

Stockholm University

Kruvelab.com

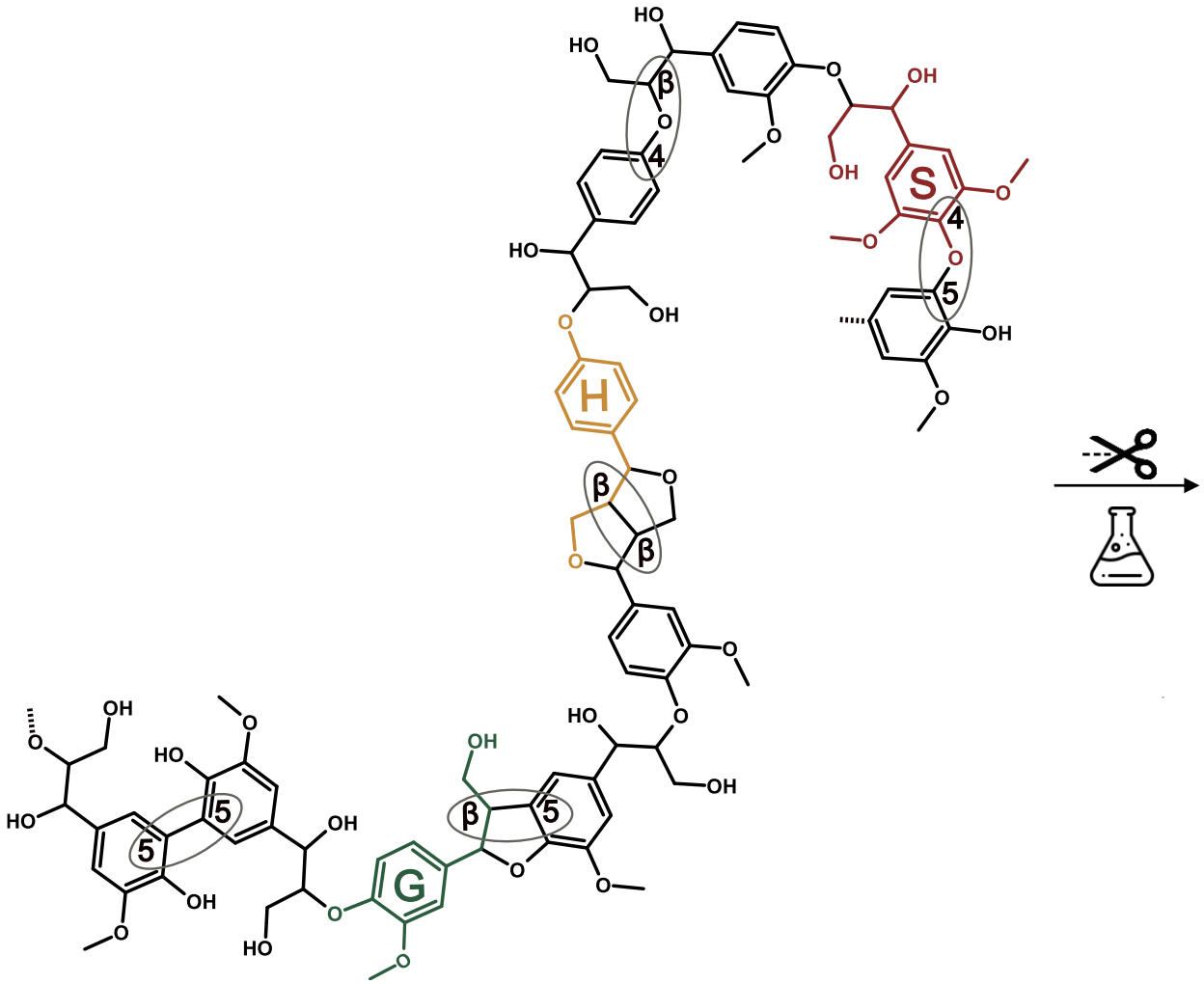
Lignin GRC: Lignin Characterisation

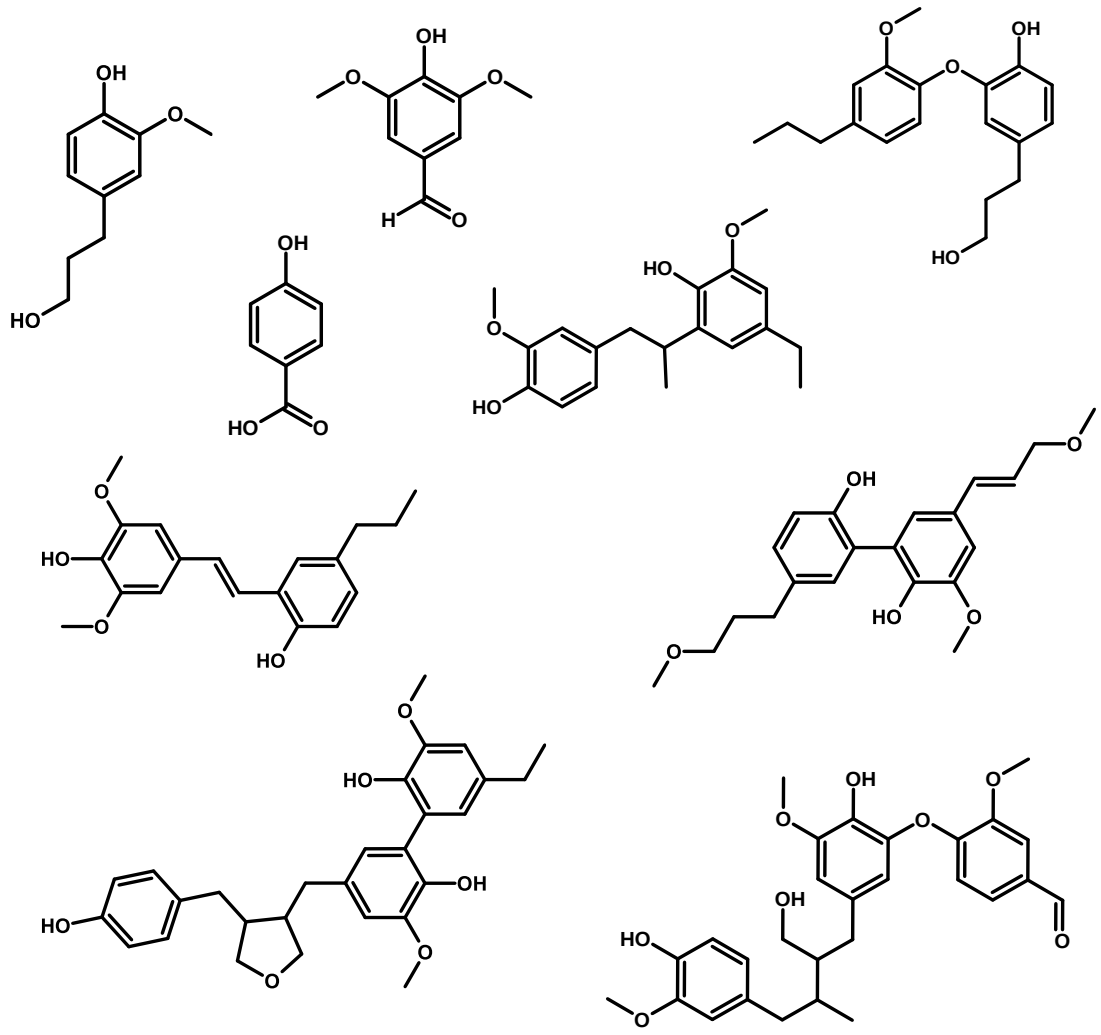
Unravelling the Structures of Depolymerised Lignin Compounds via Mass Spectrometry and Rule-Based Modelling

Eliise Tammekivi, Suthawan Muangmeesri, Joseph S. M. Samec, Anneli Kruve

Lignin GRC, 19 May 2026

DEPOLYMERISED LIGNIN



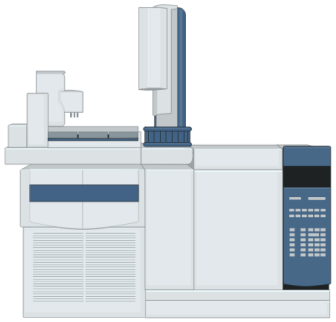


- **Hundreds to thousands** of compounds
- Ranging from monomers to oligomers
- **Structurally highly similar** compounds - including numerous **isomers**

ANALYSIS



Derivatisation



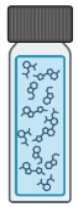
Gas Chromatography – Mass Spectrometry (GC/MS)

+

- Confident identification of monomers and smaller oligomers

-

- Needs derivatisation
- Not suitable for low-volatile or thermally labile compounds



Dissolution



Liquid Chromatography – High-Resolution Mass Spectrometry (LC/HRMS)

