# Predicting mass spectra using theoretically derived graph transformation rules

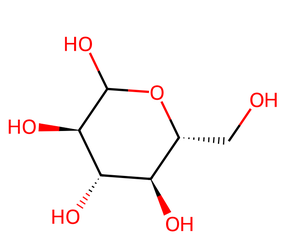
Poster-Link

Manuel Uhlir MSc BSc BSc, ao. Univ.-Prof. Mag. Dr. Christoph Flamm
Institute of theoretical chemistry, University of Vienna, Austria

## Motivation

Mass Spectrometry (MS) is a powerful tool in analytical chemistry, but its reliance on reference spectra limits its ability to identify unknown compounds. This limitation hinders MS applications in areas like drug discovery, environmental analysis, and metabolomics, where accurate identification is crucial.[1] Current spectral prediction tools are often inaccurate, particularly with complex molecular structures.[2]

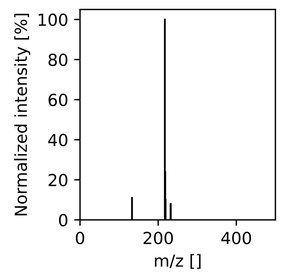
**Structure**



Source: Openbabel

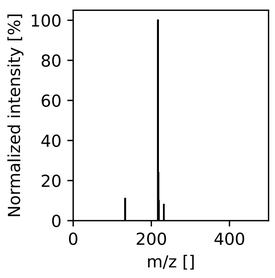
Predict

**Mass Spectrum**




Source: Chemspider, Matplotlib

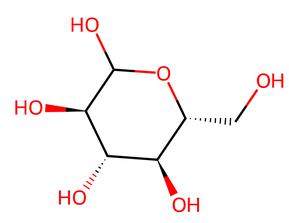
**Mass Spectrum**




Source: Chemspider, Matplotlib

Identify

**Structure**



Source: Openbabel

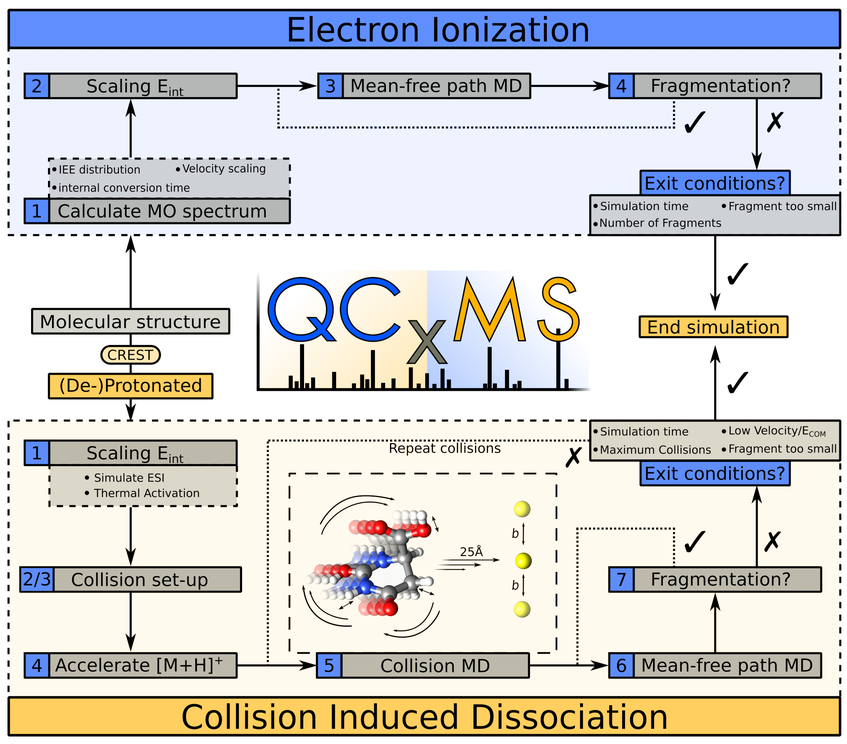
## Methodology

### Preprocessing

The molecule is prepared for analysis by identifying the most probable molecular ion, adjusting protonation if necessary, selecting the optimal protomer using CREST, and calculating the molecule's ground state.[3]

### Simulation of fragmentation trajectories

Fragmentation trajectories are simulated using the QCxMS package, employing quantum chemical methods to model both Electron Ionization (EI) and Collision Induced Dissociation (CID) under specified experimental conditions.



