Ultrahigh-throughput screening in drop-based microfluidics for directed evolution

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1073/pnas.0910781107">http://dx.doi.org/10.1073/pnas.0910781107</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>National Academy of Sciences (U.S.)</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Sun Dec 16 14:11:33 EST 2018</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/61376">http://hdl.handle.net/1721.1/61376</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
Ultrahigh-throughput screening in drop-based microfluidics for directed evolution

Jeremy J. Agresti, Eugene Antipov, Adam R. Abate, Keunho Ahn, Amy C. Rowat, Jean-Christophe Baret, Manuel Marquez, Alexander M. Klibanov, Andrew D. Griffiths, and David A. Weitz

†Harvard School of Engineering and Applied Sciences, Cambridge, MA 02138; ‡Fluid Discovery, Ltd., Emeryville, CA 94608; §Deartment of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139; ‡Department of Physics, Harvard University, Cambridge, MA 02138; ‡Institut de Science et d’Ingénierie Supramoléculaires, Université de Strasbourg, Centre National de la Recherche Scientifique Unité Mixte de Recherche 7006, 8 allée Gaspard Monge, BP 70028, F-67083 Strasbourg Cedex, France; ‡YNano LLC, 14148 Riverdowns South Drive, Midlothian, VA 23113; and §Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139

Edited by Harry L. Swinney, University of Texas at Austin, Austin, TX, and approved December 29, 2009 (received for review September 21, 2009)

The explosive growth in our knowledge of genomes, proteomes, and metabolomes is driving ever-increasing fundamental understanding of the biochemistry of life, enabling qualitatively new studies of complex biological systems and their evolution. This knowledge also drives modern biotechnologies, such as molecular engineering and synthetic biology, which have tremendous potential to address urgent problems, including developing potent new drugs and providing environmentally friendly energy. Many of these studies, however, are ultimately limited by their need for even-higher-throughput measurements of biochemical reactions. We present a general ultrahigh-throughput screening platform using drop-based microfluidics that overcomes these limitations and revolutionizes both the scale and speed of screening. We use aqueous drops dispersed in oil as picoliter-volume reaction vessels and screen them at rates of thousands per second. To demonstrate its power, we apply the system to directed evolution, identifying new mutants of the enzyme horseradish peroxidase exhibiting catalytic rates more than 10 times faster than their parent, which is already a very efficient enzyme. We exploit the ultrahigh throughput to use an initial purifying selection that removes inactive mutants; we identify ~100 variants comparable in activity to the parent from an initial population of ~10^6. After a second generation of mutagenesis and high-stringency screening, we identify several significantly improved mutants, some approaching diffusion-limited efficiency. In total, we screen ~10^9 individual enzyme reactions in only 10 h, using <150 μL of total reagent volume; compared to state-of-the-art robotic screening systems, we perform the entire assay with a 1,000-fold increase in speed and a 1-million-fold reduction in cost.

