THE THERMOSTABILITY OF THE CHEY PROTEIN IN THE BACTERIUM *THERMOTOGA MARITIMA*

by
TRACY HARDWICK

A THESIS

Presented to the Department of Chemistry and the
Honors College of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Bachelor of Arts

May 1999

APPROVED [redacted]

Frederick Dahlquist

An Abstract for the Thesis of Tracy Hardwick

for the degree of Bachelor of Arts in the Department of Chemistry to be taken May 1999.

Title: THE THERMOSTABILITY OF THE CHEY PROTEIN IN THE BACTERIUM *THERMOTOGA MARITIMA*

Approved: 

Dr. Frederick Dahlquist

The ability of many bacteria to survive at very high temperature has long been an enigma to scientists. In particular, it is perplexing that a bacterium which lives at high temperatures can have a protein which is very similar to a protein which resides in a bacterium that resides in lower temperatures. In this thesis, I will explore the CheY bacteria of the bacterium *Thermotoga maritima* in relation to the less stable CheY protein of the bacterium *Bacillus subtilis*. Although these two proteins have nearly identical three dimensional structures and very similar sequences, there is a nearly 30°C difference in the melting temperatures of these two proteins. This thesis will explore the possibilities as to why there is such a difference in stability between these two proteins.

ACKNOWLEDGEMENTS

The author expresses sincere appreciation to Professors Dahlquist, Hawley and Cogan for their assistance in the preparation of this manuscript. In addition, many thanks are due to William Deutschman without whom this project could not have been completed.

INTRODUCTION

One of the most mysterious sets of bacteria on earth are the ones which flourish in hostile environments. For years, we have been fascinated by bacteria which can live at very high temperatures, at very low temperatures, at very high salt concentrations, at high or low pH or in extremely dry conditions. Recently I have been exploring one such bacteria, *Thermotoga maritima*. *T. maritima* lives in the thermal vents near underwater volcanoes off the coast of Italy. We have been fascinated by this bacteria's ability to live at very high temperatures, with optimal growth taking place at 80 °C and with the maximal growth temperature approximately 90 °C (Swanson, Sanna, and Simon, 1996). Specifically, we are interested in why the **proteins**[1] of T. maritima do not **denature** at these extremely high temperatures. Proteins from hyperthermophilic bacteria maintain thermal stability even when they are expressed in a non-thermophilic host, suggesting that the **thermostability** of the protein is due to factors within the molecule itself rather than a stabilizing effect that the cell might give to a molecule (Sanangelantoni, Cammarano, and Tiboni, 1996). A number of explanations have been offered in the past to explain why some proteins are much more stable than others. These explanations include extra salt bridges (Perutz and Raid, 1975), an excess amount of **prolines** (which shift the conformational equilibrium toward the folded state) (Watanabe et al, 1991), more effective packing (Britton et al, 1995), and extra disulfide bonds (Goldenburg et al, 1989). Unfortunately, while most of these explanations probably contribute somewhat to thermostability, all of these possibilities have fallen short of providing a complete picture as to why some proteins are more thermostable

[1] the definitions of all boldfaced words can be found in the glossary

than other proteins. Currently, there is no strong evidence that the structures of very stable proteins are systematically different from those of less stable proteins (Daniel, Dines and Petach, 1996). When one looks at the protein CheY, which is a signal transduction protein that controls flagellar rotation in bacteria, one notices few differences between the sequence for the protein in *T. maritima*, which denatures at about 95°C[2] , and the sequence for its close cousin *Bacillus subtilis*, which denatures at a much lower temperature (between 65-67° C; W. Deutschman, unpublished results)[3] . There are only 30 amino acid differences between CheY of *T. maritima* and *B. subtilis* (out of 120 total amino acids), and only 10 of these amino acids are **buried**[4] . This is important because the buried amino acids are the ones most likely to have significant effects on the thermostability; buried amino acids are more likely to interact in significant ways with surrounding amino acids, although we are looking at a few less buried amino acids as well. Although we do not have a structural analysis of BsY, based upon the structure of *E. coli* and other bacterial CheY molecules as compared to TmY, it is likely that the amino acid differences between BsY and TmY impart little or no significant structural differences between the two molecules[5] . There is only one salt bridge in each active site, there does not appear to be any significant packing difference between the two proteins (the surface area to volume ratio is essentially the same and the cavity volumes are also the same), and there are no disulfide bonds in either protein (Usher et al, 1998). There are two extra prolines in TmY (CheY for *T. maritima*) versus BsY (CheY for *B. subtilis*), and the significance of these two prolines is as yet unclear, although it was shown that a

---

2 For a visual image of the unfolding of CheY for *B. subtilis* and *T. maritima*, please see appendix A

3 See appendix B for the sequence of CheY for *E. coli, B. subtilis,* and *T. maritima.*

4 For a summary of this information, please see appendix C.

5 For a ribbon diagram of CheY for *E. coli* and *T. maritima* please see appendix D.

few proline substitutions in an amino acid sequence can induce moderate increases in the thermostability of a protein (although it was also shown that positioning a proline at an N-cap of an **alpha helix** (α helix) can have adverse effects on thermostability) (Bogin et al, 1998). Their role in providing thermostability to TmY will be tested as described below.

The question, therefore, remains, why CheY for *T. maritima* is stable at much higher temperatures than CheY for *B. subtilis*. The lack of significant differences suggests that thermal stability, at least in this case, is brought about by subtle structural changes. Our approach to this problem is to replace single amino acids in the *T. maritima* sequence with those from *B. subtilis*, while at the same time replacing individual amino acids in BsY with those from the TmY sequence. The stability of the mutant proteins will then be measured using **circular dichroism** by monitoring unfolding as a function of temperature or the addition of denaturants to the solution. This will tell us which amino acids provide the extra stability needed to survive at the higher temperatures. Of course, this approach will yield simple answers only if each amino acid substitution behaves independently of the others. Since some of the mutation sites in TmY and BsY are surrounded by amino acids which are the same in both proteins, it is likely that those particular sites behave independently of one another. At other sites, where several amino acids which are different in TmY and BsY may interact, it will be necessary to make **double mutants** to study the thermodynamic contribution of mutations on an individual and concerted basis. It is noteworthy that cooperative interactions have been observed in other studies of thermophilic versus mesophilic bacteria (Bogin et al, 1998).

