Robust monooxygenase biocatalysts

Fürst, Maximilian

Chapter 6:

Experimental Protocols for Generating Focused Mutant Libraries and Screening for Thermostable Proteins

Maximilian J.L.J. Fürst,[a] Caterina Martin,[a] Nikola Lončar,[a] Marco W. Fraaije*[a]

[a]Molecular Enzymology Group, University of Groningen, Nijenborgh 4, 9747AG, Groningen, The Netherlands

*Corresponding author

Published in:

Methods in Enzymology; Scrutton, NS (Ed.), Academic Press: 2018; Vol. 608; pp 151–187.

Abstract

Many proteins are rapidly deactivated when exposed to high or even ambient temperatures. This can not only impede the study of a particular protein, but also is one of the major reasons why enzyme catalysis is still widely unable to compete with established chemical processes. Furthermore, differences in protein stability are a challenge in synthetic biology, when individual modules prove to be incompatible. The targeted stabilization of proteins can overcome these hurdles, and protein engineering techniques are more and more reliably supported by computational chemistry tools. Accordingly, algorithms to predict the differences in folding energy of a mutant compared to the wild-type, $\Delta\Delta G^{Fold}$, are used in the highly successful FRESCO workflow. The resulting single mutant prediction library consists typically of a few hundred amino acid exchanges, and after combining the most successful hits we so far obtained stabilized mutants which exhibited increases in apparent melting temperature of 20-35 ºC and showed vastly increased half-lives as well as resistance to cosolvents. Here we report a detailed protocol to generate these mutant libraries experimentally, covering the entire workflow from primer design, through mutagenesis, protein production and screening, to mutation combination strategies. The individual parts of the method are furthermore applicable to many other scenarios besides protein stabilization, and these protocols are valuable for any project requiring individual or semi high-throughput site-directed mutagenesis, protein expression and purification, or generation of mutant combination libraries.

Introduction

Synthetic biology builds on the modular aspects by which cells function and can be programmed. Crucial elements in cellular functions are enzymes. With the current level of knowledge in enzymology, which ranges from detailed insights into how enzymes are produced and modified by post-translational processes to the availability of databases that can be queried for known and characterized enzymes, design of novel metabolic pathways has become possible. By establishing the expression of a complementary set of enzymes, it has become feasible to create recombinant organisms capable of producing a desired target molecule. Such metabolic engineering efforts require enzymes to fulfill a wide range of criteria. [1-2] In fact, for generating effective metabolic pathways in synthetic biology, often the performance of the enzymes involved needs to be tuned. Targets of enzyme engineering often are optimizing kinetics, improved chemo-, regio- and/or enantioselectivity, altering the substrate specificity, improving thermostability, and lowering inhibition or inactivation. Moreover, it is also crucial to optimize enzymes for many other biotechnological applications. While numerous approaches in enzyme engineering have emerged in the last few decades, it can still be a challenge to engineer an enzyme towards obeying the criteria set by the target application. From the wide range of different approaches explored, it seems a consensus in enzyme engineering has emerged in recent years: a hybrid approach of random mutagenesis combined with high-throughput screening (directed evolution) and knowledge-based directed mutagenesis. Such hybrid approaches typically make use of structural and/or mechanistic knowledge to design so-called 'clever libraries' that still contain a fair amount of freedom by allowing several mutations at various positions in a protein sequence.

In this contribution we provide experimental protocols that can be used for almost any protein engineering effort that involves the creation and screening of small targeted mutant libraries. We have used these protocols for successfully generating proteins with improved (thermo)stability. The 'knowledge-based' component to guide the preparation of the mutant libraries is based on a recently developed computational method which predicts mutations that are beneficial for enzyme stability: Framework for Rapid Enzyme Stabilization by Computational libraries (FRESCO). In our hands, the computational FRESCO method in combination with the experimental protocols as described here has delivered marked improvements in thermostability of 20-35 °C in apparent melting temperature for various

proteins.[3] The FRESCO method proved to be useful not just in increasing the apparent melting temperature of the enzymes but also in giving a remarkable increase in resistance to cosolvents.[4] We have also used this approach to generate thermostable variants of flavoenzymes by using the ThermoFAD method for screening for improved mutants.[5]

The FRESCO *in silico* library can be created even by computationally inexperienced biochemists in a few weeks,[6] however, the protocols provided here can also be used for the preparation and screening of mutant libraries based by other stabilization predictions methods or even entirely different design criteria. The provided protocols can be applied to any other small library and (flavo)enzyme engineering target. We set ourselves a target to provide the reader with comprehensive and detailed, yet straightforward protocols for semi high-throughput generation, expression, production and verification of mutants in a 96-well format. Furthermore, we provide two experimental strategies for efficient combination of the best hits.

Single mutants generation

Through FRESCO, other computational predictions tools or other structure-inspired input, it is often desirable to generate a large collection of site-directed mutant proteins. From the prediction generated through FRESCO, typically a library of a few hundred mutant proteins should be prepared. Also other methods often target such a library size, as it allows medium throughput screening methods for establishing which mutant proteins perform best. We have optimized a protocol that allows the generation of such libraries, by combining established molecular biology methods to an efficient workflow, and by fine-tuning crucial steps to work seamlessly in the 96-well format (Figure 1).

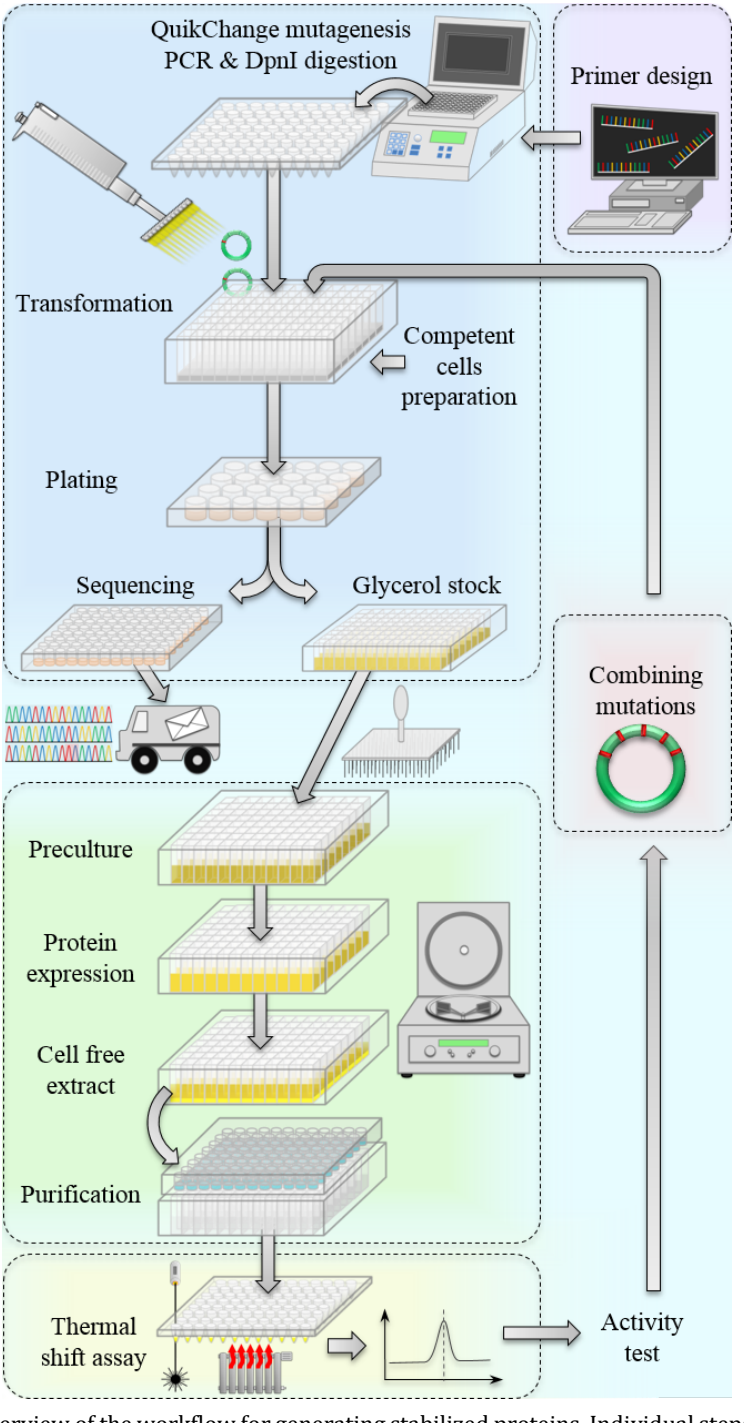

Figure 1. Overview of the workflow for generating stabilized proteins. Individual steps and some of the main equipment required are depicted. Individual sections in this chapter where the corresponding procedures are described indicated by dashed lines.

A well-established way to generate single mutants is the QuikChange™ mutagenesis method developed by Stratagene (now Agilent). In this method, a complete vector is PCR-amplified by two completely complementary primers that contain the desired mutation, and DpnI is used to digest the methylated template. The resulting linear PCR fragments contain complementary overhangs and thus circularize to a nicked plasmid, which becomes ligated by *E. coli*'s DNA repair machinery after transformation. Despite yielding relatively reliable results when using the original protocol, there was room for improvement. Limiting the failure rate in this initial experimental phase is essential to limit labor and material costs. For a typical library size of a few hundred mutants, primer synthesis, polymerase and sequencing service can become the major cost factors of the method. Consequently, a reliable primer design strategy, and a high PCR success rate are critical.

