

NMR Database of Lignin and Cell Wall Model Compounds

Sally Ralph and Larry Landucci

US Forest Products Laboratory

USDA Forest Service

One Gifford Pinchot Drive

Madison, WI 53705

email: saralph@wisc.edu

John Ralph

US Dairy Forage Research Center

USDA Agriculture Research Service

1925 Linden Drive West

Madison, WI 53706-1108

email: jralph@wisc.edu

Web: dfrc.ars.usda.gov

This database was created and is administered as a cooperative effort between the US Forest Products Laboratory and the US Dairy Forage Research Center. It was designed to provide a coherent, single source of NMR data of lignin model compounds as well as compounds modeling similar structures in grasses and other forage plants. The database exists in four different formats: an interactive HyperCard© stack for the Macintosh® computer (essentially only used in-house now), 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 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 ¹³C NMR data was collected in three common deuterated solvents (acetone, chloroform and dimethyl sulfoxide) for each compound. The ¹H NMR data was reported for 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. 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. Search routines for the software versions allow grouping the compounds with similar traits, whereas the structure index is most useful for the hardcopy version. The inclusion of many 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.

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. The intensities of the individual chemical shift signals are used for the line plots generated by the HyperCard program but they are also useful in the hardcopy version for comparison with spectra.

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.

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

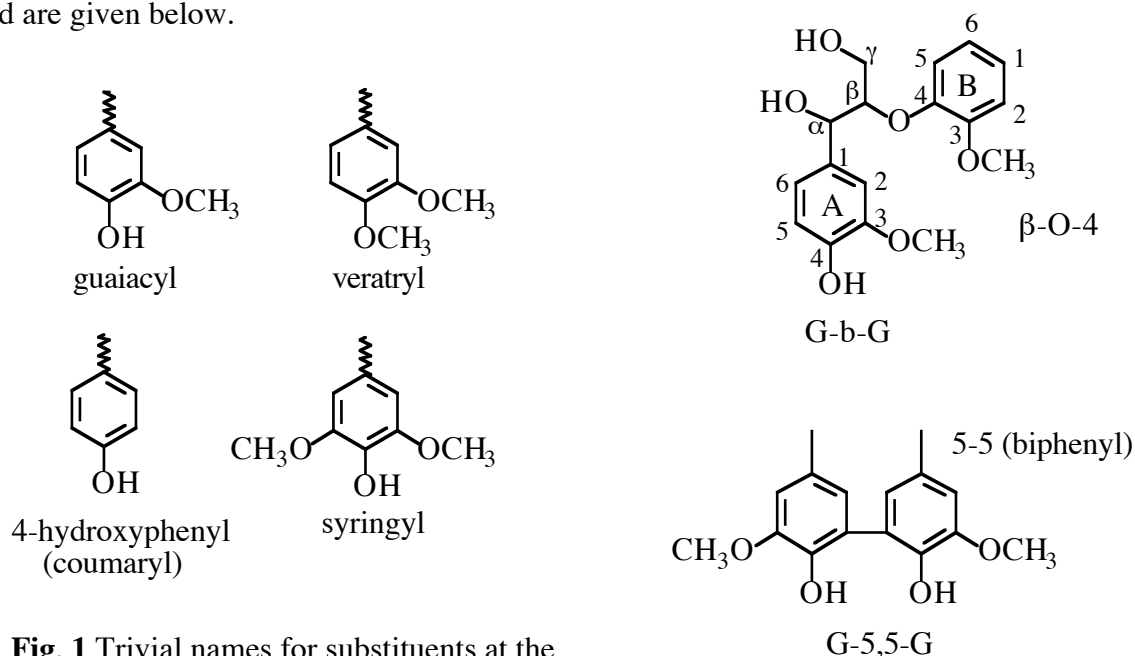


Fig. 1 Trivial names for substituents at the 3,4 and 5 positions on the aromatic ring.

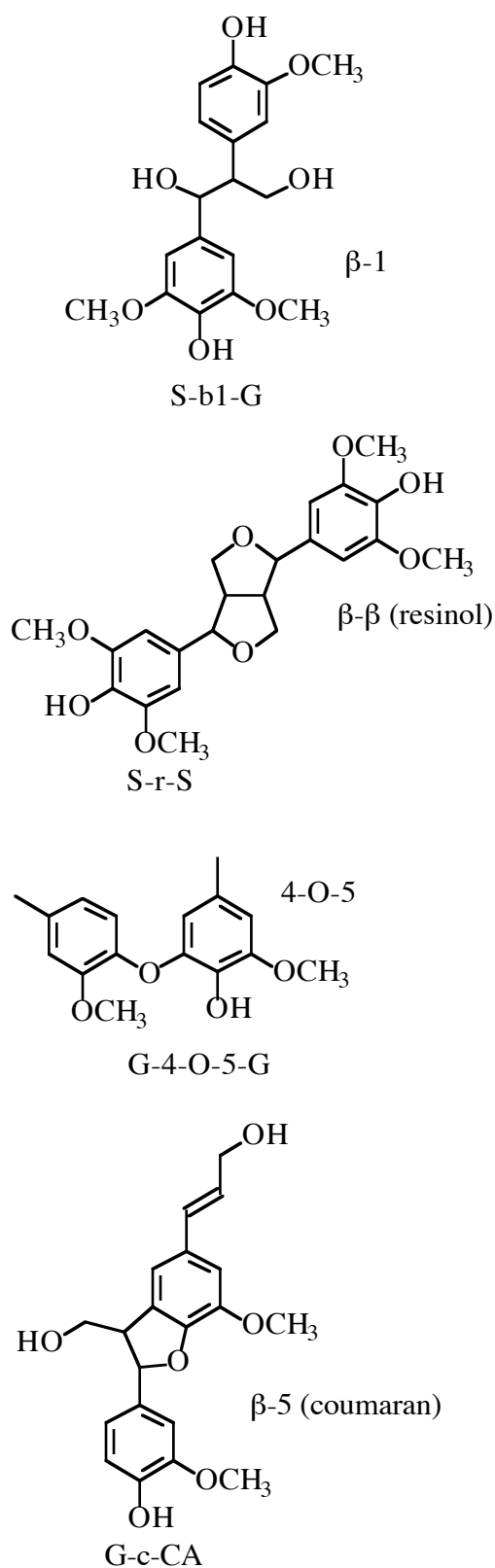


Fig. 2 Examples of linkages and abbreviated names.

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. We will also make every effort to keep those researchers with hardcopy versions supplied with updated pages. This database was written and prepared 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.

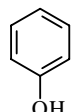
The authors gratefully acknowledge the many generous contributions made by others at both labs towards this database; William Landucci for software development, Martin Wesolowski and Kolby Hirth, NMR spectroscopists, and the frequent donation of compounds from Mike Mozuch, Noritsugu Terashima, Stéphane 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

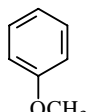
Contents

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 ferulate, <i>p</i> -coumarate, or <i>p</i> -hydroxybenzoate	xiv
Tetramers	xvi
Misc.	xvii

