

SANDIA REPORT

SAND2013-10094
Unlimited Release
November 2013

Tailoring Next-Generation Biofuels and their Combustion in Next-Generation Engines

John M. Gladden, Craig A. Taatjes, Connie Gao, Greg O'Bryan, Amy J. Powell,
Adam M. Scheer, Kevin Turner, Weihua Wu and Eizadora T. Yu

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.



Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: <http://www.osti.gov/bridge>

Available to the public from

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd.
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: <http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online>



SAND2013-10094
Unlimited Release
November 2013

Tailoring Next-Generation Biofuels and their Combustion in Next-Generation Engines

John M. Gladden and Weihua Wu
Biomass Science & Conversion Technology (Mailstop 9292),

Craig A. Taatjes and Adam M. Scheer
Combustion Chemistry Department (Mailstop 9055)

Kevin Turner and Eizadora Yu
Systems Biology Department

Greg O'Bryan
Materials Chemistry Department

Amy J. Powell
Nanobiology Department

Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185

Connie Gao
Department of Chemical Engineering,
Massachusetts Institute of Technology, Cambridge, MA 02139

Abstract

Increasing energy costs, the dependence on foreign oil supplies, and environmental concerns have emphasized the need to produce sustainable renewable fuels and chemicals. The strategy for producing next-generation biofuels must include efficient processes for biomass conversion to liquid fuels and the fuels must be compatible with current and future engines. Unfortunately, biofuel development generally takes place without any consideration of combustion characteristics, and combustion scientists typically measure biofuels properties without any feedback to the production design. We seek to optimize the fuel/engine system by bringing combustion performance, specifically for advanced next-generation engines, into the development of novel biosynthetic fuel pathways. Here we report an innovative coupling of combustion chemistry, from fundamentals to engine measurements, to the optimization of fuel production using metabolic engineering. We have established the necessary connections among the fundamental chemistry, engine science, and synthetic biology for fuel production, building a powerful framework for co-development of engines and biofuels.

ACKNOWLEDGMENTS

This work was funded under LDRD Project Number 151308, "Tailoring Next-Generation Biofuels and their Combustion in Next-Generation Engines."

Prof. Fei Qi (University of Science and Technology of China) and Prof. Ravi X. Fernandes (RWTH Aachen University) are gratefully acknowledged for their collaborative experiments characterizing the pyrolysis and ignition chemistry of 2,4-dimethylpentan-3-one.

Besides the report's authors, the following researchers also contributed to the scientific work described here:

Masood Z. Hadi (Systems Biology Department)

John E. Dec and Yi Yang (Engine Combustion Department)

Haifeng Huang, Subith S. Vasu, John D. Savee, Oliver Welz, and David L. Osborn
(Combustion Chemistry Department)

Carol Kozina and Mary Tran-Gyamfi (Biomass Science & Conversion Technology
Department)

Jose Gutierrez (Stanford University / Sandia Masters Fellow)

Lindsey Orgren (Materials Chemistry Department)

Joshua W. Allen, Shamel S. Merchant, and William H. Green (Department of Chemical
Engineering, Massachusetts Institute of Technology)

Gary A. Strobel (Department of Plant Sciences & Plant Pathology, Montana State
University)

CONTENTS

1. Introduction: The co-development of biofuels and advanced internal combustion engines ...	11
2. Characterization of VOCs from Endophytic Fungi	15
2.1 Cultivation of fungi in minimally-treated biomass feedstocks	16
2.2 VOC profiles of fungus grown in PD and biomass feedstock	17
2.3 Extracted hydrocarbon profile of strain CI-4A grown in PD	19
3. Reconstruction of the terpene pathway in <i>E.coli</i>	23
3.1 Endophyte genomes are rich in terpene synthases	23
3.2 Terpene production from the engineered <i>E. coli</i> strain	24
3.3 Optimization of terpene production	25
4. Characterization of cellulase expression in endophytes Hypoxylon and Daldinia	27
4.1 Genomic analysis of endophytic CAZy enzymes	27
4.2 <i>Beta</i> -glucosidase activities of different strains on different feedstocks	30
4.3 Exoglucanase (CBH) activities of different strains on different feedstocks	32
4.4 Endoglucanase activities of endophytes and <i>N.crassa</i> on different feedstocks	34
4.5 Zymography assay of secretome	36
5. Identification of proteins in the fungal secretome by mass spectrometry	37
6. Combustion Chemistry of Ketones and Cineole	41
6. 1 Ignition Chemistry of Ketones	42
6.1.1 Open-chain ketones	42
6.1.2 Cyclic ketones	51
6.1.3 Engine performance	55
6.2 Ignition Chemistry of Cineole	56
6.2.1 Fundamental chemistry experiments	56
6.2.2 Model generation	59
7. Genomic perspectives on “mycodiesel” fungal endophytes	67
7. 1 Sequencing and estimated evolutionary relationships of endophytic members of the fungal family Xylariaceae	67
7.2 Annotating genes and pathways potentially involved in secondary metabolism	68
8. Conclusions	71
9. References	75
Appendix A: Section 2 Materials and methods ^a	79
Appendix B: Section 3 Materials and methods ^a	83
Appendix C: Section 4 Materials and methods ^a	87
Appendix D: Section 5 Materials and methods ^a	89
Appendix E: Section 6 Materials and methods	91
Appendix F: Section 2 Supplemental tables	95

FIGURES

Figure 1. Framework for integrating combustion performance into optimization of biofuels	11
Figure 2. VOC profile of endophytic fungi cultivated on various carbon sources	17
Figure 3. Chromatogram of the extracted hydrocarbons from a culture of C14A	20
Figure 4. Conserved active domain of terpene synthases	24
Figure 5. Phylogenetic tree of terpene synthases	24
Figure 7. KEGG metabolic pathways of four genomes	28
Figure 8. Distribution of CAZy in endophytes	30
Figure 9. β -glucosidase activities of endophytes and <i>N. crassa</i> on different feedstocks	32
Figure 10. Exoglucanase activities of endophytes and <i>N. crassa</i> on different feedstocks	33
Figure 11. Endoglucanase activities of endophytes and <i>N. crassa</i> on different feedstocks	35
Figure 12. Zymography assay of the endophytic secretome	36
Figure 13. Classification of proteins within the fungal secretome with various feed stocks	39
Figure 14. General scheme of the initial steps of low-temperature oxidation	41
Figure 15. Difference mass spectra of Cl-initiated oxidation of di- <i>tert</i> -butyl ketone	43
Figure 16. Difference mass spectra of Cl-initiated oxidation of isopropyl- <i>tert</i> -butyl ketone	46
Figure 17. Absolute HO ₂ profiles for the Cl-initiated oxidation of diisopropyl ketone.	49
Figure 18. Absolute OH profiles for Cl-initiated diisopropyl ketone oxidation	50
Figure 19. Difference mass spectra at 10.2 eV for the products of Cl-initiated reaction of cyclopentanone	52
Figure 20. Photoionization spectra for $m/z = 82$ peak observed in 550 K Cl-initiated oxidation of cyclopentanone	53
Figure 21. Difference mass spectra of Cl-initiated oxidation of cineole	57
Figure 22. Time, mass, and photon-energy resolved data for Cl-initiated oxidation of cineole at 550 K	58
Figure 23. Simulated product spectrum for chlorine-initiated oxidation of cineole	64
Figure 24. Genome size and number of genes in the first sequenced members of the fungal family Xylariaceae	67
Figure 25. <i>Daldinia eschshlozii</i> secondary metabolite gene cluster size distribution	68
Figure 26. A continuous liquid-liquid extraction vessel	81
Figure 27. NMR spectrum of isolated fraction obtained through semi-preparative TLC of the fungal extracts	82
Figure 28. Vector map of plasmid JBEI3122 containing the up-pathway of terpene biosynthesis	83

TABLES

Table 1. List of endophytic fungal isolates in this study	15
Table 2. Compounds identified in the headspace of CO27-A cultures in PD and biomass feedstocks.....	21
Table 3. Summary of genome sequencing and analysis	28
Table 4. CAZys in endophytes.....	29
Table 5. Calculated CBS-QB3 relative energies at 0 K ($\Delta E_{0, \text{rel}}$) and AIEs of cyclic ethers associated with isopropyl-tert-butyl ketone oxidation.....	48
Table 6. Names and structures of key species in Cineole decomposition	61
Table 7. Computed rates for Cineole abstraction and decomposition pathways	62
Table 8. Initial species' mole fractions for reaction systems used in model generation	63
Supplementary Table 9. List of tentative compounds identified in the headspace of CI-4A cultures cultivated with PD and biomass feedstocks	95
Supplementary Table 10. List of tentative compounds identified in the headspace of EC-12 cultures cultivated with PD and biomass feedstocks	96
Supplementary Table 11. List of tentative compounds identified in the headspace of EC-38 cultures cultivated with PD and biomass feedstocks	97

NOMENCLATURE

2-CHO-yl	2-oxocyclohexyl
2-CPO-yl	2-oxocyclopentyl
2-Me-CPO	2-methyl-cyclopentanone
3-CHO-yl	3-oxocyclohexyl
3-CPO-yl	3-oxocyclopentyl
4-CHO-yl	4-oxocyclohexyl
AIE	adiabatic ionization energy
ALS	Advanced Light Source
AMDIS	Automated Mass Spectral Deconvolution and Identification System
antiSMASH	antibiotics and Secondary Metabolite Analysis Shell
AUC	area under curve
BDE	bond dissociation energy (or enthalpy)
BF	biomass feedstock
BGL	<i>beta</i> -glucosidase
CAZy	carbohydrate-active enzymes
CBH	cellobiohydrolase
CBM	carbohydrate binding module
CE	carbohydrate esterase
CHO	cyclohexanone
CPO	cyclopentanone
CRN	corn stover
DIPK	di-isopropyl ketone (2,4-dimethylpentan-3-one)
DNA	deoxyribonucleic acid
DNS	dinitrosalicylic
DOE	Department of Energy
DTbuK	di- <i>tert</i> -butyl ketone (2,2,4,4-tetramethylpentan-3-one)
EUC	arbo eucalyptus
EXPN	expansin-like protein
FT-ICR	Fourier-transform ion cyclotron resonance
FTIR	Fourier-transform infrared
GC	gas chromatography
GH	glycosyl hydrolase
GPPS	geranyl diphosphate synthase
GT	glycosyl transferase
HCCI	homogeneous-charge compression ignition
HMM	hidden Markov model
HPLC	high-performance liquid chromatography
IR	infrared
ITbuK	isopropyl <i>tert</i> -butyl ketone (2,2,4-trimethylpentan-3-one)
ITS	internal transcribed spacer
JBEI	Joint Bioenergy Institute
JGI	Joint Genome Institute
LBNL	Lawrence Berkeley National Laboratory
LC	liquid chromatography

LDRD	laboratory directed research and development
(M)PIMS	(multiplexed) photoionization mass spectrometry
MS	mass spectrometry
NASA	National Aeronautics and Space Administration
NC	<i>Neurospora crassa</i>
Nd:YAG	neodymium yttrium aluminum garnet
NIST	National Institute of Standards and Technology
NMR	nuclear magnetic resonance
PD	potato dextrose broth
PFAM	protein families
PL	polysaccharide lyase
RMG	Reaction Mechanism Generation
RRKM	Rice-Ramsperger-Kassel-Marcus
SAE	Society of Automotive Engineers
SI	spark ignition
smCOG	secondary metabolism clusters of orthologous groups
SNL	Sandia National Laboratories
SPME	solid-phase micro extraction
SWG	switchgrass
TLC	thin layer chromatography
TOF	time of flight
TS	terpene synthase
VOC	volatile organic compound

1. INTRODUCTION: THE CO-DEVELOPMENT OF BIOFUELS AND ADVANCED INTERNAL COMBUSTION ENGINES

Combustion of petroleum-based liquid fuels accounts for a large majority of the energy in the transportation sector worldwide, and liquid hydrocarbon fuels have tremendous advantages, including high volumetric energy density and easy transport and storage.¹ However, potential for supply interruptions and especially the increasing call to mitigate climate change are driving exploration of alternative liquid fuels.² Use of lignocellulosic biofuels could potentially lower greenhouse gas emissions substantially, but the recalcitrance of lignocellulosic material continues to make the broad application of such biofuels difficult. Deconstructing biomass into fermentable sugars is one of the most costly steps in biofuel production, and new methods to break down lignocellulose are being broadly investigated.³ The products that are most efficiently produced in novel biomass conversion strategies may be rather different from the compounds in petroleum distillate fuels and may have poorly known combustion characteristics. Furthermore, at the same time as the chemistry of the fuel stream is beginning to change, advanced clean, efficient combustion strategies are emerging, and these new engines are often very sensitive to fuel chemistry.^{4,5} Novel fuels may hinder the operation of advanced engines, or they may in fact be enabling: coordinated efforts towards biofuel-engine co-development are needed to navigate this landscape.

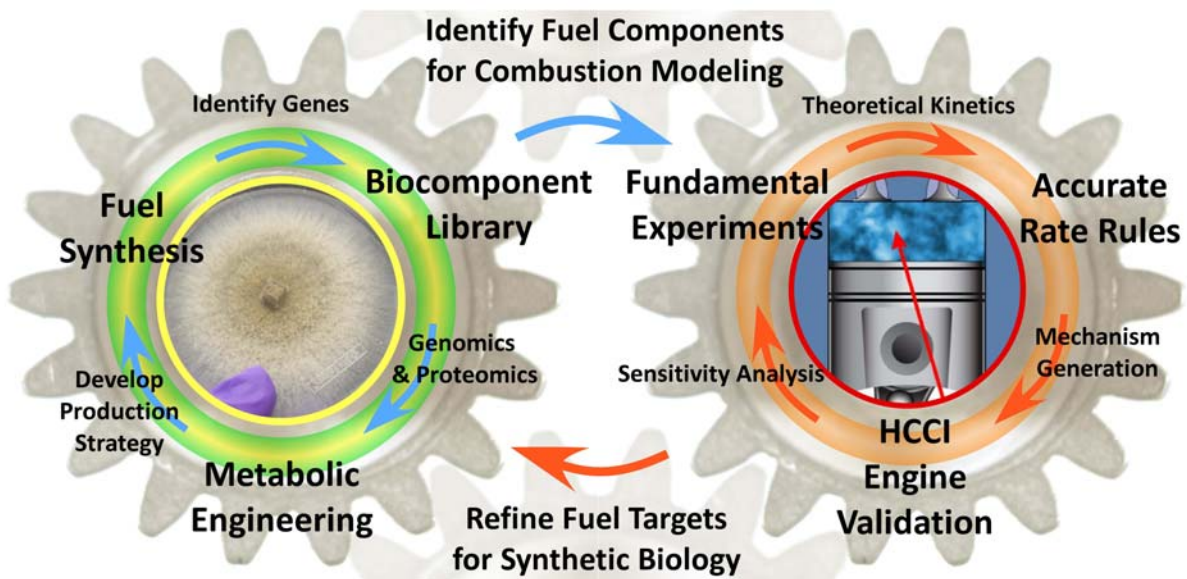


Figure 1. Framework for integrating combustion performance into optimization of biofuels

Figure 1 depicts the framework for collaborative biofuel design as developed in this project. Synthetic biologists working in close collaboration with combustion researchers develop fundamental mechanisms for the production and combustion of potential biofuels. Ignition and engine trials then provide feasibility tests for fuels and mixtures and yield recommendations for the bioengineering scale-up of specific metabolic pathways. This coupling of fundamental and applied combustion chemistry and synthetic biology is a strategy to identify and investigate the most promising fuel compounds through mutual feedback. Moreover, the development of combustion models will provide the predictive capability needed for eventual efficient utilization of the new biofuel stream.

The framework is specifically envisioned to be flexible, and to allow exploration of different biofuel production platforms and different combustion technologies. For the purposes of this project a biofuel production strategy was chosen based on metabolic pathways in endophytic fungi. These organisms naturally live in symbiosis with woody plants, consuming some of the biomass of the host plant and producing volatile organic compounds, many of which are potentially useful fuels.⁶⁻⁹ Moreover, because the endophytes produce these useful compounds “directly” from woody biomass, harnessing their metabolic pathways brings the potential for consolidation of the lignocellulose deconstruction and the sugar conversion into one process, promising substantial cost savings. The natural products of endophytic fungal metabolism include ketones, cyclic ethers and other complex oxygenates, whose combustion performance has not yet been characterized.

The combustion system chosen for study is the Homogeneous Charge Compression Ignition (HCCI) engine, a representative of the advanced combustion strategies based on low-temperature combustion. In a traditional spark-ignition (SI) or gasoline engine, a stoichiometric premixed fuel-air charge is compressed and ignited by a spark near the top of the piston travel. As the flame front propagates through the mixture the high temperatures at the flame front can break convert the nitrogen in the air to nitrogen oxides (NO_x) pollutants. However, the three-way catalysts in use in SI vehicles can remove NO_x very effectively. In a Diesel engine, air is compressed and heated as the piston moves toward the cylinder head, and combustion is initiated by injection of liquid fuel into the hot air near the top of the piston stroke. The fuel evaporates,

ignites, and burns, again propagating a high-temperature flame front through the mixture and creating NO_x . In addition, because of the injection of liquid fuel, very fuel-rich regions also occur, where particulates (soot) are formed. Moreover, the exhaust in a diesel engine still contains excess oxygen, which does not permit operation of the same effective three-way catalysts that are used for SI engines. As a result, even though diesel engines are substantially more efficient than SI engines, some of that efficiency gain must be traded away to power more involved after-treatment strategies.

In HCCI engines, a lean dilute fuel-air mixture is compressed until the heating and pressure increase autoignite the mixture near the top of the piston travel. In the dilute mixture no flame propagates and the charge burns volumetrically without reaching the temperatures at which NO_x formation occurs, and the fuel-lean combustion produces negligible particulates. Furthermore, the high compression ratio and lack of throttling losses mean that HCCI can achieve similar efficiencies to diesel combustion without requiring the costly and energy-consuming after-treatment. Some difficulties in applying such engines in consumer vehicles are obtaining sufficient power density and controlling engine operation over the entire load-speed range. The control of the combustion phasing, that is, when and how fast the heat release of combustion occurs during the piston cycle, is driven by the chemistry of autoignition, and is therefore sensitive to changes in fuel. Therefore, the HCCI engine was chosen for the combustion strategy in the present framework, because it accentuates the possible effects of novel fuel chemistry.

The conceptual optimization machinery depicted in Figure 1 operates by first identifying the spectrum of compounds that the biofuel production platform creates; in this case, that is the VOC (volatile organic compound) profile from natural fungal metabolism of biomass, as described in Section 2. From these compounds some are identified for combustion chemistry investigation. This is the first “mesh point” between biofuel production and utilization studies. The compounds are chosen based on two criteria: prominence in the product stream and lack of existing combustion chemistry knowledge. In the case of endophytic fungi, ketones and cyclic ethers were identified as fitting those criteria and were subjected to combustion chemistry studies, as described in Section 6. Fundamental measurements of representative compounds and detailed theoretical kinetics efforts are combined to develop combustion chemistry models that can

predict performance of specific molecules as fuels. These models are “validated” against HCCI engine measurements, and conclusions about combustion performance are factored into the direction of biofuel production research. This feedback is a second “mesh point.”

At the same time that the combustion studies are ongoing, the biofuel production research analyzes and dissects the metabolic machinery of the fungi. The fungal genomes are sequenced and annotated, and transcriptomic and secretomic analysis reveals the enzymes responsible for lignolytic and cellulolytic activity and for production of particular fuel molecules. These pathways can then be extracted and expressed in other tractable organisms such as *E coli* and their activity can be optimized. This research is described in Sections 3-5. The targets of this optimization are conditioned on the fuel’s combustion performance, and the success or failure of the optimization will potentially determine the course of future combustion research, two additional feedback points between the production and utilization sides of the project. For example, cineole (a cyclic ether) was identified in the optimization process as a potential scale-up target, and cineole is a focus of additional combustion research (Section 6.2).

2. CHARACTERIZATION OF VOCs FROM ENDOPHYTIC FUNGI

A promising strategy for biofuel production involves the direct conversion of lignocellulose from energy crops or agricultural wastes to value-added fuels by microbial biotechnology.

Microorganisms that possess necessary pathways for biofuel production are continuously being discovered and optimized through synthetic biology¹⁰⁻¹³. In recent years, novel endophytic fungi isolated from rainforests have been shown to produce metabolites touted to be compatible with existing fuels.^{6,7,9,14-16} Several of these fungal isolates produce short- to mid-chain aliphatic hydrocarbons, alcohols, and isoprenoids, which could be potential replacements for diesel or fuel additives.¹⁷ In addition, because of their ecologically unique lifestyle (e.g., efficient plant colonization), endophytes possess the necessary chemi-enzymatic machinery to breakdown and utilize lignocellulosic biomass. Therefore, these endophytes are attractive bioprospecting candidates as they have lignocellulose deconstruction and hydrocarbon/isoprenoid biosynthesis pathways- both of which are highly desirable in bioenergy applications.

Table 1. List of endophytic fungal isolates in this study

Fungal isolate	Genus*	Host Plant
CI-4A **	Hypoxylon	Persea indica
CO27-A	Hypoxylon	Rhizophora sp.
EC-12 **	Daldinia	Myroxylon balsamum
EC-38	Daldinia/Nodulisporium	Neea floribunda

* The taxonomic/phylogenetic classification of fungal isolates was performed by ITS sequence comparisons against reference database.

** Detailed bioactivity and GC-MS analysis of the volatile metabolites for CI-4A and EC-12 have been previously described.^{9,15}

Hydrocarbon production by *Hypoxylon sp.* and *Daldinia sp.* isolates have been demonstrated only when the fungi are cultured on standard laboratory media (e.g. potato dextrose, oatmeal, and cellulose),^{9,15} which are much simpler carbohydrates compared to plant biomass, and it is unknown how hydrocarbon production will be affected by the nature of biomass feedstock

available to the fungi. The aim of this study is to assess the effect of lignocellulosic biomass feedstocks on VOC production by candidate fungal isolates. A detailed description of the work in this section can be found at <http://ntrs.nasa.gov/search.jsp?R=20130013430>.

2.1 Cultivation of fungi in minimally-treated biomass feedstocks.

Biomass pretreatment is essential to increase hydrolysis rates of cellulase cocktails and total sugar yields. Current methods relying on thermochemical processes, although effective, tend to be very harsh and expensive, and other pretreatment methods are being explored to reduce cost and environmental impact.^{18,19} In nature, endophytic fungi obtain nutrition from the barks of trees they colonize, and thus, should have the necessary biochemical machinery to obtain sugars from lignocellulosic biomass. Initially, we wanted to determine whether the fungal isolates can make use of minimally-treated lignocellulosic biomass to support their growth. Feedstocks were milled into sub millimeter-sized chips and sterilized (by gamma radiation or autoclave) prior to use. We followed the growth of two *Hypoxylon* strains and two *Daldinia* strains (Table 1) on M9 agar plates supplied with milled corn stover, switchgrass or arbog eucalyptus as the sole carbon source. All three biomass feedstock supported the growth of all four isolates, albeit at slower growth rates and reached less fungal biomass density as compared to control PDA plates. We observed noticeable differences in the organism's ability to propagate in the feedstock media. The Hypoxylons (CI-4A, CO27-A) and EC-38 mycelium growth was fastest in corn stover. CI-4A and CO27-A growth was slowest on switchgrass, while growth was delayed the most in eucalyptus in EC-38. On the other hand, *D. eschscholzii* (EC-12) did not have any feedstock preference and achieved comparable growth rates. The same growth trends are observed when fungus was cultivated in liquid cultures. Slowest growth was expected for eucalyptus-fed fungal cultures since it is a more recalcitrant plant material (29% lignin). However, we only observed this for EC-38. FT-IR analyses on the feedstocks used in this study show that corn stover and switchgrass have comparable lignin and cellulose content (~18% lignin, ~38% cellulose), however, growth of *Hypoxylon* isolates was favored in corn stover. Further examination of this feedstock preference of the fungal isolates is in progress.

2.2 VOC profiles of fungus grown in PD and biomass feedstock.

Previous reports show prolific VOC production of endophytic fungi when grown in potato dextrose media.^{9,15} These reports were obtained from headspace SPME-GC-MS analysis of fungal cultures cultivated in agar plates or tube slants. Growing the fungi in the BF tube slants was not feasible, as the feedstock sank to the bottom of the culture tubes and was inaccessible to the fungi plug on top of the agar slants. For this study, we switched to liquid cultures to obtain higher fungal biomass densities and VOC production. We performed headspace SPME-GC-MS on all four fungal cultures in PD broth (positive control) and on uninoculated BF media (negative control) prior to profiling VOC in headspace of BF-fed fungal cultures. When cultivated in potato dextrose media, all four fungi produced primarily terpene compounds, with CI-4A generating monoterpenes (C10 terpenes) and the CO27-A and Daldinia isolates producing sesquiterpenes (C15 terpenes) (Figure 2; list of compounds in Table 2, and Supplementary Table 5-7 in appendix). Significant alcohol production in PD-fed EC-12 was also observed (Figure 2), consistent with earlier reports.