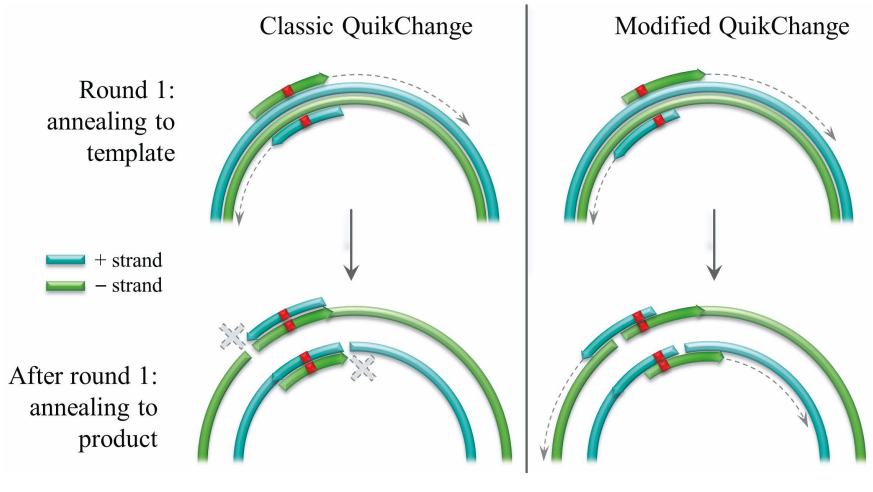

Figure 2. Comparison of the QuikChange protocols. In the modified protocol, the linearized PCR product can serve as a template, because the primer can bridge the nick. In the classic QuikChange, only the original plasmid can serve as a template.

The two most relevant limitations of the original QuikChange protocol are i) favored primer-dimer formation over template annealing, particularly for an increased number of mismatches, and ii) the insufficient accumulation of product, due to the linear rather than exponential amplification of strands (Figure 2). Adaptations to the protocol elegantly overcame both limitations,[7-8] while maintaining the simplicity and the general workflow. Accordingly, instead of using a completely complementary pair of primers, the overlapping region is confined to the 5' end. The non-overlapping 3' end is designed to have

a higher T_m than the overlapping region, allowing primer-dimers to be melted, while the annealing to the template is retained. Moreover, in this way exponential amplification is achieved, since the linearized PCR product can serve as a template because the primer is able to bridge the nick (Figure 2).
Besides primer design, one should also pay attention to other details of the (otherwise common) procedures in order to maximize the success rate. Firstly, a high-fidelity polymerase should be used for PCR, to minimize undesired and potentially unnoticed mutations. As an economically interesting alternative to Phusion, we are commonly using *Pfu* polymerase, which we and others[9] found to be similar in fidelity, despite what is commercially advertised. Assuming an error rate of 2.8×10^{-6} (see also footnote d), and two sequencing reads of 1 kb each, one would accumulate 1 unnoticed mutation per 96-well plate for a vector size of 5.7 kb $[(5700 - 2000) \cdot 2.8 \cdot 10^{-6})^{-1}$, which we find tolerable. Secondly, the in practice often overdosed amount of template DNA in the PCR should be kept at the lower end of the polymerase manufacturer's recommendations, to facilitate DpnI digestion and avoid wild-type colonies. Thirdly, high quality competent cells should be prepared in order to counter a lowered transformation efficiency in 96-well plates.

Primer design

Since reliably designing hundreds of primers manually is impractical, we recommend using automated software with adjustable parameters in order to satisfy common primer rules as well as guidelines specific to the partially overlapping design strategy discussed by Liu et al. One example is AAscan,[10] an open source tool run under Windows. Although the software has a batch mode to generate a list of primers on different positions, the input can currently not stem from a defined list of mutations, and is furthermore restricted to one amino acid at a time. There are two options to get to the desired list of primers: either by manual input of each position and mutant codon, or by generating primers for all positions and each amino acid, and subsequent filtering of the relevant primers. We describe the procedure for the latter option, including a small UNIX command line script that automatically filters the final list of primers, using as input a list of mutations generated in the computational part of the FRESCO procedure, or any other list that sticks to a common format.[a]

[a] During the visual inspection of the computational FRESCO procedure, an automatically generated spreadsheet is filled in with notes leading to acceptance or rejections of predicted

Equipment

- PC running Windows
- DNA sequence of the template plasmid
- AAscan, available from https://www.psi.ch/lbr/aascan
- A UNIX command line (shell), such as the terminal of MacOS or linux, or the Windows 10 feature known as "Windows subsystem for Linux"

Procedure

1. Download AAscan and execute it; there is no install process needed.

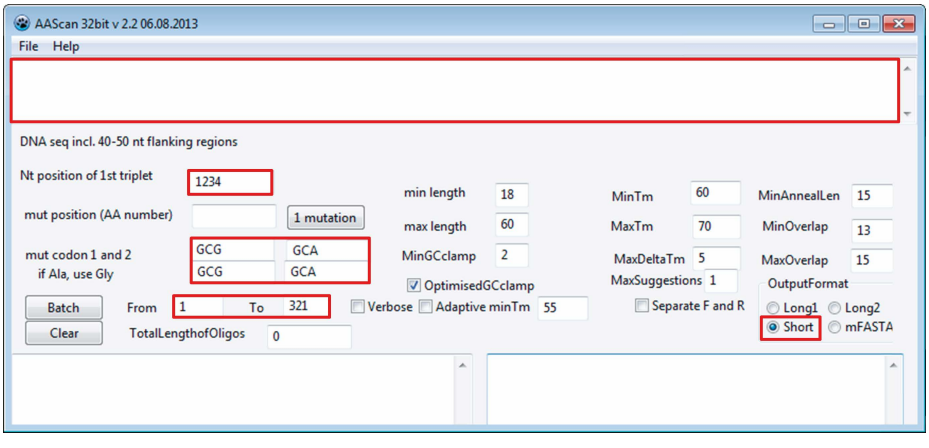

Figure 3. Screenshot of AAscan. Fields requiring adaptation are marked in red.

2. Copy and paste the unformatted sequence of the vector containing the wild-type gene in the DNA sequence field and enter the number corresponding to the sequence position of the gene's first nucleotide (usually the A of the ATG start codon).
3. Since the batch mode will be used, leave the "mut position" field empty.
4. Type the preferred codon for alanine into the "mut codon 1" field, as well as into the first "if Ala, use Gly" field below, and an alternative codon in the two fields on the right. It is important to enter the codon twice, otherwise,

stabilizing mutations. One of the columns in this file contains the mutations in an XNY format. A list compatible with the above script can be generated by simply copy and pasting the accepted mutations from this column to a plain text file (e.g. created using Notepad under Windows).

the software will mutate all alanines to glycines. Chose the codons with the help of an *E. coli* codon usage table.[b]

5. In the batch fields, type "1" in the first, and the number of the last amino acid of your protein in the second field.
6. Choose the "Short" output format and keep the rest of the parameters unchanged or adapt according to specific requirements.[c]
7. Click "Batch" and copy the output in a plain text file, named with the one letter code of the amino acid (e.g. "A.txt" for alanine).
8. Repeat steps 4-7 for all 20 amino acids.
9. On a computer with a UNIX command line (such as bash), create a folder containing these 20 files as well as a plain text file (.txt) that contains a list of all mutants in an XNY format (e.g. G129S)[a] and give the file the name "mutations.txt".
10. To get only the primers for the mutations in this list, open a command line terminal, go to the folder containing the primer files and the mutations list, and run the following script either by copy and pasting into the terminal, or by creating and executing a script file:

```
while read -r line; do
   mut=$(echo "$line" | grep -o -E '[A-Z][0-9]+[A-Z]')
   num=$(echo "$mut" | sed 's/[a-zA-Z]//g')
   res=$(echo "$mut"| grep -o -E '[A-Z]$')
   fw=$(cat "$res".txt | grep ^"$num"_F | grep -o -E [agctACGT]+)
   rev=$(cat "$res".txt | grep ^"$num"_R | grep -o -E [agctACGT]+)
   printf "%s\t%s\t%s\t%s\n" "$mut"_fw "$fw" "$mut"_rev "$rev"
done < mutations.txt > primers.txt
```

11. A file called "primers.txt" is created that contains the final list. Verify on a few examples that the primers are correct. Then order the primers from an

[b] Codon usage tables for *E. coli* can be found abundantly on the web. Considering that i) the current understanding of the interrelation of DNA sequence (including codon usage) with protein expression is still far from fully understood[11] and ii) one can assume that the impact on expression exerted by the change of a relatively small fraction of the gene's codons is minimal, in practice it is probably sufficient to just make sure to avoid *E. coli*'s very rare codons: CGA, AGA, AGG (Arg), ATA (Ile) and CTA (Leu).

[c] If AAscan cannot produce a primer that fulfills the criteria, it will output '*******', which will result in an empty field for the primer after the script. In that case, the parameters should be adapted. Increasing the MaxTm field to 75 °C can help with GC rich DNA stretches, otherwise the MinTm can also be lowered. Then the PCR annealing temperature should be adapted as well.[h]

oligonucleotide synthesis service. Ideally, order the primers in a 96-well plate and dissolved in water to 100 µM.

QuikChange library creation

The protocol below describes the procedure for the usage of the *PfuUltra* II Hotstart Master Mix for the PCR, basically following the instructions of Agilent. The protocol for the preparation of chemically competent cells in a cation-mix buffer is based on the one from Green & Rogers,[12] enabled through the research of Hanahan.[13] It is advantageous to transform the PCR product of the QuikChange reaction directly into the strain used for expression, in order to avoid the need of having to perform a second transformation step. However, many dedicated expression strains are not optimized for molecular biology applications, resulting in low transformation efficiencies. On the other hand, cloning strains are often unsuitable for protein overexpression.[d] We advise to investigate the feasibility of direct transformation for one's particular system. Test first a couple of mutants, using the high-throughput equipment, before applying the procedure to one or maximally two 96-well plates of mutants at a time. Most of the following steps can and should be performed using an (electronic) multichannel pipette, even when it is not explicitly stated.