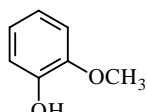
Monomers



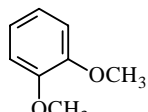
38,190*



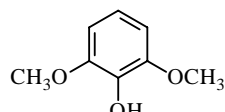
54



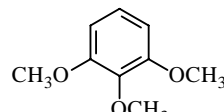
39,176*



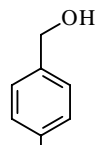
56



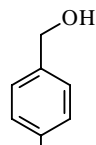
40,189*



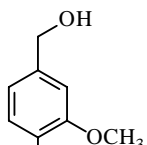
55



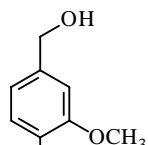
4,24*



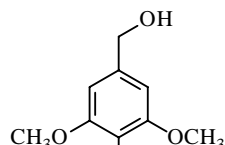
37,205*



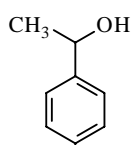
19,200*



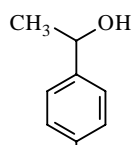
5,143*



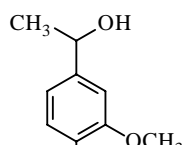
16,201*



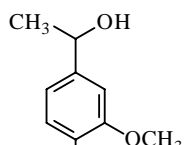
10,23*



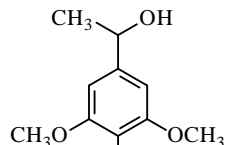
209,210*



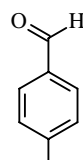
11,202*



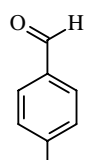
13,204*



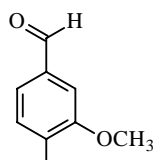
12,195*



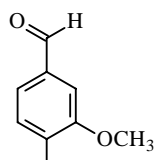
14,207*



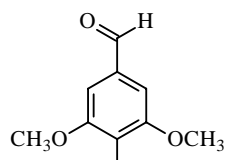
203



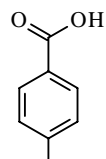
15,174*



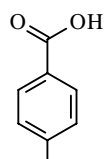
57



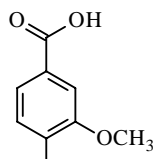
41,206*



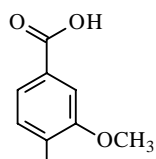
58,208*



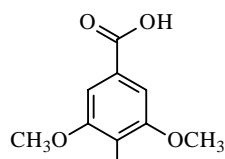
20



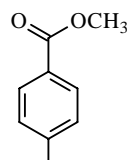
42,192*



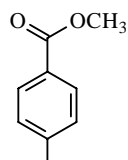
44



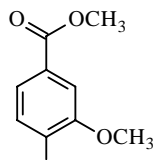
43,175*



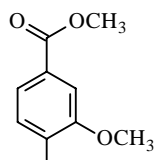
18,17*



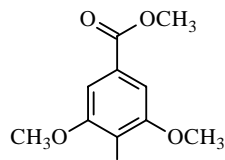
25



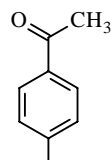
65,180*



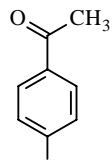
69



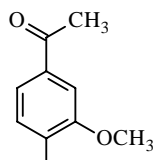
64,191*



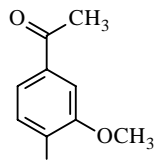
6,173*



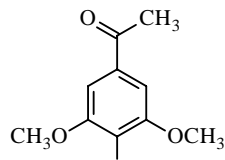
36



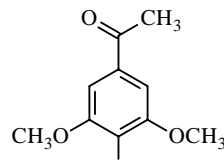
7,103*



9

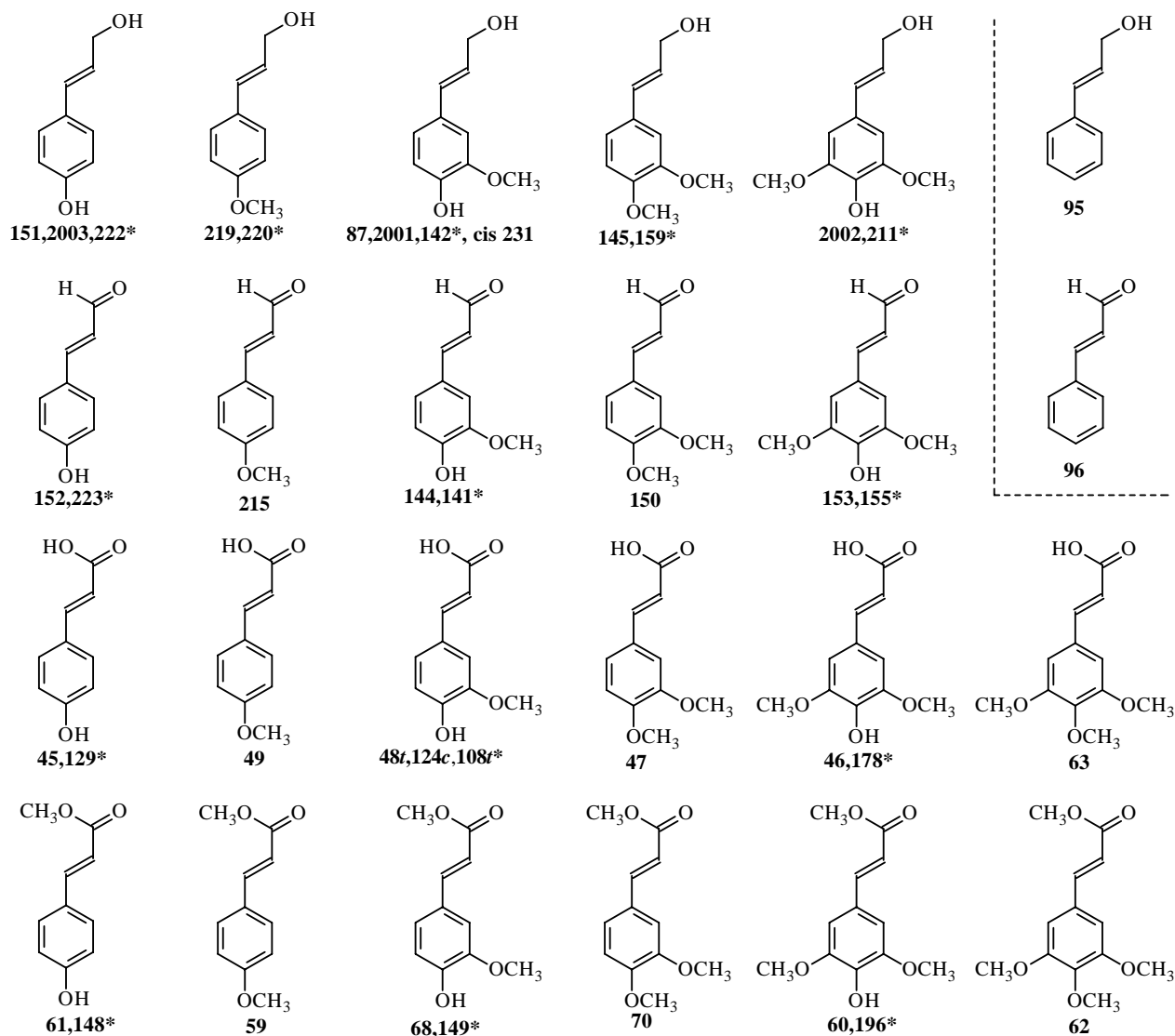


8,177*

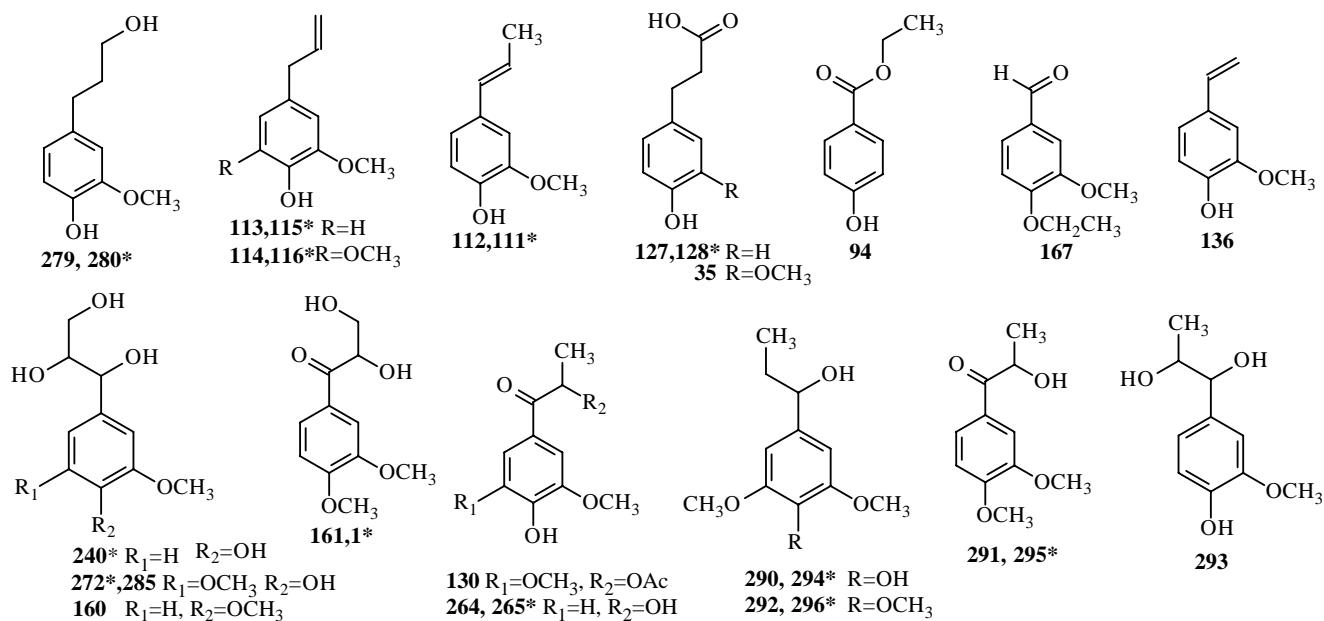


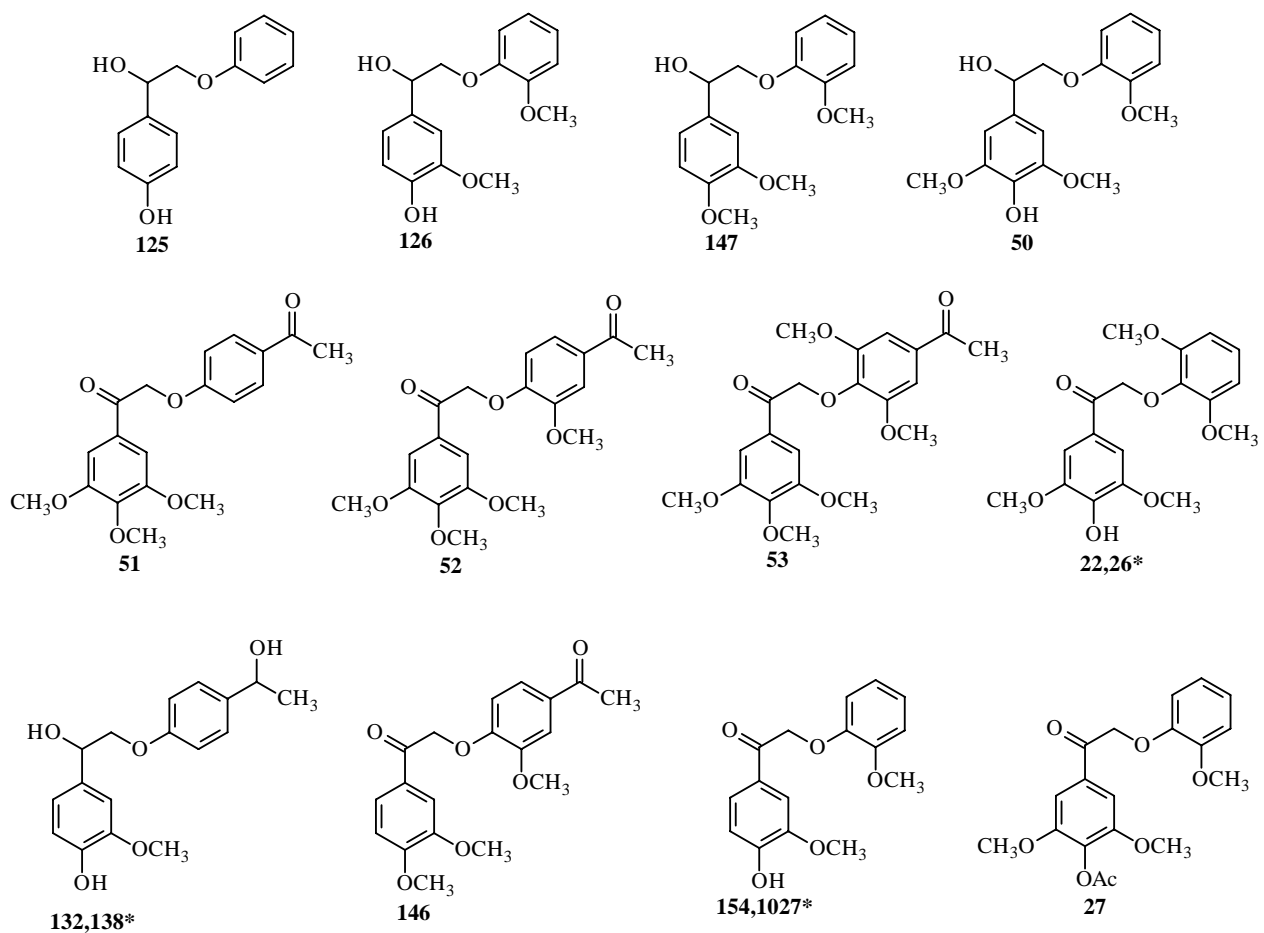
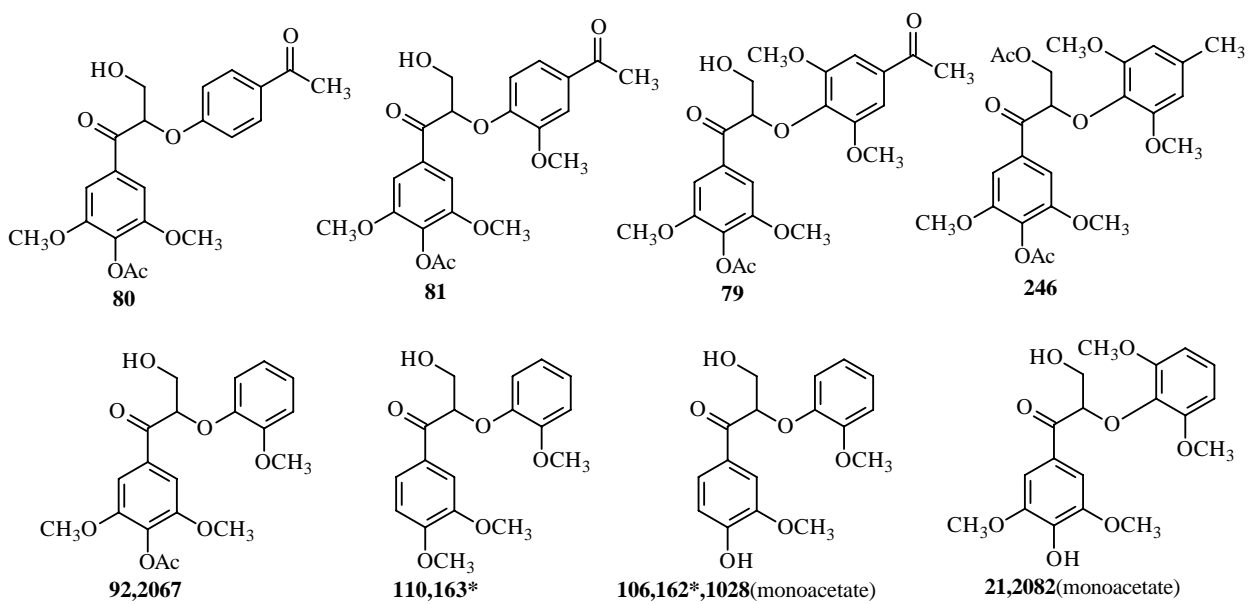
156