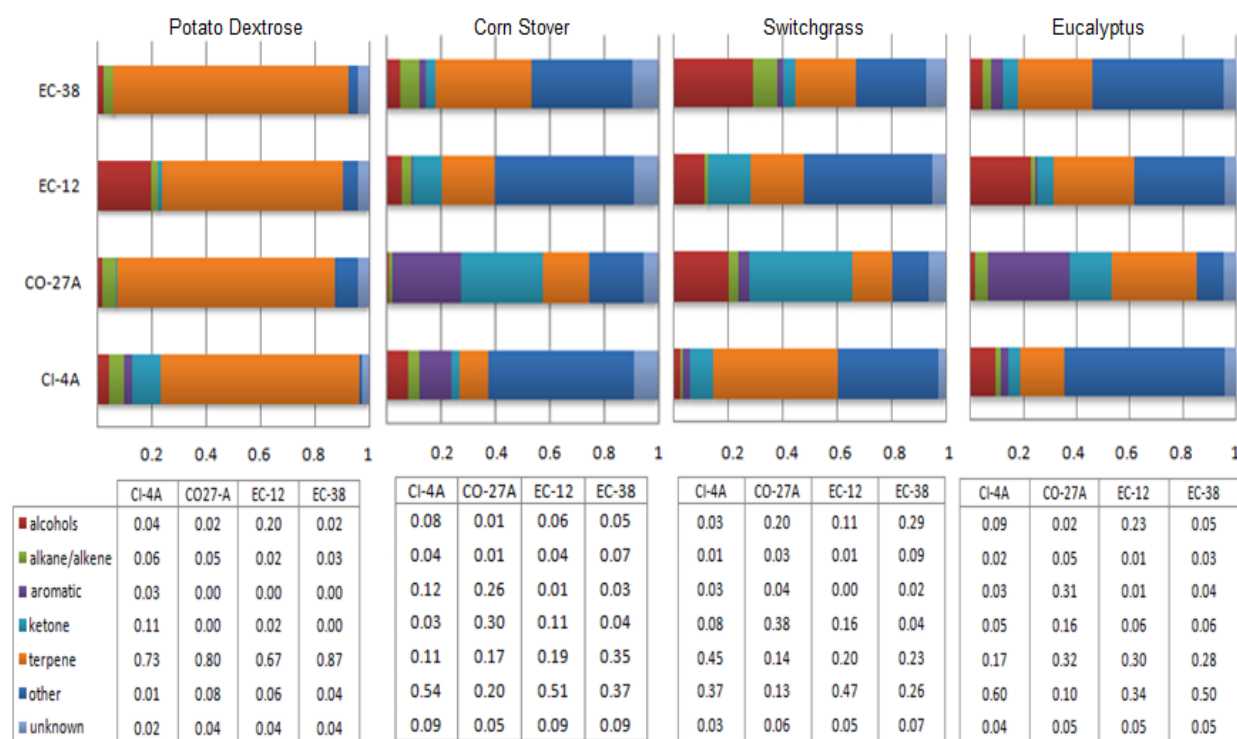


Figure 2. VOC profile of endophytic fungi cultivated on various carbon sources

Figure 2 (cont.) Compounds in the VOC profile are classified based on functional groups (see color legend) and relative abundances are reported as the sum of peak areas for each class of compounds vs total area of the peaks in chromatogram. Data shown is representative of at least three, independent GC-MS analysis of replicate cultures.

Growing the fungi in feedstocks significantly changed the VOC profile for all four isolates. Calculated Sørensen similarity indices show ~ 40% similarity between the total VOCs produced in PD and the other BF media (for figures see the link in this sections introduction). VOC profiles from BF-fed fungal cultures were comparable (similarity indices of 60-74%) between the different fungi. We also performed comparisons of the VOC profile in terms of compound functional groups (e.g. terpenes, ketones, aromatics, alcohols, etc.), to look at the changes in more detail. The similarity between VOCs from BF-fed fungi can be attributed primarily to the terpenoid compounds. In fact, the terpenoid metabolite production of EC-38 and CO27-A do not appear to be affected much by the nature of BF media, as indicated by their Sørensen indices (>73% similarity). Ketones in the VOC of BF-fed EC38 and BF-fed CO27-A cluster together and are distinct from the ketone profile of BF-fed EC12 and BF-fed CI-4A (for figures see the link in this sections introduction). Notably, CI-4A VOC profile was affected the most by changing feedstocks-just by switching from potato dextrose to BF media (24% similarity), and even between biomass feedstock (38% similarity). Not only did we observe differences in the types of terpenes detected in the headspace, but maybe more importantly, we observed a decrease in the percentage of terpenes in the headspace of the fungal cultures (Figure 2). For example, in PD-fed CI-4A cultures, we consistently detected 1,8-cineole and cyclic ketones as predominant peaks in the chromatogram. The 1,8-cineole peak is dramatically decreased in CRN- and EUC-fed CI-4A cultures. In fact, most of the monoterpenes detected in PD-fed CI-4A cultures was not observed at all in any of the BF-fed CI-4A cultures, and there appeared to be a shift to limited sesquiterpene production.

We also detected compounds in the headspace of BF-fed cultures that were not originally observed (or is a minor component, < 1% of total peak area) in the PD-fed cultures. In particular, methoxy-phenyl oxime and isothiocyanate and thiol compounds, which were minimal components in PD-fed cultures, were significant constituents in the headspace of BF-fed

cultures. The biological significance of these compounds for the fungi is unknown, but may have a connection with the fact that these compounds are effective oxygen and radical scavengers, and have been demonstrated in other systems to be potent against fungal pathogens^{20,21}. Ketones were detected in CRN- and SWG-fed isolates, and were particularly abundant in SWG-fed CO27-A (Table 2). Elevated levels of 3,7-dimethyl-1,6-octadien-3-ol or (aka linalool), was detected in the headspace of all switchgrass-fed cultures²². There was also a prevalence of aromatic compounds, mostly benzene derivatives (Figure 2) in BF-fed CO27-A cultures. However, it is more likely that these benzene derivatives are not produced by the fungi but rather, possible by-products of lignin degradation²³.

2.3 Extracted hydrocarbon profile of strain CI-4A grown in PD

The VOC profiles of the four isolates grown on different feedstocks indicates that a plethora of hydrocarbons are being produced by each strain. To better understand the actual quantities of hydrocarbons being produced, one of the four strains was selected for hydrocarbon extraction of the entire culture. The culture of strain CI-4A grown for 5 days in PD was extracted with dimethylether and the complement of hydrocarbons in the extractants was measured using LC-MS (Figure 3). The VOC profile of this culture was composed of 47% 1,8-cineole, 19% 2-Cyclohexen-1-one, 2-(2-methyl-2-propenyl)-, and 3.2% each of Cyclohexane, 1,2,4-tris(methylene)- and β -pinene (see supplementary table 5). The extracted culture had a similar complement of hydrocarbons, but the ratios were shifted: 17% 1,8-cineole, 66% 2-Cyclohexen-1-one, 2-(2-methyl-2-propenyl)-, and 9% of β -pinene. Peak 2 was not clearly resolved, but is possibly Cyclohexane, 1,2,4-tris(methylene)- at 7%. These ratios are not surprising once the vapor pressure of each compound is considered, indicating that the VOC profiles make an accurate assessment of the complement of hydrocarbons being produced by the fungi, and a simple calculation can be used to estimate the overall abundance of each compound if its vapor pressure is known.

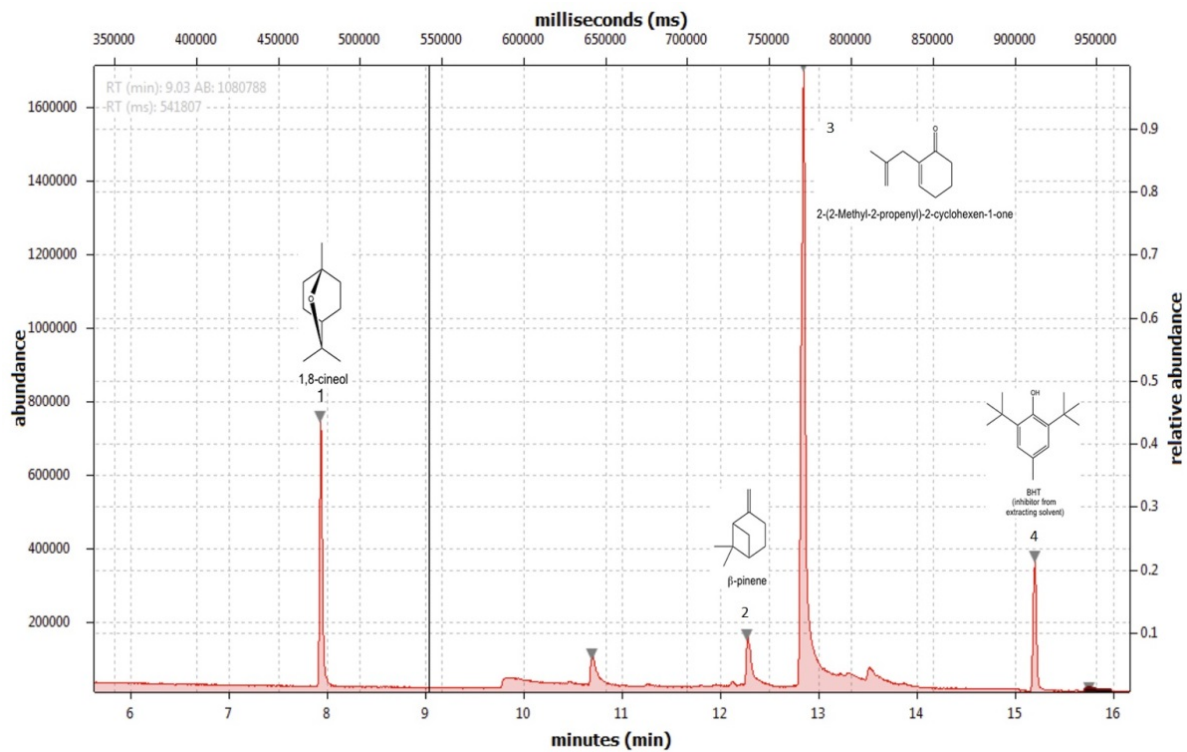


Figure 3. Chromatogram of the extracted hydrocarbons from a culture of C14A

Table 2. Compounds identified in the headspace of CO27-A cultures in PD and biomass feedstocks

Ret. time (min)	Tentative Compound	Molecular Formula	Match*	% Area PDB	% Area CRN	% Area SWG	% Area EUC
28.4	Guaia-1(10),11-diene	C15H24	87	31.5	2.9	2.6	7.7
27.9	Caryophyllene	C15H24	83	13.8	nd	<1	nd
32.0	γ -Gurjunene	C15H24	86	12.6	5	4.8	16.7
10.6	<i>Oxime-, methoxy-phenyl-</i>	<i>C8H9NO2</i>	85	7.3	1.5	1.6	2.6
25.5	β -Elemene	C15H24	87	5.9	4.8	4	1.6
28.1	Chamigrene	C15H24	87	4.8	1.5	1.3	1.8
15.0	Eucalyptol	C10H18O	88	4.8	nd	nd	nd
26.7	α -Guaiene	C15H24	88	2.4	2.2	2.4	1.1
28.6	Aromandendrene	C15H24	75	2	nd	nd	nd
17.6	Phenylethyl Alcohol	C8H10O	77	1.6	nd	nd	nd
28.9	Calamenene	C15H22	75	1.3	<1	<1	nd
11.2	Anisole	C7H8O	78	1.3	nd	nd	nd
28.2	δ -Selinene	C15H24	76	1.2	nd	nd	nd
21.3	<i>Cyclohexane, isothiocyanato-</i>	<i>C7H11NS</i>	75	<1	1.7	<1	10.2
24.0	Benzene, 3-cyclohexen-1-yl-	C12H14	75	<1	nd	nd	<1
27.6	β -Cadinene	C15H24	83	1	<1	nd	nd
28.7	δ -Cadinene	C15H24	81	<1	1	<1	nd
15.1	α -Pinene	C10H16	76	nd	nd	3.4	nd
27.3	2,5-Cyclohexadiene-1,4-dione, 2,6-bis(1,1-dimethylethyl)-	C14H20O2	77	nd	nd	nd	3
18.2	2-Heptenal, 2-propyl	C10H18O	75	nd	5.1	-	4.8
17.6	Benzene, 1-ethyl-4-methoxy-	C9H12O	88	nd	9.2	nd	1.4
16.8	3-Heptanone, 5-ethyl-4-methyl-	C10H20O	82	nd	13.3	29.5	2.6
16.5	3-Pentanone	C5H12O	75	nd	2.9	10.1	7.4
18.9	Benzene, 1-ethenyl-4-methoxy-	C9H10O	85	nd	27.1	6.6	nd
21.1	Benzothiazole	C7H5NS	84	nd	1.3	<1	2.2
14.7	Benzene, 1-methoxy-4-methyl-	C8H10O	80	nd	2.2	nd	2

*VOC profiles from potato dextrose broth (PDB), corn stover (CRN), switchgrass (SWG) and arbog eucalyptus (EUC). Percent Area is taken as AUC (area under curve) of peak of interest versus total AUC for entire chromatogram. Some compounds are not detected (nd) for certain conditions. Compounds are observed in some uninoculated media (italicized).

3. RECONSTRUCTION OF THE TERPENE PATHWAY IN *E. COLI*

Endophytes *Hypoxylon* CI-4A, EC-38, CO27-A, and *Daldinia* EC-12 produced a wide spectrum of volatile organic compounds (VOCs) when they were grown on potato dextrose and other feedstocks. Most of these VOCs are monoterpenes and sesquiterpenes, which are C10 and C15 hydrocarbons that contain a high energy density and are potential replacements for gasoline. Genomic data mining revealed that there are more than 350 genes involving in secondary metabolite biosynthesis from each endophytic genome, including isoterpenoid biosynthesis pathway genes. In this section, the putative terpene synthase (TS) genes annotated in the genome were informatically and functionally characterized.

3.1 Endophyte genomes are rich in terpene synthases

The VOC analysis of the headspace of the endophyte cultures on potato dextrose medium and other three feedstocks (switch grass, corn stover, and eucalyptus) has identified a wide spectrum of monoterpene and sesquiterpene compounds, indicating the genomes of these organisms will harbor multiple terpene synthases. Mining each of the four fungal genomes (*Hypoxylon* (CI-4A, EC-38, and CO27-A) and *Daldinia* EC-12), a total of 26 putative TS genes and 3 sesquiterpene synthase genes. A protein sequence alignment showed that these fungal TS have low identity (<30%) to plant and bacterial TS. All TS exhibit a highly conserved aspartate-rich motif (DDXXD/E) and a (N/D)DXX(S/T)XX(K/R)(D/E) consensus sequence (NSE/DTE triad) involved in binding of the Mg²⁺ cofactor the complexes the diphosphate for ionization, as shown in Figure 4. Phylogenetic tree analysis indicates that all 26 TS can be categorized into five clusters (Figure 5).

sesquiterpenes were produced, sometimes simultaneously by the same synthase. The TS in the same cluster tended to produce a similar spectrum of terpene compounds, indicating a link between structure and function. All the functional TS produced multiple compounds, including chamigrene, gurjunene, caryophyllene, selinene, pinene, limonene, ocimene, etc., except CO27-31178 which produced only cineole (see Supplementary Table 8 for representative examples).

3.3 Optimization of terpene production

To optimize expression, several different constructs were tested on five representative TS enzymes (see methods). The results showed that the optimal construct differed for each TS tested. For TS gene EC38-200002, the construct 3 (pBbE1a-TS-GPPS) didn't yield any terpene compounds. The construct 4 (pBbE1a-GPPS-TS) produced 3.5 and 5.5 times higher concentration of β -pinene and α -guaiene than construct 1, while construct 2 yielded the highest concentrations of 1S- α -pinene and β -cis-ocimene, which were 2.8 and 2.7 times higher than that from construct 1.

4. CHARACTERIZATION OF CELLULASE EXPRESSION IN ENDOPHYTES *HYPOXYLON* AND *DALDINIA*

Endophytes are a relatively unexplored class of fungi that establish a symbiotic relationship with plants. Endophytes produce a wide variety of secondary metabolites as well as cellulolytic enzymes that may assist their adaptation and survival within higher plants. Recently, the endophytes *Hypoxyton* CI-4A, EC-38, CO27-A and *Daldinia* EC-12 have been shown to produce a wide spectrum of bioactive volatile organic compounds (VOCs) that contain a high energy density, making them promising potential replacements for the fossil fuels such as gasoline. Additionally, genomic data mining identified that these four endophytes possess an extensive complement of cellulolytic enzyme machinery, which indicates they have a robust capacity for degrading cellulosic biomass. In this project, the dynamics of the cellulolytic machinery of these four strains was investigated in the presence of different feedstocks.

4.1 Genomic analysis of endophytic CAZy enzymes

The genomes of four endophytes (*Hypoxyton* CI-4A, CO27-A, EC-38, and *Daldinia* EC-12) have been sequenced. They are the first reported genomes for endophytic fungi. Transcript analysis using RNA-Seq was also conducted and covered 99 % of the genomes. In summary, the genome of the four endophytes range between 37.5 Mb and 47.3 Mb, as shown in Table 3. The numbers of gene in each genome are over 11000 and the average gene size is between 1700 bp and 1900 bp. The gene density of each genome is between 250 to 320 genes/Mbp.

Genomic analysis showed that the two most abundant gene categories are carbohydrate and lipid metabolism. In addition, all four genomes are abundant in genes of glycan biosynthesis and metabolism, polyketides biosynthesis, secondary metabolite biosynthesis, and xenobiotics biodegradation, indicating that they have a diverse potential for use in applications ranging from cellulosic enzymes cocktails to fuel production, in addition to their known application in drug discovery.

Table 3. Summary of genome sequencing and analysis

	Daldinia EC12	Hypoxylon CI4A	Hypoxylon CO27	Hypoxylon EC38
Genome size (Mb)	37.5	37.7	46.5	47.3
Contigs	641	1044	505	1168
Number of Genes	11173	11712	12256	12261
Gene Length (nt)	1811	1761	1800	1757
Transcript Length (nt)	1638	1600	1590	1576
Protein Length (aa)	467	459	459	465
Exon Length (nt)	568	564	548	552
Intron Length (nt)	94	89	112	100
Exon Frequency (no./gene)	2.89	2.84	2.9	2.86
Gene density (no./Mbp)	298	311	263	259

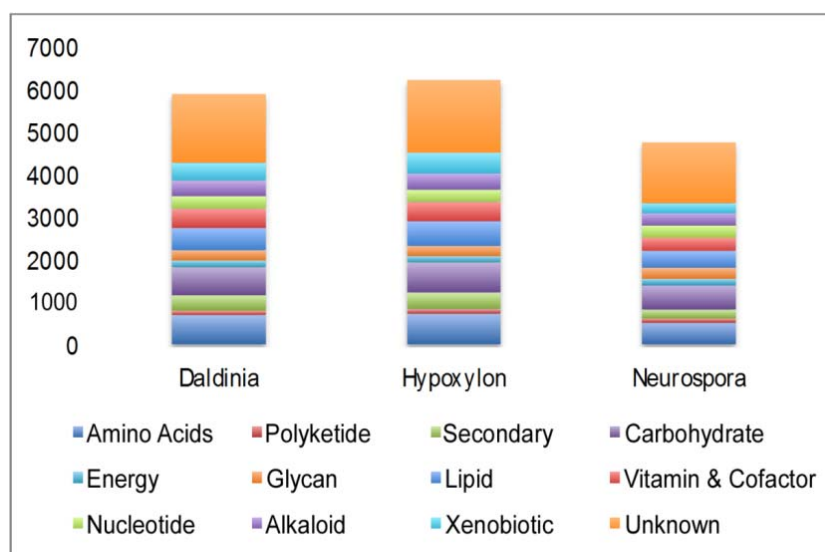


Figure 6. KEGG metabolic pathways of four genomes

Genomic analysis revealed that these four endophytes are rich in Carbohydrate Active enZymes (CAZy). The total numbers of identified putative CAZy enzymes in these fungi are 14% to 20% greater than that of the model cellulolytic fungus: *Neurospora crassa*, as shown in table 4. Among all CAZy enzymes, glycosyl hydrolase (GH) and glycosyl transferase (GT) are two most abundant enzyme super families, which account for more than 85% of all identified CAZy enzymes. In particular, the GH family enzymes are more than 60% of total CAZy enzymes.

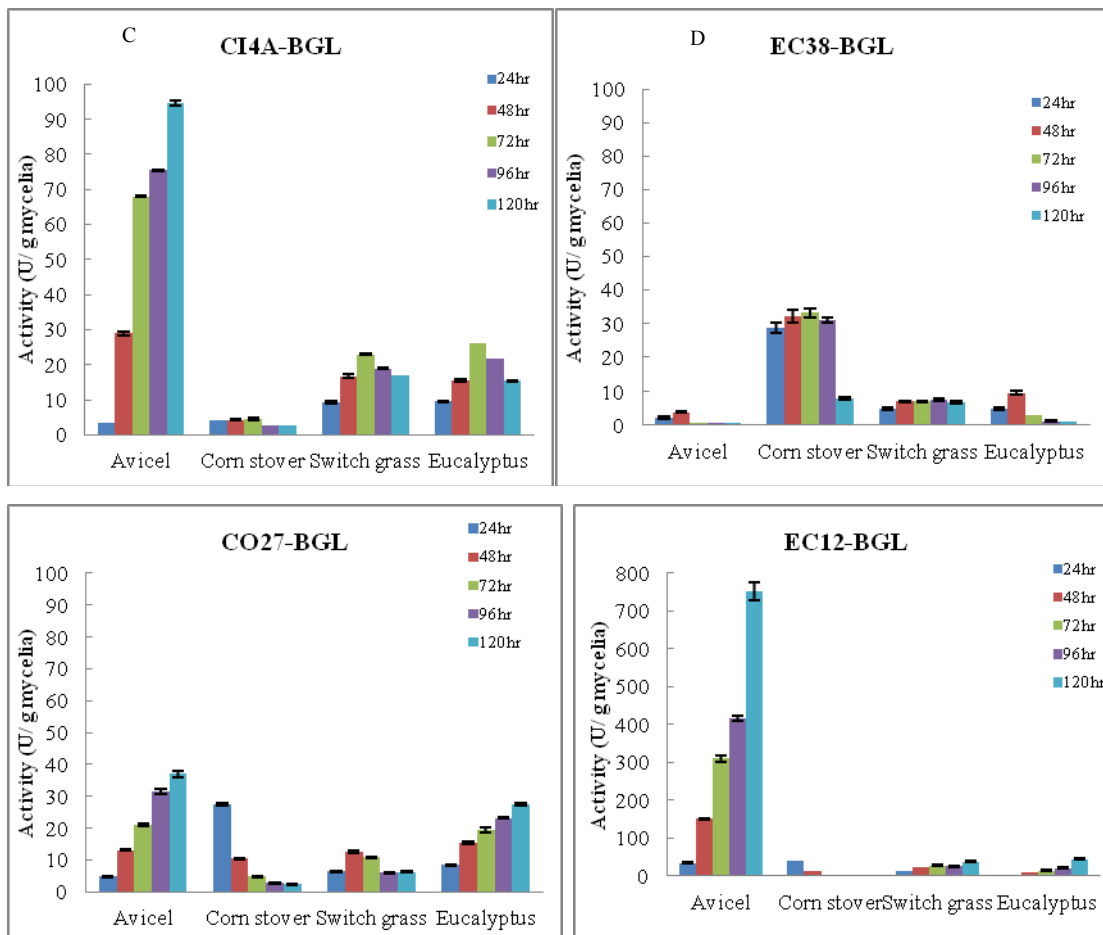
Table 4. CAZys in endophytes

	CI4A	CO27	EC12	EC38	N.crassa
GH	262	278	265	274	177
GT	92	88	88	88	76
PL	10	10	8	10	4
CE	35	38	37	38	22
CBM	33	41	35	41	42
EXPN	4	5	5	5	1
Total	417	437	417	434	365

Classification of CAZy: glycosyl hydrolases (GH), glycosyl transferase (GT), polysaccharide lyases (PL) and carbohydrate esterase (CE) and auxiliary modules- carbohydrate binding module (CBM) and plant expansin-like proteins (EXPN). *Neurospora crassa* (NC) genes are shown for comparison

In terms of glycosyl hydrolase family enzymes (Figure 8), the six most abundant categories are: GH61, which are oxidoreductase functioning as cellulase activity enhancers in the degradation of cellulosic biomass; GH43, which belong to *beta*-xylosidase family; GH3, of which most are *beta*-glucosidases; GH16, which are xyloglucanases; GH5, which are major endoglucanases; and GH18, which are chitinase family.

tested. For these three substrates, the maximal BGL activity produced by the endophytes was either achieved after the first few day of culture, except the substrate eucalyptus where strains CO27-A, EC-12, and *N. crassa* achieved the maximal BGL activities only after 5 days of culture. Strain CI-4A produced significantly higher BGL when grown on switchgrass and eucalyptus than on corn stover. The other three strains CO27-A, EC-12, and *N. crassa*, produced a similar amount of BGL activity on all three feedstocks. Interestingly, EC-38 yielded the highest BGL activity on corn stover rather than Avicel, which indicates that these fungi have different induction mechanisms for their cellulolytic machinery.



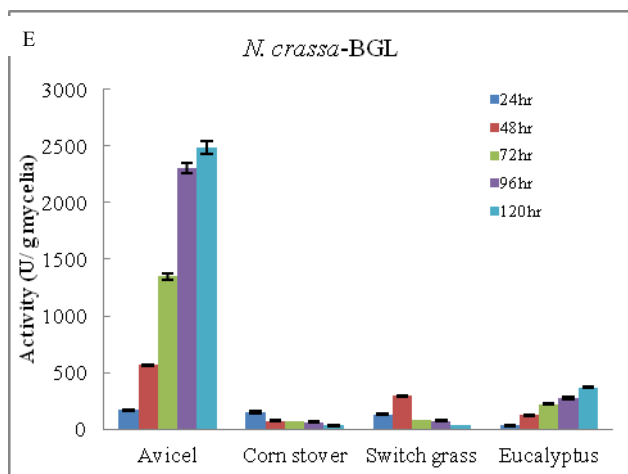


Figure 8. β -glucosidase activities of endophytes and *N. crassa* on different feedstocks

4.3 Exoglucanase (CBH) activities of different strains on different feedstocks

When all four endophytes were grown on corn stover, switchgrass, and eucalyptus, only strains *Hypoxylon* CI-4A and CO27-A produced detectable CBH activities. The maximal activities were somewhat low compared to other fungi and were below 1.1 U/ g mycelia. Surprisingly, *Daldinia* EC-12 and EC-38 produced no obvious CBH activities on these three feedstocks. However, all four endophytes produced detectable cellobiohydrolase (CBH) activities on Avicel. For all endophytes, the CBH activities increased over the duration of the culture, except for EC-38, which achieved the maximal CBH activities (0.96 U/ g mycelia) after two days. For the other three endophytes, the highest CBH activity achieved was 94.62 U/ g mycelia by strain *Daldinia* EC-12 after 5 days, followed by CO27-A (14.01 U/ g mycelia), and CI-4A (6.82 U/ g mycelia). In contrast to the endophytes, *N. crassa* produced detectable CBH activities on all four feedstocks. For the substrates Avicel and eucalyptus, the maximal CBH activity was achieved after 5 days (100.59 and 13.14 U/ g mycelia, respectively), while the highest CBH activity were reached after only two days when *N. crassa* was grown on corn stover or switchgrass.