Equipment

- Multichannel pipette (Sartorius, Biohit Picus electronic pipette 12-ch)
- 37 °C shaking incubator for test tubes (Infors, Multitron Standard)
- 96-well PCR plates with dedicated sealing (Bio-Rad, HSL9601)
- Sterile disposable or autoclavable 96-square deep well plates (Waters, 186002482)
- Sterile disposable or autoclavable 24-well plates (Greiner Bio One, 662102)
- Gas-permeable and water vapor retaining microtiter plate sealing (Excel Scientific, AeraSeal B-104)

[d] In this protocol, we describe one of our preferred expression systems making use of the pBAD vector,[14] which contains the tightly controlled and L-arabinose-inducible P_{BAD} promoter, a pBR322 origin of replication and an ampicillin resistance marker. If successful expression is dependent on constant or very well defined levels of arabinose (in practice seldom the case) the use of an *araABCD*- strain such as top10 or NEB® 10-beta is obligatory. Otherwise, any strain with an intact arabinose import system (*araEFGH*+, virtually all lab strains) can be used, as long as high enough amounts of arabinose (usually 0.1-2%) are added. In our experience, *E. coli* TOP10 is an excellent strain for both cloning/transformation, as well as expression (at least from pBAD), with the resulting advantage of avoiding the additional re-transformation step.

- Adhesive aluminium sealing foils for microtiter plates (Greiner Bio One, 676090)
- Cooling centrifuge (Eppendorf, 5804R), rotors for tubes and microtiter plates (Eppendorf, A-2-DWP)
- PCR machine (Peqlab, 96 Universal Gradient Peqstar)
- Sterile toothpicks
- Temperature controlled water bath (Julabo, TW20)
- Temperature controlled shaker, suitable for microtiter plates[e]
- Sequencing service, ideally including plasmid preparation from clones (Eurofins Genomics)

[e] A range of studies investigated expression conditions in 96-well plates mainly focusing on oxygen limitation—the main challenge in these systems.[15] General conclusions are:

- Square wells permit more effective volume than round wells, and shaking aeration is ~2x higher, since the corners act as baffles to introduce turbulences.[16]
- Sufficient head space between culture surface and seal needs to be ensured to avoid splashes, causing gas-exchange-impairing droplets on the seal and cross-contaminations.
- The larger the orbital diameter of the shaking device is, the lower needs to be the shaking frequency, in order to achieve the same oxygen transfer rates.[17]
- Oxygen transfer rates increase exponentially with the shaking frequency, but only beyond a critical value (Table 3), which is often higher than common shaking operating conditions.

The practical problem is that the proficiency of the equipment that an individual researcher has at hand is influenced by many parameters: the plate (well dimensions and shape, material), the shaking device (orbital diameter, frequency range), and the cultivation conditions (culture volume, medium, organism and strain). Since the number of variables impedes both the encounter of applicable literature, as well as empirical determination, we recommend to approach highest productivity by using 96-square deep well plates with at least 2 mL well volume, and an instrument which permits a shaking frequency range as specified in Table 3. The ideal volume/frequency pair should be determined by either using a fixed frequency for maximum oxygen transfer (Table 1), or a fixed volume of 1 mL. In both cases, the second parameter can be determined by a stepwise increase in volume/frequency to a plate sealed with paper tissue, where a sample of an expression culture was added to some representative wells. Use the volume/frequency just before the occurrence of splashing (as inferred from tissue wetting) for the expression culture.

Table 1. Shaking frequencies for instruments with different orbitals*

Orbital (mm)	50	25	12.5	6	3	1.5
Critical value freq. (rpm)	195	255	395	565	735	1070**
Maximum O_2 transfer freq. (rpm)	290	390	560	770	1000	1400**

*Values approximated from ref. 16 Figure 1, assuming a 2x increase in O_2 transfer due to square wells

**Extrapolated from fitting the data using an exponential function

Buffers and Reagents

- High fidelity polymerase and reaction components (Agilent, *PfuUltra* II Hotstart PCR Master Mix 2x)[f]
- DpnI restriction enzyme (New England Biolabs, R0176S)
- *E. coli* cells[d]
- LB Miller's high salt medium (10 g tryptone, 5 g yeast extract, 10 g NaCl in 1 L H_2O)
- LB agar plates with antibiotic
- SOC medium (20 g tryptone, 5 g yeast extract, 0.58 g NaCl, 0.19 g KCl, 2.4 g $MgSO_4$ in 1 L H_2O, pH adjusted to 7.5 with NaOH. After cooling medium to less than 50 °C, add 20 mL filter sterilized 20 % glucose solution).
- Buffer I for chemically competent cells: 100 mM RbCl, 50 mM $MnCl_2$, 30 mM potassium acetate, 10 mM $CaCl_2$, 15 % w/v glycerol; pH adjusted to 5.8 with 0.2 M acetic acid. Sterilized by filtration with 0.45 µM pore size filter.
- Buffer II for chemically competent cells: 10 mM RbCl, 75 mM $CaCl_2$, 10 mM MOPS, 15 % w/v glycerol; pH adjusted to 6.5 with 0.5 M KOH. Sterilized by filtration with 0.45 µM pore size filter.
- Sterile 70 % glycerol

Procedure

1. Preparation of chemically competent cells. Perform all steps observing sterility.
   a. Inoculate 5 mL LB in a sterile test tube with the desired *E. coli* strain.[p] Incubate overnight (o/n) at 37 °C, shaking.
   b. With this preculture, inoculate 165 mL LB (equivalent to one 96-well plate of competent cells) in a sterile culture flask and incubate shaking at 37 °C until an OD_{600} of 0.25-0.5.
   c. Transfer the culture to (a) sterile centrifuge tube(s) and place on ice for 20 min.
   d. Centrifuge at 4000 × *g* in a pre-cooled centrifuge at 4 °C for 10 min. Discard the supernatant and resuspend the pellet in 1 mL of ice-cold buffer I. Fill to 60 mL with buffer I and keep on ice for 20 min.

[f] The polymerase described here, *PfuUltra* II, is an engineered variant of the original *Pfu* polymerase, which—according to the supplier—is even higher in fidelity than the original *Pfu* which was compared to Phusion.[8]

e. Centrifuge at 4000 × *g* and 4 °C for 10 min. Discard the supernatant and resuspend in 1 mL of ice-cold Buffer II. Mix well, but carefully, fill to 6 mL and keep on ice for 15 min.
f. In a cold room, or with pre-cooled equipment and keeping reagents on ice whenever possible, aliquot 50 µL in each well of a 96-deep well plate.[g]
g. Use directly for transformation, or seal with adhesive aluminium sealing foil, flash-freeze in liquid nitrogen and store at −80 °C.

2. PCR
   a. Prepare a DNA template stock by diluting a fresh plasmid preparation of the wild-type gene-containing vector to a concentration of 1 ng/µL and store it at −20 °C.[h]
   b. If shipped otherwise, prepare primer stocks in a 96-well PCR plate by diluting all forward and reverse oligonucleotides to a concentration of 100 µM and store at −20 °C.
   c. Add 240 µL of sterile water to each well of a 96-well PCR plate with a multichannel pipette. Add 5 µL of forward primer stock and 5 µL of reverse primer stock to get a primer mix with 2 µM of each primer.
   d. In a 96-well PCR plate, mix the following reagents to a final volume of 25 µL:[i]

Table 2. PCR recipe

Component, stock concentration	Final concentration	Volume to pipet
Template stock, 1 ng/µL	0.2 ng/µL	5 µL
Primer mix, 2 µM each	0.4 µM each	5 µL
PfuUltra II Hotstart PCR Master Mix 2x	1x	12.5 µL
MiliQ water		2.5 µL

   e. Run a PCR on a thermal cycler with the following program:

[g] Use precooled pipette tips and a precooled sterile or an autoclavable multichannel reservoir on ice.

[h] Since multichannel pipettes tend to be less precise than single channel pipettes, and precision decreases with volume, we recommend the preparation of diluted stocks that require pipetting volumes of ≥ 5 µL.

[i] If cost is an issue, the volume can be decreased up to 15 µL. Lower yields due to smaller volumes are compensated by high product yields due to efficient PCRs and high-quality competent cells. If a different polymerase is used, adapt both recipe as well as thermal cycling program according to the manufacturer's instructions.

Table 3. PCR program

Temperature	Time	Cycles
95 °C	2 min	1
95 °C	20 s	
55 °C*	20 s	26
72 °C	30s/kb**	
72 °C	10 min	1

*Adapt if the default parameters of AAscan were changed[j]

**Doubled amplification time compared to manual, to increase yield[k]

3. When the PCR is finished, add 10 U DpnI to each well and incubate o/n at 37 °C.
4. Thaw a plate of competent cells on ice (or use freshly prepared cells). Add 5 µL of the DpnI-digested PCR products, but do not mix by pipetting up and down. Incubate on ice for 20 min.
5. Heat shock by placing the plate in a 42 °C water bath for 40 s.[l]
6. Put the plate on ice for 2 min, then add 500 µL of pre-warmed SOC medium.
7. Incubate on a plate shaker for 1 h at 37 °C and 1000 rpm.
8. Centrifuge at 2250 × *g* for 10 min and remove 430 µL of supernatant.
9. Resuspend the pellet with the remaining 70 µL and plate one-by-one on a 24-well plate with LB agar containing the proper antibiotic.[m] Incubate o/n at 37 °C.
10. Use a sterile toothpick to pick a colony of each well and transfer it to two 96-well plates: one with LB-agar plus antibiotic to be sent for sequencing[n] and one with liquid LB medium plus antibiotic to grow o/n.

[j] The minimum primer T_m parameter in AAscan is per default set to 60 °C. The annealing temperature should be 5 °C below the lowest T_m.

[k] Since amplification speed is one of the main competition parameters between polymerase manufacturers, we impute a possible exaggeration.

[l] The strain's commercial supplier should be consulted to infer the strain dependent optimal heat shock time. We recommend the extension of this time by ~50% for transformation in polypropylene deep well plates, due to the impaired heat transfer compared to reaction tubes.

[m] This is the only step where the 96-well format has to be abandoned. Use manual pipetting for this step. Plating can be performed by swirling motions or using sterile glass beads.

[n] Several DNA sequencing companies offer the plasmid preparation step from colonies on 96-well plates. In our case, this extra cost is reasonably low to outcompete kit prices and labor of manual plasmid preparation of 96 samples. If re-transformation into an expression strain is required, it is necessary to ensure that the sequencing company can send back the plasmid preparation. However, we recommend to request this from the company in any case.