protein engineering | compartmentalization | emulsion | horseradish peroxidase

Much of the significant progress in our ever-advancing knowledge of biology is driven by our ability to perform very large numbers of reactions and to screen large amounts of data (1). This is certainly the case for the explosion in our knowledge of the genome and of proteomes and metabolomes. It is also the case for many important emerging biotechnologies, such as molecular engineering and synthetic biology. A particularly powerful tool for these technologies is directed evolution (2–4), which is a means of producing proteins and nucleic acids with tailor-made properties, and which has led to new classes of drugs (2), as well as improved enzymes (3) and strains of microorganisms for industrial applications (4). Directed evolution is also a powerful method to study the fundamentals of evolution itself (5). Evolution is practiced in the laboratory by artificially inducing mutations to create populations of new variants. These populations are subjected to a controlled pressure allowing selection of those rare variants which exhibit improved functionality. Evolutionary information is encoded in the genome, whereas selection acts on the phenotype; this requires the genotype and phenotype to be linked. However, to be truly effective, directed evolution requires screening or selection of very large populations, or libraries, using conditions that closely match the desired functionality. Numerous technologies can link genotype and phenotype for the selection of high-affinity binding; affinity purification can select libraries as large as 10^9 nucleic acids variants by the systematic evolution of ligands by exponential enrichment (SELEX) (6) or 10^12 protein variants by phage display (7). By contrast, it is much more challenging to select for other functionalities, such as catalysis (8). Cells also can be used to link genotype and phenotype; both growth-based selections and screening cells by FACS can potentially access libraries larger than 10^8, but the range of reactions is very limited, either to survival or a cell-bound product, respectively (3, 9, 10). Genotype and phenotype can also be linked by compartmentalizing individual elements of the library in the aqueous drops of a water-in-oil emulsion (9). Bulk emulsions (3, 9) allow screening of libraries as large as 10^11, but the lack of control of reaction volumes, timing, and generality ultimately limit their flexibility. To achieve full control and flexibility, current state-of-the-art high-throughput screening methods employ microtiter plates and sophisticated robots to process up to 100,000 assays a day, or ~1/s. This is a thousand times more than were processed in an entire week in 1990 (11). However, this technology is approaching its physical limit; below the 1-μL-volumes of 1,536-well plates, evaporation and capillary forces become prohibitive (11). New developments exploiting microfluidic lab-on-a-chip technology are significantly improving screening capabilities, decreasing the reaction volume by roughly 1,000-fold while increasing the speed by ~10-fold (12). Nevertheless, these screening methods still limit the applicability of directed evolution and many of the other emerging biotechnologies (13). Any further significant extension of screening capabilities will require revolutionary improvements, and these have, as yet, not been forthcoming.

In this paper, we report the development of an ultrahigh-throughput screening platform using drop-based microfluidic devices that provides the revolutionary improvements required to significantly advance our screening capabilities. We demonstrate...
the power of this platform by using it for directed evolution to
discover variants of the enzyme horseradish peroxidase (HRP).
We use picoliter-volume aqueous drops dispersed in an inert
oil as reaction vessels; most contain no more than a single yeast
cell that displays a variant of the enzyme on its surface. We sort
these drops at rates of thousands per second, allowing us to
screen a library of $10^8$ in about 10 h, using a total reagent volume
of $<150 \mu L$. This enables us to discover variants of HRP that are
more than 10-fold faster than the parent and that approach the
diffusion-limited rate. Compared to state-of-the-art robotic
screening, this is 1,000-fold faster and uses 10-million-fold less
volume of reagent, representing a cost savings estimated to be
about 10-million-fold.

**Results and Discussion**

Our ultrahigh-throughput-screening platform consists of two
polydimethylsiloxane (PDMS) microfluidic devices (14)
connected by an incubation line (Fig. 1). In the first device, we coflow
a suspension of yeast cells displaying enzyme variants on their
surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-
duce drops with a surface (10) with a second stream containing a fluorogenic sub-
strate for the enzyme. We use a long and narrow channel to pro-

mutants (Fig. 2). The rationale of this low-stringency approach is that the resulting variants of the first generation may contain potentiating mutations, which do not show improved phenotype on their own, but instead may poise the enzyme for improved function upon acquiring other, functionally adaptive mutations (22, 23). Mutations that improve catalysis may be deleterious to other functions, such as structural stability and, to be effective, require the preexistence of stabilizing mutations to compensate (21).

To perform each round of screening, we induce the yeast to express the surface-displayed HRP, encapsulate the cells in drops, and sort drops at 2,000 s for up to 3 h, screening as many as 2 × 10^5 cells. Before round one of the screen, we create two libraries each containing >10^7 mutants of the HRP gene. Because mutations both near the active site and distant from it are known to affect enzyme function (24), we create two types of libraries: We use random mutagenesis by error-prone PCR (epPCR) to target residues throughout the protein. Mutagenesis by epPCR will tend to bias mutations distant from the active site (24), and thus we create a second library using saturation mutagenesis to target specific residues at the entrance to the active site where the substrate binds (Protein Data Bank code 1HCH) (25). No additional mutations are added between rounds. We assay the enrichment of the pool of selected yeast after each round (SI Text). Although there is very little detectable activity in the unselected libraries, by the fourth round, the selected population exhibits higher activity than that of the parent by a factor of 2 (∼8 times) (Fig. 3).

