Accepted 27th September 2016

Fully renewable polyesters *via* polycondensation catalyzed by *Thermobifida cellulosilytica* cutinase 1: an integrated approach†

Alessandro Pellis,[a] Valerio Ferrario,[b] Marco Cespugli,[b] Livia Corici,‡[c] Alice Guarneri,[b] Barbara Zartl,[a] Enrique Herrero Acero,[d] Cynthia Ebert,[b] Georg M. Guebitz[a,d] and Lucia Gardossi*[b]

The present study addresses comprehensively the problem of producing polyesters through sustainable processes while using fully renewable raw materials and biocatalysts. Polycondensation of bio-based dimethyl adipate with different diols was catalyzed by cutinase 1 from *Thermobifida cellulosilytica* (Thc_cut1) under solvent free and thin-film conditions. The biocatalyst was immobilized efficiently on a fully renewable cheap carrier based on milled rice husk. A multivariate factorial design demonstrated that Thc_cut1 is less sensitive to the presence of water in the system and it works efficiently under milder conditions (50 °C; 535 mbar) when compared to lipase B from *Candida antarctica* (CaLB), thus enabling energy savings. Experimental and computational investigations of cutinase 1 from *Thermobifida cellulosilytica* (Thc_cut1) disclosed structural and functional features that make this serine-hydrolase efficient in polycondensation reactions. Bioinformatic analysis performed with the BioGPS tool pointed out functional similarities with CaLB and provided guidelines for future engineering studies aiming, for instance, at introducing different promiscuous activities in the Thc_cut1 scaffold. The results set robust premises for a full exploitation of enzymes in environmentally and economically sustainable enzymatic polycondensation reactions.

Introduction

United Nations Environment Programme (UNEP) has calculated that the natural capital costs of polymers and plastics used in the consumer goods industry are over $75 bn per year. Interestingly, over 75% of the known and quantifiable impacts associated with plastic use are located in the upstream portion of the supply chain across all sectors, namely impacts generated from the extraction of raw materials to the manufacturing of plastic feedstock.[1]

The dimension of the problem is even more evident when considered that about 4–6% of the world's oil production is used to produce 311 million tons of polymers and plastics (http://www.plasticseurope.org) out of a global 7% of petroleum consumption ascribable to the whole chemical sector.[4] Even if global recycling rates rose from today's 14% to more than 55% – which would be higher than the rate achieved today by even the best performing countries – annual requirements for polymeric products would still double by 2050.[2] Moreover, there are several applications – within the plastic stream – where recycling strategies appear not effective because of technical or economic reasons. Examples can be found in the cosmetic and the lubricant sectors.[3] The present study addresses comprehensively the problem of producing polyesters through sustainable processes while using renewable raw materials and biocatalysts, namely enzymes. Among polymers, polyesters are a widely used class with applications ranging from clothing to food packaging and from the car industry to biomedical applications. A number of hydrolases, and lipase B from *Candida antarctica* in particular, were reported to catalyze the synthesis of polyesters *via* polycondensation or ring opening polymerization. Enzymes are attractive sustainable alternatives to toxic catalysts used in polyconden-

[a]*University of Natural Resources and Life Sciences, Institute for Environmental Biotechnology, Konrad Lorenz Strasse 20, 3430 Tulln an der Donau, Austria*

[b]*Laboratory of Applied and Computational Biocatalysis, Department of Scienze Chimiche e Farmaceutiche, Università degli Studi di Trieste, Piazzale Europa 1, 34127 Trieste, Italy. E-mail: gardossi@units.it*

[c]*SPRIN S.p.a., Via Flavia 23/1, 34148 Trieste, Italy*

[d]*Austrian Centre of Industrial Biotechnology GmbH, Division Enzymes and Polymers, Konrad Lorenz Strasse 20, 3430 Tulln an der Donau, Austria*

‡Current address: Institute of Chemistry Timisoara of Romanian Academy, Mihai Viteazul 24, 300223 Timisoara, Romania.

sation, such as metal catalysts and tin in particular. They may work also in solvent-free systems and enable the synthesis of functionalized as well biodegradable polyesters with a controlled architecture through highly selective processes at temperatures ranging between 40 °C and 90 °C, whereas conventional polycondensations are carried out at $T > 150$ °C.[4] Enzymatic synthesis generally leads to polymers with a moderate molecular weight as compared to products obtainable *via* conventional chemical synthesis, but this drawback has been circumvented by using two-step procedures, where an initial enzymatic polymerization leads to oligomers and the second step is carried out at a higher temperature and/or at a lower pressure after removal of the biocatalyst.[5] Furthermore, the synthesis of oligomers and short telechelic prepolymers with functional ends represents an effective strategy for obtaining polymers with a higher molecular weight.[6] It must be underlined that short polyesters (<2500 Da) with highly regular structures have been used in coating and adhesive applications[7] and that biodegradable oligomers can have intrinsic interesting properties as film forming in cosmetic formulations.[8]

In the present study we report on a comprehensive approach for the sustainable synthesis of polyesters, which encompasses renewability of both raw materials and immobilized enzymes, solvent use reduction, as well as energy consumption.[8,9] Since 30% of the environmental impact of polymers and plastics is ascribable to greenhouse gas emissions from raw material extraction and processing, we focused on the esters of adipic acid and 1,4-butandiol (BDO), two renewable bio-based monomers available at an industrial scale.[9–11] Indeed, there is a wide range of structurally different bio-based chemical building blocks already available *via* fermentation or chemical transformation of a renewable carbon. They are contributing to the growth of the bio-based plastic market and to the mitigation of the environmental impact of fossil-based plastics.[10]

In order to fulfill sustainability criteria, polycondensation reactions were carried out under solvent-free conditions and exploiting thin-film systems that were demonstrated to overcome viscosity by promoting optimal heat and mass transfer, thus enabling the use of lower operational temperatures.[5,12,13] Attention was paid also to the environmental impact of the immobilization procedure: hydrolases were immobilized efficiently on a renewable lignocellulosic carrier based on milled rice husk,[14] without any further chemical functionalization of the solid matrix.

Integrated experimental and computational investigations were used to study the applicability of cutinase 1 from *Thermobifida cellulosilytica* (Thc_cut1)[15,16] in polycondensations, with the objective of enlarging the portfolio of biocatalysts in polyester synthesis. Cutinases are fungal enzymes involved in plant pathogenesis caused by the depolymerization of cutin, a three-dimensional polymer of inter-esterified hydroxyl and epoxy–hydroxy fatty acids with chain lengths mostly between 16 and 18 carbon atoms.[17] Because of their ability to accept long chain substrates, cutinases are also effective in catalyzing *in vitro* polymer synthesis. Recently, a cutinase from *Fusarium solani pisi* showed consistent synthetic activity for the production of polyamides,[18] whereas cutinase from *Humicola insolens* (HiC) was used in the polycondensation of linear dicarboxylic acids and their esters (*e.g.* adipic acid, diethyl sebacate)[19] and its application in the ring opening polymerizations of lactones.[20] We have recently reported that Thc_cut1 efficiently catalyzes the polycondensation of dimethyl adipate with different polyols leading to higher M_w (~1900) and M_n (~1000)[15] when compared to the performances of other hydrolases such as CaLB or *Humicola insolens* cutinase (HiC).[19] These observations stimulated the study presented here, where experimental and bioinformatics investigations provide a more detailed understanding of the functional properties of Thc_cut1, disclosing specific features that differentiate this serine-hydrolase from other enzymes previously used in similar polycondensation reactions, namely CaLB and HiC.

Finally, the design of the experiment study sheds light on the response of Thc_cut1 under operational conditions. Molecular dynamic simulations supported the experimental study, revealing how Thc_cut1 displays optimal efficiency under milder experimental conditions as compared to CaLB, thus enabling saving energies in terms of heating and vacuum application.

Overall, the present study provides the first comprehensive analysis of the feasibility of fully renewable and economically sustainable polyesters *via* enzymatic catalysis. There is robust evidence that the next generation of sustainable polymers and plastics calls for a change of paradigm and a closer integration between chemistry and biotechnologies.

Results and discussion

BioGPS bioinformatics analysis of Thc_cut1 *versus* serine hydrolase superfamily

Recent computational investigations[15] disclosed the structural features that make Thc_cut1 readily accessible to substrates and optimally suited for covalent immobilization. As lipases and other cutinase enzymes, it presents hydrophobic superficial regions around the active site. The wide and superficial active site of Thc_cut1 makes this hydrolase a promising catalyst for enlarging the scope of enzymatic polycondensation by synthesizing new polyesters with a larger molecular weight. The scaffold of Thc_cut1 also represents a potential starting point for further engineering strategies aiming at the synthesis of novel classes of polymers.

In the present study, we analysed cutinase 1 from *Thermobifida cellulosilytica* by means of the BioGPS bioinformatics tool[21] recently developed in our group with the objective of identifying differences and similarities with respect to CaLB and cutinase from *Humicola insolens* (HiC), two hydrolases previously employed in polycondensation reactions.[19,20]

BioGPS (Global Positioning System in Biological Space) molecular descriptors[21–24] are based on GRID derived molecular descriptors and are able to describe, mathematically, the

environment of enzyme active sites both in geometrical and electrostatic terms. Active sites were mapped using the GRID force field for evaluating the energy of non-bonded interactions and then for generating pseudo-MIFs (Molecular Interaction Fields).[25] Different probes were employed for the study of four different properties: H probe, that maps the active site shape; O probe, that evaluates H-bond donor properties; N1 probe, that accounts for H-bond acceptor capabilities; and the DRY probe, that evaluates hydrophobic interactions. The magnitude of the interaction of the N1 and O probes also includes, implicitly, information about the charge contribution, since these probes already have a partially positive and negative charge respectively.

The three-dimensional structure of HiC was taken from PDB (Protein Data Bank)[26] and corresponds to the PDB code 4OYY.[27] For Thc_cut1, an already published homology model was used[15] since the crystal structure of this enzyme is not available yet. The BioGPS descriptors for the two cutinases were calculated by using BioGPS software version 2.1 using the procedure already published. For CaLB, the 1TCA[28] crystal code of PDB was used.

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 the GRID N1 probe were centered on the carbonyl oxygen as a 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. A preliminary visual comparison between the three enzymes comes from the alignment of the corresponding pseudo-MIFs reported in Fig. 1.

The visual comparison of pseudo-MIF can simply suggest a higher degree of similarity between CaLB and Thc_cut1, whereas HiC differs significantly, especially for a less extended ability to establish H-bonds. This type of mapping provides some focused information although it cannot give comprehensive analysis of functional differences, since they depend on a complex combination of all structural and electrostatic factors and, more importantly, on their interactions.

Therefore, in order to move beyond the approach based on the visual inspection and to draw an unbiased picture of structural and functional similarities, the properties of the two cutinases were statistically compared using a mathematical model calculated on a set of 41 serine-hydrolases, which also contains CaLB. The model was computed by analyzing statistically the information contained in the BioGPS descriptors of each hydrolase[21] using Unsupervised Pattern Cognition Analysis (UPCA).[29] The complete procedure used for the calculation of the UPCA model is available in the Materials and Methods section and the set of the 41 Ser hydrolases enzymes that includes amidases, peptidases, esterases and lipases is available in Table S1 in the ESI.† The approach is based on the concept that catalytic properties depend not only on structural features but especially on the ability of active sites to establish electrostatic interactions.[30,31] The BioGPS tool mapped the active sites of all the serine-hydrolases and the information

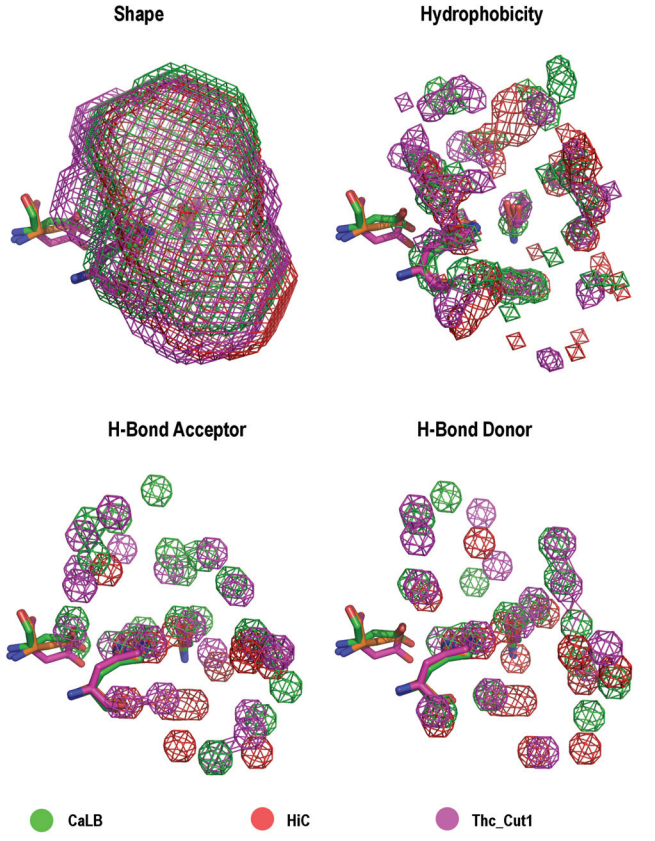

Fig. 1 Comparison of the active site properties of CaLB, HiC and Thc_cut1. Four active site properties are compared: shape, hydrophobicity, H-bond acceptor capabilities and H-bond donor capabilities. The catalytic triads (represented in stick mode) are superposed for active site orientation. PseudoMIFs are represented as wireframes and coloured in green, red and pink for CaLB, HiC and Thc_cut1 respectively.

contained in the molecular descriptors was analyzed by means of the UPCA approach. Fig. 2 reports the Global Score analysis, which encompasses all properties analyzed by the four probes. The statistical analysis sorted all hydrolases and grouped them into functionally distinct clusters, thus confirming that the structural properties explained by the BioGPS descriptors are correlated with catalytic functions of each class of serine-hydrolase.

The main advantage of this approach is represented by the visual representation of differences and similarities and it was used for "projecting" the properties of the two cutinases, calculated with the BioGPS tool, into the UPCA domain (Fig. 2). The BioGPS analysis recognizes a high level of similarity between the active sites of Thc_cut1 and CaLB (1TCA), which are projected in a region close to the esterase cluster but at the edge of the lipase group. This result supports the previous data obtained for Thc_cut1 and CaLB through molecular dynamics simulations, which indicated the absence of mobile domains responsible for interfacial activation.[15]

BioGPS information was also dissected into the UPCA plots computed for each single probe and they are reported in the ESI (Fig. S1†). The active sites of CaLB, Thc_cut1 and HiC

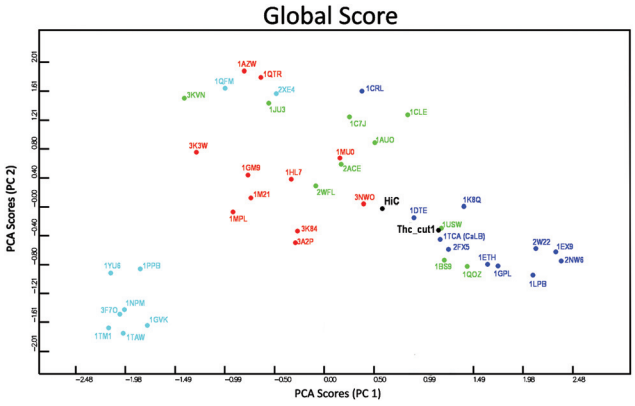

Fig. 2 Projection of cutinases from *Thermobifida cellulosilytica* (Thc_cut1) and from *Humicola insolens* (HiC) in the BioGPS model obtained through Unsupervised Pattern Cognition Analysis (UPCA). Ser hydrolases are clustered on the basis of the global similarity scores computed using the BioGPS descriptors. They are labelled according to their PDB code. Lipases are indicated in blue, esterases in green, amidases in red and proteases in cyan.

appear similar in terms of hydrophobic properties and they are positioned among esterases although close to the lipase cluster. Both Thc_cut1 and CaLB are clearly classified as esterases for their ability to establish H bonds and Thc_cut1 falls closer to the lipase group. In terms of H bond capabilities, HiC moves far away from lipases and the other two enzymes, confirming the differences observed in Fig. 1. These observations shed new light also on the behaviour of CaLB, which is recognized similar to lipases only in terms of hydrophobic properties.

Fully renewable immobilized Thc_cut1 applied in polycondensations

In our previous investigations of polycondensations of different diesters and polyols catalyzed by CaLB[12] we have demonstrated that viscosity represents a major bottleneck for solvent free enzymatic polycondensation. Optimal mass transfer and the homogeneous dispersion of the enzyme in the reaction mixture are essential for achieving a reasonable elongation of the oligomers. Although a mono-molecular dispersion of the native enzyme would lead to the highest reaction rate, immobilization of the enzyme is mandatory in order to avoid contamination of the product with the free protein and also to assure the recycling of the biocatalyst for the economics of the process.[5,7,20] We have also demonstrated that immobilized formulations characterized by a low loading of the CaLB protein are more effective as long as they are used in larger amounts to promote the dispersion of the enzyme and its accessibility.[12] However, a fully renewable and economically viable synthesis of bio-based polyesters must take into account waste production, energy consumption as well as the environmental impact of all raw materials, including immobilized biocatalysts. On that respect, it has been reported that fossil based methacrylic carriers account for a relevant part of greenhouse gas emissions due to the biocatalyst.[32] In a recent study, we have proposed rice husk (RH) as a renewable alternative carrier suitable for immobilization of CaLB and other hydrolases.[33] This inexpensive and massively available biomass (globally 100 Mt per year) is composed of SiO_2 and three different biopolymers, namely cellulose, hemicellulose and lignin.[34] A preliminary analysis of the sustainability of immobilization protocols employing rice husk has been previously reported.[15] Besides being fully renewable, at the end of its proposed applications RH can still be used as a carbon source for anaerobic digestion for biogas production, as a cellulose source for bioethanol fermentation, burnt for generating energy or exploited in different manufacturing sectors[35] in accordance with the circular economy principles.[36]

It must be noted that in our previous study all attempts of adsorbing CaLB on rice husk particles in aqueous buffer were unsuccessful (<5% adsorbed). Therefore, CaLB was covalently immobilized on rice husk particles, which were oxidized at the cellulosic component and later functionalized with a hexamethylene spacer. Looking for simpler and economical protocols for covalent immobilization, in the present study we verified that Thc_cut1 can be adsorbed in moderate yield (36% in 24 h) on rice husk particles (rThc_cut1) with a diameter in the range of 200–400 μm. In order to prevent protein leaching, the enzyme was cross-linked using glutaraldehyde. Interestingly, Table 1 shows that, despite the low loading, rThc_cut1 has a higher hydrolytic activity as compared to a reference formulation obtained by immobilizing a much higher amount of cutinase (about 7 folders) on a commercial epoxy-activated methacrylic resin.

Two different drying methods, namely freeze-drying and evaporation at environmental pressure and at 30 °C for 48 h, led to comparable enzymatic activity, although the lyophilization allows for a more extensive dehydration (ESI, Table S2†).

Table 1 Activity of Thc_cut1 immobilized on rice husk (rThc_cut1) and on epoxy-methacrylic resin Sepabeads EC-EP/M (iThc_cut1)

Enzymatic Preparation	Protein loading (mg g_{dry}^{-1})	Loaded protein (%)	Drying method	Hydrolytic activity[a] (U g^{-1})	Water content (% w/w)
rThc_cut1	4	36	Air-dried	35 ± 5	4
Lyo_rThc_cut1			Freeze drying	35 ± 3	0.2
iThc_cut1	10	99	Air-dried	13 ± 2	3
Lyo_iThc_cut1			Freeze drying	12 ± 3	0.2

[a] Calculated *via* the PNPB hydrolytic assay. All determinations were conducted in duplicates.

Table 2 Polycondensation of DMA with BDO catalyzed by different formulations of Thc_cut1 after 24 h

Entry	Enzymatic preparation	Amount (% w/w)	P (mbar)	M_w [a]	M_n [a]	PD[a]	Conv.[b] (%)
1	rThc_cut1	10	1000	594	434	1.37	64
2	rThc_cut1	10	70	230	222	1.04	18
3	rThc_cut1	30	70	487	343	1.42	62
4	Lyo_rThc_cut1	10	1000	587	384	1.53	62
5	Lyo_rThc_cut1	10	70	224	203	1.10	13
6	iThc_cut1	10	1000	1923	985	1.95	86[13]
7	iThc_cut1	10	70	480	290	1.66	48
8	Lyo_iThc_cut1	10	1000	1120	656	1.71	78
9	Lyo_iThc_cut1	10	70	319	275	1.16	43

[a] Calculated *via* GPC calibrated with low molecular weight polystyrene standards 250–70 000 Da. [b] Calculated *via* ^{1}H-NMR by comparing the ratio between the signals of methylene groups adjacent to –OH of BDO and the methylene groups of DMA (assumed as constant). All reactions were performed in duplicates.

The preparations were applied and compared in the polycondensation of dimethyl adipate (DMA) and 1,4-butanediol (BDO), two bio-based monomers widely used in polymer synthesis. Polycondensations were conducted using a thin-film solvent-free system[4] at 70 °C and working either at atmospheric pressure or at 70 mbar.

The results in Table 2 indicate that by working with 10% of formulations of rThc_cut1 with a very low loading of protein it is possible to achieve in 24 h more than 64% of conversion and oligomers made by up to 7 units. The complete characterization of the products is available in the ESI (Fig. S2–S15†). Unexpectedly, working under reduced pressure and with very dry preparations does not translate into higher conversions or elongation. Rather, there is a negative effect of reduced pressure, whereas the effect of water is less evident. When working under vacuum (entry 3) it is necessary to use a triple amount of the biocatalyst for reaching conversions observed in entry 1 that was carried out at 1000 mbar. ESI-MS analysis in Fig. 3 confirms this trend.

The same behavior was observed in the polycondensations catalyzed by Thc_cut1 immobilized on epoxy methacrylic resins, indicating that the effect is not dependent on the nature of the carrier.

The negative effect of vacuum emerged also from the study of the polycondensation of DMA with ODO (see ESI Fig. S16–S24† for complete characterization of products). Also in this case using 30% w/w of the biocatalyst at 70 mbar leads to comparable conversion (44%) with the 10% w/w 1000 mbar reaction (39%) (Table 3).

This behavior appears to be in contradiction with all previous reports on polycondensation catalyzed by CaLB,[7,12,37] where the application of vacuum had the beneficial effect of promoting the removal of the leaving nucleophile (either alcohol or water), thus preventing the reversing of the reaction and promoting higher yields and molecular weights.

Indeed, the experiments performed using CaLB immobilized on the same epoxy activated methacrylic resin (Table 4, entries 1 and 2) confirmed that it is necessary to apply vacuum for achieving satisfying yield and elongation when traces of water are still present in the enzymatic preparation (3%). Also in the case of Novozym® 435 (NZ 435®), the enzymatic preparation most widely used in previous polycondensation studies, it is evident that the reduced pressure exerts a dramatic effect

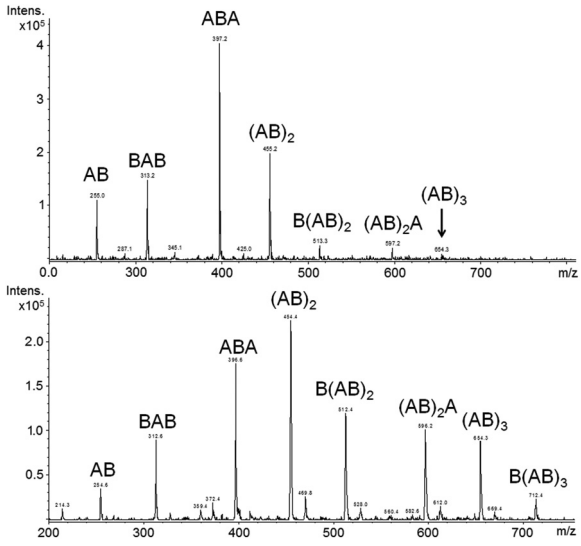

Fig. 3 ESI-MS positive ion mass spectrum of the polycondensation products of DMA with BDO catalyzed by 10% w/w rThc_cut1 at 24 h and 70 mbar (top) and 1000 mbar (bottom).

Table 3 Polycondensation of DMA with ODO using the rThc_cut1 preparation as a catalyst

Entry	Amount (% w/w)	Vacuum (mbar)	M_w [a]	M_n [a]	PD[a]	Conv.[b] (%)
1	10	1000	508	393	1.29	44
2	10	70	354	289	1.22	8
3	30	70	447	298	1.50	39

[a] Calculated *via* GPC calibrated with low molecular weight polystyrene standards 250–70 000 Da. [b] Calculated *via* ^{1}H-NMR by comparing the ratio between the polyol methylene groups adjacent to –OH area and the internal methylene groups area of DMA (assumed as constant). All reactions were performed in duplicates.

Table 4 Polycondensation of DMA with BDO catalyzed by 10% (w/w) of different formulations of CaLB after 24 h of reaction

Entry	Biocatalyst	H_2O % (w/w)	p (mbar)	M_w	M_n [a]	PD[a]	Conv.[b] (%)
1	iCaLB	3	1000	888	528	1.68	76[15]
2	iCaLB	3	70	6947	3162	2.20	90
3	NZ 435®	1	1000	1040	561	1.85	78[15]
4	NZ 435®	1	70	8357	1759	4.75	96
5	LyoNZ435®	0.1	1000	999	608	1.64	76
6	LyoNZ435®	0.1	70	8250	2438	3.38	94

[a] Calculated *via* GPC calibrated with low molecular weight polystyrene standards 250–70 000 Da. [b] Calculated *via* ^{1}H-NMR by comparing the ratio between the signals of methylene groups adjacent to –OH of BDO and the methylene groups of DMA (assumed as constant). All reactions were performed in duplicates.

on the elongation of the polyester, even when the enzyme is employed in a highly dried form (0.1% w/w). It must be noted that a higher M_n obtained using Novozym® 435 is a consequence of the detachment and dispersion of the enzyme adsorbed on the acrylic support, as also experimentally demonstrated in previous studies.[5] We previously demonstrated that both CaLB and Thc_cut1 can be recycled effectively for at least ten cycles when covalently immobilized either on RH or on epoxy methacrylic resins, with negligible protein leaching.[4,14,15]

Globally, the data here reported suggest that CaLB and Thc_cut1 are differently affected by experimental conditions. These differences show no direct dependence on the immobilization carrier, the substrate or the dehydration method.

Therefore, to the best of our knowledge, this is the first report on the complexity of effects of experimental variables on the efficiency of enzymes in enzymatic polycondensation and we considered that the issue had to be analyzed in detail by means of a design of the experiment approach.

Multivariate analysis of the effect of experimental variables on polycondensation catalyzed by immobilized Thc_cut1

In a traditional investigative approach, 'one factor at a time' is varied, with the remaining factors held constant. This method provides an estimate of the effect of the single variable under selected fixed conditions, but if the variables do no act additively (*i.e.* they interact), the resulting information is incomplete and the real optimum of a system can be identified only with difficulty or by chance.[38] By using statistical experimental designs (*i.e.* factorial design), sets of experimental conditions are selected according to a systematic plan, so that statistical analysis of data provides information about the impact of factors or their interactions on the response considered.[39]

In the present study we used a fractional factorial design to quantitatively evaluate the effect of temperature, pressure and water on diol conversion and product elongation (M_n). These experimental variables were studied in the context of the polycondensations of DMA with two different diols: BDO and ODO. A fractional factorial design was planned by setting a low and a high value for each variable, thus defining the experimental domain to be explored. Additionally, a central point – corresponding to intermediate levels of the variables – was run for each diol twice. Table 5 reports the plan of the experiments, consisting of 12 polycondensation reactions, and the obtained experimental results.

Table 5 Scheme of the experimental design developed for evaluating the effect of experimental conditions on the polycondensation (24 h) of DMA with BDO or ODO catalyzed by immobilized Thc_cut1

T (°C)	P (mbar)	H_2O (μL)	Yield (%)	M_n (Da)
		BDO		
30	70	0	85	901
70	1000	0	66	514
70	70	10	22	359
30	1000	10	50	417
50	535	5	96	1893
50	535	5	90	1266
		ODO		
70	70	0	6	271
30	1000	0	54	423
30	70	10	66	716
70	1000	10	53	434
50	535	5	85	1315
50	535	5	85	1221

The effect of moisture was studied by adding defined volumes of water to the reaction mixture. It must be noted that the water activity is the actual informative parameter affecting polycondensation and it is expected to vary as a function of the diol and also throughout the course of the reaction. In order to acquire more accurate information on the free water available in the systems and its distribution among the different phases at least at the starting of the reactions, the vapor pressure was measured in the gas phase using a humidity sensor. In principle, at equilibrium, the "free" water should be the same in all phases of the closed system, so that by measuring the vapor pressure (or RH) in the gas phase we know the a_w value of the system and, ultimately, the water available for promoting undesired hydrolytic reactions or for increasing the hydration state of the enzyme.[40]

The system containing ODO reached constant values of vapor pressure after 24 h of equilibration whereas for BDO the system took 96 h to reach equilibrium (Table 6). Most probably this behavior depends on the presence of traces of water in the highly hygroscopic BDO. The distribution of such water between the different phases occurred very slowly, as suggested by a progressive increase of RH from 34% (24 h) to 49% (96 h).

Table 6 a_w measured in the sealed vessels containing the reaction mixtures and the immobilization carrier. Values refer to systems made with 0.006 mol of each monomer in a 50 mL round-bottom flask equilibrated at 30 °C until constant readings

Monomer mixture	Water added 0 μL	Water added 5 μL	Water added 10 μL
DMA-BDO (96 h)	a_w 0.49	a_w 0.63	a_w 0.61
DMA-ODO (24 h)	a_w 0.34	a_w 0.35	a_w 0.51

Overall, the data in Table 6 indicate that at the starting of the reactions all different systems studied with the factorial design had a_w values ranging between 0.34 and 0.63.

Fig. 4 reports a global visual representation of the data obtained with BDO and ODO, with the results displaced at the edges of the cubes. The two centers report the average of values obtained from the replicates using BDO and ODO and provide information on the effect of the diol structure.

Data clearly indicate that immobilized Thc_cut1 displays the best performances at intermediate values of all experimental variables considered in the study.

The variable importance plots (VIP plots) explain the effect of each single variable and, most importantly, of their interactions (Fig. 5). They were calculated by linear regression for both responses, where the effect corresponds to the difference between the averages of responses obtained at a high and low level of each experimental factor. Details are available in the ESI, Fig. S52.†

Temperature is the factor that exerts a major and detrimental effect on the reaction, indicating that Thc_cut1 prefers milder temperatures to 70 °C. At 30 °C the enzyme is efficient but most probably the temperature is insufficient for increasing the fluidity of the reaction mixture, so that 50 °C appears to be a suitable temperature for the chosen reactions.

Interestingly, pressure does not exert a great effect on conversion and M_n but rather the interaction between pressure and temperature causes the most negative effect on the reactions. The results of the design indicate that the enzyme is severely affected by vacuum at high temperature, whereas the negative effect of vacuum is not relevant at low temperature.

It is note worthy that the chemometric approach disclosed the disruptive effect of this interaction, whereas the multivariate analysis of the preliminary data induced to attribute a negative effect on the vacuum alone.

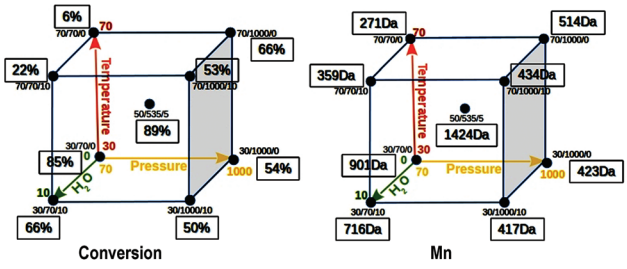

Fig. 4 Tridimensional representation of conversions and M_n obtained in the reactions of the factorial design. Numbers reported in the cubic space represent the average of values obtained from the corresponding reactions carried out with BDO and ODO.

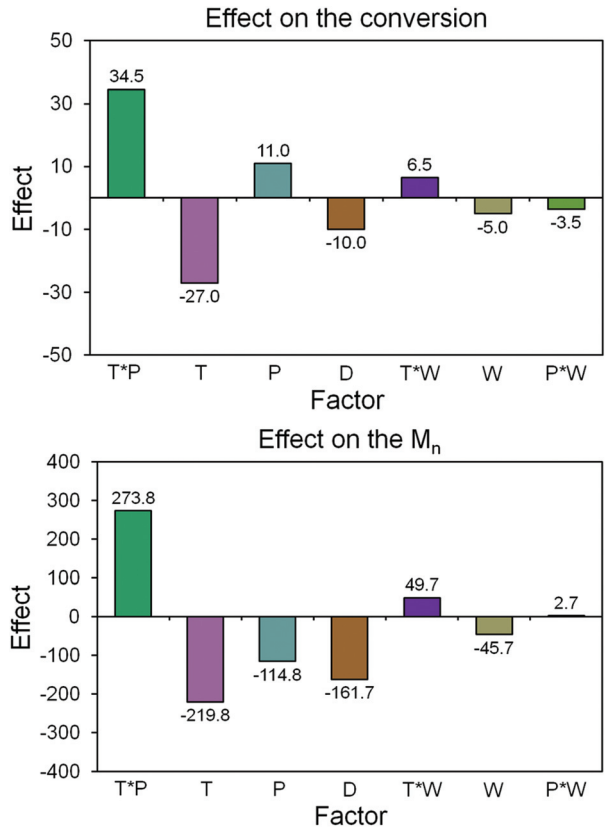

Fig. 5 Analysis of effect of variables and of their interactions on reaction conversion and M_n.

The structure of the diol is also relevant for the elongation of the reaction, whereas the conversion is not much affected. This finding suggests that the bulkiness of the substrates gains increasing importance as the reaction proceeds and the molecular weight of the oligoesters grows. It must be also taken into account that BDO is a viscous liquid, whereas ODO is a solid (melting point between 57 and 61 °C) that is suspended in the DMA and melts upon heating at 50 °C. The corresponding intermediate products of ODO are more viscous, thus hampering the elongation of the oligoesters.

Unexpectedly, the water added to the system does not exert a major effect on M_n, and polycondensation proceeds even in the presence of small volumes of water (5 μL for 0.006 moles of each monomer), at a mild temperature (50 °C) and under vacuum (535 mbar). In summary, these findings indicate that immobilized Thc_cut1 works optimally under conditions quite different from those employed so far for the polycondensation reactions catalyzed by CaLB. A comprehensive review by Gross and co-workers[20] reported a list of polycondensations catalyzed by CaLB carried out between 70 and 95 °C and most studies underlined how heating must be used in combination with high vacuum for removing traces of water or the leaving nucleophile (*i.e.* alcohol or water), thus promoting the shifting of equilibrium towards synthesis.[5,12]

Indeed, Table 4 above reports that CaLB immobilized on epoxy methacrylic resins leads to polyesters of DMA and BDO with M_ws of almost 7000 Da and M_n above 3000 Da when working at 70 °C and 70 mbar, whereas at atmospheric pressure M_n falls dramatically to 528 Da. In order to further compare the behavior of CaLB with Thc_cut1 we carried out the same polycondensation catalyzed by CaLB at 50 °C. The results are reported in the ESI (Fig. S48–S51†). Although the conversion of the diol was 90%, surprisingly ESI mass spectrometry shows the formation of polyesters with M_ws lower than 1000 Da. More importantly, the spectra pointed out the formation of hydrolytic products as well cyclic polyesters. After 72 h of reaction the elongation does not improve but rather the formation of hydrolysis products is even more evident. Apparently CaLB is more sensitive to traces of water, which affects the equilibrium of the reaction.

Molecular dynamics simulations of CaLB and Thc_cut1 at different temperatures and pressure

In the next step of the investigation we intended to investigate whether conformational changes are on the basis of the different behaviour of CaLB and Thc_cut1. Molecular Dynamics simulations were run under different conditions of temperature and pressure for 10 ns by using the software GROMACS version 4, OPLS force field definitions and explicit water as a solvent. Each enzyme was subjected to three different simulations (6 MD simulation in total) in terms of temperature and pressure conditions that are highlighted in Table 7.

Each enzyme structure was analyzed and compared on the basis of RMSF calculated on the protein Cα. The RMSF analysis of Thc_cut1 (see ESI Fig. S46†) shows that the enzyme has a similar behaviour under all conditions, with a slightly higher average enzyme mobility at 343 K, as normally expected. More relevant conformational mobility differences are observable for the terminal parts of the protein chain. The overall conformational stability of Thc_cut1 is also confirmed by the superimposition of the structures resulting from the three simulations (Fig. 6). On the other hand, Fig. 6 highlights some differences not revealed by the RMSF analysis and mainly related to side chain movements in an area flanking the active site corresponding to residues 86–91.

At 343 K and 70 mbar (blue structure) the active site appears slightly less accessible than at atmospheric pressure (red structure), as a consequence of conformational changes of residues 86–91, although it is difficult to establish a direct link between these observations and the negative effect of high temperature and vacuum on the enzyme. Nevertheless, the MD simulations did not show fast and disruptive conformational changes caused by the combination of high temperature and vacuum.

RMSF analysis performed on the Cα of CaLB (see ESI Fig. S47†) shows clear differences in the enzyme behaviour under different simulated conditions. Such differences are even more evident by superimposing the structures (Fig. 7).

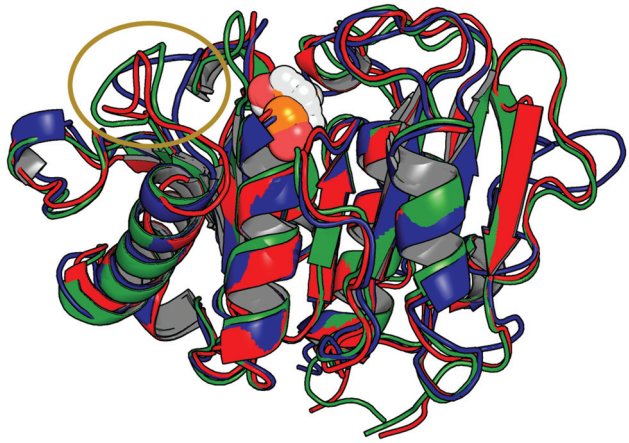

Fig. 6 Superimposition of the structures of Thc_cut1 resulting from the three different MD simulations: condition 1 (300 K, 1000 mbar) in green, condition 2 (343 K, 1000 mbar) in red; condition 3 (343 K; 70 mbar) in blue. Catalytic serines are represented in sphere mode; residues 86–91 in the proximity of the enzyme active site are highlighted in the yellow circle.

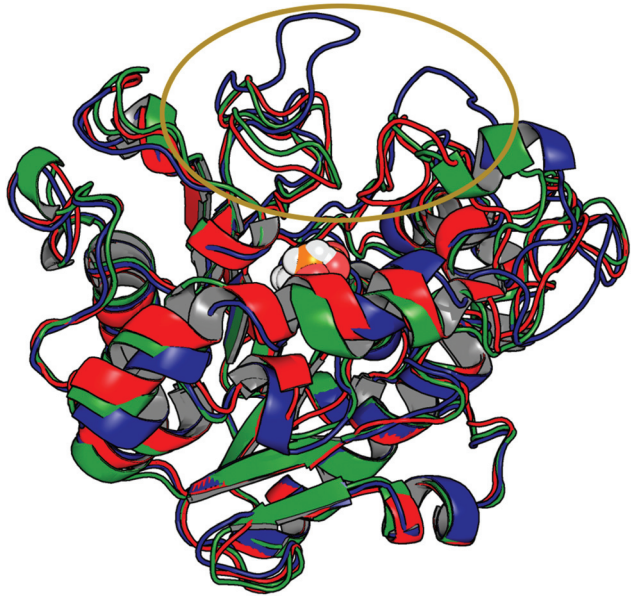

Fig. 7 Superimposition of the structures of CaLB resulting from the three different MD simulations: condition 1 (300 K, 1000 mbar) in green, condition 2 (343 K, 1000 mbar) in red; condition 3 (343 K, 70 mbar) in blue. Catalytic serines are represented in sphere mode; residues 86–91 in the proximity of the enzyme active site are highlighted in the yellow circle.

Table 7 Temperature and pressure settings of the 3 MD simulations performed on each enzyme

Simulation	Temperature	Pressure
1	300 K	1000 mbar
2	343 K	1000 mbar
3	343 K	70 mbar

Fig. 7 indicates that the protein loops at the active site entrance, corresponding to residues 138–143 and 187–193 (highlighted in the yellow circle), change their conformations on passing from physiological conditions (green structure: 300 K; atmospheric pressure) to high temperature and low pressure (343 K, 70 mbar; in blue) that widen the active site shape. Notably, no enlargement of the active site is observable in the simulation run at high temperature and atmospheric pressure (343 K, 1000 mbar; in red).

Although MD simulations provide some preliminary important hints on the stability of ThC_cut1 and the accessibility of the active sites of both hydrolases, the different effects of temperature and pressure on the two enzymes appear strictly linked to the role of water inside active sites. BioGPS analysis reported above clearly indicates that the H-bond capabilities of CaLB are different from what was observed for most lipases.

It must be noted that several authors have already discussed the possible effect of water molecules entrapped in the stereospecificity pocket of CaLB[40,41] and this issue appears of crucial importance for the optimal exploitation of this enzyme in polycondensations. Future engineering strategies of CaLB might follow that direction, also taking inspiration from the positive features displayed by Thc_cut1.

Conclusions

Enzyme catalyzed reactions, because of their selectivity and efficiency, have the potential to confer advanced and unique properties to polyesters, while boosting a sustainable innovation towards more benign polymers and synthetic processes. Here we demonstrated the feasibility of fully renewable enzymatic synthesis of polyesters by addressing comprehensively the environmental, technological and economic constraints. Bio-based monomers were used in polycondensations catalysed by a cutinase enzyme (Thc_cut1) immobilized on renewable and inexpensive carriers made by using milled rice husk following a simple and economic protocol of adsorption and crosslinking thus overcoming the environmental impact of conventional fossil-based immobilization carriers. The biocatalyst works efficiently in solvent-free systems and under very mild reaction conditions (50 °C) as compared to conventional polycondensations employing toxic metals (T > 150 °C). As compared to lipase B from *Candida antarctica*, Thc_cut1 is less sensitive to the presence of water in the system and this feature makes achieving chain elongation possible at a pressure closer to atmospheric pressure without observing hydrolytic reactions. Moreover, by operating under thin-film conditions, mass and heat transfer are optimized, with consequent energy saving.

These findings set robust premises for further exploitations of this cutinase enzyme in different environmentally and economically sustainable synthesis of biorenewables.[42]

Future investigations might address engineering strategies to enlarge the specificity and catalytic properties of Thc_cut1. The accessible and superficial active site of Thc_cut1 makes this enzymatic scaffold very attractive for polymer synthesis and transformation of bulky substrates in general. Information coming from BioGPS bioinformatics analysis provides guidelines for future engineering studies aiming, for instance, at introducing different promiscuous activities in the Thc_cut1 scaffold. On that respect, Bio_GPS analysis disclosed the similarities between Thc_cut1 and CaLB, which share the hydrophobic nature of lipases but appear to be functionally closer to esterases.

Methods

Chemicals and reagents

EC-EP/M Sepabeads were kindly donated by Resindion S.r.l., (Mitsubishi Chemical Corporation, Milan, Italy). Dimethyl adipate (DMA) was purchased from Sigma-Aldrich. 1,4-Butanediol (BDO), and 1,8-octanediol (ODO) were purchased from Merck. All other chemicals and solvents were also purchased from Sigma-Aldrich at reagent grade, and used without further purification if not otherwise specified. The samples of rice husk (carnaroli type) were kindly donated by Riseria Cusaro S.r.L. (Binasco, Italy).

Enzymes

The recombinant *Thermobifida cellulosilytica* cutinase 1 (Thc_cut1) was produced and purified as previously described.[43] The organism used for the expression was *E. coli.* Novozym® 435 was purchased from Sigma-Aldrich (product code: L4777) containing *Candida antarctica* lipase B immobilized on macroporous acrylic resin, exhibiting a synthetic activity of 11 700 PLU U per g propyl laurate units). Lipozyme CaLB (protein concentration of 8 mg mL^{-1}) was a kind gift from Novozymes (DK).

Protein structures

All the protein structures used as the dataset for BioGPS model generation (Table S1 in the ESI†) were retrieved from the Protein Data Bank (PDB)[26] and preprocessed by using the software PyMOL.[44] All molecules but the proteins were deleted (*i.e.* water molecules, inhibitors, glycosylation residues, *etc.*). The original protein structure coordinates (from the PDB) were used as inputs, without any previous superimposition.

The same approach was used for the structure of *Humicola insolens* cutinase (HiC) (PDB code 4OYY)[27] whereas the structure of cutinase from *Thermobifida cellulosilytica* (Thc_cut1) comes from an already published homology model.[15]

BioGPS model

The BioGPS (Global Positioning System in Biological Space) model was obtained as already published[21] using BioGPS software version 2.1. BioGPS is based on the "Common Reference Framework" composed by two main steps: the characterization of the protein active sites and their comparison. The active site of each enzyme was automatically detected by the FLAPsite

algorithm.[23] First the algorithm reduces the complexity of the pseudo-MIFs selecting a number of representative points using a weighted energy-based and space-coverage function. Then it generates all possible combinations of four points; each combination is termed "quadruplet". All possible quadruplets for each mapped active site were generated and stored into a biofingerprint (bitstring). Each active site was then compared within the Common Reference Framework using an "all against all" approach where each enzyme active site is compared with itself and with all the other enzyme active sites. At the end, the algorithm generates a set of Tanimoto scores[45] represented by square matrixes, namely a series of probe scores (one for each original probe) together with a global score. The set of Tanimoto scores was used as an input for the UPCA algorithm which generated the final BioGPS statistical model by generating the multidimensional space (Principal Components – PCs) which differentiated each enzyme active site on the bases of their similarities and differences.

Projection of cutinases

The two cutinases Hic and Thc_cut1, were projected into the BioGPS model. These two enzymes were processed as described in the previous section. Each cutinase was compared with itself and with all the other enzymes. Finally the Tanimoto scores[45] (see the ESI† for details) were used for the projections of the two cutinases into the BioGPS model computed for the 41 ser-hydrolases.

Milling of rice husk

Rice husk was milled using a Rotor mill ZM 200 (Retsch S.r.l., Bergamo, Italy) according to a procedure already described.[14] The raw material was separated by size using sieves of 450 and then 200 mm. The wet particles were weighed and then dried in an oven at 120 °C for 6 h. The density of RH before milling was 0.153 g mL^{-1} whereas the milled RH (size 0.2–0.4 mm) had a density of 0.437 g mL^{-1}. The full characterization of rice husk has been previously reported.[14]

Immobilization of Thc_cut1 on milled rice husk (200–400 μm diameter)

The rice husk powder (200–400 μm) was washed with an ethanol/distilled H_2O mixture (2×) 10 min each step (4 mL g^{-1} dry material). Then the particles were rinsed with distilled H_2O (2 times). The moisture content was determined as described above and an amount corresponding to 1 g of dry rice husk powder was suspended in 10 mL of 0.4 mg mL^{-1} enzyme solution in 0.1 M Tris-HCl buffer pH 7.0 at room temperature for 24 h on a rotating wheel. The progress of the immobilization was monitored by evaluating the residual protein concentration in the supernatant after 4, 8 and 24 h and data are reported in the ESI.† The supernatant containing the unbound enzyme was discharged and the cutinase adsorbed on the rice husk was subjected to a cross-linking step using 5 mL solution of glutaraldehyde 1.5% in potassium phosphate buffer 0.1 M pH 8.0, for 4 h at room temperature. The enzyme preparation was extensively washed with 0.1 M Tris-HCl buffer pH 7 in order to remove all the non-bound protein on the support, monitoring UV absorbance of washing solutions. The preparation was left to dry on a Büchner filter for 48 h at 30 °C and atmospheric pressure prior to use (if not otherwise specified).

Immobilization of Thc_cut1 on epoxy activated carrier

The epoxy-activated beads were washed with ethanol (2 times) and double distilled H_2O (2 times) prior to use. A total of 1.0 g of dry epoxy-activated beads were suspended in 10 mL of 1 mg mL^{-1} enzyme solution in 0.1 M Tris-HCl buffer pH 7 at 21 °C for 24 h on a blood rotator. The samples were withdrawn over time. The progress of the immobilization was monitored by evaluating the residual activity and the protein concentration in the supernatant. It must be noted that Tris-HCl buffer was selected as immobilization medium because native Thc_cut1 was produced in this same buffer and the exchange of buffer would cause a loss of enzymatic activity (data not shown). After the immobilization, the enzyme preparations were extensively washed with 0.1 M Tris-HCl buffer pH 7 in order to remove all the non-covalently bound protein adsorbed on the support. Finally, in order to block the unreacted epoxy groups, the enzymatic preparations were incubated in 45 mL of 3 M glycine for 24 h at 21 °C as previously reported.[15] The enzyme preparations were extensively washed with 0.1 M Tris-HCl buffer pH 7.

Moisture determination

0.2 g of immobilized enzymatic preparation was weighed in a tarred weighing bottle (A), dried for 6 h at 120 ± 5 °C, cooled down in a dessicator until a constant weight was reached and then weighed again (B). The moisture content was calculated as follows:

$$\text{Moisture content } (\%) = [(A - B)/A] \times 100$$

Activity assay for immobilized enzymes

Activity was measured at 21 °C using PNPB as a substrate. The final assay mixture was made up of 0.1 mL of the substrate solution (86 μL of PNPB and 1000 μL of 2-methyl-2-butanol), 11 mL of 100 mM Tris-HCl buffer at pH 7 and 20 mg of immobilized enzyme preparation. The increase of the absorbance at 405 nm due to the hydrolytic release of *p*-nitrophenol (ε 405 nm = 9.36 mL (μmol cm)$^{-1}$) was measured over time with a HACH Lange benchtop spectrophotometer using plastic cuvettes. A blank was included using beads where glycine was used instead of an enzyme as a blocker for the epoxy-activated beads. The activity was calculated in units (U), where 1 unit is defined as the amount of the enzyme required to hydrolyze 1 μmol of the substrate per minute under the given assay conditions.

Protein quantification

Protein concentrations were determined by using the BioRad protein assay (Bio-Rad Laboratories GmbH, Vienna, Cat. No: 500-0006). Briefly, 10 μL of the sample was added into the

wells of a 96-well micro-titer plate (Greiner 96 Flat Bottom Transparent Polystyrene). As soon as all the samples were placed into the wells, 200 μL of the prepared BioRad reaction solution were added to the wells (BioRad Reagent diluted 1 : 5 with mQH_2O). The plate was incubated for 5 min at 21 °C and 400 rpm. The buffer for protein dilution (0.1 M Tris-HCl pH 7) was used as a blank and BSA (bovine serum albumin) as the standard. The absorption after 5 min was measured at λ = 595 nm and the concentration calculated from the average of triplicate samples and blanks.

Enzymatic polycondensation using a thin-film reaction system

The thin-film enzymatic polycondensations were conducted as previously reported by Pellis *et al.*[4] Reactions based on the factorial design were carried out using equimolar amounts (0.006 mol) of diester and diol (0.006 mol) and 10% w/w of immobilized iThc_cut1. The reactions were conducted in 50 mL round-bottom flasks connected to a rotary evaporator, applying reduced pressure when requested. The reactions were operated with a Rotavapor R-215 (BÜCHI) connected to a vacuum pump Vac® V-513 (BÜCHI) and a pressure controller V-800 (BÜCHI). Temperature was controlled by means of a Waterbath B-480 (BÜCHI). After 24 h, the reaction mixture was recovered with THF and products were characterized after filtration of the biocatalyst and evaporation of THF, without any further treatment.

Measurement of water activity

Water activity of the reaction system was determined at 30 °C by using a hygrometer (DARAI-Trieste, Italy) operated and calibrated as previously reported.[39] The monomers, the porous immobilization carrier without the enzyme and different volumes of water were added, mixed and thermostated at 30 °C in the same round-bottom flasks used for carrying out the reactions. We assumed that the very small amount of protein linked to the carrier binds to a negligible percentage of water if compared with the huge bead surface and therefore does not affect the a_w of the system. The humidity sensor was maintained in contact with the gas phase within the sealed vessels and the values of relative humidity (RH) were measured until constant readings (a_w), namely the achievement of equilibrium.

The relative humidity at equilibrium is correlated with a_w by the following equation: ERH = $a_w \times 100$.

GPC

The samples were dissolved in THF (250 ppm BHT as inhibitor). Gel permeation chromatography was carried out at 30 °C on an Agilent Technologies HPLC System (Agilent Technologies 1260 Infinity) connected to a 17 369 6.0 mm ID × 40 mm L H_{HR}-H, 5 μm guard column and a 18 055 7.8 mm ID × 300 mm L GMH_{HR}-N, 5 μm TSK gel liquid chromatography column (Tosoh Bioscience, Tessenderlo, Belgium) using THF (250 ppm BHT as inhibitor) as an eluent (at a flow rate of 1 mL min^{-1}). An Agilent Technologies G1362A refractive index detector was employed for detection. The molecular weights of the polymers were calculated using linear polystyrene calibration standards (250–70 000 Da).

^{1}H-NMR

Nuclear magnetic resonance ^{1}H measurements were performed on a Bruker Avance II 400 spectrometer (resonance frequencies 400.13 MHz for ^{1}H) equipped with a 5 mm observe broadband probe head (BBFO) with *z*-gradients. $CDCl_3$ was used as an NMR solvent if not otherwise specified.

Electrospray ionization mass spectrometry (ESI-MS)

The crude reaction mixtures were analyzed on Esquire 4000 (Bruker) electrospray positive ionization by generating the ions in an acidic environment. Around 10 mg of the sample was dissolved in 2 mL of methanol containing 0.1% v/v formic acid. The generated ions were positively charged with *m*/*z* ratio falls in the range of 200–1000. The subsequent process of deconvolution allows the reconstruction of the mass peaks of the chemical species derived from the analysis of the peaks generated.

Planning of the factorial design

The factorial design was planned using the software MODDE 8.0.2 (MSK Umetrics). Polycondensations were carried out on a thin film and under solvent-less conditions using Thc_cut1 immobilized on epoxy methacrylic resin and following the procedures described above. Experiments were run following a random sequence. The independent variables considered were temperature, pressure and added water. They were studied at two levels (+1 and −1). Temperature = (70 °C; 30 °C); pressure = (1000 mbar; 70 mbar); and H_2O = (10 μL; 0 μL). The factorial design is composed by the combination of the different levels of the three variables. The effect of the diols (BDO and ODO) was evaluated by dividing the factorial design into two blocks, one for each diol. For each block, a further central point was added, corresponding to the combination of the combination of each independent variable taken at an intermediate level (level 0: T = 50 °C; p = 535 mbar; H_2O = 5 μL). The responses measured for each experiment were the conversion of the diol (calculated by NMR spectroscopy as described above) and the M_n of the product (calculated by GPC analysis). The effects of factors on responses were calculated by multiple linear regression.

Molecular dynamics simulations

The structure of CaLB (PDB 1TCA) and Thc_cut1 (homology model)[15] were protonated at pH 7.0 using the PDB2PQR server[46] based on the software PROPKA.[47] Subsequently, each protonated enzyme structure was defined into an OPLS force field.[48] Each protein was inserted in a cubic box of 216 nm^3 and solvated with an explicit solvent (TIP4 water type).[49] Thus, each enzyme system was minimized using the software GROMACS version 4[50] using the steepest descendent algorithm for 10 000 steps. Afterwards, each enzyme was simulated under three different conditions: (1) 300 K and 1000 mbar; (2) 343 K and 1000 mbar; and (3) 343 K and 70 mbar. Each MD simu-

lation was performed for 10 ns with the software GROMACS version 4 defining a NVT environment; the Particle Mesh Ewald (PME) algorithm[51] was used for the calculation of electrostatic interactions, the v-rescale algorithm64 for temperature and the Berendsen algorithm[52] for pressure were also employed.

The outcome of each MD simulation was analyzed by calculating the Root Mean Square Fluctuation (RMSF), which indicates the average movement of the protein residues during simulation. The calculation was performed on the protein Cα using the g_rmsf tool of the GROMCS 4 package.

Acknowledgements

This project (Alessandro Pellis and Livia Corici) has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 289253 (REFINE project). Valerio Ferrario is grateful to MIUR (Ministero dell'Istruzione, dell'Università e della Ricerca-Roma) and to Università degli Studi di Trieste for financial support. Lucia Gardossi acknowledges EU COST Action CM1303 System Biocatalysis for financial support. This work has been supported by the Federal Ministry of Science, Research and Economy (BMWFW), the Federal Ministry of Traffic, Innovation and Technology (bmvit), the Styrian Business Promotion Agency SFG, the Standortagentur Tirol, the Government of Lower Austria and the Business Agency Vienna through the COMET-Funding Program managed by the Austrian Research Promotion Agency FFG. We are grateful to Molecular Discovery Ltd for providing software access and to Lydia Siragusa for helpful discussions. We thank Paolo Cusaro and Riseria Cusaro S.r.l. (Binasco, Italy) for providing the rice husk and VERDER SCIENTIFIC S.r.l. (Bergamo, Italy) for milling the rice husk.

Notes and references

1 UNEP, *Valuing Plastics: The Business Case for Measuring, Managing and Disclosing Plastic Use in the Consumer Goods Industry*, United Nations Environment Programme (UNEP), Nairobi, Kenya, 2014.
2 Ellen MacArthur Foundation, *The new plastics economy: rethinking the future of plastics*, 2016.
3 UNEP, *Plastic in Cosmetics*, United Nations Environmental Programme (UNEP), Nairobi, Kenya, 2015.
4 J. H. Clark, T. J. Farmer, L. Herrero-Davila and J. Sherwooda, *Green Chem.*, 2016, **18**, 3914–3934.
5 A. Pellis, L. Corici, L. Sinigoi, N. D'Amelio, D. Fattor, V. Ferrario, C. Ebert and L. Gardossi, *Green Chem.*, 2015, **17**, 1756–1766.
6 I. Bassanini, K. Hult and S. Riva, *Beilstein J. Org. Chem.*, 2015, **11**, 1583–1595.
7 F. Binns, P. Harffey, S. M. Roberts and A. Taylor, *J. Chem. Soc., Perkin Trans. 1*, 1999, 2671–2676.
8 M. B. Ansorge-Schumacher and O. Thum, *Chem. Soc. Rev.*, 2013, **42**, 6475–6490.
9 A. A. Koutinas, A. Vlysidis, D. Pleissner, N. Kopsahelis, I. Lopez Garcia, I. K. Kookos, S. Papanikolaou, T. H. Kwanb and C. Sze Ki Lin, *Chem. Soc. Rev.*, 2014, **43**, 2587–2627.
10 A. Pellis, E. Herrero Acero, L. Gardossi, V. Ferrario and G. M. Guebitz, *Polym. Int.*, 2016, **65**, 861–871.
11 A. Pellis, E. Herrero Acero, V. Ferrario, D. Ribitsch, G. M. Guebitz and L. Gardossi, *Trends Biotechnol.*, 2016, **34**, 316–328.
12 L. Corici, A. Pellis, V. Ferrario, C. Ebert, S. Cantone and L. Gardossi, *Adv. Synth. Catal.*, 2015, **8**, 1763–1774.
13 G. Cerea, L. Gardossi, L. Sinigoi and D. Fattor, *World Patent* WO/2013110446A1, 2013.
14 L. Corici, V. Ferrario, A. Pellis, C. Ebert, S. Lotteria, S. Cantone, D. Voinovich and L. Gardossi, *RSC Adv.*, 2016, **6**, 63256–63270.
15 A. Pellis, V. Ferrario, B. Zartl, M. Brandauer, C. Gamerith, E. Herrero Acero, C. Ebert, L. Gardossi and G. M. Guebitz, *Catal. Sci. Technol.*, 2016, **6**, 3430–3442.
16 A. Pellis, A. Guarneri, M. Brandauer, E. Herrero Acero, H. Peerlings, L. Gardossi and G. M. Guebitz, *Biotechnol. J.*, 2016, **11**, 642–647.
17 P. Sieber, M. Schorderet, U. Ryser, A. Buchala, P. Kolattukudy, J. P. Métraux and C. Nawrath, *Plant Cell*, 2000, **12**, 721–737.
18 E. Stavila, R. Z. Arsyi, D. M. Petrovic and K. Loos, *Eur. Polym. J.*, 2013, **49**, 834–842.
19 D. Feder and R. A. Gross, *Biomacromolecules*, 2010, **11**, 690–697.
20 R. A. Gross, M. Ganesh and W. Lu, *Trends Biotechnol.*, 2010, **8**, 435–443.
21 V. Ferrario, L. Siragusa, C. Ebert, M. Baroni, M. Foscato, G. Cruciani and L. Gardossi, *PLoS One*, 2014, **9**, e109354.
22 M. Baroni, G. Cruciani, S. Sciabola, F. Perruccio and J. S. Mason, *J. Chem. Inf. Model.*, 2007, **47**, 279–294.
23 S. Henrich, O. M. Salo-Ahen, B. Huang, F. F. Rippmann, G. Cruciani and R. C. Wade, *J. Mol. Recognit.*, 2010, **23**, 209–219.
24 L. Siragusa, S. Cross, M. Baroni, L. Goracci and G. Cruciani, *Proteins*, 2015, **83**, 517–532.
25 P. Braiuca, L. Knapic, V. Ferrario, C. Ebert and L. Gardossi, *Adv. Synth. Catal.*, 2009, **351**, 1293–1302.
26 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.
27 D. Kold, Z. Dauter, A. K. Laustsen, A. M. Brzozowski, J. P. Turkenburg, A. D. Nielsen, H. Koldso, E. Petersen, B. Schiott, L. De Maria, K. S. Wilson, A. Svendsen and R. Wimmer, *Protein Sci.*, 2014, **23**, 1023–1035.
28 J. Uppenberg, M. T. Hansen, S. Patkar and T. A. Jones, *Structure*, 1994, **2**, 293–308.
29 P. C. Boutros and A. B. Okey, *Briefings Bioinf.*, 2005, **6**, 331–343.

30 M. H. M. Olsson, W. W. Parson and A. Warshel, *Chem. Rev.*, 2006, **106**, 1737–1756.
31 A. Warshel, P. K. Sharma, M. Kato, Y. Xiang, H. Liu and M. H. Olsson, *Chem. Rev.*, 2006, **106**, 3210–3235.
32 R. DiCosimo, J. McAuliffe, A. J. Pouloseb and G. Bohlmann, *Chem. Soc. Rev.*, 2013, **42**, 6437–6474.
33 S. Kim, C. Jimenez-Gonzalez and B. E. Dale, *Int. J. Life Cycle Assess.*, 2009, **14**, 392–400.
34 N. K. Sharma, W. S. Williams and A. Zangvil, *J. Am. Ceram. Soc.*, 1984, **67**, 715–720.
35 L. M. Contreras, H. Schelle, C. R. Sebrango and I. Pereda, *Water Sci. Technol.*, 2012, **65**, 1142–1149.
36 W. R. Stahel, *Nature*, 2016, **531**, 435–438.
37 C. Korupp, R. Weberskirch, J. J. Muller, A. Liese and L. Hilterhaus, *Org. Process Res. Dev.*, 2010, **14**, 1118–1124.
38 P. Braiuca, C. Ebert, A. Basso, P. Linda and L. Gardossi, *Trends Biotechnol.*, 2006, **24**, 419–425.
39 R. V. Ulijn, L. De Martin, P. J. Halling, A. E. M. Janssen, L. Gardossi and B. D. Moore, *Biotechnol. Bioeng.*, 2002, **80**, 509–515.
40 V. Léonard-Nevers, Z. Marton, S. Lamare, K. Hult and M. Graber, *J. Mol. Catal. B: Enzym.*, 2009, **59**, 90–95.
41 P. Spizzo, A. Basso, C. Ebert, L. Gardossi, V. Ferrario, D. Romano and F. Molinari, *Tetrahedron*, 2007, **63**, 11005–11010.
42 M. C. Franssen, P. Steunenberg, E. L. Scott, H. Zuilhofac and J. P. Sanders, *Chem. Soc. Rev.*, 2013, **42**, 6491–6533.
43 E. Herrero Acero, D. Ribitsch, G. Steinkellner, K. Gruber, K. Greimel, I. Eiteljoerg, E. Trotscha, R. Wei, W. Zimmermann, M. Zinn, A. Cavaco-Paulo, G. Freddi, H. Schwab and G. M. Guebitz, *Macromolecules*, 2011, **44**, 4632–4640.
44 *The PyMOL Molecular Graphic System, Version 1.5.0.3*, Schrodinger LLC.
45 D. J. Rogers and T. T. Tanimoto, *Science*, 1960, **132**, 1115–1118.
46 T. J. Dolinsky, J. E. Nielsen, J. A. McCammon and N. A. Baker, *Nucleic Acids Res.*, 2004, **32**, W665–W667.
47 H. Li, A. D. Robertson and J. H. Jensen, *Proteins*, 2005, **61**, 704–721.
48 G. A. Kaminski, R. A. Friesner, J. Tirado-Rives and W. L. Jorgensen, *J. Phys. Chem. B*, 2001, **105**, 6474–6487.
49 H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura and T. Head-Gordon, *J. Chem. Phys.*, 2004, **120**, 9665–9678.
50 B. Hess, C. Kutzner, D. V. D. Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.
51 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8593.
52 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Di Nola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684.