11. Add sterile glycerol to a final concentration of 30% to the o/n liquid culture plate, mix well, seal with adhesive aluminum sealing foil, flash-freeze in liquid nitrogen and store at −80 °C.

Expression and protein purification in a 96-well plate

Expression and purification should ideally be performed in a 96-well format as well. For a reliable melting temperature measurement (see the "Melting temperature screening" section), 20 µL of a solution of 0.3 mg/mL of the target protein are minimally required.[o] It should be checked in advance, whether or not enough protein is obtained from the 1 mL plate culture in each individual case. If the target amount is not achieved, an effort should be made to improve expression yields, a topic which is beyond the scope of this protocol.[e,d] Otherwise, one can prepare two 96-well plates with identical clones, and combine the cell free extracts, or switch to 24-well plates, or else express in 50 mL cultures using 250 mL flasks. Though laborious, such individual expression can be performed in a 1-2 weeks period for 50-100 mutants. When plates are used, it is important to include in each plate the wild-type as a reference, and a non-inoculated negative control for contamination in one well, or in several wells diagonal across the plate.

If expression is very high, purification from the cell free extract might be omitted for some proteins, see the "Melting temperature screening" section for details. Otherwise, many of the commonly applied purification methods are feasible in a 96-well format. We describe in this protocol the procedure for the widely applied polyhistidine-tag affinity chromatography purification using divalent nickel immobilized on beads, in our case Ni Sepharose®. Furthermore, the Tris/HCl buffer as well as the pH described here can be adapted to one's specific needs or preferences, as long as they are compatible with the purification method.

Equipment

- Temperature controlled shaker, suitable for microtiter plates[d]
- Cooling centrifuge and rotor for microtiter plates
- Sterile disposable or autoclavable 96-square deep well plates
- 96-well Filter Plates, glass polypropylene membrane (Whatman, Unifilter 2 mL, 7720-7236)

[o] The minimum value varies slightly for different proteins and should for each case be determined with the wild-type beforehand.

- 96-well Filter Plates, polysulfone membrane (Pall, AcroPrep advance, PN8130)
- Disposable 96-well microtiter plate (Greiner, 650161)
- Multichannel pipettes
- Gas-permeable and water vapor retaining microtiter plate sealing
- Adhesive aluminium sealing foils for microtiter plates

Buffer and Reagents

- LB medium
- Ni Sepharose® resin (GE Healthcare, 17526802)
- Buffer A: 50 mM Tris/HCl, pH 7.5
- Buffer B: 50 mM Tris/HCl, pH 7.5, 5 mM imidazole
- Buffer C: 50 mM Tris/HCl, pH 7.5, 500 mM imidazole
- Lysis buffer: 50 mM Tris/HCl pH 7.5, 10 mM $MgCl_2$, 0.5 mg/mL DNase I (NEB, M0303L), 1 mg/mL lysozyme (Sigma-Aldrich, L6876).

Procedure

1. Dispense 1 mL of LB medium with antibiotic in each well of a sterile 96-deep well plate and inoculate from the glycerol stock plate.[p] Grow the pre-cultures in a suitable shaker at 37 °C and use appropriate agitation o/n.
2. Inoculate the main culture plate from this preculture plate and express the protein according to your specific expression system. In order to achieve the high expression yields needed for the stability assay, consider the following guidelines:
    - Use high copy number plasmids, strongly overexpressing and inducible promoters such as P_{BAD} or T7 and induce expression after an initial growth.[d]
    - If you have a well-established expression protocol, take into consideration that you might have to modify it for the small scale expression. Typical adjustments are significantly larger inoculation volumes and higher OD_{600} induction.
    - Use a high cell density culture medium such as terrific broth (TB).
    - Use square deep well plates and vigorous shaking for high aeration.[e]

[p] This can be done with dedicated equipment (we use a cryo-replicator from EnzyScreen), or a multi-channel pipette. Since this step bears a risk of cross contamination, extra care is required, and the use of high concentrations of glycerol in the stock plate (step 11 in the "QuikChange library creation" section) avoids 'jumping' frozen debris upon penetration.

3. Centrifuge 20 min at 2250 × *g* at 4 °C and remove the supernatant by inverting the plate.
4. Thoroughly resuspend the pellet in 200 µL of lysis buffer,[q] cover the plate with aluminum seal and incubate at 20 °C with mild shaking for 30 min.[r]
5. Flash-freeze the plate in liquid nitrogen and thaw again in a water bath at room temperature.
6. Centrifuge the lysed cells in the plate for 45 min at 2250 × *g* at 4 °C to remove cells debris.
7. Equilibrate 10 mL of Ni-Sepharose resin in a gravity flow column by washing with at least 50 mL of buffer A.
8. Resuspend the beads in an equal amount of buffer A and dispense 200 µL of the resin suspension[s] in each well of a 96-well polysulfone membrane filter plate. Place the filter plate on top of a 96-deep well plate and remove excess buffer by centrifuging for 15 s at 200 × *g*.
9. Transfer the supernatant of the centrifuged lysed cells to a 96-well glass polypropylene membrane filter plate, placed on top of a 96 deep-well plate, and centrifuge at 200 × *g* for 1 min to remove residual cell debris.
10. Transfer the cell extracts to the filter plate containing the resin and resuspend very carefully. Close top and bottom with adhesive aluminum seal and incubate for 15-30 min at 4 °C in a plate shaker with mild agitation.
11. Centrifuge 15 s at 200 × *g* and collect the flow through in a 96-well microtiter plate.
12. Wash the resin with 200 µL of buffer A by centrifuging 15 s at 200 × *g*.
13. Wash with 200 µL of buffer B by centrifuging 15 s at 200 × *g*.
14. Elute with 100 µL of buffer C by centrifuging 15 s at 200 × *g* two times and collect the fractions separately in a 96-well microtiter plate.
15. Identify the most concentrated fraction (by eye or using a NanoDrop) and transfer this fraction slowly to an equilibrated desalting plate.
16. Centrifuge 15 s at 200 × *g* to load the columns, and elute by adding 100 µL of buffer A and centrifuging for 15 s at 200 × *g* and collecting in a 96-well

[q] To resuspend the pellet more easily, combine circular motions with pipetting up and down.

[r] Estimate lysis efficiency by SDS-PAGE. If necessary, increase incubation time, or—if the protein's stability is guaranteed—increase the incubation temperature to 37 °C.

[s] Be careful when pipetting pressure sensitive beads. We recommend the use of 1000 µL pipette tips where the thinnest part of the tip has been cut off.

microtiter plate (or proceed according to the desalting plate manufacturer's instructions).

17. Proceed directly with the T_m measurement, and flash-freeze the remaining purified protein in liquid nitrogen and store at −80 °C.

Melting temperature screening

Several methods to measure the thermostability of proteins have been developed. However, the feature most commonly targeted is operational stability, i.e. the stability over time under working conditions. Developing a high-throughput screen sensitive to subtle changes for this feature is, however, notoriously difficult. As an attractive alternative, a protein's apparent melting temperature is a quantitative parameter requiring minimal protein amounts for its accurate determination. Furthermore, it can be measured in less than one hour for one 96-well plate of mutant proteins. It is also easy to screen many variants for stability under different conditions, such as pH ranges and buffer additives. If dedicated equipment is not available, a real-time PCR (qPCR) machine, commonly available in many molecular biology laboratories, can be used. The T_m is then determined by following the change of the emission intensity of a fluorescent dye which is added in the assay (or by monitoring the fluorescence of a protein bound ligand, such as a flavin cofactor). Upon unfolding, proteins expose otherwise buried hydrophobic patches to which these dyes can bind, effecting a shift of the dye's fluorescence spectrum (thus the name fluorescence-based thermal shift assay, trademarked as *Thermofluor*®). The bandpass filters with which qPCR instruments are equipped are adapted to the properties of the DNA-binding dyes most commonly used in real-time PCR, but are also perfectly suitable for many protein-binding dyes (e.g. SyproOrange), which have similar fluorescent properties. Conveniently, flavins also have fluorescence emission and excitation spectra that fall in the same ranges. In the case of most flavoproteins, the release of the cofactor results in a significant increase in fluorescence, as flavins are normally quenched when bound to the protein. Thus, flavin-containing proteins can be used for the thermal shift assay without the addition of dyes, and this method was named ThermoFAD.[5] Since this method is so specific for flavoproteins, purification from the cell free extract (CFE) might even be omitted under certain circumstances. Besides the requirement for high expression levels, one has to carefully check a relevant number of mutants and the wild-type, to see whether or not the T_m measured in CFE and the T_m of purified protein is the same (or only marginally different). We have had cases

of fairly similar proteins, where in one case the melting temperature of purified protein was very different from the T_m measured in CFE (unpublished data), while it was nearly identical in the other case.[18] Care has also to be taken since this approach can result in off-target peaks, potentially overlapping with the desired signal.

Equipment

- qPCR Machine (Bio-Rad, CFX96 C1000 Touch Thermal Cycler)
- 96-Well qPCR Plates (Bio-Rad, iQ High-Profile, 2239441)
- Adhesive qPCR Plate Sealing Film (Bio-Rad, Microseal 'B', MSB1001)
- Multichannel pipette

Buffers and Reagents

- Solution of target protein with a concentration of at least 0.3 mg/mL
- Fluorescent dye (if the protein does not contain a flavin)

Procedure

1. In a qPCR machine, set up a melt curve program, where the temperature ramps from 20 °C to 99 °C with 0.5 °C increments after a 10 seconds delay and the fluorescence is measured at each interval. Choose the emission filter such that it has the highest overlap with the emission spectrum of free flavin (peak at 524 nm) or the dye used[t] (see also introduction of this section).
2. Transfer 20 µL of protein solution to a 96-well qPCR Plate.[u] Unless you have indications that the protein concentration has a noticeable effect on the T_m, there is no need to correct for concentration variations.
3. For proteins not containing flavins, add a suitable dye (see introduction of this section).
4. Eliminate bubbles by centrifuging the plate on top of a 96 deep-well plate. Seal the plate with qPCR plate sealing film and use a sheet of lint-free tissue to fix the film air-tight and remove fingerprints and other residues.
5. Place the plate in the qPCR instrument and start the melt curve program.

