

Combinatorial Peptide Library

Subjects: Biotechnology & Applied Microbiology

Contributor: Tomaz Bratkovic, Krištof Bozovičar

Completely randomizing even relatively short peptides would require a library size surpassing the capacities of most platforms. Sampling the complete mutational space for peptides exceeding 8–9 residues is therefore practically impossible, and gene diversification strategies only allow for generation and subsequent interrogation of a limited subset of the entire theoretical peptide population. Directed evolution of peptides therefore strives to ascend towards peak activity through mutational steps, accumulating beneficial mutational over several generations, resulting in improved phenotype. We briefly discuss combinatorial library platforms and take an in-depth look at diversification techniques for random and focused mutagenesis.

Keywords: combinatorial peptide library ; random mutagenesis ; focused mutagenesis

1. Introduction

A **combinatorial peptide library** is a collection of fully or partially randomized peptides of defined length. Peptide libraries are classified either as synthetic (non-biological) or biological. As the name implies, **synthetic** libraries are produced through combinatorial peptide synthesis by allowing coupling of more than one amino acid residue per position. Here, individual peptides can be coupled to (and in fact synthesized on) beads in a *one-bead-one-compound* setting ^[1] and screened against a fluorophore-labelled target using fluorescence-activated sorting (Figure. 1A). An alternative approach to synthetic library screening is *positional scanning*, where the peptide library is divided into sub-libraries, each having one of the positions defined (i.e., occupied by a specific amino acid residue), while the rest of the peptide structure is randomized (Figure. 1B). Each of the sub-libraries is assayed for activity to deduce the consensus peptide sequence with optimal activity ^[2]. Conversely, in *iterative deconvolution* strategy optimal residues at individual positions are identified in a step-by-step fashion (Figure. 1C). Sub-libraries of randomized peptides but with different defined residues at first position are screened for activity, and the optimal residue in that position is retained in the second generation sub-library which contains peptides with different defined residues at the second position. This approach is repeated until all the positions have been interrogated ^[2]. The term **biological** library signifies that the peptides are produced on ribosomes. Here, it is essential that the phenotype (the encoded peptide) and the genotype (the encoding nucleic acid sequence) are physically linked, allowing correlation of peptide sequence and genetic information during screening, and identification of hits via nucleic acid sequencing ^[3]. In library platforms such as bacterial and yeast display, the peptides are expressed from fusion genes in frame with a gene for a cell surface protein. In phage display, peptides are anchored to viral protein coat, encapsulating genetic information, through fusions with capsid structural proteins. In ribosome and mRNA display, peptides are coupled to mRNA indirectly (non-covalently via ribosome complex) or directly (covalently via puromycin), respectively.

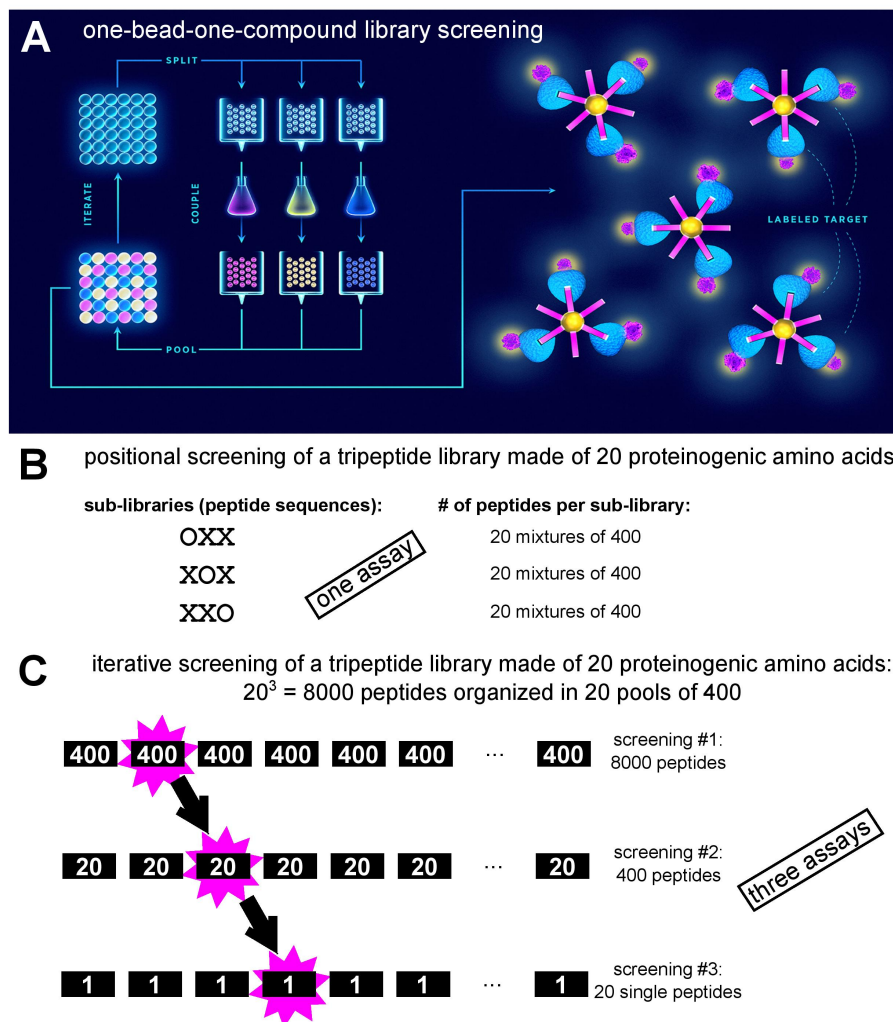


Figure 1: Depiction of various synthetic peptide library screening approaches. **A.** One-bead-one-compound library synthesis by the ‘split-and-mix’ method and screening via fluorescence-activated bead sorting after incubation with fluorophore-labelled target (adapted from [4]). **B.** Example of positional screening of a tripeptide library: 60 peptide mixtures are screened in parallel in a single assay (adapted from [2]). **C.** Example of iterative screening of a tripeptide library: three consecutive screens are performed with sub-libraries of decreasing diversity (adapted from [2]).

In the following sections, diversification and construction of biological combinatorial peptide libraries are reviewed. As the peptides are genetically encoded, the diversification step is performed at the level of DNA, either by synthesis or mutagenesis.

2. Peptide Library Design and Construction

Peptide maturation can be depicted as an ascent in a simplified fitness landscape (Figure. 2) in which the x-y coordinates denote the otherwise multidimensional genotype, and the z-axis represent the peptide’s “phenotypic” traits, e.g., target affinity. Ascending towards peak activity with mutational steps is the goal of directed evolution. Beneficial mutations accumulate over several generations upon selection pressure, resulting in improved phenotype [5].

