Published in final edited form as:

Anal Biochem. 2012 February 15; 421(2): 640–648. doi:10.1016/j.ab.2011.11.009.

Restriction enzyme-free construction of random gene mutagenesis libraries in *E. coli*

Jen C. Pai, Kevin C. Entzminger, and Jennifer A. Maynard*

Department of Chemical Engineering, University of Texas at Austin, Austin TX 78712, USA

Abstract

Directed evolution relies on both random and site-directed mutagenesis of individual genes and regulatory elements to create variants with altered activity profiles for engineering applications. Central to these experiments is the construction of large libraries of related variants. However, a number of technical hurdles continue to limit routine construction of random mutagenesis libraries in *E. coli*, in particular, inefficiencies during digestion and ligation steps. Here, we report a restriction enzyme-free approach to library generation using megaprimers, termed MegAnneal. Target DNA is first exponentially amplified using error-prone PCR and then linearly amplified with a single 3' primer to generate long, randomly mutated, single-stranded megaprimers. These are annealed to single-stranded dUTP-containing template plasmid, extended with T7 polymerase to create a complementary strand and the resulting termini ligated with T4 DNA ligase. Using this approach, we are able to reliably generate libraries of $\sim 10^7$ cfu/ μ g DNA/transformation in a single day. We have created MegAnneal libraries based on three different single-chain antibodies and identified variants with enhanced expression and ligand-binding affinity. The key advantages of this approach include facile amplification, restriction enzyme-free library generation and a significantly reduced risk of mutations outside the targeted region and wild-type contamination as compared to current methods.

Keywords

directed evolution; protein engineering; error prone PCR; kunkel mutagenesis; megaprimers

Introduction

Directed evolution, the generation of a library of mutations followed by function-based selection, is a standard tool for many protein engineering, synthetic biology and metabolic engineering applications (1,2). This method has been applied to a variety of problems, including epitope mapping of protein-protein interactions (3–5), generation of novel protein function (6,7), optimization of metabolic pathways through promoter or transcription factor engineering (8,9) and enzyme engineering, such as modulation of substrate use (10) or generation of novel catalytic activity (11). Approaches to generate libraries of variants include targeted and random mutagenesis in addition to recombination-based methods, such as DNA shuffling (12) and non-templated recombination (13), while still other methods use

© 2011 Elsevier Inc. All rights reserved.

*Correspondence Jennifer A. Maynard, MC0400, 1 University Station, University of Texas at Austin, Austin, TX, 78712. 512-471-9188 T, 512-471-7060 F, Maynard@che.utexas.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

a combination of these techniques (14). Of these, random mutagenesis is a popular and effective tool when combined with a robust function-based screen (15,16) and can be used to adjust a variety of protein biochemical or biophysical features, such as antibody affinity for ligands (17).

The goal of a directed evolution experiment is to identify clones with enhanced function; the probability of success is enhanced by generation of large libraries encoding few wild-type variants. Generation of randomly mutated libraries in *E. coli* typically proceeds with four steps: (i) amplification of the region of interest under error-prone conditions, (ii) restriction enzyme digestion, (iii) ligation into a similarly digested plasmid and (iv) transformation into competent cells. Methods to improve error-prone amplification have been extensively studied (18). For example, PCR conditions can be optimized for library quality, but this has little impact on the final library size. Similarly, as long as highly competent *E. coli* are used for transformation, this step does not typically limit library size (19). Sequential enzymatic digestion and ligation steps can limit the size of the final library (due to inefficient digestion or elimination/introduction of restriction sites during error-prone PCR (20)), often require optimization of insert:template ratios during ligation (21), and can be both time- and resource-intensive. Techniques designed to avoid this pitfall often require altered reaction conditions, specialized primer design or additional time consuming sub-cloning steps (22,23).

As a result, the cloning steps present the major bottleneck to library generation. Alternatives to enzymatic digestion and ligation typically employ PCR or *in vivo* mutagenesis (24). Direct DNA manipulation can be avoided by *in vivo* techniques, such as somatic hypermutation in B cells (25) and plasmid amplification in *E. coli* strains lacking recombination/repair enzymes (*e.g.*, XL1-Red, Stratagene). However, the use of XL1-Red suffers from a very low mutation rate, requiring propagation of plasmids for up to 100 generations to achieve sufficient error rates, and does not restrict mutations to a region of interest, but introduces them throughout the plasmid and bacterial genome (23). QuikChange (Stratagene) is a widely used method for site-directed mutagenesis and has been expanded for targeted mutagenesis libraries, including those using megaprimers (20,21,26–28). Here, primers encoding the desired mutation(s) are annealed to methylated template plasmid. A linear amplification step extends the primer to generate a replica of the entire plasmid. The wild-type template is digested with the methyl-DNA specific enzyme *DpnI* and, after transformation, endogenous nucleases (29). Almost all QuikChange-based methods require extensive primer design and optimization of PCR conditions as well as a high incidence of wild-type background in final libraries.

To avoid the high levels of wild-type plasmid complicating QuikChange mutagenesis, Kunkel mutagenesis offers an alternative approach (30). Kunkel mutagenesis relies on annealing of a mutagenic primer to uracil-containing single-stranded template DNA (dU-ssDNA) followed by primer extension to generate a mutated complementary strand without uracil (31). Upon transformation into *dut⁺ ung⁺ E. coli*, the template dU-ssDNA is digested by endogenous nucleases, allowing only the altered DNA strand to be propagated. Libraries of over 10¹¹ individual clones randomizing eight consecutive amino acids have been constructed using this method (4). Although multiple regions within a gene can be simultaneously randomized by annealing multiple primers, this method is currently limited to targeted mutations.

Here, we present a restriction enzyme-free approach to *in vitro* library generation, MegAnneal, which allows generation of large random mutagenesis libraries in a single day. This method involves five steps (Figure 1). (**1a**) Stop codons are inserted into the gene of interest and (**1b**) the resulting phagemid used to create dU-ssDNA. The stop codons in the

dU-ssDNA template do not interfere with megaprimer annealing but ensure that background clones which do not incorporate the megaprimer also are unable to produce full-length protein. **(2a)** The region of interest is amplified from plasmid under error-prone conditions. **(2b)** Using the amplified product as template, asymmetric PCR with a single nested primer generates a library of randomly mutated 3' megaprimers. **(3)** In a single reaction, the megaprimers are annealed to dUTP- and stop codon-containing, single-stranded template plasmid, **(4)** extended with T7 DNA polymerase to create double stranded DNA in which one strand incorporates a megaprimer and ligated with T4 DNA ligase. **(5)** Transformation into *dut⁺ ung⁺ E. coli* cleaves the uracil-containing wild-type template but not the megaprimer-containing plasmid DNA, resulting in the final library. MegAnneal reliably and rapidly creates large libraries ($\sim 10^7$ cfu/ μ gDNA/transformation). By omitting the traditional digestion and ligation steps, we avoid two problematic steps in library generation. Inclusion of stop codons in the dU-ssDNA template plasmid prevents functional wild-type sequences from contaminating the final library. Finally, we have used megaprimers ranging from 150–750 bp in length, to focus random mutagenesis to desired regions on the plasmid. We have successfully employed this method to generate libraries for three different single-chain antibodies (scFv) and, in conjunction with phage display, have identified variants with enhanced function from two of these libraries.