[t] Bio-Rad has a very convenient chart with numerous commercial fluorophores and their excitation and emission data. Perform a web search for “Bio-Rad fluorophore reference guide” to retrieve the pdf.

[u] It is necessary to use dedicated qPCR plates and sealing foils, permeable for the excitation and emission wavelength range.

6. Use the instruments software or manually determine the 1st derivative of the fluorescence signal as a function of temperature. The T_m is defined as the temperature where this curve has a maximum.

Combining mutations

The most thermostable single mutants should be verified for preserved catalytic activity before proceeding with the combination of mutations aimed to obtain a final highly thermostable variant. With the FRESCO approach, typically 10-20 mutations are obtained that lead to an increased stability. As the next step in the engineering approach, one would like to combine the mutations in a highly stable mutant. Often, some of the individual mutations are not compatible, though. In order to identify efficiently the most optimal combination of mutations, several strategies can be followed.

One obvious approach is to directly generate the mutant containing all the mutations that were shown to increase the T_m on a single mutant level. This approach has proven successful in a few cases where we applied the FRESCO scheme. However, it often results in an inactive enzyme. If only relatively few (<5) mutations were identified, we recommend to introduce mutations by successive rounds of QuikChange (see the "QuikChange library creation" section), and to express, purify and determine the T_m of all intermediates, in order to verify an additive effect of each mutation. We usually start with adding mutations that do not alter the charge of the protein, and generally go from highest to lowest effect on T_m. In many cases it is found that one or several mutations give a problem, once combined with others. It may be that a mutation does not add to the T_m of an already more stable combined mutant, or that it decreases the T_m, or even leads to expression of insoluble protein. In these cases, it is preferable to omit that particular mutation in the final variant. However, it is possible that a mutation causes this problem only in a certain combination with other mutations, while it would be beneficial in another. If, for example, a highly stabilizing point mutant leads to insoluble protein after already having combined 5 other single mutants, one would like to know whether the insolubility is caused by a specific interaction with only one of the previous mutations. If that was the case, and if the effect on T_m of the first mutation was lower in comparison, it would be better to remove the earlier, rather than the newly added, mutation. To investigate this, however, one would have to remove one-by-one all the previous mutations, and express, purify and compare the T_m again. If this occurs with several mutations, the effort and time spent becomes impractical.

As an alternative approach, we developed two strategies that aim to create a library of randomly combined mutations. This allows screening of all possible combinations for up to around eight mutated positions. Including more positions would result in a library size that is too big to be practicably screened in a 96-well format, unless a robotic system can be used.[19] If more than eight positions are to be combined, the mutations can be split in sets, and the same or different strategies for combinations can be applied to each set.
The described methods are based on the generation of one DNA fragment per targeted position and subsequent cloning of a shuffled pool of wild-type and mutant fragments. The methods vary with respect to the cloning techniques applied—Golden Gate and Gibson cloning—as well as the fragment generation, which is either PCR based, or through more and more affordable commercial gene synthesis. Although these techniques can in principle be combined indiscriminately, we present here two successfully applied work-flows as alternatives: Golden Gate cloning from two synthetic genes, as a currently more expensive but very easy and fast method once a suitable cloning vector is available, and Gibson cloning of PCR fragments, as a very affordable, albeit more work-intensive method.

Golden Gate gene shuffling

The first strategy proposed for randomly shuffled mutations is based on Golden Gate cloning methodology [20]. The method described and schematized in Figure 4 is a one-pot reaction that requires two donor vectors, one containing the wild-type gene sequence and the other the mutated gene sequence, plus a destination vector. Previously described methods implicated the cloning of each module, flanked by BsaI sites in different vectors and addition of type IIS restriction enzyme sites by separate PCRs for each module.[20-21] We simplified this approach developed by involving two synthetic genes: one wild-type version and one with the selected mutations, each containing BsaI restriction sites flanking sections of the gene that contain the target mutation sites. The modules are flanked by two mirrored BsaI sites with four overhang nucleotides, thus determining the end of one module and the start of the next. Restriction and ligation leads to the excision of the BsaI sites, which leads to a scarless ligation. Having the BsaI restriction sites already inside the synthetic genes allows one to obtain a shuffled library in a single restriction-ligation reaction using only three vectors. The Golden Gate restriction-ligation proved to be very efficient: 97% of the colonies contained the correct restriction pattern and we obtained 65 uniquely shuffled clones

from one 96-well plate. The main advantage of such a method is the rapidity, because the synthetic gene design takes less than a day and the library can be obtained in another day. On the other hand, although synthetic gene synthesis is becoming relatively cheap, the extra cost can remain a drawback.

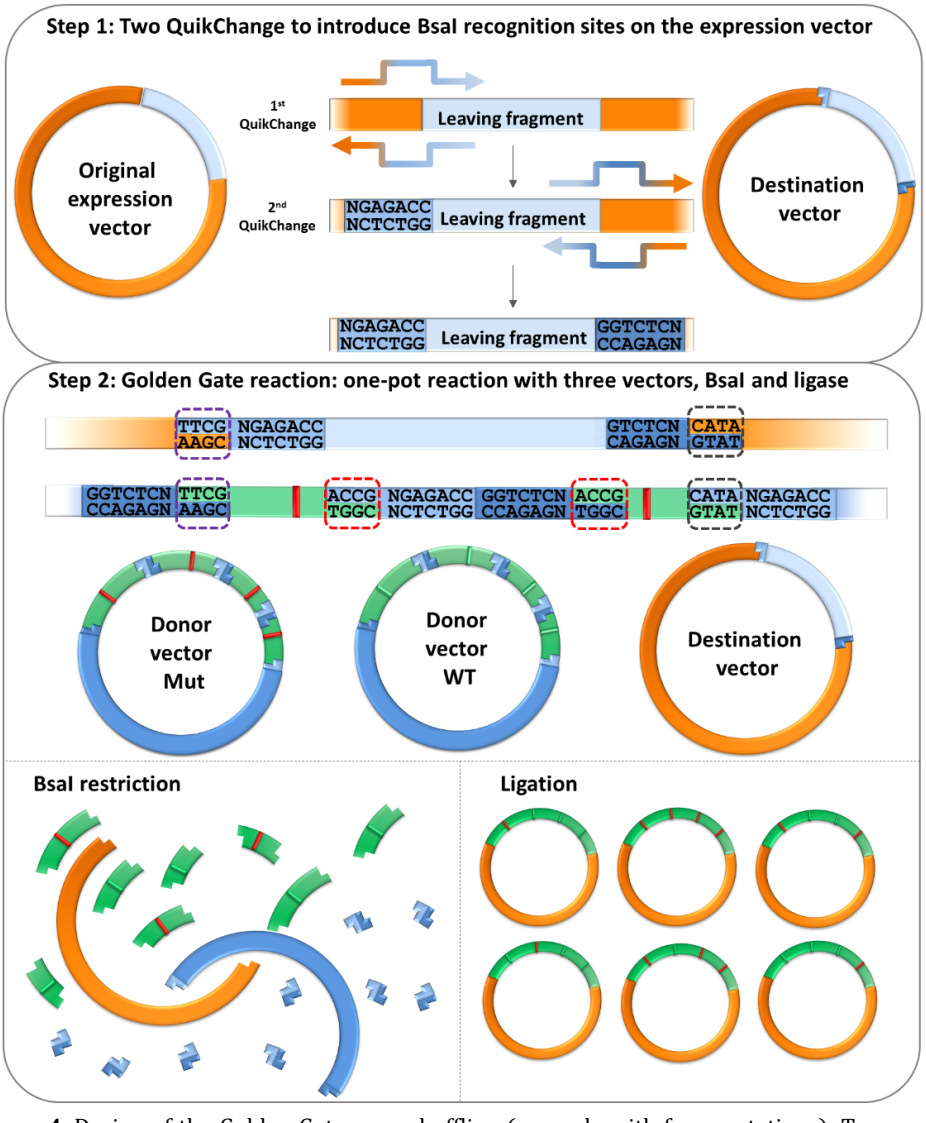

Figure 4. Design of the Golden Gate gene shuffling (example with four mutations). Top panel: generation of the destination vector. Starting from the original expression vector, two BsaI sites flanking a leaving fragment are introduced by QuikChange mutagenesis. Additional QuikChange mutagenesis may be necessary to remove unwanted BsaI recognition sites occurring in the vector. Bottom panel: one-pot Golden Gate reaction. In the case of the donor vectors, the BsaI sticky ends are designed to pair with the sticky ends of the destination vector. Moreover, in the donor vectors the BsaI recognition sites between two modules are flanked and mirrored. The 4

nucleotide overhang at the end of one fragment is replicated at the beginning of the next one to avoid nucleotide loss after the ligation. The donor vectors have the same fragment arrangement. The blue sections (donor vectors backbone, BsaI sites and leaving fragment of the destination vector) are lost after the ligation. The shuffled library consists of destination vectors with a random combination and correct order of the 4 gene modules (with or without mutations).

