

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
email: saralph@wisc.edu
608-231-9449

John Ralph

Department of Biochemistry and the
Great Lakes Bioenergy Research Center
Wisconsin Energy Institute
University of Wisconsin, Madison
1552 University Ave
Madison, WI 53726-4084
(608) 890-2429
E-mail: jralph@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 different formats: a FileMaker Pro© database for cross-platform use, an Adobe© pdf cross-platform file for viewing and printing, and a hardcopy version derived from the FileMaker Pro database. FileMaker Pro and pdf versions are available for downloading over the internet from the Dairy Forage Research Center (DFRC) web site:

<http://www.dfrc.ars.usda.gov>
(under the Software section)

A hardcopy version is available by request from the authors at the Forest Products Laboratory, but users are encouraged to print their own version. The use of trade or firm names in this publication is for reader information and does not imply endorsement by the US Department of Agriculture of any product or service.

In general ^{13}C NMR data was collected in three common deuterated solvents (acetone, chloroform and dimethyl sulfoxide) for each compound. We used the center line of the solvent peaks as our reference, 2.04 and 29.83 ppm for acetone- d_6 , 7.24 and 77.00 for CDCl_3 and 2.49 and 39.50 ppm for $\text{DMSO-}d_6$. The ^1H NMR data early on was reported for only one solvent. 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 MHz spectrometer at FPL and those compounds with an index number between 1000 and 10,000 were run at the DFRC on a Bruker 360 MHz instrument. The order of the compounds in the database reflects their arrival at the spectrometer rather than a preordained plan. The inclusion of analogous series of 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.
jralph@wisc.edu

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.

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

Table 2. Terminology of Abbreviated Structural Entities

| Entity | Abbreviation |
|-----------------------------|--------------|
| guaiacyl ring | G |
| syringyl ring | S |
| coumaryl 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 |
| p-coumaryl alcohol end unit | HA |
| ferulic acid end unit | FA |
| <i>erythro</i> | <i>e</i> |
| <i>threo</i> | <i>t</i> |

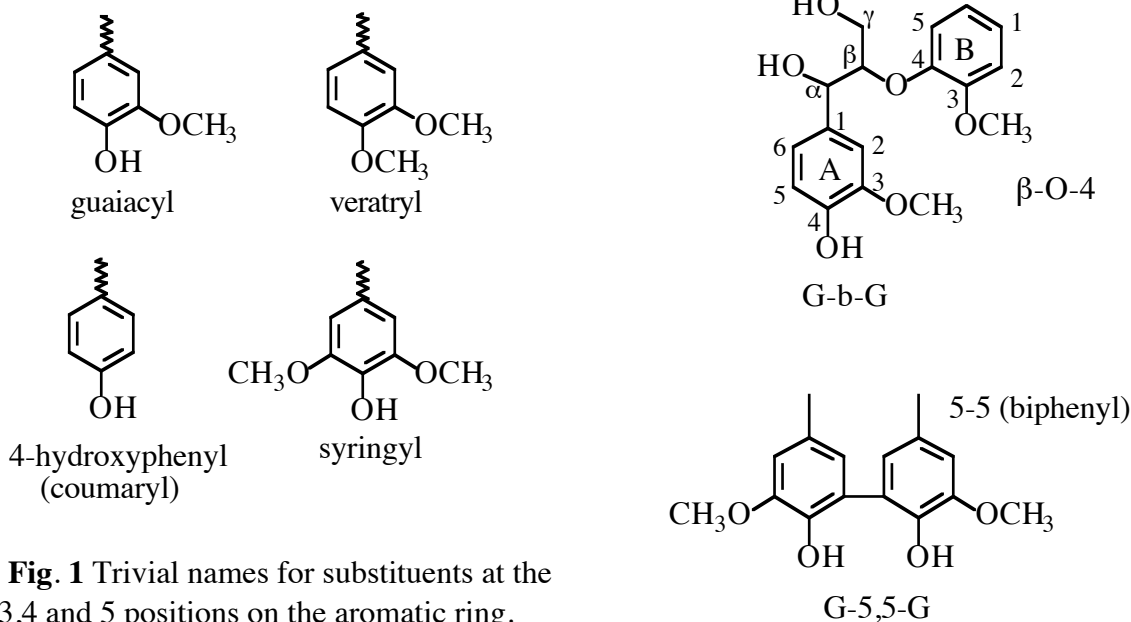


Fig. 1 Trivial names for substituents at the 3,4 and 5 positions on the aromatic ring.

The naming of the larger oligomer lignin models uses a combination of upper case letters to describe the ring structure and lower

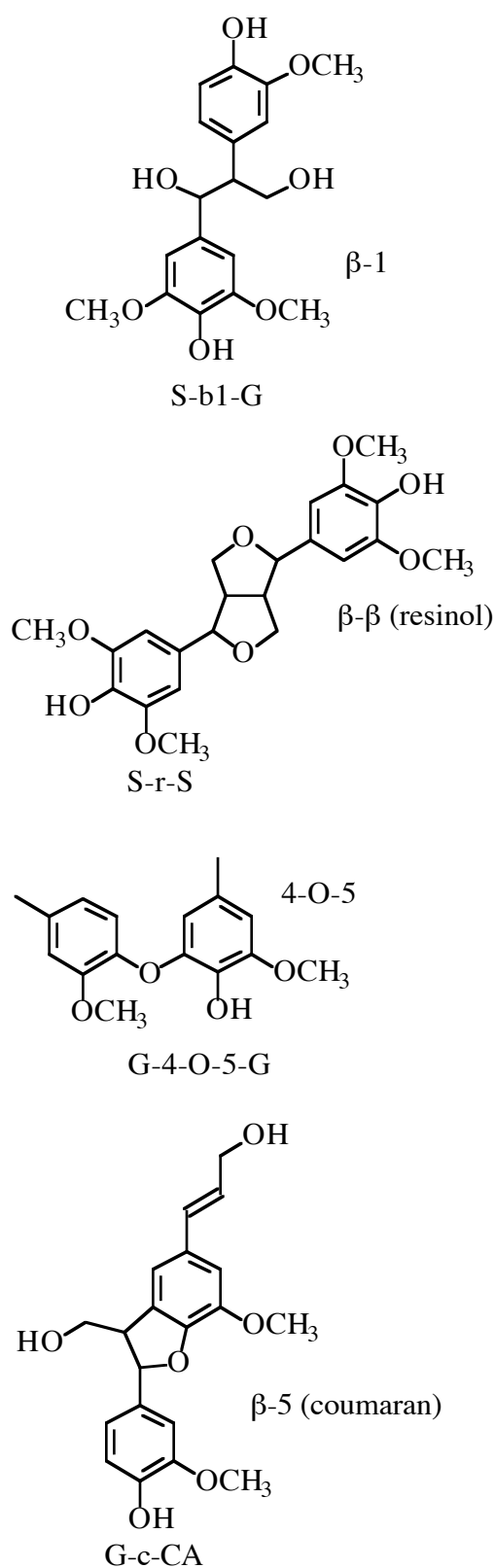


Fig. 2 Examples of linkages and abbreviated names.

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 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 data sheet. An asterisk after a number indicates the acetylated 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.

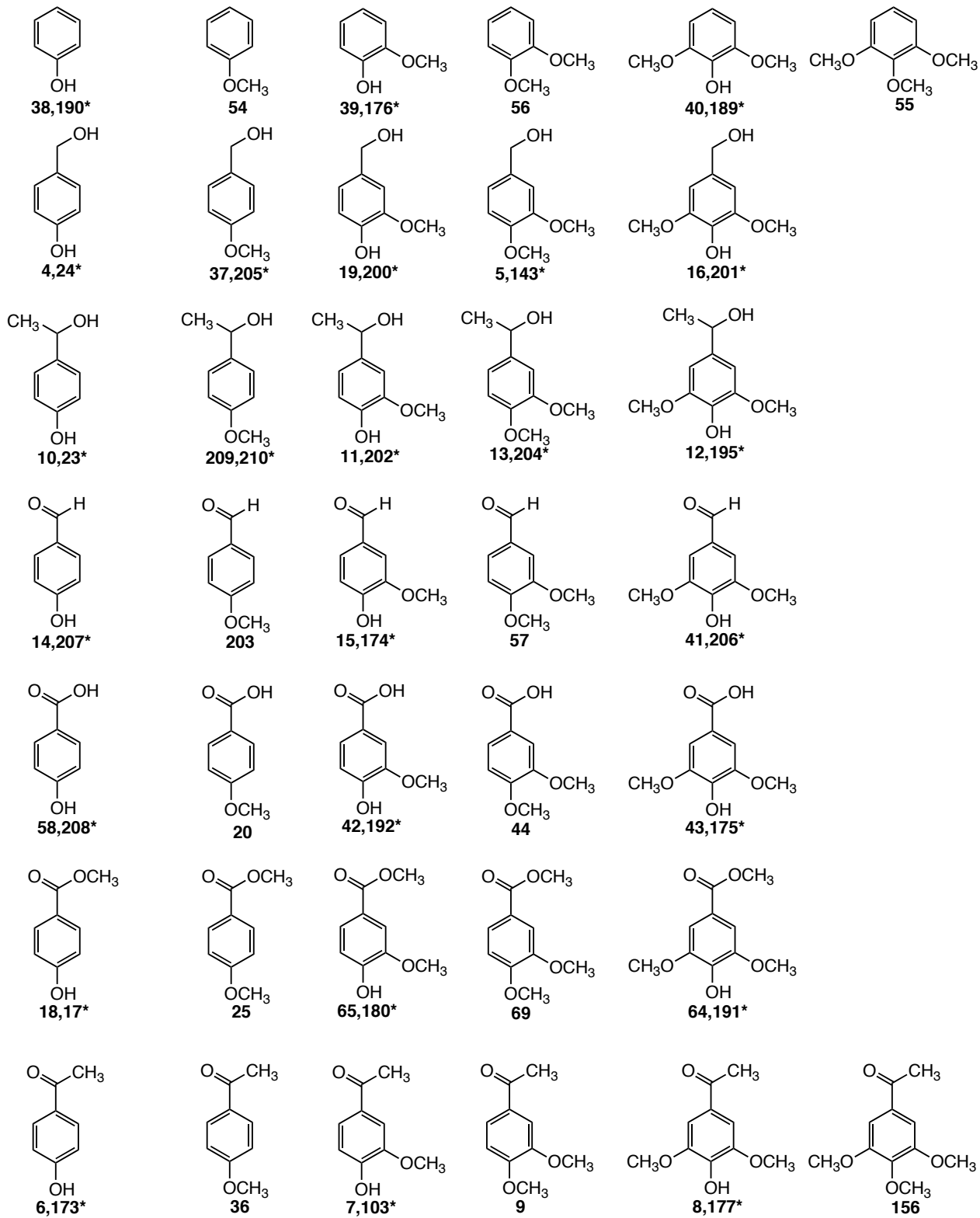
The authors gratefully acknowledge the many generous contributions made by others at both labs towards this database; Foremost Larry Landucci for compounds, patience, assignments and encouragement. William Landucci for software development, Martin Wesolowski and Kolby Hirth, NMR spectroscopists, and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

