Final Report on DE-FG03-01ER15253
PI: Faull, Kym F.

An Integrative Approach to Enegry, Carbon, and Redox Metabolism in the Cyanobacterium Synechocystis Sp. Pcc6803

This award was part of a collaborative effort to integrate studies of the structure, physiology and biochemistry of *Synechocystis* sp. PCC 6803 toward a functional model of the organism, awarded under the 'Microbial Cell Project' initiative, later incorporated into the Genomes to Life program. As indicated in the original proposal, the work at UCLA covered the proteomics sub-section of the project, complementing ultrastructural (R. Roberson, ASU) and physiological studies of mutant cell lines (W. Vermaas, ASU).

Expression proteomics.
Two primary technologies have been employed for analysis and measurement of the *Synechocystis* proteome.

1. 2D-gel electrophoresis. Currently one of the most reliable options in quantitative proteomics, typical 2D-gel experiments use isoelectric focusing (IEF) in the first dimension. In the case of membrane proteins, detergents must be added to maintain their solubility though only neutral/zwitterionic surfactants are compatible with the IEF process. We have optimized 2D gel separations for *Synechocystis* proteins extracted and separated into soluble and membrane sub-fractions. The resolution and coverage of integral membrane proteins is only marginally satisfactory and alternatives to the first dimension are being considered. Size-exclusion chromatography under non-denaturing conditions was one option that was explored but resolution was insufficient for subfractionation of the membrane-bound proteome. A more highly resolving technique, the 'Blue-native gel' has proven excellent for *Synechocystis* and we plan to set up this technology in the near future. Proteins with altered expression are being identified through standard LCMSMS technologies. The analysis of PSI, PSII and SDH deficient mutants is completed, establishing the comparative aspect of the project for integration with the ultrastructural and metabolomic experiments at ASU. We are also looking forward to receiving *ftsZ* and *VIPP1* interruption mutants to explore the effects on the proteome of cell enlargement and disruption of thylakoid biogenesis, respectively.
2. 2D liquid chromatography with mass spectrometry of intact proteins. Early experiments with total membrane protein extracts of *Synechocystis* showed that the spatial resolution of the reverse-phase separation used in front of the mass spectrometer limited detection to the one hundred or so most abundant proteins. The intact mass tags (IMTs) measured in this experiment represent the first of these measurements that will ultimately define the entire proteome. While some of the IMTs were matched to masses calculated from translations of genomic open-reading frames allowing reasonably confident identification of about half of them (hypothetical IMTs), we are currently validating identifications using a combination of peptide mass fingerprinting after cyanogen bromide cleavage and LC-MSMS after trypsin, of protein in fractions collected during LC-MS+. In

order to gain more complete proteome coverage we are applying a liquid separation in front of the LC-MS+ experiment. Size-exclusion chromatography is the first separation technology to be employed, yielding immediate benefits, while still not satisfactory for overall resolution of complexes. Total membranes were solubilized with dodecyl maltoside (1.5 %) and separated on deactivated silica (G 4000 SW). LC-MS+ analysis of less-retained chlorophyll-containing fractions, using reverse-phase and size-exclusion technologies, yielded intact protein mass spectra of the two large photosystem I subunits *PsaA* and *PsaB* as well as many other IMTs (Figures 1 & 2). These integral membrane proteins have eleven transmembrane helices and, at 81 and 83 kDa, represented one of the most significant challenges to the intact protein molecular weight approach. The identities of the proteins were confirmed by peptide mass fingerprinting and while there is good general agreement between measured and calculated masses it is noted that modest post-translational modifications are necessary to account for the measured molecular weights of the intact proteins. Whether these discrepancies are due to genuine post-translational modifications or DNA sequence errors remains to be determined. The data have been published allowing us to claim to be the first to have completed high-resolution electrospray-ionization mass spectrometry of the core subunits of Photosystem II, Photosystem I and the cytochrome b_6f complex providing effective proof-of-principle for application of the intact mass approach to the integral membrane proteome. Significantly, we reported greater integral membrane proteome coverage than a colleague studying thylakoids of *Arabidopsis* illustrating the benefits of the technique over sequential organic extraction of membrane proteins and 1D-gel analysis. The homogeneity of the *PsaA* and *PsaB* protein mass spectra attest to the quality of material grown at ASU and the viability of extraction and work up of the material after transport to UCLA.