Equipment

- Synthetic gene synthesis service (Genscript)
- PCR machine
- NanoDrop (Thermo Scientific, NanoDrop 1000)

Buffers and Reagents

- Plasmid isolation kit (Qiagen, QiaPrep, 27106)
- BsaI (NEB, R3535L)
- T4 ligase 30 WU (Thermo Scientific, EL0013)
- T4 ligase buffer (Thermo Scientific)
- Destination vector with leaving BsaI sites
- LB media with and without agar and antibiotics
- SOC medium
- Chemically competent *E. coli* cells (see the "QuikChange library creation" section)

Procedure

1. Design and construct the destination vector.
   a. The destination vector can be designed based on any vector desired for expression. First, it has to be ensured that there are no BsaI sites are in the vector. If they occur, they have to be removed by QuikChange, introducing silent mutations in coding regions, or otherwise mutations that don't alter the vector's properties.
   b. Introduce BsaI sites at the sites where the donor gene is supposed to enter. An unrelated placeholder (leaving fragment) sequence can be kept in between. The donor vectors and the destination vector should confer different antibiotic resistances. In this way the selection after the transformation will lead to the exclusion of clones with the donor vectors.
2. *In silico* design of the two synthetic genes.
   a. First identify the mutagenesis sites. The mutations must be separated by at least one codon (otherwise the sticky ends will not be the same).
   b. Design the two synthetic genes identically (one with mutations and one without).

c. Add the 4 bp of the sticky ends of the destination vector at the beginning and at the end of your gene fragment. They will have to match with the sticky ends produced after BsaI cuts the destination vector.
d. To separate all the mutation sites, select 4 freely chosen bp which will delimit the fragments, with the following precautions:
   i. The mutation sites must be inside of the modules, not on the selected 4 bp.
   ii. The sequence of 4 bp must be unique for at least 2 bp out of 4.
   iii. The 4 bp must not be palindromic.
e. Once the fragments are identified, place the BsaI recognition sequence (see Figure 4):
   i. For the first and last fragment: add NGAGACC after the first 4 bp and GGTCTCN before the last 4 bp.[v]
   ii. Next add the double BsaI recognition site NGAGACCGGTCTCN between the first and second module (and so on for the next modules). Replicate the chosen 4 bp that are before the NGAGACC after the GGTCTCN (this is to get a scarless ligation without base-pair loss).

3. Order the two synthetic genes. If you order them already cloned in a vector[w] they can be used directly in the one-pot reaction, otherwise a cloning step is needed.
4. Set the one-pot reaction with 20 µL of final volume and the following components:

Table 4. Golden Gate cloning reaction

Component	Final Concentration
Destination vector	3.75 ng/µL
Donor vector wild-type	2.5 ng/µL
Donor vector with mutations	2.5 ng/µL
T4 DNA ligase buffer	1x
T4 DNA ligase	1.5 U/µL
BsaI 1	1 U/µL

[v] The fragmentation of the two genes must be the same (BsaI cutting sites must be placed at the same position and the same sticky ends need to be used). To check the uniqueness of the sites a useful tool is provided by New England Biolabs in collaboration with Benchling (https://goldengate.neb.com/). NGAGACC and GGTCTCN represent BsaI recognition sites in different orientations (N is a random base pair that can be any of AGTC).

[w] Companies like GenScript usually offer also a cloning service. Ordering the synthetic gene already cloned in a vector will save time, because it can be used directly for the one-pot reaction.

5. In order to directly create a fully mutated variant, another assembly can be performed without the wild-type donor plasmid.
6. The thermocyler program alternates between the optimal temperatures of BsaI and the ligase, followed by a final digestion step to cut all template and BsaI containing products, and an enzyme inactivation:

Table 5. Golden Gate cloning program

Temperature	Time	Cycles
37 °C 16 °C	5 min 10 min	50
50 °C	10 min	1
80 °C	10 min	1

7. The restriction-ligation reaction (5 µL) can be used to transform 100 µL of chemically competent NEB 10β cells plated on LB agar plates with antibiotic.

Gibson shuffling

This protocol makes use of a multiple site directed mutagenesis procedure developed by Mitchell et al. which allows simultaneous mutations at several positions on a plasmid [22]. Based on this method we have developed a protocol to construct a library of mutant combinations. First, the fully mutated plasmid is created by PCR-amplifying fragments with mutated primers and subsequent Gibson assembly. In the second step, PCR is used again to generate a mix of fragments with and without the mutations, and another Gibson assembly results in the shuffled library. Although this PCR based method requires more work when compared with the Golden Gate-based protocol, it can be applied when a limited budget prohibits custom gene synthesis. Moreover, it is particularly useful if it is desired to randomize a position for more than one mutant variant. The FRESCO procedure commonly predicts several mutations for the same position, and if more than one variant proved to be stabilizing on single mutant level, this method allows the inclusion of all variations in the library, without requiring another full custom-made gene.

If two mutations are separated by 5 to 15 amino acids (or 15 to 45 bp), extra attention is required, because all of the individual steps (PCR, DNA purification, and cloning) will be less efficient for smaller fragments. The lowest limit for Gibson assembly can be considered as 150 bp—however, this length includes the length of the forward and reverse primers. Since the primer should not exceed 75 bp (the ideal is 30-50), and the mutation should always be at least 15 bp from the termini (the ideal is in the middle of the primer), the minimum

separation of two mutations is approximately 11 amino acids, or 33 bp for the smallest possible fragment. If the two desired mutations are closer together they have to be included on the same primer. Due to the same limitations, the maximum separation of two mutations on one primer is ≈14 amino acids, or 42 bp. Thus, the method in principle covers any separation of two mutations.
The maximum amount of fragments is another limitation. However, the exponential increase of possible combinations will usually prohibit a library of more than 8 target positions, and assembling 8 fragments is still feasible. Efficiency does, however, decrease with more fragments, and the general considerations for highest success rates, as pointed out before, apply here as well.

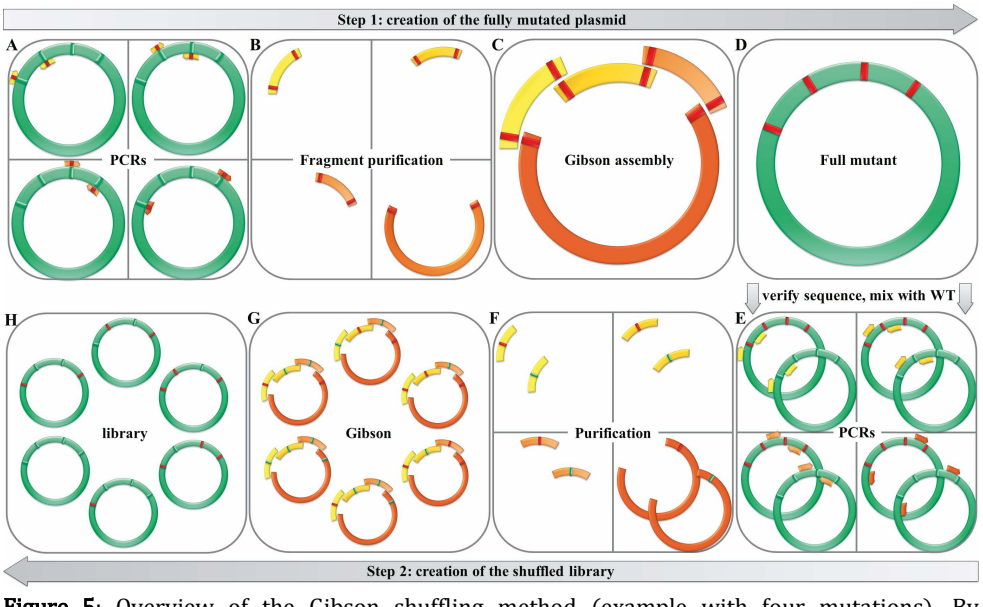

Figure 5: Overview of the Gibson shuffling method (example with four mutations). By performing PCR on the WT plasmid with primers that contain the mutations (A), fragments are generated that overlap at the site of the mutation where the primer originally bound. After purifying the fragments (B), they are assembled in a Gibson cloning step (C) to result in the fully mutated plasmid (D). This plasmid is mixed with the original WT plasmid to serve as a template in a second PCR step (E) where primers bind in between the mutated regions. This creates pools of fragments that do or do not contain the mutation, which have to be purified (F) and assembled (G) again, to yield the shuffled library (H).

Equipment

- DNA sequence analysis software (Biomatters, Geneious R8)

- RNAfold software, part of the open source ViennaRNA package, web server available at http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi [23]
- PCR machine
- Standard agarose gel electrophoresis equipment (Bio-Rad, Mini-Sub Cell GT)
- UV transilluminator, or preferably blue-light variant[y]
- NanoDrop

Buffers and Reagents

- High fidelity polymerase and reaction components
- 1% agarose gel
- Gel extraction kit, if necessary dedicated for small fragments (Qiagen, 28606)
- Gibson assembly master mix (NEB, E2611L)[z]
- Chemically competent *E. coli* cells (see the "QuikChange library creation" section)

Procedure

1. Select sites for mutagenesis and verify they are in appropriate distance (see introduction).
2. Primer design is critical to the success of the method, so strict care is required.
   a. Choose the mutagenesis codon and select around 13-25 bp upstream and downstream, ensuring on both ends to apply the GC-clamp rule [A/T][A/T/C/G][C/G] (i.e. end on one or two G or C, but not more).
   b. All primers should have a similar T_m of at least 55 °C, and a G/C content of 40-60%.
   c. Mutate the codon to the desired amino acid making sure to avoid rare codons[b]
   d. For successful Gibson assembly, it is critical to avoid secondary structures (ss) of the homologous regions (which corresponds to the primer). The most intuitive way of removing ss from the initial primer is by visual inspection, via ss prediction software, such as the RNAfold utility of the ViennaRNA package. This tool is freely available as a web server, but we commonly use its implementation in the bioinformatics software Geneious. Be sure to predict the ss at 50 °C and inspect the result (Figure 6).

