Protein Complex Analysis Project (PCAP): Microbiology Subproject

Hoi-Ying Holman[1,4], Jay Keasling[1,2,4], Aindrila Mukhopadhyay[1,4], Swapnil Chhabra[1,4], Jil T. Geller[1,4], Mary Singer[1,4], Dominique Joyner[1,4], Tamas Torok[1,4], Judy Wall[3,4], Dwayne A. Elias[3,4], **Terry C. Hazen**[1,4a] *(tchazen@lbl.gov)

[1]Lawrence Berkeley National Laboratory, Berkeley CA; [2]University of California, Berkeley CA; [3]University of Missouri, Columbia MO, [4]Virtual Institute for Microbial Stress and Survival, http://vimss.lbl.gov; *Presenting Author. [a]Subproject leader

The Microbiology Subproject of PCAP provides the relevant field experience to suggest the best direction for fundamental, but DOE relevant research as it relates to bioremediation and natural attenuation of metals and radionuclides at DOE contaminated sites. This project builds on techniques and facilities established by the Virtual Institute for Microbial Stress and Survival (VIMSS) for isolating, culturing, and characterizing *Desulfovibrio vulgaris*. The appropriate stressors for study have been identified and, using stress response pathway models from VIMSS, the relevance and feasibility for high throughput protein complex analyses is being assessed. We also produce all of the genetically engineered strains for PCAP. Two types of strain are being constructed: strains expressing affinity tagged proteins and knock out mutation strains that eliminate expression of a specific gene. We anticipated producing over 300 strains expressing affinity tagged proteins every year for complex isolation and EM labeling experiments by the other Subprojects. A much smaller number of knockout mutation strains are being produced to determine the effect of eliminating expression of components of putative stress response protein complexes. Both types of engineered strains are being generated using a two-step procedure that first integrates and then cures much of the recombinant DNA from the endogenous chromosomal location of the target gene. We are developing new counter selective markers for *D. vulgaris*. This procedure will 1) allow multiple mutations to be introduced sequentially, 2) facilitate the construction of in-frame deletions, and 3) prevent polarity in operons. The Microbiology Subproject provides high throughput phenotyping of all engineered strains to determine if any show phenotypic changes. We also determine if the tagged proteins remain functional and that they do not significantly affect cell growth or behavior. The knockout mutations are tested in a comprehensive set of conditions to determine their ability to respond to stress. High throughput optimization of culturing and harvesting of wild type cells and all engineered strains are used to determine the optimal time points, best culture techniques, and best techniques for harvesting cultures using real-time analyses with synchrotron FTIR spectromicroscopy, and other methods. Finally, we are producing large quantities of cells under different conditions and harvesting techniques for optimal protein complex analyses. To insure the quality and reproducibility of all the biomass for protein complex analyses we use extreme levels of QA/QC on all biomass production. We expect to do as many as 10,000 growth curves and 300 phenotype microarrays annually and be producing biomass for 500-1000 strains per year by end of the project. Each biomass production for each strain and each environmental condition will require anywhere from 0.1 – 400 L of culture, and we expect more than 4,000 liters of culture will be prepared and harvested every year. The Microbiology Subproject is optimizing phenotyping and biomass production to enable the other Subprojects to complete the protein complex analyses at the highest throughput possible. Once the role of protein complexes has been established in the stress response pathway, we will verify the effect that the stress response has on reduction of metals and radionuclides relevant to DOE.

During the last year, the Microbiology Subproject has supplied more than 30 sets of *D. vulgaris* cultures on biofilms for EM analysis, more than twenty 1-5 L cultures of biomass for water-soluble protein complex purification studies, and two 100 L and one 200 L cultures for membrane protein complex purification. We have designed and implemented a continuous culture system that enables us to produce more than 300 L of anaerobic mid log phase *D.*

vulgaris in as little as 5 days, including harvesting and QA techniques that maxmize reproducibility of all biomass produced. The goal of incorporating different affinity or tandem affinity (TAP) protein tags into three genes to determine the best tag for use in the PCAP project will be complete by the end of 2006. These include the Strep-tag® (IBA) for streptavidin-binding, the SPA-tag (a.k.a. CTF) that consists of a calmodulin binding motif, tobacco etch virus protease (TEV) and 3X FLAG affinity, as well as a combination of these that replaces the calmodulin binding with the Strep-tag® resulting in STF. The three genes to test are the dissimilatory sulfite reductase subunit C, pyruvate ferrodoxin oxidoreductase subunit B and ATP synthase subunit C. Additionally, several other gene targets have been identified through close collaboration with the VIMSS/ESPP group at LBNL and are currently being tagged. To determine localization of a given gene product in the cell, we have utilized both the tetracysteine and SNAP-tag™ (Covalys) in cooperation with the EM group of the PCAP project. Currently the total number of genes tagged with CTF are 3, with STF 6, with strep 20, with tetracysteine 8 and with SNAP 3. We are currently attempting to construct an ordered library for tagging in DvH. By doing so, this will allow for a relatively small number of *E. coli* clones to carry all of the genes for DvH, thereby reducing the overall workload and paving the way for higher throughput tagging. We also compared several cloning strategies for producing tagged constructs for the plasmid insertion strategy in *D. vulgaris*. The two-step TOPO-GATEWAY strategy (Invitrogen) was identified as the most economical commercially available conventional-cloning strategy amongst these. We also developed a workflow for high throughput production of tagged strains of *D. vulgaris*. This workflow was based on the TOPO-GATEWAY strategy in combination with current technology for transformation of *D. vulgaris*. We constructed custom destination vectors carrying the tags SPA and STF (to realize the GATEWAY step) - these are not available commercially. We tested the workflow through all the steps for a set of 10 randomly chosen genes from the *D. vulgaris* genome. Based on generated sequences, five of these were successfully tagged with the SPA tag. We also tested a commercially available 96-well electroporation device (BTX) for high throughput transformation of *D. vulgaris*. This system was found unsuitable. We are currently developing a custom solution for this. We also collaborated with the Subgroup D (Computational Core) for the development of automated algorithms for: 1) Primer identifications based on gene locations within operons for PCR amplifications in 96-well format, and 2) Sequence alignments for identifying errors in the amplifications or cloning steps in the workflow. All of the tagged strains constructed this year have been characterized using phenotypic microarrays (PM), and the *D. vulgaris* megaplasmid minus strain (MP(-)) being used in the electroporation studies is being aggressively characterized for all differences including stress responses by the ESPP project.