

NMR Database of Lignin and Cell Wall Model Compounds

Sally Ralph

US Forest Products Laboratory (FPL)
USDA Forest Service
One Gifford Pinchot Drive
Madison, WI 53705
+1 (608) 231-9449
email: sally.ralph@usda.gov

John Ralph

Department of Biochemistry, and the Great Lakes Bioenergy Research Center (GLBRC), Wisconsin Energy Institute (WEI)
University of Wisconsin-Madison
1552 University Avenue
Madison, WI 53726-4084
+1 (608) 890-2429
E-mail: jralph@wisc.edu

Fachuang Lu

Great Lakes Bioenergy Research Center (GLBRC), Wisconsin Energy Institute (WEI)
University of Wisconsin-Madison
1552 University Avenue
Madison, WI 53726-4084
+1 (608) 890-2552
E-mail: fachuanglu@wisc.edu

This database was designed to provide a coherent, single source of NMR data of lignin and other plant cell wall model compounds. The database exists in several parts as PDF versions that are available for downloading over the internet from DOE Great Lakes Bioenergy Research Center's web site:
https://www.glbrc.org/databases_and_software/nmrdatabase/

It is linked via its DOI at:

<https://doi.org/10.11578/2409191>

...and on OSTI.gov at:

<https://www.osti.gov/biblio/2409191>

In general NMR data was collected in three common deuterated solvents (acetone,

chloroform, and dimethyl sulfoxide) for each compound. We used the central solvent peak as our reference, 2.04 and 29.83 ppm for acetone-d₆, 7.24 and 77.00 for CDCl₃, and 2.49 and 39.50 ppm for DMSO-d₆ data. A standard set of acquisition parameters was used to acquire and process the spectra to keep the data as uniform and constant as possible. The samples were run at ambient temperature, about 298 K.

Those compounds with an index number less than 1000 were run on a Bruker 250 or 500 MHz spectrometer at FPL; those compounds with an index number between 1000 and 10,000 were run at the USDA-ARS Dairy Forage Research Center (DFRC) on a Bruker 360 MHz instrument, or on a Bruker 500 or 700 MHz instrument at the GLBRC for later compounds. The order of the compounds in the database reflects their arrival at the spectrometer rather than from a preordained plan. The inclusion of series of analogous structures with small structural differences allows calculation of substituent effects that are invaluable for chemical shift predictions of structures not included in the database.

The chemical shift assignments for most of the compounds were made by comparison with other compounds, literature values, and in some cases using the standard set of 1D and 2D NMR experiments. Every effort was made to correctly assign the chemical shifts; however, limited time and resources precluded confirming the shifts for many of the compounds. The shifts are reported to the second decimal place only to distinguish very close shifts; comparisons between spectra are practical only within ± 0.1 ppm. The authors would greatly appreciate learning of any corrections on suspect assignments.

The compounds themselves came from many sources — in-house collections, syntheses, and donations from other researchers for which we are grateful. The source of the compounds is often given in the “Notes” field along with other pertinent data.

This database was originally intended as an aid for the assignment of chemical shifts for wood and plant lignin NMR spectra. The trivial names used throughout are well known to wood chemists as is the numbering system. We have attempted to include more formal chemical names for many of the compounds and these were obtained using Beilstein’s Autonom[®] program. The chemical names for the larger 3 and 4 ring models became so cumbersome that the authors employed an abbreviated system to identify both the moieties involved as well as the linkages between the moieties. Examples of the naming, numbering and linking conventions used are given below.

letters and numbers to describe the type of linkage between the rings.

Table 2. Terminology for Abbreviated Structural Entities

Entity	Abbreviation
guaiacyl ring	G
syringyl ring	S
hydroxyphenyl ring	H
α -O-4 linkage	a
β -O-4 linkage	b
β -5 (phenylcoumaran)	c
β -1 linkage	b1
β - β (resinol)	r
5-5 (biphenyl)	5,5
coniferyl alcohol end-unit	CA
sinapyl alcohol end-unit	SA
<i>p</i> -coumaryl alcohol end-unit	HA
ferulic acid end-unit	FA
<i>erythro</i> -isomer	<i>e</i>
<i>threo</i> -isomer	<i>t</i>

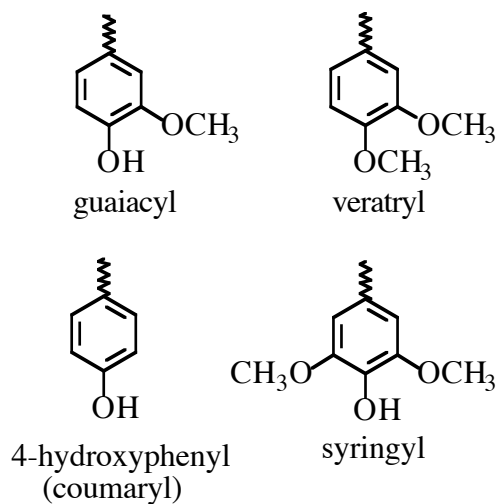
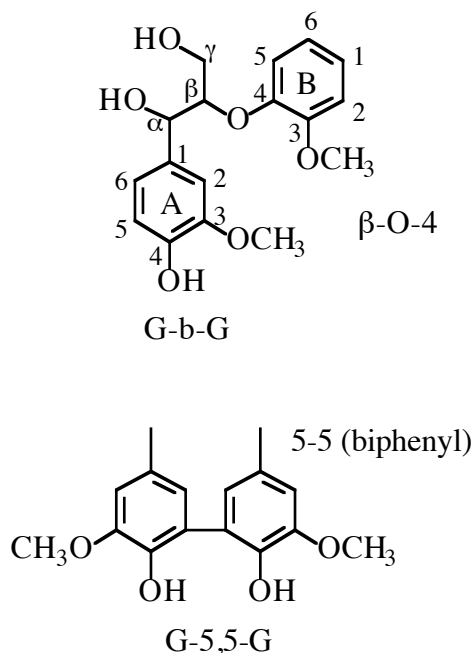


Fig. 1. Trivial names for substituents at the 3, 4 and 5 positions on the aromatic ring.

The naming of the larger oligomer lignin models uses a combination of uppercase letters to describe the ring structure and lowercase



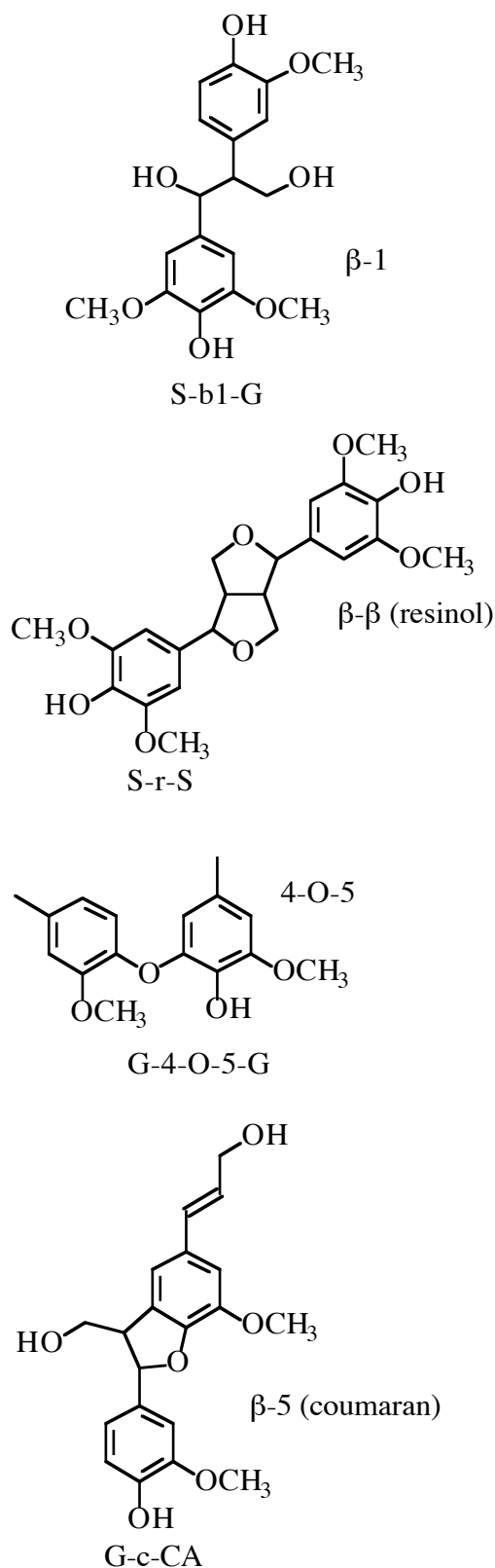


Fig. 2. Examples of linkages and abbreviated names.

With this convention the name FA-5,5-FA would represent a diferulic acid biphenyl structure. The trimer CA-a-G-b-CA would be a guaiacyl unit with two coniferyl alcohol end-groups phenol-etherified at the α - and β -positions.

The structure index is arranged based upon the number of rings in the structure. Where possible the structures are also arranged by ring type such as guaiacyl, syringyl, etc. The number under the structure refers to the index number at the top of the datasheet. An asterisk after a number indicates the peracetylated analog of that compound. In some cases only the acetylated compound is included.

We hope to continue adding to and improving this database. Regular updates will be made to the database to keep the online sources current. This database was written and prepared for the most part by U.S. Government employees on official time, and it is therefore in the public domain and not subject to copyright. Please feel free to contact the authors with suggestions or questions.

The authors gratefully acknowledge the many generous contributions made by others at three labs (FPL, DFRC, GLBRC) towards this database; foremost Larry Landucci for compounds, patience, assignments, and encouragement. William Landucci for software development, Martin Wesolowski and Kolby Hirth, NMR spectroscopists, and the frequent donation of compounds from Mike Mozuch, Noritsugu Terashima, Stéphane Quideau, Rich Helm, Fachuang Lu, Hoon Kim, Jamie Milhaupt, Susana Luque, and others. The authors also acknowledge partial support from the National Research Initiative's Competitive Grants Program/USDA (Wood Section), award #94-03465.