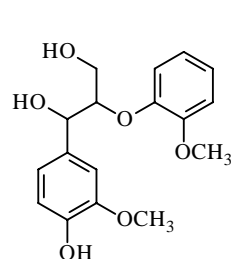
Monomers ctd



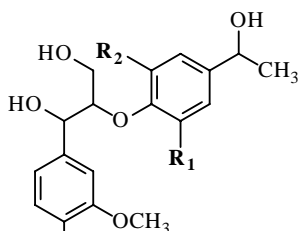
Misc. Monomers



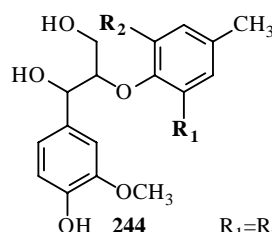
β -O-4 Dimers, 2-Carbon Sidechain **β -O-4 Dimers, 3-Carbon Sidechain, α -C=O**

β -O-4 Dimers, 3-Carbon Sidechain

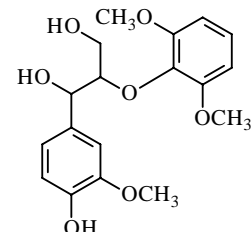
101e, 102t, 74t*, 214e*
(1029t, 1030e α, γ Ac'd)



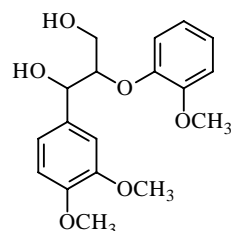
133, 137* $R_1=R_2=H$
131, 140* $R_1=OCH_3, R_2=H$
134, 139* $R_1=OCH_3, R_2=OCH_3$



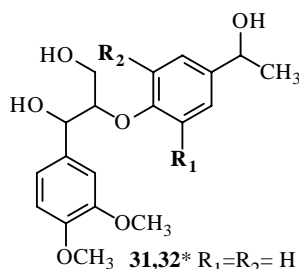
244 $R_1=R_2=H$
248t, 227e* $R_1=OCH_3, R_2=H$
242, 232* $R_1=OCH_3, R_2=OCH_3$



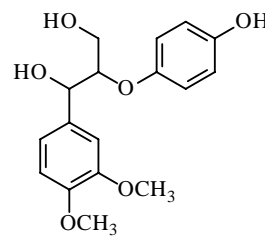
179



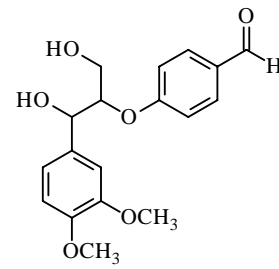
104t, 105e, 3e*



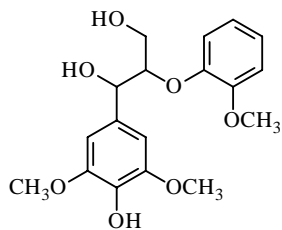
31, 32* $R_1=R_2=H$
34, 33* $R_1=OCH_3, R_2=H$
2, 29* $R_1=OCH_3, R_2=OCH_3$



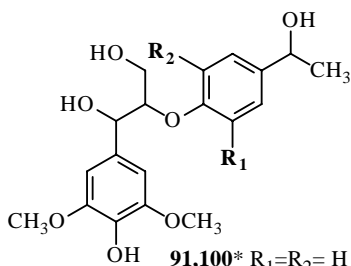
166



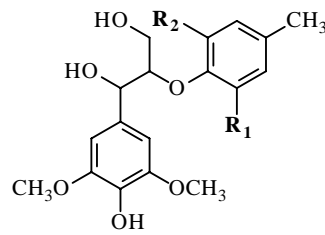
168



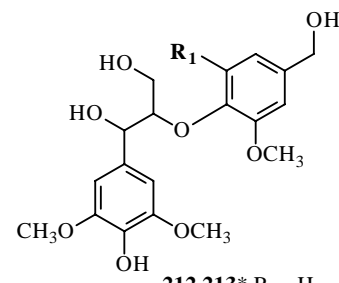
90, 99*



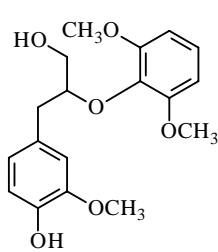
91, 100* $R_1=R_2=H$
88, 97* $R_1=OCH_3, R_2=H$
89, 98* $R_1=OCH_3, R_2=OCH_3$



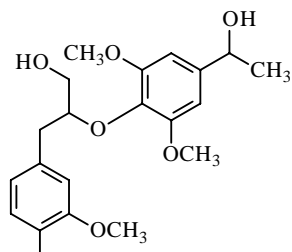
245 $R_1=OCH_3, R_2=H$
243, 230* $R_1=OCH_3, R_2=OCH_3$



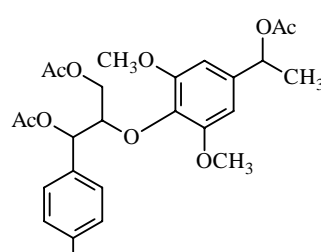
212, 213* $R_1=H$
229* $R_1=OCH_3$



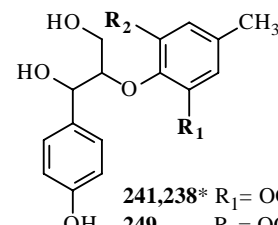
217, 218*



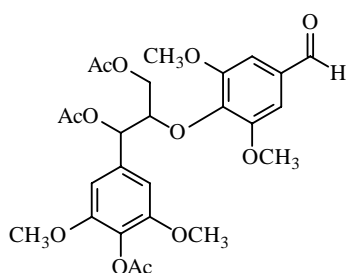
135



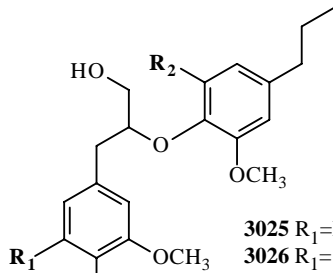
225



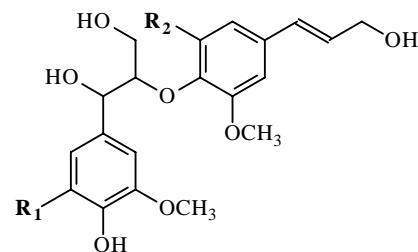
241, 238* $R_1=OCH_3, R_2=H$
249 $R_1=OCH_3, R_2=H$
268e, 269t, 270e*, 271t* $R_1=R_2=H$



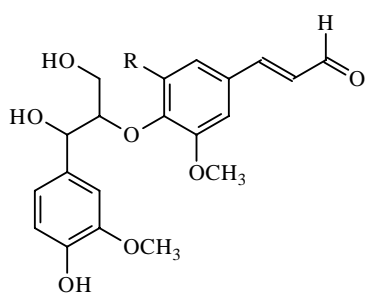
228



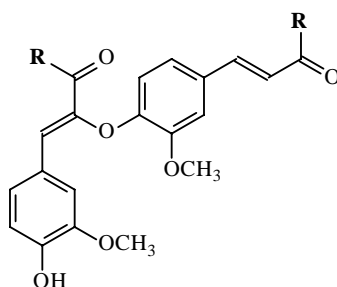
3025 $R_1=R_2=H$
3026 $R_1=H, R_2=OCH_3$
3027 $R_1=OCH_3, R_2=H$
3028 $R_1=R_2=OCH_3$



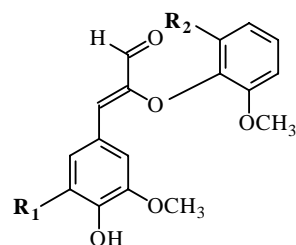
2013t, 2014e, 2012e*, 2011t*, 3008t* $R_1=R_2=H$
3067, 188* $R_1=H, R_2=OCH_3$
186* $R_1=OCH_3, R_2=H$
185* $R_1=R_2=OCH_3$

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

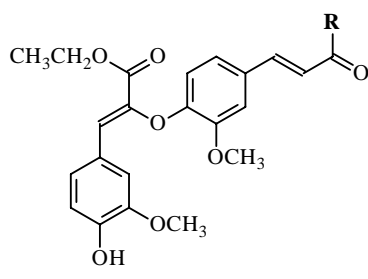
3010^f*, 3011^e* R = H
3071 R = OCH₃ (with only γ -Ac)



