monomeric (molecular weight of 29 kD), the peak would have appeared further to the right, since the protein would elute at a larger volume. The peaks are shown at different heights since one of the intensities was multiplied by 0.5, allowing for both to be shown.

Figure 9: Gel Filtration Chromatography of Sc2TIM vs. Sc2TIM



ii. Circular Dichroism Wavelength Scan

A CD wavelength scan of Sc2TIM was obtained. A sample of ScTIM wildtype was run alongside in comparison. Peaks were present at 208 nm, 215 nm and 222 nm. The circular dichroism wavelength scan shows both the wildtype yeast TIM and the mutant Sc2TIM in comparison. Sc2TIM shows signal ellipticity at 208 and 222 nm, meaning it contains alpha helices in its secondary structure. There is also signal at 215 nm, meaning beta sheets are present in its structure as well. The secondary structure of Sc2 has almost identical characteristics to the wildtype yeast TIM. With only six mutation mapped onto the wildtype yeast TIM sequence, it was not too surprising that the presence of both alpha and beta sheets remained.

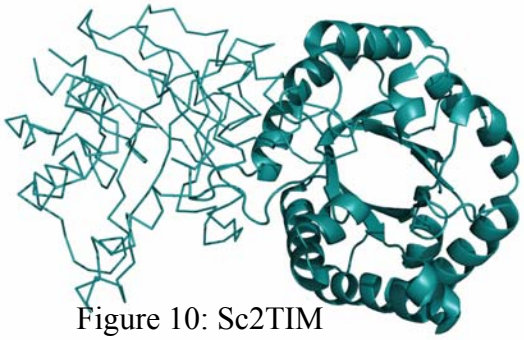


Figure 10: Sc2TIM

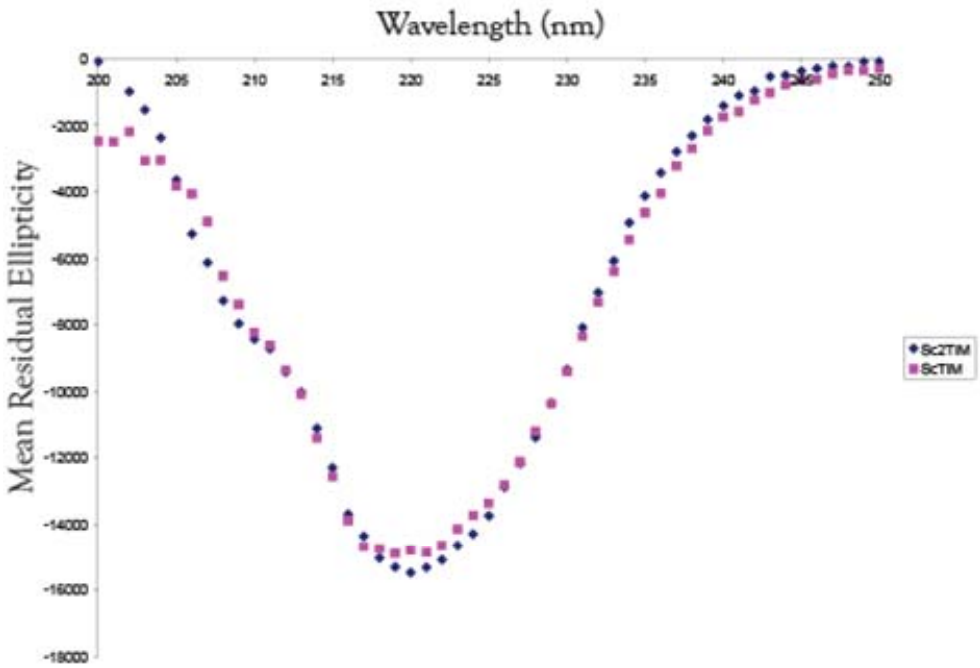
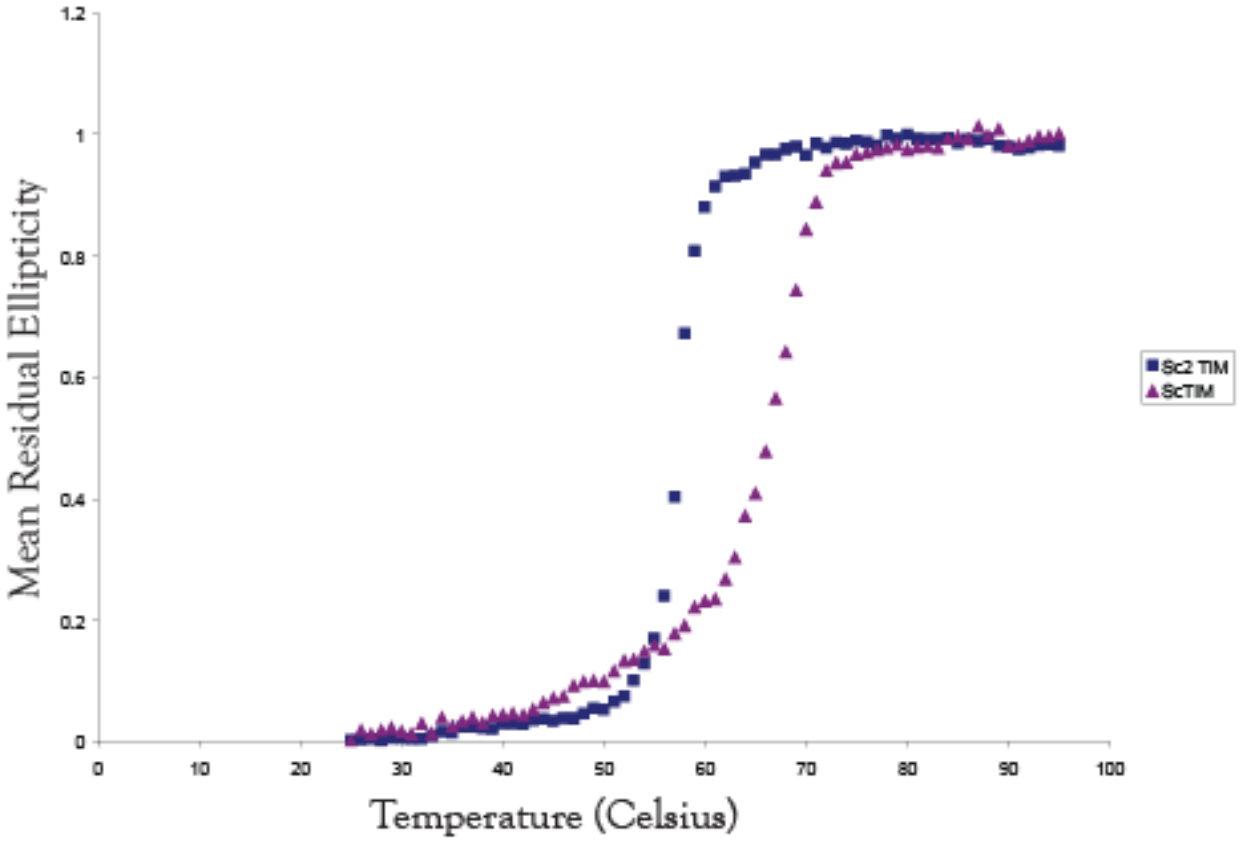


Figure 11: CD Wavelength Scan

iii. CD Thermal Melt

Thermal melt of Sc2TIM was run alongside wild type ScTIM. Sc2TIM unfolded at a slightly lower temperature than the wild type, as well as unfolded more cooperatively. Sc2TIM is not as thermodynamically stable compared to the wild type under higher temperatures. This thermal melt was taken at a wavelength of 222 nm, meaning the denaturation of the secondary structure is being read from the change in absorbance from the alpha helices.

