Check for updates

Selection platforms for directed evolution in synthetic biology

Pedro A.G. Tizei*, **Eszter Csibra***, **Leticia Torres*** and **Vitor B. Pinheiro*†[1]**

*Department of Structural and Molecular Biology, University College London, Gower Street, London, WC1E 6BT, U.K.
†Birkbeck, Department of Biological Sciences, University of London, Malet Street, WC1E 7HX, U.K.

Abstract

Life on Earth is incredibly diverse. Yet, underneath that diversity, there are a number of constants and highly conserved processes: all life is based on DNA and RNA; the genetic code is universal; biology is limited to a small subset of potential chemistries. A vast amount of knowledge has been accrued through describing and characterizing enzymes, biological processes and organisms. Nevertheless, much remains to be understood about the natural world. One of the goals in Synthetic Biology is to recapitulate biological complexity from simple systems made from biological molecules – gaining a deeper understanding of life in the process. Directed evolution is a powerful tool in Synthetic Biology, able to bypass gaps in knowledge and capable of engineering even the most highly conserved biological processes. It encompasses a range of methodologies to create variation in a population and to select individual variants with the desired function – be it a ligand, enzyme, pathway or even whole organisms. Here, we present some of the basic frameworks that underpin all evolution platforms and review some of the recent contributions from directed evolution to synthetic biology, in particular methods that have been used to engineer the Central Dogma and the genetic code.

What is directed evolution?

Directed evolution is a well-established approach for optimizing and engineering novel functions in both nucleic acids and proteins. It has been remarkably successful in isolating novel ligands and catalysts based on the natural biopolymers and it is an essential tool for exploring the potential of xenobiotic polymers – probing the boundary conditions of life itself. Directed evolution is usually compared with Darwinian selection because of similarities in their underlying principles: genetic diversity leading to diversity of phenotype, a link in selection between phenotype and genotype recovery and amplification of the selected genotypes.

Those simple principles hide a myriad of methodological details and caveats that must be considered in directed evolution experiments, including: library quality, evolutionary landscapes and evolvability, sequence spaces and selection methodologies. The only reasonable conclusion is that a range of tools are required to allow for flexibility in the evolution starting point (e.g. *de novo* compared with starting in the functional vicinity of the function sought), objective (e.g. from increased thermostability to the expansion of the substrate range beyond natural compounds), scale (e.g. from selection of a ligand to the evolution of an entire organism) and system (e.g. *in vitro* to a eukaryotic host).

Compared with rational design, a key advantage of directed evolution lies in the impact of knowledge gaps (or uncertainty). For rational design to be most effective, an accurate and complete understanding of the target system (the system to be modified) is required. Incomplete or incorrect understanding of the target system leads to a high failure rate of designs. Although failed designs can be used to improve understanding of the system being designed and of the design tools themselves – establishing a design, test, build and learn cycle characteristic of synthetic biology (see [1–4] for general reviews of synthetic biology) – this can be a lengthy and costly process.

On the other hand, evolutionary approaches can, at least in principle, bypass any knowledge requirements. Directed evolution relies only on a cycle of introducing diversity into a population followed by the partitioning of that population to isolate the desired function. Theoretically, any population can be systematically optimized towards the desired function by repeating cycles of directed evolution. In practice however, the number of variants in the population can rapidly escalate beyond the sampling capacity of any selection methodology. In addition, as a given functional variant becomes rarer in a population, there is a greater burden on the selection method to isolate them. Consequently, all available knowledge of the target system is considered when designing the directed evolution strategy to minimize the number of variants and maximize its likelihood of success.

The field of directed evolution has greatly expanded in the last 20 years and it would be impossible to discuss or even acknowledge all of our colleagues who have contributed to it in this brief review. We apologize to those whose work we have not been able to include and we highlight other

Key words: directed evolution, protein engineering, selection methodologies, synthetic biology, synthetic nucleic acid (XNA), xenobiology.

Abbreviations: aaRS, aminoacyl tRNA synthetase; CFU, colony forming units; FACS, fluorescence-associated cell sorting; IVC, *in vitro* compartmentalization; ncaa, non-canonical amino acid; SELEX, systematic evolution of ligands by exponential enrichment; XNA, synthetic nucleic acid.

[1] **To whom correspondence should be addressed (email v.pinheiro@ucl.ac.uk).**

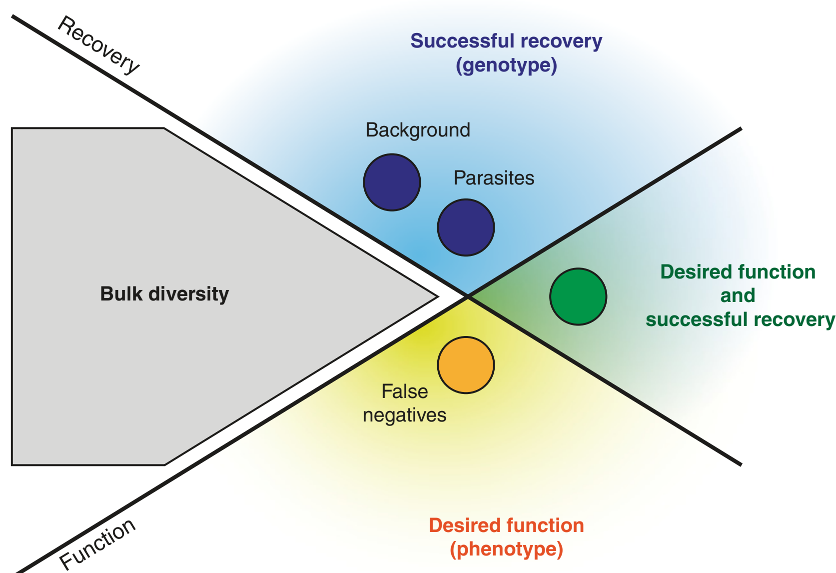
Biochem. Soc. Trans. (2016) 44, 1165–1175; doi:10.1042/BST20160076

© 2016 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution Licence 4.0 (CC BY).

Downloaded from http://portlandpress.com/biochemsoctrans/article-pdf/44/4/1165/486550/bst0441165.pdf by guest on 10 December 2022

Figure 1 | Principles of selection in directed evolution

The goal of all selection (and screening) platforms is to partition a potentially large population (shown in grey as the bulk diversity) by function (phenotype) ensuring the recovery of the genetic information that accounts for that phenotype. Strong phenotype–genotype linkages allow efficient isolation of mutants with the desired function (green). Breakdown of that linkage results in false negatives (variants that have the desired function but that are not efficiently recovered – yellow) and false positives (variants that are recovered independently of the desired function – blue), which are integral aspects of all selection strategies.



Downloaded from http://portlandpress.com/biochemsoctrans/article-pdf/44/4/1165/486550/bst0441165.pdf by guest on 10 December 2022

excellent recent reviews covering different aspects of directed evolution, including library design and diversity creation [5–8].

Here, we focus on presenting a sample of the diverse array of selection methodologies developed to date and how they may be used to alter the core of biology: changing the genetic code, changing genetic materials, changing the very chemistry of life, as well as establishing independent systems that can coexist with nature (orthogonal systems). Modification of those processes critically change biological function and have been described as unnatural molecular biology [4] or xenobiology [9,10]. Beyond improving our understanding of life and its origins, these modifications can be applied to novel biocontainment strategies [9–11], therapeutic agents [12,13] and protein engineering [14].

A strong phenotype–genotype linkage is at the heart of directed evolution

Partitioning of a diversified population is, as mentioned above, a key part of the evolutionary process and the point at which the 'fittest' are selected, where fitness is used as a quantitative description of the ability of a particular variant to perform the function being selected. This process of separating variants can be done by serially probing individual variants (screening) or by probing populations in parallel (selection). Both partitioning strategies use a measure of function (phenotype) to separate the population in a way that allows recovery of the genetic information (genotype) that encodes for the function – establishing the phenotype–genotype linkage. Breakdown of that linkage results in false negatives and false positives that undermine the evolutionary process, as illustrated in Figure 1.

False negatives are variants with the desired function that fail to be recovered during selection, undermining the evolutionary process. The loss of such variants can occur at any step along the selection process: whether through poor expression (in the case of a protein) or lower stability of an enzyme variant, due to errors or high variation in the quantification of function, or through stochastic recovery in selection.

