Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Engineering of biomolecules by bacteriophage directed evolution

Andreas K Brödel[1], Mark Isalan[1] and Alfonso Jaramillo[2,3,4,5,6,7]

Conventional *in vivo* directed evolution methods have primarily linked the biomolecule's activity to bacterial cell growth. Recent developments instead rely on the conditional growth of bacteriophages (phages), viruses that infect and replicate within bacteria. Here we review recent phage-based selection systems for *in vivo* directed evolution. These approaches have been applied to evolve a wide range of proteins including transcription factors, polymerases, proteases, DNA-binding proteins, and protein–protein interactions. Advances in this field expand the possible applications of protein and RNA engineering. This will ultimately result in new biomolecules with tailor-made properties, as well as giving us a better understanding of basic evolutionary processes.

Addresses
[1] Department of Life Sciences, Imperial College London, London SW7 2AZ, UK
[2] Warwick Integrative Synthetic Biology Centre and School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK
[3] CNRS-UMR8030, Laboratoire iSSB, Évry 91000, France
[4] Université Paris-Saclay, Évry 91000, France
[5] Université d'Évry, Évry 91000, France
[6] CEA, DRF, IG, Genoscope, Évry 91000, France
[7] Institute for Integrative Systems Biology (I2SysBio), University of Valencia-CSIC, 46980 Paterna, Spain

Corresponding author: Jaramillo, Alfonso (Alfonso.Jaramillo@synth-bio.org)

Current Opinion in Biotechnology 2018, **51**:32–38

This review comes from a themed issue on **Nanobiotechnology**

Edited by **Alfonso Jaramillo** and **Mark Howarth**

https://doi.org/10.1016/j.copbio.2017.11.004

Introduction

Protein engineering enables the development of valuable biomolecules for pharmaceutical and biotechnological purposes. There are generally two strategies to guide protein engineering: rational design or directed evolution (Figure 1). Rational design usually uses computational tools and structural considerations to identify beneficial mutations in the protein of interest [1]. Recent advances in this strategy even allow the design of proteins completely *de novo* [2–6]. In comparison, directed evolution mimics natural evolution and starts with a population of genotype(s) and then proceeds with the iterative generation of genotype diversity and a selection based on linked phenotype activity. It is applied when too little structural or biochemical information is available to guide engineering. In many cases, these two strategies can be combined in a semi-rational approach to improve the activity of biomolecules [7,8]. This illustrates how the method must be chosen to fit the particular problem.

A variety of directed evolution techniques have been developed that employ customized gene circuits [9–12]. One commonly used approach is to link the target protein's activity to cell growth, which is particularly suitable when the evolving gene directly improves cellular fitness [13–15]. Alternatively, the use of phage particles offers a convenient way to uncouple the target protein's activity from the fitness function of a cell. Instead, an artificial genetic circuit couples the evolving protein's function to increasingly efficient production of phage packaging the gene of interest [16••].

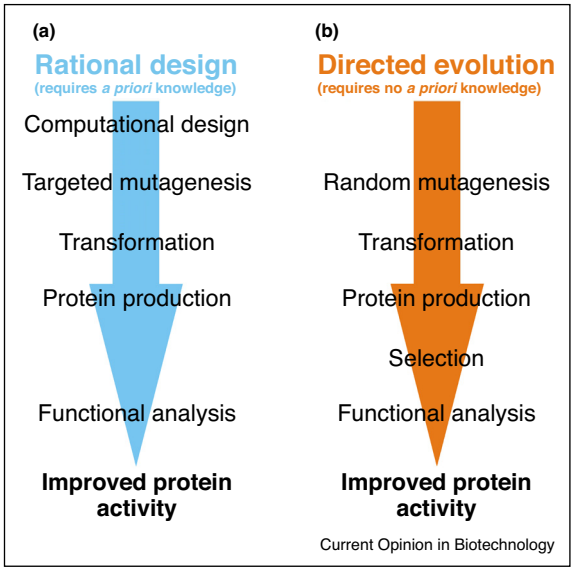
Directed evolution requires genotypic diversity in the gene of interest and this can either be achieved *in vivo* or *in vitro*. *In vivo* mutagenesis relies on intracellular modification of the target gene [17•,18,19] whereas *in vitro* mutagenesis can be achieved extracellularly by chemical modification [20], ultraviolet irradiation [21], or polymerase chain reaction (PCR) [22]. PCR-based methods generally employ an error-prone polymerase or oligonucleotides that contain randomized bases at the desired positions. Chemical mutagenesis and irradiation are less commonly-used methods because of the lack of uniform mutational spectra [20,23]. By making randomized libraries or using a progressive series of mutations, it is possible to explore the 'design space' of a target gene, ultimately enabling the engineering of new proteins.