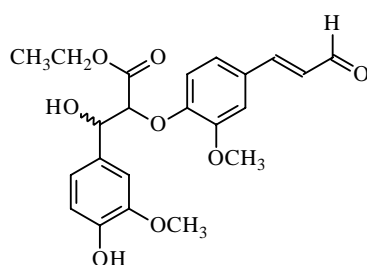
3031 R = H
2040 R = OH



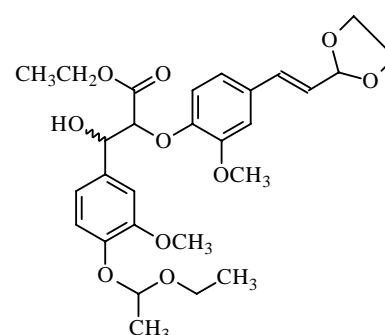
3034, 3035* R₁, R₂ = H
3036, 3037* R₁ = OCH₃, R₂ = H
3038, 3039* R₁ = H, R₂ = OCH₃
3040, 3041* R₁, R₂ = OCH₃



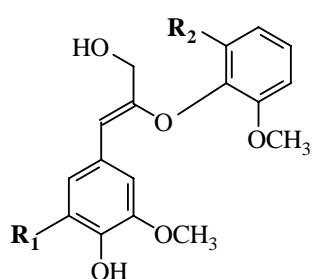
2041 R = OCH₃
2042 R = H



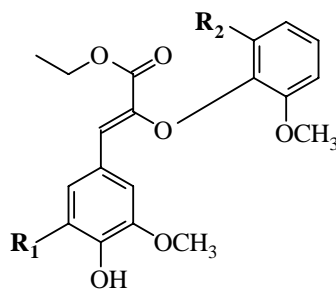
2043^e, 2044^t



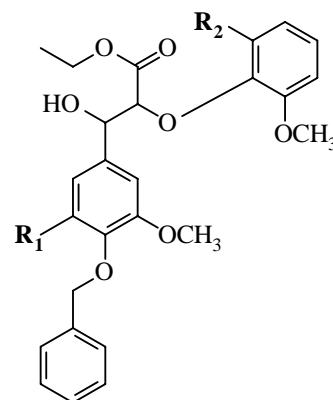
2045^e, 2046^t



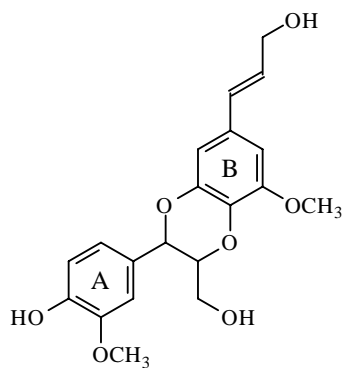
3053, R₁ = H, R₂ = OCH₃
3057, R₁ = OCH₃, R₂ = H



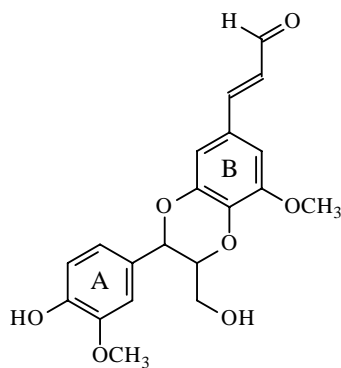
3052, R₁=H, R₂=OCH₃
3056, R₁=OCH₃, R₂=H



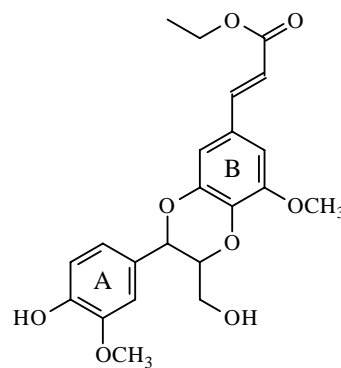
3050, 3051 isomers R₁=H, R₂=OCH₃
3054, 3055 isomers R₁=OCH₃, R₂=H

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

3068



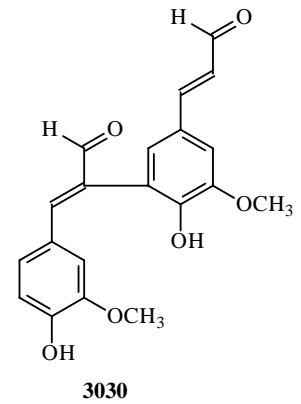
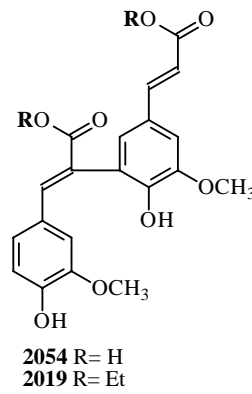
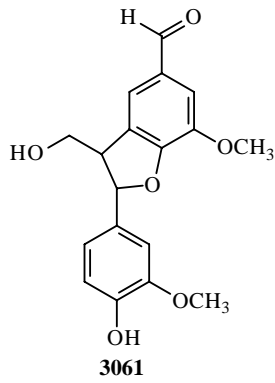
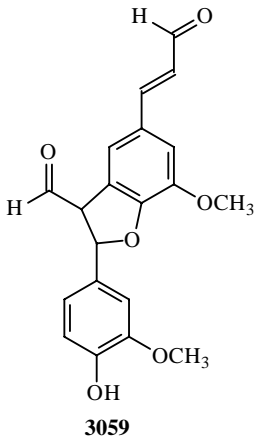
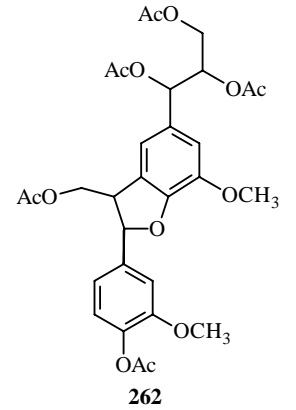
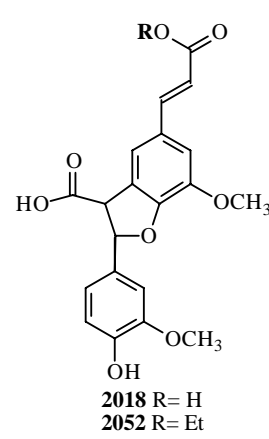
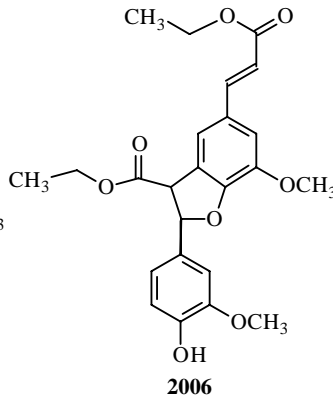
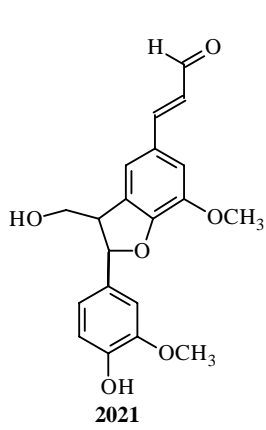
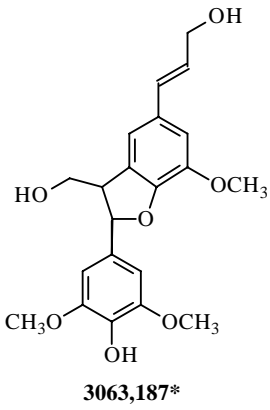
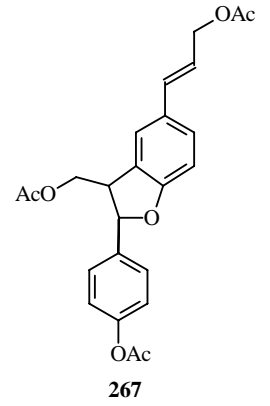
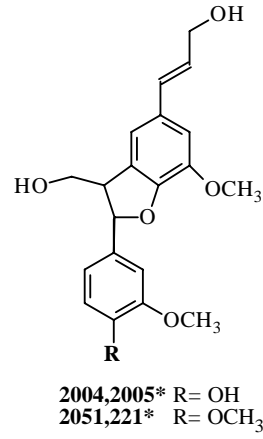
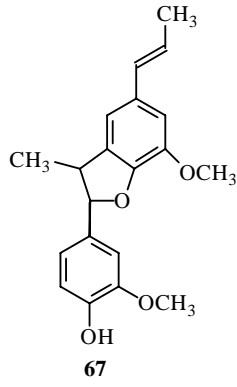
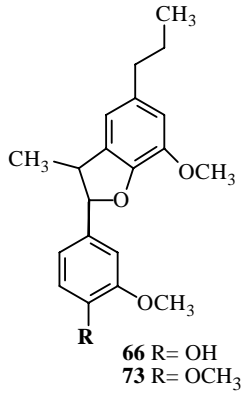
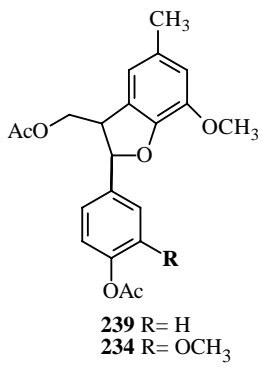
3069

