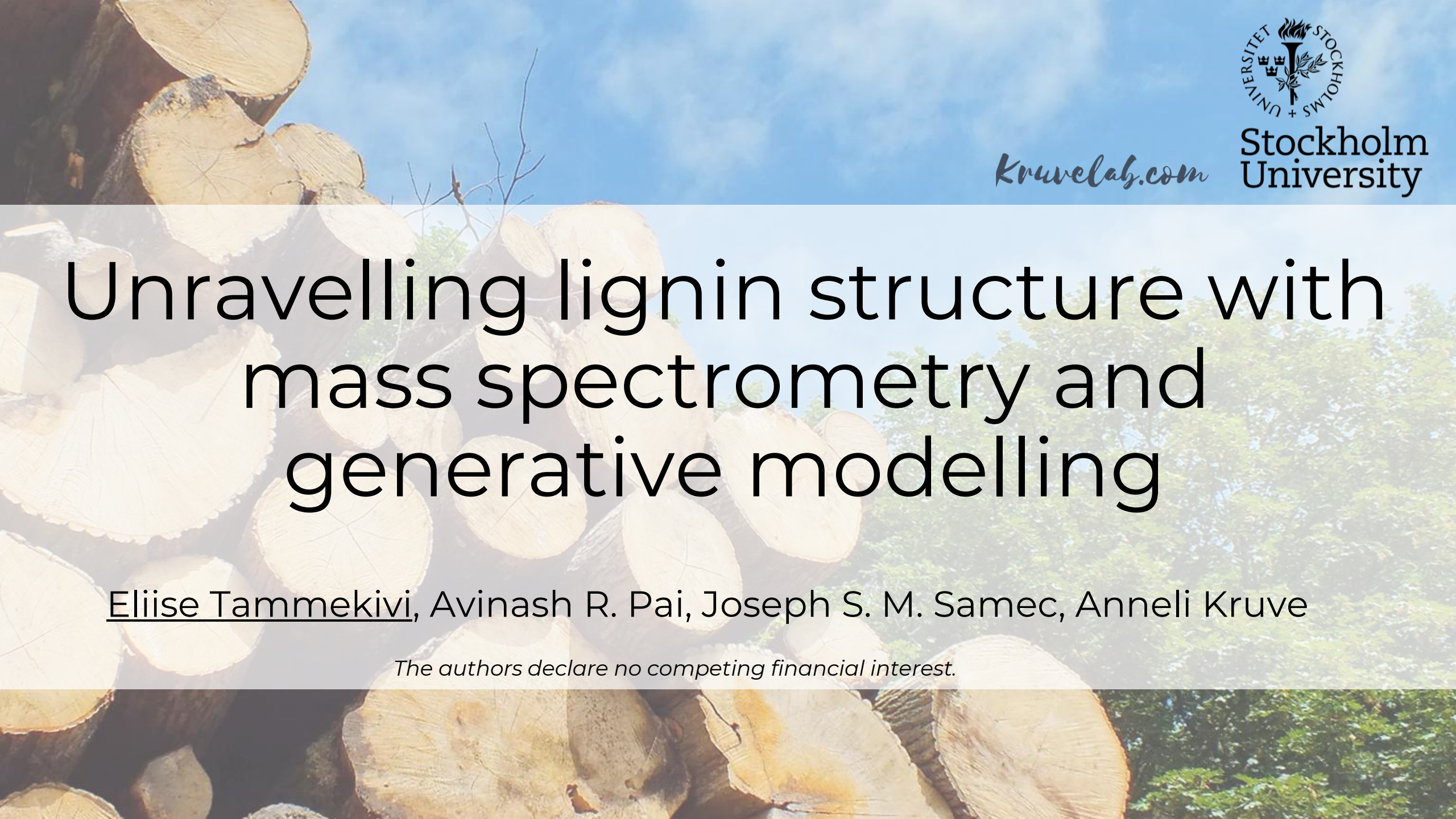


Kruvelab.com

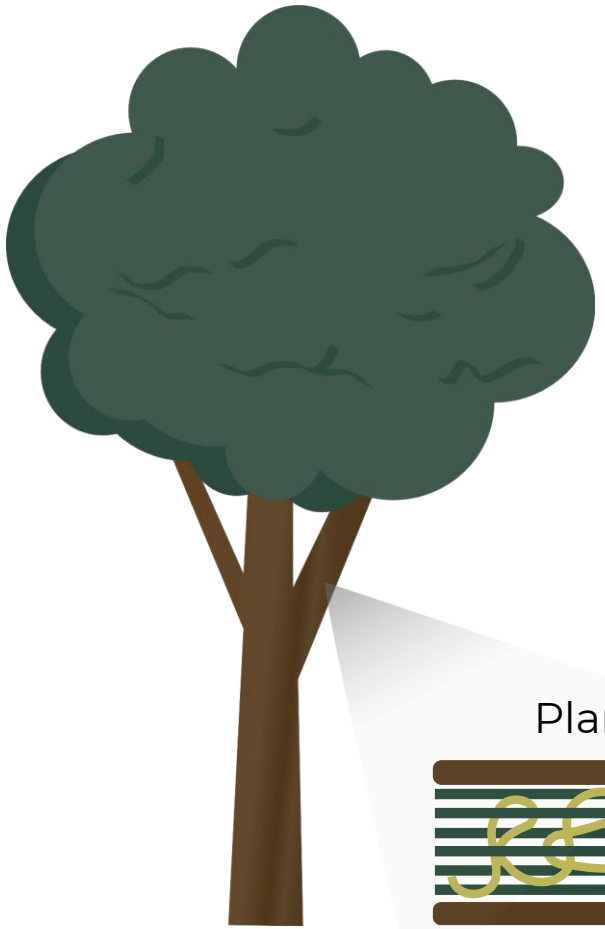
Stockholm University

Unravelling lignin structure with mass spectrometry and generative modelling

Eliise Tammekivi, Avinash R. Pai, Joseph S. M. Samec, Anneli Kruve

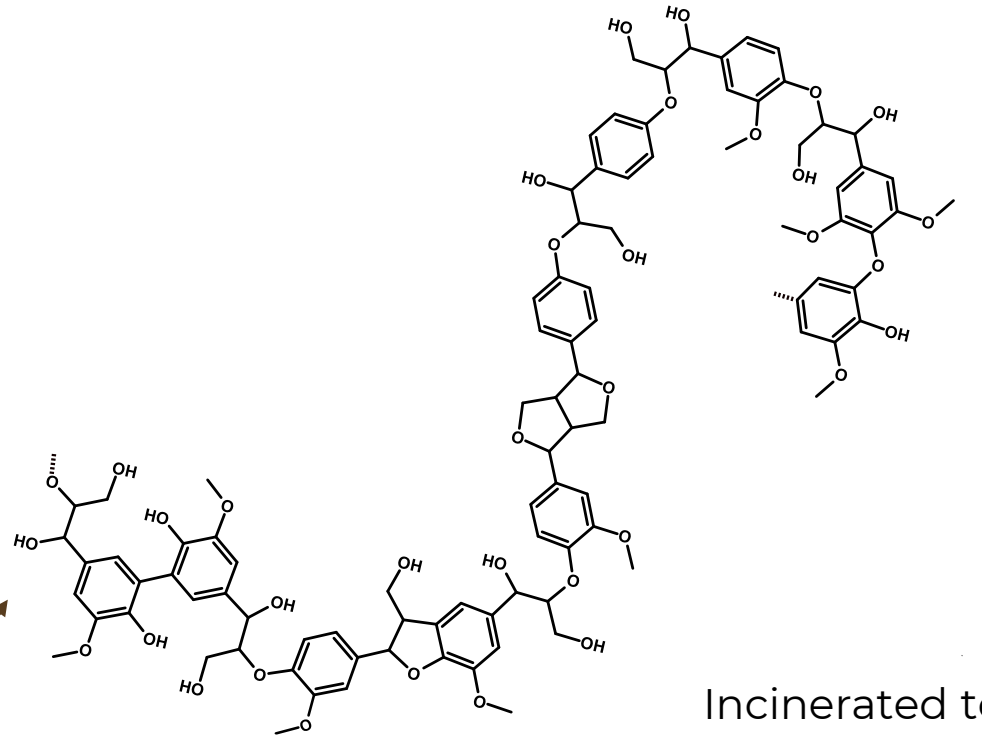
The authors declare no competing financial interest.