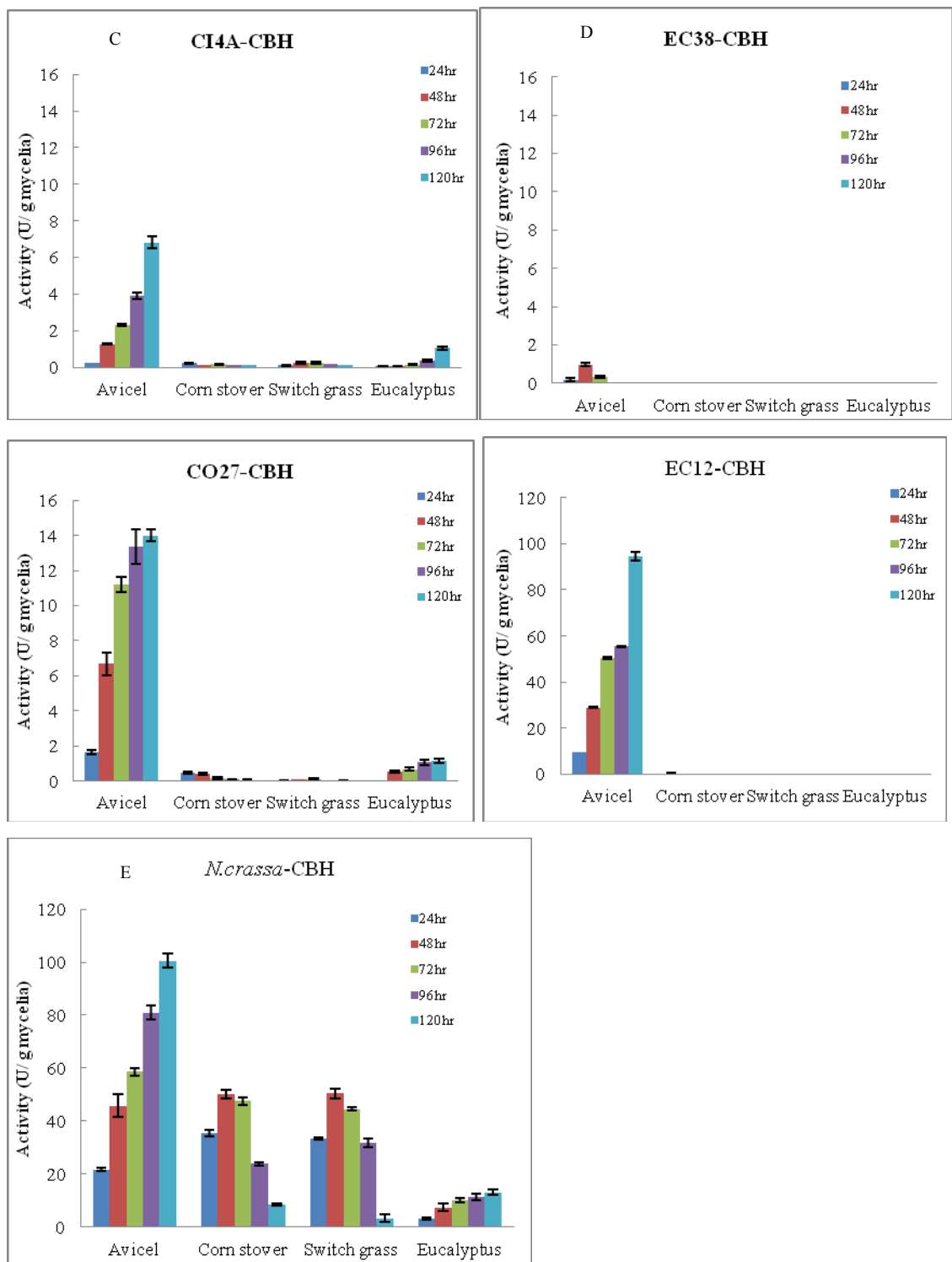


Figure 9. Exoglucanase activities of endophytes and *N. crassa* on different feedstocks

4.4 Endoglucanase activities of endophytes and *N.crassa* on different feedstocks

All endophytes produced significant endoglucanase activities on all four feedstocks, as shown in Figure 11. However, each strain had different endoglucanase expression patterns. Strains EC-38 and CI-4A produced similar levels of endoglucanase activity when grown on eucalyptus. The maximal endoglucanase activity achieved on eucalyptus was 35.85(EC-38) and 36.82(CI-4A) U/g mycelia after 5 days. When EC-38 and CI-4A were grown on Avicel, the highest values of endoglucanase occurred during the midpoint of the culture, with a maximum endoglucanase activity of 32.29 (EC-28 after 2-days) and 47.08 U/ g mycelia (CI-4A after 3-days). CI-4A produced lower endoglucanase activity when it was grown on either switchgrass (5.79 U/ g mycelia after 2-days) or corn stover (7.62 U/ g mycelia after 3-days). Compared to CI-4A, EC-38 produced four times higher endoglucanase activity on switchgrass (20.3 U/ g mycelia) and six times higher endoglucanase activity on corn stover (46.34 U/ g mycelia). CO27-A produced similar levels of endoglucanase activity on switchgrass and corn stover compared to EC-38, but when the strain was grown on Avicel and eucalyptus, it produced four (128.9 U/ g mycelia) and two (67.4 U/ g mycelia) times higher endoglucanase activity, respectively. Compared to EC-38 and CO27-A, strain EC-12 yielded similar levels of endoglucanase activity on switchgrass, corn stover, and eucalyptus. The highest endoglucanase activity achieved by strain EC-12 on Avicel was 3.5 and 14 times higher than that of CO27-A and EC-38, respectively.

N. crassa produced higher endoglucanase activity on all four feedstocks compared to the four endophytes. The highest endoglucanase activity was 1074.88 (Avicel), 528.38 (corn stover), 533.93(switch grass), and 716.05 (eucalyptus) U/ g mycelia.

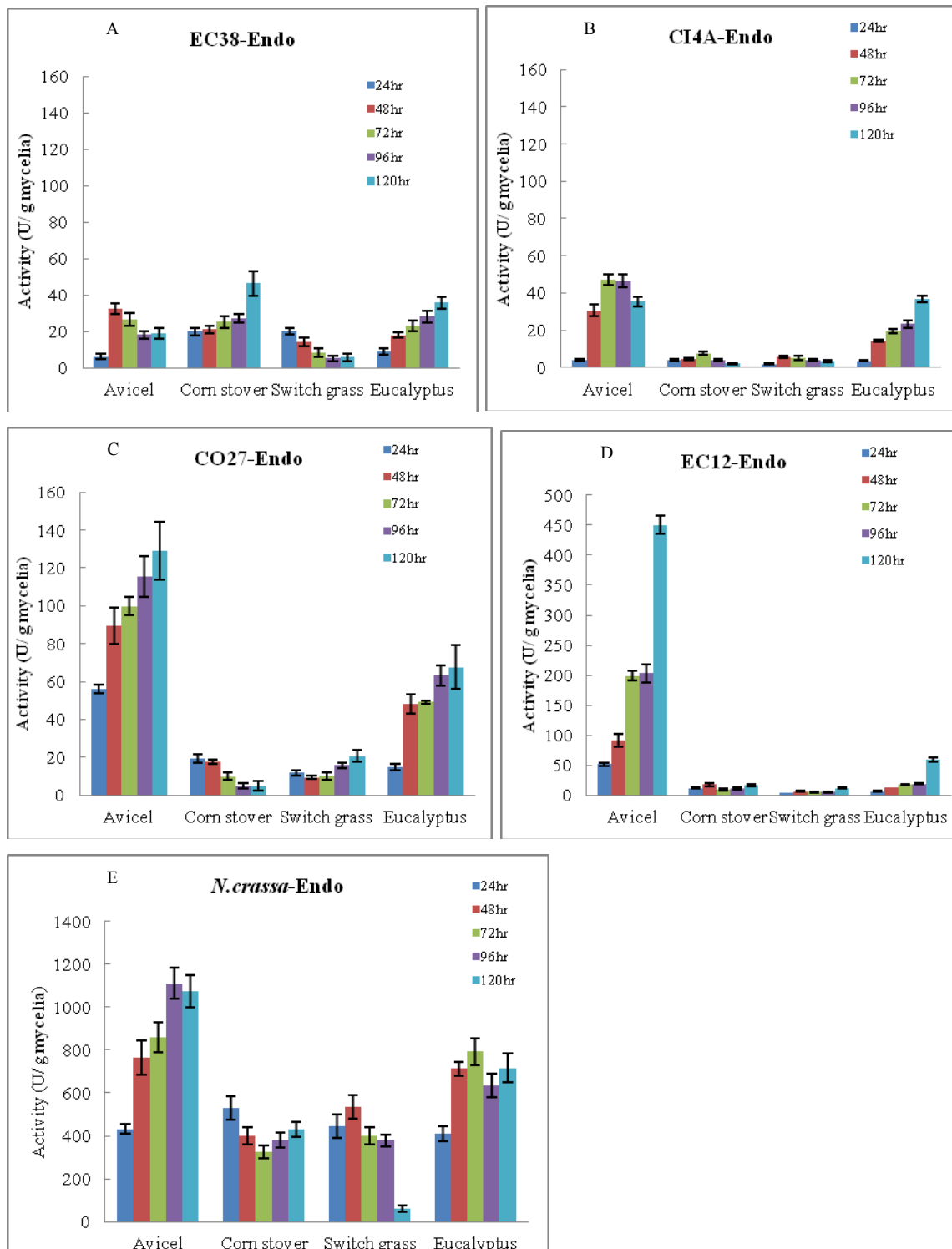


Figure 10. Endoglucanase activities of endophytes and *N. crassa* on different feedstocks

4.5 Zymography assay of secretome

To investigate the complement of endoglucanases produced by each endophytic fungi a zymogram was performed using the soluble cellulose substrate carboxymethylcellulose. The zymogram showed that the major active endoglucanase components of each secretome from the four endophytes range in size between 49KDa and 38KDa, as shown in Figure 11, except for EC-12, which has additional active enzymes between 62KDa and 98KDa. EC-12 appeared to have the highest endoglucanase activity of the four endophytes.

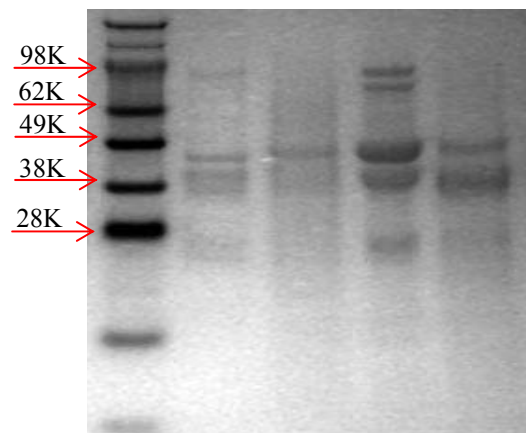


Figure 11. Zymography assay of the endophytic secretome

5. IDENTIFICATION OF PROTEINS IN THE FUNGAL SECRETOME BY MASS SPECTROMETRY

The secretome is typically defined as the combination of native secreted proteins and the cellular machinery that secretes them. Secretome-related studies are particularly relevant in understanding fungal systems, such as the four endophytes described in this project, because many fungi secrete a vast number of proteins to accommodate their saprotrophic lifestyle. Many of these proteins are of special interest in the study of pathogens or during production of recombinant proteins in the biotechnology industry. For our purpose, we sought to characterize the dynamic nature of the cellulolytic machinery of each of these fungi by culturing them on different biomass feedstocks and detecting the secretome using mass spectrometry based proteomics.

While all endophytic fungi produce secretomes, the pattern of components varies greatly depending on feedstock. In Figure 13, we can see certain patterns emerge based on classification of the identified proteins within the secretomes. Examining Avicel, it appears that a large percentage (36-16%) of the secretome has unknown function (purple) compared to the other feedstocks which have proteins of unknown function ranging from 0-13%. This is most likely due to the non-native quality of Avicel, a purified microcrystalline form of cellulose. Owing to the distinctive quality of Avicel, it induces a unique secretome (i.e. no matching proteins) compared the other feedstocks across all four fungal systems. For example, the CI-4A secretomes produced when grown on eucalyptus, potato dextrose, and switchgrass have 11 matching proteins (mostly cellulases) and an additional 7 matching proteins in at least two of the three feedstocks (mostly transporters).

Corn stover also presents with an interesting secretome pattern. For both CI-4A and EC-38, no carbohydrate active enzymes were identified (Figure 13, orange). These results are consistent with the low endoglucanase activity (7.62 U/ g mycelia after 3-days induction, 46.34 U/ g mycelia), CBH activity (6.82 U/ g mycelia, 0.96 U/ g mycelia), and BGL activities observed for CI-4A and EC-38, respectively. While EC-12 and CO27-A produce carbohydrate active enzymes, they also produce a high percentage of redox active proteins with corn stover, commenting on the variations of feed stock composition.

While some proteins were observed being produced across all feedstocks, one enzyme (cutinase) was produced only for eucalyptus and switchgrass and produced only by *Hypoxylon* CI-4A, EC-38. Aerial plant organs are protected by a cuticle composed of an insoluble polymeric structural compound, cutin, which is a polyester composed of hydroxy and hydroxyepoxy fatty acids. Cutinase, which hydrolyses cutin, facilitates fungus penetration through the cuticle and is known to be produced extracellularly by plant pathogenic fungi. Cutin monomers released from the cuticle by small amounts of cutinase on fungal spore surfaces can greatly increase the amount of cutinase secreted by the spore, the mechanism for which process is as yet unknown.

It should be noted that the classification of proteins by function is highly dependent on the quality of the genome annotation. For well characterized protein families, such as CBHs or BGLs, the current annotation methods characterization work very well. However, in the classification of *Hypoxylon* CI-4A, EC-38, CO27-A, and *Daldinia* EC-12 we still have a large number (36% of the secretome in some cases) of proteins with unknown or putative functions, so efforts are being made to improve annotations and assign function to a number of these unknowns. In summary, proteomic analysis of endophyte secretomes highlights the complex nature of the fungal secretome. While we have identified patterns in protein classification and feedstock dependent proteins, the underlying regulatory mechanism, which allows the fungus to identify a particular feedstock and secrete specific protein cocktail, remains elusive.

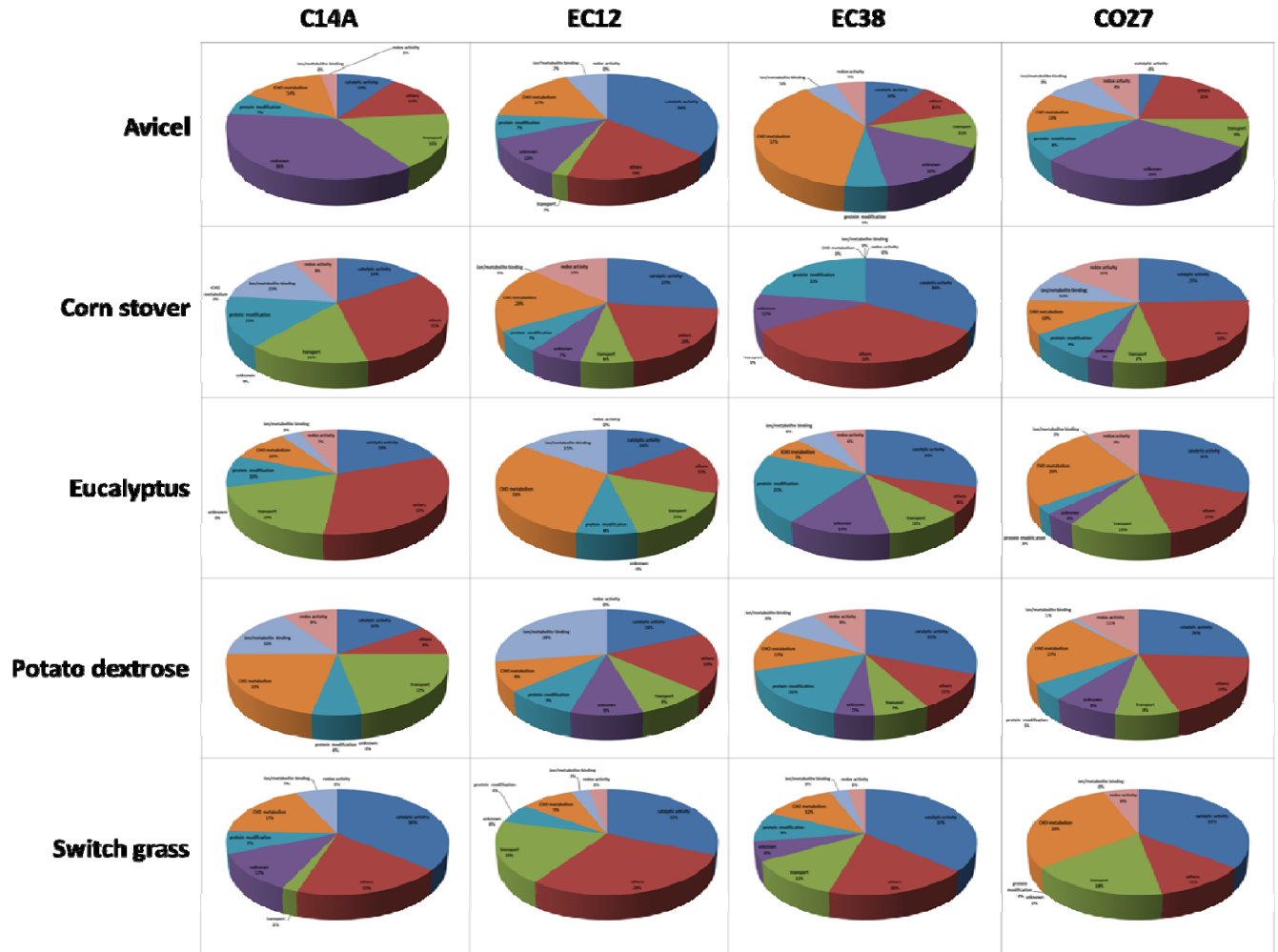


Figure 12. Classification of proteins within the fungal secretome with various feed stocks.

Classes are as follows: Unknown; Protein Modification; Carbohydrate Metabolism; Ion/metal Binding; Redox Activity; Catalytic Activity; Transport; Other

6. COMBUSTION CHEMISTRY OF KETONES AND CINEOLE

The fundamental ignition chemistry of representative compounds was studied using Cl-initiated oxidation and analyzing the product formation kinetics in the critical first steps of oxidation at low temperatures (< 800 K). The low-temperature region, where reactions of peroxy radicals are important, is where fuel-specific effects in autoignition are most prominent. Furthermore, some strategies to extend the load range of HCCI engines rely on the presence of low- and intermediate-temperature (800 – 1100 K) heat release. The low-temperature oxidation reactions are depicted in Figure 13 below.

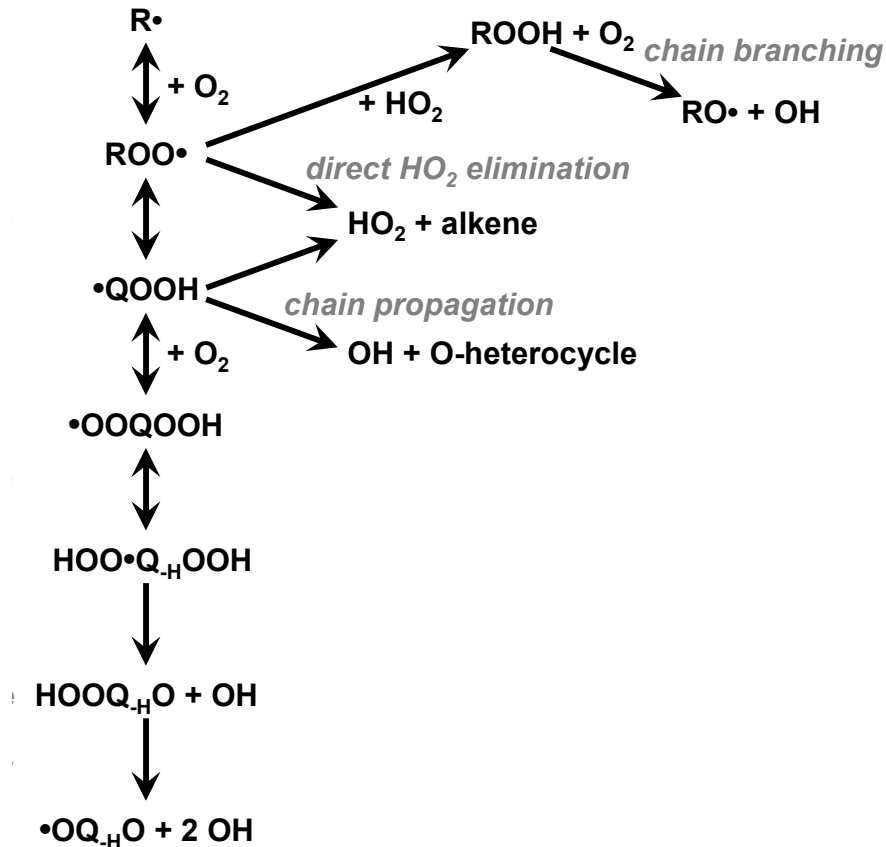


Figure 13. General scheme of the initial steps of low-temperature oxidation

The reaction of the first fuel radical R, formed when the fuel molecule loses a hydrogen atom, with molecular oxygen, O₂, sets the stage for the low-temperature combustion chemistry. The reaction proceeds through a (substituted) alkylperoxy radical ROO, which can isomerize by

internal hydrogen abstraction to a hydroperoxyalkyl radical (QOOH). The formation of HO₂ radical in this reaction system is essentially chain-terminating because of the low reactivity of HO₂. The production of the reactive OH radical is chain-carrying and is a marker for the critical chain-branching chemistry that accompanies the second O₂ addition to QOOH. In mass spectrometry experiments the co-products of HO₂ or OH are detected: these appear at 2 mass units below (for HO₂ formation) or 14 mass units above (for OH formation) the mass of the fuel molecule.

6. 1 Ignition Chemistry of Ketones

Ketones were identified as components in the natural metabolism of several of the endophytes. The carbonyl group in ketones provides vinylic resonance stabilization to radicals on the carbon atoms adjacent to the C=O group, and this resonance stabilization substantially affects the autoignition chemistry.

6.1.1 Open-chain ketones

The oxidation of representative pentanones was carried out to explore the chemical kinetic pathways for open-chain ketones. Rapid isomerization of initial 3-oxoalkyl radical species that can be formed by hydrogen abstraction from the fuel is observed,²⁴ tending to convert less stable primary radicals into secondary or tertiary radicals. The ignition chemistry of di-isopropyl ketone, 2,4-dimethylpentan-3-one, was studied in particular detail, as described in a publication under review.²⁵ That article includes development of an ignition chemistry mechanism by application of the Reaction Mechanism Generation scheme,²⁶⁻³⁰ and validation against pyrolysis and ignition delay measurements carried out by collaborators. The chemistry of pentan-3-one (diethyl ketone) will be described in a paper soon to be submitted.³¹ We describe here the chemistry of two other pentanones, 2,2,4,4-tetramethylpentan-3-one (di-*tert*-butyl ketone, DTbuK) and 2,2,4-trimethylpentan-3-one (isopropyl *tert*-butyl ketone, ITbuK), that show chemical pathways similar to those described in these other publications.

Di-*tert*-butyl ketone oxidation

Abstraction of a hydrogen atom from parent DTbuK (C₉H₁₈O, *m/z* = 142) by Cl will give HCl and a primary radical that can then further decompose unimolecularly or react with O₂. Figure 14

shows the difference mass spectra for the Cl-initiated oxidation of DTbuK at 550 K (bottom) and 700 K (top). These spectra are generated by subtracting the average pre-photolysis signal from the post-photolysis signal. This eliminates constant background peaks and allows a determination of time-resolved products. The bottom panel of Fig. 1 shows peaks at 550 K appear at $m/z = 56$, $m/z = 57$, $m/z = 58$, $m/z = 82$, $m/z = 83$, $m/z = 140$ and several small peaks, including $m/z = 156$.