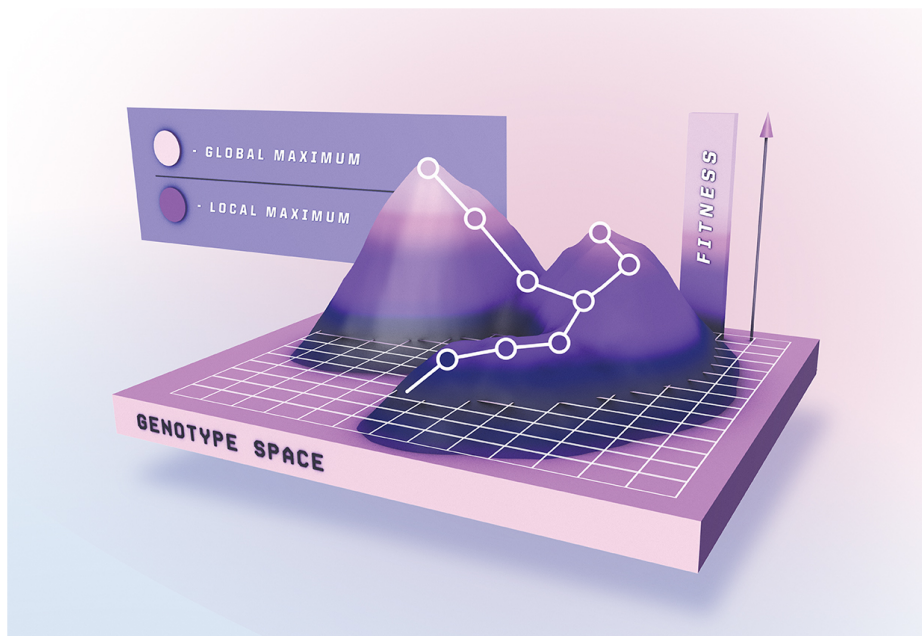


Figure 2: Maturation of a peptide depicted as ascent on a simplified fitness landscape. After each selection round, mutations are introduced into the enriched combinatorial library, and the next generation of peptides is screened for improved affinity and/or activity (adopted from [14]).

In general, library generation can be performed either through focused or random mutagenesis. The latter is usually used in the absence of structure-function relationship knowledge. In focused mutagenesis residues previously found to be essential for peptide activity are retained (or favoured over the rest of the building block set), while the others are (fully or partially) randomized. Of course, the odds that a library contains improved peptide variants are higher for those produced by focused mutagenesis. A plethora of mutagenesis methods can be used for gene diversification in library generation and we will briefly discuss them below.

2.1. Random Mutagenesis

Random mutagenesis based on physical and/or chemical mutagens is sufficient for traditional genome screening (gene inactivation), but it is not suitable for directed evolution due to limited mutational spectrum [6][7]. For library generation purposes, random mutagenesis can be performed *in vivo* in bacterial mutator strains that contain defective proofreading and repair enzymes (mutS, mutT, and mutD) [8][9][10]. Another approach in *E. coli* relies on mutagenesis plasmids (MP), which carry multiple genes for proteins affecting DNA proofreading, mismatch repair, translesion synthesis, base selection, and base excision repair, thereby enabling broad mutagenic spectra. MPs support mutation rates 322,000-fold over basal levels and are suitable for platforms based on bacterial and phage-mediated directed evolution [11]. Unfortunately, beside the library gene, mutator strains and MPs also induce deleterious mutations in host genome. In eukaryotes this was overcome by the development of orthogonal *in vivo* DNA replication apparatus, which in essence utilizes plasmid-polymerase pairs, limiting mutagenesis to a cytoplasm-only event [12]. Related phenomena are also known to occur in nature (e.g., the *Bordetella bronchiseptica* bacteriophage error-prone retroelement, which selectively introduces mutations into the gene encoding the major tropism determinant (Mtd) protein on the phage tail fibers [13]) and can be exploited for creating libraries [14].

One of the most established methods for *in vitro* random mutagenesis is the error-prone PCR (epPCR), first described in 1989 [15]. It works by harnessing the natural error rate of low-fidelity DNA polymerases, generating point mutations during PCR amplification. However, even the faulty *Taq* DNA polymerase is not erroneous enough to be useful for constructing combinatorial libraries under standard amplification conditions. The fidelity of the reaction can be further reduced by altering the amount of bivalent cations Mn^{2+} and Mg^{2+} , introduction of biased concentrations of deoxyribonucleoside triphosphates (dNTPs) [16], using mutagenic dNTP analogues [17], or adjusting elongation time and the number of cycles [18]. Random mutations can also be induced by utilizing 3'-5' proofreading-deficient polymerases [19][20][21].

Despite its popularity, epPCR suffers from limited mutational spectrum as it inclines to transitions ($A \leftrightarrow G$ or $T \leftrightarrow C$). Thus, epPCR-generated libraries are abundant in synonymous and conservative nonsynonymous mutations as a result of codon redundancy [5]. Ideally, all four transitions ($AT \rightarrow GC$ and $GC \rightarrow AT$) and eight transversions ($AT \rightarrow TA$, $AT \rightarrow CG$, $GC \rightarrow CG$, and $GC \rightarrow TA$) would occur at equal ratios, with the desired probability, and without insertions or deletions [22]. This problem has been addressed by the sequence saturation mutagenesis (SeSaM) [23] method, which utilizes deoxyinosine, a promiscuous base-pairing nucleotide that is enzymatically inserted throughout the target gene and later changed for

canonical nucleotides using standard PCR amplification of the mutated template gene. SeSaM was later improved with the introduction of SeSaM-Tv-II [124], which generates sequence space unobtainable via conventional epPCR by increasing the number of transversions. It employs a novel polymerase with increased processivity, allowing efficient read through consecutive base-pair mismatches. EpPCR has been successfully adopted for library generation in various platforms, including phage, [125], *E. coli* [126] and ribosome [127] display.

Alternatively, mutagenesis can be achieved by performing isothermal rolling circle amplification (RCA) under error-prone conditions. Using a wild-type sequence as a template, this method is able to generate a random DNA mutant library, which can be directly transformed into *E. coli* without subcloning [128]. RCA was advanced further, coupling it with Kunkel mutagenesis [129] (see below). Termed 'selective RCA' (sRCA), it operates by producing plasmids in *ung*⁻ (uracil-DNA-glycosylase deficient) *dut*⁻ (dUTP diphosphatase deficient) *E. coli* strain to introduce non-specific uridylation (dT → dU). After PCR with mutagenic primers, abasic sites are created by the uracil-DNA glycosylase in the uracil-containing template. Only mutagenized products are amplified by RCA, excluding non-mutated background sequences [130].

