

Mass Spectrometric Analysis of Protein Complexes from *Rhodopseudomonas palustris*

G.B. Hurst¹, N.C. VerBerkmoes³, T.K. Lankford², D.A. Pelletier², F.W. Larimer², R.L. Hettich¹, M.V. Buchanan¹, M.B. Strader¹, Y. Wang², L.J. Foote², S.J. Kennel²

¹Chemical Sciences Division and ²Life Sciences Division, Oak Ridge National Laboratory, and ³Genome Science and Technology Graduate School, University of Tennessee

Rhodopseudomonas palustris is a bacterial species that occurs widely in the environment and can survive in a variety of conditions ranging from light to dark, aerobic to anaerobic. This species is thus likely to have the potential for expressing different complements of proteins and protein complexes under different growth conditions. In addition to proteomics measurements under different growth conditions, a determination of the complement of protein-protein interactions, ranging from transient to stable, is vital to understanding the biology of this remarkable organism. As part of a center funded by the U.S. Department of Energy Genomes To Life program [1], we are analyzing protein complexes from *R. palustris* by expressing target proteins as fusions with affinity tags to allow subsequent isolation of other proteins associated with the target [2], followed by both “top-down” and “bottom-up” mass spectrometry analysis.

Selected *R. palustris* genes were cloned and expressed in both *E. coli* and *R. palustris* using modified pDEST vectors (Invitrogen). Fusions of these genes with His₆ and GST affinity tags, corresponding to N- and C-terminal positions on the proteins, were generated. Isolation of fusion proteins was accomplished using affinity purification with Ni-NTA or glutathione-bearing agarose beads. Expression was confirmed using 1-D PAGE and western blots. In-gel digestion of proteins in selected gel bands was followed by mass spectrometric analysis by either MALDI-TOF or LC-MS-MS using a quadrupole ion trap (QIT) mass spectrometer [2]. As a higher-throughput alternative to in-gel digestion, the entire affinity-isolated sample was analyzed by mass spectrometry [3], using both “top-down” (with FTICR MS) and, after trypsin digestion, “bottom-up” (with QIT MS) approaches. Protein identifications were accomplished by comparing tandem mass spectral data with sequence databases using Sequest or in-house tools.

As a first step, several *R. palustris* fusion proteins, expressed in *E. coli*, were verified by mass spectrometry. To complement data from western blot experiments, expression of the NirK Cu-containing nitrite reductase with an N-terminal GST tag was confirmed by MS analysis of in-gel digests of selected bands. Sequence coverage for both the GST and gene product portions of the fusion protein was obtained. Three other fusion proteins (GroEL-1 with C-terminal His₆, GroEL-2 with N-terminal GST, and GroEL-2 with C-terminal His₆) were verified by digesting the unseparated eluate from the affinity isolation, and subsequent nano-LC-MS-MS. However, several proteins from *E. coli* were also identified. Some of these *E. coli* proteins contained short runs of histidine residues that could impart some affinity toward the Ni-NTA affinity beads. Other *E. coli*

proteins presumably were isolated due to non-specific interactions with either the affinity medium or the fusion protein.

Plasmids encoding thirteen affinity-tagged fusion proteins were inserted into *R. palustris*. After growing and harvesting *R. palustris* cells, the tagged proteins were isolated using affinity chromatography. To date, five of these isolated proteins have been analyzed by mass spectrometry and anti-HIS western blotting. While all five proteins showed distinct bands in the western blot, only four were successfully identified using LC-MS-MS data, searched against the *R. palustris* database. In all cases, however, a number of additional proteins were detected, some of which were also observed in a strain not expressing any fusion protein. Use of a second affinity purification step, alternate tags, and optimized conditions will help to eliminate these background proteins.[4]

Most of the fusion proteins expressed to date in *R. palustris* are not normally produced by the bacterium under the growth conditions used, and so affinity isolation of these fusion proteins has not resulted in isolation of the protein complexes of which these fusion proteins are members. However, we isolated a C-terminally His₆ tagged chaperonin protein GroEL-2 using an ATP-containing buffer in order to stabilize the GroEL-GroES interaction. LC-MS-MS analysis of the digested isolate allowed identification of the His-tagged GroEL-2 (44 unique peptides; 76% sequence coverage) and unlabeled GroEL-1 (25 unique peptides; 64% sequence coverage), as well as their GroES-1 (2 unique peptides; 32% sequence coverage) and GroES-2 (7 unique peptides; 58% sequence coverage) counterparts in the chaperonin complex.

We are currently implementing an infrastructure for higher-throughput production and analysis of these fusion proteins for the purpose of isolating protein complexes from this bacterium under a variety of growth conditions.

References

1. <http://www.genomestolife.org>
2. (a) Gavin, A.-C. et al., "Functional organization of the yeast proteome by systematic analysis of protein complexes," *Nature* **2002**, *415*, 141-147.
(b) Ho, Y. et al., "Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry," *Nature* **2002**, *415*, 180-183.
3. Link, A.J. et al., "Direct analysis of protein complexes using mass spectrometry," *Nature Biotech.* **1999**, *17*, 676-682.
4. Puig, O. et al., "The tandem affinity purification (TAP) method: a general procedure of protein complex purification," *Methods* **2001**, *24*, 218-229.

This research sponsored by the Office of Biological and Environmental Research, U.S. Department of Energy. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.