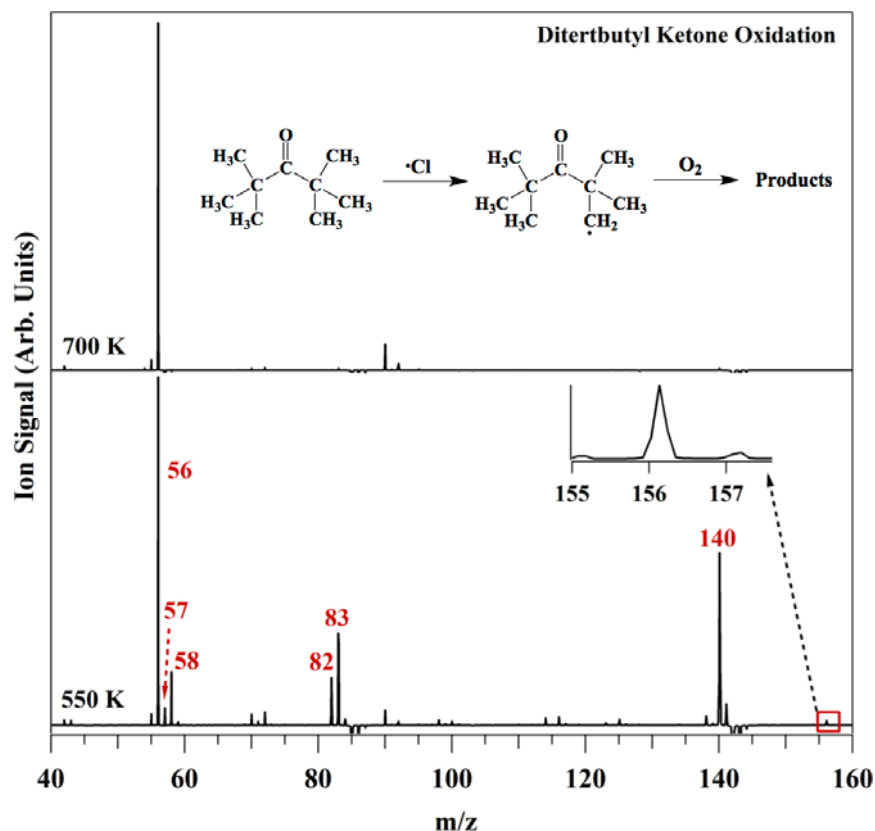
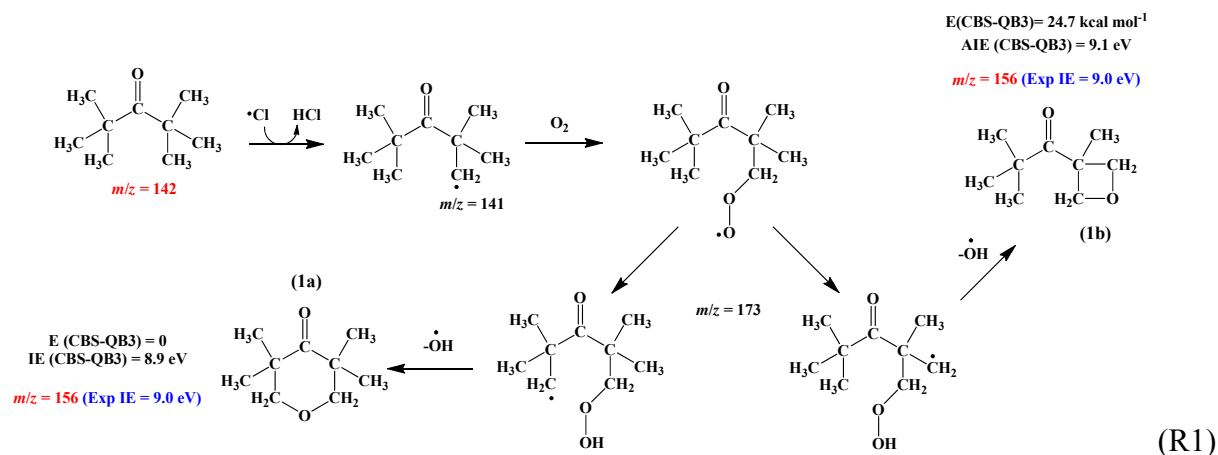


Figure 14. Difference mass spectra of Cl-initiated oxidation of di-*tert*-butyl ketone
 Bottom: 550 K, normalized to photocurrent resulting from integrating the ion signal for the 20 ms timeframe immediately following photolysis and over ionizing photon energies from 8.0-10.5 eV. Averaged pre-photolysis background signal has been subtracted. Top: same at 700 K.

i. Cyclic Ether Pathways

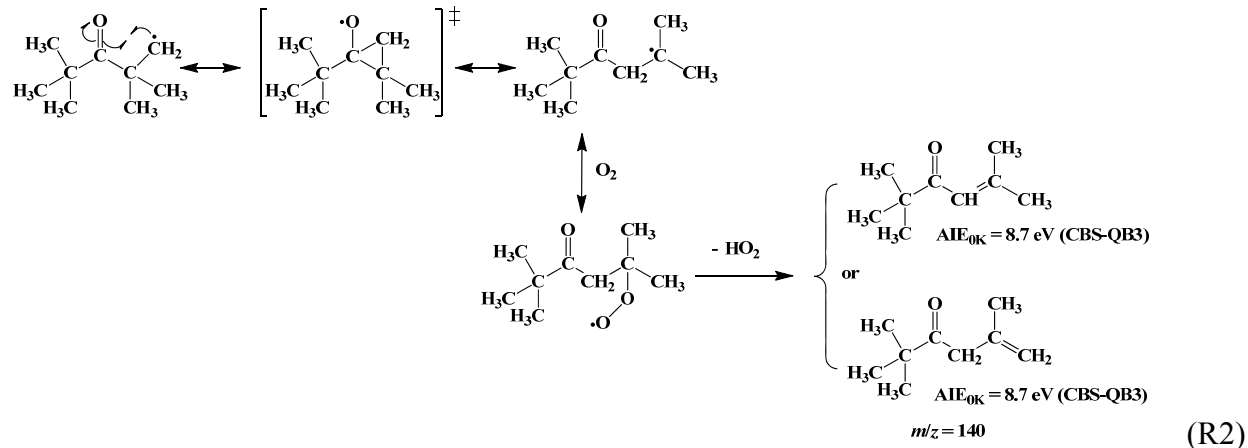
The observation of products at $m/z = 156$ ($C_9H_{18}O_2$) is consistent with formation of a cyclic ether. Reaction R1 shows the H-abstraction process, O_2 addition and two possible channels for QOOH formation followed by decomposition to cyclic ether + OH for DTbuK oxidation. Relative total energies and respective AIEs are given as calculated with the CBS-QB3 method. Both cyclic ether isomers have calculated AIEs near the experimentally observed onset (9.0 eV).



Formation of the 6-membered ring cyclic ether (1a) requires an entropically unfavorable 8-membered ring transition state to form the QOOH intermediate. Although a more favorable 6-membered ring transition state yields the QOOH on the right, the oxetane product (1b) has a total energy nearly 25 kcal mol⁻¹ higher than the 6-membered ring ether. The small signal observed at the parent mass of cyclic ethers in DTbuK oxidation is reflective of these limitations. Analogous channels in the other branched ketones are thus expected to be minor.

ii. HO₂ Elimination

The bottom panel of Figure 145 shows the *m/z* = 140 (C₉H₁₆O) product, corresponding to the stable co-product of HO₂ elimination from the peroxy (RO₂) radical. Though HO₂-elimination pathways are common in standard hydrocarbon oxidation schemes, observation of the HO₂-elimination channel from RO₂ in DTbuK raises an interesting quandary. Because the peroxy radical contains no β-H atoms, the traditional pathway in which an unsaturated product is formed upon elimination of HO₂ is not available. Rather this product arises from the reaction of a rearranged radical:²⁴



With a typical O₂ concentration of $\sim 10^{16}$ cm⁻³ and an R + O₂ second order rate coefficient in the range of $10^{-12} - 10^{-11}$ cm³ s⁻¹,³² the R + O₂ reaction has a pseudo-first order rate coefficient in our experiments on the order of $10^4 - 10^5$ s⁻¹. Using the 3-oxopropyl \rightarrow cyclopropoxy ring-closing reaction as a model system for the acyl group rearrangement in ketones, for which simple RRKM master-equation calculations yield a rate coefficient of $\sim 3 \times 10^5$ s⁻¹ at 8 Torr and 550 K, suggests the acyl group rearrangement is very likely to be important at temperatures of 550 K and higher. The secondary vinylic sites next to the carbonyl group will have lower C-H bond dissociation energies and thus be more susceptible to internal abstraction by the peroxy radical than the primary hydrogens of the terminating methyl groups. Therefore, the top channel of reaction R2 is expected to be dominant, which leads to an α,β -unsaturated ketone with a conjugated π -electron system. The calculated ionization energies for both product isomers are 8.7 eV, in agreement with the photoionization onset observed for the $m/z = 140$ product in the oxidation of DTbuK.

Isopropyl-tert-butyl ketone oxidation

Difference mass spectra for the oxidation of isopropyl-*tert*-butyl ketone at 550 and 700 K are presented in Figure 15. The ITbuK has three types of hydrogen atoms, including nine primary hydrogens on the *tert*-butyl group, six primary hydrogens on the isopropyl group and a single tertiary isopropyl hydrogen. Many of the expected decomposition and oxidation channels are analogous to those found in di-isopropyl ketone²⁵ and di-*tert*-butyl ketone (discussed above). The propene at $m/z = 42$ and isobutene at $m/z = 56$ reflect unimolecular decomposition channels of the initial radicals, which dominate as the temperature is increased from 550 to 700 K. The peaks at $m/z = 142$, $m/z = 114$ and $m/z = 59$ are assigned to cyclic ethers and associated daughter ions. The $m/z = 142$ peak displays an onset at 8.7 eV, which agrees with the calculated value of the 5-membered ring cyclic ether (8.7 eV). Other isomers cannot be excluded based on their ionization energies (Table #0%); however, the heats of formation at 0 K for the 3-membered ring and 4-membered ring cyclic ethers are 25 – 30 kcal mol⁻¹ higher and these isomers are not expected to contribute substantially to the product distribution. Acyl group rearrangement of either primary radical originating could yield a more stable secondary or tertiary radical. Either of these rearranged radical isomers could undergo O₂ addition and yield cyclic ethers, unsaturated products of HO₂ elimination, or small molecule products of β -scission pathways.

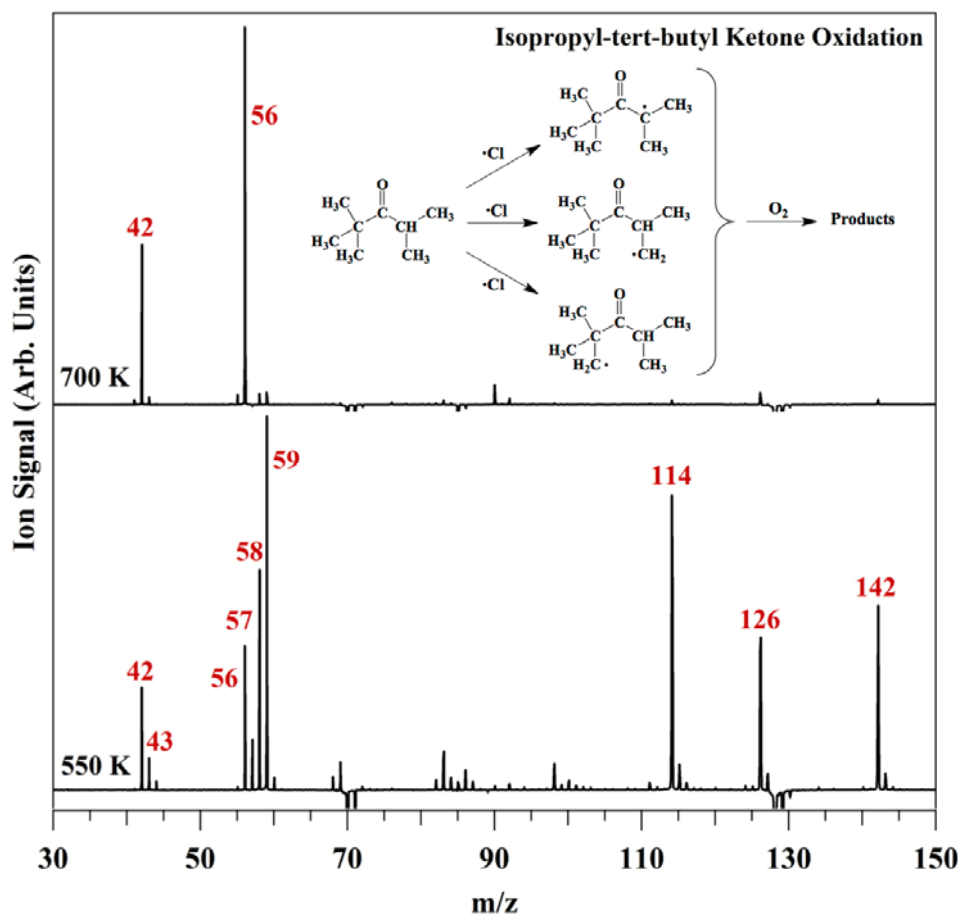
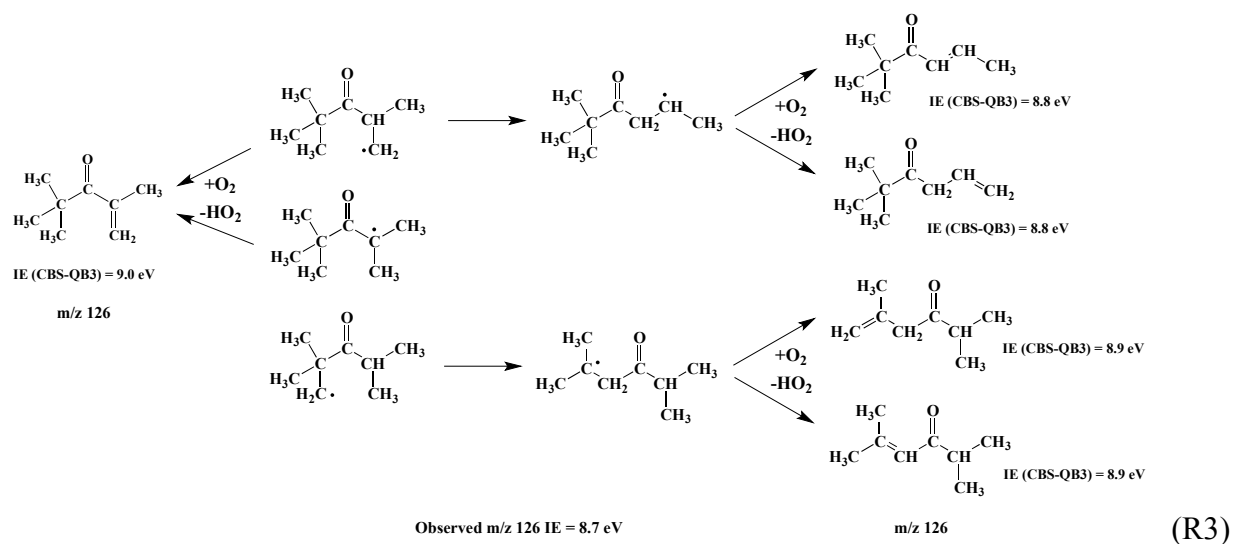


Figure 15. Difference mass spectra of Cl-initiated oxidation of isopropyl-*tert*-butyl ketone Signals normalized to photon flux and integrated for the 20 ms immediately following photolysis and over ionizing photon energies from 8.0 eV – 10.5 eV. Averaged pre-photolysis background signal has been subtracted. Bottom: 550 K; Top: same at 700 K.

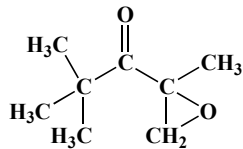
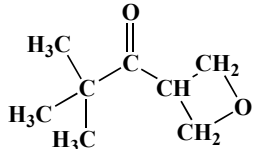
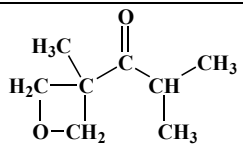
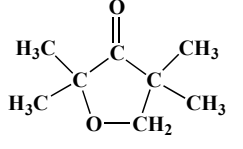
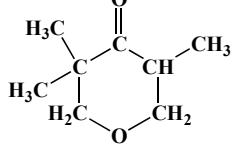
The HO₂-elimination products are again observed, resulting in a strong peak at $m/z = 126$ that displays an onset at 8.7 eV. This value is somewhat lower than the AIE of 9.0 eV calculated for 2,2,4-trimethylpent-1-en-3-one, the expected product of HO₂ elimination assuming no rearrangement of the initial ITbuK radical takes place. Reaction R3 shows this product and the others expected after rearrangements of the initial radical that are analogous to those described above, along with calculated ionization energies.



The calculated AIEs for the products of HO₂ elimination that would be expected from reaction after rearrangement of the primary radical on the isopropyl group match well that observed in experiment (8.7 eV). The calculated AIEs for the HO₂-elimination channel after rearrangement of the primary radical on the *tert*-butyl group and for HO₂ elimination from reaction of the original radical are slightly higher (Table 5) but these products may also contribute.

Table 5. Calculated CBS-QB3 relative energies at 0 K ($\Delta E_{0, \text{rel}}$) and AIEs of cyclic ethers associated with isopropyl-tert-butyl ketone oxidation.

The table does not include cyclic ethers possible after 1,2-acyl group migration from the initial primary radicals.

	$(\Delta E_{0, \text{rel}})$ (kcal mol ⁻¹)	AIE (eV)
	24.7	9.0
	28.0	9.3
	29.5	8.9
	0.0	8.7
	7.0	9.1

The dominance of β -scission pathways at 700 K in ITbuK oxidation can be used to constrain the branching ratio for primary and tertiary radical generation from H-abstraction reactions of Cl in these branched ketones. From fast thermal decomposition reactions, in concert with rapid further reactions involving the small hydrocarbon radical products, one would expect that either primary ITbuK radical would produce CO + propene + isobutene. In contrast, the tertiary radical will yield isobutene + dimethyl ketene. The absolute photoionization cross sections for propene and isobutene are known. Therefore, by taking into account the known mass discrimination factor of the TOF-MS apparatus,³³ an estimate can be made for the primary and tertiary radical formation

branching ratio in isopropyl-*tert*-butyl ketone at 700 K: $k_p/(k_p + k_t) = 0.78$; $k_t/(k_p + k_t) = 0.22$. These branching ratios compare favorably with those determined by Knox and Nelson.³⁴ Using their rate constant determinations for H-abstraction by chlorine atom of the primary and tertiary radicals in isobutane at 298 K, one arrives at an estimate of 0.75 and 0.25, respectively, for the above branching ratios in isopropyl-*tert*-butyl ketone. Note that the higher temperatures of our studies will tend to push the equilibrium toward the less-stable primary radicals. However, the tertiary site in the pentanones studied here has the benefit of vinylic stabilization with the carbonyl group, an effect which will tend to push the equilibrium toward the tertiary radical. Using our observations in isopropyl-*tert*-butyl ketone, we suggest a per hydrogen relative rate factor of 1.0 for primary radical and 4.2 for tertiary radical formation.

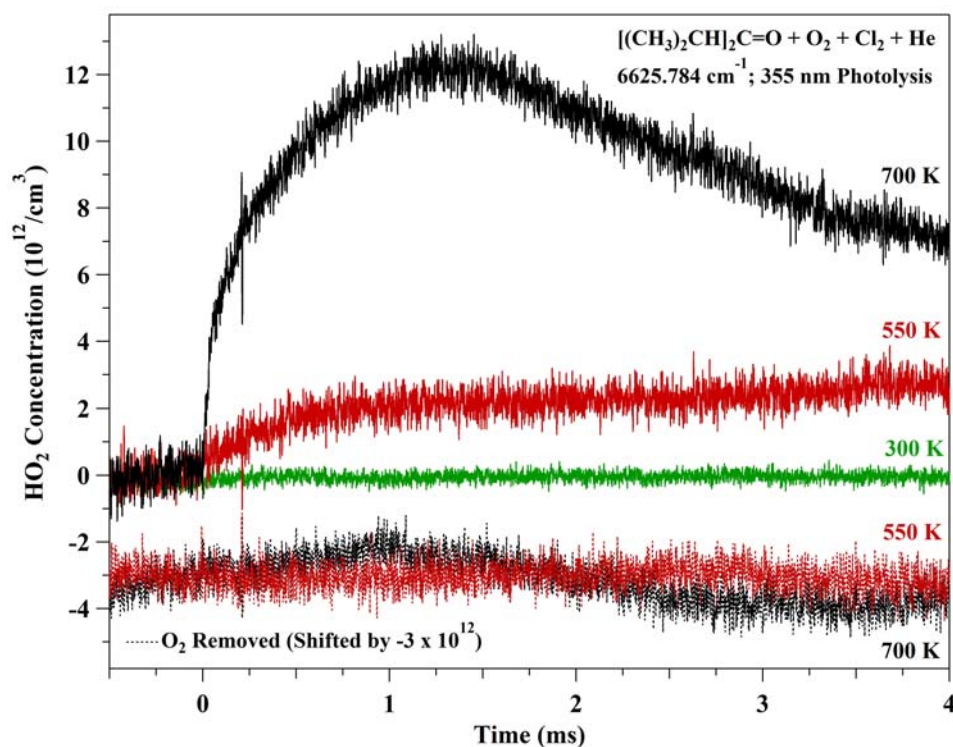


Figure 16. Absolute HO₂ profiles for the Cl-initiated oxidation of diisopropyl ketone. The bottom traces are background signals for 550 K (red) and 700 K (black), (shifted downwards by $3 \times 10^{12} \text{ cm}^{-3}$).

Direct measurement of OH and HO₂ in di-isopropyl ketone oxidation

The results of direct HO₂ absorption studies on di-isopropyl ketone (DIPK) at 20 Torr are shown in Figure 16. The HO₂ band is centered at 6625.784 cm⁻¹. Interference of broad infrared absorptions resulting from other products is removed by tuning the infrared laser slightly off-resonance and subtracting the resulting signal. At room temperature, the HO₂ yield is undetectable. More HO₂ is observed as the temperature is increased, with peak concentrations of approximately 3 × 10¹² cm⁻³ at 550 K and 1.2 × 10¹³ cm⁻³ at 700 K. The bottom curves show the background signal at these temperatures. Although the *m/z* = 112 product of HO₂-elimination clearly diminishes relative to propene as the temperature is increased from 550 K to 700 K, (Fig. 5) the overall HO₂ yield increases dramatically. This is likely due to O₂ reaction with isopropyl radical, which will produce mainly propene and HO₂.³⁵ The isopropyl radical is a major product expected from the unimolecular thermal decomposition of the initial primary DIPK radical (2,4-dimethyl-3-oxopentyl). Figure 17 shows the absolute concentration of OH radical as a function of time after photolysis. At higher temperature, the OH profile displays a much sharper rise and a faster decay. At 550 K peak concentration of 1.2 × 10¹² cm⁻³ is observed after 200 μs; at 650 K, a peak concentration of 1.6 × 10¹² cm⁻³ is observed at a much shorter reaction time, ~ 50 μs.

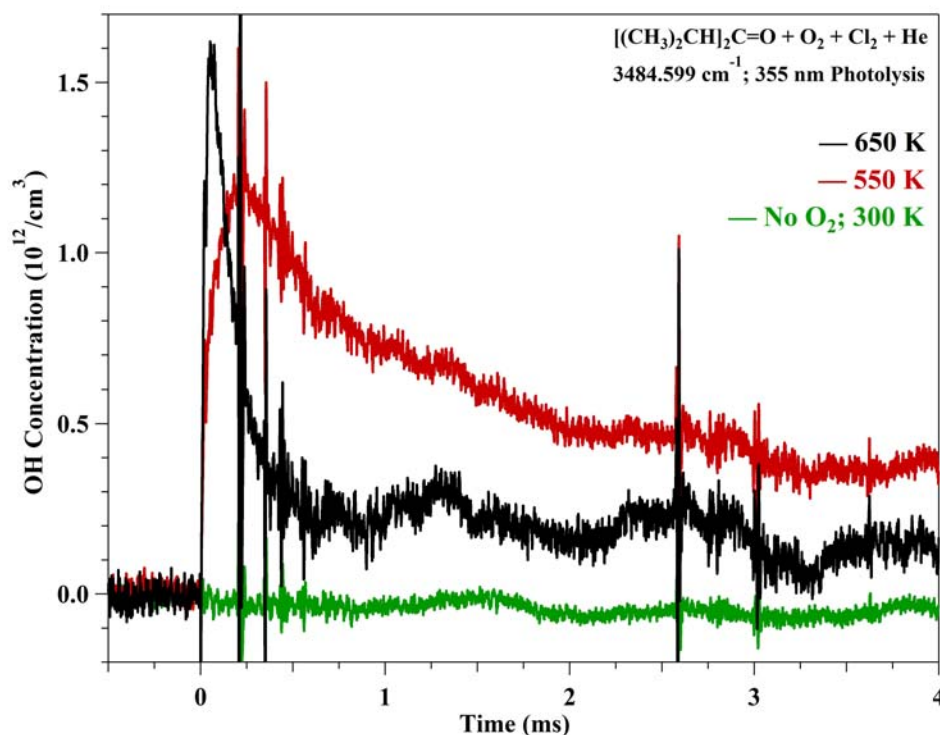
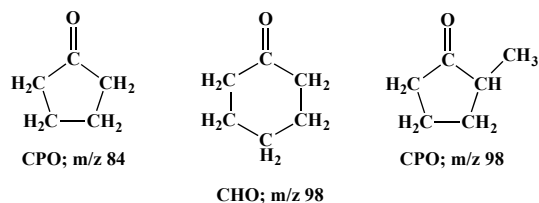


Figure 17. Absolute OH profiles for Cl-initiated diisopropyl ketone oxidation

6.1.2 Cyclic ketones

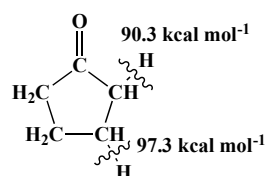
We investigated the oxidation pathways of three prototypical cyclic ketones shown in Scheme I, cyclopentanone (CPO), cyclohexanone (CHO), and 2-methyl-cyclopentanone (2-Me-CPO):



Scheme I

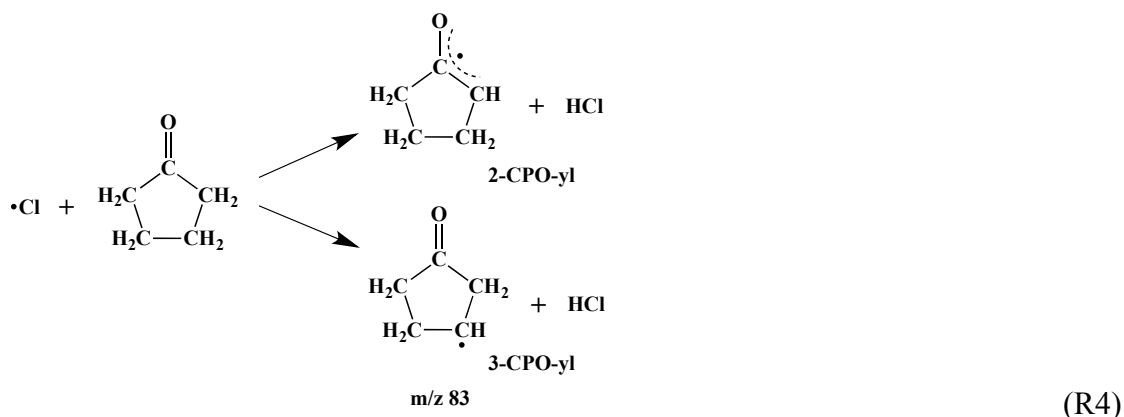
Cyclopentanone (CPO)

We discuss the autoignition of cyclopentanone first, as HCCI engine measurements (see section 6.1.3) have been carried out for CPO. Scheme II shows the calculated C-H bond dissociation energies (BDE_{0K}) for the two types of hydrogens in CPO. For reference, the H-Cl BDE has been tabulated at $102.3 \pm 0.1 \text{ kcal}\cdot\text{mol}^{-1}$.³⁶



Scheme II

Abstracting either hydrogen is exothermic and both radicals are expected in significant concentrations. Reaction R4 shows the radicals expected from Cl-initiated H-abstraction from CPO, 2-oxocyclopentyl (2-CPO-yl) and 3-oxocyclopentyl (3-CPO-yl).



