Analysis of gas chromatography/mass spectrometry data for catalytic lignin depolymerization using positive matrix factorization

Journal:	*Green Chemistry*
Manuscript ID	GC-ART-05-2018-001474
Article Type:	Paper
Date Submitted by the Author:	11-May-2018
Complete List of Authors:	Gao, Yu; Washington University, Department of Energy, Environmental & Chemical Engineering Walker, Michael; Washginton University in Saint Louis Barrett, Jacob; University of California, Santa Barbara, Department of Chemistry and Biochemistry Hosseinaei, Omid; University of Tennessee Harper, David; The University of Tennessee, ; Ford, Peter; University of California, Santa Barbara, Department of Chemistry and Biochemistry Williams, Brent; Washington University in Saint Louis Foston, Marcus; Washington University, Department of Energy, Environmental & Chemical Engineering;

Article type: Full paper

Website www.rsc.org/greenchem
Impact factor* 9.125

Journal expectations To be suitable for publication in *Green Chemistry* articles must report innovative research on the development of alternative sustainable technologies demonstrating a significant advance in green and sustainable chemistry.

Article type: Full paper Original scientific work that has not been published previously. Full papers do not have a page limit and should be appropriate in length for scientific content.

Journal scope Visit the *Green Chemistry* website for additional details of the journal scope and expectations.

Green Chemistry provides a unique forum for the publication of innovative research on the development of alternative sustainable technologies. The scope of *Green Chemistry* is based on, but not limited to, the definition proposed by Anastas and Warner (Green Chemistry: Theory and Practice, P T Anastas and J C Warner, Oxford University Press, Oxford, 1998). Green chemistry is the utilisation of a set of principles that reduces or eliminates the use or generation of hazardous substances in the design, manufacture and application of chemical products.

Green Chemistry is at the frontiers of this interdisciplinary science and publishes research that attempts to reduce the environmental impact of the chemical enterprise by developing a technology base that is inherently non-toxic to living things and the environment. Submissions on all aspects of research relating to the endeavour are welcome.

The journal publishes original and significant cutting-edge research that is likely to be of wide general appeal. Coverage includes the following, but is not limited to:

- the application of innovative technology to establish industrial procedures
- the development of environmentally improved routes, synthetic methods and processes to important products
- the design of new, greener and safer chemicals and materials
- the use of sustainable resources
- the use of biotechnology alternatives to chemistry-based solutions
- methodologies and tools for measuring environmental impact and application to real world examples
- chemical aspects of renewable energy

All items must be written so as to be widely accessible (conceptually) to chemists and technologists as well as, for example, final year undergraduates.

Reviewer responsibilities Visit the Reviewer responsibilities website for additional details of the reviewing policy and procedure for Royal of Society of Chemistry journals.

When preparing your report, please:

- Focus on the originality, importance, impact and reliability of the science. English language and grammatical errors do not need to be discussed in detail, except where it impedes scientific understanding.
- Use the journal scope and expectations to assess the manuscript's suitability for publication in *Green Chemistry*.
- State clearly whether you think the article should be accepted or rejected and include details of how the science presented in the article corresponds to publication criteria.
- Inform the Editor if there is a conflict of interest, a significant part of the work you cannot review with confidence or if parts of the work have previously been published.

Thank you for evaluating this manuscript, your advice as a reviewer for *Green Chemistry* is greatly appreciated.

Dr Katie Lim Executive Editor
Royal Society of Chemistry, UK

Professor Philip Jessop Editorial Board Chair
Queen's University, Canada

**2016 Journal Citation Reports®, (Thomson Reuters, 2017)*

Dear Prof. Buxing Han,

Enclosed is our manuscript entitled "**Analysis of the gas chromatography/mass spectrometry data for catalytic lignin depolymerization using positive matrix factorization**" for consideration as an original research article for publication in ***Green Chemistry***. Our work uses a factor analysis approach, called Positive Matrix Factorization (PMF), to analyze a large number of gas chromatography/mass spectrometry (GC-MS) datasets in a fashion that facilitates developing a deeper understanding of catalytic reaction pathways for lignin depolymerization into value-added products. This work will have broad and technical interest to the readers of ***Green Chemistry***.

Lignin valorization is of vital importance to the economic viability and positive life-cycle assessment of large-scale biorefineries. One potential approach to lignin utilization is its depolymerization into aromatic and phenolic chemicals. Thermal depolymerization of lignin generally results in such a wide product distribution that it can only be used for fuel production. However, recent research suggest that well-designed catalytic and reaction systems can be developed such that a narrow number of monomer-like compounds can be generated in high yield and selectivity. Even when applying selective catalysts that target aryl-ether bond cleavage, a complex distribution of compounds can still be generated. In this case, the resulting GC-MS chromatograms have a high-level of complexity that reflects the high-level of compositional complexity in the lignin depolymerization product. When probing large numbers of different lignin depolymerization products to explore the effect of several lignin depolymerization conditions, the quantity and complexity of these GC-MS datasets makes human analyses difficult and time-consuming. Therefore, computer-assisted signal processing can reduce GC-MS dataset complexity and transform GC-MS datasets into usable and actionable information.

In a previous effort, we applied a copper-doped porous metal oxide catalyst (CuPMO) and a methanol / dimethyl carbonate (MeOH/DMC) co-solvent to depolymerize lignin into aromatic and phenolic compounds. We used a non-precise metal catalyst along with renewable solvents that can be derived from biomass and CO_2 to process an aromatic-rich renewable resource, lignin. In this manuscript, we demonstrate the power of the application of PMF analysis to GC-MS datasets to analyze the reaction networks and product mixtures involved in lignin depolymerization reactions with our CuPMO and MeOH/DMC system under range of reaction conditions. **To our knowledge, this work represents the first application of PMF to chemical processing technology development.** We believe that these technological insights are of the caliber and interest expected by the ***Green Chemistry*** readership. We thank you very much for considering our manuscript for publication in this journal.

Sincerely,

Marcus Foston, PhD
Assistant Professor
Department of Energy, Environmental and Chemical Engineering
Washington University in St. Louis

Green Chemistry

PAPER

Analysis of gas chromatography/mass spectrometry data for catalytic lignin depolymerization using positive matrix factorization†

Received 00th January 2018,
Accepted 00th January 2018

DOI: 10.1039/x0xx00000x

www.rsc.org/

Yu Gao,[‡a] Michael J. Walker,[‡a] Jacob A. Barrett,[b] Omid Hosseinaei,[c] David P. Harper,[d] Peter C. Ford,[b] Brent J. Williams[a] and Marcus B. Foston*[a]

Various catalytic technologies are being developed to efficiently convert lignin into renewable chemicals. However, due to its complexity, catalytic lignin depolymerization often generates a wide and complex distribution of product compounds. Gas chromatography/mass spectrometry (GC-MS) is a common analytical technique to profile the compounds that comprise lignin depolymerization products. GC-MS is applied not only to determine product composition, but also to develop an understanding of catalytic reaction pathways and of relationships among catalyst structure, reaction conditions, and the resulting compounds generated. Although a very useful tool, the analysis of lignin depolymerization products with GC-MS is limited by the quality and scope of available mass spectral libraries and the ability to correlate changes in GC-MS chromatograms to changes in lignin structure, catalyst structure, and reaction condition. In this study, GC-MS data of the depolymerization products generated from organosolv hybrid poplar lignin using a copper-doped porous metal oxide catalyst and a methanol / dimethyl carbonate co-solvent was analyzed by applying a factor analysis technique, Positive Matrix Factorization (PMF). Several different solutions to the PMF model were explored. A 13-factor solution sufficiently explains the chemical changes occurring to lignin depolymerization products as a function of lignin, reaction time, catalyst, and solvent. Overall, seven factors were found to represent aromatic compounds, while one factor was defined by aliphatic compounds.

Introduction

Biorefineries have attracted widespread interest as a promising scheme for renewable energy, chemical, and material production.[1-4] If US second generation biofuel production targets are met in 2022, nearly 62 M dry tons/yr of lignin will be generated as a by-product, which is currently being under-utilized as low-grade fuel for heat and electricity.[5-8] Thus, technologies must be developed to efficiently use both carbohydrate and lignin fractions of biomass for biorefineries to be economical and have a minimal environmental footprint.

Lignin is one of three main plant cell wall components (i.e., lignin, cellulose, and hemicellulose), comprising ~15-30% of the dry weight of lignocellulosic biomass.[9] Described as a random copolymer, lignin is comprised of three major subunits [i.e., *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) monomers] linked by different types of inter-unit linkages (Figure 1). The native, aromatic-rich substructure of lignin makes it an ideal resource for the production of renewable aromatic and phenolic chemicals.[2, 10-12] Evolution has made lignin an integral component of the defensive and support structures within plant cell walls. Accordingly, lignin is highly resistant to biological and chemical deconstruction. Therefore, thermal depolymerization

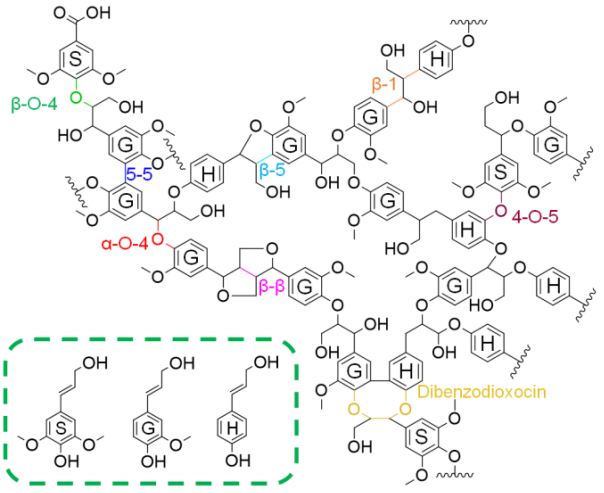

Figure 1. A structural representation of lignin, depicting various linkages and three monomers. In the bottom-left, the three monomers [i.e., coniferyl (G), sinapyl (S), and *p*-coumaryl (H) alcohol] are shown.

[a] Department of Energy, Environmental and Chemical Engineering, Washington University in St. Louis, St. Louis MO, 63130, USA.
*Email: mfoston@wustl.edu
[b] Department of Chemistry and Biochemistry, University of California, Santa Barbara, Santa Barbara CA, 93106-9510, USA.
[c] RISE Bioeconomy, PO Box 5604, 11486 Stockholm, Sweden
[d] Center for Renewable Carbon, University of Tennessee, Knoxville, Knoxville TN, 37996, USA
‡ Authors contributed equally
† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

approaches require harsh reaction conditions and lead to wide product distributions that compromise downstream processing of any particular compound for chemical production.

In lignin, 50-75% of the inter-unit linkages are comprised of substructures that contain aryl-ether bonds.[13] Therefore, selective depolymerization of lignin into its monomers for chemical production at lower temperatures, which limits secondary reactions, is possible through catalytic systems that target the cleavage of aryl-ether bonds. One such promising catalytic route is hydrogenolysis which generates aromatic and phenolic derivatives as products.

Various catalysts have been evaluated for selective aryl-ether bond cleavage and lignin depolymerization.[14-19] For example, Song et al. reported using Ni catalyst on activated carbon, alumina, or porous silica to selectively convert birch wood lignin to GC-detectable phenolics in alcohols.[20] In another contribution, Ye et al. showed the selective production of 4-ethylphenolics from hydrogenolysis of lignin via noble metals (Pt, Pd, and Ru) on an activated carbon support.[21] Ford et al. studied a copper-doped porous metal oxide catalyst (CuPMO) for lignin depolymerization via hydrogenolysis in methanol (MeOH).[22-27] In addition to catalyzing aryl-ether hydrogenolysis, CuPMO promotes alcohol reforming to provide the necessary reducing equivalents for hydrogenolysis.

There are many catalytic lignin depolymerization conditions (e.g., biomass/lignin source, reaction temperature, reaction time, catalyst structure, catalyst loading, stirring speed, etc.) and phenomena that can affect lignin depolymerization reaction kinetics and networks, and thus the resulting product distribution. Therefore, controlling lignin-derived product selectivity and yield requires the ability to analyze lignin depolymerization products. More important than enabling product analysis would be the ability to leverage analytical results to develop a deeper understanding of lignin depolymerization reactions that facilitates the design of new catalysts and reaction systems for the conversion of lignin into desired products. However, due to the compositional heterogeneity and complexity of lignin and lignin depolymerization products, such analysis and its utilization is challenging.

Gas chromatography/mass spectrometry (GC-MS) has long been applied in many catalytic lignin depolymerization studies.[28-34] Even when applying selective catalysts that target aryl-ether bond cleavage, a complex distribution of compounds can still be generated. In this case, the resulting GC-MS chromatograms have a level of complexity reflecting the compositional complexity in the lignin depolymerization product. These complex GC-MS chromatograms generally consist of numerous chromatographic features (i.e., peaks), many of which are unresolved and associated with mass spectral fragmentation patterns that are not in available mass spectral libraries. The manual comparative analysis of GC-MS datasets for a small number of lignin depolymerization products, comparing their compositional distributions, can be performed.[35-38] However, when probing large numbers of products to explore the effect of several lignin depolymerization conditions, the quantity and complexity of these GC-MS datasets make human analyses difficult and time-consuming. Hence, computer-assisted signal processing can reduce GC-MS dataset complexity and transform GC-MS datasets into usable and actionable information.

As a result, this study focuses on applying a Positive Matrix Factorization (PMF) technique to analyze a large GC-MS dataset by grouping lignin depolymerization products, or more specifically their chromatographic features, according to mass spectral (i.e., ion fragmentation) similarity. Thus, all GC-MS detected products, including products attributed to unresolved chromatographic features or chromatographic features that have mass spectral fragmentation patterns not present in the mass spectral library, can be characterized by their chemical and structural similarities, defining factors that often represent chemical homologs. PMF attempts to reduce the complexity of GC-MS datasets, and includes mechanisms that provide chemically-relevant solutions.[39] PMF has been widely used in the atmospheric chemistry community to analyze bulk MS measurements,[40, 41] and has recently been extended to more chemically-resolved GC-MS measurements of organic aerosol composition.[42, 43]

The present study was initiated to demonstrate the power of the application of PMF analysis to GC-MS datasets in the analysis of the complex reaction networks and product mixtures characteristic of lignin depolymerization. In total, 30 different reaction conditions were applied, from which, were collected 30 different sets of lignin depolymerization products. Comparing the products from 30 such samples via traditional GC-MS analysis, and more importantly, extracting a meaningful understanding of how key reaction conditions affect lignin depolymerization pathways would be very difficult, at best. To our knowledge, this is the first time that PMF analysis of GC-MS datasets has been applied in the investigation of chemical processing, and more specifically, lignin depolymerization.

Results and discussion

This study uses a methanol-soluble and methanol-insoluble fraction of lignin extracted from a hybrid poplar biomass source, respectively denoted as MS and MIS lignin. Depolymerization reactions were conducted at 300 °C using three reaction conditions: 1) without catalyst in MeOH (non-catalyzed), 2) with CuPMO in MeOH (MeOH), and 3) with CuPMO in a MeOH and dimethyl carbonate (DMC) mixture (MeOH/DMC). DMC was applied as a co-solvent with MeOH because in our previous work[26], we found that DMC would O-methylate phenolic intermediates generated from catalytic hydrogenolysis. More importantly, we also found that the resulting aromatic methyl ether products from that O-methylation are much less susceptible to aromatic ring hydrogenation pathways than their phenolic counterparts. This study was done in a time-resolved fashion, collecting products at reaction times of 1, 2, 3, 6, and 9 h. Gaseous products of MeOH reforming and lignin depolymerization were analyzed by GC with a thermal conductivity detector (GC-TCD). Solid residues remaining after lignin depolymerization (i.e., lignin that has undergone chemical modification such that it is methanol-insoluble and/or char) were analyzed by dioxane extraction, nitric acid digestion, and thermogravimetric analysis (TGA). Liquid

products (i.e., methanol-soluble lignin remaining after reaction that may or may not have undergone chemical modification and GC-detectable lignin depolymerization products) were analyzed by gel permeation chromatography (GPC) and GC-MS (see Supplemental Information for details on methods). Primarily though, GC-MS data is the main focus of this report. PMF analysis of the GC-MS data was applied to understand how lignin source (i.e., MS and MIS lignin), reaction time, and the presence of CuPMO and/or DMC affects the GC-detectable (i.e., low-molecular weight and volatile) product distribution of lignin depolymerization. We found that, compared to traditional (manual) peak integration and assignment analysis for GC-MS datasets, our PMF analysis significantly reduces the time required to complete data processing, allows compounds to be chemically classified that are not in a MS library, and facilitates the analysis of the unresolved complex mixtures (UCMs).

PMF analysis.

PMF and principal component analysis (PCA) are similar types of factorization analysis methods that seek to identify the dominant factors that cause variance within a set of data.[44] PMF is a bilinear, unmixing model in which a dataset matrix is assumed to be comprised of the linear combination of factors with constant profiles that have varying contributions across the dataset. While PMF and PCA are somewhat similar in their outcomes, PMF constrains its factor results to positive values, does not require the factors to be orthogonal, and better accounts for measurement uncertainty. As applied in this work, PMF analysis is for grouping chromatographic features into classes of compounds (i.e., factors) based on similarities in MS fragmentation patterns, or more succinctly, similarities in chemical structure. Our use of PMF analysis also involves a chromatographic binning technique that sums together sequential MS scans to reduce computational time and minimize the effects of retention time shifting. GC-MS chromatographic signal intensities therefore undergo pre-processing (i.e., retention time shift correction, background subtraction and internal standard normalization) and chromatographic binning before PMF analysis.

The combined chromatographic binning and PMF analysis of the GC-MS data from the 30 lignin depolymerization samples results in a set of PMF solutions where the number of factors within a solution is specified by the user. Therefore, the selection of a particular solution still requires a subjective choice by the user, which needs to be informed by an understanding of the PMF inputs. To settle upon a final solution and set of factors, solutions where we defined the number of factors as 2-18 were considered. Chemically distinct classes of compounds were identified in factors for solutions up to the 13-factor solution, but following that point factor splitting resulted in redundant factors that did not add additional chemical insight. Such factor splitting often results in factors that are composed of a single compound, which taken to its extreme would produce a factor for each compound detected and defeat the purpose of performing factor analysis. For solutions that are fewer than the optimal number of factors, resolvable factors are presumably superimposed. In our case, the 13-factor solution generated a set of factors that mostly did not display factor splitting yet yielded factors which provided the maximum chemical insight. Thus, a 13-factor solution (summarized in Table 1) was chosen and is the subject of further discussion. The mass spectrum for a given factor consists of a set of co-varying fragment ions that best recreates the input dataset upon a linear combination with the other factors in the solution. The electron ionization (EI) used in MS analysis produces ions from GC-separated compounds that fragment in a reproducible way, generating similar fragmentation patterns for compounds that have similar structural moieties. Therefore, the factor mass spectrum is useful in the identification of a homologous series of compounds that defines the factor.

The retention time-series data of a given factor upon further post-processing can be used to reconstruct a factor average chromatogram for a given factor eluting from the column. Factor average chromatograms can be reconstructed for either any sub-set of PMF input samples (e.g., different catalyst, lignin, and/or reaction conditions) or for all of the samples. For example, Figure S1A shows the factor average chromatograms constructed in this manner for product samples from catalyzed reactions in MeOH/DMC (green), catalyzed reactions in MeOH (blue), and non-catalyzed reactions in MeOH (black). Figure S1D shows the factor average chromatogram for the combination of all reaction conditions. The binned abundances displayed in a factor average chromatogram corresponds to the amount of the signal (i.e., output from the mass spectrometer) that the model apportions to that factor at that retention time. The factor

Table 1. Summary of the chemical assignment for individual factors with their major characteristic fragment ions.

	Defined Factor	Major Characteristic *m/z*
Factor 1	less polar and/or more volatile aromatics	39, 50, 65, 74, 93
Factor 2	air and other light contaminates	31, 40, 44
Factor 3	less polar and/or more volatile aromatics	39, 51, 63, 77, 91, 107
Factor 4	aliphatics	41, 55, 69, 83, 97, 111
Factor 5	carboxylics and carbonates	31, 47, 75
Factor 6	benzoates	39, 65, 93, 121
Factor 7	more polar and/or less volatile aromatics	39, 53, 67, 79, 95, 109, 123
Factor 8	dimethoxy benzyls	39, 65, 79, 91, 107, 119, 135, 151
Factor 9	methoxy phenyls	39, 55, 65, 79, 94, 105, 122, 137
Factor 10	trimethoxy benzyls	39, 45, 52, 66, 79, 92, 105, 120, 136, 148, 167, 181
Factor 11	unresolved compounds	31, 41, 55, 65, 77, 91, 105, 115, 135, 149, 165, 179, 191
Factor 12	Column residues and heavier contaminants	150, 165, 195, 253, 315, 393, 408, 451
Factor 13		135, 156, 179, 197, 218, 239, 255, 315, 373, 393, 451, 529

abundance (Figure S1C) of one (or multiple) factor(s) can be calculated by integrating the entire area of the factor chromatogram for a given sample (or set of samples). Due to the complexity of the products across the samples and differences in detector responses for different compounds, these factor abundances cannot quantitatively be thought of as the mass of compounds comprising different factors. The reported factor abundance is therefore in an arbitrary unit. However, considered in a more qualitative manner, a much higher abundance for a specific factor would suggest that the higher abundance factor is compositionally favored in the lignin depolymerization sample. More useful, when comparing two different lignin depolymerization samples, an increase in factor abundance for a specific factor when comparing one sample to another would suggest that one sample compositionally favors that factor relative to the other sample. Again, in that case, the intensity difference between factors is not quantitative, but suggestive in nature to the degree of difference. Thus, the comparisons of factor abundances can provide useful insights into the different processes occurring for different samples and under different reaction conditions.