False positives, or variants that are recovered but that do not display the desired phenotype, are the result of two distinct processes: one random and non-specific (background), the other the result of a viable alternative phenotype that, although not desired, can be efficiently recovered (parasites). Background is usually generated by the partitioning process itself in selection, where

© 2016 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution Licence 4.0 (CC BY).

non-specific interactions, e.g. DNA binding to nitrocellulose filters during aptamer selections [15], result in a sample of the population being recovered and taken forward to a subsequent selection round. By itself, background has little impact on the evolutionary process since further rounds of selection can be carried out until the desired function dominates the population. However, in methodologies prone to false negatives, the level of background recovery can have a significant impact on selection.

Parasites can inflict terminal damage to a directed evolution experiment by outperforming variants that display the desired phenotype – usually as a consequence of selection pressures rewarding phenotypes other than only the desired ones, or by parasitic variants outperforming the population during amplification (e.g. Spiegelman's Qβ serial dilution [16] or phage variants with high replication kinetics [17]). The impact of parasites in directed evolution is summarized by the maxim 'you get what you select for' that permeates the field [18].

The three populations (false negatives, parasites and background) are present in all selection and screening methods and this simple framework, shown in Figure 1, is a powerful tool to describe, analyse and design selection platforms – ensuring that selection pressures being used, maximize recovery of the desired mutants while taking steps to minimize or bypass the emergence of parasites.

Although a large number of selection platforms have been developed, they can be grouped into four categories: *in vivo*, *in vitro*, display systems (or *ex vivo*) and *in silico*. Each category has characteristic strengths and limitations that are the result of constraints imposed by where selection is carried out. *In silico* directed evolution, relying on computational tools to systematically generate, screen and optimize the biological function sought, is an emerging field with great potential but beyond the scope of this review [19–22].

The remaining three approaches have been remarkably successful at engineering individual biopolymers, gene circuits and genomes. In targeting core biological processes for modification, our knowledge and understanding of the systems being engineered is incomplete and limiting, making directed evolution an enabling technology and a powerful tool in the biologist's arsenal.

In vivo directed evolution

Directed evolution using *in vivo* selection platforms is possibly the strategy closest to natural evolutionary processes: the cell itself provides the physical link between genotype and phenotype, imposing the constraint that the cell must remain intact and metabolically active during key stages of selection (Figure 2). Because the entire cellular genome can contribute to a phenotype, *in vivo* selection strategies can be particularly efficient at evolving complex phenotypes (or those that require multiple steps to be observed) with diversity targeted to particular genes of interest or potentially to the entire organism genome.

Traditionally, genomic mutations have been introduced stochastically through mutagenic compounds, stressors or by inducing a higher mutation rate in the host [23]. However, some strategies are available to improve targeting and efficiency of the mutagenic process. The multiplex automated genome evolution (MAGE) process developed by Wang, Isaacs and colleagues exploits the ability of the phage lambda Red system [24] to facilitate recombination of single stranded DNA oligos with bacterial genomes [25]. It was originally validated by introducing variation to 24 genetic components involved in the heterologous production of lycopene, generating a library of over 10^9 variants partitioned by directly quantifying the synthesis of lycopene in colonies [25]. More recently, it has been used to systematically remove amber stop codons from the *Escherichia coli* genome [26,27], prior to re-introducing them to encode the incorporation of a non-canonical amino acid (ncaa) [26,27] – expanding the genetic code and creating a powerful approach towards the containment of genetically modified organisms.

Large-scale genome engineering can also be carried out by recombinases, which predictably and efficiently generate insertions, deletions and inversions. However, these enzymes usually require large (all longer than 30 base pairs) recognition sites that are naturally rare in the genome. Additional sites can be introduced in the genome, as is being done on an unprecedented scale in yeast [28–30], or by modifying recombinase target recognition [31,32].

As diversity is introduced in the system, selection begins and enrichment can be achieved by differences in cell survival, growth and replication rates or by partitioning the population based on the activity of a reporter gene (e.g. flow cytometric sorting using fluorescent reporters).

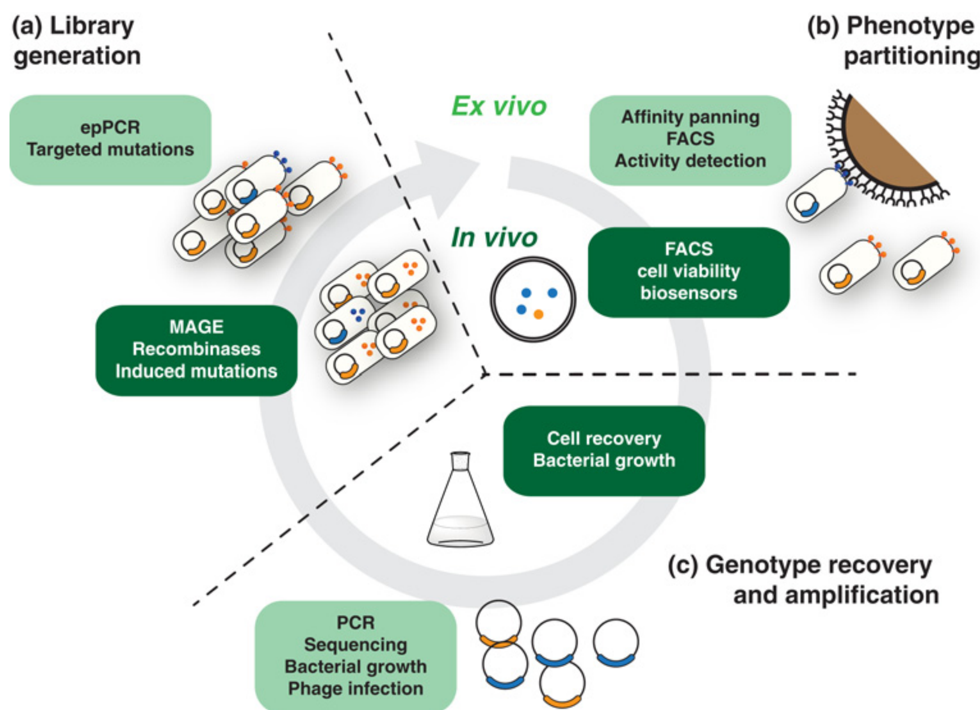
In vivo selection platforms have been successfully used in replacing natural amino acids for non-canonical ones, as well as in replacing natural nucleobases with unnatural ones in a bacterial genome. Serial passaging, in which a growing culture is diluted at an arbitrary cell density, has been successfully used by Budisa and colleagues to completely replace L-tryptophan with L-β-(thieno[3,2-b]pyrrolyl)alanine in the *E. coli* proteome [33]. Turbidostats, automated platforms that regularly adjust the rate of culture dilution based on culture density, have also been successfully used to adapt organisms to grow in substrates not naturally used in the starting organism. Using a dual turbidostat platform, Mutzel and colleagues [34] systematically evolved thymine auxotrophic *E. coli* strains that could thrive on chlorouracil, replacing thymine in its genetic material – a demonstration that genetic information storage itself can be changed if sufficiently small steps can be taken. The dual turbidostat effectively avoided biofilm-forming variants, a known parasite in continuous culture methods, but the method was still vulnerable to parasites that could synthesize thymine through uncharacterized salvage pathways, potentially reducing the selective pressure of the system. Indeed, Mutzel and colleagues identified a novel salvage pathway through tRNA U54 methyl transferase (*trmA*) that could provide up to 10 % of the genome's thymidine. Deletion of *trmA*, reduced the remaining genomic thymidine levels to below 1.5 %, the limit of detection in that particular approach.

© 2016 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution Licence 4.0 (CC BY).

Downloaded from http://portlandpress.com/biochemsoctrans/article-pdf/44/4/1165/486550/bst0441165.pdf by guest on 10 December 2022

Figure 2 | ***In vivo*** **and** ***ex vivo*** **directed evolution**

Both strategies use the cell (or phage particle) as the physical linkage between genotype and phenotype through the directed evolution process. *Ex vivo* platforms tend to focus diversity (**a**) on to a single target gene, whereas *in vivo* platforms can extend that to metabolic pathways or even whole genomes. Once generated, the diverse repertoires are partitioned (**b**) with active (blue) variants preferentially recovered over inactive variants (orange). Partition by phenotype is linked to genotype recovery and amplification (**c**) which can take place in a single step if cells are still viable (as is the norm for *in vivo* methodologies). Alternatively, as shown for the *ex vivo* selection (light green boxes), genotype recovery and amplification can be separated, introducing different limitations to the process. The amplified recovered genotypes are the starting point of a subsequent round of selection.