The 2-CPO-yl isomer has a vinyloxy type resonance stabilization. Upon O_2 addition, the resonance stabilization is lost and the RO_2 well associated with the 2-CPO-yl is shallow ($22.2 \text{ kcal mol}^{-1}$). Consequently, equilibrium favors $\text{R} + \text{O}_2$ reactants at elevated temperatures. More importantly,

the shallow well ensures that typical RO₂ exit channels, in particular HO₂ elimination and cyclic ether formation, lie above the entrance channel. Thus high concentrations of 2-oxocyclopentyl are anticipated to persist, leading to self-reaction and reactions with chlorine atoms, OH and HO₂ radicals, and alkylperoxy radicals. In contrast, the RO₂ well associated with 3-oxocyclopentyl radical is calculated to be 34.8 kcal mol⁻¹. A similar situation involving the presence of both non-resonance stabilized and resonantly-stabilized ketone radicals was also investigated for the case of diethyl ketone.³¹

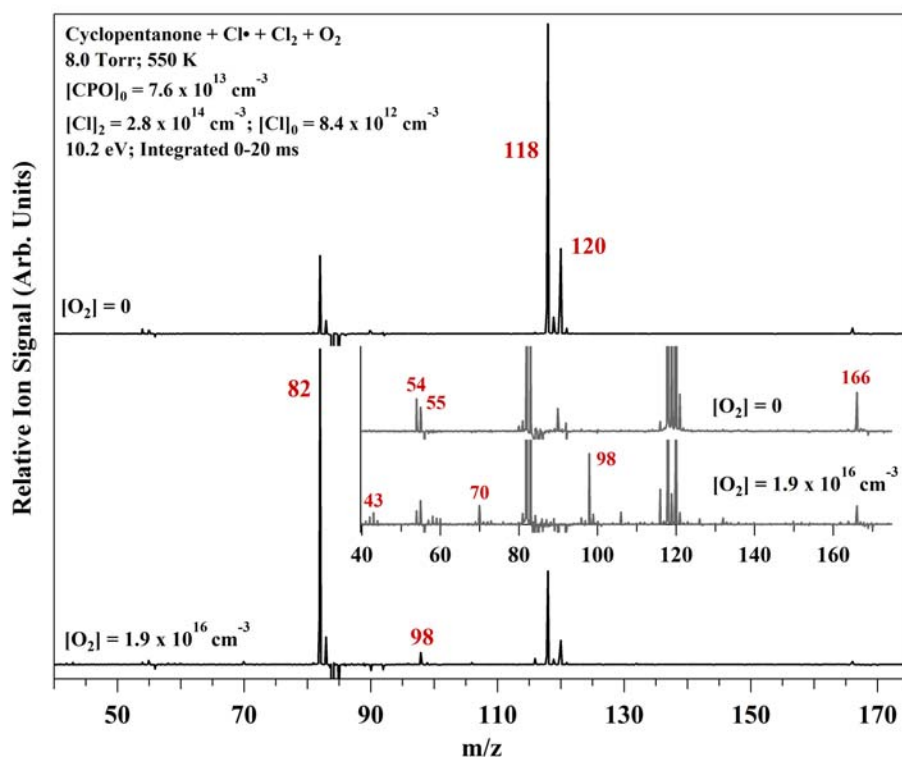


Figure 18. Difference mass spectra at 10.2 eV for the products of Cl-initiated reaction of cyclopentanone

Top: no O₂ added to the reaction mixture. Bottom: Same conditions as in the bottom spectrum but with an O₂ mole fraction of 0.133. Inset: The same product spectra with narrowed y-axis scaling to make minor products visible.

In the presence of O₂, the product at $m/z = 82$ becomes dominant, indicating the formation of RO₂ and subsequent HO₂ elimination.^{32,37} Three such channels are expected in the oxidation of CPO and these are shown in reaction R5. Even in the presence of excess oxygen, chlorinated products are still observed at $m/z = 118$ and 120 and it is thus likely that a small fraction of the $m/z = 82$ peak is due to chlorine side chemistry.

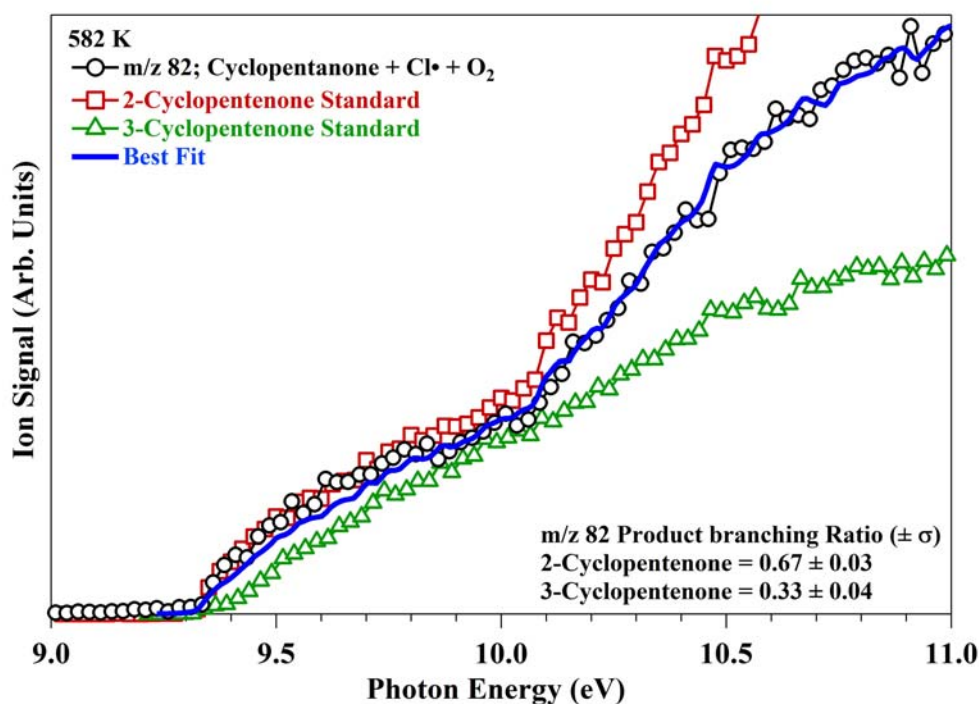
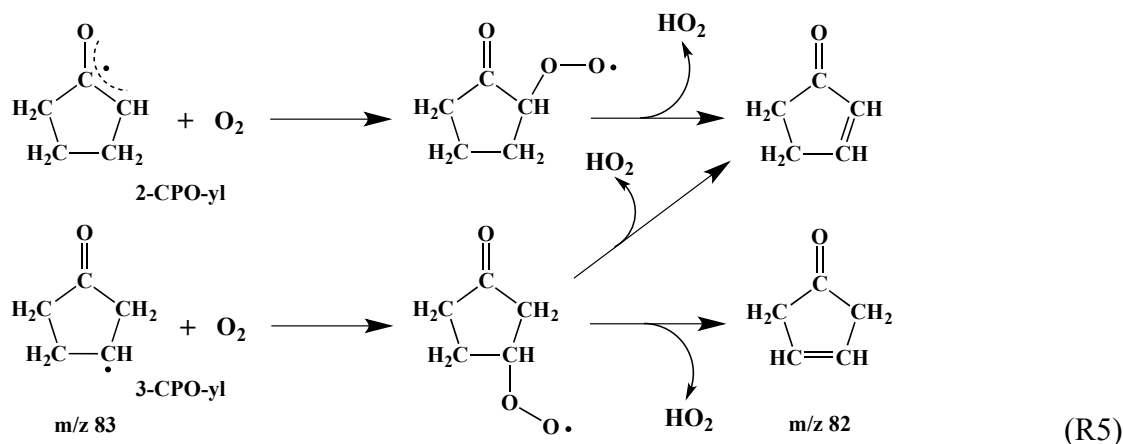


Figure 19. Photoionization spectra for $m/z = 82$ peak observed in 550 K Cl-initiated oxidation of cyclopentanone

The 2-cyclopentenone and 3-cyclopentenone standards were used as basis functions to fit the $m/z = 82$ photoionization spectrum and the solid blue line is the resulting best fit.

Pure samples of 2-cyclopentenone and 3-cyclopentenone were obtained commercially and their photoionization spectra were recorded in the same apparatus described above. By taking the resulting spectra and using them as basis functions to obtain a least-squares best fit of the $m/z = 82$ product photoionization spectrum from CPO oxidation at 550 K, one obtains the results in Figure 19. The fit returns product branching of 0.66 ± 0.03 to the 2-cyclopentenone isomer with

the remainder 3-cyclopentenone. The 2-cyclopentenone standard matches the $m/z = 82$ product of CPO oxidation very well from the onset near 9.3 eV to approximately 10.1 eV where both curves display an inflection and bifurcate. From this point, adding the 3-cyclopentenone basis function brings the best fit into better agreement with the observed product spectrum at higher energies. However, this comes at the expense of the fit at lower energies. Repeating these experiments to clarify these issues would be worthwhile.

Peaks at $m/z = 98$, 70 and 43 appear only in the presence of O_2 (Figure 18 inset, bottom). These signals are assigned to ring closure reactions of RO_2 to form cyclic ethers and OH radicals. A number of possible cyclic ethers are possible from CPO oxidation and these are shown, along with their calculated ionization energies in Table 2. We note that no stable structure was found for the cyclic ether resulting from bridging the 2 and 5 carbons and thus no energy is given.

We have calculated stationary point energies associated with these formation channels along with the HO_2 -elimination pathways discussed above. All unimolecular RO_2 exit channels for the resonance-stabilized 2-CPO-yl have calculated barriers above the entrance channel and therefore these pathways are likely limited. In contrast, the 3-CPO-yl has a deep RO_2 well and several energetically-favorable exit channels. The barrier for production of the conjugated 2-cyclopentenone isomer is significantly lower than that along the 3-cyclopentenone pathway, consistent with the experimental preference for 2-cyclopentenone. Formation of either 3-membered ring cyclic ethers is plausible, but both pathways have $RO_2 \rightarrow QOOH$ barriers of 31 – 32 $\text{kcal}\cdot\text{mol}^{-1}$, approximately 7 $\text{kcal}\cdot\text{mol}^{-1}$ greater than that for the lower-energy direct HO_2 -elimination pathway. These energy differences explain the clear preference for HO_2 elimination over cyclic ether production observed in Figure 18.

Cyclohexanone (CHO)

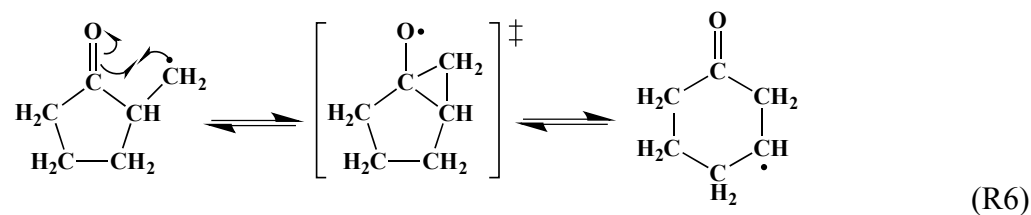
Hydrogen abstraction from CHO will yield three distinct radicals, 2-oxocyclohexyl (2-CHO-yl), 3-oxocyclohexyl (3-CHO-yl), and 4-oxocyclohexyl (4-CHO-yl). As was the case for CPO, the radical with the oxo group in the 2-position is resonance-stabilized. The co-product of HO_2 elimination ($m/z = 96$) again is the dominant signal in the mass spectrum and again two isomers are possible, 2-cyclohexenone and 3-cyclohexenone. The 2-cyclohexenone photoionization spectrum agrees very closely with the observed $m/z = 96$ product, but no commercial sample of 3-cyclohexenone was available and we have not attempted to synthesize it. The adiabatic

ionization energy of 9.09 eV calculated for 3-cyclohexenone is also close to the observed experimental onset for the $m/z = 96$ oxidation product (9.14 eV). Some 3-cyclohexenone product is expected because it represents the only HO₂-elimination channel available from the initial 4-CHO-yl radical. Relatively weak peaks at $m/z = 112$, 84 and 43 are assigned to cyclic ether formation and its daughter ions, in analogy to those seen at $m/z = 98$, 70 and 43 in CPO oxidation.

2-Methyl-Cyclopentanone (2-Me-CPO)

Abstraction of an H atom from 2-Me-CPO can yield five distinct radicals, 1-methyl-2-oxocyclopentyl, 3-methyl-2-oxocyclopentyl, 2-methyl-3-oxocyclopentyl, 3-methyl-4-oxocyclopentyl, and (2-oxocyclopentyl)methyl. The first two isomers listed are resonance-stabilized. As was the case for CPO and CHO, the co-product of HO₂-elimination is the dominant Cl-initiated oxidation product of 2-Me-CPO 550 K. Other peaks observed are assigned in direct analogy to CHO oxidation discussed above.

However, one major caveat must be addressed – a Dowd-Beckwith ring expansion³⁸ reaction of the primary (2-oxocyclopentyl)methyl radical can yield 3-CHO-yl. Such an internal isomerization, shown in reaction R6, is expected to be fast relative to O₂-addition at the conditions of this study (see above).²⁴ Equilibrium is expected to favor the secondary 3-CHO-yl, which will then undergo O₂ addition and subsequent RO₂ chemistry.



Including the Dowd-Beckwith ring expansion but excluding ring opening via direct β -scission, a total of six possible HO₂-elimination co-products are possible. However, calculations suggest that preferred HO₂-elimination channels in 2-Me-CPO oxidation are 2-methyl-2-cyclopentenone from 2-methyl-3-oxocyclopentyl, 2-methyl-4-cyclopentenone from 3-methyl-4-oxocyclopentyl, and 2-cyclohexenone from 3-CHO-yl.

6.1.3 Engine performance

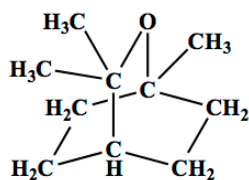
Autoignition characteristics of two representative ketone compounds, 2,4-dimethylpentan-3-one

(DIPK) and cyclopentanone (CPO), were studied in fundamental HCCI engine experiments, and compared with a conventional gasoline and neat ethanol. This work is published as an SAE paper.³⁹ In summary, both ketones showed lower autoignition reactivity and higher temperature sensitivity than gasoline or ethanol under naturally aspirated conditions. CPO shows a very small sensitivity to intake pressure and a very weak intermediate-temperature heat release, and could be useful where resistance to autoignition (i.e. knock-resistance) is important. On the other hand, the autoignition of DIPK is promoted substantially by boost (but less than gasoline), and DIPK shows an unusual low-temperature heat release for intake pressures above 1.8 bar, possibly suggesting unique chemical effects on autoignition. The combination of DIPK's autoignition enhancement with intake-pressure boost and its high temperature sensitivity allows DIPK to provide significant improvements in efficiency over gasoline for high-load intake-boosted HCCI, making DIPK a promising HCCI fuel.

6.2 Ignition Chemistry of Cineole

6.2.1 Fundamental chemistry experiments

Cineole (1,3,3-trimethyl-2-oxabicyclo[2,2,2]octane) is a saturated bicyclic compound with one of the rings containing an ether linkage (Scheme III). Endophytic fungal metabolism of cellulose has been shown to produce significant concentrations of cineole among other complex oxygenates (see Section 2). However, little is known about the fundamental oxidation chemistry governing ignition. Due to the size and complexity of the molecule, many radicals and oxidation pathways will be available in a combustion environment. Here we investigate the initial oxidation reactions of cineole via a combination of experimental and theoretical methods.



Cineole (C₁₀H₁₀O); 154 amu
1,3,3-Trimethyl-2-oxabicyclo[2,2,2]octane

Scheme III

Cineole is a large molecule prone to dissociative photoionization; daughter ions at $m/z = 139$, $m/z = 136$, $m/z = 125$, $m/z = 96$ and $m/z = 84$ are particularly intense. The co products of the chain-propagating OH-elimination channels ($\text{RO}_2 \rightarrow \text{OH} + \text{ring formation co-product}$) are expected at $m/z = 168$. Similarly, the co-products of the chain-terminating HO_2 -elimination channels ($\text{RO}_2 \rightarrow \text{HO}_2 + \text{unsaturated co-product}$) are expected at $m/z = 152$. Smaller products, particularly those resulting from β -scission of the initial radicals, are also likely.

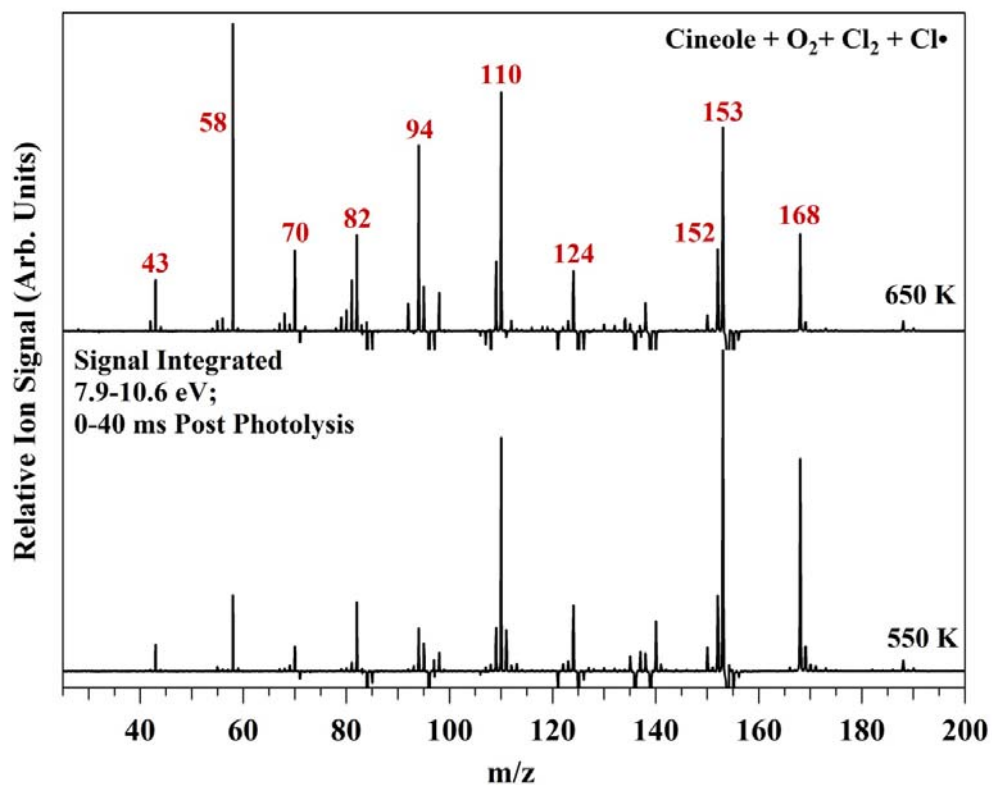


Figure 20. Difference mass spectra of Cl-initiated oxidation of cineole
 Temperature is 550 K (bottom) and 650 K (top); signals normalized to photocurrent and integrated over the 40 ms immediately following photolysis and over ionizing photon energies from 7.9 eV – 10.6 eV.

Figure 20 gives product mass spectra for the Cl-initiated oxidation of cineole at 550 K and 650 K. The spectra result from subtracting the averaged pre-photolysis signal from the post photolysis signal integrated over the first 40 ms of reaction. Negative signal from cineole and its daughters is excluded for clarity. Major peaks are observed at $m/z = 168$, 153, 152, 124, 110, 94, 82, 70, 58 and 43, along with a number of additional minor features.

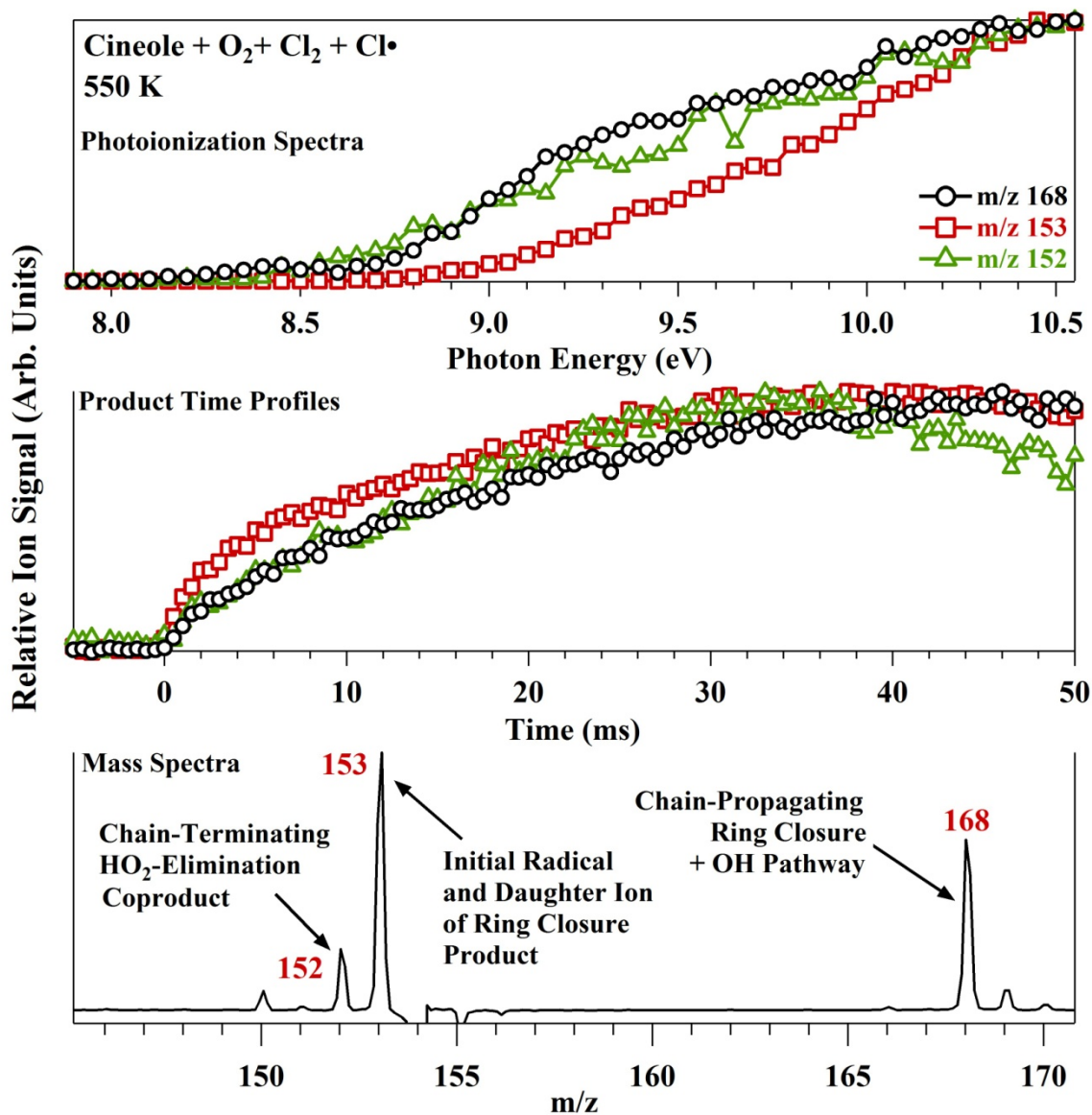


Figure 21. Time, mass, and photon-energy resolved data for Cl-initiated oxidation of cineole at 550 K

Bottom: Mass spectrum showing the coproducts of the chain-propagating and chain-terminating channels associated with the oxidation of cineole radicals. Middle: Product time profiles for $m/z = 168$, $m/z = 153$ and $m/z = 152$ signals. Top: Photoionization spectra for $m/z = 168$, $m/z = 153$ and $m/z = 152$ signals. The three sets of data are acquired simultaneously.

Figure 21 gives an example of the 3-dimensional data collection scheme highlighting signals at $m/z = 168$, $m/z = 153$ and $m/z = 152$. As discussed above, the $m/z = 168$ feature is likely due to one or more ring-closure co-products resulting from OH loss from RO_2 . This signal displays an ionization onset near 8.1 eV (Fig. 7) with additional onset appearing near 8.6 eV, a possible

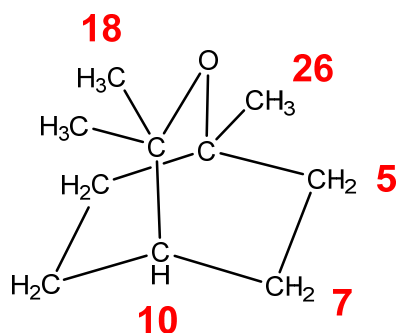
indication of two distinct products. The peak at $m/z = 152$ has an onset at 8.4 eV and is assigned to a unsaturated co-product of HO_2 -elimination. The time profile of the $m/z = 153$ signal displays a sharp rise followed by a gradual increase reflective of a stable product. Both R and RO_2 are formed rapidly but then decompose after a few ms. Therefore, $m/z = 153$ is assigned to a combination of the initial fuel radical, as well as possible daughter ions from RO_2 and daughter ion(s) resulting from loss of a neutral methyl radical from the $m/z = 168$ product(s).

6.2.2 Model generation

Quantum chemistry calculations

Prior to developing a cineole kinetic mechanism, the thermodynamics of key species and kinetics of initial abstraction reactions were calculated using quantum chemistry calculations at the CBS-QB3 level in Gaussian03.⁴⁰ Thermochemistry was computed and fitted to NASA polynomials using CanTherm, a program bundled with the open source software Reaction Mechanism Generator³⁰ for extracting thermodynamics and rate parameters from Gaussian03 log files. Rate coefficients were computed in CanTherm using conventional transition state theory. Tunneling corrections were done using Eckart theory, except for reactions involving the chlorine radical, in which Wigner tunneling was used as a first approximation.

The initial hydrogen abstraction of Cineole can occur at 5 sites, as shown in Scheme IV below. The resulting alkyl radicals are named according to the site at which hydrogen is abstracted. The names and structures of key species can be found in Table 6.



