

## **Supplementary Information**

### **Production of neo acids from biomass-derived monomers**

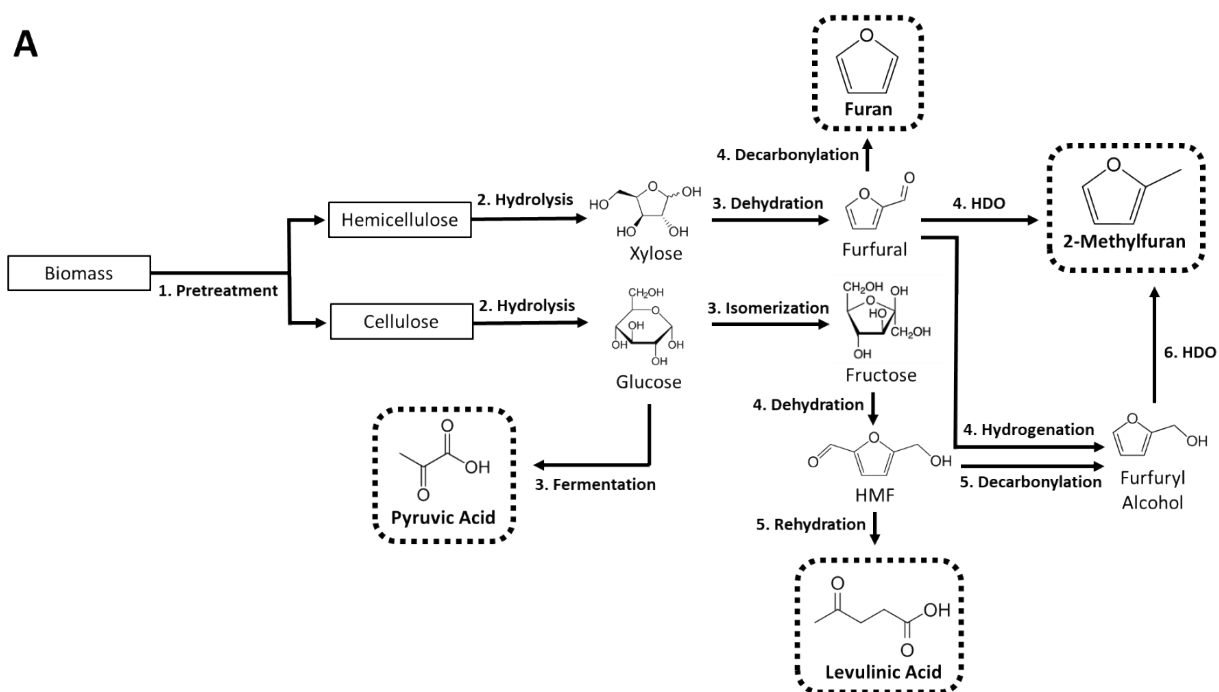
Erha Andini,<sup>1,2</sup> Jake Bragger,<sup>1</sup> Sunitha Sadula,<sup>2\*</sup> and Dionisios G. Vlachos<sup>1,2\*</sup>

<sup>1</sup>Department of Chemical and Biomolecular Engineering, 150 Academy St., University of Delaware, Newark, Delaware 19716, USA

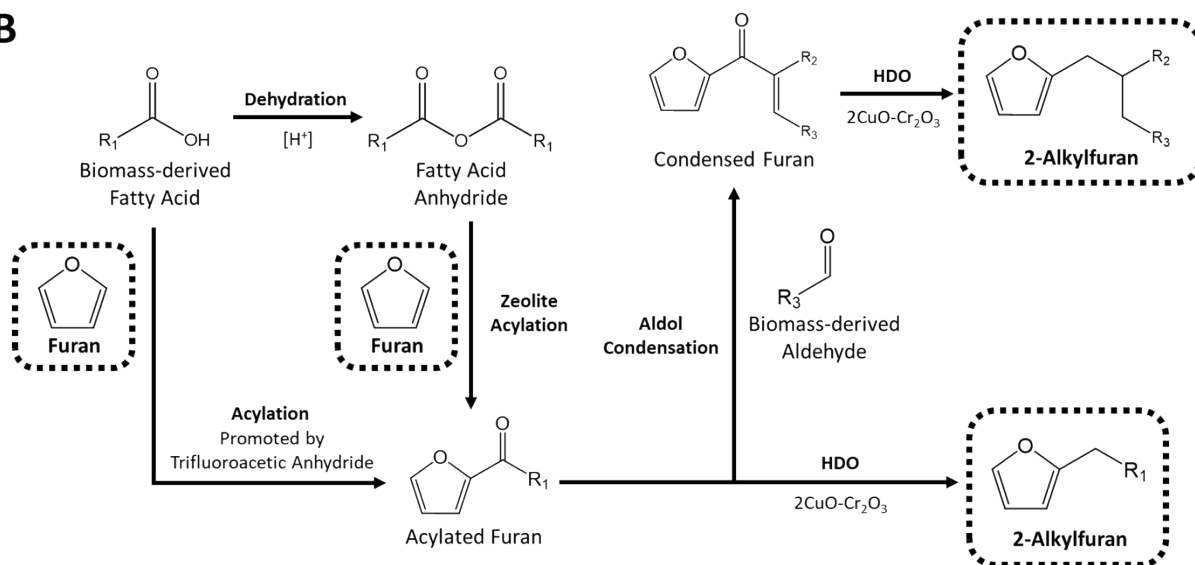
<sup>2</sup>Catalysis Center for Energy Innovation, 221 Academy St., University of Delaware, Newark, Delaware 19716, USA

\*Corresponding author: [vlachos@udel.edu](mailto:vlachos@udel.edu) and [sunithak@udel.edu](mailto:sunithak@udel.edu)