METHODS[6]

Of the thirty amino acid changes between BsY and TmY, we are working with sixteen discreet amino acids. These sixteen were chosen primarily because they are the buried amino acids. The mutants we are working with are BsY I8V, I21L, I25L, F30Y, A34G, E37T, A40R, L53I, L69I, A79P, R82K, M85V, S93A, A113P, I120L, and T123V as well as their TmY counterparts (for example, the counterpart of BsY I21L would be TmY L21I). The mutants are labeled with the single letter abbreviation of the amino acid in the original polypeptide first, the position of the amino acid in the protein[7], and the single letter abbreviation of the amino acid we wish to change it to. For example, the mutant BsY which has an leucine replacing an isoleucine at position 21 in the polypeptide chain is BsY I21L.[8]

For this experiment, we used **PCR**[9] based site directed mutagenesis. The mutated DNA is placed in XL1-Blue cells and allowed to multiply. Once the cells have grown, the DNA is extracted and sent for sequencing. If our mutation is correct, mutated cells are grown and BsY protein is harvested using a nickel column (the nickel chelates the 6x His tag, preventing our protein from running through the column). The pH is lowered[10] , and our protein dissociates from the nickel column in a fairly pure form. This is then repeated with another nickel column. The TmY protein is collected by lysing the cell, heating the solution (to

---

6 for a list of the "order of events" please see appendix E, also for a more detailed methods section, please see appendix F.

7 the number associated with the amino acid was derived from sequence alignment between *E. coli* CheY protein and TmY.

8 A table containing the letters for the various amino acids has been provided in the back.

9 For a discussion on PCR, please see appendix G.

10 Imidizole is not used here because it absorbs ultraviolet light in the 260 nanometer range, which obscures the small tyrosine peak from our protein (our protein does not contain any tryptophan residues).

unfold any non-thermostable proteins), running over a column, dialyzing the protein into sodium acetate, loading this onto an sp-sephadex c-50 column, washing the column, and then eluting with a 0 to 1 M salt gradient. The proteins are sent to another laboratory to be melted. Circular dichroism is used to determine the temperature at which these proteins unfold. At this point, we can measure the free energy of folding for the mutant versus the free energy of folding for the non-mutated protein.

RESULTS

Thus far, we have confirmed mutants for TmY V8I, L21I, Y30F, G34A, I69L, P113A, L120I, V8I/L53I (double mutant) and all of the BsY mutations[11] . Of these, we have only examined the stability of TmY Y30F, which showed no change in melting temperature. When looking only at hydrophobic interactions at the core of the protein, it is reasonable to expect that the most important contributors to thermostability are large, non-polar molecules such as Trp, Ile, Tyr, Phe, Pro, Leu, Val, Lys and Met. Using the free energy of transfer data for hydrophobic groups from Tanford, we can estimate that replacing the isoleucine from BsY with the valine from TmY, for example, should destabilize the protein by 1.28 kilocalories; replacing an alanine from BsY with a proline from TmY should stabilize the protein by about 1.87 kcals; similar values can be calculated for each substitution (Tanford, 1962) . For TmY Y30F, we would expect from this change of 0.18 kcal change from the interactions of the hydrophobic core. This should cause a small change in the melting temperature of TmY, but since we know the crystal structure of TmY and that the side chain of residue 30 is

[11] For a summary of our results, please see appendix H

mostly surrounded by buffer (rather than by other parts of the protein), it is not surprising that the change does not produce a detectable difference in melting temperature. So, we are not astonished that there is no change in the **free energy** of unfolding in the Y30F mutant. Using the figures above, we expect the biggest contributors to the stabilization of TmY to be P113A, and V123T. The largest contributor to the destabilization of TmY should be V8I.[12]

Of course, there is more going on in this protein than simple hydrophobic effects. Two amino acid changes (T37E and R40A) cause changes in the local electrostatic environment of the protein. That is, position 37 has no charge in TmY, but it has a negative charge in BsY; position 40 has a positive charge in TmY, but no charge in BsY. By looking at these mutations, we will be able to tell what effect a change in charge in the core of the the protein has on the protein as a whole.

We will also be looking specifically at the two proline substitutions in this protein. Position 79 has a proline in TmY but an alanine in BsY; this is a β turn residue, a few amino acids away from the C-terminus of an α helix. Also, position 113 in TmY has a proline which caps the N-terminus of an α helix. Thus, by evaluating these mutations, it will be possible to directly test the stabilizing or destabilizing effect of prolines which cap the N-terminus of an α helix or are β turn residues. According to Bogin et al, P113A should destabilize the protein, while P79A should stabilize it. This specific mutation will allow us to test this theory as well as allowing us to see by exactly how much this mutation affects thermostability.

---

[12] This information can be found in appendix C.

DISCUSSION

While discussing this issue with some of my non-science classmates and professors, one question seemed to continually pervade the conversation. What most people seem to be interested in is what relevance this research has for the rest of the world. So what if this bacteria can live at super high temperatures? How does this affect us?

One answer to this is to satisfy a purely human (and scientific) curiosity. Humans are innately curious about things which are unlike ourselves, and we want to understand what makes them work the way they do. Beyond this simple rationale, however, there are other explanations as to why it is important to understand these bacteria.

First of all, *Thermatoga maritima* lives in an ecosystem entirely alien to our own. Rather than using light as the source of all energy, it uses the chemicals from the thermal vent. We know next to nothing about this ecosystem. By understanding its components and the ecological niche they fill, it may be possible to better understand the system itself.

Secondly, the process of protein folding is poorly understood in the scientific community. If we understand the structural basis for thermostability, we come closer to understanding protein folding in general. When we understand protein folding in this scenario, we may be able to predict why other proteins are able to fold and maintain their folded structure, and how stable they will be once they are folded. Thus, this research could be used to enhance biochemical knowledge of protein structure, stability, and folding.

This knowledge of protein stability could prove useful in many ways including making it possible to make more stable vaccines and medications.

Currently, vaccines are so unstable that they must be stored in a refrigerator at all times and even then they must be periodically discarded because of denaturation of the proteins. If those proteins could be made more stable without diminishing the effectiveness of the vaccine, this would save the health care industry money; it would make the vaccine more widely available (for example, in places without electricity for refrigeration), and it would make life easier for health care professionals.