Structure Index

Contents

Monomers

- H, *p*-OMe-H, G, V (*p*-OMe-G), S, *p*-OMe-S SI 2
- 3-Carbon sidechain SI 3
- Hibbert's Ketones, Homovanillyl Structures, Misc. SI 4

β -O-4 Dimers

- 2-Carbon sidechain and 3-carbon sidechain with α C=O SI 5
- β -O-4 dimers, 3-carbon sidechain SI 6
- More β -O-4 dimers, and 5-hydroxy-CAlc β -O-4 dimers SI 7

β -5 Dimers

SI 8

β - β Dimers

SI 9

5-5, 5-O-4, β -1, and diphenylmethane dimers

SI 10

Trimers

- β -O-4, β -O-4- β -5, β -O-4- β - β , β -O-4- β -1, β -5- β -1, 5-O-4 SI 11
- 5-O-4, Dibenzodioxocin; β -O-4 ferulates, coumarates, and benzoates SI 12
- More β -O-4 ferulates, coumarates, and benzoates SI 13

Tetramers

SI 14

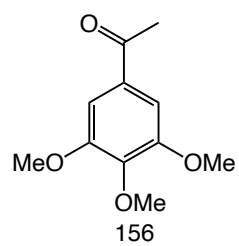
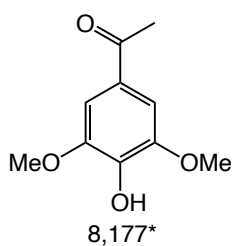
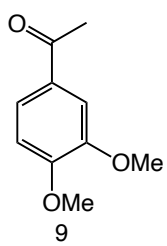
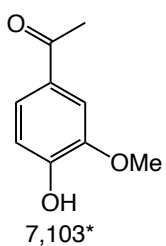
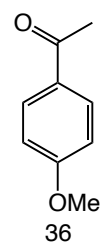
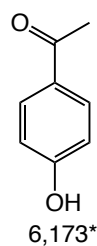
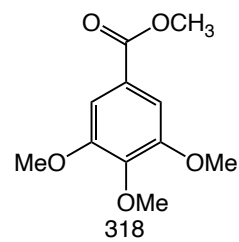
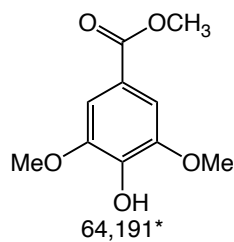
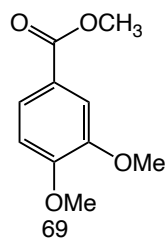
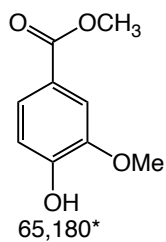
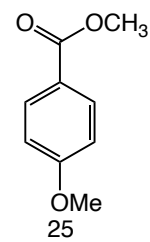
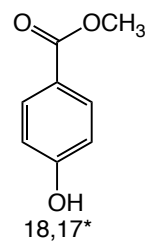
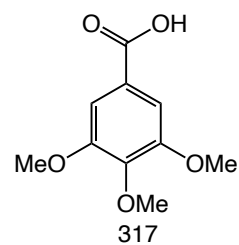
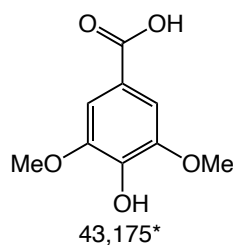
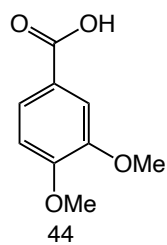
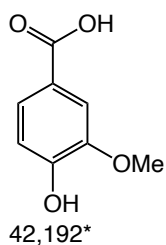
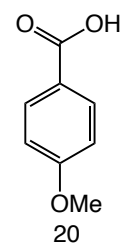
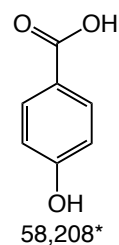
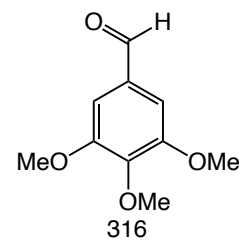
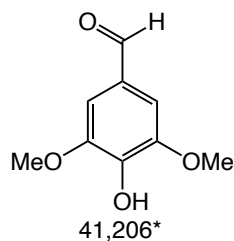
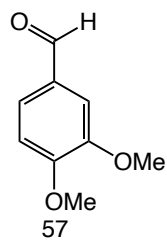
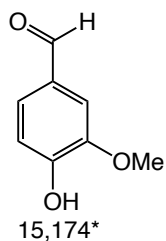
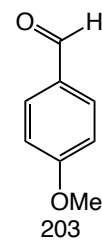
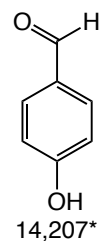
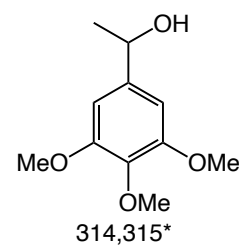
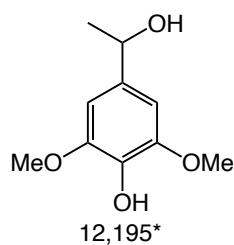
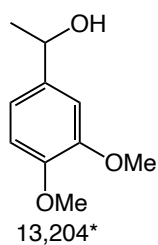
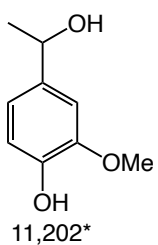
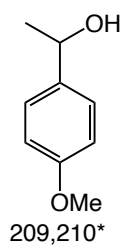
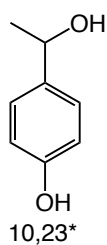
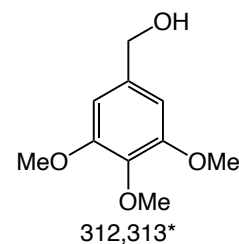
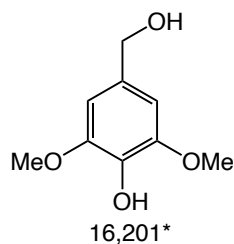
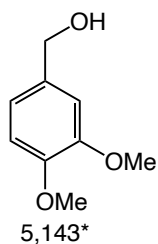
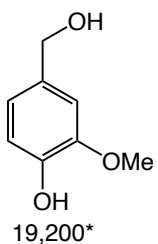
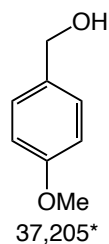
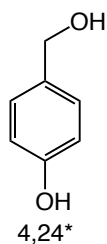
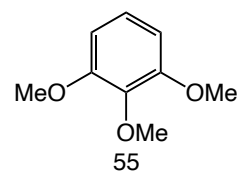
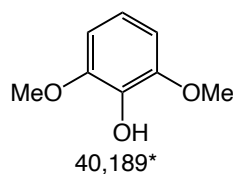
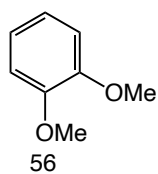
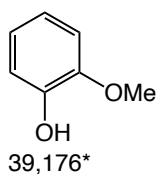
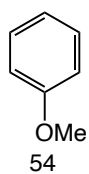
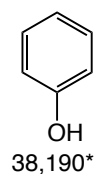
Misc.

- Including stilbenes and phenylglucosides SI 15

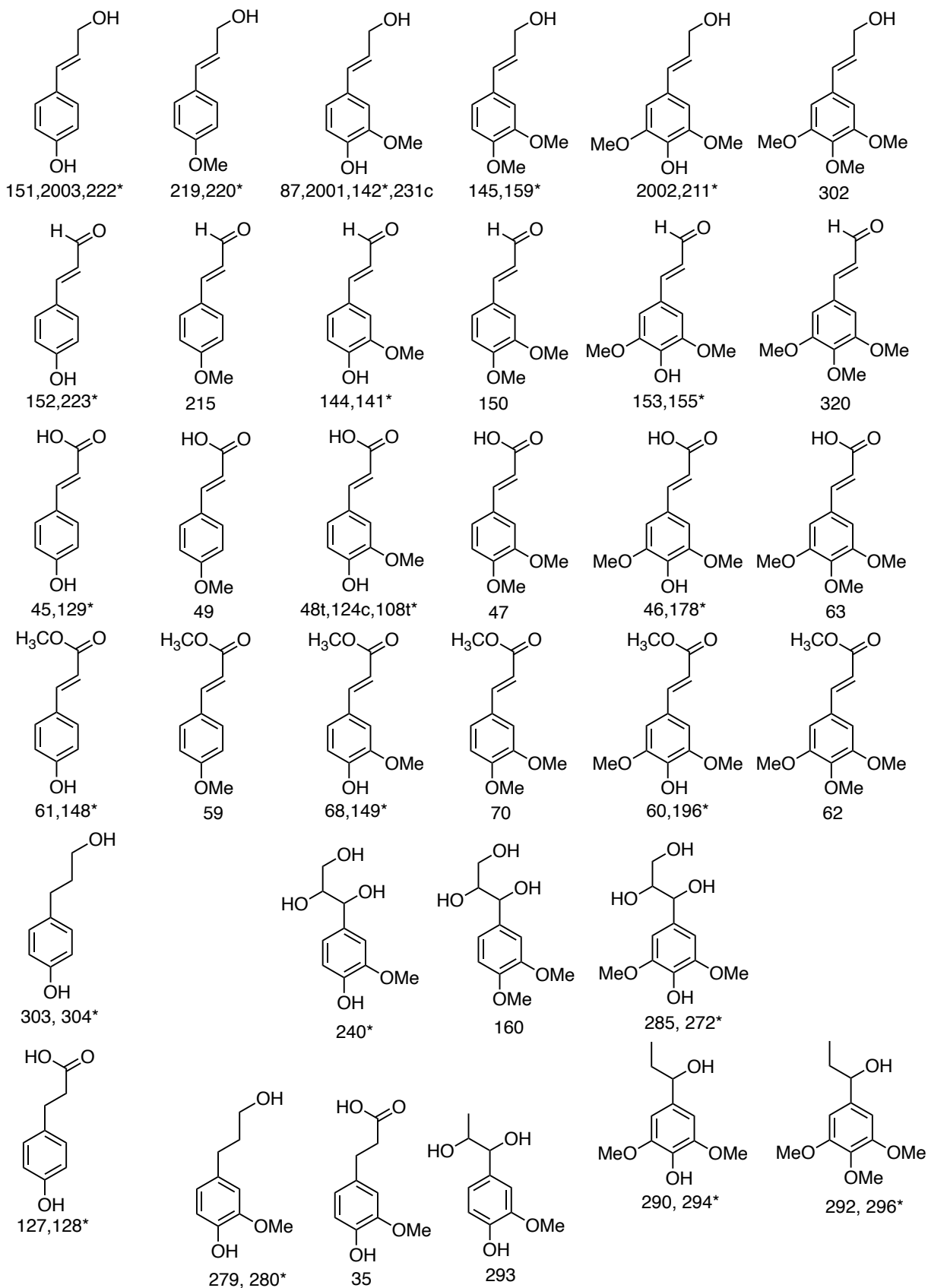
Misc. ctd.