+

- Low-volatile and thermally labile compounds can be directly analysed

-

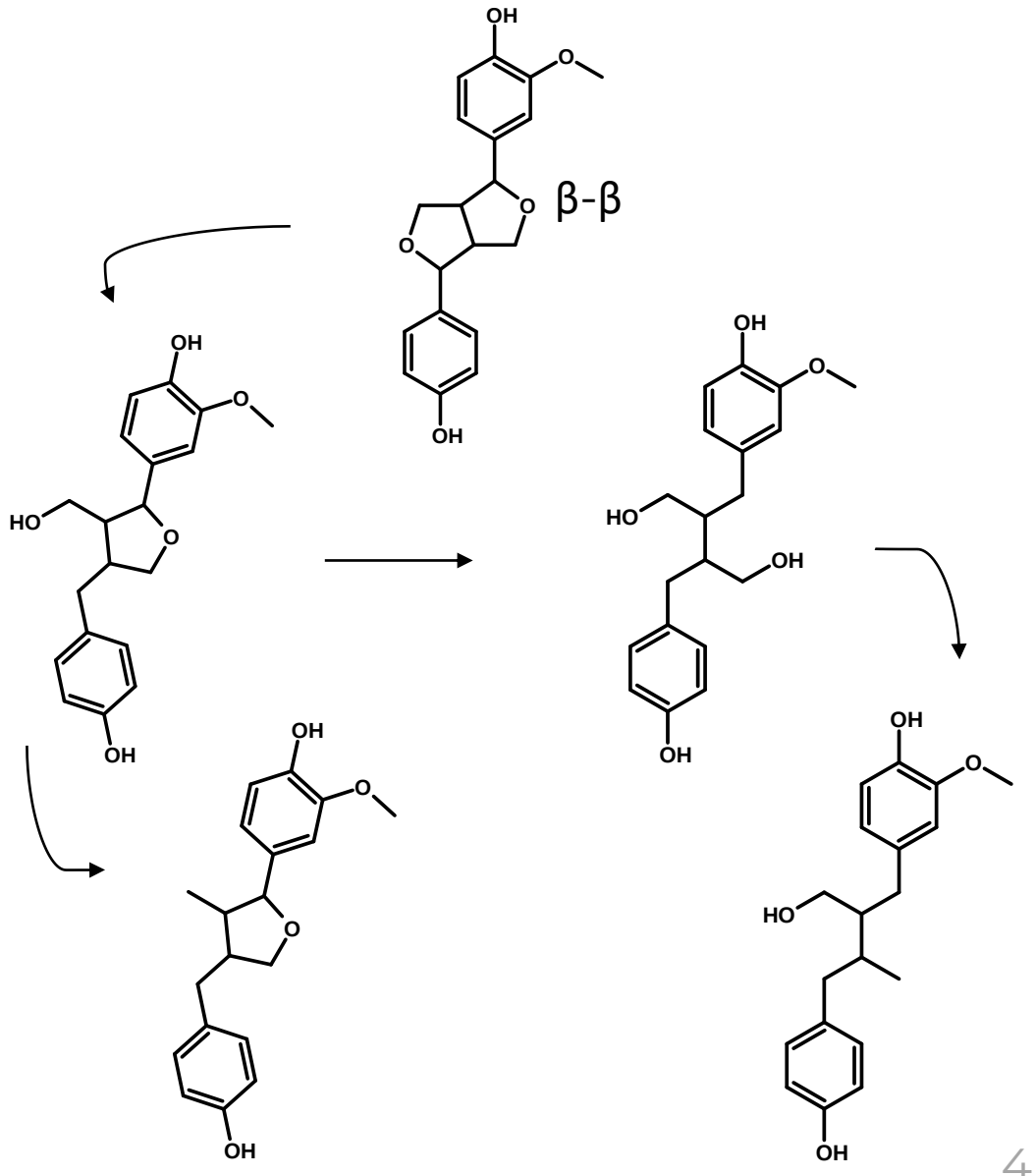
- Identification is the bottleneck

REDUCTIVE CATALYTIC FRACTIONATION (RCF)

- A lignin-first extraction and depolymerisation method
- The occurring reactions are quite well known

Reactions occurring under reductive conditions

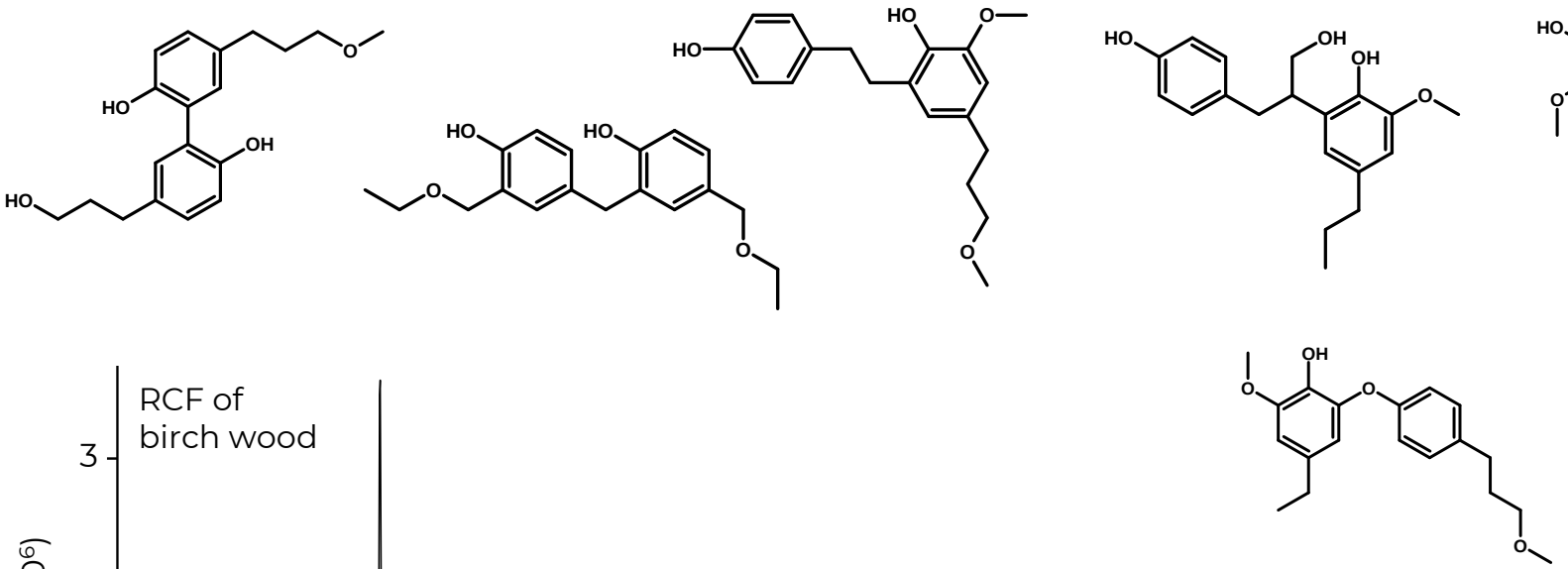
Reaction	Scheme
hydrogenation	R_1 + H_2 → R_1
hydrogenolysis	R_1–O–R_2 + H_2 → R_1–OH + CH_3–R_2
decarboxylation	R_1–C(=O)–OH + H_2 → R_1–H + CO_2
demethoxylation	R_1–O + H_2 → R_1–H + CH_3–OH
demethylation	R_1–O + H_2 → R_1–OH + CH_4
hydrodeoxygenation	R_1–OH + H_2 → R_1–H + H_2O

