Unravelling lignin structure with mass spectrometry and rule-based modelling

Eliise Tammekivi[a], Suthawan Muangmeesri[a], Joseph S. M. Samec[a], Anneli Kruve[a]
[a]Department of Chemistry, Stockholm University, Svante Arrhenius väg 16C, 114 18 Stockholm, Sweden
eliise.tammekivi@su.se

1 Background

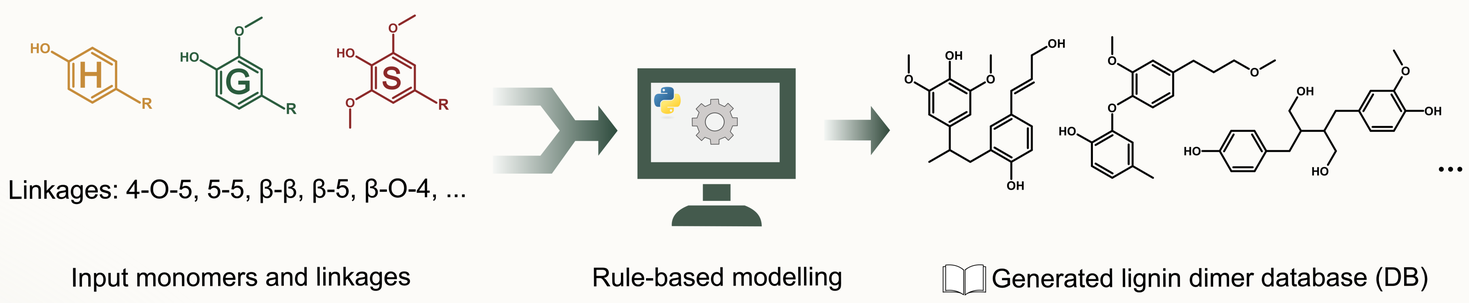
Lignin, a by-product of the paper and ethanol industries, represents a valuable resource for the production of biomaterials. A better understanding of the structures formed during depolymerisation could support more efficient lignin valorisation.[1]

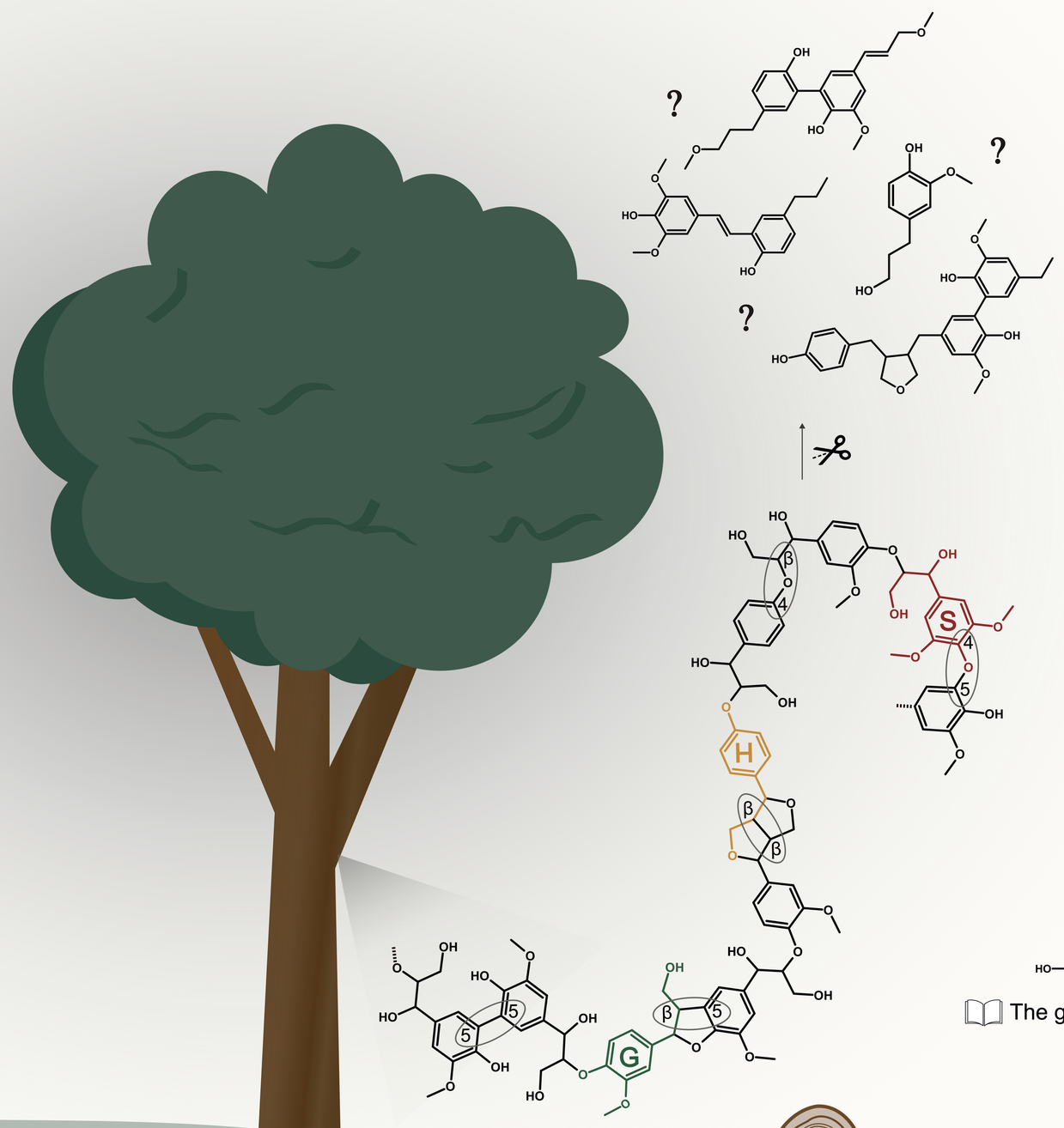
Structural characterisation of depolymerised lignin with LC/HRMS is hampared by the absence of lignin dimers, trimers and larger oligomers from databases.[2]