Although epPCR generates high mutational rates, the sequence space remains mostly untapped [131]. DNA shuffling is touted to be superior to epPCR and oligonucleotide-directed mutagenesis because it does not suffer from the possibility of introducing neutral or non-essential mutations from repeated rounds of mutagenesis [132]. DNA shuffling was the first *in vitro* recombination method and it involves random fragmentation of a pool of closely related dsDNA sequences and subsequent reassembly of fragments by PCR [133]. Such template switching generates a myriad of new sequences and improves library diversity by mimicking natural sexual recombination [134]. Meyer et al. [131] developed an approach where DNase I creates double-stranded breaks at the regions of interest, followed by denaturation and reannealing at homologous regions. Hybridized fragments then serve as templates and are subjected to repeated PCR rounds to form a whole array of new sequences. Improved methods were developed, eliminating the lengthy DNA fragmentation step. In 'staggered extension process' (StEP), polynucleotide sequences can be diversified through severely-abbreviated annealing/polymerase-catalyzed extension. In each cycle, growing fragments switch between different templates and anneal to them based on sequence complementarity. They then extend further and the cycle is repeated until full-length mosaic sequences are formed [135]. Another ingenious method for creating random customized peptide libraries by Fujishima et al. [136] works by shuffling short DNA blocks with dinucleotide overhangs, enabling efficient and seamless library assembly through a simple ligation process.

Currently, recombination methods are shifting from *in vitro* to *in vivo*. Taking advantage of the high occurrence of homologous DNA recombination events in *S. cerevisiae*, 'mutagenic organized recombination process by homologous *in vivo* grouping' (MORPHING) method was developed. MORPHING is a 'one-pot' random mutagenesis method allowing construction of libraries with various degrees of diversity. Short DNA segments are produced by epPCR, and subsequently assembled with conserved overlapping gene fragments and the linearized plasmid by *in vivo* recombination upon transformation into yeast cells [137]. Another technique for assembling linear DNA fragments with homologous ends in *E. coli* is called '*in vivo* assembly' (IVA). IVA uses PCR amplification with primers designed to substitute, delete or insert portions of DNA, and to simultaneously append homologous sequences at amplicon ends. Finally, it exploits *recA*-independent homologous recombination *in vivo*, greatly simplifying complex cloning operations. Thus, multiple simultaneous modifications (insertions, deletions, point mutations and/or site-saturation mutagenesis) are confined to a single PCR reaction, and multi-fragment assembly (library construction) proceeds in bacteria following transformation [138].

2.2. Focused Mutagenesis

Effectively exploring the sequence landscape requires structural and biochemical data (from previous random mutagenesis studies), which can be leveraged to constrain genetic variation to distinct positions of the (poly)peptide, such as regions of the peptide aptamer scaffold which can endure substitutions/insertions/deletions without affecting their overall protein fold, or those peptide residues considered not absolutely essential for specific property of interest (and whose mutation might further augment peptide's activity). Random mutagenesis results in stochastic point mutations at codons corresponding to such residues, but systematically interrogating the entire set of residues at a specific position requires a focused mutagenesis strategy. Focused libraries are typically smaller and more effective, as they only address the residues presumed to bestow the peptide with the property of interest [139].

2.2.1. Enzyme-Based Approaches

Building a library of recombinant DNA constructs is a widely adopted practice accessible to virtually all laboratories, due to the ease of oligonucleotide synthesis and availability of commercial restriction enzymes and DNA ligases. The so-called oligonucleotide-directed mutagenesis enables point or multiple mutations to the target DNA sequence [140]. Normally, a mutagenic primer is designed and synthesized, subsequently elongated by Klenow fragment of DNA polymerase I, ligated

into a vector by T4 DNA ligase and finally transformed into a competent *E. coli* strain. This process is long and includes multiple subcloning and ssDNA rescuing steps [441]. Several kits for site-specific mutagenesis based on mutagenic primers are commercially available. One of the systems works by applying a pair of forward and reverse complementary oligonucleotides with designed mutations. The primers are perfectly complementary to the template at 5' and 3' ends, but carry a changed central nucleotide sequence. A high-fidelity *Pfu* DNA polymerase is used to amplify the entire plasmid harboring the gene to be mutated, followed by the removal of the template by *DpnI* (an endonuclease specific for methylated DNA) [442]. There are numerous adaptations of this method (reviewed by Tee and Wong [221]).

An approach termed 'Kunkel mutagenesis' is commonly used for constructing libraries displayed on filamentous phage [443][444], because its genome is circular and single-stranded. In Kunkel mutagenesis, mutations are introduced with a mutagenic primer that is complementary to the circular ssDNA template. The template is propagated in an *ung⁻ dut⁻* *E. coli* strain. This enzyme handicap results in the template DNA containing uracil bases in place of thymine. The template is recovered and hybridized with the primer and extended by polymerase, followed by transformation into *ung⁺ dut⁺* host cells [29]. Upon transformation, uridylated DNA template is biologically inactivated through the action of uracil glycosylase [445] of the *ung⁺ dut⁺* host, granting a strong selection advantage to the mutated strand(s) over the template.

Overlap extension PCR is another focused mutagenesis approach. First, two DNA fragments with homologous ends (and harboring desired mutation(s)) are amplified in separate PCR reactions by using 5' complementary oligonucleotides. In a subsequent reaction, the fragments are combined; now, the overlapping 3' ends from one of the strands of each fragment anneal and serve as 'mega' primers for extension of the complementary strands. Finally, the construct is amplified with the two flanking primers [446]. Based on this strategy, the SLIM (Site-directed Ligase-Independent Mutagenesis) method, compatible with all three types of sequence modifications (insertion, deletion, and substitution), employs an inverse PCR amplification of the plasmid-embedded template by two 5' adapter-tailed long forward and reverse primers (which include modifications) and two short forward and reverse primers (identical to the long ones but lacking the 5' adapter sequences) in a single reaction, producing 4 distinct amplicons. Next, the amplicons are heat denatured and reannealed to yield 16 (hetero)duplexes, 4 of which are directly cloneable, forming circular DNA through ligation-independent pathway via complementary 5' and 3' single-stranded overhangs. All steps of the SLIM procedure are carried out in a single tube [447].

Gibson assembly is a method of combining up to 15 DNA fragments containing 20-40 bp overlaps in a single isothermal reaction. It utilizes a cocktail of three enzymes; exonuclease, DNA polymerase, and DNA ligase. The exonuclease nibbles back DNA from the 5' end, enabling annealing of homologous DNA fragments. DNA polymerase then fills in the gaps, followed by the covalent fragment joining by the DNA ligase [448]. Applications of Gibson Assembly include site-directed mutagenesis and library construction [449]. A recent adaptation, 'QuickLib', is a modified Gibson assembly method that has been used to generate a cyclic peptide library [50]. QuickLib uses two primers that share complementary 5' ends; one long partially degenerate, and the other short non-degenerate, which are then used for full plasmid PCR amplification. Subsequently, a Gibson reaction is performed which circularizes the library of linear plasmids, followed by template elimination by *DpnI* restriction.

