Chemistry & Biology, Vol. 12, 1281–1289, December, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.chembiol.2005.09.012

High-Throughput Screening of Enzyme Libraries: Thiolactonases Evolved by Fluorescence-Activated Sorting of Single Cells in Emulsion Compartments

Amir Aharoni,[1,3] Gil Amitai,[1] Kalia Bernath,[1,2] Shlomo Magdassi,[2] and Dan S. Tawfik[1,*]
[1]Department of Biological Chemistry
Weizmann Institute of Science
Rehovot 76100
Israel
[2]Casali Institute of Applied Chemistry
The Hebrew University of Jerusalem
Givat Ram
Jerusalem 91904
Israel

Summary

Single bacterial cells, each expressing a different library variant, were compartmentalized in aqueous droplets of water-in-oil (w/o) emulsions, thus maintaining a linkage between a plasmid-borne gene, the encoded enzyme variant, and the fluorescent product this enzyme may generate. Conversion into a double, water-in-oil-in-water (w/o/w) emulsion enabled the sorting of these compartments by FACS, as well as the isolation of living bacteria cells and their enzyme-coding genes. We demonstrate the directed evolution of new enzyme variants by screening >10^7 serum paraoxonase (PON1) mutants, to yield 100-fold improvements in thiolactonase activity. In vitro compartmentalization (IVC) of single cells, each carrying >10^4 enzyme molecules, in a volume of <10 femtoliter (fl), enabled detection and selection despite the fast, spontaneous hydrolysis of the substrate, the very low initial thiolactonase activity of PON1, and the use of difusable fluorescent products.

Introduction

High-throughput screening (HTS) comprises the bottleneck of many research areas, including functional genomics and directed evolution. The first step in such experiments is the creation of gene libraries, of either cDNA isolated from natural sources, or by randomization techniques that create artificial libraries comprised of up to 10^{15} gene variants. But, the screening of this wealth of diversity is challenging, as a unique screen needs to be developed for each target enzyme. While a range of HTS approaches for binding interactions is available (e.g., two-hybrid systems [1], and cell- [2–7], phage- [8], and various in vitro display technologies [9–11]), selection for enzymatic activities has proven to be more challenging. To date, the vast majority of screens for enzymatic activity are based on assaying isolated bacterial colonies on agar, or in microtiter plates. Despite advances in automation, a comparatively small number of variants, typically in the range of 10^3–10^4, and rarely above 10^5, is screened in this way [12].

FACS technology holds great promise for HTS. Modern FACSs can routinely analyze and sort >10^7 events per hour, and fluorescence is a sensitive signal, widely adopted to detect both binding and enzymatic reactions. Indeed, FACS has gained wide success in screening gene libraries to yield highly potent binding proteins such as antibodies [2, 3, 5–7, 13–15]. The potential and utility of FACS for enzyme selections has also been demonstrated [12, 16]. However, this approach has been thus far limited to cases in which diffusion of the product of the enzymatic reaction out of the cell is restricted (e.g., [17]), or cases in which the product can be entrapped on the cell surface [18, 19] or onto microbeads [20]. Likewise, single-cell phenotyping with antibody probes, or green fluorescent protein (GFP), is a standard method, but high-throughput phenotyping by enzymatic cellular markers is still a challenge. Thus, a means of compartmentalizing and sorting single genes, or cells, together with the product molecules generated by their encoded enzymes, would alleviate the need to capture the product, and thereby dramatically widen the scope of this technology. Sortable compartments have recently become available by the application of double w/o/w emulsions. A primary w/o emulsion, the aqueous droplets of which can be used to compartmentalize genes and proteins [21], was subjected to reemulsification to create a w/o/w emulsion, with a continuous, external aqueous phase. The w/o/w droplets could be sorted by FACS without compromising the content and integrity of the internal aqueous droplets [22].

Here, we describe the application of w/o/w emulsions for compartmentalization and high-throughput phenotyping of single cells, and the directed evolution of an enzyme. The target for evolution was serum paraoxonase (PON1), a mammalian enzyme that catalyzes the hydrolysis and inactivation of a broad range of substrates and has a profound impact on the onset and progression of atherosclerosis [23]. PON1, which resides on HDL plasma particles (the "good cholesterol"), was also found to hydrolyze homocysteine thiolactone (HcyT) and thereby reduce the levels of this toxic metabolite [24]. But, although PON1 is probably the only enzyme capable of hydrolyzing plasma HcyT [24], HcyT, and thiobutyrolactones (TBLs) in general, are poor substrates of PON1 ($k_{cat}/K_M \leq 100\ M^{-1}s^{-1}$). We aimed at improving the TBLase activity of PON1, and thereby providing a new potential mean of detoxification.

TBLs present a particular challenge for detection. A high background exists due to the spontaneous (nonenzymatic) hydrolysis of TBLs and the presence of thiols in the media in which the enzyme variants are expressed and screened. In addition, the signal by PON1 is very low due to its very poor catalytic efficiency with TBLs. The low signal-to-background ratio can be overcome by increasing the enzyme concentration [25] (see the Supplemental Note in the Supplemental Data available with this article online). We therefore opted to gain maximal sensitivity by compartmentalizing intact bacterial

*Correspondence: tawfik@weizmann.ac.il
[3]Present address: Department of Chemistry, University of British Columbia, Vancouver V6T 1Z4, Canada.

cells containing 10^4–10^5 enzyme molecules per cell, rather then cell-free translation that is traditionally used with IVC and yields 10–10^2 enzyme molecules per droplet [21, 26]. Compartmentalizing single cells resulted in very high enzyme concentrations within the aqueous droplets (>>1 μM; see the Supplemental Note) and enabled detection and selection despite the low signal-to-background ratio.

The overall selection scheme is depicted in Figure 1. A library of >10^6 different PON1 gene variants was transformed into *E. coli*, and individual cells, each expressing a different enzyme variant, were compartmentalized in the aqueous droplets of a w/o emulsion. After the addition of the TBL substrate together with the fluorescence thiol-detecting dye CPM [27], the emulsion was converted into a w/o/w emulsion and sorted by FACS. The phenotype (TBLase activity) was determined for each cell in the library, and cells exhibiting the highest activity were isolated. We describe the sensitivity and dynamic range of this screening system and the isolation of a wide range of improved variants with up to a 100-fold increase in catalytic proficiency for TBLs, including HcyT. We demonstrate the use of both cytoplasmic expression and surface display, which enables detection of even the wild-type's poor activity (k_{cat}/K_M = 75 $M^{-1}s^{-1}$). The application of cell-free translation in w/o/w emulsions for the directed evolution of new β-galactosidase variants is described in the accompanying manuscript [26].

