**Horizontal Gene Transfer and Temperature Adaptation in Kosmotoga olearia**  
**University Scholar Proposal Fall 2011**

My proposed University Scholar Project will be in Dr. J. Peter Gogarten's lab in the Molecular and Cell Biology Department at the University of Connecticut. My project will be in the field of microbiology and evolution. Traditionally, microbiologists have focused on understanding the biology of a particular organism, especially those that are pathogenic. This approach is useful for certain applications, but limits our understanding of microbiology to the function of single organisms, without understanding their relationships or how they change. Microbial evolution proceeds more rapidly and by different mechanisms than the evolution of multicellular organisms, a phenomenon made apparent by the recent sequencing of over 4,000 prokaryotic genomes. I believe that the next generation of microbiologists must aim to understand the processes by which evolution occurs, under what conditions and with what timing, in order to further our understanding of microbiology. A greater understanding of the process of microbial evolution has the potential to aid in the search to find and characterize microorganisms, develop antibiotics, and engineer microorganisms for bioremediation.

I have been interested in pursuing a research career in biology since I entered college, and began working in Dr. Gogarten's lab during my freshman year. This has exposed me to current research and thinking in prokaryotic and molecular evolution, the concept of evolution at the level of single genes and proteins. I believe that an understanding of evolution, obtained by both traditional and computational methods, is the most interesting question in biology today.

My project will focus on the contribution of horizontal genetic transfer to the evolution of optimal growth temperature in a group of bacteria called the Thermotogales. The aim of this project is to show how adaptation to new environments is facilitated by gene flow from other microorganisms. The scope of this project extends beyond a typical honors thesis because I have taken personal initiative in the conception and design of this experiment. This project will continue to produce results suitable for publications in academic journals, and presentations at academic conferences.
Literature Review

Traditionally, evolution has been thought of as a gradual process that proceeds mainly by vertical inheritance, the transfer of genes from parents to offspring. However, recent research has shown that evolution is also shaped by horizontal genetic transfer (HGT) [1], the exchange of genetic material between organisms that are not ancestor and descendant. HGT allows the transfer of genetic material between different species and even different phyla of microorganisms, and facilitates many evolutionary innovations, such as the rapid spread of antibiotic resistance [2]. Scientists have sought to create a tree of life that depicts the evolutionary divergence and vertical descent of all species, but it is now known that a tree depicting only the vertical descent of all prokaryotes would be uninformative about the evolutionary history of the microbial world [3]. A complete picture of microbial evolution needs to include vertical and horizontal genetic transfer, as both are processes that drive evolution [3].

An order of bacteria called the Thermotogales provides an extreme example of the pervasiveness of horizontal gene transfer, because the majority of their genes have been acquired through horizontal gene transfer [4]. These bacteria have optimal growth temperatures between 37° C and 80° C and are found in marine sediments and oil fields. A previous study suggested that the ancestor to all the extant Thermotogales had an optimal growth temperature above 80° C, based on amino acid frequencies found in reconstructed ancestral proteins [4]. Some organisms in this clade therefore have growth temperatures very close to the ancestral one, whereas others have experienced a steep drop in optimal growth temperature during their divergence. It is likely that horizontal gene transfer played a role in this process, which would indicate that not only can HGT provide new metabolic characteristics, but can facilitate the process of whole-cell adaptation to new environments.

Growth at high temperatures requires many adaptations. Cells protect their nucleic acids from denaturation through modifications in sequence, compact tertiary structures, and small ligand binding [5]. At higher temperatures, enzymatic rates change, but not necessarily in a uniform manner, and so
cells modify their enzymatic rates to keep metabolic pathways running smoothly [6]. In addition, cells employ a variety of methods to guard against accelerated protein denaturation at high temperatures[6]. Studies have analyzed the molecular sequences from thermophilic organisms, and found that they exhibit characteristic biases in their proteins and structural RNA. Higher optimal growth temperature is correlated with a higher bias of charged to polar amino acids in proteins (CvP bias) [7], as well as a higher percentage of guanine and cytosine bases in the stem regions of the 16S rRNA [8]. These biases serve to stabilize protein and RNA structures at high temperatures [7,8]. Problems of metabolic rate and molecular stability are encountered in reverse when organisms adapted to high temperatures grow below their optimum. A member of the Thermotogales growing below its optimum would experience enzymatic rate change, and a tendency of enzymes to stay bound to substrates much longer than necessary, leading to stalling of metabolic pathways and buildup of intermediates.

One previous study investigated the proteins utilized by the archaeon Pyrococcus furiosus for growth at temperatures below its optimum [9]. This study used proteomics to determine which proteins are present at 90° C, the organism’s optimal growth temperature, and at 72° C, its minimum. They found that over 100 proteins were down-regulated at lower temperatures, specifically those involved in energy metabolism and protein synthesis, whereas only 11 proteins were up-regulated at lower temperatures [9]. This study and studies like it illustrate the complexity of the response to changing growth temperature, but they do not attempt to explain the acquisition or evolution of this response.