Another approach to reduce the emergence of parasites is by combining selection strategies that differ in their vulnerabilities to parasites, such as those commonly used in the directed evolution of orthogonal aminoacyl tRNA synthetases (aaRSs) [35,36]. Orthogonality between endogenous aaRSs/tRNAs and the aaRS/tRNA pair being introduced is essential for the efficient, site-specific incorporation of ncaa, and has been successfully obtained by a combination of positive and negative selection steps. Positive selection requires ncaa incorporation at an amber stop codon placed in a selectable marker gene (e.g. antibiotic resistance marker) to ensure survival. A second selection step, a negative selection, omits the ncaa and requires that no other amino acid is incorporated at the same codon in a gene coding for a toxic product (e.g. a nuclease), which would lead to cell death [36]. Parasites that may bypass the positive selection by allowing the misincorporation of any natural amino acid are penalized in the negative selection. Similarly, inactive variants that may become parasites of the negative selection are then penalized in the positive selection – thus increasing the power of the methodology.

Biological logic circuits can also be implemented to enhance *in vivo* selection platforms, either by monitoring the overall state of the cell [37] or by responding to one or multiple inputs relevant to the selection process. These circuits can be single component circuits as recently reported by Baker and colleagues for the *in vivo* selection of progesterone synthesis in yeast [38]. The progesterone biosensor was the fusion of an unstable domain to a reporter or selectable marker: progesterone binding stabilizes the protein allowing reporter expression – fulfilling its biosensor role.

A more complex circuit was developed by Chou and Keasling for the optimization of lycopene production [39]. By coupling detection of a lycopene biosynthetic intermediate with DNA polymerase III repression, the circuit could couple low lycopene production with the host mutation rate. Increasing levels of lycopene, detected by the red colour of isolated colonies was used to guide selection towards higher producers in repeated rounds of selection.

Logic circuits do not need to be directly linked to the phenotype under selection. Ellington and colleagues

Downloaded from http://portlandpress.com/biochemsoctrans/article-pdf/44/4/1165/486550/bst0441165.pdf by guest on 10 December 2022

© 2016 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution Licence 4.0 (CC BY).

[40] established a selection platform, compartmentalized partnered replication (CPR), based on a logic circuit in which the output is the expression of the *Thermus aquaticus* thermostable DNA polymerase (*Taq*). After selection, cells are compartmentalized and successful circuits use the expressed *Taq* to amplify the genes of interest in an emulsion PCR reminiscent of compartmentalized self-replication [41]. As with other techniques that use PCR to amplify selected information, directed evolution is confined to the region bound by the final PCR.

In vitro directed evolution

In vitro directed evolution is characterized by the use of selection platforms that either bypass living cells entirely (a fully *in vitro* system) or that rely on living hosts simply for the maintenance or heterologous expression of the biopolymers to be selected (a partially *in vitro* platform). It can be used to isolate biopolymers that would otherwise be toxic to a host or that function in conditions incompatible with biology, such as in the presence of denaturants, solvents or extreme temperatures. By carrying out selection outside biology, at least some of constraints of *in vivo* selection platforms can be bypassed, such as toxicity and bottlenecks in host transformation. Fully *in vitro* systems allow cells to be bypassed altogether, and with them the limit of transformation efficiency and recovery (typically $\sim 10^8$ CFU (colony forming units) per transformation), enabling libraries of 10^{14} to be achieved and increasing the available sequence search space.

In vitro selection allows us to explore and potentially recapitulate processes thought to have occurred early in life – abiotic processes in which a single molecule retains both genotype and phenotype – through the directed evolution of nucleic acid ligands (aptamers – recent review here [42]) and enzymes (NAzymes). Systematic evolution of ligands by exponential enrichment (SELEX) [43] is well established allowing not only directed evolution of functional RNA molecules [44–46], but also functional DNA [47–49] and synthetic nucleic acids (XNAs) [50–53]. Notably, SELEX is a potential misnomer because initial libraries are sufficiently large as to include all possible variants. Since it remains technically difficult to prove an isolated functional molecule was not present in the starting library, isolation of a functional nucleic acid can be the result of selection only - and not directed evolution since no evolution would be required (where possible that distinction is maintained in the text). In most cases, selection involves isolating functional nucleic acids, converting them to DNA (to allow efficient amplification by PCR), and the regeneration of the functional nucleic acid repertoire for further rounds of selection.

Variations of SELEX allow the RNA or XNA reverse transcription to DNA to be bypassed, enabling functional nucleic acids based on chemistries that are not viable genetic materials (or for which no efficient reverse transcriptase is known). In these DNA display methodologies, the genotype (encoded in the DNA) remains physically attached to the functional nucleic acid. Recovery of functional molecules therefore retrieves the encoding DNA – the important phenotype–genotype linkage.

Liu and colleagues have extended DNA display technology to enable selection of small molecules and peptide nucleic acids *in vitro* [54–56] and even clusters of carbohydrates [57,58]. In addition, DNA display methods have also been developed for the evolution of proteins, particularly nucleic acid processing enzymes where function can be linked to binding or modification of the relevant genotype [59–62]. This can be extended to other functions but generally rely on fusions that enable a covalent link between enzyme being tested and its genotype [63].

Genotype display, however, need not imply DNA, with alternatives developed for linking mRNA (genotype) to function (phenotype). In mRNA display platforms, a covalent link is made between the protein (selected for phenotype) and its encoding mRNA (genotype) during *in vitro* translation [64]. This platform has been widely used for the isolation of protein binders from *in vitro* translated protein libraries, including selecting for *de novo* functionality, such as a protein capable of ATP binding [65]. Seelig and Szostak [66,67] demonstrated that an RNA ligase could be obtained, by mRNA display, from a naïve library without prior knowledge of mechanism or sequence landscape – highlighting the power of directed evolution. mRNA display has also been combined with the incorporation of unnatural amino acids and click chemistry to create libraries of glycopeptides, containing multiple Man_9 glycans bound to a translated peptide. Variants capable of binding with high affinity to the HIV neutralizing antibody 2G12 have been efficiently selected by this method with possible relevance for HIV vaccine development [68].

It is also possible to link protein and mRNA non-covalently, such as in ribosome display – where a stalled ribosome works as an adaptor, linking mRNA to the displayed protein. Ribosome display has been successfully applied to the selection of ligands to different molecular targets [69–72], and more recently to the selection of catalysts [73–75].

An alternative strategy for *in vitro* selection is to co-isolate genotype and phenotype in cell-like compartments, termed *in vitro* compartmentalization (IVC), through the use of emulsions. Typically, water-in-oil emulsions do not allow significant exchange of components and hence are used to isolate genotype and phenotype inside individual compartments, ensuring a robust phenotype–genotype linkage [76]. Notably, compartmentalization can be achieved through other means, such as eutectic phases [77].

A number of emulsions have been explored for directed evolution with some stable even in extreme reaction conditions, such as the high temperatures required for emulsion PCR and for the selection of thermostable enzymes [41,76]. Emulsions can be made in bulk [76,78], resulting in a polydisperse emulsion (where compartments have a distribution of sizes), or in microfluidic devices where compartment size can be tightly regulated – a monodisperse

Downloaded from http://portlandpress.com/biochemsoctrans/article-pdf/44/4/1165/486550/bst0441165.pdf by guest on 10 December 2022

© 2016 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution Licence 4.0 (CC BY).

emulsion [79]. Compartment size variation can impose experimental constraints: bulk emulsions, requiring less specialized equipment, are easier and faster to make than an equivalent emulsion made with microfluidic systems. However, compartment size variation may affect platforms where selection is carried out near the K_D of the target enzyme, and where the signal generated in a compartment is used to partition the population, since reporter signals can depend on compartment size and concentration [80] – these can undermine platforms using bulk emulsion with higher rates of false negatives and false positives.