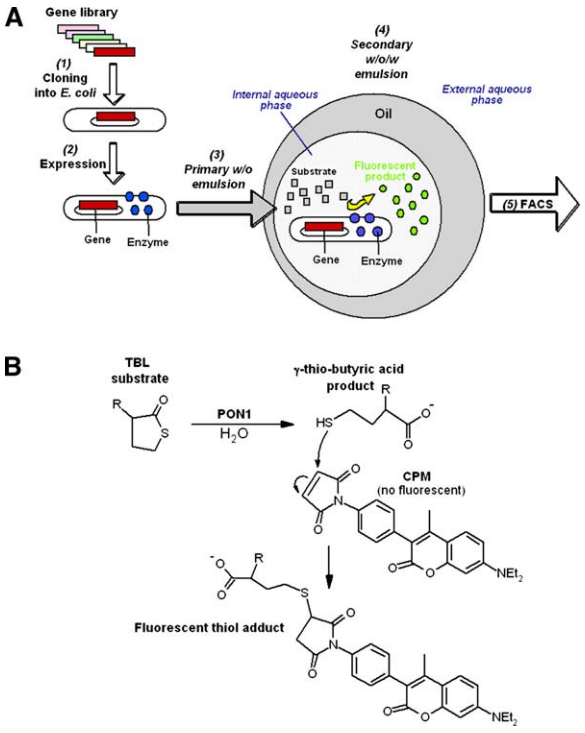

Figure 1. Single-Cell Compartmentalization and Selection by In Vitro Compartmentalization in w/o/w Emulsions

(A) (1) A gene library is transformed and cloned into *E. coli*, and (2) the encoded proteins are allowed to translate in the cytoplasm, or on the surface, of the bacteria cells. (3) Single cells are compartmentalized in the aqueous droplets of a w/o emulsion. (4) The fluorogenic substrate is added (through the oil phase), and the w/o/w emulsion is formed by emulsification of the primary w/o emulsion, enveloping the aqueous droplets with an intermediate layer of oil and providing an external aqueous phase. (5) Compartments containing the fluorescent product are sorted by FACS, and the cells imbedded in them, together with the gene encoding the enzyme of interest, are isolated.

(B) Detection of TBLase activity. Hydrolysis of γTBL (R = H) or HcyT ($R = NH_2$) releases a free thiol that reacts with the thiol-detecting reagent CPM to give a fluorescent dye product adduct [27].

Results

Compartmentalization, Detection, and Sorting of Single *E. coli* Cells

In the first instance, *E. coli* cells expressing the PON1 variants in their cytoplasms were emulsified to generate the primary w/o emulsion (Figure 1A; Step 3). Cell cultures were grown overnight, and ~5 × 10^8 cells were rinsed, resuspended in buffer, and emulsified in mineral oil containing the AbilEM90 surfactant. We found that emulsions based on this polymeric surfactant are better suited for the compartmentalization of living cells than the emulsion composition described originally [22], and that the diffusion and exchange of the fluorescent product between droplets was also slower (for an alternative emulsion composition, see [26]). The number of aqueous droplets in this emulsion (>10^{10}) was in large excess of the number of cells, rendering the vast majority of the droplets empty. However, the tendency of the *E. coli* cells to form aggregates resulted in some emulsion droplets containing multiple *E. coli* cells. We therefore introduced an internal marker by expressing GFP within the *E. coli* cells in addition to the selected enzyme. The w/o emulsion was reemulsified to generate the w/o/w double emulsion (Figure 1A; Step 4), in which the TBL substrate, the detecting dye CPM, and individual *E. coli* cells were cocompartmentalized in a continuous phase of water that is amenable to FACS. The FACS triggering threshold was set on GFP emission (530 nm), and an appropriate gate was chosen corresponding to the level of emission of single cells (Figure 2A). In this way, the sort completely ignored droplets with no cells, and it avoided the isolation of droplets containing more than one cell. This approach allowed for >10-fold higher enrichment factors, and 20 times faster sorting rates, than those obtained by triggering on the standard forward and side scatter parameter (droplet size).

Detection of the TBLase activity of the compartmentalized cells was via the UV fluorescence signal (450 nm) generated by the hydrolysis of γ-TBL to give γ-thiobutyric acid and coupling to the CPM probe (Figure 1B). To demonstrate the sensitivity and dynamic range of detection, we compare the recombinant wild-type (wt) PON1 [28] (k_{cat}/K_M = 75 s^{-1} M^{-1}) with variant 1E9, which was isolated from the library selections described below and exhibits ~100-fold higher TBLase activity. Cells expressing these two PON1 variants were separately emulsified and analyzed by FACS (Figure 2C). The GFP emission was used for triggering the sort and restricting it to droplets containing single cells (Gate R1, Figure 2A). Events were further gated by the forward and side scattering parameters to obtain the middle-sized droplets (as in Gate R2, Figure 2B) and maximal enrichment [22]. Both R1 and R2 gates were required to obtain maximal enrichment and purity.

