Directed evolution of a *para*-nitrobenzyl esterase for aqueous-organic solvents

Jeffrey C. Moore and Frances H. Arnold*

Division of Chemistry and Chemical Engineering 210-41, California Institute of Technology, Pasadena, CA 91125.
**Corresponding author (e-mail: frances@cheme.caltech.edu).*

Received 13 December 1995; accepted 24 January 1996.

Through sequential generations of random mutagenesis and screening, we have directed the evolution of an esterase for deprotection of an antibiotic *p*-nitrobenzyl ester in aqueous-organic solvents. Because rapid screening directly on the desired antibiotic (loracarbef) nucleus *p*-nitrobenzyl ester was not feasible, the *p*-nitrophenyl ester was employed. Catalytic performance on the screening substrate was shown to reasonably mimic enzyme activity toward the desired ester. One *p*-nitrobenzyl esterase variant performs as well in 30% dimethylformamide as the wildtype enzyme in water, reflecting a 16-fold increase in esterase activity. Random pairwise gene recombination of two positive variants led to a further two-fold improvement in activity. Considering also the increased expression level achieved during these experiments, the net result of four sequential generations of random mutagenesis and the one recombination step is a 50–60-fold increase in total activity. Although the contributions of individual effective amino acid substitutions to enhanced activity are small (<2-fold increases), the accumulation of multiple mutations by directed evolution allows significant improvement of the biocatalyst for reactions on substrates and under conditions not already optimized in nature. The positions of the effective amino acid substitutions have been identified in a pNB esterase structural model developed based on its homology to acetylcholinesterase and triacylglycerol lipase. None appear to interact directly with the antibiotic substrate, further underscoring the difficulty of predicting their effects in a 'rational' design effort.

Keywords: random mutagenesis, antibiotic synthesis, enzymatic deprotection

We have used sequential cycles of random mutagenesis and screening to direct the evolution of enzymes to adapt to new demands and to exhibit features never required in nature. In an earlier study, a subtilisin was 'evolved' to be hundreds of times more active than its wildtype ancestor in high concentrations of dimethylformamide (DMF)[1,2]. Here we describe the directed evolution of an enzyme for a reaction never required in nature, the efficient hydrolysis of an antibiotic *para*-nitrobenzyl ester intermediate, also in aqueous-organic solvent mixtures.

Para-nitrobenzyl alcohol (pNB-OH) is used to protect carboxylic acid functionalities during the synthesis of cephalosporin-derived antibiotics[3,4]. Deprotection usually involves catalytic zinc in organic solvents. On an industrial scale this process generates large amounts of solvent and zinc-containing waste material. Interested in pursuing more cost-effective alternatives, scientists began searching for an esterase for the deprotection reaction[3]. Screening whole cell preparations of numerous bacterial and fungal cultures resulted in the identification of a *Bacillus subtilis* '*para*-nitrobenzyl esterase' moderately effective for the hydrolysis of pNB-protected cephalosporins[3]. Although the reaction yield was high, the enzyme preparation could not compete with the speed, economy, and small reaction volumes of the zinc-catalyzed deprotection, and interest in enzymatic deprotection waned. In recent years increasing concern over environmental issues has forced industry to reevaluate the types and amounts of wastes produced. Thus, the *B. subtilis* gene coding for pNB esterase was cloned and overexpressed in *Escherichia coli* to produce large quantities of enzyme for use in antibiotics synthesis[5].

The targeted cephalosporin-derived antibiotics evolved during this time from first-generation cephalexin (one of the two original cephalosporin substrates used to screen for pNB esterase), second-generation cefaclor, third-generation cefixime, and fourth-generation loracarbef. These antibiotics were developed to be readily absorbed (generation one), more potent (generation two), much more potent (generation three), and, finally, much more stable in the body (generation four)[6]. All are synthesized via the pNB ester[5], and the protected intermediates are all only sparingly soluble in water. Fourth-generation loracarbef is synthesized chemically with no microbial fermentation steps. An efficient pNB esterase enzyme was needed to catalyze deprotection (Fig. 1A) toward the end of the loracarbef synthesis.

The pNB group, however, makes the resulting *p*-nitrobenzyl ester of the loracarbef nucleus (LCN-pNB) virtually insoluble in water. Wildtype pNB esterase suffers from a limitation common to a large number of enzymes evaluated for use in chemical synthesis: While the substrates are only sparingly soluble in water, the enzyme's catalytic capability is drastically reduced by even small quantities of water-miscible nonaqueous solvents. In this study, the pNB esterase gene has been subjected to sequential generations of random mutagenesis and screening followed by recombination, in order to dramatically increase pNB esterase's activity toward the pNB-protected loracarbef nucleus. In addition, by screening for enhanced activity in aqueous DMF, new enzymes that are useful tools for the hydrolysis of pNB esters in aqueous-organic solvents have been created.

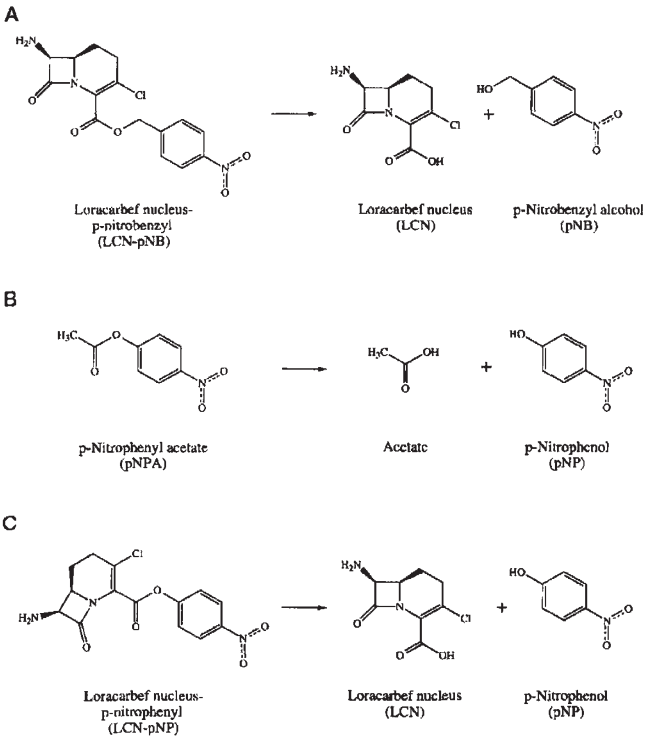

Figure 1. Substrates and products of reactions catalyzed by pNB esterase. (A) Desired hydrolysis of *para*-nitrobenzyl loracarbef nucleus. Screening reactions: (B) esterase-catalyzed hydrolysis of *p*-nitrophenyl acetate, (C) hydolysis of *p*-nitrophenyl loracarbef nucleus.

Results and discussion

Experimental design and strategy for directed evolution. For most problems of practical interest, mutant enzyme libraries must be screened rather than selected, one enzyme at a time. This effectively limits the library that can be searched to 10^4 to 10^6 variants, thereby precluding large evolutionary leaps in a single generation of random (point) mutagenesis. Most variations in amino acid sequence are either neutral or deleterious; beneficial mutations are relatively rare. Therefore, an effective strategy for directing the evolution of an enzyme involves creating and screening libraries of enzyme variants that differ from the parent by only one or two amino acids[7]. The screening method used to identify useful variants should therefore ensure that the expected small enhancements of the rare beneficial mutations can be measured, and the generation of new, useful enzymes relies on accumulating many such small improvements. Our strategy was to use sequential generations of random mutagenesis and screening, in which a single variant was chosen in each generation as the parent for the next. This 'asexual' evolutionary approach can also be augmented with 'sexual' or recombination methods for rapid accumulation of beneficial mutations[7–10]. Here we used a very simple method to randomly recombine beneficial mutations.

Directed enzyme evolution requires a rapid and sensitive screen for the properties of interest. Visual (colorimetric) assays are optimal in this regard. The desired reaction (Fig. 1A) is problematic because the reactant and products do not absorb in the visible region and their UV spectra are very similar. *Para*-nitrophenyl acetate (pNPA) is a general esterase substrate (Fig. 1B). The ability of the alcohol oxygen to form resonance structures that participate in conjugation with the phenyl ring gives rise to the characteristic yellow color of the nitrophenol product. The pNPA substrate is membrane permeable and is hydrolyzed rapidly by whole *E. coli* cells expressing pNB esterase intracellularly. Cells without the expression plasmid do not catalyze the conversion. Because this substrate is sterically and chemically very different from the loracarbef-pNB ester (LCN-pNB) it is not well suited for optimizing esterase activity on LCN-pNB. It is useful, however, for screening to optimize enzyme expression in new hosts.

To generate enzymes with improved activity toward the LCN-pNB substrate, a new substrate was designed to include the *p*-nitrophenol chromophore and as much of the loracarbef nucleus as possible. The resulting "hybrid" LCN-pNP substrate shown in Figure 1C is also membrane permeable, obviating the need for cell lysis during screening of the bacterial colonies for esterase activity. Also important is the choice of screening conditions. These should mimic as closely as possible the ultimate desired reaction conditions (temperature, pH, solvent, substrate concentration, etc.)[2]. Screening for improved pNB esterases was carried out at room temperature, in pH 7.0 buffer, increasing the concentration of DMF (up to 25% v/v) as the activity improved in successive generations.