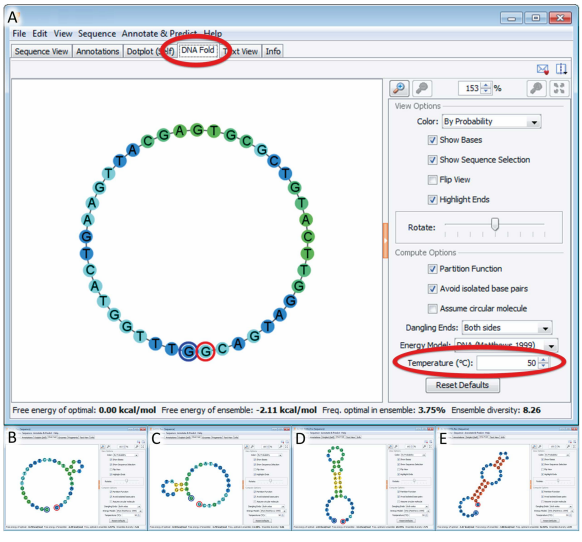

Figure 6. Secondary structure of primers/homologous regions. In Geneious, go to the DNA-fold tab, and select 50 °C. Structure A) is ideal, B) good, C) tolerable, D) critical and E) unacceptable.

e. Weak ss (probability < 0.5, color code green/blue), especially when distant from the termini, can be accepted.
f. Stronger ss (yellow/red), especially when close to the termini, should be removed by i) change of mutant codon; ii) introduction of (not more than a few) silent mutations for other codons, not too close to the termini; and/or iii) removal/addition of nucleotides at the 5' and/or 3' end of the primer (keeping the GC-clamp).
g. In this way, the best possible compromise has to be found between avoiding rare codons, removing strong ss and avoiding too many silent mutations, especially close to the termini.
h. The reverse complement of the generated sequence serves as the reverse primer for the previous fragment.
i. Repeat steps 2a–2h for all mutant positions.
j. Design a second set of primers which bind in between the mutated positions (Figure 6E). They should again be 30-50 bp long, contain a GC-clamp on both sides and the exact position may be chosen freely with the intention to again keep the ss minimal. As before, the reverse complements act as primers for the adjacent fragments.

3. Using the wild-type plasmid as a template, perform individual PCRs to generate fragments containing the mutations by combining primers as color-coded in Figure 6A: the forward primer of position 1 with the reverse

primer of position 2, etc. Make 40 µL reactions, and stick to the recommendations of the polymerase's manufacturer. When using *PfuUltra* II, see step 2 in the "QuikChange library creation" section for recipe and program recommendations.

4. Purify the fragments by gel purification after agarose gel electrophoresis[x]
   a. Prepare a 1% agarose gel with pockets large enough to hold 40 µL. Load an appropriate ladder and the PCR products with loading buffer and run the gel at 85 V, 30 mins.
   b. Avoid taking an image of the gel, to minimize the exposure time of the DNA to UV.[y] Verify the correct size of the fragments and the PCR success directly while excising the bands from the gel with a scalpel. Then purify the DNA using a gel extraction kit. Elute in 20 µL (intense band) or 10 µL (faint band).
   c. Measure DNA concentration using a NanoDrop.
5. Gibson assembly reaction
   a. Thaw two 15 µL aliquots of Gibson assembly master mix[z] on ice.
   b. Calculate the correct amounts of vector to insert **molar ratios**, e.g. by using the NEB ligation calculator (http://nebiocalculator.neb.com/#!/ligation). This tool gives the required ng of insert for different ratios, upon input of a certain insert and vector length. For optimal assembly, use 100 ng of vector (i.e. the largest fragment). For fragments <200 bp, use 1:5, for 200-400 bp use 1:3, and for >400 bp use a 1:2 molar ratio.

[x] In principle it is possible to replace this step with DpnI digestion. We recommend a gel extraction anyway, because i) DpnI digestion is never 100% efficient in removing template, while gel purification is; ii) it will immediately show success or failure of the PCR, iii) a clean-up step is also required after DpnI digestion; iv) even if yields are low, the required amounts in the assembly are very small; and v) it is the faster method (gel 30 mins + extraction 15 min, vs. ≥ 2 h DpnI digest + clean-up).

[y] DNA damage occurs very rapidly under UV light. A blue-light transilluminator prevents UV damage and allows band excision without time pressure.

[z] An alternative to the commercial Gibson mix, is to prepare it in-house:

   a. Prepare first a 5x reaction buffer containing 25 % w/v PEG-8000, 500 mM Tris/HCl pH 7.5, 50 mM $MgCl_2$, 50 mM DTT, 1 mM dATP, 1 mM dCTP, 1 mM dGTP, 1 mM dTTP, and 5 mM NAD. Sterilize by filtration and freeze at −20 °C in aliquots of 100 µL.
   b. d Prepare next the Gibson assembly master mix containing 1x reaction buffer, 0.0053 U/µL T5 exonuclease, 0.033 U/µL Phusion polymerase, and 5.33 U/µL Taq DNA ligase. Freeze at -20 °C in aliquots of 15 µL.

c. Next, calculate the volumes corresponding to this mass, considering the concentrations of your fragments. The total volume of the reaction mix is 20 µL, leaving 5 µL for fragments. If the sum of the volumes exceeds 5 µL, decrease the vector mass to 50 ng and use half of the fragments. If it still exceeds 5 µL, increase the total volume by using two Gibson MM aliquots, or try to get higher concentrations by optimizing PCR, extraction yield, and using lower elution volumes. Fill the Gibson assembly mix up to 20 µL with sterile MilliQ water.
d. As a negative control, prepare an identical sample excluding the smallest fragment.
e. Run the isothermal assembly at 50 °C for 1 h.

6. For transformation, use 3.5 µL of the assembly mix (no clean-up required) and follow the recommendations outlined in step 4-7 of the "QuikChange library creation" section.
7. Pick several colonies, and grow them overnight in 5 mL LB medium. Make a plasmid preparation and send for sequencing.
8. Once verified that the full mutant was generated, perform another round of individual PCRs.
   a. For the template, create a mix of the wild-type plasmid with the fully mutated plasmid. Check the concentration with the NanoDrop and make a precisely 1:1 mixed stock solution. Volumes that require reasonable pipetting amounts (> 10 µL) should be used for improved precision.
   b. Prepare PCR reactions with primer combinations as color-coded in Figure 1E: the forward primer binding before position 1 with the reverse primer before position 2, etc. Make 40 µL reactions, and stick to the recommendations of the polymerase's manufacturer. When using *PfuUltra* II, see step 2 in the "QuikChange library creation" section for recipe and program recommendations.
9. Proceed in exactly the same way as described in steps 4-6 to purify and assemble the fragments, resulting in a shuffled library of mutant combinations.

Final stabilized mutant

After creating the shuffled library by either of the above methods, the success and shuffling efficiency should be verified by sequencing. In principle it is sufficient to sequence a mix of the entire library. In that case, individual clones are screened for thermostability without specific sequence information, and only the most stable variants are then sequenced again. A greater amount of

information is gained if a statistically relevant portion of individual clones (e.g. 10 % of the library size) or even the entire library is sequenced. In that case, one can draw more conclusions about the effect of different combinations and determine incompatible patterns as well as beneficial arrangements.
After determination of the mutant with the combination that yields the highest melting temperature, a verification of preserved catalytic activity is self-evident. This confirmation has to be adapted to one's specific requirements and is not described in this protocol. It is important to notice, however, that a common observation is an upshift of the temperature for optimal enzyme activity. Thus, even though one could observe a lowered activity at the assay temperature that is standard for the wild-type, a higher activity is likely to be observed at a higher temperature. Nevertheless, if the activity is generally lowered, one can test the mutants next in rank with regard to the T_m to find the ideal compromise. Since the computational part of the FRESCO procedure already avoids mutations near the enzyme's active site, we usually do not observe an impaired catalytic activity for the thermostable mutants.

Equipment

- Sterile disposable or autoclavable 96-square deep well plates
- Sterile 15 mL tube (Greiner)
- Sterile toothpicks
- DNA sequence analysis software

Buffers and Reagents

- LB medium

Procedure

1. Analyze the transformed combination library generated by either of the two methods outlined the "Combining mutations" section. By using the highly competent cells (step 1 of the "QuikChange library creation" section), one should obtain a few hundred colonies. Transform more of the shuffled library assembly mix, if necessary.
2. Estimate the amount of clones that have to be analyzed to achieve a good coverage of the library. By rule of thumb, analyze three times the size of the library. E.g. for one mutant amino acid at seven positions, the number of possible combinations is $2^7 = 128$. Threefold oversampling gives 384 clones, which corresponds to exactly four 96-well plates.
3. Prepare the 96-well plates containing LB medium and appropriate antibiotics and a sterile 15 mL tube filled with 14 mL LB.

4. Pick individual colonies from the library plate with sterile toothpicks and place them in the 96-well plate to inoculate one well per colony. Keep some wells for the wild-type, the fully mutated variant, and non-inoculated contamination controls.
5. Remove the toothpicks one by one, and dip each of them in the tube with LB medium before discarding them.
6. Mix the tube inoculated with all mutants, and grow 5 mL of culture overnight, to isolate the mix of plasmids and send it for Sanger sequencing.
7. Analyze the sequencing chromatogram, and verify that the peaks for the mutated nucleotides are approximately of the same intensity as the wild-type nucleotides (Figure 7). If either WT or mutant signals seem to dominate, sequence a significant portion of individual clones to determine if there is a bias in the library. If that is the case, repeat the assembly reaction of the library generation, and pay extra care to using exactly the same amounts of wild-type and mutant plasmid.

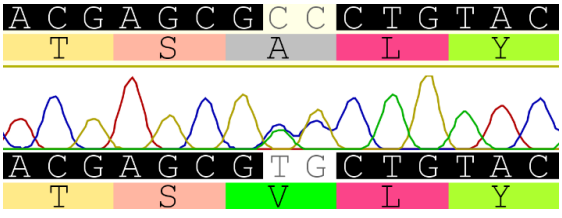

Figure 7. Example of a sequencing chromatogram obtained from a successful mix of WT and mutant. The third amino acid position shows an even mix of peaks for GCC (Ala) and GTG (Val).

8. To produce and screen the individual clones of the library, proceed in exactly the same way as outlined in the "Expression and protein purification in a 96-well plate" and "Melting temperature screening" sections.

Summary and conclusions

The fourth wave of protein engineering is being announced as the synergy of directed evolution with rational and computational design to create more robust enzymes with better specificities or novel reactivities. Yet, one often still faces the quite tedious experimental work in generating and evaluating the corresponding libraries. With the protocols described above, we hope to contribute to efficient strategies to create improved enzymes.