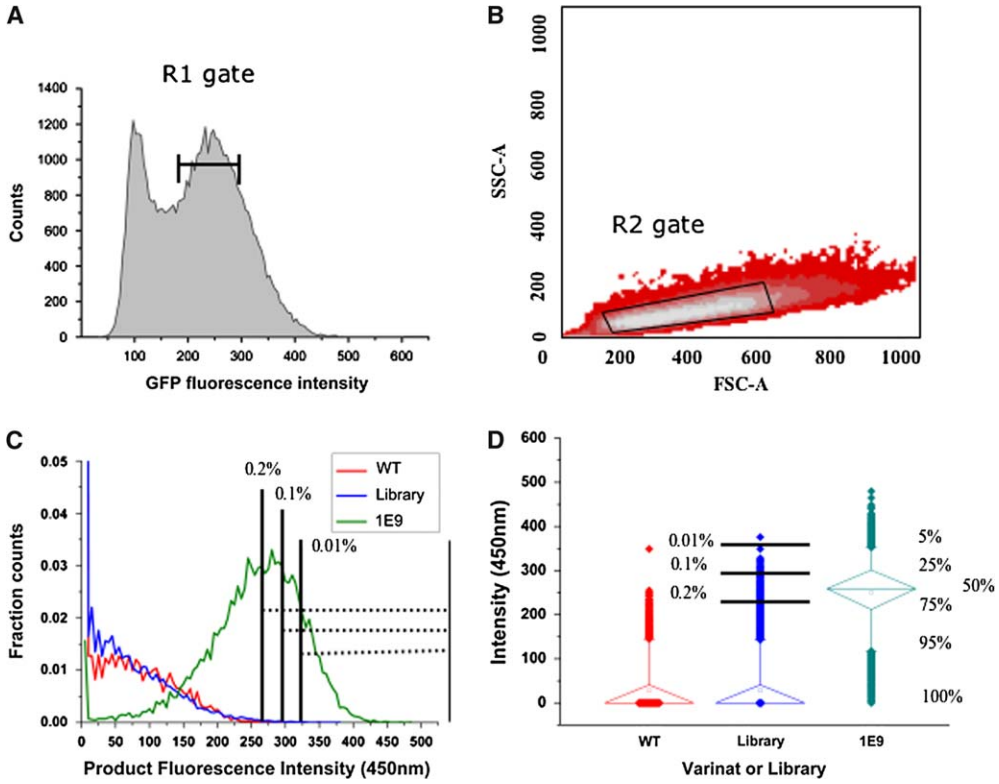

Figure 2. FACS Detection and Selection of the TBLase Activity of PON1-Carrying *E. coli* Cells in w/o/w Emulsion Droplets

Cells expressing in their cytoplasm a particular PON1 variant, together with the γTBL substrate and the thiol-detecting dye, were emulsified.

(A) For increased sorting rate and enrichment, cells were labeled by GFP expression. Shown is a representative histogram of the GFP emission for the entire population of droplets. Events gated in R1 (~50% of total events) correspond to single cells.

(B) Representative density plot FSC-H (forward scatter) and SSC-H (side scatter) analysis of the R1 population and the R2 gate (50% of the R1 events).

(C) Under these two gates (R1 + R2), comprising ~25% of the original population, droplets that exhibit high product-related fluorescence emission (450 nm) were sorted from the PON1 library. Different gates corresponding to different upper percentiles of TBLase activity within the library were chosen for sorting.

(D) Box chart of the UV emission exhibited by single cells expressing wt PON1, improved variant 1E9, and the gene library. The 25^{th}–75^{th} percentiles of single-cell emissions are found within the diamonds, the whiskers show the 5^{th}–95^{th} percentiles, and the events above and below the whiskers represent the upper and lower, respectively, 5% of the population. The average emission values exhibited by each cell population are marked by a circle. It can be seen that the resolution of the improved variant 1E9 over wt PON1 resulted in ~75% of 1E9 gene-carrying cells exhibiting UV emission levels at a range at which no wt gene-carrying cells emit. In contrast, a small upper percentage of library variants that fall within the range of TBLase activity exhibited by the variant 1E9 is clearly visible. Indeed, collection and growth of the 0.2% top percentile led to the isolation of variants exhibiting TBLase activity that is comparable to 1E9.

Droplets could then be categorized by the level of UV fluorescence (Figure 2C). Significant differences between the fluorescence intensities of these samples were observed in accordance with their enzymatic activities. The high amounts of PON1 (assays of the enzymatic activity in lysed cells indicated ~10^5 active PON1 molecules per cell) contained within the small volume of the emulsion droplets yield a local concentration of ~10 μM (Supplemental Note), which appears to allow for the detection of low enzymatic rates, as wt PON1 shows a modest separation from the inactive PON1-H115Q/H177R mutant (Figure S1C). The resolution of the improved variant 1E9 over wt PON1 resulted in ~75% of 1E9 gene-carrying cells exhibiting UV emission levels at a range at which no wt-gene-carrying cells emit (Figure 2D). Thus, the sensitivity of detection is high, and its dynamic range spans over more than two orders of magnitude. Longer incubation times (≤2 hr on ice) increased the UV emission signals of both wt PON1 and 1E9, but the ratio between the two remained largely unchanged. Diffusion of the product, and instability of the double emulsion, rendered longer incubations impossible.

We also performed spiking experiments (model selections) in which cells expressing the improved TBLase variant 1E9 (the isolation of which is described below) were mixed with a large excess (100- to 1000-fold) of cells carrying wt PON1. Cell mixtures were emulsified and analyzed by FACS by using the criteria of GFP emission, forward and side scattering, and UV fluorescence intensity, as described above. Multiple "positive" events (200–10^5) were collected into growth medium, and were then plated on agar. Isolated colonies were picked into the individual wells of 96-well plates and grown in liquid media, and the crude cell lysates were assayed for TBLase activity. Clones carrying 1E9 were easily distinguished from wt PON1 by virtue of exhibiting ~100-fold higher TBLase activity, allowing for the determination of

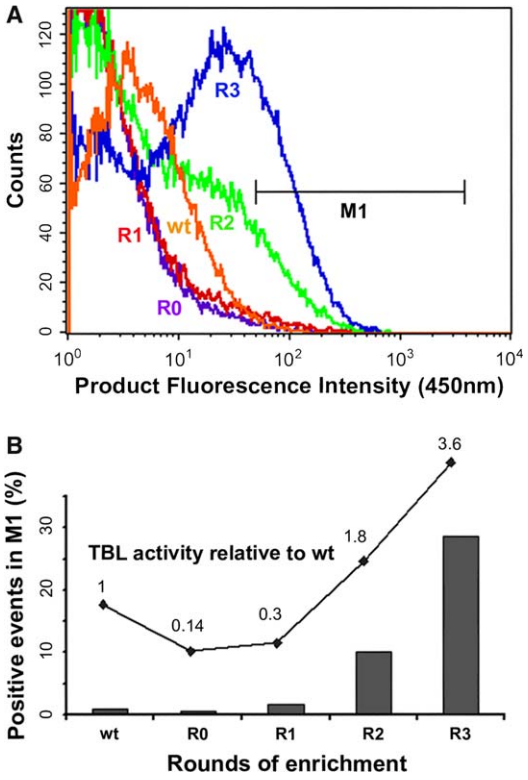

Figure 3. Library Selections by Three Iterative Rounds of Sorting by FACS

(A) FACS histogram analysis of the TBLase activity, detected by the 450 nm fluorescence intensity, observed in w/o/w emulsions prepared with *E. coli* cells expressing: wt PON1 (wt, orange), the unselected PON1 library (R0, purple), and the library after one (R1, red), two (R2, green), and three (R3, blue) rounds of FACS enrichment The library was sorted by using the criteria of GFP fluorescence (Figure 2A; R1 gate), size and shape (Figure 2B; R2 gate), and 450 nm fluorescence intensity corresponding to the TBLase activity (gate M1).