For this study, mutations were made in the entire pNB esterase gene sequence by error-prone polymerase chain reaction (PCR) methods. Earlier directed evolution efforts to enhance the activity of subtilisin E in aqueous DMF generated beneficial amino acid substitutions distributed over the entire primary sequence, although they were generally clustered in the 3-D structure around the active site and substrate binding pocket[1,2]. Although targeting the mutagenesis to selected regions of the gene sequence can dramatically reduce the screening requirements, it imposes limitations on the possible solutions that we deemed unnecessarily restrictive.

The simple and effective error-prone PCR method[11–13] yields DNA mutation frequencies that are ideally suited to directed enzyme evolution[1]. If the mutation frequency is too high, most of the enzymes generated will be inactive. If it is too low, the wild-type background will be too high and little diversity will be sampled. Because approximately one-third of the DNA base substitutions are silent with respect to amino acid sequence, the ideal number of DNA base substitutions per targeted sequence is greater than one. When more than three DNA substitutions are generated per sequence, more than two amino acid substitutions per enzyme are produced on average. The enzyme's performance is a function of all the substitutions contained within; the result becomes a competition between the rare beneficial substitutions and the much more common deleterious substitutions[14]. Thus, the likelihood of finding an improved enzyme (the frequency of positive variants) in a given generation usually decreases with mutation frequency, because beneficial mutations will occur in a background of deleterious ones. Given the severe search constraints imposed by screening[7], a reasonable target for the mutation frequency is greater than one and not too much larger than three DNA base substitutions per sequence. The ideal substitution rate, and therefore the optimal conditions for the error-prone PCR method, depends on the length of the DNA targeted for random mutagenesis. PCR conditions which generate substitution rates from 0.25 to 20 substitutions per 1000 base pairs have been described[11–13]. Appropriate conditions can be identified to accommodate most gene lengths of interest.

Random mutagenesis and screening of pNB esterase. PCR random mutagenesis was performed in a 2000 base pair DNA segment encoding the entire 1500 bp pNB esterase open reading frame, as described in the Experimental Protocol. The conditions used for error-prone PCR were based on a desired substitution frequency of 1 to 3 substitutions per 1000 bases (1.5 to 4.5 substitutions per 1500 bp gene). Screening consisted of resuspending

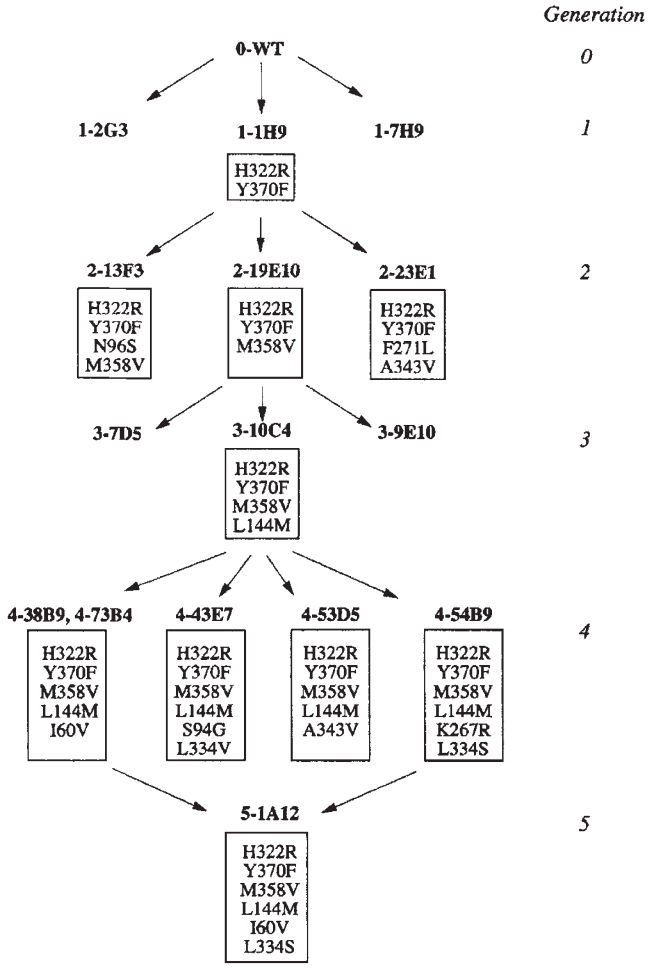

Figure 2. Lineage of pNB esterase variants generated by directed evolution. The esterase variants are labeled in bold-type; also listed are the amino acid substitutions found in each variant.

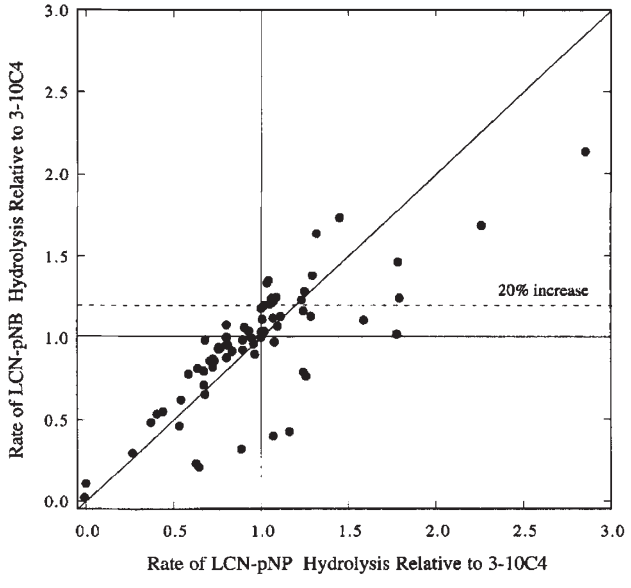

Figure 3. Hydrolysis rates of fourth-generation variants on pNB-LCN and pNP-LCN substrates, measured in whole-cell screening assay, 15% DMF, pH 7.0, 0.8 mM substrate. Rates are normalized to those of parent 3-10C4.

individual bacterial colonies in a small volume of buffer and solvent in 96-well plates. The turbidity of the bacterial suspension was used to estimate cell concentration. A small volume of the bacterial suspension was added to a buffered solution of a pNP substrate (pNPA or LCN-pNP) and DMF, and the rate of product appearance (yellow color) was normalized to the cell concentration in order to obtain an estimate of enzyme activity per bacterial cell. Those colonies that generated higher activity to turbidity ratios were retested. The variant pNB esterases of the best clones were then purified and tested again on the screening substrate and on LCN-pNB.

In the first generation of directed evolution, 1000 colonies were screened for activity on the pNPA substrate in 20% DMF. Of these, 33 were rescreened as potential positive variants. The three colonies with the highest activity to cell density ratios were grown, along with the bacteria expressing wildtype pNB esterase, in 1 L cultures. The partially purified enzymes were assayed for activity on the pNPA and LCN-pNB substrates. All showed higher total activity than wildtype on the pNPA substrate; only one, 1-1H9, showed significant total activity increase over wildtype on the actual LCN-pNB substrate. (1H9 designates the variant; the initial 1- indicates generation. 1-1H9 should be read "variant 1H9 of generation 1.") This variant was therefore used as the parent for the second round of mutagenesis. Figure 2 summarizes the lineages of all the enzyme variants characterized during this study.

In the second generation, 2800 colonies were screened on the LCN-pNP hybrid substrate and 15% DMF. Sixty-five potential positives were rescreened, and again the best three were grown in 1 L cultures along with the wildtype and the 1-1H9 parent. The purified pNB esterases (2-13F3, 2-19E10, and 2-23E1, 1-1H9 and wildtype) were assayed on all three ester substrates (pNPA, LCN-pNP, and LCN-pNB). While these second-generation variants had lost some of their ability to hydrolyze pNPA, all three exhibited increased activity (up to 2-fold higher) on LCN-pNP. For variants 2-13F3 and 2-19E10, the increased activity also applied to the desired *p*-nitrobenzyl substrate, LCN-pNB. However, the variant showing the most activity on LCN-pNP, 2-23E1, did not show marked improvement on LCN-pNB. 2-19E10 was used as the parent for the third round of mutagenesis (Fig. 2).

The third generation of directed evolution involved screening 1500 colonies on the LCN-pNP hybrid substrate and 20% DMF. Forty potential positives were rescreened, and the enzymes from the three best clones (3-7D5, 3-9E10, and 3-10C4) were purified. Of these three showing improved activity on LCN-pNP, only one, 3-10C4, was more active toward the LCN-pNB substrate. Third-generation 3-10C4 exhibited improvements of 40% over 2-19E10 in 2.5% DMF and 50% in 15% DMF for the hydrolysis of the *p*-nitrobenzyl ester.

