Multi-omics Sample Preparation Workflow for Proteins and DNA Using the Reversible Protein Tag ProMTag

Stephanie Biedka[1], Jonathan S Minden[1], Amber Lucas[1]

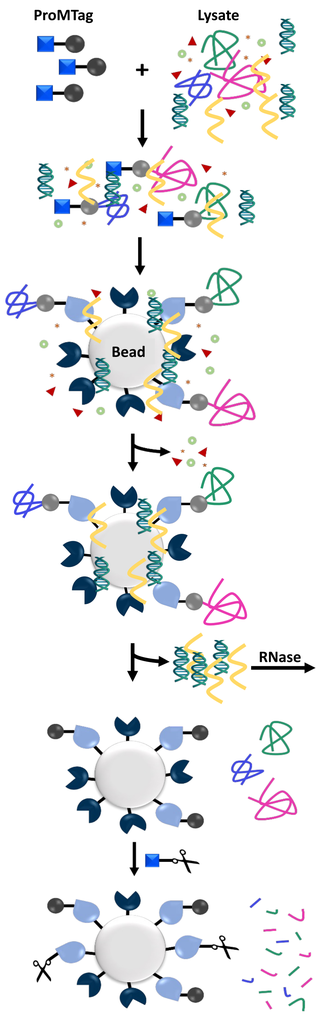
[1]*Impact Proteomics, Pittsburgh, PA*

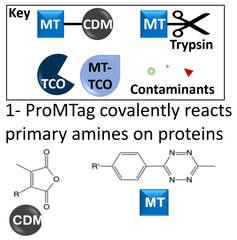
Abstract

Sample preparation is a crucial first step for both genomics and proteomics workflows. Removal of contaminants such as salts, detergents, and other biologics while maintaining high yields of the desired product is key to reproducible and informative results from these analyses. More and more frequently, these -omics technologies are being used in tandem to gain deeper insights into biological processes. However, multi-omics sample preparation remains tedious and usually requires many steps in multiple different sample preparation workflows.

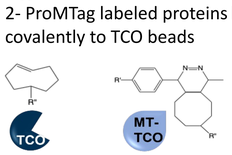
In this study, we present a new multi-omics workflow for the simultaneous preparation of DNA and protein samples from a single starting cell lysate. We accomplished this using the ProMTag reversible click chemistry technology that allows for reversible modification of the surface of proteins. Using ProMTag we were able to tag proteins in a cell lysate, bind them to ProMTag capture resin, and then precipitate nucleic acids so they also stay with the resin. With the nucleic acids and proteins bound to the resin, we were then able to wash away detergents, salts, and other contaminants. We then eluted the nucleic acids by resolubilizing in a nucleic acid elution buffer. We then were able to reverse the ProMTag by adding the protein elution buffer and eluted the sample in a mass spectrometry (MS) compatible buffer ready for proteomic analysis. Using this workflow we got yields >75% for protein and >90% for DNA. Gel electrophoresis showed a genomic DNA band free of degradation and the 260/280 absorption ratio indicated a pure DNA sample. Whole genome sequencing and MS proteomics analysis were performed and compared to traditional separate sample preparation workflows for DNA and proteins. This work establishes a new, high yield, reproducible workflow for the simultaneous preparation of DNA and proteins for genomics and proteomic analysis from a single starting sample.

ProMTag Multiprep DNA and protein extraction and cleanup workflow





1- ProMTag covalently reacts with primary amines on proteins



2- ProMTag labeled proteins bind covalently to TCO beads

3- Extensive washing and/or buffer exchange removes contaminants

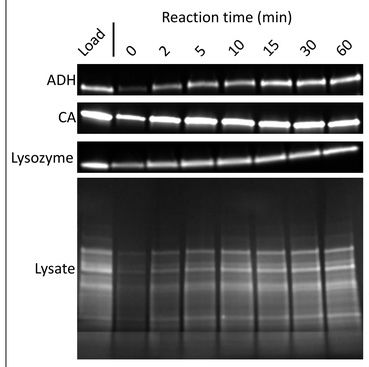
4- DNA and RNA are resolubilized in a nucleic acid elution buffer