In this review, we first discuss the requirements for using phages to evolve biomolecules. We then focus on new directed evolution methods based on conditional phage replication that have been developed thanks to advances in molecular and synthetic biology.

Re-engineering phage–host genetic interactions to select functional biomolecules

When evolving a target gene from either a gene library or mutation system, the phenotype selection can either be performed outside a living cell (*in vitro*) or inside (*in vivo*). *In vivo* evolution systems allow selecting for more

Figure 1



Protein engineering by rational design or directed evolution. **(a)** Rational design uses computational tools as well as structural or other biochemical knowledge to identify beneficial mutations in the protein of interest. These mutations are inserted into the gene of interest (targeted mutagenesis) which is then expressed in host cells. Functional analysis for each protein variant is performed to confirm improved activity. **(b)** Directed evolution is applied when too little structural or biochemical information is available to guide engineering. Mutations in the gene of interest are inserted randomly or by targeting specific positions in the gene sequence leading to a library of gene variants. Functional library members are then selected via a suitable selection system (e.g. phage-assisted evolution) against a target function. The activity of the selected protein is finally confirmed by functional analysis. Rational design and directed evolution are often combined to obtain the best results (semi-rational approach).

complex functions than *in vitro* methods (e.g. phage display) which are only suitable for binary protein–molecule interactions [24] (Figure 2a). By contrast, intracellular evolution potentially allows selection for multi-step processes, as long as they can be linked to genotype survival [25•]. For example, intracellular processes can facilitate the simultaneous mutation and selection of the gene of interest. Furthermore, it enables the use of counterselections against an undesired biomolecule function [26]. Another advantage of intracellular evolution is the subsequent compatibility of evolved genes or complex gene networks with the entire host cell machinery, as these have to function in a host cell context. To exploit these advantages, alternative phage-assisted directed evolution platforms have been developed.

To allow enrichment of functional genes, phage selection systems require a link between the desired phenotype and conditional phage replication. This can be achieved by removing an essential gene required for phage replication from the phage genome and linking its expression to the function of the evolving biomolecule. Alternatively, this gene (or genes) could be a host co-factor required by the phage replication but dispensable to the cell (to allow cell survival in the uninfected cells that are required as a host reservoir). However, the only approach developed so far rely on moving essential genes from the phage to the host cell or its associated plasmids [16••,27•] (Figure 2b,c). These systems may be classified according to the degree of phage engineering involved, where only a single gene may be moved or practically all of them.

The evolving biomolecule has to be encoded in the phage and a genetic system has to be designed to allow a functional molecule to activate the expression of the essential gene (*positive selection*). When the evolving biomolecule is able to induce the expression of the missing gene, infectious virions will package the DNA encoding the biomolecule, promoting its survival. The conditional expression of the essential gene can be done at the transcriptional or post-transcriptional levels, depending on the biomolecule to be evolved (e.g. a transcription factor or a riboregulator).

Alternatively, selection may consist in designing a conditional interference with phage replication if a biomolecule is functional (*negative selection*). This is used to penalize any unwanted activity such as the original parental function of the biomolecule. The selection can also be complex or variable, where the stringency of positive and negative selection can be modulated exogenously [26].

Many alternative phage–host systems can in principle be chosen for the evolution of biomolecules depending on the application. For instance, if one wanted to evolve a photosynthetic protein, one might choose a cyanobacterium and one of its known phages. The disadvantage of such approaches is that the phage biology is not well characterized. Consequently, in this article we will focus on *Escherichia coli* due to the lack of reported works with other organisms. The *E. coli* phages M13 [28], T4 [29,30], T7 [31] or λ [32] have been used to optimize protein function and stability with phage display, although M13 has been the only phage vector used to evolve biomolecules *in vivo* thus far.

Evolving biomolecules through positive selection

Recently, a new method to evolve biomolecules using M13 was developed, using a redesign of the host to implement a positive selection: Phage-Assisted Continuous Evolution (PACE) describes a general approach for the directed evolution of proteins *in vivo* [16••]. Using PACE, new T7 RNA polymerase (RNAP) variants against a T3 promoter have been evolved, which are not bound by the wild-type T7 RNAP. For this, the minor coat protein pIII is replaced by the evolving gene