In combination, the factor mass spectrum and average chromatogram are used to assign the chemical identity of a given factor. For example, in the 13-factor solution, the Factor 1 mass spectrum contains fragment ions with *m/z* values of 39, 50, 65, 74, and 93 (Table 1). These fragment ions match common diagnostic ions that originate from and represent fragments of compounds that contain aromatic/phenolic moieties. The factor average chromatogram for Factor 1 indicates that compounds assigned to Factor 1 elute at early retention times, which corresponds to aromatics that are less polar and/or more volatile (due to the GC column type and GC oven heating ramp program). To verify the factor assignments, peaks in the factor average chromatograms were assigned by identifying compounds eluting at the same retention time for GC-MS datasets from the 30 lignin depolymerization samples. Due to the limited number of known compounds in the mass spectral library and the number of unresolved peaks, not all peaks could be assigned with a high level of certainty. However, the majority of identified peaks suggested that the classifications of the factors based on the factor mass spectrum are reliable. A complete list of MS library-identified compounds from the GC chromatogram of all 30 samples is provided in Table S1.

The 13 identified factors include both resolved and unresolved GC-detectable products. Due to the reliance on differences in mass spectral fragmentation patterns in separating the factors, it is important to note that a single compound may contribute to more than a single factor. Table 1 defines each factor's chemical identity based on the characteristic *m/z* values in the factor mass spectrum and the identified compounds associated with the factor. Table S2 provides further details about fragment ions and their chemical assignment. Compounds in Factors 1 and 3 are primarily low polarity aromatic/phenolic compounds (Table 1, Figure S1, and Figure S3). Factor 3 compounds generate fragment ions that are associated with benzyl moieties (Table 1 and Figure S3). The differences between Factors 1 and 3 are driven by the association of fragments related to phenol and methyl 4-hydroxybenzoate with Factor 1, which are absent from Factor 3. Similarly, Factor 7 compounds are also aromatic but tend to be more polar and/or less volatile than compounds in Factors 1 and 3 (Table 1 and Figure S7). Factor 6 compounds have fragment ions that are associated with benzoate moieties (Table 1 and Figure S6). The Factor 6 average chromatogram is dominated by two compounds: methyl 4-hydroxybenzoate and methyl 4-hydroxy-3-methoxybenzoate. Other classes of aromatic compounds are found in Factors 8, 9, and 10. Specifically, Factors 8,

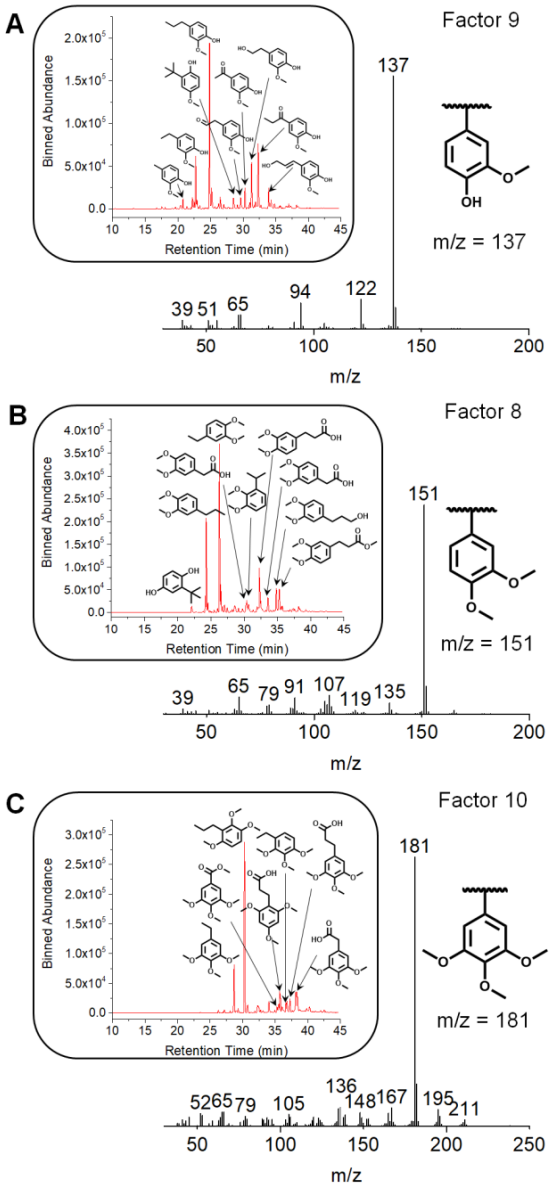

Figure 2. Factor average chromatograms and mass spectra for Factor 8 (dimethoxy benzylic), Factor 9 (methoxy phenolic), and Factor 10 (trimethoxy benzylic) of the 13-factor PMF solution for GC-MS datasets of lignin depoylmerization samples from MS/MIS lignin that have udergone depolymerization for 1-9 h using non-catalyzed, MeOH, and MeOH/DMC conditions.

9, and 10 contain compounds that have fragment ions associated with dimethoxy benzylic, methoxy phenolic, and trimethoxy benzylic moieties, respectively (Table 1, Figure S8, Figure S9, and Figure S10). The mass spectra for Factors 8, 9, and 10 are dominated by fragment ions at *m/z* 151, 137, and 181, respectively (Figure 2). Similarities in Factors 1 and 3 due to their association with fragments related to phenol and Factors 1 and 6 due to their association with fragments related to methyl 4-hydroxybenzoate suggest these factors may be the result of splitting. However, the 13-factor solution was the first solution to generate separate factors (i.e., Factors 8, 9, and 10) related to compounds that resemble lignin monomers. Though not done, we could have also visualized the aggregated factor abundances of Factors 1 and 3; Factors 1 and 6; and Factors 1, 3, and 6 to account for this splitting and further detect trends that inform our understanding of lignin depolymerization.

The mass spectrum of Factor 5 is dominated by a fragment ion at *m/z* 75, suggesting compounds in Factor 5 contain moieties that have a carboxylate and two additional carbons (e.g., methyl acetate, ethyl formate, propionate, etc.). Factors 4 and 11 contain unresolved chromatographic features within their factor average chromatograms. Factor 4 has primarily a low-retention time (low polarity and/or more volatile compounds) unresolved complex mixture (UCM) with mass spectral features indicative of substituted aliphatics (Table 1 and Figure S4). Conversely, Factor 11 features a high-retention time UCM (higher polarity and/or less volatile compounds) that includes mass spectral features indicative of both aliphatics and aromatics (Table 1 and Figure S11). UCMs appear as a hump or background feature in a chromatogram and represent a large number of co-eluting compounds. UCMs are commonly observed for petroleum or biomass-derived pyrolysis oils.[45, 46] Factors 4 and 11 suggest that our lignin depolymerization samples contain both polar and non-polar UCMs. The remaining factors can be assigned as measurement artifacts, which persist even after the pre-processing. Factor 2 comes from air within the GC-MS, with *m/z* 32, 40, and 44 attributed to oxygen, argon, and carbon dioxide, respectively. The variation in abundance of Factor 2 across the different samples is largely driven by the scaling differences introduced with the internal standard normalization. The column bleed from the GC column is the defining feature of Factors 12 and 13, and the variation in abundance across samples is driven by how similar the subtracted blank sample was to a given sample. Details of the factor mass spectra and average chromatograms for all 13 solution factors are given in Figures S1-S13.

Analysis of trends in GC-detectable products as grouped by PMF.

The overall GC-detectable aromatic compound production was monitored by combining the aromatic factors (Factors 1, 3, 6, 7, 8, 9, and 10). Figure 3A is the factor average chromatogram for the combined aromatic factors and the aggregated factor abundance for these aromatic factors across reaction conditions are in Figure 3B. Similarly, the aliphatic factor average chromatogram (Factor 4) is in Figure 3C with the factor abundance for the aliphatic factor across reaction conditions in Figure 3D.

As shown in Figure 3, catalyzed depolymerizations (MeOH and MeOH/DMC) generated higher factor abundances of GC-detectable products (i.e., aromatic and aliphatic) compared to non-catalyzed depolymerizations. Additionally, the overall production of aromatic compounds in catalyzed depolymerizations was significantly higher than that without catalyst. This increase for catalyzed depolymerizations was due to the catalytic activity[26] of CuPMO producing H_2 (Figure S16) and performing aryl-ether hydrogenolysis. Over the course of the entire reaction, the production of GC-detectable products remained relatively low for depolymerizations without catalyst, which indicated that most of the depolymerized lignin fragments in liquid products are larger molecular weight species (i.e., non- GC-detectable) that were unreacted or subject to condensation reactions.

Aside from our work[26], Bernt et al. also found that anisole and ethoxybenzene are much less reactive (i.e., slower rate of conversion) over CuPMO in MeOH than are phenol or guaiacol.[24] Phenol conversion was attributed primarily to reaction pathways: 1) reduction to cyclohexanol (k_{obs} ~0.5 h^{-1}), 2) methylation of the aromatic ring to give cresols (k_{obs} ~0.1 h^{-1}), and 3) O-methylation at phenolic alcohols to form anisole (k_{obs} ~0.1 h^{-1}). The primary product of anisole over CuPMO in MeOH was benzene, a very stable product, which resulted from (C_{aryl}-$O_{methoxy}$) hydrogenolysis of the methoxyl group; although, at longer reaction times some ring hydrogenation to methoxy-cyclohexane occurred. These results suggest that DMC and O-methylation of fragments from lignin depolymerization increases pathways to products that are less susceptible to certain reactions, such as ring hydrogenation reactions, when compared to their phenolic counterparts. We consider hydrogenation of phenolics as undesirable due to the loss of aromaticity and broadening of product distribution. With increasing reaction time, the abundance of the aliphatic factor increased for MeOH samples (Figure 2D), while a relatively lower abundance of the aliphatic factor was detected for MeOH/DMC samples. Observed relative increases for the abundances of Factors 8 and 10, which represent increases in O-methylated aromatics, are the highest for MeOH/DMC samples.

Factors 1 and 3 represent low polarity aromatic compounds (Figure S1 and S3), while Factor 11 compounds are associated with a more polar and/or less volatile UCM (Figure S11). The abundance of both Factors 3 and 11 in non-catalyzed samples are relatively lower when compared to their abundance in MeOH and MeOH/DMC samples. Whereas, the abundance for Factors 3 and 11 increase with reaction time for MeOH and MeOH/DMC samples.

Factors 5, 6, and 7 represent compounds that contain carbonyl substituents that have been correlated with the production of reactive compounds similar to those that require stabilization and upgrading in lignin pyrolysis oil.[4, 47] Factor 5 compounds are aromatic and aliphatic compounds that generate carboxylate moieties upon EI, and are much more abundant in non-catalyzed samples than MeOH or MeOH/DMC samples (Figure S5). The average factor chromatogram of Factor 5 for non-catalyzed samples shows two major chromatographic features at retention times of 32.33 min [3-(3,4-

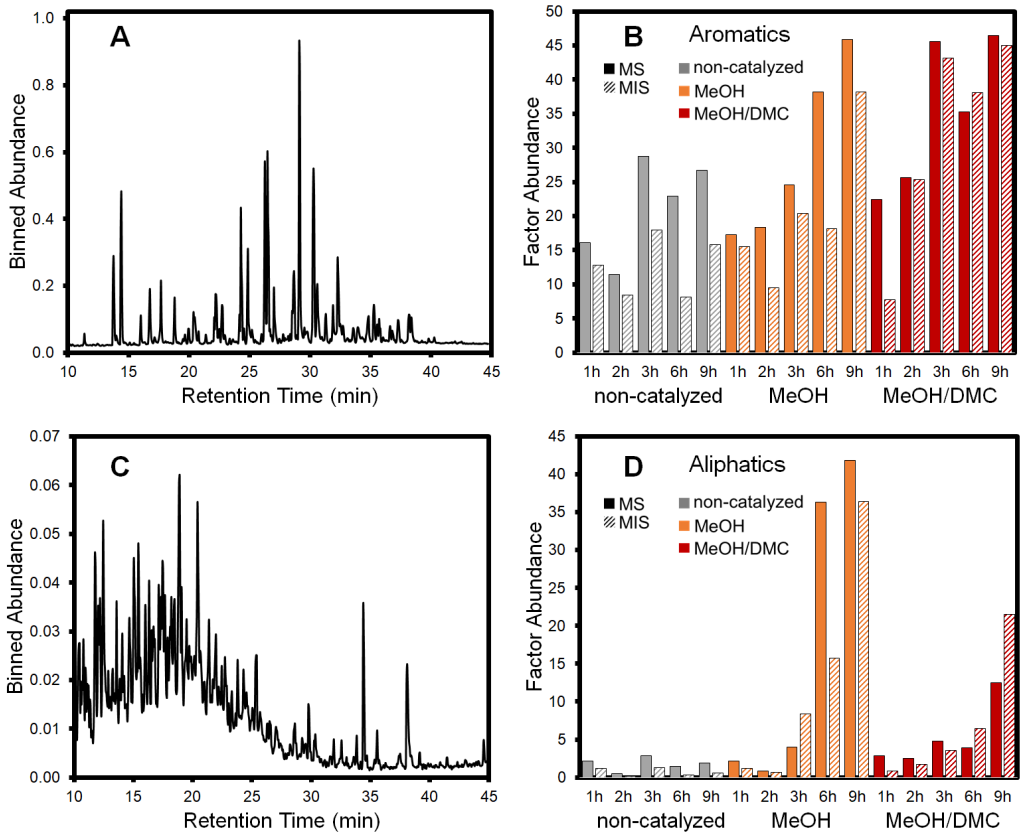

Figure 3. A) Factor average chromatograms of the combined aromatic factors (i.e., Factors 1, 3, 6, 7, 8, 9 and 10); B) combined aromatic factor abundance; C) factor average chromatograms of the aliphatic factor (i.e., Factor 4); and D) aliphatic factor abundance for lignin depoylmerization samples from MS/MIS lignin that have udergone depolymerization for 1-9 h using non-catalyzed, MeOH, and MeOH/DMC conditions.

dimethoxyphenyl)propanoic acid] and 38.23 min [methyl octadecenoate or methyl stearate]. Factor 6 compounds, defined as having benzoate moieties, are more abundant in non-catalyzed samples than MeOH samples and then more abundant in MeOH samples than MeOH/DMC samples. The average factor chromatogram of Factor 6 for non-catalyzed and MeOH samples shows two major chromatographic features at retention times of 29.11 min [methyl 4-hydroxybenzoate] and 30.61 min [methyl 4-hydroxy-3-methoxybenzoate], which are not present in the DMC/MeOH samples (Figure S6). The average chromatogram of Factor 6 for MeOH/DMC samples suggests that the DMC shifts reaction pathways in such a fashion that the production of compounds with benzoate moieties is reduced, in particular, methyl 4-hydroxybenzoate and methyl 4-hydroxy-3-methoxybenzoate. Factor 7 compounds are most prevalent in MeOH samples at reaction times of 6 and 9 h and MeOH/DMC samples at a reaction time of 9 h. For Factor 7, resolved compounds begin to elute from the column after 22 min that are primarily associated with non-catalyzed and MeOH/DMC samples, whereas ~ 5 – 32 min the average factor chromatogram presents as a UCM as a result of contributions from MeOH samples (Figure S7). The 2D ^{13}C-^{1}H heteronuclear single quantum correlation (HSQC) nuclear magnetic resonance (NMR) spectra of MS and MIS lignin substrates show they contain α-oxidized syringyl, α-oxidized guaiacyl and *p*-hydroxybenzoate monolignols (Figure S14). Factor 6 and 7 compound production could be, in part, linked to the release those monomers.

Factors 8, 9, and 10 most resemble the monomeric substructures of lignin (Figure 2 and 4). Factors 8 and 10 are individually defined by characteristic features of di- and tri-methoxylated compounds, which most likely result from O-methylation of phenolic intermediates from G and S lignin monomers. Consequently, these two factors have higher abundance in MeOH/DMC samples (Figure 4). The average chromatograms of Factors 8 and 10 for MeOH and MeOH/DMC samples each have two major chromatographic peaks at retention times of 24.27 min [4-ethyl-1,2-dimethoxybenzene] and 26.26 min [1,2-dimethoxy-4-propylbenzene] for MeOH and as 28.68 min [5-ethyl-1,2,3-trimethoxybenzene] and 30.29 min [1,2,3-trimethoxy-5-propylbenzene] for MeOH/DMC (Figure S8 and S10). Bernt el al. showed that 1,2-dimethoxybenzene was more reactive than anisole over CuPMO in MeOH; however, the primary product of 1,2-dimethoxybenzene was anisole resulting from the (C_{aryl}-$O_{methoxy}$) hydrogenolysis of a methoxyl group.[24] As a result, we expect that the di- and tri-methoxybenzene compounds resulting from DMC mediated lignin depolymerization would similarly be susceptible to methoxyl group hydrogenolysis. The

average chromatograms for Factors 8 and 10 suggest that the presence of the catalyst does promote phenolic O-methylation producing di- and tri-methoxylated aromatic compounds but that, even without catalyst, some phenolic O-methylation occurs. Notably, non-catalyzed reactions tend to produce a wide distribution of di- and tri-methoxylated compounds that are more polar than 1,2-dimethoxy-4-propylbenzene and 1,2,3-trimethoxy-5-propylbenzene due to the presence of various of functionalities on the fragment at the monolignol propyl substituent. The abundances for Factors 8 and 10 suggest that at longer reaction times (> 3 h), di- and tri-methoxylated aromatic compounds are converted into some other species. Conversely, Factor 9 represents characteristic features of compounds that contain aromatic rings with both hydroxyl and methoxyl substituents, which are generated from G lignin monomers that did not undergo O-methylation. The average chromatogram of Factor 9 for MeOH and MeOH/DMC samples each have a common major chromatographic peak at retention times of 24.87 min [2-methoxy-4-propylphenol]. Compounds related to Factor 9 show significantly higher abundance in the MeOH samples at short reaction times between 1-3 h than other samples. The decrease in Factor 9 abundance for MeOH samples beyond a reaction time of 3 h is likely due to reaction pathways shown to be prevalent for phenol conversion over CuPMO. The average factor chromatogram of Factor 9 for non-catalyzed samples has three additional major chromatographic peaks not observed in MeOH or MeOH/DMC sample at retention times of 29.64 min [2-(4-hydroxy-3-methoxyphenyl)acetaldehyde], 31.31 min [4-(2-hydroxyethyl)-2-methoxyphenol], and 32.33 min [1-(4-hydroxy-3-methoxyphenyl)propan-1-one] (Figure S9).

Comparing the lignin depolymerization products generated from MS and MIS lignin, overall GC-detectable aromatic compound production as well as Factors 8 and 10 abundances for MS and MIS samples were similar. This suggests there was little effect on product distributions caused by room temperature solubility or by chemical differences that existed for MS and MIS, which are highlighted in Table S3, Table S4, and Figure S14. However, Factor 6 abundances suggest that under non-catalyzed and MeOH conditions MS-derived samples contain more benzoates compounds, while Factor 4 abundances suggest that at reaction times of 6 and 9 h in MeOH, MS is more susceptible to aliphatic forming reaction pathways. Factor 5 abundances suggest that under non-catalyzed conditions and at a reaction times of 1 and 2 h, MIS-derived samples contain more carboxylate compounds.

Non-GC-detectable products.

Lignin oligomers are both intermediates and products of lignin depolymerization; however, these lignin oligomers are not detectable by GC-MS. Hence, the production of lignin oligomers was examined by GPC analysis. Untreated lignin and their depolymerized liquid products were directly injected into the GPC. Relative molecular weight values, including number-average molecular weight (M_n), weight-average molecular weight (M_w), and polydispersity index (PDI = M_w/M_n), were determined based on GPC retention times and a polystyrene standard calibration curve (Table S5 and S6). A higher PDI means a broader distribution of molecular

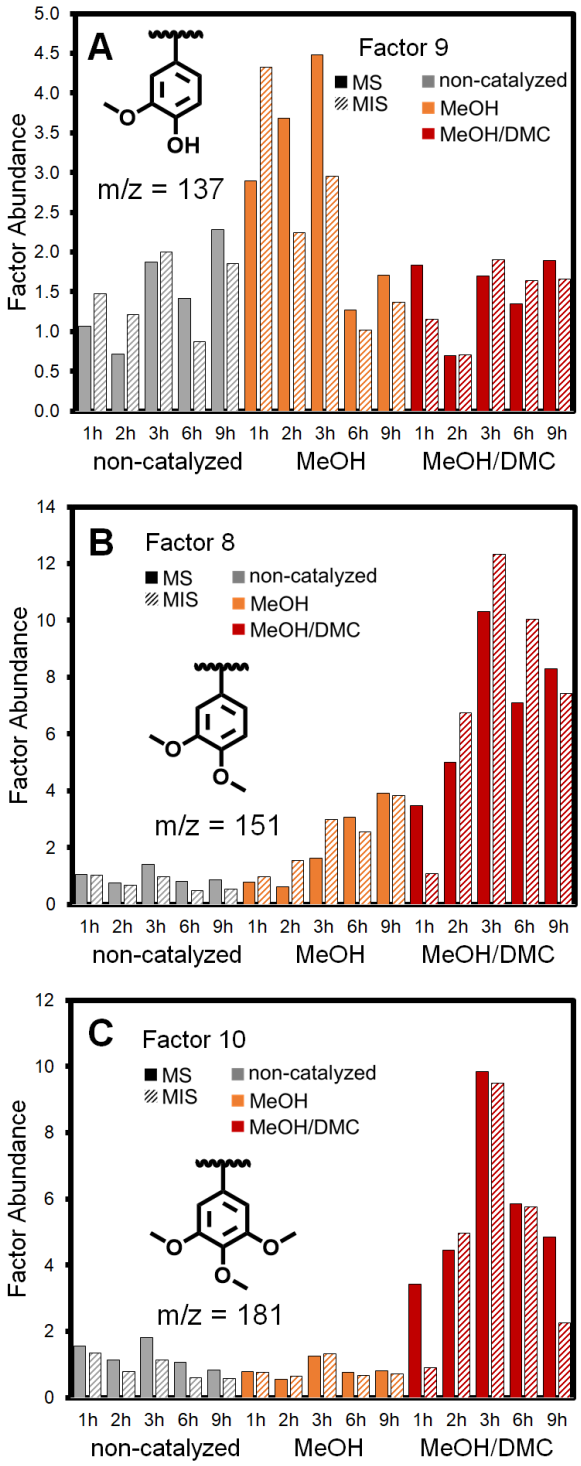

Figure 4. A) Factor 8 (dimethoxy benzylic); B) Factor 9 (methoxy phenolic); and C) Factor 10 (trimethoxy benzylic) abundance for lignin depoylmerization samples from MS/MIS that have udergone depolymerization for 1-9 h using non-catalyzed, MeOH, and MeOH/DMC conditions.

weights. Lignin depolymerization, as signified by chromatographic shifts toward longer retention times, occurred faster for catalyzed depolymerizations. After reaction times of 2 h for catalyzed depolymerizations, the broad peak from ~22-30 min representing untreated lignin almost completely disappears (Figure 5). However, this peak persists or even shifts toward shorter retention times for non-catalyzed depolymerizations at long reaction times. This suggest either that effective depolymerization is not occurring or that re-condensation of GC-detectable compounds is occurring. High abundance of Factors 5 and 6 seem to correlate with samples that are susceptible to GPC chromatographic shifts toward longer retention times; however, well-defined, co-varying trends between GPC behavior and factor abundance were difficult to extract.

