PROTEIN ENGINEERING

Beating the odds

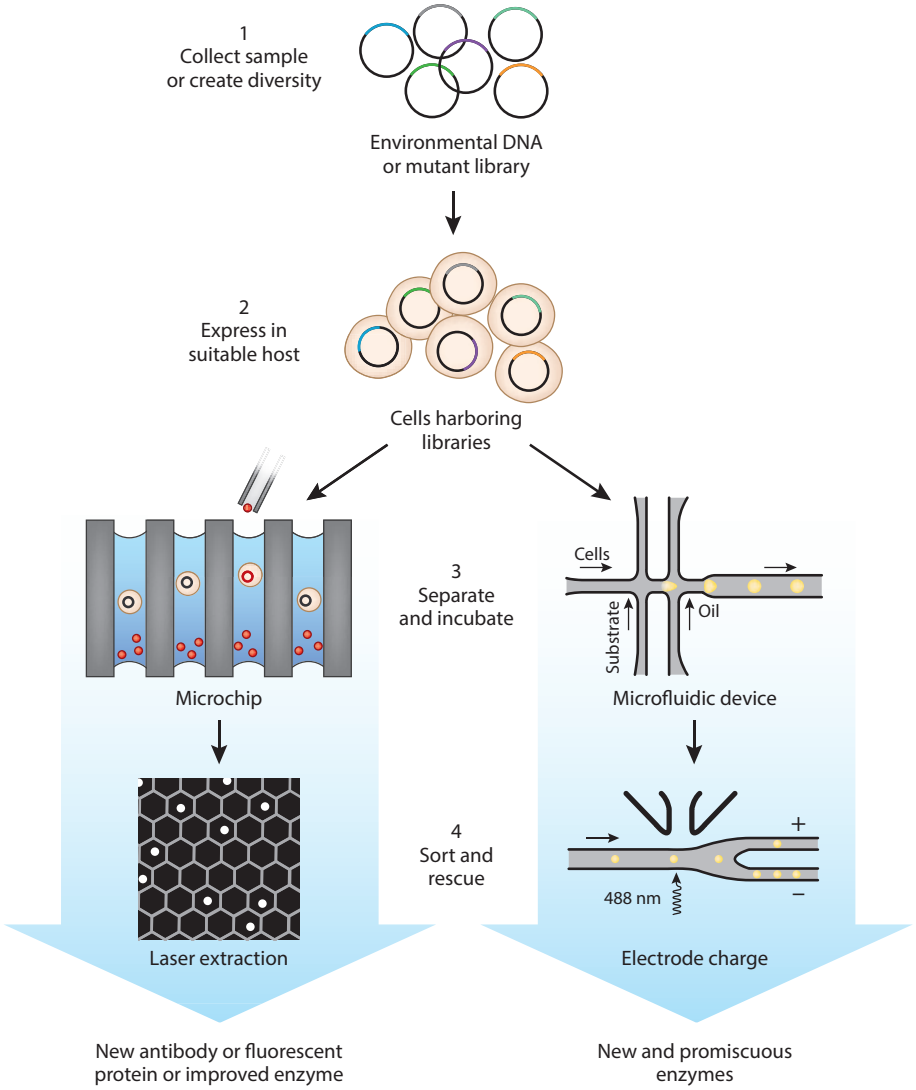
Recently developed advanced microfluidics-based systems have outperformed known screening tools with respect to throughput, flexibility, sensitivity and tricks for hit recovery. This has enabled the discovery of novel and improved proteins from random mutagenesis libraries or metagenome-based sources.

Uwe T Bornscheuer

Millions of novel protein sequences are currently being discovered, at an incredible pace. For instance, a recent article[1] described the identification of 8 complete and 789 draft genomes within a single metagenome library from an aquifer water sample. Unfortunately, the analysis of this huge pool of sequences remains challenging, given that the function of ~30–40% of the genes remains unknown and that initial annotations have often been proven to be incorrect. Similarly, protein engineering efforts using directed evolution create large mutant libraries (commonly in the range of 10^6–10^9 variants) that must be interrogated to identify desired improved proteins and, in particulaar, enzymes. Consequently, the availability of rapid, reliable screening methods for the functional characterization of natural and man-made diversity remains the primary bottleneck in this field.