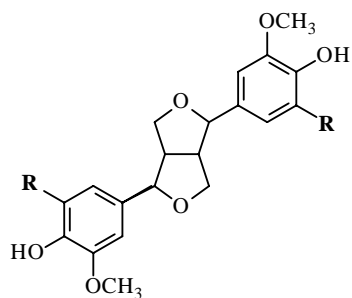


3070

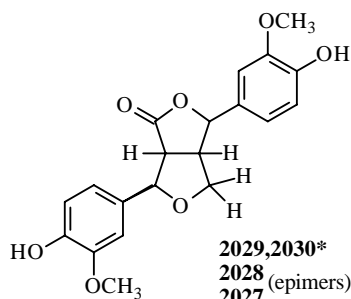
X

β -5 Dimers

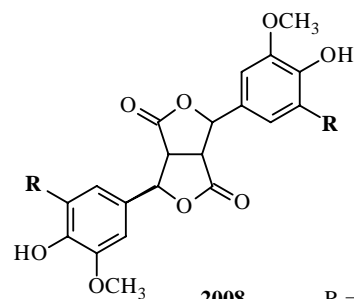


β - β Dimers

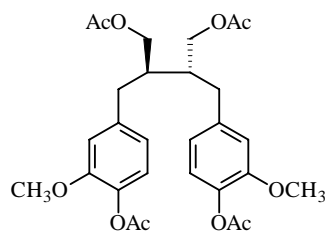
2020,109* R = H
117,123* R = OCH₃



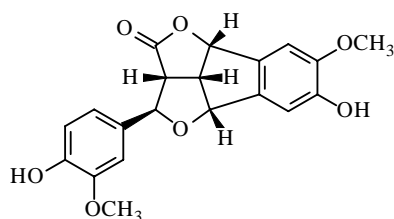
2029,2030*
2028 (epimers)
2027



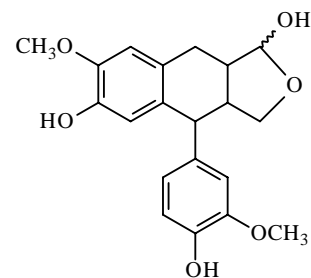
2008 R = H
3002, 3042* R = OCH₃



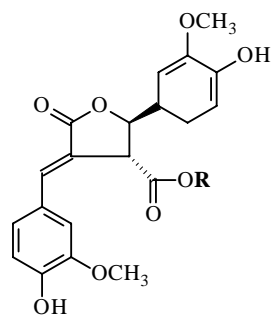
2070



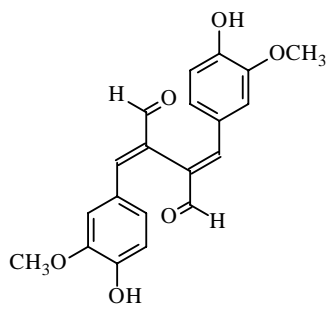
2026



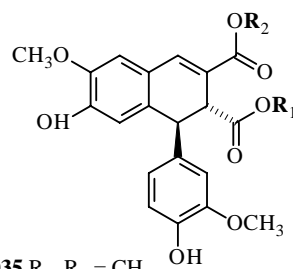
3014, 3015 isomers



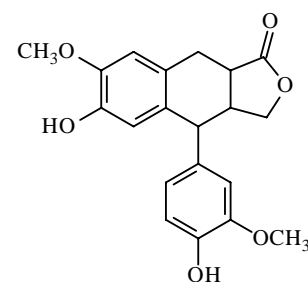
2033 R = H
2034 R = OCH₃



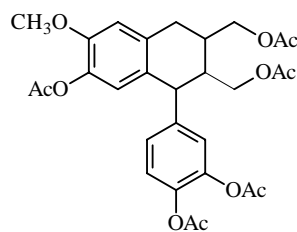
3032



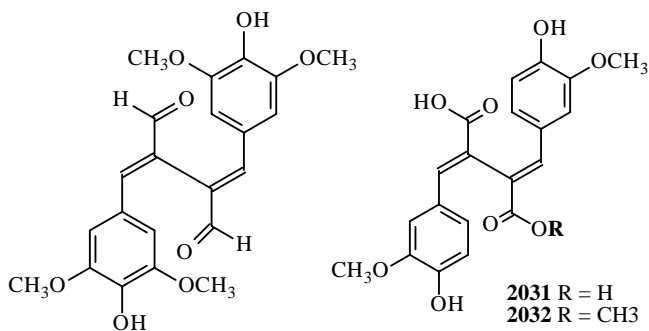
2035 R₁, R₂ = CH₃
2036 R₁, R₂ = H
2062 R₁ = CH₃, R₂ = H



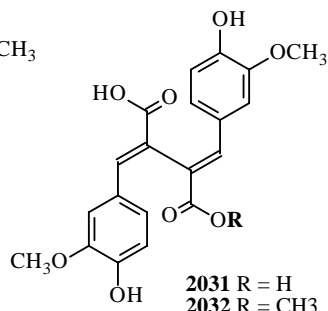
3017, 3016*



2069,3018

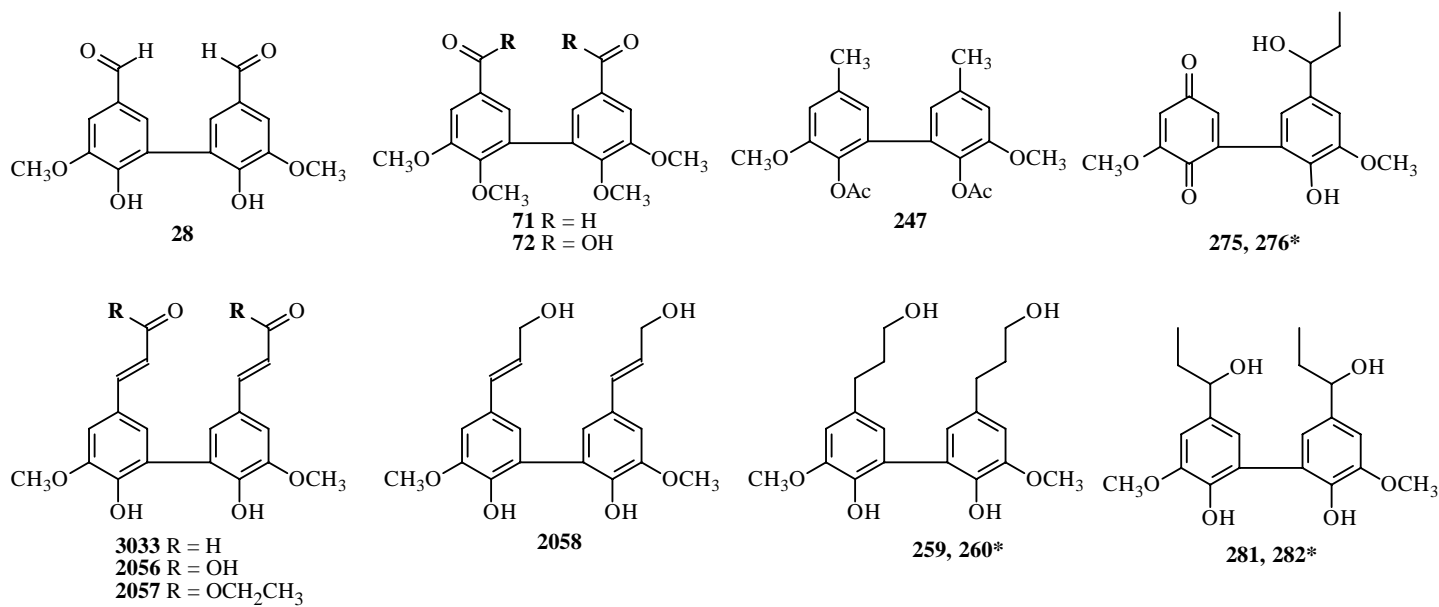
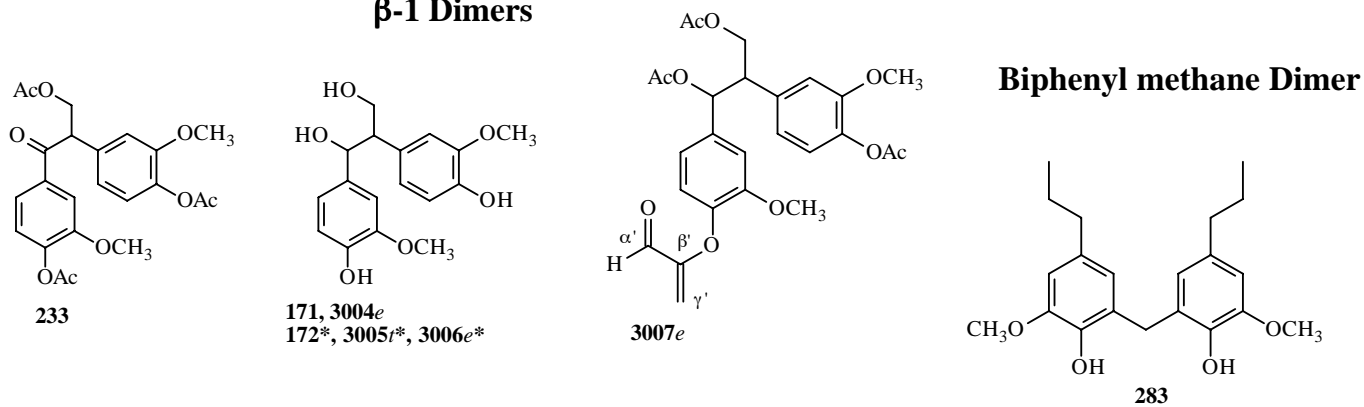