Finally, it is possible that once we understand what makes proteins more thermostable, it may be possible to extend that knowledge to industrial purposes. In other words, there are some industries which have a need for thermostable enzymes; our research might be able to show them how to engineer enzymes to make them more stable at high temperatures.

Perhaps a more interesting question than the one listed above might have to do with why there are so few bacteria with the ability to withstand high temperatures. Surely it would not hurt *E. coli*, for example, to be able to withstand higher temperatures, especially given that when it infects a body, the body automatically raises its temperature so as to destroy the invader. In fact, logical as it may sound that a bacterium would always benefit from having more thermostable proteins, this is not the case.

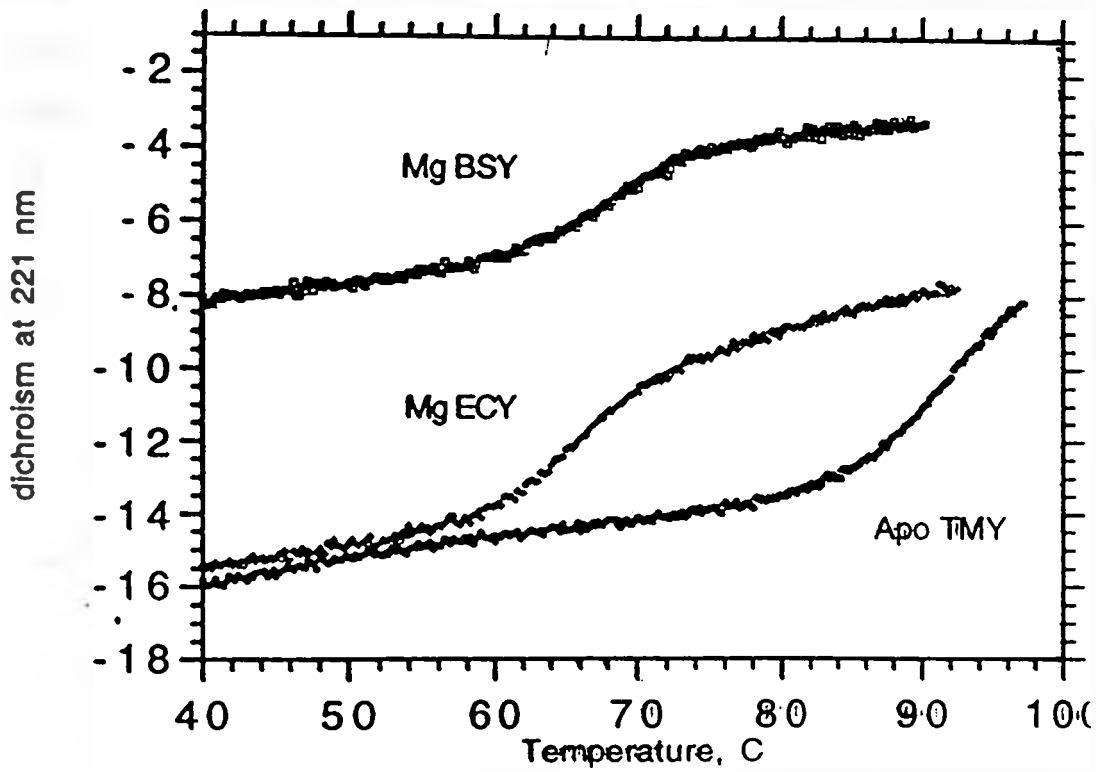
It was shown in 1996 that at a given temperature, thermostable enzymes are less active than enzymes designed to live in cooler environments (Daniel, Dines, and Petach, 1996). However, while in the environment each protein was designed for, the activity of enzymes from mesophiles and extreme thermophiles is similar (Daniel, Dines, and Petach, 1996). So, if a thermophilic enzyme was contained in a non-thermophilic bacterium, that enzyme would be less able to perform its normal cellular functions. It is not yet understood why this is true.

Finally, the study of these bacteria may give us a clue as to what life was like when the earth was young. It is likely that when life began it was mainly adapted to high temperature environments. How did it evolve to live at lower temperatures? Were there reasons other than the one discussed above why this was advantageous? Perhaps the study of these ultrathermostable proteins may give us a clue as to how life began and evolved.

One possible scenario is that over the course of evolution, as the earth cooled, fewer and fewer bacteria found it necessary or beneficial to keep their hyperthermophilic proteins. So, these ultra-stable proteins were gradually replaced by proteins which were folded and functional at more reasonable temperatures. While it may still be nice for some bacteria to be resistant to heat in certain instances (such as when the body raises its temperature to fight off infection), the vast majority of the time it is more beneficial to have enzymes which are not thermophiles and so that the cell can work more easily with its proteins. So, ultrathermostable enzymes are now sequestered off to isolated regions of the earth and it is only in those places that we find these hardy bacteria from a billion years ago.

APPENDICES

APPENDIX A


dichroism at 221 nm
-2
-4
-6
-8
-10
-12
-14
-16
-18
Mg BSY
Mg ECY
Apo TIMY
40
50
60
70
80
90
100
Temperature, C

APPENDIX B

```
ECY:    1 MADKELKFLVVDDFSTMRRIVRNLLKELGFNNVEEAEDGVDALNKLQAGG
          .: :.|:||| . ||.:...:: .:|:. | :|:.| :|:.|..
TMY:    1 ...MGKRVLIVDDAAFMRMMLKDIITKAGYEVAGEATNGREAVEKYKEFK
            |: |:|||||||||||||:|||| | |:|||:||:|| .|||||||  |
BSY:    1 ...MAHRILIVDDAAFMRMMIKDILVKNGFEVVAEAENGAQAVEKYKEHK

II              sssss   hhhhhhhhhhhhh   ssssss  hhhhhhhhhhh


ECY:   51 YGFVISDWNMPNMDGLELLKTIRADGAMSALPVLMVTAEAKKENIIAAAQ
            | .|::||:|:|::.:|.|    . .:.:::. :| . .  :| |.
TMY:   48 PDIVTMDITMPEMNGIDAIKEIMKIDP..NAKIIVCSAMGQQAMVIEAIK
          ||:|||||||||||-||.|:|||  ||   :|:||.|||||||  |||:||
BSY:   48 PDLVTMDITMPEMDGITALKEIKQIDA..QARIIMCSAMGQQSMVIDAIQ

II          sssss        hhhhhhhhhhh       sssss    hhhhhhhhh


ECY: 101 AGASGYVVKPFTAATLEEKLNKIFEKLGM  129
           |||.::|||||. . : | |||:
TMY:  96 AGAKDFIVKPFQPSRVVEALNKVSK....  120
           |||||||||||||:.|||||:||
BSY:  96 AGAKDFIVKPFQADRVLEAINKTLN....  120

II         h   ssss    hhhhhhhhhhh
```