- Including benzoates, quinones, 4-O-benzyl models and β -O-4/ α -5 SI 16

Monomers

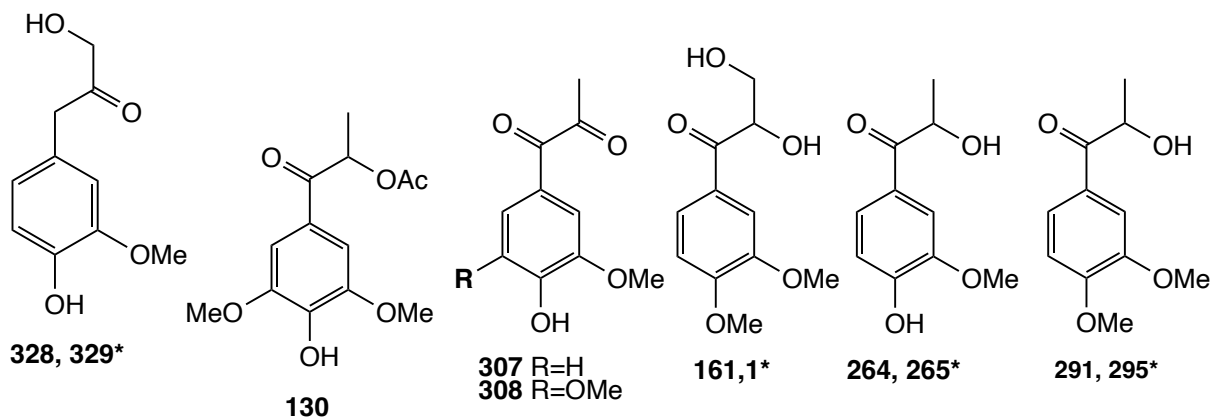


Monomers - 3C SC

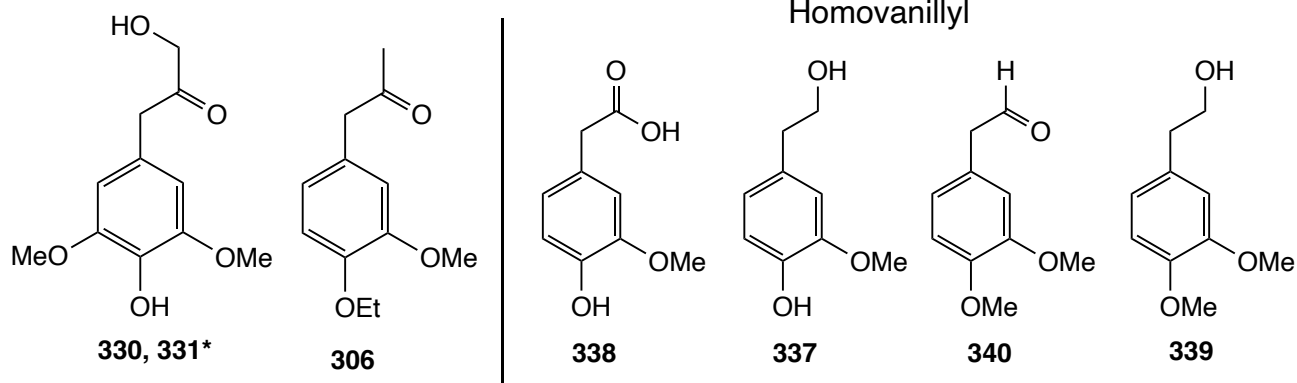


Monomers - lignin degradation

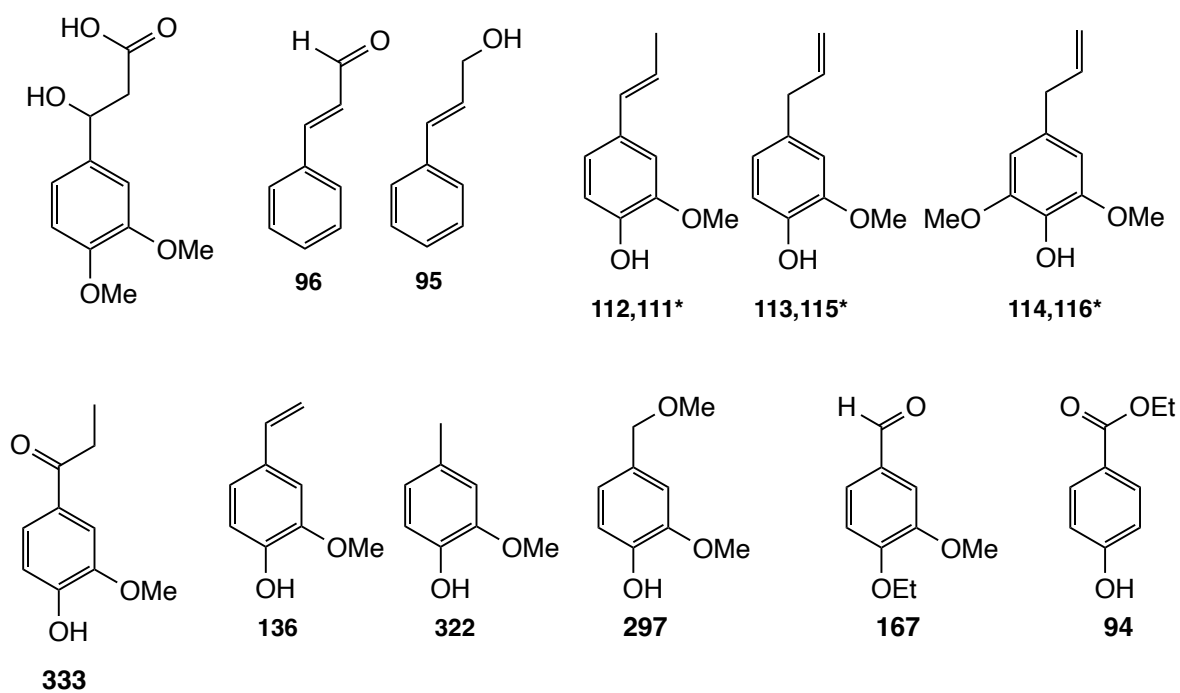
Hibbert's ketones



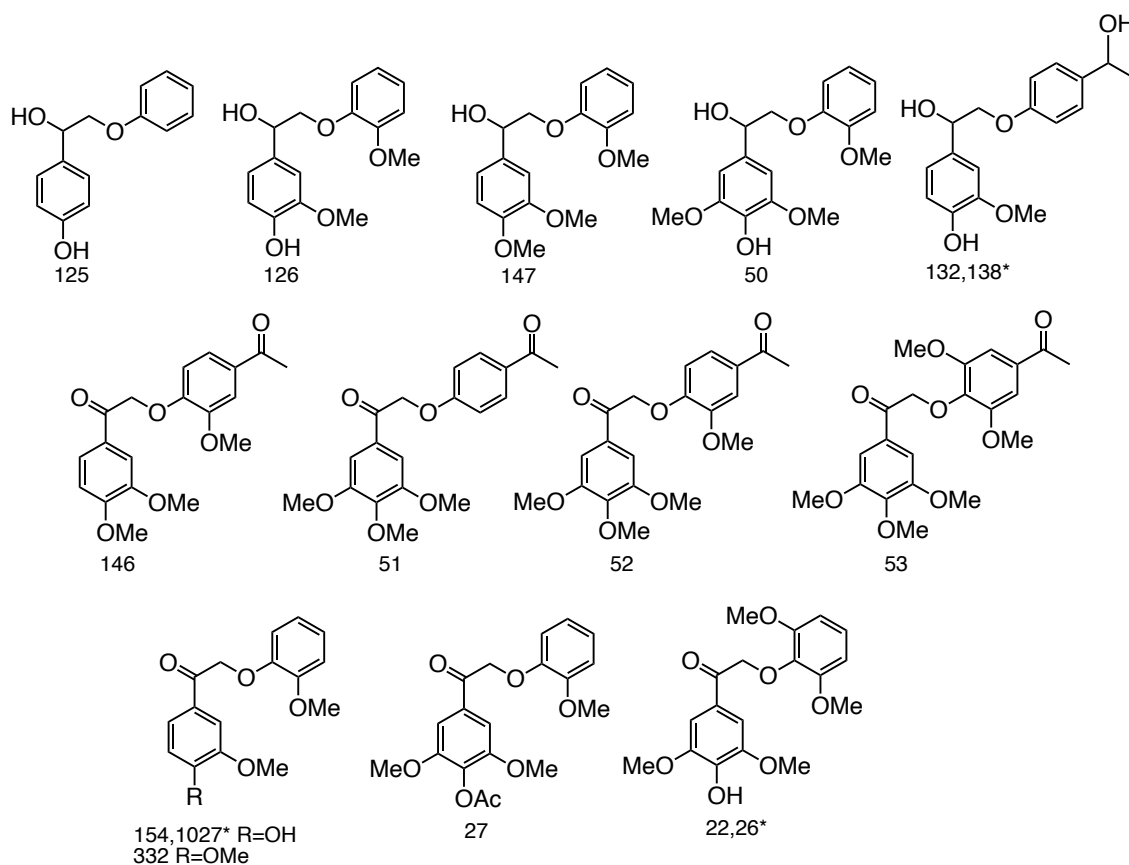
Homovanillyl



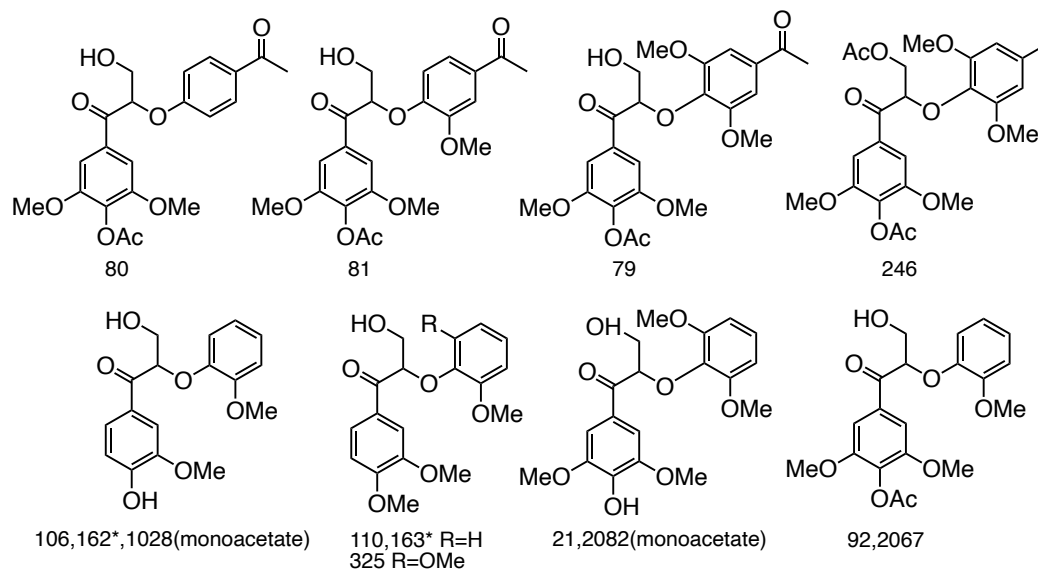
Monomers - Misc.