Materials and Methods

Template dU-ssDNA plasmid preparation

This method requires uracil-containing template plasmid with an M13 origin of replication for production of single-stranded DNA, the gene of interest, and a compatible display platform to select enhanced variants. ScFv genes (750 bp, orientation V_L-[G₄S]₄-V_H) were digested using *Sfi*I and ligated into the phagemid vector pMoPac24 (~4800 bp), which contains an M13 origin of replication and generates a fusion protein between the scFv and phage coat protein gpIII, allowing library selection by phage display (32). Three scFv genes were used: 3D5/EE scFv, with initially poor expression and weak affinity for the EE peptide (sequence: EYMPME) (33,34); M2 scFv, with poor expression and poor affinity for the FLAG peptide (sequence: DYKDDDDK) and hu1B7 scFv, with poor expression and moderate affinity for the pertussis toxin (35). To minimize the presence of wild-type scFv in libraries, three tandem stop codons were inserted in the complementarity determining region 3 of the heavy chain (CDR H3) using standard QuikChange mutagenesis (Stratagene, primers provided in Table 3). For control library V, which uses two primers targeting separate regions, an additional three stop codons were inserted at CDR H2 in order to select for full-length scFvs resulting from the annealing of both primers. Stop codons were inserted near the 3' end of the scFv; because megaprimers are synthesized from the 3' end of the gene, this permits even short megaprimers to anneal to the template dU-ssDNA and replace the stop codons, allowing full-length scFv to be expressed. After sequencing, (University of Texas at Austin Core Facility), plasmids were transformed into the *dut⁻ ung⁻ E. coli* strain CJ236 (NEB) which substitutes uracil for thymine in DNA. Phage were produced by infection with M13 phage and dU-ssDNA purified as described previously (31). This dU-ssDNA was used as the template during the megaprimer annealing and extension steps.

Generation of randomly mutated megaprimers

Randomly mutated megaprimers were produced via (i) amplification of scFvs under error-prone conditions followed by (ii) asymmetric PCR to generate 3' megaprimers containing randomly inserted mutations. First, the entire scFv gene (10 ng) was amplified using the low fidelity Mutazyme II DNA Polymerase (Stratagene, La Jolla, CA) with flanking primers 5' scback and 3' huCk (3D5/EE and hu1B7 scFvs) or 5' pAKpel and 3' pAK400 (M2 scFv) according to commercial instructions to obtain a mutation rate of 1%. The 50 μ l reactions

were incubated at 94°C for 4 min followed by 25 cycles of 94°C for 30 sec, 58°C (47°C M2 library) for 30 sec and 72°C for 1 min and a final round at 72°C for 4 min. After amplification, PCR product (~800 bp) was purified from a 1% TAE agarose gel using the QIAquick gel extraction kit (Qiagen). To provide an initial estimate of mutation frequency, purified 3D5/EE PCR product was cloned into the pTopo vector (Invitrogen), transformed into *E. coli* and 10 independent colonies sequenced with M13 Forward.

In the asymmetric PCR step, error-prone PCR products were used as template for generation of 3' megaprimers containing randomly inserted mutations. Approximately 400 ng PCR product was linearly amplified in 50 µl reactions containing Vent polymerase and 3' scforlong (3D5/EE scFv) or 3' scM2_Mut (M2 scFv) by incubating at 94°C for 4 min followed by 30 cycles of 94°C for 30 sec, 60°C (52°C M2 library) for 30 sec and 72°C for 1 min and a final round at 72°C for 4 min. Megaprimers of varying lengths were purified from a 1% TAE agarose gel as above and phosphorylated in 20 µl reactions containing ~2 µg megaprimer, TM Buffer (0.05 M Tris, 0.01 M MgCl₂, pH 7.5), 1 mM ATP, 5 mM DTT and 5 units T4 polynucleotide kinase and incubated at 37°C for 1 hour (31). To serve as a check that MegAnneal is compatible with alternative methods of megaprimer generation, the initial Mutazyme reaction was omitted for the hu1B7 scFv and the megaprimers produced in a single step by asymmetric amplification of template dU-ssDNA under biased nucleotide conditions. Approximately 500 ng error-prone product was linearly amplified in 100 µl reactions containing either Vent or Taq polymerase, biased nucleotides (0.2 mM dATP, 0.2 mM dGTP, 1 mM dCTP and 1 mM dTTP), 7 mM MgSO₄ total and 3' Mo54. Amplification proceeded by incubation at 95°C 3 min followed by 30 cycles of 94°C 30 sec, 52°C 30 sec and 72°C 1 min and a final round at 72°C for 4 min. Megaprimers were then purified and enzymatically phosphorylated as above.

Generation of randomly mutated plasmids with megaprimers

To generate libraries for directed evolution, megaprimers were annealed to dU-ssDNA and extended in a modified Kunkel procedure (31). dU-ssDNA (20 µg) was combined with a threefold molar excess of phosphorylated megaprimer in TM buffer at a final volume of 250 µl. The mixture was heated to 90°C for 2 min to disrupt secondary structure, followed by cooling at 50°C for 3 min, then 20°C for 5 min to facilitate megaprimer-template annealing. The mixture was adjusted to contain 0.34 mM ATP, 0.85 mM dNTPs, 5.1 mM DTT, 240 U T4 DNA ligase and 3 U T7 DNA polymerase, and the reaction was allowed to proceed overnight at room temperature. Extension and ligation yielded covalent, closed circular double-stranded plasmid (ccc-DNA) comprising a dUTP-containing template strand and an anti-sense dTTP-containing strand incorporating the mutagenic megaprimer. The reaction was desalted using the QIAquick gel extraction kit (Qiagen), and DNA was quantified by UV absorbance (1 OD_{260nm} = 50 µg/ml). Successful generation of ccc-DNA was confirmed with a 1% TAE agarose gel (500 ng each component), and band intensities were measured using ImageJ for library quality checks (36).