Besides conventional enzymes involved in cumbersome digestion and ligation steps, other enzymes can be utilized for mutagenesis. In nature, lambda exonuclease aids viral DNA recombination. It progressively degrades the 5'-phosphoryl strand of a duplex DNA from 5' to 3', producing ssDNA and (mono)nucleotides [51]. To exploit this property, first, a PCR amplification using template ssDNA and phosphorylated primers with overlapping regions is performed. The PCR product is then treated with lambda exonuclease, generating ssDNA fragments that are subsequently annealed via overlap regions. Afterwards, Klenow fragment is employed to create dsDNA. In this manner, site-specific mutagenesis can be performed using primers that contain degenerate bases [52].

One of the most broadly used approaches for characterization of individual amino acid residues of a (poly)peptide with regards to their contribution to binding affinity or activity is the alanine-scanning mutagenesis. As the name implies, the technique is based on systematic substitution of residues with alanine, and assessing ligand's activity in a biochemical assay. Alanine eliminates the influence of all side chain atoms beyond the beta-carbon, thus exploring the role of side chain functional groups at interrogated positions [53]. For example, a conventional single-site alanine-scanning was used to assess the contribution of individual amino acid residues of a Fc fragment binding peptide displayed on filamentous phage [54]. Since this type of approach is laborious, methods have been developed for multiple alanine substitutions in a high-throughput manner [55]. One such approach builds on the codon-based mutagenesis, analyzing multiple positions, applying split and mix synthesis to produce degenerate oligonucleotides (one pool for the alanine codon and another for the wild-type codon) [56]. An alternative to alanine-scanning is serine-scanning, which follows the logic that, sometimes, substitutions with the hydrophobic alanine side chains may be more detrimental to the peptide's affinity compared to the

slightly larger but hydrophilic serine side chain. Similarly, homolog-scanning (substitutions at individual positions with similar residues) may be employed with the goal of minimizing structural disruption and identifying residues essential for maintaining a function [57].

Another site-directed mutagenesis type is the cassette mutagenesis. It works by replacing a section of genetic information with an alternative, synthetic sequence – a ‘cassette’ [58]. Different from other approaches that target short regions of a gene, this method is convenient for sequences up to 100 bp in size [59][60]. A prerequisite for this method to be practical is that the gene cassette must be flanked by two restriction sites that are complementary and unique with digest sites on the targeted vector. Restriction enzymes excise the targeted fragment from a vector that can then be replaced with DNA sequences carrying desired mutations. If a larger fragment is to be cloned, the ‘megaprimer’ approach is applied by amplification with a series of oligonucleotides [61]. This method can also benefit from using ‘spiked’ synthetic oligonucleotides, allowing randomization at multiple sites [62][63]. Cassette mutagenesis is based on Kunkel mutagenesis, which is time-consuming, so researchers developed an improved version termed “PFunkel”, a conflation of *Pfu* DNA polymerase and Kunkel mutagenesis, that can be performed in a day’s work [64][65]. To overcome the main constraint of site-directed mutagenesis, which is the tedious primer design, rational design techniques can be utilized to introduce desired mutations at precise positions. Researchers can leverage readily available tools such as AAScan, PCRdesign, and MutantChecker to simplify and boost the mutagenesis process [66].

2.2.2. Chemical-Based Mutagenesis

Chemical-based mutations involve various chemical methods to produce desired mutants. To chemically synthesize fully randomized oligonucleotides, a mixture of nucleotides must be applied at each coupling step [67]. A calamitous problem with this strategy is the pronounced bias resulting from the uneven incorporation frequency of the 4 nucleotide building blocks due to their inherent reactivity differences, rendering statistical random mutations inaccessible. Avoiding incorporation of stop codons is practically unattainable and the system is inclined towards amino acid residues encoded by redundant codons [68]. This problem can be tackled by adjusting the mutational frequency with ‘spiked oligonucleotides’ [62], taking into account the differences in reactivity of mononucleotides and the redundant genetic code. The essence of DNA spiking is that non-equimolar ratio of bases at targeted positions are applied during oligonucleotide synthesis, meaning each wild type nucleotide can be custom ‘doped’, achieving either ‘soft’ (high incidence of a certain nucleotide) or ‘hard’ (equal incidence of all four nucleotides) randomization, manually tuning the occurrence of certain amino acids at defined positions in the (poly)peptide chain.

Site-saturation mutagenesis seeks to achieve mutation at a maximal capacity by examining substitutions of a given residue against all possible amino acids. A fully randomized codon NNN (where N = A/C/G/T) gives rise to all possible 64 variant combinations (also known as 64-fold degeneracy) and codes for all 20 amino acids and 3 stop codons. This causes difficulties during library screening and risks enrichment of non-functional clones due to the random introduction of termination codons [69]. Operating with NNK, NNS, and NNB codons (where K = G/T, S = C/G, and B = C/G/T) minimizes the degeneracy in the third position of each codon, consequently lowering codon redundancy and the frequency of terminations [70]. However, such degenerate primers are expensive to synthesize, and using a single degenerate primer to completely eliminate codon redundancy while providing all 20 amino acids is unattainable, due to disproportional representation of certain amino acids [71][72]. Other strategies have to be employed to circumvent these constraints.

To synthesize redundancy-free mutagenic primers, mono [73], di [74], or trinucleotide phosphoramidite [75] solutions (or combinations [76]) can be used. This way, mixtures of oligonucleotides encoding all possible amino acid substitutions within a defined stretch of peptide or a limited number of amino acids (i.e., ‘tailored’ randomization) can be synthesized. This fine-tuning gives complete control over amino acid prevalence at defined positions in the corresponding (poly)peptide sequence, achieving ‘soft’ or ‘hard’ randomization. With this approach codon redundancy and stop codons are completely eliminated [68]. Another randomization strategy labeled ‘MAX’ eliminates genetic redundancy by using a collection of 20 primers containing only codons for each amino acid with the highest expression frequency in *E. coli* [77]. These primers are annealed to a template strand with completely randomized codons (NNN or NNK) at the targeted position. Any misannealing is trivial, since only the ligated selection strand is amplified by a subsequent PCR. The produced random cassettes are then enzyme-digested for cloning. Further development of this strategy gave birth to an upgraded version dubbed ProxiMAX in which multiple contiguous codons are randomized in a non-degenerate manner [78]. Here, a donor blunt-end dsDNA with terminal MAX codons and an upstream *MlyI* restriction site is ligated to an acceptor blunt-end dsDNA. The product strands are amplified, analyzed, and combined at desired ratios in the next randomization cycle. After each ligation cycle, endonuclease *MlyI* is applied to remove the donor DNA strand, making only the randomized sequences available for the successive ligation cycle.