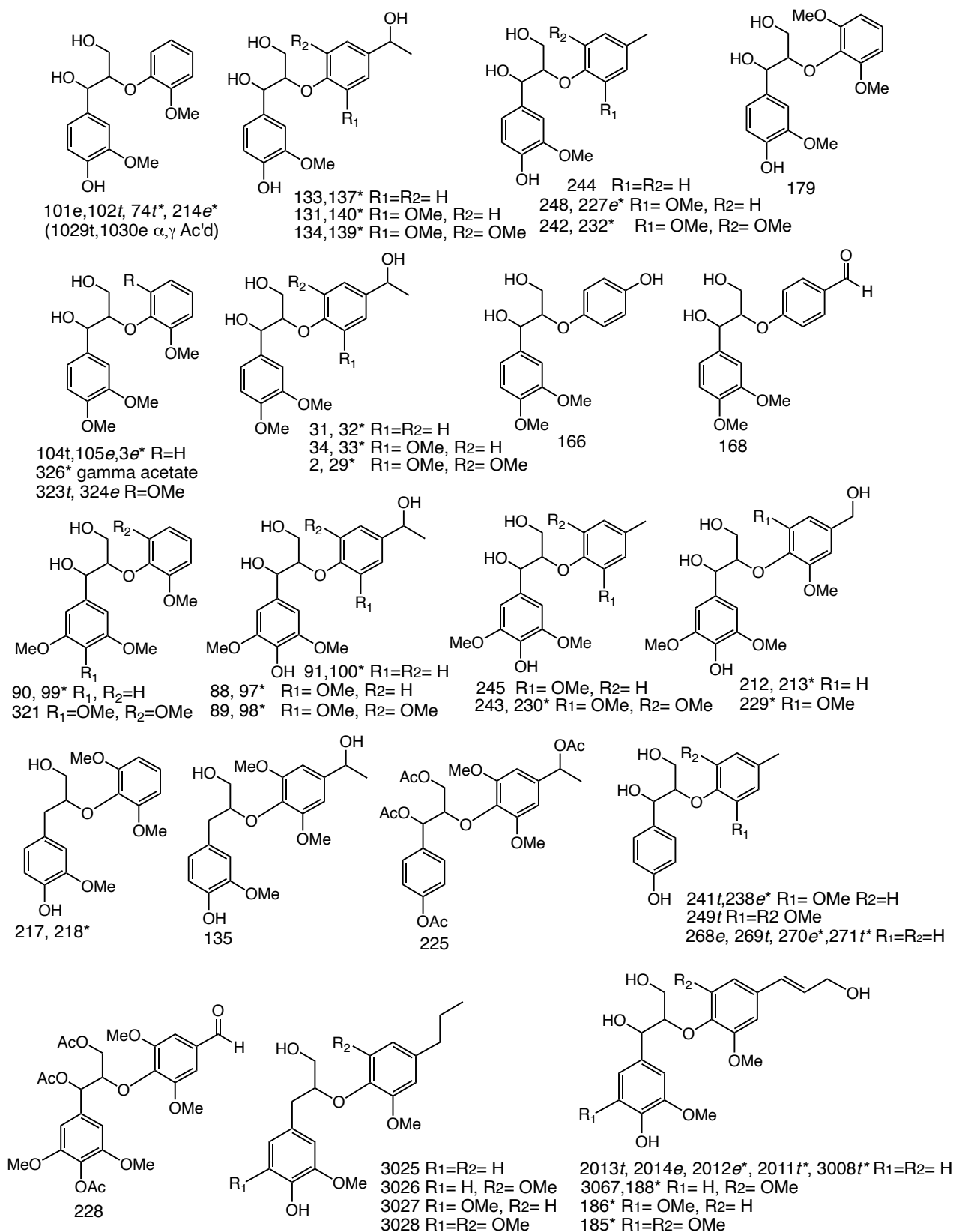
β -O-4 Dimers, 2C SC



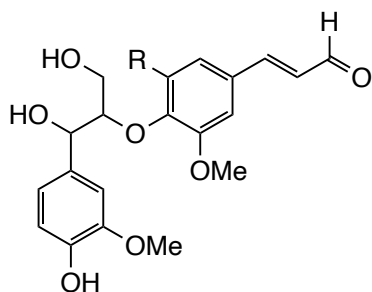
β -O-4 Dimers, 3-Carbon Sidechain, α -C=O



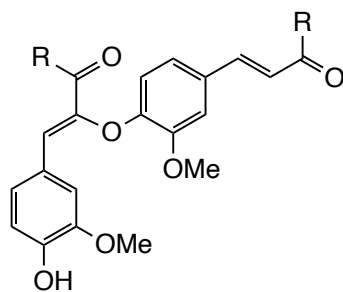
β-O-4 Dimers, 3-Carbon Sidechain



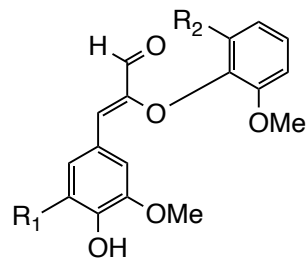
More β -O-4 Dimers, 3-Carbon Sidechain



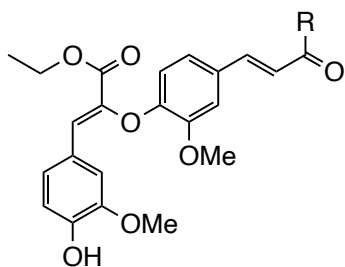
3010*t**, 3011*e** R = H
3071 R = OMe (with only γ -Ac)



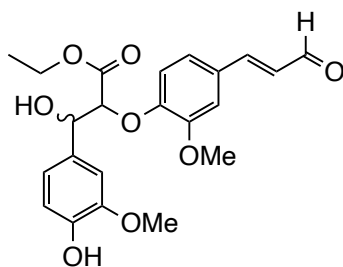
3031 R = H
2040 R = OH



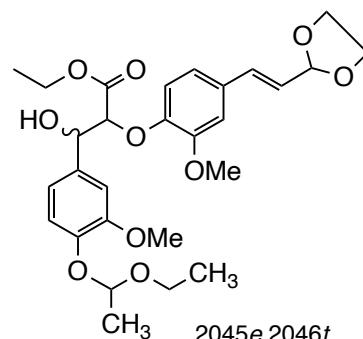
3034, 3035* R₁, R₂ = H
3036, 3037* R₁ = OMe, R₂ = H
3038, 3039* R₁ = H, R₂ = OMe
3040, 3041* R₁, R₂ = OMe



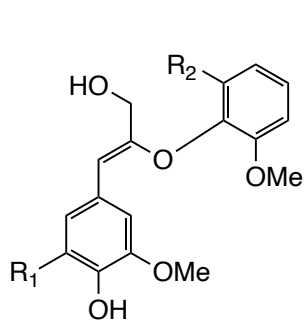
2041 R = OMe
2042 R = H



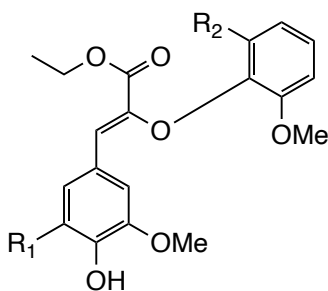
2043*e*, 2044*t*



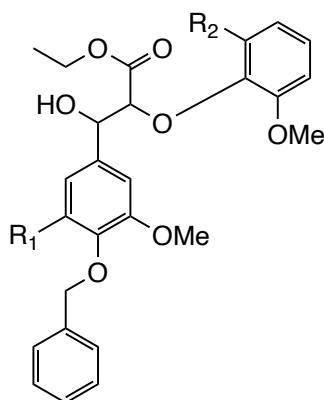
2045*e*, 2046*t*



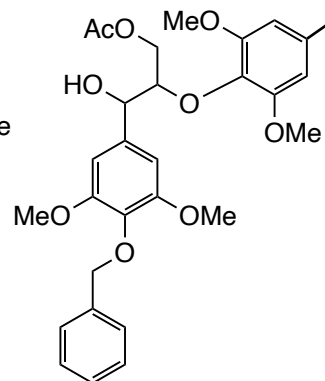
3053, R₁ = H, R₂ = OMe
3057, R₁ = OMe, R₂ = H



3052, R₁ = H, R₂ = OMe
3056, R₁ = OMe, R₂ = H

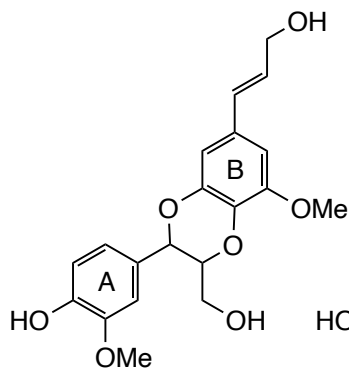


3050, 3051 isomers R₁ = H, R₂ = OMe
3054, 3055 isomers R₁ = OMe, R₂ = H

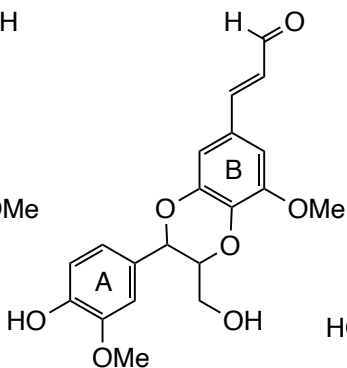


3079

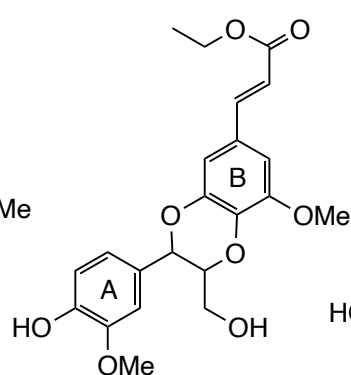
5-Hydroxyconiferyl alcohol β -O-4 Dimers