Trainees.

Dr. Rodrigo Aguilera was recruited for proteomics of *Synechocystis* and made significant contributions before moving to USC as associate director of proteomics within the school of pharmacy. Rodrigo made a lateral move from a local doping laboratory where he was an acknowledged leader in the use of isotope ratio mass spectrometry to detect testosterone abuse and was given a thorough training in modern proteomics.

Mr Puneet Souda was recruited at the later stages of the project, bringing top-class expertise in 2D-gel electrophoresis from Dr. Frank Witzman's laboratory. Pete is also an information technology expert and has revolutionized the computing systems of the laboratory. He is being trained in modern proteomics and we hope to recruit him to the PhD program.

Infrastructure.

Since the original proposal we have made further instrument acquisitions relevant to the project. The new MALDI-TOF (DE-STR, Applied Biosystems) has been installed and is proving to be extremely sensitive for proteomics applications, most notably peptide mass fingerprinting. A 7 Tesla Fourier transform ion cyclotron resonance mass spectrometer has been installed and, while outside of the scope of the original proposal, will be

employed where necessary on the 6803 project. Recently we demonstrated the first application of Fourier-transform mass spectrometry to an integral membrane protein (26783 Da bacteriorhodopsin at 8ppm and 64000 resolution) opening this option for ultra-high accuracy and resolution measurements of *Synechocystis* proteins. We purchased a hybrid quadrupole-time-of-flight instrument (QSTAR XL) with nano-chromatography system (LC Packings). This instrument has become the workhorse LC-MSMS system in the laboratory and is employed for identification of 6803 proteins. A linear ion-trap instrument has been installed since January 2005 and is proving highly sensitive (attomole level for short peptides).

Publications directly supported by DE-FG03-01ER15253.

Whitelegge JP, Gómez SM, Aguilera R, Roberson RW, Vermaas WF, Crother TR, Champion CI, Nally JE, Blanco DR, Lovett MA, Miller JN, Faull KF. (2002) Identification of Proteins and Intact Mass Measurements in Proteomics. Applied Genomics and Proteomics **1**, 85-94.

This invited paper was written as an introduction to the proteomics approaches being employed at UCLA, including 2D-gel electrophoresis and liquid chromatography mass spectrometry of intact proteins. A preliminary list of hypothetical *intact mass tags* from *Synechocystis* sp. PCC 6803 was presented.

Whitelegge, J.P., R. Aguilera, H. Zhang, R. Taylor, W.A. Cramer (2002) Full subunit coverage liquid chromatography electrospray-ionization mass spectrometry (LCMS+) of an Oligomeric Membrane Protein: Cytochrome b_6f Complex from Spinach and the Cyanobacterium, *M. laminosus*. (2002). Molecular and Cellular Proteomics **1**, 816-827.

This paper continued an existing collaboration with Bill Cramer and established the importance of liquid chromatography with mass spectrometry and fraction collection (LC-MS+) as a powerful tool for studies of integral membrane protein complexes. Based on the hypothetical intact mass tags observed for the cytochrome *b* subunit we concluded the presence of a covalently bound heme molecule. This was confirmed when the crystal structures of the cyanobacterial complex as well as one from *Chlamydomonas* were solved, revealing a third *c*-type cytochrome attached to Cys 35 of the subunit. The structure papers, published in Science and Nature, both cited the mass spectrometry work in this paper, providing a significant jump-start to the citation index of *Molecular and Cellular Proteomics* in its first year and election of Dr. Whitelegge to the editorial board in 2004.

Whitelegge, J.P. (2003) Thylakoid membrane proteomics. Photosynthesis Research **78**, 265-277.