To identify the amino acid changes in the selected variants, we sequence a total of 82 individual mutants and find 50 unique sequences. All of them are full-length genes with no internal stop codons. There are between one and five amino acid substitutions per mutant gene, and all encode active enzymes with catalytic rates similar to that of the parent (Fig. 3 and SI Text). From the distributions of the numbers of identical sequences (26), we estimate there are a total of ∼200 unique active sequences in the starting library of 2 × 10^7 (Fig. 3), thus only one in 10^5 of the original mutants is active (17).

The results of this purifying selection can be used to find mutants with enhanced catalytic activity. To accomplish this, we further mutate the 18 highest-activity clones from the first generation to create a new library again containing ∼10^7 variants. We screen as before, but sort the brightest 5—10% of the drops, which contain the enzymes with the highest catalytic activity. The activity of the selected population reaches a plateau of ∼8 times that of the parent by the fifth round of screening (Fig. 3). We then plate the cells, measure the activity of 184 mutants, and sequence the top 50%. Of these 92 sequences, only 31 are unique (SI Text); because many sequences are identified multiple times, this indicates that the population is beginning to converge on a smaller set of highly active sequences. In contrast to the variants selected in the first generation of evolution, the second-generation mutants are all at least five times faster than the wild type, with the best being 12 times faster (Fig. 3 and SI Text). An analysis of the turnover number, kcat, and the Michaelis constant, KM, shows that all of the top 13 mutants from the second generation have significantly increased catalytic efficiency (kcat/KM). All are more than twice as efficient as the wild type, and the best mutant is nearly seven times more efficient, with kcat/KM ∼ 2.5 × 10^7 M^−1 s^−1; this is approaching the maximum rate allowed by diffusion-limited encounters of ∼10^8 M^−1 s^−1 (SI Text). Even though the screen was performed at a substrate concentration of 100 μM, near the wild-type KM, we find that the KM increases for all mutants; however, this is compensated for by even larger increases in kcat, of up to 11-fold, yielding the higher catalytic efficiencies observed.

Fourteen out of 30 amino acid substitutions found in mutants from the neutral selection are inherited in mutants selected from...
the second-generation libraries. In addition, there are 27 new substitutions in the second-generation sequences, giving a total of 41 substitutions. Every second-generation mutant inherits at least one substitution from the first generation, as expected, none of the first-generation sequences occur exactly in the second generation, because a much higher threshold is used for sorting. Of the 31 unique sequences, 23 contain at least one new substitution, while the remaining eight contain only recombinations of the first-generation substitutions. Interestingly, some substitutions that give only moderate increases in catalytic rates in first-generation mutants are also found in many of the fastest second-generation mutants. Seven of the 14 inherited substitutions cluster around sites of known structural importance, such as the two structural calcium ions and Phe221, which are known to be involved in heme organization (19) (Fig. 3A). This supports the hypothesis that many of the neutral substitutions from the first generation act as potentiating mutations when either recombined with each other or with new substitutions in the second generation. Because only one in $10^5$ mutants were active in the first round, obtaining such a large reservoir would be impossible using a robotic screen, where the maximum throughput is ~$10^5$ samples per day. Indeed, a plate-based directed evolution study to improve HRP (27) finds only a single non-wild-type active mutant in the first round of screening of ~$10^4$ reactions, in accord with these observations. The effectiveness of the large reservoir of potentiating mutations in bringing about adaptive change underscores the advantage of the ultrahigh-throughput microfluidic screening platform.

We quantify the advantages of the drop-based microfluidic platform by comparing requirements for the full screen to a conservative estimate for those of a robot (Table 1). A reasonable estimate for the throughput of the robot gives a total time for the screen of nearly 2 years; by comparison, the microfluidic device requires only 5 h for the full screen. This is well over a 1,000-fold reduction. Similarly, using a reaction volume of 100 μL per assay with the robot, the total volume of reagent is 5,000 L; by comparison, the microfluidic device uses only 150 μL of reagents. This is more than a 10-million-fold reduction. Including all supplies and amortization, the total cost for screen with the robot would be ~$1.5 million; by comparison, the cost for the microfluidic screen is under $4. This is a 4-million-fold reduction.

