A method for multi-codon scanning mutagenesis of proteins based on asymmetric transposons

Jia Liu and T.Ashton Cropp¹

Department of Chemistry, Virginia Commonwealth University, Richmond,
VA 23284, USA

¹To whom correspondence should be addressed.
E-mail: tacropp@vcu.edu

Received August 13, 2011; revised November 8, 2011;
accepted November 22, 2011

Edited by Jason Chin

Random mutagenesis followed by selection or screening is a commonly used strategy to improve protein function. Despite many available methods for random mutagenesis, nearly all generate mutations at the nucleotide level. An ideal mutagenesis method would allow for the generation of ‘codon mutations’ to change protein sequence with defined or mixed amino acids of choice. Herein we report a method that allows for mutations of one, two or three consecutive codons. Key to this method is the development of a Mu transposon variant with asymmetric terminal sequences. As a demonstration of the method, we performed multi-codon scanning on the gene encoding superfolder GFP (sfGFP). Characterization of 50 randomly chosen clones from each library showed that more than 40% of the mutants in these three libraries contained seamless, in-frame mutations with low site preference. By screening only 500 colonies from each library, we successfully identified several spectra-shift mutations, including a S205D variant that was found to bear a single excitation peak in the UV region.

Keywords: codon mutagenesis/directed evolution/transposon

Introduction

Directed evolution has now become a common laboratory tool to improve protein function. In these experiments the quality and diversity of mutant libraries is crucial for success. Most current random mutagenesis methods, such as error-prone polymerase chain reaction (PCR) (Cirino *et al.*, 2003) generate mutations at the nucleotide level. One disadvantage of such methods is that the protein sequence diversity is compromised due to codon redundancy and the rarity of consecutive nucleotide mutations. Scanning amino acid mutagenesis takes a more rational approach in which single amino acids are individually mutated across a segment or the length of protein sequence (Cunningham and Wells, 1989). For evolution purposes, an ideal mutagenesis method would allow one to substitute every amino acid of a protein sequence in a statistically equivalent manner (Wong *et al.*,

2006). Random insertion and deletion (RID) mutagenesis (Murakami *et al.*, 2002) is one of the early attempts to generate mutations at the amino acid level. RID allows random deletions of an arbitrary number of nucleotides from a gene of interest followed by insertions of any desired sequences. The method, however, requires multiple inter-molecular DNA manipulations and can be difficult to employ in generating large libraries. Recently, an *in vitro* transposon mutagenesis method has been developed to introduce random pentapeptide insertions into a protein sequence (Cao *et al.*, 1997; Hallet *et al.*, 1997; Hayes *et al.*, 1997). This method relies on the integrations of Mu transposons into target DNA sequences with relatively low site preference (Haapa *et al.*, 1999; Poussu *et al.*, 2004). Subsequent removal of the transposon generates a nucleotide ‘scar’ sequence resulting in protein mutations. Interestingly, the transposon sequences outside of the transposase recognition sites (R sites) can be modified to carry essentially any sequence. Notably, Jones and co-workers described a Mu transposon variant in which the flanking sequences were modified to contain the recognition sites of MlyI, a type IIS restriction endonuclease (Jones, 2005). Upon MlyI digestion, the transposon sequence can be removed from the target gene along with a triplet nucleotide that can be subsequently replaced by three new nucleotides (Baldwin *et al.*, 2008) or even a new protein domain (Edwards *et al.*, 2008). This approach, referred to as triplet nucleotide exchange (TriNEx), is perhaps the closest to achieve a non-redundant mutagenesis method (Baldwin *et al.*, 2008).