Another strategy that has been developed by Tang et al. [74] is cost-effective and uses degenerate codons to eliminate or achieve near-zero redundancy. A mixture of four codons, NDT, VMA, ATG, and TGG (where D = A/G/T, V = A/C/G, M = A/C) with a molar ratio of 12:6:1:1 at each randomized position results in an equal theoretical distribution for each of the 20 amino acids, without occurrences of stop codons. Following a similar rationale, Kille et al. [72] developed the “22c-trick” which uses only three codons per randomized position; NDT, VHG, and TGG (where H=A/C/T), at 12:9:1 molar ratio. The name sprung from the usage of 22 unique codons, achieving near uniform amino acid distribution (i.e., 2/22 for Leu and Val, and 1/22 for each of the remaining 18 amino acids). Other sophisticated primer mixing strategies have been reported [79][80][81], although picking the best approach is mostly dependent on the size and quality of the library to be prepared, and the lab's operating budget [82].

References

1. Kit S. Lam; Sydney E. Salmon; Evan M. Hersh; Victor J. Hruby; Wieslaw M. Kazmierski; Richard J. Knapp; A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* **1991**, 354, 82-84, [10.1038/354082a0](https://doi.org/10.1038/354082a0).
2. Daniela Marasco; Giuseppe Perretta; Marco Sabatella; Menotti Ruvo; Past and future perspectives of synthetic peptide libraries.. *Current Protein & Peptide Science* **2008**, 9, 447-467, [10.2174/138920308785915209](https://doi.org/10.2174/138920308785915209).
3. Asier Galán; Lubos Comor; Anita Horvatić; Josipa Kuleš; Nicolas Guillemín; Vladimir Mrljak; Mangesh Bhidé; Library-based display technologies: where do we stand?. *Molecular BioSystems* **2016**, 12, 2342-2358, [10.1039/c6mb00219f](https://doi.org/10.1039/c6mb00219f).
4. Krištof Bozovičar; Tomaž Bratkovič; Evolving a Peptide: Library Platforms and Diversification Strategies. *International Journal of Molecular Sciences* **2019**, 21, 215, [10.3390/ijms21010215](https://doi.org/10.3390/ijms21010215).
5. Michael S. Packer; David R. Liu; Methods for the directed evolution of proteins. *Nature Reviews Microbiology* **2015**, 16, 379-394, [10.1038/nrg3927](https://doi.org/10.1038/nrg3927).
6. Yu-Ping Lai; Jing Huang; Lin-Fa Wang; Jun Li; Zi-Rong Wu; A new approach to random mutagenesis in vitro. *Biotechnology and Bioengineering* **2004**, 86, 622-627, [10.1002/bit.20066](https://doi.org/10.1002/bit.20066).
7. R. Myers; L. Lerman; T. Maniatis; A general method for saturation mutagenesis of cloned DNA fragments. *Science* **1985**, 229, 242-247, [10.1126/science.2990046](https://doi.org/10.1126/science.2990046).
8. E C Cox; Bacterial Mutator Genes and the Control of Spontaneous Mutation. *Annual Review of Genetics* **1976**, 10, 135-156, [10.1146/annurev.ge.10.120176.001031](https://doi.org/10.1146/annurev.ge.10.120176.001031).
9. Alan Greener; Marie Callahan; Bruce Jerpseth; Michael K. Trower; An Efficient Random Mutagenesis Technique Using an E. coli Mutator Strain. *In Vitro Mutagenesis Protocols* **2003**, 57, 375-386, [10.1385/0-89603-332-5:375](https://doi.org/10.1385/0-89603-332-5:375).
10. Scheuermann, R.; Tam, S.; Burgers, P.M.J. Identification of the ε-subunit of Escherichia coli DNA polymerase III holoenzyme as the dnaQ gene product: A fidelity subunit for DNA replication. *Proc. Natl. Acad. Sci. U. S. A.* 1983, 80, 7085–7089.
11. Ahmed H. Badran; David R. Liu; Development of potent in vivo mutagenesis plasmids with broad mutational spectra. *Nature Communications* **2015**, 6, 8425, [10.1038/ncomms9425](https://doi.org/10.1038/ncomms9425).
12. Arjun Ravikumar; Adrian Arrieta; Chang C Liu; An orthogonal DNA replication system in yeast. *Nature Chemical Biology* **2014**, 10, 175-177, [10.1038/nchembio.1439](https://doi.org/10.1038/nchembio.1439).
13. Xing Zhang; Huatao Guo; Lei Jin; Elizabeth Czornyj; Asher Hodes; Wong H Hui; Angela W Nieh; Jeff F Miller; Z Hong Z hou; A new topology of the HK97-like fold revealed in Bordetella bacteriophage by cryoEM at 3.5 Å resolution. *eLife* **2013**, 2, 2013, [10.7554/eLife.01299](https://doi.org/10.7554/eLife.01299).
14. Tom Z. Yuan; Cathie M. Overstreet; Issa S. Moody; Gregory A. Weiss; Protein Engineering with Biosynthesized Libraries from Bordetella bronchiseptica Bacteriophage. *PLOS ONE* **2013**, 8, e55617, [10.1371/journal.pone.0055617](https://doi.org/10.1371/journal.pone.0055617).
15. Leung, D.W.; Chen, E.; Goeddel, D. V.; A Method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* **1989**, 1, 11–15, .
16. Juili L. Lin-Goerke; David J. Robbins; John D. Burczak; PCR-Based Random Mutagenesis Using Manganese and Reduced dNTP Concentration. *BioTechniques* **1997**, 23, 409-412, [10.2144/97233bm12](https://doi.org/10.2144/97233bm12).
17. Manuela Zaccolo; David M. Williams; Daniel M. Brown; Ermanno Gherardi; An Approach to Random Mutagenesis of DNA Using Mixtures of Triphosphate Derivatives of Nucleoside Analogues. *Journal of Molecular Biology* **1996**, 255, 589-603, [10.1006/jmbi.1996.0049](https://doi.org/10.1006/jmbi.1996.0049).
18. Elizabeth O. McCullum; Berea A. R. Williams; Jinglei Zhang; John C. Chaput; Random Mutagenesis by Error-Prone PCR. *Advanced Structural Safety Studies* **2010**, 634, 103-109, [10.1007/978-1-60761-652-8_7](https://doi.org/10.1007/978-1-60761-652-8_7).