Figure 12: CD Thermal Melt

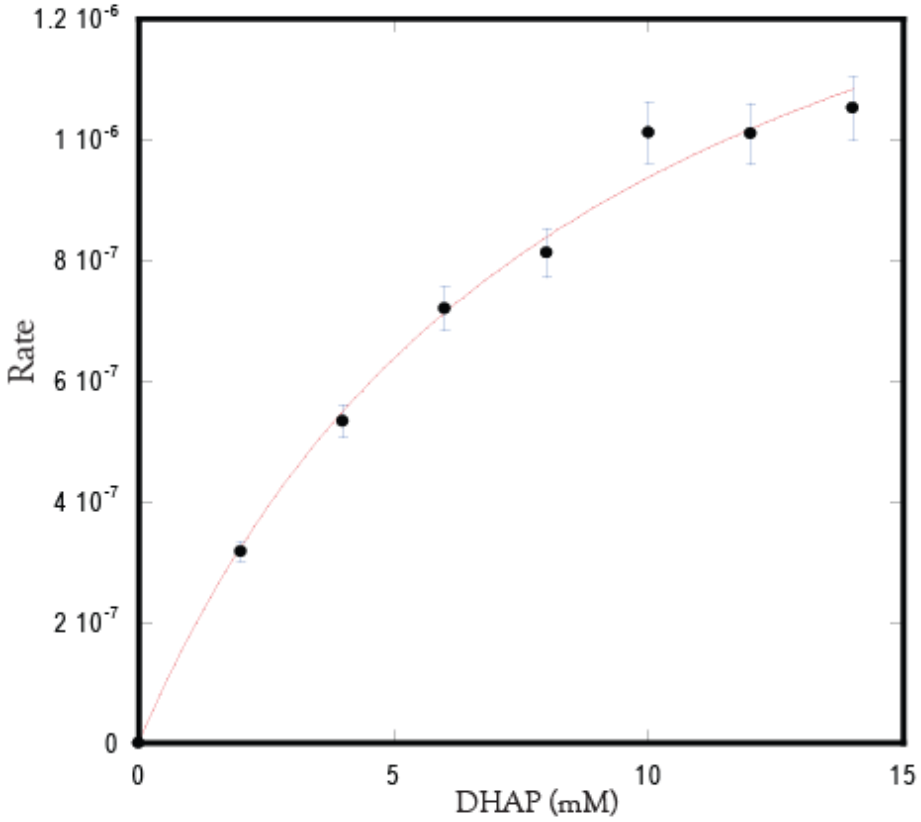


D. Activity

i. Michaelis-Menten Kinetics

The results for this experiment showed that Sc2TIM had a higher K_M value than the literature value for wild type yeast TIM. The K_M value gives information about the binding efficiency between the enzyme and substrate. Although the K_M appears to be higher, it is within error and of the same value for both Sc and Sc2TIM. The k_{cat} and k_{cat}/K_M value of both wildtype and Sc2TIM were in about the same magnitude of each other. The similar k_{cat} value meant that the turnover rate of substrate into product was about the same per molecule of enzyme. The k_{cat}/K_M value also shows that the rate constant of the wildtype is similar to Sc2TIM. Michaelis- Menten curve is shown in Figure 13.

Figure 13: Michaelis –Menten Curve



	Source	K_M (mM)	k_{cat} (min^{-1})	k_{cat}/K_M
Sc2TIM		8.9 (±1.5)	(7.16±0.02) x 10^ -4	(8.01±0.17) x 10^- 6
ScTIM	Plaut & Knowles, Biochem J., 1972	3.8 (±0.03)	(2.92 ±0.04) x 10^-4	(7.68±0.01) x 10^ -6

III. Discussion

The stability of wildtype yeast TIM was predicted to increase once highly conserved residues were mapped onto its genome. This hypothesis was based on the fact that nature selects for function, thus if many organisms possess a characteristic: it must be beneficial for its function. In previous studies, using consensus design successfully created more stable variants. To further test the degree of conservation, triosephosphate isomerase, a ubiquitous enzyme, was chosen for the study.

