Project Name:
Proteomic and Genomic Analysis of Lycopene-Overproducing *Escherichia coli* Strains

Professor Charles L. Cooney, Department of Chemical Engineering, MIT

Project Overview:
Systems biology represents a powerful method to describe and manipulate phenotypes of interest by analyzing and incorporating biological information from various levels of cellular organization. Such an approach is illustrated from a library of both rationally-directed and combinatorial gene knockout strains of *E. coli* which have been shown to produce various levels of the small molecule lycopene when transformed with the pAC-LYC plasmid. Lycopene is an important neutraceutical of the diverse and valuable isoprenoid class, and therefore an improved description of its recombinant production has significant implications. A systems biology strategy was followed to discover global proteomic and genomic expression changes associated with increased lycopene production of mutant *E. coli* constructs using whole-genome *E. coli* DNA microarrays and a novel LC-MS technique, respectively.

Faculty Involved:
• Prof. Charles L. Cooney (MIT)
PhD Student Involved:
• Brian E. Mickus (MIT)

Expected deliverables:
• 2 Publications
• Systems biology model for applying proteomic and genomic data to the designing of experiments for further increasing yield of a desired product
• Elucidation of the mechanism by which such overproduction is achieved

Results:
Following validation of the DNA microarray method by an ANOVA analysis, transcriptional profiling was applied to the five high lycopene-producing knockout mutants. One observation was greater than 4 times more differential expression for strains with a deletion in *hnr*, an important regulator of s_S factor, as compared to mutants with metabolic targets. Additionally, the conservation of expression trends across a majority of distantly-related mutants for specific genes was seen.
Within a specific mutant family of D*gdhA*, D*gdhA* D*aceE*, and D*gdhA* D*aceE* D*pyjid* strains, *hisH* of the histidine biosynthetic pathway was down-regulated more than 10-fold while a gene encoding for a critical component of ATP synthase, *atpE*, was up-regulated more than 3-fold in each of the mutants. Based upon these array results and a link between histidine and purine biosynthesis, the pre-engineered and mutant strains were supplemented with

histidine and tryptophan; however, it was found that lycopene production was not significantly affected for any of the strains.
To supplement the genomic data, proteomic expression changes between the strains were analyzed using a novel LC-MS technique for simultaneous identification and label-free quantitation of proteins. The combined database search results from the LC-MS analysis of the eight samples provided a total of greater than 500 protein identifications. While a majority of proteins showed little expression change in the mutants relative to the pre-engineered strain, some key proteins ranged from 10-fold up-regulated to 10-fold down-regulated across all proteins and mutants. In particular, WrbA, a NADPH quinone oxidoreductase which reduces the quinone pool, was observed up-regulated while MdoG, a periplasmic protein which may affect the membrane storage capacity for lycopene, was down-regulated.
A simultaneous examination of both genomic and proteomic data revealed that the TCA cycle may be generally down-regulated while the PEP-glyoxylate cycle may be generally up-regulated across the various mutants. A possible explanation for the importance of the PEP-glyoxylate cycle is its effect upon the cellular NADPH/NADH balance since lycopene biosynthesis requires large amounts of NADPH. Accordingly, a repressor of the glyoxylate pathway, IclR, and a membrane-bound pyridine nucleotide transhydrogenase, PntB, were targeted as potentially important to lycopene production.
Current work is progressing to overexpress and delete the *hisH*, *wrbA*, *iclR*, and *pntB* genes and determine resulting effects upon lycopene production.

Timeline (through August 2008):
• February-April 2008: Molecular biology work to overexpress and delete the *hisH*, *wrbA*, *iclR*, and *pntB* genes based upon previous projects and to determine the effects upon lycopene production
• May 2008: Further examine proteomic data for additional key proteins that are differentially-expressed with lycopene production and investigate enrichment of specific gene ontologies in expression data sets
• June-August 2008: Additional work following-up on results