LIGNIN



Plant cell wall

Lignin



Hemicellulose

Cellulose

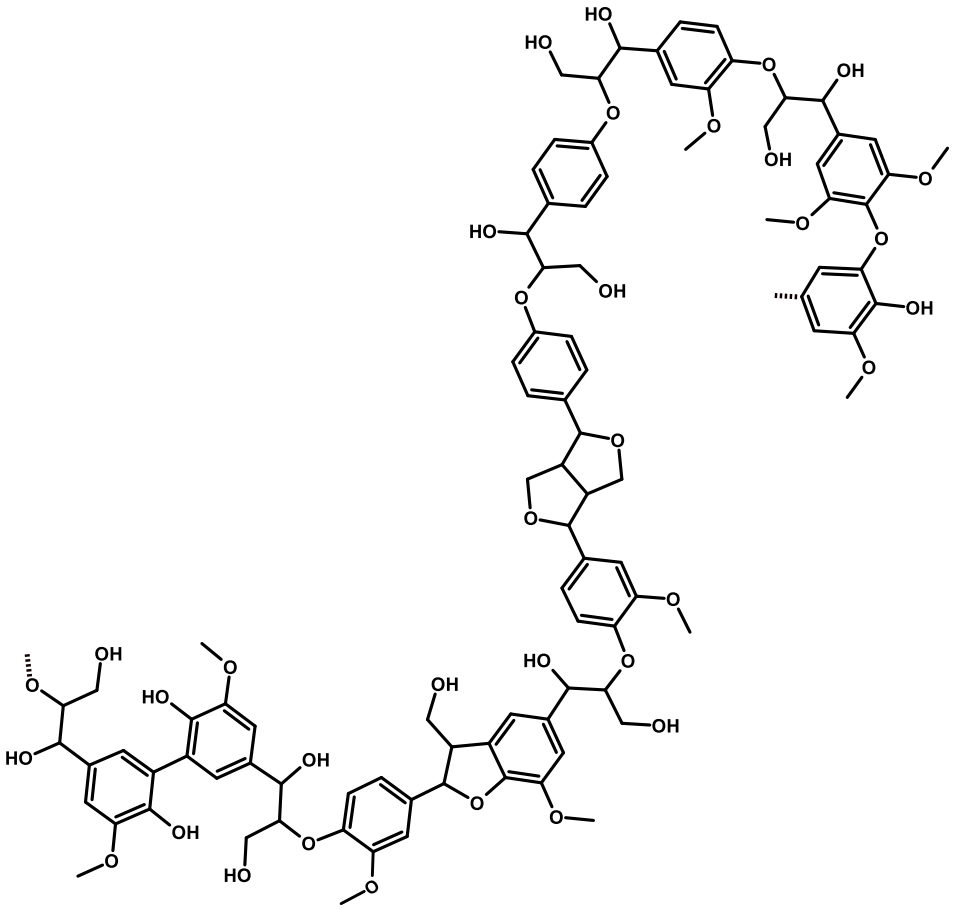
Paper, textile, food, and pharmaceutical industries

...

Incinerated to

- recover chemicals
- produce energy

LIGNIN



Extraction of compounds

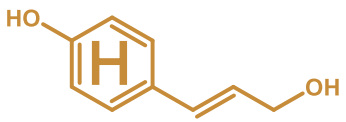
Biofuel

Biomaterials

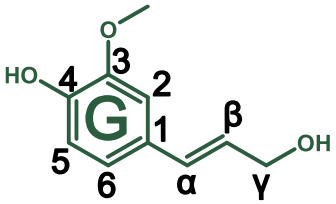
- The most abundant natural aromatic material

- Around ~2% is commercialised
- Develop new valorisation ways

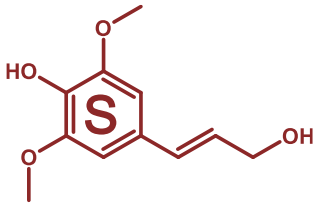
LIGNIN



p-coumaryl alcohol (H)

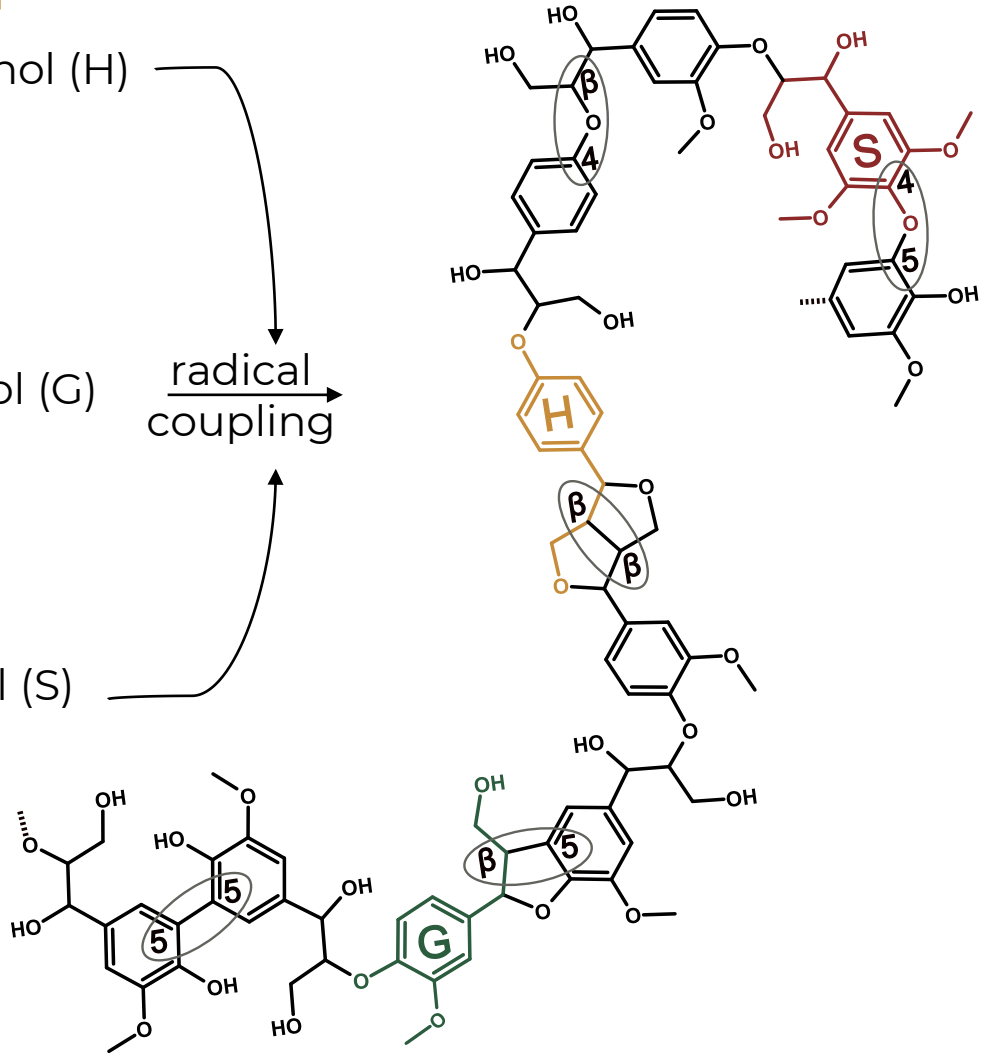


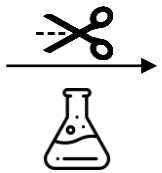
coniferyl alcohol (G)

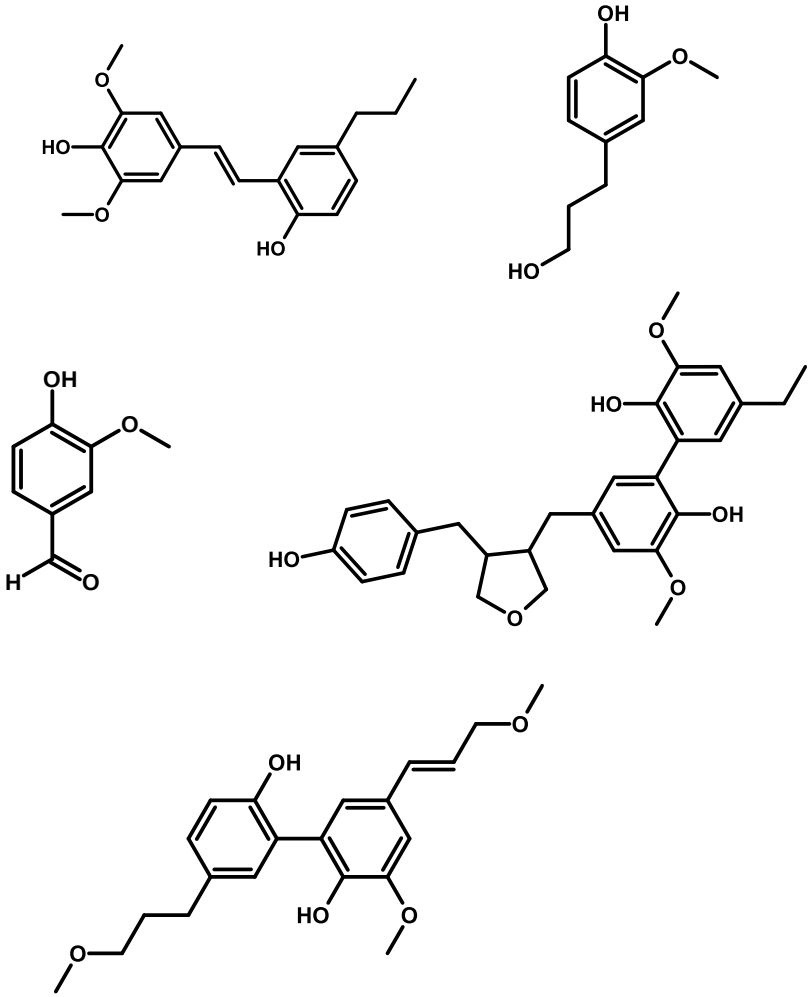


sinapyl alcohol (S)