The desalted polymerase reaction (~10 µg total DNA) was combined with electrocompetent XL1-Blue *E. coli* cells (500 µl, 10⁸ cfu/µg DNA competency), electroporated in a 2 mm gap electrocuvette (Biorad, Hercules, CA) at 2.5 kV, 25 µF, 200 Ω and immediately recovered in 2 ml TB media at 37°C for 45 min with shaking before transfer to 500 ml TB with 200 µg/ml ampicillin. To determine library size, serial dilutions were plated on selective agar plates. Three scFv libraries using the 3D5/EE gene (I-III) were produced using megaprimers of varying lengths (Table 1) to focus mutagenesis to different regions of the scFv gene as well as determine the effect of megaprimer length on final library size and diversity. Library VII was created using M2 scFv to ensure that results were not template-dependent, while library VI using the hu1B7 gene was created to show the versatility of this method to interface with multiple error-prone PCR protocols. Two control libraries were created using either one (IV)

or two (V) 50 bp oligonucleotides (Table 3) to provide a calibration point for library size with respect to standard Kunkel mutagenesis (31).

Analysis of library diversity and quality

After transformation, 30–50 clones were sequenced using 5' pakpel primer to determine the scFv mutation rate, percent of clones incorporating the megaprimer, and rates of mispriming and frame-shifting. To assess the rate of unintended mutations introduced during linear amplification, a region of the backbone plasmid was sequenced from 10 clones from each library using the 5' skp primer.

To demonstrate the ability of MegAnneal to engineer desired protein characteristics, recombinant scFv-displaying M13 phage were selected for enhanced solubility and ligand affinity from libraries I and II. After growth with shaking at 37°C to mid-log phase, expression was induced with 1 mM IPTG at 25°C for 3 hours. M13K07 helper phage (Sigma) were added (MOI of 10) and the culture was allowed to rest for 30 min at 25°C. Cultures were incubated at 25°C for 2 hours with shaking followed by addition of kanamycin (25 µg/ml) and incubation overnight with shaking (~12 hours at 25°C). Phage were purified by double-precipitation with 1/5 volume 20% PEG-8000/2.5 M NaCl, resuspended in PBS and quantified by absorbance at 260 nm.

Phage panning was performed in high-binding polystyrene 96-well plates (Costar), coated with 4 µg/ml anti-c-myc antibody (9E10, Sigma) or 10 µg/ml target ligand (MBP-EE, maltose binding protein containing a C-terminal EE tag). Panning consisted of three selection rounds: one with immobilized anti-c-myc to capture full-length scFvs with a c-terminal c-myc tag, followed by two rounds with the MBP-EE ligand. For each round, 10¹² colony forming units (cfu) were added to a well coated and blocked with 5% non-fat milk in PBS. After equilibration for 1 hour at 37°C, wells were washed 10 times with PBS/0.05% Tween. Bound phage were eluted with 0.1 M glycine/HCl pH 2.2 and neutralized with 2 M Tris base. Exponential phase *E. coli* ER2738 (2 ml; NEB) were infected with eluted phage to amplify eluted clones for subsequent panning rounds.

To characterize individual clones selected during panning, single colonies were inoculated into 150 µl TB with ampicillin (200 µg/ml) in a sterile 96 well plate and phage were produced as above. Cells were pelleted, and 50 µL of phage-containing supernatant were used in each of two ELISA activity assays. First, incorporation of full-length scFv into phage particles was monitored via binding to immobilized anti-c-myc. Next, phage presenting active scFv were monitored via binding to immobilized MBP-EE. In each assay, phage were detected using anti-M13-HRP (1:5000, GE Healthcare), with the signal being developed with TMB substrate (Thermo Scientific) and quenched with 1N HCl and the absorbance being measured at 450 nm. Negative controls included uncoated, blocked ELISA wells and phage harboring plasmid containing the scFv-stop gene; positive controls included scFvs known to present well on phage. Plots of ELISA data were prepared using GraphPad Prism 5 (www.graphpad.com), and error bars equal to one standard deviation are included.

Results

Error-prone scFv gene amplification

As a first test of our method, we focused on the 3D5/EE scFv, aiming to identify variants with improved EE peptide affinity, expression and solubility. Our ultimate goal is that well-behaved scFvs with EE peptide specificity could be used as crystallization chaperones for membrane proteins presenting the EE peptide (33). Amplification of the scFv genes produced a sharp band of the expected size (~800 bp) and a yield of ~2 µg total DNA (data not shown). Mutation rates during amplification with the Mutazyme II polymerase are

controlled by the quantity of template plasmid introduced into the PCR reaction. With the 10 ng input used here and 20 amplification cycles, 3–4 mutations per 1000 bp were expected (Stratagene). Sequencing of 40 3D5/EE clones after amplification and TopoTA cloning identified up to 10 mutations per scFv gene with an average of three. While the mutations were distributed throughout the scFv gene, they were more frequently observed in the heavy chain.

In principal, any random mutagenesis procedure is compatible with MegAnneal. Mutazyme II DNA polymerase was selected due to its ability to introduce mutations with minimal nucleotide bias, in contrast with other error-prone PCR methods (21,24). In addition, the simple control of mutation rates from 2–18 mutations/kb by varying the amount of DNA template from 0.001–100 ng without compromising PCR product yields is attractive (37). Lastly, the mutagenesis rate was more predictable than amplification with biased nucleotides or in the *mutS* deficient *E. coli* strain XL1-Red, which yielded few mutations after 15 rounds of cell growth (data not shown), similar to previous reports (37).

Generation of error-prone megaprimers

Purified, error-prone PCR product was used as template for an asymmetric PCR reaction using a single 3' primer annealing to the scFv terminus. Under these conditions, incorrect primer annealing or poor polymerase processivity can result in premature termination of the growing DNA strand; as a result the reaction can yield products with a range of sizes. Using the same 3' primer for both error-prone gene amplification and asymmetric PCR produced poor yields with multiple products (Figure 2a), while use of a nested primer annealing 46 bp upstream resulted in a strong band corresponding to the expected size of the scFv gene for both the 3D5/EE (Figure 2a) and M2 scFv genes (Figure 2b). Since the hu1B7 gene amplified poorly with Mutazyme II, we produced random megaprimers in a single step, using dU-ssDNA as template with a single 3' primer and biased nucleotide conditions. This reaction yielded a dominant product of the expected size (Figure 2b). Subsequent analysis indicated that optimal libraries resulted from megaprimers with single, strong bands of the expected size based on primer positions. Poor quality PCR products, including multiple bands, smears and bands larger than the expected gene size, resulted in libraries with smaller sizes and insertions/deletions in the gene.