In the fourth generation, 7400 colonies were screened for activity on LCN-pNP in 20% DMF. Two hundred fifty were rescreened as potential positives. Sixty-four clones, consisting of those either most active in 5% DMF, most active in 20% DMF, or the best ratio of activities in 20% to 5% DMF, were screened along with wildtype, 1-1H9, 2-19E10, and 3-10C4 on the LCN-pNB target substrate in 96-well plates. The activities of these clones on LCN-pNP and LCN-pNB normalized to the activities of the parent 3-10C4 are plotted in Figure 3. Of the 64 colonies, five show activity increases of 50% or more over 3-10C4 on the *p*-nitrobenzyl ester, while 16 show increases of 20% or more. The best variant was 4-54B9, with more than a 2-fold improvement over 3-10C4. The remaining four best variants all demonstrated approximately 60-65% improvement over 3-10C4; these were labeled 4-38B9, 4-43E7, 4-53D5, and 4-73B4 (Fig. 2).

Figure 3 allows us to evaluate the rapid screening assay using

the LCN-pNP substrate. If the enzymes' ability to hydrolyze this substrate perfectly mimicked their activity on the pNB ester, then all the points would lie on a line of slope 1.0. The scatter of the points about this line indicates that use of the pNP ester results in false positives as well as false negatives during screening. The two-level screening method, in which a much smaller number of positives are rescreened on the pNB substrate before choosing the parent for the next generation, is useful for avoiding backward steps in the evolution of pNB esterase activity. Furthermore, the line of best fit through the points has a slope of 0.7 (not shown), indicating that activities on the pNB substrate are generally not quite as high as activities toward the screening pNP substrate. Thus, there is a small, but finite, cost to substituting the pNP reaction for the pNB reaction during rapid screening. This cost is easily compensated, however, by the relative ease of screening large numbers of colonies with the *p*-nitrophenyl ester.

Kinetic characterization of evolved pNB esterases. Seven pNB esterases from the four generations of directed evolution, 0-WT, 1-1H9, 2-19E10, 3-10C4, 4-38B9, 4-43E7, and 4-54B9, were purified and further studied in order to follow the evolutionary progress of specific (as opposed to total) catalytic activity toward the various substrates. Table 1 summarizes the catalytic constants k_{cat} and K_M measured for hydrolysis of the three esters in various concentrations of DMF. Because the limited solubility of the pNPA substrate did not permit the high substrate concentrations required to accurately determine k_{cat} and K_M individually, only k_{cat}/K_M is reported for this substrate.

Evolutionary progressions of the catalytic efficiencies k_{cat}/K_M are displayed in Figure 4A–C for the three different substrates. The dramatic negative effect of the polar organic solvent DMF on esterase activity is apparent in these figures. Figure 4A shows the catalytic efficiency toward pNPA for the series of evolved pNB esterases. Only 1-1H9 was chosen based on its performance on this substrate; purification and analysis indicate that the majority of improvement in activity demonstrated by this clone during screening is due to an approximate 4-fold increase in amount of enzyme produced. The wildtype enzyme outperforms 1-1H9 in terms of specific activity in purely aqueous environments. The screening was carried out in the presence of DMF, however, and 1-1H9 exhibits higher specific activity towards pNPA than wildtype in both 15% and 30% DMF. The remaining variants behave similarly when assayed on this substrate. With increasing DMF, the 'evolved' enzyme variants compare more favorably with wildtype, and at 30% DMF, all the variants are at least somewhat more efficient. Variants 4-38B9 and 4-43E7, presumably by virtue of having been screened in DMF for four generations, are the best performers in 30% DMF.

Table 1 show the results of kinetic analyses performed using the hybrid LCN-pNP substrate with which three out of the four rounds of screening were carried out. The positive results of the directed evolution are clearly discernible. In 1%, 15%, and 30% DMF, the two least active enzymes are wildtype and 1-1H9. The catalytic rate constant of 2-19E10, the first variant to have been screened on the hybrid substrate, is increased by 50% over wildtype and a factor of three over its parent 1-1H9 in 1% DMF. The activity increase is sensitive to the presence of DMF; the k_{cat} enhancement is essentially lost in 30% DMF. This trend continues with the 3-10C4 variant, whose k_{cat} is again higher than its 2-19E10 parent in 1% DMF. 3-10C4, however, does not lose this in 30% DMF, where k_{cat} is twice that of 2-19E10. 3-10C4 is the parent of the remaining variants, all of which show further enhanced activity. 4-38B9 shows the smallest improvement, 4-43E7 shows a consistent nearly 2-fold increase in k_{cat} across all DMF ranges, and 4-54B9 is the most active of all the variants with a nearly 4-fold increase in k_{cat} over its parent. 4-54B9 exhibits nearly 16 times the specific activity of wildtype pNB esterase. At high substrate concentrations in 30% DMF, 4-54B9 hydrolyzes LCN-pNP as fast as wildtype pNB esterase does in water.

DMF dramatically increases K_M while also decreasing k_{cat}. As observed previously for subtilisin[1], the mutations accumulated during directed evolution mitigate the increased K_M. The effects on k_{cat}, however, are more prominent: k_{cat} increases more than 9-fold from wildtype to 4-54B9 in 15% DMF, while K_M decreases by less than a factor of two. This result reflects the relatively high substrate concentration used during screening (0.8 mM). At substrate concentrations on the order of K_M, increased specific activity will result mainly from improvements in k_{cat}. These improvements are obviously the most useful for enzymes intended to be used for transformations at high substrate concentrations. The overall catalytic efficiencies k_{cat}/K_M are plotted in Figure 4B, which clearly shows the evolutionary progression from one generation to the next.

When the pNB esterase variants were assayed on the LCN *p*-nitrobenzyl ester, similar trends were observed (Table 1 and Fig. 4C). Again the wildtype enzyme is the poorest performer, in terms of catalytic efficiency as well as maximum activity at high substrate concentrations. As was seen with the pNP screening substrate, the fourth-generation variants are more active than the third-generation variant, which is more active than the second-generation variant, etc., with only minor variations. The first is that 1-1H9 no longer lags wildtype in specific activity. The second is that many of the evolved variants, especially those in the fourth generation, exhibit slightly lower increases in activity with respect

Table 1. Kinetic parameters k_{cat}, K_m, and k_{cat}/K_m for pNB esterase variants, as derived from the curve fits of kinetic data[15]. All k_{cat}/K_M data are in units of s^{-1} M^{-1}. All data were collected in 0.1 M PIPES, pH 7.0 at 30°C.

p-Nitrophenyl acetate

Variant	0% DMF k_{cat}/K_M (s^{-1} M^{-1})	15% DMF k_{cat}/K_M (s^{-1} M^{-1})	30% DMF k_{cat}/K_M (s^{-1} M^{-1})
0-WT	5700	940	240
1-1H9	4100	1100	320
2-19E10	3200	890	280
3-10C4	2500	860	330
4-38B9	3000	1000	360
4-43E7	2400	810	290
4-54B9	2600	640	270

Loracarbef nucleus *p*-nitrophenyl ester

Variant	1% DMF k_{cat} (s^{-1})	1% DMF K_M (mM)	1% DMF k_{cat}/K_M	15% DMF k_{cat} (s^{-1})	15% DMF K_M (mM)	15% DMF k_{cat}/K_M	30% DMF k_{cat} (s^{-1})	30% DMF K_M (mM)	30% DMF k_{cat}/K_M
0-WT	0.14	0.07	2100	0.10	2.36	43	0.024	7.55	3
1-1H9	0.08	0.03	2500	0.07	1.84	38	0.021	6.05	3
2-19E10	0.25	0.09	2700	0.12	1.50	82	0.023	4.67	5
3-10C4	0.35	0.11	3200	0.21	1.77	118	0.044	5.18	9
4-38B9	0.39	0.06	6200	0.25	1.41	174	0.067	5.68	12
4-43E7	0.67	0.08	8500	0.40	1.27	313	0.094	5.09	18
4-54B9	1.44	0.12	12000	0.93	1.27	735	0.141	2.92	48

Loracarbef nucleus *p*-nitrobenzyl ester

Variant	1% DMF k_{cat} (s^{-1})	1% DMF K_M (mM)	1% DMF k_{cat}/K_M	15% DMF k_{cat} (s^{-1})	15% DMF K_M (mM)	15% DMF k_{cat}/K_M
0-WT	0.027	0.04	640	0.009	0.88	10
1-1H9	0.032	0.04	850	0.012	0.85	14
2-19E10	0.077	0.04	1700	0.030	0.87	35
3-10C4	0.11	0.07	1600	0.027	0.40	68
4-38B9	0.14	0.04	3700	0.031	0.31	100
4-43E7	0.15	0.03	4500	0.050	0.37	135
4-54B9	0.32	0.05	6600	0.11	0.46	246