While both RID (Murakami *et al.*, 2002) and TriNEx can generate mutations in the unit of triplet nucleotides, our interest was to generate ‘codon mutations’ in which the reading frame of the mutation is defined. In a previous study, we reported a process for replacing single, random in-frame codons with a new codon of choice (Daggett *et al.*, 2009). This method employs the same MlyI transposon described in the TriNEx method (Jones, 2005; Baldwin *et al.*, 2008) and integrates an intein-based reading-frame selection system (Gerth *et al.*, 2004). This codon-based random mutagenesis limits unwanted mutations resulting in a ‘pure’ library (Daggett *et al.*, 2009). In this study, we report an advanced method that can generate up to three consecutive, random codon mutations. Importantly, this improved method uses an asymmetric Mu transposon that dramatically simplifies the DNA manipulation. As a demonstration of the power of this approach, we generated three libraries that scan single, di- or tri-NDT degenerate codon (where N = A, T, G or C and D = A, G or T) (Reetz *et al.*, 2008), respectively, throughout the gene encoding superfolder GFP (sfGFP) (Pedelacq *et al.*, 2006). By screening only 500 clones from each library, we successfully identified mutants that showed altered spectral properties.

Materials and methods

Construction of asymmetric mini-Mu transposon and target plasmid pITsfGFP

Asymmetric Mu transposons were constructed by amplifying the described MlyI TAG linker (Daggett *et al.*, 2009) by PCR with primers 5'-AAAGGATCcttagCGGCGCACGAAAAACGCGAAAGCGTTTCACGATAAATGCGAAAACGG AAGGTATTCGCAATAATCTTAAT-3' and 5'-AAAAGATCt ccgaCGGCGCACGAAAAACGCGAAAGCGTTTCACGATA AATGCGAAAATGAGTAAACGGGTCTGACTC-3' (BglII and BamHI sites underlined and five nucleotide difference shown in lowercased letters). This PCR was performed using *Taq* DNA polymerase with the following cycle conditions: 95°C for 3 min, 11 cycles of 95°C for 30 s, 66°C (−1°C/cycle) for 30 s and 72°C for 2 min, followed by 20 cycles of 95°C for 30 s, 66°C for 30 s and 72°C for 2 min and final extension at 72°C for 10 min. This method allowed for optimal annealing of the relatively long oligonucleotides. The PCR product was blunted and cloned in the pUC18 vector. Prior to transposition reactions, the transposon DNA was released from the vector by BamHI/BglII digestion and gel-purified using a QIAquick gel-extraction kit (Qiagen, Valencia, CA, USA). The pITsfGFP plasmid was generated by cloning the sfGFP gene into the pIT vector (Daggett *et al.*, 2009). The sfGFP gene used in this study is a synthetic gene that has been optimized for expression in *Escherichia coli* and is void of required restriction sites (Liu *et al.*, 2010). Sequence details of all plasmids are contained in the Supplementary Information section.

Generation of a sfGFP library with random in-frame transposon insertions

Transposition reactions were performed in 20 µl solutions containing 450 ng pITsfGFP, 125 ng gel-purified transposon DNA, 1 unit of HyperMu MuA transposase (Epicentre Biotechnologies, Madison, WI, USA), 50 mM Tris-acetate (pH 7.5), 150 mM potassium acetate, 10 mM magnesium acetate and 4 mM spermidine. The reaction was incubated at 30°C for 4 h and halted by the addition of 2 µl of 0.1% sodium dodecyl sulfate followed by heat inactivation at 70°C for 10 min. The reaction was cooled on ice and 1 µl of the solution was transformed into 50 µl electro-competent *E. coli* cells (DH10B). The transformants were recovered at 37°C for 1 h and plated on Luria Bertani (LB) agar supplemented with 50 µg/ml kanamycin and 40 µg/ml ampicillin. The plates were incubated at room temperature for 48 h. More than 10 000 colonies resulted from the 20 µl transposon reaction and were collected to build a pITsfGFP-transposon library. Twenty-three colonies were randomly picked from this library and digested with FastDigest MlyI and EcoRI (Fermentas Inc., Glen Burnie, MD, USA) to verify the randomness of transposon insertions. The library DNA was extracted and digested with NheI/EcoRI (Fermentas). This yielded four DNA fragments (pIT backbone-transposon, pIT backbone, sfGFP-transposon and sfGFP). The sfGFP-transposon fragment was isolated and ligated into the NheI/EcoRI site of a modified pTrcHisA expression vector (Invitrogen Corporation, Carlsbad, CA, USA) in which the BsgI and BpmI sites had been removed (see Supplementary data). A total of 30 000 colonies were collected from this ligation, ensuring the full coverage of library diversity. The