3060

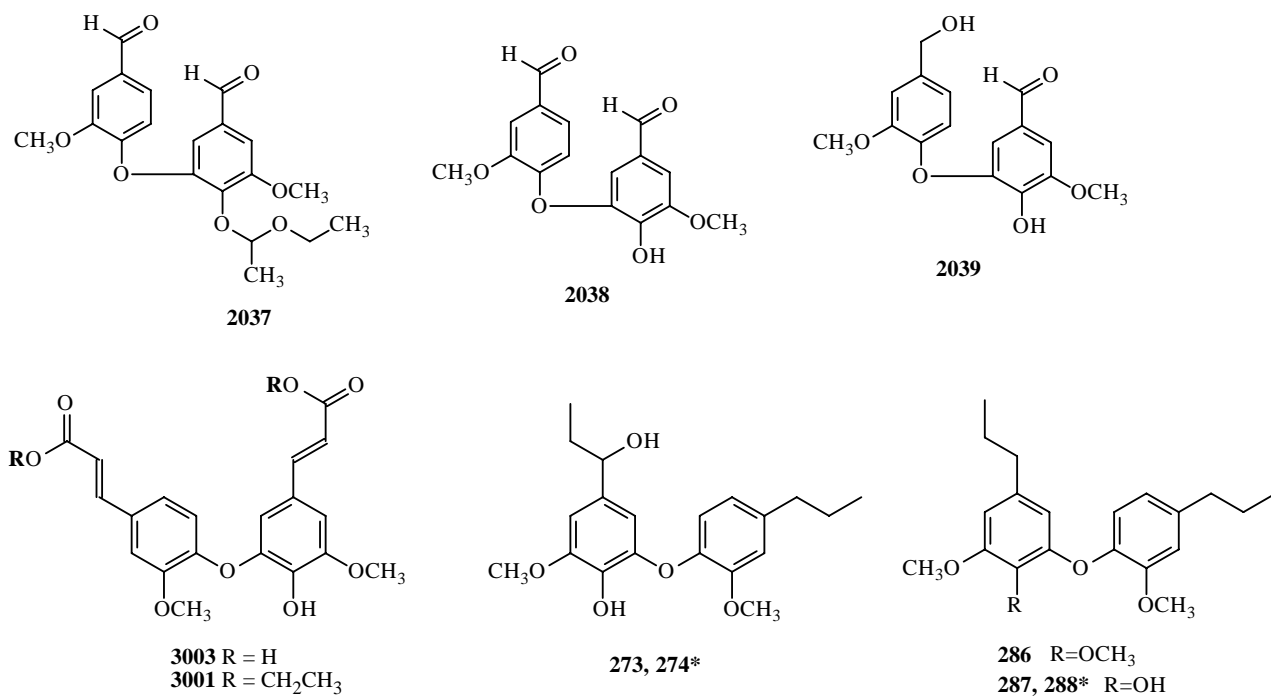


2031 R = H
2032 R = CH₃

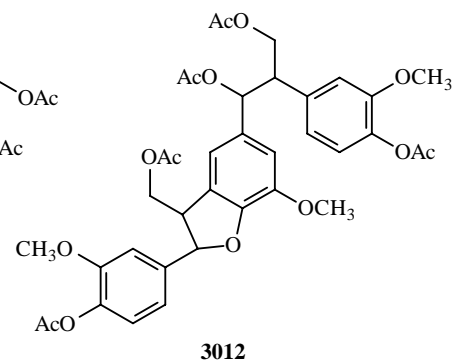
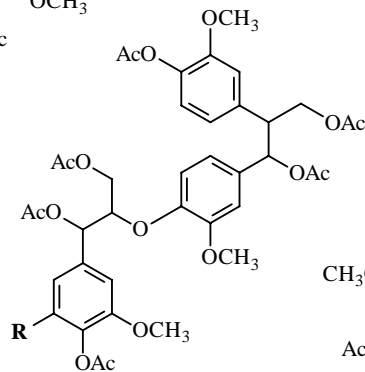
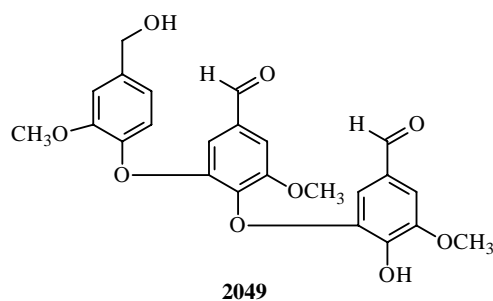
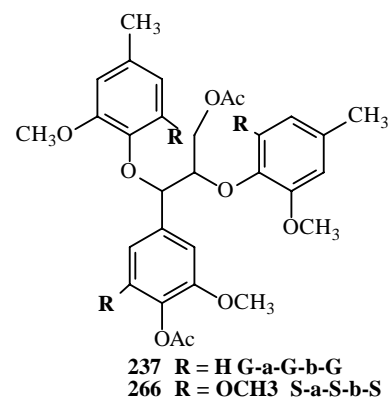
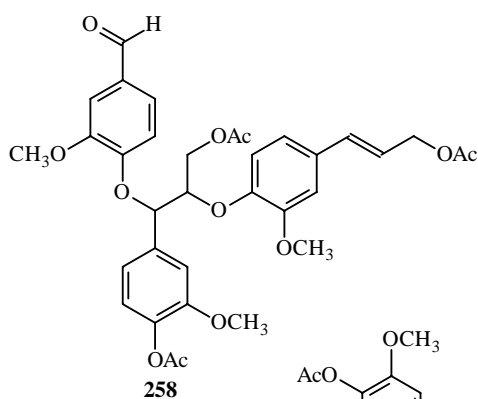
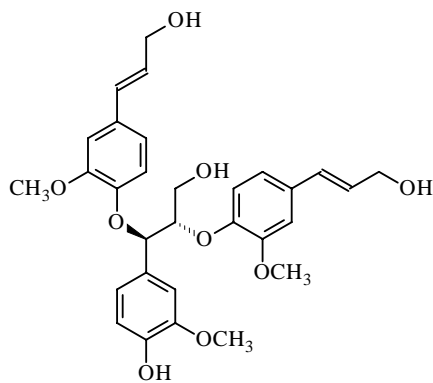
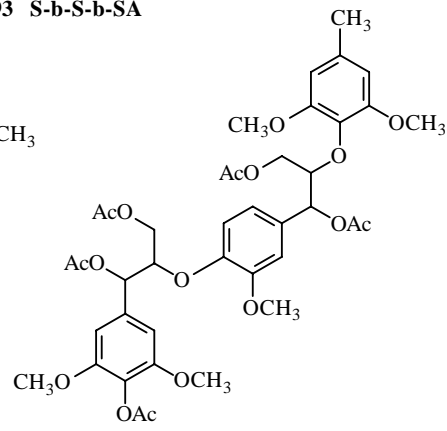
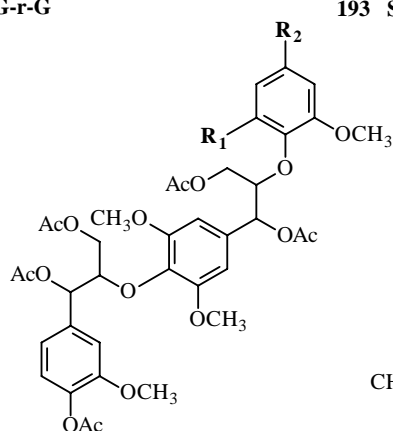
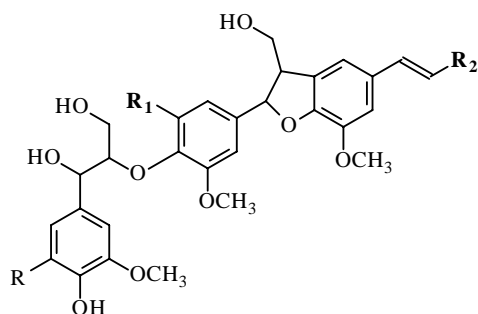
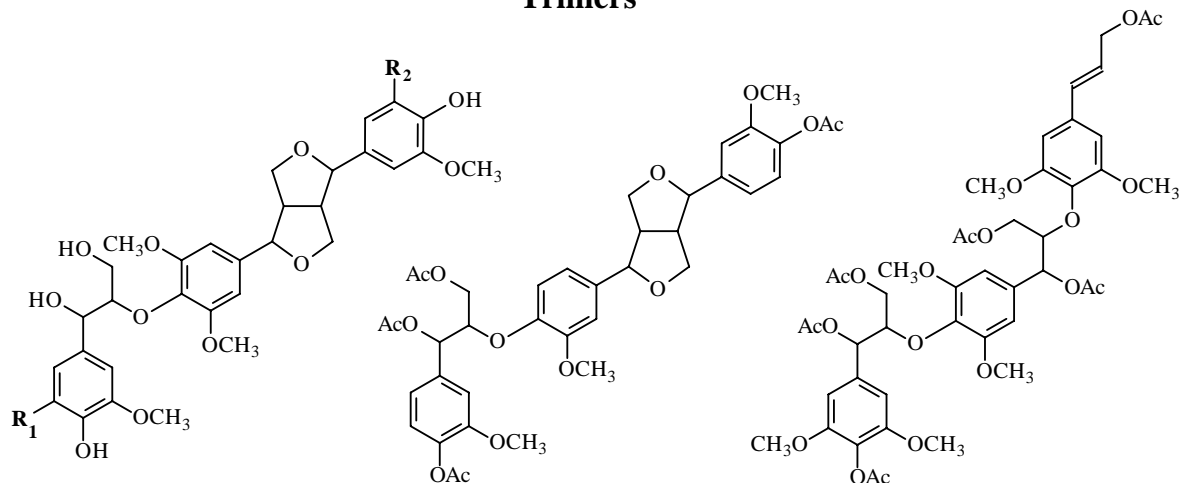
5-5 Dimers

 β -1 Dimers

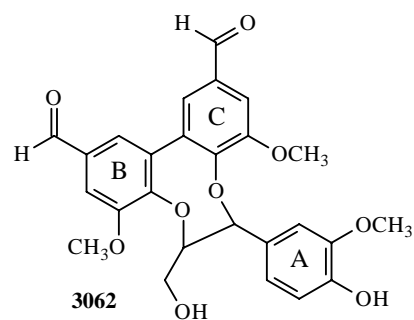
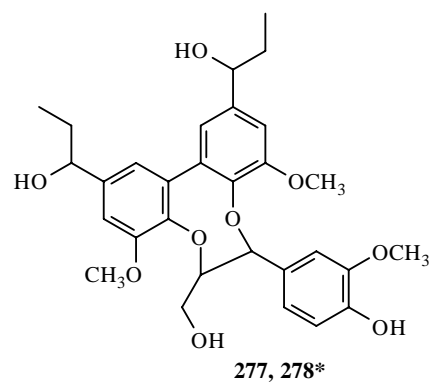
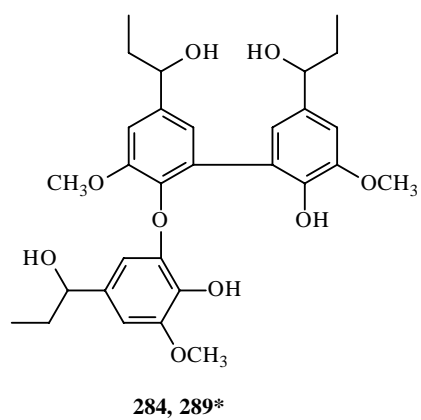
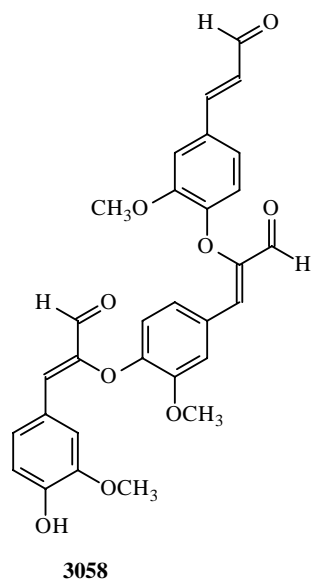
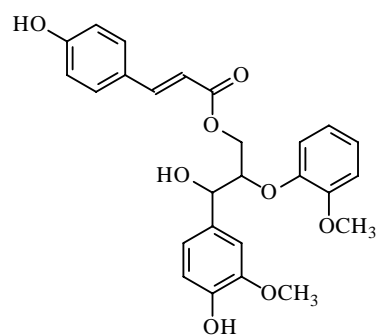
5-O-4 Dimers