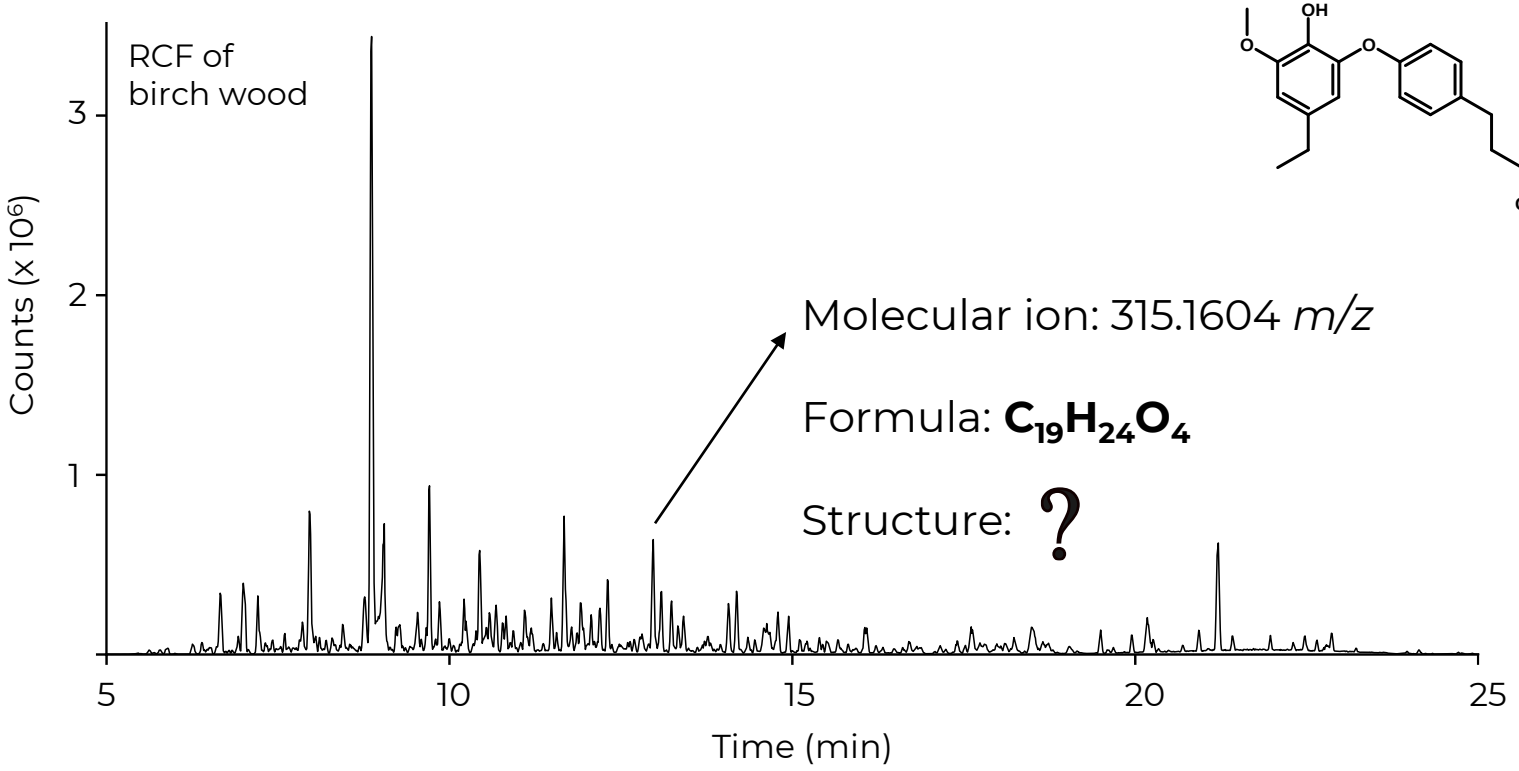


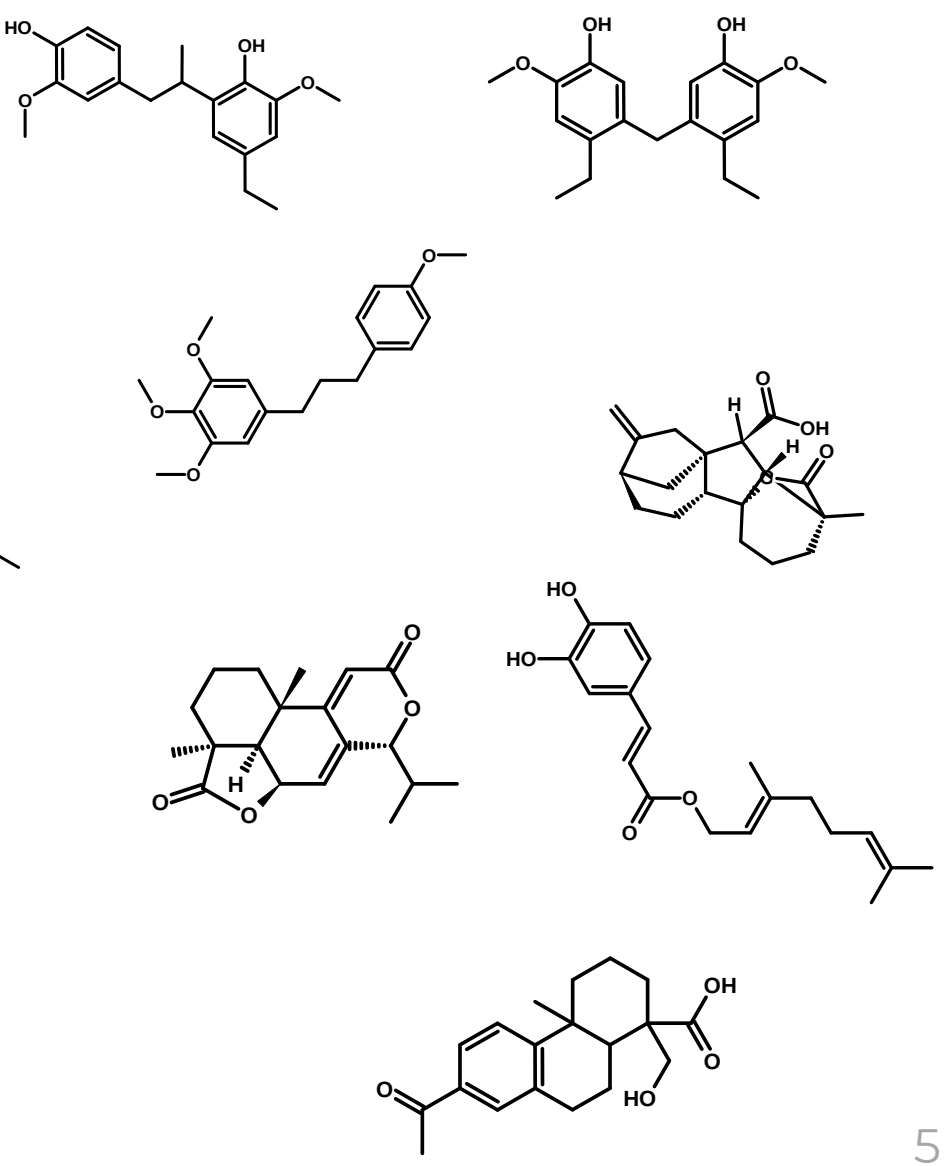
Bourbiaux et al., *Catalysis Today* (**2021**)