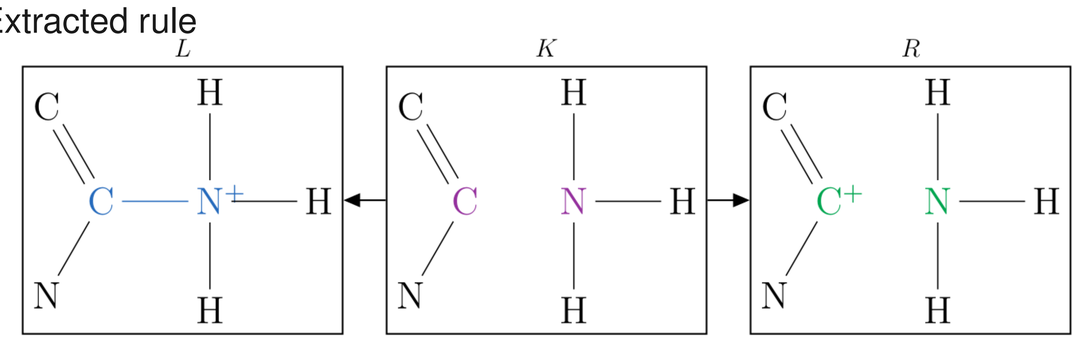

Source: QCxMS [4]

### Analysis of fragmentation trajectory

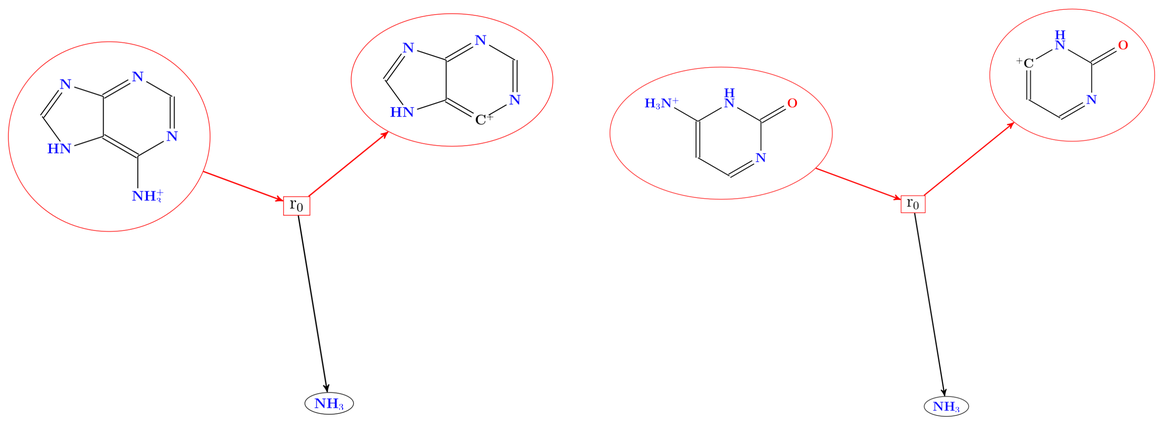
The fragmentation trajectories which consist of coordinate files are then analyzed using our analysis pipeline. This allows for both the extraction of the molecular structure of the fragments and the fragmentation chemistry in the form of double pushout reaction rules.