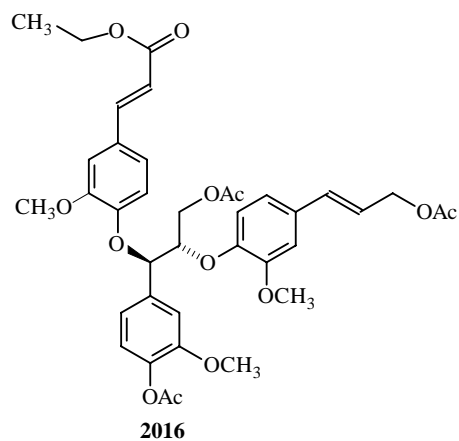
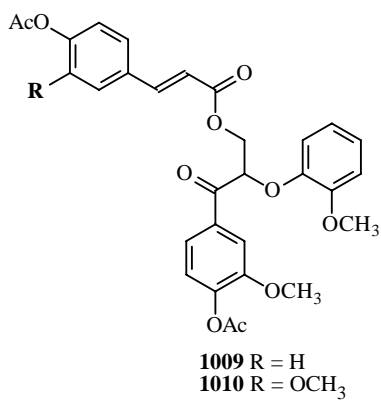
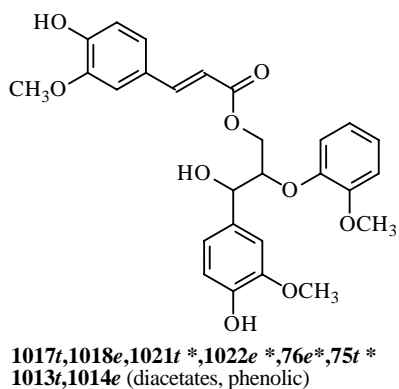
Trimers



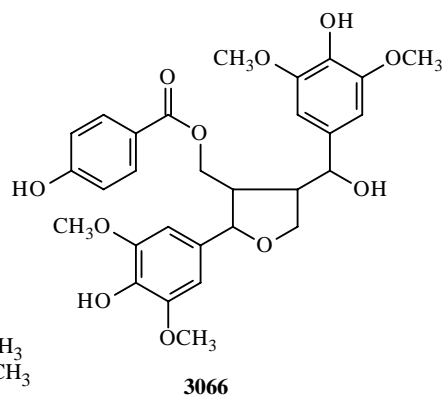
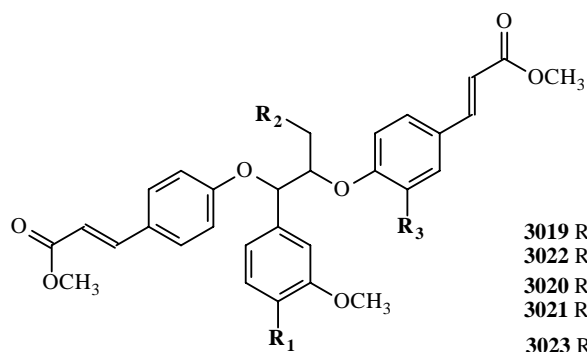
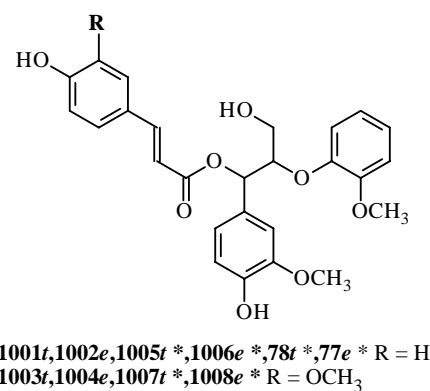
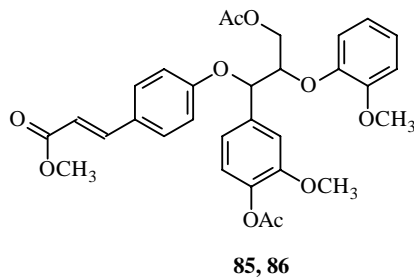
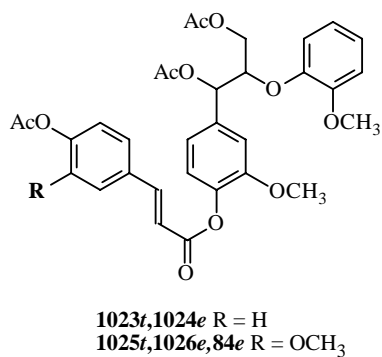
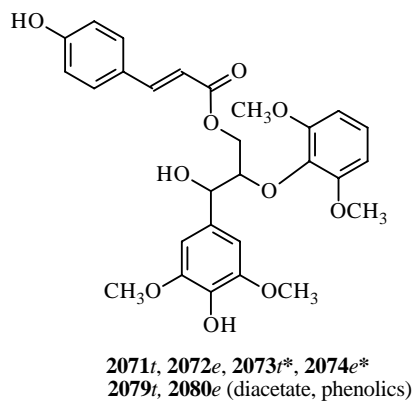
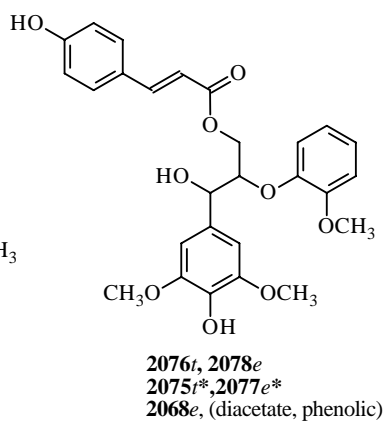
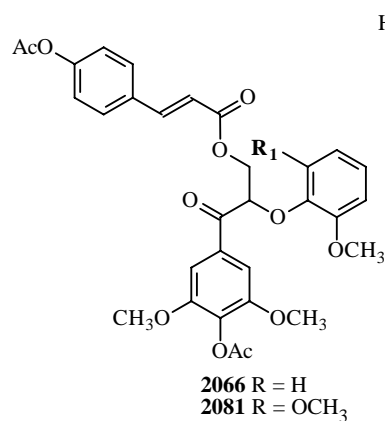
Trimers Ctd

Trimers Containing Ferulate or *p*-Coumarate

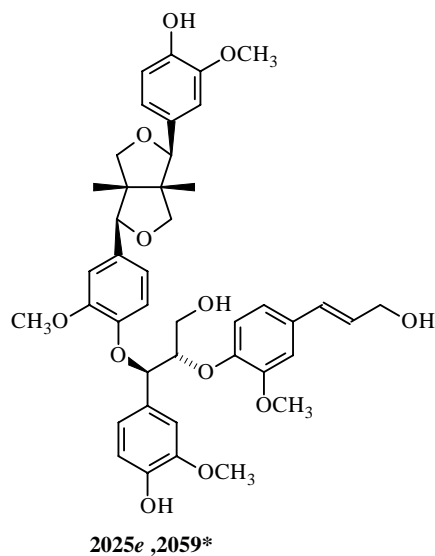
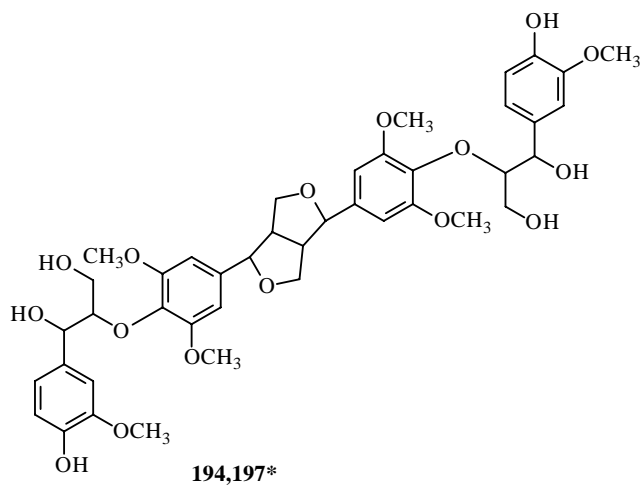
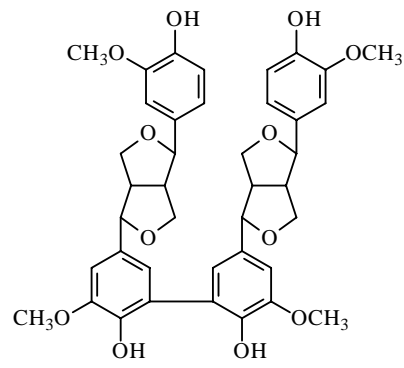
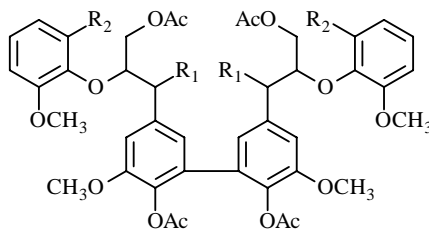
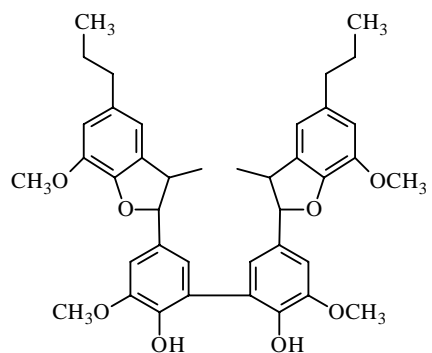
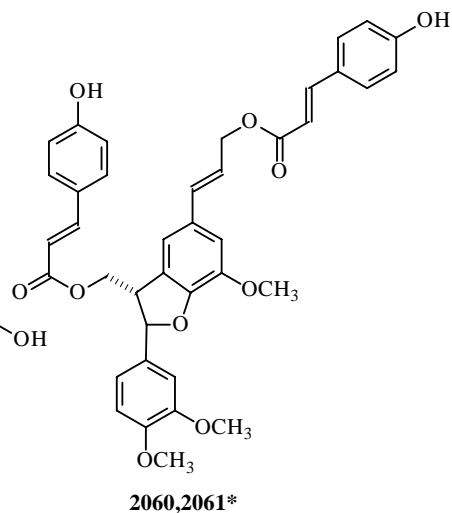
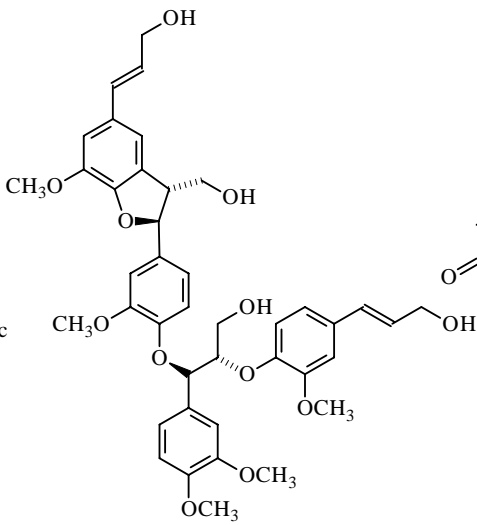
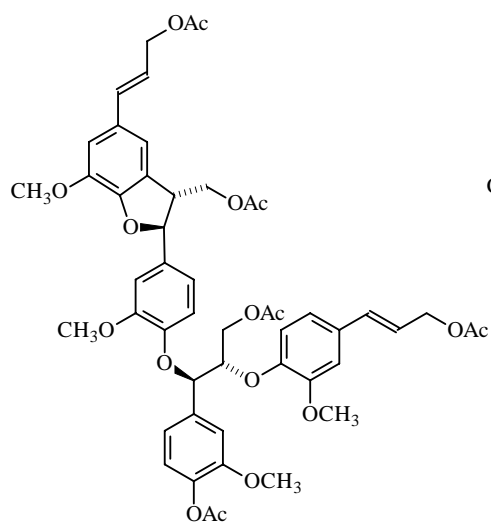
1015t, 1016e, 1019t, 1020e, 82t, 83e, 1011t, 1012e (diacetates, phenolic)