Emulsions have become the basis of several selection and screening strategies. Microfluidics platforms have been developed to introduce single cells (or beads) per compartment [81,82] with compartments being individually sorted [83–85], fused, split, stored and disrupted [86] – providing a more versatile range of methods than it is possible in bulk emulsions. For instance, Chaput and co-workers developed a microfluidic-based platform for the selection of polymerases that employs an optical detection of the enzymatic activity followed by fluorescence-associated cell sorting (FACS) enrichment [87].

Holliger and colleagues, in developing compartmentalized bead tagging (CBT) [88], demonstrated that it is possible to change the contents of the compartments in a bulk emulsion. In CBT, ribozyme genes were bound to paramagnetic beads and transcribed *in vitro* in a first emulsion. The transcribed ribozymes were ligated to the bead, allowing the first emulsion to be disrupted without breaking the phenotype–genotype linkage (between gene and ribozyme). Recovered beads could then be re-emulsified for ribozyme selection in a second bulk emulsion.

In addition to recapitulating processes from early biology, such as RNAzyme-based RNA replication, *in vitro* selection platforms have also been used to expand the Central Dogma. Holliger and colleagues, used an IVC selection strategy termed compartmentalized self-tagging (CST) to isolate thermostable DNA polymerase variants capable of synthesizing a number of different XNAs [50]. Together with a rationally designed reverse transcriptase, this demonstrated that the natural nucleic acids are not unique in being able to store genetic information. Although the polymerases isolated could synthesize XNAs, they retained DNA polymerase activity limiting their use towards introducing XNAs *in vivo*.

Compartmentalized systems have few but significant advantages over other *in vitro* selection platforms. The added compartmentalization minimizes cross-reactivity (or cross-catalysis) that can negatively affect *in vitro* platforms creating phenotype–genotype linkages on inactive variants. In addition, compartmentalization allows for changes in the topology of selection that make the platform more robust (Figure 3). For instance, in ribosome display a single mRNA molecule is linked to a single protein (a one-to-one mapping). Any RNA degradation or amino acid misincorporation destroys the phenotype–genotype linkage of that molecule (creating a false negative). On the other hand, in compartmentalized systems, multiple enzymes can act on multiple substrates to create a robust phenotype–genotype linkage (a many-to-many mapping) such that degradation of one substrate, catalyst or one molecule of the genotype does not undermine the phenotype–genotype linkage. Robustness in one-to-one platforms is achieved through redundancy in the initial library whereas a many-to-many selection strategy can achieve a significant level of robustness even in low redundancy repertoires.

Ex vivo directed evolution

Ex vivo selection, more commonly referred to as surface display, groups platforms and methodologies in which the biopolymer under selection (usually a protein) is accessible (i.e. outside) but still attached to a host, be it a bacteriophage or whole cell (Figure 2). Like *in vivo* platforms, a key advantage of *ex vivo* systems is that selection can be carried out using live hosts (or infective viable phages), bypassing costly intermediate steps and streamlining the evolution process. Crucially, *ex vivo* platforms bypass the challenge of bringing reagents (or targets) into the host, and extend the reaction conditions available for selection; though the latter is still limited to conditions that do not disrupt the host or the link between displayed biopolymer and host, both of which would undermine the phenotype–genotype linkage.

Phage display, in which gene fusions allow the display of a protein on the surface of a phage particle, is the earliest and by far the most successful *ex vivo* platform developed to date [89]. It has been extensively used in the development of antibody-based therapeutics [90] as well as for isolating a range of other ligands and enzymes [90–93].

Being a mature technology has allowed researchers to probe its shortcomings and biases in a series of carefully controlled experiments [17] showing, for instance, that abundance does not correlate with binding affinity in phage display selections. This can be rationalized by conceptualizing partitioning as the result of two selective processes: binding and amplification [17]. The latter is also effectively a selection step, and isolated variants with low amplification kinetics will be selectively lost (false negatives). A subsequent deep sequencing analysis quantified the diversity drop from a 10^6 library to enrichment of ~150 clones, which dominated 20 % of the selected library [94]. Some of those biases can be avoided by the use of emulsions or droplet PCR [95–97] or careful functional variant identification (via deep sequencing, [94]). Further, the use of bioinformatics analyses or selection databases [98–100] may allow the exclusion of parasites. Akin to the development of orthogonal aaRSs by multiple selections that differ in how they can be exploited by parasites, such problems can be circumvented by the use of different hosts (e.g. M13 compared with lambda) for each round of selection.

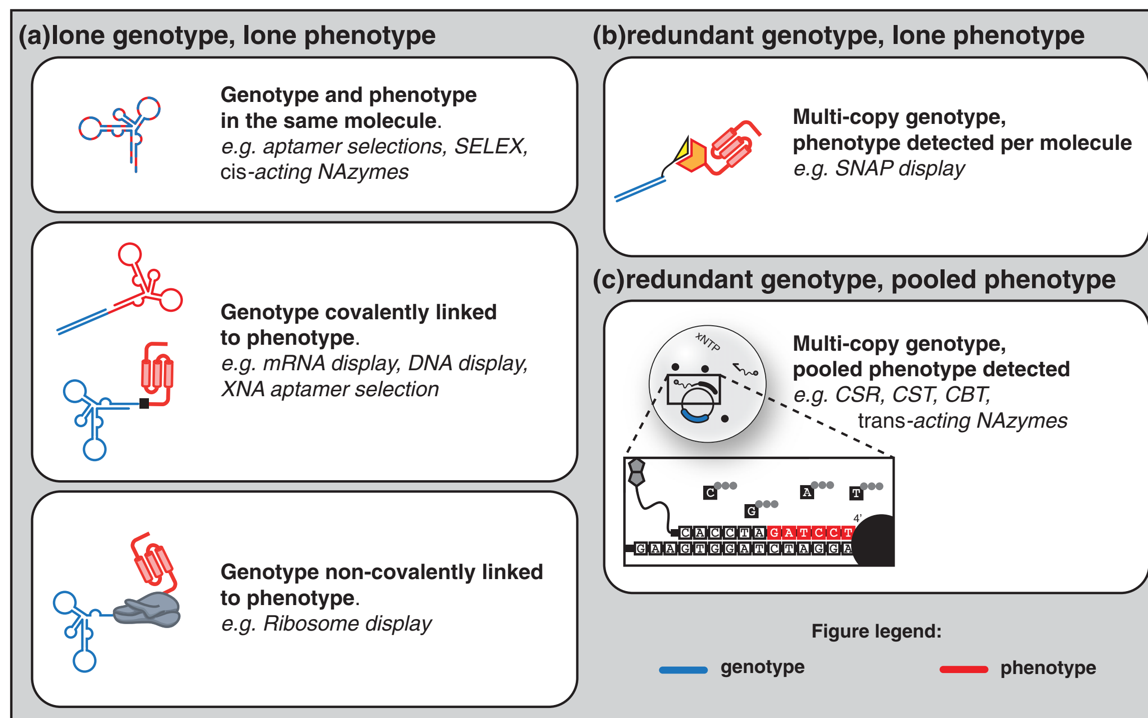
Currently, bar a few exceptions, surface display methods remain an under-exploited technology in the directed evolution toolbox for synthetic biology. For instance, phage display has been used to engineer thermostable DNA

© 2016 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution Licence 4.0 (CC BY).

Downloaded from http://portlandpress.com/biochemsoctrans/article-pdf/44/4/1165/486550/bst0441165.pdf by guest on 10 December 2022

Figure 3 | *In vitro* selection

Platforms for *in vitro* selection can be broadly divided by the available redundancy of phenotype and genotype linkages. In a number of selection strategies, the link is unique – a lone genotype molecule is linked to a lone molecule that may have the phenotype being selected (**a**). Compartmentalization strategies enable redundancy in the system with one-to-many [redundant genotype to lone phenotype (**b**) or lone genotype to pooled phenotype (not shown)] and many-to-many [redundant genotype to pooled phenotype (**c**)] mappings between phenotype and genotype available.



polymerases with extended substrate spectra [101–103], novel ligands using expanded genetic codes [104] and even to test the impact of the correlation between chemical diversity in the genetic code and fitness [105]. Key to those developments is the *ex vivo* localization of the biopolymers being selected, which enables their access to non-biological materials (e.g. xNTPs, oligonucleotides).

Ex vivo selection, however, extends much further than bacteriophages with display platforms established for Gram-negative and Gram-positive bacteria [106–108], as well as alternative platforms such as display on the surface of *Bacillus subtilis* spores [109]. Display platforms in eukaryotic cells, both yeast and cultured mammalian cells [110–112] have also been demonstrated and are of particular interest as they allow the incorporation of post-translational modifications in the passenger, particularly relevant to therapeutic antibody engineering.