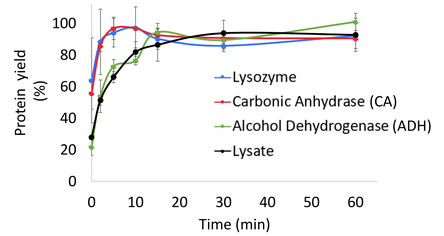
5- Proteins are released by addition of elution buffer to reverse CDM reaction

6- (Optional) MT-Trypsin is added to digest proteins into peptides. MT-Trypsin is captured on the TCO beads during digestion

The ProMTag workflow allows cleanup of intact proteins or peptides

Elution of unmodified proteins from TCO beads after extensive washing





Model proteins and a yeast lysate were labeled with ProMTag and bound to TCO beads. After extensive washing, proteins were eluted from the TCO beads by reversal of the protein-CDM linkage. This releases proteins in their original, unmodified state. After 10 minutes >80% of the proteins were released, increasing to >90% by 30 minutes.

MS analysis of a yeast proteome cleaned up with the ProMTag workflow vs a conventional in-solution digest

Conventional

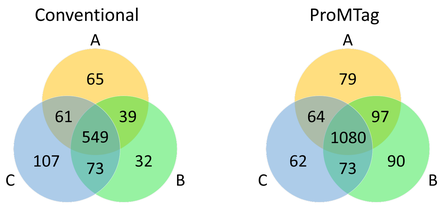
	Proteins	Peptides	Spectra
A	877	6349	14937
B	819	5902	14988
C	873	5760	15664
Total	856 ± 26	6004 ± 251	15196 ± 331

ProMTag

	Proteins	Peptides	Spectra
A	1456	13855	31214
B	1462	13982	30580
C	1435	13710	28933
Total	1451 ± 12	13849 ± 111	30242 ± 961

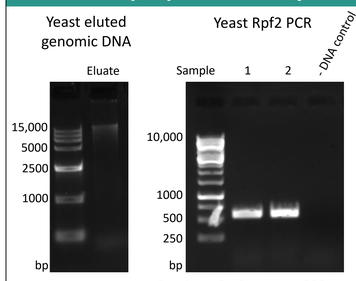
Triplicate peptide cleanups of yeast lysate were carried out with either a conventional in-solution digest or the ProMTag workflow. Peptide samples were analyzed by LC-MS/MS.

Proteins identified



Overlap of proteins identified in each triplicate based on the *p*-value score.

The Multiprep workflow yields high quality DNA suitable for PCR and WGS

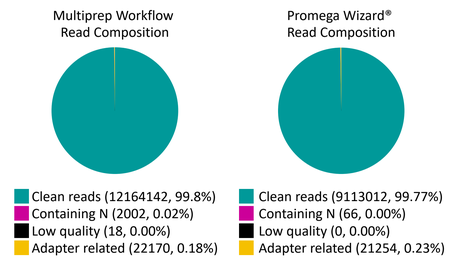


Predicted Rpf2 product: 580 bp
Yeast lysate cleanup input:
Sample 1- 20 µg
Sample 2- 100 µg

Input yeast lysate (µg)	DNA yield (ng)
20	45.6
100	256

A preliminary study was done to assess the quality of DNA recovered via the Multiprep workflow. Nucleic acids were cleaned up from yeast lysates starting with 20 µg or 100 µg of protein. This DNA was used for PCR (Rpf2 PCR shown as a representative) and amplification of the desired gene was successful.