Table 1. Comparison of time and costs* for the complete screen using traditional methods and in microfluidic emulsions

<table>
<thead>
<tr>
<th></th>
<th>Robot</th>
<th>Microfluidic drops</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total reactions</td>
<td>$5 \times 10^7$</td>
<td>$5 \times 10^7$</td>
</tr>
<tr>
<td>Reaction volume</td>
<td>100 μL</td>
<td>6 μL</td>
</tr>
<tr>
<td>Total volume</td>
<td>5,000 L</td>
<td>150 μL</td>
</tr>
<tr>
<td>Reactions/day</td>
<td>73,000</td>
<td>$1 \times 10^8$</td>
</tr>
<tr>
<td>Total time</td>
<td>~2 years</td>
<td>~7 h</td>
</tr>
<tr>
<td>Number of plates/devices</td>
<td>260,000</td>
<td>2</td>
</tr>
<tr>
<td>Cost of plates/devices</td>
<td>$520,000$</td>
<td>$1,000$</td>
</tr>
<tr>
<td>Cost of tips</td>
<td>$10 million</td>
<td>$0.30</td>
</tr>
<tr>
<td>Amortized cost of instruments</td>
<td>$280,000$</td>
<td>$1.70</td>
</tr>
<tr>
<td>Substrate</td>
<td>$4.75 million</td>
<td>$0.25</td>
</tr>
<tr>
<td>Total cost</td>
<td>$15.81 million</td>
<td>$2.50</td>
</tr>
</tbody>
</table>

*Details in SI Text
The ability to screen libraries of \(>10^6\) in just a few hours at a cost of only a few dollars will be of enormous benefit for directed evolution. There has already been some success screening small libraries that yield only modest improvements, and then performing repeated rounds of mutation and screening (27). However, when selecting for the binding activity of proteins, a clear relationship between library size and the affinity of the selected proteins is observed experimentally (7): Using antibody V-genes from nonimmunized donors, small phage-antibody libraries of \(<10^6\) genes yield only antibodies with \(K_d \sim 10^{-6}\) affinities, whereas larger libraries of \(>10^{10}\) yield \(K_d \sim 10^{-9}\). Similar improvements in the catalytic efficiency of enzymes should be possible with the use of larger libraries hitherto impossible using traditional robotic screening systems.

The drop-based microfluidic platform described here represents a unique class of screening system. When used with cells, the system operates as a drop-based FACS in that it interrogates individual cells and sorts them based on the results. However, unlike a traditional FACS, the cells remain encapsulated in drops and the entire reaction vessel is assayed and sorted. Prior to sorting, drops can be fused (28) to add additional reagents or even other cells, further increasing the versatility. In addition, this methodology is not limited to reactions with cells (15) or even to biological reactions. In this sense, drop-based microfluidics has many of the merits of both screening using microtiter plates and FACS: The assays are compartmentalized but the compartments themselves are screened and sorted at high speed based on fluorescence. This has the potential to significantly enhance the utility and applicability of high-throughput screening, not only for directed evolution, but also in many other areas of biology and chemistry where large parameter spaces need to be explored quickly and affordably. Moreover, the improvements in the ultra-high-throughput capabilities of the drop-based microfluidics are somewhat analogous to those of microprocessors; where transistor density has increased by \(\sim 100,000\)-fold in the past few decades (29). These improvements have led to completely new and unexpected applications for microprocessors; so too should the improvements in screening technology lead to unexpected new applications.

Materials and Methods

Construction of HRP Libraries. For the first generation of evolution, we use two previously created libraries (16, 30). In the first, the HRP gene is randomly mutated using epPCR at a rate of one to three mutations per gene. In the second, we combinatorially mutate, using saturation mutagenesis, five residues near the mouth of the active site where the substrate binds. We target Phe68, Gly69, Asn72, Ser73, and Ala74 in one reaction and positions Asn137, Leu138, Ala140, Phe142, and Phe143 in a second reaction. We then combine these two saturation libraries. We transform the randomized DNA from the library into electrocompetent EBY-100 cells and grow the resulting \(\sim 10^7\) transformants under selective conditions.