Two publications now report novel tools and concepts (**Fig. 1**) to address this challenge, going significantly beyond the previous state of the art. In this issue, Chen *et al.*[2] developed a technology platform they dubbed "μSCALE." First, millions of protein variants produced by bacteria (*Escherichia coli*) or yeast (*Saccharomyces cerevisiae*) were spatially separated and mixed with magnetic microparticles in a microcapillary array. These variants were monitored by a camera, allowing up to 10,000 samples to be distinguished per second. The hits were then retrieved by a very smart and precise laser-based extraction system (**Fig. 1**, bottom left). This laser technique has the major advantage of isolating viable cells, allowing subsequent verification of initial hits through the necessary genotype-phenotype linkage. The authors demonstrated the usefulness of this system for three important areas of application for proteins: the identification of (i) a new antibody against a clinical target, (ii) a novel orange fluorescent protein for biosensor applications and (iii) an improved alkaline phosphatase with reduced inhibition.

Colin *et al.*[3] describe a microfluidics technology to interrogate a metagenome library (>10^6 variants) to identify novel enzymes with promiscuous activities, such that they can catalyze more than one distinct chemical reaction[4]. Their system is composed of monodisperse water-in-oil picoliter droplets, each, statistically, presumed to contain one *E. coli* cell



Kim Caesar/Nature Publishing Group

Figure 1 | Principle of the two microfluidic-based ultra-high-throughput systems. Top: biodiversity—either from metagenomic sources (environmental DNA) or random mutagenesis—is subjected to screening followed by isolation and characterization of hits. Bottom left: cells—up to 4 million individual *E. coli* cells per square inch—are spatially separated in a microarray and analyzed by imaging for the desired fluorescence readout, and finally hits are retrieved through the use of a unique laser-based extraction system[2]. Bottom right[3]: alternatively, monodisperse water-in-oil droplets (of picoliter volume)—each (statistically) containing an *E. coli* cell harboring a different enzyme variant—are incubated with the substrate of interest, and this is followed by rapid sorting of desired hits[3]. The numbers highlight the steps required. The author thanks D. Last (Institute of Biochemistry, Greifswald) for his help in designing the figure.

harboring a different library member. They screened these droplets for variants able to hydrolyze sulfate monoesters or phosphate triesters (**Fig. 1**, bottom right). In total, 20 million droplets could be assayed by fluorescence measurements and sorted within 2 hours. After further verification, the authors characterized 14 final hits biochemically, resulting in the identification of a set of completely novel hydrolytic enzymes. These enzymes have the potential to degrade xenobiotics, provide unprecedented promiscuity and span three enzyme superfamilies. This exploration of sequence space also provided detailed insights into the routes of evolution. A further charm of the method is that the substrate scope of most of the novel enzymes could have not been predicted by bioinformatic tools from their sequences. Overall, the microfluidic system was sensitive enough to identify enzymes with very low activities and low natural abundance as compared to common hydrolases.

The discovery of novel enzymes and the use of protein engineering represent a vibrant area of research, especially for biocatalysis. Enzymes catalyzing novel chemistry, as well as improved ones with optimal performance under process conditions, are highly desired, with an impressive number of success stories being reported[5]. Since the early 1990s, when directed evolution "in the test tube" started to become the standard tool in protein engineering, scientists have struggled to deal with the large sizes of mutant libraries that are generated. To address this, various concepts for ultra-high-throughput screening (uHTS) have been developed, such as *in vitro* compartmentalization[6], fluorescence-activated cell sorting for screening or selection[7] and, more recently, microfluidic systems[8]. A common limitation of these approaches is that the conditions under which these assays yielding fluorescent products are run are usually far from reality. Thus, the putative hits from uHTS may not show the expected performance when exposed to (physiologically or chemically) relevant substrates and conditions ("You get what you screen for"[9]). Both Chen *et al.* and Colin *et al.* now offer advantages over established uHTS methods in determining the binding, single-turnover and real-time kinetics of an enzyme to be measured, despite the necessity of using fluorescence readouts. However, confirmation of these hits—commonly accomplished through microtiter-plate assays followed by classical biochemical characterization of purified protein—is still required and dictates the overall time scale for protein discovery and engineering projects. Thus a potential approach to deal with massive uHTS using the recently described microfluidic tools is to reduce the library size and at the same time increase the number of well-performing initial hits. This can be achieved through the use of bioinformatic tools to analyze and filter enzyme superfamilies by, for example, correlated mutation analysis for better predictions of function and key mutations. Overall, the microfluidic systems described offer the potential to increase the success rate in protein discovery and engineering even further. ■