LC/HRMS IDENTIFICATION

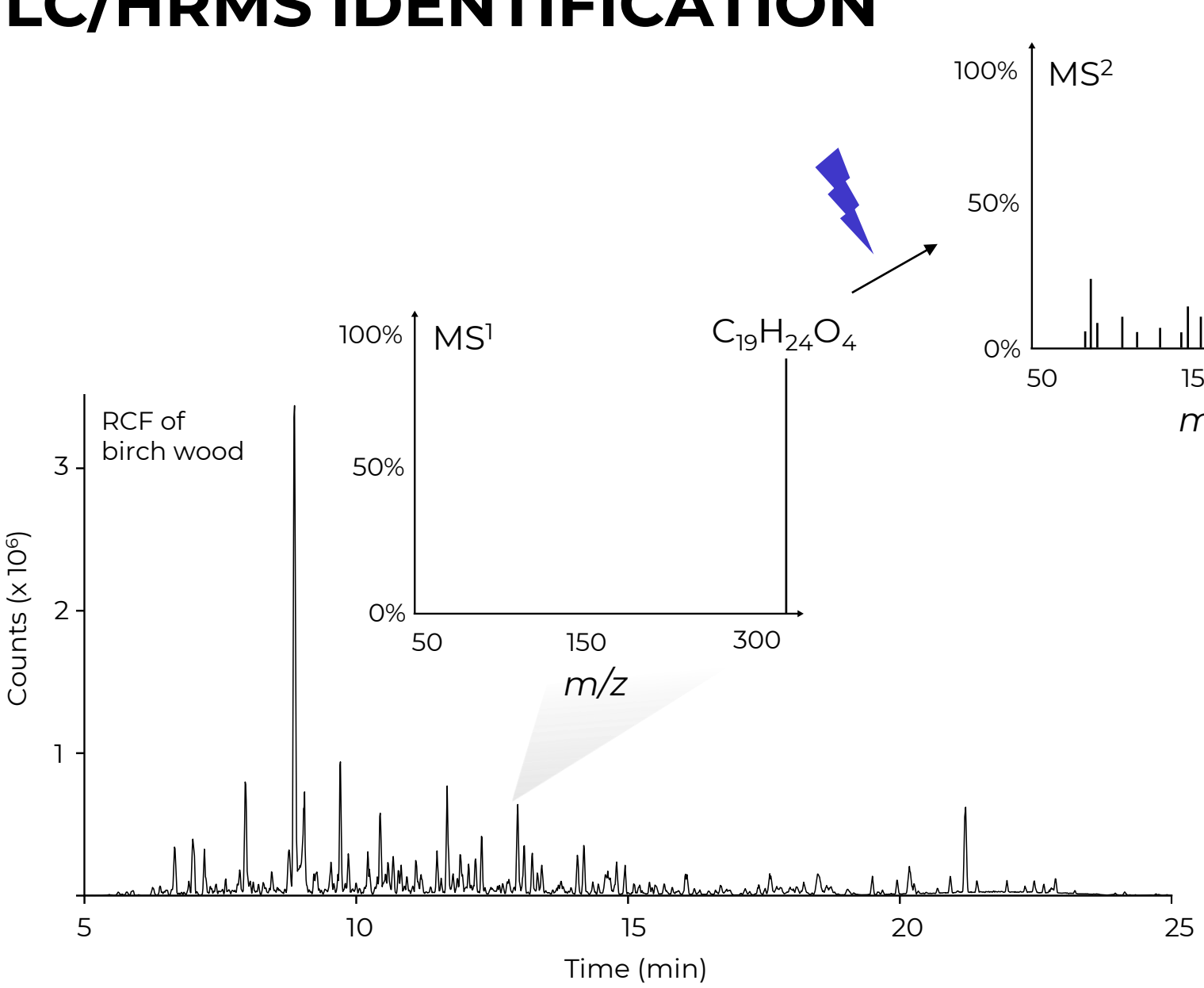
PubChem: 3,069 structures





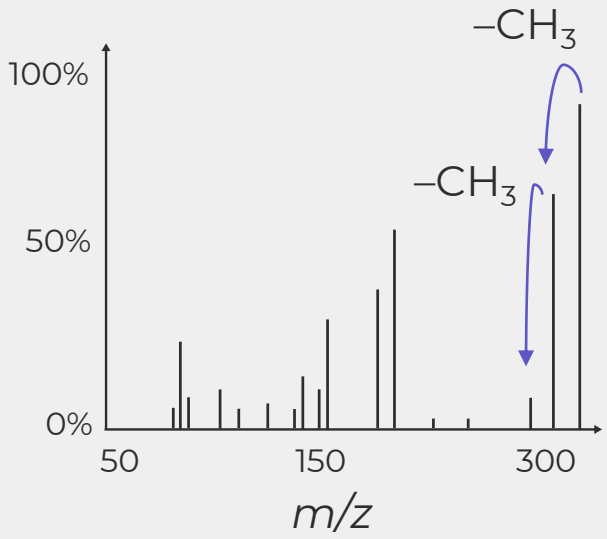


LC/HRMS IDENTIFICATION

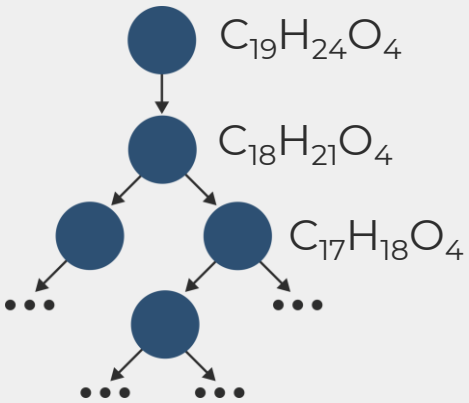


- MS library matching
  In-house library, MassBank, GNPS
- Manual annotation
- *In-silico* tools
  SIRIUS, CFM-ID, MetFrag

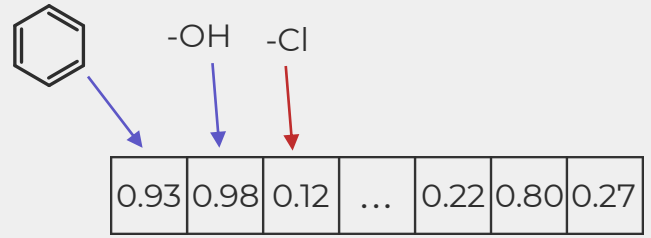
SIRIUS+CSI:FINGER ID



Fragmentation tree

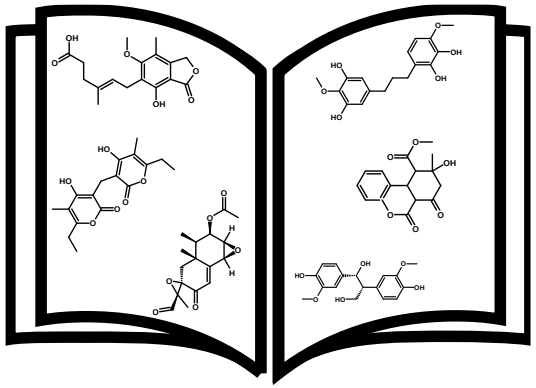


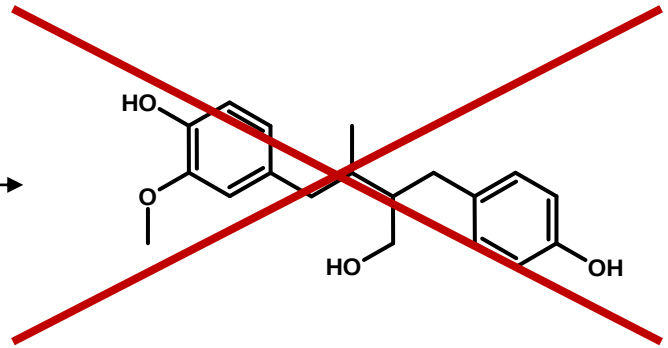
Probabilistic fingerprint



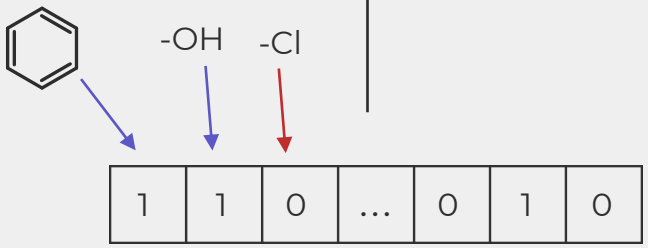
Top 1 match

PubChem, COCONUT, GNPS, LOTUS, Plantcyc, PubMed ...





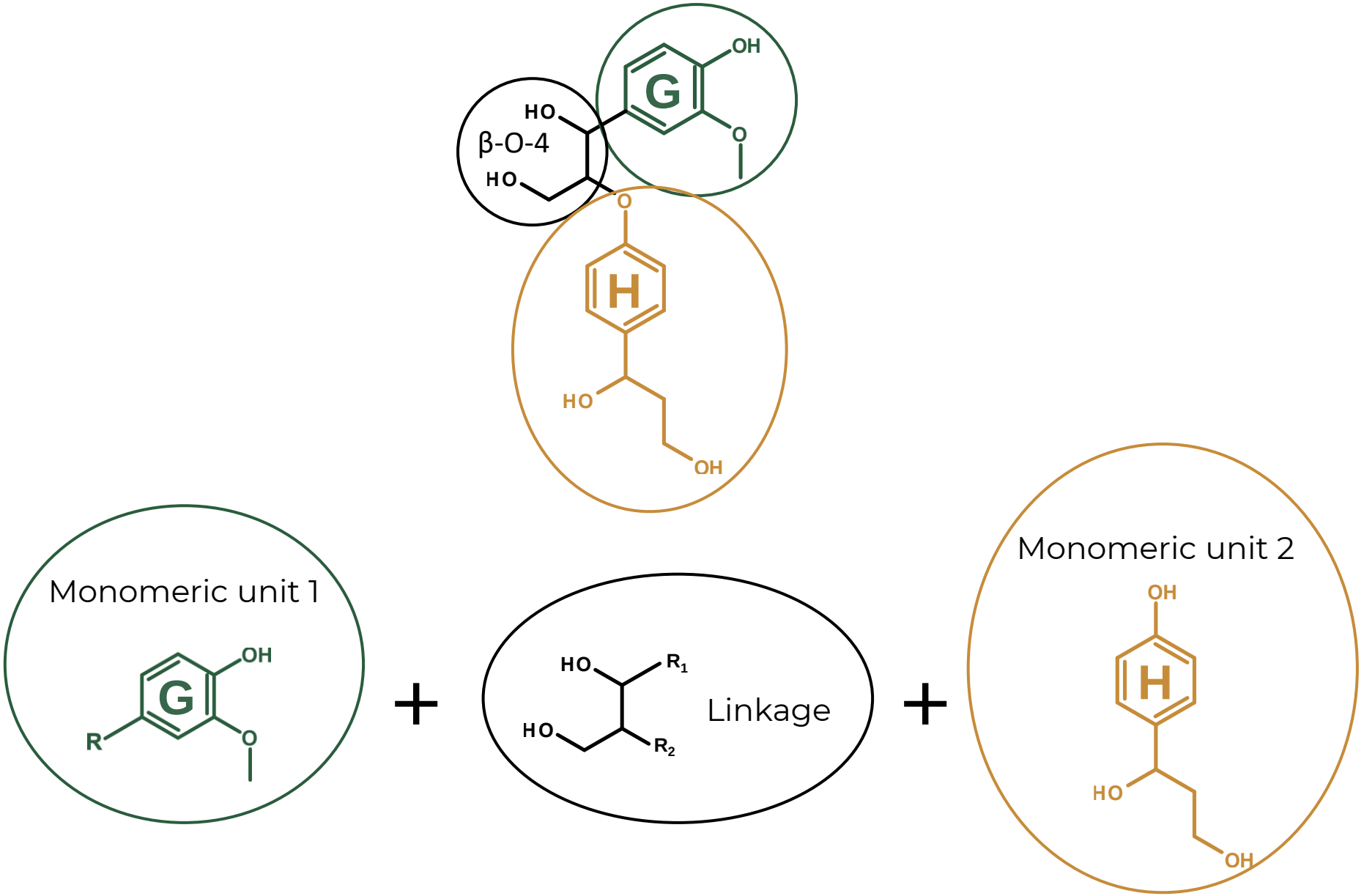
Oc1ccc(CC(C)C(CO)Cc2ccc(O)cc2)cc1OC



Binary fingerprint

Dührkop et al., *Nat. Methods* (**2019**)

