Comparison of conserved protein complexes across multiple microbial species to evaluate high-throughput approaches for mapping the microbial interactome

D. Pelletier[1] (pelletierda@ornl.gov), G. Hurst[1], S. Kennel[1], L. Foote[1], P. Lankford[1], T. Lu[1], W. McDonald[1], C. McKeown[1], J. Morrell-Falvey[1], D. Schmoyer[1], E. Livesay[3], F. Collart[2], D. Auberry[3], K. Auberry[3], Y. Gorby[3], B. Hooker[3], E. Hill[3], C. Lin[3], P. Moore[3], R. Moore[3], R. Saripalli[3], K. Victry[3], V. Kery[3], S. Wiley[3], and M. Buchanan[1]

[1]Oak Ridge National Laboratory, Oak Ridge, TN; [2]Argonne National Laboratory, Chicago, IL; [3]Pacific Northwest National Laboratory, Richland, WA

The ever-increasing number of available complete genome sequences for bacteria that have been less extensively studied than model systems such as *E. coli* has led to the need for additional methods for system wide functional genomic analysis. Methods (2-hydrid or affinity isolation) for the systematic identification of protein-protein interaction networks have been applied to a number of organisms (e.g., yeast, *C. elegans*, *E. coli*, etc.). It is clear that no single method can be used to fully describe all the interactions in an organism. These data when combined with other genome scale analyses, the transcriptome, the proteome and phenotypic observations, can lead to a better understanding of molecular interactions involved in a biological process.

We have developed a general, scaleable methodology for mapping protein-protein interaction networks, based on affinity isolation of protein complexes, which can be extended to bacterial systems beyond well-characterized model organisms. Specific open reading frames from the annotated genome are identified, cloned and expressed as fusion protein "baits" bearing affinity tags in the target organism (endogenous approach) or in *E. coli* followed by purification then exposed to the proteins from wild type organisms (exogenous approach). Following affinity isolation, protein interactors are identified using mass spectrometry. The metadata are managed using Nautilus LIMS. MS data analyzed by SEQUEST and then undergo computational and statistical analysis. We are using this approach, as part of the Center for Molecular and Cellular Systems, on two phylogenetically distinct bacteria of interest to DOE--the alpha-proteobacterium *Rhodopseudomonas palustris*, and the gamma-proteobacterium *Shewanella oneidensis*. Initial experiments have focused on a number of protein complexes conserved across multiple microbial species to validate this approach for mapping microbial interactomes.

As an example of these studies, the results of mapping the interaction subnetworks for homologs of open reading frames annotated to be part of DNA-directed RNA polymerase, F1F0-ATP synthase and the RNA degradosome will be presented. Results from *R. palustris* and *S. oneidensis* will be compared to results obtained for homologous subnetworks in other organisms that are available in public protein interaction databases (BIND, DIP, etc.). Our data demonstrate the feasibility of our high-throughput approaches for mapping microbial protein interaction networks.

An internal project database and website tracks and summarizes data from all experiments. Publicly accessible views of selected pages of this website show a summary of the status of targeted proteins for the endogenous approach (http://maple.lsd.ornl.gov/cgi-bin/gtl_demo/public_target_status.cgi) and for the exogenous approach (http://maple.lsd.ornl.gov/cgi-bin/gtl_demo/public_ex_target_status.cgi,), and example interactor identifications for a selected subnetwork (http://maple.lsd.ornl.gov/gtl_demo/index.html).