Table 1. Oligonucleotides in inverse PCR reactions

Name	Sequence	Use
FF1	TACTTTTTCTGGAGCCGTCGGA	First PCR, Fwd, remove one codon
FF2	TTTTTCGTGTGCAGTCGGA	First PCR, Fwd, remove two codons
FR1	AHNAATCAACGACTTT GCGCCGCTAAG	First PCR, Rev, introduce one codon
FR2	AHNAHNAATCAACGAC TTTGCGCCGCTAAG	First PCR, Rev, introduce two codons
FR3	AHNAHNAHNAATCAACGA CTTTGCGCCGCTAAG	First PCR, Rev, introduce three codons
SF0	TAACGTGCAGTTACAAGTCGTTGATT	Second PCR, Fwd, cleave to new codons
SF1	CATCGTGCAGATGCGCCGCTAAG	Second PCR, Fwd, remove one codon
SR	CTTCGTGCAGTAAATGCGCCGCTAAG	Second PCR, Rev, cleave to target gene

Note: BsgI and BpmI sites are underlined.

library DNA was then extracted and used as template in the following PCRs.

Generation of in-frame mutations

Single, di- and tri-NDT libraries (in which N = A, T, G or C and D = A, G or T) were constructed using the same pTrcHisA-sfGFP-transposon library DNA as the template but different oligonucleotide pairs in the PCR amplification. The optimum template concentrations and annealing temperatures were individually determined for each reaction. The PCRs were performed in four tubes of 100 µl solution each containing X pg/µl library DNA (see below), 0.5 µM each of corresponding forward and reverse oligonucleotides, 0.2 mM dNTPs, 0.005 U/µl of Phusion DNA polymerase (New England Biolabs, Ipswich, MA, USA), 1× Phusion HF buffer (NEB, providing 1.5 mM MgCl₂) and additional 2 mM MgSO₄. X indicates the optimum template concentrations and were 100, 125 and 200 pg/µl for single, di- and tri-NDT library, respectively. The PCR cycle conditions were: initial denaturation at 98°C for 2 min, 18 cycles of 98°C for 10 s, Y°C for 30 s and 72°C for 2 min followed by a final extension at 72°C for 10 min. Y represents the optimum annealing temperatures and were 57, 59 and 65°C for single, di- and tri-NDT library, respectively. The oligonucleotide pairs were FF1/FR1, FF2/FR2 and FF2/FR3 for single codon, di-codon and tri-codon libraries, respectively (see Table 1 for sequences). The PCR product was purified using QIAquick PCR purification kit (Qiagen). The purified PCR product of single codon library was digested with BpmI in a 50 µl reaction mixture containing 500 ng DNA, 1 U FastDigest BpmI (Fermentas) and 1× FastDigest buffer (Fermentas). The reaction was incubated at 30°C for 4.5 h followed by heat inactivation at 65°C for 20 min. The purified PCR product of di-codon or tri-codon library was digested with BsgI in a 50 µl reaction mixture containing 500 ng DNA, 6 U BsgI (NEB), 80 µM S-adenosylmethionine (SAM) and 1× buffer 4 (NEB). The reaction was incubated at 37°C for 4.5 h followed by heat inactivation at 65°C for 20 min. The digestion product of each library was treated with Klenow fragment by adding 1.7 µl of 1 mM dNTPs and 0.5 U of Klenow fragment (NEB) into the heat-inactivated reaction solution. The Klenow treatment was performed at