Through the various methods of analysis on Sc2, much can be inferred about its biophysical characteristics and activity. The gel filtration of Sc2TIM illustrated that it is a dimeric protein at 4 degrees Celsius. This shows that mutating six of the most highly conserved residues did not significantly change its quaternary structure. This was expected since there were only six residues mutated onto the wild type gene.

Just as the quaternary structure of Sc2TIM was similar to the wild type, so was the secondary structure. The circular dichroism wavelength scan illustrated the presence of both alpha and beta sheets present in Sc2TIM. Also, the path of ellipticity follows a very similar shape as wild type TIM, as seen in figure 10. Once again, with only six mutation mapped onto the wild type yeast TIM sequence, it was not too surprising that the presence of both alpha and beta sheets remained.

The circular dichroism thermal melt illustrated the one biophysical characteristic which deviated from the wild type-thermostability. Figure 11 shows that Sc2TIM unfolds at a slightly lower temperature than wildtype, meaning Sc2TIM is not as stable under higher temperatures compared to the wild type TIM. It is hypothesized that the presence of one of the highly conserved mutations is destabilizing. This was somewhat surprising in that past studies have

shown how consensus residues increase the T_m of a protein, and these consensus mutations decreased the T_m. When the consensus protein of triosephosphate isomerase was studied, it also demonstrated a higher T_m, as well as less cooperatively in its unfolding. Sc2TIM, consisting of only the most highly conserved residues, illustrated not only unfolding at a lower temperature, but more cooperativity in its unfolding as well.

Michaelis-Menten kinetics also illustrated Sc2TIM's behavior deviating from the wildtype. The K_M value of the Sc2TIM was slightly higher than the wild type, but within error. This means the binding efficiency of the enzyme was very similar to the wild type properties of yeast TIM.. This was not surprising since only six mutations were mapped onto the wild type gene. However, it was not expected that the K_M would be considerably lower than the wild type either. The Sc2TIM k_{cat} and k_{cat}/K_M values were very similar to the wild type which was also expected.

Overall, it seems that Sc2TIM has many overlapping similarities regarding the physical characteristics of the wild type ScTIM. This was part of the aim for this project was to create a protein that maintains the same wild type properties and is functionally active. However, the main goal was to create a more stable variant. Unfortunately, the weaker binding efficiency and decrease in stability under higher temperatures, illustrated the presence of destabilizing mutations that were mapped onto the wild type TIM. This goes against our hypothesis, which predicted Sc2TIM would become a more thermostable protein compared to the wild type. A likely reason causing the instability may be due to two of the mutations being adjacent to the active site. The two mutations near the active site involve a switch from phenylalanine to tryptophan and a leucine to methionine. The first mutation (phenylalanine to tryptophan) leads to a change in size, and a very small change in polarity within that site. The second mutation from