Common to current *ex vivo* platforms is that display is achieved through gene fusion between a protein that naturally localizes to the surface of the host and the protein of interest (or passenger). This imposes some of the key constraints of the technology: the protein of interest and its fusion partner have to remain active once fused and be successfully exported to the host surface. For instance, export to the host surface is generally not an issue in display platforms based on lytic bacteriophages (e.g. T7), as the phage capsid is synthesized and assembled in the bacterial cytoplasm where most proteins fold efficiently. However, the fused host-selection protein cannot significantly affect the capsid assembly process or the function of the capsid in virulence; the latter is a constraint if phages are being recovered by infection of a susceptible host.

The use of cells, or other sizeable particles such as liposomes or other double emulsions [113,114] and beads [88,115,116], enables the partitioning of the population by FACS [90,117]. Although FACS is a high-throughput screening tool rather than selection, it provides an unparalleled level of flexibility, allowing display levels to be normalized and quantification of multiple parameters of a population. It also enables methods for coupling enzyme function to fluorescence. A particularly powerful method was developed recently in which hydrolase function on the cell surface is linked to the covalent attachment of biotinylated tyramide via horseradish peroxidase [118,119]. The success of this technique to quantitatively couple catalytic activity to fluorescent labelling for evolution [120] suggests that tools such as this one will be invaluable to allow the more widespread adoption of cell display for enzyme engineering for a wider range of functionalities [121].

Downloaded from http://portlandpress.com/biochemsoctrans/article-pdf/44/4/1165/486550/bst0441165.pdf by guest on 10 December 2022

© 2016 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution Licence 4.0 (CC BY).

Directed evolution as a tool for synthetic biology

Despite the diversity and versatility of selection platforms available, novel ones are regularly being developed – delivering custom solutions to ever growing challenges. As molecular biology methods and technologies develop, novel strategies to diversify and partition a biopolymer population become available, increasing experimental control, throughput and pace.

Directed evolution performs the design, build and test cycle of synthetic biology on a scale that is unnatural in engineering: it would be the equivalent of building millions (or even trillions) of slightly different machines (e.g. watches) in search of a specific improvement (e.g. more precise time keeping). On an engineering scale, such approach would be prohibitive, if even possible. However, on a biological scale, millions are still small numbers, barely able to cover the immediate sequence neighbourhood of even a small protein.

Directed evolution has successfully been used to isolate novel and optimize existing function on natural and synthetic biopolymers. But its key strength lies on how it deals with uncertainty. Even in the absence of complete understanding of complex biological systems, directed evolution is a powerful tool to re-engineer even the most central truths of life on our planet – that life is based on DNA and RNA, and that life requires (or is optimal with only) 20 amino acids.

Funding

This work was supported by the CAPES Foundation [grant number BEX 8985-13-8 (to P.A.G.T.)]; the European Research Council [grant number ERC-2013-StG project 336936 (HNAepisome) (to E.C., L.T. and V.B.P.)]; and the Biotechnology and Biological Sciences Research Council [grant number BB/K018132/1 (to V.B.P.)].

References

1 Way, J.C., Collins, J.J., Keasling, J.D. and Silver, P.A. (2014) Integrating biological redesign: where synthetic biology came from and where it needs to go. Cell **157**, 151–161 CrossRef PubMed

2 Purnick, P.E. and Weiss, R. (2009) The second wave of synthetic biology: from modules to systems. Nat. Rev. Mol. Cell Biol. **10**, 410–422 CrossRef PubMed

3 Cameron, D.E., Bashor, C.J. and Collins, J.J. (2014) A brief history of synthetic biology. Nat. Rev. Microbiol. **12**, 381–390 CrossRef PubMed

4 Deplazes, A. (2009) Piecing together a puzzle. An exposition of synthetic biology. EMBO Rep. **10**, 428–432 CrossRef PubMed

5 Lane, M.D. and Seelig, B. (2014) Advances in the directed evolution of proteins. Curr. Opin. Chem. Biol. **22**, 129–136 CrossRef PubMed

6 Dalby, P.A. (2014) Chapter 3: The role of directed protein evolution in synthetic biology. In (Ryadnov, M., Brunsveld, L. and Suga, H., eds), pp. 79–105, The Royal Society of Chemistry, London, Synthetic Biology: Volume 1 PubMed

7 Packer, M.S. and Liu, D.R. (2015) Methods for the directed evolution of proteins. Nat. Rev. Genet. **16**, 379–394 CrossRef PubMed

8 Badran, A.H. and Liu, D.R. (2015) *In vivo* continuous directed evolution. Curr. Opin. Chem. Biol. **24**, 1–10 CrossRef PubMed

9 Budisa, N. (2014) Xenobiology, new-to-nature synthetic cells and genetic firewall. Curr. Org. Chem. **18**, 936–943 CrossRef

10 Schmidt, M. (2010) Xenobiology: a new form of life as the ultimate biosafety tool. Bioessays **32**, 322–331 CrossRef PubMed

11 Schmidt, M. and de Lorenzo, V. (2016) Synthetic bugs on the loose: containment options for deeply engineered (micro)organisms. Curr. Opin. Biotechnol. **38**, 90–96 CrossRef PubMed

12 Lundin, K.E., Gissberg, O. and Smith, C.I. (2015) Oligonucleotide therapies: the past and the present. Hum. Gene. Ther. **26**, 475–485 CrossRef PubMed

13 Sharma, V.K., Rungta, P. and Prasad, A.K. (2014) Nucleic acid therapeutics: basic concepts and recent developments. RSC Adv. **4**, 16618–16631 CrossRef

14 Young, T.S. and Schultz, P.G. (2010) Beyond the canonical 20 amino acids: expanding the genetic lexicon. J. Biol. Chem. **285**, 11039–11044 CrossRef PubMed

15 Vant-Hull, B., Gold, L. and Zichi, D.A. (2000) Theoretical principles of *in vitro* selection using combinatorial nucleic acid libraries. Curr. Protoc. Nucleic Acid Chem. **Chapter 9**, Unit 9.1

16 Eigen, M. (1971) Selforganization of matter and the evolution of biological macromolecules. Naturwissenschaften **58**, 465–523 CrossRef PubMed

17 Derda, R., Tang, S.K., Li, S.C., Ng, S., Matochko, W. and Jafari, M.R. (2011) Diversity of phage-displayed libraries of peptides during panning and amplification. Molecules **16**, 1776–1803 CrossRef PubMed

18 Zhao, H. and Arnold, F.H. (1997) Combinatorial protein design: strategies for screening protein libraries. Curr. Opin. Struct. Biol. **7**, 480–485 CrossRef PubMed

19 Russ, W.P., Lowery, D.M., Mishra, P., Yaffe, M.B. and Ranganathan, R. (2005) Natural-like function in artificial WW domains. Nature **437**, 579–583 CrossRef PubMed

20 Damborsky, J. and Brezovsky, J. (2009) Computational tools for designing and engineering biocatalysts. Curr. Opin. Chem. Biol. **13**, 26–34 CrossRef PubMed

21 Siegel, J.B., Zanghellini, A., Lovick, H.M., Kiss, G., Lambert, A.R., St Clair, J.L., Gallaher, J.L., Hilvert, D., Gelb, M.H., Stoddard, B.L. et al. (2010) Computational design of an enzyme catalyst for a stereoselective bimolecular Diels–Alder reaction. Science **329**, 309–313 CrossRef PubMed

22 Penchovsky, R. (2014) Computational design of allosteric ribozymes as molecular biosensors. Biotechnol. Adv. **32**, 1015–1027 CrossRef PubMed

23 Yuan, L., Kurek, I., English, J. and Keenan, R. (2005) Laboratory-directed protein evolution. Microbiol. Mol. Biol. Rev. **69**, 373–392 CrossRef PubMed

24 Murphy, K.C. (1998) Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. J. Bacteriol. **180**, 2063–2071 PubMed

25 Wang, H.H., Isaacs, F.J., Carr, P.A., Sun, Z.Z., Xu, G., Forest, C.R. and Church, G.M. (2009) Programming cells by multiplex genome engineering and accelerated evolution. Nature **460**, 894–898 CrossRef PubMed