radical coupling







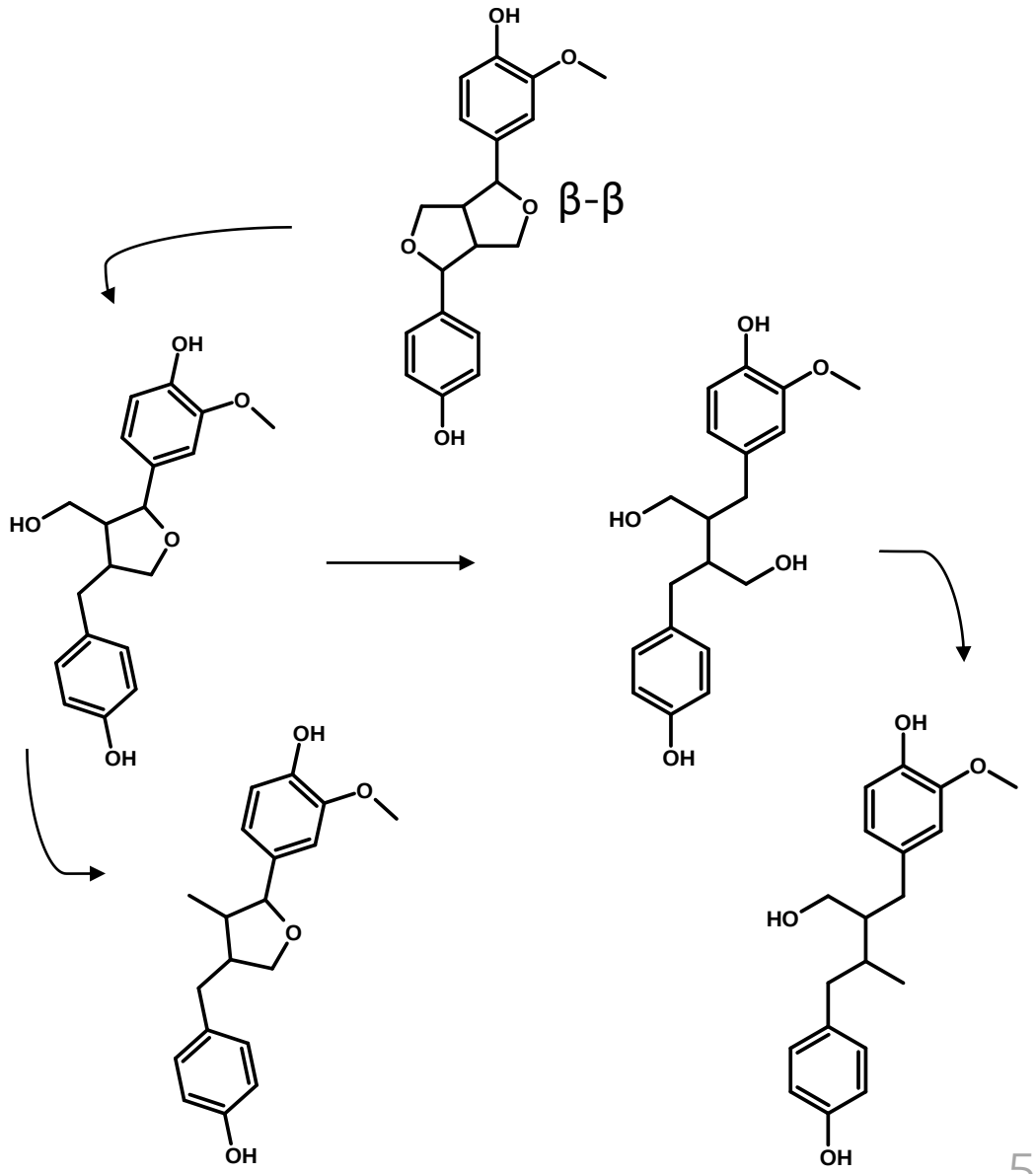
- Mixture containing hundreds of compounds
- Structurally highly similar compounds (including numerous isomers)

REDUCTIVE CATALYTIC FRACTIONATION (RCF)

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

Reactions occurring under reductive conditions

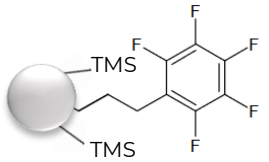
Reaction	Scheme
hydrogenation	R_1 + H_2 → R_1
hydrogenolysis	R_1–O–R_2 + H_2 → R_1–OH + CH_3–R_2
decarboxylation	R_1–C(=O)–OH + H_2 → R_1–H + CO_2
demethoxylation	R_1–O + H_2 → R_1–H + CH_3–OH
demethylation	R_1–O + H_2 → R_1–OH + CH_4
hydrodeoxygenation	R_1–OH + H_2 → R_1–H + H_2O

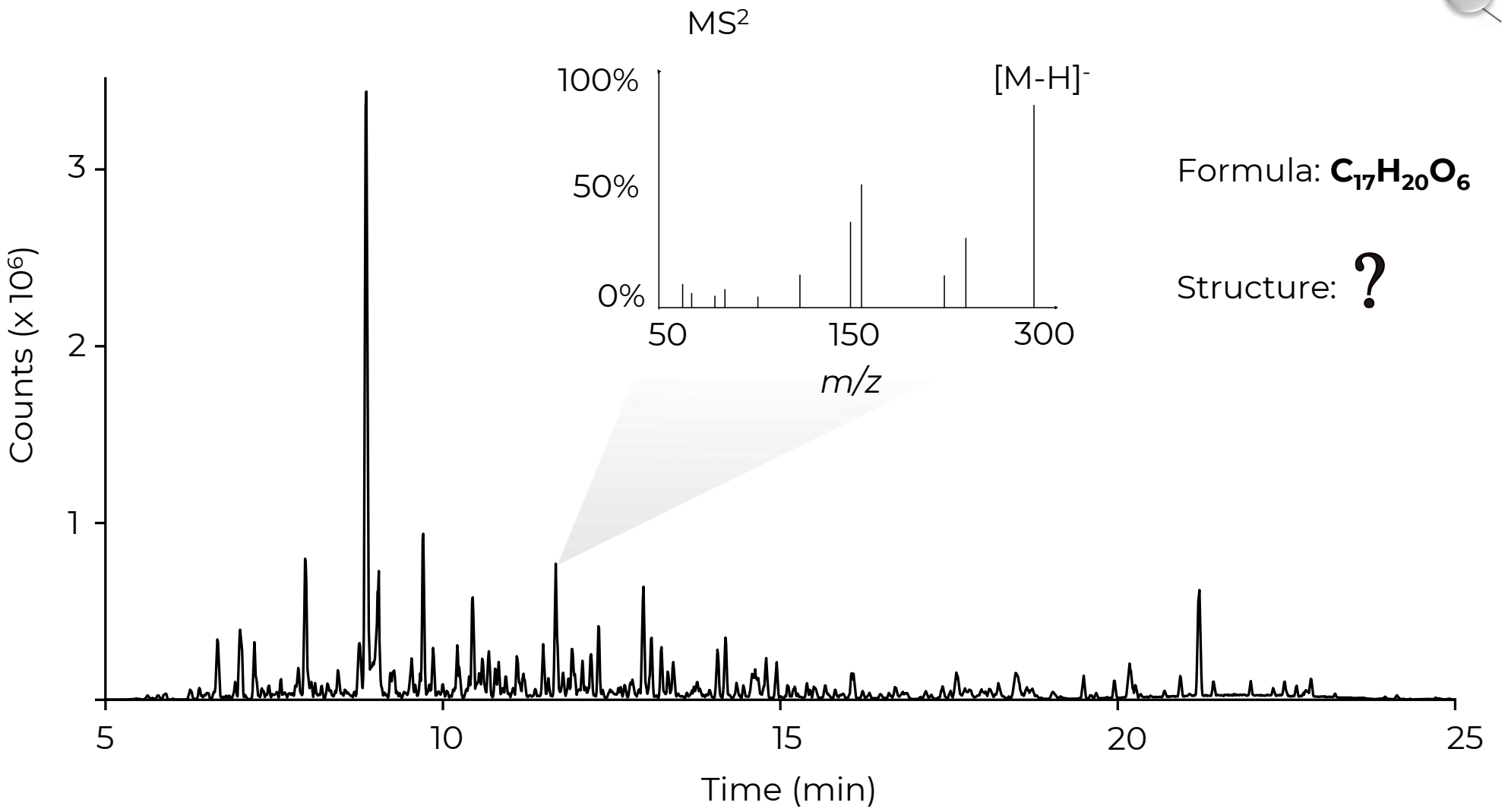


Bourbiaux et al., *Catalysis Today* (**2021**)