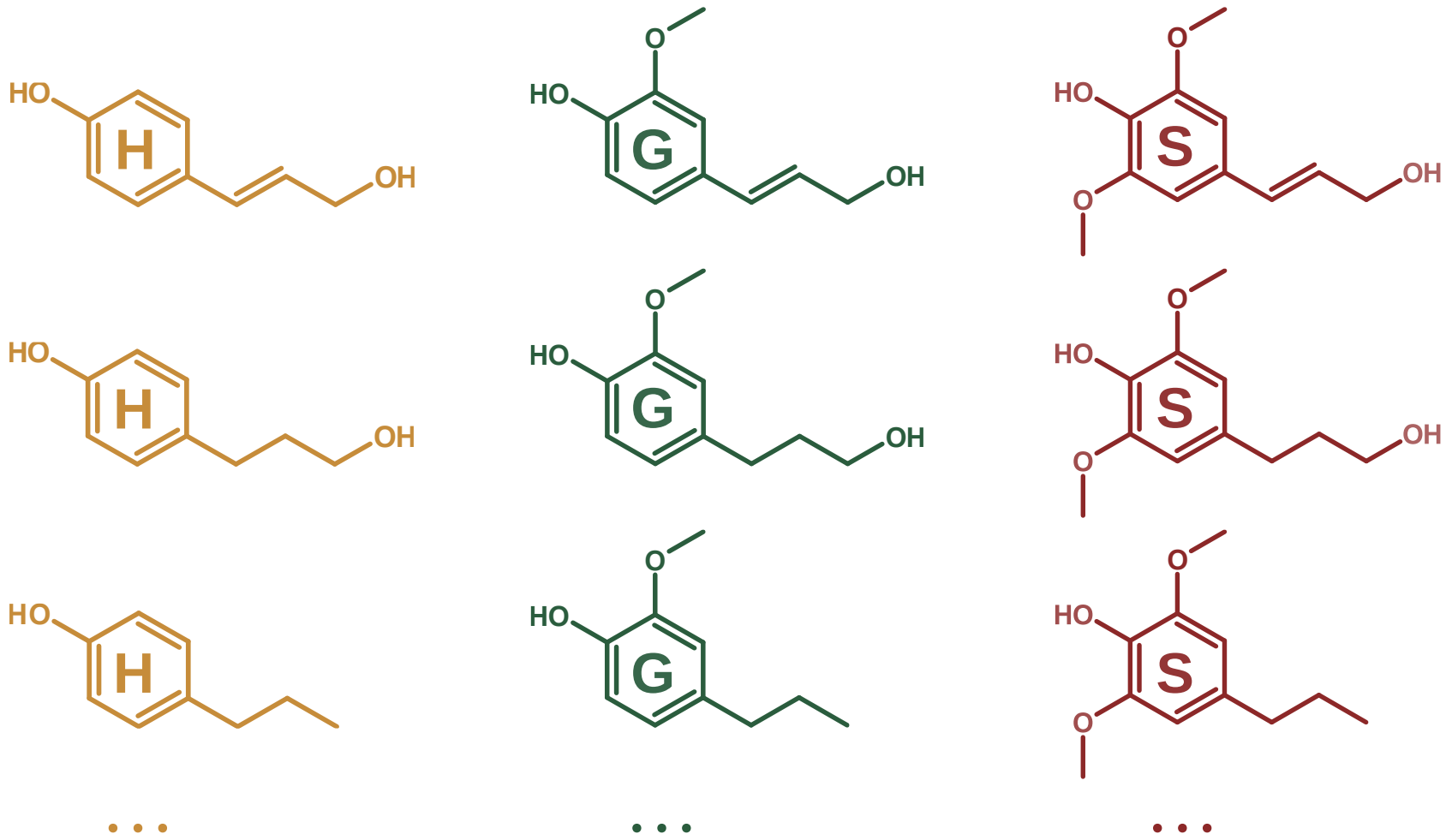
GENERATING LIGNIN DIMER STRUCTURES



GENERATING LIGNIN DIMER STRUCTURES

- Monomeric units in the potential RCF lignin dimers

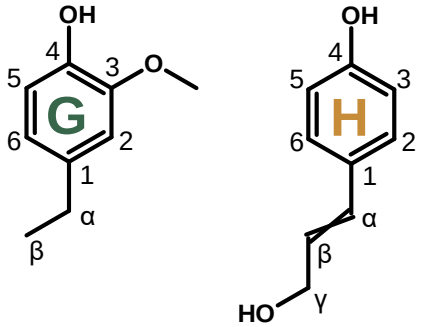
In total: **26 monomeric units**



Dao Thi et al., *Green Chem.* (**2022**)
Aelst et al., *Chem. Sci.* (**2020**)
Anderson et al., *Nat. Commun.* (**2019**)