26 Mandell, D.J., Lajoie, M.J., Mee, M.T., Takeuchi, R., Kuznetsov, G., Norville, J.E., Gregg, C.J., Stoddard, B.L. and Church, G.M. (2015) Biocontainment of genetically modified organisms by synthetic protein design. Nature **518**, 55–60 CrossRef PubMed

27 Rovner, A.J., Haimovich, A.D., Katz, S.R., Li, Z., Grome, M.W., Gassaway, B.M., Amiram, M., Patel, J.R., Gallagher, R.R., Rinehart, J. and Isaacs, F.J. (2015) Recoded organisms engineered to depend on synthetic amino acids. Nature **518**, 89–93 CrossRef PubMed

28 Dymond, J.S., Richardson, S.M., Coombes, C.E., Babatz, T., Muller, H., Annaluru, N., Blake, W.J., Schwerzmann, J.W., Dai, J., Lindstrom, D.L. et al. (2011) Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. Nature **477**, 471–476 CrossRef PubMed

29 Jovicevic, D., Blount, B.A. and Ellis, T. (2014) Total synthesis of a eukaryotic chromosome: redesigning and SCRaMbLE-ing yeast. Bioessays **36**, 855–860 CrossRef PubMed

30 Shen, Y., Stracquadanio, G., Wang, Y., Yang, K., Mitchell, L.A., Xue, Y., Cai, Y., Chen, T., Dymond, J.S., Kang, K. et al. (2016) SCRaMbLE generates designed combinatorial stochastic diversity in synthetic chromosomes. Genome Res. **26**, 36–49 CrossRef PubMed

31 Gaj, T., Mercer, A.C., Sirk, S.J., Smith, H.L. and Barbas, 3rd, C.F. (2013) A comprehensive approach to zinc-finger recombinase customization enables genomic targeting in human cells. Nucleic Acids Res. **41**, 3937–3946 CrossRef PubMed

© 2016 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution Licence 4.0 (CC BY).

Downloaded from http://portlandpress.com/biochemsoctrans/article-pdf/44/4/1165/486550/bst0441165.pdf by guest on 10 December 2022

32 Sirk, S.J., Gaj, T., Jonsson, A., Mercer, A.C. and Barbas, 3rd, C.F. (2014) Expanding the zinc-finger recombinase repertoire: directed evolution and mutational analysis of serine recombinase specificity determinants. Nucleic Acids Res. **42**, 4755–4766 CrossRef PubMed
33 Hoesl, M.G., Oehm, S., Durkin, P., Darmon, E., Peil, L., Aerni, H.R., Rappsilber, J., Rinehart, J., Leach, D., Soll, D. and Budisa, N. (2015) Chemical evolution of a bacterial proteome. Angew. Chem. Int. Ed. Engl. **54**, 10030–10034 CrossRef PubMed
34 Marliere, P., Patrouix, J., Doring, V., Herdewijn, P., Tricot, S., Cruveiller, S., Bouzon, M. and Mutzel, R. (2011) Chemical evolution of a bacterium's genome. Angew. Chem. Int. Ed. Engl. **50**, 7109–7114 CrossRef PubMed
35 Santoro, S.W., Wang, L., Herberich, B., King, D.S. and Schultz, P.G. (2002) An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. Nat. Biotechnol. **20**, 1044–1048 CrossRef PubMed
36 Chin, J.W., Martin, A.B., King, D.S., Wang, L. and Schultz, P.G. (2002) Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **99**, 11020–11024 CrossRef PubMed
37 Ceroni, F., Algar, R., Stan, G.B. and Ellis, T. (2015) Quantifying cellular capacity identifies gene expression designs with reduced burden. Nat. Methods **12**, 415–418 CrossRef PubMed
38 Feng, J., Jester, B.W., Tinberg, C.E., Mandell, D.J., Antunes, M.S., Chari, R., Morey, K.J., Rios, X., Medford, J.I., Church, G.M. et al. (2015) A general strategy to construct small molecule biosensors in eukaryotes. Elife **4**, e10606 PubMed
39 Chou, H.H. and Keasling, J.D. (2013) Programming adaptive control to evolve increased metabolite production. Nat. Commun. **4**, 2595 CrossRef PubMed
40 Ellefson, J.W., Meyer, A.J., Hughes, R.A., Cannon, J.R., Brodbelt, J.S. and Ellington, A.D. (2014) Directed evolution of genetic parts and circuits by compartmentalized partnered replication. Nat. Biotechnol. **32**, 97–101 CrossRef PubMed
41 Ghadessy, F.J., Ong, J.L. and Holliger, P. (2001) Directed evolution of polymerase function by compartmentalized self-replication. Proc. Natl. Acad. Sci. U.S.A. **98**, 4552–4557 CrossRef PubMed
42 Blind, M. and Blank, M. (2015) Aptamer Selection Technology and Recent Advances. Mol. Ther. Nucleic Acids **4**, e223 CrossRef
43 Tuerk, C. and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science **249**, 505–510 CrossRef PubMed
44 Ellington, A.D. and Szostak, J.W. (1990) *In vitro* selection of RNA molecules that bind specific ligands. Nature **346**, 818–822 CrossRef PubMed
45 Robertson, D.L. and Joyce, G.F. (1990) Selection *in vitro* of an RNA enzyme that specifically cleaves single-stranded DNA. Nature **344**, 467–468 CrossRef PubMed
46 Bartel, D.P. and Szostak, J.W. (1993) Isolation of new ribozymes from a large pool of random sequences [see comment]. Science **261**, 1411–1418 CrossRef PubMed
47 Bock, L.C., Griffin, L.C., Latham, J.A., Vermaas, E.H. and Toole, J.J. (1992) Selection of single-stranded DNA molecules that bind and inhibit human thrombin. Nature **355**, 564–566 CrossRef PubMed
48 Breaker, R.R. and Joyce, G.F. (1994) A DNA enzyme that cleaves RNA. Chem. Biol. **1**, 223–229 CrossRef PubMed
49 Baum, D.A. and Silverman, S.K. (2008) Deoxyribozymes: useful DNA catalysts *in vitro* and *in vivo*. Cell. Mol. Life Sci. **65**, 2156–2174 CrossRef PubMed
50 Pinheiro, V.B., Taylor, A.I., Cozens, C., Abramov, M., Renders, M., Zhang, S., Chaput, J.C., Wengel, J., Peak-Chew, S.Y., McLaughlin, S.H. et al. (2012) Synthetic genetic polymers capable of heredity and evolution. Science **336**, 341–344 CrossRef PubMed
51 Taylor, A.I., Arangundy-Franklin, S. and Holliger, P. (2014) Towards applications of synthetic genetic polymers in diagnosis and therapy. Curr. Opin. Chem. Biol. **22**, 79–84 CrossRef PubMed
52 Sczepanski, J.T. and Joyce, G.F. (2014) A cross-chiral RNA polymerase ribozyme. Nature **515**, 440–442 CrossRef PubMed
53 Taylor, A.I., Pinheiro, V.B., Smola, M.J., Morgunov, A.S., Peak-Chew, S., Cozens, C., Weeks, K.M., Herdewijn, P. and Holliger, P. (2015) Catalysts from synthetic genetic polymers. Nature **518**, 427–430 CrossRef PubMed
54 Rosenbaum, D.M. and Liu, D.R. (2003) Efficient and sequence-specific DNA-templated polymerization of peptide nucleic acid aldehydes. J. Am. Chem. Soc. **125**, 13924–13925 CrossRef PubMed
55 Brudno, Y., Birnbaum, M.E., Kleiner, R.E. and Liu, D.R. (2010) An *in vitro* translation, selection and amplification system for peptide nucleic acids. Nat. Chem. Biol. **6**, 148–155 CrossRef PubMed
56 Yu, H., Zhang, S. and Chaput, J.C. (2012) Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. Nat. Chem. **4**, 183–187 CrossRef PubMed
57 Temme, J.S., Drzyzga, M.G., MacPherson, I.S. and Krauss, I.J. (2013) Directed evolution of 2G12-targeted nonamannose glycoclusters by SELMA. Chemistry **19**, 17291–17295 CrossRef PubMed
58 Temme, J.S., MacPherson, I.S., DeCourcey, J.F. and Krauss, I.J. (2014) High temperature SELMA: evolution of DNA-supported oligomannose clusters which are tightly recognized by HIV bnAb 2G12. J. Am. Chem. Soc. **136**, 1726–1729 CrossRef PubMed
59 Stein, V., Sielaff, I., Johnsson, K. and Hollfelder, F. (2007) A covalent chemical genotype–phenotype linkage for *in vitro* protein evolution. Chembiochem **8**, 2191–2194 CrossRef PubMed
60 Odegrip, R., Coomber, D., Eldridge, B., Hederer, R., Kuhlman, P.A., Ullman, C., FitzGerald, K. and McGregor, D. (2004) CIS display: *in vitro* selection of peptides from libraries of protein–DNA complexes. Proc. Natl. Acad. Sci. U.S.A. **101**, 2806–2810 CrossRef PubMed
61 Reiersen, H., Lobersli, I., Loset, G.A., Hvattum, E., Simonsen, B., Stacy, J.E., McGregor, D., Fitzgerald, K., Welschof, M., Brekke, O.H. and Marvik, O.J. (2005) Covalent antibody display – an *in vitro* antibody–DNA library selection system. Nucleic Acids Res. **33**, e10 CrossRef PubMed
62 Bertschinger, J. and Neri, D. (2004) Covalent DNA display as a novel tool for directed evolution of proteins *in vitro*. Protein Eng. Des. Sel. **17**, 699–707 CrossRef PubMed
63 Doi, N. and Yanagawa, H. (1999) STABLE: protein–DNA fusion system for screening of combinatorial protein libraries *in vitro*. FEBS Lett. **457**, 227–230 CrossRef PubMed
64 Liu, R., Barrick, J.E., Szostak, J.W. and Roberts, R.W. (2000) Optimized synthesis of RNA–protein fusions for *in vitro* protein selection. Methods Enzymol. **318**, 268–293 CrossRef PubMed
65 Keefe, A.D. and Szostak, J.W. (2001) Functional proteins from a random-sequence library. Nature **410**, 715–718 CrossRef PubMed
66 Seelig, B. and Szostak, J.W. (2007) Selection and evolution of enzymes from a partially randomized non-catalytic scaffold. Nature **448**, 828–831 CrossRef PubMed
67 Seelig, B. (2011) mRNA display for the selection and evolution of enzymes from *in vitro*-translated protein libraries. Nat. Protoc. **6**, 540–552 CrossRef PubMed
68 Horiya, S., Bailey, J.K., Temme, J.S., Guillen Schlippe, Y.V. and Krauss, I.J. (2014) Directed evolution of multivalent glycopeptides tightly recognized by HIV antibody 2G12. J. Am. Chem. Soc. **136**, 5407–5415 CrossRef PubMed
69 Hanes, J. and Pluckthun, A. (1997) *In vitro* selection and evolution of functional proteins by using ribosome display. Proc. Natl. Acad. Sci. U.S.A. **94**, 4937–4942 CrossRef PubMed
70 He, M. and Taussig, M.J. (1997) Antibody–ribosome–mRNA (ARM) complexes as efficient selection particles for *in vitro* display and evolution of antibody combining sites. Nucleic Acids Res. **25**, 5132–5134 CrossRef PubMed
71 Irving, R.A., Coia, G., Roberts, A., Nuttall, S.D. and Hudson, P.J. (2001) Ribosome display and affinity maturation: from antibodies to single V-domains and steps towards cancer therapeutics. J. Immunol. Methods **248**, 31–45 CrossRef PubMed
72 Matsuura, T. and Plückthun, A. (2003) Selection based on the folding properties of proteins with ribosome display. FEBS Lett. **539**, 24–28 CrossRef PubMed
73 Amstutz, P., Pelletier, J.N., Guggisberg, A., Jermutus, L., Cesaro-Tadic, S., Zahnd, C. and Pluckthun, A. (2002) *In vitro* selection for catalytic activity with ribosome display. J. Am. Chem. Soc. **124**, 9396–9403 CrossRef PubMed
74 Quinn, D.J., Cunningham, S., Walker, B. and Scott, C.J. (2008) Activity-based selection of a proteolytic species using ribosome display. Biochem. Biophys. Res. Commun. **370**, 77–81 CrossRef PubMed