Scheme IV

Hydrogen abstraction reaction kinetics of cineole with chlorine radical and with methyl radical were computed. The rate parameters of all computed reactions can be found in Table 7. The

Table 6. Names and structures of key species in Cineole decomposition

Name	SMILES	Structure
R5	<chem>C1(C2CCC(C)([CH]C2)O1)(C)C</chem>	
R7	<chem>C12(CCC(C(C)(C)O2)[CH]C1)C</chem>	
R10	<chem>C12(CC[C](C(C)(C)O2)CC1)C</chem>	
R18	<chem>C12(CCC(C(C)([CH2])O2)CC1)C</chem>	
R26	<chem>C1(C2CCC(CC2)([CH2])O1)(C)C</chem>	
R5a	<chem>CC1(C)OC2=CCC1CC2</chem>	
R5b	<chem>[CH2]CC1CC=C(C)OC1(C)C</chem>	
R5c	<chem>CC1=CCC(C(C)(C)[O])CC1</chem>	
R5d	<chem>C=CC1(C)CC[CH]C(C)(C)O1</chem>	
R7a (same as R5e)	<chem>CC12C=CC(CC1)C(C)(C)O2</chem>	
R7b	<chem>C=CC1CCCOC1(C)C</chem>	
R7c	<chem>[CH2]CC1(C)CC=CC(C)(C)O1</chem>	
R7d	<chem>CCOC1(C)CC=CCC1</chem>	
R7e	<chem>CC1(C)OC2(C)CC=C1CC2</chem>	

Table 7. Computed rates for Cineole abstraction and decomposition pathways

Reaction	A (mol-cm ³ -s units)	N	E (kcal mol ⁻¹)
Cineole + Cl <=> HCl + R5	9.250e+04	2.555	-9.939
Cineole + Cl <=> HCl + R7	1.513e+05	2.552	-11.253
Cineole + Cl <=> HCl + R10 (estimate based on relative rate of hydrogen abstractions with CH ₃)	3.026e+05	2.552	-11.253
Cineole + Cl <=> HCl + R18	1.605e+05	2.604	-4.231
Cineole + Cl <=> HCl + R26	2.524e+05	2.559	-4.289
Cineole + CH ₃ <=> CH ₄ + R5	5.260e-04	4.661	5.155
Cineole + CH ₃ <=> CH ₄ + R7	1.517e-03	4.623	5.304
Cineole + CH ₃ <=> CH ₄ + R10	3.914e-03	4.580	5.401
Cineole + CH ₃ <=> CH ₄ + R18	1.588e-04	4.913	7.398
Cineole + CH ₃ <=> CH ₄ + R26	1.670e-04	4.898	7.638
R5 <=> R5b	6.140e+11	0.500	25.798
R5 <=> R5c	1.181e+13	0.071	19.363
R7 <=> R7b	1.348e+12	0.323	32.163
R7 <=> R7c	4.251e+11	0.596	28.921
R7 <=> R7d	3.497e+12	0.261	26.670

Mechanism Generation

A preliminary kinetic model for the chlorine-initiated oxidation experiments was generated using the automated mechanism generation software package Reaction Mechanism Generator (RMG),³⁰ using the rates and thermochemistry calculated by quantum chemistry as libraries. Chlorine and HCl thermochemistry values were taken from the Burcat thermochemistry database.⁴¹ The thermochemistry of other species was estimated using the group additivity values from the RMG database,⁴² which consist of mostly Benson-type extrapolations from quantum chemical calculations on small molecules. The thermochemistry of fused cyclic species is not estimated reliably by the Benson groups, so for those species the software automatically spawns semi-empirical PM3

calculations. These calculations are performed using MOPAC2012, RDKit, and SYMMETRY software packages. For cyclic radical species, HBI (hydrogen bond increment) corrections are applied to PM3 calculations for the parent saturated compound.⁴³

RMG automatically constructs a kinetic model using a flux-based algorithm which incorporates the subset of important species and reactions with fluxes above a user's specified tolerance. Two simulated reaction systems were used to generate the model at 550 K and 650 K, both at a pressure of 4 Torr. The mole fractions match those of the experimental conditions and can be found in Table 8.

Table 8. Initial species' mole fractions for reaction systems used in model generation

Temperature (K)	Mole fractions				He
	Cineole	Cl ₂	Cl	O ₂	
550	3.7e-4	8.4e-4	2.565e-5	0.256	0.742
650	3.7e-4	8.41e-4	2.52e-5	0.252	0.746

The model was generated with a tolerance of 0.05, resulting in a mechanism containing 90 species and 819 reactions.

Simulation of the model at ALS conditions generates the mass spectrum shown in Figure 22. by integrating the species' mole fractions over the 40 ms experimental reaction time window. The spectra are very similar for both 550 K and 650 K, with minor differences in integrated mole fractions.

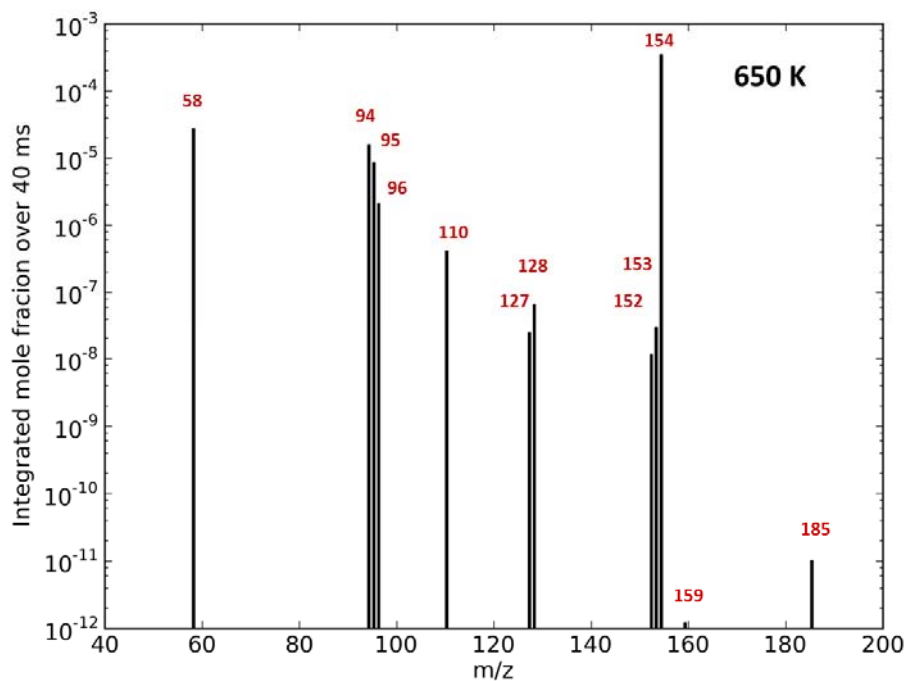
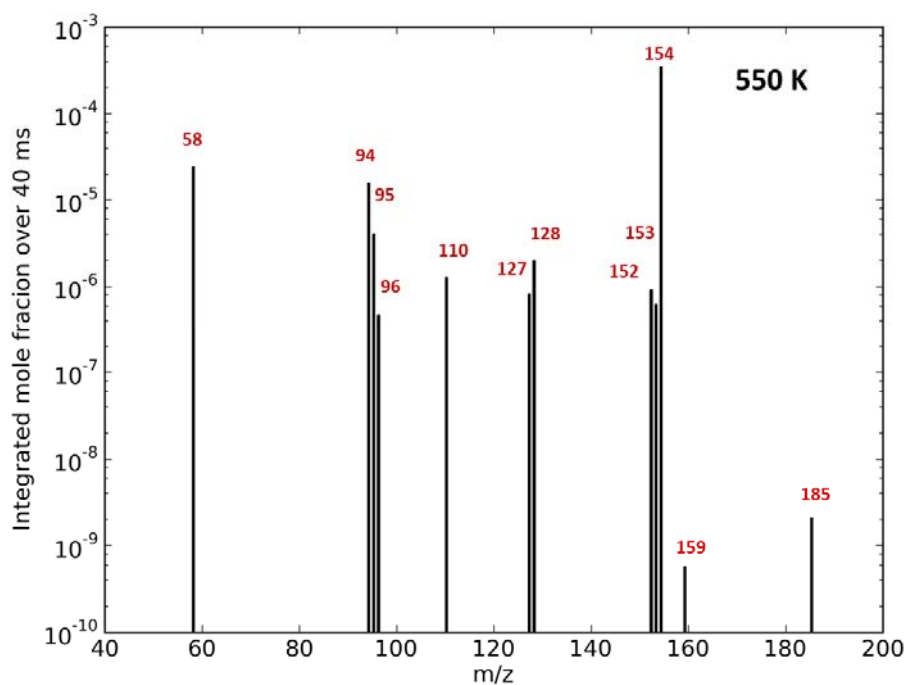
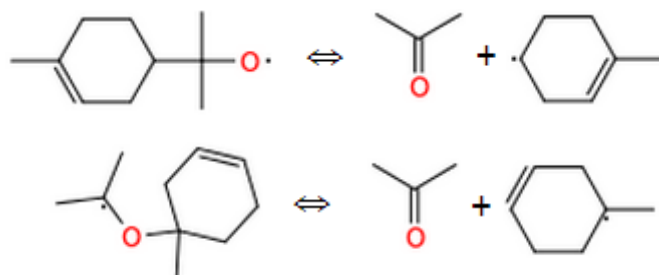


Figure 22. Simulated product spectrum for chlorine-initiated oxidation of cineole
 Temperature is 550 K (top) and 650 K (bottom). Non-carbon containing species are excluded.

The parent cineole molecule gives the main peak at $m/z = 154$. The alkyl radicals from initial chlorine abstraction contribute to $m/z = 153$. The primary contributors to $m/z = 152$ are the R7a and R7e alkene products. The major peak at $m/z = 58$ found in both the model and experiment is acetone, which results from the beta scission of R5c and R7d radicals, shown below:



The co-products of these reactions produce the peak at $m/z = 95$, another peak found in experiment. The HO_2 elimination channels for these products lead to an additional peak at $m/z = 94$. The peak at $m/z = 127$ results from the $m/z = 95$ product adding O_2 .

The predicted peak at $m/z = 185$ corresponds to the ROO radicals associated with the initial alkyl radical + O_2 reactions, which are not expected to be observed experimentally at the parent mass.⁴⁴ However, the $\text{ROO} \leftrightarrow \text{QOOH}$ pathways remain missing in the model, which may be due to the estimated kinetics being too slow and below the threshold flux tolerance set for model generation. Due to this, cyclic ether formation pathways originating from the initial alkyl radical $\text{R} \rightarrow \text{ROO} \rightarrow \text{QOOH} \rightarrow \text{RO} + \text{OH}$ pathways were not incorporated in the model. The cyclic ethers resulting from this pathway lead to an $m/z = 168$ peak found in experiment. Therefore, it is a challenge for RMG to pick up these pathways. The model used to generate the results shown here is a first-guess model built using a relatively loose tolerance. The next steps are to tighten the tolerances (e.g. to pick up the peroxy pathways leading to cyclic ethers) and to refine the sensitive numbers in this model using quantum chemistry.

7. GENOMIC PERSPECTIVES ON “MYCODIESEL” FUNGAL ENDOPHYTES

7. 1 Sequencing and estimated evolutionary relationships of endophytic members of the fungal family Xylariaceae

Draft genomes (~ 90-140x coverage) for four endophytic taxa belonging to the fungal family Xylariaceae were generated to identify genes and biochemical pathways potentially involved in secondary metabolism. Previous work suggested that certain endophytic fungi are potentially useful in producing “drop in” fuel-like precursors by either re-configuring secondary metabolism, and/or porting promising biochemical “parts” to chassis organisms for production of target molecule(s).⁴⁵ First views of these organisms’ genomes show similarity to other sequenced members of the Ascomycota, in terms of genome size, number of genes, gene size and structure (Figure 23)⁴⁶.

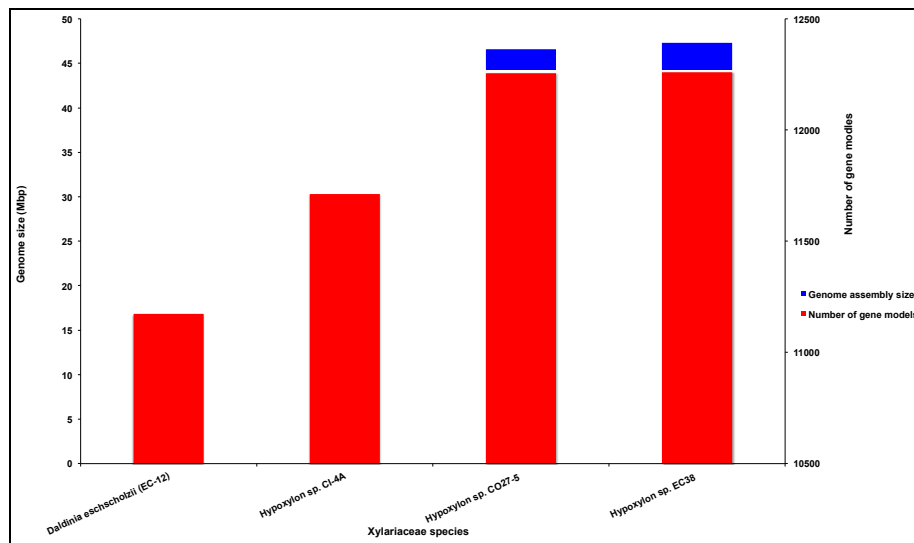


Figure 23. Genome size and number of genes in the first sequenced members of the fungal family Xylariaceae

Draft assemblies and initial gene predictions suggest similarity to other sequenced ascomycetous taxa, possessing ~ 10 thousand genes, and linear estimated genome sizes ~ 40 megabases. Portions of the genome that do not encode recognizable polypeptides are represented in blue. The *Daldinia eschscholzii* genome appears to have a substantially lower density of protein-coding genes than the *Hypoxylon* species.

7.2 Annotating genes and pathways potentially involved in secondary metabolism

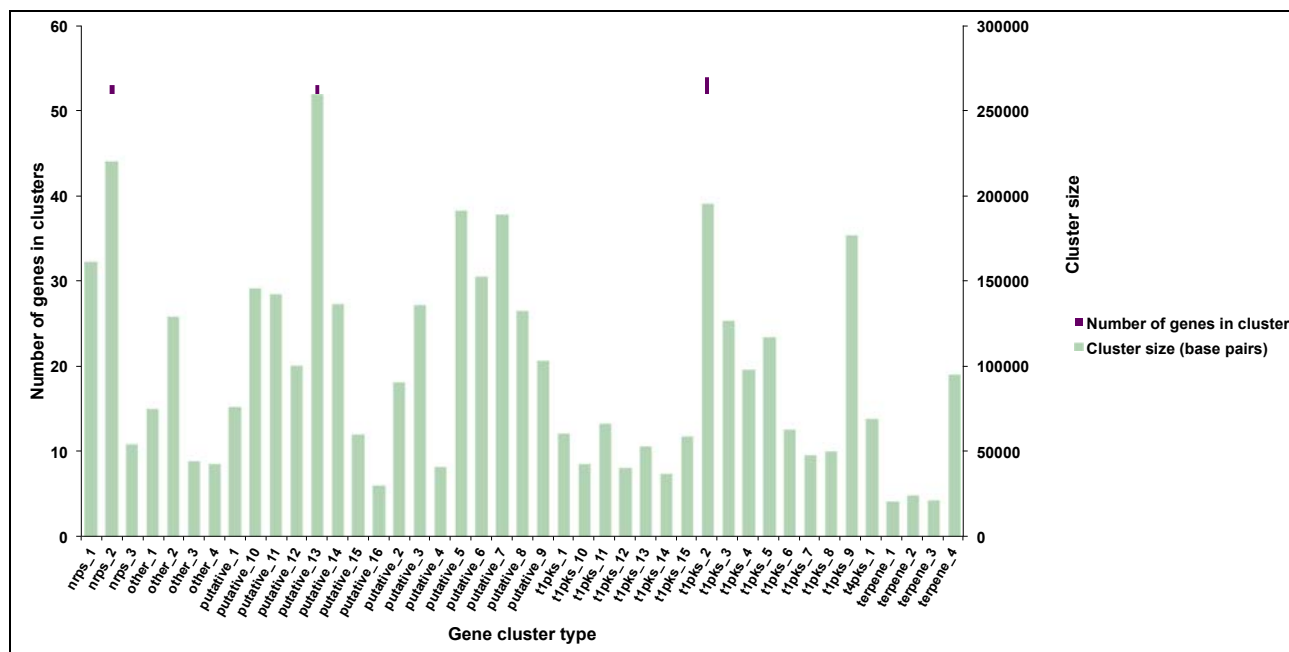


Figure 24. Daldinia eschschlozii secondary metabolite gene cluster size distribution. Forty-three potential secondary metabolic gene clusters were identified. These clusters contain a total of 978 genes, suggesting that ~11% of the genome is in cluster-like architecture. Cluster types include non-ribosomal peptide synthases, terpenes and diverse polyketide synthases. On average, clusters in this organism are 97, 074 basepairs in length, and contain ~23 genes.

Genomic scaffolds (representing draft assemblies) for each of the sequenced organisms were submitted to the antibiotics and Secondary Metabolite Analysis Shell (antiSMASH) using following options: 1) nucleotide input; 2) DNA of Eukaryotic origin; 3) all 24 cluster types; 4) detect putative gene clusters based on PFAM domain probabilities; 5) smCOG analysis for functional prediction and phylogenetic analysis of genes; 6) gene cluster blast analysis; 7) whole-genome PFAM analysis. antiSMASH supports genome-wide identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in microbial (*i.e.*, bacterial and fungal) genomes⁴⁷. Default threshold settings were used for antiSMASH genome scans. Diverse genes and (physically clustered) pathways encoding potential secondary metabolites were identified in each organism (*e.g.*, Figure 24), and annotation associated with these loci intersect with the volatile organic compounds (VOCs) identified in recent high performance liquid chromatography (HPLC) analyses⁴⁸. Output from antiSMASH analyses was further parsed with custom Perl scripts to enumerate potential gene clusters in a genome, and these data were used to

derive summary information about the type, size and number of genes in putative secondary metabolic gene clusters.

Additional resources for annotation of genes and pathways involved in secondary were also preliminarily established. Fasta-formatted flat files were generated for all sequence records in the “Molecular Polyketide Synthase Database,” with a view towards performing HMM-based genome scans to identify remote homologues of polyketide biosynthesis genes⁴⁹. A JGI-developed web-crawler was implemented to extract records from this public database.

In conclusion, the “Mycodiesel” endophytes are the first organisms sequenced from the fungal family Xylariaceae. The species in this family share a recent common ancestor, and are closely-related to the cellulolytic model fungi *Neurospora crassa* and *Trichoderma reesei*. Candidate secondary metabolic pathways were identified in “Mycodiesel” genomes, and include polyketide synthases, non-ribosomal peptide synthases and putative terpene biosynthetic pathways. The candidate pathways potentially encode polypeptides producing the VOCs measured in HPLC analyses. Annotation results obtained for the Xylariaceae taxa suggest potential operon-like architectures for numerous secondary metabolic pathways, a phenomenon that has been observed to varying degrees in all sequenced members of the Ascomycota (pathogens and non-pathogens alike)². Neither the “wet laboratory” investigations nor the *in silico* efforts indicated that the secondary metabolite production potential of the Xylariaceae taxa was necessarily different than other sequenced members of the Pezizomycotina. *Hypoxylon sp.* (EC-38) shows the greatest number of potential secondary metabolite-encoding clusters (65) and *Daldinia eschshlozii* (EC-12) the fewest (43), with the other *Hypoxylon* taxa possessing intermediate numbers [*Hypoxylon sp.* (C14A) has 52 potential secondary metabolic gene clusters, and *Hypoxylon sp.* (CO27-5) 58].

8. CONCLUSIONS

The co-development of biofuels and advanced combustion strategies will be a key tool in building a sustainable and secure transportation infrastructure. This project has created a framework for such a co-development to proceed, focusing on a particular biofuel production platform and a specific kind of advanced engine technology. For substantial impact on the transportation field, a co-development program will need to encompass more potential sources of fuels, more combustion strategies, and more aspects of combustion; however the framework and the strategy from this project are transferable to new production and utilization methods.

Moreover, the project has made new discoveries both in methods of biofuel production that build on metabolic pathways of endophytic fungi and in the chemistry of potential fungal fuels for advanced internal combustion engines.

Consolidating lignocellulosic processing and fuel synthesis capabilities in a single microbe is an ideal platform for biofuel production. *Hypoxyton* and *Daldinia* endophytes have demonstrated the ability to produce hydrocarbons when cultivated in prototypical media. In this study, we have shown that while the endophytes can be cultivated on lignocellulosic biomass, it also altered and diminished their capacity to produce hydrocarbons. This effect is most significant in *Hypoxyton* C1-4A, where 1,8-cineole production was significantly reduced in BF-fed cultures. Further research is needed (e.g. metabolic engineering) to optimize growth in lignocellulosic biomass and modulate VOC production in these endophytic fungi. These findings are important when developing industrial hosts (microbes or fungi) with both biomass deconstruction and fuel synthesis pathways.

“Mycodiesel” endophytes are the first organisms sequenced from the fungal family Xylariaceae. The species in this family share a recent common ancestor, and are closely-related to the cellulolytic model fungi *Neurospora crassa* and *Trichoderma reesei*. Candidate secondary metabolic pathways were identified in “Mycodiesel” endophyte genomes, and include polyketide synthases, non-ribosomal peptide synthases and terpene biosynthesis loci. The candidate pathways potentially encode polypeptides producing the VOCs measured by HPLC. Annotation results obtained for the Xylariaceae taxa suggest potential operon-like architectures for numerous secondary metabolic pathways, which has been observed to some degree in all sequenced

members of the Ascomycota (pathogens and non-pathogens alike). The secondary metabolite production potential of the Xylariaceae taxa is not dissimilar to that of other sequenced members of the Pezizomycotina. *Hypoxylon sp.* (EC-38) shows the greatest number of potential secondary metabolite-encoding clusters (65) and *Daldinia eschshlozii* (EC-12) the fewest (43), with the other *Hypoxylon* taxa possessing intermediate [*Hypoxylon sp.* (CI4A) has 52 potential secondary metabolic gene clusters, and *Hypoxylon sp.* (CO27-5) has 58.

The sequencing of these endophytes has enabled the identification and characterization of some of the VOC pathways in these organisms, and has laid the groundwork for an iterative loop between the biology of fuel production and fuel combustion chemistry. In this project, the monoterpene cineole was identified in the VOC complement of several of the endophytic fungi screened, and an examination of its fuel chemistry suggested that it's a potentially viable biofuel. Therefore efforts on the biology side were focused on characterizing the terpene biosynthesis pathway of these fungi and porting it into a biofuel chassis. The terminal enzyme in the terpene biosynthesis pathway is the terpene synthase enzyme, which produces the unique complement of terpenes produced by any given organism. When examining the four endophytic fungi, a total of twenty five potential TS were identified. Each TS was screened and they were determined to yield a wide range of terpenes, including cineole produced by a single enzyme.

We have investigated the combustion chemistry of the fuels that could be produced by endophytic fungal metabolism of lignocellulosic biomass. Ketones and cyclic ethers were identified in the VOC product spectrum from endophytes grown on a range of substrates. Fundamental chemistry studies and engine measurements on ketones uncovered some unusual and beneficial characteristics in their ignition chemistry. The cyclic ketones, exemplified by cyclopentanone (CPO), are resistant to autoignition and show low reactivity in HCCI engines. The fundamental chemistry experiments show that the initial steps of low-temperature cyclopentanone oxidation are dominated by chain-terminating formation of HO₂. On the other hand, low-temperature oxidation of open-chain ketones such as di-isopropyl ketone (DIPK) displays significant chain-propagating formation of OH radicals, and DIPK shows low-temperature heat release in HCCI operation above 1.8 bar intake pressure, and, unusually, retains significant sensitivity to temperature even under highly boosted conditions. Consequently, DIPK is a promising fuel for high-load HCCI and CPO may be a useful as a knock-resistant spark-

ignition (SI) fuel. Future work may elucidate the chemical reasons for the unusual pressure and temperature dependence of the DIPK autoignition and further explore the possible uses of cyclic ketones in SI engines.

Cineole, another potential biofuel among the fungal metabolism products, is a saturated bicyclic ether. Our initial fundamental chemistry studies suggest significant formation of OH that indicates the possibility of low-temperature heat release in compression-ignition engines, although this has not yet been directly investigated experimentally. Further exploration of cineole production and combustion would be one logical area for continuation of the co-development framework that this project has established.

9. REFERENCES

- 1 Energy Information Administration. *International Energy Outlook*. (2011).
- 2 Chu, S. & Majumdar, A. Opportunities and challenges for a sustainable energy future. *Nature* **488**, 294-303, doi:10.1038/nature11475 (2012).
- 3 Peralta-Yahya, P. P., Zhang, F., Cardayre, S. B. d. & Keasling, J. D. Microbial engineering for the production of advanced biofuels. *Nature* **488**, 320-328, doi:10.1038/nature11478 (2012).
- 4 Manley, D. K., McIlroy, A. & Taatjes, C. A. Research needs for future internal combustion engines. *Phys. Today* **61**, 47-52, doi:10.1063/1.3027991 (2008).
- 5 Kohse-Höinghaus, K. *et al.* Biofuel Combustion Chemistry: From Ethanol to Biodiesel. *Angew. Chem. Int. Ed.* **49**, 3572-3597, doi:10.1002/anie.200905335 (2010).
- 6 Singh, S. K. *et al.* An Endophytic *Phomopsis* sp. Possessing Bioactivity and Fuel Potential with its Volatile Organic Compounds. *Microb. Ecol.* **61**, 729-739 (2011).
- 7 Strobel, G. A. *et al.* The production of myco-diesel hydrocarbons and their derivatives by the endophytic fungus *Gliocladium roseum* (NRRL 50072). *Microbiol-Sgm* **154**, 3319-3328, doi:10.1099/mic.0.2008/022186-0 (2008).
- 8 Strobel, G. A. *et al.* Synergism among volatile organic compounds resulting in increased antibiosis in *Oidium* sp. *Fems Microbiol Lett* **283**, 140-145, doi:10.1111/j.1574-6968.2008.01137.x (2008).
- 9 Mends, M. T. *et al.* An Endophytic *Nodulisporium* sp. Producing Volatile Organic Compounds Having Bioactivity and Fuel Potential. *J. Pet. Environ. Biotechnol.* **3**, 117, doi:10.4172/2157-7463.1000117 (2012).
- 10 Connor, M. R. & Liao, J. C. Engineering of an *Escherichia coli* strain for the production of 3-methyl-1-butanol. *Appl Environ Microbiol* **74**, 5769-5775, doi:10.1128/AEM.00468-08 (2008).
- 11 Fischer, C. R., Klein-Marcuschamer, D. & Stephanopoulos, G. Selection and optimization of microbial hosts for biofuels production. *Metab. Eng.* **10**, 295-304, doi:10.1016/j.ymben.2008.06.009 (2008).
- 12 Larsson, S., Cassland, P. & Jonsson, L. J. Development of a *Saccharomyces cerevisiae* strain with enhanced resistance to phenolic fermentation inhibitors in lignocellulose hydrolysates by heterologous expression of laccase. *Appl. Environ. Microbiol.* **67**, 1163-1170, doi:10.1128/AEM.67.3.1163-1170.2001 (2001).
- 13 Peralta-Yahya, P. P. *et al.* Identification and microbial production of a terpene-based advanced biofuel. *Nat Commun* **2**, doi:10.1038/Ncomms1494 (2011).
- 14 Strobel, G. *et al.* An endophytic/pathogenic *Phoma* sp. from creosote bush producing biologically active volatile compounds having fuel potential. *FEMS Microbiol. Lett.* **320** 87-94 (2011).
- 15 Tomscheck, A. R. *et al.* *Hypoxylon* sp., an Endophyte of *Persea indica*, Producing 1,8-Cineole and Other Bioactive Volatiles with Fuel Potential. *Microb Ecol* **60**, 903-914, doi:10.1007/s00248-010-9759-6 (2010).
- 16 Griffin, M. A., Spakowicz, D. J., Gianoulis, T. A. & Strobel, S. A. Volatile organic compound production by organisms in the genus *Ascocoryne* and a re-evaluation of myco-diesel production by NRRL 50072. *Microbiology* **156**, 3814-3829, doi:10.1099/mic.0.041327-0 (2010).