To assess the effect of megaprimer length on library size and diversity, megaprimers were separated by size on an agarose gel and purified prior to generation of ccc-DNA (Figure 2b). Megaprimers generated with a non-nested primer generated three distinct bands, which were each purified and used to construct three separate 3D5/EE scFv libraries: 250 bp (Library I), 750 bp (Library II), and pooled primers from 150–250 bp (Library III). Library VII, based off M2 scFv, used a 750bp megaprimer. For hu1B7, single-step megaprimer synthesis produced discrete bands as a function of the extension time. A dominant band at ~700 bp was purified, used to produce ccc-DNA and then library VI. For comparison, two control libraries were constructed via targeted Kunkel mutagenesis (31) randomizing heavy chain CDR residues using either a single 50 bp primer (3' HCDR3, library IV), or two 50 bp primers (3' HCDR2 and 3' HCDR3, library V; Table 1).

Generation of megaprimer-containing ccc-DNA

Gel-purified megaprimers were annealed to dU-ssDNA, extended and ligated in a single reaction to generate a heteroduplex consisting of a dUTP-containing strand and a complementary dTTP-containing strand (ccc-DNA). The larger double-stranded ccc-DNA reaction products migrate more slowly than dU-ssDNA during electrophoresis and typically separate into three bands, allowing an easy check on reaction progress (Figure 3). The ccc-DNA separates into a large strand-displaced product resulting from unwanted T7 DNA

polymerase activity; unligated heteroduplex dsDNA and small correctly extended and ligated, supercoiled ccc-DNA. Each megaprimer preparation produced ccc-DNA of similar yield and quality. Control libraries using 50 bp randomized oligonucleotides to directly produced ccc-DNA of a similar quality, as judged by DNA gel electrophoresis.

The ratio of megaprimer to dU-ssDNA used during ccc-DNA synthesis and the intensity of the supercoiled ccc-DNA band each influenced the resulting library size. We found that increasing this megaprimer:template ratio from 3:1 to 6:1 doubled the resulting library size, while further increases up to 9:1 actually reduced library size. This may be a result of a high megaprimer concentration driving the annealing reaction balanced by the final ratio of ligated, supercoiled ccc-DNA to non-replicative megaprimer DNA. Similarly, strand-displaced product is also unable to replicate in bacterial cells. Thus, the ratio of ccc-DNA to strand-displaced product, analyzed by band intensity with ImageJ, serves as a measure of library DNA quality (38). The ccc-DNA generated with megaprimers produced higher quality ratios than the control libraries (Table 1). For the 3D5/EE libraries, the average quality ratio for MegAnneal libraries was 1.90 ± 0.48 , while the average quality ratio for control libraries was 0.79 ± 0.15 . A possible explanation is that 150–750 bp megaprimers have a higher annealing free energy than the short 50 bp nucleotides and are less easily displaced by T7 polymerase.

MegAnneal produces libraries of large size and high quality

During library generation, four metrics can be monitored as a measure of the final library quality: (1) megaprimer size and yield; (2) ccc-DNA quality and yield; and after transformation, (3) the percent of full-length, non-template genes; and (4) the total number of independent clones. Megaprimer generation was optimized by use of a nested primer relative to that used during error-prone amplification as the use of identical 3' primers for both randomization and megaprimer generation produces a range of PCR product sizes (Figure 2a). Additionally, primer design to reduce non-specific annealing, such as low GC content, the absence of hairpins, or repetitive sequences, has been shown to enhance library size and transformation efficiency (38). Analysis of ccc-DNA quality by gel electrophoresis shows stronger and more distinct product bands for smaller megaprimers (Table 1). Non-specific megaprimer annealing was observed at low frequencies, as monitored by the presence of truncated scFvs with only V_H domains.

After confirmation of successful production, ccc-DNA was transformed into electrocompetent *E. coli* XL1-Blue, with aliquots plated to estimate library size, which ranged from $0.93\text{--}4.1 \times 10^7$ cfu/ug DNA/transformation. Sequencing confirmed the presence of full-length scFv in about half the transformants (range, 37–79%), consistent with standard Kunkel results (Table 1). Control primers (50 bp) and shorter megaprimers (up to 250 bp; libraries I, IV and V) tended to produce larger libraries than did longer megaprimers (over 250 bp; libraries II and VI; Table 1).

Additionally, we have generated libraries using single megaprimers <100 bp in size to produce very large libraries ($\sim 10^8$ cfu/ug DNA/transformation; data not shown). The correlation between megaprimer length and library size may reflect the more rapid kinetics or sequence-specific binding of smaller oligonucleotides. Finally, use of multiple megaprimers to simultaneously randomize multiple regions of a gene biases the library for the shorter megaprimer and reduced the overall library quality. Generating a library with two 50 bp megaprimers (library V) did not significantly affect library size as compared to a library using only one 50 bp megaprimer (IV), though it did increase background wild-type levels, potentially due to a larger amount of mis-priming (Table 1).

To determine the library mutation rate and assess whether megaprimer extension introduces additional mutations within the plasmid, 30–50 clones per library were sequenced at both the scFv gene and a region of the plasmid distal from that targeted for mutation. The scFv mutagenesis rates were calculated accounting for megaprimer size. Since a 250 bp megaprimer covers only one-third of the scFv gene, a library produced using this megaprimer is expected to have a lower overall gene mutation rate than one generated with a 750 bp megaprimer. Mutagenesis rates within megaprimer-encoded regions were statistically similar to that observed after error-prone amplification by Mutazyme II (0.1–0.9%; Table 1), revealing that mutation rates can be tightly controlled during error-prone amplification and are not significantly affected by later steps. All libraries generated contained similar mutation rates, with mutations distributed across the entire megaprimer-encoded region, and comprised of transitions and transversions as expected. Furthermore, no mutations were observed in a region of the plasmid not targeted for mutagenesis. Megaprimer mis-priming was observed exclusively in the variable length megaprimer library III, in which 7.8% of sequenced clones (4 of 51), possessed frame shifts and one non-target gene insertion.