RCF LIGNIN ANALYSIS WITH LC/HRMS

Column: Kinetex **F5** Core-shell

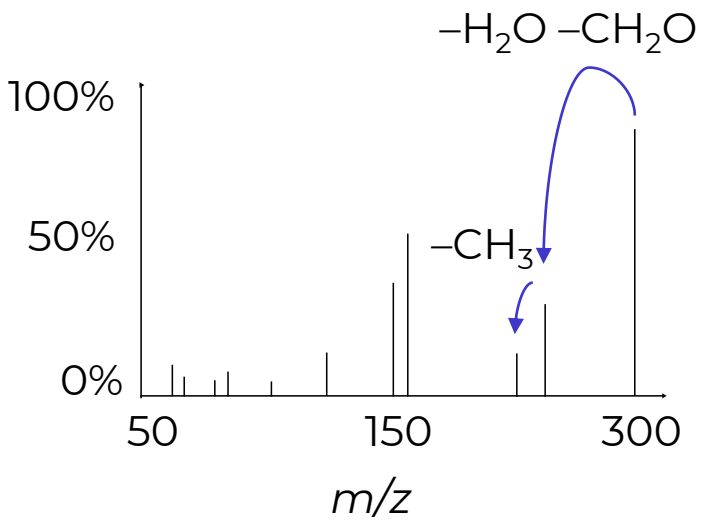




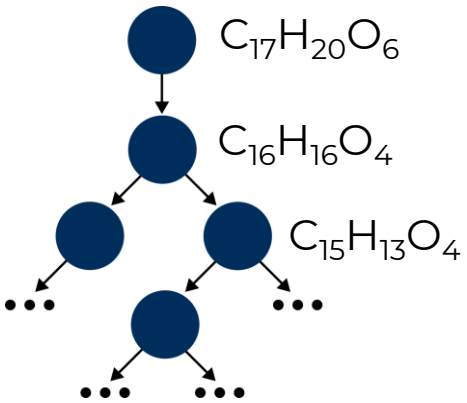
Formula: $\mathbf{C_{17}H_{20}O_6}$

Structure: ?