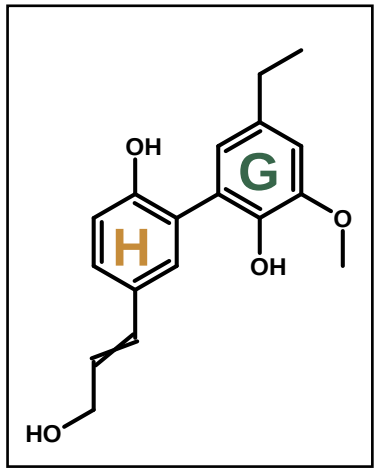
GENERATING LIGNIN DIMER STRUCTURES

Exemplary monomeric units to show the linkages

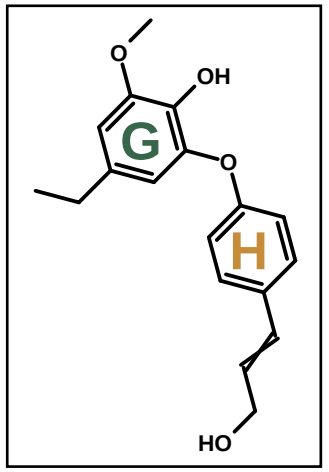


In total: **33 linkages**

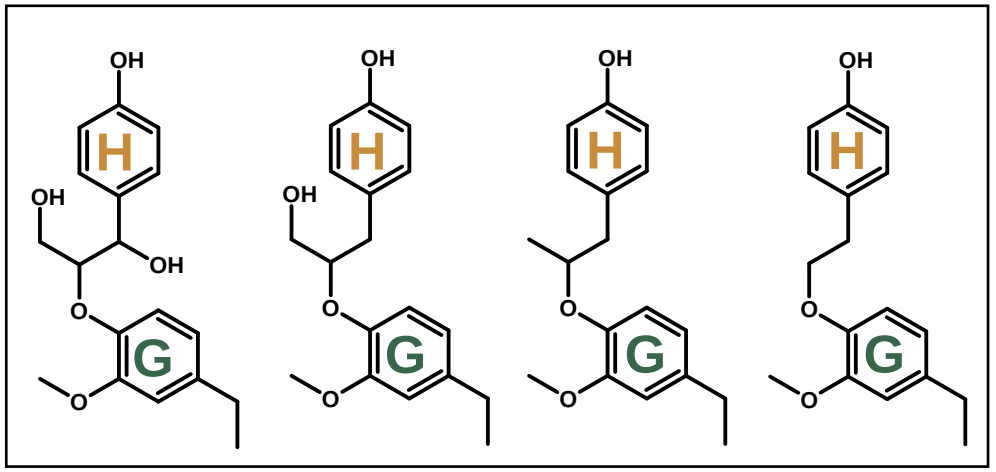
5-5: 1 linkage



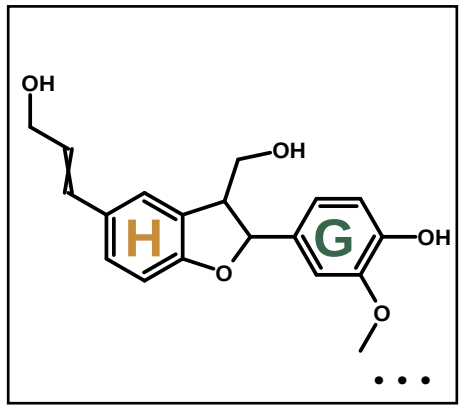
4-O-5: 1 linkage



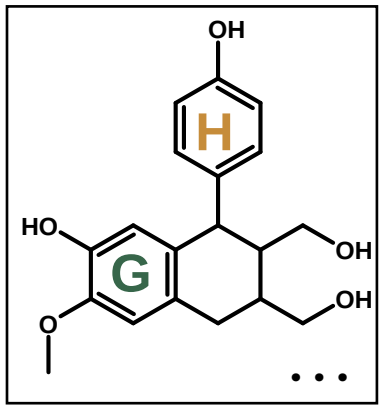
β-O-4: 4 linkages



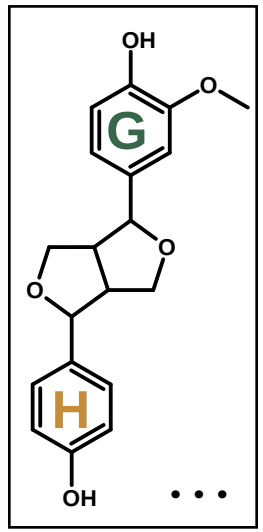
β-5:
6 linkages



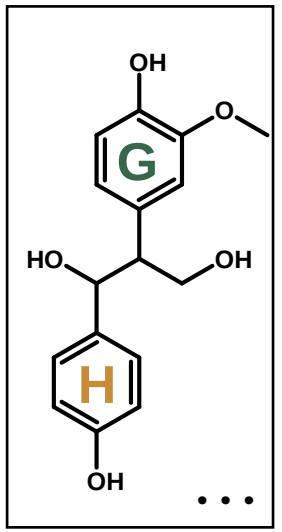
β-β α-2:
8 linkages



β-β:
9 linkages



β-1:
4 linkages

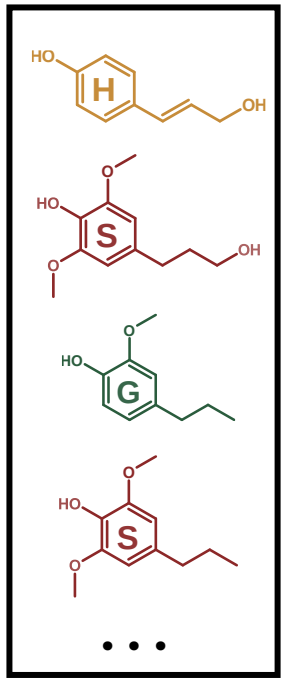


Dao Thi et al., *Green Chem.* (**2022**)
Aelst et al., *Chem. Sci.* (**2020**)
Anderson et al., *Nat. Commun.* (**2019**)

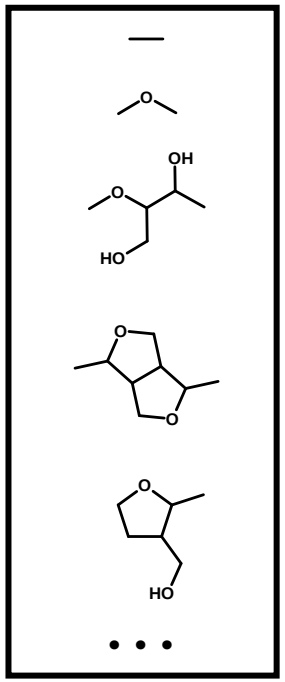
GENERATING LIGNIN DIMER STRUCTURES

- Python script using **rule-based modelling**
- Implementing RDKit package

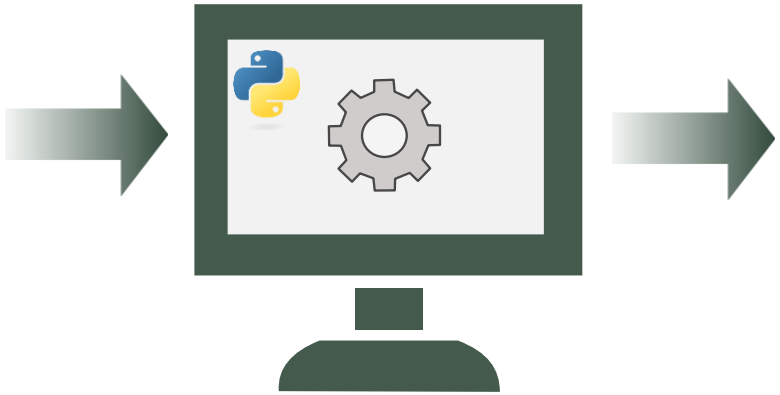
26 monomers



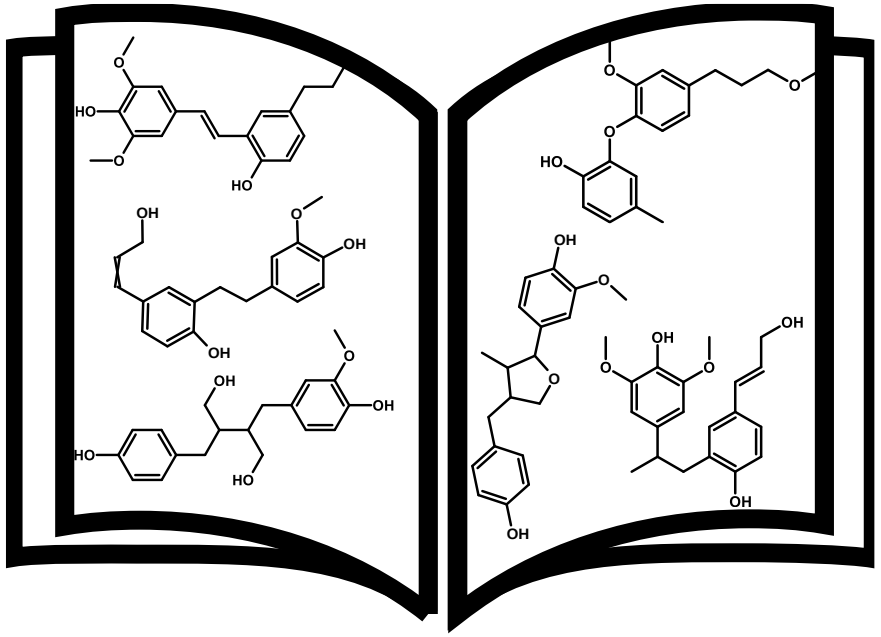
33 linkages



Generating all possible combinations



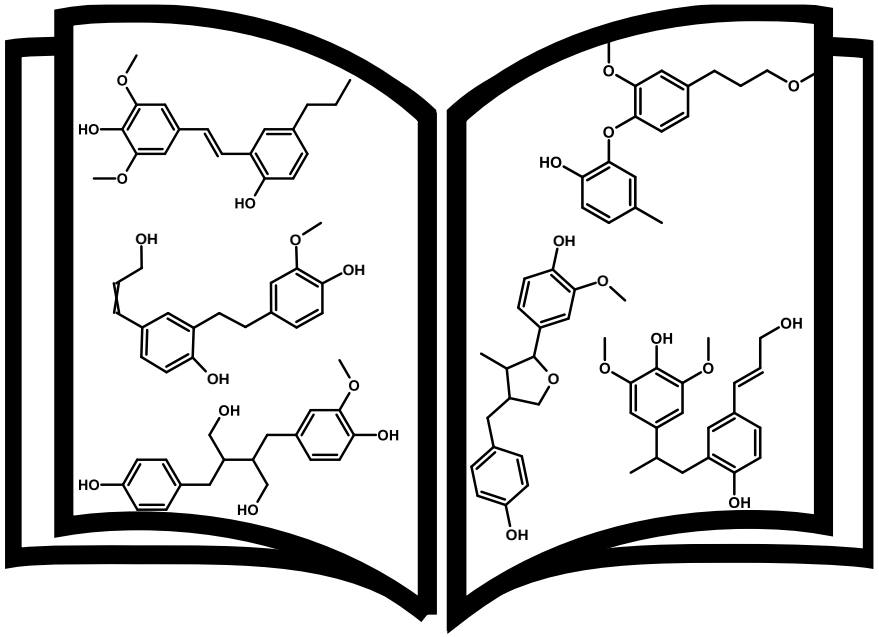
Library of generated RCF **lignin dimer structures**



GENERATING LIGNIN DIMER STRUCTURES

- Contains **1945 dimer** structures
- Structures of **potential** lignin RCF dimers
- Only **117 molecular formulas**
- Only **228 (11.7%)** of them are in **PubChem**
- **SMILES** for each structure

Library of generated RCF **lignin dimer structures**



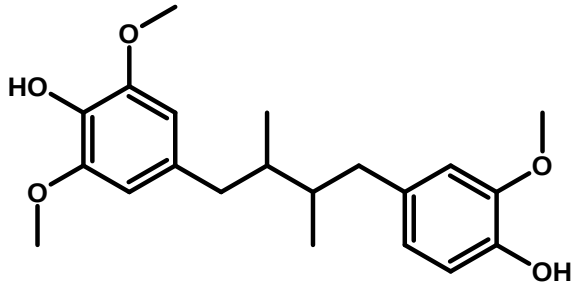
Generated Dimer	ID	Bond Type	Monomer 1	Monomer 2
C1=CC(=CC=C1CCC2=CC=C(C=C2)O)O	ID-1	B1-27	Oc1ccc(cc1)C=CCO	Oc1ccc(cc1)C=CCO
COC1=C(C=CC(=C1)CCC2=CC=C(C=C2)O)O	ID-2	B1-27	Oc1ccc(cc1)C=CCO	Oc1ccc(cc1OC)C=CCO
COC1=CC(=CC(=C1O)OC)CCC2=CC=C(C=C2)O	ID-3	B1-27	Oc1ccc(cc1)C=CCO	COc1cc(cc(OC)c1O)C=CCO
C1=CC(=CC=C1CCC2=CC(=C(C=C2)O)O)O	ID-4	B1-27	Oc1ccc(cc1)C=CCO	Oc1cc(C)ccc1O
COC1=CC(=CC(=C1O)O)CCC2=CC=C(C=C2)O	ID-5	B1-27	Oc1ccc(cc1)C=CCO	Oc1cc(C)cc(OC)c1O