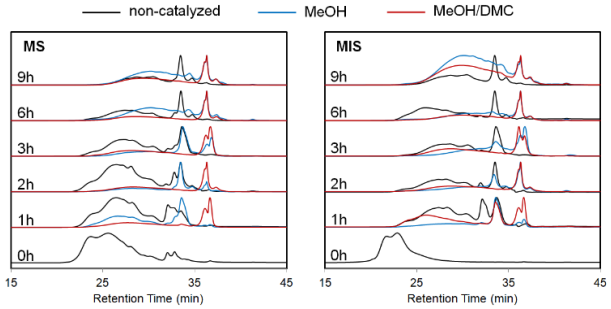

Figure 5. GPC chromatograms for untreated lignin and for product samples from MS and MIS liignin that have udergone depolymerization for 1-9 h using non-catalyzed, MeOH, and MeOH/DMC conditions.

Solid products.

Raw solid residues were also analyzed to study the leftover lignin and char formation as shown in Figure S15. Solid residues from each reaction were first separated from liquid products by filtration. Dioxane was used to extract dioxane-soluble solid products after lignin depolymerization. Note that both MS and MIS lignin are soluble in dioxane at room temperature. For the reactions without catalyst, the leftover solids after dioxane extraction were composed of dioxane-insoluble lignin and char. In this case, dioxane-insoluble solids were treated with nitric acid to determine char yields. Char formation increased with reaction time and was higher for MIS. For the reactions with catalyst, char formation was never observed and dioxane-insoluble solid content were higher than reactions without catalyst. In addition, dioxane-insoluble solid content decreased as a function of increasing reaction time. Dioxane-insoluble solids formation most likely result from reactions that modify the chemical structure of the lignin such that it is no longer soluble in dioxane. Dioxane-insoluble solids contents were higher for MIS lignin depolymerization samples when compared to MS lignin depolymerization samples.

Gas products.

Gas products were collected and analyzed to track the formation of H_2 production (Figure S16). No gas or H_2 formed for non-catalyzed depolymerizations. Gas products in both MeOH and MeOH/DMC catalyzed depolymerizations are mainly composed of H_2 and CO_2, with small amounts of CO and CH_4. The production of H_2 corresponds well with the increase overall aromatic compound production monitored by aromatic factors (Factors 1, 3, 6, 7, 8, 9, and 10) and the aliphatic compound production monitored by Factor 4.

Conclusions

In summary, upon PMF analysis of GC-MS datasets from 30 different lignin depolymerization products that were depolymerized as a function of lignin, reaction time, catalyst, and solvent, we determined a 13-factor solution sufficiently explains the chemical changes occurring. These 13 factors represent various classes of compounds based on similarities in chemical structure that best reconstruct the original lignin depolymerization PMF inputs. Factors include low and high polarity aromatic compounds, compounds with carbonyl moieties, compounds that resemble lignin monomers, and aliphatic compounds. Overall catalyzed depolymerizations generated higher factor abundances of GC-detectable products compared to non-catalyzed depolymerizations. In addition, we found that with increasing reaction time, the abundance of the aliphatic factor increased for MeOH samples while MeOH/DMC samples remained at a relatively low abundance. Thus, catalyzed depolymerization in the latter medium was superior at preserving product aromaticity. The products generated by reaction in MeOH in the absence of catalyst seemed to contain more compounds with carbonyl substituents. Lastly, we determined that there was little discernable difference in the GC-detectable products generated from MS and MIS lignin.

The complexity of lignin makes conducting fundamental research into its catalytic reactivity difficult. As a result, the development of analytical tools that can effectively capture this complexity and data processing methods that can interpret those analytical results is critical to the success of such fundamental research. Our results show that PMF analysis, as a computer-assisted signal processing tool to reduce GC-MS dataset complexity, can be applied to GC-MS datasets not only for the purposes of understanding lignin depolymerization but also a broad range of other chemical processes that involve complex reaction networks and product distributions.

Experimental

Materials.

Analytical grade methanol, n-decane, and reagent grade dimethyl carbonate (DMC) were used as purchased from Sigma-Aldrich. Lignin used in this work was extracted from *Populus spp*. biomass (see Supplementary Information). The CuPMO catalyst used in this work was synthesized by following the same procedure as reported by Ford et al. (see Supplementary Information).[26]

Lignin depolymerization.

The lignin depolymerization reactions have been conducted in stainless steel bomb reactors with internal volume of ~ 10 mL. The bomb reactor design used was from previous work by Ford et al.[22, 23, 26, 48, 49] Each reactor was charged with 100 mg of lignin (MS or MIS) and 100 mg of CuPMO catalyst. Either methanol (3 mL) only or pre-mixed methanol and dimethyl carbonate (2:1 ratio, 3 mL) solution with n-decane (1.76 µL) as internal standard was added into the reactor as solvent. The reactor was heated in an isotherm muffle furnace (Thermo Fisher Scientific) at 300 °C for reaction times of 1, 2, 3, 6, and 9 h. A series of reactions on the same lignin substrates were conducted without catalyst in methanol. Reactors were quenched within an ice water bath. Gas products were collected by an inverted graduate cylinder, which is pre-filled with water. The volume of gas products was measured by the displacement of water by gas collected in the cylinder. Gas composition was determined by GC-TCD. Solid residues and liquid products were separated by a vacuum filtration apparatus with a 0.45 µm nylon membrane filter. Solids were further washed by analytical grade methanol portion by portion until the total liquid products volume is 20 mL. Liquid products were collected for GPC and GC-MS analysis. Solid residues were further washed by dioxane (5 mL) three times to extract leftover untreated lignin (that was soluble in dioxane) from dioxane-insoluble lignin, char, and catalyst. The solids following dioxane extraction were subjected to TGA and nitric acid digestions to determine the amount of catalyst and char present.

Product characterization.

GC-MS was used to characterize the GC-detectable products from lignin depolymerization. GC-MS samples (1 µL) were injected on an Agilent GC system 7890A coupled with an Agilent 5975C mass spectroscopy with triple-axis detector. Triplicate injections were performed for the 6h MeOH/DMC samples derived from MS and MIS lignin. GC analysis was performed using a Restek fused silica RTX-50 capillary column (ID: 0.25 mm, film thickness: 0.5 µm, and length: 30 m) with the following program: 2 min at 35 °C and then ramped at 5 °C/min up to 300 °C for 5 min with helium as a carrier gas (splitting ratio: 10:1). The mass spectral scan rate was 1.6 scans/s with acquisition from *m/z* 30-600. GC-MS data was exported and analyzed through ChemStation Software. Identification of the compounds was carried out by comparing the mass spectra obtained with these from Palisade Complete Mass Spectral Database (600 K edition, Palisade Mass Spectrometry, Ithaca, NY). Non-GC-detectable products (gas, solid, and non-GC-detectable liquid products) characterization details are described further in the Supplementary Information section.

PMF analysis.

Positive matrix factorization (PMF) takes an input data matrix, $\boldsymbol{X}$(n x m), and separates the data into a time series matrix, $\boldsymbol{G}$ (n x p), and factor profile matrix, $\boldsymbol{F}$ (p x m), where p is the user-specified number of factors in the solution. A residual matrix, $\boldsymbol{E}$ (n x m), consists of the portion of the input data that cannot be captured by the factors to ensure mathematical continuity (equation 1).

$$\boldsymbol{X} = \boldsymbol{GF} + \boldsymbol{E} \quad [1]$$

The determination of the factors is achieved through the minimization of a function, $\boldsymbol{Q}$, which is the sum of uncertainty-weighted squared residuals:

$$Q = \sum_{n}^{i=1} \sum_{m}^{j=1} \frac{e_{ij}^2}{\sigma_{ij}^2}, \text{such that } g_{ik} \geq 0 \; and \; f_{kj} \geq 0 \quad [2]$$

where e_{ij} is the residual for a given value x_{ij}, and this is weighted by σ_{ij}, which corresponds to the uncertainty in the measured value. Constraining g_{ik} and $f_{kj}$to positive values ensures that nonsensical, negative solutions are not obtained.

Prior to PMF analysis, preprocessing of the GC-MS data is required, and was carried out in a custom software package developed within Igor Pro (version 6.37, Wavemetrics, Inc.), which is available upon request and will be more broadly available following further development. The pre-processing package includes: 1) loading in the GC-MS datasets, 2) correcting for shifts in chromatographic retention times across samples, 3) subtracting mass spectral background contributions identified by blank samples, 4) binning sequential MS scans, 5) generating uncertainty estimates for each bin, and 6) scaling a given sample's abundance by a user-defined factor. For the retention time shift corrections, a relatively simple, linear shift was applied to each based upon the change in retention time of the n-decane internal standard. Blank samples, both with and without DMC in addition to MeOH solvent with internal standard, were used to attempt to remove the influence of instrumental artifacts. However, this approach is often insufficient to remove all artifacts (e.g. the presence of air or column bleed), due to sample-to-sample variation.[42] A chromatogram binning approach, described in detail previously,[42] was used to decrease the computational burden of solving the PMF model, and bins were composed of 5 sequential mass spectral scans. In total, 667 bins for each of the 30 reaction conditions were constructed from 3335 mass spectral scans, corresponding to retention times of 8.97 - 44.76 minutes for each sample. The included mass spectra, which comprise the columns of the input data matrix, ranged from 30-600 Th.

One of the most challenging aspects of conducting PMF analysis of an entire chromatogram's dataset is deriving appropriate uncertainty estimates (σ_{ij}) for all input data values. Building upon previous efforts to use PMF on datasets from GC-MS work,[42, 50] the uncertainties were calculated as:

$$\sigma_{ij} = \begin{cases} 2 \times MDL_{ij}, if \; x_{ij} < MDL_{ij} \\ \sqrt{\left(x_{ij} \times precision\right)^2 + \left(MDL_{ij}\right)^2}, if \; x_{ij} > MDL_{ij} \end{cases} \quad [3]$$

where MDL_{ij} is a retention time and mass-to-charge (*m/z*) ratio dependent detection limit estimate based on a sample blank chromatogram, and precision is an estimate of the reproducibility of the instrument (10% for this study).

Finally, the data for each sample was scaled based upon the integrated abundance of the n-decane internal standard. The peak integrations were performed in the Igor-based "TERN" software (version 2.1.6), which utilizes a peak fitting approach for

quantification.[51] The PMF calculations were carried out in another custom software package (PMF Evaluation Tool version 3.00A) within Igor Pro, which utilizes the PMF2 solver.[41] To prevent an oversized impact from low abundance data within the matrix, *m/z* values with a signal-to-noise ratio (SNR) of less than 2 had their uncertainty values increased by a factor of 2, and values with SNR < 0.2 were excluded from the analysis entirely, as has been reported previously.[39] While the various Igor packages used are freely available, both the Igor Pro software and PMF2 solver require licenses for use.

Acknowledgements

Funding for this research was provided by the National Science Foundation Catalysis Program under award number CBET-1603692 and by the USDA National Institute of Food and Agriculture, Hatch project 1012359. We would also like to thank the Washington University's Engineering Communication Center for their help in revising this manuscript.

References

1 A. J. Ragauskas, C. K. Williams, B. H. Davison, G. Britovsek, J. Cairney, C. A. Eckert, W. J. Frederick, J. P. Hallett, D. J. Leak and C. L. Liotta, *Science*, 2006, **311**, 484-489.

2 A. J. Ragauskas, G. T. Beckham, M. J. Biddy, R. Chandra, F. Chen, M. F. Davis, B. H. Davison, R. A. Dixon, P. Gilna and M. Keller, *Science*, 2014, **344**, 1246843.

3 M. P. Pandey and C. S. Kim, *Chemical Engineering & Technology*, 2011, **34**, 29-41.

4 Y. Gao, M. Beganovic and M. B. Foston, in *Valorization of Lignocellulosic Biomass in a Biorefinery: From Logistics to Environmental and Performance Impact*, eds. R. Kumar, S. Singh and V. Balan, Nova Science Publishers, Inc., USA, 2016, ch. 8, pp. pp. 245-292.

5 W. O. Doherty, P. Mousavioun and C. M. Fellows, *Industrial Crops and Products*, 2011, **33**, 259-276.

6 C. A. Gasser, G. Hommes, A. Schäffer and P. F.-X. Corvini, *Applied Microbiology and Biotechnology*, 2012, **95**, 1115-1134.

7 G. W. Huber, S. Iborra and A. Corma, *Chemical Reviews*, 2006, **106**, 4044-4098.

8 A. G. Vishtal and A. Kraslawski, *BioResources*, 2011, **6**, 3547-3568.

9 W. Boerjan, J. Ralph and M. Baucher, *Annual review of plant biology*, 2003, **54**, 519-546.

10 M. Fache, B. Boutevin and S. Caillol, *ACS Sustainable Chemistry & Engineering*, 2016, **4**, 35-46.

11 C. Li, X. Zhao, A. Wang, G. W. Huber and T. Zhang, *Chemical Reviews*, 2015, **115**, 11559-11624.

12 R. Rinaldi, R. Jastrzebski, M. T. Clough, J. Ralph, M. Kennema, P. C. Bruijnincx and B. M. Weckhuysen, *Angewandte Chemie International Edition*, 2016, **55**, 8164-8215.

13 R. B. Santos, P. Hart, H. Jameel and H.-m. Chang, *Wood Based Lignin Reactions Important to the Biorefinery and Pulp and Paper Industries*, 2013.

14 M. Zaheer and R. Kempe, *ACS Catalysis*, 2015, **5**, 1675-1684.

15 P. J. Deuss and K. Barta, *Coordination Chemistry Reviews*, 2016, **306**, 510-532.

16 J. Zhang, H. Asakura, J. van Rijn, J. Yang, P. Duchesne, B. Zhang, X. Chen, P. Zhang, M. Saeys and N. Yan, *Green Chemistry*, 2014, **16**, 2432-2437.

17 P. J. Deuss, M. Scott, F. Tran, N. J. Westwood, J. G. de Vries and K. Barta, *Journal of the American Chemical Society*, 2015, **137**, 7456-7467.

18 L.-P. Xiao, S. Wang, H. Li, Z. Li, Z.-J. Shi, L. Xiao, R.-C. Sun, Y. Fang and G. Song, *ACS Catalysis*, 2017, **7**, 7535-7542.

19 J. Zhang, J. Teo, X. Chen, H. Asakura, T. Tanaka, K. Teramura and N. Yan, *ACS Catalysis*, 2014, **4**, 1574-1583.

20 Q. Song, F. Wang, J. Cai, Y. Wang, J. Zhang, W. Yu and J. Xu, *Energy & Environmental Science*, 2013, **6**, 994-1007.

21 Y. Ye, Y. Zhang, J. Fan and J. Chang, *Bioresource Technology*, 2012, **118**, 648-651.

22 K. Barta, T. D. Matson, M. L. Fettig, S. L. Scott, A. V. Iretskii and P. C. Ford, *Green Chemistry*, 2010, **12**, 1640-1647.

23 T. D. Matson, K. Barta, A. V. Iretskii and P. C. Ford, *Journal of the American Chemical Society*, 2011, **133**, 14090-14097.

24 C. M. Bernt, G. Bottari, J. A. Barrett, S. L. Scott, K. Barta and P. C. Ford, *Catalysis Science & Technology*, 2016, **6**, 2984-2994.

25 C. M. Bernt, H. Manesewan, M. Chui, M. Boscolo and P. C. Ford, *ACS Sustainable Chemistry & Engineering*, 2018, **6**, 2510-2516.

26 J. A. Barrett, Y. Gao, C. M. Bernt, M. Chui, A. T. Tran, M. B. Foston and P. C. Ford, *ACS Sustainable Chemistry & Engineering*, 2016, **4**, 6877-6886.

27 M. Chui, G. Metzker, C. M. Bernt, A. T. Tran, A. C. B. Burtoloso and P. C. Ford, *ACS Sustainable Chemistry & Engineering*, 2017, **5**, 3158-3169.

28 P. Chen, Q. Zhang, R. Shu, Y. Xu, L. Ma and T. Wang, *Bioresource technology*, 2017, **226**, 125-131.

29 A. D. Brittain, N. J. Chrisandina, R. E. Cooper, M. Buchanan, J. R. Cort, M. V. Olarte and C. Sievers, *Catalysis Today*, 2018, **302**, 180-189.

30 R. Shu, Y. Xu, L. Ma, Q. Zhang, C. Wang and Y. Chen, *Chemical Engineering Journal*, 2018.

31 J. Long, Y. Xu, T. Wang, Z. Yuan, R. Shu, Q. Zhang and L. Ma, *Applied Energy*, 2015, **141**, 70-79.

32 S. Van den Bosch, W. Schutyser, S.-F. Koelewijn, T. Renders, C. Courtin and B. Sels, *Chemical Communications*, 2015, **51**, 13158-13161.

33 W. Wanmolee, N. Laosiripojana, P. Daorattanachai, L. Moghaddam, J. Rencoret, J. C. del Río and W. O. Doherty, *ACS Sustainable Chemistry & Engineering*, 2018, **6**, 3010-3018.

34 H. Luo and M. M. Abu-Omar, *Green Chemistry*, 2018.

35 I. Klein, B. Saha and M. M. Abu-Omar, *Catalysis Science & Technology*, 2015, **5**, 3242-3245.

36 T. Ouyang, L. Wang, F. Cheng, Y. Hu and X. Zhao, *BioResources*, 2018, **13**, 3880-3891.

37 I. Kumaniaev, E. Subbotina, J. Sävmarker, M. Larhed, M. V. Galkin and J. S. Samec, *Green Chemistry*, 2017, **19**, 5767-5771.

38 J. Long, R. Shu, Z. Yuan, T. Wang, Y. Xu, X. Zhang, Q. Zhang and L. Ma, *Applied Energy*, 2015, **157**, 540-545.

39 P. Paatero and P. K. Hopke, *Analytica Chimica Acta*, 2003, **490**, 277-289.

40 I. M. Ulbrich, M. R. Canagaratna, Q. Zhang, D. R. Worsnop and J. L. Jimenez, *Atmos. Chem. Phys.*, 2009, **9**, 2891-2918.

41 Q. Zhang, J. L. Jimenez, M. R. Canagaratna, I. M. Ulbrich, N. L. Ng, D. R. Worsnop and Y. Sun, *Analytical and Bioanalytical Chemistry*, 2011, **401**, 3045-3067.
42 Y. Zhang, B. J. Williams, A. H. Goldstein, K. Docherty, I. M. Ulbrich and J. L. Jimenez, *Aerosol Science and Technology*, 2014, **48**, 1166-1182.
43 Y. Zhang, B. J. Williams, A. H. Goldstein, K. S. Docherty and J. L. Jimenez, *Atmos. Meas. Tech.*, 2016, **9**, 5637-5653.
44 P. Paatero and U. Tapper, *Environmetrics*, 1994, **5**, 111-126.
45 G. S. Frysinger, R. B. Gaines, L. Xu and C. M. Reddy, *Environmental Science & Technology*, 2003, **37**, 1653-1662.
46 C. Lindfors, E. Kuoppala, A. Oasmaa, Y. Solantausta and V. Arpiainen, *Energy & Fuels*, 2014, **28**, 5785-5791.
47 Q. Bu, H. Lei, A. H. Zacher, L. Wang, S. Ren, J. Liang, Y. Wei, Y. Liu, J. Tang, Q. Zhang and R. Ruan, *Bioresource Technology*, 2012, **124**, 470-477.
48 G. S. Macala, T. D. Matson, C. L. Johnson, R. S. Lewis, A. V. Iretskii and P. C. Ford, *ChemSusChem*, 2009, **2**, 215-217.
49 T. S. Hansen, K. Barta, P. T. Anastas, P. C. Ford and A. Riisager, *Green Chemistry*, 2012, **14**, 2457-2461.
50 B. J. Williams, A. H. Goldstein, N. M. Kreisberg, S. V. Hering, D. R. Worsnop, I. M. Ulbrich, K. S. Docherty and J. L. Jimenez, *Atmos. Chem. Phys.*, 2010, **10**, 11577-11603.
51 G. Isaacman-VanWertz, D. T. Sueper, K. C. Aikin, B. M. Lerner, J. B. Gilman, J. A. de Gouw, D. R. Worsnop and A. H. Goldstein, *Journal of Chromatography A*, 2017, **1529**, 81-92.