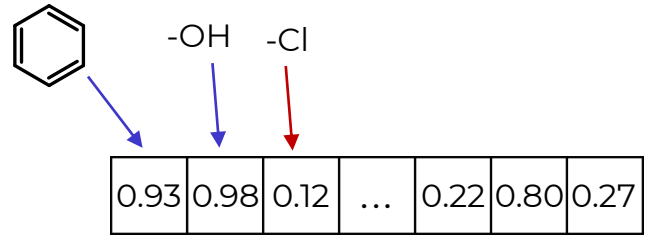
SIRIUS+CSI:FINGERID



Fragmentation tree

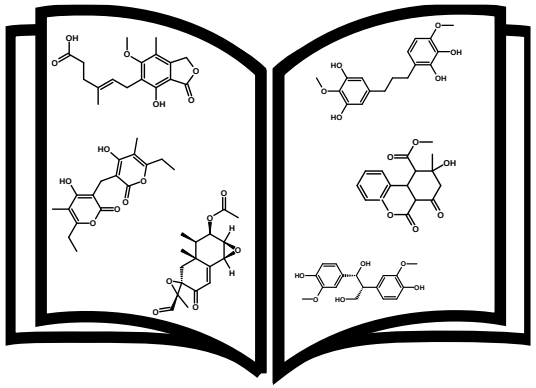


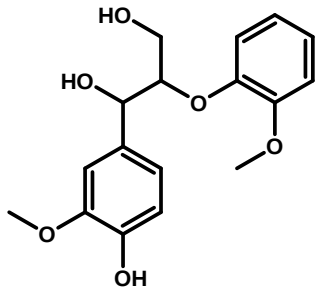
Probabilistic fingerprint



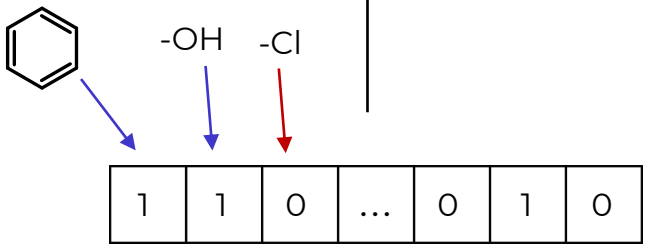
Top 1 match

PubChem, COCONUT, …





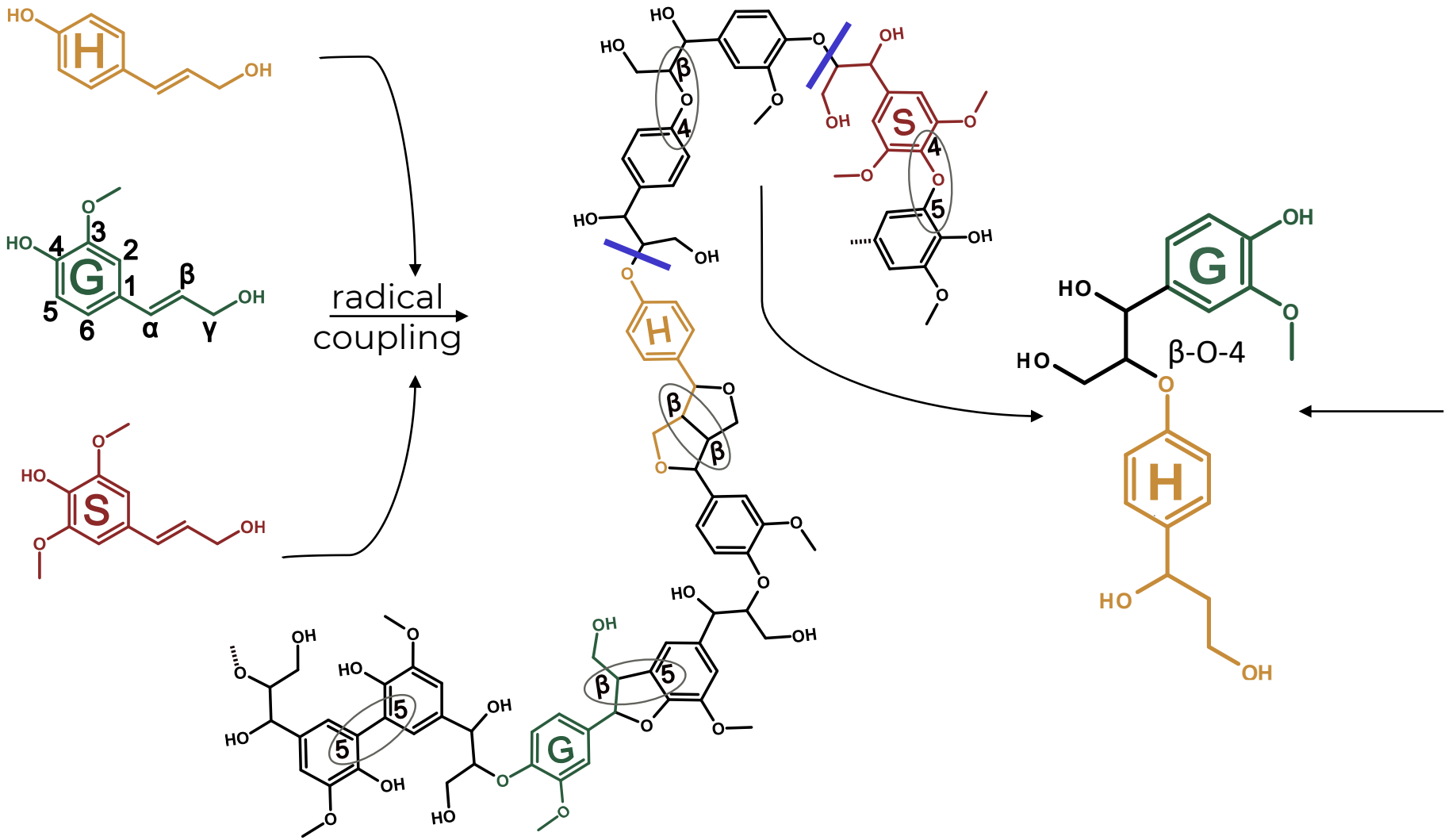
OC(C(Oc1ccccc1OC)CO)c1cc(OC)c(O)cc1



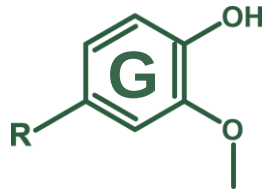
Binary fingerprint

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

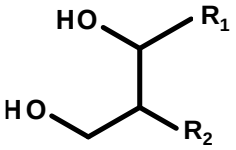
LIGNIN DIMER STRUCTURES



Monomeric unit 1



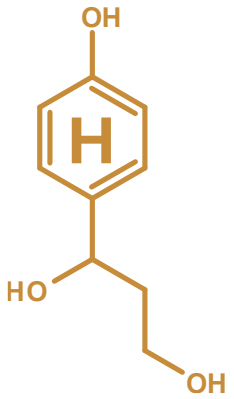
+



Linkage

+

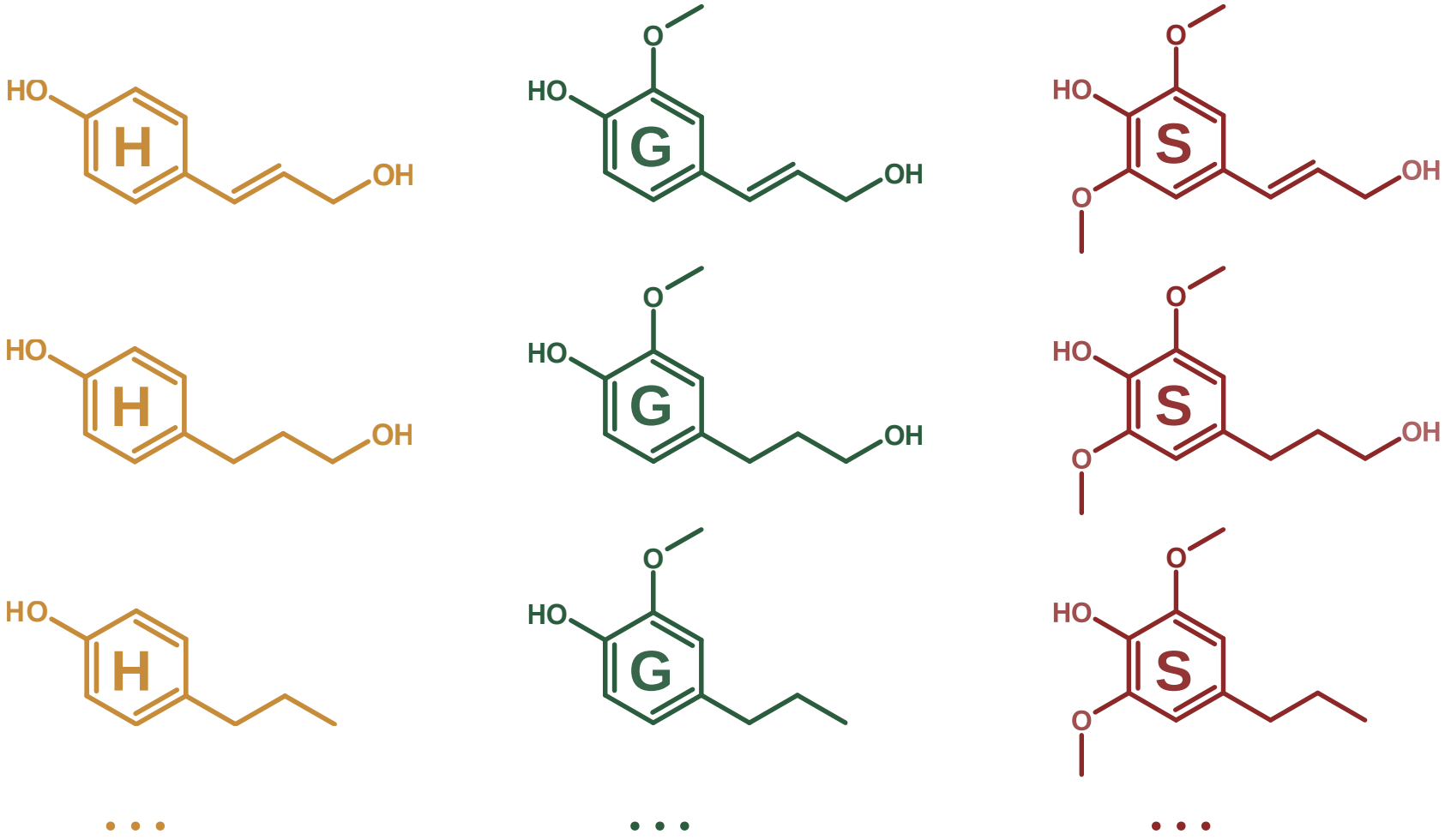
Monomeric unit 2



GENERATING LIGNIN DIMER STRUCTURES

- Monomeric units in the potential RCF lignin dimers

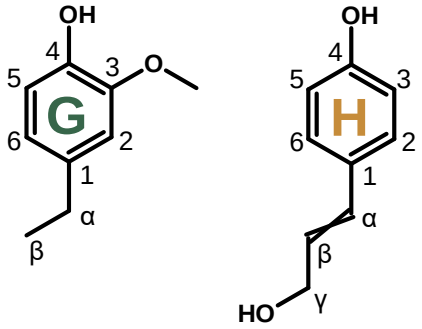
In total: **26 monomeric units**



Dao Thi et al., *Green Chem.* (**2022**)

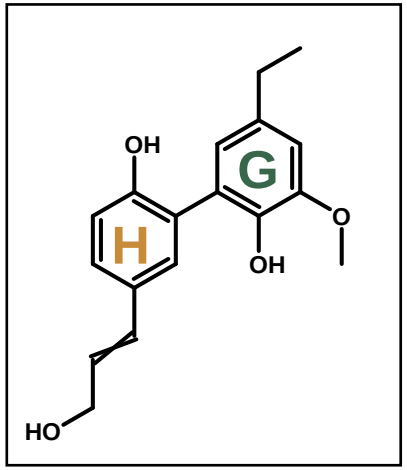
GENERATING LIGNIN DIMER STRUCTURES

Exemplary monomeric units to show the linkages

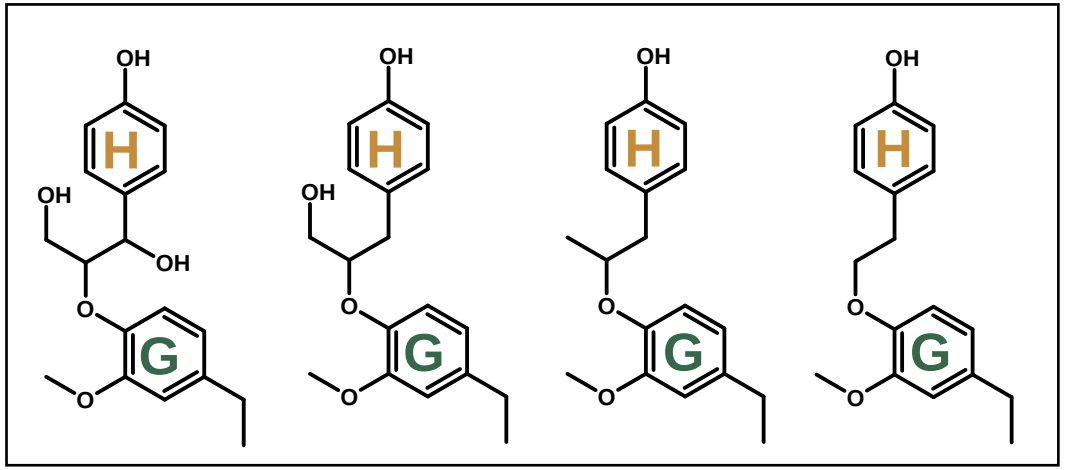


In total: **24 linkages**

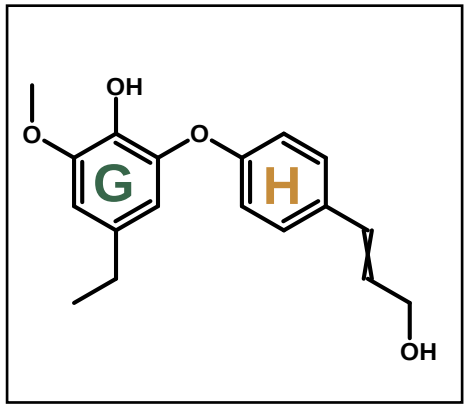
5-5: 1 linkage



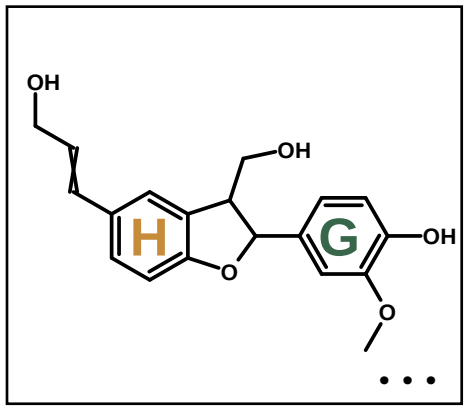
β-O-4: 4 linkages



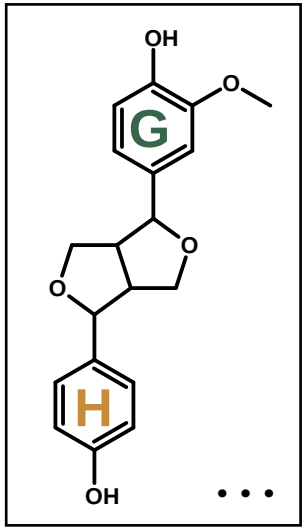
4-O-5: 1 linkage



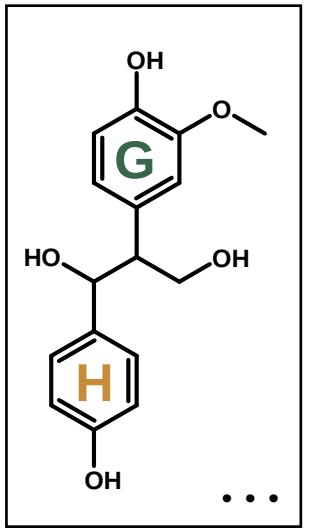
β-5: 6 linkages



β-β: 8 linkages



β-1: 4 linkages

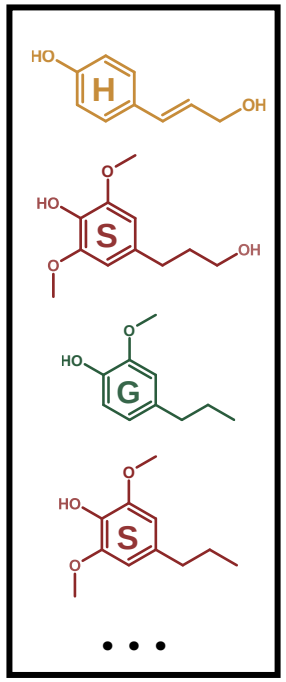


Dao Thi et al., *Green Chem.* (**2022**)