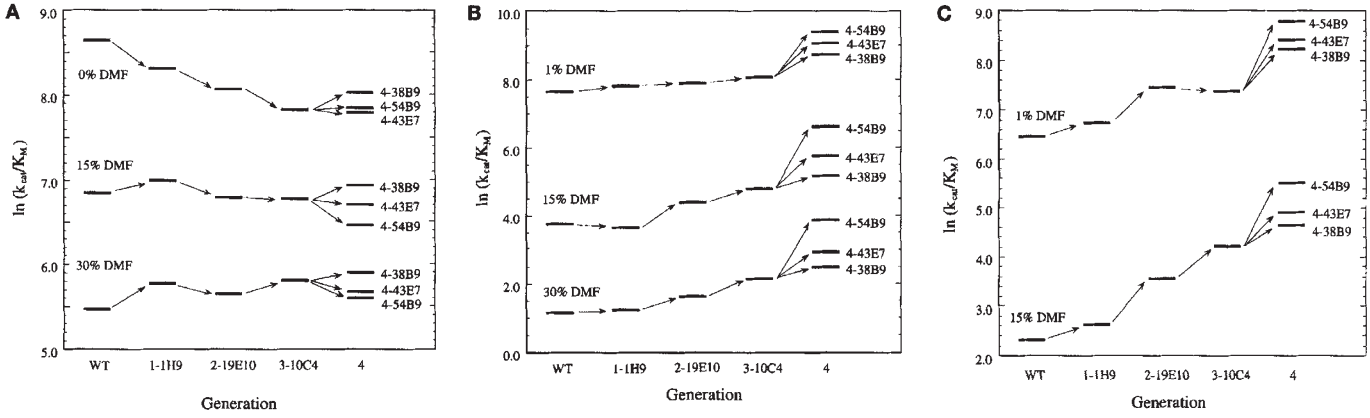

Figure 4. Evolutionary progression of catalytic efficiencies for the evolved pNB esterases on (A) pNPA, (B) pNP-LCN, and (C) pNB-LCN substrates. pNPA substrate was used to screen first generation. Subsequent generations were screened on pNP-LCN.

to the ancestral, wildtype enzyme. Finally, DMF has a greater negative effect in the LCN-pNB hydrolysis reaction than in the hydrolysis of LCN-pNP. Hydrolysis data on LCN-pNB in 30% DMF are not reported because the reaction rates were measured by the change in absorbance at 289 nm, and DMF absorbs strongly in this range. This, combined with the high substrate concentrations required in 30% DMF, made accurate rate measurements very difficult. The best pNB esterase, 4-54B9, is ten times more efficient than wildtype in 1% DMF; it is 24 times more efficient in 15% DMF.

Sequence analysis. Complete DNA sequences were obtained for all the pNB esterase genes listed in Table 2[15]. This table summarizes the positions of the DNA base changes in each gene with respect to the wildtype pNB esterase sequence, as well as new substitutions not present in the previous generation parent gene. All together, the sequences contain 31 base substitutions, of which 29 are unambiguously unique. The two substitutions which may not be unique are those where identical substitutions were found in two different variants of the same generation: the A to G substitutions at position 1075 in 2-13F3 and 2-19E10, and the A to G substitutions at position 181 in 4-38B9 and 4-73B4. In Figure 5, which plots the cumulative DNA substitutions as a function of position within the gene sequence, the DNA mutations are rather

Table 2. DNA mutations found in evolved pNB esterases. Bold-faced horizontal lines mark the beginning and end of the open reading frame (bp 1 and 1470). Mutations indicated in reverse type are those new to that generation. Normal type indicates mutations accumulated in previous generations.

DNA Pos	Gen. 1	Generation 2			Gen. 3	Generation 4					Gen. 5
	1H9	13F3	19E10	23E1	10C4	38B9	43E7	53D5	54B9	73B4	1A12
-12			A → T		A → T	A → T	A → T	A → T	A → T	A → T	A → T
87	C → T	C → T	C → T	C → T	C → T	C → T	C → T	C → T	C → T	C → T	C → T
102									T → C		
181						A → G				T → C	A → G
255								T → C			
283							A → G				
290		A → G									
291					T → C	T → C	T → C	T → C	T → C	T → C	T → C
333			T → A		T → A	T → A	T → A	T → A	T → A	T → A	T → A
399				T → C							
433					T → A	T → A	T → A	T → A	T → A	T → A	T → A
720								T → C			
803									A → G		
814			T → G								
968	A → G	A → G	A → G	A → G	A → G	A → G	A → G	A → G	A → G	A → G	A → G
1003							T → G				
1004									T → C		T → C
1031				C → T				C → T			
1075		A → G	A → G		A → G	A → G	A → G	A → G	A → G	A → G	A → G
1112	A → T	A → G	A → G	A → G	A → G	A → G	A → G	A → G	A → G	A → G	A → G
1122		A → G									
1239			A → G		A → G	A → G	A → G	A → G	A → G	A → G	A → G
1302		A → G									
1485	A → G	A → G	A → G	A → G	A → G	A → G	A → G	A → G	A → G	A → G	A → G
1568										T → A	
1618				T → C							
1658	A → G	T → C	T → C	T → C	T → C	T → C	T → C	T → C	T → C	T → C	T → C
1678	T → C	T → C	T → C	T → C	T → C	T → C	T → C	T → C	T → C	T → C	T → C
1745			C → T		C → T	C → T	C → T	C → T	C → T	C → T	C → T

evenly distributed along the targeted sequence in the evolved enzymes. The types of substitutions generated, however, are not well distributed. Of the 29 unique substitutions, 25 change an A or T, while four change C or G. These bases were changed almost half of the time to G, with A to G changes making up the majority (9/10) of these. Transitions outnumber transversions 24 to 5. This behavior is consistent with that reported by Leung for the PCR conditions used[11].

PCR conditions that result in a less biased mutation pattern tend to produce mutations at a frequency that is too high for genes of this length (1500 bp). The mutation bias could be reduced by using other PCR conditions that lead to a higher mutagenic rate[13] while decreasing the number of PCR cycles, since the accumulation of mutations is cycle-dependent[12]. Our results, however, show that the nonrandom nature of the mutations obtained using the current conditions do not preclude the production of substantially improved enzyme variants. Of course, the diversity of sequences that can be accessed is substantially reduced. Given that only a limited library of sequences can be screened, however, this may not become a severe practical limitation until evolutionary performance optima are approached.

The amino acid substitutions in the pNB esterase variants were determined by translation from the DNA sequences and are indicated in Figure 2. Because several of the DNA mutations lie outside the open reading frame, and even more are silent, the mutations at the amino acid level are far fewer than at the DNA level. For example, 1-1H9 contains six DNA mutations, of which only three lie within the pNB esterase open reading frame (Table 2). Of these, two lead to amino acid changes and one is silent. This mutated gene served as the parent sequence for the second generation. All three clones sequenced from the second generation contain the base substitutions observed in the 1-1H9 sequence as well as a few new substitutions. One silent and one translated DNA substitution are added to the evolutionary sequence in 3-10C4; a T to A substitution at position 433 gives rise to Leu 144 substituted by Met. The next five sequences are progeny of 3-10C4 from the fourth generation. All have the eight DNA base substitutions of 3-10C4 in common as the cumulative result of the three previous rounds of evolution.

When the locations of the DNA mutations that are translated into amino acid substitutions are investigated (highlighted with circles in Fig. 5), it becomes clear that the positions of the amino acid changes in the evolved pNB esterases are distinctly less random. Half of the translated DNA mutations lie within a 144 bp stretch of DNA (less than 10% of the open reading frame), which suggests that the corresponding region of the enzyme may be particularly important for substrate recognition and enzyme activity in DMF. This section of the sequence could clearly be targeted in a more intensive campaign involving random or site-directed mutagenesis. It is important to remember, however, that significant effective mutations also lie outside this region.

Recombination of positive mutations. Numerous protein engineering studies have shown that positive mutations, whether identified by screening of randomly mutagenized colonies or by so-called "rational design approaches," can often be combined for additional beneficial effects. Generally, this has been accomplished by site-directed mutagenesis, a laborious task when multiple mutations are involved. Application of the evolutionary strategy of sequential generations of random mutagenesis and screening used here automatically searches for those mutations that work in concert with mutations from previous generations. This approach, however, leaves behind potentially useful and laboriously obtained information in the form of all the other positive variants identified but not chosen to parent a subsequent generation. In order to identify the best possible combination of multiple beneficial mutations, especially if complete sequences are not yet available, a method for random gene recombination is useful. Screening the resulting library of recombined variants can rapidly identify the most effective combinations of mutations.

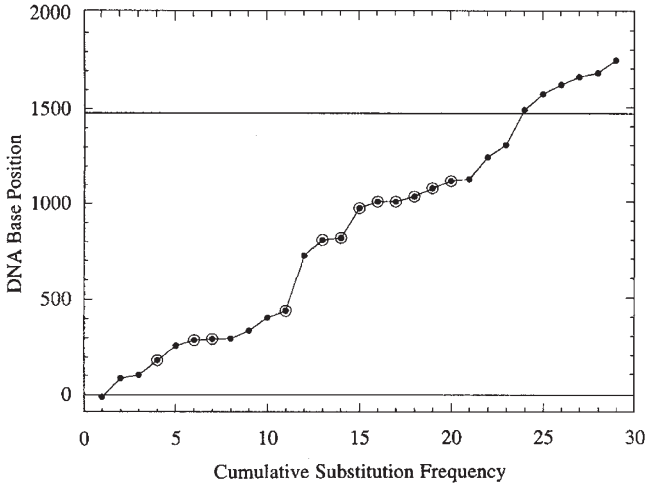

Figure 5. Locations and cumulative frequency of DNA mutations reported in Table 2. Horizontal lines mark the open reading frame. Circles indicate nonsilent DNA mutations.