~ 90% SIMILAR (13/120 NON-SIMILAR)

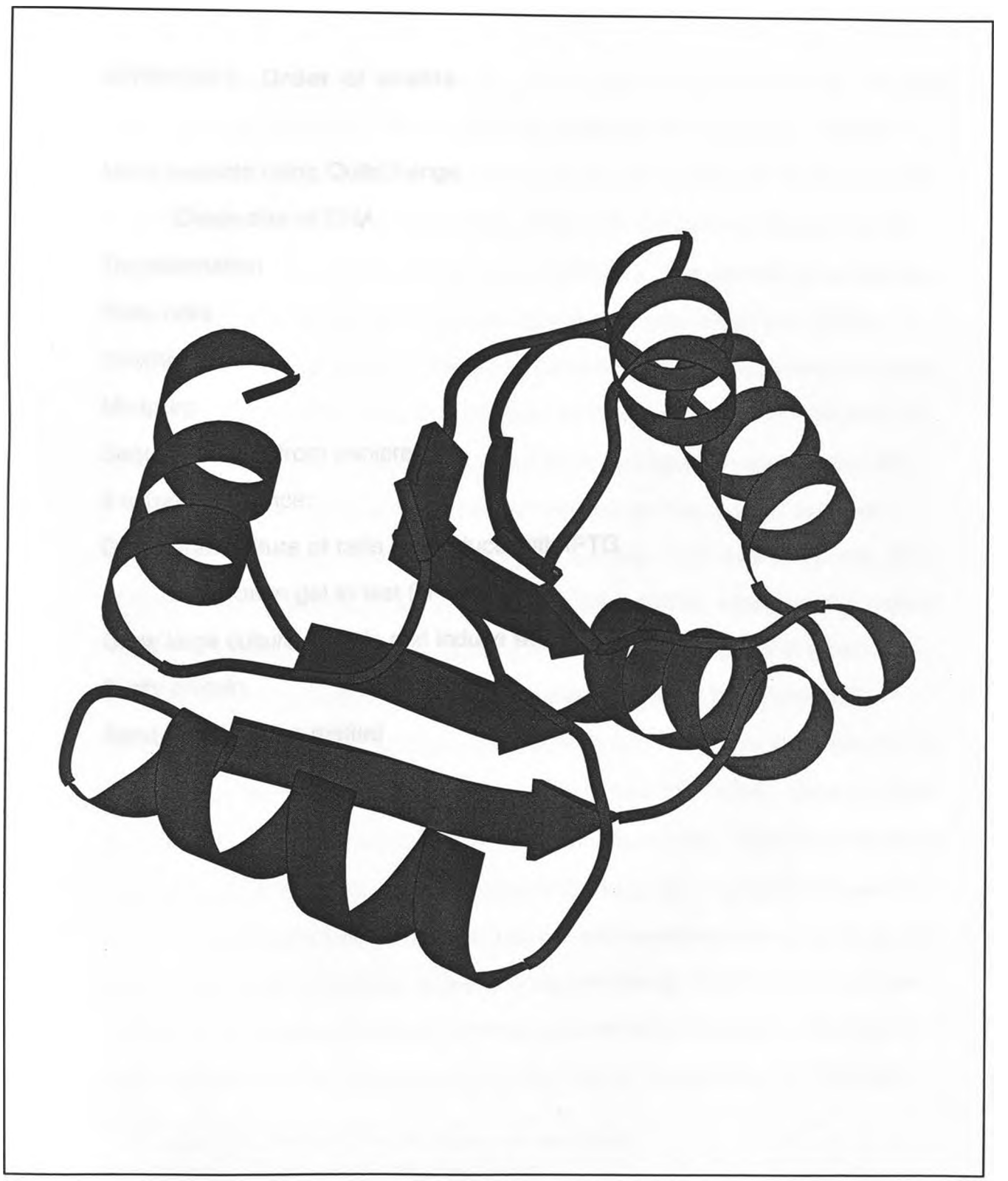
~ 75% IDENTICAL (31/120 NON-IDENTICAL)

Appendix C

TmY mutati	surface accessability	ΔΔG(transfer)Kcal/r	ionic effect
V8I	0	-1.25	-
L21I	5	-0.55	-
I25L	0	0.55	-
Y30F	9	0.2	-
G34A	14	-0.75	-
T37E	54		+
R40A	73	0	+
I53L	2	0.55	-
I69L	0	0.55	-
P79A	87	1.85	-
K82R	67	0.75	-
V85M	0	0.4	-
A93S	72		-
P113A	77	1.85	-
L120I	2	-0.55	-
V123T	18	1.25	+
		sum ΔΔG = 4.85	
		sum ΔΔG (buried) = 1.15	
		* assuming complete	
		burial	

APPENDIX D

EcY



TmY

APPENDIX E. **Order of events**.

Make mutants using QuikChange

Check size of DNA

Transformation

Plate cells

Overnight

Miniprep

Sequence DNA from miniprep

If correct sequence:

Grow small culture of cells and induce with IPTG

run protein gel to test for expression

Grow large culture of cells and induce with IPTG again

Purify protein

Send protein to be melted

APPENDIX F **Methods explanation**

The first step in the procedure is, of course, making the mutants. For this step, the QuikChange™ Site-Directed Mutagenesis Kit was used. This kit allows a specific mutant to be made at precisely the location at which we need it. This kit works by first **annealing** a primer with the desired mutation to the site which must be mutated (this is done by PCR[1]). Our primers were typically 30-35 base pairs long with at least 12-15 base pairs on either side of the mutation (so that it anneals to the correct spot and makes the desired mutation). The primers terminate with a **GC** pair and an ideal primer has at least 40% GC content[2] (because the GC pairs contain three hydrogen bonds whereas **AT** pairs only contain two hydrogen bonds. This means that the GC pairs will anneal more tightly to the DNA). Using the existing **DNA** as a **template**, **DNA polymerase** makes an exact replica of the original DNA except for the point at which the mutation was made. Since the parental DNA existed in a cell originally, it will be **methylated**. At this point, we digest the methylated (parental) DNA so that only the mutated DNA is left. Once this is completed, the DNA ligated into a vector; our choice of vectors was PET-22b[3] . Once the DNA is ligated, it is transformed into XL1-Blue[4] competent cells, which then repair the nick in the DNA by using their own copy of DNA ligase. The transformation is done by electroporation. The basic idea behind electroporation is to stress the cells (in this case by running an electric current through them), which causes them to take in things from their external environment. By placing the DNA in their environment, we are encouraging the cells to incorporate the DNA into

---

1 for a description of how PCR works, please see appendix G

2 statistics from QuikChange™ instruction manual

3 for a discussion on vectors, please see appendix I

4 For a note on the choice of competent cells, please see appendix J

themselves.