19. Mondon, P.; Grand, D.; Souyris, N.; Emond, S.; Bouayadi, K.; Kharrat, H. MutagenTM: A random mutagenesis method providing a complementary diversity generated by human error-prone DNA polymerases. In *Methods in Molecular Biology*; 2010; Vol. 634, pp. 373–386 ISBN 9781607616511.
20. Thomas Vanhercke; Christophe Ampe; Luc Tirry; Peter Denolf; Reducing mutational bias in random protein libraries. *Analytical Biochemistry* **2005**, 339, 9-14, [10.1016/j.ab.2004.11.032](https://doi.org/10.1016/j.ab.2004.11.032).
21. Jianqiang Ye; Feng Wen; Yifei Xu; Nan Zhao; Liping Long; Hailiang Sun; Jialiang Yang; Jim Cooley; G. Todd Pharr; Richard Webby; et al. Error-prone pcr-based mutagenesis strategy for rapidly generating high-yield influenza vaccine candidates.. *Virology* **2015**, 482, 234-43, [10.1016/j.virol.2015.03.051](https://doi.org/10.1016/j.virol.2015.03.051).
22. Kang Lan Tee; Tuck Seng Wong; Polishing the craft of genetic diversity creation in directed evolution. *Biotechnology Advances* **2013**, 31, 1707-1721, [10.1016/j.biotechadv.2013.08.021](https://doi.org/10.1016/j.biotechadv.2013.08.021).
23. Tuck Seng Wong; Kang Lan Tee; Bernhard Hauer; Ulrich Schwaneberg; Sequence saturation mutagenesis (SeSaM): a novel method for directed evolution. *Nucleic Acids Research* **2004**, 32, e26-e26, [10.1093/nar/gnh028](https://doi.org/10.1093/nar/gnh028).
24. Hemanshu Mundhada; Jan Marienhagen; Andreea Scacioc; Alexander Schenk; Danilo Roccatano; Ulrich Schwaneberg; SeSaM-Tv-II Generates a Protein Sequence Space that is Unobtainable by epPCR. *ChemBioChem* **2011**, 12, 1595-1601, [10.1002/cbic.201100010](https://doi.org/10.1002/cbic.201100010).
25. Li Fang; Zhi Xu; Guan-Song Wang; Fu-Yun Ji; Chun-Xia Mei; Juan Liu; Guo-Ming Wu; Directed Evolution of an LBP/C D14 Inhibitory Peptide and Its Anti-Endotoxin Activity. *PLOS ONE* **2014**, 9, e101406, [10.1371/journal.pone.0101406](https://doi.org/10.1371/journal.pone.0101406).
26. Tania Selas Castifeiras; Steven G. Williams; Antony Hitchcock; Jeffrey A. Cole; Daniel C. Smith; Tim W. Overton; Development of a generic β -lactamase screening system for improved signal peptides for periplasmic targeting of recombinant proteins in Escherichia coli.. *Scientific Reports* **2018**, 8, 6986, [10.1038/s41598-018-25192-3](https://doi.org/10.1038/s41598-018-25192-3).
27. C. Zahnd; S. Spinelli; B. Luginbuhl; P. Amstutz; C. Cambillau; Andreas Plückthun; Directed in Vitro Evolution and Crystallographic Analysis of a Peptide-binding Single Chain Antibody Fragment (scFv) with Low Picomolar Affinity. *Journal of Biological Chemistry* **2004**, 279, 18870-18877, [10.1074/jbc.m309169200](https://doi.org/10.1074/jbc.m309169200).
28. Ryota Fujii; Motomitsu Kitaoka; Kiyoshi Hayashi; One-step random mutagenesis by error-prone rolling circle amplification. *Nucleic Acids Research* **2004**, 32, e145-e145, [10.1093/nar/gnh147](https://doi.org/10.1093/nar/gnh147).
29. T. A. Kunkel; Rapid and efficient site-specific mutagenesis without phenotypic selection.. *Proceedings of the National Academy of Sciences* **1985**, 82, 488-492, [10.1073/pnas.82.2.488](https://doi.org/10.1073/pnas.82.2.488).
30. Tuomas Huovinen; Eeva-Christine Brockmann; Sultana Akter; Susan Perez-Gamarra; Jani Ylä-Pelto; Yuan Liu; Urpo Lamminmäki; Primer Extension Mutagenesis Powered by Selective Rolling Circle Amplification. *PLOS ONE* **2012**, 7, e31817, [10.1371/journal.pone.0031817](https://doi.org/10.1371/journal.pone.0031817).
31. Adam J. Meyer; Jared W. Ellefson; Andrew D. Ellington; Library Generation by Gene Shuffling. *Current Protocols in Molecular Biology* **2014**, 105, 15.12.1-15.12., [10.1002/0471142727.mb1512s105](https://doi.org/10.1002/0471142727.mb1512s105).
32. Chia Chiu Lim; Yee Siew Choong; Theam Soon Lim; Cognizance of Molecular Methods for the Generation of Mutagenic Phage Display Antibody Libraries for Affinity Maturation. *International Journal of Molecular Sciences* **2019**, 20, 1861, [10.3390/ijms20081861](https://doi.org/10.3390/ijms20081861).
33. W. P. Stemmer; DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution.. *Proceedings of the National Academy of Sciences* **1994**, 91, 10747-10751, [10.1073/pnas.91.22.10747](https://doi.org/10.1073/pnas.91.22.10747).
34. Alexandra J. Reid; DNA shuffling: Modifying the hand that nature dealt. *In Vitro Cellular & Developmental Biology - Plant* **2000**, 36, 331-337, [10.1007/s11627-000-0060-0](https://doi.org/10.1007/s11627-000-0060-0).
35. Huimin Zhao; Lori Giver; Zhixin Shao; Joseph A. Affholter; Frances H. Arnold; Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nature Biotechnology* **1998**, 16, 258-261, [10.1038/nbt0398-258](https://doi.org/10.1038/nbt0398-258).
36. Kosuke Fujishima; Chris Venter; Kendrick Wang; Raphael Ferreira; Lynn J. Rothschild; An overhang-based DNA block shuffling method for creating a customized random library. *Scientific Reports* **2015**, 5, 9740, [10.1038/srep09740](https://doi.org/10.1038/srep09740).
37. David Gonzalez-Perez; Patricia Molina-Espeja; Eva Garcia-Ruiz; Miguel Alcalde; Mutagenic Organized Recombination Process by Homologous In Vivo Grouping (MORPHING) for Directed Enzyme Evolution. *PLOS ONE* **2014**, 9, e90919, [10.1371/journal.pone.0090919](https://doi.org/10.1371/journal.pone.0090919).
38. Javier García-Nafría; Jake F. Watson; Ingo H. Greger; IVA cloning: A single-tube universal cloning system exploiting bacterial In Vivo Assembly. *Scientific Reports* **2016**, 6, 27459, [10.1038/srep27459](https://doi.org/10.1038/srep27459).
39. Chung, D.H.; Potter, S.C.; Tanomrat, A.C.; Ravikumar, K.M.; Toney, M.D.; Site-directed mutant libraries for isolating minimal mutations yielding functional changes. *Protein Engineering, Design and Selection* **2017**, 30, 347-357, [10.1093/proteng/gzx013](https://doi.org/10.1093/proteng/gzx013).