- 17 Jovanovic Tews, I. *et al.* A Survey of Opportunities for Microbial Conversion of Biomass to Hydrocarbon Compatible Fuels. (Pacific Northwest National Laboratory, Richland, WA, 2010).
- 18 Kumar, P., Barrett, D. M., Delwiche, M. J. & Stroeve, P. Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production. *I&E Chem. Res.* **48**, 3713-3729, doi:10.1021/ie801542g (2009).
- 19 Keller, F. A., Hamilton, J. E. & Nguyen, Q. A. Microbial pretreatment of biomass: potential for reducing severity of thermochemical biomass pretreatment. *Appl Biochem Biotechnol* **105 -108**, 27-41 (2003).
- 20 Sarwar, M., Kirkegaard, J. A., Wong, P. T. W. & Desmarchelier, J. M. Biofumigation potential of brassicas - III. In vitro toxicity of isothiocyanates to soil-borne fungal pathogens. *Plant Soil* **201**, 103-112 (1998).
- 21 Parthiban, P., Kabilan, S., Ramkumar, V. & Jeong, Y. T. Stereocontrolled facile synthesis and antimicrobial activity of oximes and oxime ethers of diversely substituted bispidines. *Bioorg. Med. Chem. Lett.* **20**, 6452-6458, doi:10.1016/j.bmcl.2010.09.079 (2010).
- 22 Graus, M. *et al.* in *American Geophysical Union, Fall Meeting 2011* (2011).
- 23 Pareek, S., Azuma, J. I., Matsui, S. & Shimizu, Y. Degradation of lignin and lignin model compound under sulfate reducing condition. *Water Sci Technol* **44**, 351-358 (2001).
- 24 Scheer, A. M., Welz, O., Sasaki, D. Y., Osborn, D. L. & Taatjes, C. A. Facile Rearrangement of 3-Oxoalkyl Radicals is Evident in Low-Temperature Gas-Phase Oxidation of Ketones. *J. Am. Chem. Soc.* **135**, 14256-14265, doi:10.1021/ja405892y (2013).
- 25 Allen, J. W. *et al.* A Coordinated Investigation of the Combustion Chemistry of Diisopropyl Ketone, a Prototype for Biofuels Produced by Endophytic Fungi. *Combust. Flame* **in press** doi:10.1016/j.combustflame.2013.10.1019 (2013).
- 26 Van Geem, K. M. *et al.* Automatic reaction network generation using RMG for steam cracking of n-hexane *AIChE Journal* **52**, 718-730 (2006).
- 27 Green, W. H. *et al.* Computer Construction of Detailed Chemical Kinetic Models for Gas-Phase Reactors. *Ind. Eng. Chem. Res.* **40**, 5362-5370 (2001).
- 28 Matheu, D. M., Green Jr., W. H. & Grenda, J. M. Capturing Pressure-Dependence in Automated Mechanism Generation: Reactions Through Cycloalkyl Intermediates. *Int. J. Chem. Kinet.* **35**, 95-119 (2003).
- 29 Susnow, R. G., Dean, A. M., Green, W. H., Peczak, P. & Broadbelt, L. J. Rate-Based Construction of Kinetic Models for Complex Systems. *J. Phys. Chem. A* **101**, 3731-3740 (2001).
- 30 W. H. Green *et al.* *Reaction Mechanism Generator (RMG-Py)*, <<https://github.com/GreenGroup/RMG-Py>> (2013).
- 31 Scheer, A. M., Welz, O., Zador, J., Osborn, D. L. & Taatjes, C. A. Low-Temperature Combustion Chemistry of Novel Biofuels: Resonance-Stabilized QOOH in the Chlorine-Initiated Oxidation of Diethyl Ketone. *In Preparation* (2013).
- 32 Zádor, J., Taatjes, C. A. & Fernandes, R. X. Kinetics of elementary reactions in low-temperature autoignition chemistry. *Prog. Energy Combust. Sci.* **37**, 371-421, doi:10.1016/j.peccs.2010.06.006 (2011).
- 33 Welz, O. *et al.* Low-temperature combustion chemistry of biofuels: pathways in the initial low-temperature (550 K-750 K) oxidation chemistry of isopentanol. *Phys. Chem. Chem. Phys.* **14**, 3112-3127, doi:10.1039/c2cp23248k (2012).

- 34 Karl, C. L., Maas, E. J. & Reusch, W. Acyl rearrangements in radical reactions. *J. Org. Chem.* **37**, 2834-2840, doi:10.1021/jo00983a009 (1972).
- 35 Goldsmith, C. F., Green, W. H. & Klippenstein, S. J. Role of O₂ + QOOH in Low-Temperature Ignition of Propane. 1. Temperature and Pressure Dependent Rate Coefficients. *J. Phys. Chem. A* **116**, 3325-3346, doi:10.1021/jp210722w (2012).
- 36 Darwent, B. D. Bond Dissociation Energies in Simple Molecules: B. deB. Darwent. (U.S. National Bureau of Standards, 1970).
- 37 Taatjes, C. A. Recent Developments in the Coupling of Theory and Experiment to Study the Elementary Chemistry of Autoignition. *Journal of the Combustion Society of Japan* **50**, 29-38 (2008).
- 38 Wang, Z. in *Comprehensive Organic Name Reactions and Reagents* Ch. 201, 939-941 (John Wiley and Sons, 2010).
- 39 Dec, J. E. & Yang, Y. Bio-Ketones: Autoignition Characteristics and Their Potential as Fuels for HCCI Engines. *SAE paper* **2013-01-2627** (2013).
- 40 Gaussian 03 (2004).
- 41 Burcat, A. & Ruscic, B. Third Millennium Ideal Gas and Condensed Phase Thermochemical Database for Combustion with Updates from Active Thermochemical Tables. (2005).
- 42 W. H. Green et al. *RMG-database*, <<https://github.com/GreenGroup/RMG-database>> (2013).
- 43 Lay, T. H., Bozzelli, J. W., Dean, A. M. & Ritter, E. R. Hydrogen-Atom Bond Increments for Calculation of Thermodynamic Properties of Hydrocarbon Radical Species. *J. Phys. Chem.* **99**, 14514-14527 (1995).
- 44 Meloni, G. *et al.* Energy-resolved photoionization of alkyl peroxy radicals and the stability of their cations. *J. Am. Chem. Soc.* **128**, 13559-13567 (2006).
- 45 Gianoulis, T. A. *et al.* Genomic Analysis of the Hydrocarbon-Producing, Cellulolytic, Endophytic Fungus *Ascocoryne sarcoides*. *PLoS Genet* **8**, e1002558, doi:10.1371/journal.pgen.1002558 (2012).
- 46 Galagan, J. E., Henn, M. R., Ma, L.-J., Cuomo, C. A. & Birren, B. Genomics of the fungal kingdom: Insights into eukaryotic biology. *Genome Research* **15**, 1620-1631, doi:10.1101/gr.3767105 (2005).
- 47 Medema, M. H. *et al.* antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Research* **39**, W339-W346, doi:10.1093/nar/gkr466 (2011).
- 48 Gladden, J. M., Powell, A. J., Yu, E., Hadi, M. Z. & Taatjes, C. A. Understanding secondary metabolic potential in endophytic members of the fungal family Xylariaceae. *in preparation* (2013).
- 49 Tae, H., Kong, E.-B. & Park, K. ASMPKS: an analysis system for modular polyketide synthases. *BMC Bioinformatics* **8**, 327 (2007).
- 50 Stein, S. E. An Integrated Method for Spectrum Extraction and Compound Identification from GC/MS Data. *J. Am. Soc. Mass Spectrom.* **10**, 770-781 (1999).
- 51 Sørensen, T. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Kongelige Danske Videnskabernes Selskab* **5**, 1-34 (1957).

- 52 Osborn, D. L. *et al.* The multiplexed chemical kinetic photoionization mass spectrometer: A new approach to isomer-resolved chemical kinetics. *Rev. Sci. Instrum.* **79**, 104103, doi:10.1063/1.3000004 (2008).
- 53 Taatjes, C. A. *et al.* "Imaging" combustion chemistry via multiplexed synchrotron-photoionization mass spectrometry. *Physical Chemistry Chemical Physics* **10**, 20-34, doi:10.1039/b713460f (2008).
- 54 Herriott, D., Kompfner, R. & Kogelnik, H. Off-axis paths in spherical mirror interferometers. *Appl. Optics* **3**, 523-526, doi:10.1364/ao.3.000523 (1964).
- 55 Pilgrim, J. S., Jennings, R. T. & Taatjes, C. A. Temperature controlled multiple pass absorption cell for gas phase chemical kinetics studies. *Rev. Sci. Instrum.* **68**, 1875-1878, doi:10.1063/1.1147960 (1997).
- 56 Huang, H., Merthe, D. J., Zador, J., Jusinski, L. E. & Taatjes, C. A. New experiments and validated master-equation modeling for OH production in propyl + O-2 reactions. *Proc. Combust. Inst.* **33**, 293-299, doi:10.1016/j.proci.2010.06.039 (2011).
- 57 DeSain, J. D., Klippenstein, S. J. & Taatjes, C. A. Time-resolved measurements of OH and HO₂ product formation in pulsed-photolytic chlorine atom initiated oxidation of neopentane. *Physical Chemistry Chemical Physics* **5**, 1584-1592, doi:10.1039/b211452f (2003).
- 58 Rothman, L. S. *et al.* The HITRAN 2004 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.* **96**, 139-204, doi:10.1016/j.jqsrt.2004.10.008 (2005).
- 59 Montgomery, J. A., Frisch, M. J., Ochterski, J. W. & Petersson, G. A. A complete basis set model chemistry. VI. Use of density functional geometries and frequencies. *J. Chem. Phys.* **110**, 2822-2827, doi:10.1063/1.477924 (1999).
- 60 Montgomery, J. A., Frisch, M. J., Ochterski, J. W. & Petersson, G. A. A complete basis set model chemistry. VII. Use of the minimum population localization method. *J. Chem. Phys.* **112**, 6532-6542, doi:10.1063/1.481224 (2000).
- 61 Gaussian 09, Revision A.02 (Gaussian, Inc., Wallingford CT, 2009).

APPENDIX A: SECTION 2 MATERIALS AND METHODS^a

Chemicals and standards. 2-Phenylethanol (>99%), 1-Pentanol(>99.8%), 2-ethyl-1-hexanol (>99/5%), Cyclohexyl isothiocyanate (98%); 1-methyl-1,4-cyclohexadiene(>96%), 3-pentanone(>99%), 2,6-Di-tert-butyl-1,4-benzoquinone(98%), and all terpene and alkane standard solution (mixture of C8-C20 in hexane) were purchased from Sigma-Aldrich. Potato dextrose and M9 media were purchased from Difco Laboratories.

Organisms and Culture Conditions. This study was based on several endophytic fungi isolates (listed in Table 1) obtained from the Mycological Collection of the Department of Plant Sciences at Montana State University. Fungal cultures were grown on potato dextrose agar (PDA, 39g/L) at room temperature and passaged every 7 days. For investigating fungal growth in biomass feedstock (BF), fungal isolates were grown in agar plates containing M9 media supplemented with 2mM MgSO₄, 0.1mM CaCl₂, trace elements, and 0.25-0.5% of sterile feedstock (irradiated milled corn stover, CRN; switchgrass, SWG; or arbog eucalyptus, EUC). PDA and BF plates were inoculated with 5 mm plugs (placed in the center of plate) containing fungal mycelium and allowed to grow at room temperature for up to several weeks. Growth measurements were taken from four replicate plates for PDA and each feedstock biomass (at 0.25% and 0.5% w/v biomass concentration). Radial growth measurements were taken every 3 days, marking four points along the circumference of fungal mycelia, until growth reached the edge of the plate. An average growth was determined by taking 16 points from four individual plates. For VOC analysis, fungi were grown in 60 mL of potato dextrose (PD) broth or feedstock-supplemented M9 media (as described above). The 250 ml culture flasks were sealed with screw caps and incubated at room temperature on platform shakers (~100 rpm) for time periods of up to 25 days.

VOC Extraction and GC-MS analysis. The volatiles were extracted from the culture headspace after 7 days (for PD cultures) or 15 days (for BF cultures). Briefly, a small hole was drilled in the cap for insertion of a Solid Phase Micro Extraction (SPME) fiber holder (Supelco) containing a pre-conditioned Stableflex SPME fiber with 50/30 μm DVB/CAR/PDMS coating. Fiber was exposed to headspace for 5 minutes (up to 15 minutes for the BF cultures) and analyzed immediately. The volatiles bound to the fiber were desorbed for 3 minutes in a split-/splitless injector (splitless mode, 250°C) of a Varian 3800 gas chromatograph coupled with a Varian Saturn 2000 ion trap mass spectrometer. The column was a DB5-HT (30m x .25 mm x 0.25 μm).

Helium was used as carrier gas at constant flow rate 1 ml/min. The oven program was 30°C (hold 3 min), 5°C/min to 220°C, 10°C/min to 250°C (hold 5 min). The MSD parameters were EI at 70eV, mass range was 30-500 Da, and the scan speed was 2 scans/sec. Headspace SPME-GC-MS was performed on at least three replicate cultures for each fungi-feedstock combination.

MS Data analysis. GC-MS data deconvolution was performed using the Automated Mass Spectral Deconvolution and Identification System (AMDIS) spectral deconvolution software package (v. 2.70, NIST Gaithersburg)⁵⁰. AMDIS deconvolution settings were as follows: resolution (medium), sensitivity (low), shape requirement (medium), and component width at 10. Spectral components were searched against the NIST 2011 mass spectral library, and only components with mass spectra match factors > 70% were reported as tentatively identified compounds. Compounds with peak areas > 1% of the total peak area in the chromatogram are reported. Sørensen similarity indices (based on the presence/absence of VOC in culture headspace) was measured as a means to estimate the differences among samples from different fungal species and carbon/feedstock sources⁵¹.

Liquid-liquid extraction of the CI-4A culture. Seven day old mycelia of strain *Hypoxylon* CI4A grown on potato dextrose (PD) agar plates were collected and re-suspended into 5 ml of PD broth. The re-suspended mycelia were inoculated into 250 ml of PD broth and incubated at 25°C for 5 days to develop the mycelia. This culture was then inoculated into a BioFlo3000 fermenter with the working volume of 6L PD broth. The fermenter was sparged at rate of 2.0 ml/min compressed air (5 psi), stirred at 200 rpm for 7 days. Once the fermentation finished, the culture broth was collected by filtrating the mycelia through 30 µm Nylon membrane. The collected fermentation broth was used for liquid-liquid extraction with dimethyl ether.

Initial attempts to isolate the organic content of the fungal fermentation broth utilized a combination of steam distillation and liquid-liquid extraction. The concept was to first concentrate the organic components in the aqueous layer in an effort to separate materials by phase separation. Unfortunately, the concentration of organic matter was too low to allow phase separation to occur. Therefore a liquid-liquid extraction of the aqueous layer with diethyl ether was performed. In later separations the steam distillation process was deemed unnecessary and only an extraction was performed to isolate materials. To improve the efficiency of the

separation a process of using a continuous liquid-liquid extraction vessel was employed, Figure 25. This allowed for a longer extraction period and provided a minorly improved yield from the extraction. Diethyl ether was chosen due to the low volatility (easier to remove while retaining the less volatile extracts) and also because less heat is required to perform the operation. Since the extracts are subjected to some heat during this process this also mitigates concerns regarding stability of the extracts during the processing. The process is conducted in this manner: the extraction vessel is filled $\frac{3}{4}$ full of the fungal broth to be extracted and the remaining volume of the vessel is filled with diethyl ether. A stir bar is added to keep the solutions agitated and a cooling condenser is attached. A collection flask with fresh diethyl ether is brought to a boil (~ 40 °C). The distilled ether cools in the condenser and collects in the extraction vessel. Since ether is lighter than the broth it stays as the upper layer and any excess solvent volume (containing extracted components) flows down the connecting tube to the collection flask. Since the organic extracted components have a lower volatility than diethyl ether they remain in the collecting flask while ether distills back into the extraction vessel. After a period of 16-18 hours the diethyl ether layer containing the organic extracts is separated from the broth and then dried over magnesium sulfate to remove excess water. The mixture is then concentrated using a rotary evaporator to collect the less volatile organic components.

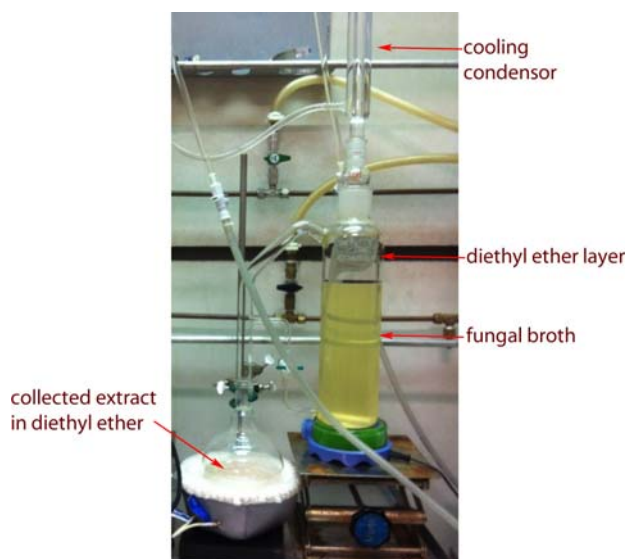


Figure 25. A continuous liquid-liquid extraction vessel

The individual constituents of the collected samples were identified using gas chromatography with a mass spectrometer as the detector (GC-MS). Typically, between 0.25-0.50 μL of the

collected materials is added to 1 mL of methanol and then injected into the instrument. Figure 3 shows a representative GC-MS chromatogram of one of the first amounts of extracted organics. Integration was performed on the chromatogram to determine the relative percentages of the constituents. The results vary between fermentation batches, common percentages included (as labeled on the chromatogram) 1) 1.88% 1,8-cineole, 2) 8.33% [beta]-pinene, 3) 22.48% 2-(2-methyl-2-propenyl)-2-cyclohexen-1-one, and 4) 8.49% BHT (a stabilizer present in the solvent that is concentrated through processing).

Semi-preparative TLC experiments were run with 3% ethyl acetate in hexanes in order to isolate the organic compounds. The separate spots on the plate were identified under 254 nm light and silica gel with the adsorbed components was then scraped off with a razor blade. The silica gel was washed with deuterated chloroform (CDCl_3) to remove the organics from the gel. This CDCl_3 solution was then run through proton NMR to produce the spectrum shown in Figure 26. This was done in order to confirm the structures identified by the GC-MS results. The mixture of components makes it difficult to determine the exact nature of the collected fractions and this will be investigated in future research of these materials.

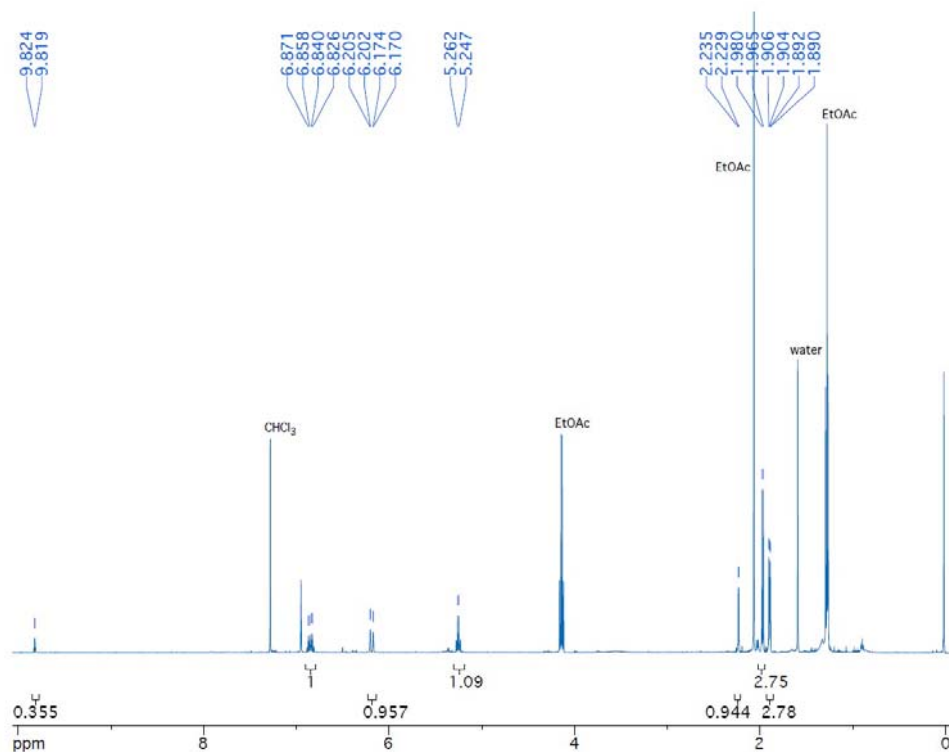


Figure 26. NMR spectrum of isolated fraction obtained through semi-preparative TLC of the fungal extracts.

APPENDIX B: SECTION 3 MATERIALS AND METHODS^a

Expression Host and Vector. The *E.coli* strain DH10B and DH1 were obtained from Jay Keasling group in Joint BioEnergy Institute (JBEI). The plasmid JBEI3122 (Figure 27) was provided courtesy of Dr. Jorge Alonso-Gutierrez. The plasmid JBEI3122 contains the genes encoding all the enzymes in the mevalonate pathway except the geranyl diphosphate synthase (GPPS) and terpene synthase (TS). The protein sequences of putative terpene synthases were downloaded from endophyte genomes published by Joint Genome Institute. All the terpene synthase protein sequences and a protein sequence of GPPS (GenBank: AF513112.1, GPPS_{Ag}) from *Abies grandis* with the chloroplast signal peptide truncated were converted into DNA sequences and codon optimized based on *E. coli* codon bias. All the gene sequences were synthesized by Genscript.

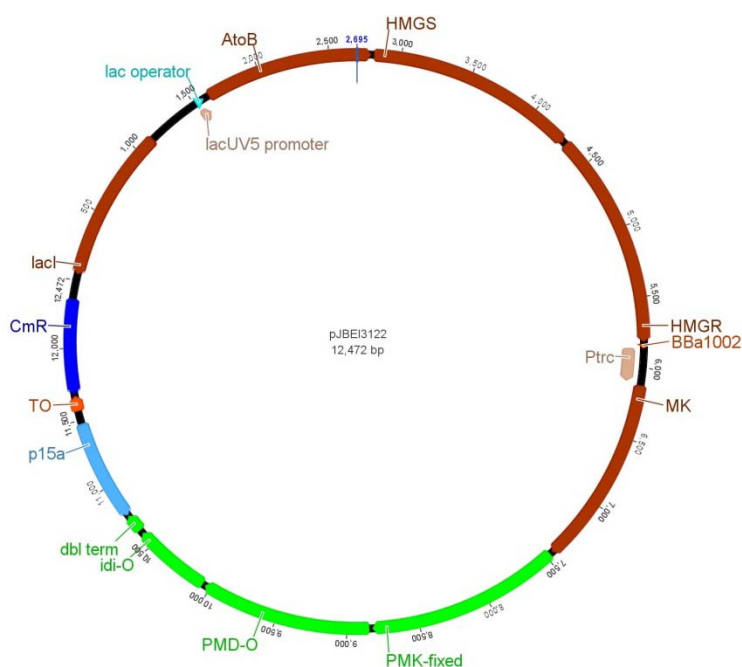


Figure 27. Vector map of plasmid JBEI3122 containing the up-pathway of terpene biosynthesis

The up-pathway converts acetyl-CoA into isoprenyl diphosphate in seven enzymatic steps: acetyl-CoA acetyltransferase(AtoB); HMG-CoA synthase(HMGS);HMG-CoA reductase(HMGR); mevalonate kinase(MK); phosphomevalonate kinase(PMK); Mevalonate diphosphate decarboxylase(PMD); isoprenyl diphosphate isomerase (idi).

Reconstruct the terpene synthetic pathway into *E.coli* strain DH1

Four strategies were applied to screen terpene synthase and optimize the terpene production in the engineering *E.coli* strain.

Construct 1: The each terpene synthase ORF including the sequences of corresponding ribosome binding site was sub-cloned into plasmid pBbE1a under EcoRI and BamHI cutting site to obtain vector pBbE1a-TS in Genescript. The GPPS was sub-cloned into vector pBbE2k under EcoRI and BamHI cutting site to obtain plasmid pBbE2k-GPPS_{Ag} as well. The plasmids pJBEI3122, pBbE1a-TS, and pBbE2k-GPPS_{Ag} were co-electroporated into expression host DH1 for the screening of terpene synthase. The plasmids pJBEI3122 and pBbE2k-GPPS_{Ag} were co-electroporated into strain DH1 as well as negative control.

Construct 2: the gene GPPS_{Ag} was amplified from vector pBbE2k-GPPS_{Ag} with the primers pBbE1a-GPPS-AG-F2 (5-GAG TCT GAATTCAAAAAGATCT GAG ATT TAT CAC ATA GAG CGG -3) and pBbE1a-GPPS_LS-R3(5-GAG TCT GAA TTC CTC GAG GAT CC TCA ATT TTG TCT GAA TGC CAC-3). The amplicon of GPPS_{Ag} was digested with BglII and XhoI and sub-cloned into plasmid pJBEI3122 right downstream of gene isoprenyl diphosphate isomerase (*idi*) to obtain plasmid pJBEI3122-GPPS_{Ag}. The plasmids pJBEI3122-GPPS_{Ag} and pBbE1a-TS were co-electroporated into *E.coli* strain DH1 for terpene production. The plasmid pJBEI3122-GPPS_{Ag} was transformed into strain DH1 as negative control as well.

Construct 3: The amplicon of gene GPPS_{Ag} was digested BglII and XhoI and cloned into plasmid pBbE1a-TS to obtain plasmid pBbE1a-TS-GPPS_{Ag}. The plasmids pJBEI3122 and pBbE1a-TS-GPPS_{Ag} were co-electroporated into *E.coli* strain DH1 for the production of terpene.