**A**



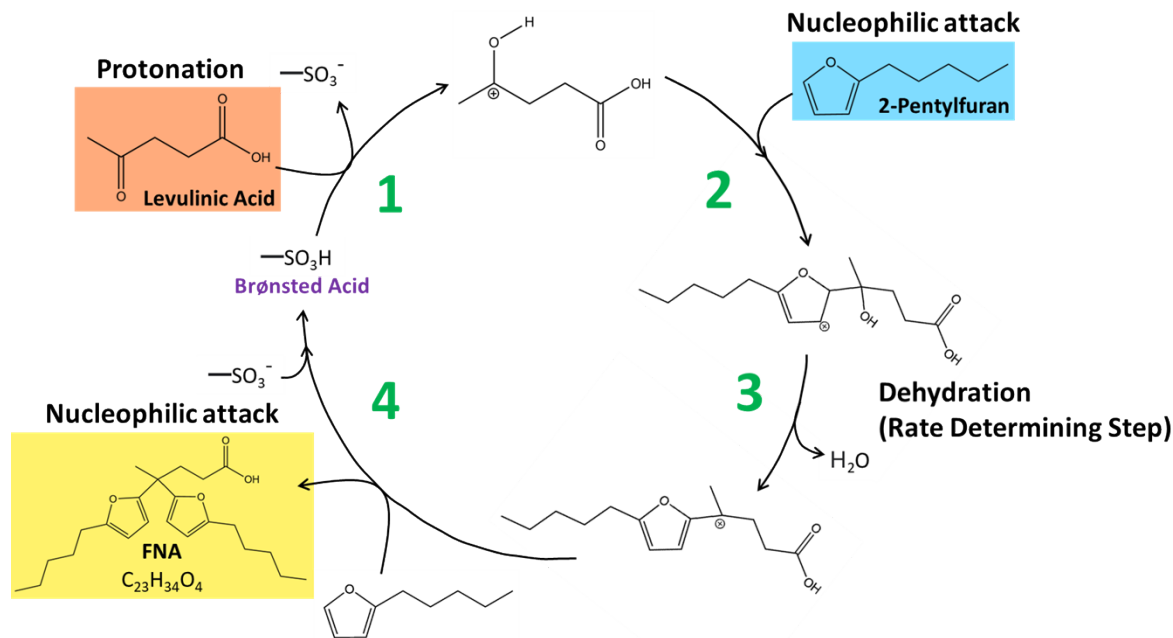
**B**



**Figure S1.** The synthesis of furan, 2-methylfuran, 2-alkylfuran, levulinic acid, and pyruvic acid.

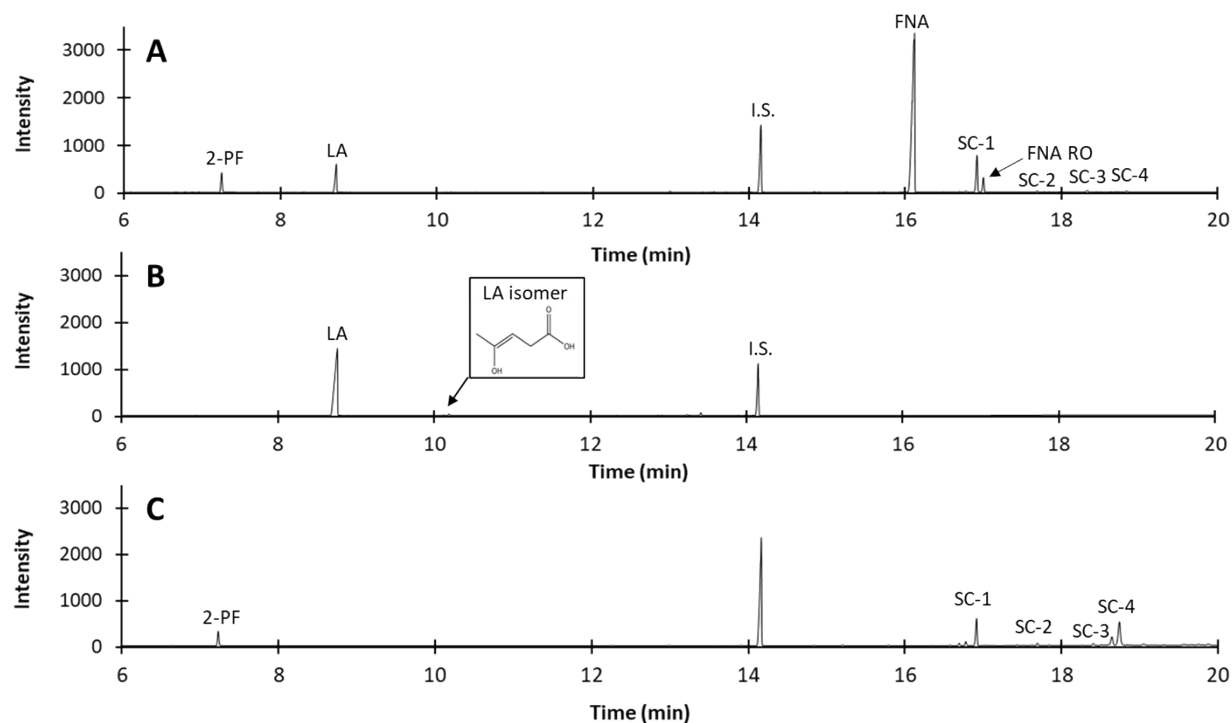
**Table S1.** Summary of catalyst screening. Reaction conditions: 20 mmol 2-pentylfuran (2-PF), 10 mmol levulinic acid (LA), 0.107 mmol H<sup>+</sup> catalyst, 65°C, 6 h, 800 rpm.

Compound	Catalyst									
	APW79S	SiO <sub>2</sub> - Al <sub>2</sub> O <sub>3</sub> (Si/Al=5)	H-Y (Si/Al=30)	Al-MCM- 41 (Si/Al=25)	APW98	PTA	PMA	H <sub>2</sub> SO <sub>4</sub>	CH <sub>3</sub> SO <sub>3</sub> H	P-TSA
2-PF conversion (%)	99.0	18.9	23.0	19.6	81.8	99.9	97.0	76.2	89.7	91.2
LA conversion (%)	82.1	9.4	3.8	53.7	66.3	84.0	77.6	65.8	66.2	66.4
FFA yield (%)	87.3	0.1	1.4	0.3	59.8	76.1	65.5	62.8	59.9	60.8
PFSCs yield (%)	6.2	0.9	1.0	0.6	5.8	7.8	15.8	5.2	5.2	6.9
FFA RO yield (%)	2.8	0.1	0.0	0.1	1.0	6.0	2.1	2.0	6.4	7.5
Total C Balance (%)	100.9	84.2	83.5	73.9	88.2	93.4	90.7	96.1	86.9	89.4



**Scheme S1.** Reaction mechanism for the condensation of 2-pentylfuran with levulinic acid.

The mechanism of HAA has been well studied and the HAA reaction mechanism of 2-alkylfuran with ketone and levulinic acid with phenol has been reported.<sup>1,2</sup> A plausible HAA reaction mechanism of 2-PF with LA (**Scheme S1**) over Aquivion PW79S was deduced considering the properties of the catalyst and the products. The hydroxyalkylation step is initiated by the protonation of the ketone group on LA by the catalyst, resulting in a carbocation intermediate (1). Then, the carbocation intermediate is attacked by a 2-PF molecule, forming the monomer (2). The alkylation step proceeds with the dehydration of the monomer, which entails proton transfer from furan ring to the carbonyl oxygen of the ketone and is the rate-determining step, resulting in another carbocation (3). Subsequently, the carbocation is alkylated by another 2-PF molecule, yielding the dimer or FFA (4). The carboxyl group on LA does not participate in the reaction as the presence of extra oxygen allows for delocalization of electrons, which provides extra stability.