Structure Index

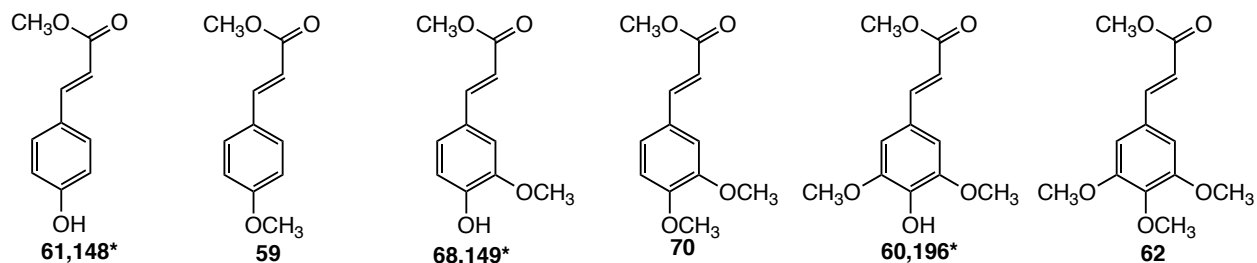
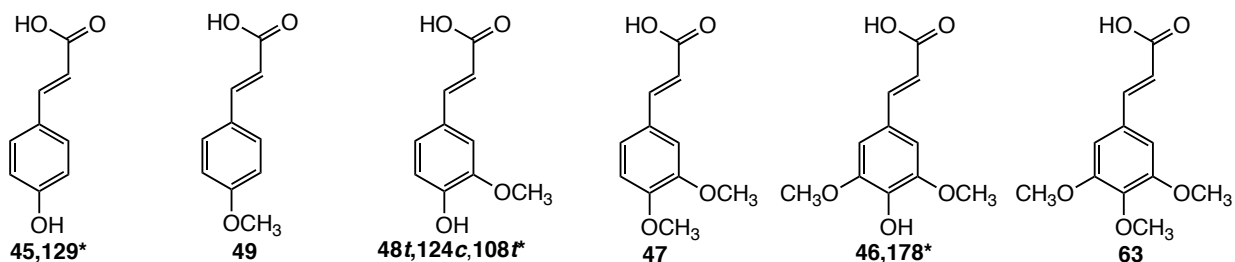
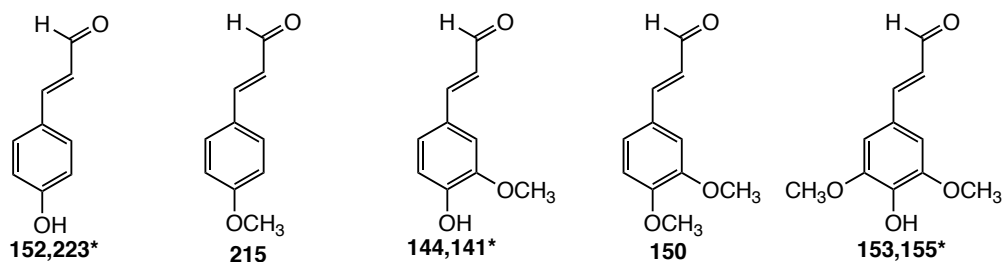
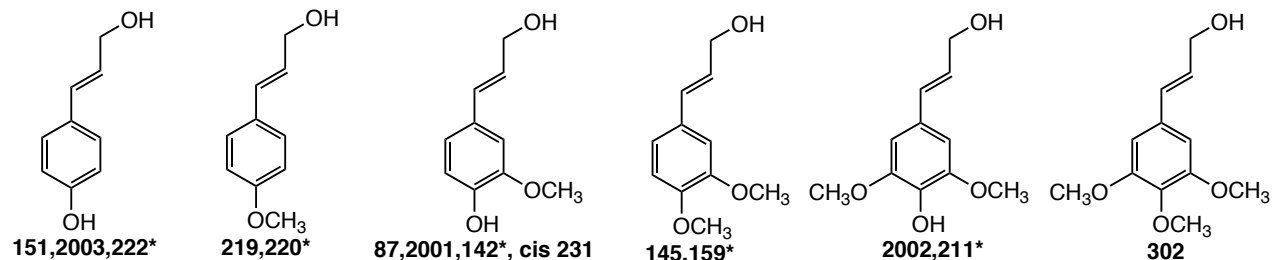
Contents

| | |
|---|-------|
| Monomers | v |
| β -O-4 dimers, 2-carbon sidechain and 3 carbon sidechain with α -C=O | vii |
| β -O-4 dimers, 3-carbon sidechain | viii |
| β -5 dimers | x |
| β - β dimers | xi |
| 5-5 dimers, β -1 dimers , 4-O-5 dimers | xii |
| Trimers | xiii |
| Trimers containing ferulic or coumaric acid | xiv |
| Tetramers | xvi |
| Misc. | xvii |
| Misc. ctd. | xviii |