(B) Bar graph showing, for the various rounds of enrichment, the increase in the percentage of positive events (M1 gate), and TBLase activity measured in crude lysate of the selected pool, for the various rounds of enrichment. The TBLase activity was measured in lysates prepared from the pool of cells obtained after each round and was normalized to the activity exhibited by wt PON1 under the same conditions.

the ratio of 1E9 to wt PON1 clones in the selected pool. Alternatively, single positive events were sorted directly into individual wells of 96-well plates, grown, and assayed for TBLase activity. We observed enrichments of 40- up to 290-fold depending on the stringency of gating (Table S1). These correlated well with the FACS enrichment factors calculated from the number of "positive" events in the 1E9 versus wt samples. The recovery of living *E. coli* cells was quite good: ~50% of the positive events, including those emerging from sorting a single cell per well, yielded a colony.

Selection of PON1 Libraries for TBLase Activity

Directed evolution through random mutagenesis often identifies mutations far from the active site, perhaps due to the simple fact that active site residues comprise a comparatively small fraction of the entire protein [29]. We therefore decided to create libraries that specifically target active site residues that dictate substrate selectivity [29]. Structure and sequence analysis led to the classification of 16 residues that are located within and around PON1's active site and appear to have led to the divergence of the PON family in nature and the alteration of its substrate selectivity in directed evolution experiments (Table S2) [30]. However, a simultaneous diversification of all 16 positions would result in an impossibly high library size, and an extremely high mutation rate, rendering almost all library variants inactive. We therefore applied a protocol for spiking in of randomizing oligos, so that each library variant carries 3 mutated residues, on average, and the entire repertoire of 16 residues is explored in the complete library. Briefly, the PON1 gene was randomly digested to yield 50–125 bp fragments that were reassembled, as in DNA shuffling [31], in the presence of a mixture of 16 short oligos. Each oligo encoded one randomized codon and 3′ and 5′ flanking regions matching the wt PON1 gene. The concentration of the randomizing oligo mixture was adjusted to obtain the desired mutation rate (~3 randomized codons per gene).

The library plasmid DNA was transformed to BL21 (DE3) cells carrying the GFP expression vector. Approximately 5×10^8 cells, grown from 5×10^6 individual transformants, were emulsified, and $\sim 5 \times 10^7$ individual bacteria were analyzed by FACS (Figure 2). Positive events were sorted by using the criteria of GFP fluorescence, size, and shape, and UV fluorescence intensity corresponding to the TBLase activity. Two sorting strategies were applied. The first strategy aimed at high stringency and single round selections. Individual events falling within different upper percentiles of TBLase activity (0.2%–0.01%; Figures 2C and 2D) were sorted directly into the single wells of a microplate, grown in liquid medium, lysed, and assayed for TBLase activity. Single events were taken in total of 408, out of which 224 individual clones grew (55% recovery). The lysates derived from 20 out of these 224 clones showed >2-fold higher TBLase activity than wt PON1; the best ones exhibited >20-fold higher rates relative to wt PON1. The second strategy involved three iterative rounds of enrichment by FACS performed with gates that afford lower stringency and higher recovery. Positive events (2×10^4–10^5) were collected into one tube and plated on agar. The resulting colonies were pooled, and the plasmid DNA was extracted and transformed for a second round of sorting. Three rounds of sorting were performed, and, in each round, an increase in the number of positive events and the TBLase activity of the selected pool was observed (Figure 3). The plasmid DNA extracted from the third round of sorting was retransformed, and 360 colonies were picked and individually grown in 96-well plates. The cells were lysed, and the cleared lysates were assayed for TBLase activity. About a third of the clones exhibited significantly higher TBLase activity, and the best ones exhibited >20-fold higher rates relative to wt PON1.

Subsequent to the selections described above, crude lysates from all clones exhibiting high TBLase activity were assayed with a range of PON1 substrates, including γTBL, DEPCyC, and paraoxon (phosphotriesters) and 7AcC (an acetyl ester), with wt PON1 serving as

Table 1. The Evolved TBLases: Kinetic Parameters of Two Representative Variants

Variant[a]	Mutations	γTBL (γ-Thiobutyrolactone)			L-HcyT (L-Homocysteine Thiolactone)
		k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M ($M^{-1}s^{-1}$)	k_{cat}/K_M ($M^{-1}s^{-1}$)
rePON1 (wild-type)[b]	—	n.d.	>20	75	≈3
1E9	K192A, I291F, T332S	12.0	1.72	7.0×10^3 (**93**)[c]	83 (**27**)[c]
2B3	I74L, K192G, I291L, T332S	7.15	0.94	7.6×10^3 (**101**)[c]	190 (**61**)[c]

[a] All PON1 variants were fused to a thioredoxin (Trx) via a 6× histidine tag to facilitate expression and purification [28].
[b] rePON1 refers to recombinant PON1 variant G3C9 evolved for functional expression in *E. coli*; it is almost identical in sequence to wild-type rabbit PON1 (95% identity; 98% similarity) and exhibits essentially the same kinetic parameters as wild-type PON1 isolated from sera [28].
[c] Noted with boldface in parentheses is the fold increase in catalytic efficiency relative to wild-type PON1.

reference. It appeared that the selected variants exhibited considerable phenotypic diversity (representative results are included in Table S3). It is also noteworthy that no significant differences, either in phenotype or in sequence, were observed between variants isolated in the high-stringency, single round selections described in Figure 2C (e.g., variants B7, E5, G8, H10) and by iterative enrichments (Figure 3; e.g., variants 1E9 or 2B3). This suggests that the selection has largely exhausted the sequence space afforded by this library.

The variants exhibiting the highest TBLase activity were overexpressed in *E. coli*, purified, and analyzed in detail (Table 1). The improvements in TBLase catalytic efficiency (k_{cat}/K_M) were found to be in the range of 20- to 100-fold, for both γTBL and HcyT, as expected from their structural homology (Figure 1B).

Detection of Surface-Displayed PON1

To examine the generality of this methodology, and to open the road to more challenging selections, we have attempted the detection and sorting of surface-displayed enzymes. Various PON1 variants were displayed on the surface of *E. coli* by fusion to the outer membrane protein A (OmpA) [4]. The recombinant wt PON1 was displayed alongside the previously identified variant 1HT that exhibits 93-fold higher TBLase activity [30] and a heavily mutated PON1 library that exhibits almost no TBLase activity. The enzyme-displaying bacteria were compartmentalized and analyzed by FACS as described above. We observed excellent separation between the three variants (Figure 4). As is the case with cytoplasmic expression, the fluorescence signal of the surface-displayed variants was stable after several hours of storage of the emulsion on ice, and no mixing of product between the droplets was observed. These results indicate that enzyme detection is also possible when the activity takes place outside the cell, and that the diffusion of the product is restricted by compartmentalization in the droplets of the water-in-oil emulsion. This conclusion was further supported by the compartmentalization and ample detection of purified PON1 enzyme variants in buffer (data not shown). We observed, however, severe biased growth rates of cell lines displaying different PON1 variants on their surfaces, and takeover by rapid growth of clones that do not express PON1 at all (due to plasmids harboring a stop codon or the complete absence of the PON1 insert). This phenomenon of clonal competition was previously observed with cell display [32]. It resulted in very low enrichment factors (>10-fold lower than expected from the FACS enrichment factors described in Figure 4), and it prevented the selection of surface-displayed PON1 libraries.