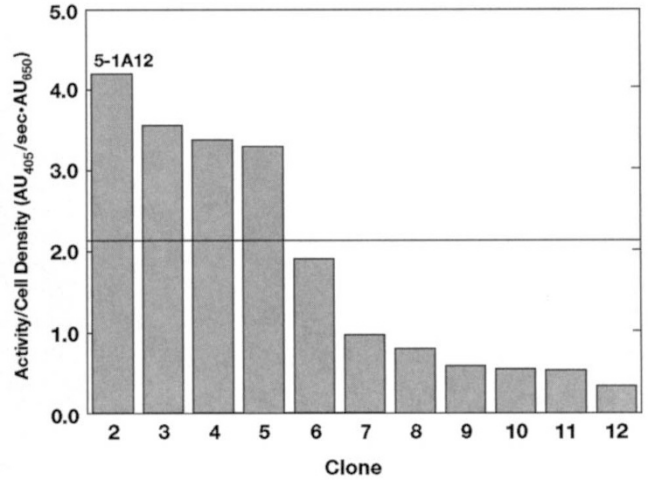

Figure 6. Activities of clones obtained from random ligation mixture of 4-38B9 and 4-54B9. Horizontal line indicates the activity of 4-54B9.

To demonstrate the utility of gene recombination in the context of the directed evolution of pNB esterase, a small, biased library of fifth-generation variants was generated by recombining the genes from the fourth-generation variants by restriction and religation. The purified plasmids from the five variants from generation four were each cut with restriction enzymes Xba I, Bam HI, and Xho I and purified, as described in the Experimental Protocol. The DNA fragments were mixed with the DNA fragments from 4-54B9, the best variant from the fourth generation, in pairwise fashion (e.g., one tube contained the fragments from 4-38B9 and 4-54B9, a second tube contained fragments from 4-43E7 and 4-54B9, etc.). *E. coli* transformed with the ligation products of these mixtures

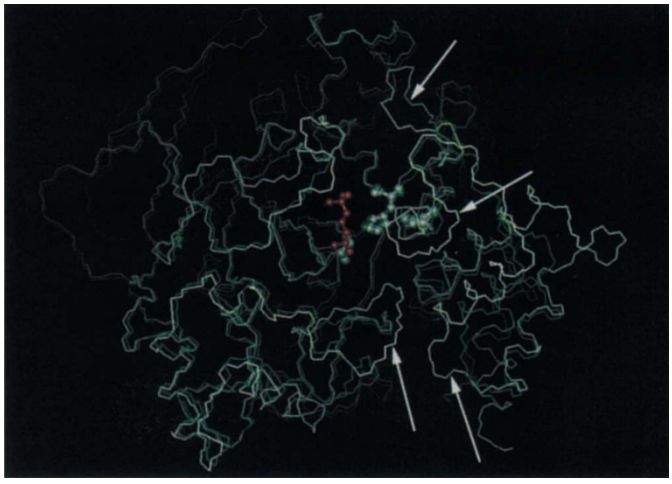

Figure 7. Superposition of acetylcholinesterase crystal structure (grey)[17] and pNB esterase model[15]. Active site His, Ser, and Glu residues are highlighted with ball-and-stick structures. Acetylcholine substrate is shown in red. Sequences which serve to define the size and shape of the substrate binding cavity and are present in acetylcholinesterase, but absent in pNB esterase, are indicated by the arrows. The superposition was performed using InsightII from Biosym.

and the expression plasmid was screened on the LCN-pNP substrate in 25% DMF. Of the four recombination attempts, only the mixture recombining 4-54B9 and 4-38B9 resulted in an enhancement in activity over 4-54B9. According to the screening results shown in Figure 6, this clone (labeled 5-1A12) displays approximately twice the activity of 4-54B9. We estimate that this clone expresses 50–60-fold more total esterase activity than the wildtype system.

Variant 4-38B9 contains one new amino acid substitution not found in its parent enzyme, I60V, while 4-54B9 contains two, K267R and L334S (Fig. 2). The DNA mutation responsible for I60V is at bp 181 (Table 2), upstream of the XhoI cleavage site (bp 991). The DNA mutation responsible for K267R is also located upstream of the cleavage site, at bp 803, while that for L334S is downstream (bp 1004). Thus, clone 5-1A12 is believed to contain the I60V and L334S amino acid substitutions (Fig. 2). In addition, the K267R mutation has been removed.

Further effective combinations of mutations could obviously be found by random recombination methods or, now that complete DNA sequences are available, by site-directed mutagenesis. For example, we have used the PCR recombination method of Stemmer[8–10] to recombine a variety of improved pNB esterases to obtain further enhancements in activity[7] (unpublished).

Homology studies and structural modeling. A search of the major protein databases (Protein Information Resource, Swiss Protein, translated GenBank and Protein Data Bank) revealed that *B. subtilis* pNB esterase shares significant homology with a number of esterases in the α/β hydrolase family, including acetylcholinesterase, butyrylcholinesterase,carboxylesterase,thioesterase, several lipases and cholesterol esterases, and carbamate hydrolase[15]. The carbamate hydrolase was discovered by screening organisms for activity on phenmedipham, an aromatic carbamate herbicide[16]. Discovered in an *Arthrobacter oxidans* strain from soil samples of phenmedipham-treated fields, the enzyme's natural function is not known. Although these enzymes all share significant sequence homology (>30% identity), and therefore presumably also share a common three-dimensional fold, they are responsible for hydrolyzing a wide range of substrates. The pNB esterase hydrolysis mechanism can be inferred as proceeding through an acyl enzyme intermediate common to this serine hydrolase family. Sequence alignment indicates that the catalytic triad of pNB esterase consists of Ser 189, Glu 310, and His 399[15].

Sequence alignments of 11 homologous enzymes and x-ray crystal structures of *Torpedo californica* acetylcholinesterase[17] and *Geotrichum candidum* triacylglycerol lipase[18] were used to generate an approximate three-dimensional structure for pNB esterase[15]. This was accomplished using the program MODELLER, which used the alignments, crystal structures, and a large array of probabilistic distance, dihedral angle, and spatial restraints derived from a database of three-dimensional structures[19,20]. The resulting pNB esterase model is compared to the acetylcholinesterase x-ray structure in Figure 7.

Although an extensive library of bacteria and fungi were screened to find pNB esterase, the search was by no means exhaustive[3]. The possibility still exists that other enzymes might also have

Table 3. Amino acid substitutions in evolved pNB esterases.

Amino acid substitution	pNB esterase variant	Location	Distance from substrate[a] (Å)	Element of secondary structure	Original residue conserved[b]	Replacement residue found[c]	Most common residues at this position
Ile 60 Val	4-38B9, 4-73B4	near surface	13	none	no	2	M/S
Ser 94 Gly	4-43E7	surface	30	surface loop	no	4, 8, 9, 10	G/K
Asn 96 Ser	2-13F3	surface	30	surface loop	no	10	N/K
Leu 144 Met	3-10C4	buried	20	surface loop	yes	5	L
Lys 267 Arg	4-54B9	near surface	12	none	no	none	I (6/11)[d]
Phe 271 Leu	2-23E1	surface	13	surface loop	no	8, 9	L (6/11)[d]
Leu 334 Val	4-43E7	surface	16	helix	yes	1, 2	L
Leu 334 Ser	4-54B9	surface	16	helix	yes	none	L
Ala 343 Val	2-23E1, 4-53D5	near surface	20	helix	no	1, 2, 6, 7	V/L
Met 358 Val	2-13F3, 2-19E10	surface	7	helix	no	1, 2, 3, 10	V/L
(His 322 Arg)	1-1H9	surface	24	surface loop	no	none	E/D (6/11)[d]
(Tyr 370 Phe)	1-1H9	buried	15	helix	no	1, 3	L/V

[a]Shortest distance from this residue to the substrate. [b]Yes = at least half of the homologous proteins contain the residue substituted at the corresponding location. [c]Indicates the homologous proteins from following list which contain the substituted amino acid residue. 1 - Acetylcholinesterase, *Torpedo californica*, Ref. 17. 2 - Acetylcholinesterase, *Oryctolagus cuniculus*, Ref. 24. 3 - Butyrylcholinesterase, *Oryctolagus cuniculus*, Ref. 25. 4 - Carboxylesterase, *Oryctolagus cuniculus*, Ref. 26. 5 - Carboxylesterase, *Dictyostelium discoideum*, Ref. 27. 6 - Thioesterase, *Anas platyrhynchos*, Ref. 28. 7 - Triacylglycerol lipase, *Geotrichum candidum*, Ref. 18. 8 - Triacylglycerol lipase, *Candida rugosa*, Ref. 29. 9 - Cholesterol esterase, *Candida rugosa*, Ref. 30. 10 - Carbamate hydrolase, *Arthrobacter oxidans*, Ref. 18. [d]Number of sequences (out of eleven) that contain amino acid residues at this location.