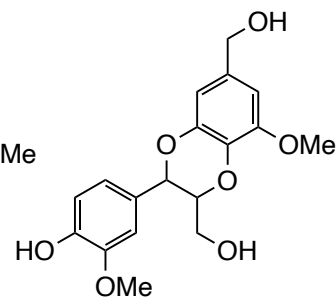
3068



3069

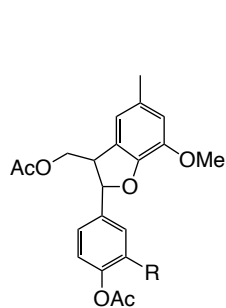


3070

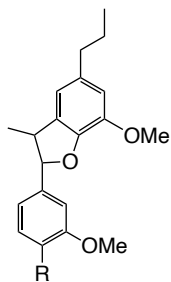


3043, 3044*

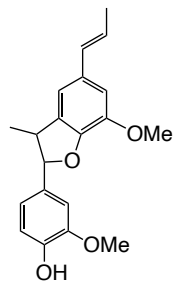
β-5 Dimers



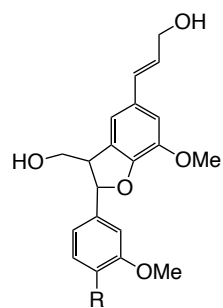
239 R= H
234 R= OMe



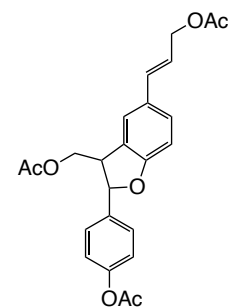
66 R= OH
73 R= OMe



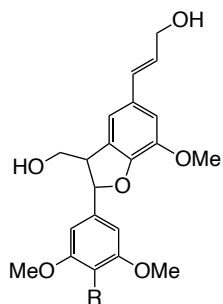
67



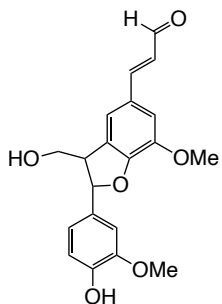
2004,2005* R= OH
2051,221* R= OMe



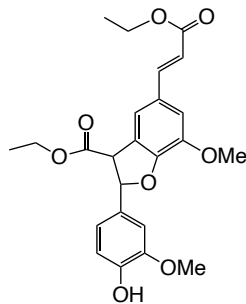
267



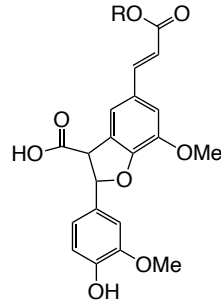
3063,187* R= OH
3074, 3075* R= OMe



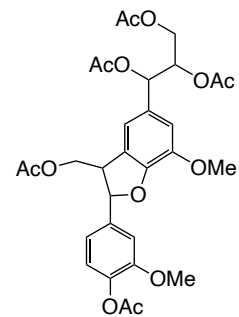
2021



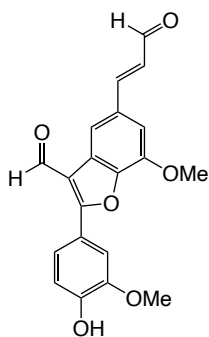
2006



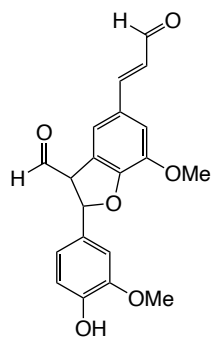
2018 R= H
2052 R= Et



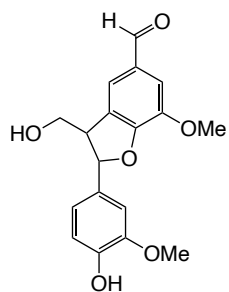
262



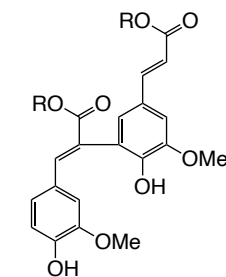
3073



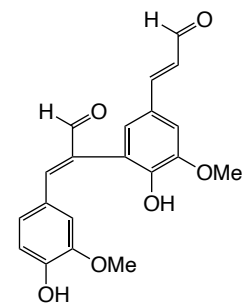
3059



3061

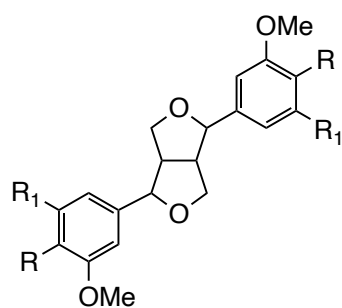


2054 R= H
2019 R= Et

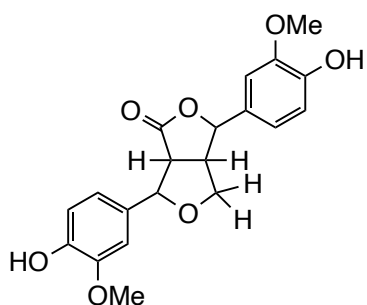


3030

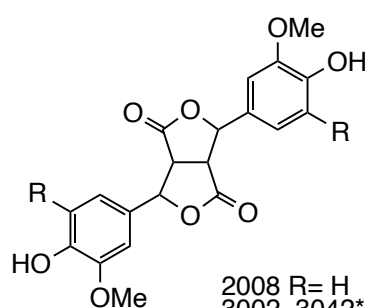
β-β Dimers



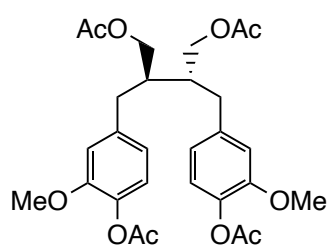
2020, 109* $R_1 = H$
 117, 123* $R_1 = OMe$
 305 $R = OMe$ $R_1 = H$



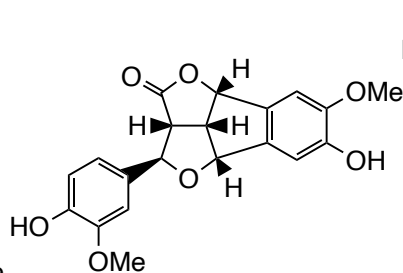
2029, 2030*
 2028 (epimers)
 2027



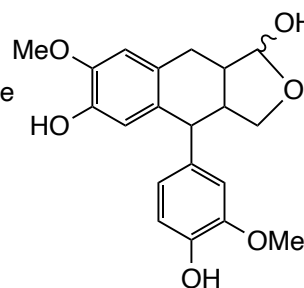
2008 $R = H$
 3002, 3042* $R = OMe$



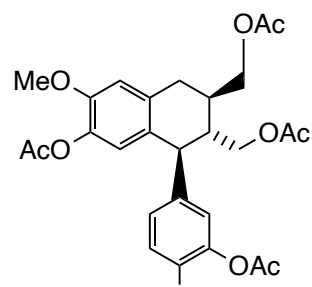
2070



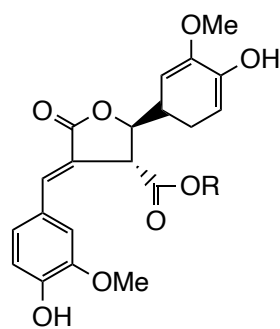
2026



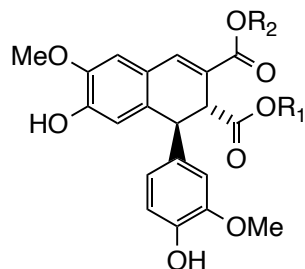
3014, 3015 isomers



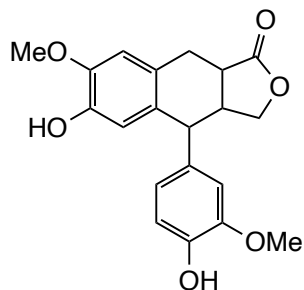
2069



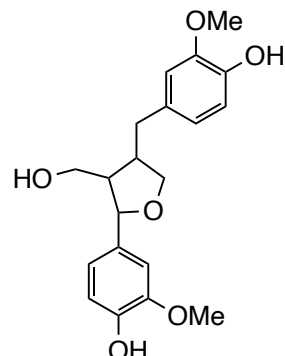
2033 $R = H$
 2034 $R = OMe$



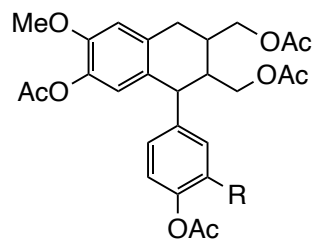
2035 $R_1, R_2 = CH_3$
 2036 $R_1, R_2 = H$
 2062 $R_1 = CH_3, R_2 = H$