**Figure S2.** GC chromatogram overlay of the A) HAA chemistry between 2-PF and LA and blank experiments of B) 2-Pentylfuran and C) Levulinic Acid.

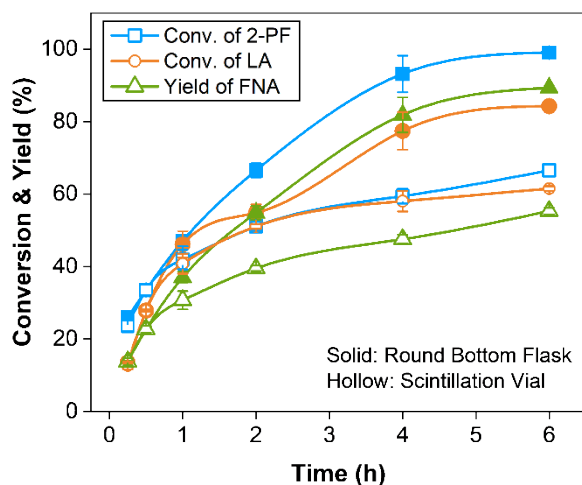
**Table S2.** Initial HAA reaction rate data.  $E_a = 34.3$  kJ/mol (M = molar). The formation rate refers to FNA.

Temperature (°C)	$r_{\text{consumption}}$ (M/s)		$r_{\text{formation}}$ (M/s)
	2-PF	LA	
35	0.051	0.057	0.036
45	0.130	0.038	0.062
55	0.196	0.043	0.081

**Table S3.** Properties of commercial solid acid catalysts.

Entry	Catalyst	Acid Density (H <sup>+</sup> mmol/g)	Surface Area (m <sup>2</sup> /g)	Pore Diameter (nm)	Form	Reference
1	Aquivion PW79S	1.26	<1	-	Powder	3
2	Aquivion PW98	1	<1	-	Powder	4
3	Phosphotungstic Acid	1.04	<1	-	Powder	5,6
4	Phosphomolybdic	1.5	1-5	-	Powder	7,8

	Acid					
5	HY	0.31	780	~0.7	Powder	<sup>9</sup>
6	Silica-Alumina	0.34	569	5.4	Powder	<sup>9</sup>



**Figure S3.** Effect of geometry. Reaction conditions: 20 mmol 2-pentylfuran (2-PF), 10 mmol levulinic acid (LA), 0.107 mmol H<sup>+</sup> (0.085 g) Aquivion PW79S, 65°C, 800 rpm. (RBF = round bottom flask, SV = scintillation vial).

### ***Reaction optimization using active learning***

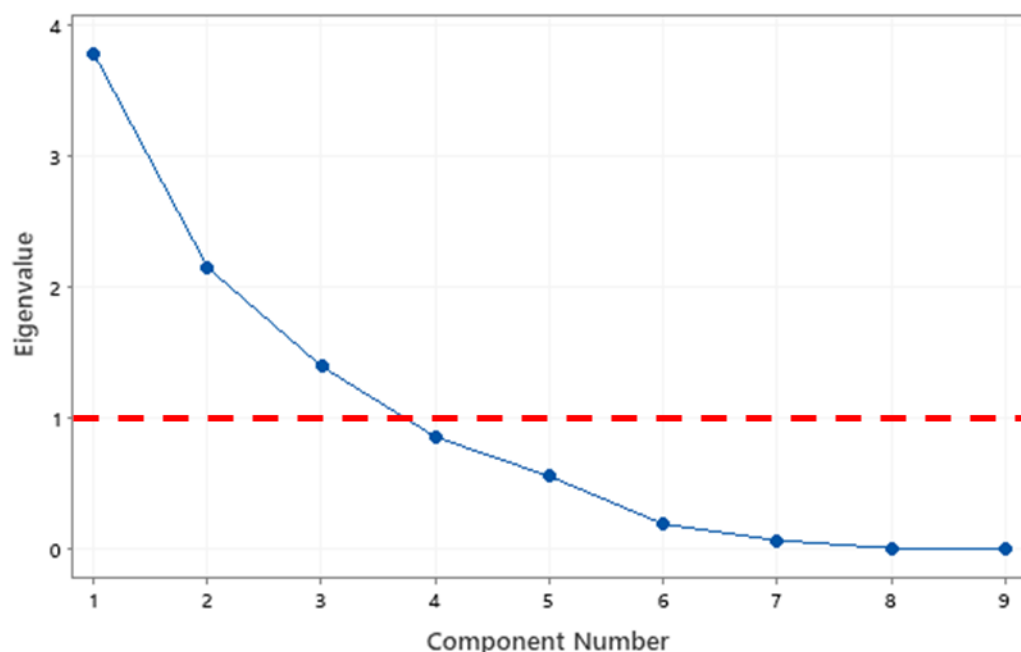
The yield of the products was optimized using active learning (see Methods). 16 points using a Latin hypercube sampling (LHS) were selected with varying temperature, reaction time, molar ratio, and catalyst loading. The LHS was selected due to its ability to sample efficiently a multidimensional parameter space. NEXTorCh then processes the data and predicts optima where 4 subsequent sampling points are generated per iteration. The optimum overall FNA and FNA RO yield found in experiments vs. iteration number is plotted in **Figure S5A**. The data is listed in **Table S4**. The yield increased at the first iteration but decreased in the third and fourth iterations. This occurs as the algorithm suggests an optimum around the edge points where one of the input parameters takes its extreme value. In the following iterations, since the algorithm has sampled enough in the region (i.e., exploitation), it explores other parts to improve the model's overall accuracy. The decrease in the optimum yield of iterations three and four suggests an exploration with no increased product. The optimum neo-acid precursors on the first iteration require more catalyst and longer reaction time compared to the initial LHS point (**Table S4**). This process highlights the utility of a data-driven approach for optimizing product yield. If a traditional central composite DOE with four factors were used, at least 80 experiments would have been needed, which might still not cover the true optimum. NEXTorCh efficiently reduced experimental time (<30 runs to identify an optimum) and consumables.