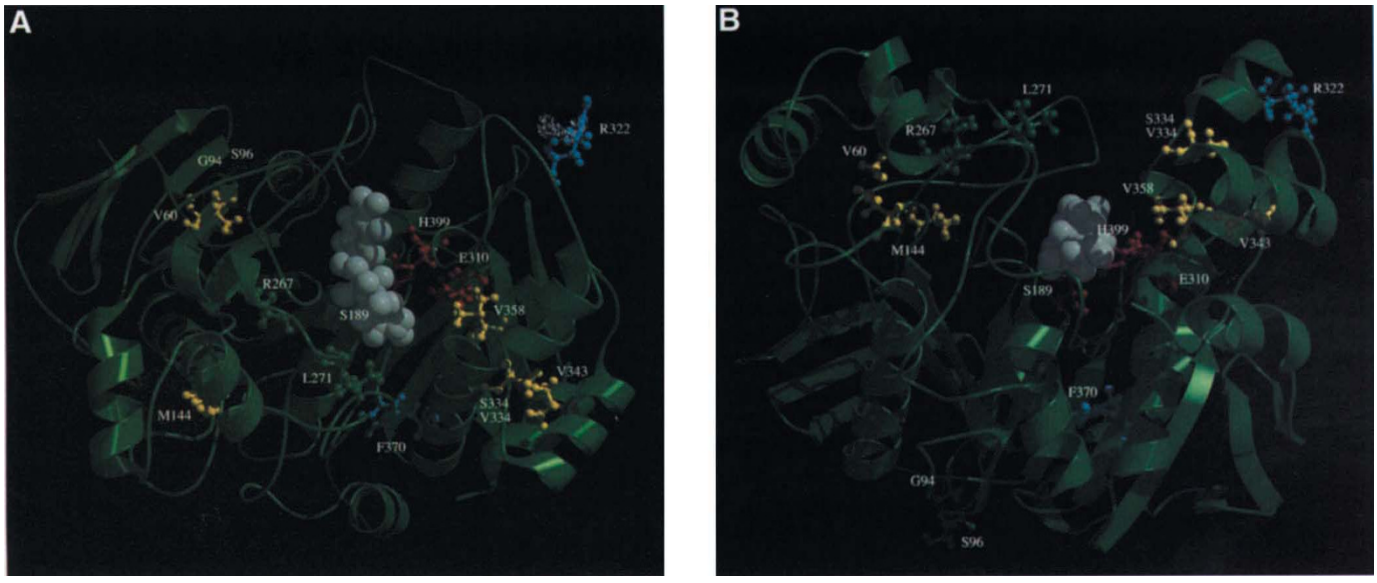

Figure 8. Structure of pNB esterase with the LCN-pNB substrate (grey)[15]. Active site residues are shown in red. Amino acid substitutions found in evolved pNB esterase variants are highlighted: substitutions found in 1-1H9 are shown in cyan, substitutions believed to be neutral are indicated in green, while substitutions that enhance specific esterase activity are shown in yellow. (A) Top view highlights the radial distribution of mutations around the active site. (B) 'Side' view shows that the effective mutations are clustered in the top half of the enzyme.

substantial activity towards this substrate. These sequences could be used to accelerate a search for a more highly active pNB esterase. This possibility is further enhanced by the fact that many of the enzymes showing the greatest homology with pNB esterase are mammalian enzymes, which were not tested in the original search for a pNB esterase. Six enzymes chosen based on their homology with pNB esterase or their known activity on a broad range of ester substrates were tested for the ability to hydrolyze the LCN-pNB substrate. Acetylcholinesterase (*Electrophorus electricus*), butyrylcholinesterase (horse serum), *Chromobacterium viscosum* lipase, cholesterol esterase (porcine pancreas), as well as the broadly specific subtilisin A (*B. licheniformis*) and porcine pancreatic elastase demonstrated no measurable activity toward LCN pNB[15].

The fact that the other esterases do not catalyze the LCN-pNB hydrolysis reaction can be rationalized based on the structures in Figure 7. A small number of amino acid differences between acetylcholinesterase and pNB esterase significantly alter the binding pocket. The remaining homologous enzymes show similar variability in the sizes and residue compositions of loops responsible for defining the substrate binding pocket[15]. The wide array of substrates recognized by the different esterases reflects sequence and structural variations in this region. It also means that significant changes in substrate selectivity can be expected as a result of accumulated alterations in the amino acids surrounding the binding pocket.

Structural analysis of the evolved pNB esterases. Figure 8 shows top and side views of wild-type pNB esterase with the LCN-pNB substrate (grey)[15]. Residues in the catalytic triad are indicated in red. A total of 12 amino acid substitutions at 11 different sites were found in the variants with improved activity towards various substrates and in the presence of DMF. These substitutions are highlighted in Figure 8 and summarized in Table 3, which also indicates whether the amino acid is buried or surface-accessible, its position in secondary structure, and whether or not the particular amino acid is conserved among the homologous enzymes. Similar to what was observed during the directed evolution of subtilisin for activity in dimethylformamide[1,2], the majority of the amino acid substitutions lie on or near the surface of the enzyme and/or are clustered around the substrate binding pocket.

Of the 12 amino acid substitutions identified, two were found in 1-1H9, an enzyme variant expressed at higher levels but without increased specific activity (His 322 Arg and Tyr 370 Phe, listed in parentheses in Table 3 and indicated in blue in Fig. 8). Expression level is not necessarily a property of the amino acid sequence; alterations in the DNA sequence may play an important role in transcriptional regulation of expression. The two amino acid changes or any of the six DNA mutations could be responsible for increasing the total activity of this clone.

Of the remaining 10 amino acid substitutions, the four at positions 94, 96, 267, and 271 are believed to be neutral or deleterious (shown in green in Fig. 8). The Asn 96 Ser substitution occurs in variant 2-13F3. 2-13F3 and 2-19E10 share Met 358 Val, which is the only mutation found in the more active 2-19E10 variant. It can therefore be argued that position 358 is responsible for the improved activity, while the substitution at position 96 is neutral, or even possibly deleterious. Similarly, since variants 2-23E1 and 4-53D5 share Ala 343 Val, and this is the only substitution found in going from 3-10C4 to 4-53D5, the additional Phe 271 Leu substitution in variant 2-23E1 is probably incidental. The substitutions at residues 94 and 267 were found accompanied by one additional amino acid substitution each in 4-43E7 and 4-54B9, respectively. These two variants share substitutions at position 334, and this residue is likely to be responsible for the enhanced activity.

Most mutations are neutral or deleterious. The size of the library containing all double amino acid substitutions is larger than the sequence length squared. Screening on the order of a thousand colonies covers only a tiny fraction of this mutant space, and is, therefore, very unlikely to uncover a sequence containing two mutations, both of which positively contribute to activity enhancement. Thus, it is argued that only one of the two mutations in 4-43E7 and 4-54B9 is responsible for these variants' increase in activity, and that is the one they have in common, at

position 334. This is further supported by the recombination experiment (*vide infra*) which combined Val 60 from 4-38B9 and Ser 334 from 4-54B9 while omitting Arg 267, to create 5-1A12, a variant with twice the apparent activity of 4-54B9.

The neutral substitutions at residues 94, 96, 267, and 271 are all located at or near the surface of the enzyme in flexible surface loops and outside discernible elements of secondary structure. In addition, these positions are not conserved among the homologous enzymes, suggesting that the pNB esterase can tolerate variability at these locations.

The six remaining substitutions are those believed to be responsible for enhancements in specific esterase activity. Four of these (at residues 334, 343, and 358) are positioned in α-helices, located on the LCN end of the elongated substrate binding pocket and on one side of the entrance to the binding pocket, as seen in Figure 8. Val 358 is the closest to the substrate, at 7 Å from the nearest point of the substrate, while the others are up to 20 Å away (Table 3). Thus, it appears that none of these four mutations can directly interact with bound substrate. The amino acids at these locations in the wildtype and homologous enzymes are typically hydrophobic; the large, hydrophobic amino acid Leu predominates at these positions in the homologous enzymes. The evolved pNB esterase variants contain smaller valine and serine. Ala 343 Val in variant 2-23E1 is responsible for an increase in activity on the screening substrate LCN-pNP but does not improve activity toward LCN-pNB. Why this mutation should affect activity toward one substrate and not the other is unclear.

The remaining amino acid substitutions that enhance esterase activity are at positions 60 and 144. Ile 60 Val (in 4-38B9) is responsible for a significant improvement in catalytic efficiency, as indicated in Figure 4. This residue is located near the surface of the enzyme, 13 Å from the closest point of the substrate. This is the only activity-enhancing mutation that is closer to the leaving group (pNP or pNB) than to the LCN portion of the substrate. Leu 144 Met is the only one of the six that is buried deep beneath the solvent-accessible enzyme surface. It is also located relatively far from the substrate, at 20 Å from the LCN. This substitution's contribution to the activity of 3-10C4 is greatly enhanced in the presence of DMF (Table 1, Fig. 4). The only other variant showing similar behavior is 1-1H9 toward pNPA. 1-1H9 is also the only other variant to have a completely buried substitution (Tyr 370 Phe).