For the second generation, we begin by using the staggered extension process (STEP) (31) to recombine fastest 18 first-generation mutants. All 18 plasmids were mixed in equimolar amounts and then 1.25 ng of this mixture was used as a template for the STEP reaction with 200 μM dNTP, 150 nM each of the primers for the epPCR libraries described above, and 2.5 U of Taq polymerase (New England Biolabs). The reactions were cycled 99 times at 94 °C for 30 s followed by 55 °C for 5 s yielding a sharp band of the expected size at \(\sim 1, 100\) bp on an agarose gel. We used the product of the STEP reaction as a template for epPCR using the Mutazyme II kit following the standard protocol.

Preparation of Yeast for Sorting. HRP libraries were grown, induced, washed twice with PBS containing 0.5% BSA and a mM Mann, and then filtered through a 10 um nylon mesh syringe filter (Millipore). The optical density of the cells was measured, and the cells were centrifuged and resuspended at the appropriate density for loading into drops in a PBS solution containing 0.2% BSA and 1 mM mannose. The final substrate solution was PBS with 100 μM AUR (Invitrogen) and 300 μM H$_2$O$_2$.

Fabrication of Microfluidic Devices. We make PDMS microfluidic devices using standard soft lithographic methods (14) using SU-8-on-Si wafer masters, and PDMS-on-glass devices. We design electrodes as channels in the PDMS device and fill them with a low melting point metal alloy by running the wire into the punched holes [InAlloy 19 (52In, 32.5Sn, 16.5Sb) 0.020 in. diameter]. We first treat the electrode channels with 3-mercaptopropyltrimethoxysilane (gelest) diluted 1:10 in acetonitrile. On a 70 °C hot plate, we flow the silane into the channel with a syringe and then immediately blow it out with air. We then flow the metal into the channels while the device is still on the hot plate. We make electrical connections using eight-pin terminal blocks (digkey) glued (foctite UV cured) to the surface of the device for strain relief.

Coflow Drop Maker. To form monodisperse aqueous drops in a fluorocarbon oil, we use a flow-focusing geometry (33) to form aqueous drops in a fluorinated oil (Fig. 1). We first coflow two aqueous streams, one containing enzyme-displaying cells and the other containing substrates for the enzyme, to the drop-making junction. The drop-making geometry and flow rates have both been optimized to produce a high volume fraction (67%) emulsion of 23 μm drops at a rate of 2 kHz. This achieved for one third of the flows (two aqueous and one oil) at 20 μL h$^{-1}$ through a 10 μm square nozzle. The incubation line is 360 μm i.d. polyetheretherketone tubing. We can change the length or diameter of the incubation line that connects the drop maker and the sorter to adjust the reaction time prior to sorting.

Sorting Device. After the incubation delay line, the close-packed drops then flow into a constriction where the flow is split, and from where the drops are uniform in size, they are spaced evenly and thus periodically reintegrated into the device. The drops then come to a T junction, where more oil is added to the drop stream (400 μL h$^{-1}$ for the drop-making flow rates described above) to space the drops when they enter the sorting junction. Drops flow to the asymmetric “Y” sorting junction, where they can take one of two paths. In the absence of sorting, drops preferentially flow to the keep channel, as it has a lower fluidic resistance. The relative fluidic resistance of the top and bottom branches is set by the length of the narrow part of the channels at the sorting junction (Fig. 1 and SI Text). The constriction in the top branch is half the length of the lower path. Therefore, at first approximation, this allows roughly two-thirds of the flow (and all of the drops) to go to the upper channel and one-third to go to the lower channel.

When a drop passes the laser, its fluorescence is collected by a microscope objective and focused on a photomultiplier tube (Hamamatsu) connected to a computer running a custom software program (NI LabView). We used five cycles at 1 kHz. The output of the drops is bright enough to cross the voltage threshold set in the program, the software sends several cycles of a 20 kHz single-ended square wave to the sorting electrodes after being amplified by a factor of 1,000 by a high-voltage amplifier (Trek). We vary the number of cycles and voltage depending on the flow rates, drop size, and drop rate. In this study (460 μL h$^{-1}$, 23 μm, 1 kHz), we used five cycles at 1 kV. The sorted drops move up the field gradient created by the electrodes by dielectrophoresis (18) and are pulled into the keep channel. Movie S2 and Movie S3 are recorded at 4,000 frames s$^{-1}$ and show the result of sorting a demonstration emulsion containing fluorescein (light drops) or bromophenol blue 1% by weight in water.

