# Evaluation of the Effect of Varying Cellular Lysis and Biomass Quantities on the Fractionation and Proteome Measurement of Microbial Isolates


Ritin Sharma[1,2], Brian Dill[1], Nathan VerBerkmoes[1] and Robert L. Hettich[1]

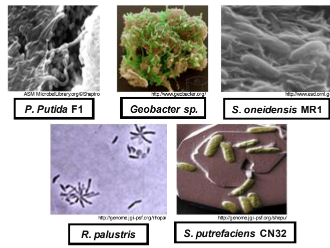
[1]Oak Ridge National Laboratory, Oak Ridge, TN; [2]Genome Science and Technology, University of Tennessee, Knoxville, TN

## OVERVIEW

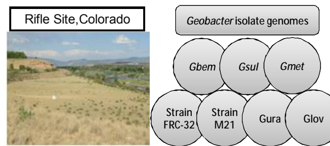
- Efficient cellular lysis and minimal protein loss are important factors in shotgun proteomic measurements. Sodium Dodecyl Sulfate (SDS), while effective for lysis, has been problematic because of its incompatibility with ES-MS.
- Recent work has shown the use and removal of SDS for deep proteome characterizations (Wisniewski *et al.*, Nature Methods 2009; Bothelo *et al.,* JPR 2010). We undertook a systematic study to evaluate the efficacy of four different SDS clean-up methods at varying biomass quantities (10 µg to 1 mg total protein) on *E. coli*.
- The best detergent clean-up method was coupled to GELFrEE (Gel-Eluted Liquid Fraction Electrophoresis) fractionation for deep proteome measurements of three different biological samples with increasing microbial complexity. The GELFrEE fractionation separates intact proteins based on molecular weight and produce liquid fractions which are well suited for downstream LC-MS/MS experiments. We report that by employing MS on fractionated and crude extracts simultaneously, proteome coverage can be enhanced for diverse range of biological samples.

## INTRODUCTION

- In this project, we have compared existing detergent clean-up methods for proteomics studies against modified FASP method and our in-house developed SDS-TCA protocol. For our preliminary studies, we have chosen a very well characterized microbe *E. coli*.
- Following this study, we used three biological samples with increasing complexity to investigate suitability of GELFrEE fractionation to enhance proteome coverage. For the lowest complexity sample we continued with *E. coli*. For the intermediate complexity sample we made an artificial mix of 5 microbial isolates in known ratio ( *P. putida* F1, *S. oneidensis* MR1, *S. putrefaciens* CN32, *R. palustris* and *Geobacter* sp.) For the high complexity sample we used ground water sample from Rifle site, CO.




- Rifle, Colorado is a former uranium and vanadium milling site which is now managed by U.S. Department of Energy for bioremediation.
- This highly contaminated site has several species of microbes which thrive efficiently and also degrade highly toxic soluble heavy metals to their less toxic insoluble form.
- The metagenome and metaproteome (Wilkins, AEM, 2009) from Rifle site revealed presence of several species of Geobacter which play a very important role in metal sequestration.




## TWO-STAGE STRATEGY FOR PROTEOMICS MEASUREMENT

- Using *E. coli* lysate at three different protein concentrations, we evaluated four different detergent (SDS) clean-up methods as shown in Figure 1.
- The best cleaned-up method was employed in preparing fractions obtained by fractionation of increasingly complex biological samples via GELFrEE technology as depicted in Figure 2.

**Figure 1.** Experimental flowchart employed in current study

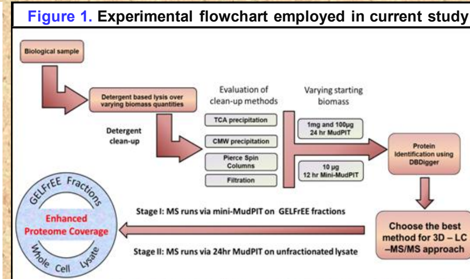
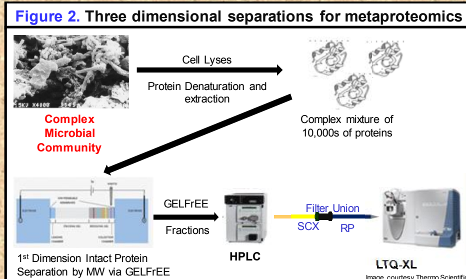