In this project, we aim to expand the structural databases by generating lignin oligomer structures *in silico*.

2 Structure generator

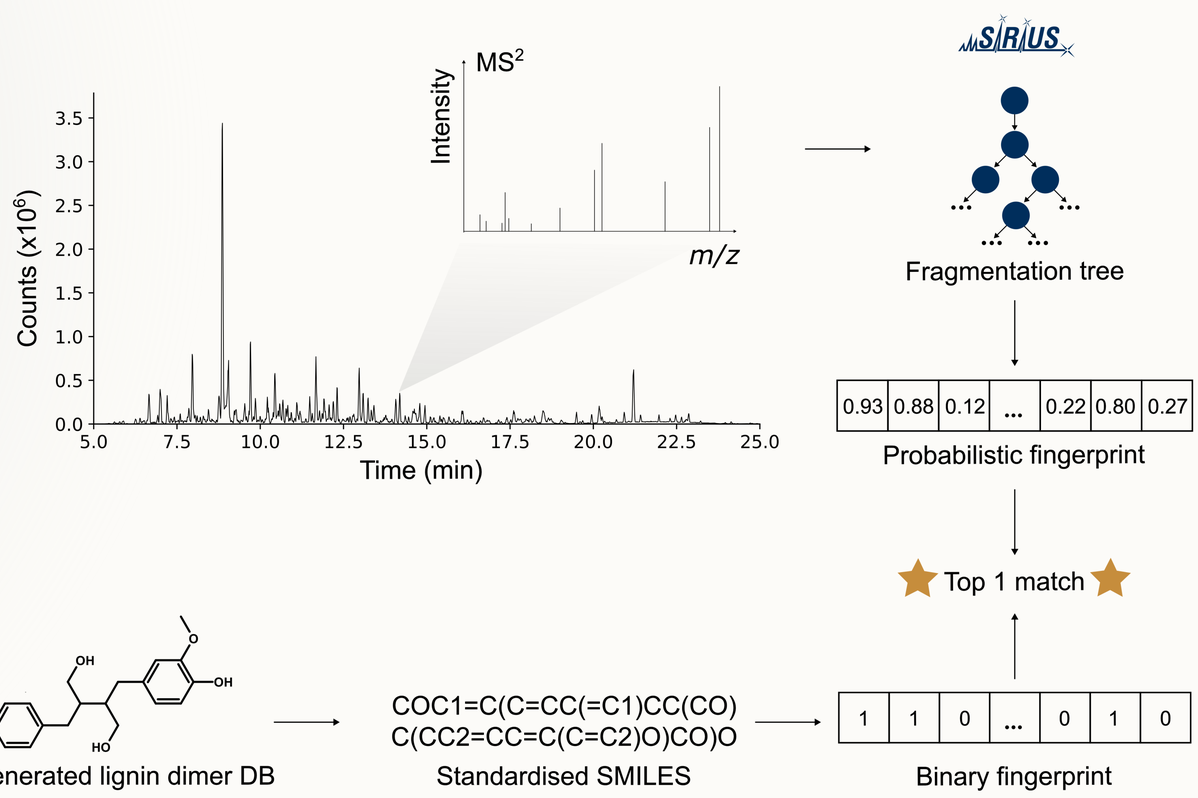
The reactions taking place during the reductive catalytic fractionation (RCF) of lignin were compiled.[3] Based on these reactions, a Python code utilising RDKit was written to generate lignin dimer structures.
ChatGPT (OpenAI, 2025) was used for coding assistance.