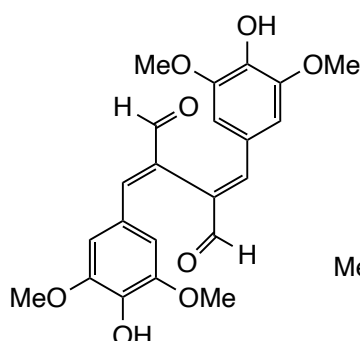
3017, 3016*



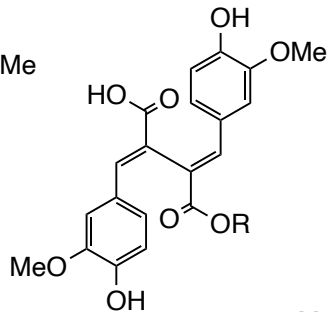
334, 335



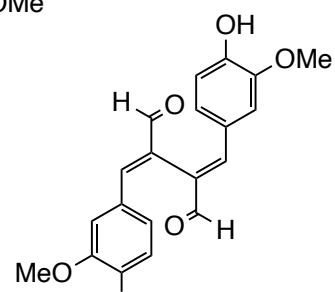
2069 $R = OAc$
 3018 $R = OMe$



3060



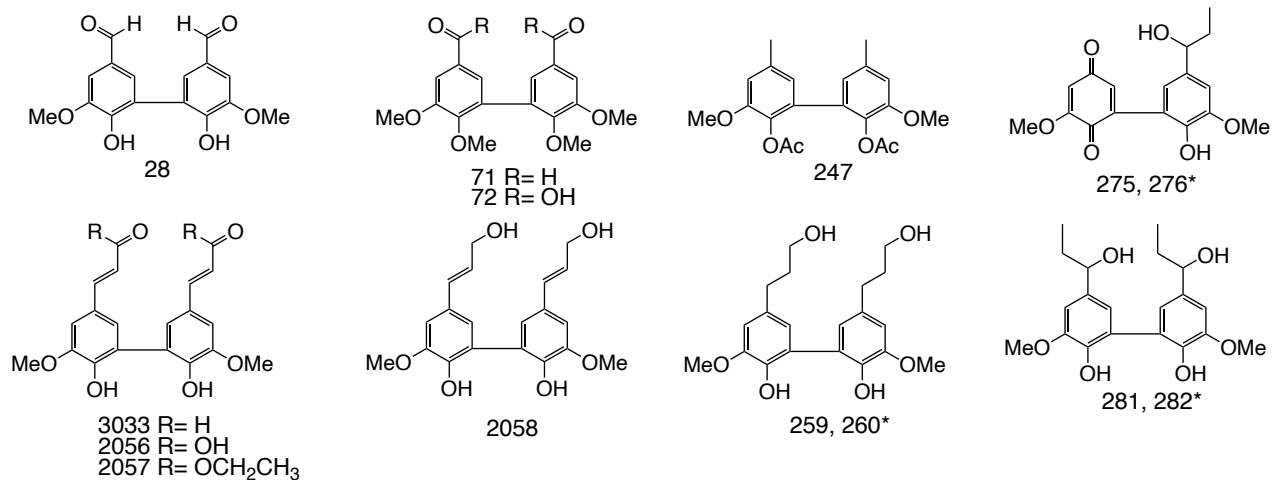
2031 $R = H$
 2032 $R = CH_3$



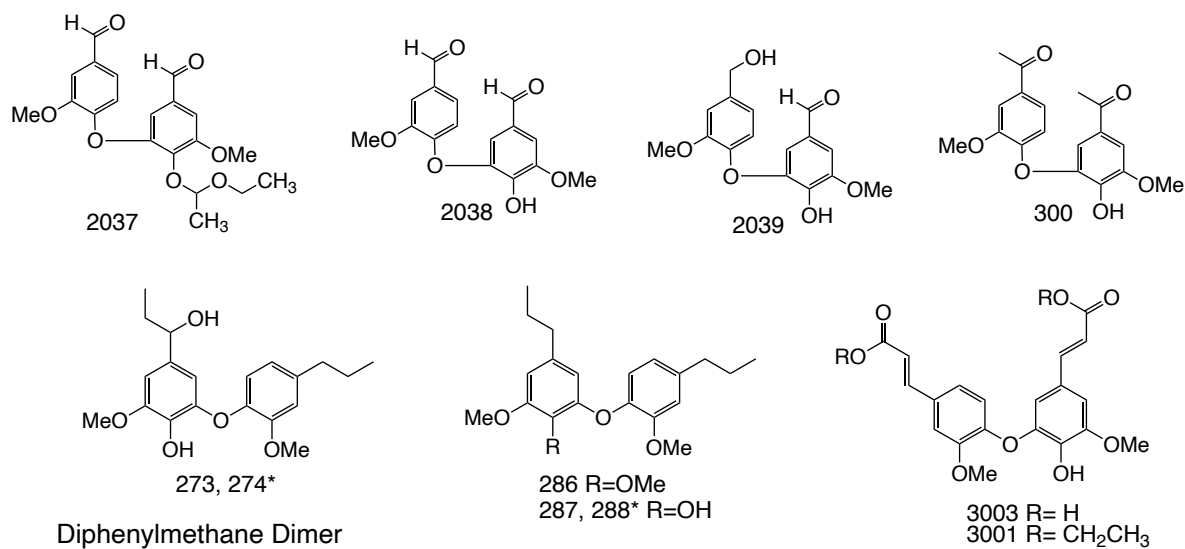
3032

5-5, 5-O-4, β -1, Di-Ph

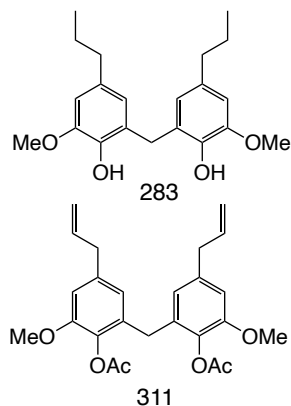
5-5 Dimers



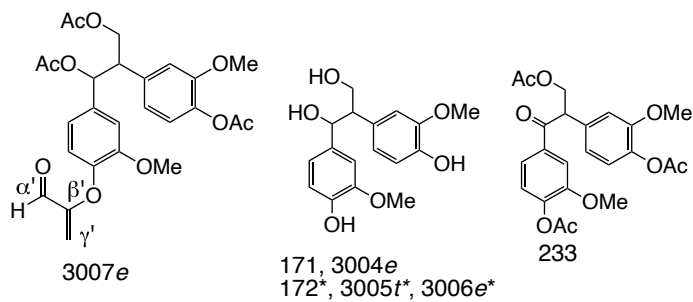
5-O-4 Dimers



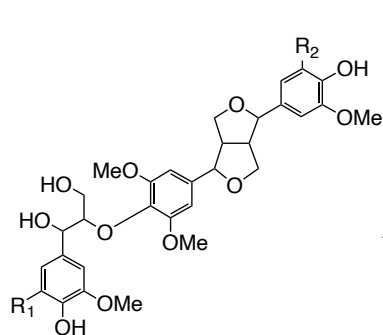
Diphenylmethane Dimer



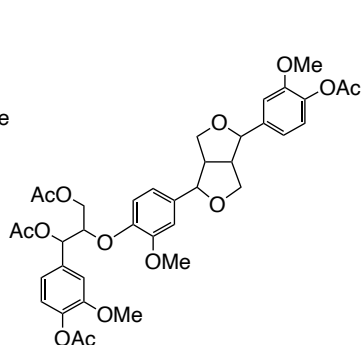
β -1 Dimers



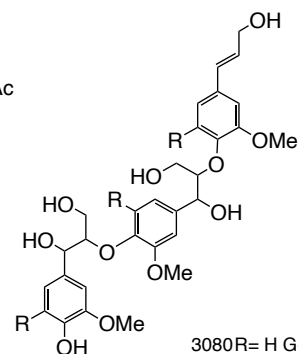
Trimers



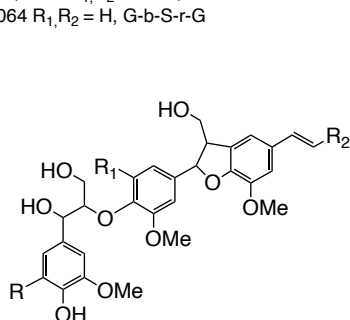
183,184* R₁ = H, R₂ = OMe, G-b-S-r-S
198,199* R₁,R₂ = OMe, S-b-S-r-S
3064 R₁,R₂ = H, G-b-S-r-G



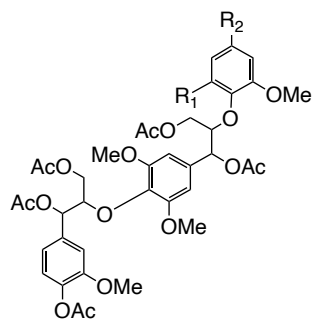
263 G-b-G-r-G



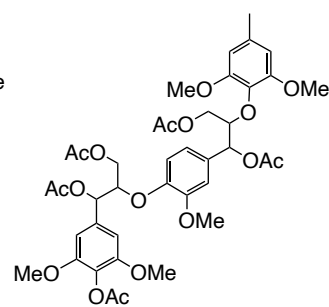
3080 R = H G-b-G-b-CA
193* R = OMe S-b-S-b-SA



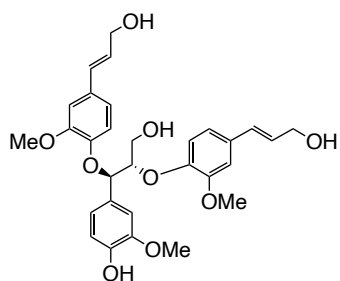
181,182* R = H, R₁ = OMe, R₂ = CH₂OH G-b-S-c-CA
216* R = H, R₁ = H, R₂ = CH₂OH G-b-G-c-CA
3065 R = H, R₁ = OMe, R₂ = CHO G-b-G-c-CAld
3072 R = OMe, R₁ = H, R₂ = CH₂OH S-b-G-c-CA



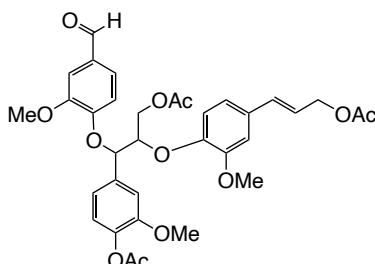
226 R₁ = R₂ = H G-b-S-b-G
250 R₁ = OMe R₂ = CH₃ G-b-S-b-S



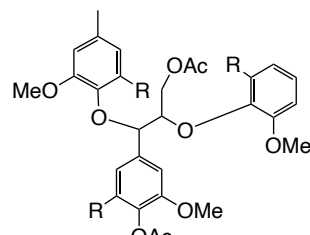
261 S-b-G-S



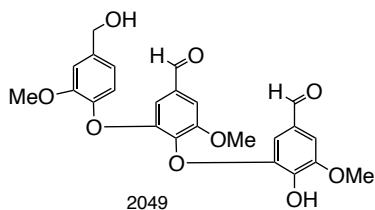
2015e,2022e* CA-a-G-b-CA



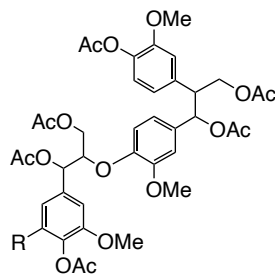
258



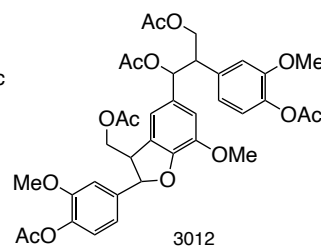
237 R = H G-a-G-b-G
266 R = OMe S-a-S-b-S