40. Mark J. Zoller; Michael Smith; Oligonucleotide-Directed Mutagenesis: A Simple Method Using Two Oligonucleotide Primers and a Single-Stranded DNA Template. *DNA* **1984**, 3, 479-488, [10.1089/dna.1.1984.3.479](https://doi.org/10.1089/dna.1.1984.3.479).
41. Walker, K.W.; Site-directed mutagenesis. *Encycl. Cell Biol.* **2015**, 1, 122–127., .
42. Jeffrey Braman; Carol Papworth; Alan Greener; Ralph Rapley; Site-Directed Mutagenesis Using Double-Stranded Plasmid DNA Templates. *Nucleic Acid Protocols Handbook, The* **2003**, 9, 835-844, [10.1385/1-59259-038-1:835](https://doi.org/10.1385/1-59259-038-1:835).
43. Renhua Huang; Pete Fang; Brian K. Kay; Improvements to the Kunkel mutagenesis protocol for constructing primary and secondary phage-display libraries.. *Methods* **2012**, 58, 10-7, [10.1016/j.ymeth.2012.08.008](https://doi.org/10.1016/j.ymeth.2012.08.008).
44. Michael D Scholle; John W Kehoe; Brian K Kay; Efficient construction of a large collection of phage-displayed combinatorial peptide libraries.. *Combinatorial Chemistry & High Throughput Screening* **2005**, 8, , .
45. Tomas Lindahl; DNA Glycosylases, Endonucleases for Apurinic/Apyrimidinic Sites, and Base Excision-Repair. *Progress in Nucleic Acid Research and Molecular Biology* **1979**, 22, 135-192, [10.1016/s0079-6603\(08\)60800-4](https://doi.org/10.1016/s0079-6603(08)60800-4).
46. S N Ho; H D Hunt; R M Horton; J K Pullen; L R Pease; Site-directed mutagenesis by overlap extension using the polymerase chain reaction.. *Gene* **1989**, 77, 51-59, .
47. Joyce Chiu; Paul E. March; Ryan Lee; Daniel Tillett; Site-directed, Ligase-Independent Mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4 h. *Nucleic Acids Research* **2004**, 32, e174-e174, [10.1093/nar/gnh172](https://doi.org/10.1093/nar/gnh172).
48. Daniel Gibson; One-step enzymatic assembly of DNA molecules up to several hundred kilobases in size. *Protocol Exchange* **2009**, 6, 343-345, [10.1038/nprot.2009.77](https://doi.org/10.1038/nprot.2009.77).
49. Steven Thomas; Nathaniel D Maynard; John Gill; DNA library construction using Gibson Assembly®. *Nature Methods* **2015**, 12, i–ii, [10.1038/nmeth.f.384](https://doi.org/10.1038/nmeth.f.384).
50. Pierre Galka; Elisabeth Jamez; Gilles Joachim; Patrice Soumillion; QuickLib, a method for building fully synthetic plasmid libraries by seamless cloning of degenerate oligonucleotides. *PLOS ONE* **2017**, 12, e0175146, [10.1371/journal.pone.0175146](https://doi.org/10.1371/journal.pone.0175146).
51. Paul G. Mitsis; Jae G. Kwagh; Characterization of the interaction of lambda exonuclease with the ends of DNA.. *Nucleic Acids Research* **1999**, 27, 3057-3063, [10.1093/nar/27.15.3057](https://doi.org/10.1093/nar/27.15.3057).
52. Bee Nar Lim; Yee Siew Choong; Asma Ismail; Jörn Glöckler; Zoltan Konthur; Theam Soon Lim; Directed evolution of nucleotide-based libraries using lambda exonuclease. *BioTechniques* **2012**, 53, 357-364, [10.2144/000113964](https://doi.org/10.2144/000113964).
53. Gregory A. Weiss; Colin K. Watanabe; Alan Zhong; Audrey Goddard; Sachdev S. Sidhu; Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proceedings of the National Academy of Sciences* **2000**, 97, 8950-8954, [10.1073/pnas.160252097](https://doi.org/10.1073/pnas.160252097).
54. Nika Kruljec; Peter Molek; Vesna Hodnik; Gregor Anderluh; Tomaž Bratkovič; Development and Characterization of Peptide Ligands of Immunoglobulin G Fc Region. *Bioconjugate Chemistry* **2018**, 29, 2763-2775, [10.1021/acs.bioconjchem.8b00395](https://doi.org/10.1021/acs.bioconjchem.8b00395).
55. Kim L Morrison; Gregory A Weiss; Combinatorial alanine-scanning.. *Current Opinion in Chemical Biology* **2001**, 5, 302-307, [10.1016/s1367-5931\(00\)00206-4](https://doi.org/10.1016/s1367-5931(00)00206-4).
56. J. Chatellier; A. Mazza; R. Brousseau; T. Vernet; Codon-Based Combinatorial Alanine Scanning Site-Directed Mutagenesis: Design, Implementation, and Polymerase Chain Reaction Screening. *Analytical Biochemistry* **1995**, 229, 282-290, [10.1006/abio.1995.1414](https://doi.org/10.1006/abio.1995.1414).
57. Gábor Pál; Shun-Yin Fong; Anthony A. Kossiakoff; Sachdev S. Sidhu; Alternative views of functional protein binding epitopes obtained by combinatorial shotgun scanning mutagenesis. *Protein Science* **2005**, 14, 2405-2413, [10.1110/ps.051519805](https://doi.org/10.1110/ps.051519805).
58. J A Wells; M Vasser; D B Powers; Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites.. *Gene* **1985**, 34, 315-323, .
59. Deena M. Kegler-Ebo; Catherine M. Docktor; Daniel DiMaio; Codon cassette mutagenesis: a general method to insert or replace individual codons by using universal mutagenic cassettes. *Nucleic Acids Research* **1994**, 22, 1593-1599, [10.1093/nar/22.9.1593](https://doi.org/10.1093/nar/22.9.1593).
60. M Smith; In Vitro Mutagenesis. *Annual Review of Genetics* **1985**, 19, 423-462, [10.1146/annurev.ge.19.120185.002231](https://doi.org/10.1146/annurev.ge.19.120185.002231).
61. R. Lai; A. Bekessy; C.C. Chen; T. Walsh; R. Barnard; Megaprimer Mutagenesis Using Very Long Primers. *BioTechniques* **2003**, 34, 52-56, [10.2144/03341bm07](https://doi.org/10.2144/03341bm07).
62. Edson Cárcamo; Abigail Roldán-Salgado; Joel Osuna; Iván Bello-Sanmartín; Jorge A. Yáñez; Gloria Saab-Rincón; Héctor Viadiu; Paul Gaytán; Spiked Genes: A Method to Introduce Random Point Nucleotide Mutations Evenly throughout