Uwe T. Bornscheuer is at the Institute of Biochemistry, Department of Biotechnology and Enzyme Catalysis, Greifswald University, Greifswald, Germany.
e-mail: uwe.bornscheuer@uni-greifswald.de

Published online 7 December 2015
doi:10.1038/nchembio.1989

References

1. Brown, C.T. *et al. Nature* **523**, 208–211 (2015).
2. Chen, B. *et al. Nat. Chem. Biol.* doi:10.1038/nchembio.1978 (7 December 2015).
3. Colin, P.Y. *et al. Nat. Commun.* doi:10.1038/ncomms10008 (7 December 2015).
4. Kazlauskas, R.J. & Bornscheuer, U.T. *Nat. Chem. Biol.* **5**, 526–529 (2009).
5. Bornscheuer, U.T. *et al. Nature* **485**, 185–194 (2012).
6. Griffiths, A.D. & Tawfik, D.S. *Nat. Biotechnol.* **16**, 652–656 (1998).
7. Fernández-Álvaro, E. *et al. Angew. Chem. Int. Edn. Engl.* **50**, 8584–8587 (2011).
8. Colin, P.Y., Zinchenko, A. & Hollfelder, F. *Curr. Opin. Struct. Biol.* **33**, 42–51 (2015).
9. Schmidt-Dannert, C. & Arnold, F.H. *Trends Biotechnol.* **17**, 135–136 (1999).

Competing financial interests

The author declares no competing financial interests.

SYNTHETIC BIOLOGY

Building genetic containment

Since the 1980s, scientists have worked on designing genetic codes to reinforce containment and control of genetically engineered microbes. New mechanistic studies of "deadman" and "passcode" gene circuits provide a flexible platform to build new safety switches.

Karmella A Haynes

Currently, it is unclear whether the unmonitored release of every genetically engineered microbe (GEM) into the environment is absolutely harmful or absolutely safe. Decades of scientific research and changes in government policies may be needed to determine and define the danger or safety of every type of GEM, as current information is sparse[1]. Many GEMs provide valuable immediate benefits to human society, but these benefits could be delayed unnecessarily by drawn-out vetting processes. Today, GEMs are being used in closed industrial environments to generate scents, flavors, fragrances, solid materials, fuels and pharmaceuticals on demand[2]. Biocontainment, the prevention of unintended environmental release, is an approach that allows the use of beneficial GEMs while avoiding uncertain risks. Physical containment within flasks and bioreactors and regulated disposal of decontaminated GEM cultures constitute the standard approach in industry and research lab settings. If the use of GEMs continues to scale up and these organisms are used in open settings in the future, external physical containment measures will no longer be sufficient. We will need genetically programmed controls that operate within the GEMs themselves. Chan *et al.* report the construction and validation of DNA-based genetic safeguards[3] that prevent escaped microbes from proliferating unchecked. Two customizable designs are composed of 'nuts and bolts' sub-parts that potentially allow scientists to tune the function of the gene circuits to suit different microbial strains and conditions.

An ideal genetic safeguard, sometimes referred to as a 'kill switch', simultaneously harbors a neutral state that allows billions of microbes to happily thrive and an activatable state that swiftly kills the entire population. As a second requirement, the engineered