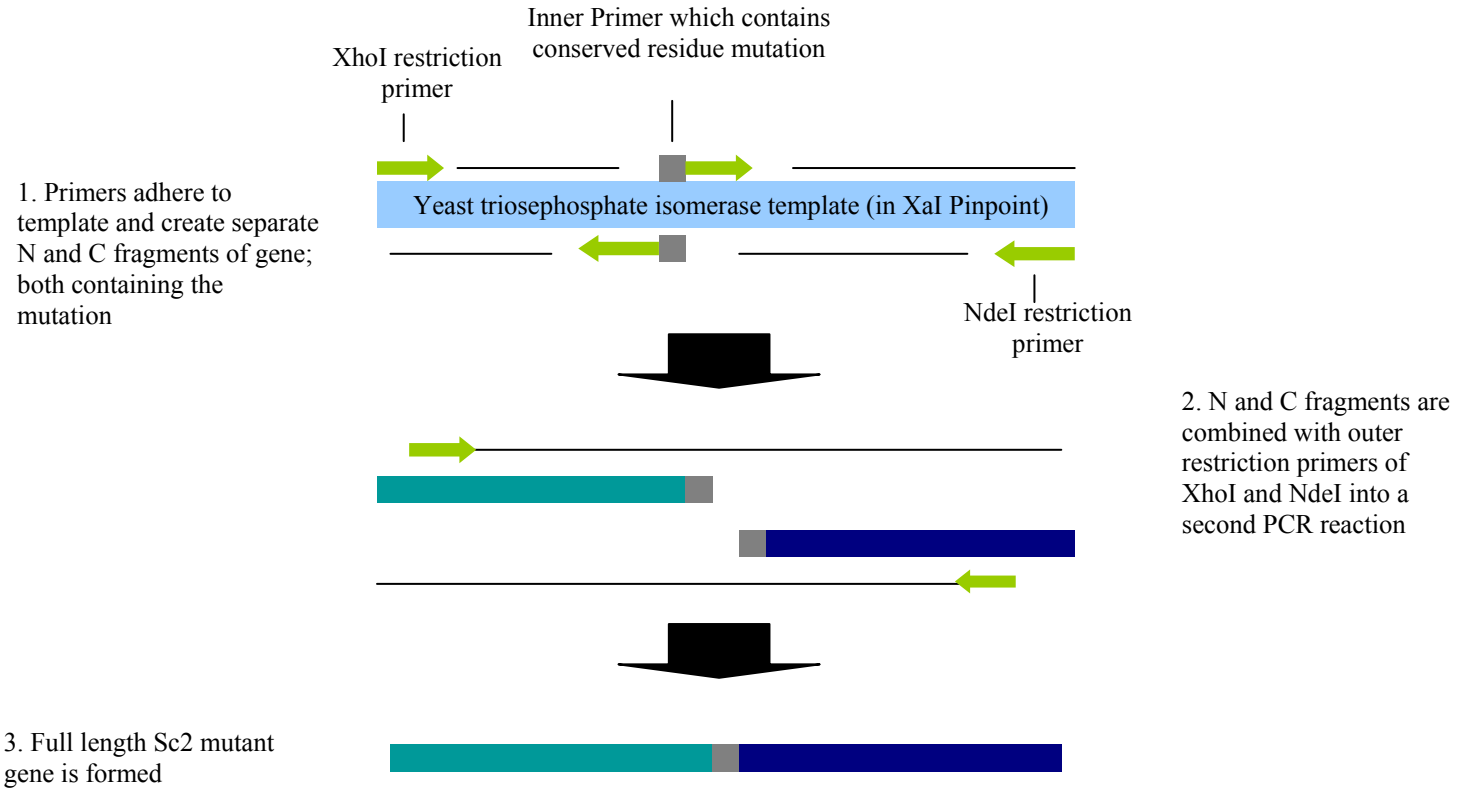
leucine to methionine, also involves a negligible change in size and polarity. It is difficult to predict which of those mutations, if it is either of the two, is destabilizing. Another likely reason causing the instability is due to the mutations affecting correlated residues. Mutating one can significantly alter its association with another. To determine which mutations are causing the instability, six individual mutants of Sc2TIM will be created, each containing one highly conserved residue. This project will be discussed in the next chapter.

Chapter 2: Construction of Sc2 Mutants

I. Material and Methods

Sc2TIM will be constructing using overlap PCR

Schematic diagram of synthesis:



First Overlap Reagents

For N Terminal, yeast TIM gene in XaI Pinpoint (1:1000 dilution), 1/100 dilution of dNTPs, 10X Thermopol buffer, 1 mM 3' end Mutant, 1 mM XhoI 5' Restriction Primer, 1/100 of total volume used for Taq polymerase and ddH2O are added into a PCR tube.

For C Terminal Fragment, yeast TIM gene in XaI Pinpoint (1:1000 dilution), 1/100 dilution of dNTPs, 10X Thermopol buffer, 1 mM 5' end Mutant,1 mM NdeI 3' Restriction Primer, 1/100 of total volume used for Taq polymerase and ddH2O are added into PCR tube.
PCR conditions for the reaction are set for 20 cycles at 94 ºC for 2 minutes, 94 ºC for 30 seconds, 61 ºC for 45 seconds, 72ºC for 1 minute and 72 ºC for 4 minutes.