Preparation method	Raw reads	Effective (%)	Error (%)	Q20 (%)
Multiprep workflow	12188332	99.8	0.03	97.92
Promega Wizard® kit	9134332	99.77	0.03	97.76

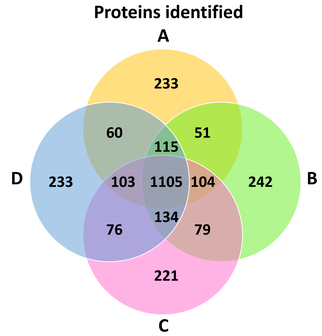


Yeast DNA was purified by the Multiprep workflow or the Promega Wizard® Genomic DNA Purification kit and sent to Novogene for whole genome sequencing. The Multiprep workflow produced high quality results that matched or surpassed the Promega kit.

UPECK Multiprep cleanup of DNA and peptides from HEK cells

MS analysis of a HEK cell proteome cleaned up with the ProMTag Multiprep workflow

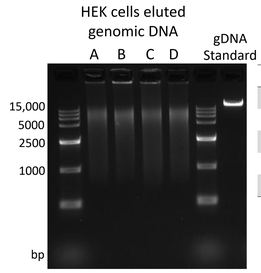
Proteins identified



Replicate	Proteins	Peptides	Spectra
A	1891	13489	25150
B	1999	13631	25730
C	1954	13415	25038
D	1989	13368	25126
Total	1958 ± 49	13475 ± 115	25261 ± 316

Quadruplicate DNA and peptide ProMTag Multiprep cleanups were carried out on a HEK cell lysate. The prepared peptides were sent for LC-MS/MS analysis. Shown is the overlap of proteins identified in each of the replicates based on the *p*-value score.

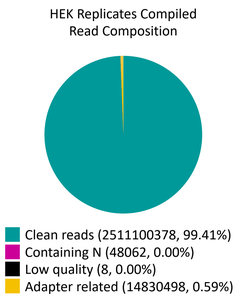
Whole genome sequencing produced high quality, reproducible results



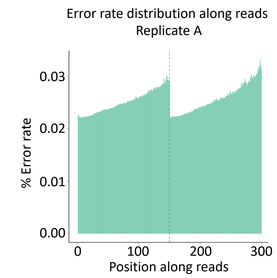
Replicate	DNA yield (ng)
A	286
B	395.2
C	483.6
D	587.6
Average	438.1 ± 128.3

HEK cell DNA was purified from 100 µg of HEK cell lysate via the Multiprep workflow in quadruplicate. These samples were sent to Novogene for whole genome sequencing. All 4 samples yielded high quality WGS data with effective %, error %, and Q20 % similar to the values obtained with our yeast samples.

Replicate	Raw reads	Effective (%)	Error (%)	Q20 (%)
A	687790614	99.38	0.03	97.51
B	615728084	99.44	0.03	97.65
C	607275118	99.36	0.03	97.48
D	615185130	99.47	0.03	97.57
Average	631494737 ± 37728885	99.41 ± 0.05	0.03 ± 0	97.55 ± 0.08



Across all replicates, 99.41% of reads were clean, while the majority of the remainder were discarded due to containing adapters.



The error rate along reads for all replicates is lowest at the beginning of the read and higher at the end, as expected. Replicate A is shown as a representative.

Conclusions

The ProMTag Multiprep workflow allows for cleanup of DNA and protein or peptide. The cleaned-up DNA and peptides are suitable for MS and WGS, allowing multiomics analysis of a single starting sample. We are currently working to expand the Multiprep workflow to allow purification of RNA as well as DNA and protein or peptides.

Acknowledgements

This work was supported by NSF SBIR Phase II Award 2036199. This work was also supported in part by NIH P30 CA124435 utilizing the Stanford Cancer Institute Proteomics/Mass Spectrometry Shared Resource at the Vincent Coates Foundation Mass Spectrometry Laboratory, Stanford University Mass Spectrometry (RRID:SCR_017801). We thank Ryan Leib and the Stanford University Mass Spectrometry Facility for their help with mass spectrometry analysis, data analysis, and statistical analysis.