We then plate the transformed XL1-Blue cells onto an LBH/ampicillin plate[5], pick off the colonies the next day, and grow them up over night. The next day we miniprep the DNA.

For the minipreps, we use Promega's Wizard™ Plus Miniprep DNA Purification Systems kit. After spinning down the cells grown the night before, the cells are resuspended in resuspension solution (**Tris**, **EDTA**[6] , and **RNase A**). We then add lysis solution to break open the cells and expose the DNA (the lysis solution is **sodium hydroxide** and **SDS**). After **lysing** the cells, neutralization solution is added (potassium acetate). This creates a white precipitate consisting of **genomic DNA** and SDS, which we then **centrifuge** to create a pellet of precipitate. The now clear supernatent is placed in a syringe with glass beads and the slurry is pushed through a minicolumn. The DNA sticks to the glass rather than being pushed out with the remainder of the liquid. We then wash the DNA with Column Wash Solution (potassium acetate, Tris-HCl, EDTA, and ethanol). This rids the glass beads of any impurities, such as some proteins, which might have stuck to them.[7] The minicolumn is then centrifuged to rid it of excess liquid which might be left. At this point we add water to the minicolumn, which allows the DNA to dissociate from the glass and return into solution. The minicolumn is spun one last time, and the liquid which we get from it is water containing our desired DNA fragment.[8]

At this point the DNA is sent to a different lab for sequencing. If we do indeed have the correct mutation inserted into our DNA, we grow some mutated

---

[5] For a visual representation of this, please see appendix K (provided by Stratagene).

[6] EDTA chelates metals to keep them from interfering with the action of the DNAse.

[7] Proteins might transiently stick to the glass at high salt concentrations. The idea of the washes is to lower the salt concentration and make these proteins return to the solution. This effectively rids the glass beads of the impurities without ridding it of the plasmid DNA.

[8] For a visual representation of this procedure, please see appendix L

cells and test to see if whether the engineered DNA will produce protein (to do this we **induce** with **IPTG** and take several samples at different times after induction). If expression is good, more cells are grown and induced and protein is harvested from the solution. Protein is harvested from solution by first putting them in a solution consisting of guanadinium-HCl, sodium phosphate, and Tris/HCl, which denatures proteins and breaks apart the cell walls. We then spin this suspension, and discard a pellet which consists mostly of cell walls and a few non-soluble proteins. Next, we pour the solution into a nickel column. Since our protein has a 6x histidine tag at the end, it sticks firmly to this column because the His tag **chelates** the Ni^{2+} ion. In a series of steps, we lower the pH by increments. This allows other proteins to slough off the column first, but since ours has the 6x His tag, it sticks well to the column. Finally, we drop the pH enough to make even our tagged protein come off the column, and at that point we have essentially pure mutated BsY protein. To purify the protein further, we run this over a Ni^{2+} again, this time using a different buffer.

For TmY, the procedure is the same through the IPTG step, but after we induce, grow and spin down the cells, we resuspend them in Tris and magnesium chloride and French press them. The French press essentially physically breaks open the cell, exposing the protein (among other things) inside. At this point we spin the solution and heat the supernatent at 80 °C. This serves to denature most of the proteins except ours (since ours is heat stable), thus making them come out of solution. We then spin this solution once again and load the supernatent onto a DE-52 column containing tris and magnesium chloride. Next, we wash the column with tris and collect this wash. Next, the wash is **dialyzed** into sodium acetate at low pH. The protein is then loaded onto an sp-sephadex c-50 column and washed with sodium acetate loading buffer. Finally, TmY is **eluted** in a 0 to 1 **molar** salt gradient.

This protein is then sent to another laboratory to be unfolded and the free energy of folding for the mutant can then be compared to the free energy of folding for the wild type protein. The measurement of the unfolding is done by a method called circular dichroism. Circular dichroism is a spectroscopic technique which is used when an optically active chemical absorbs right and left hand circular **polarized** light differently. The protein is placed in a device which measures the absorbance of light, which is a function of the absorbtion of the polarized light (right and left hand circular polarized light cancel each other out, so if there are equal amounts of each, then the emergent light is said to be depolarized, whereas is more of one or the other is absorbed, the emergent light is polarized). In this case, the absorbance is determined by the helical residues in the protein. As temperature increases, the protein becomes less and less stable until finally the protein denatures and the alpha helices (along with every other part of the protein) unfold. As the alpha helices unfold, absorbance drops, and thus we can determine at what point the temperature has gotten too high for the protein to maintain its folded structure.

APPENDIX G **PCR**

The polymerase chain reaction (PCR) offers a fast and convenient method of amplifying a specific segment of DNA. In this procedure, a denatured DNA sample is incubated with DNA polymerase, **dNTPs**, and two **oligonucleotide** primers whose sequences flank the DNA segment of interest so that the DNA polymerase will bind the DNA and synthesize new complementary strands. When this process is repeated, many new strands are made of the DNA segment. In each sequence in the PCR reaction, the two strands of DNA are thermally separated, the primers anneal to their complementary segments of the DNA, and the polymerase polymerizes the dNTPs to form complementary strands with the two separated strands. Thus, each cycle of the reaction approximately doubles the amount of DNA preset.