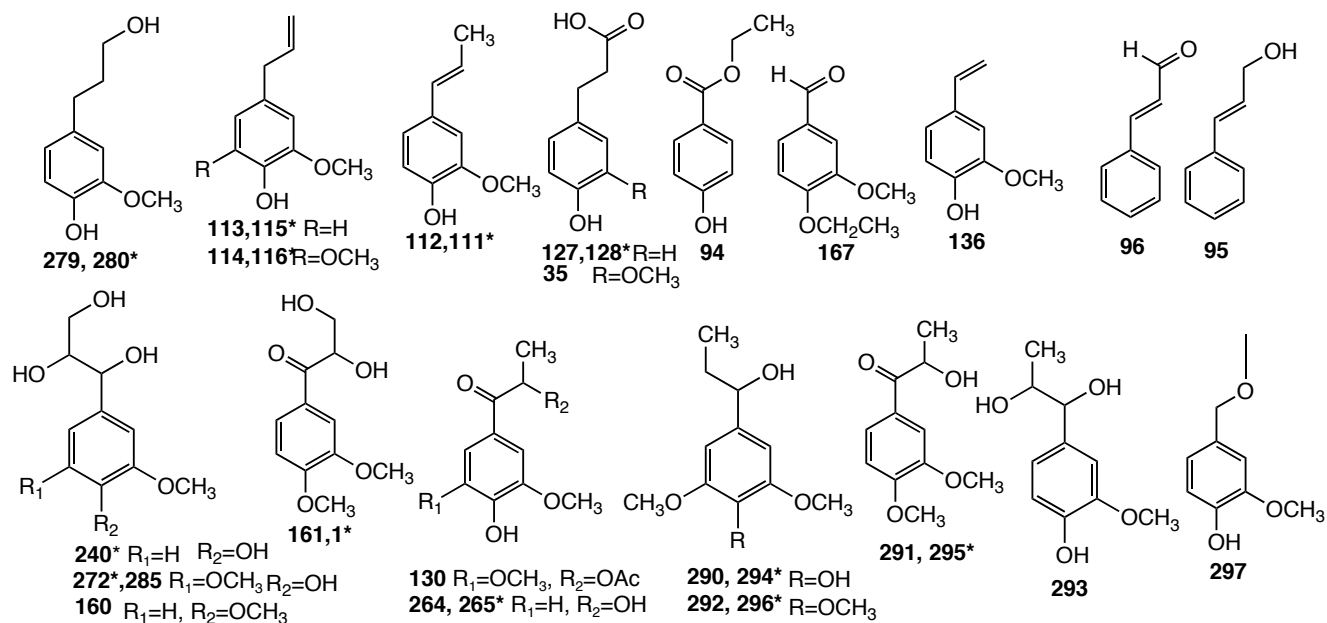
Monomers

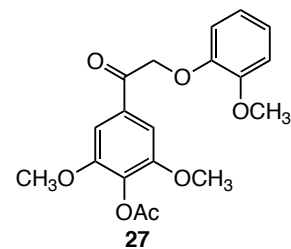
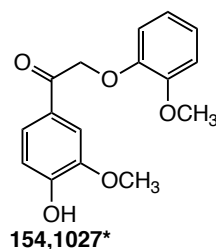
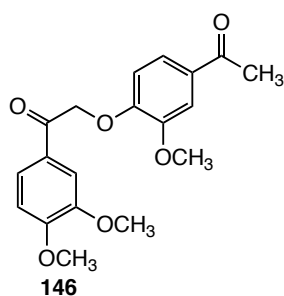
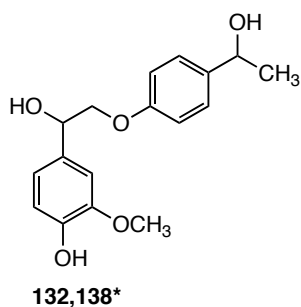
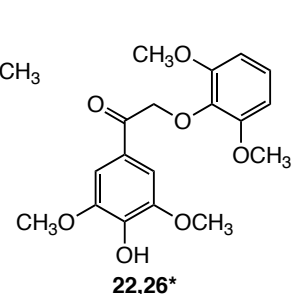
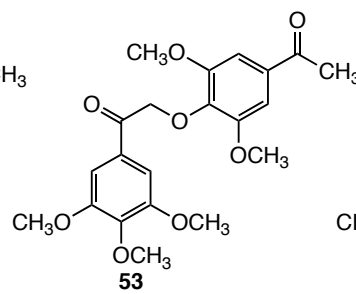
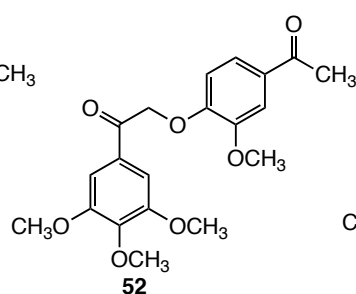
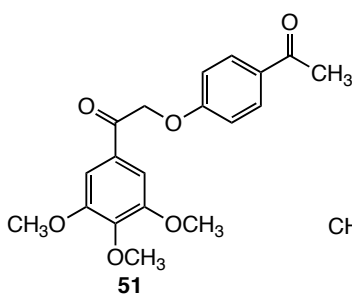
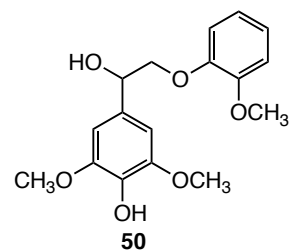
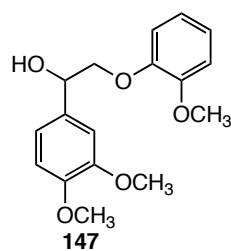
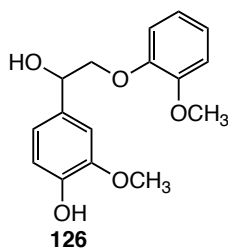
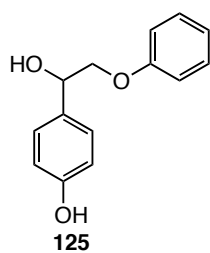
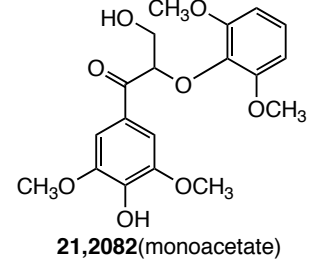
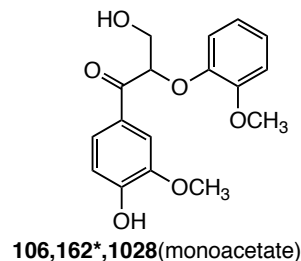
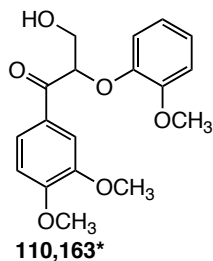
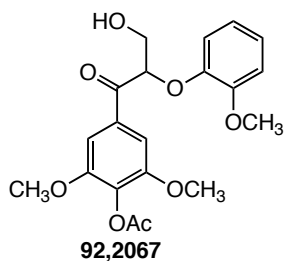
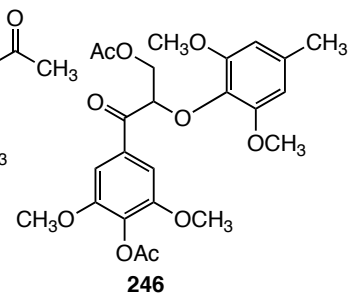
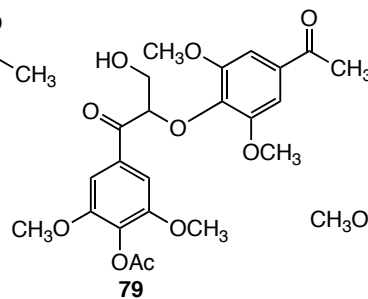
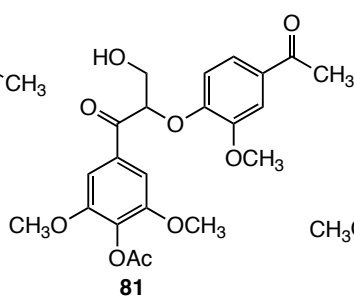
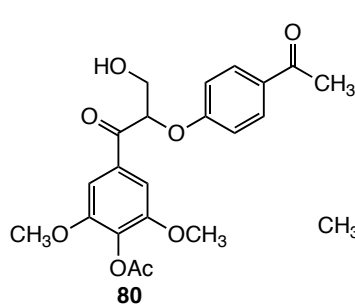


Monomers ctd

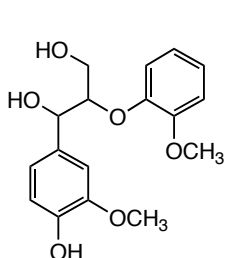


Misc. Monomers

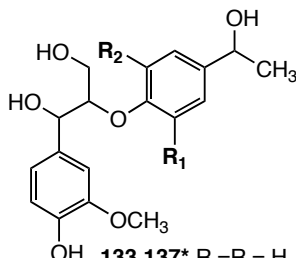


b-O-4 Dimers, 2-Carbon Sidechain**b-O-4 Dimers, 3-Carbon Sidechain, a-C=O**

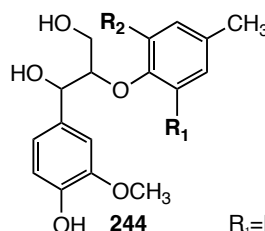
b-O-4 Dimers, 3-Carbon Sidechain



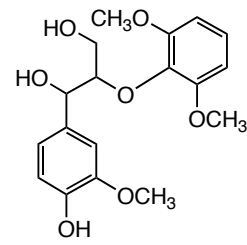
101e,102t,74f*,214e*
(1029t,1030e a,g Ac'd)



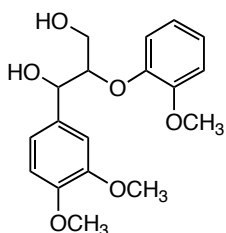
133,137* R₁=R₂=H
131,140* R₁=OCH₃, R₂=H
134,139* R₁=OCH₃, R₂=OCH₃



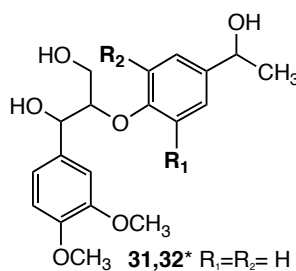
244 R₁=R₂=H
248t,227e* R₁=OCH₃, R₂=H
242,232* R₁=OCH₃, R₂=OCH₃



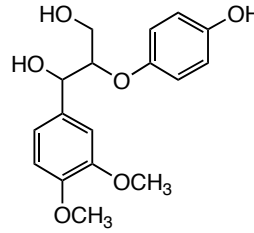
179



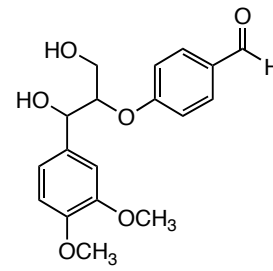
104t,105e,3e*



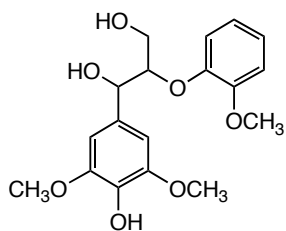
31,32* R₁=R₂=H
34,33* R₁=OCH₃, R₂=H
2,29* R₁=OCH₃, R₂=OCH₃



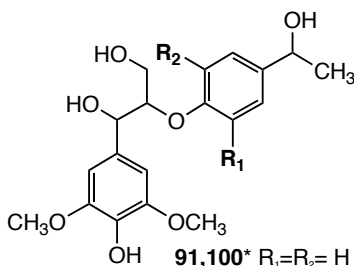
166



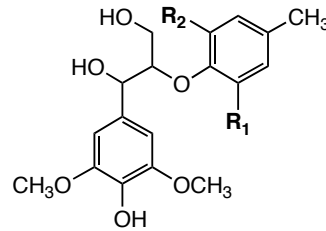
168



90,99*

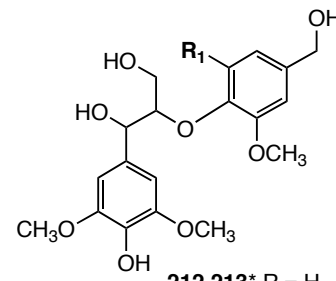


91,100* R₁=R₂=H
88,97* R₁=OCH₃, R₂=H
89,98* R₁=OCH₃, R₂=OCH₃

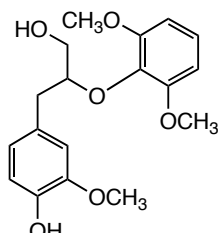


245

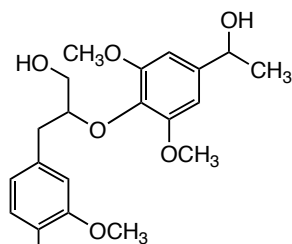
R₁=OCH₃, R₂=H
243,230* R₁=OCH₃, R₂=OCH₃



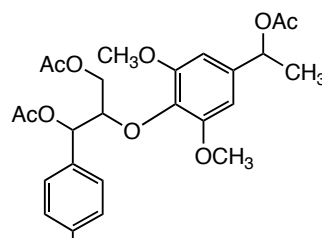
212,213* R₁=H
229* R₁=OCH₃



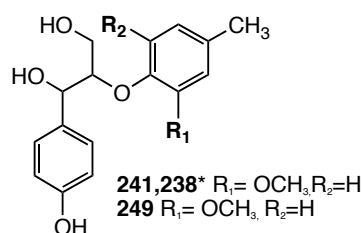
217,218*



135

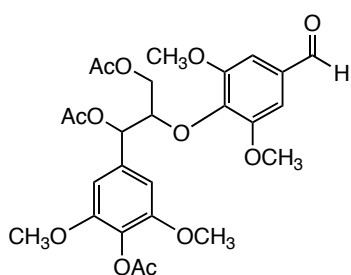


225

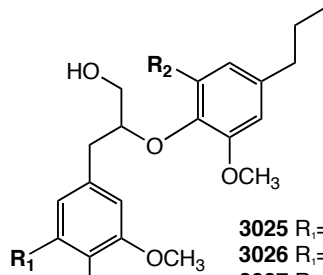


241,238* R₁=OCH₃, R₂=H
249 R₁=OCH₃, R₂=H

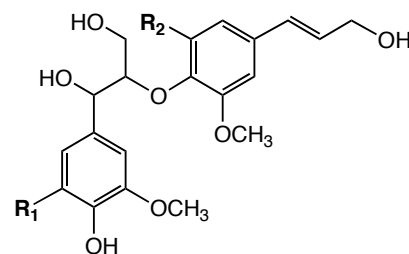
268e,269t,270e*,271f* R₁=R₂=H



228

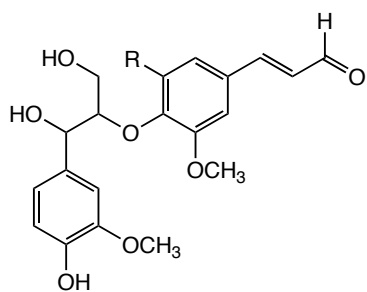


3025 R₁=R₂=H
3026 R₁=H, R₂=OCH₃
3027 R₁=OCH₃, R₂=H
3028 R₁=R₂=OCH₃

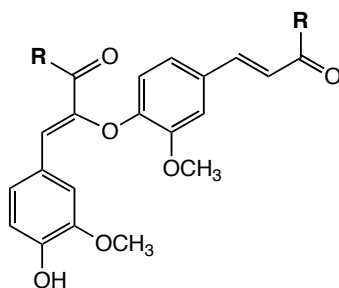


2013t,2014e,2012e*,2011f*,3008f* R₁=R₂=H
3067,188* R₁=H, R₂=OCH₃
186* R₁=OCH₃, R₂=H
185* R₁=R₂=OCH₃