**Table S4.** Summary of multi-parameter optimization using NEX Torch. First 15 runs are the initial dataset. 4 additional runs added at each iteration. Parameters varied = temperature, time, catalyst loading, and molar ratio.

Exp	Trial	T (C)	Catalyst Loading (mmol H+)	Time (h)	Molar Ratio (mol/mol)	2-PF Conv. (%)	LA Conv. (%)	FNA Yield (%)	FNA RO Yield (%)	FNA + FNA RO Yield (%)	PFSCs Yield (%)	Total C Balance (%)
0	0	76	0.196	4	2.77	99.1	98.2	77.8	5.6	83.3	10.4	94.8
1	0	69	0.13	11	1.80	99.6	77.7	84.7	5.2	90.0	5.21	100.7
2	0	63	0.174	1	1.96	57.9	49.9	49.9	0.3	50.2	3.41	97.5
3	0	96	0.253	11	2.36	99.8	93.9	65.8	8.1	74.0	8.9	84.2
4	0	86	0.094	10	2.23	99.6	88.9	79.2	6.1	85.3	6.3	94.1
5	0	57	0.123	7	1.55	98.1	63.4	83.0	1.5	84.5	4.5	99.9
6	0	98	0.148	10	2.43	99.7	92.4	68.5	7.3	75.9	7.9	85.4
7	0	75	0.071	5	1.48	99.8	62.4	80.4	4.5	85.0	4.59	100.0
8	0	83	0.103	8	1.10	99.9	54.7	61.0	8.8	69.8	3.2	88.2
9	0	52	0.226	3	1.63	79.9	59.1	67.7	0.5	68.1	4.2	97.6
10	0	93	0.052	2	2.88	96.9	89.1	69.8	3.7	73.5	11.1	89.0
11	0	56	0.179	9	1.38	99.4	61.1	81.8	2.4	84.2	4.3	100.0
12	0	70	0.258	6	2.57	99.4	98.0	80.8	5.9	86.7	8.7	96.2
13	0	60	0.294	4	1.16	99.8	54.7	78.2	3.7	81.9	4.9	101.6
14	0	88	0.21	7	2.70	99.7	96.3	72.7	8.0	80.7	9.7	91.3
15	0	65	0.107	6	2.00	99.0	82.1	87.3	2.8	90.1	6.2	100.9
16	1	65	0.122	8	2.09	99.5	82.7	83.9	3.4	87.4	5.4	96.8
17	1	66	0.051	7	1.99	98.0	72.2	79.4	1.5	80.9	5.6	94.0
18	1	66	0.170	7	1.98	99.8	81.8	86.2	4.9	91.1	6.0	101.3
19	1	58	0.120	7	2.34	91.8	81.4	78.4	1.3	79.8	5.3	95.2
20	2	67	0.182	6	1.82	99.9	72.8	79.19	4.4	83.6	6.2	96.2
21	2	63	0.209	10	2.06	99.9	84.2	81.36	5.9	87.3	6.7	97.4
22	2	59	0.263	7	1.76	99.9	72.3	81.19	4.3	85.5	5.9	98.1
23	2	62	0.083	11	2.23	99.2	85.2	84.25	2.7	87.0	6.1	96.7
24	3	55	0.114	11	1.75	97.0	63.6	75.3	1.3	76.6	4.9	92.5
25	3	73	0.069	12	2.17	99.4	80.2	77.5	3.3	80.7	5.7	90.9
26	3	73	0.115	6	2.23	99.6	83.6	78.4	3.9	82.3	6.4	92.3
27	3	70	0.126	10	2.28	99.7	84.5	76.7	5.2	81.9	6.9	92.1

### ***Understanding the correlation between parameters via machine learning***

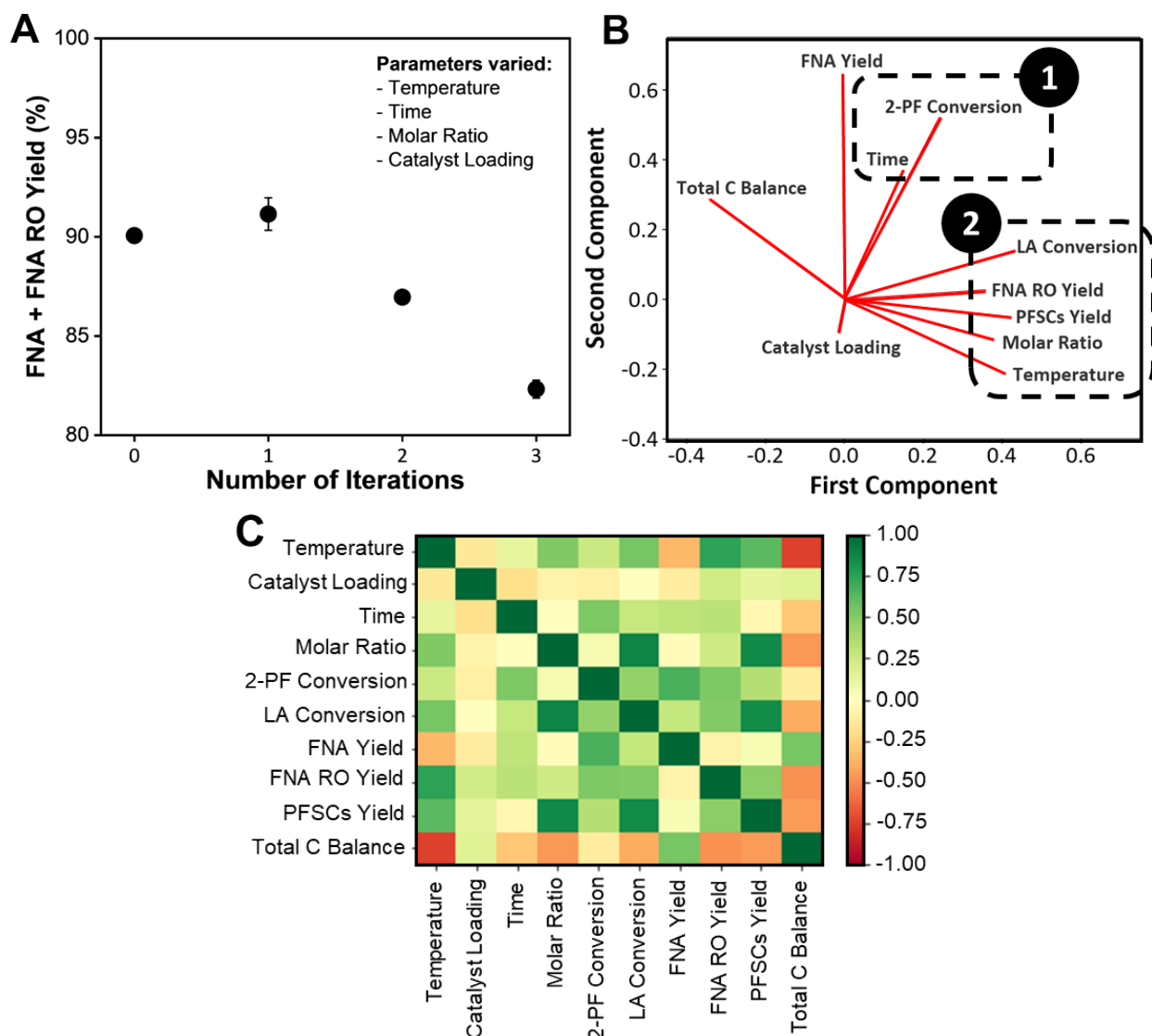
Principal component analysis (PCA) was performed to understand the effect of reaction conditions on product distribution. PCA reduces dimensionality, identifies independent factors, and reveals correlations and significance of factors.<sup>10,11</sup> Our dataset consists of 10 features; four are reaction conditions (temperature, time, molar ratio, and catalyst loading) and the rest entail conversions and yields (2-PF conversion, LA conversion, FNA yield, FNA RO yield, PFSCs yield, and total carbon balance). The scree plot (**Figure S4**) shows that three principal components explain 75% of the variation and are sufficient for exploratory analysis. Generally, factors clustered together show a strong positive correlation, and those orthogonal to each other show little or no correlation. Factors on opposite sides are inversely correlated.