Discussion

We have isolated new PON1 variants with TBLase catalytic efficiencies approaching 10^4 $M^{-1}s^{-1}$. While these are far from optimum, they comprise, to our knowledge, the most potent homocysteine thiolactonases known to date. One mutation, Thr332Ser, appeared in all selected clones (Table S3). This mutation is in a residue located ~6 Å from the catalytic calcium ion that lies at the very bottom of PON1's deep active site [30], and it appears to be the key for increasing TBLase activity. Mutations in Ile291 (also in the active site wall, and ~10 Å away from the calcium) to either Ala or Phe appear in most of the improved variants. Finally, the mutation of Lys192 into Gly or Arg is of interest, as natural polymorphism is observed in this residue that is related to susceptibility to organophosphates (OPs) and increased risk for atherosclerosis [33]. Interestingly, we have previously observed different mutations in both Thr332 and Ile291 in variants isolated by screening of PON1 libraries generated by error-prone PCR by using conventional colorimetric screens on agar and in 96-well plates. The two different mutations were initially observed in two separate clones (Thr332Ala, and Ile291Leu), and they were then combined by DNA shuffling to give a variant carrying both of these mutations (1HT). As shown here, selecting from a much larger repertoire enabled the *simultaneous* selection of mutations in both of these residues (332, 291) as well as other mutations (Table 1; Table S3), although the TBLase activity of these variants is very similar to that of 1HT ($k_{cat}/K_M = 7 \times 10^3$ $M^{-1}s^{-1}$; [30, 34]). It seems, therefore, that in this particular case, both focused mutagenesis of active site residues, and random mutations all along the gene followed by recombination, converged into a similar combination of mutations. Further improvements in PON1's TBLase activity may require changes in additional residues outside its active site that will be selected from completely random libraries.

Another notable outcome of the selection of a large repertoire of mutants is the isolation of a range of new phenotypes; for example, the paraoxonase activity of the various clones varied from a 2-fold increase (1H1) to a 9-fold decrease (2B3) despite their similar TBLase activity (Table S3). This diversity may enable the generation of PON1 variants with other tailor-made substrate

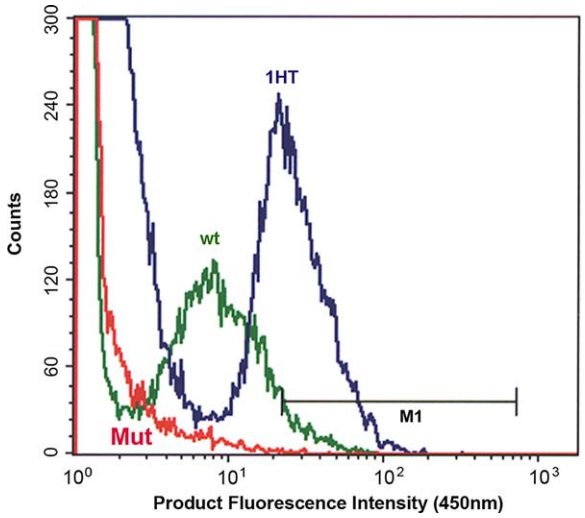

Figure 4. FACS Detection of the TBLase Activity of Surface-Displayed PON1 Variants Compartmentalized in w/o/w Double Emulsions

E. coli cells displaying different PON1 variants were separately emulsified and analyzed. Shown is a histogram for fluorescence at 450 nm corresponding to the thiol-derivatized dye, for a subpopulation gated by droplet size, as in region R2 of Figure 2B. Indicated are: a highly mutated PON1 gene library exhibiting essentially no TBLase activity (Mut, in red), wt PON1 (k_{cat}/K_M < 100 $M^{-1}s^{-1}$; in green), and a 93-fold improved variant 1HT (in blue) [30]. The percentage of "positive" events in M1 (out of R2) was found to be: 0.01% for the mutated PON1 library, 0.26% for wt PON1, and 2.3% for the improved 1HT variant. The calculated enrichment factor is therefore: 26-fold for enrichment of wt PON1 from the mutated library, and 230-fold for the enrichment of the 1HT variant.

specificities, as well as the further evolution of TBLase activity. This kind of phenotypic variation was not observed in many previous selections of PON1, performed on small libraries (10^3–10^4 variants) by traditional colony screens [30, 34], and is likely to be the result of screening a library of >10^6 variants. By directing the library to positions that dictate PON1's substrate selectivity, and tuning the rate of mutagenesis to ~3 mutations per gene, we avoided a high load of mutations that would have resulted in complete loss of activity, yet maintained the ability to select double and triple mutants in a single round of evolution.

Using w/o/w emulsions for compartmentalizing and sorting enzymatic activities represents a major advance. First, rather then tailoring each substrate and enzyme for a particular selection mode [3, 12, 16, 19, 20, 35], this technology allows for the use of a huge variety of existing fluorogenic enzyme substrates. Having special substrate/product modifications is not only a technical hurdle; one usually gets what one selects for [36], and the evolved enzyme is likely to be tailored for the modified substrate rather than the substrate of interest. In this respect, the CPM thiol-detecting dye is an ideal reporter for any reaction generating a free thiol (for oxo-lactone substrates that are detectable with CPM, see [37]). Second, analysis and sorting by FACS is a genuinely high-throughput method: our library selections of >10^7 bacteria required less than 2 hr. Unlike bulk selection methodologies such as panning on immobilized ligands, FACS allows for fine tuning of the selection threshold, enrichment, and recovery [2–7, 14–16, 18–20, 22, 26, 32, 38]. These capabilities are demonstrated here in the harvesting of the top 0.01% of TBLase activity from a library of >10^6 variants, and in isolating, in a single round of sorting, the most active variants. Other demonstrated advantages of FACS, for example in performing a parallel selection on two different substrates and exerting positive and negative selective pressures for the isolation of highly selective enzyme variants (i.e., with high activity toward the target substrate, and lower activity toward an undesired substrate, e.g., the original substrate) [18], could be harnessed by using compartmentalization in w/o/w emulsion.