Supplementary Information (SI) for

Analysis of gas chromatography/mass spectrometry data for catalytic lignin depolymerization using positive matrix factorization

Yu Gao,[‡a] Michael J. Walker,[‡a] Jacob A. Barrett,[b] Omid Hosseinaei,[c] David P. Harper,[d] Peter C. Ford,[b] Brent J. Williams[a] and Marcus B. Foston*[a]

[a]Department of Energy, Environmental and Chemical Engineering, Washington University in St. Louis, St. Louis MO, 63130, USA.
*Email: mfoston@wustl.edu

[b]Department of Chemistry and Biochemistry, University of California, Santa Barbara, Santa Barbara CA, 93106-9510, USA.

[c]RISE Bioeconomy, PO Box 5604, 11486 Stockholm, Sweden

[d]Center for Renewable Carbon, University of Tennessee, Knoxville, Knoxville TN, 37996, USA

Table of Contents

Experimental Procedures:

Figures:

Tables:

Experimental Procedures.

Lignin extraction. Pulp grade chips (about 4 cm^2 and thicknesses of 0.5–1 cm) of hybrid poplar (*Populus spp.*) were used for organosolv fractionation. The biomass sample was treated in a flow-through reactor with a 16:34:50 wt% mixture of methyl isobutyl ketone (MIBK), ethanol, and water. Sulfuric acid (0.05 M) was used as catalyst and fractionation performed at a temperature of 150 °C for 120 min. The black liquor fraction containing dissolved lignin and hemicellulose was separated into a hemicellulose rich aqueous phase and a lignin rich organic phase by adding solid NaCl (10 g per 100 mL of deionized water in the initial solvent mixture) in a separation funnel. After observation of phase separation between the aqueous and organic phases, the aqueous phase that rests at the bottom of the separatory funnel was drained. The organic phase was washed twice by adding 30% v/v deionized water to remove residual sugars and ethanol from the organic phase. Lignin from organic phase phases was isolated by rotary evaporation, followed by trituration of the solid residue with diethyl ether (5 times ~200 mL) and final washing with deionized water (3 L deionized water, room temperature, 12 h). The final lignin was filtered through a paper filter and dried in a vacuum oven at temperature of 80°C for 12 h. Organosolv lignin samples were extracted with methanol at room temperature and solid to liquid ratio of 1:10. Extraction was for about 15 h and then the solution was filtered under vacuum. Methanol insoluble fraction recovered as solid on the filter paper and methanol soluble fraction isolated by rotary evaporation of the filtrates. All lignin samples were dried in a vacuum oven at temperature of 80 °C for 12 h.

Catalyst synthesis. A solution of 250 mL deionized water containing $Mg(NO_3)_2 \cdot 6H_2O$ (30.8 g, 0.12 mol), $Cu(NO_3)_2 \cdot 3H_2O$ (7.25 g, 0.03 mol), and $Al(NO_3)_3 \cdot 9H_2O$ (18.76 g, 0.05 mol), was slowly added to a Na_2CO_3 buffer (5.3 g, 0.05 mol in 375 mL) at 65 °C with vigorous stirring. The pH of the mixture was maintained at approximately 10 by alternating aliquots of 1 M NaOH to the reaction mixture. After the addition of the metal solution was complete, the reaction slurry was stirred overnight. The light blue precipitate was isolated by filtration and washed with a sodium carbonate solution (0.05 mol in 1 L distilled water) for a minimum of four hours, then filtered and washed with deionized water. The precipitate was dried overnight at 110 °C resulting in Cu_{20}HTC.[4]

GC-TCD analysis. To quantify the gas contents, 100 µL of raw gas products was manually injected into the gas chromatography system (GC, 7890B, Agilent Technologies) with thermal conductivity detector (TCD). Inlet temperature was set to 250 °C. Supelco Carboxen-1010 PLOT column (ID: 0.32 mm, average thickness: 15 µm, and length: 30 m) was used with an isotherm method at 75 °C for 10 mins. Helium was used as a carrier gas. Gas products were identified and quantified by the standard gas mixture comprising CO, CO_2, H_2, N_2, and O_2 in helium (custom-mixed by scott specialty gases, Plumsteadville, PA).

GPC analysis. GPC analysis was performed to determine the molecular weight distribution of the liquid products. In a Waters e2695 system with a 2489 UV detector (260 nm), a three-column sequence of WatersTM Styragel columns (HR0.5, HR1, and HR3) was used for the analysis. Tetrahydrofuran (THF) was used as eluent, and the flow rate was 1.0 ml/min. 1 mL of raw liquid product was first filtered to a 2 mL HPLC vial through a 0.45-µm nylon membrane filter, and 50 µL of this sample was injected into the instrument. Molecular weights (M_n and M_w) were calibrated against a polystyrene calibration curve. A calibration curve was constructed by fitting a third-order polynomial equation to the retention volumes obtained from six narrow polystyrene standards and two small molecules (diphenylmethane and toluene) ranging in molecular weight from 92 to 3.4 × 104 g/mol. The curve fit had an R^2 value of 0.99.

Thermal gravimetric analysis. Gravimetric analysis was conducted on the solids residues after dioxane washes. The thermogravimetric analyses were carried out in a Q5000 TGA instrument

(TA instrument). ~5-10 mg dry solid samples were placed onto a platinum TGA pan. Furnace was programed to heat to 900 °C in 2 mins with air (ultrazero grade) flowrate of 10 L/h and nitrogen flowrate of 25 L/h. Furnace was held at 900 °C for additional 10 mins until no further changes in sample weight observed. Weight loss percentages were recorded to calculate the catalyst content in the solid residues.

Quantitative ^{31}P NMR. ^{31}P NMR was performed after derivatization of the untreated MS and MIS lignin with 2–chloro– 4,4,5,5–tetramethyl–1,3,2–dioxaphospholane (TMDP). N-Hydroxy-5-norbornene-2,3-dicarboxylic acid imide and chromium (III) acetylacetonate were used as an internal standard and a relaxation agent, respectively. The quantitative ^{31}P NMR spectra were recorded using a Varian 400-NMR spectrometer at frequency of 162 MHz using a 90° pulse angle, 25 s pulse delay, and 256 transients at room temperature.

Quantitative ^{13}C NMR. For quantitative ^{13}C NMR spectroscopy, 150 mg of lignin was dissolved in 0.75 mL of DMSO-d_6. Chromium (III) acetylacetonate were used as relaxation agent (0.01 M). An inverse-gated decoupling pulse sequence was used with a 90° pulse angle, 1.7 s relaxation delay and an acquisition time of 1.40 s. A total of 20,000 scans were recorded.

^{13}C–^{1}H (HSQC) NMR. HSQC NMR was carried out using a Varian 400-MR spectrometer operating at frequency of 399.78 MHz for proton and 100.54 MHz for carbon. For 2D (HSQC) spectroscopy, 100 mg of lignin were dissolved in 0.75 mL of DMSO-d6. NMR spectra were recorded at 25°C using the (HC)bsgHSQCAD pulse program. The experiment used 32 transients and 512 time increments in the ^{13}C dimension. A 90° pulse with a pulse delay of 1.5 s, an acquisition time of 0.15 s and a $^{1}J_{CH}$ of 147 Hz were employed. DMSO was used as an internal reference.

Nitric acid digestion. Solid residues from depolymerization were treated with 5 ml of 70% HNO_3 at room temperature for 16 h. Mixture was heated to 50 °C for another hour, then, cooled to room temperature. Leftover solids (Char) were recovered by filtration through glass fiber filters. Leftover solids were washed with excessive amount of water DI water for three time and then dried for gravimetric analysis.

Figures:

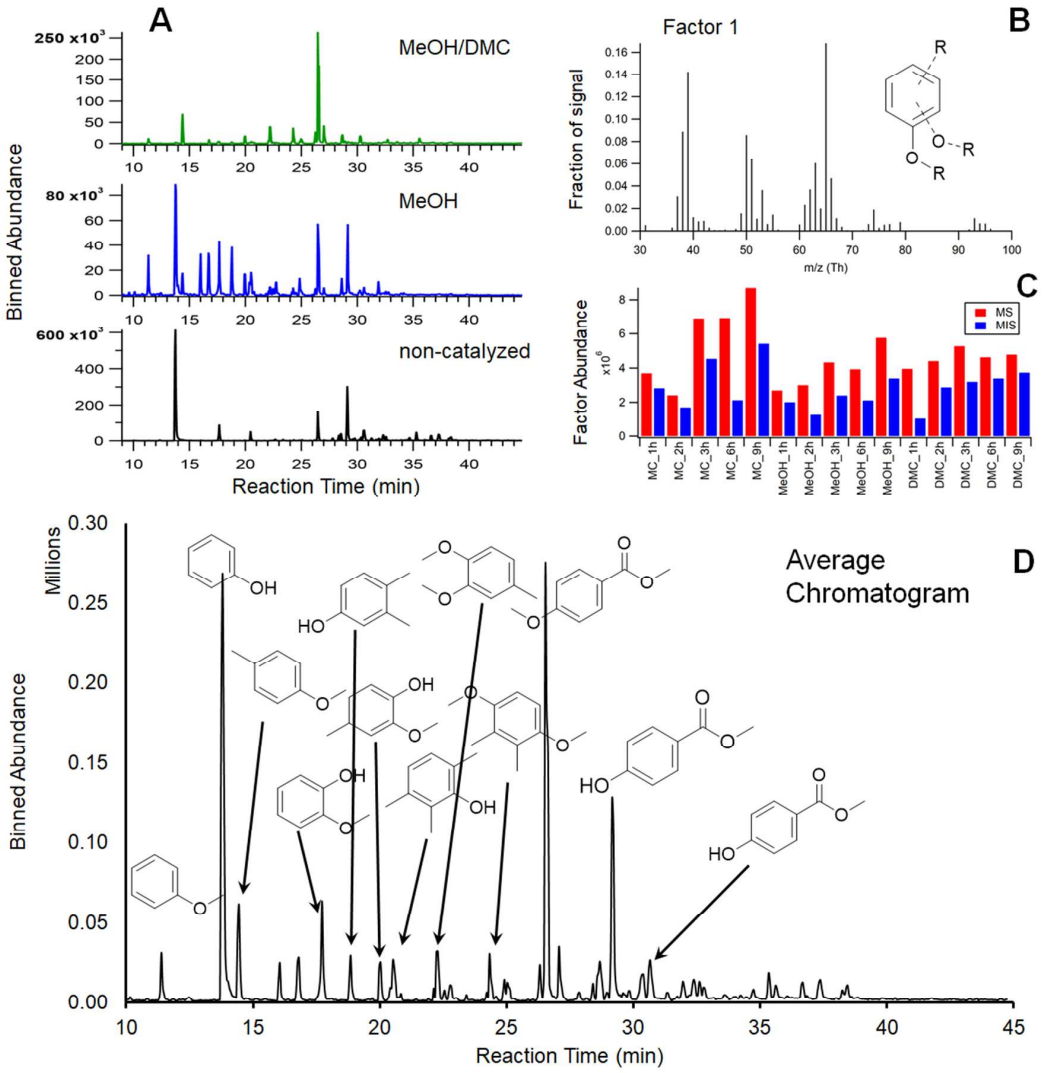

Figure S1. Factor 1 is defined by compounds that generate mass spectral fragments that are less polar and/or more volatile aromatics (13-factor solution). A) PMF-reconstructed Factor 1 average chromatograms for product samples from catalyzed reactions in MeOH/DMC (green), catalyzed reactions in MeOH (blue), and non-catalyzed reactions in MeOH (black). B) The Factor 1 mass spectrum. C) Factor 1 abundance in the product samples generated from non-catalyzed reactions in MeOH (MC), catalyzed reactions in MeOH (MeOH), and catalyzed reactions in MeOH/DMC (DMC) each undergoing reaction for 1, 2, 3, 6, and 9 h. The Factor 1 abundance is shown for the product generated from methanol-soluble (MS; red) and methanol-insoluble (MIS; blue) fraction of lignin extracted from a hybrid poplar biomass source. D) PMF-reconstructed Factor 1 average chromatograms for product samples from the combination of all reaction conditions. Individual compound structures identified in (D) were verified by Palisade Complete Mass Spectral Database mostly with more than 90 % matching.

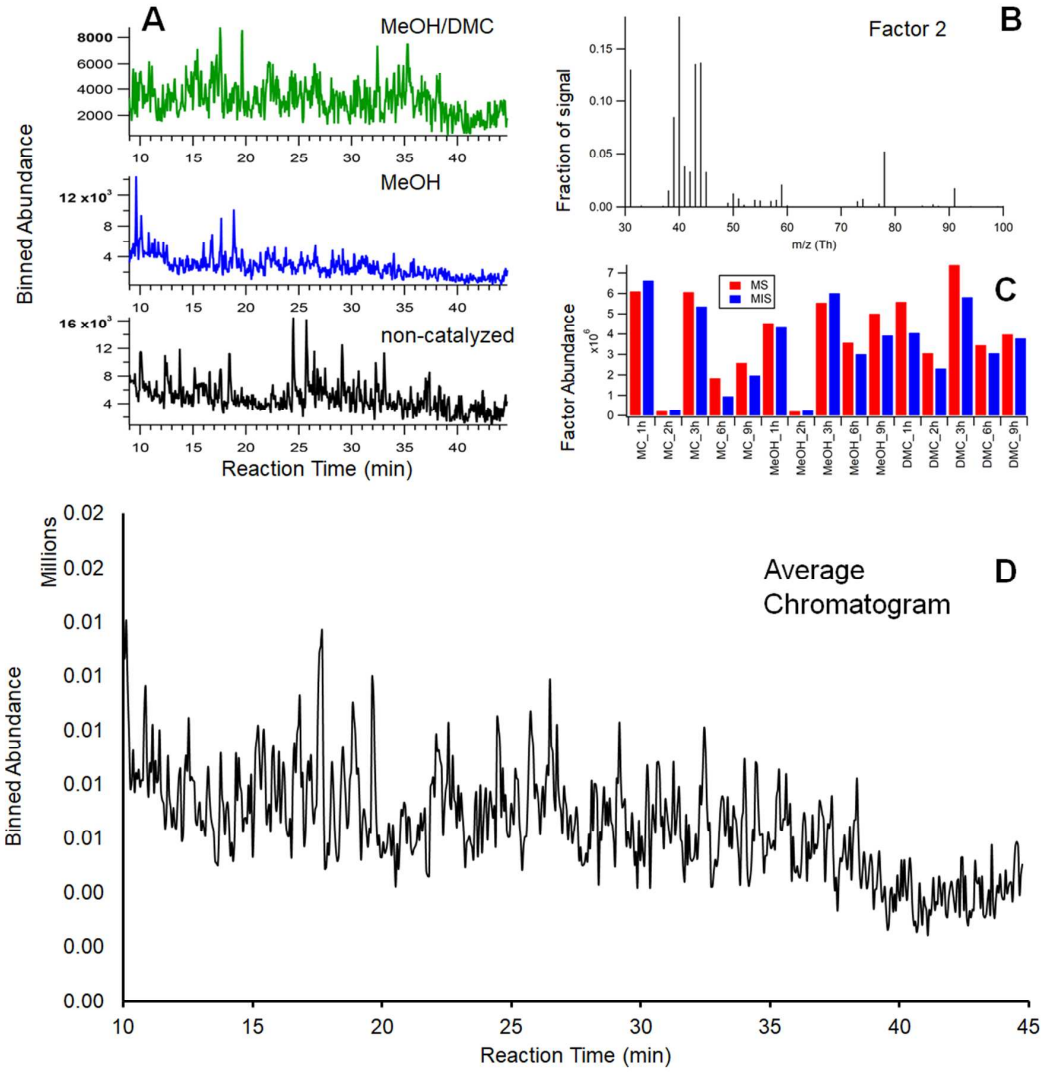

Figure S2. Factor 2 is defined by compounds that generate mass spectral fragments that are air and light molecular weight contaminates (13-factor solution). A) PMF-reconstructed Factor 2 average chromatograms for product samples from catalyzed reactions in MeOH/DMC (green), catalyzed reactions in MeOH (blue), and non-catalyzed reactions in MeOH (black). B) The Factor 2 mass spectrum. C) Factor 2 abundance in the product samples generated from non-catalyzed reactions in MeOH (MC), catalyzed reactions in MeOH (MeOH), and catalyzed reactions in MeOH/DMC (DMC) each undergoing reaction for 1, 2, 3, 6, and 9 h. The Factor 2 abundance is shown for the product generated from methanol-soluble (MS; red) and methanol-insoluble (MIS; blue) fraction of lignin extracted from a hybrid poplar biomass source. D) PMF-reconstructed Factor 2 average chromatograms for product samples from the combination of all reaction conditions.

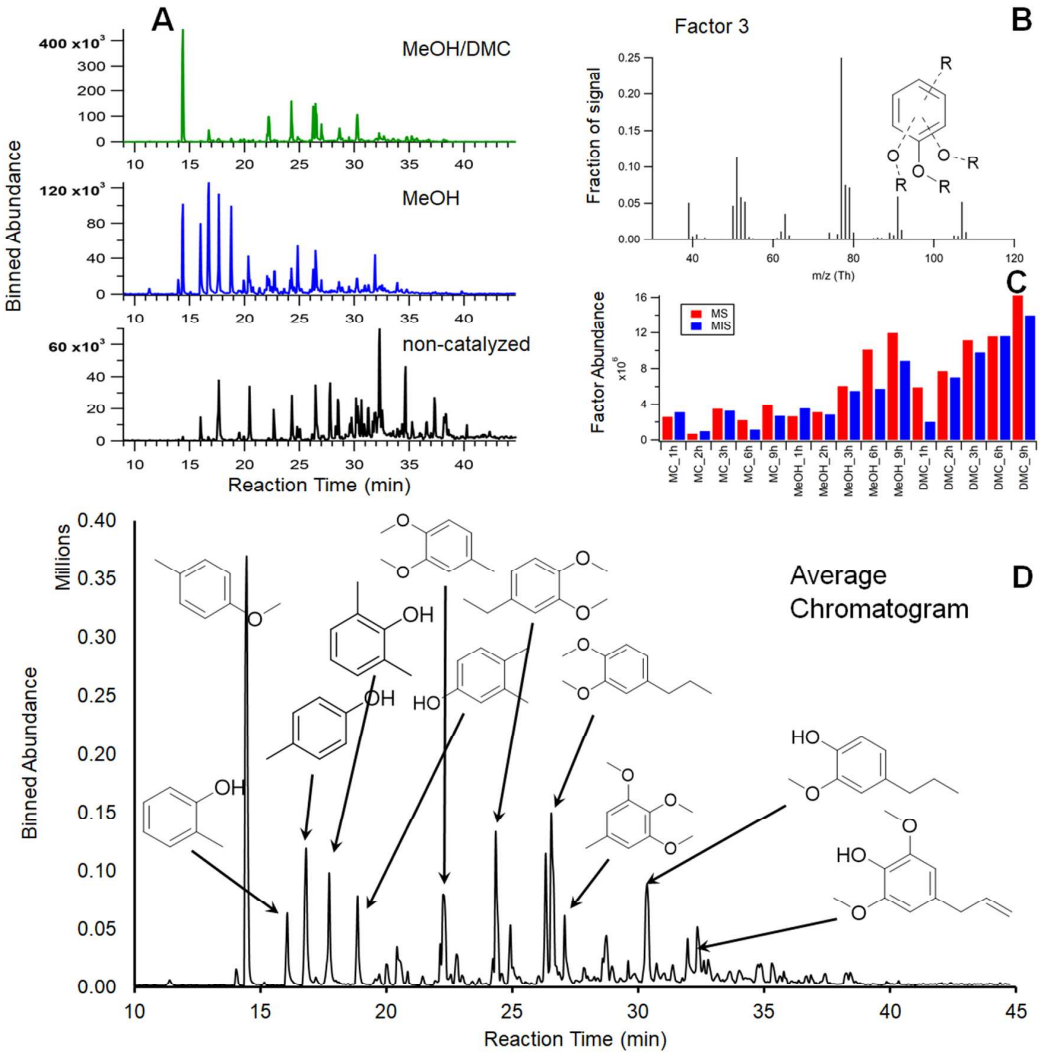

Figure S3. Factor 3 is defined by compounds that generate mass spectral fragments that are less polar and/or more volatile aromatics (13-factor solution). A) PMF-reconstructed Factor 3 average chromatograms for product samples from catalyzed reactions in MeOH/DMC (green), catalyzed reactions in MeOH (blue), and non-catalyzed reactions in MeOH (black). B) The Factor 3 mass spectrum. C) Factor 3 abundance in the product samples generated from non-catalyzed reactions in MeOH (MC), catalyzed reactions in MeOH (MeOH), and catalyzed reactions in MeOH/DMC (DMC) each undergoing reaction for 1, 2, 3, 6, and 9 h. The Factor 3 abundance is shown for the product generated from methanol-soluble (MS; red) and methanol-insoluble (MIS; blue) fraction of lignin extracted from a hybrid poplar biomass source. D) PMF-reconstructed Factor 3 average chromatograms for product samples from the combination of all reaction conditions. Individual compound structures identified in (D) were verified by Palisade Complete Mass Spectral Database mostly with more than 90 % matching.