**Figure 2.** Three dimensional separations for metaproteomics




## EVALUATION OF CLEAN-UP METHODS

**Figure 3.** Four different methods of detergent clean-up evaluated with *E. coli* isolate

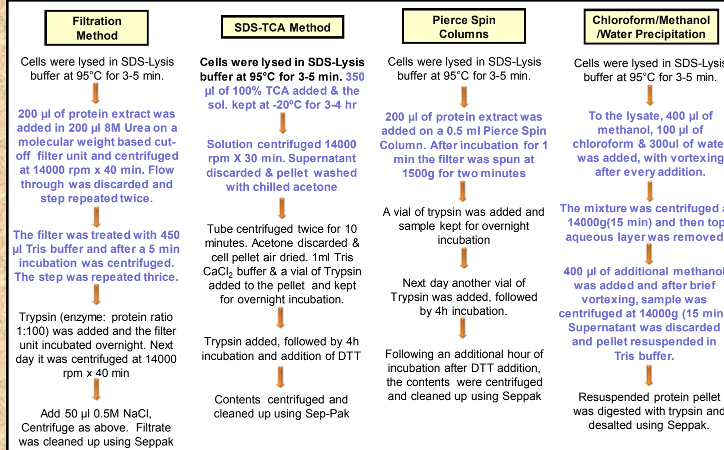



**Table 1.** Summary of MS measurements of each detergent clean-up method on 1 mg *E. coli* isolate

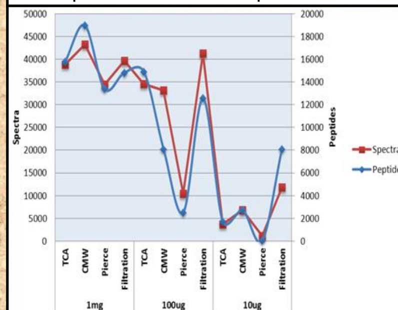
| Method | Number of Proteins | | | % of Predicted Proteome | Average % Sequence Coverage | Molecular Weight Range |
|---|---|---|---|---|---|---|
| | Replicate1 | Replicate2 | Combined Runs | | | |
| SDS-TCA | 1405 | 1334 | 1587 | 36.14% | 30.05% | 4.3kDa-181.5kDa |
| Pierce | 1395 | 1289 | 1513 | 34.5% | 27.47% | 4.3kDa-181.5kDa |
| CMW | 1449 | 1245 | 1546 | 35.20% | 34.13% | 4.2kDa-181.5kDa |
| Filtration | 1517 | 1535 | 1682 | 38.30% | 26.40% | 4.2kDa-181.5kDa |

- Table 1 shows that Filtration method gives comparatively higher number of protein identifications compared to other three clean-up methods.
- Figure 4 shows that the starting biomass is critical for efficient working of detergent clean-up methods. The decrease in starting biomass greatly affect all the clean-up methods except for filtration method.
- From the Table 1 and Figure 4 we conclude that the Filtration method is superior in giving higher protein identification at various starting biomass amounts.

### LC-MS/MS and Informatics

- All *E. coli* lysates were analyzed via two-dimensional (2D) nano LC-MS/MS system with a SCX back column connected to RP front column on a LTQ-Orbitrap-XL(Thermo Fischer Scientific) with 24 (1 mg, 100 µg total protein) and 12 hour (10 µg total protein) run per sample.
- All artificial community mix samples, Rifle ground water samples for the unfractionated whole cell study were analyzed via two-dimensional (2D) nano LC-MS/MS system with a SCX back column connected to RP front column on a LTQ-instrument (Thermo Fischer Scientific) with 24 hour runs per sample.
- All the GELFrEE fractions were analyzed via 2D nano LC-MS/MS system with a SCX back column connected to RP front column on LTQ (Thermo Fischer Scientific) instrument with 6 hours run per sample.
- The MS/MS spectra obtained from *E. coli* runs were searched against the *E. coli* predicted proteome, artificial community mix runs against the predicted proteome of 5 species used in mix, Rifle ground water sample runs against the Rifle metagenome and 7 *Geobacter* isolates using the in-house program DBDigger (Tabb DL *et al.*, Anal Chem, 2005).
- Only proteins identified with two fully tryptic peptides were considered for further biological study.

