Session I – Chemical Engineering
(Alphabetical)

Determining the Cytotoxicity of and Cell Tolerance to Bioorthogonal Non-canonical Amino Acid Tagging
Yuanyi Dong
Mentor: David A. Tirrell

Non-canonical amino acids are amino acids which normally do not exist in proteins. A variety of non-canonical amino acids containing azide or alkyne functional groups have previously been developed. These engineered non-canonical amino acids can be readily incorporated into newly synthesized proteins in both bacterial and mammalian cells. Non-canonical amino acid incorporation may cause protein instability. Here we focus on changes in protein stability and the associated toxicity to mammalian cells. We use a broad range of non-canonical amino acid concentrations to detect if there is an associated cytotoxicity and changes in cellular proteome as a result of protein misfolding and degradation. This is measured using protein gels and western blots. Cellular morphology is investigated using fluorescent microscopy.

The Creation of a Highly Thermo-stable and Highly Active P450 from a Chimeric Parent
Lucas A Hartsough
Mentors: Frances Arnold and Martina Carbone

Our goal is to create a highly stable and highly active cytochrome P-450_{BM3} mutant from a chimeric parent through directed evolution for use on an analog of the natural substrate, and another for use on human drugs with industrial application in mind. We have already completed the first cursory rounds of evolution, which have yielded a mutant that has a 29% improvement over the parental chimera on our test criteria, the percent of protein surviving a 10 minute incubation at 61°C. This mutant then became the next parent in a series of optimizations for both thermo-stability and residual activity after incubation, yielding a highly active progeny that maintained comparable activity to that seen in highly stable proteins produced using point-mutant parents.

Engineering New Strategies for Improving Cellulose Hydrolysis: A Multifaceted Approach
Arvind Kannan
Mentor: Frances Arnold

In this paper, we investigate several approaches for improving the degradation of cellulose by cellulases in industrially relevant conditions. First, we determined the tolerance of a diverse set of cellobiohydrolase class II (CBHII) cellulases to the ionic liquid 1-ethyl-3-methylimidazolium (EMIM) acetate, which is used in chemical cellulose treatments. We found that neither differences in thermostability nor sequence diversity produced measurable differences in IL tolerance across the CBHII enzyme set. Further mechanistic studies revealed that cellulase inactivation by ILs is not caused by irreversible denaturation of the protein, which limits the applicability of enzyme engineering strategies to improve IL tolerance. Second, we designed a protocol for the magnetic capture of cellulose, which can be combined with cell surface display vectors to screen for novel cellulose-binding proteins. We successfully bound the cellulose to a biotinylated scaffoldin, allowing complete capture of the cellulose with streptavidin-coated magnetic beads. Third, pilot experiments revealed that high throughput screens for CBHII activity can be improved by using smaller cellulose particles and that cellulase binding can be measured via competitive inhibition by non-catalytic scaffoldins. Finally, we are screening saturation mutagenesis libraries for sites in the cellobiohydrolase class I (CBHI) catalytic tunnel so as to determine the relationship between the enzyme's catalytic activity and its processivity.

Enzymatic Ordering in Cellulosomes
Aaron M. Levine
Mentors: Jamie Cate, Frances Arnold, and Veronica Zepeda

A major impediment to the economical conversion of plant matter into bio-fuels is the energy intensive breakdown of lignocellulosic matter into simpler hexose and pentose sugars that can then be fermented into usable fuels such as ethanol. Produced by various soil bacteria, cellulosomes, complexes of multiple cellulose degrading enzymes attached to a common protein scaffold, have been observed to degrade cellulosic matter with higher efficiency than their constituent enzymes alone. To understand how this increased efficiency is achieved we studied the cellulosomes produced by *Clostridium papyrosolvens* (*CpC7*). An outstanding question is whether cellulosomes are random assemblies, or have preferential ordering of proteins bound to the enzymatic attachment sites on the scaffold. To observe any ordering of the over seventy cellulosomal proteins *CpC7* is capable of producing, we made a *CpC7/E. coli* shuttle vector containing the gene encoding the protein scaffold of the *CpC7* cellulosome. Through PCR, the linker regions connecting the enzyme binding sites on the scaffold were replaced with protease cleavage sites so

that after expression in *CpC7* the resulting cellulosomes could be purified, cut, and analyzed by mass spectrometry to determine any specificity.