Construct 4: The amplicon of gene GPPS_{Ag} was digested EcoRI and cloned into plasmid pBbE1a-TS to obtain plasmid pBbE1a-GPPS_{Ag}-TS. The right orientation of GPPS_{Ag} was confirmed by PCR using orientating primers: pBbE1a-GPPS-TS-F(5-CAT CCG GCT CGT ATA ATG TGT GG-3) and pBbE1a-GPPS-TS-R (5-GCTC CTC GGT TCC TCC AAC AAG-3). The plasmids pJBEI3122 and pBbE1a-GPPS_{Ag}-TS were co-electroporated into *E.coli* strain DH1 for the production of terpene.

Screen the terpene synthases and characterize the terpene production by the engineered *E.coli*

Production of terpene: the colonies of construct1 containing each terpene synthase were streaked into 15 ml of LB medium with 100 mM of Ampicillin, 34 µg/L chlorophenicol, and 25 mM of kanamycin. The cultures were incubated at 37°C at 220 rpm overnight. Then, 2 ml of overnight culture (OD around) were transferred into 20 ml of EZ-rich medium (Teknova, CA) containing 20 g/L glucose as sole carbon source. Then, the culture were continued to incubate at 37°C, 220 rpm until the OD_{600nm} reached 0.8. The cultures were induced by adding isopropyl-β-D-1-thiogalactopyranoside (IPTG) at the final concentration 500 mM for another 20 hours at 30°C, 160 rpm.

GC/MS analysis of terpene: Gas analysis of the volatile organic compounds in the headspace of each culture was done on a preconditioned solid-phase micro-extraction (SPME) syringe consisting of 50/30 divinylbenzene/carboxen on polydimethylsiloxane on a Stable Flex fiber for an hour. Then, the syringe was inserted into the injection port of a Varian 3800 gas chromatograph containing a 30m x 0.25mm i.d DB wax capillary column with a film thickness 0.25 µm. The column was temperature programmed as follows: 60°C for 4 min, increasing to 120°C at 10°C/min and holding for 5 min, then increasing to 220°C at 20°C/min and holding for 2 min, then increasing to 250°C at 50°C/min and holding for 4 min. The carrier gas was ultra-high purity helium at constant flow rate 1 ml/min, and the initial column head pressure was 50Kpa. A 120s injection time was used to desorb the VOC from the sample fiber into a split/splitless injection (splitless mode, 220°C) of the chromatograph coupled with a Saturn 2000 ion trap mass spectrometer. The MSD parameters were EI at 70eV, mass range was 30-500 Da, and the scan speed was 2 scans/sec.

MS Data analysis: GC-MS data deconvolution was performed using the Automated Mass Spectral Deconvolution and Identification System (AMDIS) spectral deconvolution software package (v. 2.70, NIST Gaithersburg). AMDIS deconvolution settings were as follows: resolution (medium), sensitivity (low), shape requirement (medium), and component width at 10. Spectral components were searched against the NIST 2011 mass spectral library, and only components with mass spectra match factors > 80% were reported as tentatively identified compounds. Compounds with peak areas > 0.5% of the total peak area in the chromatogram are reported.

APPENDIX C: SECTION 4 MATERIALS AND METHODS^a

Strains, culture medium and conditions. The endophytes *Hypoxylon* CI4A, EC38, CO27 and *Daldinia* EC12 grown on barley were transferred onto potato dextrose agar plates to develop mycelia for 5 days. The mycelia were collected and re-suspended into 2 ml of potato dextrose medium. 1 ml of potato dextrose with mycelia were transferred into 200 ml of fresh potato dextrose medium and cultured at room temperature for 5 days. The grown-up mycelia were collected by filtering through 30 μ m Nylon membrane and rinsed with 1X M9 salt solution for three times to remove all oligosaccharide residues. The collected mycelia were transferred to 200 ml of 1X M9 medium containing 20g/L Avicel, switch grass, corn stover, or eucalyptus mill powder as sole carbon source, respectively, for the inductions of cellulase. The samples were taken out every 24 hours and culture supernatants were prepared for following cellulase activity assay.

A *Neurospora crassa* strain (FGSC #2489) was characterized for the expression of cellulase in comparison with the endophytes *Hypoxylon* and *Daldinia* as well. The 10-days' spores of *N. crassa* strain were collected and mycelia were removed by filtrating through four layers of cheese cloth. The spores were inoculated into 1X Vogel's medium containing 20 g/L glucose as sole carbon at 10^6 spore/ ml medium and cultured at 25°C for 16 hours. The mycelia were collected by filtrated through 100 μ m Nylon membrane and rinsed with 1X Vogel's salts three times to remove glucose residue. Then, the mycelia were transferred into 200 ml of 1X Vogel's medium containing 20 g/L Avicel, switch grass, corn stover, or eucalyptus mill powder as sole carbon source for the induction of cellulase.

Cellulase activity assay. Endoglucanase activity was measured by dinitrosalicylic (DNS) colorimetric method as described below. A mixture of 1.5 ml of culture supernatant, 0.5 ml of carboxymethylcellulose (CM-cellulose) in 50 mM citrate buffer (pH 4.8) was incubated at 50°C for 1 hour. The reactions were quenched by adding 100 μ l of reaction mixture into 600 μ l of DNS reagent and the resulted mixtures were incubated at 100°C for exactly 10 minutes. One unit of endoglucanase activity is defined as the amount of enzyme that produced 1 μ mol of glucose in one minute under the assay conditions described above.

Exoglucanase activities were measured using *p*-nitrophenyl- β -D-lactoside (*p*NPL) as the substrate described as below. The enzyme reaction solution contains 1.5 ml of culture supernatant, 0.5 ml of 1 mg/ml *p*NPL in 50mM citrate buffer (pH4.8). The reaction mixture was incubated at 50°C for 30 minutes and quenched by adding 0.9 ml of 50 mM NaOH into 0.6 ml reaction mixture immediately. *Beta*-glucosidase (BGL) activity was measured using *p*-nitrophenyl- β -D-glucopyranoside (*p*NPG) as the substrate. Similar to the exoglucanase activity assay, the enzyme reaction mixture containing 1ml of culture supernatant, 0.5 ml of 50 mM citrate butter (pH4.8), and 0.5 ml of 5 mM *p*NPG in 50 mM citrate buffer (pH4.8) was incubated at 50°C for 10 minutes and quenched by adding 0.9 ml of 50 mM NaOH into 0.6 ml of reaction mixture as well. One unit of exoglucanase and beta-glucosidase activities are defined as the amount of enzyme that produced 1 μ mol of nitrophenol in one minute under the assay conditions described above. All the activities of three enzymes were normalized into U/ g mycelia.

APPENDIX D: SECTION 5 MATERIALS AND METHODS^a

Microorganism Cultivation Conditions and Secretome Extraction: Fungal endophytes C14A, CO27, EC12, and EC38 have been grown on solid and liquid potato dextrose media, as well as minimal media supplemented with Avicel, corn, eucalyptus, or switchgrass feedstocks (obtained from JBEI, Emeryville, CA). Minimal media was prepared by making a master mix as follows: 100mL of 5X M9 Salts, 1mL 1M MgSO₄, 50μL 1M CaCl₂, 3mL 170X trace elements. For 1L of 170X trace elements, the following was added: Nitrioloacetic acid (15g), MgSO₄·7H₂O (30g), MnSO₄·H₂O (5g), NaCl (10g), FeCl₂ (0.7g), CoSO₄ (1g), CaCl₂·2H₂O (1g), ZnSO₄·7H₂O (1g), CuSO₄·5H₂O (0.1g), Alk(SO₄)₂·12H₂O (0.1g), H₃BO₃ (1g), Na₂MoO₄·H₂O (0.1g). For secretome preparation, 60mL liquid cultures were grown for 5 days at 27°C with gentle agitation (approximately 85rpm). Fungal cultures were filtered with 0.22μM Stericups (Millipore), then concentrated to 200μL using Vivaspin columns (GE Healthcare).

Protein Digestion, Peptide Extraction, and Mass Spectrometric Analysis: In brief, protein from each sample was run in duplicate on a 4-12% Bis-Tris SDS-PAGE gel (Life Technologies) at 170V for 60 minutes. Gels were stained using Simply Blue Safe Stain (Life Technologies) and were destained overnight. Resulting bands were manually excised. Bands were destained, lyophilized, and digested with Trypsin Gold (Promega) according to manufacturer protocol, with the exception that the digests were carried out for 6 hours. Digested samples were then analyzed with by direct infusion mass spectrometry.

Mass spectrometry and MS/MS Analysis: Peptides were reconstituted in 10 mM ammonium acetate, 85% acetonitrile, 0.1% formic acid and analyzed by direct infusion electrospray ionization coupled to a Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR-MS) (Bruker Daltonics). Each sample was injected twice at two equal volumes and independently analyzed. The mass spectrometer was set in positive ion mode and data were acquired with a selected mass range of 400–2200 *m/z*, with observed peptides with +1 to +5 charge states. With the high resolving power of the FT-ICR-MS most peptide assignments were made by mass alone. To verify this approach MS/MS was employed to confirm the sequence of several peptides.

Mass Spectrometric Data Search and Analysis: The data acquiring was carried out using mass software (Bruker Daltonics). The peak list generation was performed using DataAnalysis 2.0 software (Bruker Daltonics). Protein identification was performed using ProteinProspector MS-Fit against user generated databases annotated by JGI. Matching criteria included up to 1 missed trypsin cleavage and required at least 4 peptides within a 20ppm error for identification (in most cases at least 7 peptides were actually matched). Identified proteins were than grouped into classes based on annotated activities or structural properties .

APPENDIX E: SECTION 6 MATERIALS AND METHODS

Experiment

The experimental apparatus used for this study has been described in detail elsewhere^{52,53} and only a synopsis will be given here. The three compounds were studied using pulsed-photolytic chlorine-initiated oxidation at elevated temperatures (550 – 700 K). Sample gas (approximately 1%) seeded in helium is mixed with excess oxygen, molecular chlorine and pure helium and introduced into a 1.05 i.d. quartz tube. The 351 nm line of an excimer laser operating at 4 Hz photolyzes approximately 1.5% of the initial Cl₂, generating Cl• that reacts with the parent molecule via H-abstraction reactions. Enough chlorine is used to deplete roughly 10% of the parent ketone. Calibrated mass flow controllers are used to introduce reactants and carrier gas into the flight tube. A bubbler held at 20 °C, backed by He, provides a steady source of the ketone of interest. Depending on the vapor pressure of the ketone, initial concentrations of 1 - 7 × 10¹³ cm⁻³ in the flow tube are estimated at 550 K. Approximately 10% depletion of the parent is observed within 5 ms after photolysis. The consumption of the parent ketone molecule by Cl atoms yield HCl and a radical corresponding to parent –H atom. Pressure in the tube is maintained at 8.0 Torr by feedback control of a butterfly valve in the exhaust line. Typical initial mole fractions are estimated to be 5 × 10⁻⁴ (7.1 × 10¹³ cm⁻³) for the parent ketone, 2 × 10⁻³ (2.8 × 10¹⁴ cm⁻³) for Cl₂, 0.2 (2.8 × 10¹⁶ cm⁻³) for O₂ and the balance He. Gas escapes the reaction mixture through a 650 μm orifice in the sidewall of the tube and a skimmer provides a collimated molecular beam. The molecular beam enters the ionization region of an orthogonal extraction linear time-of-flight mass spectrometer where it is crossed by tunable VUV radiation supplied by the Advanced Light Source synchrotron at Lawrence Berkeley National Laboratory. Ion optics are pulsed at 40 or 50 kHz and product cations are detected as a function of mass, time, and photon energy in a multiplexed fashion. The 3-dimensional data can be sliced and integrated to obtain photoionization spectra and time profiles for individual species.^{33,53} Alternately, a hydrogen discharge lamp provides a continuous source of 10.2 eV ionizing radiation for experiments performed with the same apparatus at Sandia National Laboratories.

Direct Measurement of OH and HO₂

Time resolved concentrations of OH and HO₂ radicals in the Cl-initiated oxidation of DIPK were directly monitored with Herriott-type multi-pass differential absorption experiments using a slow-flow reactor cell.^{54,55} A detailed description of the experiment is given elsewhere⁵⁶ and only a synopsis is provided here. The apparatus consists of a quartz reactor tube 150 cm in length and 5 cm in diameter. Temperature is maintained between room temperature and 800 K with feedback control. The third harmonic of a Nd:YAG laser (355 nm) firing at 0.5 Hz is used as the photolysis source. As in the MPIMS experiment, molecular chlorine is used as the precursor for Cl radicals. The pressure inside the cell is maintained at 20 Torr by a feedback-controlled butterfly valve connected to a roots blower. The time resolution of the [OH] and [HO₂] measurements is 1.0 μs. Each time profile consists of 100,000 data points taken over 5 ms.

To monitor HO₂, a continuous-wave probe laser at 1.51 μm is tuned to a known overtone transition line of the OH-stretch in HO₂ at 6625.784 cm⁻¹. Raw HO₂ absorption signal is normalized to HO₂ signal from the reaction of CH₂OH + O₂ taken under conditions otherwise identical to the DIPK experiment. This former proceeds with essentially 100% yield to CH₂O + HO₂.⁵⁷ Similar to HO₂, in the OH absorption experiment, an OPO IR laser (Linios OS4000) beam probes OH by exciting the fundamental OH-stretch at 3484.59923 cm⁻¹.⁵⁸ The probe lasers complete a total of 29 passes inside the cell.

Sample Preparation

Diethyl ketone, di-isopropyl ketone, isopropyl-tert-butyl ketone and di-tert-butyl ketone were obtained commercially. 2,5-dimethylcyclopentanone and 2,2,5,5-tetramethylcyclopentanone were also obtained commercially at stated purities of 98%. Each sample was freeze pump thawed to remove high vapor pressure impurities before use. Bubblers maintained at constant backing pressures with He carrier gas were used to introduce ketone samples. The photoionization calibration spectra of the cyclopentanones were obtained by making gaseous mixtures of approximately 0.2% sample with 1000 Torr of He in 4 L stainless steel cylinders.

Quantum Chemical Calculations

Ab initio calculations were performed with the CBS-QB3 composite method^{59,60} on the Gaussian 09 program suite⁶¹ to predict adiabatic ionization energies (AIEs) for product isomers as well as their relative energies at 0 K. Product assignments are aided by comparison of the calculated AIEs with the onsets of the respective photoionization spectra obtained at the ALS. Neutral and cation geometries were optimized independently. Additionally, investigations of certain potential energy surfaces relevant to oxidation of ketones were carried out at the CBS-QB3 level. Lowest energy conformers of minima and first-order saddle points (transition states) were systematically searched for at B3LYP/6-31G(d,p) level of theory. The optimized geometries were then used as input for the CBS-QB3 calculations. Transition states were confirmed by the presence of one imaginary frequency in the normal mode analysis, and the normal mode displacements associated with the imaginary frequency were visually inspected to confirm that the transition state connects to the correct reactant and product. These calculations are reserved for diisopropyl ketone because this compound provides a good model for analogous pathways expected in each of the branched ketones studied here.

APPENDIX F: SECTION 2 SUPPLEMENTAL TABLES

Supplementary Table 9. List of tentative compounds identified in the headspace of CI-4A cultures cultivated with PD and biomass feedstocks

Ret. time (min)	Tentative Compound	Molecular Formula	Match*	% Area PDB	% Area CRN	% Area SWG	% Area EUC
15.2	Eucalyptol	C10H18O	84	47.4	2.7	14.6	7.6
24.0	2-Cyclohexen-1-one, 2-(2-methyl-2-propenyl)-	C10H14O	76	19.2	4	<1	nd
15.5	Cyclohexane, 1,2,4-tris(methylene)-	C9H12	84	3.2	-	nd	<1
13.6	<i>β</i> -pinene	C10H16O	83	3.2	nd	nd	nd
13.3	1,3,6-Cyclooctatriene	C8H10	81	3	nd	nd	nd
17.7	Phenylethyl Alcohol	C8H10O	84	2.2	nd	nd	nd
15.9	<i>γ</i> -Terpinene	C10H16	85	1.8	nd	nd	nd
22.8	Umbellulone	C10H14O	75	1.8	nd	nd	nd
17.8	1-Decen-3-yne	C10H16	75	1.7	nd	nd	nd
11.4	<i>β</i> -Thujene	C10H16	79	1.4	nd	nd	nd
15.8	Cyclohexene, 1-(1-propynyl)-	C9H12	80	1	nd	nd	nd
20.1	<i>α</i> -Terpineol	C10H18O	84	1	nd	<1	nd
21.3	<i>Cyclohexane, isothiocyanato-</i>	<i>C7H11NS</i>	80	<1	32.4	17.9	54.6
28.1	Chamigrene	C15H24	81	<1	1	1.2	1
28.4	Guaia-1(10),11-diene	C15H24	85	<1	3	3.9	3.9
32.0	<i>γ</i> -Gurjunene	C15H24	75	<1	2	10.1	7.4
10.7	<i>Oxime-, methoxy-phenyl-</i>	<i>C8H9NO2</i>	81	nd	17.2	1.7	10.9
17.6	Benzene, 1-ethyl-4-methoxy-	C9H12O	91	nd	8.2	nd	nd
14.9	1-Hexanol, 2-ethyl-	C8H18O	84	nd	7.4	nd	<1
18.9	Benzene, 1-ethenyl-4-methoxy-	C9H10O	90	nd	1.9	1.7	nd
21.1	Benzothiazole	C7H5NS	90	nd	2.2	<1	1.3
5.7	2-Pentenal, 2-methyl-	C6H10O	75	nd	2	17.9	nd
26.0	Valencene	C15H24	75	nd	1	nd	nd
13.5	3-Octanone	C8H16O	85	nd	nd	6.5	2.8
17.2	1,6-Octadien-3-ol, 3,7-dimethyl-	C10H18O	83	nd	nd	3	nd
25.5	<i>β</i> -Elemene	C15H24	75	nd	nd	2.6	2.2
26.7	<i>α</i> -Guaiene	C15H24	89	nd	nd	1	1
24.3	2,2,6,7-Tetramethyl-10-oxatricyclo[4.3.0.1(1,7)]decan-5-one	C13H20O2	80	nd	nd	2	1.4
18.6	Benzene, 1,2-dimethoxy-	C8H10O2	91	nd	nd	nd	1.1

*VOC profiles from potato dextrose broth (PD), corn stover (CRN), switchgrass (SWG) and arbo eucalyptus (EUC). % Area is taken as AUC (area under curve) of peak of interest versus total AUC for entire chromatogram. Some compounds are not detected (nd) for certain conditions. Compounds are observed in some uninoculated media (italicized).

Supplementary Table 10. List of tentative compounds identified in the headspace of EC-12 cultures cultivated with PD and biomass feedstocks

Ret. time (min)	Tentative Compound	Molecular Formula	Match*	% Area PD	% Area CRN	% Area SWG	% Area EUC
5.4	1-Butanol, 3-methyl- or 1-Pentanol	C5H12O	80	23.5	nd	nd	nd
10.6	<i>Oxime-, methoxy-phenyl-</i>	<i>C8H9NO2</i>	<i>84</i>	4.7	30.2	<1	2.7
17.6	Phenylethyl Alcohol	C8H10O	90	12	nd	nd	nd
25.3	Alloaromadendrene	C15H24	78	0.9	nd	1	nd
25.5	<i>β</i> -Elemene	C15H24	86	8.8	nd	11.1	11.4
26.3	Caryophyllene	C15H24	84	0.7	-	1	<1
26.7	<i>α</i> -Guaiene	C15H24	88	4.1	nd	4.4	5.9
27.7	2H-1-Benzopyran-2-one, 4,7-dihydroxy-	C9H6O4	75	3.8	nd	nd	nd
28.1	Valencene	C15H24	88	2.9	3	nd	3.6
28.2	Chamigrene	C15H24	83	1.5	-	nd	nd
28.3	Guaia-1(10),11-diene	C15H24	86	15.7	nd	10.1	15
32.0	<i>γ</i> -Gurjunene	C15H24	83	15.5	nd	22.3	18.8
19.3	6-Butyl-1,4-cycloheptadiene	C11H18	82	nd	3.4	nd	nd
23.9	<i>γ</i> -Elemene	C15H24	78	nd	1	nd	nd
28.4	Phenol, 2,4-bis(1,1-dimethylethyl)-	C14H22O	87	nd	5.1	nd	nd
28.8	trans-calamenene	C15H22	70	nd	2.9	nd	nd
28.7	<i>β</i> -Cadinene	C15H24	76	nd	2.9	nd	nd
25.1	<i>α</i> -Copaene	C15H24	75	nd	2.9	nd	nd
27.3	Aromadendrene	C15H24	77	nd	2.7	<1	nd
25.3	Eremophilene	C15H24	75	nd	1	nd	nd
26.2	4a(2H)-Naphthalenol, octahydro-4,8a-dimethyl-	C12H22O	75	nd	1.6	nd	nd
18.8	2,4,6-Cycloheptatrien-1-one	C7H6O	75	nd	1	nd	nd
14.9	1-Hexanol, 2-ethyl-	C8H18O	82	nd	nd	2.6	1
21.3	Cyclohexane, isothiocyanato-	C7H11NS	78	nd	nd	3.7	19.4
21.9	2(5H)-Furanone, 4-methyl-5,5-bis(2-methyl-2-propenyl)-	C13H18O2	78	nd	nd	7.9	7
17.2	1,6-Octadien-3-ol, 3,7-dimethyl-	C10H18O	80	nd	nd	16.4	nd
13.5	3-Heptanone, 6-methyl	C8H16O	58	nd	nd	7	nd
28.1	2-Isopropenyl-4a,8-dimethyl-1,2,3,4,4a,5,6,8a-octahydronaphthalene	C15H24	88	nd	nd	3.6	2.2
27.3	2,5-Cyclohexadiene-1,4-dione, 2,6-bis(1,1-dimethylethyl)-	C14H20O2	83	nd	nd	1	1.3
21.1	Benzothiazole	C7H5NS	94	nd	nd	nd	1
19.9	Creosol	C8H10O2	70	nd	nd	nd	1
13.3	Phenol	C6H6O	83	nd	nd	nd	2.2

*VOC profiles from potato dextrose broth (PD), corn stover (CRN), switchgrass (SWG) and arbog eucalyptus (EUC). % Area is taken as AUC (area under curve) of peak of interest versus total AUC for entire chromatogram. Some compounds are not detected (nd) for certain conditions. Compounds are observed in some uninoculated media (italicized).

Supplementary Table 11. List of tentative compounds identified in the headspace of EC-38 cultures cultivated with PD and biomass feedstocks

Ret. time (min)	Tentative Compound	Molecular Formula	Match*	% Area			
				PD	CRN	SWG	EUC
28.4	Guaia-1(10),11-diene	C15H24	86	35	7.7	5.3	9.7
32.1	γ -Gurjunene	C15H24	83	13.3	22.2	10.3	nd
15.0	Eucalyptol	C10H18O	86	11.6	2.6	nd	<1
28.1	Chamigrene	C15H24	82	9.2	3	1.6	2.8
25.5	β -Elemene	C15H24	87	4.2	7.2	2.4	2.7
28.3	2-Isopropenyl-4a,8-dimethyl-1,2,3,4,4a,5,6,8a-octahydronaphthalene	C15H24	75	3.5	nd	nd	nd
26.7	α -Guaiene	C15H24	89	3	2.7	3.2	2
5.4	1-Pentanol or 1-Butanol, 3-methyl-	C5H12O	75	2.7	nd	nd	nd
14.9	Cyclohexene, 1-methyl-4-(1-methylethenyl)-	C10H16	78	2.1	nd	nd	nd
13.6	β -Pinene	C10H16	86	1.8	nd	nd	nd
17.6	Phenylethyl Alcohol	C8H10O	92	1	nd	nd	nd
10.6	<i>Oxime-, methoxy-phenyl-</i>	<i>C8H9NO2</i>	83	1	19.2	11.6	1
26.9	Himachala-2,4-diene	C15H24	79	1	nd	nd	nd
28.8	Calamenene	C15H22	75	<1	1	<1	1.9
28.7	β -Cadinene	C15H24	86	<1	1.3	<1	1.7
21.3	Cyclohexane, isothiocyanato-	C7H11NS	85	nd	1	<1	7.2
14.9	<i>1-Hexanol, 2-ethyl-</i>	<i>C8H18O</i>	75	nd	3.6	2.5	nd
16.5	3-Pentanone	C5H10O	89	nd	2.9	2.9	<1
16.8	3-Heptanone, 5-ethyl-4-methyl-	C9H18O	91	nd	3	4.4	2.5
18.6	Benzene, 1,2-dimethoxy-	C8H10O2	75	nd	<1	nd	4
18.9	Benzene, 1-ethenyl-4-methoxy-	C9H10O	86	nd	2.1	1.4	nd
6.2	1,3,5-Cycloheptatriene	C7H8	85	nd	2	nd	nd
17.2	1,6-Octadien-3-ol, 3,7-dimethyl-	C10H18O	87	nd	nd	31.2	nd
33.1	1H-Indene, 2,3-dihydro-1,1,3-trimethyl-3-phenyl-	C18H20	75	nd	nd	nd	1
25.8	Methyleugenol	C11H14O2	84	nd	nd	nd	1
23.6	Benzene, 4-ethyl-1,2-dimethoxy-	C10H14O2	89	nd	nd	nd	1.2
21.1	Benzothiazole	C7H5NS	93	nd	nd	nd	1.4
18.5	Cyclohexanol, 1-methyl-4-(1-methylethyl)-, trans-	C10H20O	78	nd	nd	nd	4
16.2	Ethanone, 2,2-dihydroxy-1-phenyl-	C8H8O3	75	nd	nd	nd	1.2
13.3	Phenol	C6H6O	80	nd	nd	nd	7.8

*VOC profiles from potato dextrose broth (PD), corn stover (CRN), switchgrass (SWG) and arbo eucalyptus (EUC). % Area is taken as AUC (area under curve) of peak of interest versus total AUC for entire chromatogram. Some compounds are not detected (nd) for certain conditions. Compounds are observed in some uninoculated media (italicized).

Distribution

4 Lawrence Livermore National Laboratory
Attn: N. Dunipace (1)
P.O. Box 808, MS L-795
Livermore, CA 94551-0808

1	MS0359	D. Chavez, LDRD Office	1911
1	MS0899	Technical Library	9536 (electronic copy)
1	MS9054	Art Pontau	8360 (electronic copy)
1	MS9055	Thomas B. Settersten	8353 (electronic copy)
1	MS9291	Blake A. Simmons	8630 (electronic copy)



Sandia National Laboratories