After promising pioneering work done using FRESCO for thermostabilization of limonene epoxide hydrolase (T_m increase of 35 °C) [3], haloalkane dehalogenase (T_m increase of 23 °C) [24] and halohydrin dehalogenase (T_m increase of 28 °C) [4], we have decided to further expand the application of FRESCO to flavin-containing enzymes. Thereby, the FRESCO approach was proven to be useful in engineering stability in a flavoenzyme: a more robust variant of 5-(hydroxymethyl)furfural oxidase (HMF oxidase) was engineered [18]. As described in the protocol above, we could take advantage of the fluorescent properties of the flavin cofactor to rapidly identify improved HMF oxidase variants. HMF oxidase was shown to be able to convert HMF into the polymer building block FDCA [25]. By engineering an HMF oxidase that is efficient in converting HMF into FDCA and is thermostable, a valuable biocatalytic tool is now available to be used in strategies for producing renewable-based polymers. It demonstrates that the FRESCO methodology combined with the protocols described above is a powerful combination for engineering tuned (flavo)enzymes. Because many structures are available for various biocatalytically interesting flavoenzymes which display an astonishing catalytic flexibility, they are ideally suited in this methodology or other knowledge-based enzyme engineering approaches. There is a growing list of studies in which a flavoenzyme is key for the design of a new metabolic pathway or biocatalytic cascade reaction [26]. Examples of newly engineered or proposed metabolic pathways that include flavin-containing enzymes are: i) the recent illustrative example that employs a bacterial flavin-containing monooxygenase and a glucosyltransferase for a sustainable indigo dyeing strategy [27], ii) the whole cell conversion of limonene into chiral carvolactone with the help of a flavin-containing Baeyer-Villiger monooxygenase [28], iii) the redesign of cyclohexanone monooxygenase to enable production of methyl propanoate [29] and iv) the use of flavoprotein oxidases fused to a peroxidase for biocatalytic cascade reactions [30]. Except for FRESCO, the protocols described here can also be used as part of other (flavo)enzyme engineering efforts. Therefore, we hope that they will be adopted by other laboratories for the generation of improved proteins for biocatalysis.

Acknowledgements

Part of the research for this work has received funding from the European Union (EU) project ROBOX (grant agreement 635734) under EU's Horizon 2020 Programme Research and Innovation actions H2020-LEIT BIO-2014-1.

References

1 Marcheschi, RJ; Gronenberg, LS; Liao, JC. Protein Engineering for Metabolic Engineering: Current and Next-Generation Tools. *Biotechnology Journal* **2013** (8) 545-555.

2 Liu, Y; Shin, H-d; Li, J; Liu, L. Toward Metabolic Engineering in the Context of System Biology and Synthetic Biology: Advances and Prospects. *Appl. Microbiol. Biotechnol.* **2015** (99) 1109-1118.

3 Wijma, HJ; Floor, RJ; Jekel, PA; Baker, D; Marrink, SJ; Janssen, DB. Computationally Designed Libraries for Rapid Enzyme Stabilization. *Protein Eng. Des. Sel.* **2014** (27) 49-58.

4 Arabnejad, H; Dal Lago, M; Jekel, PA; Floor, RJ; Thunnissen, A-MWH; Terwisscha van Scheltinga, AC; Wijma, HJ; Janssen, DB. A Robust Cosolvent-Compatible Halohydrin Dehalogenase by Computational Library Design. *Protein Engineering, Design and Selection* **2017** (30) 175-189.

5 Forneris, F; Orru, R; Bonivento, D; Chiarelli, LR; Mattevi, A. *Thermo*fad, a *Thermofluor*-Adapted Flavin *ad hoc* Detection System for Protein Folding and Ligand Binding. *FEBS J.* **2009** (276) 2833-2840.

6 Wijma, HJ; Fürst, MJLJ; Janssen, DB. A Computational Library Design Protocol for Rapid Improvement of Protein Stability: Fresco. In: *Protein Engineering: Methods and Protocols*; Bornscheuer, UT, Höhne, M (Eds.), Springer: New York, NY; **2018**; Vol. 1685; pp 69-85.

7 Liu, H; Naismith, JH. An Efficient One-Step Site-Directed Deletion, Insertion, Single and Multiple-Site Plasmid Mutagenesis Protocol. *BMC Biotechnol.* **2008** (8) 91.

8 Zheng, L; Baumann, U; Reymond, J-L. An Efficient One-Step Site-Directed and Site-Saturation Mutagenesis Protocol. *Nucleic Acids Res.* **2004** (32) e115.

9 McInerney, P; Adams, P; Hadi, MZ. Error Rate Comparison During Polymerase Chain Reaction by DNA Polymerase. *Mol. Biol. Int.* **2014** (2014) 1-8.

10 Sun, D; Ostermaier, MK; Heydenreich, FM; Mayer, D; Jaussi, R; Standfuss, J; Veprintsev, DB. Aascan, Pcrdesign and Mutantchecker: A Suite of Programs for Primer Design and Sequence Analysis for High-Throughput Scanning Mutagenesis. *PLoS One* **2013** (8) e78878.

11 Boël, G; Letso, R; Neely, H; Price, WN; Wong, K-H; Su, M; Luff, JD; Valecha, M; Everett, JK; Acton, TB; Xiao, R; Montelione, GT; Aalberts, DP; Hunt, JF. Codon Influence on Protein Expression in *E. coli* Correlates with Mrna Levels. *Nature* **2016** (529) 358-363.

12 Green, R; Rogers, EJ. Transformation of Chemically Competent *E. coli*. *Changes* **2013** (29) 329-336.

13 Hanahan, D; Jessee, J; Bloom, FR. Plasmid Transformation of *Escherichia coli* and Other Bacteria. In: *Methods in Enzymology*; Academic Press; **1991**; Vol. 204; pp 63-113.

14 Guzman, LM; Belin, D; Carson, MJ; Beckwith, J. Tight Regulation, Modulation, and High-Level Expression by Vectors Containing the Arabinose P_{bad} Promoter. *J. Bacteriol.* **1995** (177) 4121-4130.

15 Duetz, WA. Microtiter Plates as Mini-Bioreactors: Miniaturization of Fermentation Methods. *Trends Microbiol.* **2007** (15) 469-475.

16 Duetz, WA; Witholt, B. Oxygen Transfer by Orbital Shaking of Square Vessels and Deepwell Microtiter Plates of Various Dimensions. *Biochem. Eng. J.* **2004** (17) 181-185.

17 Hermann, R; Lehmann, M; Büchs, J. Characterization of Gas–Liquid Mass Transfer Phenomena in Microtiter Plates. *Biotechnol. Bioeng.* **2003** (81) 178-186.

18 Martin, C; Ovalle Maqueo, A; Wijma, HJ; Fraaije, MW. Creating a More Robust 5-Hydroxymethylfurfural Oxidase by Combining Computational Predictions with a Novel Effective Library Design. *Biotechnology for Biofuels* **2018** (11) 56.

19 Reetz, MT; Kahakeaw, D; Lohmer, R. Addressing the Numbers Problem in Directed Evolution. *ChemBioChem* **2008** (9) 1797-1804.

20 Engler, C; Gruetzner, R; Kandzia, R; Marillonnet, S. Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type Iis Restriction Enzymes. *PLoS one* **2009** (4) e5553.

21 Zhang, F; Cong, L; Lodato, S; Kosuri, S; Church, GM; Arlotta, P. Efficient Construction of Sequence-Specific Tal Effectors for Modulating Mammalian Transcription. *Nat. Biotechnol.* **2011** (29) 149-154.
22 Mitchell, LA; Cai, Y; Taylor, M; Noronha, AM; Chuang, J; Dai, L; Boeke, JD. Multichange Isothermal Mutagenesis: A New Strategy for Multiple Site-Directed Mutations in Plasmid DNA. *ACS Synth. Biol.* **2013** (2) 473-477.
23 Lorenz, R; Bernhart, SH; Höner zu Siederdissen, C; Tafer, H; Flamm, C; Stadler, PF; Hofacker, IL. Viennarna Package 2.0. *Algorithms for Molecular Biology* **2011** (6) 26-40.
24 Floor, RJ; Wijma, HJ; Colpa, DI; Ramos-Silva, A; Jekel, PA; Szymanski, W; Feringa, BL; Marrink, SJ; Janssen, DB. Computational Library Design for Increasing Haloalkane Dehalogenase Stability. *ChemBioChem* **2014** (15) 1660-72.
25 Dijkman, WP; Groothuis, DE; Fraaije, MW. Enzyme-Catalyzed Oxidation of 5-Hydroxymethylfurfural to Furan-2,5-Dicarboxylic Acid. *Angew. Chem. Int. Ed. Engl.* **2014** (53) 6515-8.
26 Wang, J; Yang, Y; Zhang, R; Shen, X; Chen, Z; Wang, J; Yuan, Q; Yan, Y. Microbial Production of Branched-Chain Dicarboxylate 2-Methylsuccinic Acid Via Enoate Reductase-Mediated Bioreduction. *Metab. Eng.* **2018** (45) 1-10.
27 Hsu, TM; Welner, DH; Russ, ZN; Cervantes, B; Prathuri, RL; Adams, PD; Dueber, JE. Employing a Biochemical Protecting Group for a Sustainable Indigo Dyeing Strategy. *Nat. Chem. Biol.* **2018** (14) 256-261.
28 Oberleitner, N; Ressmann, A; Bica, K; Gärtner, P; Fraaije, M; Bornscheuer, U; Rudroff, F; Mihovilovic, M. From Waste to Value–Direct Utilization of Limonene from Orange Peel in a Biocatalytic Cascade Reaction Towards Chiral Carvolactone. *Green Chemistry* **2017** (19) 367-371.
29 van Beek, HL; Romero, E; Fraaije, MW. Engineering Cyclohexanone Monooxygenase for the Production of Methyl Propanoate. *ACS Chem. Biol.* **2017** (12) 291-299.
30 Colpa, DI; Lončar, N; Schmidt, M; Fraaije, MW. Creating Oxidase–Peroxidase Fusion Enzymes as a Toolbox for Cascade Reactions. *ChemBioChem* **2017** (18) 2226-2230.