2nd Overlap Reagents

1 mM N Terminal Fragment, 1 mM C Terminal Fragment, 1/100 dilution of 25 mM dNTPs, 10X Thermopol Buffer, DeepVent polymerase and dH2O are then combined into another PCR tube. After 8 PCR cycles of the 25, 1 uL NdeI restriction primer and 1 uL of XhoI primer are added into the PCR reaction and the cycle is continued.

II. Results and Discussion

The first overlap PCR step is successful in showing the correct amount of bands:

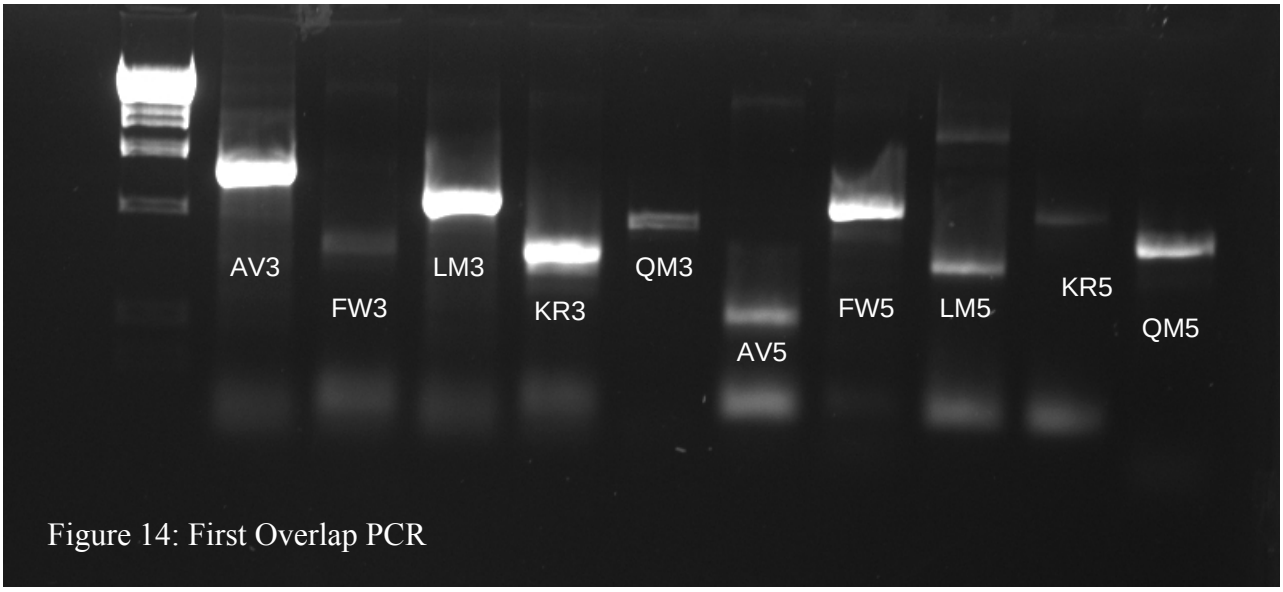
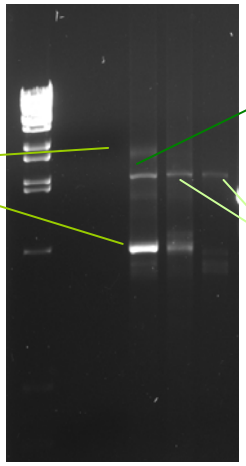


Figure 14: First Overlap PCR

The second PCR cycle is the more challenging step. For months there were several problems in optimizing for the correct conditions. Issues such as multiple bands showing up or no bands at all were the main issues. Background problems, with template appearing during the second round of PCR, were seen. This is caused with the restriction primers adhere to the plasmid and form product. Methods to rid the template were to use a dilute amount of plasmid during the first PCR cycle and to perform a DpnI digest after each PCR reaction to rid the template present. After this was corrected, the mutants were ligated and transformed into DH10B cells. Several mutations were shown in the sequencing results, likely meaning that the low fidelity polymerase, Taq, giving problems. When switching to a higher fidelity polymerase, Phusion, fewer mutations appeared, but unfortunately none of the samples sequenced were perfect.