- an Entire Gene Using a Complete Set of Spiked Oligonucleotides for the Assembly. *ACS Omega* **2017**, 2, 3183-3191, [10.1021/acsomega.7b00508](https://doi.org/10.1021/acsomega.7b00508).
63. J D Hermes; S M Parekh; S C Blacklow; H Köster; J R Knowles; A reliable method for random mutagenesis: the generation of mutant libraries using spiked oligodeoxyribonucleotide primers.. *Gene* **1989**, 84, , .
 64. Elad Firnberg; Marc Ostermeier; PFunkel: Efficient, Expansive, User-Defined Mutagenesis. *PLOS ONE* **2012**, 7, e52031, [10.1371/journal.pone.0052031](https://doi.org/10.1371/journal.pone.0052031).
 65. Francesca Valetti; Gianfranco Gilardi; Improvement of Biocatalysts for Industrial and Environmental Purposes by Saturation Mutagenesis. *Biomolecules* **2013**, 3, 778-811, [10.3390/biom3040778](https://doi.org/10.3390/biom3040778).
 66. Dawei Sun; Martin K. Ostermaier; Franziska M. Heydenreich; Daniel Mayer; Rolf Jaussi; Joerg Standfuss; Dmitry B. Ve printsev; AAscan, PCRdesign and MutantChecker: A Suite of Programs for Primer Design and Sequence Analysis for High-Throughput Scanning Mutagenesis. *PLOS ONE* **2013**, 8, e78878, [10.1371/journal.pone.0078878](https://doi.org/10.1371/journal.pone.0078878).
 67. Keith M. Derbyshire; Joseph J. Salvo; Nigel D.F. Grindley; A simple and efficient procedure for saturation mutagenesis using mixed oligodeoxynucleotides. *Gene* **1986**, 46, 145-152, [10.1016/0378-1119\(86\)90398-7](https://doi.org/10.1016/0378-1119(86)90398-7).
 68. Tamil Selvi Arunachalam; Claudia Wichert; Bettina Appel; Sabine Müller; Mixed oligonucleotides for random mutagenesis: best way of making them. *Organic & Biomolecular Chemistry* **2012**, 10, 4641, [10.1039/c2ob25328c](https://doi.org/10.1039/c2ob25328c).
 69. Rodrigo M.P. Siloto; Randall J. Weselake; Site saturation mutagenesis: Methods and applications in protein engineering. *Biocatalysis and Agricultural Biotechnology* **2012**, 1, 181-189, [10.1016/j.bcab.2012.03.010](https://doi.org/10.1016/j.bcab.2012.03.010).
 70. Yuval Nov; When Second Best Is Good Enough: Another Probabilistic Look at Saturation Mutagenesis. *Applied and Environmental Microbiology* **2011**, 78, 258-262, [10.1128/AEM.06265-11](https://doi.org/10.1128/AEM.06265-11).
 71. Lixia Tang; Hui Gao; Xuechen Zhu; Xiong Wang; Ming Zhou; Rongxiang Jiang; Construction of "small-intelligent" focused mutagenesis libraries using well-designed combinatorial degenerate primers. *BioTechniques* **2012**, 52, 149-158, [10.2144/000113820](https://doi.org/10.2144/000113820).
 72. Sabrina Kille; Carlos G. Acevedo-Rocha; Loreto P. Parra; Zhi-Gang Zhang; Diederik J. Opperman; Manfred T. Reetz; Juan Pablo Acevedo; Sabrina Hoebenreich; Reducing Codon Redundancy and Screening Effort of Combinatorial Protein Libraries Created by Saturation Mutagenesis. *ACS Synthetic Biology* **2012**, 2, 83-92, [10.1021/sb300037w](https://doi.org/10.1021/sb300037w).
 73. Paul Gaytán; Abigail Roldán-Salgado; Elimination of Redundant and Stop Codons during the Chemical Synthesis of Degenerate Oligonucleotides. Combinatorial Testing on the Chromophore Region of the Red Fluorescent Protein mKate. *ACS Synthetic Biology* **2013**, 2, 453-462, [10.1021/sb3001326](https://doi.org/10.1021/sb3001326).
 74. P Neuner; Codon-based mutagenesis using dimer-phosphoramidites. *Nucleic Acids Research* **1998**, 26, 1223-1227, [10.1093/nar/26.5.1223](https://doi.org/10.1093/nar/26.5.1223).
 75. A Ono; A Matsuda; J Zhao; D V Santi; The synthesis of blocked triplet-phosphoramidites and their use in mutagenesis.. *Nucleic Acids Research* **1995**, 23, 4677-4682, .
 76. Paul Gaytán; Casandra Contreras-Zambrano; Mónica Ortiz-Alvarado; Alfredo Morales-Pablos; Jorge Yáñez; TrimerDimer: an oligonucleotide-based saturation mutagenesis approach that removes redundant and stop codons.. *Nucleic Acids Research* **2009**, 37, e125-e125, [10.1093/nar/gkp602](https://doi.org/10.1093/nar/gkp602).
 77. Marcus D. Hughes; David A. Nagel; Albert F. Santos; Andrew J. Sutherland; Anna V. Hine; Removing the redundancy from randomised gene libraries.. *Journal of Molecular Biology* **2003**, 331, 973-979, [10.1016/s0022-2836\(03\)00833-7](https://doi.org/10.1016/s0022-2836(03)00833-7).
 78. Mohammed Ashraf; Laura Frigotto; Matthew E. Smith; Seema Patel; Marcus D. Hughes; Andrew J. Poole; Husam R.M. Hebaishi; Christopher G. Ullman; Anna V. Hine; ProxiMAX randomization: a new technology for non-degenerate saturation mutagenesis of contiguous codons. *Biochemical Society Transactions* **2013**, 41, 1189-1194, [10.1042/bst20130123](https://doi.org/10.1042/bst20130123).
 79. Gur Pines; Assaf Pines; Andrew D. Garst; Ramsey I. Zeitoun; Sean A. Lynch; Ryan T. Gill; Codon Compression Algorithms for Saturation Mutagenesis. *ACS Synthetic Biology* **2014**, 4, 604-614, [10.1021/sb500282v](https://doi.org/10.1021/sb500282v).
 80. Lixia Tang; Xiong Wang; Beibei Ru; Hengfei Sun; Jian Huang; Hui Gao; MDC-Analyzer: A novel degenerate primer design tool for the construction of intelligent mutagenesis libraries with contiguous sites. *BioTechniques* **2014**, 56, , [10.2144/000114177](https://doi.org/10.2144/000114177).
 81. Aitao Li; Carlos G. Acevedo-Rocha; Manfred T. Reetz; Boosting the efficiency of site-saturation mutagenesis for a difficult-to-randomize gene by a two-step PCR strategy.. *Applied Microbiology and Biotechnology* **2018**, 102, 6095-6103, [10.1007/s00253-018-9041-2](https://doi.org/10.1007/s00253-018-9041-2).
 82. Carlos G. Acevedo-Rocha; Manfred T. Reetz; Yuval Nov; Economical analysis of saturation mutagenesis experiments. *Scientific Reports* **2015**, 5, 10654, [10.1038/srep10654](https://doi.org/10.1038/srep10654).