Enhanced scFvs recovered from MegAnneal libraries

After three rounds of phage panning and selection, clones from libraries I and II were screened for scFv expression and EE peptide affinity. Randomly selected colonies were grown in 96 well plates and infected with M13 phage to produce scFv-displaying phage; the phage were used in two ELISA binding assays. Binding to anti-c-myc antibody indicates the presence of a full-length scFv on the phage surface (as the c-myc peptide is encoded at the scFv c-terminus) and implies megaprimer replacement of the stop codons that were present in the template plasmid. Binding to MBP-EE ligand reflects scFv activity, while the EE/c-myc ratio provides a measure of specific binding activity. A tightly binding but poorly expressed scFv will have a high ratio; an scFv which is both tightly binding and well-expressed will have a ratio near one. Of the 93 colonies screened from each library, 25 displayed both EE and c-myc signal above background (Figure 4). Of these clones, 3 exhibited significantly higher specific activity than the wild-type scFv. Further analysis of soluble protein variants selected from this library identified ones with 38-fold higher EE peptide affinity, six-fold greater solubility (2.3 vs 12.8 mg/ml) and enhanced expression levels as compared to wild-type, 3.1 vs 8.5 mg/L culture (33).

Discussion

Kunkel mutagenesis was developed as an early method to introduce site-directed mutations in genes (30) and has been used to produce site-directed libraries as large as 10^{11} independent clones randomizing eight sequential amino acid residues (4). We have adapted Kunkel mutagenesis to generate randomized libraries of select regions or an entire gene, termed MegAnneal (Figure 1). This method allows rapid production of large and diverse random mutagenesis libraries with mutations targeted to one or more locations on a plasmid and covering regions of 150 to 750 bp in size. Production of dU-ssDNA template plasmid containing stop codons within the region of interest requires an initial time investment, but is crucial to minimize the presence of functional wild-type genes in the final library. With the dU-ssDNA template in hand, we are able to rapidly create large libraries (10^7 cfu/ μ g ccc-DNA/transformation) within 24 hours. Importantly, ccc-DNA can be stored indefinitely and simply re-transformed to produce a fresh library for screening, reducing the effect of growth differences among clones upon repeated growth of frozen libraries.

MegAnneal and competing technologies

When generating libraries for directed evolution studies in *E. coli*, a primary consideration is whether to employ targeted or random mutagenesis (Table 2). In mutator strains (23) and during whole plasmid PCR under error-prone conditions (20), mutations are introduced throughout the plasmid. Library generation is relatively easy using these methods, but mutation frequency can be variable and mutation of essential regions of the plasmid complicates screening. Alternatively, mutations can be restricted to a gene fragment, as in DNA shuffling (12) and PCR-based mutagenesis (19), but here, the sequential digestion and ligation steps present bottlenecks to generation of large libraries.

Primer-based methods allow for precise control over the location of the randomized region, with the number and length of primers being important factors. QuikChange using primers with degenerate codons is widely used for targeted mutagenesis (29) but is compatible with only shorter primers. Modifications to QuikChange employ larger primers to either introduce multiple targeted mutations, as in megaprimer QuikChange (27,28), or to perform random mutagenesis throughout a gene, as in MEGAWHOP (21). All QuikChange-based methods rely on *DpnI* digestion to remove wild-type template from the final library, but even efficient digestion can result in a low rate of wild-type background (<10%) that can quickly overrun a library (19). Elimination of wild-type background is ensured in Kunkel mutagenesis by incorporating stop codons into the template (31), preventing expression of the wild-type protein during selection. However, to date, this method has used only smaller primers for targeted mutagenesis.

MegAnneal combines the versatility of megaprimer-based QuikChange methods with the elimination of wild-type background inherent to Kunkel mutagenesis. A large range of megaprimer lengths, from 150–750 bp as shown here, allows flexibility for many directed evolution applications. Random mutagenesis can be introduced by amplifying under error-prone conditions, with tight control over mutation frequency. The primer used for megaprimer generation can also encode targeted mutations (data not shown). Notably, the limitations of Kunkel mutagenesis also apply to MegAnneal. The vector chosen must be compatible with not only the method of screening but also the method of dU-ssDNA generation, which requires an M13 origin of replication. Similarly, as the *dut-ung* deletions in strain CJ236, used to produce dU-ssDNA, are marked with chloramphenicol resistance, this antibiotic marker should be avoided.

We directly compared library quality from both QuikChange and MegAnneal methods by gel electrophoresis (Figure 5). Two major disadvantages to QuikChange include primer dimer formation and mis-priming during whole plasmid amplification, leading to a smear of bands during electrophoresis. Primer-dimer formation occurs frequently in QuikChange and competes with plasmid annealing, reducing yields of altered plasmid (39). MegAnneal uses a single primer to initiate megaprimer amplification, largely eliminating the risk of intermolecular dimer formation. QuikChange requires an optimized annealing temperature to reduce mis-priming during each of the 20–30 amplification cycles (40). In contrast, MegAnneal employs slow cooling during the single megaprimer annealing step, reducing the frequency of mis-priming, and facilitating simultaneous annealing of multiple primers to the dU-ssDNA template.

For exponential PCR-based methods, unintended mutations (*i.e.* in the plasmid origin of replication, promoter or antibiotic resistance gene) are exponentially propagated in successive cycles and may result in the loss of clones from the library or diminished fitness during selection, inadvertently limiting library quality. MegAnneal has a minimal risk of introducing mutations outside the intended region because the T7 polymerase error rate is low (1.5×10^{-5} per bp, NEB) and each template is amplified only once. In fact, no

unintended mutations were observed in the template plasmid when sequencing an ~700 bp region from 30–50 clones from each library.

Because the wild-type phenotype may contain low activity of the desired characteristic, such as low affinity for a ligand, even a small background level of wild-type sequence can quickly dominate a library. Most site-directed and random mutagenesis protocols only have one mechanism to limit the presence of wild-type genes in the library. Many methods, including QuikChange, use methylated template DNA, which is digested *in vitro* with *DpnI* prior to transformation. In our hands, this results in recovery of 25% wild-type template (data not shown). MegAnneal instead uses uracil-containing dU-ssDNA, which is selectively degraded *in vitro* and *in vivo* in *dut⁺ ung⁺ E. coli*, yielding 20–50% stop-codon template in all libraries (Table 1). Additionally, inserting stop codons into the template plasmid eliminates the risk of wild-type selection. The use of two orthogonal mechanisms to prevent contamination by functional wild-type sequence ensures the generation of productive libraries.