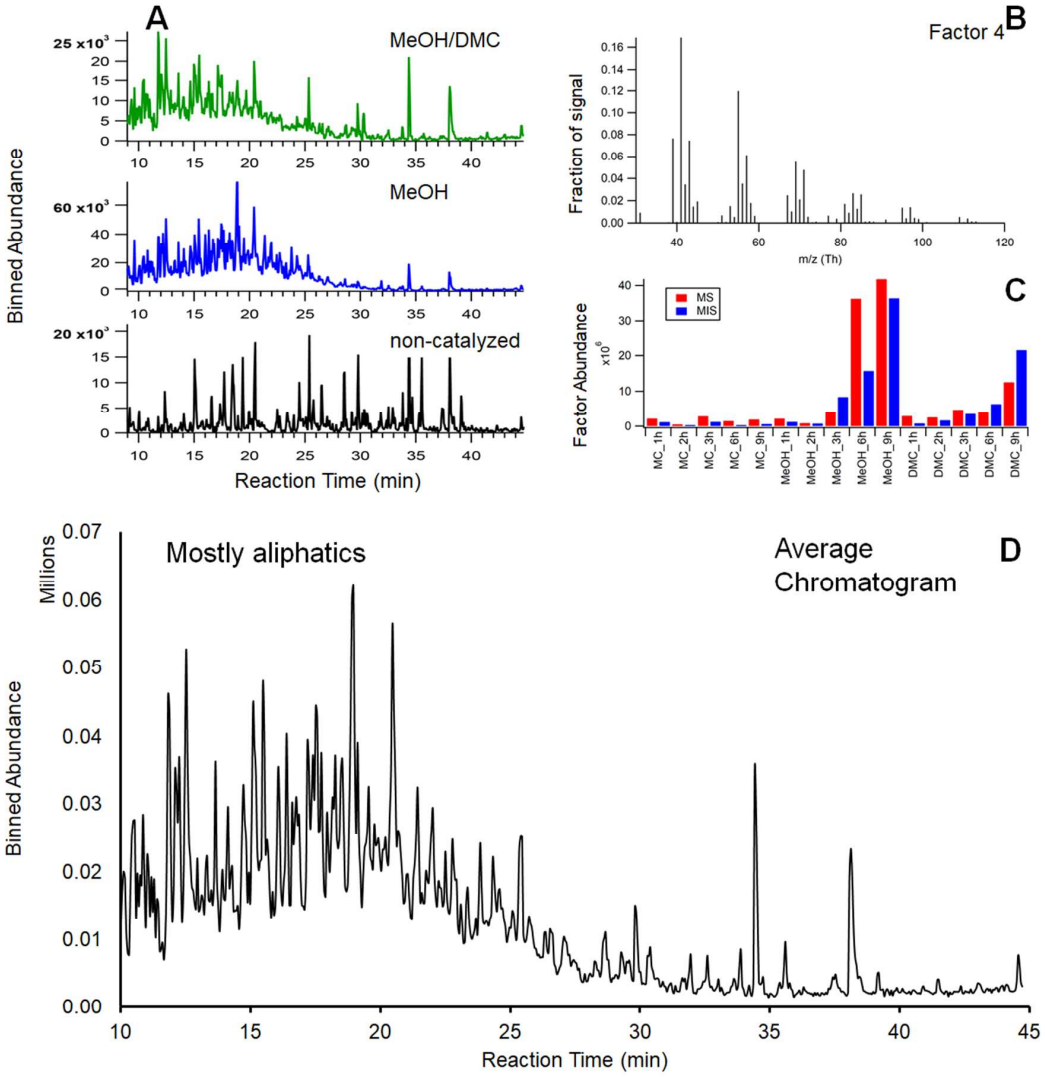

Figure S4. Factor 4 is defined by compounds that generate mass spectral fragments that are aliphatics (13-factor solution). A) PMF-reconstructed Factor 4 average chromatograms for product samples from catalyzed reactions in MeOH/DMC (green), catalyzed reactions in MeOH (blue), and non-catalyzed reactions in MeOH (black). B) The Factor 4 mass spectrum. C) Factor 4 abundance in the product samples generated from non-catalyzed reactions in MeOH (MC), catalyzed reactions in MeOH (MeOH), and catalyzed reactions in MeOH/DMC (DMC) each undergoing reaction for 1, 2, 3, 6, and 9 h. The Factor 4 abundance is shown for the product generated from methanol-soluble (MS; red) and methanol-insoluble (MIS; blue) fraction of lignin extracted from a hybrid poplar biomass source. D) PMF-reconstructed Factor 4 average chromatograms for product samples from the combination of all reaction conditions. Individual compound structures identified in (D) were verified by Palisade Complete Mass Spectral Database with average 56 % matching.

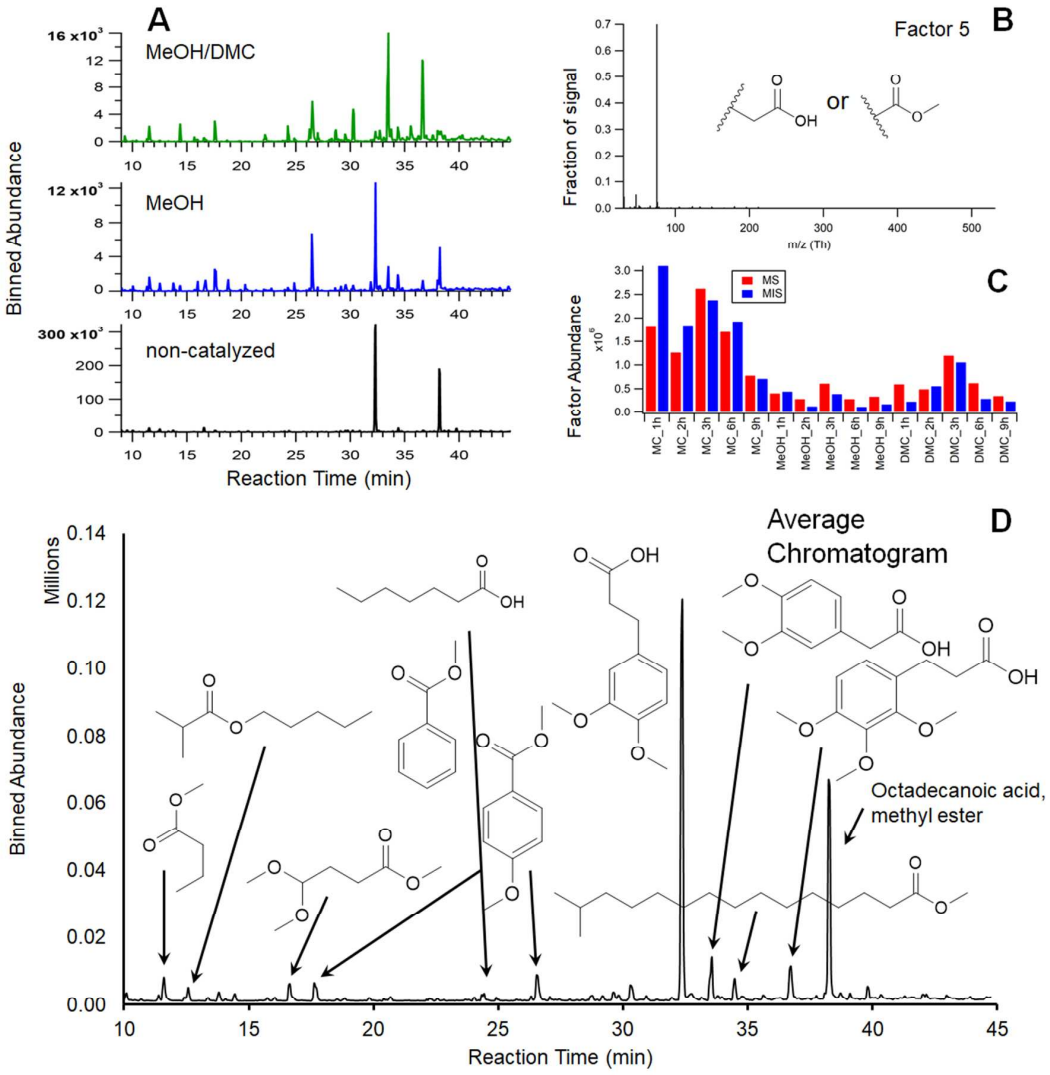

Figure S5. Factor 5 is defined by compounds that generate mass spectral fragments that are carboxylates (13-factor solution). A) PMF-reconstructed Factor 5 average chromatograms for product samples from catalyzed reactions in MeOH/DMC (green), catalyzed reactions in MeOH (blue), and non-catalyzed reactions in MeOH (black). B) The Factor 5 mass spectrum. C) Factor 5 abundance in the product samples generated from non-catalyzed reactions in MeOH (MC), catalyzed reactions in MeOH (MeOH), and catalyzed reactions in MeOH/DMC (DMC) each undergoing reaction for 1, 2, 3, 6, and 9 h. The Factor 5 abundance is shown for the product generated from methanol-soluble (MS; red) and methanol-insoluble (MIS; blue) fraction of lignin extracted from a hybrid poplar biomass source. D) PMF-reconstructed Factor 5 average chromatograms for product samples from the combination of all reaction conditions. Individual compound structures identified in (D) were verified by Palisade Complete Mass Spectral Database mostly with more than 90 % matching.

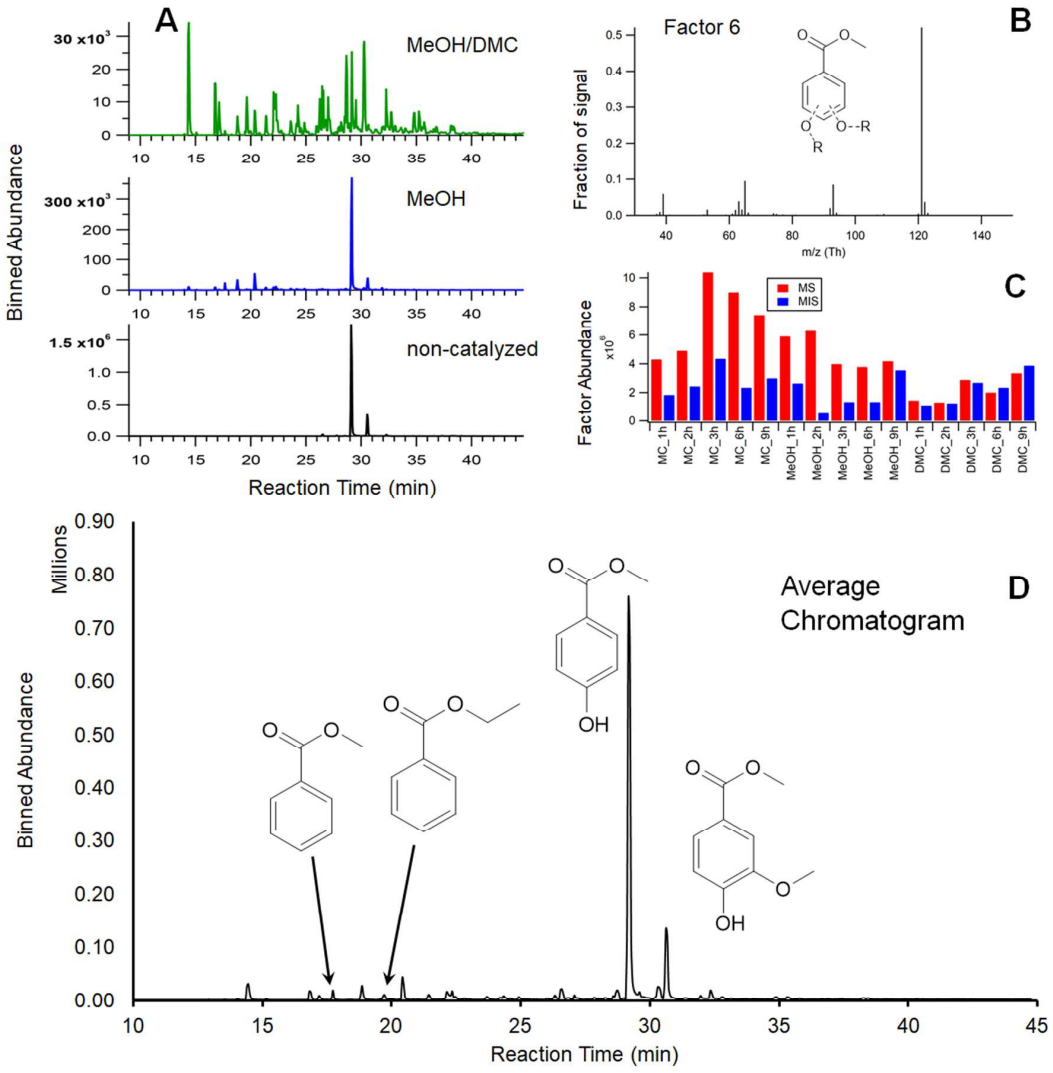

Figure S6. Factor 6 is defined by compounds that generate mass spectral fragments that are benzoates (13-factor solution). A) PMF-reconstructed Factor 6 average chromatograms for product samples from catalyzed reactions in MeOH/DMC (green), catalyzed reactions in MeOH (blue), and non-catalyzed reactions in MeOH (black). B) The Factor 6 mass spectrum. C) Factor 6 abundance in the product samples generated from non-catalyzed reactions in MeOH (MC), catalyzed reactions in MeOH (MeOH), and catalyzed reactions in MeOH/DMC (DMC) each undergoing reaction for 1, 2, 3, 6, and 9 h. The Factor 6 abundance is shown for the product generated from methanol-soluble (MS; red) and methanol-insoluble (MIS; blue) fraction of lignin extracted from a hybrid poplar biomass source. D) PMF-reconstructed Factor 6 average chromatograms for product samples from the combination of all reaction conditions. Individual compound structures identified in (D) were verified by Palisade Complete Mass Spectral Database mostly with more than 90 % matching.

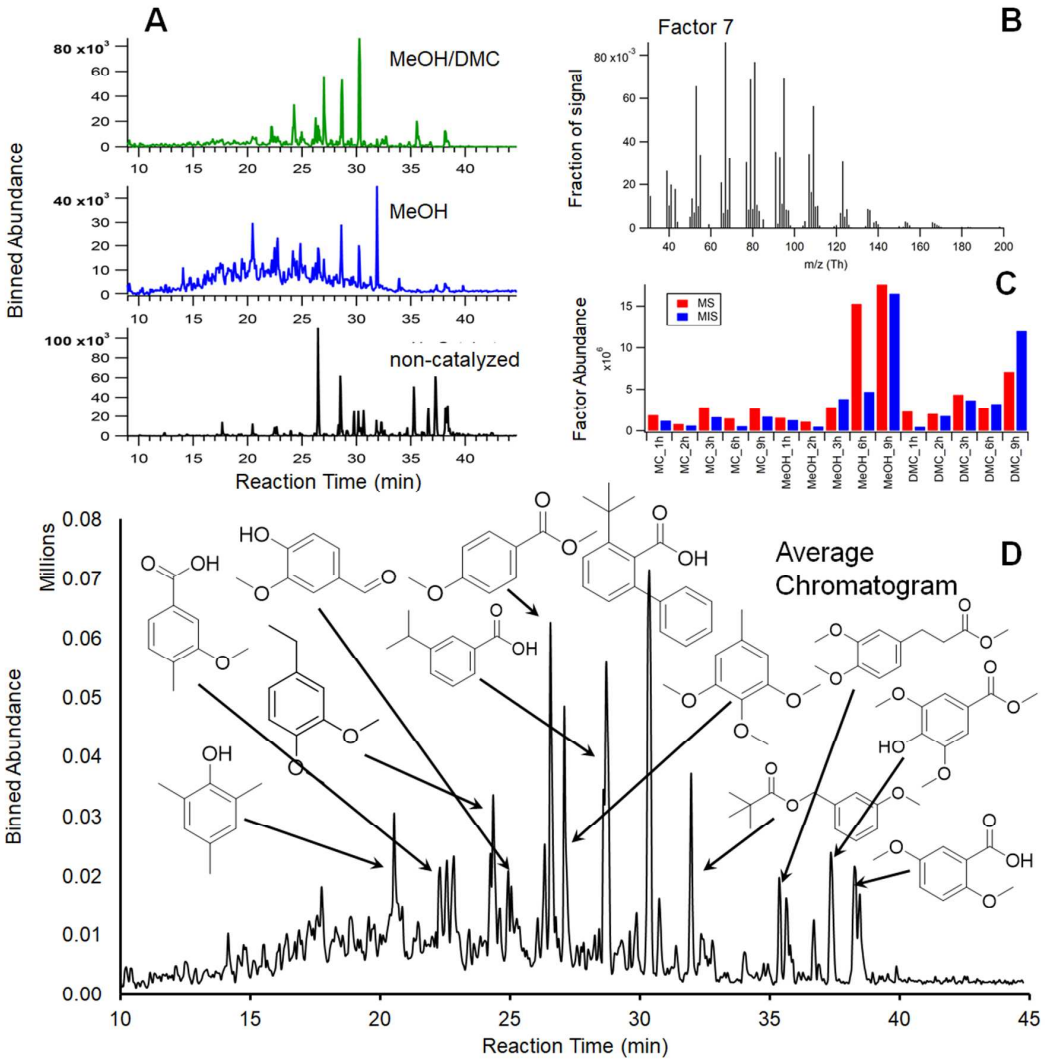

Figure S7. Factor 7 is defined by compounds that generate mass spectral fragments that are more polar and/or less volatile aromatics (13-factor solution). A) PMF-reconstructed Factor 7 average chromatograms for product samples from catalyzed reactions in MeOH/DMC (green), catalyzed reactions in MeOH (blue), and non-catalyzed reactions in MeOH (black). B) The Factor 7 mass spectrum. C) Factor 7 abundance in the product samples generated from non-catalyzed reactions in MeOH (MC), catalyzed reactions in MeOH (MeOH), and catalyzed reactions in MeOH/DMC (DMC) each undergoing reaction for 1, 2, 3, 6, and 9 h. The Factor 7 abundance is shown for the product generated from methanol-soluble (MS; red) and methanol-insoluble (MIS; blue) fraction of lignin extracted from a hybrid poplar biomass source. D) PMF-reconstructed Factor 7 average chromatograms for product samples from the combination of all reaction conditions. Individual compound structures identified in (D) were verified by Palisade Complete Mass Spectral Database mostly with more than 70 % matching.

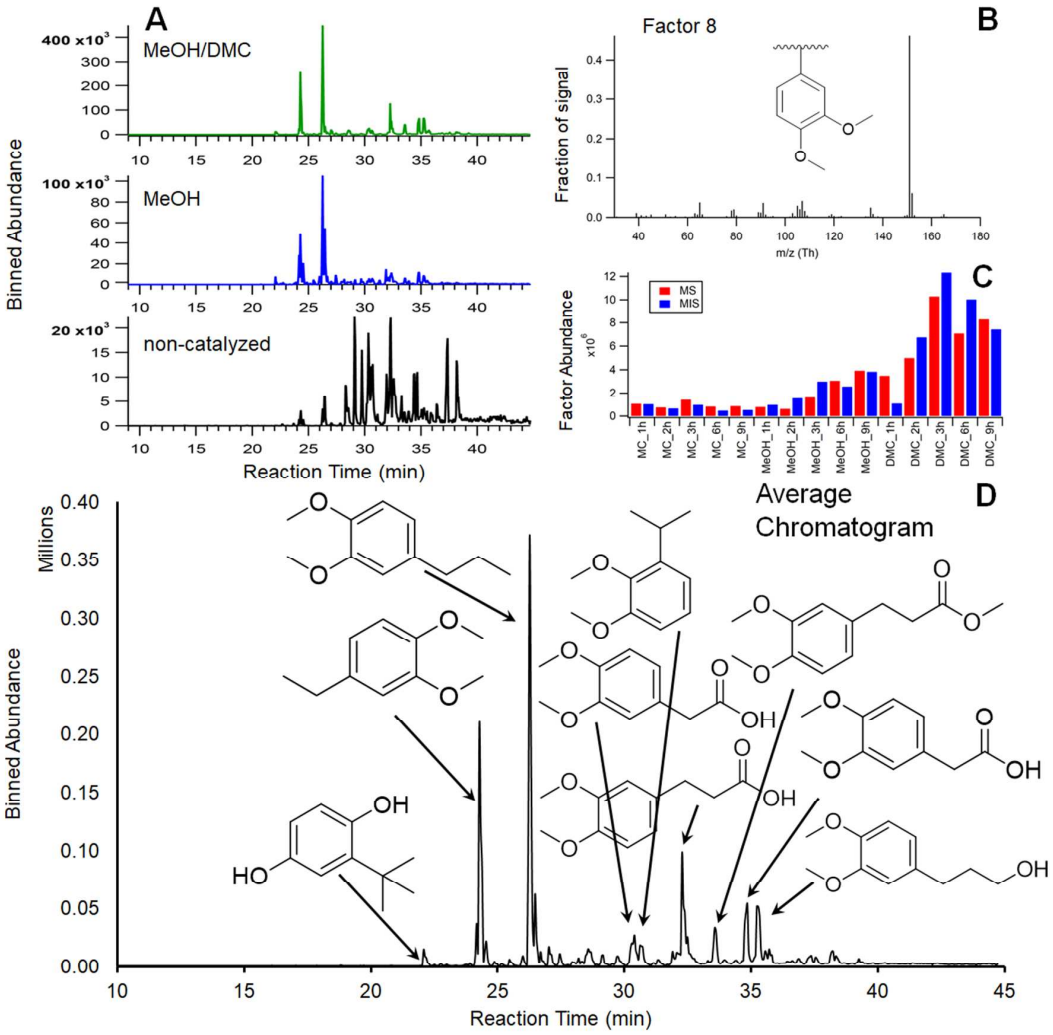

Figure S8. Factor 8 is defined by compounds that generate mass spectral fragments that are dimethoxy benzyls (13-factor solution). A) PMF-reconstructed Factor 8 average chromatograms for product samples from catalyzed reactions in MeOH/DMC (green), catalyzed reactions in MeOH (blue), and non-catalyzed reactions in MeOH (black). B) The Factor 8 mass spectrum. C) Factor 8 abundance in the product samples generated from non-catalyzed reactions in MeOH (MC), catalyzed reactions in MeOH (MeOH), and catalyzed reactions in MeOH/DMC (DMC) each undergoing reaction for 1, 2, 3, 6, and 9 h. The Factor 8 abundance is shown for the product generated from methanol-soluble (MS; red) and methanol-insoluble (MIS; blue) fraction of lignin extracted from a hybrid poplar biomass source. D) PMF-reconstructed Factor 8 average chromatograms for product samples from the combination of all reaction conditions. Individual compound structures identified in (D) were verified by Palisade Complete Mass Spectral Database mostly with more than 90 % matching.