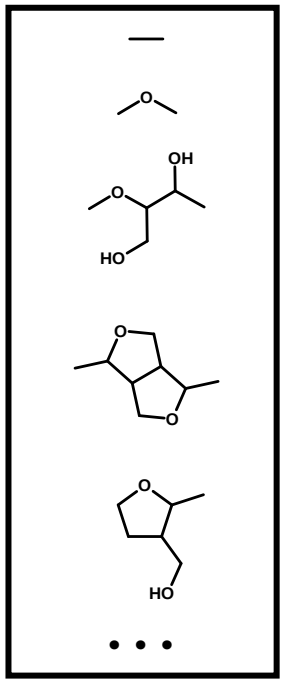
GENERATING LIGNIN DIMER STRUCTURES

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

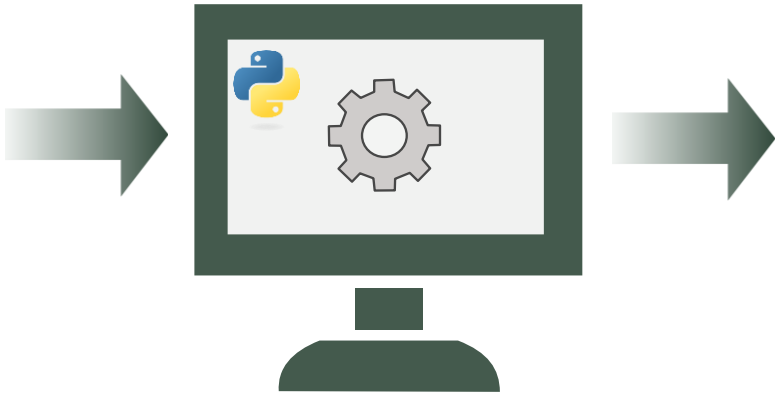
26 monomers



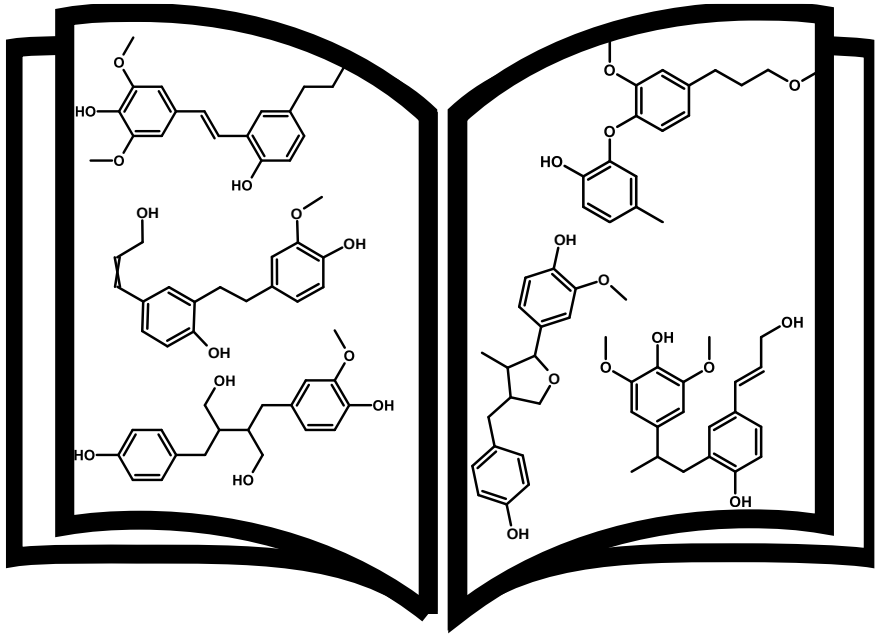
24 linkages



Generating all possible combinations



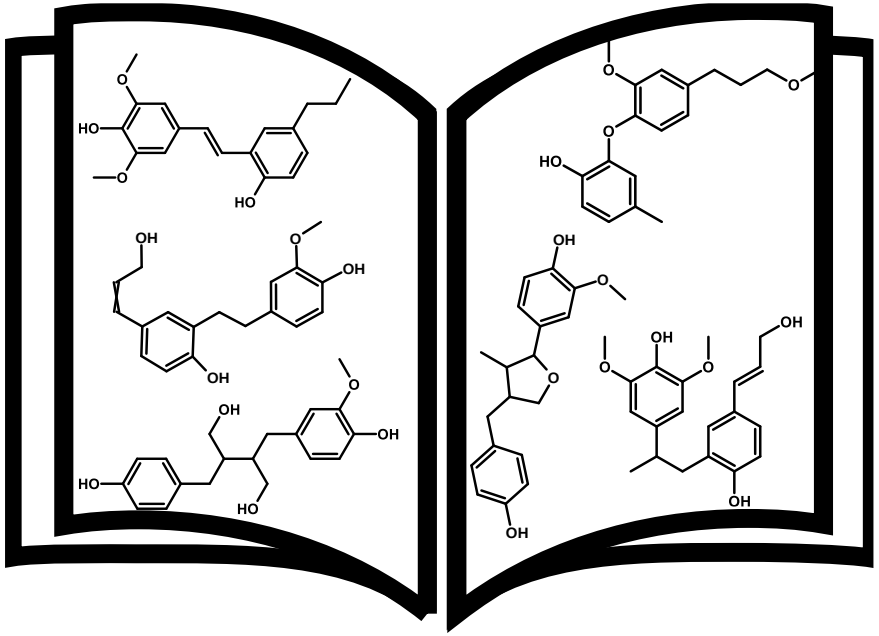
Library of generated RCF **lignin dimer structures**



GENERATING LIGNIN DIMER STRUCTURES

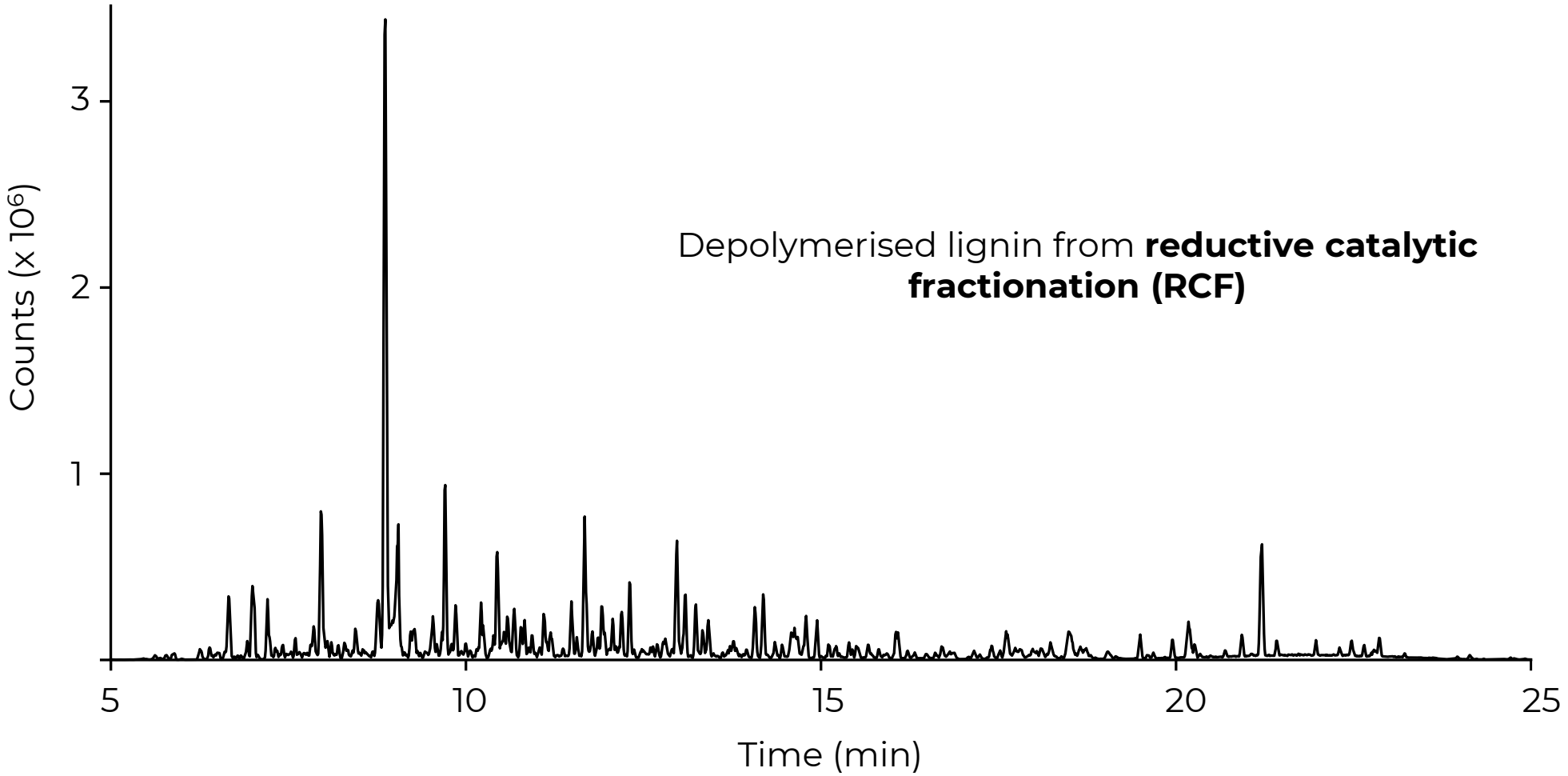
- Contains **1720 dimer** structures
- **SMILES** for each structure
- Only **117 molecular formulas**
- Only **195 (12%)** of them are in **PubChem**
- Structures of **potential** lignin RCF dimers