APPENDIX H

TmY	Oligos ordered	Mutant Confirmed	His-tag	Cells Grown	Protein Purified	Archived	Archived as His tag	CD Done	Delta Melt
V81	X	X				X			
L21I	X	X		X		X			
Y30F	X	X		X	X	X		X	0
G34A	X	X				X			
T37E	X								
R40A	X								
I53L	X								
I69L	X	X		X	X	X		X	-5.2
P79A									
K82R	X								
V85M									
A93S	X								
P113A	X	X							
L120I	X	X		X	X	X		X	-6.5
V123T	X								
V81/I53L		X		X		X			
BsY									
I8V	X	X	X	X	X		X		
I21L	X	X	X	X	X	X	X		
F30Y	X								
A34G	X	X	X	X	X	X	X		
E37T	X	X	X	X		X	X		
A40R	X	X	X	X	X	X	X		
L53I	X								
L69I	X	X	X	X	X	X	X		
A79P	X								
R82K	X								
M85V	X	X	X	X		X	X		
S93A	X	X	X			X	X		
A113P	X								
I120L	X								
T123V	X								
I120F		X				X			
A81P		X	X			X	X		
A81P/R110C			X				X		

APPENDIX I VECTORS

The vector we decided to use to express BsY was PET-22b. This vector was chosen primarily because it it has a T7 promoter, which means that it expresses well. This vector was also chosen because it coded for six histidines at the spot just after where we inserted our plasmid. This enabled us to use a nickel column to which histidine would stick in order to purify our protein. This means that PET-22b makes purification of proteins much simpler than if we had decided to use another vector.

APPENDIX J COMPETENT CELLS

Throughout this experiment, we use two main strains of competent cells: XL1-Blue cells, and BL-21 cells. These two cells are both strains of *E. coli* cells. BL-21 was chosen because it has λDE3 activity, which is necessary to express proteins under a T7 promoter (which is what we use for BsY). XL1-Blue was chosen because it is deficient in exonuclease activity, so it will not chew up the nicked DNA that we transform into it. It also expresses proteins under lactose expression control (which is necessary for TmY).

APPENDIX k

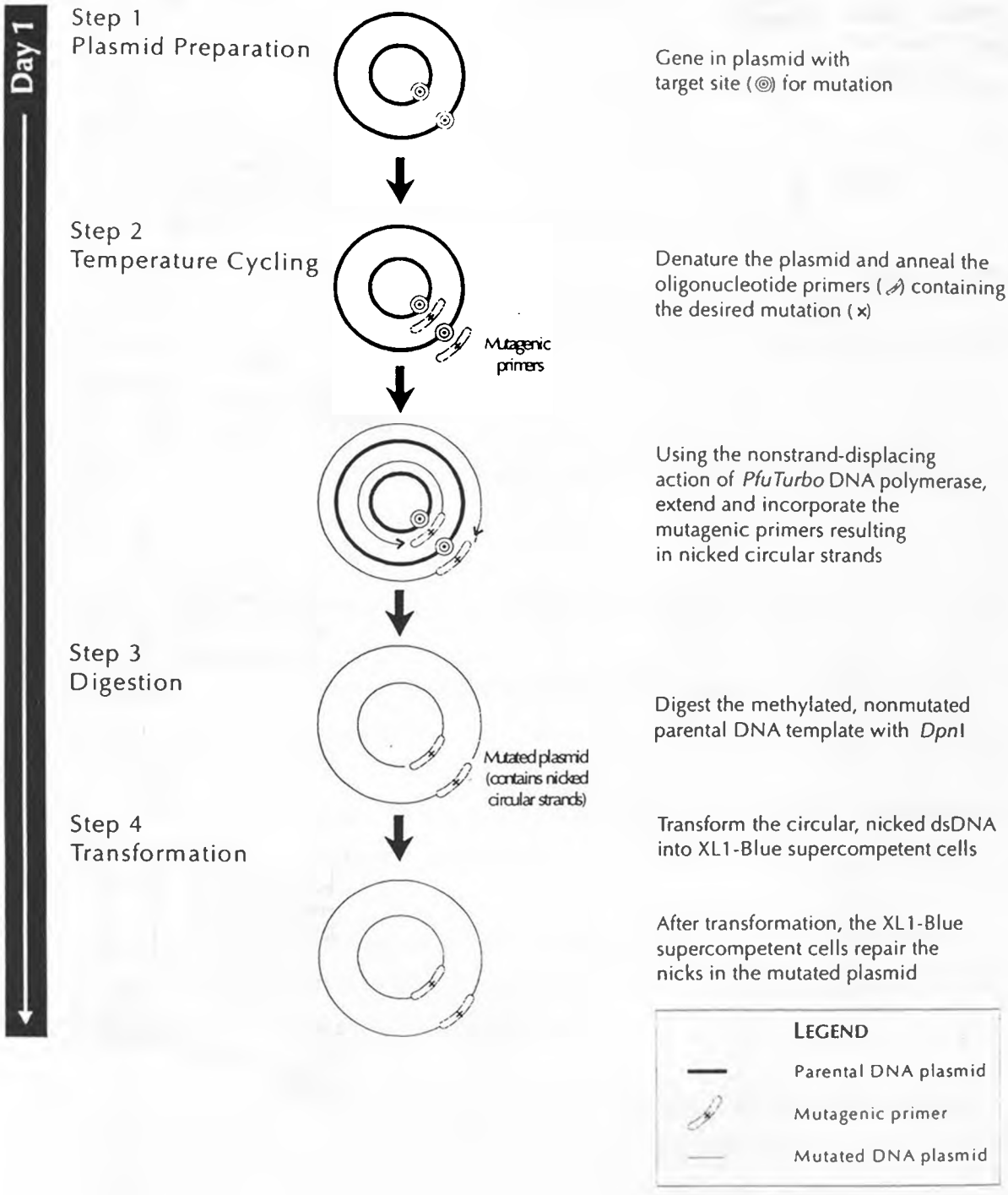
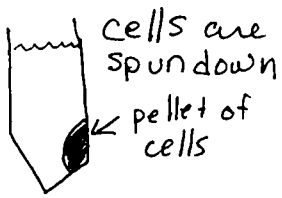
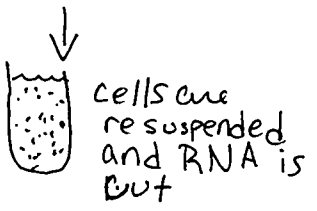


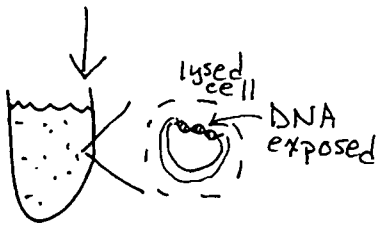
FIGURE 1 Overview of the QuikChange™ site-directed mutagenesis method.