A Model System for the Chemical Synthesis of Multivalent Antibody Structures
Helen Luo
Mentor: David A. Tirrell

Antibodies and antibody fragments form an integral part of protein diagnostic and therapeutic agents that may increase our ability to combat infectious agents and cancer. Many efforts have been undertaken to create structures composed of multiple antibody fragments that can bind to two or more antigens simultaneously. Previous methods have created multivalent structures using recombinant DNA technology, but these approaches are limited by proper protein expression and folding. We are developing an alternative strategy where chemistry can be used to covalently link fragments. Our model system consists of a single-chain variable fragment (scFv) capable of high affinity binding to digoxin. By incorporating reactive azide- and alkyne-containing methionine analogs into separate scFv samples, we can link scFvs together using "click chemistry." In order to control the location of the linking, we are constructing a high-affinity methionine-free scFv using rational design and high-throughput screening approaches. Subsequently, we will reintroduce methionine codons into the scFv at locations calculated to be surface accessible based on available crystal structures. Click chemistry experiments will be performed to study the efficiency of various linking locations. We will characterize the binding kinetics of our multimeric products to determine the viability of this modular approach to synthesizing multivalent structures.

Directed Evolution of Fungal Cellulase toward Improved Hydrolysis of Cellulose
Shruti Mishra
Mentor: Frances Arnold

Cellulose is the most abundant carbon compound on Earth, and as such is a very plentiful resource. Recent studies have also shown that cellulosic biofuels have the potential to replace fossil fuels in meeting energy needs. However, before cellulosic biofuels can be used commercially, we need efficient and thermostable cellulases, the enzymes responsible for the hydrolysis of cellulose. This is achieved by evolving the cellulases found in nature to have these desired properties. In nature, the source of improvement is random mutations that cause one mutant to be more evolutionarily fit. We modeled this in the laboratory by using the error-prone polymerase chain reaction to generate random mutations in the gene encoding the natural enzyme. In order to determine which mutant was best in the laboratory we screened each enzyme with substrate and used a sugar assay to compare their overall activities. So far, we have screened 3000 and 4500 mutants from two libraries with different mutation rates, respectively, and have narrowed our search to the top 18 mutants. Once we identify the top mutants, we will purify them for characterization. This will complete one round of directed evolution, and the best mutant of this generation can be used as the parent of the next generation, in order to facilitate further evolution.

Creating a Diverse Group of Thermostable CBHI Cellulases for Use in Biofuel Applications
Shannon R Mohler
Mentors: Frances Arnold and Pete Heinzelman

Structure-guided recombination involves using protein crystal structures to decompose related enzymes, or "parents," into constituent stretches of continuous amino acids, or "blocks," and then reassembling these blocks to produce novel enzymes with a high probability of retaining functionality. We have applied structure-guided recombination to the challenge of creating a collection of thermostable CBH I enzymes with diverse and novel properties for use in biomass-to-fuel conversion processes. Already, three parent CBH I enzymes and a number of chimeras created by recombining blocks from these enzymes have been expressed in yeast. We have already created one CBH I chimera that is more thermostable than the three parents. Further recombination could lead to CBH Is with even greater thermostability and will result in our having a collection of functionally diverse CBH I enzymes. This family of CBH I enzymes can be combined with other cellulases, such as the family of CBH II enzymes that our laboratory has created, in formulating application specific enzyme mixtures for biomass degradation applications.

Exploring Evolutionary Pathways of Improved Enzymatic Activity in Different Amino Acid Compositions
Karthik Narsimhan
Mentor: David A. Tirrell

Engineers have exploited protein promiscuity to evolve proteins with enhanced activity towards new substrates. The Arnold lab has employed directed evolution to evolve Cytochrome P450, a monoxygenase, to oxidize linear alkanes and gaseous alkanes into alcohols. Alteration of protein specificity has not been extensively studied when incorporating non-canonical amino acids into proteins. Chloramphenicol

acetyltransferase (CAT) was evolved to be active towards thiamphenicol, a methyl-sulfonyl chloramphenicol analogue. CAT was evolved in two environments: using homopropargylglycine (hpg), a non-canonical methionine analogue, and nineteen amino acids; and the twenty natural amino acids. Evolution of both environments was done in parallel. Global quantitative incorporation (> 95%) of hpg allowed direct comparison of the effects of incorporation. Error prone PCR was employed to diversify CAT sequences, which were expressed in methionine auxotrophic *E. coli*. The media shift technique allowed for the expression of a CAT library with hpg and variants were assayed for improved activity. Screening will yield CAT variants with mutations that facilitate methionine analogues at methionine positions or eliminate methionine positions. CAT's evolution in various environments will demonstrate how different sets of available amino acids could drastically affect the optimal enzymatic structure and limit or enhance enzyme function towards altered substrates.