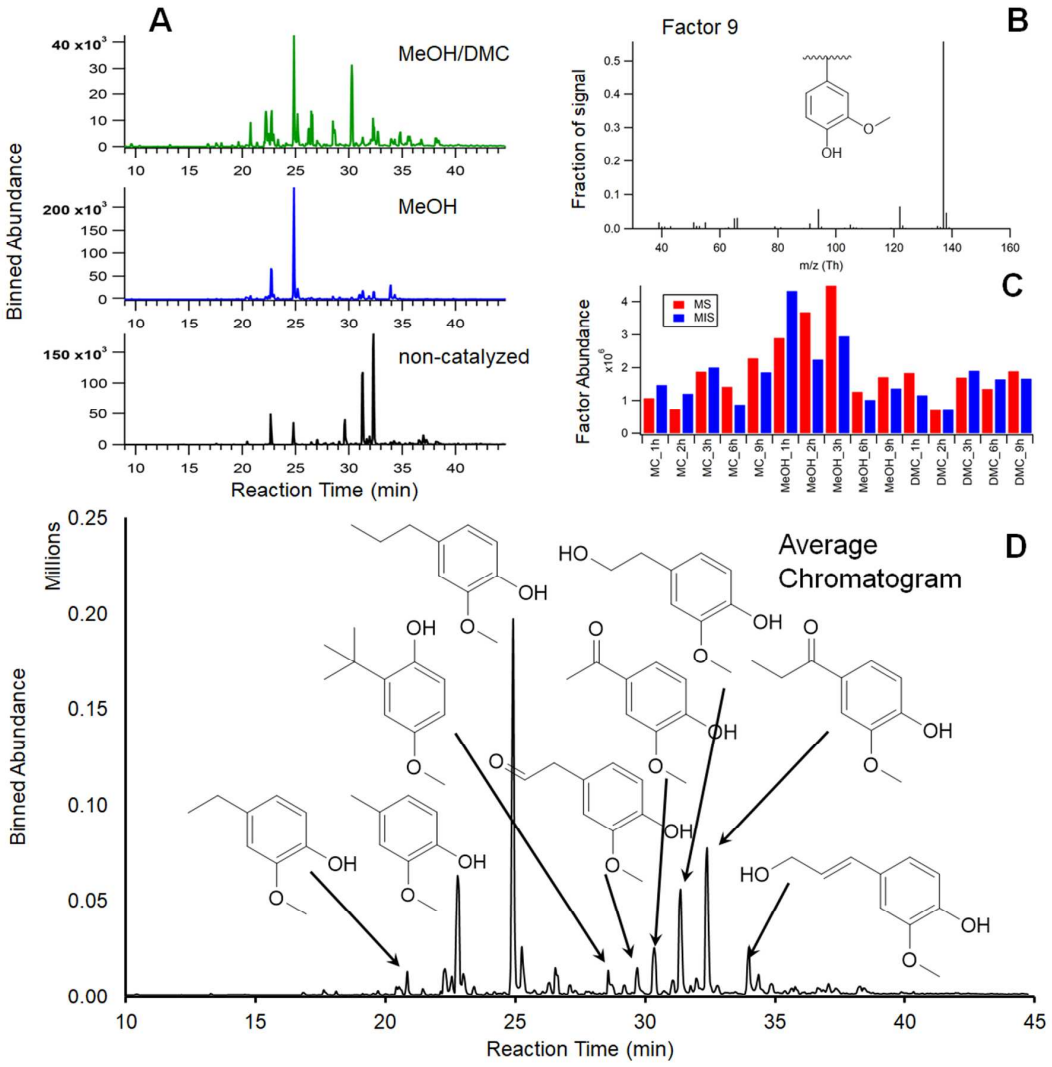

Figure S9. Factor 9 is defined by compounds that generate mass spectral fragments that are methoxy phenyls (13-factor solution). A) PMF-reconstructed Factor 9 average chromatograms for product samples from catalyzed reactions in MeOH/DMC (green), catalyzed reactions in MeOH (blue), and non-catalyzed reactions in MeOH (black). B) The Factor 9 mass spectrum. C) Factor 9 abundance in the product samples generated from non-catalyzed reactions in MeOH (MC), catalyzed reactions in MeOH (MeOH), and catalyzed reactions in MeOH/DMC (DMC) each undergoing reaction for 1, 2, 3, 6, and 9 h. The Factor 9 abundance is shown for the product generated from methanol-soluble (MS; red) and methanol-insoluble (MIS; blue) fraction of lignin extracted from a hybrid poplar biomass source. D) PMF-reconstructed Factor 9 average chromatograms for product samples from the combination of all reaction conditions. Individual compound structures identified in (D) were verified by Palisade Complete Mass Spectral Database mostly with more than 90 % matching.

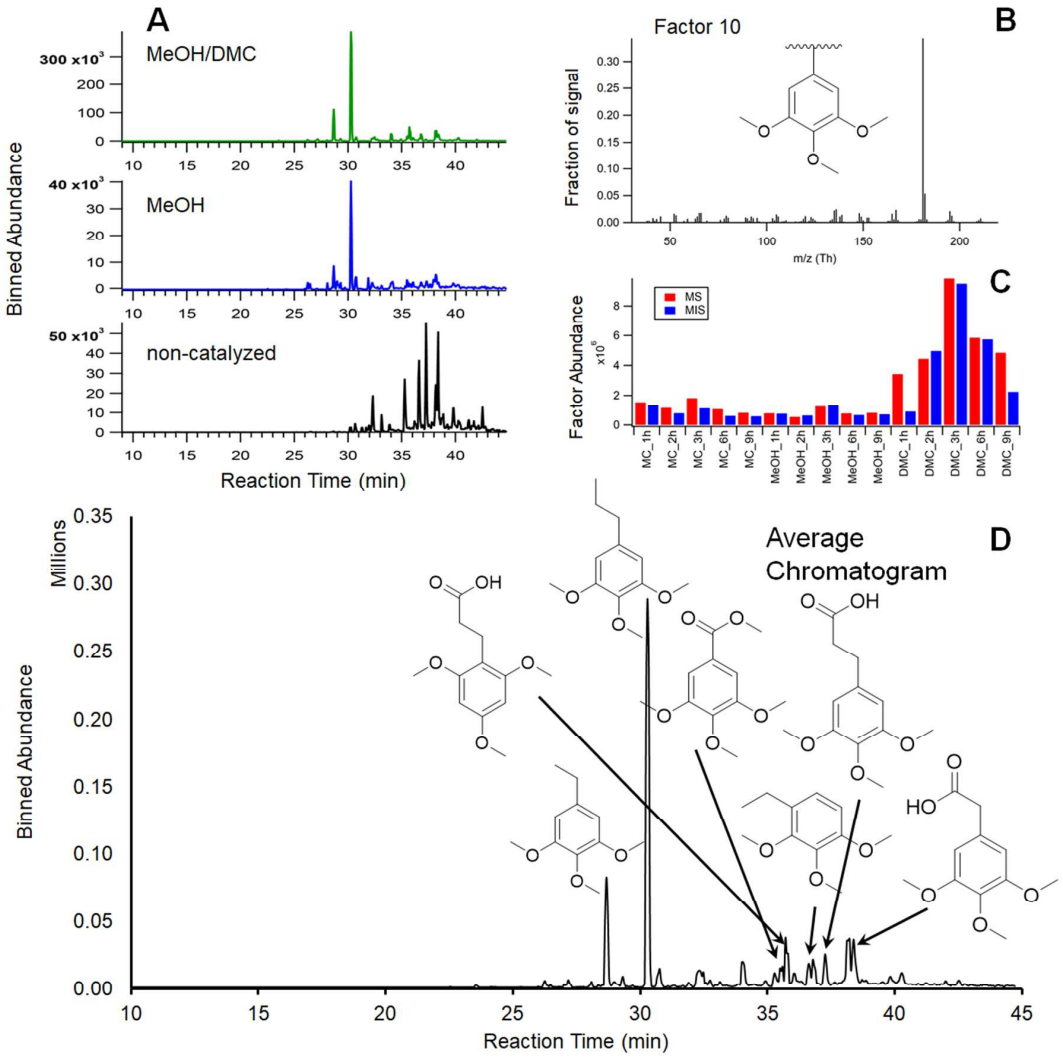

Figure S10. Factor 10 is defined by compounds that generate mass spectral fragments that are trimethoxy benzyls (13-factor solution). A) PMF-reconstructed Factor 10 average chromatograms for product samples from catalyzed reactions in MeOH/DMC (green), catalyzed reactions in MeOH (blue), and non-catalyzed reactions in MeOH (black). B) The Factor 10 mass spectrum. C) Factor 10 abundance in the product samples generated from non-catalyzed reactions in MeOH (MC), catalyzed reactions in MeOH (MeOH), and catalyzed reactions in MeOH/DMC (DMC) each undergoing reaction for 1, 2, 3, 6, and 9 h. The Factor 10 abundance is shown for the product generated from methanol-soluble (MS; red) and methanol-insoluble (MIS; blue) fraction of lignin extracted from a hybrid poplar biomass source. D) PMF-reconstructed Factor 10 average chromatograms for product samples from the combination of all reaction conditions. Individual compound structures identified in (D) were verified by Palisade Complete Mass Spectral Database mostly with more than 90 % matching.

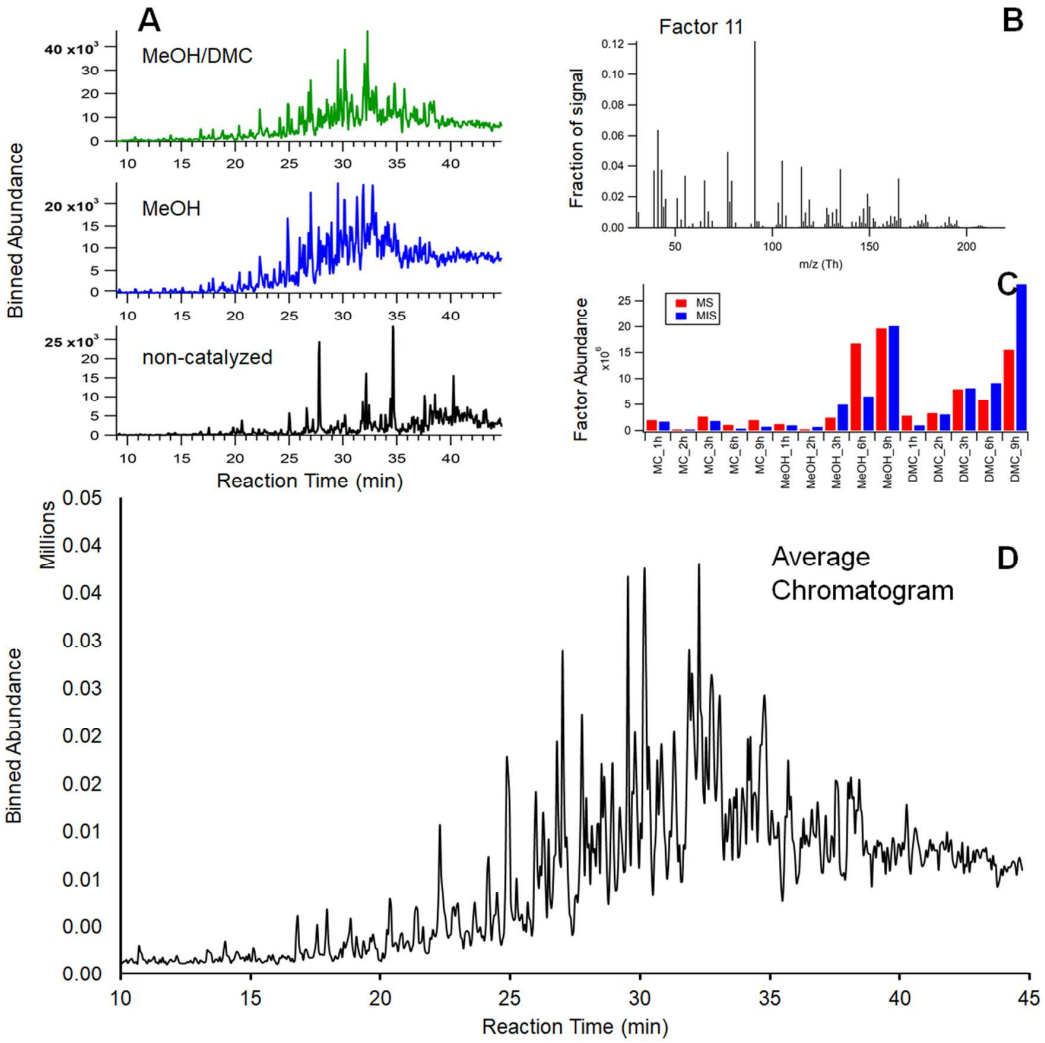

Figure S11. Factor 11 is defined by compounds that generate mass spectral fragments that are unresolved complex mixtures (13-factor solution). A) PMF-reconstructed Factor 11 average chromatograms for product samples from catalyzed reactions in MeOH/DMC (green), catalyzed reactions in MeOH (blue), and non-catalyzed reactions in MeOH (black). B) The Factor 11 mass spectrum. C) Factor 11 abundance in the product samples generated from non-catalyzed reactions in MeOH (MC), catalyzed reactions in MeOH (MeOH), and catalyzed reactions in MeOH/DMC (DMC) each undergoing reaction for 1, 2, 3, 6, and 9 h. The Factor 11 abundance is shown for the product generated from methanol-soluble (MS; red) and methanol-insoluble (MIS; blue) fraction of lignin extracted from a hybrid poplar biomass source. D) PMF-reconstructed Factor 11 average chromatograms for product samples from the combination of all reaction conditions.

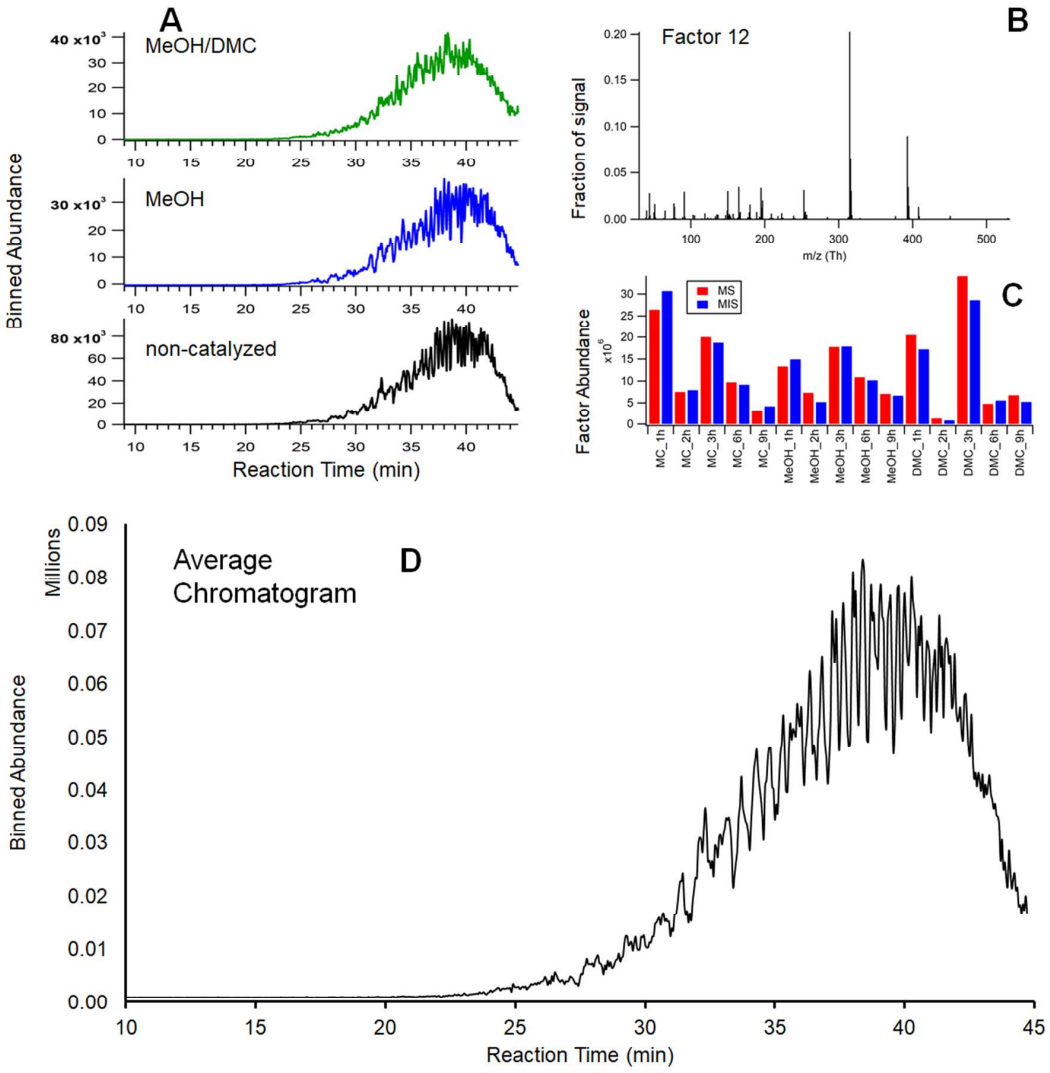

Figure S12. Factor 12 is defined by compounds that generate mass spectral fragments that are column bleed residues (13-factor solution). A) PMF-reconstructed Factor 12 average chromatograms for product samples from catalyzed reactions in MeOH/DMC (green), catalyzed reactions in MeOH (blue), and non-catalyzed reactions in MeOH (black). B) The Factor 12 mass spectrum. C) Factor 12 abundance in the product samples generated from non-catalyzed reactions in MeOH (MC), catalyzed reactions in MeOH (MeOH), and catalyzed reactions in MeOH/DMC (DMC) each undergoing reaction for 1, 2, 3, 6, and 9 h. The Factor 12 abundance is shown for the product generated from methanol-soluble (MS; red) and methanol-insoluble (MIS; blue) fraction of lignin extracted from a hybrid poplar biomass source. D) PMF-reconstructed Factor 12 average chromatograms for product samples from the combination of all reaction conditions.

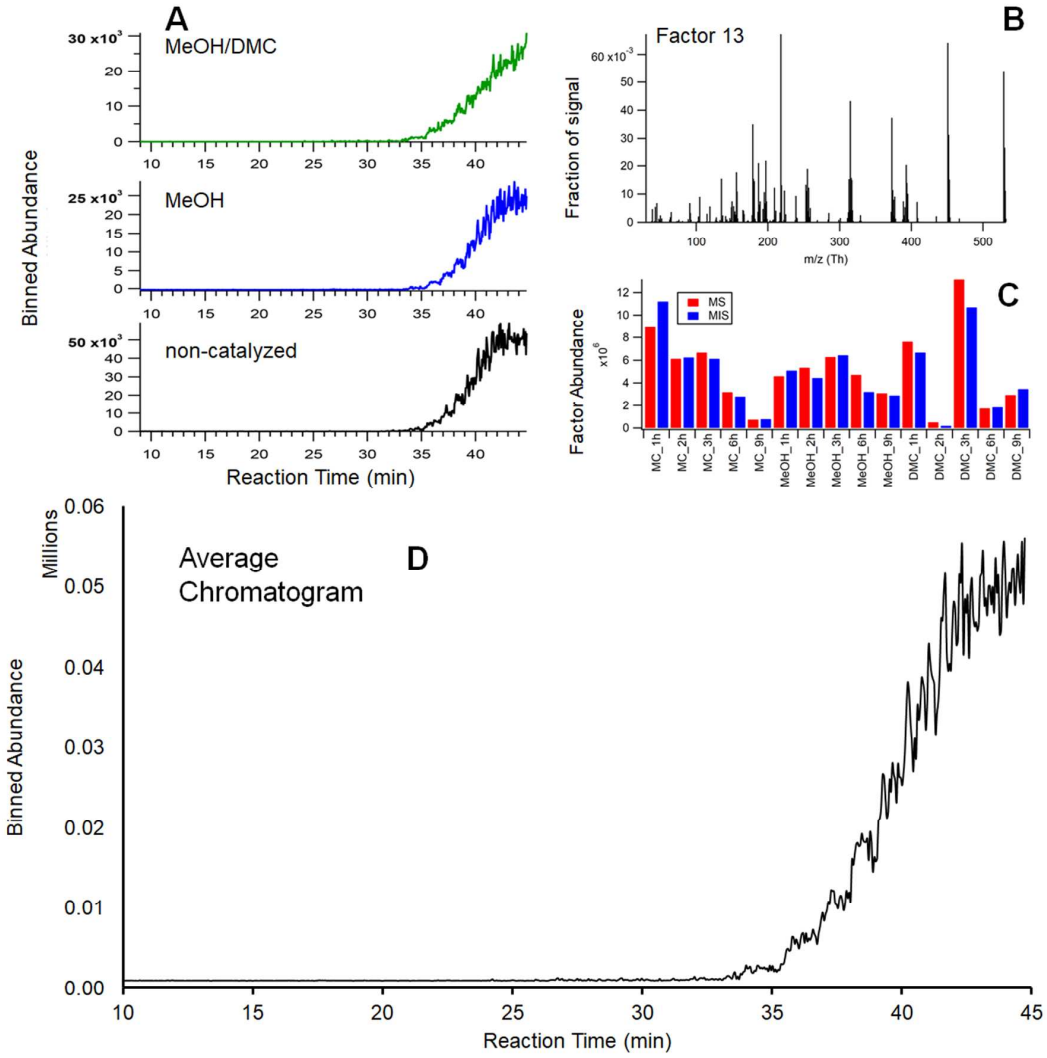

Figure S13. Factor 13 is defined by compounds that generate mass spectral fragments that are column bleed residues (13-factor solution). A) PMF-reconstructed Factor 13 average chromatograms for product samples from catalyzed reactions in MeOH/DMC (green), catalyzed reactions in MeOH (blue), and non-catalyzed reactions in MeOH (black). B) The Factor 13 mass spectrum. C) Factor 13 abundance in the product samples generated from non-catalyzed reactions in MeOH (MC), catalyzed reactions in MeOH (MeOH), and catalyzed reactions in MeOH/DMC (DMC) each undergoing reaction for 1, 2, 3, 6, and 9 h. The Factor 13 abundance is shown for the product generated from methanol-soluble (MS; red) and methanol-insoluble (MIS; blue) fraction of lignin extracted from a hybrid poplar biomass source. D) PMF-reconstructed Factor 13 average chromatograms for product samples from the combination of all reaction conditions.