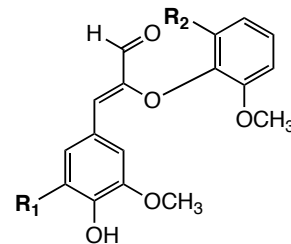
More b-O-4 Dimers, 3-Carbon Sidechain



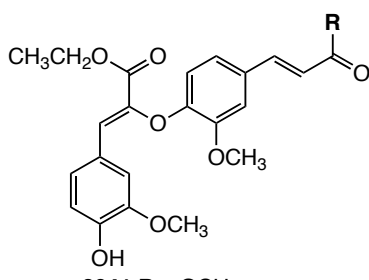
3010^f, 3011^e* R = H
3071 R = OCH₃ (with only g-Ac)



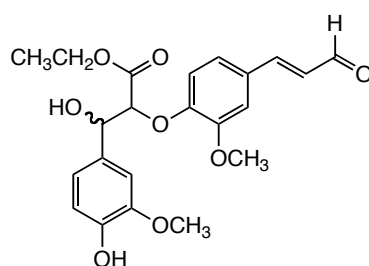
3031 R = H
2040 R = OH



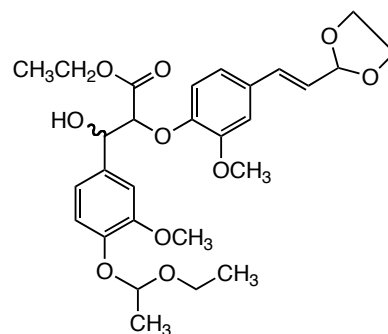
3034, 3035* R₁, R₂ = H
3036, 3037* R₁ = OCH₃, R₂ = H
3038, 3039* R₁ = H, R₂ = OCH₃
3040, 3041* R₁, R₂ = OCH₃



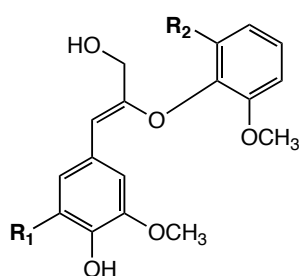
2041 R = OCH₃
2042 R = H



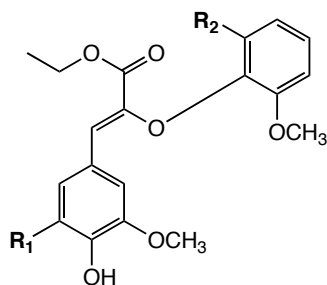
2043e, 2044f



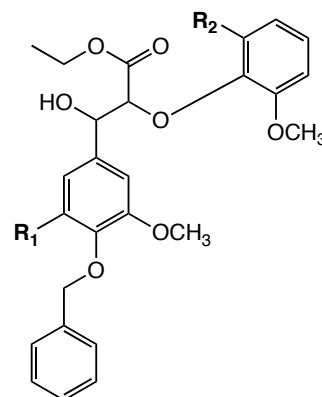
2045e, 2046f



3053, R₁ = H, R₂ = OCH₃
3057, R₁ = OCH₃, R₂ = H

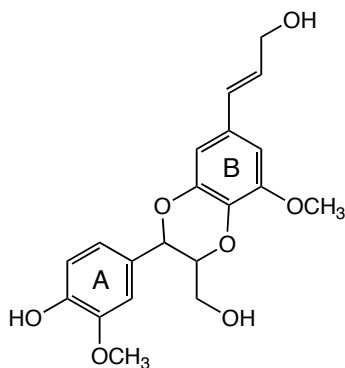


3052, R₁ = H, R₂ = OCH₃
3056, R₁ = OCH₃, R₂ = H

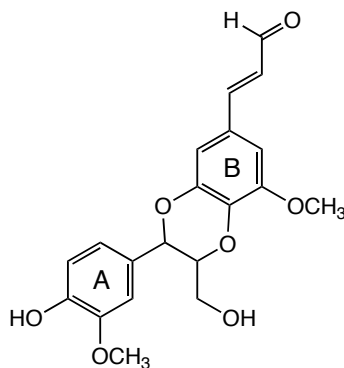


3050, 3051 isomers R₁ = H, R₂ = OCH₃
3054, 3055 isomers R₁ = OCH₃, R₂ = H

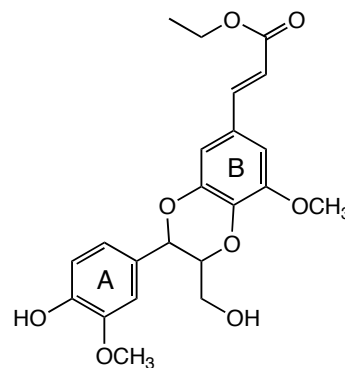
5-Hydroxyconiferyl alcohol b-O-4 Dimers, 3-Carbon Sidechain



3068



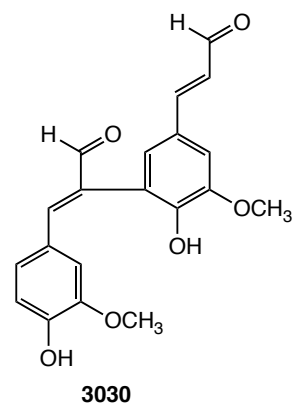
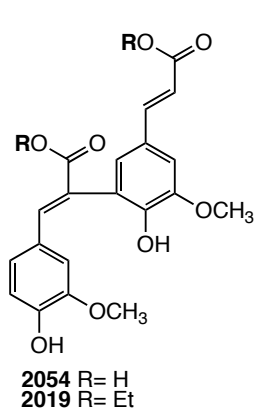
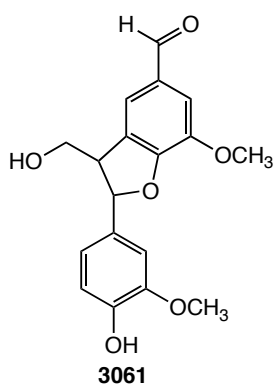
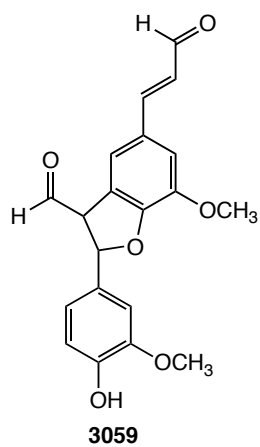
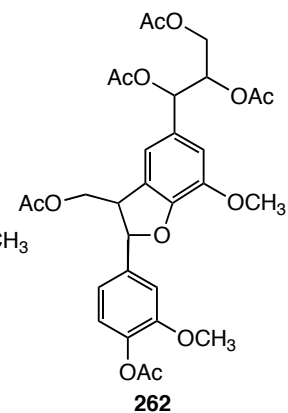
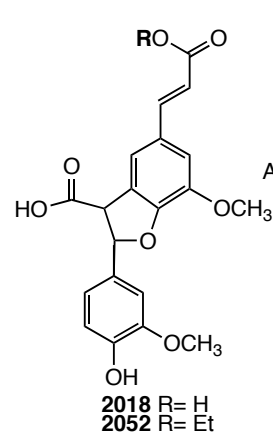
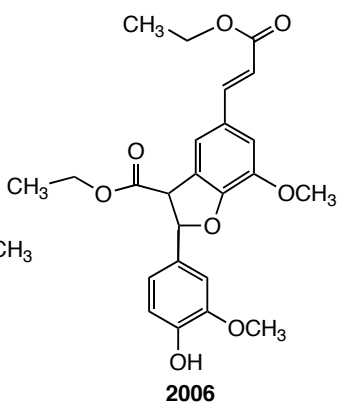
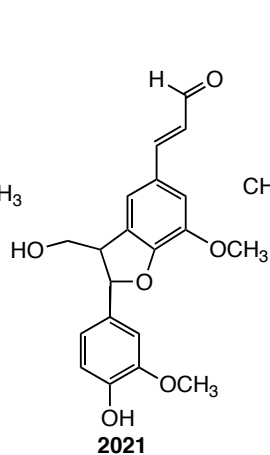
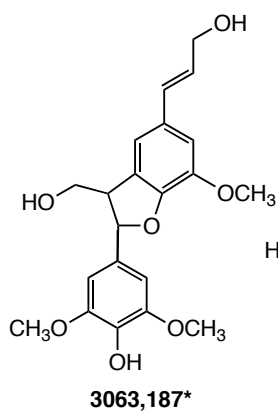
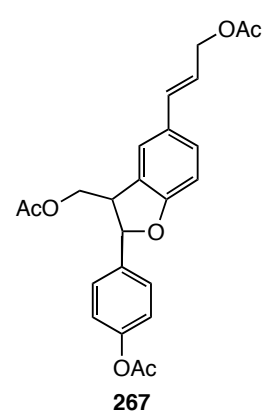
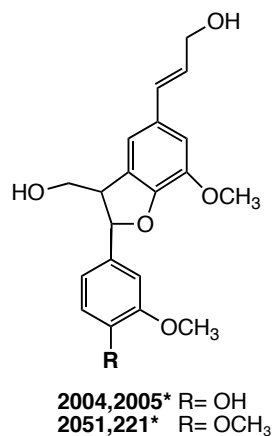
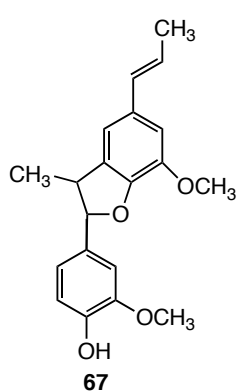
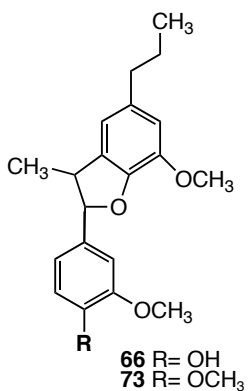
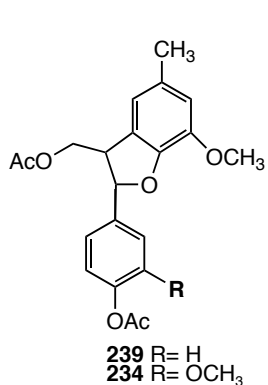
3069