3 LC/HRMS^2 analysis

The MS^2 spectra were processed with SIRIUS+CSI:FingerID v. 6.1.0 for tentative structural annotation.[4] The generated lignin DB and publicly available structural databases, including PubChem, were utilised.

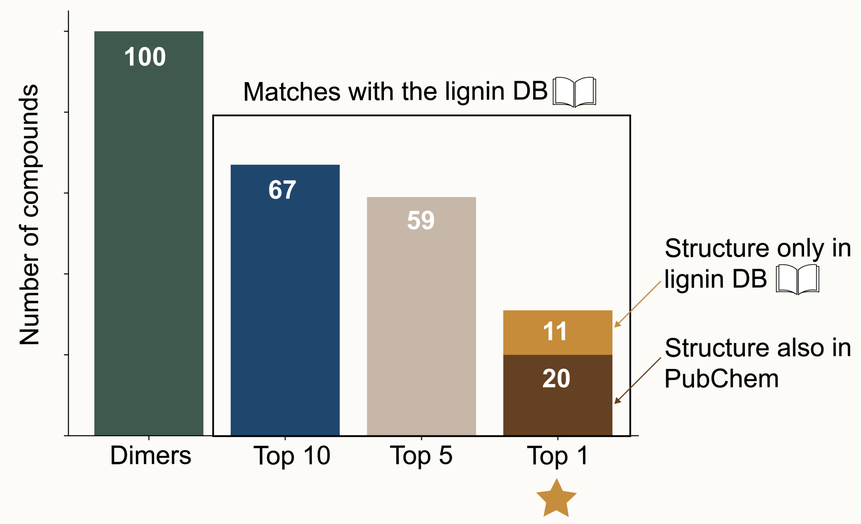


4 Results

The generated lignin dimer DB contains 1720 compounds with 117 molecular formulas. From these, only 195 were found in PubChem.

For the depolymerised lignin sample, 331 compounds were detected using DDA.

Out of the 100 potential dimers, 11 had a SIRIUS match with a structure only present in the lignin DB.



5 Conclusions

Lignin structural DB was created *in silico*, containing structures absent from PubChem.

The lignin DB can be imported to SIRIUS for tentative identification of lignin oligomers.

Next, the lignin DB will be expanded to include larger oligomers and structures generated by other depolymerisations.

[1] S. Bertella, J. S. Luterbacker, Trends Chem. 2 (5) (**2020**) 440-453.
[2] E. Tammekivi et al., Anal. Chim. Acta. 1288 (**2024**) 342157.
[3] H. Dao Thi et al., Green Chem. 24 (**2022**) 191-206.
[4] K. Dührkop et al., Nat. Methods 16 (4) (**2019**) 299-302.

Acknowledgment
The project received funding from the Wenner-Gren Foundations project nr UPD2023-0118 and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program grant agreement No. 101124488, LearningStructurE: Machine Learning and Mass Spectrometry for Structural Elucidation of Novel Toxic Chemicals.

Kruve Lab