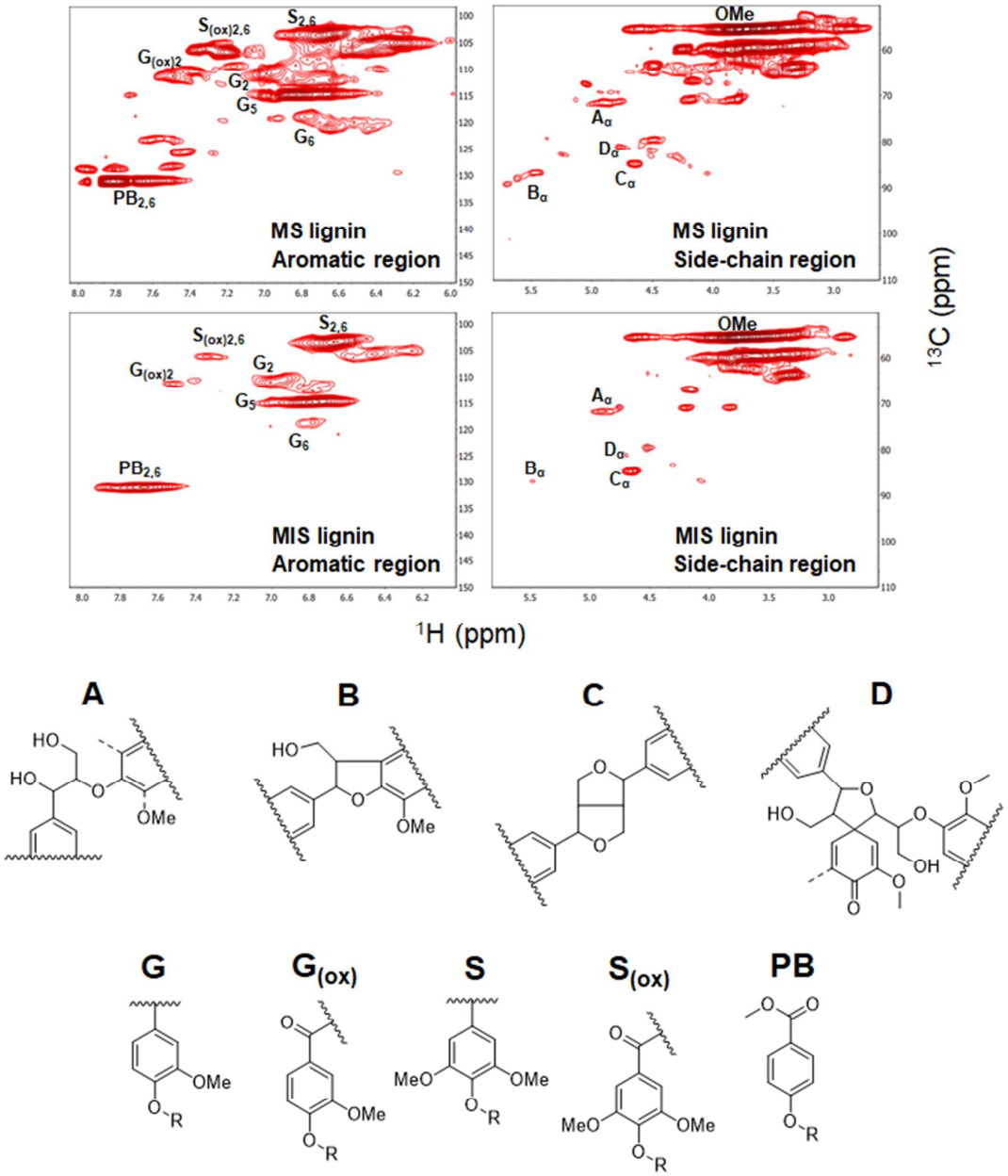

Figure S14. $^{13}C–^{1}H$ HSQC NMR spectra of untreated MS and MIS lignin: (A) β–O–4′ aryl ether, (B) phenylcoumaran, (C) resinol, and (D) spirodienone linkages; (OMe) methoxyl groups; and (S) syringyl, (G) guaiacyl, (S_{ox}) α-oxidized syringyl, (G_{ox}) α-oxidized guaiacyl, (H) *p*-hydroxyphenyl, and (PB) *p*-hydroxybenzoate monomers.

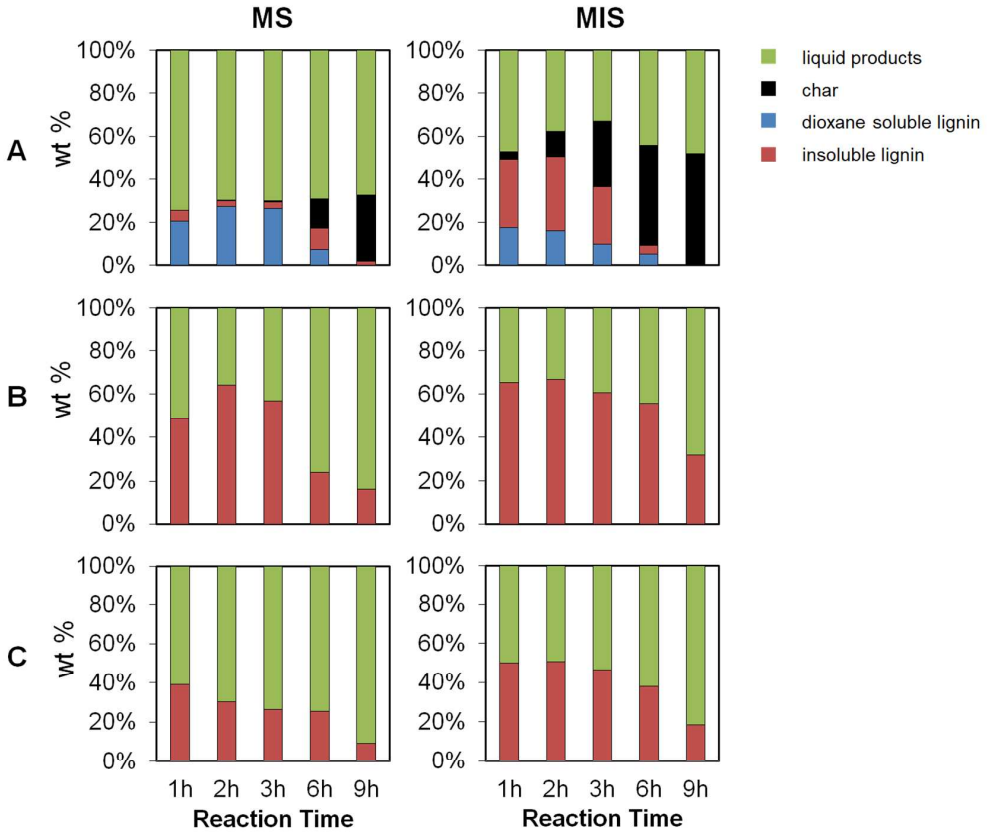

Figure S15. Carbon balance of all liquid and solid products from MS/MIS lignin depolymerization for non-catalyzed (A), MeOH (B), and MeOH/DMC (C) conditions for 1-9 h.

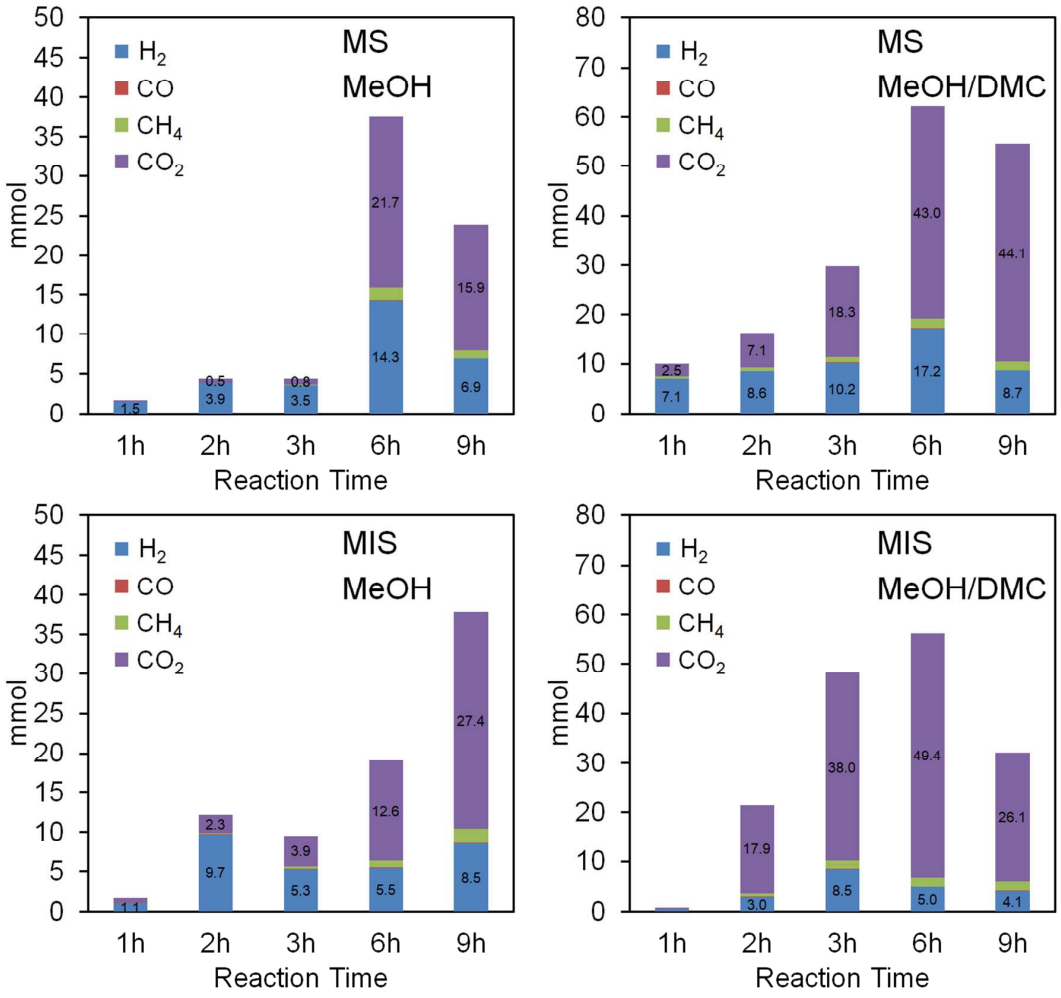

Figure S16. Yields of gas in mmol from MS/MIS lignin depolymerization in non-catalyzed, MeOH, and MeOH/DMC conditions for 1-9 h.

Tables.

Table S1. GC-MS detected peak assignments for compounds in samples from MS/MIS lignin depolymerization in non-catalyzed, MeOH, and MeOH/DMC conditions for 1-9 h. Assignments are based on mass spectral database searches using the Palisade Complete Mass Spectral Database (600 K edition, Palisade Mass Spectrometry, Ithaca, NY)

Number	Compound List	Retention Time (min)	Matching %
1	hydroxy acetic acid methyl ester	2.21	63
2	1-propanol	2.267	90
3	1-butanol	2.385	18
4	formic acid ethyl ester	2.396	28
5	2-butanol	2.561	83
6	2,2-dimethoxy propane	2.7	86
7	tetrahydro-6,6-dimethyl-2H-Pyran-2-one	2.836	78
8	1-butanol	2.857	83
9	3-buten-1-ol	2.965	68
10	1-ethoxy-2-propanol	3.18	78
11	2-methoxy ethanol	3.351	81
12	benzoic acid, 3-pyridyl ester	3.652	23
13	1,2-butylene glycol	3.845	72
14	trimethoxy methane	4.296	78
15	3-methoxy-1-butanol	4.393	72
16	3-pentanol	4.415	74
17	3-heptanol	4.457	45
18	2,2-dimethyoxybutane	4.51	76
19	2-methylbutan-1-ol	4.682	59
20	2-(2-methoxyethoxy) ethanol	4.822	40
21	1,2,3-trimethyl cyclopentene	4.854	93
22	4-methyl-2-pentanone	5.026	64
23	1,4-dioxane	5.06	96
24	toluene	5.423-5.477	94
25	2-(2-ethoxyethaoxy) ethanol	5.61	64
26	alpha-methyl-1,4-benzenedimethanol	5.981	50
27	trans-2,5,5-trimethyl-1,3-hexadiene	6.175	83
28	isopropyl butanoate	6.42	38
29	3,5,5-trimethyl cyclohexene	6.475	43
30	trans-2,5,5-trimethyl-1,3-hexadiene	6.507	49
31	1,3-dimethyl-2-methylene cyclohexane	6.69	43
32	methoxy acetic acid, methyl ester	6.883	86
33	propyl hydrazine	6.905	83
34	3-hexanol	7.141	64

35	1,1-dimethoxy ethane	7.13	78
36	1,2-butanediol	7.152	76
37	5-methyl-3-hexanone	7.227	49
38	1,1-dimethoxy-2-propanone	7.31	93
39	methyl dimethoxyacetate	7.699	9
40	cis-2-methyl-cyclopentanol	7.946	90
41	2-hexen-1-ol	7.989	35
42	xylene	8.249-8.281	88
43	5-methyl-2-(1-methylethyl) cyclohexanol	8.944	53
44	carbamic acid, methyl ester	9.009	43
45	4-methyl -1-heptene	9.127	59
46	1-heptyne	9.138	56
47	1- butanol	9.148	9
48	vinyl-2-(ethoxy)ethyl ether	9.32	45
49	2octyl cyclopropanetetradecanoic acid, methyl ester	9.331	40
50	heptadecane	9.395	50
51	1,1,3-trimethoxypropane	9.535	50
52	cyclohexanol	9.631	76
53	4,5-diethyl-1,2-dimethyl cyclohexene	9.642	83
54	2,5-dimethyl-2-(1-methylethenyl) cyclohexanone	9.653	64
55	cyclohexanol	9.685	50
56	Furfural	10.05	87
57	1,3-pentadiene	10.157	72
58	2,3-dimethyl-3-undecanol	10.372	42
59	1,6-hexanediol	10.383	42
60	pantolactone	10.48	53
61	3-hydroxy-3-methylpent-4-enal	10.49	53
62	3,5-dimethyl cyclohexanol	10.608	38
63	d-siomenthol	10.63	40
64	propane	10.705	4
65	4-pentenal	10.747	59
66	2-methyl cyclohexanol	10.834	95
67	2,6-dimethyl-2-heptanol	10.89	68
68	trans-2-methyl cyclohexanol	11.016	91
69	4-methyl cyclohexanol	11.145	50
70	4-methylcyclohexene	11.242	50
71	methoxy benzene	11.361-11.414	97
72	1-cyclopropyl-2-propen-1-one	11.321	53
73	isopropenyl allyl acetylene	11.381	9
74	1.4-cyclohexanedimethanol	11.457	47
75	1-heptyne	11.542	59
76	trimethoxymethane	11.56	50

77	butanoic acid, 4-methoxy, methyl ester	11.735	64
78	2,2-dimethylcyclohexanone	11.8	64
79	4-methyl-1-heptanol	11.811	50
80	(S)-2-hexen-4-ol	11.832	50
81	3-acetyl-2,6-heptanedione	12.047	53
82	methoxy cyclheptane	12.122	43
83	3-penten-2-ol	12.133	52
84	3,4-dimethylcyclohexanol	12.143	50
85	2,4-dimethylcyclohexanol	12.24	64
86	4-oxo-5-methoxy-2-penten-5-olide	12.39	94
87	3-methylpent-2-ene-1,5-diol	12.466	36
88	4-pentenal	12.466	25
89	3,3-dimethyl cyclohexanol	12.476	46
90	2-methyl propanoic acid pentyl ester	12.51	63
91	1-hexene	12.906	49
92	1-ethoxy-octane	13.067	47
93	2-methyl-3-pentanol	13.174	47
94	E-1,5,9-decatriene	13.271	64
95	1,2-dimethyl-cyclopent-2-enecarboxylic acid	13.464	42
96	2-methyl-1-octene	13.582	47
97	3,3,4-trimethylcyclohexanone	13.593	53
98	4-pentenal	3.603	49
99	phenol	13.79	99
100	1-methoxy-2-methylbenzene	14.012	93
101	2-heptenal	14.022	50
102	phenol acetate	14.087	43
103	2-methyl-2-oxiranyl-cyclobutanone	14.108	53
104	2-methyl-1-buten-3-yne	14.376	53
105	1-methoxy-4-methylbenzene	14.444	99
106	methyl furoate	14.473	38
107	3-methyl cyclohexene	14.677	50
108	1,5-heptadiene	14.806	38
109	1-tetradecanol	15.01	80
110	15-tetracosenoic acid, methyl ester	15.03	80
111	2-hexenal	15.074	59
112	4-pentyn-1-ol	15.106	49
113	4-oxo-pentanoic acid, methyl ester	15.139	80
114	2-ethyl hydrazinecarboxylic acid, methyl ester	15.18	78
115	2-methyl-1-pentene	15.214	46
116	1H-pyrrole-2,5-dione	15.41	39
117	2,3-bis(methylene)-1,4-butanediol	15.439	45
118	1(2-methylbutyl) cyclopentane	15.461	42

119	dodecanal	15.482	47
120	2-ethoxy-2-(2-furyl)ethanol	15.6	43
121	2-isopropyl-5-methyl-1-heptanol	15.622	38
122	bis(2-butoxyethyl) ether	15.76	40
123	2-methyl phenol	16.03	98
124	4-methyl phenol	16.083	97
125	butanedioic acid, dimethyl ester	16.126-16.169	83
126	butanedioic acid, dimethyl ester	16.158	83
127	(2,4,6-trimethylcyclohexyl) methanol	16.255	38
128	2-heptyne	16.341	58
129	3-heptadecenal	16.352	47
130	3,3,5-trimethyl cyclohexanol	16.577	35
131	4,4-dimethoxy-butanoic acid, methyl ester	16.599	48
132	5,5-dimethoxy-3-methyl-2-penten-3-ol	16.61	50
133	4-methylphenol	16.695-16.747	96
134	2,4-dimethylanisole	16.805	96
135	2,3-dimethylanisole	16.846	86
136	4-oxo-pentanoic, ethyl ester	17.08	38
137	p-cumenol	17.157	56
138	heptyl isobutyl ketone	17.189	93
139	3-(1-methylethyl)-phenol	17.21	86
140	1-methyl-1-(2-methyl-2-propenyl) cyclopentane	17.221	68
141	3-cyclopropylcarbonyloxydodecane	17.297	47
142	2-penten-1-ol	17.38	47
143	2-methyl-2-cyclopenten-1-one	17.49	49
144	Benzoic acid, methyl ester	17.576-17.623	98
145	2,6-dimethyl phenol	17.704	97
146	2-methoxy phenol	17.741	96
147	2-methylene cyclohexanol	17.94	70
148	3,3-dimethyl-2-methylene-4,7-oxo-cyclopentane[a]cyclohept-5-ene	18.07	52
149	5-hexyl-2-furaldehyde	18.112	46
150	(1,3-dimethyl-2-methylene-cyclopentyl) methanol	18.145	48
151	9-octadecen-1-ol	18.25	47
152	1-dodecanol	18.37	38
153	3-butyn-1-ol	18.424	62
154	4-cyclohexyl-3-(methoxycarbonyl)-2-methyl-4-butanolide	18.52	50
155	(trimethyl-butyl)-cyclohexane	18.63	56
156	1,3-dioxolane-2-methanol, 2,4-dimethyl	18.714	40
157	3,4-dimethyl phenol	18.81	98

158	2,4-dimethylphenol	18.853	97
159	(methoxymethyl) benzene	18.864	76
160	8-hydroxyocta-1,2-diene-4-one	18.982	50
161	(E)-1-(benzyloxy)-2,3-epoxyocatane	19.09	46
162	1-(2,2-dimethylcyclobutyl)ethanone	19.132	47
163	1-methyl-3-vinyl-3-cyclohexen-1-ol	19.218	43
164	1,4,4-trimethylcyclohexa-2-en-1-ol	19.422	68
165	2-cyclohexen-1-ol, 3,5,5-trimethyl	19.454	47
166	4-methyl benzenemethanol	19.529	47
167	benzoic acid ethyl ester	19.561	86
168	2-ethenyl-2-butenal	19.647	64
169	1-methoxy-4-propyl benzene	19.658	93
170	2-(4-methoxyphenyl)ethanol	19.701	68
171	1-methyl-6-propyl phenol	19.712	80
172	7-[(tetrahydro-2H-pyran-2-yl)oxy]-2-octen-1-ol	19.723	38
173	5-hexyn-1-ol	19.776	42
174	2-methoxy-4-methyl phenol	19.97	99
175	(2S,6S)-(2,6-dimethylcyclihexylidene) methanone	20.012	74
176	1,2-dimethoxy benzene	20.023	97
177	(2S,4S)-5,5-dimethyl-2,4-hexanediol	20.173	14
178	4-methoxy-2-methyl phenol	20.302	76
179	2-ethyl-2,5-dimethylcyclopent-2-enone	20.313	74
180	2,3,4-trimethyl phenol	20.388	98
181	2,6-dimethyl-2,4-heptadiene	20.485	64
182	2-methyl cyclododecanone	20.506	80
183	4-methyl-2-methoxy phenol	20.496-20.540	98
184	1-furyl-1-ethoxy-ethanol	20.657	50
185	methyl-4-pentynoate	20.753	38
186	3,4-dihydroxyacetophenone	20.786	72
187	4-ethyl-2-methoxy phenol	20.807	94
188	2,6,6-trimethyl-1-cyclohexene-1-carboxaldehyde	20.839	58
189	4-hydroxy-benzoic acid methyl ester	21.419	5
190	5,5-dimethyl-1-propyl-1,3-cyclopentadiene	21.891	9
191	endo, exo-3,7-dioxatetracyclodeca-9-ene	21.945	83
192	trans-2-nonadecene	20.968	47
193	2,3,6-trimethyl phenol	21.419	95
194	Cis-4-(tetrahydropyran-2-yloxy)cyclohex-2-enol	21.44	47
195	4-propyl phenol	21.88	74
196	1-formyl-2,2,6-trimethyl-3-cis-(3-methylbut-2-enyl)5-cyclohexene	21.966	46
197	2-(1,1-dimethylethyl)-1,4-benzenediol	22.106	76
198	2-(4-methoxyphenyl)ethanol	22.154	90