25°C for 16 min and stopped by addition of 1.1 µl of 0.5 M ethylenediaminetetraacetic acid, pH 8.0, followed by heat inactivation at 75°C for 20 min. The Klenow-treated digestion product was gel-purified using QIAquick gel-extraction kit (Qiagen) and then subjected to an intra-molecular ligation that contained 2.5 ng/µl DNA, 0.5 Weiss unit/µl T4 DNA ligase (Fermentas), 40 mM Tris-HCl (pH 7.8 at 25°C), 10 mM MgCl₂, 10 mM DTT and 0.5 mM ATP in a 10 µl reaction. The ligation reaction was incubated at 16°C overnight and then heat inactivated at 70°C for 7 min. Several 10 µl ligation reactions were performed for each library, pooled and then purified by ethanol precipitation. The purified ligation product was transformed into electro-competent DH10B *E. coli* cells. The transformants were recovered at 37°C for 1 h and then plated on LB agar supplemented with 100 µg/ml ampicillin. In all, 2×10^5 colonies were collected for each library and the DNA was extracted to serve as the template in the second inverse PCRs. The conditions for the second inverse PCR were generally the same as the first one with, again, individually determined optimum template concentrations and annealing temperatures. The template concentrations were 200, 125 and 200 pg/µl for single, di- and tri-NDT library, respectively. The annealing temperatures were 67, 63 and 66°C for single, di- and tri-NDT library, respectively. The oligonucleotide pairs for single, di- and tri-NDT libraries were SF0/SR, SF0/SR and SF1/SR, respectively (see Table 1 for the sequence). The PCR product was purified, digested with BsgI, treated with Klenow fragment, gel-purified, re-ligated and transformed as described above. More than 2×10^5 colonies were collected from each library and the library DNA was extracted for further characterization.

Screening for spectra-shift mutants

Five hundred colonies were randomly chosen from each library and inoculated into 96-well plates. Each well contained 1 ml LB media supplemented with 100 µg/ml ampicillin and 1 mM isopropyl-β-D-1-galactopyranoside (IPTG). The plates were grown at 37°C for 12 h. The expression culture (50 µl) from each well was diluted with 250 µl distilled water and then loaded into 96-well fluorescence plates. The OD₆₀₀ for each sample is generally between 0.15 and 0.2 under these conditions. The plates were screened using a SpectraMax M5 microplate reader (Molecular Devices, Sunnyvale, CA, USA) with 2 nm bandpass. Screening was performed using both 400 nm excitation wavelength with an emission range from 430 to 600 nm and 460 nm excitation wavelength with an emission range from 490 to 540 nm. The background fluorescence of DH10B cells transformed with empty pTrcHisA vector was subtracted from each sample reading. The calibrated spectra data were then normalized and compared with the original sfGFP spectrum. Candidate spectra-shift mutants were confirmed by three replicates as described above. The spectra-shift mutants were then diluted to an OD₆₀₀ of 0.2000 ± 0.0050 and the whole-cell spectra were recorded using a model F4500 fluorescence spectrophotometer (Hitachi, Tokyo, Japan). The obtained excitation and emission spectra were calibrated, normalized and compared with the normalized spectrum of original sfGFP. These spectra-shift mutants were sequenced by DNA analysis facility on Science Hill at Yale University (New Haven, CT, USA).

Results and discussion

Construction of a functional asymmetric Mu transposon

Previously, transposons have been used successfully to mutate sequence with foreign segments of DNA (Baldwin *et al.*, 2008; Daggett *et al.*, 2009). We hypothesized that they could also be used to generate new annealing sites for oligonucleotides within a plasmid. The Mu transposon is known to accommodate nucleotide changes outside of the MuA transposase recognition sequences (Poussu *et al.*, 2004; Jones, 2005). These sequences have always been reported as symmetrical inverted repeats, however, and to the best of our knowledge there has been no attempt generating asymmetric transposons. Herein, we designed an asymmetric Mu transposon that contains five differing nucleotides outside of the R sites (Fig. 1B). Upon transposon insertion, the asymmetric terminal sequences are delivered into a random position of a target gene which can then serve as unique oligonucleotide binding regions, or ‘handles’ in PCR for the installation of asymmetric type IIS or IIG restriction endonuclease sites. When located at the 3′-end of the oligonucleotide, a five-nucleotide difference is sufficient for the specific annealing of carefully designed oligonucleotides, despite the homology of the remaining sequence.