2049

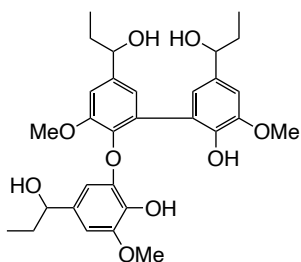


236, 3013 R = H G-b-G-b1-G
235 R = OMe S-b-G-b1-G

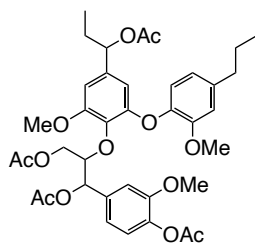


3012

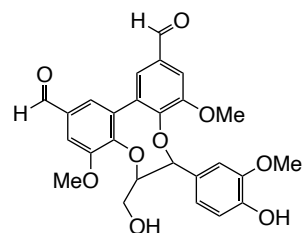
Trimers Cont'd



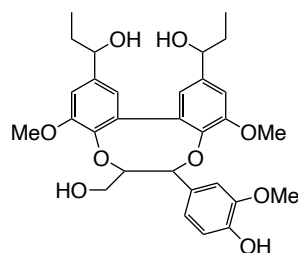
284, 289*



299

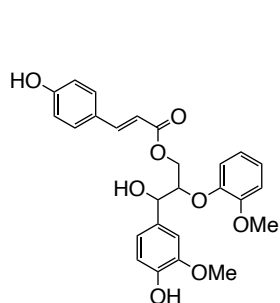


3062

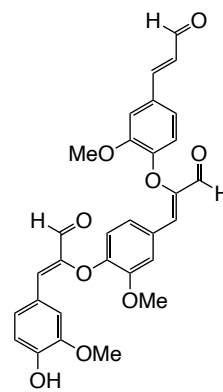


277, 278*

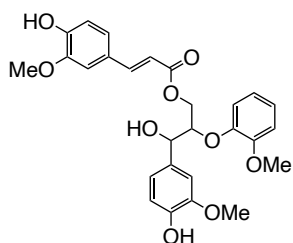
Trimers Containing Ferulate, Coumarate or Benzoate



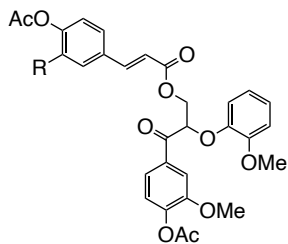
1015t, 1016e, 1019t*, 1020e*, 82t*, 83e*
1011t, 1012e (diacetates, phenolic)



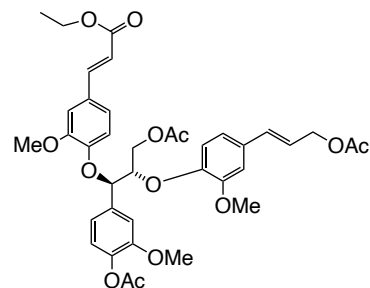
3058



1017t, 1018e, 1021t*, 1022e*, 76e*, 75t*
1013t, 1014e (diacetates, phenolic)

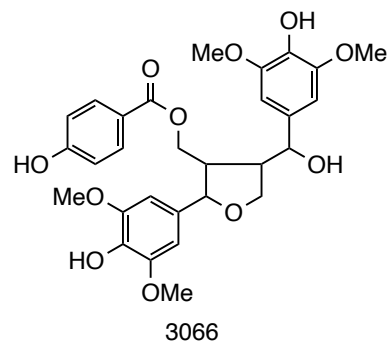
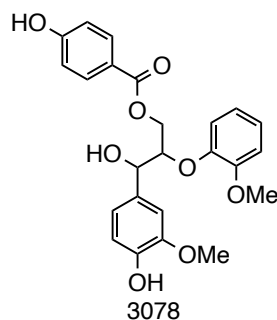
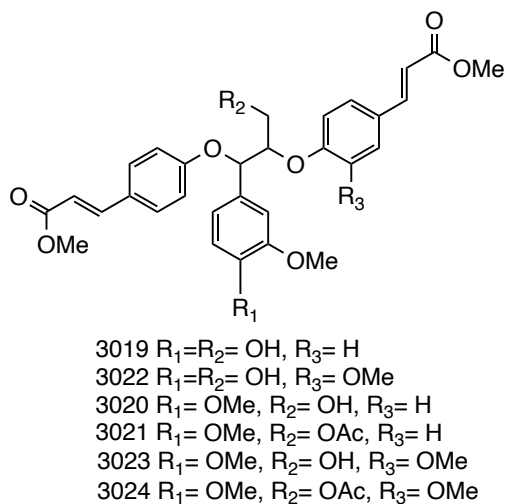
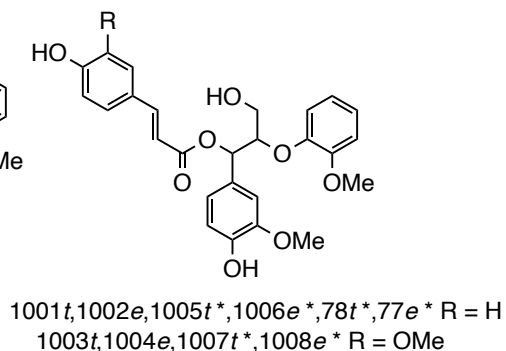
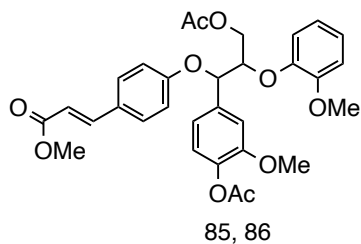
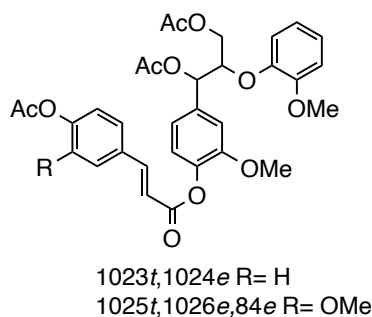
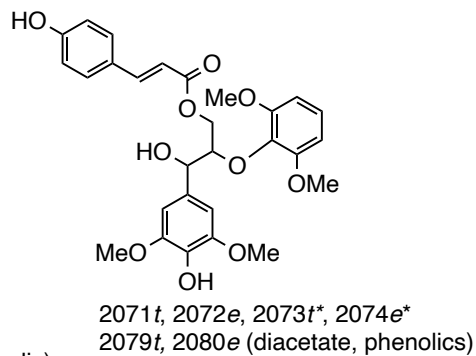
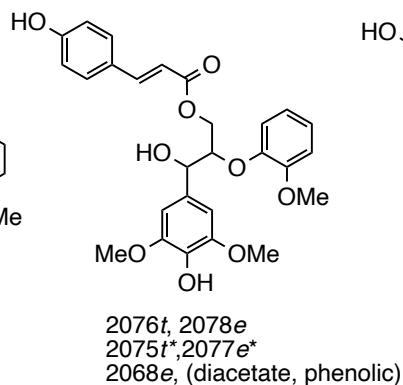
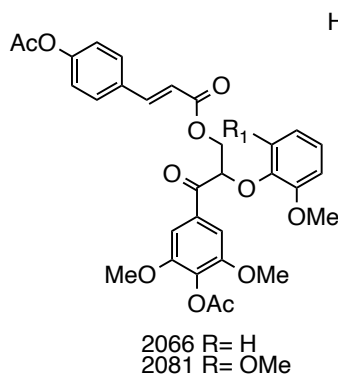