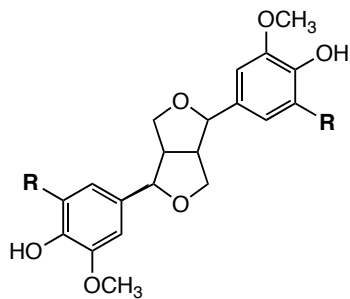
3070

X

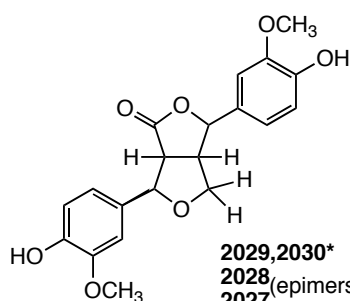
b-5 Dimers



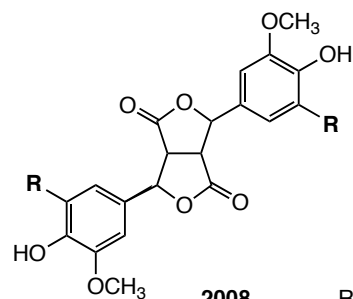
b-b Dimers



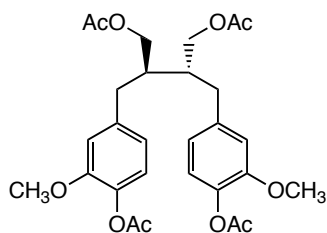
2020,109* R = H
117,123* R = OCH₃



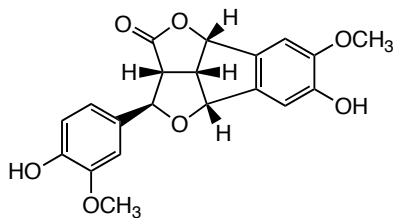
2029,2030*
2028, (epimers)
2027



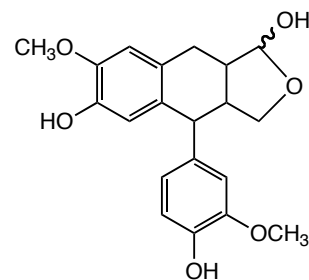
2008
3002, 3042* R = H
R = OCH₃



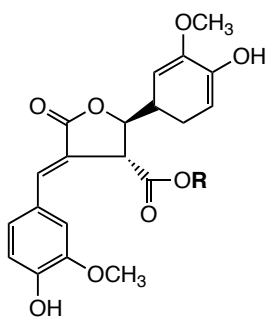
2070



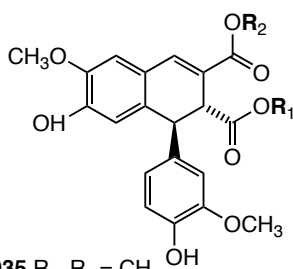
2026



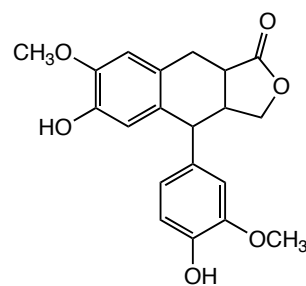
3014, 3015 isomers



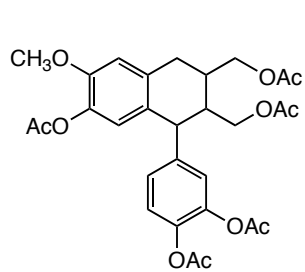
2033 R = H
2034 R = OCH₃



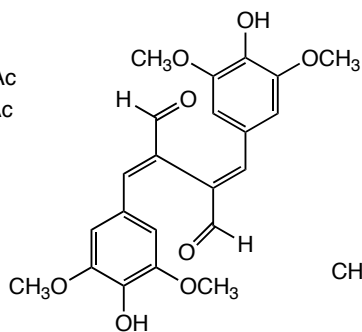
2035 R₁, R₂ = CH₃
2036 R₁, R₂ = H
2062 R₁ = CH₃, R₂ = H



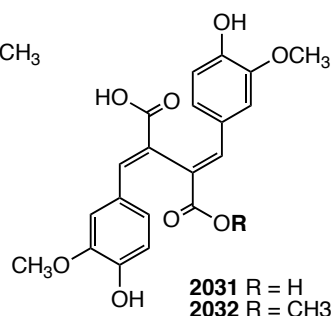
3017, 3016*



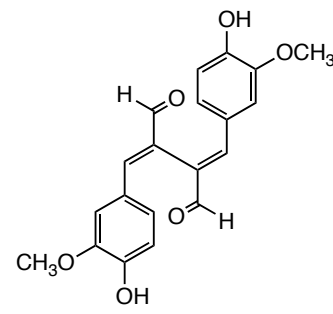
2069,3018



3060

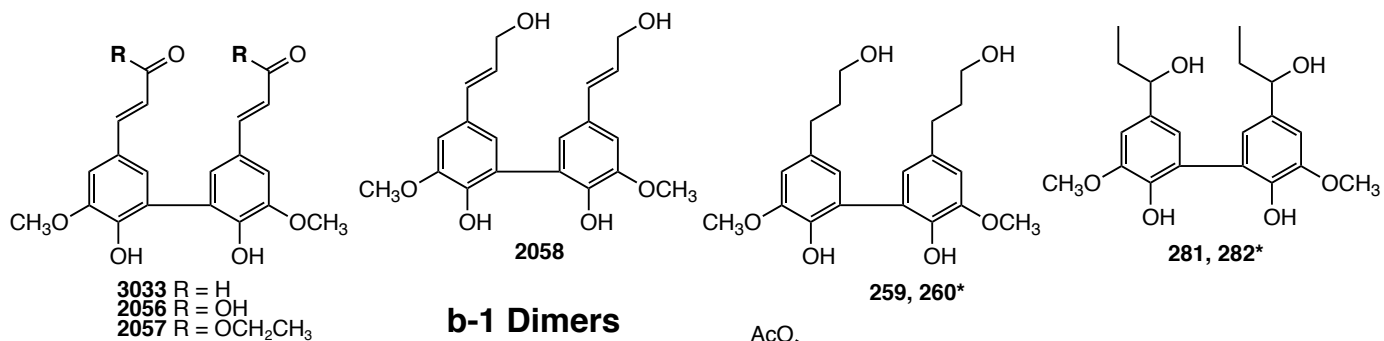
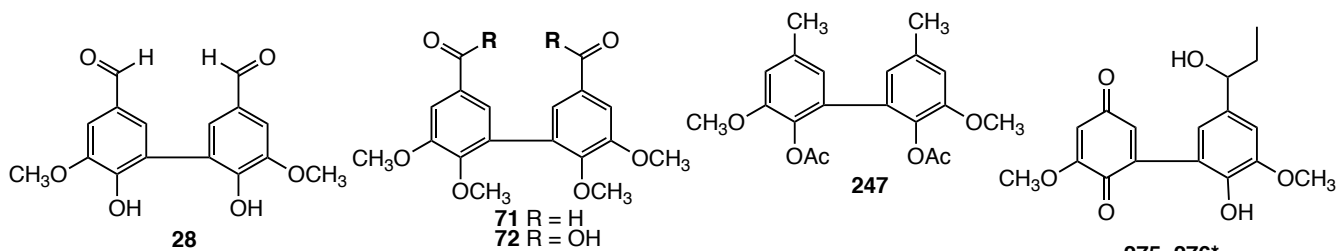


