Fully renewable polyesters *via* polycondensation catalyzed by *Thermobifida cellulosilytica* cutinase: an integrated approach

ELECTRONIC SUPPLEMENTARY MATERIAL

Complete protocol for the calculation of the BioGPS model for the 42 Ser Hydrolases using Unsupervised Pattern Cognition Analysis (UPCA)

Active sites were mapped using the GRID force field [1] for evaluating the type and the energy of non-bonded interactions and then for generating the pseudo-MIFs (Molecular Interaction Fields). Four different probes were employed: H probe takes into account the active site shape; O probe that evaluates H-bond donor properties; N1 probe that evaluates the H-bond acceptor capabilities; the DRY probe accounting for hydrophobic interactions. The magnitude of the interaction of the N1 and O probes includes, implicitly, also information about the charge contribution, since these probes have already a partial positive and negative charge respectively.

With the pseudo-MIF procedure, the mapped properties are considered as electron-density like fields centered on each atom, corresponding to specific probe types (i.e. the interaction energies coming from GRID N1 probe were centered on carbonyl oxygen as H-bond acceptor). Afterwards, the algorithm reduces the complexity of the pseudo-MIFs selecting a number of representative points using a weighted energy-based and space-coverage function.

For each active site the algorithm generated all possible combinations of four points; each combination is termed ‘‘quadruplet’’ (in mathematics, a tuple is a finite group of objects and a quadruplet is written as 4-tuple,). Moreover, the function includes the geometrical information into each quadruplet. All possible quadruplets for each mapped active site were generated and stored into a bio-fingerprint (bitstring) that constitutes the Common Reference Framework. For catching similarities and differences between two or more active sites, the algorithm compares their Common Reference Frameworks using an ‘‘all against all’’ approach where each enzyme active site is compared with itself and with all the other enzyme active sites; the algorithm searches for similar quadruplets and then overlaps the corresponding 3D structures. At the end, the algorithm generates a set of Tanimoto scores [2] (BioGPS descriptors) represented by square matrixes, namely a series of probe scores (one for each original GRID probe) together with a global score. The descriptors are calculated for a given superposition by directly comparing the overlapping volumes of the pseudo-MIFs.

The output is represented by different square matrixes which represent the BioGPS descriptors, namely a series of probe scores (one for each original GRID probe) together with a global score. The information contained in the BioGPS descriptors was then statistically analyzed by means of Unsupervised Pattern Cognition Analysis (UPCA) [3].

Ser hydrolases were sorted and grouped into clusters where structural properties, explained by the BioGPS descriptors, are correlated with catalytic functions (e.g. protease, lipase, esterase and amidase catalytic activity).

Unsupervised Pattern Cognition Analysis (UPCA), a well established algorithmic platform for performing systematic analysis of data sets, was used to reduce the dimensionality of data. The UPCA algorithm converts a set of correlated descriptors into a new set of linearly independent variables (orthogonal transformation) called Principal Components (PCs). Principal Components are simply a linear combination of the original correlated variables. The first Principal Component

---

[1] Goodford PJ (1985) A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. Journal of Medicinal Chemistry 28: 849–857.

[2] Rogers DJ, Tanimoto TT (1960) A computer program for classifying plants. Science 132: 1115–1118.

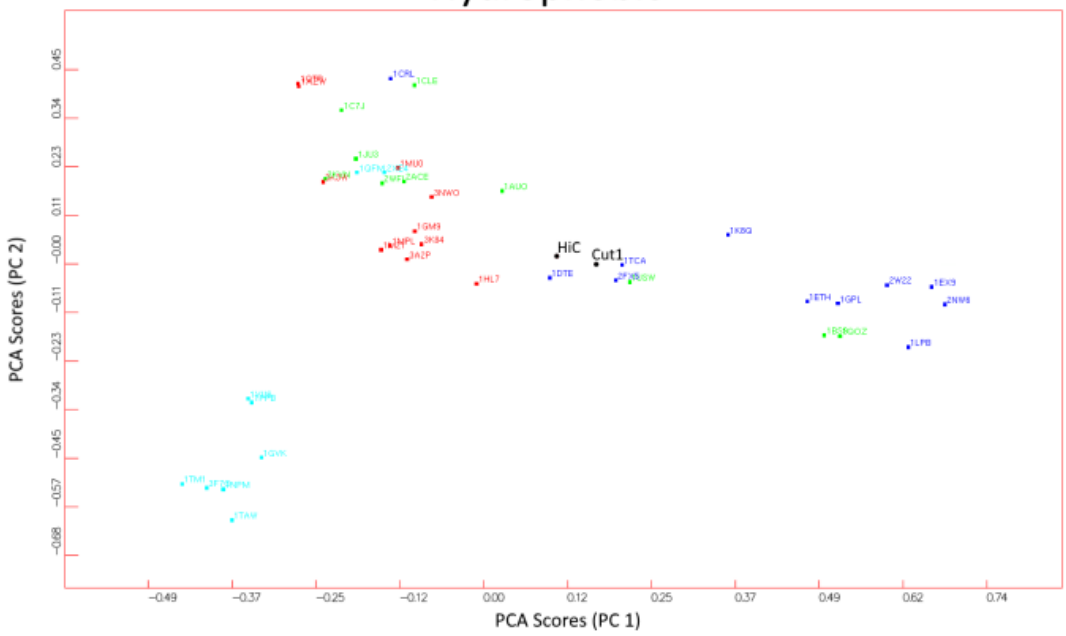
[3] Boutros PC, Okey AB (2005) Unsupervised pattern recognition: an introduction to the whys and wherefores of clustering microarray data. Briefings in Bioinformatics 6: 331–343.

(PC1) is calculated in order to maximize the variance of the object in the dataset. The following principal components are calculated to maximize the variance in the data that is not explained by the previous PC yet. UPCA can easily detect clusters of different active sites for capturing and quantifying differences between protein classes.

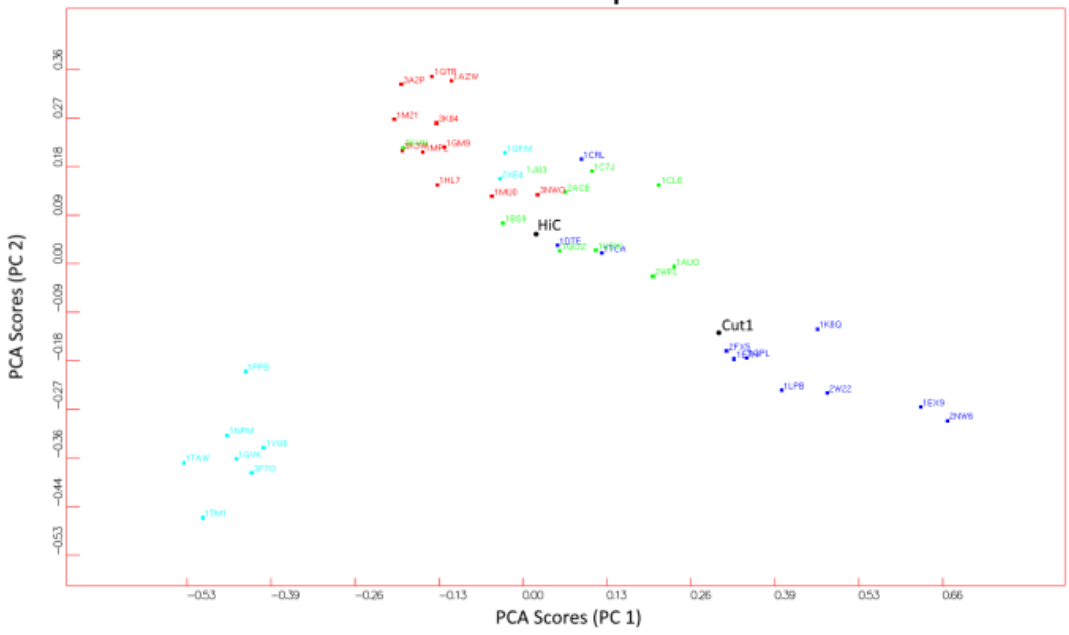
Table S1: Data set used for the computation of the BioGPS-UPCA model.

Enzyme class	PDB code	Source	Substrate
Lipases	1CRL	*Candida rugosa*	triacylglycerol
	1DTE	*Humicola lanuginosa*	triacylglycerol
	1ETH	*Sus scrofa*	triacylglycerol
	1EX9	*Pseudomonas aeruginosa*	triacylglycerol
	1GPL	*Cavia porcellus*	triacylglycerol
	1K8Q	*Canis lupus familiaris*	triacylglycerol
	1LPB	*Homo sapiens*	triacylglycerol
	1TCA	*Candida antarctica*	triacylglycerol
	2FX5	*Pseudomonas mendocina*	triacylglycerol
	2NW6	*Burkholderia cepacia*	triacylglycerol
	2W22	*Geobacillus thermocatenulatus*	triacylglycerol
Esterases	1AUO	*Pseudomonas fluorescens*	broad specificity
	1BS9	*Penicillium purpurogenum*	xylanes acetates
	1C7J	*Bacillus subtilis*	p-nitrobenzyl esters
	1CLE	*Candida cylindracea*	cholesterol esters
	1JU3	*Rhodococcus sp.*	cocaine
	1QOZ	*Tricoderma reesei*	xylanes acetates
	1USW	*Aspergillus niger*	feroloyl-polysaccharide
	2ACE	*Torpedo californica*	acetylcoline
	2WFL	*Rauvolfia serpentine*	polyneuridine aldehyde
	3KVN	*Pseudomonas aeruginosa*	rhamnolipids
Proteases	1GVK	*Sus scrofa*	Ala-|-Xaa
	1NPM	*Mus musclus*	Lys/Arg-|-Xaa
	1PPB	*Homo sapiens*	Arg-|-Gly fibrinogen
	1QFM	*Sus scrofa*	Pro-|-Xaa (~30aa)
	1TAW	*Bos Taurus*	Lys/Arg-|-Xaa
	1TM1	*Bacillus amyloliquefaciens*	uncharged P1
	1YU6	*Bacillus licheniformis*	uncharged P1
	2XE4	*Leshmania major*	olygopeptides
	3F7O	*Peacelomyces lilacinus*	peptides
Amidases	1AZW	*Xantomonas campestris*	NH-Pro-|-Xaa
	1GM9	*Escherichia coli*	penicillin
	1HL7	*Microbacterium sp.*	γ-lactam
	1M21	*Stenotrophomonas maltophilia*	C terminal amide
	1MPL	*Streptomyces sp.*	L-Lys-D-Ala-|-D-Ala
	1MU0	*Thermoplasma acidophilum*	NH-Pro-|-Xaa
	1QTR	*Serratia marcescens*	NH-Pro-|-Xaa
	3A2P	*Arthrobacter sp.*	6-amino exanoate dimer
	3K3W	*Alcaligens faecalis*	penicillin
	3K84	*Rattus norvegicus*	fatty acid amide
	3NWO	*Mycobacterium smegmatis*	NH-Pro-|-Xaa

Hydrophobic



H-Bond Acceptors



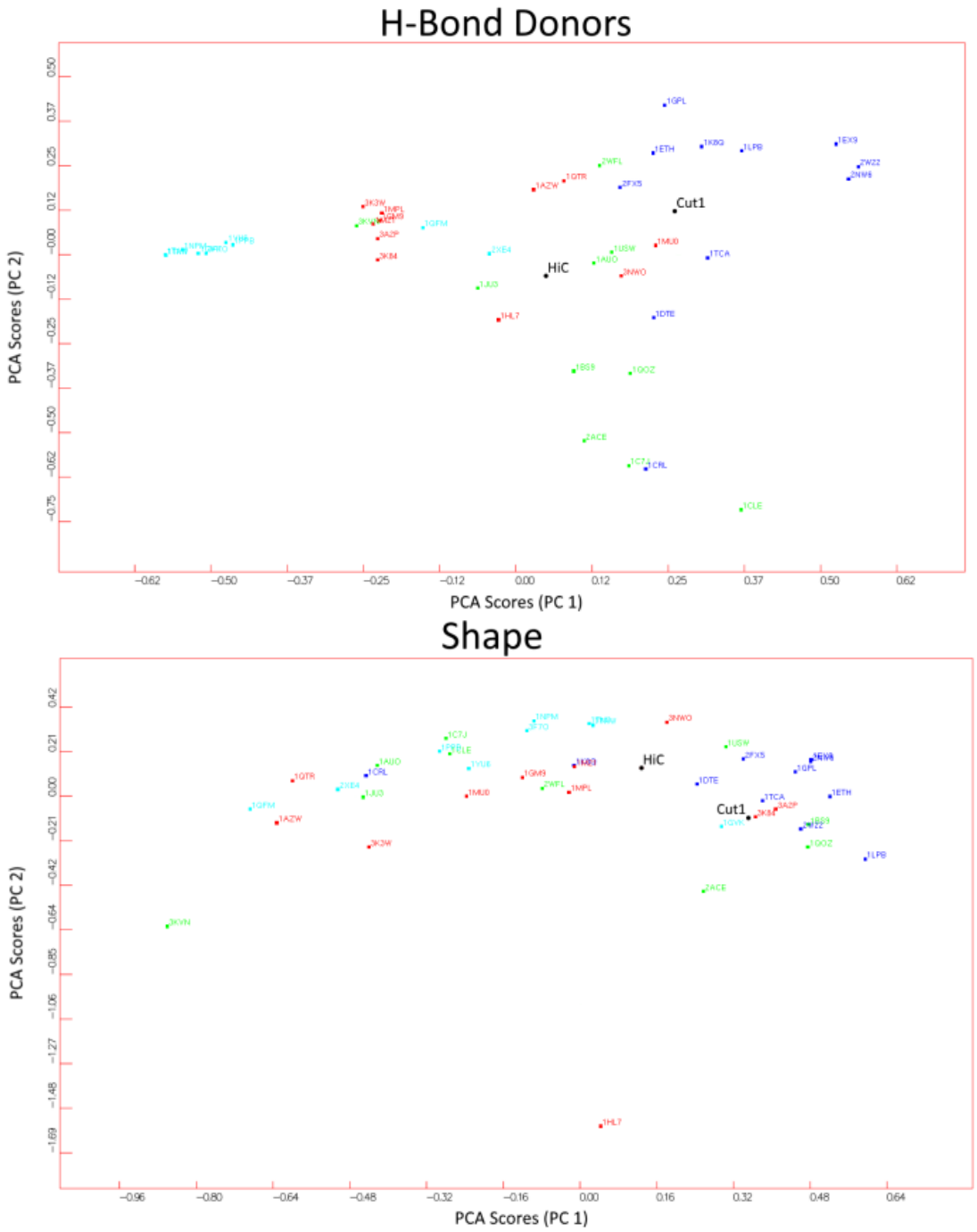


Figure S1: Projection of cutinases from *Thermobifida cellulosilytica* (Thc_cut1) and from *Humicola insolens* (HiC) in the BioGPS models calculated for each single probe through Unsupervised Pattern Cognition Analysis (UPCA). Ser hydrolases are clustered on the basis of BioGPS descriptors considering each single mapped property (single probe score) and labelled according to their PDB code. Lipases are indicated in blue, esterases in green, amidases in red and proteases in cyan.

Characterization of products obtained from the polycondensation of DMA and BDO catalyzed by Thc_cut1 immobilized on rice husk by adsorption and cross-linking (final water content: 4 % w/w).

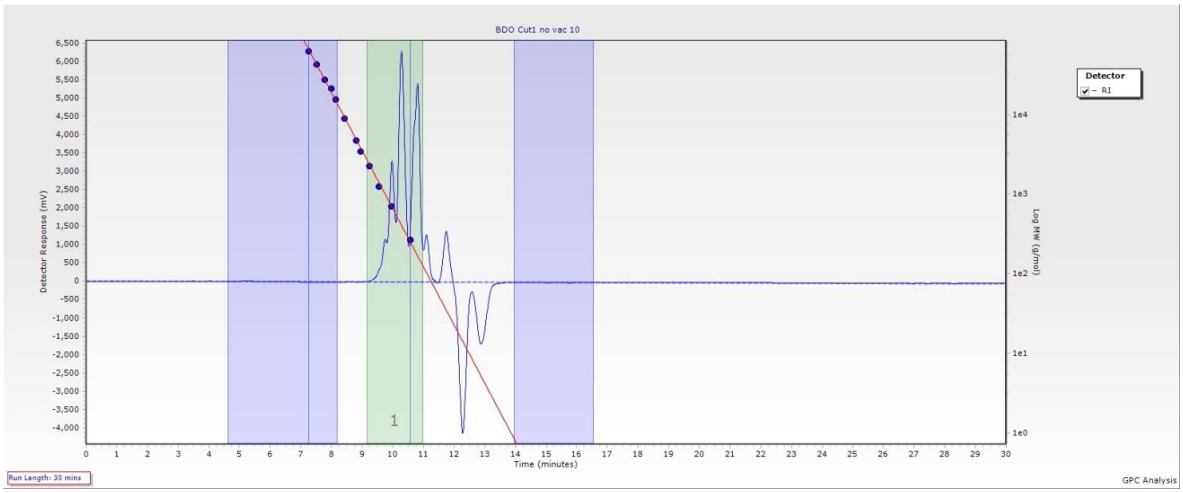


Figure S2. GPC chromatogram of the polycondensation products of DMA with BDO catalyzed by 10% w/w rTHC-Cut1 at 24 h and 1000 mbar. Mw= 594, Mn= 434, PD= 1.369. Entry: 1, Table 2.

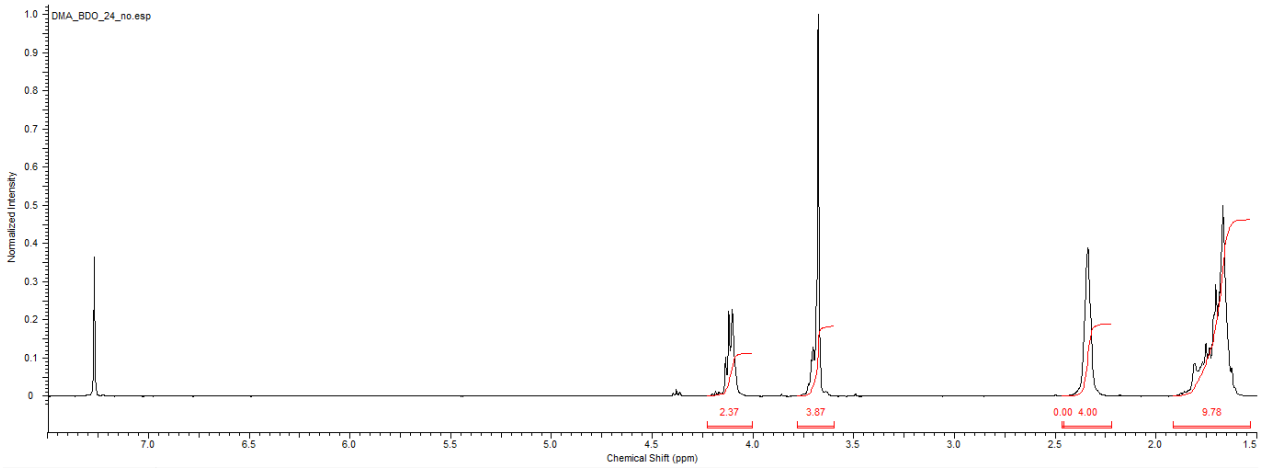


Figure S3. ^{1}H-NMR spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w rTHC-Cut1 at 24 h and 1000 mbar. Conversion: 64% . Entry 1, Table 2.

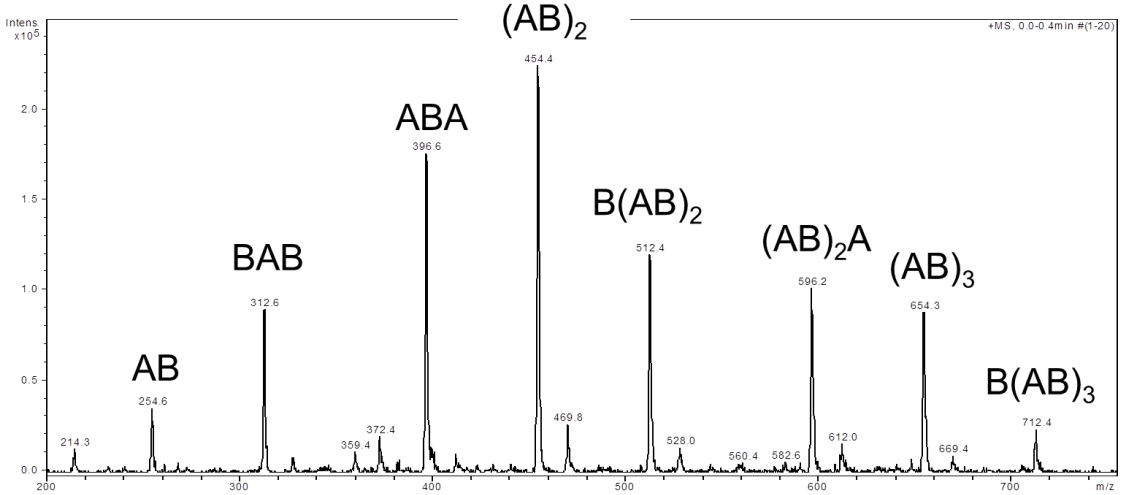


Figure S4. ESI-MS positive ion mass spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w rTHC-Cut1 at 24 h and 1000 mbar. Entry 1, Table 2.

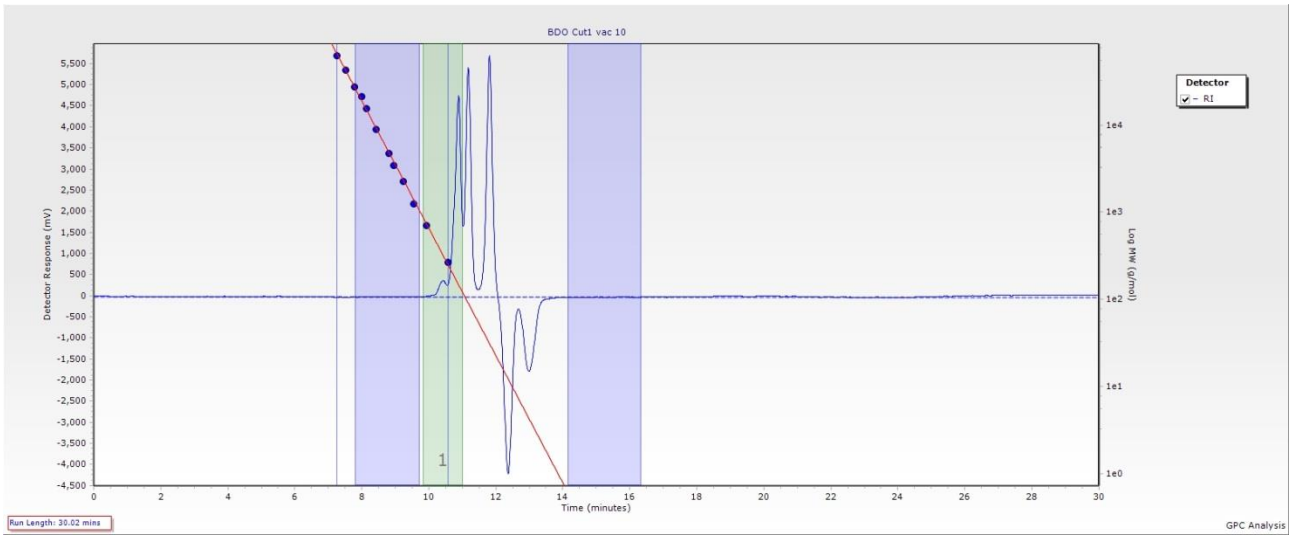


Figure S5. GPC chromatogram of the polycondensation products of DMA with BDO catalyzed by 10% w/w rTHC-Cut1 at 24 h and 70 mbar. Mw= 230, Mn= 222, PD= 1.037. Entry 2, Table 2.

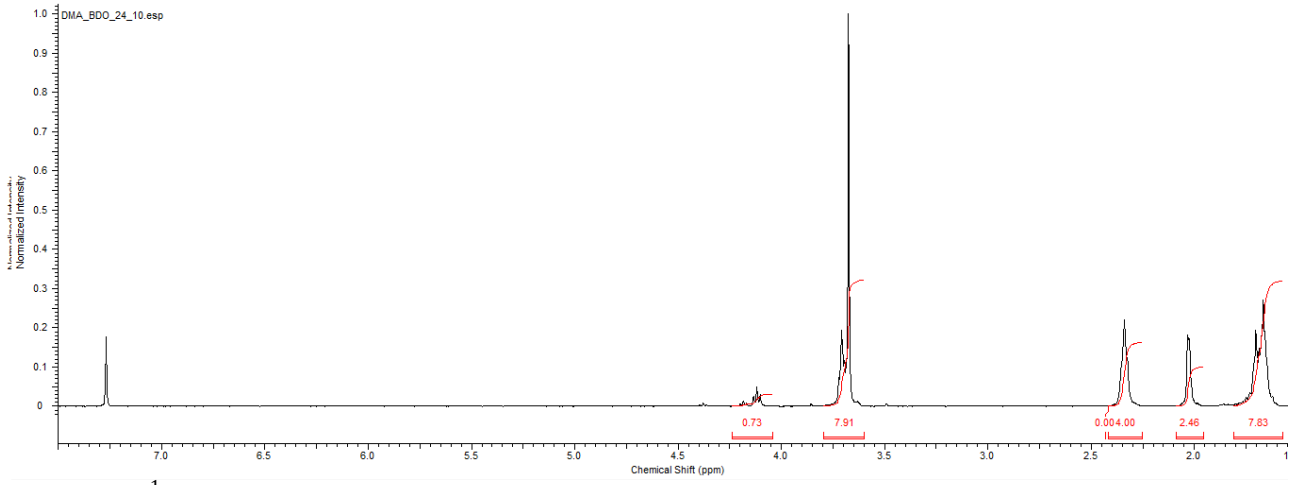


Figure S6. ^{1}H-NMR spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w rTHC-Cut1 at 24 h and 70 mbar.Conversion: 18% . Entry 2, Table 2.

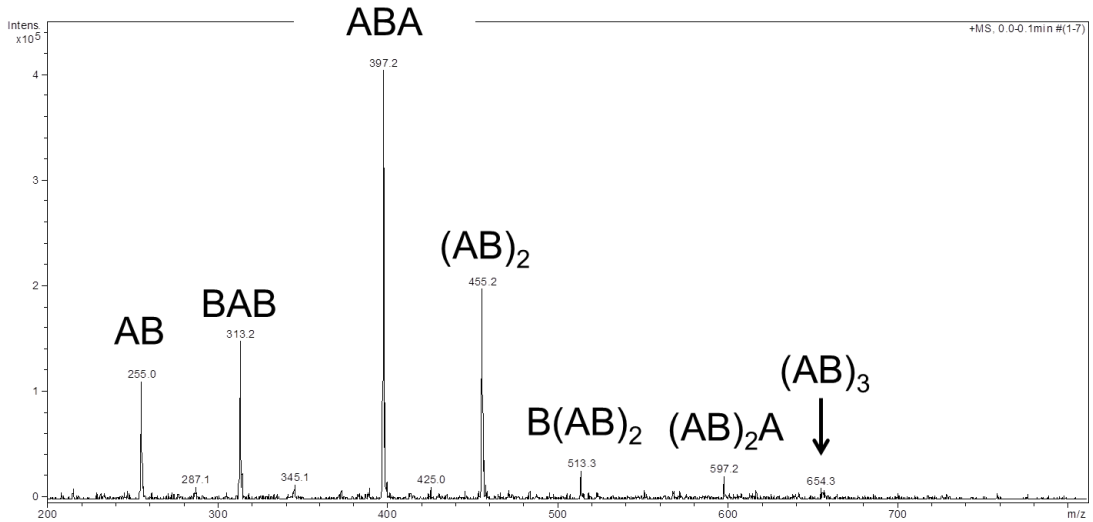


Figure S7. ESI-MS positive ion mass spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w rTHC-Cut1 at 24 h and 70 mbar. Entry 1, Table 2.

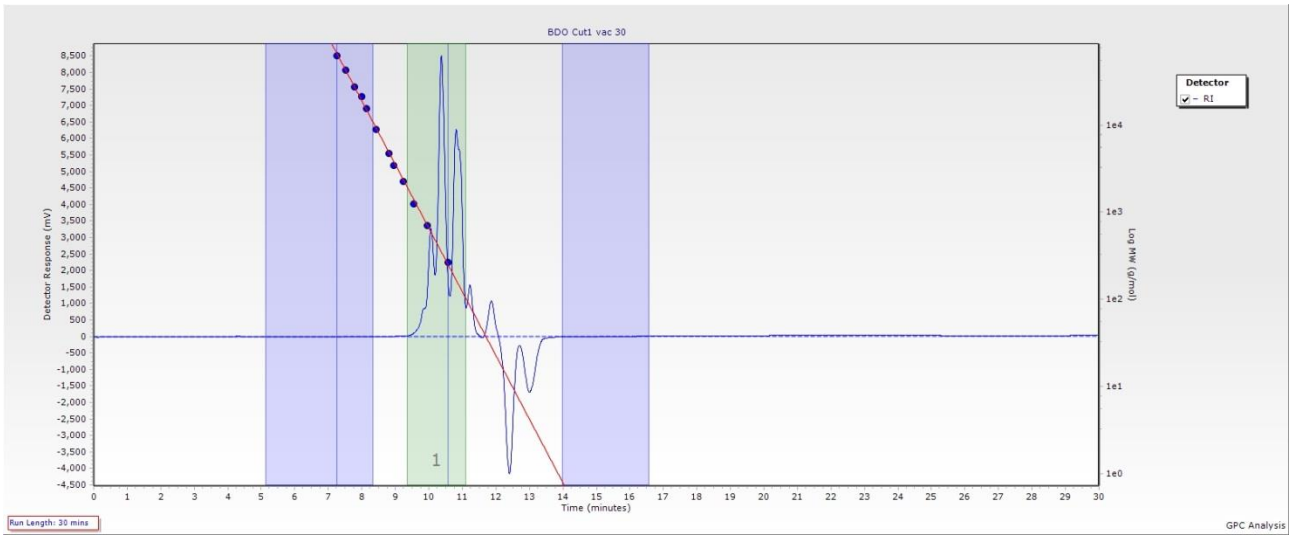


Figure S8. GPC chromatogram of the polycondensation products of DMA with BDO catalyzed by 30% w/w rTHC-Cut1 at 24 h and 70 mbar. Mw= 487, Mn= 343, PD= 1.268. Entry 3, Table 2.

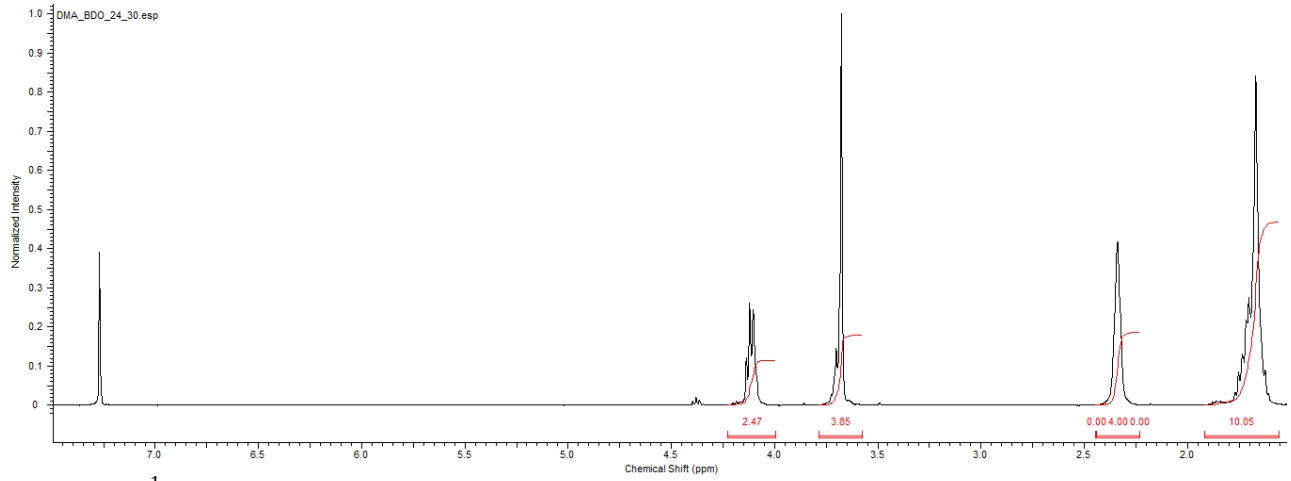


Figure S9. ^{1}H-NMR spectrum of the polycondensation products of DMA with BDO catalyzed by 30% w/w rTHC-Cut1 at 24 h and 70 mbar. Conversion: 62%. Entry 3, Table 2.

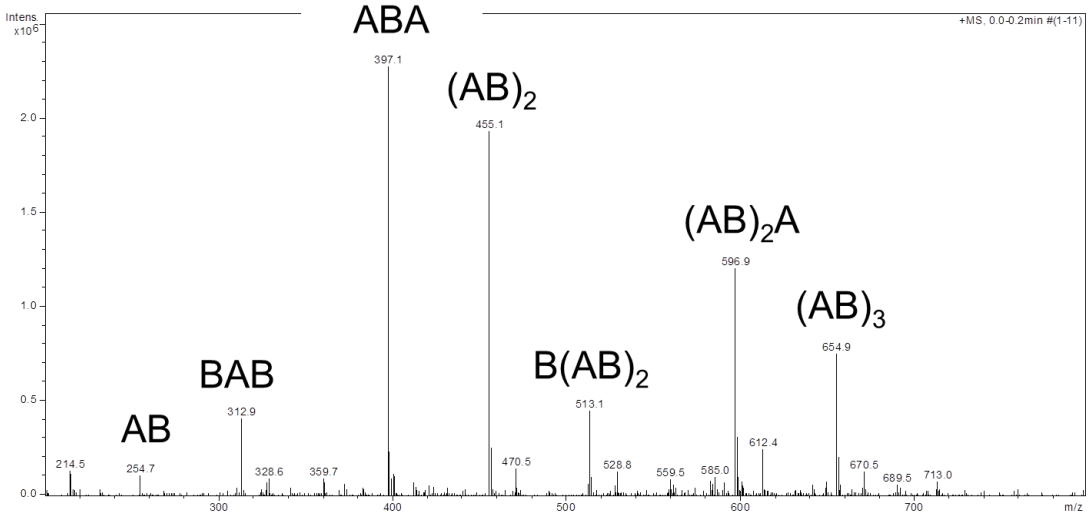


Figure S10. ESI-MS positive ion mass spectrum of the polycondensation products of DMA with BDO catalyzed by 30% w/w rTHC-Cut1 at 24 h and 70 mbar. Entry 3, Table 2.

Characterization of products obtained from the polycondensation of DMA and BDO catalyzed by Thc_cut1 immobilized on rice husk by adsorption and cross-linking (final water content: 0.2 % w/w).

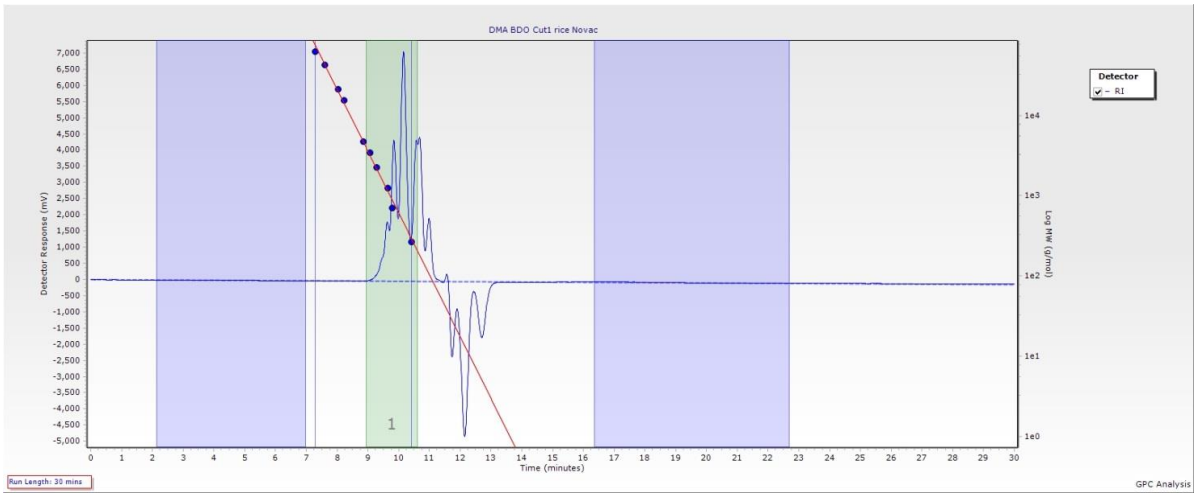


Figure S11. GPC chromatogram of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyo_rTHC-Cut1 at 24 h and 1000 mbar. Mw= 587, Mn= 384, PD= 1.268. Entry 4, Table 2.

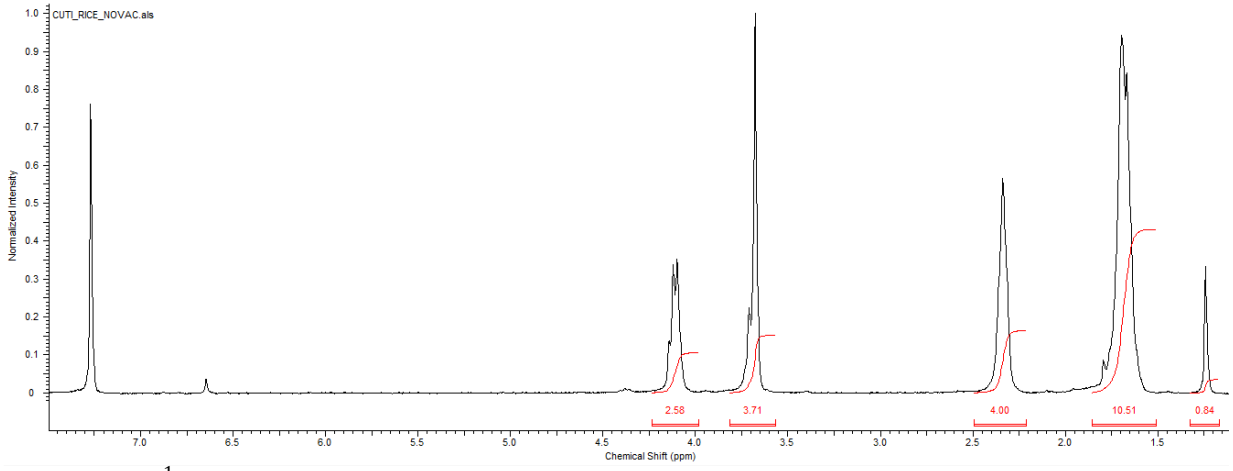


Figure S12. ^{1}H-NMR spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyo_rTHC-Cut1 at 24 h and 1000 mbar. Conversion: 62% . Entry 4, Table 2.

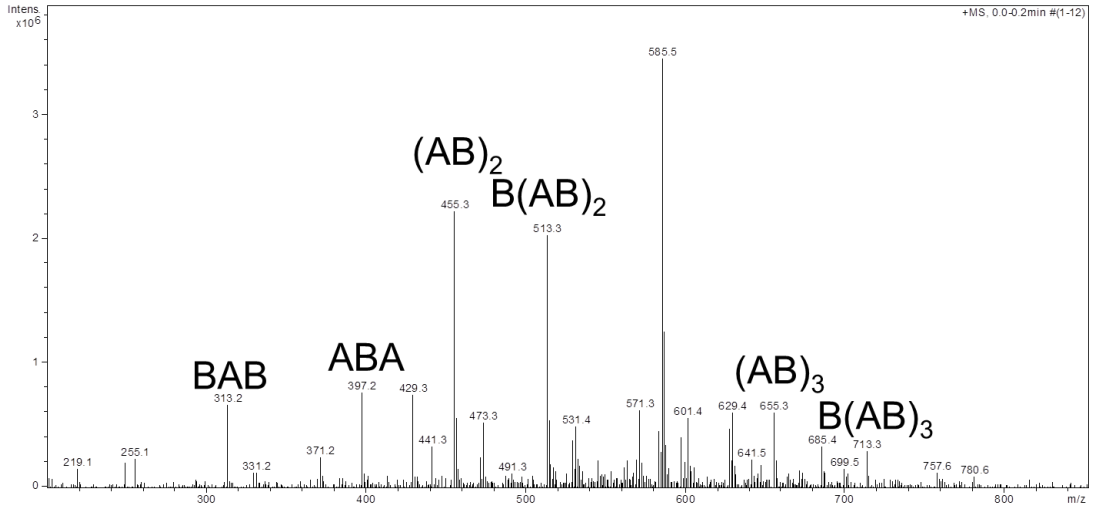


Figure S13. ESI-MS positive ion mass spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyo_rThc_cut1 at 24 h and 1000 mbar. Entry 4, Table 2.

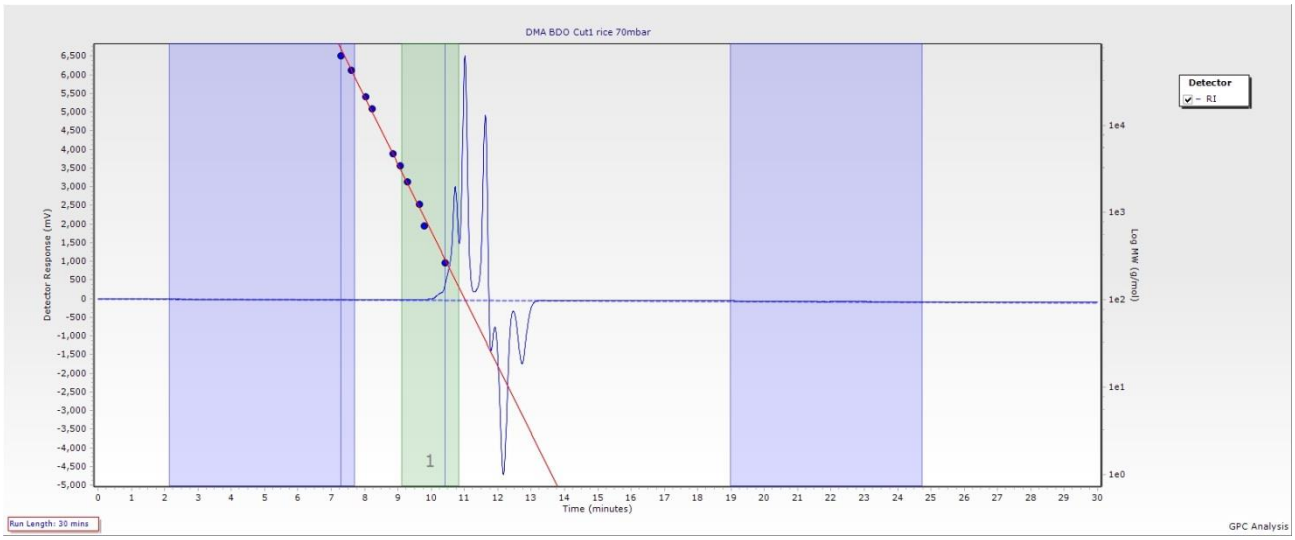


Figure S14. GPC chromatogram of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyo_rTHC-Cut1 at 24 h and 70 mbar. Mw= 224, Mn= 203, PD= 1.136. Entry 5, Table 2.

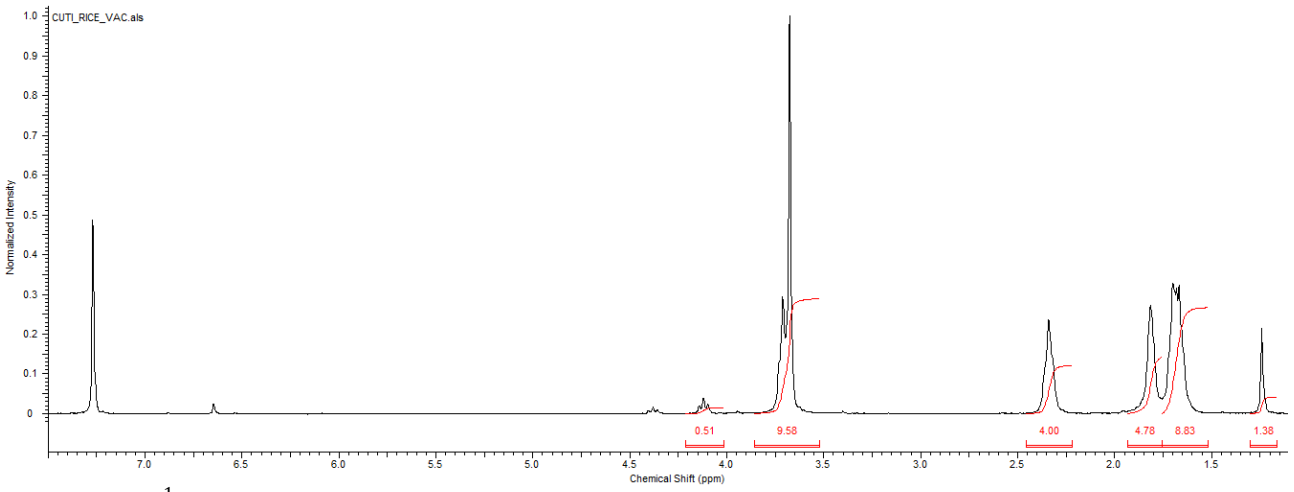


Figure S15. ^{1}H-NMR spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyo_rTHC-Cut1 at 24 h and 70 mbar. Conversion: 13% . Entry 5, Table 2.

Characterization of products obtained from the polycondensation of DMA and ODO catalyzed by Thc_cut1 immobilized on rice husk by adsorption and cross-linking (final water content: 4 % w/w).

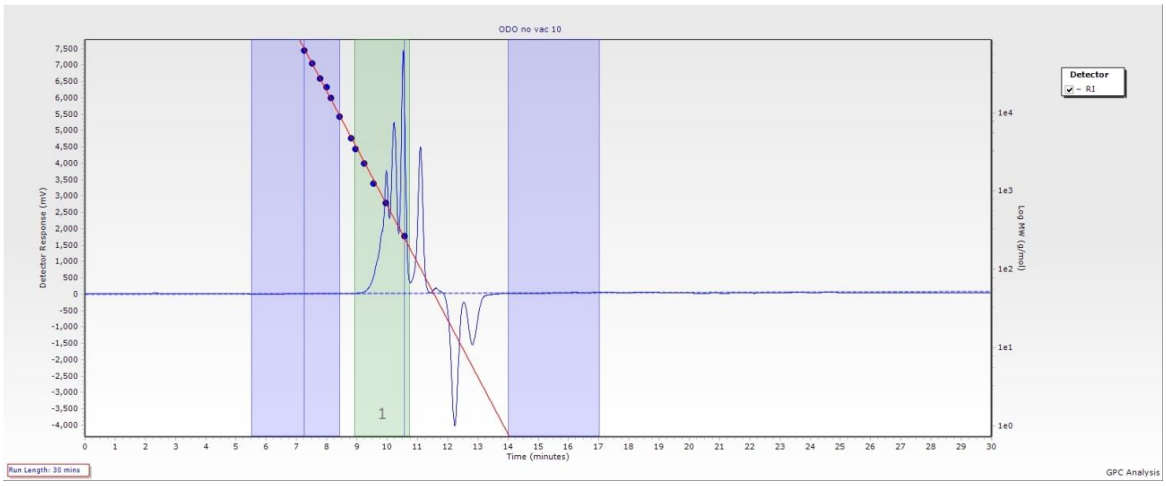


Figure S16. GPC chromatogram of the polycondensation products of DMA with ODO catalyzed by 10% w/w rTHC-Cut1 at 24 h and 1000 mbar. Mw= 508, Mn= 393, PD= 1.293. Entry 10, Table 3

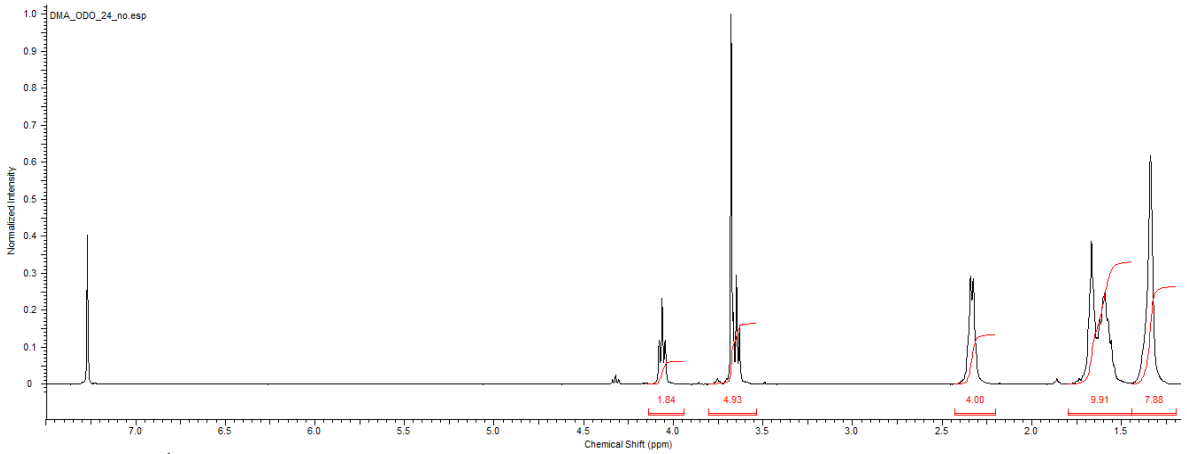


Figure S17. ^{1}H-NMR spectrum of the polycondensation products of DMA with ODO catalyzed by 10% w/w rTHC-Cut1 at 24 h and 1000 mbar. Conversion: 44% Entry 10, Table 3.

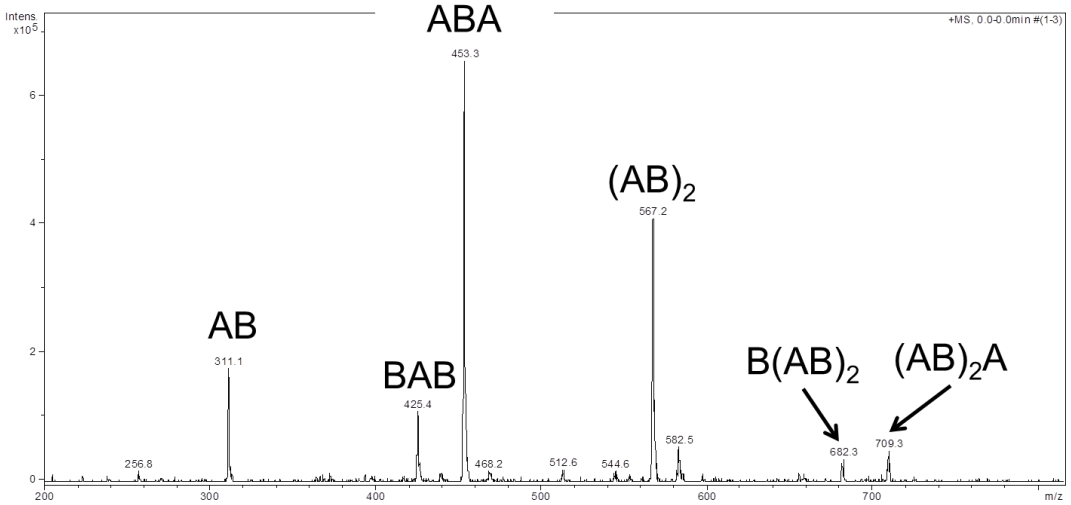


Figure S18. ESI-MS positive ion mass spectrum of the polycondensation products of DMA with ODO catalyzed by 10% w/w rTHC-Cut1 at 24 h and 1000 mbar. Entry 10, Table 3

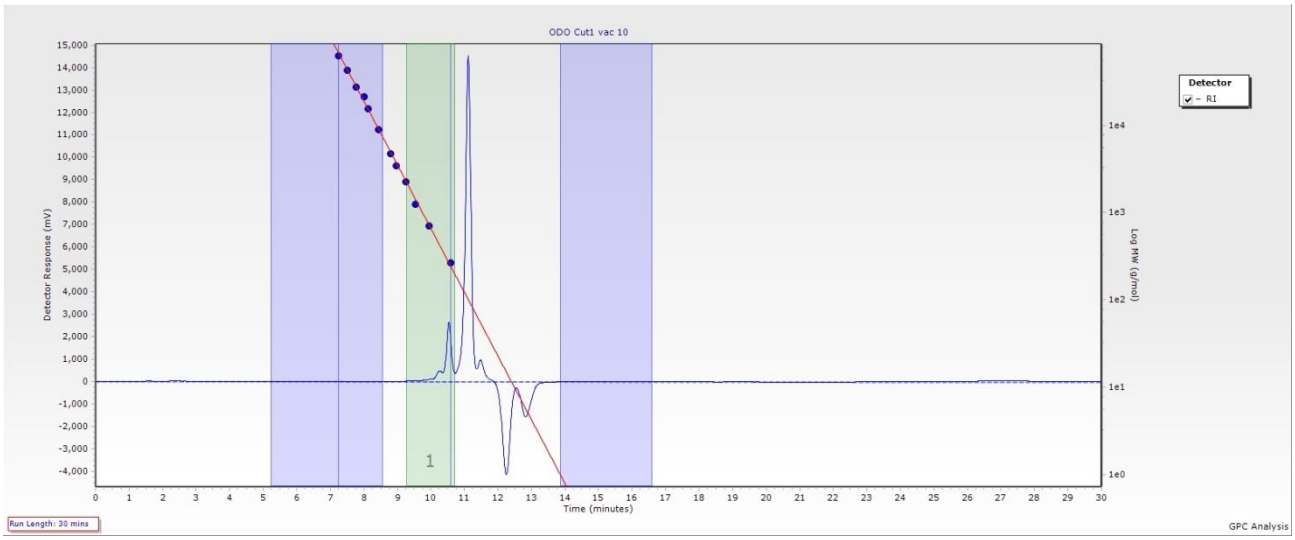


Figure S19. GPC chromatogram of the polycondensation products of DMA with ODO catalyzed by 10% w/w rTHC-Cut1 at 24 h and 70 mbar. Mw= 354, Mn= 298, PD= 1.188. Entry 11, Table 3.

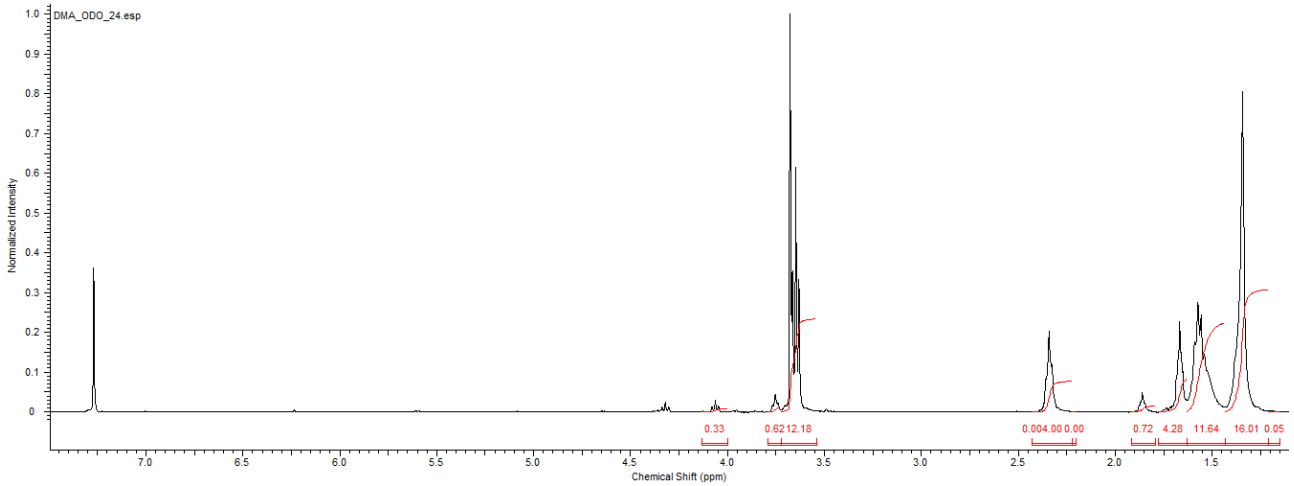


Figure S20. ^{1}H-NMR spectrum of the polycondensation products of DMA with ODO catalyzed by 10% w/w rTHC-Cut1 at 24 h and 70 mbar.Conversion: 8%. Entry 11, Table 3.

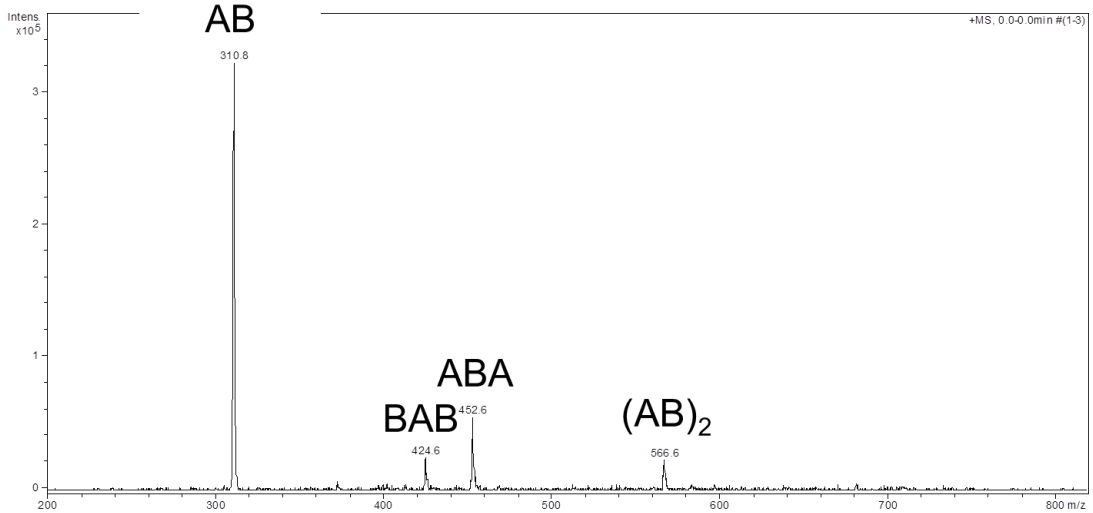


Figure S21. ESI-MS positive ion mass spectrum of the polycondensation products of DMA with ODO catalyzed by 10% w/w rTHC-Cut1 at 24 h and 70 mbar. Entry 11, Table 3.

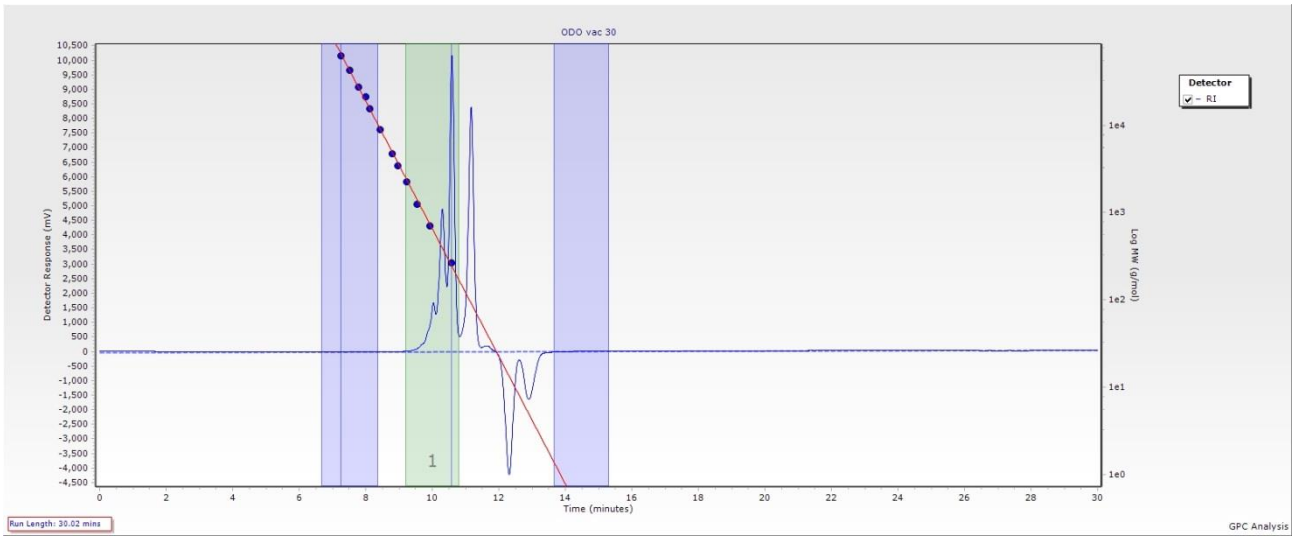


Figure S22. GPC chromatogram of the polycondensation products of DMA with ODO catalyzed by 30% w/w rTHC-Cut1 at 24 h and 70 mbar Mw= 447, Mn= 298, PD= 1.500. Entry 12, Table 3

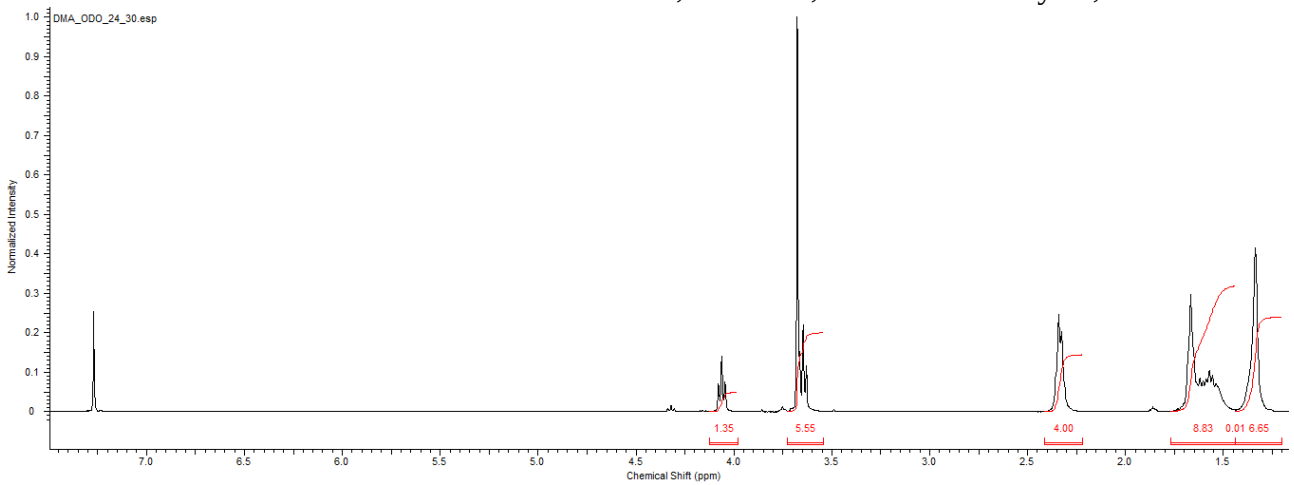


Figure S23. ^{1}H-NMR spectrum of the polycondensation products of DMA with ODO catalyzed by 30% w/w rTHC-Cut1 at 24 h and 70 mbar. .Conversion: 39%. Entry 12, Table 3

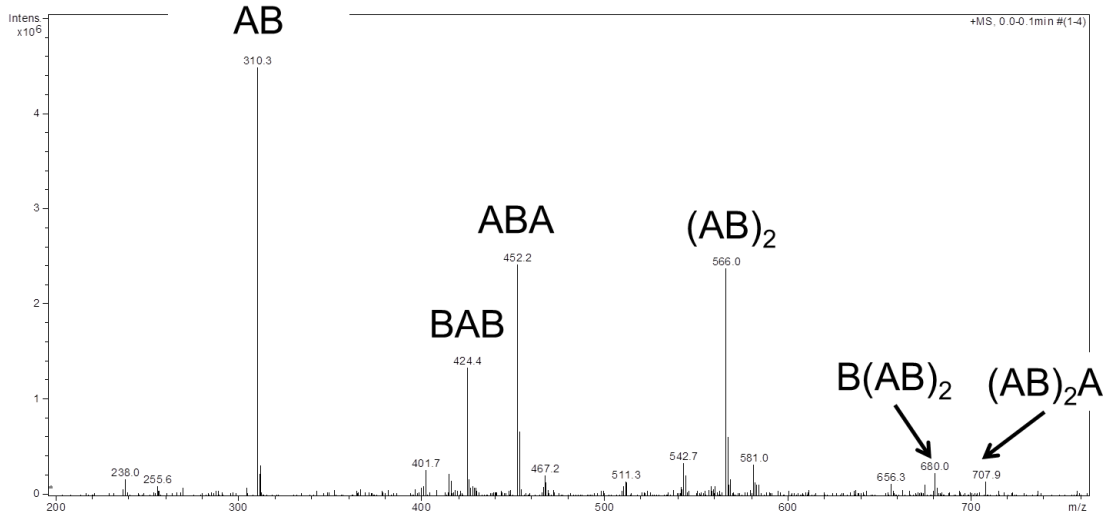


Figure S24. ESI-MS positive ion mass spectrum of the polycondensation products of DMA with ODO catalyzed by 30% w/w rTHC-Cut1 at 24 h and 70 mbar. Entry 12, Table 3.

Characterization of products obtained from the polycondensation of DMA and BDO catalyzed by Thc_cut1 immobilized on epoxy methacrylic resin (EC-EP)..

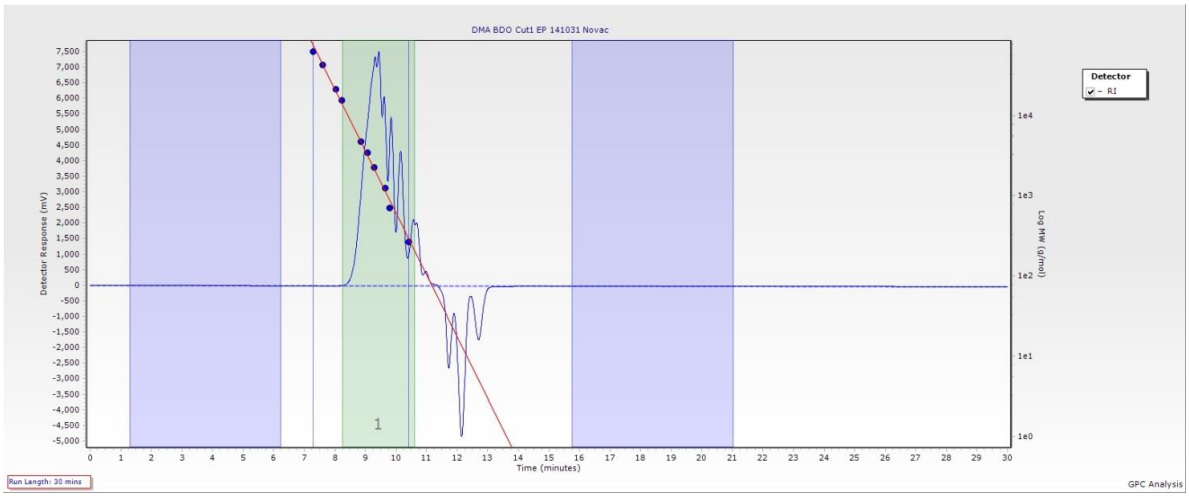


Figure S25. GPC chromatogram of the polycondensation products of DMA with BDO catalyzed by 10% w/w iTHC-Cut1 at 24 h and 1000 mbar 141031 Mw= 1923, Mn= 985, PD= 1.952. Entry 6, Table 2

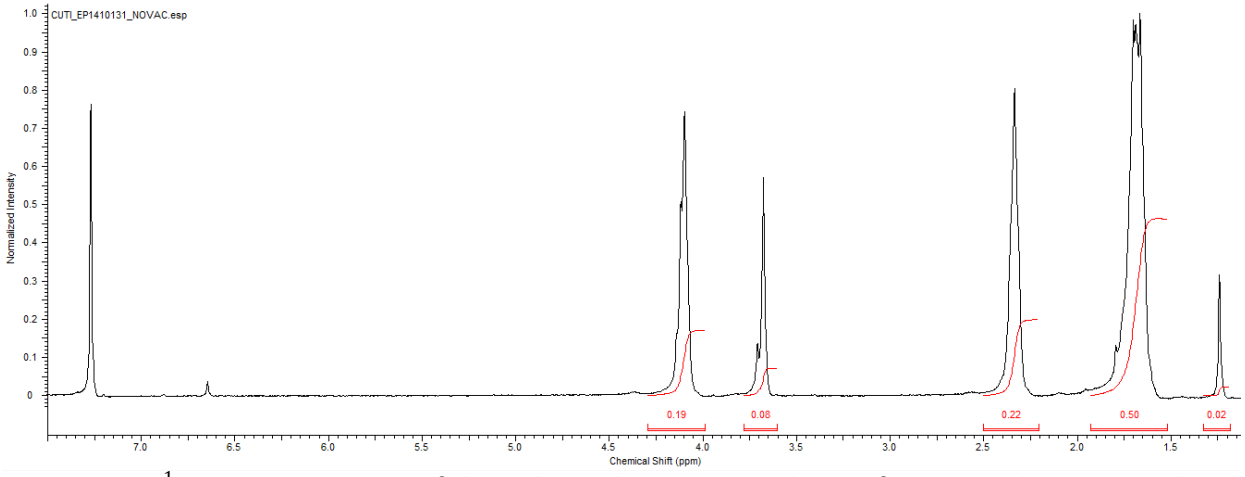


Figure S26. ^{1}H-NMR spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w iTHC-Cut1 at 24 h and 1000 mbar.Conversion: 86% Entry 6, Table 2.

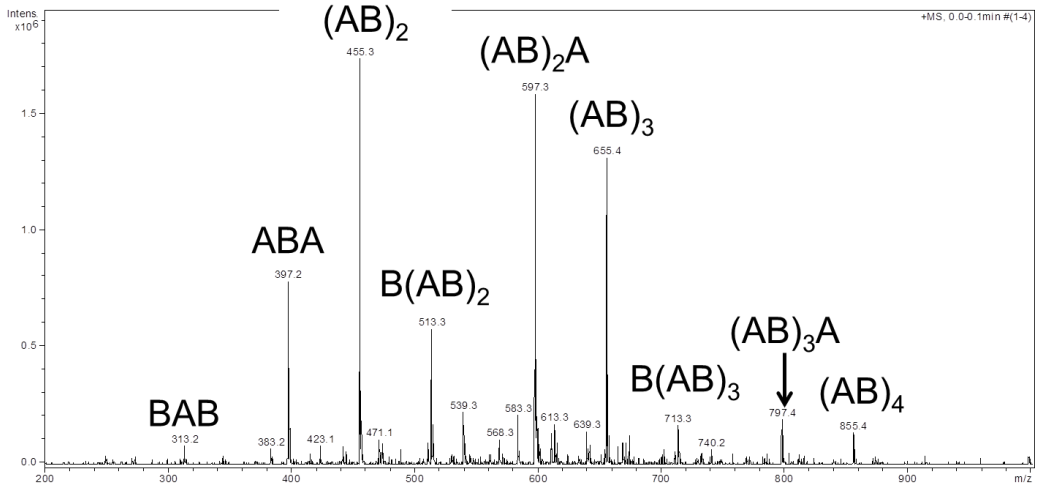


Figure S27. ESI-MS positive ion mass spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w iTHC-Cut1 at 24 h and 1000 mbar. Entry 6, Table 2

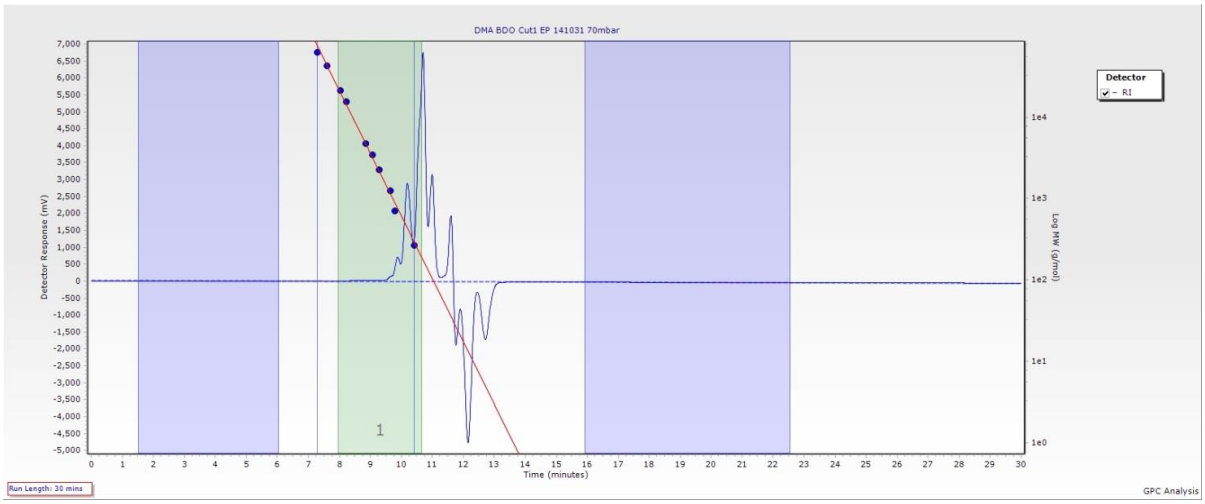


Figure S28. GPC chromatogram of the polycondensation products of DMA with BDO catalyzed by 10% w/w iTHC-Cut1 at 24 h and 70 mbar 141031 Mw= 480, Mn= 290, PD= 1.655. Entry 7, Table 2.

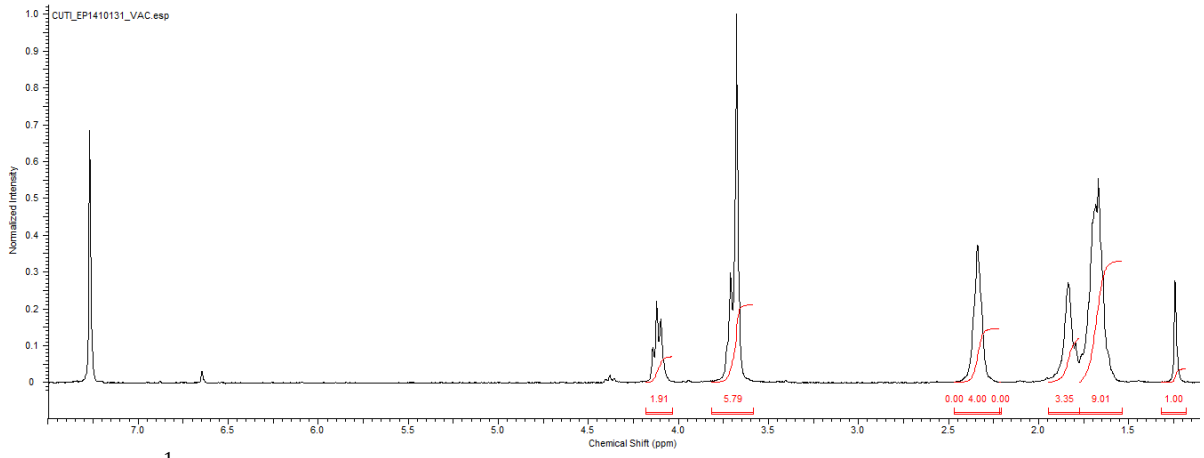


Figure S29. ^{1}H-NMR spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w iTHC-Cut1 at 24 h and 70 mbar.Conversion: 48% Entry 7, Table 2.

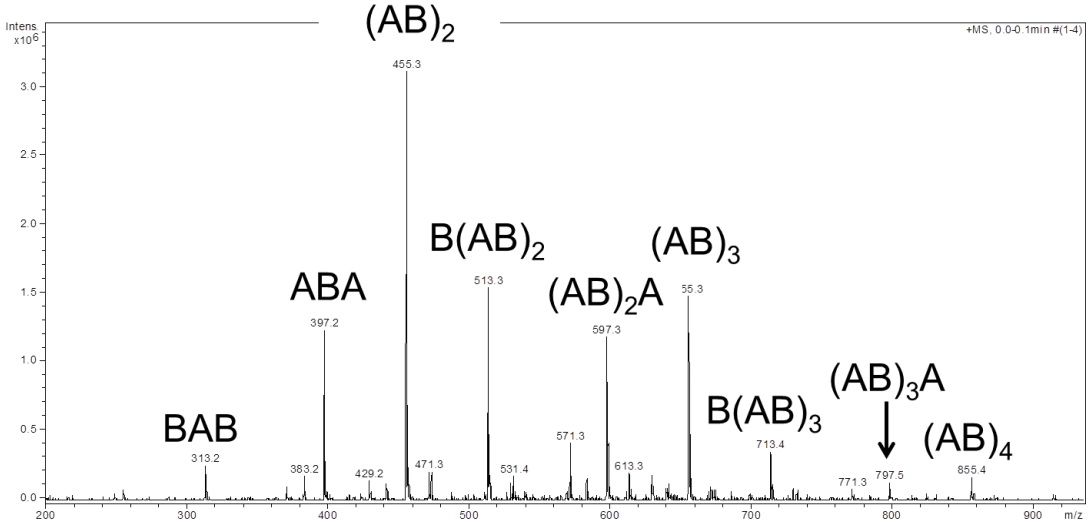


Figure S30. ESI-MS positive ion mass spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w iTHC-Cut1 at 24 h and 70 mbar. Entry 7, Table 2

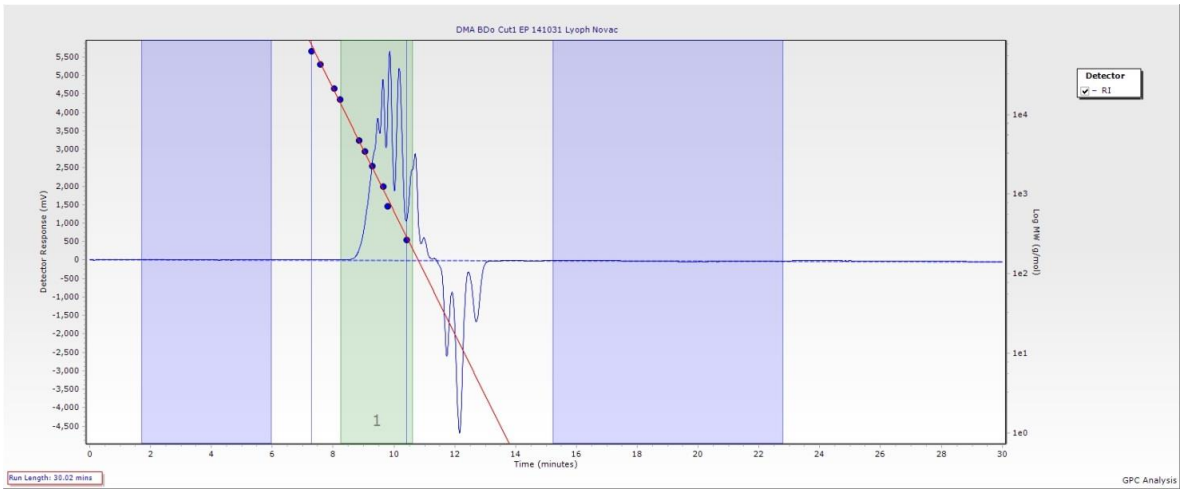


Figure S31. GPC chromatogram of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyo_iTHC-Cut1 at 24 h and 1000 mbar 141031 Mw= 1120, Mn= 656, PD= 1.707. Entry 8, Table 2

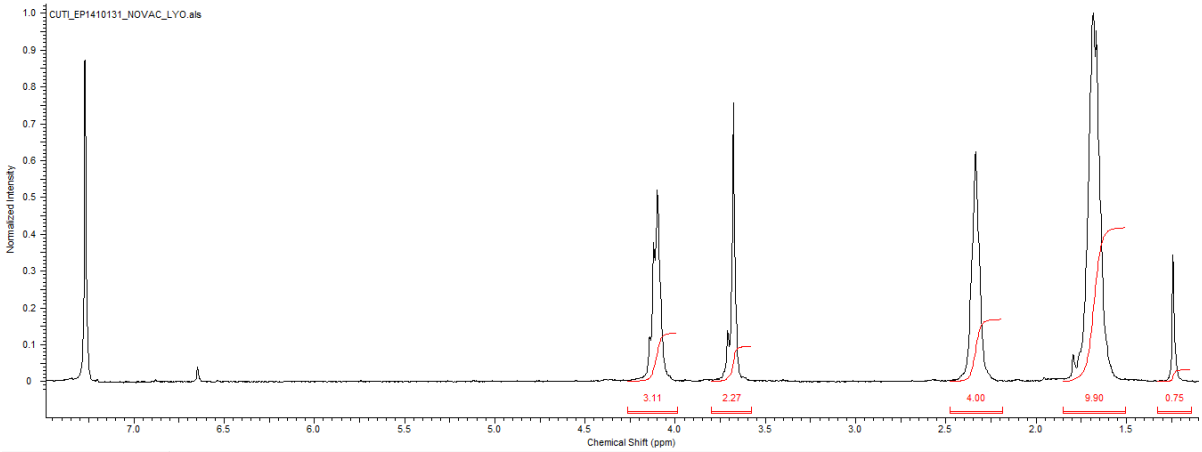


Figure S32. ^{1}H-NMR spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyo_iTHC-Cut1 at 24 h and 1000 mbar.Conversion: 78% . Entry 8, Table 2

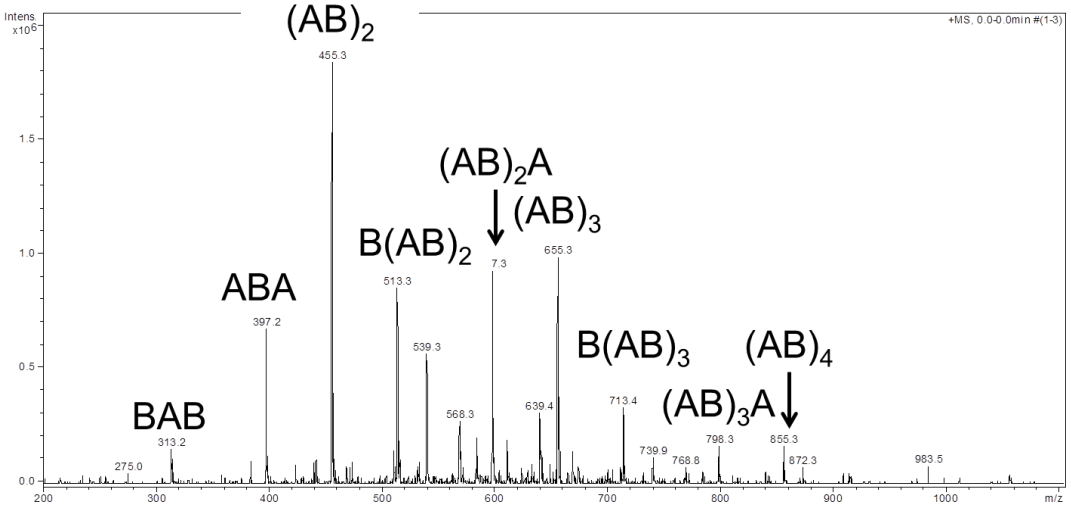


Figure S33. ESI-MS positive ion mass spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyo_iTHC-Cut1 at 24 h and 1000 mbar. Entry 8, Table 2

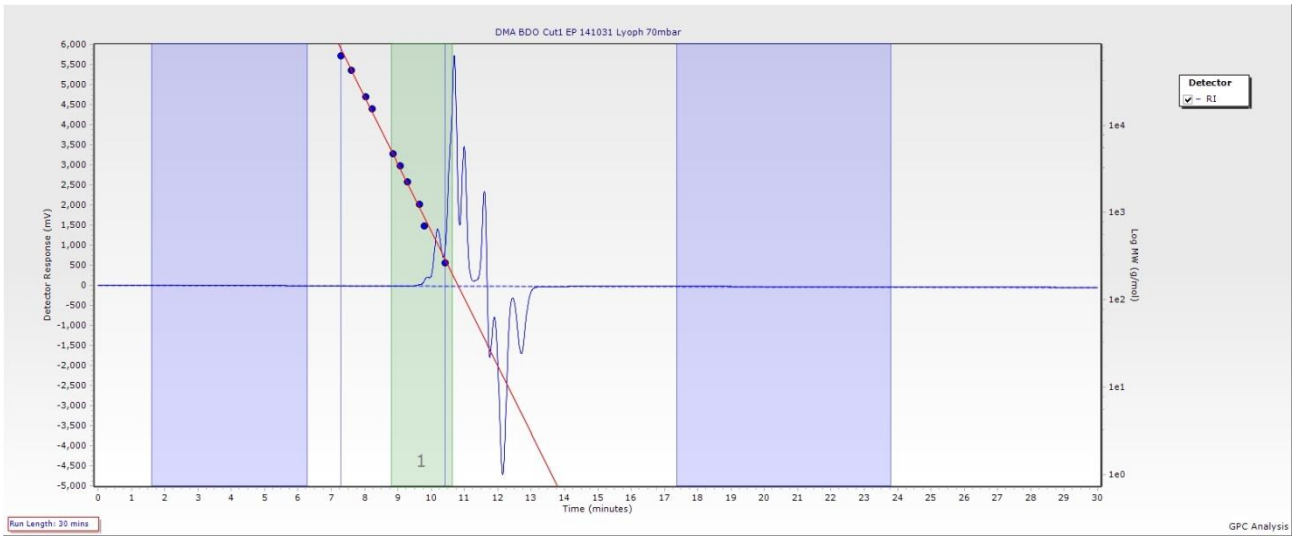


Figure S34. GPC chromatogram of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyo_iTHC-Cut1 at 24 h and 70 mbar 141031 Mw= 319, Mn= 275, PD= 1.160. Entry 9, Table 2

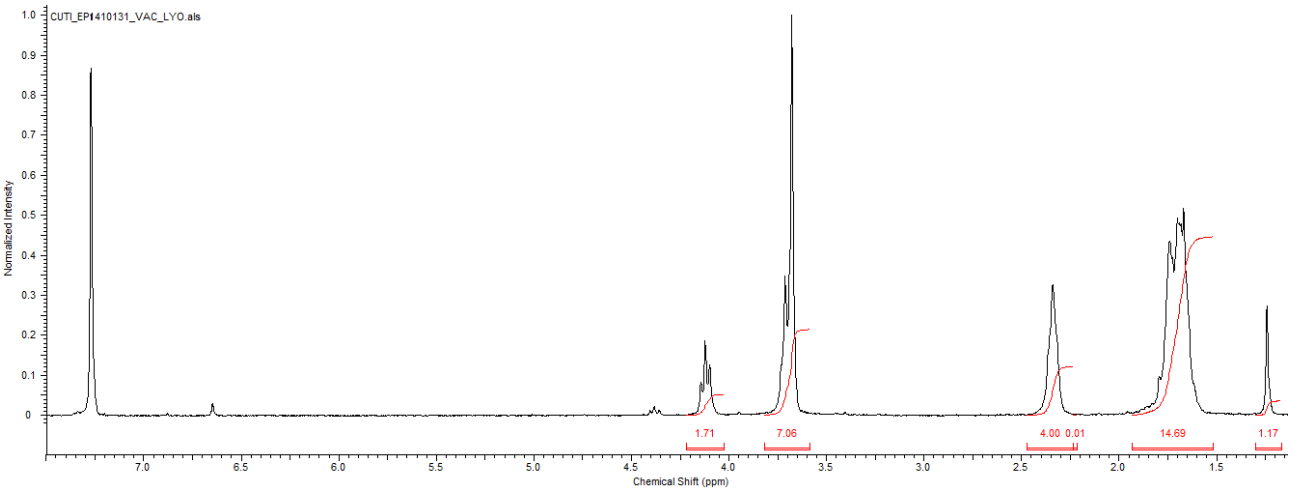


Figure S35. ^{1}H-NMR spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyo_iTHC-Cut1 at 24 h and 70 mbar.Conversion: 43%. Entry 9, Table 2

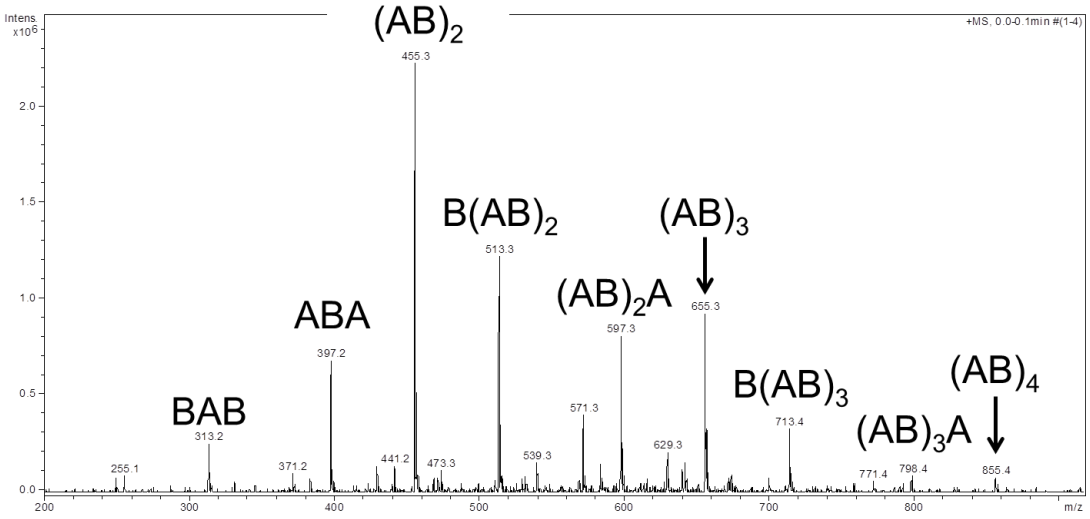


Figure S36. ESI-MS positive ion mass spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyo_iTHC-Cut1 at 24 h and 70 mbar. Entry 9, Table 2

Characterization of products obtained from the polycondensation of DMA and BDO catalyzed by Novozym 435, a commercially available formulation of CaLB physically adsorbed on organic resin

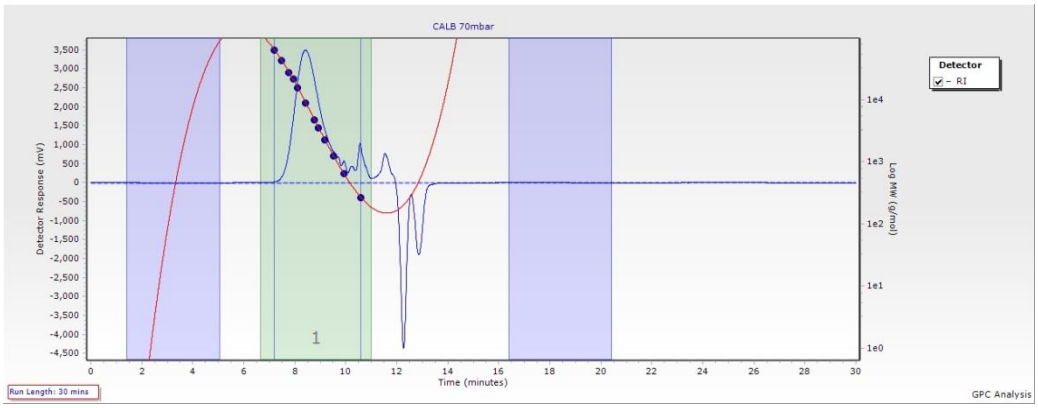


Figure S37. GPC chromatogram of the polycondensation products of DMA with BDO catalyzed by 10% w/w Novozym 435® at 24 h and 70 mbar Mw= 8357, Mn= 1759, PD= 4.751. Entry 16, Table 4.

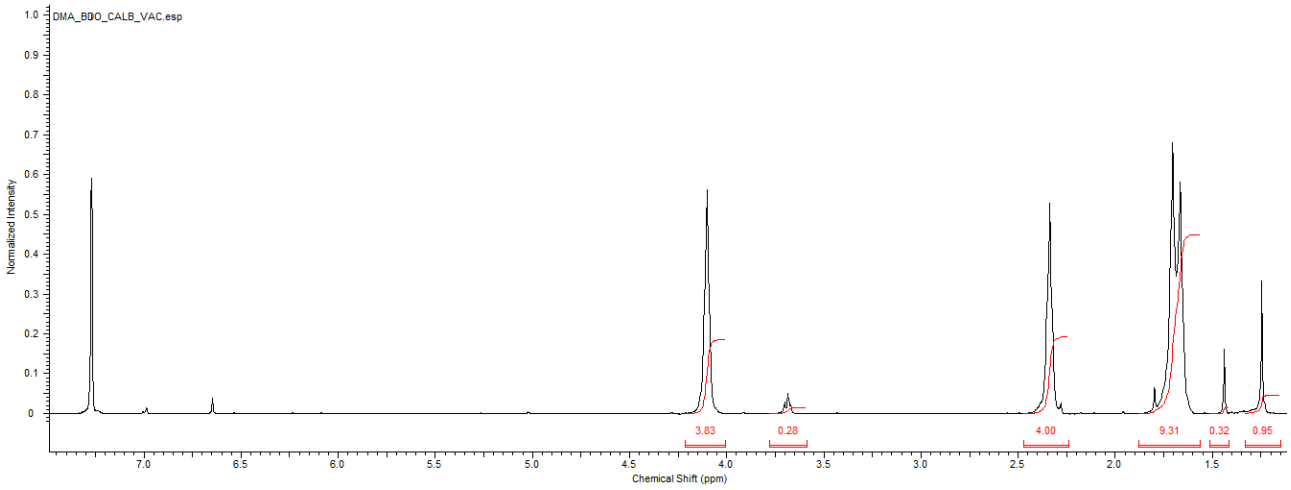


Figure S38. ^{1}H-NMR spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w Novozym 435® at 24 h and 70 mbar. Conversion: 96% Entry 16, Table 4

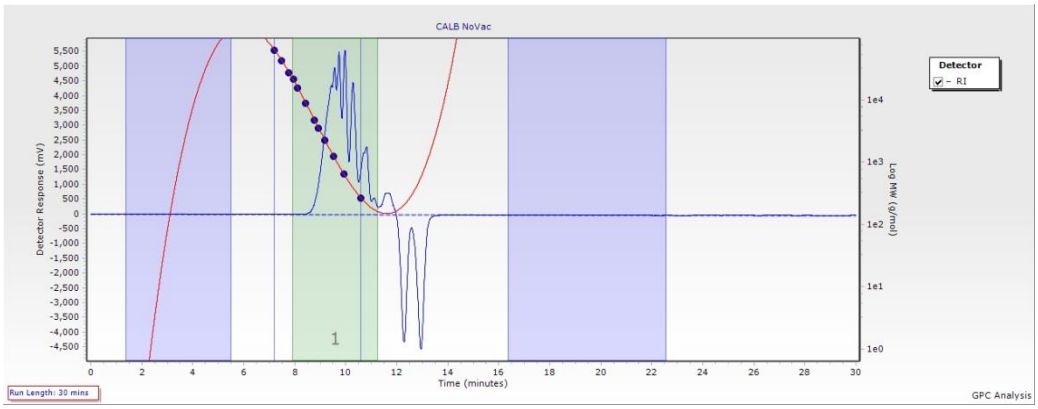


Figure S39. GPC chromatogram of the polycondensation products of DMA with BDO catalyzed by 10% w/w Novozym 435® at 24 h and 1000 mbar Mw= 1040, Mn= 561, PD= 1.854. Entry 15, Table 4

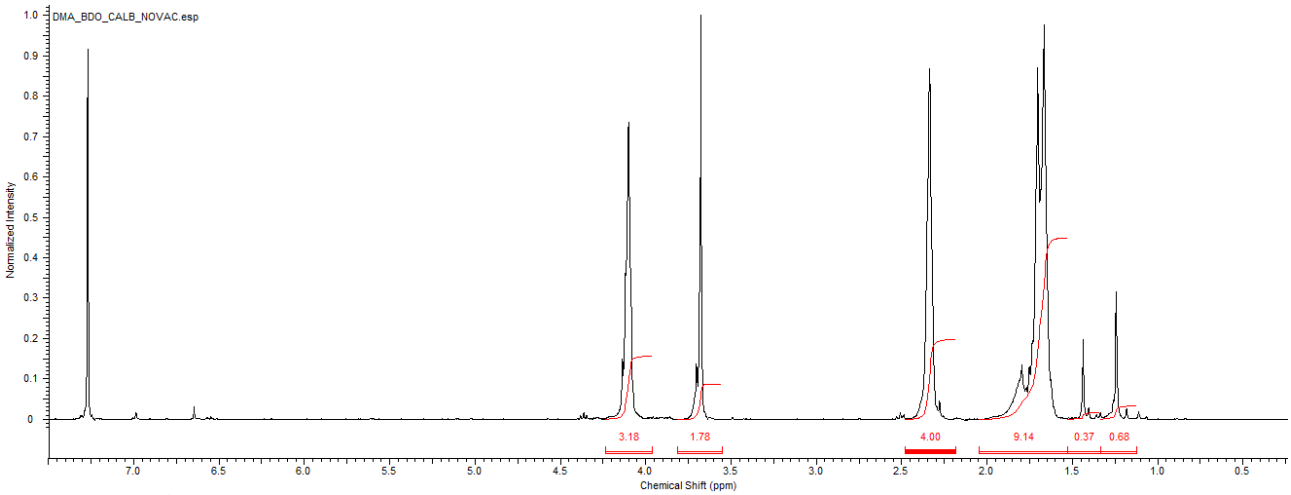


Figure S40. ^{1}H-NMR spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w Novozym 435® at 24 h and 1000 mbar.Conversion: 78% Entry 15, Table 4

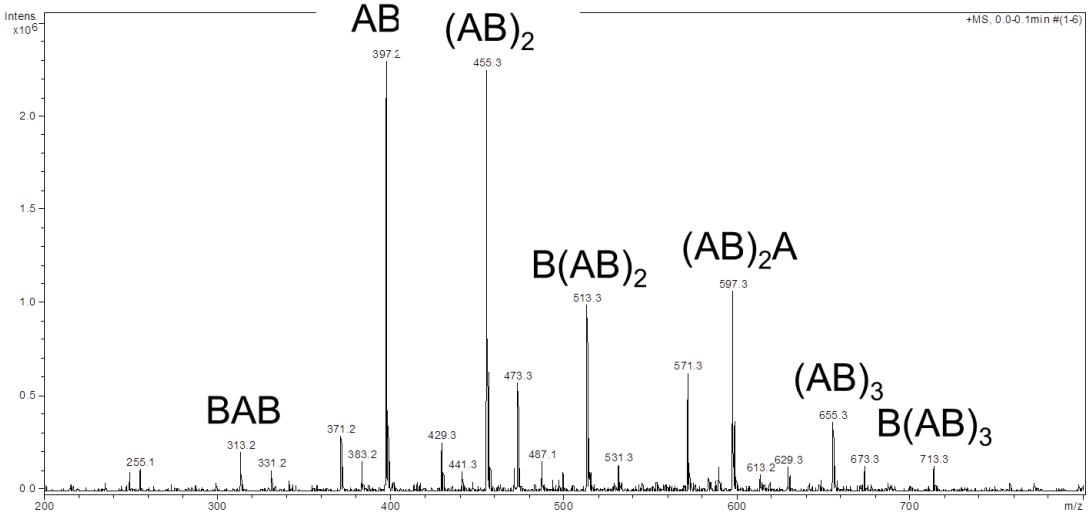


Figure S41. ESI-MS positive ion mass spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w Novozym 435® at 24 h and 1000 mbar. Entry 15, Table 4

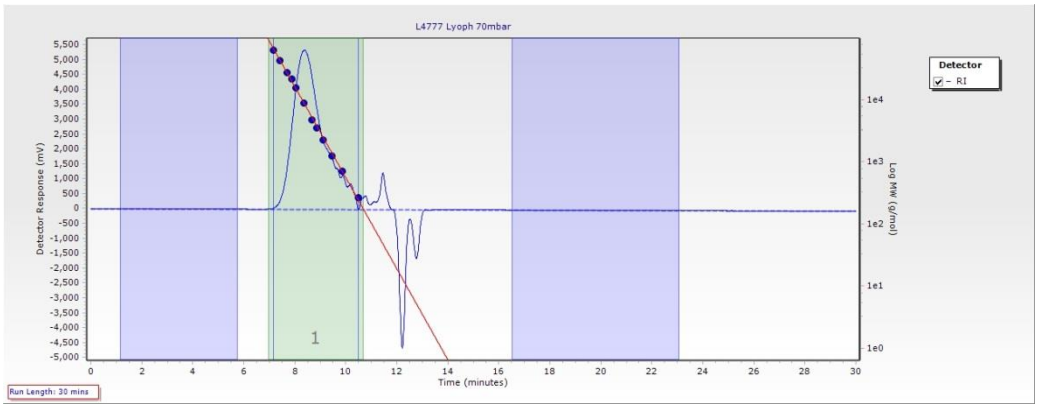


Figure S42. GPC chromatogram of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyophilized Novozym 435® at 24 h and 70 mbar Mw= 8250, Mn= 2438, PD= 3.384. Entry 18, Table 4

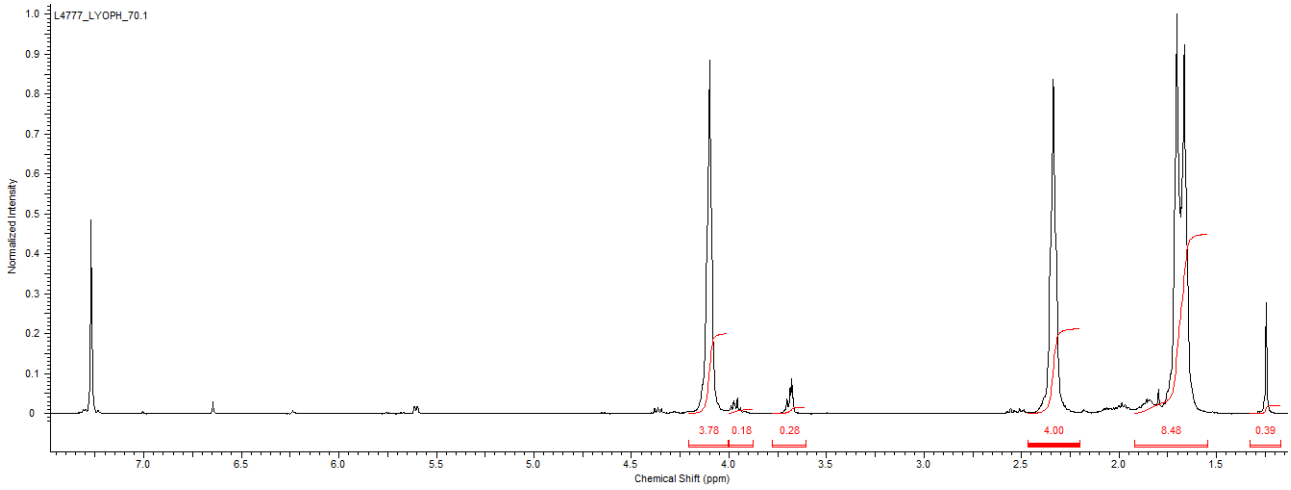


Figure S43. ^{1}H-NMR spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyophilized Novozym 435® at 24 h and 70 mbar.Conversion: 94% Entry 18, Table 4

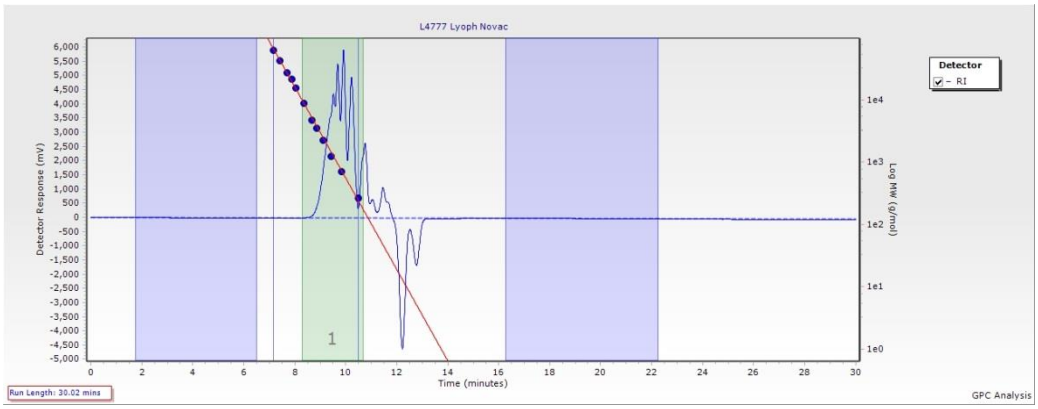


Figure S44. GPC chromatogram of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyophilized Novozym 435® at 24 h and 1000 mbar Mw= 999, Mn= 608, PD= 1.643. Entry 17, Table 4.

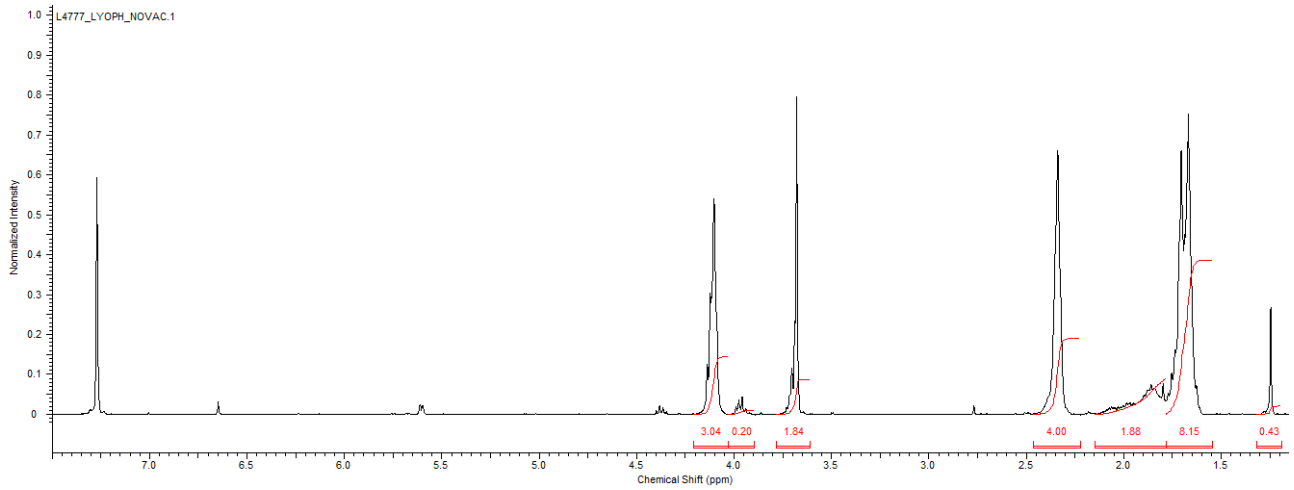


Figure S45. ^{1}H-NMR spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w lyophilized Novozym 435® at 24 h and 70 mbar.Conversion: 76% . Entry 17, Table 4

Molecular Dynamics Simulations of CaLB and Thc_cut1

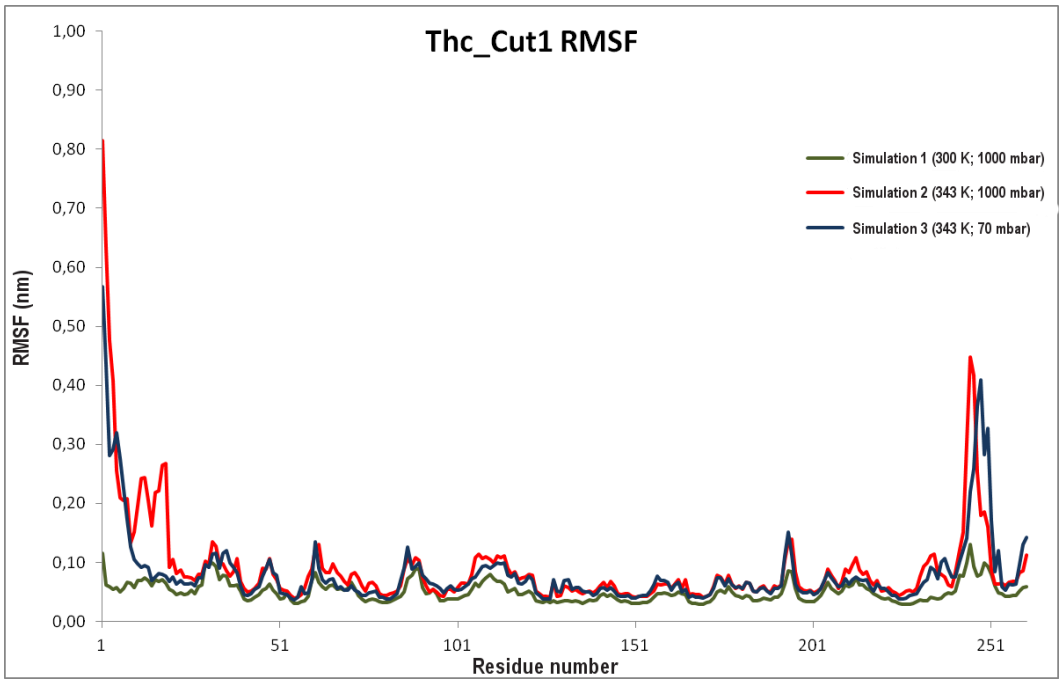


Figure S46. Thc_Cut1 RMSF analysis and comparison. The 3 different simulation conditions are highlighted in different colours which are indicated, together with the simulation conditions in the chart legend.

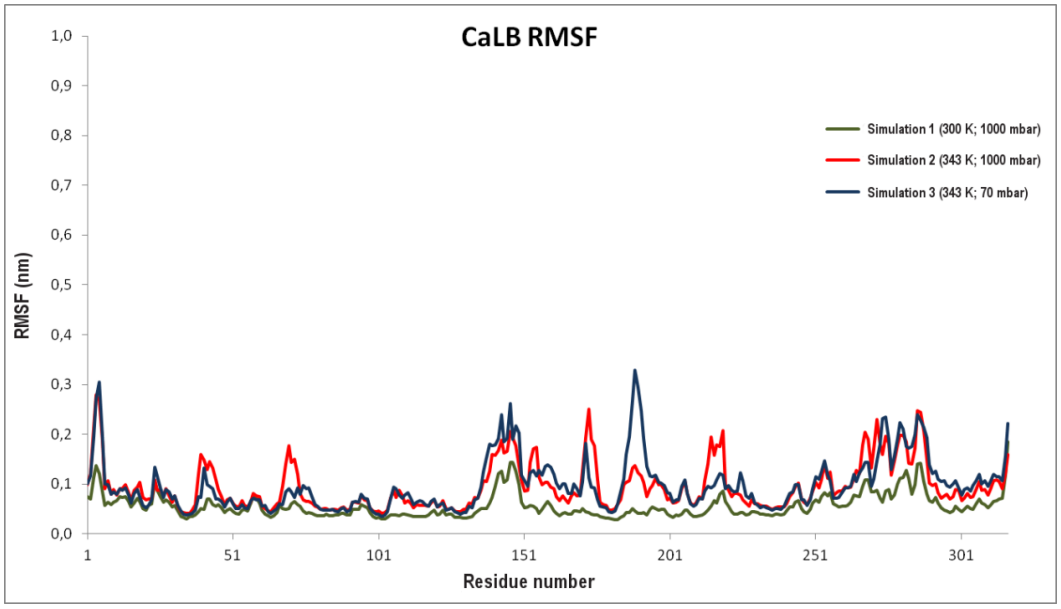
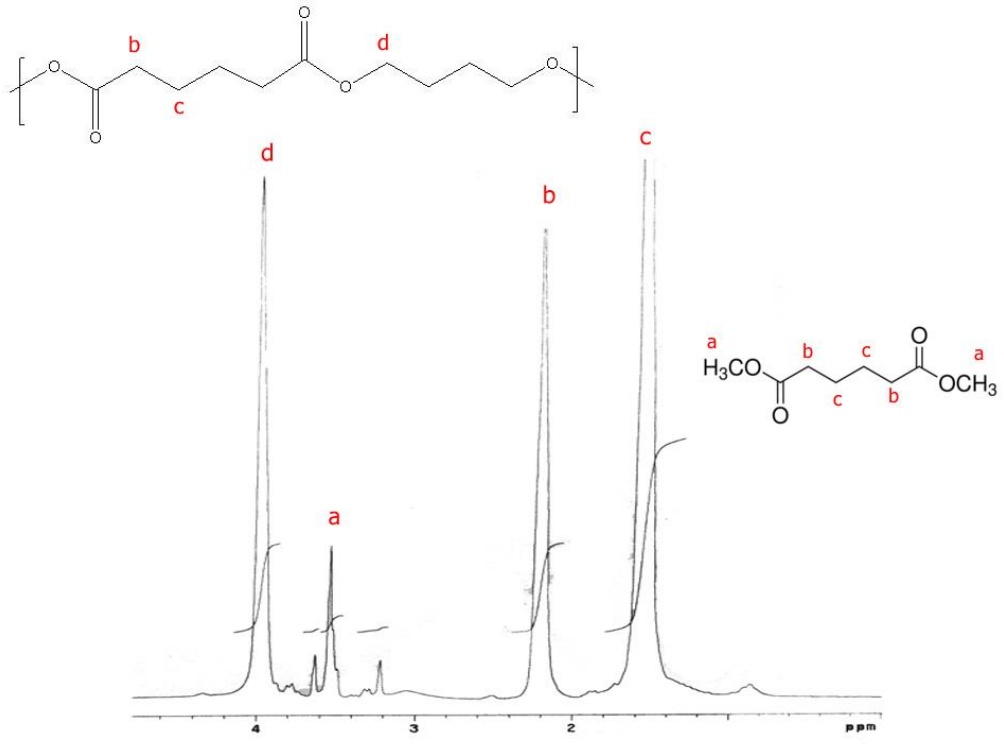
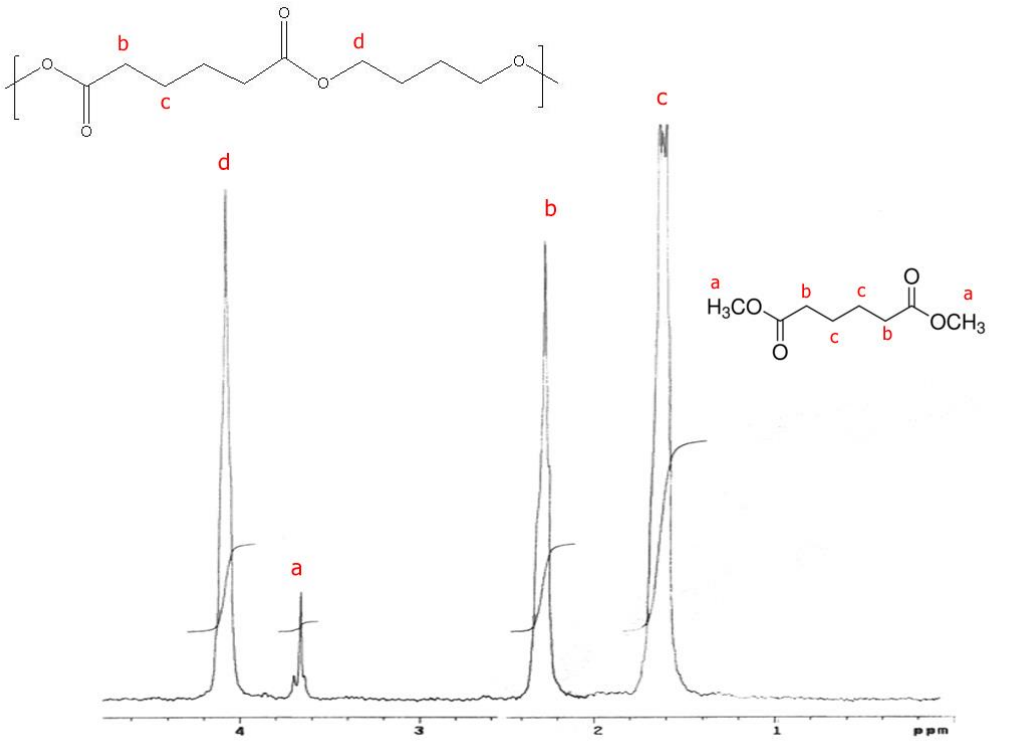


Figure S47. CaLB RMSF analysis and comparison. The 3 different simulation conditions are highlighted in different colours which are indicated, together with the simulation conditions in the chart legend.

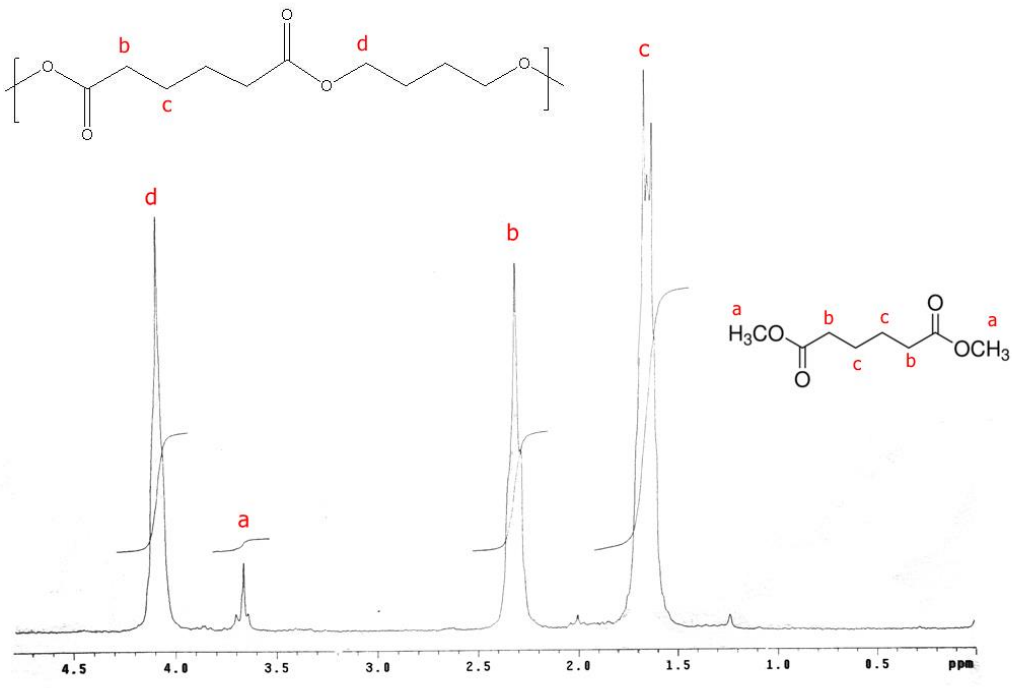
Characterization of products obtained from the polycondensation of DMA and BDO caylyzed by CaLB immobilized on epoxy methacrylic resin (EC-EP), (50 °C; 70 mbar, no water added to the system).



a



b



c

Figure S48 . [1]H-NMR spectrum of products of polycondensation between BDO and DMA after a) 24 h, b) 48 h, c) 72 h.

^{1}H-NMR ($CDCl_3$), δ 3.95 (2H, d, cis, -C<u>H</u>$_2$-OCO-), 3.86 (2H, d, trans, -C<u>H</u>$_2$-OCO-), 3.64 (3H, s, CH_2-CO-OC<u>H</u>$_3$), 3.50 (2H, d, cis, CH-C<u>H</u>$_2$-OH), 3.42 (2H, d, trans, CH-C<u>H</u>$_2$-OH), 2.30 (2H, t, -CH_2-C<u>H</u>$_2$-CO-), 1.84-1.4 (1H, m, CH_2-C<u>H</u>-CH_2), 1.58 (4H, quin, -CH_2-C<u>H</u>$_2$-CH_2-), 1.0 (2H, m, CH_2-C<u>H</u>$_2$-CH_2).

The presence of ester functional group of the formed oligomers P(BA) and P(CA) was confirmed by ^{1}H-NMR spectroscopy (Figure 12, 13). The spectra were collected in $CDCl_3$. The methoxy groups of unreacted diester are presented at ~ 3.6 ppm. The methylene group adjacent to ester moiety he ester peaks are presented in the range of 3.8-4.0 ppm, with the rest of the spectrum resulting from hydrophobic protons of linear or cyclic diol.

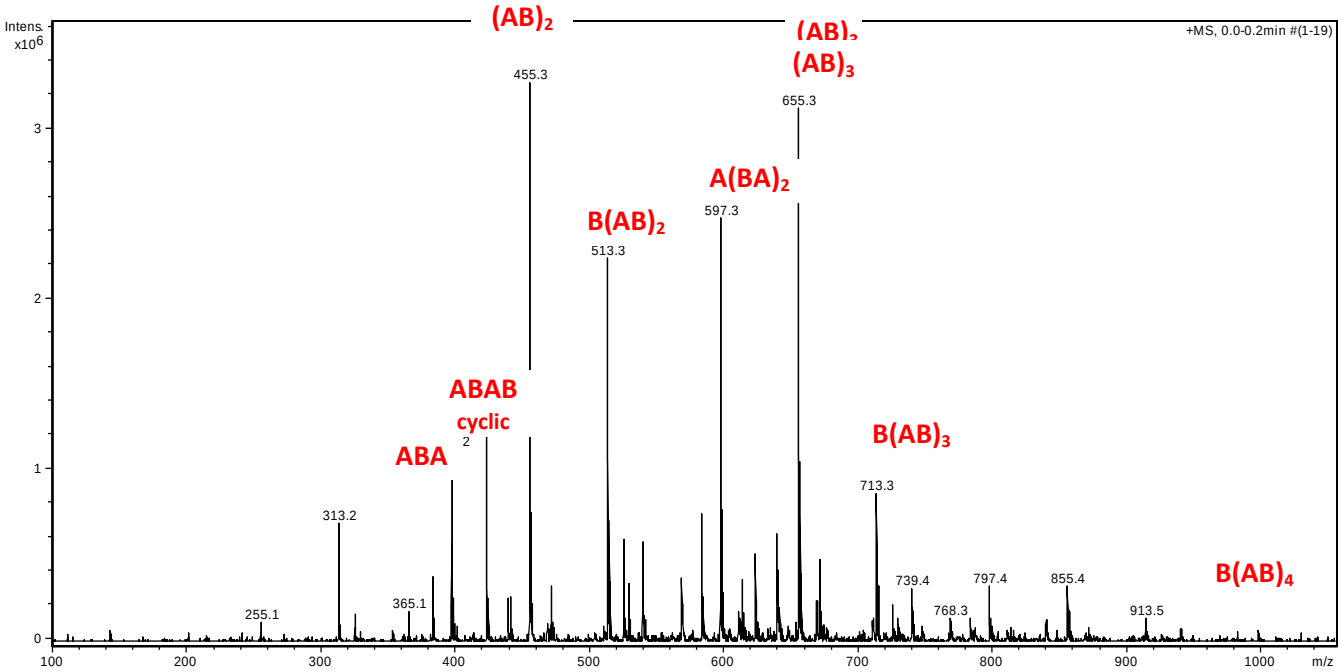


Figure S49. ESI-MS positive ion mass spectrum of polycondensation products of DMA (A) and BDO (B) catalyzed by CALB immobilized on epoxy methacrylic resin at 50°C, 70mbar, no water added to the system and 24 h of reaction.

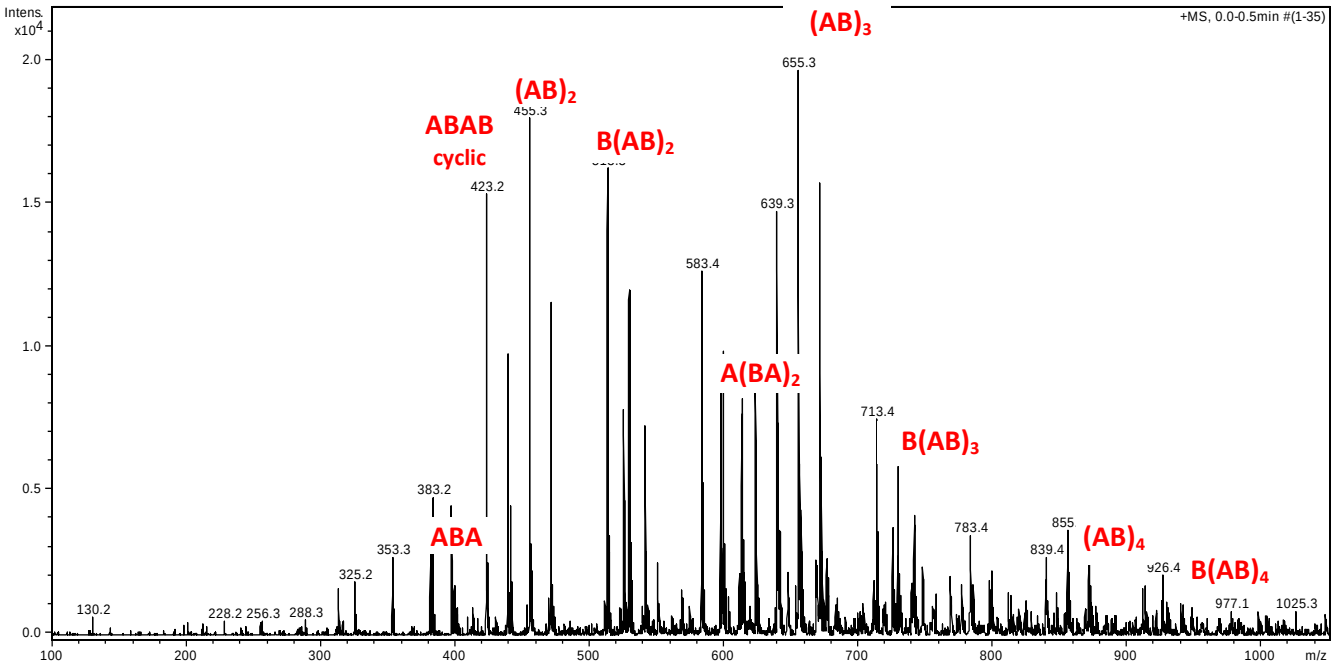


Figure S50. ESI-MS positive ion mass spectrum of polycondensation products of DMA (A) and BDO (B) catalyzed by CALB immobilized on epoxy methacrylic resin at 50°C, 70mbar, no water added to the system and 48 h of reaction.

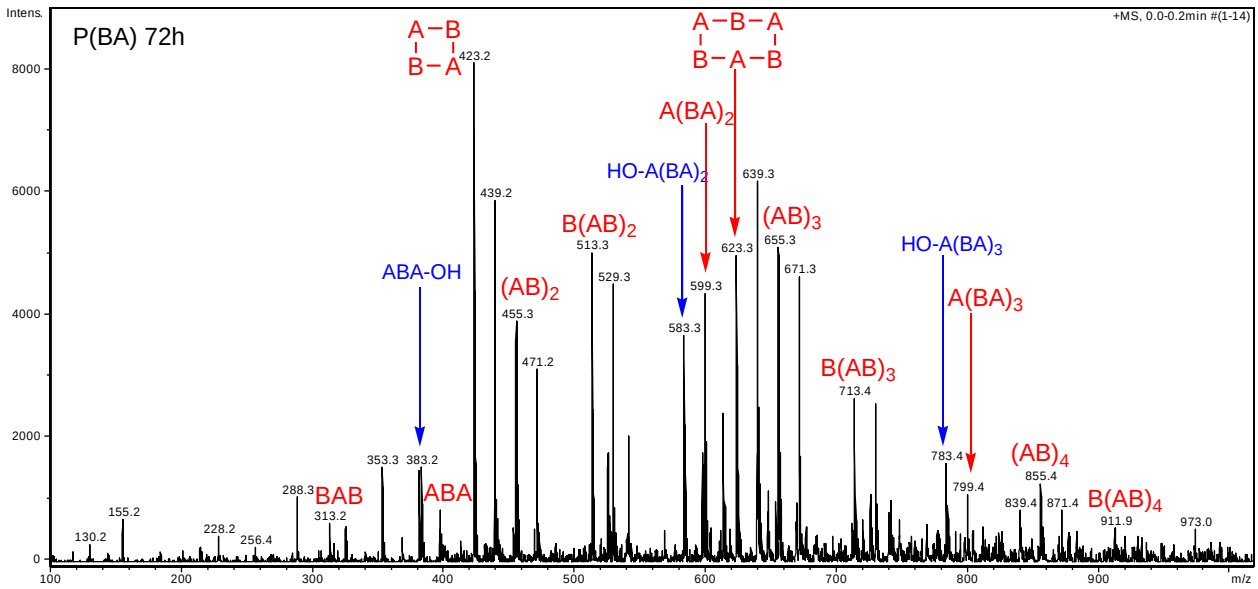
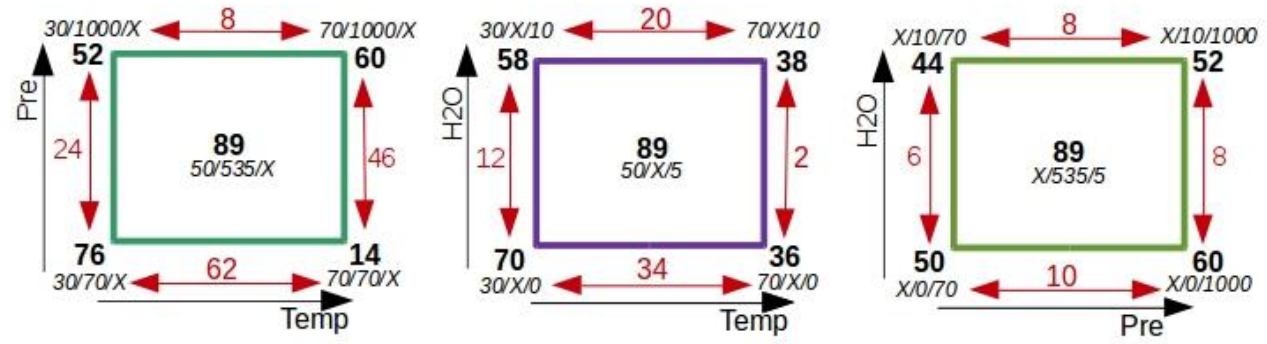
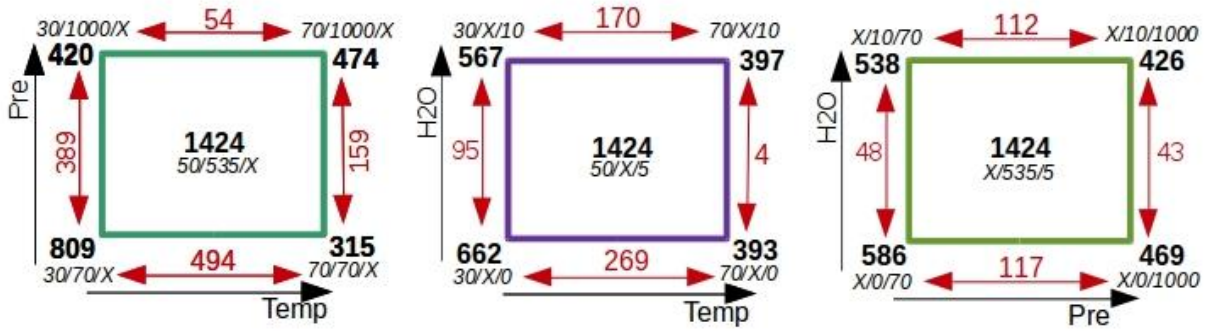


Figure S51. ESI-MS positive ion mass spectrum of polycondensation products of DMA (A) and BDO (B) catalyzed by CALB immobilized on epoxy methacrylic resin at 50°C, 70mbar, no water added to the system and 72 h of reaction.



Conversion



Mn

Figure S52: Details of interactions between temperature and pressure and effects on conversion and Mn. Each edge of the cube corresponds to a possible combination between the two factors (T and p) taken into consideration. The values reported at cube's edge are the average of the response measured under those conditions. The difference between the "edge's values" are highlighted in red. Such difference can be interpreted as an indication of how the change of level of one variable (e.g. T or p) affect the influence of the other variable on the response (i.e. conversion or Mn).

Table 2. Water content of all immobilized enzymes employed in the study

Enzymatic preparation	Drying method	Water content (w/w %)*
Novozym 435®	Air-drying	1
Novozym 435®	Lyophilization	0.1
iCaLB	Air-drying	3
iCaLB	Lyophilization	0.2
iThc_cut1	Air-drying	3
iThc_cut1	Lyophilization	0.2
rThc_cut1	Air-drying	4
rThc_cut1	Lyophilization	0.2

* All determinations were conducted in duplicates.

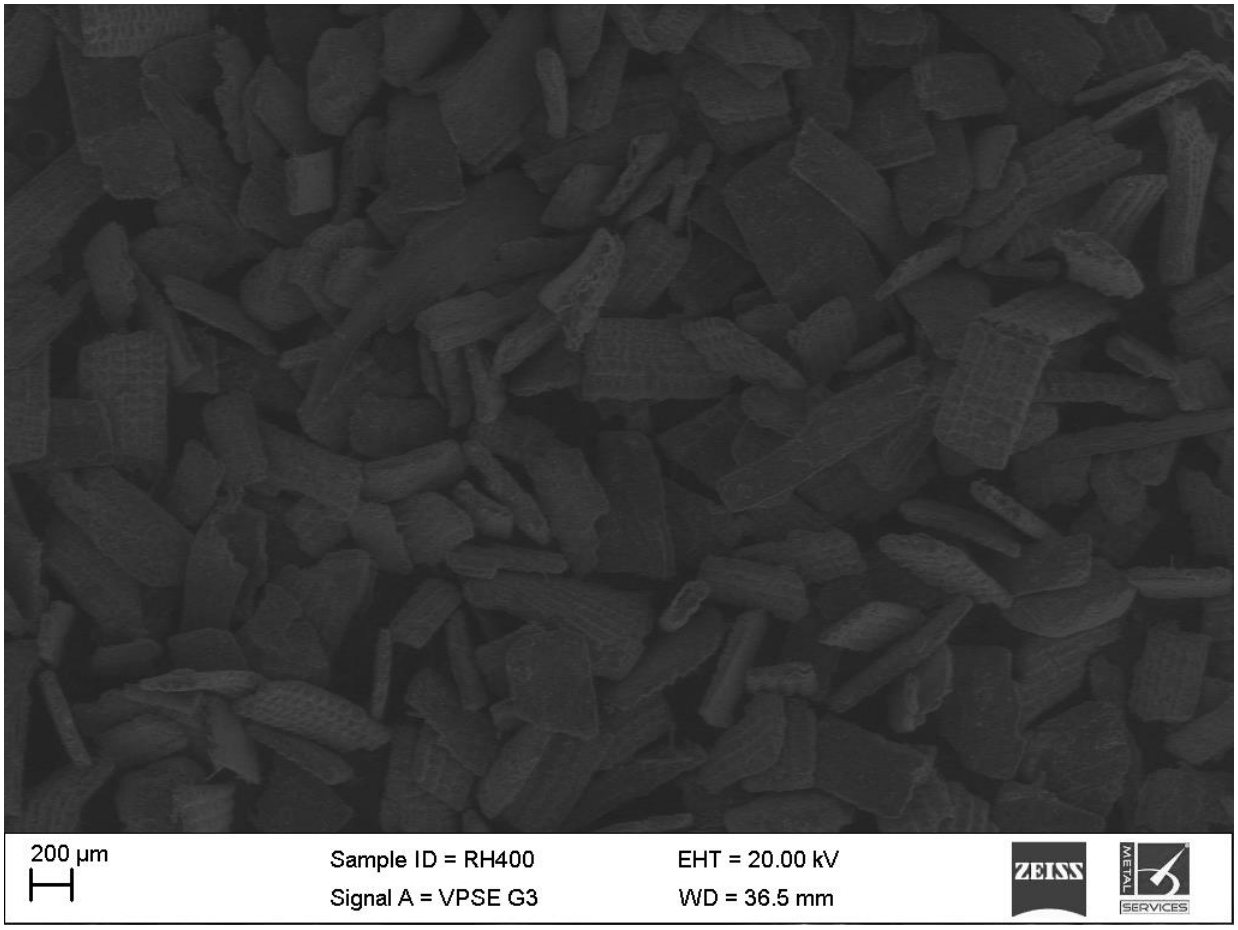


Figure S52: Scanning electron micrographs of milled rice husk fibres used for immobilization of ThC_cut1.