Other methodologies have been developed for screening large enzyme libraries [34]. These are based on different display technologies (e.g., phage display), followed by isolation of active variants by binding transition state analogs and suicide inhibitors [35], or by covalent labeling with the product [39]. Another notable example is the display of enzyme variants on the surface of bacteria, followed by detection and sorting by FACS of a fluorescent, positively charged product captured onto the cell surface [18, 19]. Compartmentalization by w/o/w emulsions enables the selection of similar bacterial displayed libraries, as well as libraries cloned in *E. coli* for cytoplasmic expression, for rate enhancement, and for efficient turnover. The restriction of product diffusion by compartmentalization affords a sensitive and general mode of detection. Due to product diffusion, the same thiol-detecting dye, when diffused through a layer of soft agar, would not allow the detection of TBLase activity in colonies of wt PON1, and even the 100-fold improved variants could be barely seen. The same product was easily and quantitatively detected when compartmentalized in emulsion droplets, regardless of if the enzyme was expressed within the bacteria (Figure 2), or on their surface (Figure 4). The high local enzyme concentration afforded by the emulsion droplets increased the signal-to-background ratio (see the Supplemental Note) and provided a way by which very weak enzymatic activities can be detected against high backgrounds. Particularly notable is the ample detection of surface-displayed wt PON1 (Figure 4) despite its extremely low catalytic efficiency (k_{cat}/K_M = 75 $M^{-1}s^{-1}$). Surface-displayed PON1 variants could be detected not only with γTBL as substrate, but also with HcyT (data not shown), which is highly labile to hydrolysis ($t_{1/2}$ at pH 8 with no enzyme <1 hr). Although, due to biased growth rates, our enrichment factors for surface-displayed PON1 variants were poor, our results do indicate the potential utility of surface display for screening of substrates that do not penetrate through the *E. coli* membrane. Periplasmic expression is another alternative [3, 13], and so are cell-free expression systems, as described in the accompanying article [26]. The latter allow for the sampling of very large libraries while avoiding the complications that arise from selecting living organisms.

Finally, we have demonstrated the use of this technology in directed enzyme evolution, but its applicability to other areas, e.g., functional genomics and high-throughput single-cell phenotyping, is rather obvious. To this end, several aspects of the technology, such as droplet size homogeneity and consistency (having ideally, one aqueous droplet per oil droplet), may require improvement. Yet, the ability to compartmentalize intact

cells opens the road to single-cell analyses of large populations. This methodology could be used to directly monitor the actual levels of endogenous cellular enzymes rather than analyzing expression levels through a reporter protein such as GFP [38, 40], and of identifying rare phenotypes and genotypes with large cell populations [41]. Several different phenotypic variations could be detected simultaneously by using different fluorescent colors [18], and cells exhibiting a desirable phenotype could be subsequently isolated and their phenotype and genotype analyzed in detail.

Significance

We present a new, to our knowledge, high-throughput screening approach (>10^7 variants per hour) that enables the isolation of single cells by using FACS. Contrary to other FACS-based methods, the fluorescent probe for selection need not be physically linked to the sorted cells. Instead, the enzyme-expressing cells were individually compartmentalized in the water droplets of a water-in-oil (w/o) emulsion that restricts the diffusion of the fluorescent product. Conversion into a double water-in-oil-in-water (w/o/w) emulsion enabled the sorting of the compartments by FACS, and the isolation of living bacterial cells and their enzyme-coding genes. The restriction of product diffusion by compartmentalization affords a sensitive and general mode of detection and selection, for both rate enhancement and efficient turnover. It enables the selection of enzyme libraries by using ordinary, soluble, and difusable fluorogenic substrates, and either cell-free translation [26] or cellular expression in the cytoplasm, periplasm, or on the surface. Compartmentalization of single cells in emulsion droplets also provides unusually high enzyme concentrations (>10^4 enzyme molecules in <10 femtoliter [fL]), thus enabling detection and selection at extremely low signal-to-noise ratios. Using this approach, we screened a library of >10^7 serum paraoxonase (PON1) mutants for improvement in thiolactonase activity. We isolated a range of new PON1 variants with TBLase catalytic efficiencies approaching 10^4 $M^{-1}s^{-1}$ (~100-fold higher than wild-type PON1). These probably comprise the most potent mean of enzymatic detoxification of thiolactones, including homocysteine thiolactonase, available to date.

Experimental Procedures

DNA Manipulations

The gene library was prepared from PON1 variant G3C9 amplified off its pET vector [28] by using primers pET-Nes1-bc and pET-Nes1-fo [28]. The resulting DNA was randomly digested with DNaseI to yield 50–125 bp fragments as previously described [28, 31]. The fragments were reassembled, as in DNA shuffling [31], in the presence of a mixture of 16 short oligos. Each oligo encoded one randomized codon and 3′ and 5′ flanking regions matching the wt PON1 gene (Table S2). The reaction mixture was assembled and amplified by nested PCR as described [28]. It was found that adding the 16 oligos at a final concentration of 4 fmol/μl (or 4 nM) each to 120 ng purified DNA fragments in a 50 μl PCR mixture resulted in an average of 3 mutated positions per gene. The assembled genes were ligated into a modified pET 32b(+) vector by using NcoI/NotI restriction sites [28] and were transformed by electroporation to *E. coli* SS320 cells to give ca. 1.3 × 10^6 individual transformants, from which the plasmid DNA was extracted and purified. For surface display, the PON1-G3C9 gene, variant 1HT, and a highly mutated PON1 library prepared by error-prone PCR [28] were recloned into the pIB vector [42] through the NcoI/NotI restriction sites.