In addition to primer annealing sequences, the transposon contains a DNA segment for reading-frame selection (Fig. 1A). The selection of reading frame is based on the self-splicing of a VMA (vacuolar ATPase subunit in *Saccharomyces cerevisiae*) intein and functional TEM-1 β-lactamase (Gerth *et al.*, 2004). When and only when transposon insertions occur in the same reading frame with the gene of interest, will the leading peptide from pIT vector and the reading-frame selection element from the asymmetric transposon produce the full-length β-lactamase that can confer resistance to ampicillin. The control of transposon insertions ensures that all the subsequent mutations occur in the frame of the target gene as these mutations are generated from the oligonucleotides that bind to the asymmetric ends of transposon.

The efficiency of insertion of the asymmetric transposon was estimated to be ~1% by comparison to a control plate containing only kanamycin selection (for plasmid maintenance). This efficiency is comparable to other Mu transposon variants (Daggett *et al.*, 2009), suggesting that the nucleotides flanking the transposase recognition sites need not be identical. Wild-type Mu transposons are known to insert into target DNA sequences with relatively low site preference (Haapa *et al.*, 1999; Poussu *et al.*, 2004). To ensure that this is still the case, 23 colonies were randomly chosen from the pITsfGFP-transposon library and digested with EcoRI/MlyI to verify randomness of insertions. Indeed, insertions occurred throughout the entire sfGFP gene (Fig. S1), suggesting that the asymmetric terminal sequences do not alter the randomness of transposon insertions.

Generation of in-frame mutations

The overall process for mutation generation is shown in Fig. 1A. Because transposons insert randomly into the entire pIT plasmid, we chose to isolate those insertions inside the sfGFP gene by digesting the transposon library using NheI and EcoRI restriction sites that flank the sfGFP gene. The DNA fragment corresponding to sfGFP with transposon insertions can be transferred into any user-defined vector

Fig. 1. Multi-codon scanning mutagenesis. (A) pTsfGFP contains the Tat signal peptide and VMA intein N-terminal domain and the target gene fused in frame. The asymmetric transposon contains a region for reading-frame selection consisting of a fusion of the VMA intein C-terminal domain and TEM-1 β -lactamase (black). (B) The new sequence containing the (i) asymmetric annealing sites, (ii) introduced codons and (iii) ‘anchor’ sequences are introduced in the first PCR as shown. (C) A second PCR removes the anchor sequence resulting in seamless mutations. The lowercased letters in oligonucleotides indicate non-paired nucleotides.

such as pTrcHisA. The cleared transposon library was used as the template in an inverse PCR (Ochman *et al.*, 1988). In this PCR, two oligonucleotides anneal distinguishably to each transposon end and amplify the entire plasmid in an inverse direction (Fig. 1B) with the product being a library of linearized plasmids. The forward oligonucleotide introduces a *BsgI* or *BpmI* restriction site that can cleave one or two in-frame codons from the gene of interest. The reverse oligonucleotide supplies up to three new codons and an ‘anchor’ sequence that can serve as the oligonucleotide binding region in a second PCR. *BsgI* or *BpmI* digestion of this PCR product followed by re-ligation yields a library of plasmids each containing a random codon ‘scar’ that is filled with new codons and an ‘anchor’ sequence. The resultant library installed with the ‘anchor’ sequence is then subjected to a second inverse PCR in which two oligonucleotides anneal to the ‘anchor’ (Fig. 1C). Subsequent processing by *BsgI* digestion therefore generates a mutant library in which each clone has up to three consecutive, random in-frame codons replaced by codons of choice.

In comparison with previous methods (Daggett *et al.*, 2009), one of the most important advances is that the number of codons to be mutated is easily adjusted by using different oligonucleotide pairs with the same transposon library template. This is achieved by precise placement of the *BsgI* or *BpmI* sites in oligonucleotides such that a certain number of nucleotides can be removed from the target sequence. Moreover, the identity of the codon replacement(s) is easily dictated by the oligonucleotide used. In the examples shown here, we have demonstrated our method by generating in-frame mutations for up to three consecutive codons; however, mutations of four and five codons may be possible.