199	3,4-dimethoxy toluene	22.224-22.275	99
200	2,4,6-trimethyl phenol	22.299	93
201	3,4-dimethylanisole	22.31	60
202	5-ethoxymethyl furfural	22.4	64
203	2,3,5-trimethyl phenol	22.439	95
204	4 ethyl-4-methyl-2-cyclohexen-1-one	22.492	60
205	3-ethyl guaiacol	22.503-22.514	66
206	1,4-dimethoxy-2-methyl benzene	22.514	87
207	1-(2-furanyl)-3-pentanone	22.535	70
208	2-butynedioic acid, dimethyl ester	22.621	37
209	2-methoxy benzeneethanol	22.718	87
210	4-ethyl-2-methoxy phenol	22.747	98
211	3,4-dimethoxy toluene	22.75	91
212	4-methoxy acetophenone	22.825	72
213	4-ethyl-2-methoxy phenol	22.943	93
214	(2-phenethylcarbamoyl-ethyl)-carbamic acid, benzyl ester	22.965	64
215	ethenyl benzene	23.008	38
216	3,4-diethyl-2,5-dimethyl-2,4-hexadiene	23.029	49
217	3,5-dihydroxy acetophenone	23.34	63
218	2,3,5-trimethyl-1,4-benzenediol	23.351	72
219	4-hydroxy-2,4,5-trimethyl-2,5-cyclohexadien-1-one	23.383	47
220	4-methoxy-1,2-benzenediol	23.394	83
221	2 methoxy-1,4-benzenediol	23.448	78
222	8,8-dimethyl-1,9-diazabicyclo[5.5.0]decane-5,10-dione	23.566	82
223	4-methyl-2-propylphenol	23.63	81
224	2-methylocta-2,4,6-trienedial	23.652	64
225	methyl-8-oxooctanoate	23.759	35
226	2-(2-methyl-2-propenyl)cyclohexanone	23.834	68
227	1-cyclohexene-1-carboxylic acid	24.006	43
228	1,4-dimethoxy-2,3-diemthylbenzene	24.188	86
229	4-ethyl-1,2-dimethoxy phenol	24.307	99
230	2-methoxy-4-vinylphenol	24.348	97
231	heptanoic acid	24.5	64
232	1,4-dimethoxy-2,3-dimethylbenzene	24.535	92
233	5-Allyl-6-methyl-3,3a,4,6-tetrahydropyrolo[3,4-c]isoxazole	24.543	83
234	2-(3-methyl-2-butenylidene)cyclohexanone	24.586	64
235	2-methoxy-4-ethyl-6-methyl phenol	24.596	86
236	2-methoxy-4-propyl phenol	24.852-24.872	96

237	2-methyl-5-(1-methylethyl) phenol	24.876	93
238	2,3,5,6-tetramethyl phenol	24.994	89
239	1,2,3-trimethoxy benzene	25.015	97
240	Cis-1-hydroxy-2-methoxy-4-propenyl benzene	25.079	64
241	5-methylnicotinic acid	25.187	59
242	2-methoxy-4-propyl phenol	25.23	87
243	2-methoxy-4-propyl phenol	25.262	76
244	1-(3,4-dimethoxyphenyl) ethanone	25.283	72
245	8-oxa-9-azabicyclo[3.2.2]non-6-ene	25.316	43
246	cyclotetradecane	25.348	46
247	cis-1-ethyl-2-methyl cyclopentane	25.4	76
248	2,3-dimethyl-4-methoxy phenol	25.477	64
249	5-methoxy-2,3,4-trimethyl phenol	25.53	64
250	4-(3-hydroxy-1-propenyl)-2-methoxy phenol	25.595	59
251	hexanoic acid	25.67	52
252	4-D-2-methyl-3-pentanol	25.745	53
253	4,4-dimethyl heptanedioic dimethyl ester	25.83	86
254	3,4-dimethoxy propiophenone	25.981	53
255	1-(2,4-dihydroxy-3-propylphenyl)ethanone	25.992	83
256	3,4-dimethoxy propiophenone	26.014	52
257	1-(2-hydroxy-5-methoxy-4-methylphenyl) ethanone	26.067	43
258	1,4-dimethoxy-2,3,5-trimethyl benzene	26.164	74
259	1,2,3-trimethoxy benzene	26.196	86
260	1,2-dimethoxy-4-n-propyl benzene	26.28-26.322	98
261	2,6 dimethoxy phenol	26.538	97
262	4-methoxybenzoic acid, methyl ester	26.497-26.552	99
263	3-methoxychromene	26.679	62
264	5-methoxy-2,3,4-trimethyl phenol	26.7	80
265	5-hepten-3-yn-2-ol, 6-methyl-5-(1-methylethyl)	26.711	76
266	1-(2-hydroxy-6-(methoxymethyl)phenyl) ethanone	26.829	52
267	ethyl-2-methyl-5-cyanopenta-2,4-dienoate	26.862	50
268	4-(2-methyl-cyclohex-1-enyl)-but-3-en-2-one	26.904	46
269	2-methoxy-5-(2'hydroxyethyl) phenol	27.055	57
270	1,2,3-trimethoxy-5-methyl benzene	27.093	99
271	1,2-dimethyl-2-(1-naphthyl) cyclopropane	27.184	72
272	5-ethyl-1,2,3-trimethoxy benzene	27.194	50
273	di-t-butyl phenol	27.216	60
274	8,8-dimethyl-1,9-diazabicyclo[5.3.0]decane-5,10-dione	27.248	80
275	1,2-dimethoxy-4-n-propyl benzene	27.463	66
276	anisaldehyde dimethyl acetal	27.509	52
277	1-methyl-2-(phenylmethyl) benzene	27.527	52

278	2-hydroxy-5-methoxy benzaldehyde	27.699	52
279	Isopropylidenecyclobutenone	27.742	50
280	3,4-dimethoxy propiophenone	27.774	68
281	2-(phenylethynyl) phenol	27.785	72
282	2-methoxy-4-(1-propenyl) phenol	27.836	96
283	1-(2,4-dimethoxyphenyl)-1-propanone	27.957	81
284	3,4-diethoxy benzaldehyde	28.021	47
285	4-methoxybenzoic acid, ethyl ester	28.107	94
286	5-methoxy-2,3,4-trimethyl phenol	28.182	94
287	4-methoxy-2,4,6-trimethyl cyclohexa-2,5-dienone	28.204	74
288	4-hydroxy-3-methoxy benzaldehyde, vanillin	28.364	98
289	m-isopropylbenzoic acid	28.418	49
290	(1,1-dimethylethyl)-4-methoxy-phenol	28.547	87
291	1,2,3-trimethoxy benzene	28.59	83
292	2,6-dimethoxy-4-methyl phenol	28.611	95
293	5-ethyl-1,2,3-trimethoxy benzene	28.676	98
294	5-ethyl-1,2,3-trimethoxybenzene	28.729	98
295	3-[3,4-(methylenedioxy)phenyl]propan-1-ol	28.88	78
296	methyl 3-methoxy-4-methyl benzoate	28.901	87
297	(5R0-1-methyl-5-(1-methyl-1-ethenyl)2,3-diazabicyclo[3.3.0.]octane	28.966	53
298	1,2-dimethyl-4-(phenylmethyl) benzene	29.02	50
299	4-hydroxylbenzoic acid, methyl ester	29.152	98
300	3-hydroxy benzoic acid, methyl ester	29.233	87
301	methyl-3-(5-acetyl-2-tienyl)-2-propenoate	29.298	53
302	(2-methoxyethoxy) benzene	29.417	32
303	4-(1,1-dimethylethyl) benzenemethanol	29.545	87
304	homo-vanillin (4-hydroxy-3-methoxy-phenyl) actaldehyde	29.653	72
305	3-isopropyl-1,2-dimethoxybenzene	29.717	82
306	3,4-dimethoxy benzaldehyde	29.797	96
307	trans-methyl iso-eugenol	29.814	95
308	1-(2-ethenyl-1-cyclohexenyl)-2-methyl-2-propen-1-ol	30.05	35
309	1,2,3,4-tetrahydro-9-propyl anthracene	30.173	52
310	2-ethoxy-3,4,6,7,8,9-hexahydro-8,8-dimethyl-6-oxo-2H-chromene	30.19	59
311	4-ethyl syringol	30.2	91
312	tert-butyl biphenyl carboxylic acid	30.254	89
313	1,2,3-trimethoxy-5-propylbenzene	30.28-30.332	96
314	1-(4-hydroxy-3-methoxyphenyl) ethanone, acetovanillone	30.393	96
315	methyl 3-(5-formyl-2-furyl_)-2-propenoate	30.404	43
316	3,4-dimethoxy benzeneacetic acid	30.415	87

317	3-(3,4-dimethoxyphenyl) propionic acid	30.479	58
318	4-hydroxybenzoic acid, methyl ester	30.594	98
319	1,3-dimethoxy-2-(prop-2-enyl) benzene	30.683	60
320	4-hydroxy-3-methoxybenzoic acid, methyl ester	30.701	97
321	propio-syringone	30.768	62
322	3,5-bis(1-methylethyl) phenol	30.823	64
323	2-methoxy benzoic acid ethyl ester	30.941	87
324	4-(ethoxymethyl)-2-methoxy phenol	30.995	38
325	2-hydroxy-5-methoxy benzaldehyde	31.199	51
326	Homovanillyl alcohol	31.287	96
327	2-methoxy-4-propyl-phenol	31.338	90
328	3,4-diethoxy benzaldehyde	31.349	53
329	2-(2,5-dimethoxy-phenyl) propionaldehyde	31.36	49
330	3,4-diethoxy benzaldehyde	31.392	46
331	4-vinyl syringol	31.692	64
332	2-acetyl-3,6-dimethyl benzoic acid	31.821	49
333	4-propyl syringol	31.864	86
334	syringyl aldehyde	31.908	89
335	3-methoxybenzyl-2,2-dimethyl propanoate	31.918	64
336	ethyl vanillate	31.982	47
337	1-(2,5-dimethoxy-4-methylphenyl)-2-propanone	32.047	46
338	trans-isoelemicin	32.058	76
339	1,2,3-trimethoxy-5-(2-propenyl) benzene	32.111	95
340	2,6-dimethoxy-4-(2-propenyl) phenol	32.197	68
341	3-(3,4-dimethoxyphenyl) propionic acid	32.294	53
342	2,2-dimethoxyethoxy benzene	32.336	78
343	3-(3,4-dimetoxyphenyl)propionic acid	32.347	49
344	(7,7-dimethyl-1-oxo-2,3,4,5,6,7-hexahydro-1H-inden02-yl)acetic acid, ethyl ester	32.42	50
345	benzofuran-4(5H)-one, 6,7-dihydro-, oxime	32.433	48
346	4-hydroxy-3-methoxy benzoic acid, methyl ester	32.476	48
347	2-(2,4,5-trimethylphenyl)propylene oxide	32.573	93
348	4-hydroxy-3-methoxy propiophenone	32.585	96
349	3,4-dimethoxy benzoic acid, methyl ester	32.734	92
350	2-(1,1-dimethyl-2-propenyl)-3,6-dimethyl pehnol	32.734	76
351	3,4-dimethoxybenzoic acid, methyl ester	32.767	98
352	2-phenoxyethyl-beta-pehnylpropionate	32.809	48
353	2,6-dimethoxy benzoic acid, methyl ester	33.077	93
354	1,2,4-triethyl-5-methyl benzene	33.142	56
355	methyl syringate	33.163	49
356	3,4-dimethoxy benzaldehyde	33.303	47
357	6-methoxy-2,2-dimethyl-1-indanone	33.313	38
358	(2,2-dimethoxyethyl) benzene	33.464	52

359	trans-4-propenyl syringol	33.496	53
360	alpha, 4-dihydroxy-3-methoxy benzeneacetic acid methyl ester	33.555	87
361	3,4-dimethoxy benzeneacetic acid	33.603	98
362	4-hydroxy-3-methoxy-benzeneacetic acid	33.883	76
363	coniferyl alcohol	33.916	91
364	o-Methylmaleimycin	33.979	68
365	2-(2-formylvinyl)azulene-1-carbaldehyde	34.022	46
366	2,2-diphenylpropionic acid	34.086	52
367	dimethyl 4-(2'-furyl)-1-methyl-2,3-dihydro-1H-indole-6,7-dicarboxylate	34.183	55
368	4-[(4-hydroxy-3-methoxyphenoxy)methyl)]-3-methoxybenzaldehyde	34.263	46
369	4-hydroxy-3-methoxy-phenylacetylformic	34.301	60
370	8-(biphenyl-2-ylmethyl)-5-ethyl-2,3,5,6-tetrahydroimidazo[1,2-a] pyridine	34.333	47
371	tridecanoic acid, methyl ester	34.398	97
372	14-methyl-pentadecanoic acid, methyl ester	34.441	89
373	2,6-dimethoxy-4-(2-propenyl) phenol	34.678	91
374	3-(3,4-dimethoxyphenyl)-1-propanol	34.795	93
375	3,4-dimethoxy benzenepropanol	34.842	86
376	3,4-dimethoxy benzenepropanoic acid, methyl ester	35.246	92
377	3,4-dimethoxy benzenepropanoic acid, methyl ester	35.299	94
378	syringaldehyde	35.309	97
379	hexadecanoic acid, ethyl ester	35.568	99
380	3,4,5-benzoic acid, methyl ester	35.611	94
381	2,5-dimethoxybenzoic acid	35.74	52
382	4-hydroxy-3,5-dimethoxy benzaldehyde	35.793	50
383	2-ethyldiphenyl methane	36.029	34
384	methyl-2-oxo-1-propyl cycloheptanecarboxylate	36.094	38
385	1-(4-hydroxy-3,5-dimethoxyphenyl) ethanone	36.652	96
386	(2,2-dimethoxyethyoxy) benzene	36.695	54
387	(3-methoxyphenyl) carbamic acid, methyl ester	36.802	42
388	3-(2,3,4-trimethoxypehnyl)propionic acid	36.845	50
389	alpha, hydroxy-3-methoxy benzeneacetic acid, methyl ester	37.017	35
390	(Z)-7-phenyl-1,4-heptadien-6-yne	37.168	46
391	4-hydroxy-3,5-dimethoxybenzoic acid, methyl ester	37.305	97
392	syringyl acetone	37.386	92
393	Octadecanoic acid, methyl ester	38.048	95
394	2,5-dimethoxy benzoic acid	38.155	72
395	2,2-dimethoxyethoxy benzene	38.23	76
396	ethyl 4-hydroxyphenylcarbamate	38.4	68
397	o-2-benzimidazolyl phenol	38.584	30

Table S2. List of major characteristic *m/z* values.[1-3]

Characteristic m/z	Factors	Likely Molecular Formula	Fragment Identity
39	1, 3, 6-10	C_3H_3	aromatic
41	4, 11	C_3H_5	aliphatic
45	10	C_2H_5O	aliphatic alcohol
50	1	C_4H_2	aromatic
51	3	C_4H_3	aromatic
52	10	C_4H_4	aromatic
53	7	C_3HO or C_4H_5	aliphatic alcohol or aliphatic
55	4, 9, 11	C_4H_7	aliphatic
63	3	C_5H_3	aromatic
65	1, 6, 8, 9, 11	C_5H_5	aromatic
69	4	C_5H_9	aliphatic
74	1	CH_2=C(OH)OCH	methyl ester
75	5	C_2H5O-C=O+2H[a] or C_2H_5COO+2H[a]	carboxylate or carboxylic
77	3, 11	C_6H_5	aromatic
79	8-10	C_6H_5+2H[a]	aromatic
83	4	C_6H_{11}	aliphatic
91	3, 8	$C_6H_5CH_2$	benzylic
92	10	$C_6H_5CH_2$+H[a]	benzylic
93	1, 6	C_6H_5O	phenolic
94	9	C_6H_5O+H[a]	phenolic
95	7	C_6H_5O+2H[a]	phenolic
97	4	C_7H_{13}	aliphatic
105	9-11	C_6H_5C=O or C_6H_5-CH_2CH_2	benzocarbonyl or ethylbenzyl
107	3, 8	$CH_2C_6H_4OH$ or $C_6H_5CH_2O$	benzylic alcohol or benzylic ether
109	7	C_6H_5-CH_2O+2H[a]	benzylic ether
111	4	C_8H_{15}	aliphatic
119	8	C_6H_5-$C(CH_3)_2$	isopropylbenzyl
121	6	C_6H_5COO	benzoate
122	9	CH3OC6H4CH2	methoxy benzyl
137	9	$CH_3OC_6H_3OHCH_2$	methoxy phenyl
151	8	$(CH_3O)_2C_6H_3CH_2$	dimethoxy benzyl
167	10	$(CH_3O)_2$C6H2OHCH_2	dimethoxy phenyl
181	10	$(CH_3O)_3C_6H_2CH_2$	trimethoxy benzyl
195	10, 12	$(CH_3O)_3C_6H_2CH_2CH_2$	trimethoxy benzyl

[a]The "+H" notation means that the ion was formed by a rearrangement that involved the transfer of a hydrogen atom from some other part of the molecule.

Table S3. Distribution of hydroxyl group contents (mmol/g) based on quantitative ^{31}P NMR data for untreated MS and MIS lignin.

Sample	Carboxylic acid	Phenolic OH				Total Phenolic OH	Aliphatic OH
		H	C_5 substituted*	G	S		
MS	0.18	0.56	0.16	0.84	2.19	3.75	1.37
MIS	0.01	0.29	0.42	0.53	1.21	2.45	2.15

* C_5 substituted: β-5, 4-O-5, and 5-5 substructures

Table S4. Distribution of carbons functional group contents (percent) based on quantitative ^{13}C NMR data for untreated MS and MIS lignin.

% Carbon	MS	MIS
Aliphatic C	6.51	1.96
Methoxyl-aromatic C	27.0	15.6
Aliphatic C-O	3.17	11.6
Aromatic C-H	15.4	17.3
Aromatic C-C	8.74	26.6
Aromatic C-O	37.8	26.0
Carbonyl C	1.42	0.84

Table S5. GPC detected number-average molecular weight (M_n), weighted-average molecular weight (M_w), and polydispersity index (PDI) of untreated and depolymerized MS lignin in non-catalyzed, MeOH, and MeOH/DMC conditions for 1-9 h. Molecular weights were determined based on a polystyrene standard calibration curve.

MS	M_n	M_w	PDI
untreated	883	1734	1.96
non-catalyzed 1h	625	1038	1.66
non-catalyzed 2h	461	987	2.14
non-catalyzed 3h	476	793	1.67
non-catalyzed 6h	346	671	1.94
non-catalyzed 9h	280	415	1.48
MeOH 1h	445	745	1.67
MeOH 2h	238	398	1.67
MeOH 3h	236	372	1.58
MeOH 6h	250	443	1.78
MeOH 9h	245	370	1.51
MeOH/DMC 1h	223	495	2.22
MeOH/DMC 2h	238	454	1.91
MeOH/DMC 3h	228	491	2.15
MeOH/DMC 6h	231	421	1.82
MeOH/DMC 9h	234	363	1.55

Table S6. GPC detected M_n, M_w, and PDI of untreated and depolymerized MIS lignin in non-catalyzed, MeOH, and MeOH/DMC conditions for 1-9 h. Molecular weights were determined based on a polystyrene standard calibration curve.

MIS	Mn	Mw	PDI
untreated	2923	7867	2.69
non-catalyzed 1h	480	779	1.62
non-catalyzed 2h	361	732	2.03
non-catalyzed 3h	430	637	1.48
non-catalyzed 6h	408	892	2.18
non-catalyzed 9h	296	463	1.56
MeOH 1h	298	496	1.67
MeOH 2h	217	377	1.74
MeOH 3h	208	318	1.52
MeOH 6h	228	452	1.98
MeOH 9h	270	417	1.55
MeOH/DMC 1h	317	771	2.43
MeOH/DMC 2h	241	445	1.85
MeOH/DMC 3h	247	513	2.08
MeOH/DMC 6h	231	441	1.9
MeOH/DMC 9h	266	410	1.54

References:

1. M. Hamming, *Interpretation of mass spectra of organic compounds*, Elsevier, 2012.
2. F. W. McLafferty and F. Turecek, *Interpretation of mass spectra*, University science books, 1993.
3. R. M. Silverstein, F. X. Webster, D. J. Kiemle and D. L. Bryce, *Spectrometric identification of organic compounds*, John wiley & sons, 2014.
4. J. A. Barrett, Y. Gao, C. M. Bernt, M. Chui, A. T. Tran, M. B. Foston and P. C. Ford, *ACS Sustainable Chemistry & Engineering*, 2016, **4**, 6877-6886.