In the case of the Thermotogales, it may have been the horizontal acquisition of new regulatory genes or metabolic enzymes that allowed certain lineages to begin to tolerate and adapt to life at lower temperatures. We hope to demonstrate that this has occurred by examining the evolutionary history of the bacterium Kosmotoga olearia, whose wide growth temperature range and relatively low optimal growth temperature make it an ideal candidate for our experiment. Ultimately we hope to show that HGT can facilitate evolution of a whole-cell trait.
Methods

My research will require a combination of laboratory and computational skills. It involves collaboration between Dr. Peter Gogarten’s lab, which has experience with developing and implementing new bioinformatic techniques, and Dr. Kenneth Noll’s lab, which has experience with culturing the Thermotogales and with protein extraction for proteomics. Computational resources are provided by Dr. Gogarten’s lab, and funding for proteomics comes from a collaborative grant between Dr. Noll’s lab and a research group at Pacific Northwest National Laboratories.

As preliminary work, we have predicted which proteins will be present at different temperatures in *K. olearia* based on bioinformatic methods. Previously I demonstrated the robustness of the correlation between CvP bias and optimal growth temperature in the Thermotogales [10], and from this I created a correlation line between the CvP bias for each protein universal to the Thermotogales and optimal growth temperature. I then used that line to determine an “optimal temperature” for the universal proteins found in *K. olearia*. Of the three proteins with the lowest predicted optimum in *K. olearia*, one is a glutaredoxin, which was found to be present in *Pyrococcus furiosus* when grown at 72° C [9], one is a thioredoxin, which participates in redox regulation in consort with glutaredoxin [11], and the third is a molecular chaperone. I have also determined which genes are duplicated in *Kosmotoga olearia* that are not duplicated in *T. maritima*, using the BLAST search algorithm [12] with an e-value cutoff of 10e-20. I detected 59 such genes, and about 50% of them are involved in metabolic processes. It is possible that the reason *K. olearia* maintains multiple copies of genes that its relative *T. maritima* does not is that one of these copies is activated at lower temperatures.

We began the process of protein extraction for proteomics this summer. We grew large volumes of *Kosmotoga olearia* cells at 40°, 50°, and 65° C. The cells were isolated by centrifugation and lysed by bead beating and sonication. We separated the soluble (cytosolic) and insoluble (membrane) fractions of the cells by centrifugation. We then further separated the membrane fraction into inner and
outer cell membrane by centrifugation on a sucrose gradient. The samples will be shipped to the PNNL facility for the whole cell proteomic analyses, and from that we obtain information about what proteins are present in *Kosmotoga olearia* at three different temperatures.

From there, we aim to answer specific questions using the proteomics data. First, we will determine what proteins are expressed exclusively at each different temperature, and whether these are proteins involved in specific cellular functions. This can be done using the NCBI's Clusters of Orthologous Groups (COG) database [13]. It is likely that these results will be similar to the previous study on *P. furiosus*, and we will find changes in presence of proteins involved in carbohydrate metabolism and transcription.

Next, we will elucidate the evolutionary history of the proteins expressed at different temperatures, to determine if the proteins were acquired via HGT, and when the transfer occurred during the divergence of the Thermotogales. This will involve BLAST searches to gather related sequences, followed by phylogenetic reconstruction to determine the evolutionary relationship of the sequences. From these tree reconstructions, we can determine the nearest phylogenetic neighbor outside the group Thermotogales, indicating if, from where, and when the gene was horizontally transferred. I will write scripts using PERL and R to automate this process, so that it can be performed quickly on a large number of genes of interest. In addition, a presence/absence analysis will be used to determine which genes are present in which lineage within the Thermotogales. For example, if a certain gene is only found in one clade, the most parsimonious explanation is that the gene was transferred to the ancestor of that clade before its divergence. This analysis in combination with the HGT analysis will show us when and from where the proteins involved in temperature adaptation were acquired.

Our study will seek to explain the evolutionary history of temperature adaptation for *Kosmotoga olearia*, demonstrate the contribution of horizontal gene transfer to evolution, and allow for greater understanding of the process of microbial evolution.
Plan of Study: If awarded University scholar, I will be able to earn minors in Mathematics and Bioinformatics, and take MCB courses at the graduate level in my areas of interest. I have already completed two courses that will be critical to the completion of my project, MCB 5472: Computer Methods in Molecular Evolution, and MCB 3610W: Physiology and Biochemistry of Prokaryotes, as well as most of the courses required for my major. For this reason, I would relish the ability to take graduate courses, as there are many graduate courses in MCB and EEB that would allow me to further my intellectual development and contribute to the completion of my project. These graduate courses will also strengthen my resume when applying to graduate schools, as I will demonstrate my ability to successfully complete graduate level coursework. I intend to seek a PhD in molecular biology, so I have chosen not to stay an extra year to complete a master’s degree. My application to PhD programs will be strengthened more by extensive research experience documented by scientific publications than by a master’s degree, so I intend to focus my energies on research.

In the spring of 2012, I will be able to enroll in MCB 5240: Virology and MCB 5896: Frontiers in Microbial Systems Analysis, both graduate level courses. University scholar status will allow me to complete my final content area 2 Gen-Ed during the summer, allowing me to make time for Math 2210Q: Linear Algebra during the school year. This will put me on track to earning a math minor.

During the fall of 2012, I will be able to take graduate level MCB 5621: Molecular Biology and Genetics of Prokaryotes, a relevant and interesting course, as well as complete the MCB major requirements with MCB 5100: Biochemistry. The statistics course I will complete that semester, STAT 3445, will be the final course that I need to complete my minor in bioinformatics.

University Scholar will be the most important during my final semester. I will be concentrating on research and completing my thesis, but I will be very excited to take EEB 5349: Phylogenetics, which will inform me about tree reconstruction techniques, as well as finish my math minor with MATH 3170: Fundamentals of stochastic processes.
Works Cited