**Figure 4.** Peptide and Spectral Count for 4 different SDS clean up methods at three different protein concentration




## INTACT PROTEIN FRACTIONATION COUPLED TO FILTRATION/FASP-LC-MS/MS

**Figure 5.** Unbiased *E. coli* proteome measurement by GELFrEE fractionation

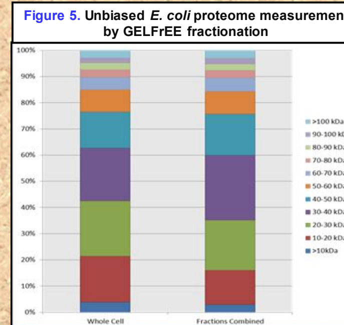
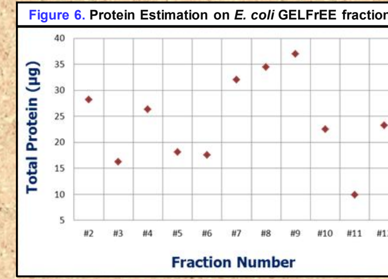



**Figure 6.** Protein Estimation on *E. coli* GELFrEE fractions




### GELFrEE Method

- For the Rifle ground water, artificial Community mix and *E. coli* isolate samples, GELFrEE 8100 10% Cartridge was employed.
- 150 µl of sample was loaded into each channel of the cartridge and fractions collected after pausing the runs at specific time intervals distributed across ~3 hours.
- At each pause, the recovered solution was combined into a single vial for each fraction.
- All the fractions were prepped using FASPKits (Protein Discovery, Inc.) in lieu of Filtration method described before.
- All the protein estimation experiments were carried out using Bio-Rad RC-DC protein assay.

- The GELFrEE fractionation approach is unbiased in fractionating proteome as shown in Figure 5. The *E. coli* proteome as per MW is well represented by both fractionation and whole cell lysate approach.
- Protein yield per fraction is on average 25 µg as shown in Figure 6 for *E. coli*. The total protein recovery is ~60%. This amount of protein is most suited to prep with filtration based detergent clean-up.
- As can be seen from Figure 7, the two stage approach results in increase in protein identification by 220 for low complexity sample of *E. coli*, 1191 for medium complexity sample of artificial community and 1688 for high complexity sample of Rifle Ground water.
- The "Unfractionated Lysate" protein identifications number in Figure 7 represent total number of proteins identified by running 24 hr MudPIT on three technical replicates for each biological samples.
- The "Fractions combined" protein identifications number in Figure 7 represent total number of proteins identified in a single run of 12 GELFrEE fractions via 6-hr mini-MudPIT.

**Figure 7.** Enhancement in proteome coverage by coupling fractionation+whole cell lysate LC-MS/MS

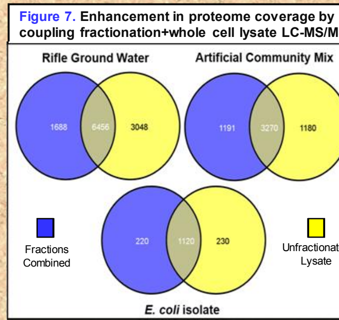
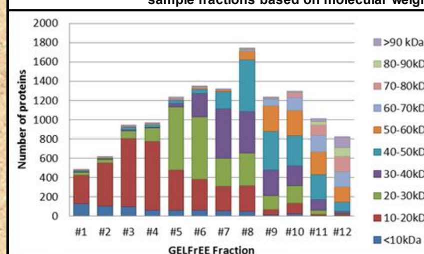



**Figure 8.** GELFrEE fractionation enriches artificial community sample fractions based on molecular weight




- Figure 8 illustrates that higher fraction number in a GELFrEE run represents fractions collected at later time points, and therefore are enriched in high molecular weight proteins.
- Figure 9 shows that each GELFrEE fraction contributes unique proteins to the total protein identification.
- Figure 10 illustrates that the unique proteins within each GELFrEE fraction of artificial community mix are also enriched according to molecular weight.