ANNOTATION OF RCF DIMERS

Lignin DB

SIRIUS results for compound **A**

Top 1

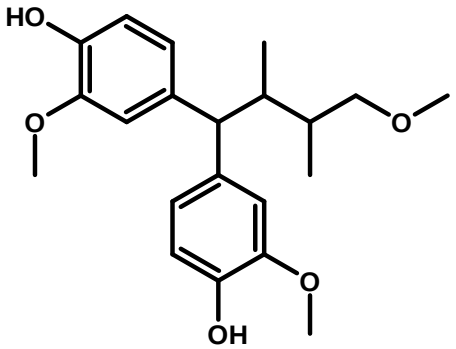


Sources:

Lignin DB

PubChem CHEBI

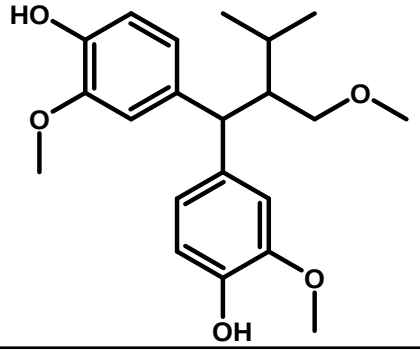
Top 2



Sources:

PubMed PubChem

Top 3

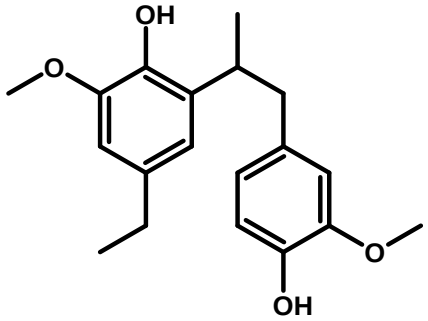


Sources:

PubMed PubChem

SIRIUS results for compound **B**

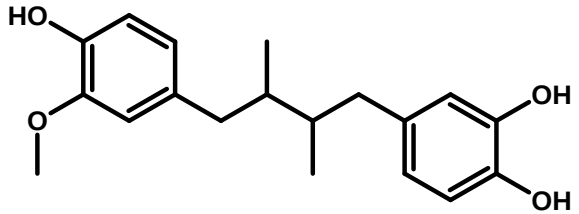
Top 1



Sources:

Lignin DB

Top 2

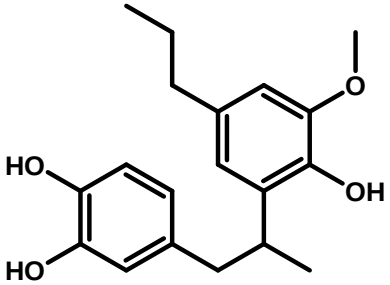


Sources:

Lignin DB

SuperNatural PubChem COCONUT

Top 3

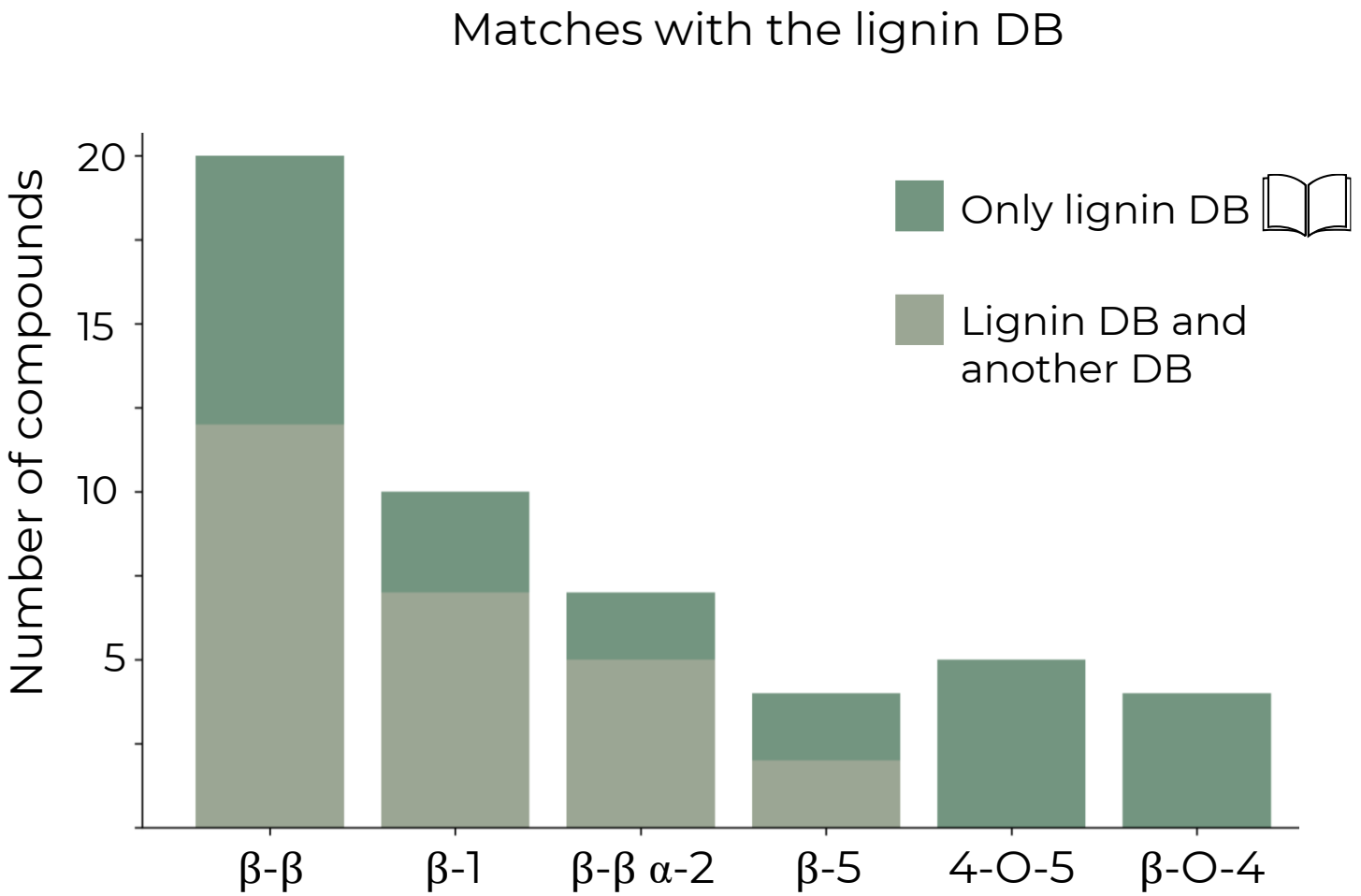


Sources:

Lignin DB

HIGH-RANKING STRUCTURAL MATCHES

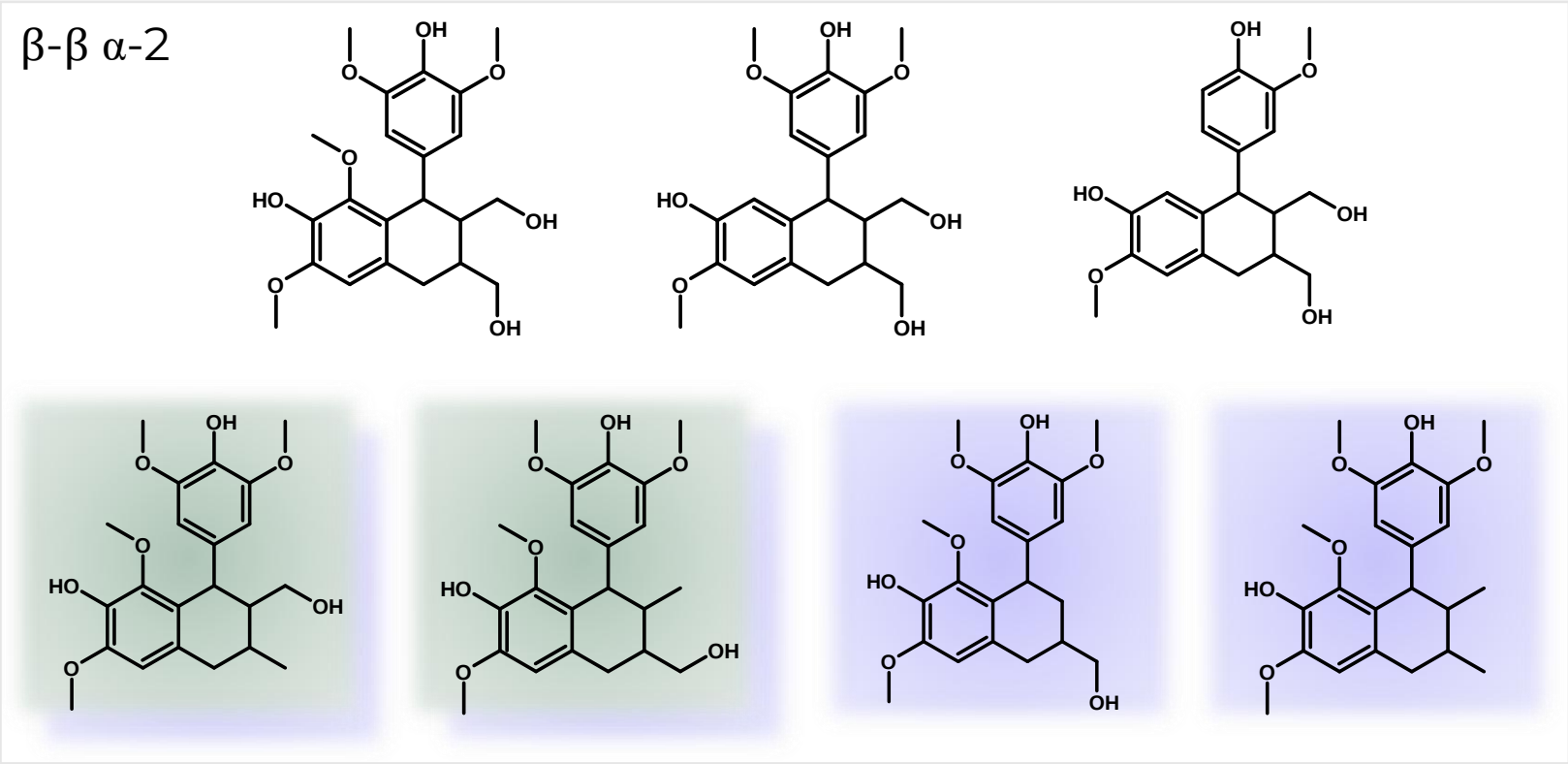
- 102 potential dimers
- High ranking structural match for 50 dimers
- All of the structures consist of G or S units, no H units
  - Consistent with birch wood



HIGH-RANKING STRUCTURAL MATCHES

Only lignin DB

Not previously identified in RCF lignin*

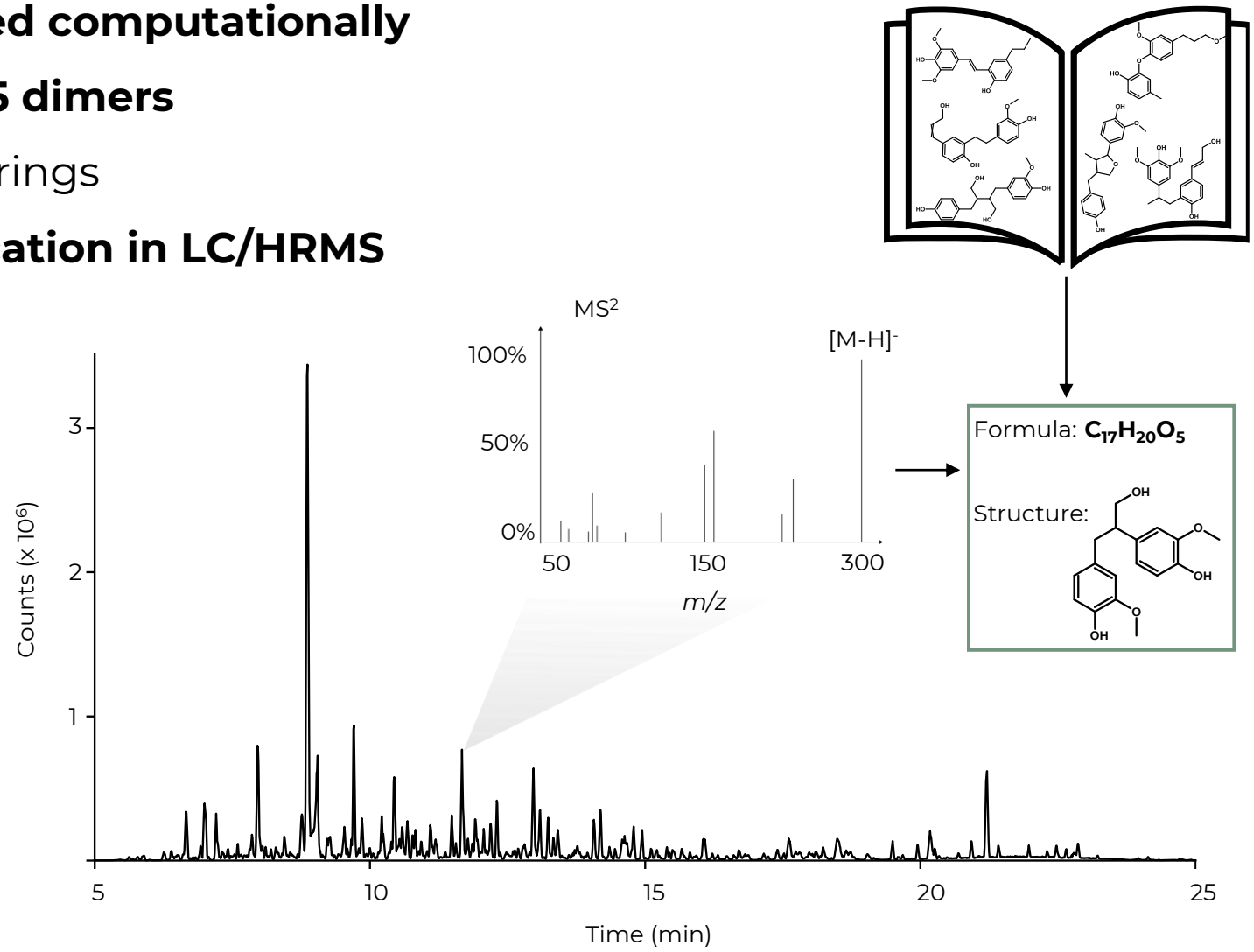


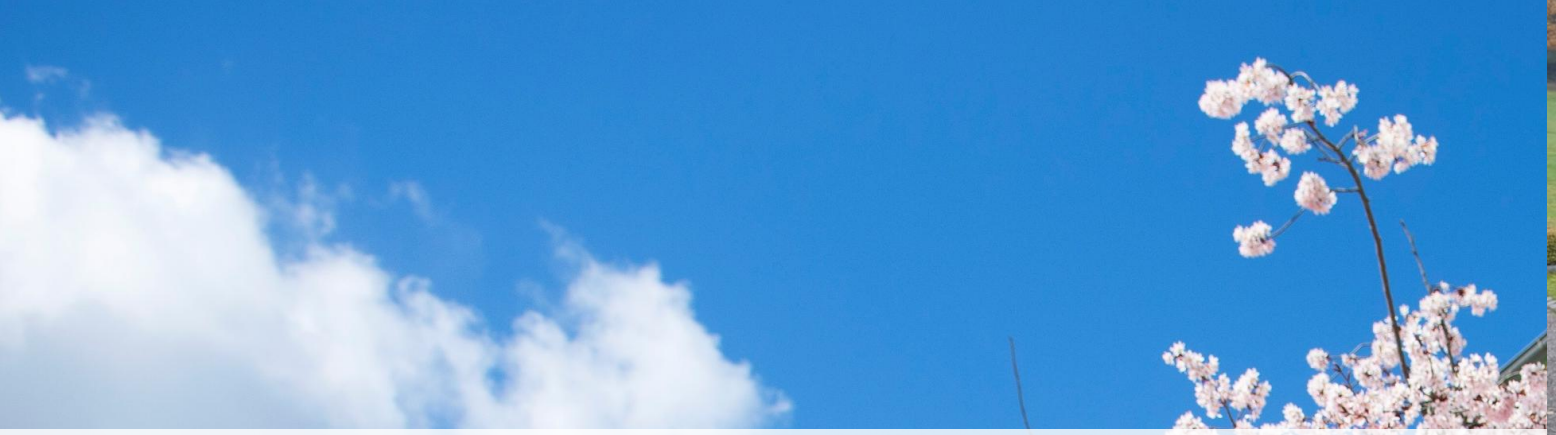
*

Dao Thi et al., *Green Chem.* (**2022**)
Aelst et al., *Chem. Sci.* (**2020**)
Anderson et al., *Nat. Commun.* (**2019**)

SOME TAKE-HOME MESSAGES

- **Lignin structures** can be **generated computationally**
  - RCF lignin library contains **1945 dimers**
  - Will be published as **SMILES** strings
- Can be used for **tentative identification in LC/HRMS**
  - **new** structures in databases
  - **confirm** existing structures







THANK YOU!

Unravelling the Structures of Depolymerised Lignin Compounds via Mass Spectrometry and Rule-Based Modelling

Eliise Tammekivi, Suthawan Muangmeesri, Joseph S. M. Samec, Anneli Kruve

eliise.tammekivi@su.se



Kruvelab.com

Wenner-Gren Foundations
Wenner-Gren Stiftelserna

erc
European Research Council
Established by the European Commission