Exploring the Role of Higher-Order Energetic Effects in Protein Structure
Albert H. Ng
Mentors: Frances Arnold and Christopher Snow

Due to the intimate relationship between protein structure and function, detailed models of folded protein conformations are essential to the field of molecular biology. To this end, numerous computational algorithms and molecular mechanical force fields have been developed to predict native protein three-dimensional structures. Many of these optimization algorithms are only suitable for energy functions that consist of pairwise decomposable effects. Our goal is to improve upon the accuracy of current side-chain repacking algorithms by introducing higher order energy effects using the Arnold Lab's *CHOMP* package. Specifically, the AMOEBA force field is used to examine the role of polarization in the packing of buried side-chains. Third-order energy terms are calculated for the 1PGB test case via parallel computing, and the inclusion of these terms is shown to improve the approximation of the energy landscape compared to that with only pairwise terms. A distance-dependent cutoff metric is introduced to reduce the number of third-order calculations while still capturing the most significant terms. The FASTER combinatorial optimization algorithm is modified to work with the new energy terms, and repacking performance of the third-order AMOEBA approximation is compared to those of the second-order approximation as well as the OPLS-aa and Rosetta energy functions.

Coarse-Graining Of Ionic Interactions With Amino Acids
Gregory J Rubinstein
Mentor: Garegin Papoian

Coarse-grained molecular dynamics are necessary for processes with long time scales and large length scales, which are computationally impractical for all-atom simulations. Including explicit ions in coarse-grained models significantly improve their accuracy in simulating certain systems, especially those that are highly charged. In this study, the interactions of ions with amino acids are added to the coarse-grained model. Atomistic simulations of each amino acid in solvents containing water and ions are conducted in order to determine the ionic radial distribution functions (RDF). Coarse graining of the amino acid system is then achieved by translating the atomistic RDFs into parameter values for the Hamiltonian of the interactions between the ions and amino acid. Optimization of the Hamiltonian parameter values is performed with renormalization group theory. The ionic interactions with the amino acids are then implemented into the coarse-grained models with the fully optimized parameter values. Such improved models can then be used to more accurately model complex biological systems, such as those containing histones and histone tails, which are essential in the accessibility of specific DNA sequences.

The Effect of Underlying Substrate Compliance on the Maturation of Cell-cell Contacts in Multi-cellular Epithelial Aggregation
Dongying Shen
Mentors: Anand Asthagiri and Jin-Hong Kim

The compliance of extra-cellular matrix (ECM) has been shown to influence many cell behaviors including cell differentiation and migration. In fact, stiffening of surrounding tissue has often implicated in cancer development by deregulating tissue homeostasis. In our study, cells were plated on polyacrylamide gels, whose stiffness can be systematically varied by changing the extent of cross linking with different bis-acrylamide concentrations to obtain various substrate compliances. To visualize and measure the extent of cell-cell contact, the immunofluorescence staining of contact-related proteins, E-cadherin and vinculin, is performed on normal rat kidney (NRK) and Madin-Darby canine kidney (MDCK) epithelial cells. A more continuous E-cadherin pattern is observed in NRK cells plated on softer gels, indicating more mature cell-cell contact. Vinculin is preferentially recruited to cell-cell contacts, signifying contact maturation, on softer gels, whereas it localizes to cell adhesion to substrate on stiffer ones. Our study will be completed by quantitatively measuring the extent of contact maturation for different substrate compliances through automated image processing.

Cellulase Engineering: Multi Carbohydrate Binding Modules to Improve Activity of HJPLUS
Xinlin Yu
Mentors: Frances Arnold and Pete Heinzelman

Because petroleum, which common everyday materials are made from, is steadily running out, biorefineries are using biomass which is mostly composed of cellulose to make glucose which can then be fed to the microbes that produce the polymers to make the everyday items. Unfortunately this process is very costly due to the price of cellulase, the enzyme that catalyzes the hydrolysis of cellulose. This project is focused on engineering a cellulase with improved activity by adding a second carbohydrate binding module (CBM) to the N or C terminal of HJPLUS, a CBHII. After creating the various constructs, binding assays where the enzymes bind to Avicel at various temperatures (60-80°C) and time periods are performed to determine where or not there are improvements. From the data presently collected multi CBMs do not appear to improve activity.