**Figure 9.** Distribution of unique proteins with respect to GELFrEE fraction from Rifle Ground Water

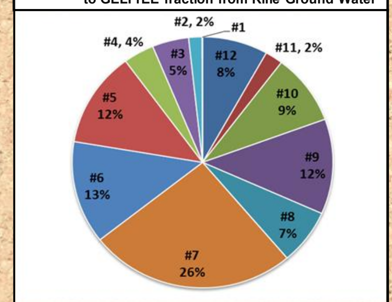



**Figure 10.** Distribution of unique proteins with respect to molecular weight in each GELFrEE fraction

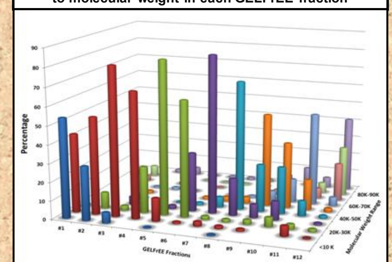



**Table 2.** Representative unique and abundant proteins identified by fractionationed or unfractionated protocols on Rifle groundwater sample

| Rifle groundwater Fractionated Runs | | | Rifle groundwater UnFractionated Runs | | |
|---|---|---|---|---|---|
| Gene Name | Organism | Function | Gene Name | Organism | Function |
| GbemDRAFT_3062 | *Geobacter bemidjiensis* | cold-shock DNA-binding domain protein | GlovDRAFT_0067 | *Geobacter lovleyi SZ* | Polyribonucleotide nucleotidyltransferase |
| Daro_0053 | *Dechloromonas aromatica RCB* | Nitrogen regulatory protein P-II | Gura_4238 | *Geobacter uraniumreducens RF4* | molybdopterin oxidoreductase |
| geob_sp_m18:GM1 | *Geobacter sp. M18* | H+transporting two-sector ATPase B/B' subunit | GbemDRAFT_3241 | *Geobacter bemidjiensis* | ribosomal protein L16 |
| GSU2843 | *Geobacter sulfurreducens* | rpsH: ribosomal protein S8 | Gmet_2638 | *Geobacter metallireducens GS-15* | Phosphoenolpyruvate carboxykinase (GTP) |
| GM18DRAFT_1843 | *Geobacter sp. M18* | outer membrane chaperone Skp (OmpH) | Gmet_3249 | *Geobacter metallireducens GS-15* | Biotin/lipoyl attachment |
| GbemDRAFT_1016 | *Geobacter bemidjiensis* | flagellar FlbT family protein | GSU0344 | *Geobacter sulfurreducens* | NADH dehydrogenase I, G subunit, putative |
| Gura_3731 | *Geobacter uraniumreducens* | uridylate kinase | GlovDRAFT_2185 | *Geobacter lovleyi SZ* | uroporphyrinogen decarboxylase |

## DISCUSSION

- Both the fractionated and unfractionated 2D-LC-MS/MS are able to provide deep proteome coverage of complex environmental samples.
- Both the fractionation and unfractionation schemes identified unique set of proteins for all the biological samples. An example of this is shown in Table 2 where unique and high abundance protein from Rifle ground water are listed.
- The GELFrEE method identified unique proteins across each liquid fraction, with majority of unique proteins within each fraction were limited to a distinct molecular weight cut-off.

## CONCLUSIONS

- When the starting biomass is high, all the detergent clean-up methods perform almost equally with filtration method slightly better with respect to total protein identification.
- When the sample amount is limited, the filtration method outperforms other three detergent cleaning-up method tested and gives more than twice the protein identifications compared to second best.
- The initial intact protein fractionation strategy provides an additional resolution level to existing 2D-LC MS/MS protocol.
- For the Rifle ground water sample, we were able to identify 11,192 proteins by combining information from 6hr 2d-LC-MS/MS runs on GELFrEE fractions along with 24-hour unfractionated runs.
- The amount of starting biomass sample available dictates the fractionation that can be envisioned. This can be circumvented by use of multiple cartridges of GELFrEE or pushing the loading limit of the GELFrEE system.


## ACKNOWLEDGMENTS

- R. Sharma is supported by UT/ORNL graduate program in Genome Science & Technology.
- This research was supported by the U.S. Department of Energy, Office of Biological and Environmental Research, Environmental Remediation Sciences Program.
- We would like to thank the entire Proteome working group from the DOE funded Rifle IFC research project.
- Special thanks to Jeremy L. Norris, Chuck Witkowski and Jay Harkins from Protein Discovery, Inc. for technical assistance and advice.
- Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy.