microbial biotechnology

Microbial Biotechnology (2010) **3**(6), 677–690 doi:10.1111/j.1751-7915.2010.00185.x

Mining bacterial genomes for novel arylesterase activity

Lijun Wang,[1] Valentina Mavisakalyan,[2] Elisabeth R. M. Tillier,[3,4,5] Greg W. Clark,[3] Alexei V. Savchenko,[2] Alexander F. Yakunin[2] and Emma R. Master[1]*

[1]*Department of Chemical Engineering and Applied Chemistry, University of Toronto, 200 College Street, Toronto, ON, M5S 3E5, Canada.*

[2]*Banting and Best Department of Medical Research, University of Toronto, Toronto, ON, M5G 1L6, Canada.*

[3]*Department of Medical Biophysics, University of Toronto, 101 College Street, Toronto, ON, M5G 1L7, Canada.*

[4]*OCI/PMH Campbell Family Institute for Cancer Research, Toronto, ON, Canada.*

[5]*Ministry of Health and Long-Term Planning, Toronto, ON, Canada.*

Summary

One hundred and seventy-one genes encoding potential esterases from 11 bacterial genomes were cloned and overexpressed in *Escherichia coli*; 74 of the clones produced soluble proteins. All 74 soluble proteins were purified and screened for esterase activity; 36 proteins showed carboxyl esterase activity on short-chain esters, 17 demonstrated arylesterase activity, while 38 proteins did not exhibit any activity towards the test substrates. Esterases from *Rhodopseudomonas palustris* (RpEST-1, RpEST-2 and RpEST-3), *Pseudomonas putida* (PpEST-1, PpEST-2 and PpEST-3), *Pseudomonas aeruginosa* (PaEST-1) and *Streptomyces avermitilis* (SavEST-1) were selected for detailed biochemical characterization. All of the enzymes showed optimal activity at neutral or alkaline pH, and the half-life of each enzyme at 50°C ranged from < 5 min to over 5 h. PpEST-3, RpEST-1 and RpEST-2 demonstrated the highest specific activity with *p*NP-esters; these enzymes were also among the most stable at 50°C and in the presence of detergents, polar and non-polar organic solvents, and imidazolium ionic liquids. Accordingly, these enzymes are particularly interesting targets for subsequent application trials. Finally, biochemical and bioinformatic analyses were compared to reveal sequence features that could be correlated to enzymes with arylesterase activity, facilitating subsequent searches for new esterases in microbial genome sequences.

Received 18 January, 2010; accepted 4 March, 2010. *For correspondence. E-mail emma.master@utoronto.ca; Tel. (+1) 416 946 7861; Fax (+1) 416 978 8605.

Introduction

Carboxylic ester hydrolases (3.1.1.x) comprise a valuable source of enzymes for biotechnological applications since many can catalyse both hydrolytic and synthetic reactions depending on the reaction condition. Further, these enzymes frequently exhibit regio- and stereo-specificity under mild reaction conditions, making them ideal tools for green chemistry particularly when applied for the production of chiral pure compounds, and more recently biodiesel (reviewed in Sharma *et al.*, 2001; Parawira, 2009). Most known carboxyl esterases belong to the α/β hydrolase fold superfamily of enzymes (Ollis *et al.*, 1992; Hotelier *et al.*, 2004), which use water to hydrolyse ester bonds in aqueous solutions forming an alcohol and a carboxylic acid. Their catalytic mechanism involves the coordinated activity of a conserved nucleophilic serine, typically present in a GXSXG motif (where X can be any amino acid residue), and an acidic residue (aspartate or glutamate) that hydrogen bonds with a conserved histidine residue (Jaeger *et al.*, 1999). Examples of carboxylic ester hydrolases are carboxylesterases (3.1.1.1), arylesterases (3.1.1.2) and lipases (3.1.1.3) (Chang *et al.*, 2009).

Carboxylesterases and arylesterases typically catalyse the hydrolysis of water-soluble and short- to medium-length aliphatic esters, and can be distinguished by the latter's ability to preferentially hydrolyse aromatic esters. In contrast, lipases typically display high activity towards water-insoluble long-chain esters. Interfacial activation and atypical Michaelis–Menten kinetic curves have also been used to distinguish lipases from esterases (reviewed in Jaeger *et al.*, 1994; Lotti and Alberghina, 2007). These phenomena have been correlated to the presence of a flexible hydrophobic α-helix domain (lid or flap) in many lipases, which are thought to expose the active site and direct substrate binding at the interface between water and solvent phases (Winkler *et al.*, 1990). Lipases also typically contain comparatively high numbers of non-polar residues, including Val, Leu and Ile, which have been

implicated in increased binding to aggregates of hydrophobic substrates (Kim *et al*., 1997). Finally, lipases have been distinguished from esterases by having optimal activity at comparatively alkaline pH, and by distinct surface electrostatic potential distributions at corresponding pH optima (Arpigny and Jaeger, 1999).

With increasing motivation to transition from petroleum-based products to renewable fuels and chemicals, investment in biotechnologies that convert biomass to liquid fuels has been invigorated in recent years. While most of these research activities focus on the production of fermentable sugars from the carbohydrate fraction of biomass, it is clear that value-added compounds from the lignin fraction is an important means to reducing biofuel costs (Ragauskas *et al*., 2006). Several examples of enzyme catalysed *trans*-esterification of sugars and fatty acids have been reported, and illustrate the potential of using this approach to synthesize high-value products from plant-derived polysaccharides (Ferrer *et al*., 1999; Plou *et al*., 2002; Yang and Wang, 2003). However, there are only a few examples where value-added products from lignin were synthesized using esterases. These include esterification of phenolic acid with *n*-butanol by a feruloyl esterase from *Fusarium oxysporum*, and esterification of glycerol with sinapic acid by a feruloyl esterase from *Aspergillus niger* (Topakas *et al*., 2003; Tsuchiyama *et al*., 2007; Vafiadi *et al*., 2009). Phenolic acids, including sinapic acids, are derived from lignin. By increasing the arsenal of arylesterases with activity towards phenolic acids and stability in solutions that promote *trans*-esterification activity, the development of new, high-value chemicals from lignin can be facilitated.

At the time of writing, well over 2000 bacterial genome sequences were publicly available at various levels of completion. These sequences, as well as the increasing number of metagenome sequences, constitute an enormous opportunity for enzyme discovery and application development (Lorenz and Eck, 2005). However, many of the genes identified through genome and metagenome sequencing were annotated based on sequence similarity to other proteins. Although this approach is quick and inexpensive, over 40% of sequences typically fail to be assigned a function, and many open reading frames are incorrectly annotated (Green and Karp, 2004). Alternatively, a more definitive approach to assigning a molecular function to a predicted open reading frame is to isolate and biochemically characterize the corresponding protein (Kuznetsova *et al*., 2006). Given the industrial significance of arylesterases and the difficulty of distinguishing these enzymes from carboxylesterases and lipases based on sequence comparisons alone (Fojan *et al*., 2000), the objective of this study was to search publicly available bacterial genome sequences for potential arylesterases, and then confirm the activity through biochemical characterization of the recombinantly expressed protein. A similar approach was successfully used for the characterization of novel microbial carboxylesterases and lipases (Kim *et al*., 2004; Ro *et al*., 2004). In the current study, biochemically characterized esterases were also evaluated for stability in a broad range of organic solvents and ionic liquids (ILs) that could support *trans*-esterification activity. Bioinformatic analysis of the biochemically characterized enzymes was also performed in an effort to improve sequence-based approaches to arylesterase discovery.

Results and discussion

Selection and purification of enzymes with arylesterase activity

To identify novel arylesterases, we selected over 170 genes from 11 bacterial genomes (Table S1) based on sharing 30–40% sequence identity to a biochemically characterized arylesterase, as well as presence of conserved catalytic residues and sequence motifs (Bornscheuer, 2002; Cheeseman *et al*., 2004). The genomic DNA samples were selected based on their having originated from microorganisms involved in wastewater treatment (ex. *Nitrosomonas* sp.), bioremediation of aromatic compounds (ex. *Rhodopseudomonas* sp.) and plant pathogenesis (ex. *Ralstonia solanacearum*). Common soil bacteria that participate in the decay and cycling of plant biomass (ex. species of *Streptomyces* and *Pseudomonas*) were also included in our analyses.

While lipases and esterases are difficult to distinguish at the sequence level, differences in catalytic properties can be used to differentiate these enzymes. Therefore, each candidate gene was cloned into a pET15b-based vector for heterologous expression in *Escherichia coli*. A total of 171 genes were cloned and 74 of them were expressed as soluble proteins in *E. coli* (Table S1). All 74 soluble proteins were purified to homogeneity and then screened using colorimetric assays; 36 proteins showed activity on *p*NP-esters at pH 8 while 38 proteins did not exhibit any activity towards the test substrates (Fig. 1, Table S2). Phenyl acetate is a commonly used substrate to identify bacterial arylesterases (Shaw *et al*., 1994; Fenster *et al*., 2000). Seventeen purified enzymes demonstrated arylesterase activity; none of these hydrolysed olive oil to detectable levels (Fig. 2).

Many arylesterases characterized to date were isolated from mammals, and are also known as serum paraoxonases given their ability to hydrolyse paraoxon and other organophosphorus esters (http://www.brenda-enzymes.org). While none of the purified bacterial enzymes showed paraoxonase activity in our work, 6 of the 17 arylesterases were also active on medium-chain aliphatic substrates (Table S2). These included three

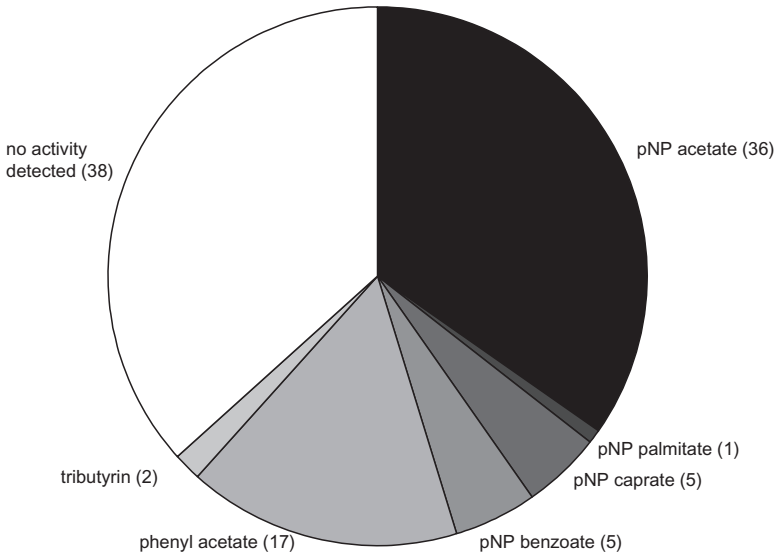

Fig. 1. The distribution of enzyme activities that were detected using a variety of esterase and lipase substrates. All assays were performed using standard conditions and were incubated for 30 min at 37°C. In total, 74 purified proteins were screened; the number shown in parentheses represents the number of proteins that demonstrated activity with each substrate, or that did not demonstrate activity with any of the test substrates.

esterases from *Rhodopseudomonas palustris* (RpEST-1, RpEST-2 and RpEST-3), two esterases from *Pseudomonas putida* (PpEST-2, PpEST-3) and an esterase from *Pseudomonas aeruginosa* (PaEST-1). As described above, our objective was to isolate enzymes with potential to modify lignin-derived phenolic compounds, and in particular, increase the lipophilicity of these compounds to extend their antioxidant properties in nutraceutical and pharmaceutical applications (Chalas *et al.*, 2001). Accordingly, enzymes with arylesterase activity that were also active on medium-length *p*NP-esters were selected for further biochemical characterization. An enzyme from *Streptomyces avermitilis* (SavEST-1) and enzyme from *P. putida* (PpEST-1) that hydrolysed *p*NP-acetate but not phenyl acetate were also included in this characterization for comparison.

pH optimum and stability of purified esterases

The 'electrostatic catapult' repulsion model for carboxylic ester hydrolases predicts the rapid release of negatively charged free fatty acid products from a negatively charged active site (Neves Petersen *et al.*, 2001). Consistent with this model, others have shown that many lipases are distinguished from carboxyl esterases by having a negative surface potential at alkaline pH. Likewise, by comparing the pH optima of the eight esterases isolated in this study to the pI values calculated from corresponding

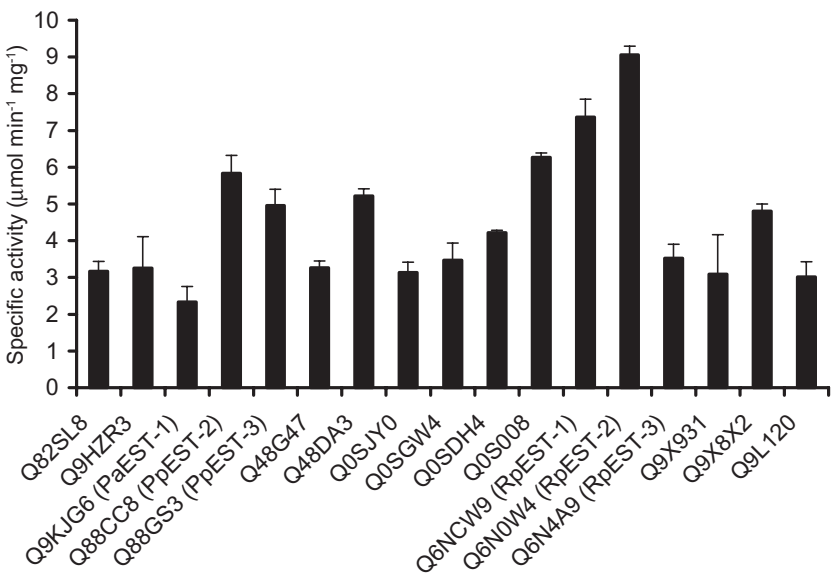

Fig. 2. Specific activity of 17 novel arylesterases against phenyl acetate. Reactions were performed using 30 mM of substrate and 1 μg 100 μl^{-1} of protein. SWISSPROT accession numbers are given. $n = 3$; error bars correspond to standard deviation from the mean.

Table 1. Biochemical properties of purified carboxylic ester hydrolases.

Property	PaEST-1	PpEST-1	PpEST-2	PpEST-3	RpEST-1	RpEST-2	RpEST-3	SavEST-1
UniProt ID	Q9KJG6	Q88QX0	Q88CC8	Q88GS3	Q6NCW9	Q6N0W4	Q6N4A9	Q82QJ4
MW[a]	34.8	26.3	30	31.1	32.4	26.2	27.2	28.5
pI[a]	6.1	5.0	4.9	5.1	5.3	5.3	5.6	5.4
pH optimum	7–9	8–9	9	8–9	7	9	9	8
pH stability[b,c]	6–8	5–11	6–8	4–10	4–9	5–10	8–10	7–10
Half-life at 50°C[b]	< 5 min	< 10 min	4 h	5 h	5 h	> 5 h[d]	< 5 min	2 h

a. Predicted from primary amino acid sequence.
b. *n* (number of replicates) = 6.
c. Residual activity was > 80%.
d. Half-life at 55°C was 4 h.
The first two letters of enzyme abbreviations correspond to bacterial genus and species names, respectively.

protein sequences, it can be predicted that these enzymes are most active when their overall charge is negative (Table 1). Notably, the pH optima of the arylesterases were neutral or alkaline, consistent with biochemical data for arylesterases listed in the BRENDA enzyme database (http://www.brenda-enzymes.info). Still, with the exception of RpEST-3 and SavEST-1, all of the esterases tested retained at least 80% activity after pre-incubation in both acidic and alkaline buffers (Table 1, Fig. S1).

Thermostability of purified esterases

RpEST-1, PpEST-2 and PpEST-3 exhibited the highest half-lives at 50°C (Table 1; Fig. S2). While RpEST-2 was fully active at 50°C after 5 h incubation, the half-life of the enzyme at 55°C was 4 h. In contrast, 10 min at 50°C reduced RpEST-3 and PaEST-1 activity by 80%, while PpEST-1 was entirely inactivated. Since the enzymes characterized in this study were obtained from mesophilic bacteria, the lack of correlation between the enzyme source and thermal stability is not surprising.

Temperature stability experiments using bacterial arylesterases are lacking. Still, a review of currently available data indicates that the thermostability of RpEST-2 is lower than that of thermophilic arylesterases from *Archaeoglobus fulgidus* and *Sulfolobus solfataricus* (Kim *et al*., 2008; Park *et al*., 2008), but higher than many carboxylesterases isolated from mesophilic bacteria (http://www.brenda-enzymes.org). In an effort to predict amino acid residues that were correlated to temperature stability of bacterial arylesterases, we looked for preferential substitutions in the thermal stable enzymes characterized in this study (relative to P22862) that were previously found to occur in other thermophilic enzymes; the abundance of amino acids that typically occur at increased frequency in thermal stable enzymes was also determined (Argos *et al*., 1979; Berezovsky *et al*., 2007; Zhou *et al*., 2008). Unfortunately, neither of these analyses led to convincing predictions that could explain the relative temperature stability of PpEST-2, PpEST-3, RpEST-1 and RpEST-2.

Effect of detergents on enzyme stability

Stability in the presence of detergents is an important characteristic of enzymes that might be used in detergent formulations as well as biotransformations where lipophilic substrates are solubilized by the addition of surfactant. Accordingly, we determined the effect of various detergents on the stability of esterases isolated in this study. Esterases from *R. palustris* showed high stability in selected detergents compared with esterases from *P. putida*, *P. aeruginosa* and *S. avermitilis*, and pre-incubation of RpEST-1 with detergents resulted in higher residual activity compared with pre-incubation in buffer alone (Table 2). Increased enzyme activity after

Table 2. Stability of carboxylic ester hydrolases in detergents.

		Residual activity (%)							
Detergents	CMC at 22°C[a]	PaEST-1	PpEST-1	PpEST-2	PpEST-3	RpEST-1	RpEST-2	RpEST-3	SavEST-1
1% v/v Tween 20	0.042 mM	24 ± 4	31 ± 8	17 ± 1	82 ± 24	131 ± 5	96 ± 14	64 ± 8	78 ± 9
1% v/v Tween 60	0.022 mM	101 ± 15	40 ± 10	89 ± 2	78 ± 26	136 ± 39	93 ± 9	88 ± 7	90 ± 5
1% v/v Tween 80	0.028 mM	32 ± 8	64 ± 9	80 ± 3	50 ± 3	135 ± 9	100 ± 4	91 ± 8	70 ± 1
1% v/v Triton X-100	0.20 mM	25 ± 7	27 ± 8	58 ± 3	110 ± 10	174 ± 8	93 ± 16	90 ± 9	67 ± 16
1 mM SDS	5.29 mM	40 ± 4	30 ± 8	58 ± 6	53 ± 6	177 ± 3	95 ± 1	73 ± 9	28 ± 7

a. From Jacquier and Desbene (1995) and Patist and colleagues (2000).
Residual activities were measured after 5 h at 37°C and compared with enzyme samples pre-incubated for the same time in 50 mM potassium phosphate buffer (pH 7). *n* = 6; errors indicate standard derivation.

pre-incubation in various detergents was recently observed for a thermostable arylesterase from *S. solfataricus* (Park *et al*., 2008). In that case, partial unfolding was thought to increase the flexibility and activity of the enzyme. Similarly, RpEST-1 is among the most thermostable protein characterized in the present study, suggesting that partial unfolding might also increase the activity of this enzyme. Alternatively, the activating effect of detergents might be associated with the partial refolding of misfolded recombinant protein. Although not activated by pre-incubation in detergent, the thermal stability of other enzymes characterized in this study could be correlated to stability in the detergents tested.

Effect of organic solvents on enzyme stability

The stability of purified esterases in organic solvents is important to assessing the potential of each enzyme to be used as synthetic catalysts in non-aqueous media. In addition, esterases that are resistant to organic solvents can be applied in industrial processes where solvents are used to increase the thermal stability of the enzyme or minimize biocontamination (Klibanov, 2001). Overall, the bacterial esterases characterized in this study were more stable in non-polar solvents (log $P > 2.5$) than in polar solvents (log $P < 0.89$) (Table 3). This result is consistent with the general agreement that water miscible polar solvents destabilize enzymes by stripping away an essential water layer at the protein surface (Gupta, 1992; Klibanov, 1997). By comparison, non-polar solvents can stabilize enzymes by preventing the loss of this water layer and restricting the conformational mobility of the protein (Klibanov, 2001). Consistent with the stabilization effect of non-polar solvents, nearly all of the enzymes tested showed higher residual activity after pre-incubation in increasing concentrations of solvents with log $P \geq 2.5$ compared with pre-incubation in buffer alone. Similar to the impacts observed using detergents, RpEST-1, RpEST-3 and PpEST-3 also showed comparatively high activity after pre-incubation in polar solvents, including DMSO, methanol and isopropanol. As discussed below, these enzymes also demonstrated the highest activity on medium-length *p*NP-esters. The correlation between solvent stability and hydrolysis of *p*NP-esters with comparatively long acyl chain length is intriguing, and might be explained by differences in the electrostatic potential of the molecular surface of these esterases (Neves Petersen *et al*., 2001).

Since enzyme catalysed *trans*-esterification reactions are typically performed in up to 40% DMSO, the standard activity assay was repeated in the presence of 0–95% DMSO. While RpEST-3 and PpEST-3 were activated by pre-incubation with 15–30% or 30–50% DMSO, respectively, the specific activity of these enzymes decreased rapidly when measured in the presence of more than 10% DMSO. In contrast, PpEST-2, RpEST-1, RpEST-2 and SavEST-1 exhibited 70%, 66%, 46%, 39% of activity in 40% DMSO compared with reactions in buffer alone. The performance of these enzymes could be improved by lyophilization or immobilization (Klibanov, 2001). However, since DMSO is used as a co-solvent to prevent the hydrolysis of water-soluble substrates, the relative stabilities and activities of arylesterases in non-polar solvents including hexane, butanol and isopropanol might be more relevant to predicting their ability to catalyse the esterification of lignin-derived aromatic compounds.

Effect of ILs on enzyme stability

Ionic liquids offer new options for retaining enzyme activity and stability in synthetic reactions performed in the absence of water (Yang and Pan, 2005). Due to their low vapour pressure and ability to be recovered and recycled, ILs are generally considered as green solvents. Enzyme activity in IL has been correlated to the viscosity of the medium and not log P values, and although enzymes are typically stabilized by ILs with increasing hydrophobicity, polar ILs have been shown to stabilize enzymes compared with organic media with similar polarity (Yang and Pan, 2005). Numerous examples of lipase-catalysed *trans*-esterification of simple sugars in ILs have been published, and emphasize the benefit of using ILs to solubilize polar substrates while retaining enzyme activity (Park and Kazlauskas, 2001). Given the polarity of lignin-derived phenolic acids, it is anticipated that ILs will increase the efficiency of *trans*-esterification reactions involving these compounds.

Accordingly, esterase stability was evaluated after pre-incubation in three typical imidazolium-based ILs and one phosphonium-based IL where each IL was prepared as a co-solvent system containing DMSO (Table 4). While none of the recombinant esterases or commercial lipases were active after pre-incubation in more than 40% Cyphos 109, increasing the proportion of phosphonium-based ILs in the co-solvent system generally increased the residual activity of each enzyme (Table 4). An exception was [BMIm]PF_6, which significantly destabilized esterases isolated from *P. putida* and *S. avermitilis* at all test concentrations. Similar to stability studies using detergents and organic solvents, RpEST-3 displayed better or equivalent stability in ILs compared with the commercial lipases PS-SD and AY-30. The stability of RpEST-2 in ILs was also comparable to the commercial lipases. It is interesting to note that among the esterases characterized in this study, RpEST-3 was the least thermal stable and RpEST-2 was the most thermal stable, suggesting that in this case, thermal stability was poorly correlated to solvent stability and stability in ILs.

Table 3. Stability of carboxylic ester hydrolases in organic solvents.

Organic solvents	log P	Concentration (v/v%)[a]	PS-SD	AY-30	PaEST-1	PpEST-1	PpEST-2	PpEST-3	RpEST-1	RpEST-2	RpEST-3	SAV_est1
DMSO	−1.22	15	105 ± 10	85 ± 5	77 ± 11	42 ± 10	78 ± 3	100 ± 14	112 ± 25	83 ± 6	138 ± 12	103 ± 17
		30	105 ± 8	69 ± 4	28 ± 7	32 ± 7	54 ± 2	111 ± 11	118 ± 18	94 ± 6	171 ± 7	103 ± 4
		50	116 ± 12	18 ± 7	25 ± 6	0 ± 3	0 ± 0	121 ± 11	16 ± 2	110 ± 3	70 ± 44	4 ± 2
Methanol	−0.76	15	86 ± 5	96 ± 4	35 ± 7	92 ± 53	75 ± 6	87 ± 18	112 ± 14	83 ± 6	72 ± 7	37 ± 13
		30	100 ± 12	104 ± 9	22 ± 4	0 ± 1	3 ± 0	70 ± 31	132 ± 17	84 ± 8	66 ± 4	1 ± 1
		50	105 ± 12	75 ± 14	20 ± 3	0 ± 2	0 ± 1	103 ± 9	139 ± 26	3 ± 0	68 ± 10	1 ± 1
1,6-Dixaone	−0.42	15	102 ± 12	57 ± 9	35 ± 8	0 ± 3	0 ± 0	97 ± 8	86 ± 2	85 ± 7	98 ± 8	6 ± 2
		30	111 ± 26	35 ± 5	33 ± 4	1 ± 2	0 ± 0	78 ± 6	12 ± 2	7 ± 2	91 ± 11	1 ± 1
		50	125 ± 26	4 ± 5	38 ± 7	5 ± 4	0 ± 0	1 ± 1	0 ± 2	5 ± 1	114 ± 5	5 ± 1
Acetonitrile	−0.34	15	87 ± 4	29 ± 46	21 ± 2	0 ± 2	47 ± 2	86 ± 29	109 ± 8	85 ± 2	72 ± 4	123 ± 8
		30	70 ± 12	98 ± 22	20 ± 2	0 ± 2	0 ± 1	110 ± 38	111 ± 7	3 ± 0	84 ± 13	2 ± 2
		50	95 ± 13	12 ± 9	23 ± 5	0 ± 21	0 ± 5	0 ± 3	0 ± 1	3 ± 1	84 ± 12	0 ± 12
Isopropanol	−0.28	15	94 ± 5	66 ± 11	18 ± 6	0 ± 3	12 ± 7	115 ± 14	94 ± 11	98 ± 5	70 ± 5	53 ± 24
		30	110 ± 16	38 ± 8	19 ± 3	0 ± 3	0 ± 1	95 ± 14	79 ± 17	6 ± 2	72 ± 10	0 ± 1
		50	116 ± 22	3 ± 5	15 ± 4	0 ± 3	0 ± 1	21 ± 22	2 ± 1	3 ± 1	88 ± 12	0 ± 1
Ethanol	−0.24	15	91 ± 3	63 ± 2	20 ± 2	0 ± 1	56 ± 5	103 ± 15	104 ± 5	87 ± 5	76 ± 5	26 ± 10
		30	104 ± 16	49 ± 2	16 ± 4	0 ± 2	1 ± 0	89 ± 12	124 ± 11	67 ± 2	80 ± 4	0 ± 1
		50	108 ± 15	14 ± 7	25 ± 7	0 ± 2	0 ± 0	34 ± 8	3 ± 1	3 ± 1	73 ± 11	1 ± 2
Acetone	−0.23	15	89 ± 5	111 ± 12	36 ± 11	6 ± 3	53 ± 10	105 ± 11	99 ± 6	84 ± 9	71 ± 4	118 ± 15
		30	91 ± 14	109 ± 32	22 ± 3	0 ± 2	7 ± 16	117 ± 19	75 ± 24	94 ± 4	66 ± 5	128 ± 18
		50	96 ± 30	72 ± 10	19 ± 4	0 ± 2	0 ± 0	86 ± 8	172 ± 12	3 ± 0	72 ± 9	4 ± 3
Tetrahydro-furan	0.49	15	89 ± 3	105 ± 12	24 ± 3	0 ± 1	0 ± 1	108 ± 21	118 ± 4	81 ± 3	79 ± 9	119 ± 16
		30	87 ± 3	115 ± 8	24 ± 4	0 ± 2	0 ± 1	89 ± 7	119 ± 8	3 ± 0	84 ± 12	6 ± 1
		50	88 ± 15	50 ± 51	32 ± 6	0 ± 1	0 ± 0	0 ± 0	15 ± 4	4 ± 1	104 ± 10	2 ± 2
1-Butanol	0.8	15	61 ± 13	17 ± 3	22 ± 3	2 ± 1	4 ± 2	111 ± 16	2 ± 0	86 ± 4	70 ± 4	1 ± 1
		30	79 ± 11	13 ± 7	23 ± 3	0 ± 1	1 ± 1	125 ± 27	1 ± 1	90 ± 4	62 ± 7	1 ± 1
		50	114 ± 27	20 ± 6	24 ± 4	2 ± 2	1 ± 0	94 ± 20	2 ± 1	88 ± 28	68 ± 12	1 ± 1
Tert-amyl alcohol	0.89	15	96 ± 8	34 ± 11	26 ± 3	0 ± 3	0 ± 0	125 ± 11	86 ± 14	94 ± 4	80 ± 4	0 ± 2
		30	86 ± 19	4 ± 5	25 ± 4	0 ± 3	0 ± 0	133 ± 23	4 ± 1	99 ± 4	69 ± 7	0 ± 1
		50	106 ± 8	0 ± 6	24 ± 4	0 ± 6	0 ± 0	164 ± 26	3 ± 0	87 ± 33	72 ± 16	0 ± 3
Toluene	2.5	15	95 ± 51	119 ± 4	105 ± 3	38 ± 5	15 ± 12	44 ± 57	101 ± 7	83 ± 6	101 ± 9	111 ± 16
		30	100 ± 55	131 ± 13	130 ± 7	94 ± 18	99 ± 5	81 ± 62	122 ± 7	83 ± 8	119 ± 6	127 ± 18
		50	97 ± 36	182 ± 20	160 ± 54	148 ± 15	115 ± 5	174 ± 12	181 ± 14	82 ± 17	132 ± 39	167 ± 31
p-Xylene	3.1	15	81 ± 9	137 ± 11	107 ± 5	176 ± 45	51 ± 17	88 ± 31	120 ± 7	73 ± 2	87 ± 35	104 ± 18
		30	64 ± 14	250 ± 90	121 ± 49	120 ± 11	94 ± 6	135 ± 9	118 ± 6	69 ± 9	127 ± 10	106 ± 33
		50	65 ± 28	323 ± 70	178 ± 54	320 ± 183	117 ± 7	194 ± 34	169 ± 14	79 ± 14	116 ± 40	135 ± 31
Cyclohexane	3.2	15	73 ± 8	154 ± 10	126 ± 17	11 ± 15	42 ± 47	63 ± 56	112 ± 5	72 ± 8	125 ± 8	116 ± 16
		30	85 ± 7	190 ± 30	142 ± 21	176 ± 74	78 ± 35	56 ± 41	140 ± 6	77 ± 8	131 ± 18	147 ± 15
		50	77 ± 25	20 013	166 ± 75	356 ± 127	126 ± 9	166 ± 38	187 ± 17	74 ± 7	162 ± 70	208 ± 16
Hexane	3.9	15	89 ± 9	93 ± 3	95 ± 20	32 ± 36	62 ± 41	43 ± 58	105 ± 10	78 ± 6	115 ± 8	132 ± 8
		30	85 ± 12	142 ± 20	138 ± 12	300 ± 101	66 ± 32	85 ± 64	116 ± 10	76 ± 9	130 ± 6	143 ± 8
		50	97 ± 17	195 ± 14	143 ± 74	285 ± 70	120 ± 14	202 ± 22	165 ± 12	79 ± 14	147 ± 55	187 ± 19

a. Organic solvents were mixed with 50 mM potassium phosphate buffer (pH 7).
Residual activities were measured after 5 h incubation at 37°C; residual activity of enzyme samples pre-incubated for the same time in 50 mM potassium phosphate buffer (pH 7) was considered 100%. PS-SD and AY-30 are commercial lipases that function in organic solvents; their solvent stabilities were measured for comparison with the carboxyl esterases isolated in the current study. $n = 6$; errors indicate standard derivation.

Table 4. Stability of carboxylic ester hydrolases in ionic liquids.

Ionic liquids	Concentration (v/v%)[a]	PS-SD	AY-30	PaEST-1	PpEST-1	PpEST-2	PpEST-3	RpEST-1	RpEST-2	RpEST-3	SavEST-1
100% DMSO	–	0 ± 0	16 ± 3	22 ± 2	0 ± 1	8 ± 1	0 ± 4	5 ± 2	4 ± 2	59 ± 8	85 ± 3
$[BMIm]BF_4$	40	3 ± 0	70 ± 11	36 ± 1	13 ± 8	9 ± 1	11 ± 5	9 ± 2	6 ± 4	67 ± 3	14 ± 1
	50	3 ± 0	72 ± 14	31 ± 6	13 ± 8	8 ± 1	13 ± 8	9 ± 3	6 ± 4	74 ± 7	12 ± 2
	60	3 ± 0	78 ± 14	41 ± 5	14 ± 7	11 ± 0	18 ± 12	10 ± 3	5 ± 3	84 ± 9	15 ± 2
	70	24 ± 16	73 ± 19	47 ± 8	17 ± 12	13 ± 1	14 ± 14	11 ± 3	20 ± 19	102 ± 11	15 ± 2
	85	60 ± 19	91 ± 15	47 ± 8	18 ± 10	16 ± 1	25 ± 17	13 ± 4	78 ± 20	184 ± 16	24 ± 4
	100	12 ± 3	105 ± 19	60 ± 3	30 ± 19	23 ± 2	39 ± 27	20 ± 6	84 ± 14	205 ± 53	37 ± 8
$[BMIm]PF_6$	40	0 ± 1	19 ± 8	20 ± 18	0 ± 6	0 ± 2	0 ± 13	1 ± 2	3 ± 4	43 ± 18	0 ± 2
	50	0 ± 1	14 ± 5	22 ± 21	0 ± 8	0 ± 7	0 ± 12	0 ± 5	4 ± 3	20 ± 12	0 ± 4
	60	23 ± 2	7 ± 14	23 ± 11	0 ± 27	0 ± 8	0 ± 39	0 ± 3	22 ± 18	41 ± 21	1 ± 5
	70	35 ± 4	16 ± 18	23 ± 16	2 ± 13	0 ± 8	0 ± 2	5 ± 4	61 ± 8	81 ± 12	0 ± 1
	85	20 ± 7	33 ± 28	20 ± 2	0 ± 28	0 ± 4	42 ± 28	41 ± 19	61 ± 11	77 ± 12	6 ± 4
	100	12 ± 5	32 ± 40	10 ± 15	0 ± 8	0 ± 4	0 ± 4	2 ± 3	10 ± 5	11 ± 18	0 ± 4
$[BMIm]CF_3SO_3$	40	3 ± 0	82 ± 14	34 ± 9	16 ± 9	13 ± 1	13 ± 53	11 ± 2	6 ± 4	84 ± 10	13 ± 1
	50	4 ± 0	93 ± 8	50 ± 7	20 ± 11	15 ± 1	18 ± 10	11 ± 2	6 ± 6	94 ± 17	14 ± 1
	60	6 ± 1	100 ± 8	54 ± 7	25 ± 14	18 ± 2	56 ± 22	13 ± 2	9 ± 2	108 ± 22	16 ± 1
	70	44 ± 9	116 ± 7	61 ± 11	27 ± 16	20 ± 2	84 ± 44	21 ± 3	35 ± 12	119 ± 13	19 ± 3
	85	40 ± 3	135 ± 14	81 ± 13	42 ± 21	37 ± 6	120 ± 44	51 ± 9	44 ± 19	176 ± 19	28 ± 4
	100	34 ± 6	145 ± 11	64 ± 19	67 ± 16	96 ± 10	121 ± 25	41 ± 10	63 ± 10	177 ± 14	27 ± 1

a. Ionic liquids were mixed with DMSO.
Residual activities were measured after 5 h incubation at 37°C and compared with enzyme samples pre-incubated for the same time in 50 mM potassium phosphate buffer (pH 7). $n = 6$; errors indicate standard derivation.

Kinetic analysis of purified esterases

Lipase hydrolysis of soluble substrates often displays atypical Michaelis–Menten kinetics exemplified by a dramatic increase in activity when the substrate approaches its solubility limit and a substrate emulsion forms (Jaeger *et al.*, 1994). In contrast, kinetic data collected in this study obeyed the Michaelis–Menten equation, further supporting the esterase classification given to the enzymes characterized in this study. Table 5 summarizes the kinetic parameters obtained for each esterase with a range of *p*NP-esters. With the exception of PpEST-3 activity on *p*NP-acetate, the catalytic efficiencies of the esterases from *R. palustris* and *P. putida* generally decreased with increasing length of the acyl chain. The kinetic data were also consistent with primary screens, which indicated that PpEST-1 and SavEST-1 were most active on *p*NP-acetate, and that PpEST-2 hydrolysed phenyl acetate as well as *p*NP-acetate but demonstrated negligible activity on other *p*NP-esters (Table S2). While the K_m of most of the esterases increased with increasing acyl chain length, the K_m of PaEST-1 decreased slightly with increasing length of the *p*NP-ester. Two esterases from *R. palustris* (RpEST-1 and RpEST-2) demonstrated the highest specific activity on all *p*NP-esters. *Rhodopseudomonas palustris* is a soil bacterium known for its ability to degrade lignin-derived aromatic compounds (Larimer *et al.*, 2004). Since RpEST-1 and RpEST-2 were also among the most stable esterases in organic solvents and ILs, respectively, these enzymes are particularly attractive candidates for *trans*-esterification of lignin-derived phenolic acids and long-chain alcohols.

Sequence analyses for improved prediction of arylesterases

In an attempt to improve future predictions of enzymes with arylesterase activity based on primary protein sequence information, a phylogenetic tree was constructed containing the 36 active bacterial enzymes purified in this study and the characterized arylesterase from *Pseudomonas fluorescens* P22862 (Fig. S3). Overall sequence comparison did not cluster the proteins according to measured enzyme activities, nor did comparison of sequences to the Lipase Engineering Database, although like P22862, sequence similarity between RpEST-1 and haloperoxidases was revealed (Fischer and

Table 5. Comparison of kinetic parameters of carboxylic ester hydrolases at their optimal pH with *p*NP-esters.

Enzyme	pH	Substrate	Specific activity[a] (μmol min^{-1} mg^{-1})	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} mM^{-1})
PaEST-1	8	*p*NP-acetate	3.2 ± 0.4	1.9 ± 0.2	1.4 ± 0.2	1.3
		*p*NP-propionate	15.1 ± 2.4	8.8 ± 1.4	2.5 ± 0.5	3.5
		*p*NP-butyrate	38.6 ± 9.6	22.4 ± 5.6	5.4 ± 1.7	4.1
		*p*NP-caproate	22.2 ± 4.1	12.9 ± 2.4	1.4 ± 0.4	9.2
		*p*NP-caprate	15.8 ± 1.8	9.2 ± 1.0	1.2 ± 0.2	7.6
		*p*NP-laurate	0.4 ± 0.1	0.2 ± 0.05	0.7 ± 0.4	0.3
PpEST-1	8	*p*NP-acetate	8.8 ± 0.5	3.9 ± 0.2	1.4 ± 0.1	2.8
		*p*NP-propionate	7.1 ± 0.8	3.1 ± 0.3	3.0 ± 0.4	1.0
PpEST-2	9	*p*NP-acetate	25 ± 1.0	12.5 ± 0.5	1.1 ± 0.1	11.4
		*p*NP-propionate	6.0 ± 0.3	3 ± 0.1	1.0 ± 0.1	3.0
PpEST-3	8	*p*NP-acetate	11.9 ± 0.6	6.2 ± 0.3	1.0 ± 0.1	6.2
		*p*NP-propionate	13.1 ± 0.2	6.8 ± 0.1	0.1 ± 0.0	67.9
		*p*NP-caproate	23.3 ± 0.5	12.1 ± 0.3	0.2 ± 0.0	60.4
		*p*NP-caprate	16.8 ± 0.7	8.7 ± 0.4	0.8 ± 0.1	10.9
		*p*NP-laurate	9.8 ± 1.4	5.1 ± 0.7	2.5 ± 0.4	2.0
RpEST-1	7	*p*NP-acetate	121.5 ± 2.9	65.6 ± 1.5	0.3 ± 0.0	219.0
		*p*NP-propionate	568.2 ± 28.4	306.8 ± 15.3	1.5 ± 0.1	205.0
		*p*NP-butyrate	139.8 ± 2.7	75.5 ± 1.6	0.5 ± 0.0	151.0
		*p*NP-caproate	401.9 ± 21.8	217.0 ± 11.8	1.9 ± 0.1	114.0
RpEST-2	9	*p*NP-acetate	120.6 ± 9.9	52.7 ± 4.3	1.4 ± 0.2	37.6
		*p*NP-propionate	55.9 ± 17.9	24.4 ± 7.8	4.9 ± 1.8	4.9
		*p*NP-butyrate	76.2 ± 29.2	33.2 ± 12.8	8.8 ± 3.6	3.8
		*p*NP-caproate	103.0 ± 42.9	45.0 ± 18	8.3 ± 3.8	5.4
		*p*NP-caprate	16.7 ± 7.9	7.3 ± 3.4	8.4 ± 4.3	0.9
RpEST-3	9	*p*NP-acetate	ND			
		*p*NP-propionate	ND			
		*p*NP-butyrate	15.8 ± 1.6	7.2 ± 0.7	3.6 ± 0.4	2.0
		*p*NP-caproate	11.1 ± 1.0	5.0 ± 0.4	2.1 ± 0.3	2.4
SavEST-1	8	*p*NP-acetate	6 ± 0.2	2.9 ± 0.1	0.3 ± 0.0	9.5

a. Calculated from V_{max}.

Reactions were performed at 37°C. ND, initial reaction rates were not detectable. All kinetic data obeyed the Michaelis–Menten equation with $R^2 \geq 0.96$. $n = 3$; errors indicate standard deviation.

Pleiss, 2003; Cheeseman *et al*., 2004). Further analyses were performed to identify sequence motifs that are correlated to substrate preference but might span a short region of the overall protein sequence. These analyses were performed using: (i) covariation analysis tools that identify the functional association of residues (Tillier and Lui, 2003), (ii) Pratt (Jonassen *et al*., 1995), a tool for discovering conserved motifs from unaligned sequences, and (iii) Sequence Harmony (Feenstra *et al*., 2007), a tool designed for finding amino acid residues responsible for enzyme specificity. Although these primary sequence analyses did not reveal amino acid positions whose substitution patterns correlated with enzyme specificity, sequence comparisons that referenced the solved structure of P22862 revealed conservation of residues predicted to participate in substrate binding and stabilization of reaction intermediates. The acyl-binding pocket of P22862 comprises Trp-28, Val-121 and Phe-198, and the alcohol-binding pocket comprises Phe-93 and Phe-162; Met-95 and Trp-28 in P22862 are predicted to form an oxyanion hole that participates in the stabilization of reaction intermediates (Cheeseman *et al*., 2004). While all of the active enzymes characterized in this study contained the conserved catalytic triad and GXSXG motif, a sequence logo clearly revealed the conservation of Met95 in enzymes with arylesterase activity that like P22862, also hydrolysed short-chain esters (Fig. 3). PpEST-1, PpEST-2, PpEST-3 and RpEST-3 also contained Phe or an alternate aromatic amino acid at positions corresponding to the acyl-binding pocket in P22862. Notably, enzymes isolated in this study that lacked arylesterase activity contained functionally similar amino acids at positions corresponding to the acyl-binding pocket in P22862, but generally lacked aromatic amino acids at positions corresponding to the alcohol-binding pocket of that enzyme.

Conclusions

Owing to the synthetic potential and limited commercial availability of enzymes with arylesterase activity, the enzymes characterized in this study constitute an important set of new, industrially relevant biocatalysts. To our best knowledge, RpEST-1, RpEST-2, RpEST-3, PpEST-2, PpEST-3 and PaEST-1 represent the first enzymes with arylesterase activity to be isolated and characterized from *R. palustris*, *P. putida* and *P. aeruginosa* respectively. RpEST-1, RpEST-2 and PpEST-3 demonstrated the best combination of activity, thermal stability and solvent stability, suggesting that they are the best bacterial candidates for subsequent *trans*-esterification trials. Presently, the ability of these enzymes to catalyse hydrolysis reactions in the presence of organic solvents, and esterification of lignin-derived compounds, is being evaluated. Finally, the bioinformatic analyses performed in our study emphasize the limits of primary sequences alone to accurately predict genes that encode arylesterase activity, emphasizing the importance of structural and biochemical studies in the discovery of enzymes for new biotechnological applications.

Experimental procedures

Materials

All bacterial genomic DNA samples were purchased from the ATCC. Paraoxon, phenyl acetate, olive oil, tributyrin and *p*NP-ester substrates (*p*NP-acetate, *p*NP-propanioate, *p*NP-butyrate, *p*NP-caproate, *p*NP-caprate, *p*NP-laurate, *p*NP-palmitate, *p*NP-benzoate), Tween 20, Tween 60, Tween 80, Triton X-100, SDS, EDTA and 2-mercaptoethanol, along with solvents, including methanol, ethanol, isopropanol, 1-butanol, *tert*-amyl alcohol, acetone, acetonitrile, tetrahydrofuran, 1,6-dixaone, DMSO, toluene, *p*-xylene, hexane, cyclohexane, were purchased from Sigma (Canada). Ionic liquids, including [BMIm]BF_4, [BMIm]PF_6, [BMIm]CF_3SO_3, were purchased from Solvent Innovation GmbH (Germany) and Cyphos 109 was kindly provided by Cytec Industries (Canada). PS 'Amano' SD and AY 'Amano' 30 were kindly supplied by AMANO Enzyme. The *p*NP-ester stock solutions (10 mM) were prepared in DMSO.

Selection of potential esterases from microbial genomes

The biochemically and structurally characterized arylesterase PFE from *P. fluorescens* (P22862) was used as a query to search the SWISSPROT sequences for uncharacterized homologous proteins. The BLASTP algorithm was used to compare the P22862 sequence to the following 11 bacterial genomes: *Agrobacterium tumefaciens* C58, *Nitrosomonas europaea*, *P. aeruginosa*, *P. putida* KT2440, *Pseudomonas syringae* pv. *phaseolicola* (strain 1448A/Race 6), *Rhodococcus* sp. RHA1, *R. palustris* CGA009, *Ralstonia solanacearum* 11696D-S, *S. avermitilis*, *Streptomyces coelicolor* and *Thermotoga maritima*. Sequences identified in each genome were aligned to P22862 to confirm the presence of conserved catalytic residues and sequence motifs (Fojan *et al*., 2000).

Gene cloning and protein purification

Genes were PCR amplified using the genomic DNA purchased from the ATCC. PCR was performed using the Pfx DNA polymerase (Invitrogen) and the following PCR cycles: denaturation at 95°C for 15 s, annealing at 53°C for 30 s and elongation at 68°C for 1 min. PCR products were flanked by BseRI restriction sites to facilitate cloning into p15Tv-L (GenBank accession EF456736). Ligation products were transformed into *E. coli* DH5α by electroporation. *Escherichia coli* transformants were cultured at 37°C in autoinduction media (Kuznetsova *et al*., 2006). Autoinduction medium (ZYP 5052) contained 1 mM $MgSO_4$, metal mix (50 μM $FeCl_3 \bullet 6H_2O$, 20 μM $CaCl_2$, 10 μM

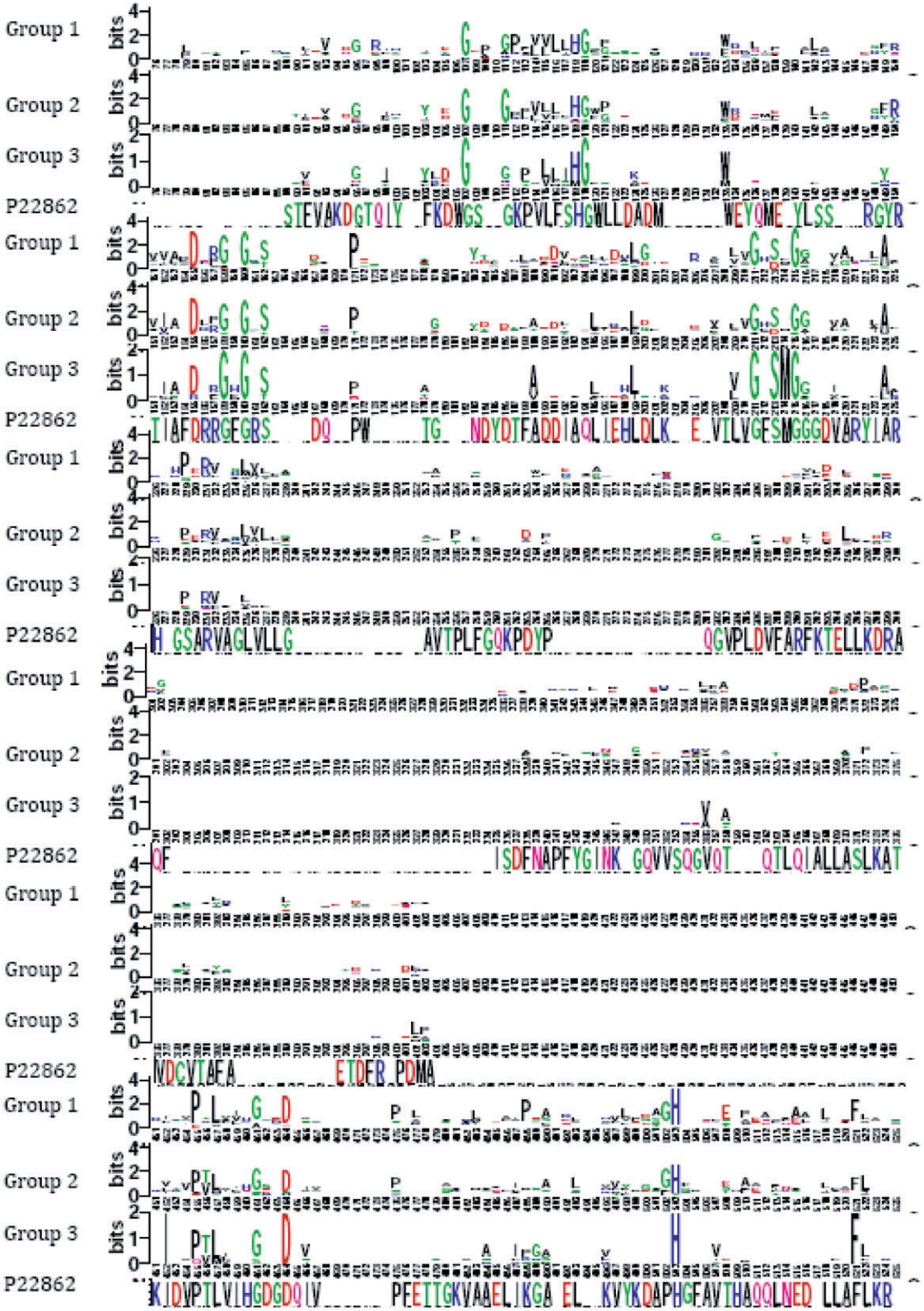

Fig. 3. Sequence logo of active esterases. The MAFFT alignment of the 36 active enzymes isolated in this study was separated into three groups based on their specificity (see Table S2) and their sequence logos are shown aligned along with the sequence for P22862 for which a crystal structure is available. The following list indicates in parentheses, the position of amino acids in the gapped alignment of P22862 that are discussed in the text: W28 (W120), F93 (F212), M95 (M214), V121 (V254), F162 (F342), F198 (F382).

$ZnSO_4$•$7H_2O$, 10 μM $MnCl_2$•$4H_2O$, 2 μM $CoCl_2$•$6H_2O$, 2 μM $CuCl_2$•$2H_2O$, 2 μM $NiCl_2$•$6H_2O$, 2 μM Na_2MoO_4•$5H_2O$, 2 μM Na_2SeO_3•$5H_2O$ and 2 μM H_3BO_3), 20× NPS [0.5 M $(NH_4)_2SO_4$, 1 M KH_2PO_4, 1 M Na_2HPO_4], GGL (250 g l^{-1} glycerol, 25 g l^{-1} glucose, 100 g l^{-1} lactose), 10× TY (100 g l^{-1} tryptone, 50 g l^{-1} yeast extract) and 100 μg ml^{-1} of ampicillin. Protein expression was induced by switching the temperature to 20°C; cells were harvested after overnight incubation, suspended in binding buffer (300 mM NaCl, 50 mM Hepes, pH 7.5, 5% Glycerol, 5 mM imidazole), passed through one freeze–thaw cycle, and then lysed by sonication. Cell extracts were collected by centrifugation and incubated with Ni resin (Qiagen) for 2 h. The resin was then washed with 200 ml of washing buffer (300 mM NaCl, 50 mM Hepes, pH 7.5, 5% Glycerol, 30 mM imidazole) and eluted with approximately 10 ml of elution buffer (300 mM NaCl, 50 mM Hepes, pH 7.5, 5% Glycerol, 250 mM imidazole). Protein concentrations were measured using the Bradford assay and their purities were evaluated by 15% SDS-PAGE. Purified proteins were dialysed using pre-crystallization buffer [300 mM NaCl, 10 mM Hepes (pH 7.5)], flash frozen in liquid nitrogen, and then stored at −80°C.

Primary enzyme screens

The standard assay used to screen purified enzymes was performed in a 96-well plate format at 37°C; final reaction volumes were 200 μl and contained 1 μg of purified protein unless otherwise specified. For *p*NP-esters, reactions were prepared by mixing 1 vol. of 10 mM substrate with 9 vols of 100 mM potassium phosphate buffer (pH 7) containing 2.5% Triton X-100, followed by immediate vortexing. Liberation of *p*NP was measured at 400 nm, and 0.002–0.1 μmol of *p*NP was used to generate a standard curve. All purified proteins were initially screened at pH 6, 7 and 8 on *p*NP-acetate, *p*NP-caprate, *p*NP-palmitate and *p*NP-benzoate.

The release of acidic products from phenyl acetate, olive oil and tributyrin was detected using bromothymol blue (BTB) as described by Martínez-Martínez and colleagues (2007), with minor modifications. In brief, 30 mM of each substrate was prepared in 5 mM potassium phosphate buffer (pH 7.3) containing 0.01% of BTB; reactions were followed by measuring absorbance at 616 nm. A standard curve using acetic acid was prepared to calculate enzyme activity on phenyl acetate.

Optimal pH and pH stability

The effect of pH was investigated using the Britton and Robinson's universal buffer system (50 mM phosphoric acid, acetic acid and boric acid; pH adjusted from 4 to 11 using NaOH). The pH stability of each enzyme was determined by measuring the activity that remained after 24 h of incubation at 37°C and pH 4–11. The concentrations of protein in each pre-incubation were 5 μg ml^{-1} RpEST-1, 25 μg ml^{-1} RPEst-2, 50 μg ml^{-1} RpEST-3, 50 μg ml^{-1} PpEST-1, 50 μg ml^{-1} PpEst-2, 50 μg ml^{-1} PpEST-3, 50 μg ml^{-1} PaEST-1 and 100 μg ml^{-1} SavEST-1, which allowed initial reaction rates to be measured when diluting the pre-incubated enzyme sample 1:10 in the reaction buffer containing 0.5 mM *p*NP-acetate. Residual enzyme activities were measured at the pH optimum of each enzyme.

Temperature stability

Residual enzyme activity was measured after incubation at 22°C, 30°C, 37°C, 50°C, 55°C and 70°C for up to 5 h. Enzyme samples were incubated in 50 mM potassium phosphate buffer (pH 7) at concentrations indicated above. After the pre-incubation period, the enzyme sample was diluted 1:10 in the reaction buffer containing 0.5 mM *p*NP-acetate.

Effect of detergents on enzyme stability

The effect of several detergents on enzyme stability was analysed by incubating each enzyme in 50 mM potassium phosphate buffer (pH 7), containing 1% (v/v) Tween 20, Tween 60, Tween 80 or Triton X-100, or 1 mM SDS, EDTA or 2-mercaptoethanol. The concentration of protein in each pre-incubation was as indicated above. Following 5 h of pre-incubation at 37°C, the enzyme sample was diluted 1:10 in the reaction buffer containing 0.5 mM *p*NP-acetate. Residual enzyme activities were measured at the pH optimum of each enzyme. The small amount of detergent that was transferred with the enzyme into the reaction mixture had a negligible effect on absorbency.

Effect of organic solvents on enzyme stability

Residual enzyme activity was measured using the standard assay following 5 h of pre-incubation at 37°C in the following solvents: methanol, ethanol, isopropanol, 1-butanol, *tert*-amyl alcohol, acetone, acetonitrile, tetrahydrofuran, 1,6-dixaone, DMSO, toluene, *p*-xylene, hexane, cyclohexane. Organic solvents were prepared in 15%, 30% or 50% (v/v) in 50 mM potassium phosphate buffer (pH 7). The concentration of protein in each pre-incubation was as indicated above. Following the 5 h pre-incubation, the enzyme sample was diluted 1:10 in the reaction buffer containing 0.5 mM *p*NP-acetate to measure residual activity. The effect of DMSO on enzyme activity was studied in more detail. Here, each enzyme was incubated in potassium phosphate buffer containing 0–40% of DMSO with 5% increments. Residual enzyme activities were measured after 24 h of pre-incubation at 37°C. Enzyme activities were also measured in the presence of 5–95% DMSO. In both cases, *p*NP-acetate was used as the substrate.

Effect of ILs on enzyme stability

Three imidazolium-based ILs ([BMIm]BF_4, [BMIm]PF_6, [BMIm]CF_3SO_3), and one phosphonium-based IL (Cyphos 109) were used to study the effect of ILs on enzyme stability. The imidazolium ILs were selected since they are most commonly used in lipase-catalysed synthetic reactions (van Rantwijk *et al.*, 2003). The concentration of protein in each pre-incubation was as indicated above. Following 3 h of pre-incubation at 37°C in 40%, 50%, 60%, 70%, 85% or 100% of each solvent mixed with DMSO, the enzyme sample was

diluted 1:10 in the reaction buffer containing 0.5 mM *p*NP-acetate to measure residual activity.

Enzyme kinetics

Kinetic parameters were obtained at 37°C and using the optimal pH and *p*NP substrates. Substrate concentrations ranged from 0.05 mM to 1 mM with 50 μM increments (i.e. 20 different substrate concentrations). The amount of enzyme used in each 100 μl of reaction was 0.05 μg for RpEST-1, 0.1 μg for RpEST-2, 0.5 μg for RpEST-3, 0.5 μg for PpEST-1, 0.5 μg for PpEST-2, 0.25 μg for PpEST-3, 0.5 μg for PaEST-1 and 0.5 μg for SavEST-1. Initial rates were obtained by measuring reaction products every minute for 10 min. Kinetic parameters were calculated using the Michaelis–Menten equation (GraphPad Prism5 Software).

Sequence analysis

The phylogenetic tree was obtained using proml [PHYLIP (Phylogeny Inference Package) version 3.7] (Felsenstein, 1989) and the following parameters: PMB model, gamma parameter 0.7 and three HMM categories. The 37 protein sequences were aligned using Mafft version 6 (E-INS-i) (Katoh *et al*., 2005). The sequence logos of sub-alignments of three enzyme groups were obtained using WebLogo (Schneider and Stephens, 1990; Crooks *et al*., 2004).

Acknowledgements

We thank all members of the Structural Proteomics in Toronto Centre (SPiT) for help in gene cloning and protein purification. This work was supported by grants from the Natural Sciences and Engineering Research Council to E.R.M., A.F.Y., A.V.S. and E.R.M.T., from Genome Canada and the Ontario Genomics Institute to A.F.Y., and from NIH to A.V.S. G.W.C. is partially supported by the C.I.H.R. training programme in Protein Folding and E.R.M.T. is a recipient of the C.R.C. in analytical genomics.

References

Argos, P., Rossman, M.G., Grau, U.M., Zuber, H., Frank, G., and Tratschin, J.D. (1979) Thermal stability and protein structure. *Biochemistry* **18:** 5698–5703.

Arpigny, J.L., and Jaeger, K.E. (1999) Bacterial lipolytic enzymes: classification and properties. *Biochem J* **343** (Part 1)**:** 177–183.

Berezovsky, I.N., Zeldovich, K.B., and Shakhnovich, E.I. (2007) Positive and negative design in stability and thermal adaptation of natural proteins. *PLoS Comput Biol* **3:** e52.

Bornscheuer, U.T. (2002) Microbial carboxyl esterases: classification, properties and application in biocatalysis. *FEMS Microbiol Rev* **26:** 73–81.

Chalas, J., Claise, C., Edeas, M., Messaoudi, C., Vergnes, L., Abella, A., and Lindenbaum, A. (2001) Effect of ethyl esterification of phenolic acids on low-density lipoprotein oxidation. *Biomed Pharmacother* **55:** 54–60.

Chang, A., Scheer, M., Grote, A., Schomburg, I., and Schomburg, D. (2009) BRENDA, AMENDA and FRENDA the enzyme information system: new content and tools in 2009. *Nucleic Acids Res* **37:** D588–D592.

Cheeseman, J.D., Tocilj, A., Park, S., Schrag, J.D., and Kazlauskas, R.J. (2004) Structure of an aryl esterase from *Pseudomonas fluorescens*. *Acta Crystallogr D Biol Crystallogr* **60:** 1237–1243.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res* **14:** 1188–1190.

Feenstra, K.A., Pirovano, W., Krab, K., and Heringa, J. (2007) Sequence harmony: detecting functional specificity from alignments. *Nucleic Acids Res* **35:** W495–W498.

Felsenstein, J. (1989) PHYLIP – phylogeny inference package. *Cladistics* **5:** 164–166.

Fenster, K.M., Parkin, K.L., and Steele, J.L. (2000) Characterization of an arylesterase from *Lactobacillus helveticus* CNRZ32. *J Appl Microbiol* **88:** 572–583.

Ferrer, M., Cruces, M.A., Bernabe, M., Ballesteros, A., and Plou, F.J. (1999) Lipase-catalyzed regioselective acylation of sucrose in two-solvent mixtures. *Biotechnol Bioeng* **65:** 10–16.

Fischer, M., and Pleiss, J. (2003) The Lipase Engineering Database: a navigation and analysis tool for protein families. *Nucleic Acids Res* **31:** 319–321.

Fojan, P., Jonson, P.H., Petersen, M.T., and Petersen, S.B. (2000) What distinguishes an esterase from a lipase: a novel structural approach. *Biochimie* **82:** 1033–1041.

Green, M.L., and Karp, P.D. (2004) A Bayesian method for identifying missing enzymes in predicted metabolic pathway databases. *BMC Bioinformatics* **5:** 76.

Gupta, M.N. (1992) Enzyme function in organic solvents. *Eur J Biochem* **203:** 25–32.

Hotelier, T., Renault, L., Cousin, X., Negre, V., Marchot, P., and Chatonnet, A. (2004) ESTHER, the database of the alpha/beta-hydrolase fold superfamily of proteins. *Nucleic Acids Res* **32:** D145–D147.

Jacquier, J.C., and Desbene, P.L. (1995) Determination of critical micelle concentration by capillary electrophoresis. Theoretical approach and validation. *J Chromatogr* **718:** 167–175.

Jaeger, K.E., Ransac, S., Dijkstra, B.W., Colson, C., van Heuvel, M., and Misset, O. (1994) Bacterial lipases. *FEMS Microbiol Rev* **15:** 29–63.

Jaeger, K.E., Dijkstra, B.W., and Reetz, M.T. (1999) Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications of lipases. *Annu Rev Microbiol* **53:** 315–351.

Jonassen, I., Collins, J.F., and Higgins, D. (1995) Finding flexible patterns in unaligned protein sequences. *Protein Sci* **4:** 1587–1595.

Katoh, K., Kuma, K., Toh, H., and Miyata, T. (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* **33:** 511–518.

Kim, K.K., Song, H.K., Shin, D.H., Hwang, K.Y., and Suh, S.W. (1997) The crystal structure of a triacylglycerol lipase from *Pseudomonas cepacia* reveals a highly open conformation in the absence of a bound inhibitor. *Structure* **5:** 173–185.

Kim, H.K., Jung, Y.J., Choi, W.C., Ryu, H.S., Oh, T.K., and Lee, J.K. (2004) Sequence-based approach to finding functional lipases from microbial genome databases. *FEMS Microbiol Lett* **235:** 349–355.

Kim, S.B., Lee, W., and Ryu, Y.W. (2008) Cloning and characterization of thermostable esterase from *Archaeoglobus fulgidus*. *J Microbiol* **46:** 100–107.

Klibanov, A.M. (1997) Why are enzymes less active in organic solvents than in water? *Trends Biotechnol* **15:** 97–101.

Klibanov, A.M. (2001) Improving enzymes by using them in organic solvents. *Nature* **409:** 241–246.

Kuznetsova, E., Proudfoot, M., Gonzalez, C.F., Brown, G., Omelchenko, M.V., Borozan, I., *et al*. (2006) Genome-wide analysis of substrate specificities of the *Escherichia coli* haloacid dehalogenase-like phosphatase family. *J Biol Chem* **281:** 36149–36161.

Larimer, F.W., Chain, P., Hauser, L., Lamerdin, J., Malfatti, S., Do, L., *et al*. (2004) Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris*. *Nat Biotechnol* **22:** 55–61.

Lorenz, P., and Eck, J. (2005) Metagenomics and industrial applications. *Nat Rev Microbiol* **3:** 510–516.

Lotti, M., and Alberghina, L. (2007) Lipases. In *Industrial Enzymes*. Polaina, J., and MacCabe, A. (eds). Dordrecht, the Netherlands: Springer, pp. 263–281.

Martínez-Martínez, I., Montoro-García, S., Lozada-Ramírez, J.D., Sánchez-Ferrer, A., and García-Carmona, F. (2007) A colorimetric assay for the determination of acetyl xylan esterase or cephalosporin C acetyl esterase activities using 7-amino cephalosporanic acid,cephalosporin C, or acetylated xylan as substrate. *Anal Biochem* **369:** 210–217.

Neves Petersen, M.T., Fojan, P., and Petersen, S.B. (2001) How do lipases and esterases work: the electrostatic contribution. *J Biotechnol* **85:** 115–147.

Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., *et al*. (1992) The alpha/beta hydrolase fold. *Protein Eng* **5:** 197–211.

Parawira, W. (2009) Biotechnological production of biodiesel fuel using biocatalysed transesterification: a review. *Crit Rev Biotechnol* **29:** 82–93.

Park, S., and Kazlauskas, R.J. (2001) Improved preparation and use of room-temperature ionic liquids in lipase-catalyzed enantio- and regioselective acylations. *J Org Chem* **66:** 8395–8401.

Park, Y.J., Yoon, S.J., and Lee, H.B. (2008) A novel thermostable arylesterase from the archaeon *Sulfolobus solfataricus* P1: purification, characterization, and expression. *J Bacteriol* **190:** 8086–8095.

Patist, A., Bhagwat, S.S., Penfield, K.W., Aikens, P., and Shah, D.O. (2000) On the measurement of critical micelle concentrations of pure and technical-grade non-ionic surfactants. *J Surfact Deterg* **3:** 53–58.

Plou, F.J., Cruces, M.A., Ferrer, M., Fuentes, G., Pastor, E., Bernabe, M., *et al*. (2002) Enzymatic acylation of di- and trisaccharides with fatty acids: choosing the appropriate enzyme, support and solvent. *J Biotechnol* **96:** 55–66.

Ragauskas, A.J., Williams, C.K., Davison, B.H., Britovsek, G., Cairney, J., Eckert, C.A., *et al*. (2006) The path forward for biofuels and biomaterials. *Science* **311:** 484–489.

van Rantwijk, F., Madeira Lau, R., and Sheldon, R.A. (2003) Biocatalytic transformations in ionic liquids. *Trends Biotechnol* **21:** 131–138.

Ro, H.S., Hong, H.P., Kho, B.H., Kim, S., and Chung, B.H. (2004) Genome-wide cloning and characterization of microbial esterases. *FEMS Microbiol Lett* **233:** 97–105.

Schneider, T.D., and Stephens, R.M. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* **18:** 6097–6100.

Sharma, R., Chisti, Y., and Banerjee, U.C. (2001) Production, purification, characterization, and applications of lipases. *Biotechnol Adv* **19:** 627–662.

Shaw, J.F., Chang, R.C., Chuang, K.H., Yen, Y.T., Wang, Y.J., and Wang, F.G. (1994) Nucleotide sequence of a novel arylesterase gene from *Vibro mimicus* and characterization of the enzyme expressed in *Escherichia coli*. *Biochem J* **298** (Part 3)**:** 675–680.

Tillier, E.R.M., and Lui, T.W.H. (2003) Using multiple interdependency to separate functional from phylogenetic correlations in protein alignments. *Bioinformatics* **19:** 750–755.

Topakas, E., Stamatis, H., Biely, P., Kekos, D., Macris, B.J., and Christakopoulos, P. (2003) Purification and characterization of a feruloyl esterase from *Fusarium oxysporum* catalyzing esterification of phenolic acids in ternary water–organic solvent mixtures. *J Biotechnol* **102:** 33–44.

Tsuchiyama, M., Sakamoto, T., Tanimori, S., Murata, S., and Kawasaki, H. (2007) Enzymatic synthesis of hydroxycinnamic acid glycerol esters using type A feruloyl esterase from *Aspergillus niger*. *Biosci Biotechnol Biochem* **71:** 2606–2609.

Vafiadi, C., Topakas, E., Nahmias, V.R., Faulds, C.B., and Christakopoulos, P. (2009) Feruloyl esterase-catalysed synthesis of glycerol sinapate using ionic liquids mixtures. *J Biotechnol* **139:** 124–129.

Winkler, F.K., D'Arcy, A., and Hunziker, W. (1990) Structure of human pancreatic lipase. *Nature* **343:** 771–774.

Yang, Z., and Pan, P. (2005) Ionic liquids: green solvents for nonaqueous biocatalysis. *Enzyme Microb Technol* **37:** 19–28.

Yang, K., and Wang, Y. (2003) Lipase-catalyzed cellulose acetylation in aqueous and organic media. *Biotechnol Prog* **19:** 1664–1671.

Zhou, X.X., Wang, Y.B., Pan, Y.J., and Li, W.F. (2008) Differences in amino acids composition and coupling patterns between mesophilic and thermophilic proteins. *Amino Acids* **34:** 25–33.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. pH stabilities of purified esterases. (A) *Pseudomonas putida* [PpEST-1 (◆), PpEST-2 (■), PpEST-3 (▲)], (B) *Rhodopseudomonas palustris* [RpEST-1 (◆), RpEST-2 (■), RpEST-3 (▲)], and (C) *Pseudomonas aeruginosa* [PaEST-1(◆)] and *Streptomyces avermitilis* [SavEST-1 (■)]. Enzyme activity was measured after 24 h of incubation at 37°C using 0.5 mM *p*NP-acetate. $n = 6$; error bars correspond to standard deviation from the mean.

Fig. S2. Temperature stability of purified esterases. (A) RpEST-1, (B) RpEST-2, (C) RpEST-2, (D) PpEST-1, (E) PpEST-2, (F) PpEST-3, (G) PaEST-1, (H) SavEST-1. Enzyme activity was measured after 1 h (◆), 2 h (■), 3 h (▲), 4 h (X) and 5 h (●) of incubation at the specified temperature using 0.5 mM *p*NP-acetate. $n = 6$; error bars correspond to standard deviation from the mean.

Fig. S3. Phylogenetic tree of the 36 active enzymes isolated in this study. The number of Bootstrap samples with over 50% support of 100 samples are shown. The leaf labels are the SWISSPROT accession for the sequence, then the specificity group number, and the species code, separated by dashes. In red are the least specific enzymes (group 3), in blue are enzymes active on phenyl acetate and *p*NP-acetate (group 2), and in black are enzymes only active on *p*NP-acetate.

Table. S1. Summary of gene targets and expression trials.

Table. S2. Activity screens of carboxylic ester hydrolases purified in this study.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