**Figure S4.** Scree plot of PCA analysis.

The two principal components that account for most of the variability are shown in the loading plot (**Figure S5B**). The results clearly show two clusters: group 1 includes 2-PF conversion and time, and group 2 consists of the LA conversion, FNA RO yields, PFSCs yield, molar ratio, and temperature. 2-PF conversion is positively correlated with the reaction time. Additionally, there is a clear correlation between FNA yield and 2-PF conversion with reaction time; the longer the reaction time, the more 2-PF converts and the more FNA forms. The reaction time (group 1) is nearly antiparallel with catalyst loading and total carbon balance, indicating an inverse relation between them, *i.e.*, as the catalyst loading increases, less reaction time is required and as the reaction time increases, the reactants and products participate in side-reactions forming byproducts, resulting in the lower carbon balance. This fact indicates a tradeoff between increasing catalyst loading and conversion of 2-PF.





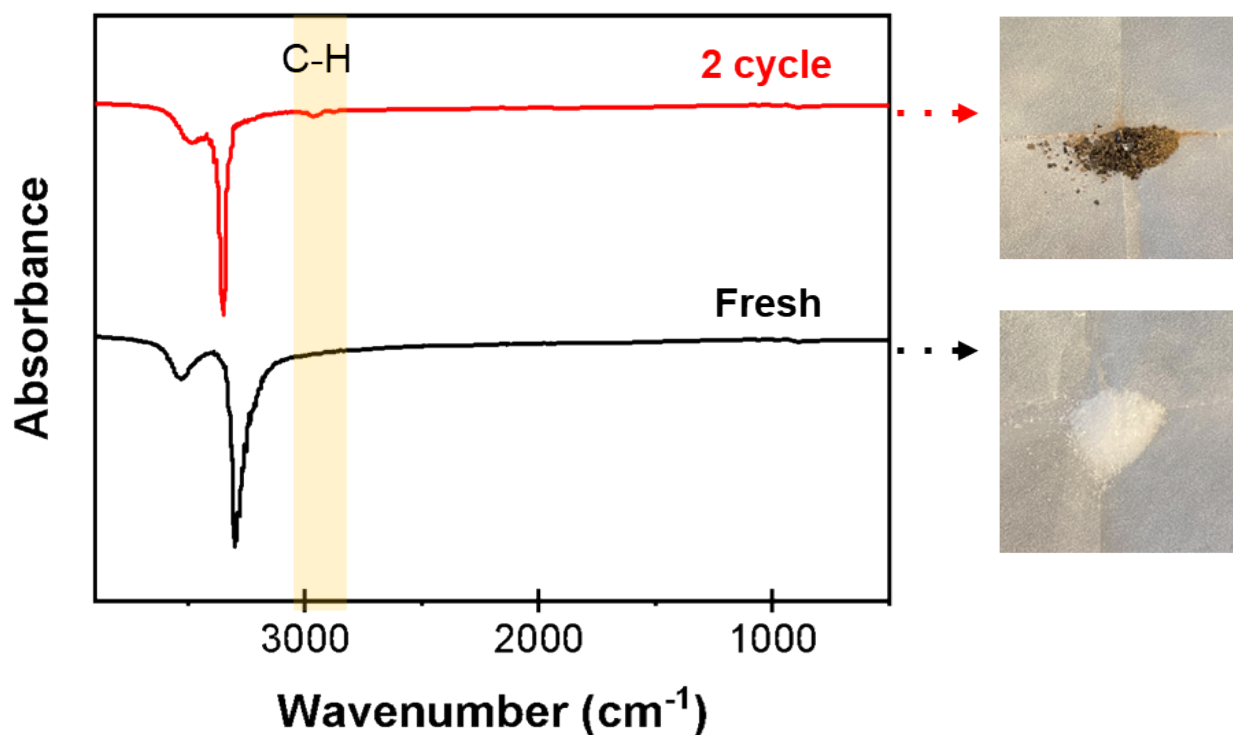
**Figure S5. Multi-parameter optimization using NEX Torch for the HAA reaction.** A) Optimal overall FNA and FNA RO yield (%) from experiments in each iteration. The error bar indicates the standard deviation from 2 measurements. B) Principal component analysis of the correlations between reaction conditions and products. C) Correlation matrix of reaction conditions and outcomes. The green and red colors indicate strong positive and negative correlations between two features, respectively.

The reaction time is almost orthogonal with LA conversion and FNA RO and PFSCs yields, *i.e.*, it does not affect them. These phenomena occur due to (a) LA is an electrophile and its conversion is dependent on the reactivity of the nucleophile, (b) the formation of FNA RO is due to participation of the water by-product, and (c) the formation of PFSCs is due to 2-PF's tendency to self-condense due to its reactivity. Catalysts that could suppresses the self-condensation of 2-PF will be important for future work. LA conversion and FNA RO and PFSCs yields are positively correlated with temperature and reactant's molar ratio. Regarding LA, the high conversion could be due to forming angelica lactones<sup>12</sup> rather than FNA. Maximov *et al.* reported the formation of angelica lactones from levulinic acid in over an acidic catalyst at 80 – 120°C. FNA yield is

inversely correlated with temperature, and temperature is inversely correlated (strongly) to total C balance, as well as the rest of cluster 2 (weakly). As temperature increases, FNA participates in undesirable side reactions resulting in the lower carbon balance. This fact suggests that a moderate temperature is preferable for HAA chemistries. With the data from multi-parameter optimization, we generated a correlation matrix to further clarify interactions between variables (**Figure S5C**). The resulting matrix agrees with the observations deduced from the PCA results shown in **Figure S5B**.

**Table S5.** Detailed experimental results of HAA reaction between different 2-alkylfurans and keto acid of varying molecular sizes. Reaction Conditions: 20 mmol 2-alkylfuran, (2-alkylfuran/keto acid) (mol/mol) = 2, 0.107 mmol H<sup>+</sup> Aquivion PW79S, 6 h, 65°C, 800 rpm.

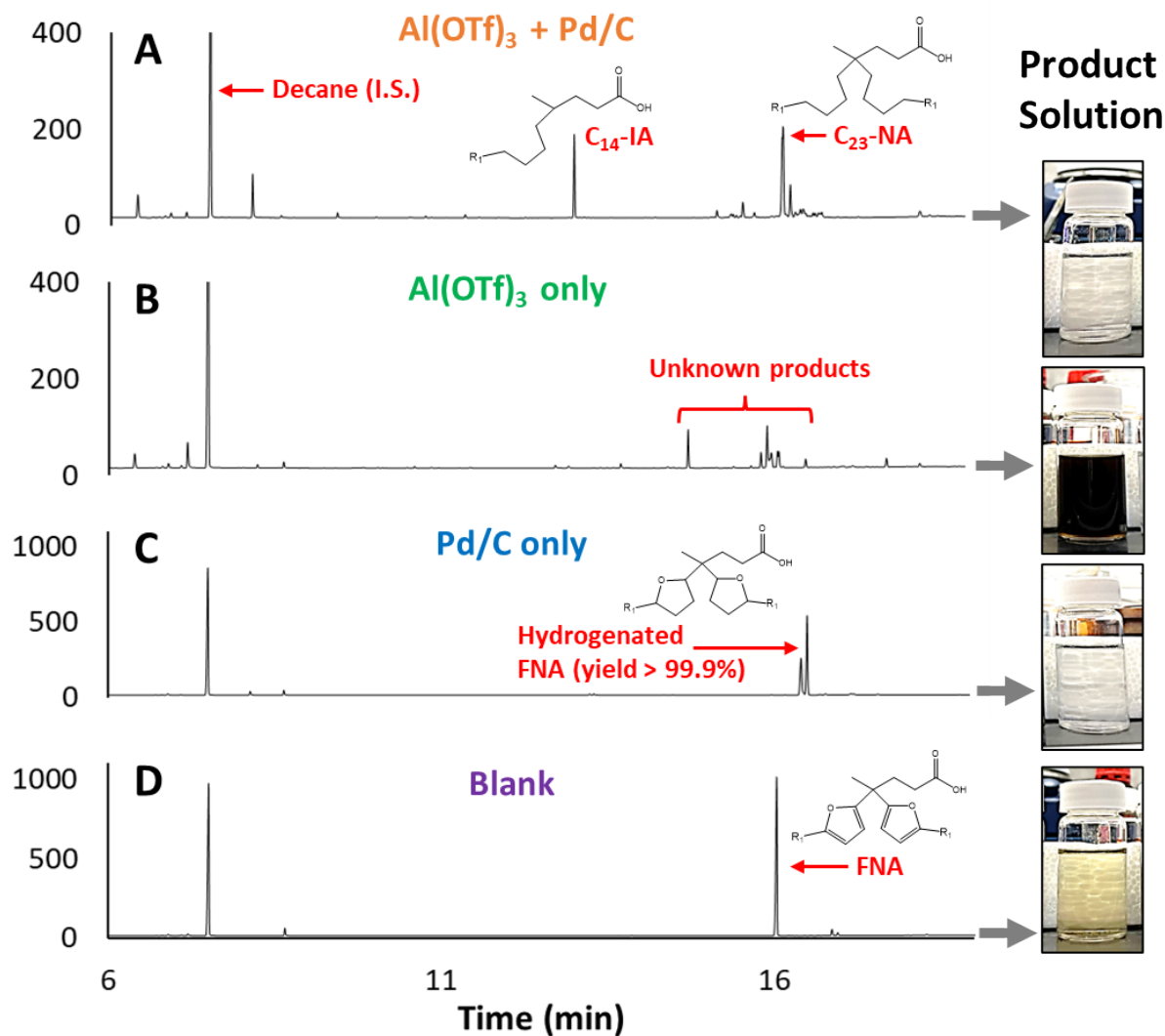
Entry	Reagents		#C	R Conv. (%)	Keto Acid Conv. (%)	Yield FNA & FNA RO (%)	Total C Balance (%)	% Error of FNA & FNA RO Yield
	R	Keto Acid						
1	Methyl	LA	C15	96.5	58.4	47.3	75.1	0.9
2	Ethyl	LA	C17	99.8	69.4	76.7	94.4	1.2
3	n-Propyl	LA	C19	95.7	78.6	86.9	101.9	0.6
4	n-Butyl	LA	C21	99.9	78.1	74.9	87.1	2.0
5	n-Pentyl	LA	C23	99.5	83.6	90.1	100.9	0.1
7	n-Heptyl	LA	C27	99.3	83.0	82.0	89.4	0.2
8	Methyl	PA	C13	97.4	95.8	85.6	93.2	0.4
9	Ethyl	PA	C15	99.2	99.7	87.2	89.6	2.9
10	n-Propyl	PA	C17	97.9	99.6	91.9	95.1	0.4
11	n-Butyl	PA	C19	99.9	99.6	89.6	92.2	0.9
12	n-Pentyl	PA	C21	98.4	99.8	85.5	89.0	2.5
14	n-Heptyl	PA	C25	96.9	99.2	79.2	84.0	0.2



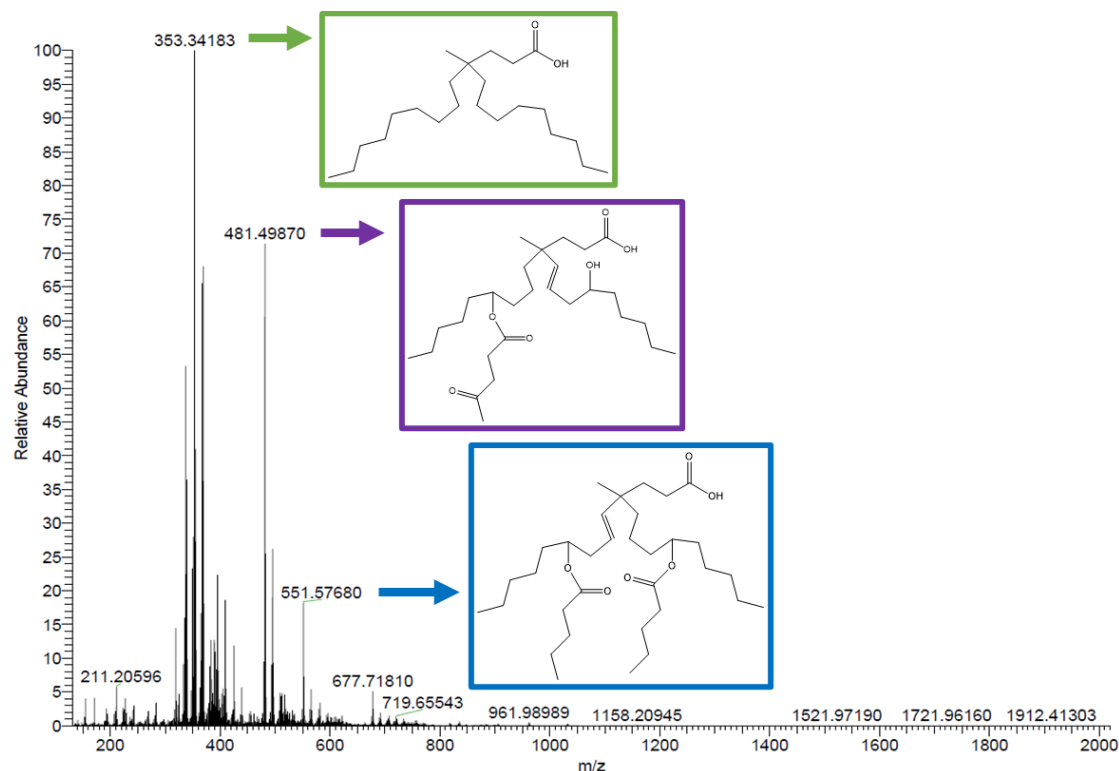
**Figure S6.** FT-IR spectra of fresh and used Aquivion PW79S catalysts.

**Table S6.** The effect of catalysts on product distribution. Reaction conditions: 1 mmol FNA, 6 mol% Al(OTf)<sub>3</sub>, 2 mol% Pd/C, 10 mL cyclohexane, 3 MPa H<sub>2</sub>, 180°C, 30 min and 750 rpm.

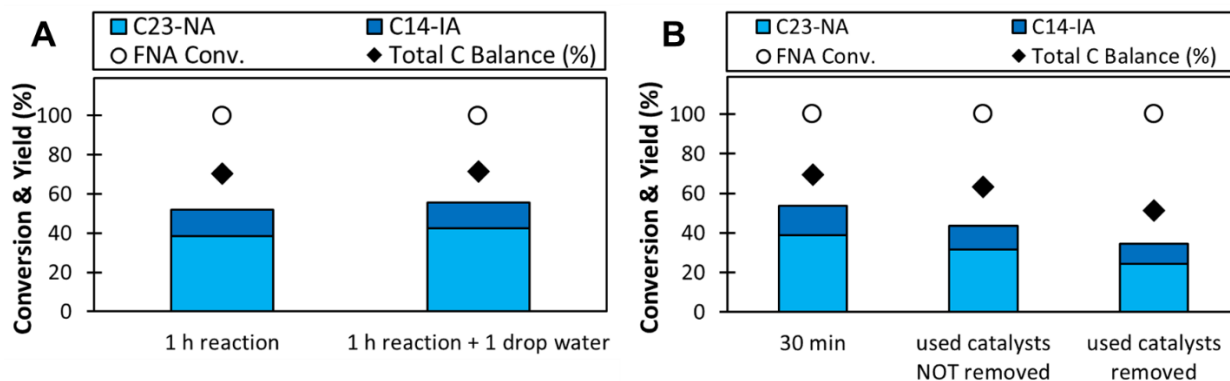
Catalyst(s)	FNA + FNA RO Conversion (%)	C23-NA yield (%)	C14-IA yield (%)	Hydrogenated FNA yield (%)
Al(OTf) <sub>3</sub> + Pd/C	100	40	15	0
Al(OTf) <sub>3</sub> only	91	0	0	0
Pd/C only	100	0	0	100
Blank (no catalysts)	0	0	0	0



**Figure S7.** GC overlay for the HDO product mixture obtained with and without a catalyst. A) HDO reaction over  $\text{Al}(\text{OTf})_3 + \text{Pd/C}$  catalysts, B) HDO reaction over  $\text{Al}(\text{OTf})_3$  only (without hydrogenation catalyst), C) HDO reaction over  $\text{Pd/C}$  only (without metal triflate), D) HDO reaction without any hydrogenation catalyst and metal triflate (Blank reaction). Reaction conditions: 1 mmol FNA, 6 mol%  $\text{Al}(\text{OTf})_3$ , 2 mol%  $\text{Pd/C}$ , 10 mL cyclohexane, 3 MPa  $\text{H}_2$ , 180°C, 1 h, 500 rpm.

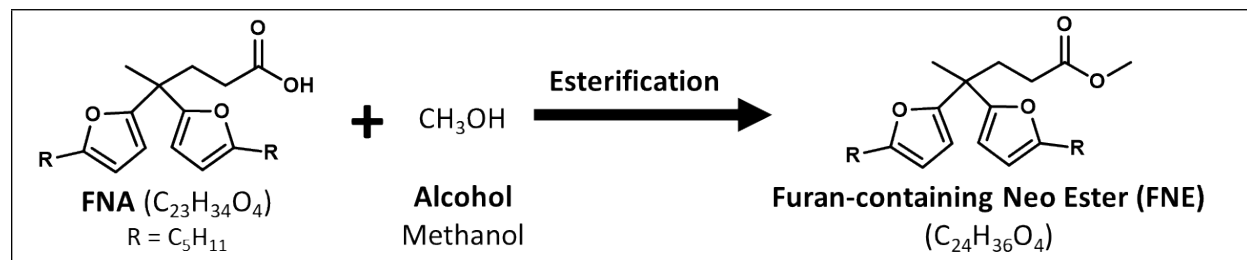


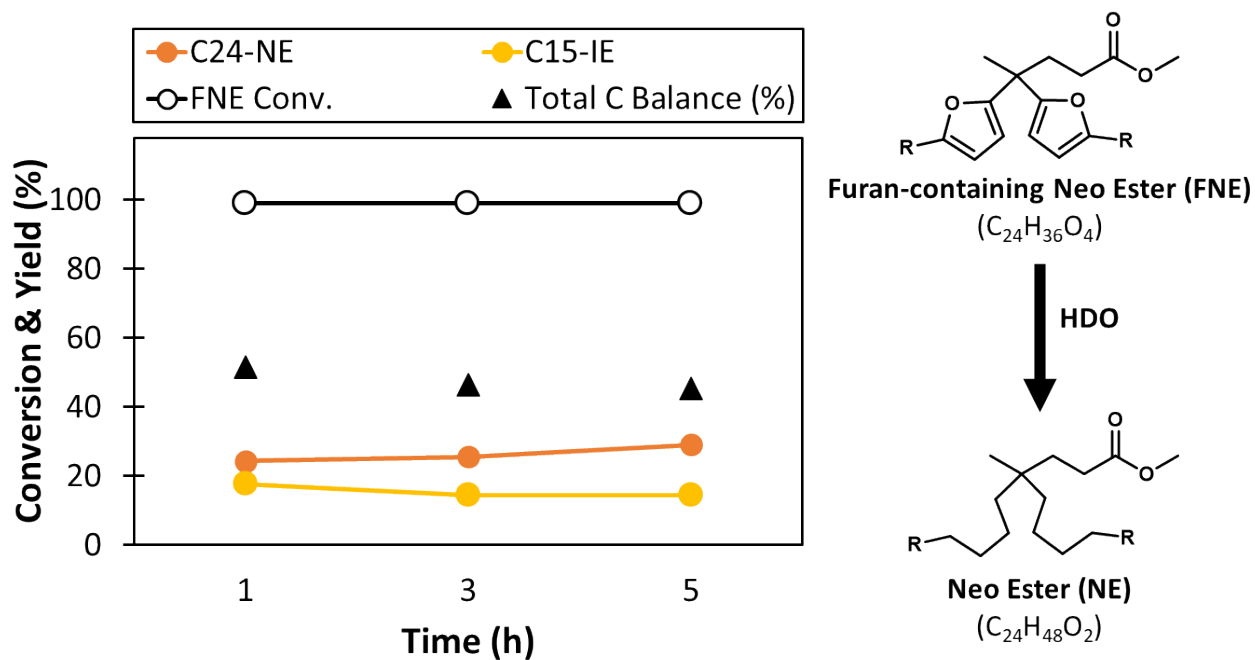
**Figure S8.** LCMS of HDO product solution. Reaction conditions: 1 mmol FNA, 6 mol% Al(OTf)<sub>3</sub>, 2 mol% Pd/C, 10 mL cyclohexane, 3 MPa H<sub>2</sub>, 180°C, 1 h, 500 rpm.



**Figure S9.** Effect of A) water addition. Reaction conditions: 1 mmol FNA, 6 mol% Al(OTf)<sub>3</sub>, 2 mol% Pd/C, 10 mL cyclohexane, 3 MPa H<sub>2</sub>, 180°C, 1 h, and 750 rpm. 1 drop of water = ~40 mg, and B) used catalyst removal. Reaction conditions: 1 mmol FNA, 6 mol% Al(OTf)<sub>3</sub>, 2 mol% Pd/C, 10 mL cyclohexane, 3 MPa H<sub>2</sub>, 180°C, 30 min, and 750 rpm. The reaction mixture was then added into a new reactor with fresh catalysts and reacted for 30 min.

**Table S7.** Esterification of FNA with alcohol (methanol) making FNE. FNA conversion and FNE purity upon esterification of FNA with alcohol. Reaction conditions: 4 g FNA, 16 mL methanol, 2 MPa N<sub>2</sub>, 200°C, 4 h.

		
<b>Compound</b>	<b>Conversion (%)</b>	<b>Purity (%)</b>
FNA	>99.9	-
FNE	-	93.5



**Figure S10.** Time-dependent study of FNE ring-opening. Reaction conditions: 1 mmol FNE, 6 mol% Al(OTf)<sub>3</sub>, 2 mol% Pd/C, 10 mL cyclohexane, 3 MPa H<sub>2</sub>, 180°C, and 750 rpm.

## References

1. S. Liu, R. Bhattacharjee, S. Li, A. Danielson, T. Mazal, B. Saha and D. G. Vlachos, *Green Chem.*, 2020, **22**, 7896–7906.
2. S. Van de Vyver, S. Helsen, J. Geboers, F. Yu, J. Thomas, M. Smet, W. Dehaen, Y. Román-Leshkov, I. Hermans and B. F. Sels, *ACS Catal.*, 2012, **2**, 2700–2704.
3. S. Liu, B. Saha and D. G. Vlachos, *Green Chem.*, 2019, **21**, 3606–3614.
4. A. Karam, K. De Oliveira Vigier, S. Marinkovic, B. Estrine, C. Oldani and F. Jérôme, *ACS Catal.*, 2017, **7**, 2990–2997.
5. A. Kuvayskaya, *Electronic Theses and Dissertations*, 2020, Paper 3738.
6. S. He, Y. Ai, W. Dai, S. Zhai, H. Song and J. Lin, *Polym. Test.*, 2021, **100**, 107246.
7. R. S. Salama, S. M. Hassan, A. I. Ahmed, W. S. A. El-Yazeed and M. A. Mannaa, *RSC Adv.*, 2020, **10**, 21115–21128.
8. H. C. Zhou, J. L. Song, Q. L. Meng, Z. H. He, Z. W. Jiang, B. W. Zhou, H. Z. Liu and B. X. Han, *Green Chem.*, 2016, **18**, 220–225.
9. R. Weingarten, G. A. Tompsett, W. C. Conner and G. W. Huber, *J. Catal.*, 2011, **279**, 174–182.
10. J. Lever, M. Krzywinski and N. Altman, *Nat. Methods*, 2017, **14**, 641–642.
11. A. J. Medford, A. J., M. R. Kunz, S. M. Ewing, T. Borders and R. Fushimi, *ACS Catal.*, 2018, **8**, 7403–7429.
12. A. L. Maximov, A. V. Zolotukhina, A. A. Mamedli, L. A. Kulikov and E. A. Karakhanov, *ChemCatChem*, 2017, **10**, 222–233.