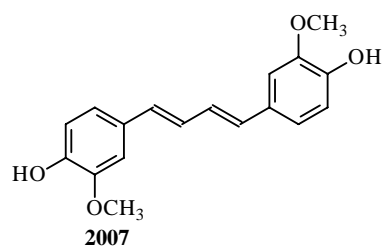
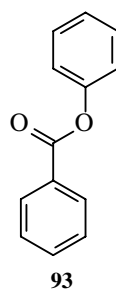
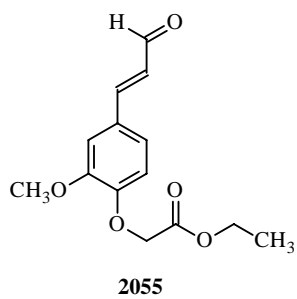
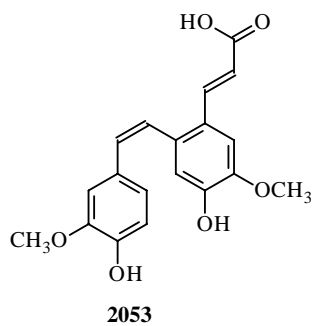
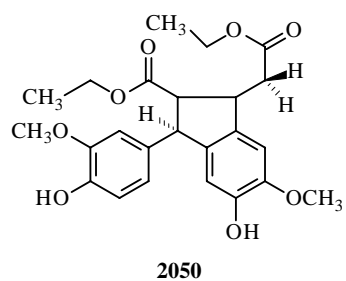
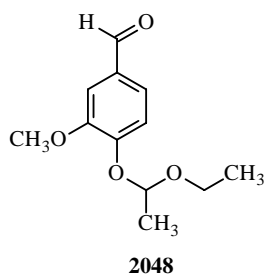
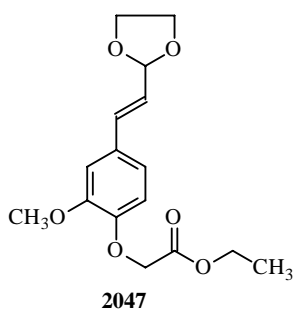
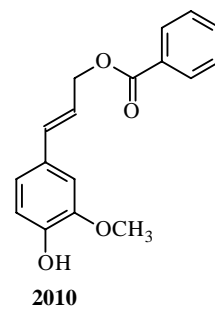
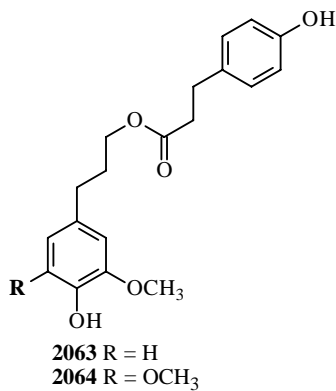
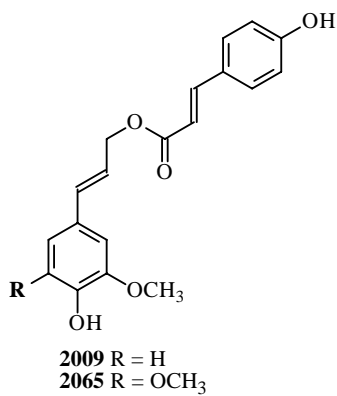
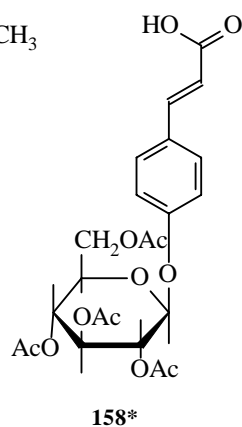
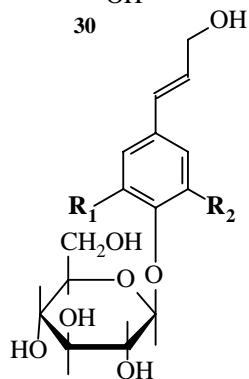
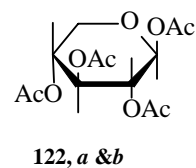
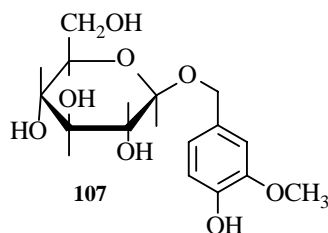
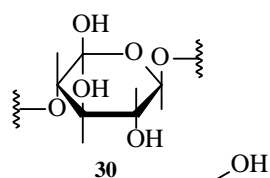
More Trimers Containing Ferulate, *p*-Coumarate or *p*-OH-Benzoate



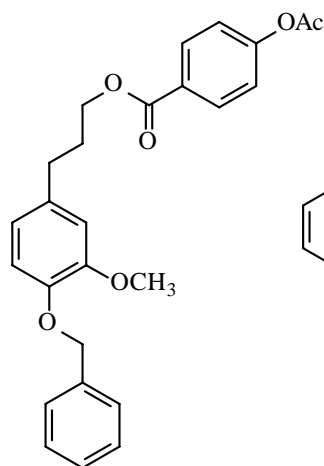
Tetramers



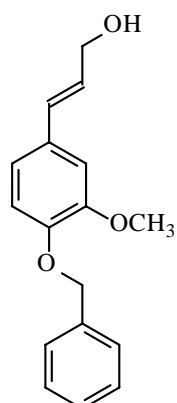
Misc. Compounds



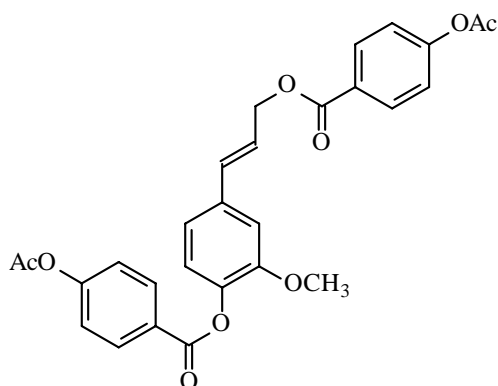
Misc. Compounds



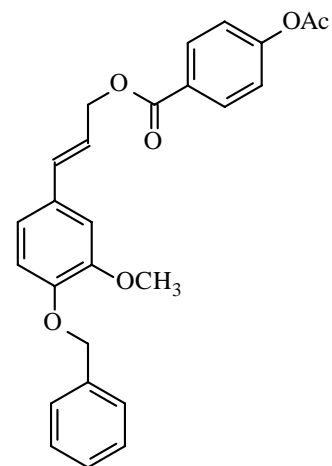
251



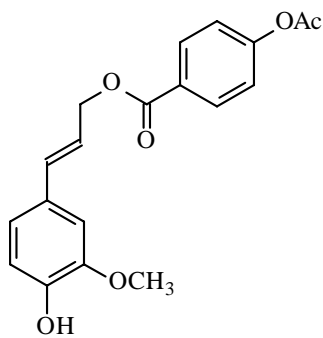
252



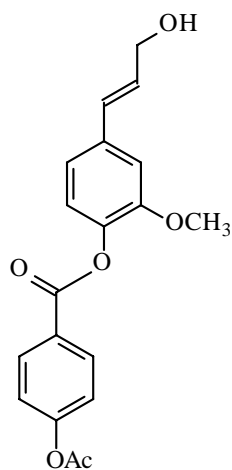
253



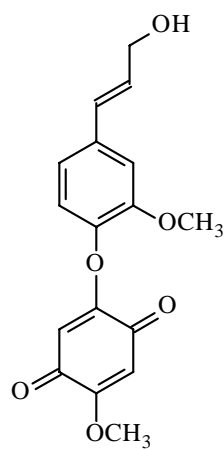
254



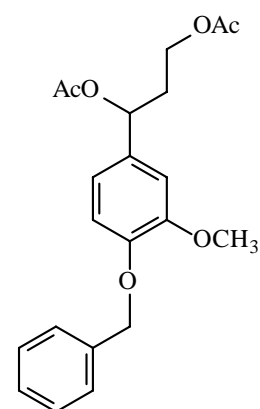
255



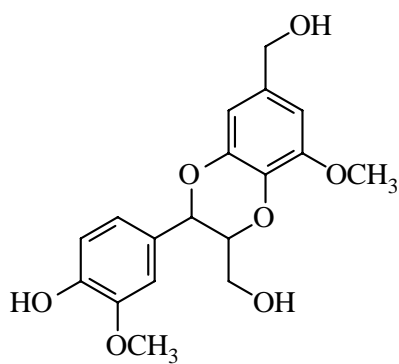
256



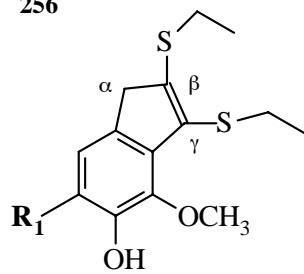
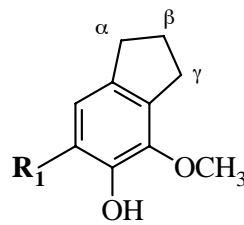
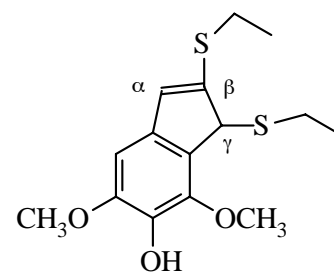
257



3029



3043, 3044*

3046 $R_1 = \text{OCH}_3$
3049 $R_1 = \text{H}$ 3045 $R_1 = \text{OCH}_3$
3048 $R_1 = \text{H}$ 

3047