Library of generated RCF **lignin dimer structures**



APPLYING THE LIGNIN LIBRARY: SUSPECT

- **DDA (Top 5)** using the **suspect list** of the lignin dimer molecular formulas
- **mzMine** (4.3.0) was used for data processing
- The lignin library was imported into SIRIUS (6.0.4) as a **custom database**



APPLYING THE LIGNIN LIBRARY: SUSPECT

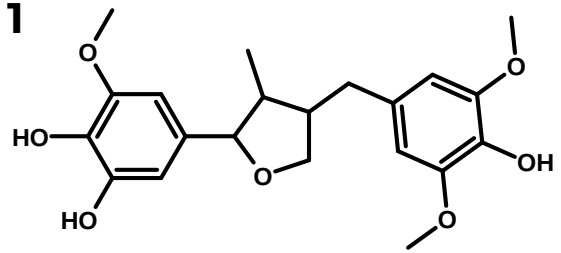
- **DDA (Top 5)** using the **suspect list** of the lignin dimer molecular formulas
- **mzMine** (4.3.0) was used for data processing
- The lignin library was imported into SIRIUS (6.0.4) as a **custom database**

SIRIUS results for one MS^2

Top 1

Sources:

Lignin DB

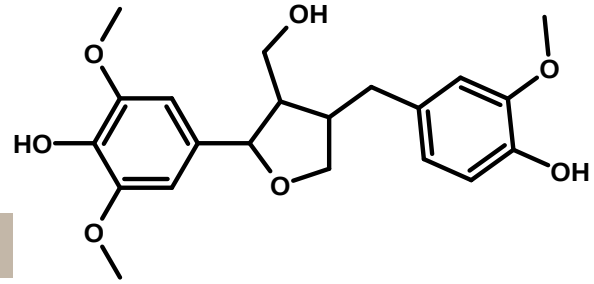


Top 2

Sources:

Lignin DB

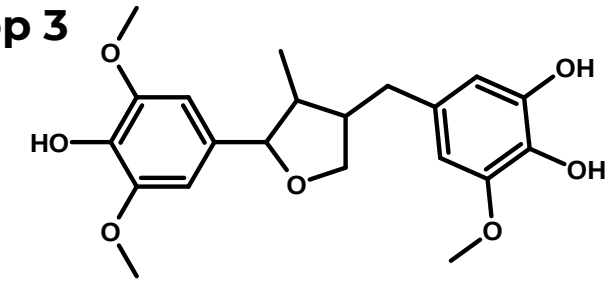
SuperNatural | LOTUS | PubChem | COCONUT



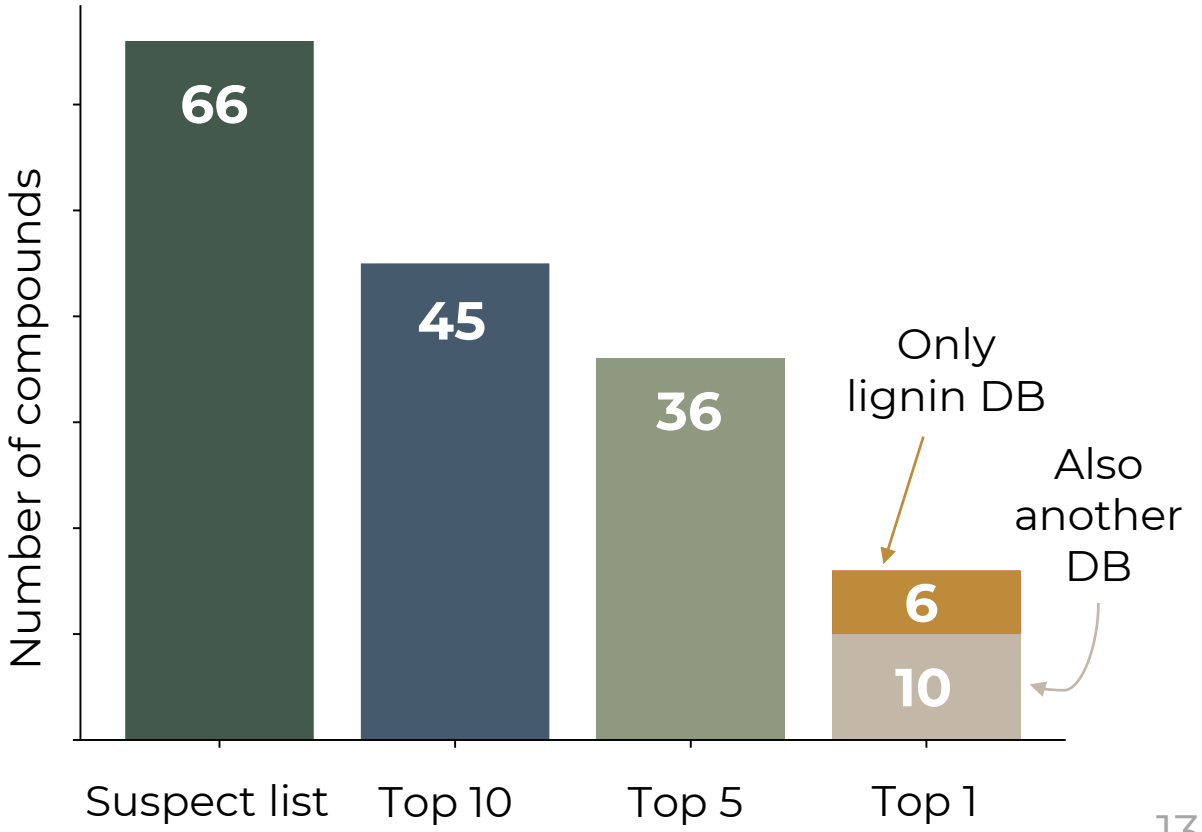
Top 3

Sources:

Lignin DB

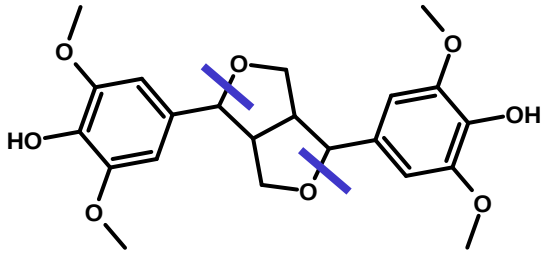


Matches with the lignin library



APPLYING THE LIGNIN LIBRARY: SUSPECT

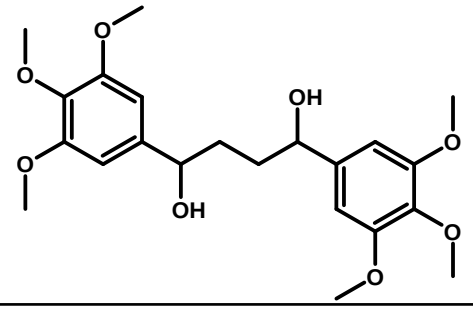
Original linkage:



Top 1

Sources:

PubChem



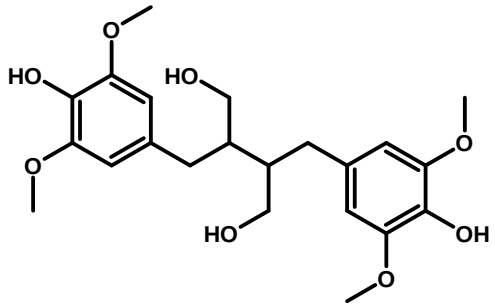
Top 2

Sources:

Lignin DB

PubChem | DSSTox

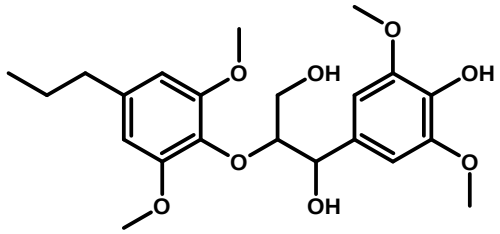
Blood Exposome | HMDB



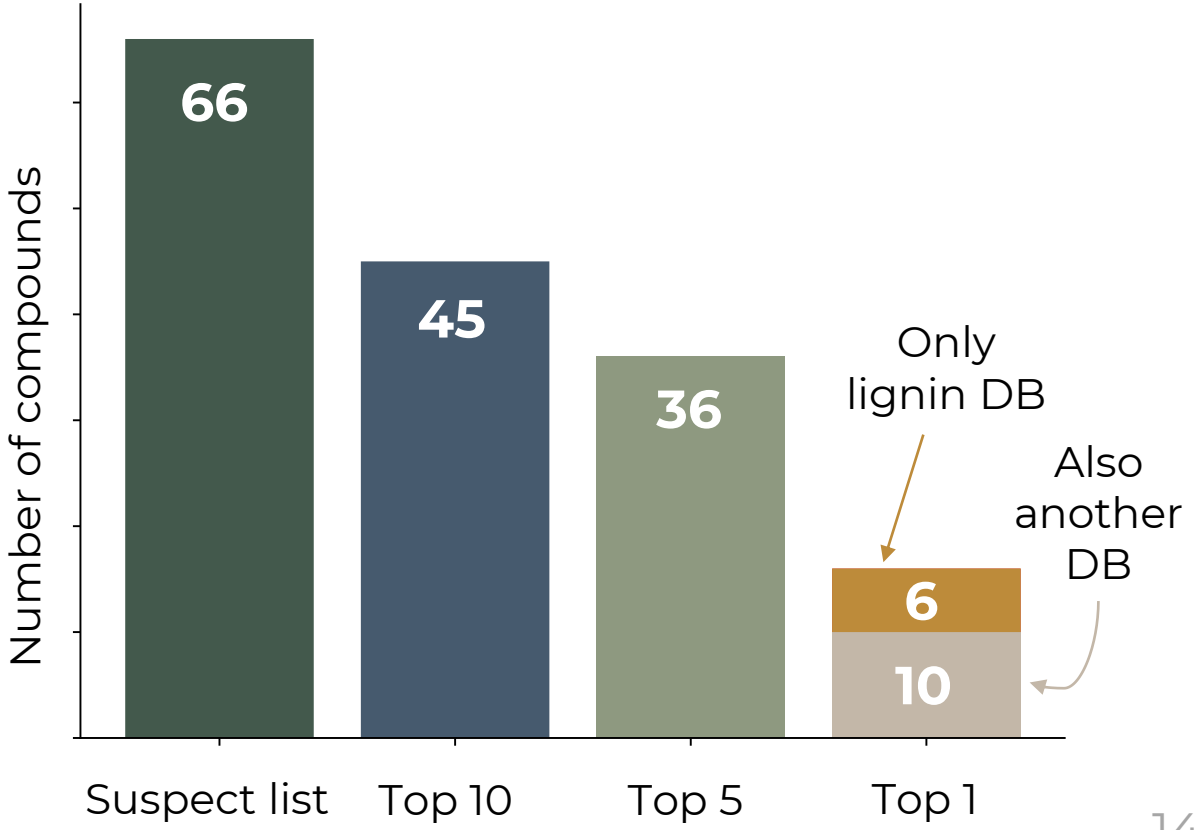
Top 3

Sources:

Lignin DB



Matches with the lignin library



APPLYING THE LIGNIN LIBRARY: NON-TARGET

- Two HRMS instruments were used for analysis:
- Thermo **Orbitrap MS** (Exploris 480)
  - 0.1% HCOOH in both H_2O and ACN
  - 5 mM NH_4HCO_3 in H_2O
- Waters **Cyclic IMS-TOF**
  - 0.1% HCOOH in both H_2O and ACN

	Suspect	Non-target
Method	**Orbitrap NH_4HCO_3**	**Orbitrap NH_4HCO_3**
Acquisition	DDA Top 5	DDA Top 5
Features	449	428
Features with MS^2	66	415
Top 10 match	45 (68%)	45 (11%)
Top 1 match	16 (24%)	16 (4%)

*Percentage (%) is calculated taking features with MS^2 as 100%

APPLYING THE LIGNIN LIBRARY: NON-TARGET

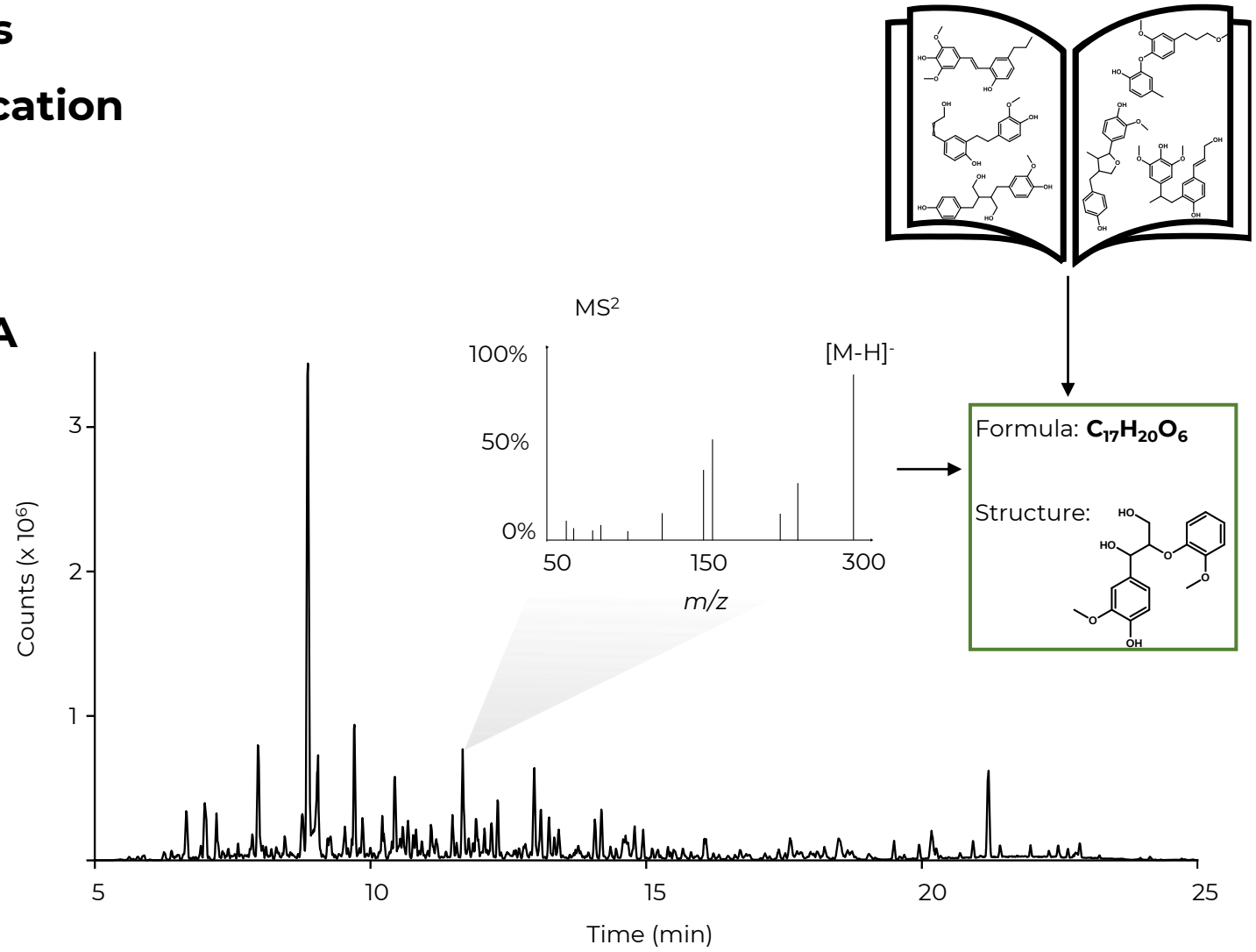
- Two HRMS instruments were used for analysis:
- Thermo **Orbitrap MS** (Exploris 480)
  - 0.1% HCOOH in both H_2O and ACN
  - 5 mM NH_4HCO_3 in H_2O
- Waters **Cyclic IMS-TOF**
  - 0.1% HCOOH in both H_2O and ACN

	Non-target		
Method	**Orbitrap NH_4HCO_3**	**Orbitrap HCOOH**	**IMS HCOOH**
Acquisition	DDA Top 5	DDA Top 5	DIA
Features	428	903	448
Features with MS^2	415	822	442
Top 10 match	45 (11%)	72 (9%)	58 (13%)
Top 1 match	16 (4%)	37 (5%)	20 (5%)

*Percentage (%) is calculated taking features with MS^2 as 100%

CONCLUSIONS

- **Lignin structures** can be **generated computationally**
- Lignin library contains **1720 dimers**
- Can be used for **tentative identification**
  - **new** structures in databases
  - **confirm** existing structures
- Can be used for both **DIA and DDA**



THANK YOU!

Unravelling lignin structure with mass spectrometry and generative modelling

Eliise Tammekivi, Avinash R. Pai, Joseph S. M. Samec, Anneli Kruve

Kruvelab.com

Wenner-Gren Foundations
Wenner-Gren Stiftelserna

European Research Council
Established by the European Commission