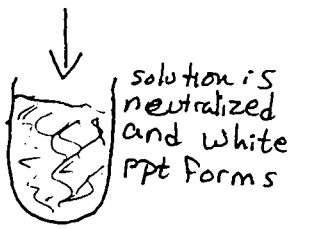
APPENDIX L

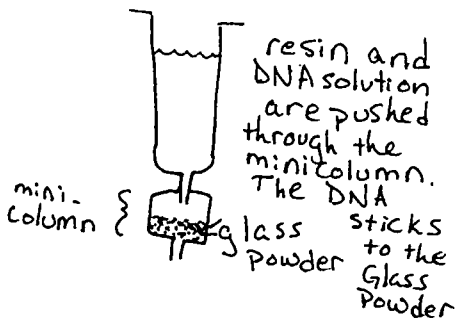
Minipreps

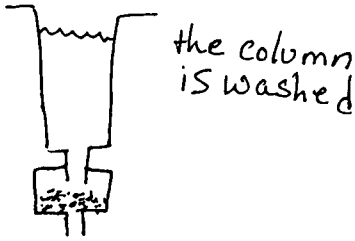




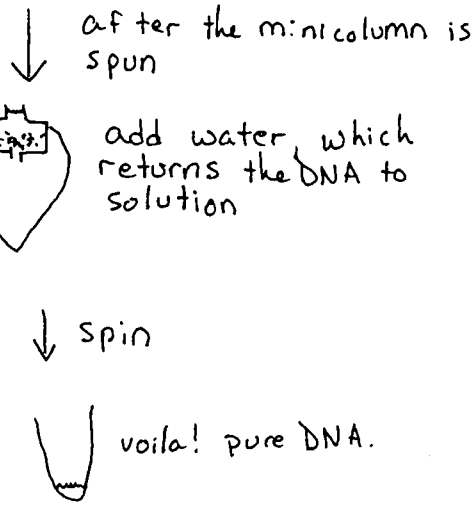








(cont. at top.)



Glossary

A - adenine. This is one of the four bases that makes up DNA. Adenine and thymine join by two hydrogen bonds.

alpha helix - an energy efficient structure present in proteins.

anneal - stick together. When we anneal two DNA strands, we are "gluing" them together, with a cytosine sticking to a guanine and an adenine sticking to a thymine.

BL-21 cells - an *E. coli* strain of cells which has λDE3 activity, which is necessary to express BsY.

buffer - this is a system which is resistant to changes in pH.

buried - A buried amino acid is one which is not accessable to the solvent; it is on the inside of the protein.

C - cytosine. This is one of the four bases that makes up DNA. Cytosine and guanine join by three hydrogen bonds.

catalyst - a molecule (in our case, usually an enzyme) which helps the reaction proceed. It is neither a starting reagent nor an end product, but it is something which makes the reaction "go".

centrifuge - the idea behind a centrifuge is to spin a tube so fast that the heavier matter (usually solid particles) are compacted at the bottom of the tube while the lighter matter is left at the top of the tube. see "spin down"

chelate - this is when a protein or other large molecule binds a metal ion to bring it either into or out of solution.

chemotaxis - a cell's sensing of and movement toward, or away from, a specific chemical agent.

circular dichroism - this measures the α helix content in a protein. As the protein denatures, the α helix content drops.

cuvette - a container which holds a solution and allows a laser to pass through

denature - unfold. When one denatures a protein one renders it inactive and also insoluble.

dialysis - Removal of small molecules from a solution of a macromolecule (in our case, a protein), by allowing them to diffuse through a semi-permeable membrane

DNA - deoxyribonucleic acid. This is the foundation of life. DNA tells the cell what proteins to produce.

DNA polymerase - This enzyme utilizes single-stranded DNA as a template on which to catalyze the synthesis of the complementary strand of DNA.

dNTPs - ATP, GTP, CTP, TTP. These are the four bases which make up any given stretch of DNA.

double mutant - this is where change more than one amino acid. For example, we might make a mutant of BsY I21L / F30Y with both these mutations in the same DNA strand.

EDTA - ethylenediamminetetraacetate. This molecule binds to metal ions to bring them out of solution.

elution - a way of washing a column in which those proteins with low affinities for the ion exchanger move through to column faster than the proteins that bind to the ion exchanger

free energy - the component of the total energy of a system that can do work at a constant temperature and pressure.

G - guanine. This is one of the four bases that makes up DNA. This base is complementary to cytosine.

genomic DNA - The DNA intrinsic to the cell.

induce - make the cell produce lots of proteins.

IPTG - Isopropylthiogalactoside. This induces DNA to produce protein.

lysis - the cell bursts, spilling the DNA into the supernatent.

methylate - put a methyl on each adenine. This is what allows the cell to tell what is its own DNA and what is an intruder. It is also what allows us to get rid of the parental DNA without harming our mutated DNA.

molar - moles per liter (a unit of concentration)

mole - 6×10^{23} molecules. This is a convenient way of noting the number of molecules of a particular substance present.

oligonucleotide - a short nucleic acid, usually containing 50 or fewer nucleotides.

peptide bond - a link between the amino group of one amino acid and the carboxyl group of another amino acid. A water molecule is lost when this link is made.

pH - a measure of the acidity or bacisity of a solution. The lower the pH, the more acid the solution. The higher the pH, the more basic the solution.

plasmid - a DNA molecule distinct from the bacterial chromosome replicated by the cell. Plasmids contain the requisite genetic machinery to permit their autonomous propagation in a bacterial host.

polarized light - light that oscillates only in one plane. A molecule which is easily polarized (by either left or right hand polarized light) has an electron cloud around one or more atoms which is easily distorted.

polypeptide - a series of amino acids joined together in a polymer by peptide bonds.

proline - an amino acid which shifts the conformational equilibrium toward the folded state (in effect it destabilizes the unfolded state by allowing fewer conformations of the protein).

protein - a series of polypeptides joined together in a macromolecule.

proteolysis - the digestion or degradation of proteins.

