DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Reference herein to any social initiative (including but not limited to Diversity, Equity, and Inclusion (DEI); Community Benefits Plans (CBP); Justice 40; etc.) is made by the Author independent of any current requirement by the United States Government and does not constitute or imply endorsement, recommendation, or support by the United States Government or any agency thereof.

Pacific Northwest
NATIONAL LABORATORY

PNNL-37377

Algorithms and file structures to extend and enhance liquid chromatography and ion mobility mass spectrometry workflows (CRADA 465) Final Report

February 2025

Aivett Bilbao Pena

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

NOTICE

Printed in the United States of America

Available to DOE and DOE contractors from the Office of Scientific and Technical Information, P.O. Box 62, Oak Ridge, TN 37831-0062
www.osti.gov
ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312 ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov

PNNL-37377

Algorithms and file structures to extend and enhance liquid chromatography and ion mobility mass spectrometry workflows (CRADA 465) Final Report

February 2025

Aivett Bilbao Pena

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Cooperative Research and Development Agreement (CRADA) Final Report

Report Date: 02/18/2025

In accordance with Requirements set forth in the terms of the CRADA, this document is the CRADA Final Report, including a list of Subject Inventions, to be provided to PNNL Information Release who will forward to the DOE Office of Scientific and Technical Information as part of the commitment to the public to demonstrate results of federally funded research. **PNNL acknowledges that the CRADA parties have been involved in the preparation of the report or reviewed the report.**

Parties to the Agreement:

PNNL/Battelle Memorial Institute

Agilent Technologies, Inc.

CRADA number: 465 (Project No. 76584)

CRADA Title: Algorithms and file structures to extend and enhance liquid chromatography and ion mobility mass spectrometry workflows

Responsible Technical Contact at DOE Lab (PNNL):

Dr. Aivett Bilbao Pena
Pacific Northwest National Laboratory/Battelle
P.O. Box 999, MSIN K8-98
Richland, WA 99352
Tel: 509-371-7741. Email: Aivett.Bilbao@pnnl.gov

Name and Email Address of POC at Partner Company(ies):

Dr. John Fjeldsted
Agilent Technologies, Inc.
5301 Stevens Creek Blvd
Santa Clara, CA 95052
Tel: 408-553-7751. Email: john_fjeldsted@agilent.com

Sponsoring DOE Program Office(s):

Not applicable – 100% funds-in CRADA; however, both Todd Anderson and Michael Riches (Office of Science) are informed of this effort.

Joint Work Statement Funding Table showing DOE funding commitment:

	Funding Amounts			
CRADA Parties	**DOE Funding**	**Funds-In**	***In-kind**	**Total**
Participant(s)		$160,000	$50,000	$210,000
DOE Funding to PNNL		N/A	N/A	
Total of all Contributions				$210,000

Federal Admin Charge on funds-in (if applicable): $3,495 under base award + $1,165 under Amendment No. 1 for a total of $4,660

Provide a list of publications, conference papers, or other public releases of results, developed under this CRADA:

Software: PNNL-PreProcessor (new release)

https://pnnl-comp-mass-spec.github.io/PNNL-PreProcessor

Publication: A preprocessing tool for enhanced ion mobility–mass spectrometry-based omics workflows. Authors: Aivett Bilbao, Bryson C Gibbons, Sarah M Stow, Jennifer E Kyle, Kent J Bloodsworth, Samuel H Payne, Richard D Smith, Yehia M Ibrahim, Erin S Baker, John C Fjeldsted. https://pubs.acs.org/doi/10.1021/acs.jproteome.1c00425.

Provide a detailed list of all subject inventions, to include patent applications, copyrights, and trademarks:

Software: 2DLC-MS to LC-IM-MS data file converter, IPIDs 32136-E (DOE S# S-166,852)

Software: IMS data file transformations, IPID 32137-E (DOE S# S-166,853)

Executive Summary of CRADA Work

The purpose of this project was to continue supporting customizations of algorithms and raw data file structures to enhance software workflows for liquid chromatography (LC), mass spectrometry (MS) and ion mobility mass spectrometry (IM-MS)-based protein and metabolite characterization. PNNL worked with Agilent to design, implement, evaluate, and demonstrate new algorithms and integrated them as functionalities into the PNNL-PreProcessor software. The project augmented PNNL's capabilities to analyze complex proteomics and metabolomics samples. These capabilities are directly beneficial to DOE and PNNL efforts to characterize and analyze these compounds in microbial and plant communities. The project assisted Agilent in further developing improved instrument-software solutions combining liquid chromatography and ion mobility with mass spectrometry for widespread applications in life sciences and other fields.

Summary of Research Results

During this project, we have implemented and evaluated several improvements in the PNNL PreProcessor software. These improvements enhance LC-IM-MS, 2D LC-MS, and LC-MS software workflows for molecular characterization in complex samples. Specifically:

- We completed a converter for parsing multi-file 2D LC-MS analyses into single LC-IM-MS data files. Besides the mode parsing multiple 2D-LC files into a single LC-IM-MS file, an additional mode was implemented to parse a single 2D-LC file using a user-specified modulation or length of the first LC separation. Initial testing of feature detection and targeted data extraction was performed resulting in detection of expected ions.
- We completed the implementation of a function for interpolation of the ion mobility dimension. Interpolation and demultiplexing for All-Ions data was also implemented.
- Per mutual agreement, the task to develop a prototype for an IM-MS/MS targeted data extraction tool was replaced by the implementation of a prototype converter from single frame 4D All Ions to 3D DDA MS/MS data file. The prototype was completed.
- We implemented improvements in speed using multi-threaded processing, GUI, and preprocessing for QTOF (3D) files.
- We implemented code to make all the algorithms accessible through the command line to support auto processing. Tests were performed to ensure proper function.
- We implemented two new functionalities in the PreProcessor for processing data from developmental ion mobility devices: compression (i.e. accumulation or sum) of frames grouped by a user-specified number and compression of drift bins grouped by a user-specified number. Improved detection of low-intensity signals was observed during evaluation of these algorithms. Initial evaluation of using longer bit sequences for multiplexing was performed.
- We implemented a data filter to keep a user-selected RT range of the data.
- We implemented the conversion of single-frame files to QTOF (3D) files with 3 options for time scaling.
- We implemented a prototype to convert single frame IM-MS files parsing the CCS as a substitute of IM with linear a non-linear functions, which can further be converted to 3D format. We also implemented a function to convert LC-IM-MS files parsing the CCS as a substitute of IM.

Per mutual agreement, the task to investigate methods for visualization of 2D-LC-IM-MS was replaced by the work to implement an initial DDA file structure for IM. An initial prototype code was developed and demonstrated

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

www.pnnl.gov