2031 R = H
2032 R = CH₃

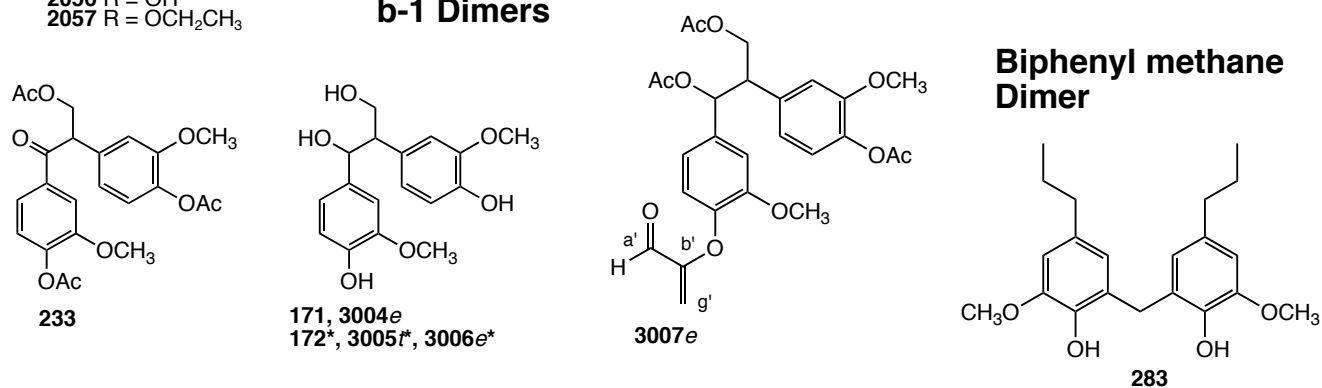


3032

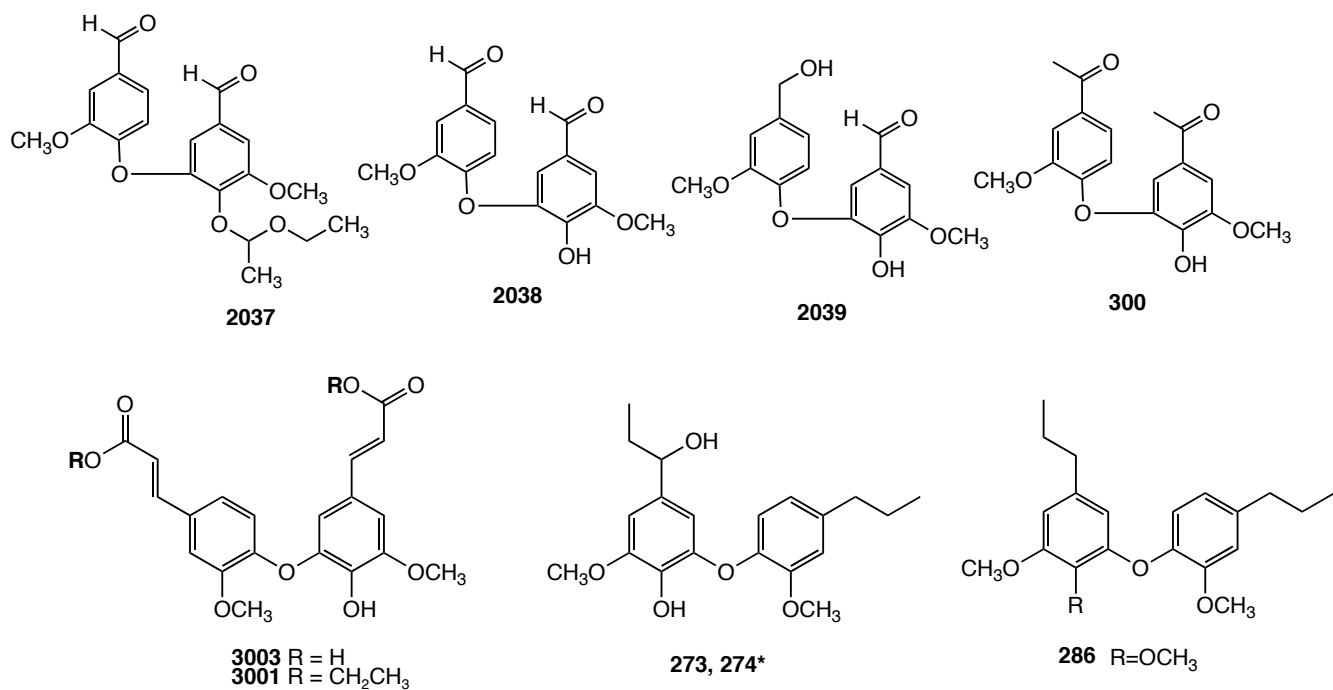
5-5 Dimers



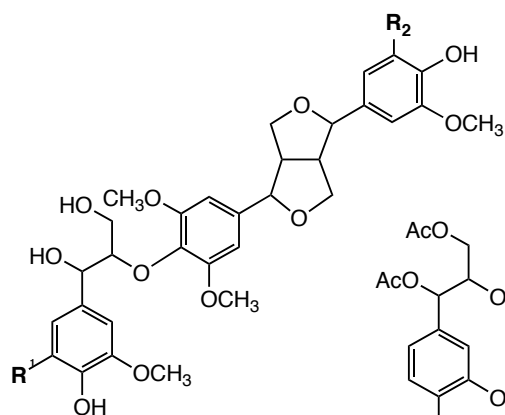
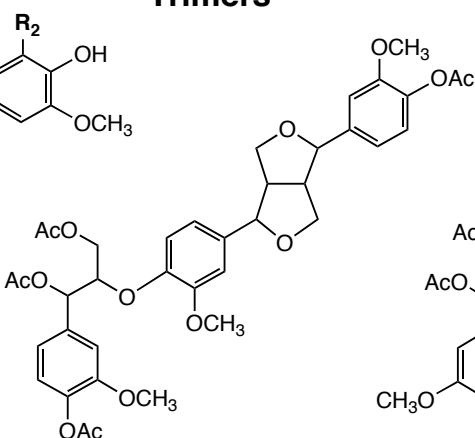
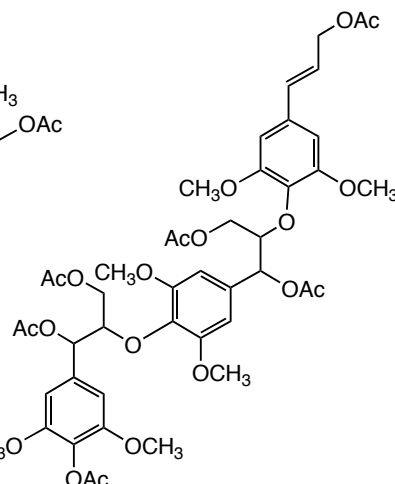
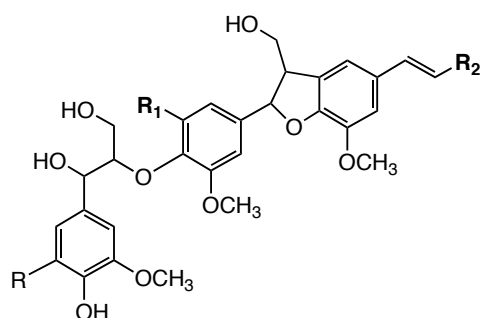
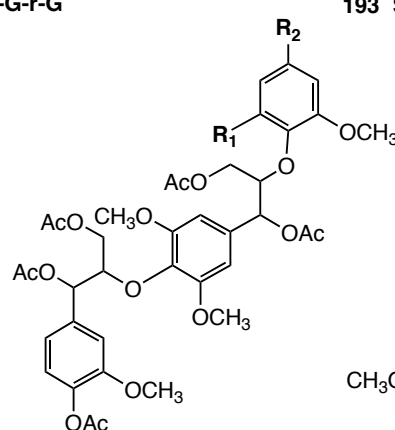
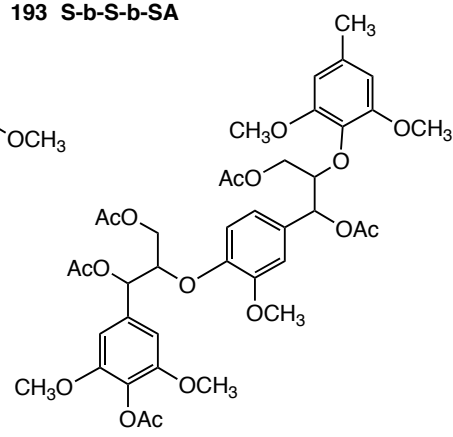
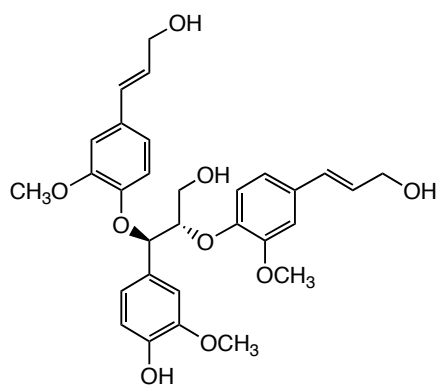
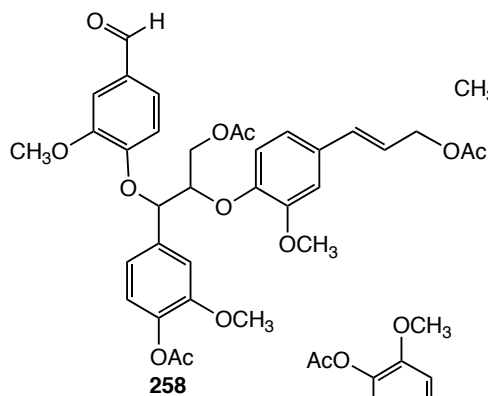
b-1 Dimers



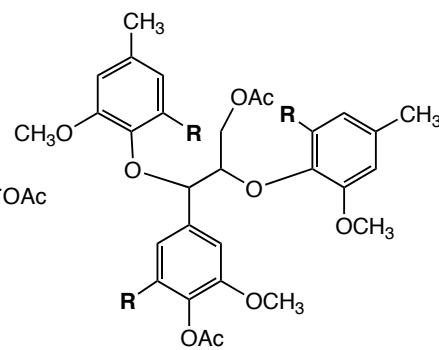
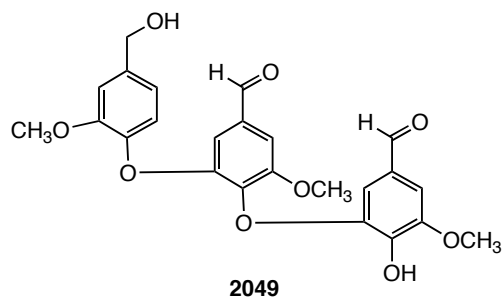
5-O-4 Dimers



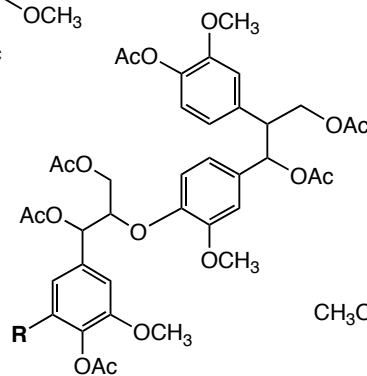
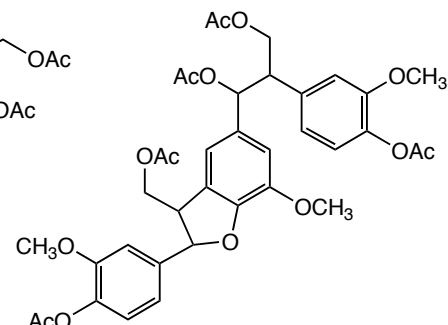
Trimers