Emulsification of *E. coli* Cells

BL21(DE3) cells were transformed with a plasmid encoding GFP under the *Arg* promoter (kindly provided by Uri Alon, Weizmann Institute) and the kanamycin-resistance gene. Electrocompetent cells of a subsequent clone were prepared. Plasmid DNA encoding PON1 variants was transformed to these cells, and the electroporated cells were used to directly inoculate (without plating on agar) growth in liquid medium. The transformation efficiency was determined (by plating diluted aliquots of the original electroporation mix on agar plates) to be >5 × 10^6. For surface display, the PON1-pIB plasmids were similarly transformed to *E. coli* TG1 cells. Cultures of both display and cytoplasmic-expressing cells were grown overnight with shaking at 30ºC in LB medium supplemented with ampicillin (100 μg/ml) and kanamycin (20 μg/ml). The cells were spun down at 2500 RCF for 5 min and were rinsed twice with ice-cold activity buffer (50 mM Tris [pH 8], 1 mM $CaCl_2$, 100 mM NaCl). Approximately 3 × 10^9 cells were resuspended in 0.8 ml activity buffer and passed through a 5 μm syringe filter prior to emulsification. The TBLase activity of the cell samples was determined with 10 μl rinsed cells, 0.5 mM DTNB, and 1 mM γTBL in 0.2 ml activity buffer, by measuring absorbance at 412 nm. Filtered cell suspensions (80 μl) were added immediately to 0.8 ml ice-cold oil mix of 2.9% AbilEM90 (Tego, Germany) in light mineral oil (Sigma). The two phases were homogenized on ice, in a 2 ml round-bottom cryotube (Corning), for 5 min at 9500 rpm by using the IKA T-25 homogenizer. The thiol-detecting dye N-(4-(7-diethylamino-4-methylcoumarin-3-yl)phenyl)maleimide (CPM; Molecular Probes) (20 mM in DMF) and γTBL (200 mM in DMSO) were added to this w/o emulsion, 4 μl each, to a final concentration of 50 μM and 0.5 mM, respectively. The second water phase was then added (0.8 ml 1.5% carboxy methyl cellulose [CMC, medium viscosity, Sigma] and 1% Triton-X102 [Sigma] in activity buffer), and the mixture was homogenized, on ice, for 3 min at 8000 rpm to give the double w/o/w emulsion.

FACS Analysis and Sorting

w/o/w emulsions were diluted ~200-fold in activity buffer, and they were run in the FACSAria flow cytometer (Becton-Dickinson) with PBS as sheath fluid. The threshold for event detection was set to forward scatter and GFP emission (except for analysis of surface-displayed PON1 in which a GFP marker was not used). The average sort rate was ~4000 events per second, when using a 70 μm nozzle, exciting argon ion (488 nm) and 405 nm lasers, and measuring emissions passing the 530 ± 20 nm (FITC) band-pass filter for GFP emission and the 450 nm (violet 1) filter for the product-dye emission. Single, unaggregated droplets carrying one *E. coli* cell were gated by using forward and side scatter and GFP emission criteria (Figures 2A and 2B). Cells were sorted into glass tubes or 96-well plates containing 200 μl LB medium. Pools of sorted positives were plated on LB agar plates containing ampicillin. Colonies were allowed to grow overnight at 37ºC, and the cells were removed from the agar plates. The TBLase activity of the pooled cells was determined as described above, and plasmid DNA was extracted and retransformed for further rounds of sorting. FACS data were processed by using the CellQuest Pro Software (Becton Dickinson).

Screening, Purification, and Analysis of PON1 Variants

Plasmid DNA extracted from the third round of FACS was transformed to Origami B (DE3) *E. coli* cells. Isolated colonies were picked from agar plates and grown in 96-well plates as described [28, 30]. The clarified cell lysates were assayed for TBLase activity with γTBL (1 mM, 5 μl lysate) and HcyT (1 mM, 40 μl lysate), phosphotriesterase activity (0.1 mM DEPCyC, 30 μl lysate), aryl esterase activity (0.12 mM 7AcC, 2 μl lysate), and lactonase activity (1 mM DHC, 0.5 μl lysate) as described [28, 30]. Selected variants were recloned into pET32(b) to facilitate protein expression and purification [28]. The enzymes were purified, and their kinetic parameters were determined as described [28, 30].

Supplemental Data

Supplemental data including two figures, three tables, and a note are available at http://www.chembiol.com/cgi/content/full/12/12/1281/DC1/.

Acknowledgments

We are very grateful to Eitan Ariel and Ayala Sharp for their devoted assistance with the FACS. Financial support from the Israel Science Foundation (ISF) through its Bikura grant, and the Israeli Ministry of Science (IMOS), are gratefully acknowledged. D.T. is the incumbent of the Elaine Blond Career Development Chair.

Received: July 8, 2005
Revised: September 20, 2005
Accepted: September 22, 2005
Published: December 16, 2005