To confirm that MegAnneal could produce libraries for directed evolution experiments, we selected variants from libraries I–III with improved expression and EE peptide affinity. Selected 3D5/EE scFv mutants exhibited 38-fold higher EE peptide affinity (Figure 4), six-fold higher solubility and nearly three-fold higher expression levels (33). We are currently using MegAnneal to engineer a FLAG peptide-binding scFv, and have improved expression and affinity from undetectable to low levels; we expect sequential rounds of library generation followed by selection to improve levels further (data not shown). Not only is MegAnneal an effective method to generate random mutagenesis libraries, but it is expected to be consistent with simultaneous randomization of distinct locales and combined random and targeted mutagenesis. For instance, using the error-prone PCR product from M2 scFv, a nested primer containing four or six consecutive randomized codons was used during asymmetric amplification, and the megaprimer produced was used to successfully synthesize ccc-DNA (data not shown).

Conclusions

Directed evolution experiments frequently use targeted or random mutagenesis to identify variants with enhanced characteristics or novel functions. Current methods typically require sequential enzyme digestion and ligation steps to incorporate altered sequences into plasmids. Here, we present a restriction enzyme-free approach termed MegAnneal, capable of generating large, diverse libraries with random mutations distributed across all or part of a gene. Use of an *in vitro* generated megaprimer controls mutation frequencies, while linear amplification steps reduce mutation frequency outside the intended gene. We have constructed six independent libraries with varied megaprimer size and number, and have demonstrated its ability to generate increased expression, solubility and antigen affinity for an scFv using phage selection.

Acknowledgments

FUNDING

This work was supported by grants from the National Institutes of Health [AI022439] and [GM095638] to J.A.M.

The authors thank Jeong-min Hyun and Carolyn Smith for their assistance in analyzing library sequences.

Abbreviations

dU-ssDNA	dUTP-containing single-stranded DNA
EE tag	EYMPME peptide
ssDNA	single stranded DNA
ccc-DNA	covalently closed circular double-stranded DNA
MBP-EE	maltose binding protein with a C-terminal EE tag

References

1. Carr PA, Church GM. Genome Engineering. *Nature Biotechnology*. 2009; 27:1151–1162.
2. Shivange AV, Marienhagen J, Mundhada H, Schenk A, Schwaneberg U. Advances in generating functional diversity for directed protein evolution. *Current Opinion in Chemical Biology*. 2009; 13:19–25. [PubMed: 19261539]
3. Weiss GA, Watanabe CK, Zhong A, Goddard A, Sidhu SS. Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proceedings of the National Academy of Sciences*. 2000; 97:8950–8954.
4. Murase K, Morrison KL, Tam PY, Stafford RL, Jurnak F, Weiss GA. EF-Tu binding peptides identified, dissected, and affinity optimized by phage display. *Chemistry and Biology*. 2003; 10:161–168. [PubMed: 12618188]
5. Fazelinia H, Cirino PC, Maranas CD. Extending iterative protein redesign and optimization (IPRO) in protein library design for ligand specificity. *Biophysical Journal*. 2007; 92:2120–2130. [PubMed: 17208966]
6. Petrounia IP, Arnold FH. Designed evolution of enzymatic properties. *Current Opinion in Biotechnology*. 2000; 11:325–330. [PubMed: 10975451]
7. Rubin-Pitel SB, Zhao H. Recent advances in biocatalysis by directed enzyme evolution. *Combinatorial Chemistry and High Throughput Screening*. 2006; 9:247–257. [PubMed: 16724916]
8. Yu H, Tyo K, Alper H, Klein-Marcuschamer D, Stephanopoulos G. A high-throughput screen for hyaluronic acid accumulation in recombinant *Escherichia coli* transformed by libraries of engineered sigma factors. *Biotechnology and Bioengineering*. 2008; 101:788–796. [PubMed: 18500764]
9. Alper H, Fischer C, Nevoigt E, Stephanopoulos G. Tuning genetic control through promoter engineering. *Proceedings of the National Academy of Sciences*. 2005; 102:12678–12683.
10. Tao H, Cornish VW. Milestones in directed enzyme evolution. *Current Opinion in Chemical Biology*. 2002; 6:858–864. [PubMed: 12470742]
11. Ang EL, Obbard JP, Zhao H. Directed evolution of aniline dioxygenase for enhanced bioremediation of aromatic amines. *Applied Microbiology and Biotechnology*. 2009; 81:1063–1070. [PubMed: 18813921]
12. Crameri A, Raillard S, Bermudez E, Stemmer WPC. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature*. 1998; 391:288–291. [PubMed: 9440693]
13. Zhao H, Giver L, Shao Z, Affholter JA, Arnold FH. Molecular evolution by staggered extension process (StEP) *in vitro* recombination. *Nature Biotechnology*. 1998; 16:258–261.
14. Lutz S, Ostermeier M, Moore GL, Maranas CD, Benkovic SJ. Creating multiple-crossover DNA libraries independent of sequence identity. *Proceedings of the National Academy of Sciences*. 2001; 98:11248–11253.
15. Arnold, FH.; Georgiou, G. Directed enzyme evolution: screening and selection methods. Vol. 230. Humana Press Inc; Totoawa: 2003.
16. Eijssink VGH, Gåseidnes S, Borchert TV, van den Burg B. Directed evolution of enzyme stability. *Biomolecular Engineering*. 2005; 22:21–30. [PubMed: 15857780]
17. Zahnd C, Spinelli S, Luginbühl B, Amstutz P, Cambillau C, Plückthun A. 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. [PubMed: 14754898]
18. Cirino PC, Mayer KM, Umeno D. Generating mutant libraries using error-prone PCR. *Methods in Molecular Biology*. 2003; 231:3–9. [PubMed: 12824595]
 19. Tobias, AV. Directed evolution library creation: methods and protocols. Arnold, FH.; Georgiou, G., editors. Vol. 231. Humana Press; Totowa, New Jersey: 2003. p. 11-16.
 20. Koyanagi T, Yoshida E, Minami H, Katayama T, Kumagai H. A rapid, simple, and effective method of constructing a randomly mutagenized plasmid library free from ligation. *Bioscience, Biotechnology and Biochemistry*. 2008; 72:1134–1137.
 21. Miyazaki K, Takenouchi M. Creating random mutagenesis libraries using megaprimer PCR of whole plasmid. *BioTechniques*. 2002; 33:1033–1038. [PubMed: 12449380]
 22. Qin W, Zhao A, Han Y, Wen W, Li Y, Chen G, Zhang Z, Wang H. A novel technique for efficient construction of large scFv libraries. *Molecular Biotechnology*. 2007; 37:201–205. [PubMed: 17952665]
 23. Ling MM, Robinson BH. Approaches to DNA mutagenesis: an overview. *Analytical Biochemistry*. 1997; 254:157–178. [PubMed: 9417773]
 24. Sen S, Dasu VV, Mandal B. Developments in directed evolution for improving enzyme functions. *Applied Biochemistry and Biotechnology*. 2007; 143:212–223. [PubMed: 18057449]
 25. Wang CL, Yang DC, Wabl M. Directed molecular evolution by somatic hypermutation. *Protein Engineering, Design and Selection*. 2004; 17:659–664.
 26. Zheng L, Baumann U, Reymond J. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Research*. 2004; 32:e115. [PubMed: 15304544]
 27. Tseng W, Lin J, Wei T, Fang T. A novel megaprimered and ligase-free, PCR-based, site-directed mutagenesis method. *Analytical Biochemistry*. 2008; 375:376–378. [PubMed: 18198125]
 28. Sanchis J, Fernández L, Carballeira JD, Drone J, Gumulya Y, Höbenreich H, Kahakeaw D, Kille S, Lohmer R, Peyralans JJ, et al. Improved PCR method for the creation of saturation mutagenesis libraries in directed evolution: application to difficult-to-amplify templates. *Applied Microbiology and Biotechnology*. 2008; 81:387–397. [PubMed: 18820909]
 29. Hogrefe HH, Cline J, Youngblood GL, Allen RM. Creating randomized amino acid libraries with the QuikChange Multi Site-Directed Mutagenesis Kit. *Biotechniques*. 2002; 33:1158–1165. [PubMed: 12449398]
 30. Kunkel TA. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proceedings of the National Academy of Sciences*. 1985; 82:488–492.
 31. Sidhu, SS.; Weiss, GA. Phage display: a practical approach. Clackson, T.; Lowman, HB., editors. Vol. 266. Oxford University Press; New York: 2004. p. 27-41.
 32. Hayhurst A, Happe S, Mabry R, Koch Z, Iverson BL, Georgiou G. Isolation and expression of recombinant antibody fragments to the biological warfare pathogen *Brucella melitensis*. *Journal of Immunological Methods*. 2003; 276:185–196. [PubMed: 12738372]
 33. Pai JC, Culver JA, Drury JE, Motani RS, Lieberman RL, Maynard JA. Conversion of scFv peptide-binding specificity for crystal chaperone development. *Protein Engineering, Design and Selection*. 2011; 24:419–428.
 34. Pai, JC. PhD dissertation. University of Texas at Austin; Austin, TX: 2010. Engineering peptide specific hyper-crystallizable antibody fragments (scFv) as potential chaperones for co-crystallization.
 35. Sutherland JN, Maynard JA. Characterization of a key neutralizing epitope on pertussis toxin recognized by monoclonal antibody 1B7. *Biochemistry*. 2009; 48:11982–11993. [PubMed: 19899804]
 36. Abrámoff MD, Magalhães PJ, Ram SJ. Image processing with ImageJ. *Biophotonics International*. 2004; 11:36–42.
 37. Rasila TS, Pajunen MI, Savilahti H. Critical evaluation of random mutagenesis by error-prone polymerase chain reaction protocols, *Escherichia coli* mutator strain, and hydroxylamine treatment. *Analytical Biochemistry*. 2009; 388:71–80. [PubMed: 19454214]