In general, it is difficult to explain the mechanism(s) by which these amino acid substitutions enhance the catalytic activity of the evolved pNB esterases toward any of these substrates. Although the effective amino acid substitutions all appear to cluster in the 'top' half of the enzyme containing the active site (Fig. 8B), none lie in segments predicted to interact directly with the bound substrate. To have predicted that these substitutions would enhance pNB esterase activity would have been impossible.

Experimental protocol

pNB esterase expression system. Plasmid pNB106R containing the pNB esterase gene under the control of an altered λ promoter, *pL*106[5,21] was kindly provided by Stephen Queener at Eli Lilly & Co. (Indianapolis, IN). The plasmid also contains a temperature-sensitive λ CI repressor, which inactivates the *pL*106 promoter below 35°C.

Restriction of DNA. Plasmid and fragment DNA when required were cut with Bam HI, Xba I, and where indicated, XhoI (Boehringer Mannheim, Indianapolis, IN) in restriction buffer B at 37°C for 1 h. The resulting linear DNA was then run on a 1% agarose gel and separated into bands according to size. Appropriate bands were excised and extracted using either the GeneClean (Bio101, La Jolla, CA) or Qiagen (Chatsworth, CA) methods and eluted in Tris-EDTA buffer.

Random mutagenesis. The pNB esterase gene (1470 bp) in pNB106R is flanked by an XbaI restriction site 51 bp before the start of the ORF and by a BamHI site 313 bp downstream from the stop codon5. PCR primers (3′-GAGCACATCAGATCTATTAAC-5′ and 3′-GGAGTGGCTCACAGTCG-GTGG-5′) were synthesized to complement regions 25 bp upstream of the XbaI site and 143 bp downstream of the BamHI site to allow random mutagenesis over a 2000 bp region including the entire pNB esterase ORF. A 100 μl volume of solution containing 1 mM dNTPs, 16.6 mM $(NH_4)_2SO_4$, 67 mMTris-HCl (pH 8.8), 6.1 mM $MgCl_2$, 6.7 μM EDTA (pH 8.0), 10 mM β-mercaptoethanol, 10 μl DMSO, 30 ng of forward and reverse primers, 10 ng of plasmid pNB106R and 2.5 units of Taq DNA polymerase (Perkin Elmer-Cetus, Emeryville, CA) was covered with 2–3 drops of light mineral oil (Sigma, St. Louis, MO) and placed in a Precision Scientific (Chicago, IL) thermal cycler well containing 2–3 drops of mineral oil. The thermal cycler repeated the following steps for 25 cycles: 1 minute at 94°C, 1 minute at 42°C, and 2 minutes at 72°C, designed to generate an error frequency of approximately one substitution per 1000 bases[11]. The fragment of DNA amplified by this technique was then subjected to a phenol/chloroform extraction and ethanol precipitation. The DNA was restricted and purified as described above.

Competent cell preparation. Competent TG1 cells were prepared using the $CaCl_2$ method[22]. TG1 cells were grown overnight at 37°C in a 3 ml culture of LB broth. The cells were diluted 1:200 in fresh LB and allowed to grow to an OD_{600} of 0.35 to 0.40. They were placed on ice for 1 h and spun at maximum speed in a 4°C Beckman tabletop centrifuge. The cell pellet was resuspended in 0.5 volumes of 0.1 M $CaCl_2$ and allowed to sit on ice for 30 min to 1 h and recentrifuged as before. The cell pellets were resuspended in sterile 0.02 volumes of 0.1 M $CaCl_2$, 10% v/v glycerol and frozen at –70°C until use.

Ligation and transformation. Ligation reactions were performed using T4 DNA ligase (Boehringer Mannheim). Vector DNA (the entire pNB106R plasmid excluding the pNB esterase gene between XbaI and BamHI), insert DNA (the pNB esterase gene between XbaI and BamHI or, where indicated, the two fragments XbaI to XhoI and XhoI to BamHI), 10X ligation buffer, water and enzyme were combined and incubated at 4°C overnight (12–16 h). The solution was incubated with previously prepared competent cells on ice for 1 h. The cells were then heat shocked at 42°C for 1 min, supplied with an equal volume of LB media, and incubated at 30°C for 45 min. This solution was then plated onto LB plates containing tetracycline to 20 μg/ml (LB Tet plates).

Screening. TG1 transformants arising from ligations of pNB106R vector and randomly mutagenized inserts were allowed to grow for 36 to 48 h before shifting to 42°C to induce expression of the pNB esterase gene. After an 8-h induction period, each colony was picked with a sterile toothpick and resuspended in a unique well of a 96-well plate containing 200 μl of 0.1 M Tris-HCl, pH 7.0. The turbidity of each well was measured as the absorbance at 620 nm adjusted by a cell-free reference well by a SInstruments 96-well plate reader. A 20 μl aliquot from each well was pipetted into a second 96-well plate, to which was added 200 μl of a substrate solution containing 0.8 mM pNPA and 0.4% (v/v) acetonitrile or 0.3 to 0.8 mM LCN-pNP, 0.1M Tris-HCl pH 7.0, and 0% to 30% v/v DMF (EM Science Guaranteed Reagent grade). The resulting reaction was monitored using the 96-well plate reader at 405 nm. Reactions were typically monitored for 11 data points varying from 15 sec between data points for 0% DMF measurements to 180 sec between data points for 30% DMF. The slopes of the best-fit lines through the 11 data points for each of the 96 wells were normalized by the corresponding absorbance at 620 nm. These normalized values were compared, and the wells exhibiting the highest activity to turbidity ratios were plated onto LB Tet plates. Two single colonies from these plates were restreaked on LB Tet plates to provide single-colony isolates for further testing. Two single colonies from each of these second plates (four colonies total) were then arrayed onto LB Tet plates using sterile toothpicks. This collection of potential positive variants was then rescreened using the activity to turbidity ratio assay. Crude screening was performed on the fourth-generation variants with the LCN-pNB substrate using a similar whole-cell assay. One hundred-μl samples of the resuspended colonies were removed from each well of the 96-well plate and added to a quartz cuvette containing a 1 ml reaction solution consisting of 2.5% DMF, 0.1 M Tris-HCl pH 7.0, and 0.25 mM LCN-pNB. The absorbance at 291 nm of each sample was measured for 2.5 min using a UV spectrophotometer. Initial rates were measured for both the LCN-pNP substrate in the 96-well plate assay (above) and the LCN-pNB substrate in quartz cuvettes. Both sets of slopes generated from the initial rate data were normalized to the turbidity measurements at 620 nm.

Cell growth. Single colonies were inoculated into 5 ml LB Tet culture tubes and allowed to grow overnight at 30°C. The contents of these tubes were then used to inoculate a 1-L culture of LB Tet and allowed to grow to maximum turbidity. These 1-L cultures were decanted into sterile centrifuge bottles and spun at 6000 rpms in a JA-10 rotor for 15 min in a Beckman (Palo Alto, CA) centrifuge. The cell pellets were resuspended in LB Tet pre-warmed to 42°C. The flasks were placed in a 42°C incubator and allowed to shake for 8 h°. The cells were harvested by centrifugation, resuspended in a centrifuge tube in 25 ml of Buffer A (10 mM potassium phosphate, 1 mM β-mercaptoethanol, and 0.5 mM EDTA, pH 7.0) and placed on ice.