ACKNOWLEDGMENTS. We thank J. Blouwolff for electrode technique, D. Marquardt for help with optical design, L. Frenz for software consultation, D. Aubrey for modeling electric fields, and Darren Link, Jeff Branciforte, Michael Lamsa, and Howard Shapiro for discussions. This work was supported by the Human Frontier Science Program (HFSP) Grant RGP0004/2005-C102, the National Science Foundation Grants DMR-0602684 and DMR-0649865, the Harvard Materials Research Science and Engineering Center DMR-0824084, Centre National de la Recherche Scientifique, the Massachusetts Life Sciences Center, and the Agence National de la Recherche (ANR-05-BLAN-5937). J.-C. B. was supported by an European Molecular Biology Organization long-term fellowship and by the European Commission Framework Programme 6 MiFem Network. A.C.R. is an HFSP cross-disciplinary fellow.

Corrections and Editorial Expression of Concern

ENGINEERING

The authors note that due to a printer’s error, the panels in Fig. 2 were mislabeled. The corrected figure and its legend appear below.

Fig. 2. Schematic of the directed evolution experiment. (A) The wild-type HRP gene is encoded on a plasmid as a C-terminal fusion to the Aga2 gene to allow surface display, and expression is driven by an inducible GAL (10) promoter. (B) We create two libraries for each generation. Each library has \( \sim 10^7 \) variants. For the first generation, we use one error-prone PCR (epPCR), and one active-site-targeted saturation mutagenesis library. For the second generation, we make both a high- and a low-mutation-rate epPCR library after recombination of the fastest first-generation sequences (SI). (C) The libraries are transformed into yeast (strain EBY100). Upon induction with galactose, each cell displays on its surface \( \sim 10,000 \) copies of a single mutant HRP protein (\( \mu \text{HRP} \)). (D) The yeast and nonfluorescent substrate are coencapsulated into drops on the microfluidic platform (Fig. 1). In the first round of each generation, we maximize the number of mutants screened by using a higher loading, \( \sim 1 \) cell per drop. In subsequent rounds, to minimize coencapsulation and ensure the highest enrichment possible, we load cells at 0.3 cells per drop (17). (E) Active mutants convert the AUR (gray) to its fluorescent oxidation product (pink) in an incubation line (F), and then flow into the sorter (G), where the brightest drops are sorted. We break the emulsion to release the cells from the drops, allow the cells to replicate, and then repeat the growth, induction, and sorting process.

www.pnas.org/cgi/doi/10.1073/pnas.1002891107
MEDICAL SCIENCES

The authors note that all columns and error bars in their figures represent means and SDs.

www.pnas.org/cgi/doi/10.1073/pnas.1002415107

NEUROSCIENCE

The authors note that the author name Christine Fenema-Notestine should have appeared as Christine Fennema-Notestine. The corrected author line appears below. The online version has been corrected.

Dominic Hollanda,1, James B. Brewera,b, Donald J. Haglerc, Christine Fennema-Notestineb,c, Anders M. Dalea,b, and the Alzheimer’s Disease Neuroimaging Initiative2

www.pnas.org/cgi/doi/10.1073/pnas.1001505107

EDITORIAL EXPRESSION OF CONCERN. PNAS is publishing an Editorial Expression of Concern regarding the following two articles:


The editors wish to note that we have received a report from the University of Alabama at Birmingham (UAB) that has investigated allegations of falsified or fabricated protein crystallographic structures including PDB codes 1RID and 2A01, which were published in the PNAS papers noted above. The UAB committee has forwarded their findings to the US Office of Research Integrity (ORI). We are awaiting the findings of ORI to determine the appropriate next steps.

Randy Schekman
Editor-in-Chief

www.pnas.org/cgi/doi/10.1073/pnas.1003210107