38. Wassman CD, Tam PY, Lathrop RH, Weiss GA. Predicting oligonucleotide-directed mutagenesis failures in protein engineering. *Nucleic Acids Research*. 2004; 32:6407–6413. [PubMed: 15585664]
39. Edelheit O, Hanukoglu A, Hanukoglu I. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnology*. 2009; 9:61. [PubMed: 19566935]
40. Rychlik W, Spencer WJ, Rhoads RE. Optimization of the annealing temperature for DNA amplification *in vitro*. *Nucleic Acids Research*. 1990; 18:6409–6412. [PubMed: 2243783]

Figure 1. Restriction enzyme-free generation of random mutation libraries using MegAnneal
 Template plasmid for du-ssDNA production is prepared by introducing three stop codons into the region of the gene targeted for mutagenesis; this prevents contamination by functional wild-type genes in the final library (step 1a). This template is then transformed into CJ236 *dut*⁻ *ung*⁻ *E. coli* followed by M13 phage super-infection to produce dU-ssDNA (step 1b). Simultaneously, random mutations are incorporated into the genetic region of interest using error-prone PCR and plasmid containing the complete, active gene (step 2a). The resulting product is purified and linearly amplified with a nested 3' primer to generate randomized megaprimers (step 2b). These are annealed to the template dU-ssDNA, and prime T7 polymerase to synthesize a dTTP- and megaprimer- containing copy of the template plasmid, whose ends are then covalently joined by T4 DNA ligase (step 3). Complementary base pairing between template and copied strands results in covalent, closed, circular double-stranded DNA (ccc-dsDNA), whose yield and quality can be readily monitored by gel electrophoresis (step 4). Upon transformation into *dut*⁺ *ung*⁺ *E. coli*, the uracil-containing plasmid strand is efficiently restricted by native nucleases while the megaprimer-containing strand is replicated and propagated (step 5), with each unique plasmid comprising a member of the final library.

Figure 2. Megaprimer amplification

A. Nested primers produce higher quality megaprimers than non-nested. The region of interest was first amplified under error-prone conditions with 5' and 3' flanking primers from plasmid harboring the full-length 3D5/EE scFv gene. Next, megaprimers were produced by asymmetric amplification of the error-prone PCR product with either the same 3' primer used in the initial error-prone amplification reaction (lane 1) or a nested 3' primer annealing 46 bp upstream of the original site (lane 2). **B.** Agarose gel analysis of megaprimers used to create select MegAnneal libraries. For libraries I, II and VII, the megaprimers were all produced with nested primers, resulting in sharp bands of the expected size, based on the initial error-prone amplification step, while library VI megaprimers were produced in a single step by asymmetric PCR under error-prone conditions. More detailed analysis of the megaprimer size and quality are tabulated in Table 1.

Figure 3. Error-prone megaprimers produce high-quality ccc-DNA

DNA gel analysis of **A**, the stop codon-containing dU-ssDNA template plasmid produced in step 1b and used to anneal with megaprimers, and **B**, the purified ccc-DNA, produced in step 4, for all libraries. While the dU-ssDNA purified from phage migrates as a single species, inefficiencies in the megaprimer annealing, extension and ligation processes result in multiple bands for the ccc-DNA product: the fastest-migrating band is the correct ccc-DNA product (*), the middle band represents fully amplified but unligated (or “nicked”) product while the slowest-migrating band is the undesirable strand-displaced product (**).