183,184* $R_1 = H, R_2 = OCH_3$, **G-b-S-r-S**198,199* $R_1, R_2 = OCH_3$, **S-b-S-r-S**3064 $R_1, R_2 = H$, **G-b-S-r-G**263 **G-b-G-r-G**193 **S-b-S-b-SA**181,182* $R = H, R_1 = OCH_3, R_2 = CH_2OH$ **G-b-S-c-CA**216* $R = H, R_1 = H, R_2 = CH_2OH$ **G-b-G-c-CA**3065 $R = H, R_1 = OCH_3, R_2 = CHO$ **G-b-G-c-CAld**3072 $R = OCH_3, R_1 = H, R_2 = CH_2OH$ **S-b-G-c-CA**226 $R_1 = R_2 = H$ **G-b-S-b-G**250 $R_1 = OCH_3, R_2 = CH_3$ **G-b-S-b-S**261 **S-b-G-S**2015e,2022e* **CA-a-G-b-CA**

258

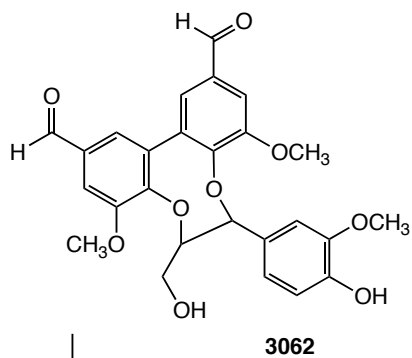
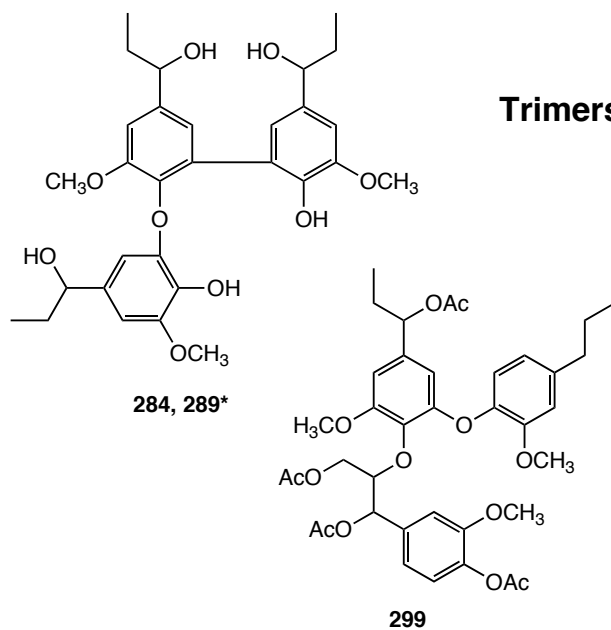
237 $R = H$ **G-a-G-b-G**266 $R = OCH_3$ **S-a-S-b-S**

2049

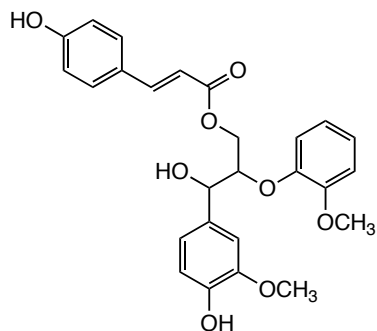
236, 3013 $R = H$ **G-b-G-b1-G**235 $R = OCH_3$ **S-b-G-b1-G**

3012

Trimers Cont'd

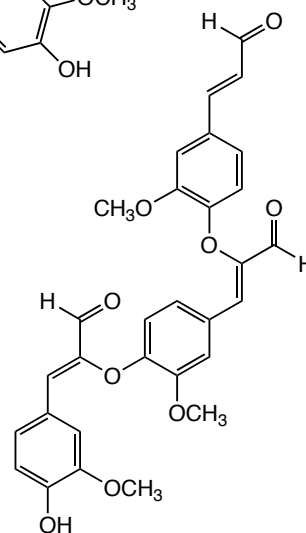


Trimers Containing Ferulic or Coumaric Acid

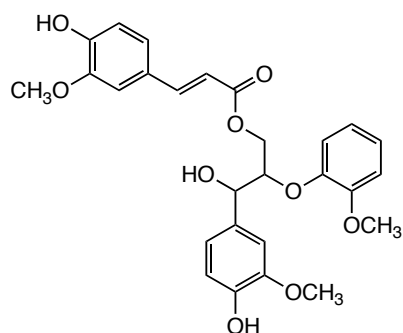


1015t, 1016e, 1019t*, 1020e*, 82t*, 83e*

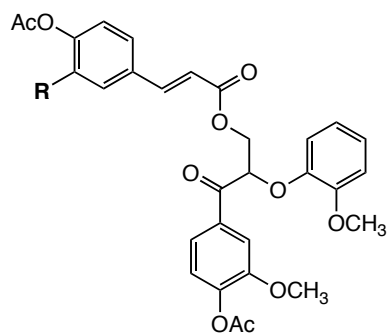
1011t, 1012e (diacetates, phenolic)



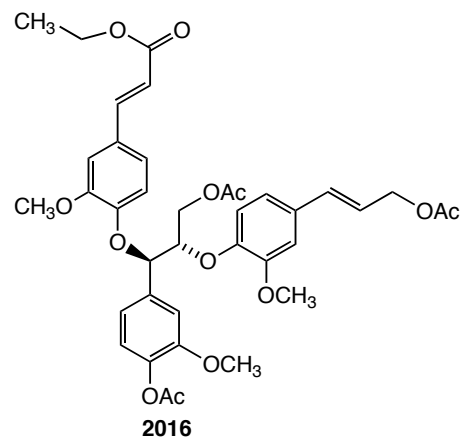
3058



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

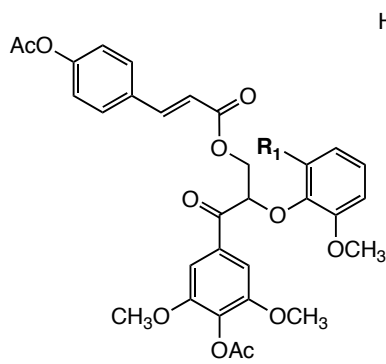


1009 R = H
1010 R = OCH₃

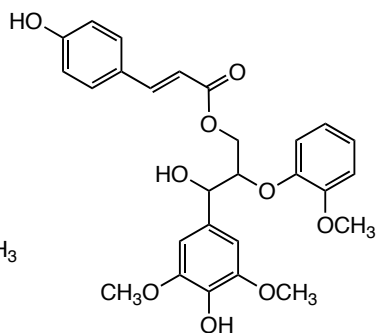


2016

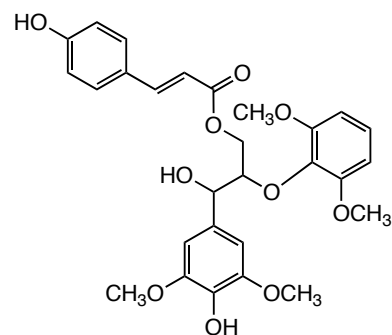
More Trimers Containing Ferulic, Coumaric or *p*-OH-Benzoic Acid



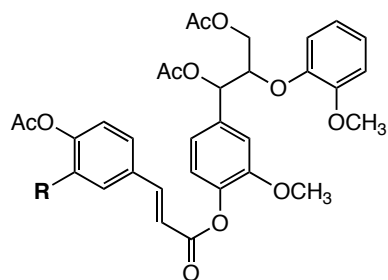
2066 R = H
2081 R = OCH₃



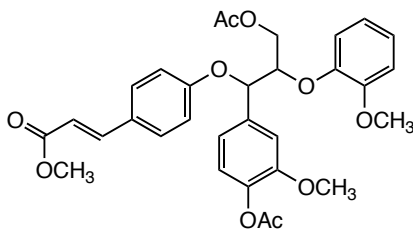
2076t, 2078e
2075t*, 2077e*
2068e, (diacetate, phenolic)



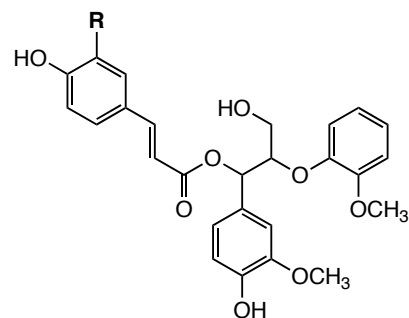
2071t, 2072e, 2073t*, 2074e*
2079t, 2080e (diacetate, phenolics)



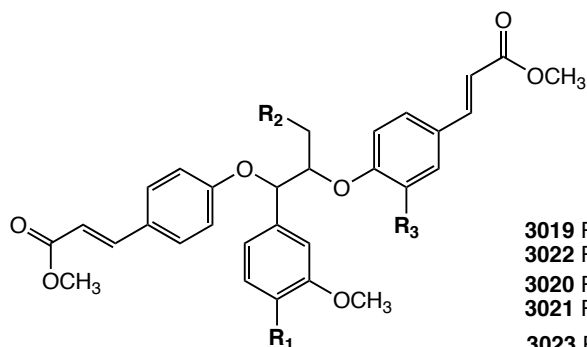
1023t, 1024e R = H
1025t, 1026e, 84e R = OCH₃



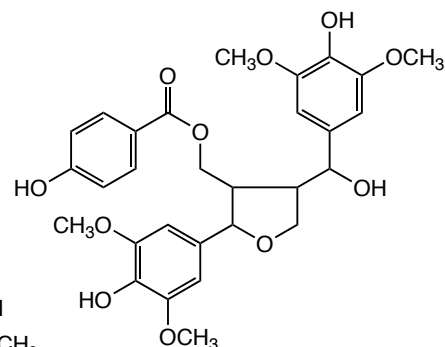
85, 86



1001t, 1002e, 1005t*, 1006e*, 78t*, 77e* R = H
1003t, 1004e, 1007t*, 1008e* R = OCH₃