RNase - ribonuclease. This is a nuclease that catalyzes the hydrolysis of certain internucleotide linkages of RNA. RNase A cleaves RNA after pyrimidine residues (after a thymine or a cytosine).

S D S - sodium dodecyl sulfate. This is a detergent commonly used in biochemical preparations because it binds strongly to most types of proteins. Detergents are protein denaturants.

sodium hydroxide - (NaOH) This is a strong base, frequently used as an agent to lyse cells.

spin down- centrifuge. This is exactly what it sounds like - we put the solution into a machine and it pellets whatever solid particles may be in it.

supernatent - the fluid which bathes the cells.

T - thymine. One of the four bases that makes up DNA. Thymine and adenine join with two hydrogen bonds.

template - the strand of DNA which is used as a "mold" upon which another strand of DNA can be built.

thermostable - the ability of a particular protein to undergo high temperatures while not denaturing. The more thermostable a particular protein is, the higher the temperature it can withstand.

Tris - a common nitrogen containing buffer.

vector - an autonomously replicating DNA molecule into which the plasmid is inserted.

XL1-Blue cells - an *E. coli* strain of cells which is deficient in exonuclease activity, so it will not chew up nicked DNA.

BIBLIOGRAPHY

Alber, T. (1989) Mutational effects on protein stability. *Annual Review of Biochemistry.* Volume 58,765-798.

Bai, Y., Milne, J.S., Mayne, L. and Englander, S.W. (1994) Protein stability parameters measured by hydrogen exchange. *Proteins* volume 20, 4-14.

Barak, Rina and Eisenbach, Michael (1992) Correlation between Phosphorylation of the Chemotaxis Protein CheY and its Activity at the Flagellar Motor. *Biochemistry* volume31, 1821-1826.

Bogin, Oren, Peretz, Moshe, Hacham, Yael, Korkhin, Yakov, Frolow, Felix, Kalb(Gilboa), A. Joseph, and Burstein, Yigal. (1998) Enhanced thermal stability of *Clostridium beijerinckii* alcohol dehydrogenase after strategic substitution of amino acid residues with prolines from the homologous thermophilic *Thermoanaerobacter brockii* alcohol dehydrogenase. *Protein Science* volume 7, 1156-1163.

Bowie, J.U., Reidhaar-Olson, J.F., Lim, W.A. and Sauer, R.T. (1990) Deciphering the message in protein sequences: tolerance to amino acid substitutions. *Science* volume 247, 1306-1310.

Britton, K.L., Baker, P.J., Borges, K.N., Engel, P.C., Pasquo, A., Rice, D.W., Robb, F.T., Scandurra, R., Stillman, T,J, and Yip, K.S. (1995) Insights into thermal stability from a comparison of the glutamate dehydrogenases from *Pyrococcus furiosus* and *Thermococcus litoralis. European Journal of Biochemistry.* volume 229, 688-695.

Daniel, Roy M., Dines, Mark, Petach, Helen H. (1996) The denaturation and degradation of stable ensymes at high temperatures. *Journal of Biochemistry.* volume 317, 1-11.

Goldenberg, D.P., Frieden, R.W., Haack, J.A. and Morrison, T.B. (1989) Mutational analysis of a protein folding pathway. *Nature* volume 338, 127-132.

Kawamura, Shunsuke, Abe, Yoshito, Ueda, Tadashi, Masumoto, Kiyonari, Imoto, Taiji, Yamasaki, Nobuyyki, and Kimura, Makoto. (1998) Investigation of the Structural Basis for Thermostability of DNA-binding Progein HU from *Bacillus stearothermophilus. The Journal of Biological Chemistry.* volume 273, 19982-19987.

Klump, H.H., Adams, M.W.W. and Robb, F.T. (1994) Life in the pressure cooker: the thermal unfolding of proteins from hyperthermophiles. *Pure and Applied Chemistry* volume 66, 485-489.

Korolev, S., Nayal, M., Barnes, W.M., DiCera, E. and Waksman, G. (1995) Crystal structures of the large fragment of *Thermus aquaticus* DNA polymerase 1 at 2.5 Å resolution: structural basis for thermo-stability. *Proceedings of the National Acadamy of Science USA* volume 92, 9264-9268.

Perutz, M. and Raidt, H. (1975) Stereochemical basis of heat stability in bacterial ferrodoxins and in hemoglobin A2. *Nature* volume 255, 256-259.

Sanangelantoni, A.M., Cammarano, P. and Tiboni, O. (1996) Manipulation of the *tuf* gene provides clues to the localization of sequence element(s) involved in the thermal stability of *Thermotoga maritima* elongation factor Tu. *Microbiology* volume 142, 2525-2532.

Serrano, Luis, Day, Anthony G., and Fersht, Alan R. (1993) Step-wise Mutation of Barnase to Binase: A Procedure for Engineering Increased Stability of Proteins and an Experimental Analysis of the Evolution of Protein Stability. *Journal of Molecular Biology* volume 233, 305-312.

Swanson, R.V., Sanna, M.G. and Simon, M.I. (1996) Thermostable chemotaxis proteins from the hyperthermophilic bacterium *Thermotoga maritima*. *Journal of Bacteriology* volume 178, 484 - 489.

Tanford, Charles. (1962) Contribution of Hydrophobic Interactions to the Stability of the Globular Conformation of Proteins. *Journal of the American Chemical Society* volume 84, 4240-4247.

Usher, K.C., De La Cruz, A.F.A, Dahlquist, F.W., Swanson,R.V., Simon, M.I., and Remington, S.J. (1998) Crystal structures of Che Y from Thermotoga maritima do not support conventional explanations for the structrual basis of enhanced thermostability. *Protein Science* volume7, 403-411.

Volz, K. (1994) Structural conservation in the Che Y superfamily. *Biochemistry* volume 32, 11741-11753.

Watanabe, J., Chishiro, K,. Kitamura, K. and Suzuki, Y. (1991) Proline residues responsible for thermostability occur with high frequency in the loop regions of an extremely thermostable oligo-1,6-glucosidase from *Bacillus thermoglucosidasius* KP1006. *Journal of Biological Chemistry* volume 266, 24287-24294.

Zhu, X., Rebello, J., Matsumura, P., Volz, K. (1997) Crystal structures of Che Y Mutants Y106W and T87I/Y106W. *The Journal of Biological Chemistry* volume 272, 5000-5006.