Downloaded from http://portlandpress.com/biochemsoctrans/article-pdf/44/4/1165/486550/bst0441165.pdf by guest on 10 December 2022

© 2016 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution Licence 4.0 (CC BY).

75 Skirgaila, R., Pudzaitis, V., Paliksa, S., Vaitkevicius, M. and Janulaitis, A. (2013) Compartmentalization of destabilized enzyme–mRNA–ribosome complexes generated by ribosome display: a novel tool for the directed evolution of enzymes. Protein Eng. Des. Sel. **26**, 453–461 CrossRef PubMed
76 Tawfik, D.S. and Griffiths, A.D. (1998) Man-made cell-like compartments for molecular evolution. Nat. Biotechnol. **16**, 652–656 CrossRef PubMed
77 Attwater, J., Wochner, A. and Holliger, P. (2013) In-ice evolution of RNA polymerase ribozyme activity. Nat. Chem. **5**, 1011–1018 CrossRef PubMed
78 Diehl, F., Li, M., He, Y., Kinzler, K.W., Vogelstein, B. and Dressman, D. (2006) BEAMing: single-molecule PCR on microparticles in water-in-oil emulsions. Nat. Methods **3**, 551–559 CrossRef PubMed
79 Paegel, B.M. and Joyce, G.F. (2010) Microfluidic compartmentalized directed evolution. Chem. Biol. **17**, 717–724 CrossRef PubMed
80 Kaltenbach, M., Devenish, S.R. and Hollfelder, F. (2012) A simple method to evaluate the biochemical compatibility of oil/surfactant mixtures for experiments in microdroplets. Lab Chip **12**, 4185–4192 CrossRef PubMed
81 Huebner, A., Srisa-Art, M., Holt, D., Abell, C., Hollfelder, F., deMello, A.J. and Edel, J.B. (2007) Quantitative detection of protein expression in single cells using droplet microfluidics. Chem. Commun. (Camb). 1218–1220 CrossRef PubMed
82 Theberge, A.B., Courtois, F., Schaerli, Y., Fischlechner, M., Abell, C., Hollfelder, F. and Huck, W.T. (2010) Microdroplets in microfluidics: an evolving platform for discoveries in chemistry and biology. Angew. Chem. Int. Ed. Engl. **49**, 5846–5868 CrossRef PubMed
83 Baret, J.C., Miller, O.J., Taly, V., Ryckelynck, M., El-Harrak, A., Frenz, L., Rick, C., Samuels, M.L., Hutchison, J.B., Agresti, J.J. et al. (2009) Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. Lab Chip **9**, 1850–1858 CrossRef PubMed
84 Fallah-Araghi, A., Baret, J.C., Ryckelynck, M. and Griffiths, A.D. (2012) A completely *in vitro* ultrahigh-throughput droplet-based microfluidic screening system for protein engineering and directed evolution. Lab Chip **12**, 882–891 CrossRef PubMed
85 Kintses, B., Hein, C., Mohamed, M.F., Fischlechner, M., Courtois, F., Laine, C. and Hollfelder, F. (2012) Picoliter cell lysate assays in microfluidic droplet compartments for directed enzyme evolution. Chem. Biol. **19**, 1001–1009 CrossRef PubMed
86 Schaerli, Y. and Hollfelder, F. (2009) The potential of microfluidic water-in-oil droplets in experimental biology. Mol. Biosyst. **5**, 1392–1404 CrossRef PubMed
87 Larsen, A.C., Dunn, M.R., Hatch, A., Sau, S.P., Youngbull, C. and Chaput, J.C. (2016) A general strategy for expanding polymerase function by droplet microfluidics. Nat. Commun. **7**, 11235 CrossRef PubMed
88 Wochner, A., Attwater, J., Coulson, A. and Holliger, P. (2011) Ribozyme-catalyzed transcription of an active ribozyme. Science **332**, 209–212 CrossRef PubMed
89 Smith, G.P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science **228**, 1315–1317 CrossRef PubMed
90 Pande, J., Szewczyk, M.M. and Grover, A.K. (2010) Phage display: concept, innovations, applications and future. Biotechnol. Adv. **28**, 849–858 CrossRef PubMed
91 Huse, W.D., Sastry, L., Iverson, S.A., Kang, A.S., Alting-Mees, M., Burton, D.R., Benkovic, S.J. and Lerner, R.A. (1989) Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. Science **246**, 1275–1281 CrossRef PubMed
92 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 CrossRef PubMed
93 Krumpe, L.R. and Mori, T. (2007) Potential of phage-displayed peptide library technology to identify functional targeting peptides. Expert Opin. Drug Discov. **2**, 525 CrossRef PubMed
94 Matochko, W.L., Chu, K., Jin, B., Lee, S.W., Whitesides, G.M. and Derda, R. (2012) Deep sequencing analysis of phage libraries using Illumina platform. Methods **58**, 47–55 CrossRef PubMed
95 Griffiths, A.D. and Tawfik, D.S. (2006) Miniaturising the laboratory in emulsion droplets. Trends Biotechnol **24**, 395–402 CrossRef PubMed
96 Williams, R., Peisajovich, S.G., Miller, O.J., Magdassi, S., Tawfik, D.S. and Griffiths, A.D. (2006) Amplification of complex gene libraries by emulsion PCR. Nat. Methods **3**, 545–550 CrossRef PubMed
97 Derda, R., Tang, S.K. and Whitesides, G.M. (2010) Uniform amplification of phage with different growth characteristics in individual compartments consisting of monodisperse droplets. Angew. Chem. Int. Ed. Engl. **49**, 5301–5304 CrossRef PubMed
98 Rodi, D.J., Soares, A.S. and Makowski, L. (2002) Quantitative assessment of peptide sequence diversity in M13 combinatorial peptide phage display libraries. J. Mol. Biol. **322**, 1039–1052 CrossRef PubMed
99 Makowski, L. and Soares, A. (2003) Estimating the diversity of peptide populations from limited sequence data. Bioinformatics **19**, 483–489 CrossRef PubMed
100 Huang, J., Ru, B., Li, S., Lin, H. and Guo, F.B. (2010) SAROTUP: scanner and reporter of target-unrelated peptides. J. Biomed. Biotechnol. **2010**, 101932 PubMed