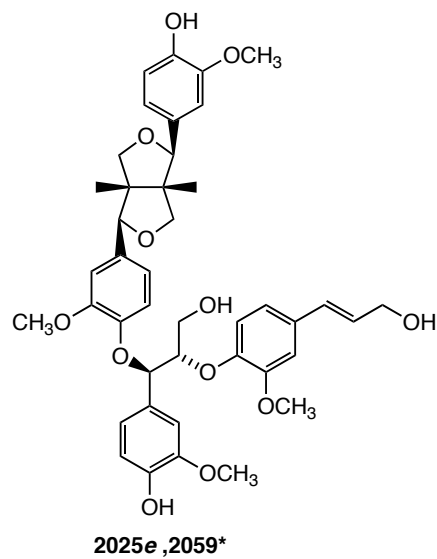
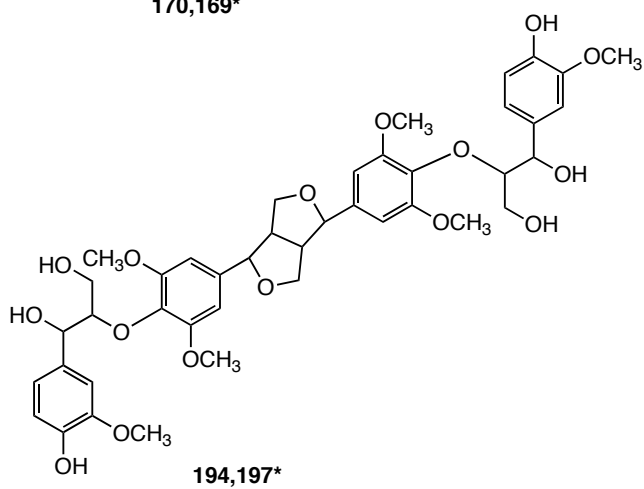
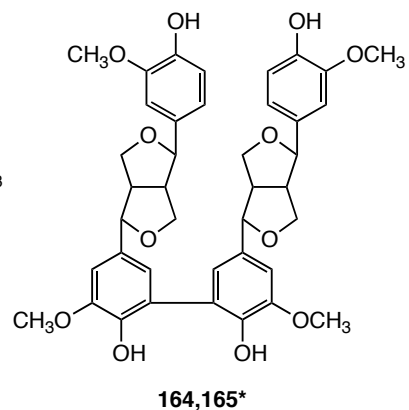
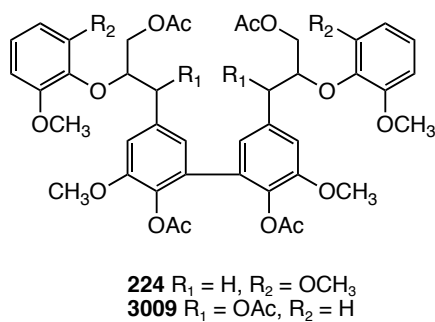
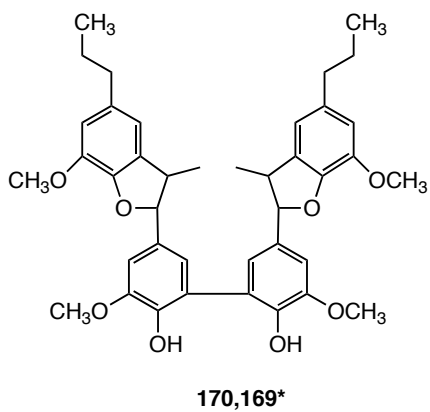
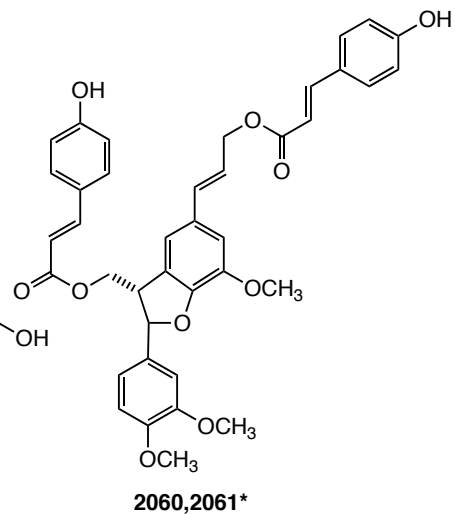
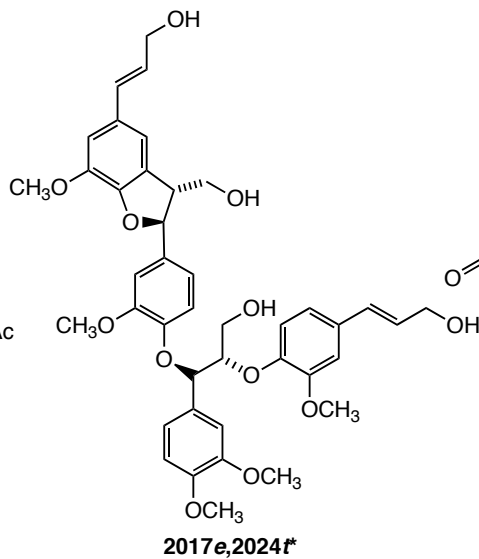
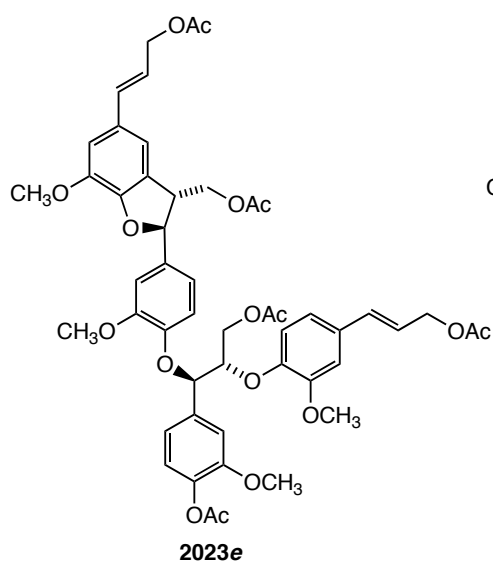


3019 R₁=R₂= OH, R₃= H
3022 R₁=R₂= OH, R₃= OCH₃
3020 R₁= OCH₃, R₂= OH, R₃= H
3021 R₁= OCH₃, R₂= OAc, R₃= H
3023 R₁= OCH₃, R₂= OH, R₃= OCH₃
3024 R₁= OCH₃, R₂= OAc, R₃= OCH₃

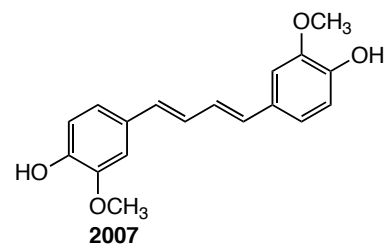
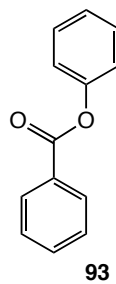
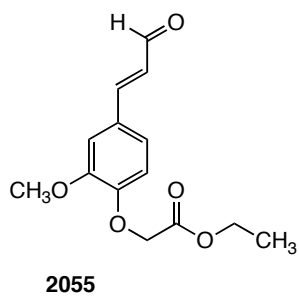
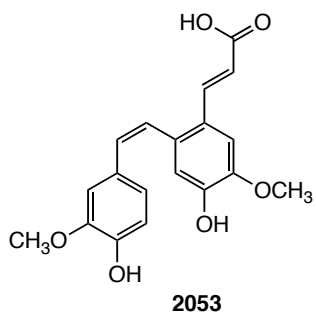
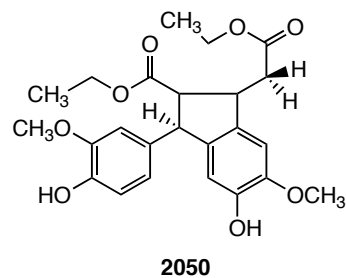
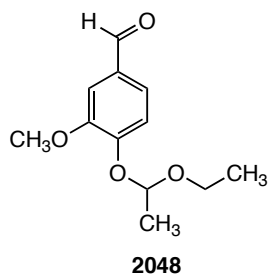
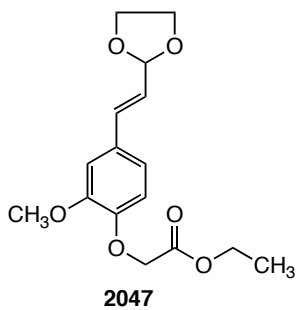
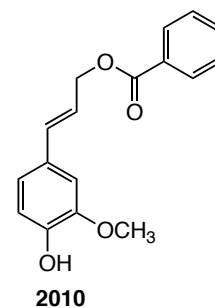
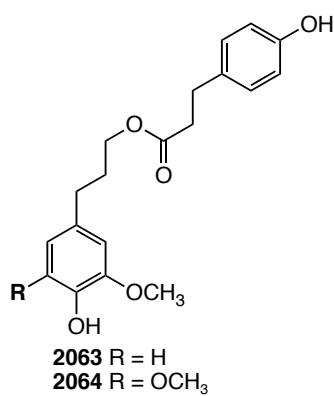
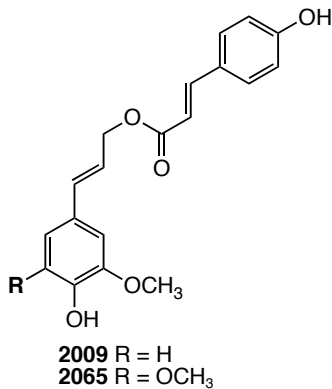
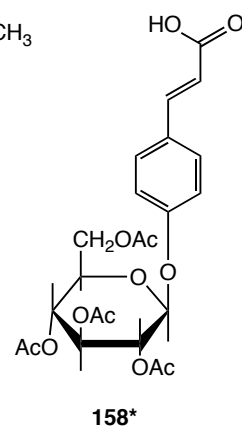
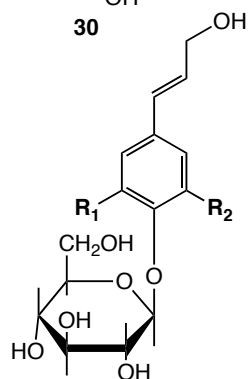
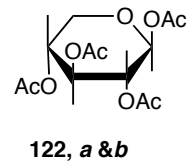
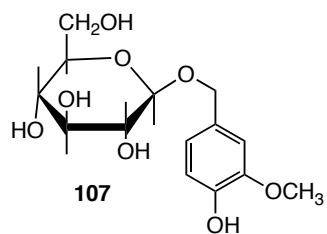
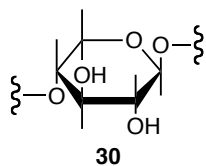


3066

Tetramers



Misc. Compounds



Misc. Compounds