1009 R= H
1010 R=OMe

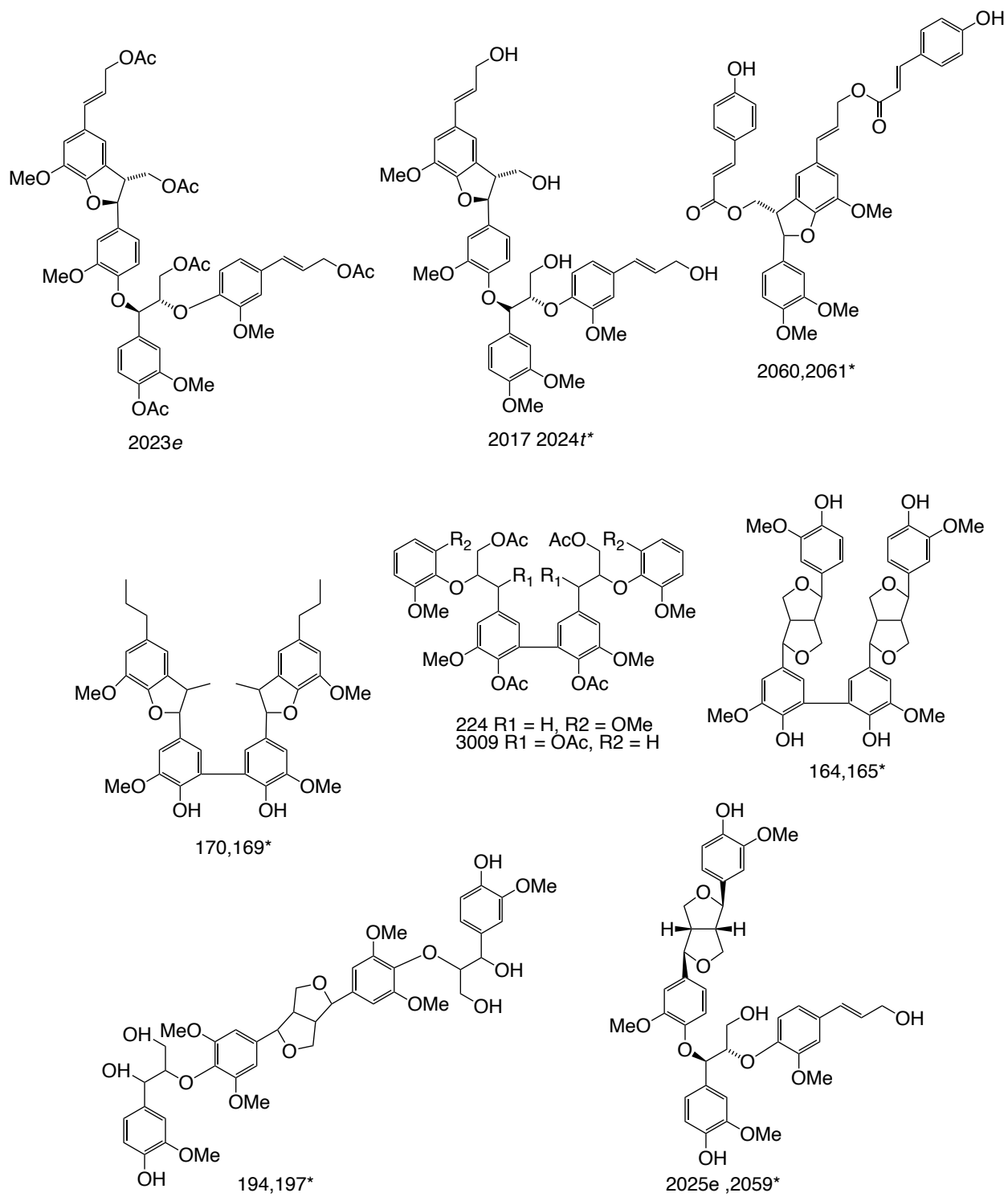


2016

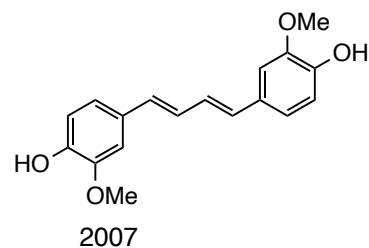
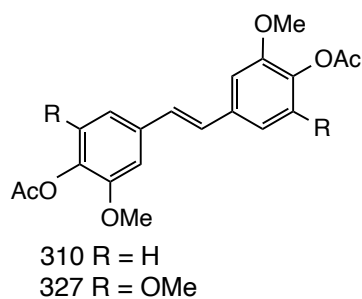
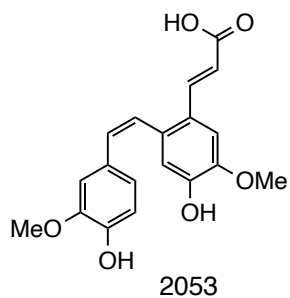
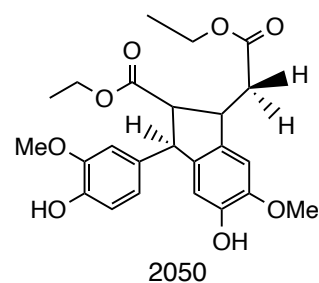
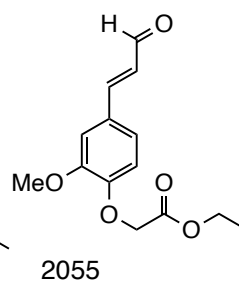
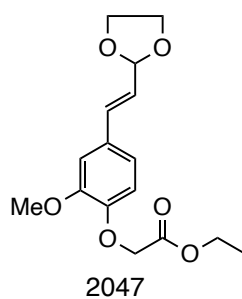
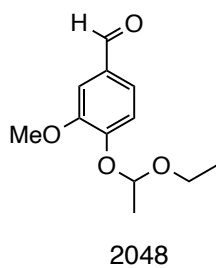
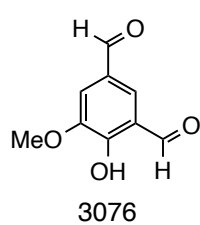
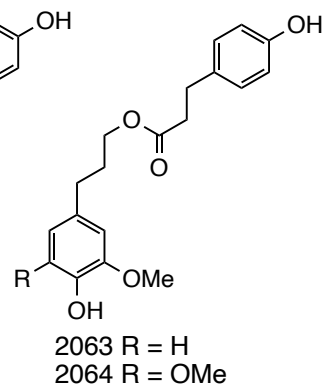
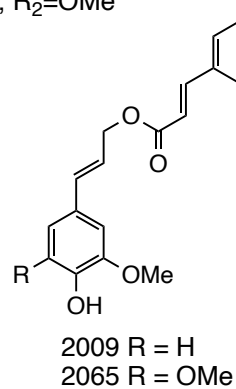
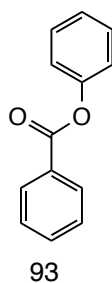
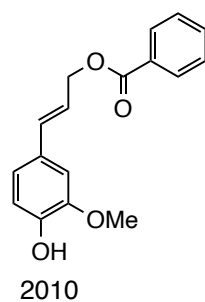
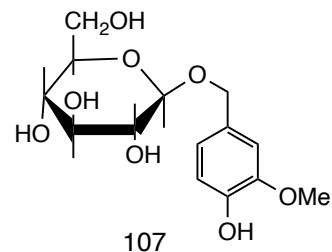
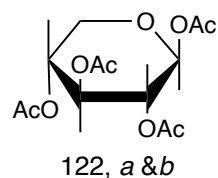
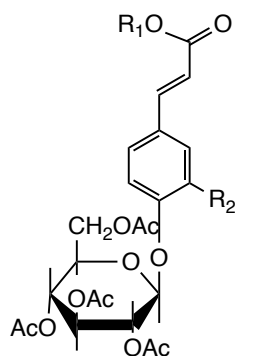
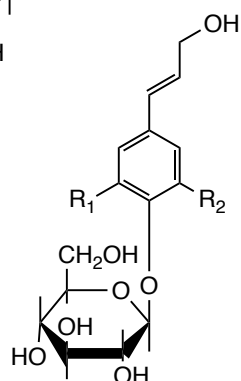
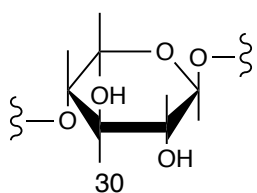
More Trimers Containing Ferulic, Coumaric or *p*-OH-Benzoic Acid



Tetramers

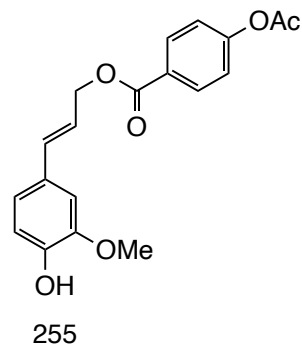
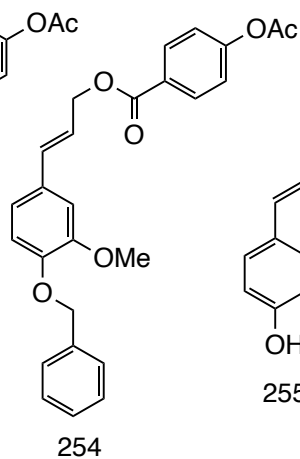
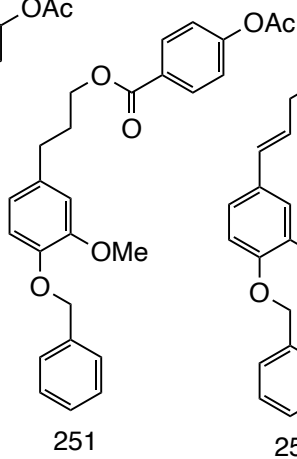
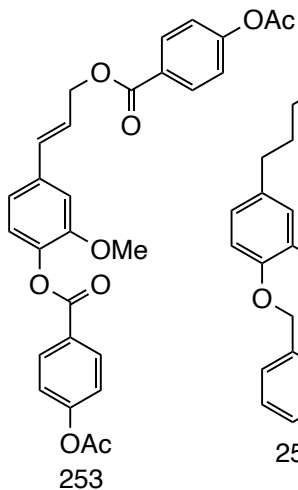
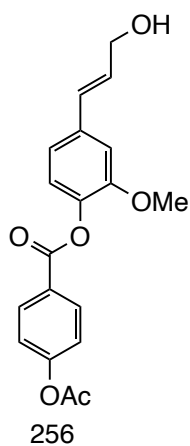


Misc. Compounds

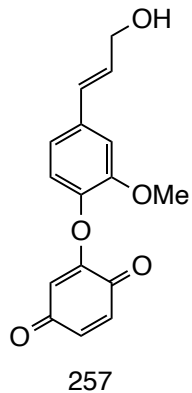
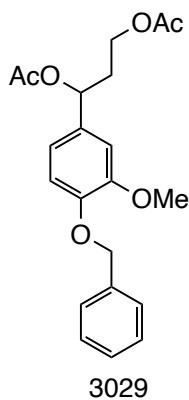
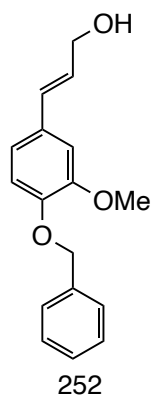


Misc. Compounds

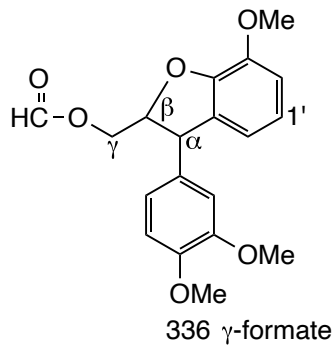
Benzoates



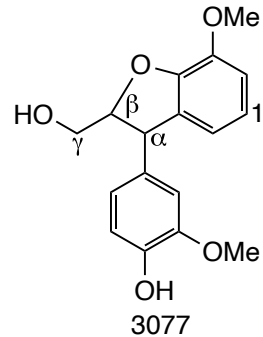
benzylated



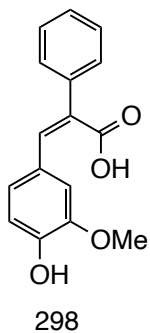
p-quinone



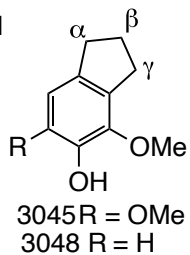
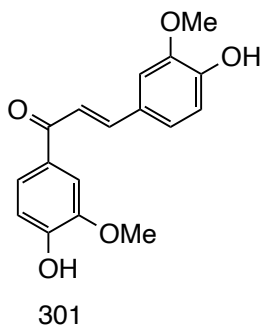
α -5- β -O-4



Stilbene carboxylic acid



Chalcone



Thioacidolysis products

