New CCS Prediction Workflow to Extend MS/MS Spectral Libraries On The Fly with CCS Information

Sofie Weinkouff[1]; Nikolas Kessler[1]; Aiko Barsch[1]; Lucy Woods[1]; Heiko Neuweger[1], Bruker Daltonics GmbH & Co. KG, Fahrenheitstraße 4, 28359 Bremen, Germany

Introduction

Confident annotation and identification of small molecule compounds remains a major challenge in Metabolomics research.

Trapped Ion Mobility Mass Spectrometry (TIMS) enables the determination of the Collision Cross Section (CCS) of ionized compounds. The CCS of compounds depends on their three-dimensional composition and is therefore an ideal additional, orthogonal qualifier for confident compound annotation, along with mass accuracy, retention time fit, isotopic pattern fit and MS/MS spectra matching.

However, many publicly available databases and MS/MS libraries still lack reference CCS values. The new automatic CCS-Predict workflow in the Spectral Library annotation presented here addresses this issue by extending MS/MS spectral libraries with predicted CCS values. These are generated on the fly using a machine learning based CCS prediction model.

Methods

- CCS-Predict, a machine learning CCS prediction tool in MetaboScape® was integrated into automated MS/MS spectral library annotation.
- To validate the workflow, human urine samples were measured with LC-TIMS-MS/MS using a timsTOF Pro 2 MS instrument (Bruker) in combination with a 23 minutes reversed phase UHPLC gradient
- Processing and annotation was performed with a preliminary version of the MetaboScape 2023 software (Bruker)
- Untargeted compound annotation was performed using the automatic CCS prediction for spectral libraries in MetaboScape with the Bruker NIST 2020 MS/MS spectral library [1]
- To confirm selected predicted CCS values, the data was additionally matched to retention times from Bruker HMDB Metabolite Library [2] and reference CCS values from the CCS Compendium [3]

Results

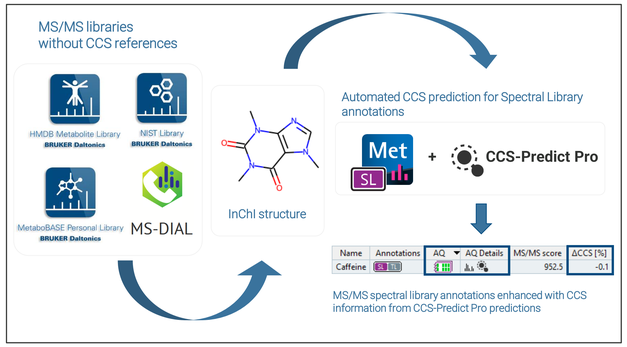

Fig. 1 The novel CCS-Predict workflow in MetaboScape turns any MS/MS spectral library with structure information into a CCS-enabled library. It automatically calculates CCS-based annotation quality scores and labels predicted values in the Annotation Quality (AQ) Details column.

- CCS-Predict is fully integrated into the spectral library annotation workflow of MetaboScape (Fig. 1).
- For the prediction of CCS values, CCS-Predict computes molecular descriptors from the input structure information and uses the CCS-Predict Pro model as a basis to predict CCS values.
- The CCS-Predict Pro model was trained on data consisting of a broad range of small molecules structures and CCS references. These were used to calculate feature vectors from molecular descriptors to perform Support Vector Regression.
- CCS-Predict is triggered automatically during spectral library annotation and first predicts a CCS value from provided structure information, second calculates a deviation between measured and predicted CCS and lastly assigns a CCS-based annotation confidence on-the fly, without modifying the original library.

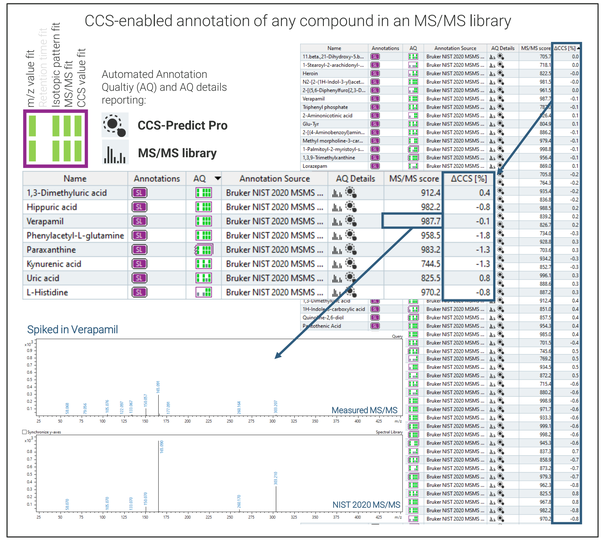

Fig. 2 MS/MS Spectral library annotations enhanced with CCS prediction and automatic annotation confidence reporting using the automatic workflow and the new machine learning model CCS-Predict Pro

- Any MS/MS spectral library with structure information can be imported into MetaboScape and used as a CCS-enabled library.
- All annotations receive automatic Annotation Quality (AQ) scoring, AQ details and CCS deviation between measured and predicted.
- Annotation using the NIST 2020 MS/MS library resulted in 145 unique annotations with relative CCS deviations below 3 % (Fig.2), which was recently reported as the acceptable deviation between measured and predicted CCS values by Das *et al.*, 2022 [4].
- 40 unique annotations based on mass accuracy, isotope pattern matching, MS/MS spectra and predicted CCS value matching showed **overall maximum annotation confidence** (mass dev. < 2 ppm, MS/MS score > 900, mSigma < 20, CCS deviations measured vs. prediction < 2 %).
- Tentative annotations of L-Histidine and Paraxanthine as well as spiked in Verapamil were confirmed with both reference retention times from HMDB and CCS references from CCS Compendium.

Summary

Here, we present a novel workflow for automated CCS prediction from any MS/MS spectral library with structure information for additional confidence in large-scale small molecule annotations.

References

[1] https://store.bruker.com/products/bruker-nist-mass-spectral-library

[2] https://www.bruker.com/en/products-and-solutions/mass-spectrometry/ms-software/metabolomics-spectral-libraries.html

[3] https://doi.org/10.1039/C8SC04396E

[4] https://doi.org/10.1021/jasms.1c00315

Conclusion

- Novel workflow for automated CCS prediction based on machine learning for MS/MS spectral library annotations
- Provides CCS-enabled annotation of any user-created or publicly available MS/MS library.
- This new annotation workflow unlocks higher annotation confidence for existing MS/MS libraries even in the absence of CCS reference values.

4D-Metabolomics