Figure 4. MegAnneal libraries yield EE binding scFvs with a range of activities

Monoclonal scFv-displaying phage were characterized after selection for EE-binding activity from MegAnneal libraries. After phage production in 96 well plates, the supernatant was used in two ELISAs to screen variants for scFv expression and activity and to select lead candidates for further characterization. The specific activity of each clone is shown, calculated as the ratio of EE-peptide to c-myc ELISA signal. Screen results of clones after selection from **A**. library I, constructed with 250 bp megaprimers targeting only the variable heavy chain, and **B**. library II, constructed with 750 bp megaprimers, targeting the entire scFv gene. Phage displaying the wild-type 3D5/EE scFv and the truncated scFv-stop gene are included as controls; error bars indicate the range of two ELISAs.

Figure 5. Comparison of QuikChange vs MegAnneal for targeted mutagenesis

A. PCR product using the QuikChange method with overlapping primers at varying annealing temperatures from 47 to 52°C for stop codon insertion into HCDR3 of 3D5/EE scFv (primers provided in Table 3). **B.** Product after primer annealing to dU-ssDNA and complementary strand synthesis: lane 1, ccc-DNA from 3' HCDR2_Stop primer; lane 2, ccc-DNA from 3' HCDR3_Stop; lane 3: ccc-DNA from simultaneous mutagenesis using both 3' HCDR2_Stop and 3' HCDR3_Stop. Arrow indicates the desired band.

Table 1

Library size and quality generated using MegAnneal

Libraries I-III were based off the 3D5/EE scFv gene, library VI was based off the hULB7 scFv gene, and library VII was based off the M2 scFv gene. Libraries IV and V serve as controls and were constructed using a modified Kunkel method (31).

Library	# primers	megaprimer length (nt)	cfu/ μ g DNA/transformation	% clones encoding full-length scFv*	error rate per scFv (%)	quality ratio (ccc-DNA:strand displaced DNA)
I	1	250	1.0×10^7	79.2	1–2 bp (0.1–0.2%)	1.00 ± 0.012
II	1	750	9.3×10^6	41.3	4–7 bp (0.5–0.9%)	1.02 ± 0.0048
III	1	150 – 250	1.1×10^6	54.9	1–4 bp (0.1–0.5%)	0.84 ± 0.036
IV	1 (50 bp)	-	4.1×10^7	50.0	-	0.96 ± 0.026
V	2 (50 bp)	-	1.1×10^7	37.0	-	0.70 ± 0.053
VI	1	750	1.0×10^6	62.5	2–5 bp (0.2–0.7%)	1.16 ± 0.013
VII	1	750	4.8×10^7	23.0	1–7 bp (0.1–0.9%)	0.76 ± 0.037

*. Based on 30–50 clones sequenced per library.

Table 2

Methods used to construct libraries in *E. coli*.

Method	# Mutagenic primers	Primer length (bp)	Controllable mutation frequency	Targeted Random mutagenesis	Random mutagenesis	Limitations	Reference
Mutator Strain	N/A	N/A	No	No	Yes	Mutations occur throughout vector, very low mutagenesis rate	(23)
Whole Plasmid PCR	N/A	N/A	Yes	No	Yes	Mutations occur throughout vector	(20)
DNA Shuffling	N/A	N/A	No	No	Yes	Homologous genes required, library size limitations from sequential cloning steps	(12)
Random mutagenesis with ligation	N/A	N/A	Yes	No	Yes	Library size limitations from sequential cloning steps	(19)
QuikChange	1	~50	N/A	Yes	No	Extensive primer design, optimization of PCR, high incidence of mis-priming, wild-type background contamination	(29)
Megaprimered QuikChange	1	450–5500	N/A	Yes	No	Moderate primer design, PCR optimization, wild-type background contamination	(27,28)
MEGAWHOP	1	200–1000	Yes	No	Yes	Moderate primer design, PCR optimization,	(21)

Method	# Mutagenic primers	Primer length (bp)	Controllable mutation frequency	Targeted Random mutagenesis	Random mutagenesis	Limitations	Reference
Kunkel Mutagenesis	multiple	~50	N/A	Yes	No	wild-type background contamination Choice of vector (M13 origin)	(31)
MegAnneal	multiple	50 – 1500	Yes	Yes	Yes	Moderate primer design, PCR optimization, choice of vector	-

Table 3

Oligonucleotides used in this study

Primer name	Primer sequence (5'-3')	Purpose
HCDR2	CACGGTGAGTGTGGCCCTTNNCTTVNHHMNYSBNGTTATASNNNSNNNSNNAYYSNNARDMYDAAATSNNTCCGATCCACTCCAGACC	Library V generation
HCDR3	CCTTGACCCCAAGTAATCCATAGCSNNNSNNNGCTSNNSNNNSNNABNTGCACAGTAGTATACG	Library IV-V generation
5' scback	GCATTACTCGCGGCCAGCCGGCCATGGCG	Libraries I-III, VI error-prone amplification
3' huck	GGCGGGAAGATGAAGACAGATG	Libraries I-III, VI error-prone amplification
5' pAKpel	ATGAAATACCTATTGCCTAC	M2 library error-prone amplification, Sequencing
3' pAK400	CCATTTTTCACCTTCACAGGTC	M2 library error-prone amplification
3' scforlong	GTATTAGGAATTCGGCCCCCGAGGCC	Library I-III megaprimer synthesis
3' Mo54	TGGCGGGAAGATGAAGACAGATGG	Library IV megaprimer synthesis
3' scM2_Mut	CGAGGAGACGGTGAGTGT	M2 megaprimer synthesis
5' 3STP_3D5_CDR2for	AGGGCCACACTCACCGTGGATAAGTCCAGCAGC	Stop codon insertion for library V template
3' HCDR2_Stop	GTGTGGCCCTGCCCTAGAACTTTTAGTTTAGCTCGTGGCCGCCATTATTTCG	Stop codon insertion for library V template
5' 3STP_3D5_CDR3for	TCAAGGAACCACAGTCACCGTCTCGGCCTCG	Stop codon insertion for library I-IV template
3' HCDR3_Stop	GGTTCCTTGACCCCCATTAAGCTCAGCTTTAACTTTCACAGTAGTATACG	Stop codon insertion for library I-IV template
5' scM2_Stop	GTGCCGTATATACTACTGTGCTCGATAATAAAGCTTAAGGTTATGACTAC	Stop codon insertion for M2 library template
3' scM2_Stop2	CAGTAGTATACGGCACTATCTTCGCTAGTCAG	Stop codon insertion for M2 library template
M13 Forward	GTAAACGACGGCCAG	Sequencing