First Lane: Lambda
Fourth Lane: WY mutant
Fifth Lane: NTerm Negative
Sixth Lane: CTerm Negative

The next method to solve this problem was to clone right into the expression vector, pHLIC, rather than clone into the screening vector, XaI Pinpoint. The reason why this method was preferred over the other was because XaI Pinpoint has a tac promoter located on its plasmid. This promoter allows the cell to constantly produce protein with no necessity of an inducer to control it (in a strain lacking large amounts of LacI). This allows the DH10B cells to be capable of producing proteins toxic to the cell and cause unwanted mutations in the gene. Cloning into the expression vector, pHLIC, differs in that the vector contains a T7 promoter which only produces protein in the presence of an inducer (such as lactose or IPTG). This allows the desired gene not to be transcribed and translated in the cell when cloning, and prevents interferences that cause unwanted mutations.

Once this gene is successfully formed, it can be protein expressed and further characterization can be done. The individual mutants will inform which conserved mutations are destabilizing, as well as which ones are stabilizing. Once this is determined, a new mutant can be formed containing only stabilizing mutations mapped onto the wild type. This will help answer the question how consensus residues affect yeast triosephosphate isomerase.

References

1. Steipe, Boris. Evolutionary Approaches to Protein Engineering.

2. Lehmann, Martin. (2000) The consensus concept for thermostability engineering of proteins. *Biochimica et Biophysica Acta.* 408-415.

3. Stemmer, Willem P.C. (1995). Single-Step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene.* 49-53.

4. Plaut, Barbara & Knowles, J.R. (1972). pH-Dependence of the Triose Phosphate Isomerase Reaction. *Journal of Biochemistry.* 311-320.

5. Lehmann, Martin. (2002). The consensus concept for thermostability engineering of proteins: further proof of concept. *Protein Engineering.* 403-411.

6. Lehmann, Martin & Wyss, Markus. (2001) Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution. *Science.* 371-375.

7. Kohl, Andreas & Binz, H. Kaspar. (2002). Designed to be stable: Crystal structure of a consensus ankyrin protein. *PNAS.* 1700-1705.

8. Magliery, Thomas J. & Regan, Lynne. (2004). Beyond Consensus: Statistical Free Energies Reveal Hidden Interactions in the Design of a TPR Motif. *Journal of Molecular Biology.* 731-745.

9. Nagono, Nozomi. (2002). One Fold with Many Functions: The Evolutionary Relationships between TIM Barrel Families Based on their Sequences, Structures and Functions. *Journal of Molecular Biology.* 741-765.

10. Main, Ewan RG. (2004). The folding and design of repeat proteins: reaching a consensus. *Current Opinion on Structural Biology.* 482-489.

11. Mosavi, Leila K. (2004) The ankyrin repeat as molecular architecture for protein recognition. *Protein Science.* 1435-1448.

12. Lehmann, Martin. (2000). From DNA sequence to improved functionality: using protein sequence comparisons to rapidly design a thermostable consensus phytase. *Protein Engineering.* 49-57.

13. Binz, H. Kaspar. (2003). Designing Repeat Proteins: Well-Expressed, Soluble and Stable Proteins from Combinatorial Libraries of Consensus Ankyrin Repeat Proteins. *Journal of Molecular Biology.* 489-503.

14. Steipe, Boris. (1994). Sequence Statistics Reliably Predict Stabilizing Mutations in a Protein Domain. *Journal of Molecular Biology.* 188-192.

15. Lolis, Elias & Alber, Tom. (1989). Structure of Yeast Triosephosphate Isomerase at 1.9-Å Resolution.