Enzyme purification. PNB esterases were purified using a modification of the procedure of Chen et al.[24]. Cells were lysed using a French Press (SLM Instruments, Urbana, IL) by placing the chilled sample into a steel housing, which was compressed to 20,000 atmospheres and then released to ambient conditions, causing the cells to rupture. This process was repeated three times. The steel housing was kept chilled prior to use at 4°C and the samples were stored before and after on ice. After lysis the cell debris was pelleted by centrifugation at 12,000g in a JA-20 rotor for 15 min at 4°C. The cell lysate supernatant was adjusted to pH 5.0 with HCl, and the newly formed precipitate was removed by centrifugation at 12,000g in a JA-20 rotor for 30 min at 4°C. The supernatant volume was measured, and ammonium sulfate was dissolved to 45% saturation at 0°C (18.5 g/100 ml). The solution was chilled on ice for 5 min and centrifuged in a JA-20 rotor at 12,000g for 30 min at 4°C. The supernatant was transferred to a new centrifuge tube, where ammonium sulfate was added to bring the final amount to 85% saturation at 0°C. Centrifugation was performed as before, and the supernatant discarded. The pellet was redissolved in Buffer B (10 mM Tris-HCl, 50 mM NaCl, 1 mM β-mercaptoethanol, and 0.5 mM EDTA, pH 8.5), placed in an Amicon (Beverly, MA) spin filtration unit (Centricon-10) and buffer exchanged three times with Buffer B to remove the ammonium sulfate. The resulting protein sample was applied to a DEAE-sepharose column (2.5 cm ID × 10 cm high) precquilibrated in Buffer B. The column was rinsed with Buffer B until the baseline was restored. The column was then rinsed with Buffer C (10 mM Tris-HCl, 50 mM NaCl, pH 7.0) until the pH reached 7.0. An NaCl gradient from 50 to 500 mM in Buffer C (300 ml total volume) was passed through the column and fractions collected. Those fractions containing activity were pooled and then applied to an immobilized metal affinity chromatography (IMAC) column (2.5 cm ID × 10 cm, Fast-flow Chelating Sepharose, Pharmacia, Uppsala, Sweden) loaded with Cu^{2+} and prepared as per the manufacturer's instructions. The column was first pretreated by rinsing with three column volumes of 0.5 M NaCl, 50 mM EDTA, pH 8.5, to remove all chelated metal ions, 2 M NaCl to remove any ionically bound material, and 1 M NaOH to remove any denatured protein. Copper as 100 mM copper sulfate in 100 mM sodium acetate, pH 4.6, was loaded onto the IMAC column, washed with 20 mM sodium phosphate, 0.5 M NaCl, 50 mM imidazole pH 7.2 until pH 7.2, and finally equilibrated with five column volumes of 20 mM sodium phosphate, 0.5 M NaCl, 1 mM imidazole for sample loading. The sample was applied to the column and the column washed with the 1 mM imidazole solution until the baseline was restored. A linear gradient formed by 100 ml of 1 mM and 10 mM imidazole solutions (200 ml total volume) was applied, and fractions were collected. All tubes demonstrating higher than background activity were pooled, concentrated, and buffer exchanged into 0.1 M Tris-HCl, pH 7.0, in the Amicon Centricon-10 units as before. Protein purity was verified by SDS-PAGE, and concentration was determined using the Bio-Rad (Hercules, CA) Protein Assay Reagent.

Kinetic assays. Kinetic assays were performed on the pNPA, LCN-pNP, and LCN-pNB substrates. For pNPA and LCN-pNP, concentrations varying from 0.0625 mM to 16 mM and 0, 15, and 30% DMF in 0.1 M PIPES (Sigma), pH 7.0, were combined with equal volumes of enzyme samples in a 96-well plate. The samples were mixed simultaneously using an eight-channel pipetteman, and the reaction was monitored using the 96-well plate reader. The absorbance values were recorded, and initial slopes were obtained by linear regression and used for calculating k_{cat}/K_M. For the LCN-pNB substrate, concentrations varying from 0.0156 mM to 8.0 mM and 0%, 15% and 30% DMF in PIPES, pH 7.0, were combined with enzyme samples in a quartz cuvette and were measured in a spectrophotometer at 289 nm. All assays were performed in triplicate. K_M and k_{cat} values were obtained by non-linear regression of the reaction data using the program Kaleidograph (Abelbeck Software). k_{cat} values were corrected for the effects of DMF on the extinction coefficients of *p*-nitrophenol and *p*-nitrobenzyl alcohol.

Acknowledgments

The authors are grateful to Stephen Queener and Cathleen Cantwell of Eli Lilly & Co. for materials and for many helpful discussions. They also acknowledge the able assistance of Tao Long, Kejian Wu, Janice Lau and Roopa Ramamoorthi. This work was supported by the US Office of Naval Research (N00014-91-J-1397) and the US Department of Energy's program in Biological and Chemical Technologies Research within the Office of Industrial Technologies, Energy Efficiency and Renewables. J. M. is grateful to Merck & Co. for the support of a Merck Graduate Fellowship.

1. Chen, K. and Arnold, F. 1993. Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtillsin E for catalysis in dimethylformamide. *Proc. Natl. Acad. Sci. USA* **90:**5618–5622.
2. You, L. and Arnold, F. H. 1995. Directed evolution of *Subtilisin E* in *Bacillus subtilis* to enhance total activity in aqueous dimethylformamide. *Protein Eng.* **9:**77-83.
3. Brannon, D. R., Mabe, J. A., and Fukuda, D. S. 1976. De-esterification of cephalosporin *para*-nitrobenzyl esters by microbial enzymes. *J. Antibiotics* **29:**121–124.
4. U. S. Patent 3,725,359 [1975].
5. Zock, J., Cantwell, C., Swartling, J., Hodges, R., Pohl, T., Sutton, K., Rosteck Jr., P., McGilvray, D., and Queener, S. 1994. The *Bacillus subtilis* pnbA gene encoding *p*-nitrobenzyl esterase—cloning, sequence and high-level expression in *Escherichia coli*. *Gene* **151:**37–43.
6. Cooper, R. D. G. 1992. The carbacephems: a new beta-lactam antibiotic class. *Am. J. Med.* **92** Supplement 6A:S2–S6.
7. Arnold, F. H. 1996. Directed evolution: creating biocatalysts for the future. *Chem. Eng. Science. In press.*
8. Stemmer, W. P. C. 1994. DNA shuffling by random fragmentation and reassembly: *in vitro* recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* **91:**10747–10751.
9. Stemmer, W. P. C. 1994. Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature* **370:**389–391.
10. Crameri, A., Whitehorn, E., Tate, E., and Stemmer, W. P. C. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nature Biotechnology* **14:**315–319.
11. Leung, D. W., Chen, E., and Goeddel, D. V. 1989. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* **1:**11–15.
12. Eckert, K. A. and Kunkel, T. A. 1991. DNA polymerase fidelity and the polymerase chain reaction. *PCR Methods Applic.* **1:**17–24.
13. Cadwell, R. C. and Joyce, G. F. 1992. Randomization of genes by PCR mutagenesis. *PCR Methods Applic.* **2:**28–33.
14. Chen, K. and Arnold, F. 1991. Enzyme engineering for nonaqueous solvents: random mutagenesis to enhance activity of subtilisin E in polar organic media. *Bio/Technology* **9:**1073–1077.
15. Moore, 1996. PhD thesis, Cal. Inst. of Tech., Pasadena, CA.
16. Pohlenz, H. D., Boidol, W., Schuttke, I., and Streber, W. R. 1992. Purification and properties of an arthrobacter-oxydans p52 carbamate hydrolase specific for the herbicide phenmedipham and nucleotide-sequence of the corresponding gene. *J. Bacteriol.* **174:**6600–6607.
17. Sussman, J. L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L., and Silman, I. 1991. Atomic structure of acetylcholinesterase from *Torpedo californica*: a prototypic acetylcholine-binding protein. *Science* **253:**872–879.
18. Schrag, J. D. and Cygler, M. 1993. 1.8 angstroms refined structure of the lipase from *Geotrichum candidum*. *J. Mol. Biol.* **230:**575–591.
19. Sali, A. and Blundell, T. L. 1993. Comparative modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234:**779–815.
20. Sali, A. and Overington, J. P. 1994. Derivation of rules for comparative modeling from a database of protein structure alignments. *Prot. Sci.* **3:**1582–1596.
21. U.S. Patent Application No. 07,739,2801.
22. Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. pp. 1.82–1.84 in *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
23. Chen, Y., Usui, S., Queener, S. W., and Yu, C. 1995. Purification and properties of *p*-nitrobenzyl esterase from *Bacillus subtilis*. *J. Ind. Micro.* **15:**10–18.
24. Jbilo, O., L'Hermite, Y., Talesa, V., Toutant, J. P., and Chatonnet, A. 1994. Acetylcholinesterase and butyrylcholinesterase expression in adult-rabbit tissues and during development. *Eur. J. Biochem.* **225:**115–124.
25. Jbilo, O. and Chatonnet, A. 1990. Complete sequence of rabbit butyryl-cholinesterase. *Nucleic Acids Res.* **18:**3990.
26. Ozols, J. 1989. Isolation, properties, and the complete amino-acid sequence of a 2nd form of 60-kda glycoprotein esterase—orientation of the 60-kda proteins in the microsomal membrane. J. Biol. Chem. **264:**12533–12545.
27. Bomblies, L., Biegelmann, E., Doering, V., Gerisch, G., Krafft-Czepa, H., Noegel, A. A., Schleicher, M., and Humbel B. M. 1990. Membrane-enclosed crystals in *Dictyostelium discoideum* cells, consisting of developmentally regulated proteins with sequence similarities to known esterases. *J. Cell Biol.* **110:**669–679.
28. Hwang, C-S. and Kolattukudy, P. E. 1993. Molecular cloning and sequencing of thioesterase B cDNA and stimulation of expression of the thioesterase B gene associated with hormonal induction of peroxisomal proliferation. *J. Biol. Chem.* **268:**14278–14284.
29. Lotti, M., Grandori, R., Fusetti, F., Longhi, S., Brocca, S., Tramontano, A., and Alberghina, L. 1993. Cloning and analysis of *Candida cylindracea* lipase sequences. *Gene* **124:**45–55.
30. Kaiser, R., Erman, M., Duax, W. L., Ghosh, D., and Joernvall, H. 1994. Monomeric and dimeric forms of cholesterol esterase from *Candida cylindracea*. Primary structure, identity in peptide patterns, and additional microheterogeneity. *FEBS Lett.* **337:**123–127.