## Applying the learned rules

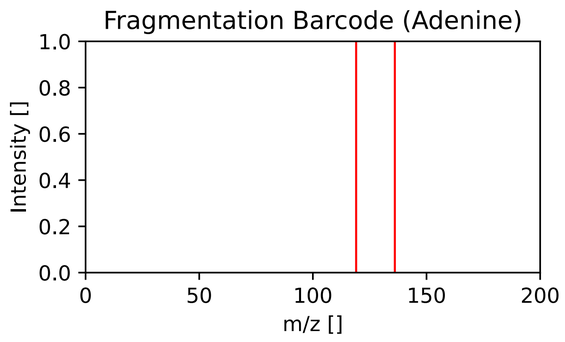
Extracted rule

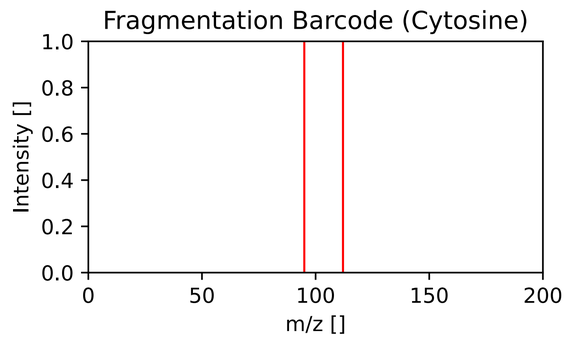



Rules are applied using mød.[5]



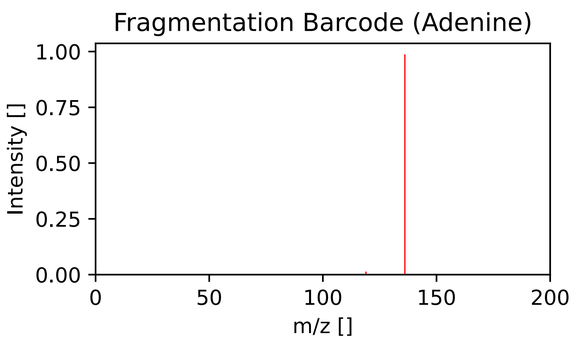
This results in a fragmentation barcode.

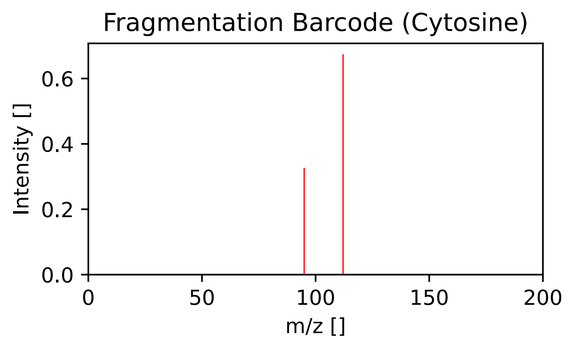



## Estimating fragmentation frequencies

We can train a machine learning model, trained on reference spectra, to estimate the probability of applying specific fragmentation rules based on the molecular context. This allows for pruning of the fragmentation tree when the probability is low and also to estimate the frequency of a given fragment by multiplying the resulting probabilities.




## Summary and Outlook

This project aims to enhance the interpretation of fragmentation trajectories in mass spectrometry. We will achieve this by comparing generated fragmentation frees with SIRIUS annotation trees, extracting rules from various predictors, and estimating the probability of applying these rules in specific contexts. Additionally, we will experiment with machine-learned force fields to improve the accuracy and speed of the quantum chemical simulations in predicting molecular fragmentation. This comprehensive approach will provide a deeper understanding of the underlying mechanisms driving fragmentation processes.

## KEY REFERENCES

[1] Augustin Scalbert et al. DOI: 10.1007/s11306-009-0168-0. URL: https://doi.org/10.1007/s11306-009-0168-0.

[2] Sebastian Böcker et al. DOI: 10.1093/bioinformatics/btn603. URL: https://doi.org/10.1093/bioinformatics/btn603.

[3] Philipp Pracht, Christoph Alexander Bauer, and Stefan Grimme. DOI: https://doi.org/10.1002/jcc.24922. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.24922. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.24922.

[4] Christoph Alexander Bauer and Stefan Grimme. DOI: 10.1021/jp5096618. URL: https://doi.org/10.1021/jp5096618.

[5] Jakob L. Andersen et al. "A Software Package for Chemically Inspired Graph Transformation". In: Springer International Publishing, 2016, pp. 73–88. ISBN: 0302-9743. DOI: 10.1007/978-3-319-40530-8_5. URL: https://dx.doi.org/10.1007/978-3-319-40530-8_5.

## ACKNOWLEDGEMENT AND PARTNERS

universität wien
MATOMIC
novo nordisk foundation
tbi

## MORE INFORMATION

**Uhlir Manuel**
University of Vienna,
Faculty of chemistry,
Institute for theoretical chemistry,
Theoretical Biochemistry Group,

manuel.uhlir@univie.ac.at

https://github.com/um-univie/