This invited paper was published in the journal's first special edition on proteomics and included new data from the *Synechocystis* project. 2D-chromatography combined with LC-MS+ was demonstrated to be suitable for analysis of PsaA and PsaB of the Photosystem I reaction center for the first time showing that electrospray-ionization mass spectrometry of large integral membrane proteins was feasible, in this case for eleven transmembrane helix molecules up to 83 kDa. The paper also

included the first example of the application of Fourier-transform mass spectrometry to an integral membrane protein, in this case Bacteriorhodopsin with seven transmembrane helices.

Whitelegge, J.P., Katz, J., Pihakari, K., Hale, R., Aguilera, R., Gómez, S.M., Faull, K.F., Vavilin, D, Vermaas, W. (2004) Subtle modifcation of isotope ratio proteomics (SMIRP); a new strategy for expression proteomics. Phytochemistry, **65(11)**, 1507-1515.

This paper was published in the first special edition of Phytochemistry on proteomics, for which Dr. Whitelegge was a guest editor. The paper outlines the basis for a new approach to quantitative proteomics in which a single isotopic envelop replaces two. Thus isotope ratio is used to code samples for relative expression, as well as turnover, measurements. The paper was reviewed and accepted for publication immediately before the 2004 meeting of the American Society for Mass Spectrometry, thereby establishing intellectual property to cover the concept presented there. Subtle modification of isotope ratio proteomics represents the best strategy available for coding of living humans, whose isotope ratio already varies depending upon dietary origin, and NIH have recently funded a proposal to develop the idea.

Whitelegge, J.P. (2005) Mass spectrometry for high throughput quantitative proteomics in plant research: lessons from thylakoid membranes. Plant Physiol. Biochem. **42(12)**, 919-27.

This invited paper reviews the current state of proteomics and presents new data for the first special issue on proteomics of this prominent European journal. Results of 2D-chromatography experiments on *Synechocystis* membrane preparations were presented along with a top-down experiment on the membrane proteolipid AtpH. This protein oligomer makes up the transmembrane proton turbine of the Fo domain of ATP synthase, and is the first two transmembrane alpha helix domain protein to be analyzed by tandem mass spectrometry.

Other publications acknowledging support of DE-FG03-01ER15253.

Whitelegge, J.P., Gómez, S.M., Faull, K.F. (2003) Proteomics of Membrane Proteins. Proteome Characterization and Proteomics. R.D. Smith and T. Veenstra, eds. Advances in Protein Chemistry 65: 271-307.

Gómez SM, Bil' KY, Aguilera R, Nishio JN, Faull KF, Whitelegge JP. (2003) Transit peptide cleavage sites of integral thylakoid membrane proteins. Mol Cell Proteomics **2 (10)**, 1068-85.

Whitelegge, J.P. (2004) HPLC and Mass Spectrometry of Intrinsic Membrane Proteins. Methods in Molecular Biology vol. 251, HPLC of Peptides and Proteins. M. Aguilar ed. 323-339.

Nally, J.E., Whitelegge, J.P., Aguilera, R., Pereira, M.M., Blanco, D.R., Lovett, M.A. (2005) Purification and proteomic analysis of outer membrane vesicles from a

clinical isolate of *Leptospira interrogans* serovar Copenhageni. Proteomics **5(1)**, 144-52.

Patents.

The SMIRP concept is the subject of a US patent application; “System and Method for Expression Proteomics based on Isotope Ratio Modification”, mailed 13th January 2005.

Future plans.

Current research continues to build upon progress made throughout the granting period, specifically in two areas.

Firstly, we are undertaking a through evaluation of quantitation through stable isotope labeling and quantitative mass spectrometry. There is little doubt that the future development of proteomics as a useful tool for the biologist is absolutely dependent upon accurate and reproducible quantitation. Further, this must be measured on the same signal that allowed protein identification, so that ***direct*** quantitation is achieved. That is, the quantitative measurement is extracted from the peptide mass spectrum, either by comparison of a pair of signals for relative quantitation, or by comparison to an internal standard for absolute quantitation. The role of SMIRP in this arena will be thoroughly explored under new R21 support from NIH (PI: Whitelegge).

Secondly, we are proceeding with development of 2D-chromatography of intact integral membrane proteins with non-denaturing technologies in the first dimension and LC-MS+ in the second.