101 Jestin, J.L., Kristensen, P. and Winter, G. (1999) A method for the selection of catalytic activity using phage display and proximity coupling. Angew. Chem. Int. Ed. Engl. **38**, 1124–1127 CrossRef PubMed
102 Xia, G., Chen, L., Sera, T., Fa, M., Schultz, P.G. and Romesberg, F.E. (2002) Directed evolution of novel polymerase activities: mutation of a DNA polymerase into an efficient RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. **99**, 6597–6602 CrossRef PubMed
103 Delespaul, W., Peeters, Y., Herdewijn, P. and Robben, J. (2015) A novel helper phage for HaloTag-mediated co-display of enzyme and substrate on phage. Biochem. Biophys. Res. Commun. **460**, 245–249 CrossRef PubMed
104 Liu, C.C., Mack, A.V., Tsao, M.L., Mills, J.H., Lee, H.S., Choe, H., Farzan, M., Schultz, P.G. and Smider, V.V. (2008) Protein evolution with an expanded genetic code. Proc. Natl. Acad. Sci. U.S.A. **105**, 17688–17693 CrossRef PubMed
105 Hammerling, M.J., Ellefson, J.W., Boutz, D.R., Marcotte, E.M., Ellington, A.D. and Barrick, J.E. (2014) Bacteriophages use an expanded genetic code on evolutionary paths to higher fitness. Nat. Chem. Biol. **10**, 178–180 CrossRef PubMed
106 Strauss, A. and Gotz, F. (1996) *In vivo* immobilization of enzymatically active polypeptides on the cell surface of Staphylococcus carnosus. Mol. Microbiol. **21**, 491–500 CrossRef PubMed
107 Lee, S.Y., Choi, J.H. and Xu, Z. (2003) Microbial cell-surface display. Trends Biotechnol. **21**, 45–52 CrossRef PubMed
108 van Bloois, E., Winter, R.T., Kolmar, H. and Fraaije, M.W. (2011) Decorating microbes: surface display of proteins on *Escherichia coli*. Trends Biotechnol. **29**, 79–86 CrossRef PubMed
109 Isticato, R., Cangiano, G., Tran, H.T., Ciabattini, A., Medaglini, D., Oggioni, M.R., De Felice, M., Pozzi, G. and Ricca, E. (2001) Surface display of recombinant proteins on Bacillus subtilis spores. J. Bacteriol. **183**, 6294–6301 CrossRef PubMed
110 Sergeeva, A., Kolonin, M.G., Molldrem, J.J., Pasqualini, R. and Arap, W. (2006) Display technologies: application for the discovery of drug and gene delivery agents. Adv. Drug Deliv. Rev. **58**, 1622–1654 CrossRef PubMed
111 Gai, S.A. and Wittrup, K.D. (2007) Yeast surface display for protein engineering and characterization. Curr. Opin. Struct. Biol. **17**, 467–473 CrossRef PubMed
112 Zhou, C., Jacobsen, F.W., Cai, L., Chen, Q. and Shen, W.D. (2010) Development of a novel mammalian cell surface antibody display platform. MAbs **2**, 508–518 CrossRef PubMed
113 Zinchenko, A., Devenish, S.R., Kintses, B., Colin, P.Y., Fischlechner, M. and Hollfelder, F. (2014) One in a million: flow cytometric sorting of single cell-lysate assays in monodisperse picolitre double emulsion droplets for directed evolution. Anal. Chem. **86**, 2526–2533 CrossRef PubMed
114 Uyeda, A., Watanabe, T., Kato, Y., Watanabe, H., Yomo, T., Hohsaka, T. and Matsuura, T. (2015) Liposome-based *in vitro* evolution of aminoacyl-tRNA synthetase for enhanced pyrrolysine derivative incorporation. Chembiochem **16**, 1797–1802 CrossRef PubMed
115 Diamante, L., Gatti-Lafranconi, P., Schaerli, Y. and Hollfelder, F. (2013) *In vitro* affinity screening of protein and peptide binders by megavalent bead surface display. Protein Eng. Des. Sel. **26**, 713–724 CrossRef PubMed

© 2016 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution Licence 4.0 (CC BY).

Downloaded from http://portlandpress.com/biochemsoctrans/article-pdf/44/4/1165/486550/bst0441165.pdf by guest on 10 December 2022

116 Fischlechner, M., Schaerli, Y., Mohamed, M.F., Patil, S., Abell, C. and Hollfelder, F. (2014) Evolution of enzyme catalysts caged in biomimetic gel-shell beads. Nat. Chem. **6**, 791–796 CrossRef PubMed
117 Daugherty, P.S. (2007) Protein engineering with bacterial display. Curr. Opin. Struct. Biol. **17**, 474–480 CrossRef PubMed
118 Becker, S., Michalczyk, A., Wilhelm, S., Jaeger, K.E. and Kolmar, H. (2007) Ultrahigh-throughput screening to identify *E. coli* cells expressing functionally active enzymes on their surface. Chembiochem **8**, 943–949 CrossRef PubMed
119 Lipovsek, D., Antipov, E., Armstrong, K.A., Olsen, M.J., Klibanov, A.M., Tidor, B. and Wittrup, K.D. (2007) Selection of horseradish peroxidase variants with enhanced enantioselectivity by yeast surface display. Chem. Biol. **14**, 1176–1185 CrossRef PubMed
120 Becker, S., Hobenreich, H., Vogel, A., Knorr, J., Wilhelm, S., Rosenau, F., Jaeger, K.E., Reetz, M.T. and Kolmar, H. (2008) Single-cell high-throughput screening to identify enantioselective hydrolytic enzymes. Angew. Chem. Int. Ed. Engl. **47**, 5085–5088 CrossRef PubMed
121 Maass, A., Heiseler, T., Maass, F., Fritz, J., Hofmeyer, T., Glotzbach, B., Becker, S. and Kolmar, H. (2014) A general strategy for antibody library screening via conversion of transient target binding into permanent reporter deposition. Protein Eng. Des. Sel. **27**, 41–47 CrossRef PubMed

Received 24 March 2016
doi:10.1042/BST20160076

Downloaded from http://portlandpress.com/biochemsoctrans/article-pdf/44/4/1165/486550/bst0441165.pdf by guest on 10 December 2022

© 2016 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution Licence 4.0 (CC BY).