References

1. Fields, S., and Song, O.K. (1989). A novel genetic system to detect protein protein interactions. Nature *340*, 245–246.
2. Feldhaus, M.J., Siegel, R.W., Opresko, L.K., Coleman, J.R., Feldhaus, J.M.W., Yeung, Y.A., Cochran, J.R., Heinzelman, P., Colby, D., Swers, J., et al. (2003). Flow-cytometric isolation of human antibodies from a nonimmune *Saccharomyces cerevisiae* surface display library. Nat. Biotechnol. *21*, 163–170.
3. Chen, G., Hayhurst, A., Thomas, J.G., Harvey, B.R., Iverson, B.L., and Georgiou, G. (2001). Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS). Nat. Biotechnol. *19*, 537–542.
4. Georgiou, G. (2001). Analysis of large libraries of protein mutants using flow cytometry. In Advances in Protein Chemistry, Volume 55, F.M. Richards, D.S. Eisenberg, P.S. Kim, and F.H. Arnold, eds. (San Diego: Academic Press), pp. 293–315.
5. Hayhurst, A., and Georgiou, G. (2001). High-throughput antibody isolation. Curr. Opin. Chem. Biol. *5*, 683–689.
6. Wittrup, K.D. (2001). Protein engineering by cell-surface display. Curr. Opin. Biotechnol. *12*, 395–399.
7. Wentzel, A., Christmann, A., Adams, T., and Kolmar, H. (2001). Display of passenger proteins on the surface of *Escherichia coli* K-12 by the enterohemorrhagic E-coli intimin EaeA. J. Bacteriol. *183*, 7273–7284.
8. Smith, G.P., and Petrenko, V.A. (1997). Phage display. Chem. Rev. *97*, 391–410.
9. Lipovsek, D., and Pluckthun, A. (2004). In-vitro protein evolution by ribosome display and mRNA display. J. Immunol. Methods *290*, 51–67.
10. Roberts, R.W. (1999). Totally in vitro protein selection using mRNA-protein fusions and ribosome display. Curr. Opin. Chem. Biol. *3*, 268–273.
11. Leemhuis, H., Stein, V., Griffiths, A.D., and Hollfelder, F. (2005). New genotype-phenotype linkages for directed evolution of functional proteins. Curr. Opin. Struct. Biol. *15*, 472–478.
12. Aharoni, A., Griffiths, A.D., and Tawfik, D.S. (2005). High-throughput screens and selections of enzyme-encoding genes. Curr. Opin. Chem. Biol. *9*, 210–216.
13. Harvey, B.R., Georgiou, G., Hayhurst, A., Jeong, K.J., Iverson, B.L., and Rogers, G.K. (2004). Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from *Escherichia coli*-expressed libraries. Proc. Natl. Acad. Sci. USA *101*, 9193–9198.
14. Shusta, E.V., Holler, P.D., Kieke, M.C., Kranz, D.M., and Wittrup, K.D. (2000). Directed evolution of a stable scaffold for T-cell receptor engineering. Nat. Biotechnol. *18*, 754–759.
15. Nguyen, A.W., and Daugherty, P.S. (2005). Evolutionary optimization of fluorescent proteins for intracellular FRET. Nat. Biotechnol. *23*, 355–360.
16. Becker, S., Schmoldt, H.-U., Adams, T.M., Wilhelm, S., and Kolmar, H. (2004). Ultra-high throughput screening based on cell-surface display and fluorescence-activated cell sorting for the identification of novel biocatalysts. Curr. Opin. Biotechnol. *15*, 323–329.
17. Kawarasaki, Y., Griswold, K.E., Stevenson, J.D., Selzer, T., Benkovic, S.J., Iverson, B.L., and Georgiou, G. (2003). Enhanced crossover SCRATCHY: construction and high-throughput screening of a combinatorial library containing multiple non-homologous crossovers. Nucleic Acids Res. *31*, e126.
18. Varadarajan, N., Gam, J., Olsen, M.J., Georgiou, G., and Iverson, B.L. (2005). Engineering of protease variants exhibiting high catalytic activity and exquisite substrate selectivity. Proc. Natl. Acad. Sci. USA *102*, 6855–6860.
19. Olsen, M.J., Stephens, D., Griffiths, D., Daugherty, P., Georgiou, G., and Iverson, B.L. (2000). Function-based isolation of novel enzymes from a large library. Nat. Biotechnol. *18*, 1071–1074.
20. Griffiths, A.D., and Tawfik, D.S. (2003). Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization. EMBO J. *22*, 24–35.
21. Tawfik, D.S., and Griffiths, A.D. (1998). Man-made cell-like compartments for molecular evolution. Nat. Biotechnol. *16*, 652–656.
22. Bernath, K., Hai, M.T., Mastrobattista, E., Griffiths, A.D., Magdassi, S., and Tawfik, D.S. (2004). In vitro compartmentalization by double emulsions: sorting and gene enrichment by fluorescence activated cell sorting. Anal. Biochem. *325*, 151–157.
23. Lusis, A.J. (2000). Atherosclerosis. Nature *407*, 233–241.
24. Jakubowski, H. (2000). Calcium-dependent human serum homocysteine thiolactone hydrolase - A protective mechanism against protein N-homocysteinylation. J. Biol. Chem. *275*, 3957–3962.
25. Griffiths, A.D., and Tawfik, D.S. (2000). Man-made enzymes - from design to in vitro compartmentalisation. Curr. Opin. Biotechnol. *11*, 338–353.
26. Mastrobattista, E., Abecassis, V., Chanudet, E., Treacy, P., Kelly, B.T., and Griffiths, A. (2005). High-throughput screening of enzyme libraries: in vitro evolution of β-galactosidase by fluorescence-activated sorting of double emulsions. Chem. Biol. *12*, this issue, 1291–1300.
27. Parvari, R., Pecht, I., and Soreq, H. (1983). A microfluorometric assay for cholinesterases suitable for multiple kinetic determinations of picomoles released thiocholine. Anal. Biochem. *133*, 450–456.
28. Aharoni, A., Gaidukov, L., Yagur, S., Toker, L., Silman, I., and Tawfik, D.S. (2004). Directed evolution of mammalian paraoxonases PON1 and PON3 for bacterial expression and catalytic specialization. Proc. Natl. Acad. Sci. USA *101*, 482–487.
29. Park, S., Morley, K.L., Horsman, G.P., Holmquist, M., Hult, K., and Kazlauskas, R.J. (2005). Focusing mutations into the *P. fluorescens* esterase binding site increases enantioselectivity more effectively than distant mutations. Chem. Biol. *12*, 45–54.
30. Harel, M., Aharoni, A., Gaidukov, L., Brumshtein, B., Khersonsky, O., Meged, R., Dvir, H., Ravelli, R.B.G., McCarthy, A., Toker, L., et al. (2004). Structure and evolution of the serum paraoxonase family of detoxifying and anti-atherosclerotic enzymes. Nat. Struct. Mol. Biol. *11*, 412–419.
31. Stemmer, W.P.C. (1994). Rapid evolution of a protein in-vitro by DNA shuffling. Nature *370*, 389–391.
32. Daugherty, P., Iverson, B., and Georgiou, G. (2000). Flow cytometric screening of cell-based libraries. J. Immunol. Methods *243*, 211–227.
33. Draganov, D.I., and La Du, B.N. (2004). Pharmacogenetics of paraoxonases: a brief review. Naunyn Schmiedebergs Arch. Pharmacol. *369*, 78–88.
34. Aharoni, A., Gaidukov, L., Khersonsky, O., Gould, M.S., Roodeveldt, C., and Tawfik, D.S. (2005). The 'evolvability' of promiscuous protein functions. Nat. Genet. *37*, 73–77.
35. Fernandez-Gacio, A., Uguen, M., and Fastrez, J. (2003). Phage display as a tool for the directed evolution of enzymes. Trends Biotechnol. *21*, 408–414.
36. Schmidt-Dannert, C., and Arnold, F.H. (1999). Directed evolution of industrial enzymes. Trends Biotechnol. *17*, 135–136.
37. Khersonsky, O., and Tawfik, D.S. (2005). Chromogenic and fluorogenic assays for the lactonase activity of serum paraoxonases (PONs). ChemBioChem, in press.
38. Lahav, G., Rosenfeld, N., Sigal, A., Geva-Zatorsky, N., Levine, A.J., Elowitz, M.B., and Alon, U. (2004). Dynamics of the p53-Mdm2 feedback loop in individual cells. Nat. Genet. *36*, 147–150.

39. Cesaro-Tadic, S., Lagos, D., Honegger, A., Rickard, J.H., Partridge, L.J., Blackburn, G.M., and Pluckthun, A. (2003). Turnover-based in vitro selection and evolution of biocatalysts from a fully synthetic antibody library. Nat. Biotechnol. *21*, 679–685.
40. Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene expression in a single cell. Science *297*, 1183–1186.
41. Dressman, D., Yan, H., Traverso, G., Kinzler, K.W., and Vogelstein, B. (2003). Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. Proc. Natl. Acad. Sci. USA *100*, 8817–8822.
42. Benhar, I., Azriel, R., Nahary, L., Shaky, S., Berdichevsky, Y., Tamarkin, A., and Wels, W. (2000). Highly efficient selection of phage antibodies mediated by display of antigen as Lpp-OmpA′ fusions on live bacteria. J. Mol. Biol. *301*, 893–904.