Table II. Mutagenesis efficiency

Library	One	Two	Three
Total sequenced	50	50	50
Incorrect	21	20	28
Correct	29	30	22
% Correct	58	60	44

Overall efficiency of library construction

To assess the quality of the library, 50 randomly chosen colonies from each library were sequenced. The sequencing results are summarized in Table 2 (see Table SI for detailed mutation types). The frequency of clones containing seamless, in-frame mutations in single, di- and tri-NDT library were 58, 60 and 44%, respectively. This suggests that the described method has sufficient mutagenesis efficiency for the construction of large-scale libraries. The incorrect mutations generally arose from incomplete *BsgI* or *BpmI* digestion (resulting in uncut oligonucleotide sequence) or overdigestion (resulting in one additional nucleotide deletion).

To analyze if these mutations are distributed uniformly along the sfGFP, the cumulative fraction of correct mutations were plotted to amino acid position of sfGFP (Fig. 2). Again the quality of the three libraries is comparable. While mutations occur in all regions of the sfGFP protein there is a mild preference for the N-terminal mutations, likely the result of transposon target sequence bias. Next we examined the amino acid composition of the introduced codons. The degenerate sequence ‘NDT’ (12 codons/12 amino acids) is a balanced mix of codons that can increase the chance of positive hits in screening (Reetz *et al.*, 2008). Ideally, these

Fig. 2. Distribution of identified mutations. The mutations are plotted to the amino acid position of sfGFP. For consecutive mutations, the positions of the first mutations are chosen for plotting the curve.

Fig. 3. Amino acid composition of introduced codons. The amino acids encoded by NDT codons are Gly (GGT), Val (GTT), Leu (CTT), Ile (ATT), Cys (TGT), Ser (AGT), Arg (CGT), His (CAT), Asp (GAT), Asn (AAT), Phe (TTT) and Tyr (TAT).

12 codons can be delivered into a protein sequence with equal probability. In this study, however, we observed bias of the introduced codons (Fig. 3). The expected average number of each of the 12 amino acids is 13. The mutants in the current study are rich in Arg and His mutations and short of Ile and Phe mutations. This codon bias most likely arises from the synthesis of oligonucleotides using automated mixing and could be easily controlled using altered nucleotide mixes.

Identification of spectra-shift mutants of sfGFP

As a proof of principle we decided to perform multi-codon scanning on sfGFP and search for variants that exhibited altered spectral properties. We successfully identified several spectra-shift mutants by screening only 500 colonies from each library (Fig. 4). The excitation peaks of these mutants shifted up to 15 nm to blue and up to 25 nm to red as compared with the 490 nm peak in original sfGFP. For the emission peaks, up to 6 nm blue shift and up to 12 nm red shift were observed. Interestingly, among those identified spectra-shift mutations is a red-shift mutation S203Y which has been previously reported in yellow fluorescent protein (Ormo *et al.*, 1996). This tyrosine mutation introduces an aromatic ring that can overlay with the tyrosine residue in chromophore and therefore stabilize the excited π -electrons of chromophore. More interestingly, an S205D mutant was found to be 7-fold more fluorescent than the original sfGFP when excited under UV light but non-fluorescent when excited under visible light (Fig. 5). Characterization of the

Fig. 4. Normalized excitation and emission of spectra-shift mutants. For excitation spectra, the emission wavelengths are 530 nm for S203Y mutant, 520 nm for S205C and Q204L/S205V/V206D mutants and 515 nm for all other mutants. For emission spectra, the excitation wavelengths are 505 nm for S203Y and 490 nm for all other mutants.

excitation spectra showed that the S205D mutation generated a single peak at 395 nm as compared with a minor peak at 395 nm and a major peak at 490 nm in the original sfGFP. The 395 and 490 nm excitation peaks are known to arise from the different ionization states of the chromophore (Brejc *et al.*, 1997). The neutral/protonated chromophore is responsible for the 395 nm peak. Upon light irradiation, the chromophore becomes deprotonated and can absorb light in the visible region (Brejc *et al.*, 1997). We speculate that the enhanced fluorescence under UV light and the elimination of fluorescence under visible light in the S205D mutant are due to the change of ionization states of the chromophore. The S205 residue in the sfGFP structure (PDB entry 2B3P) (Pedelacq *et al.*, 2006) projects towards the chromophore and the S205D mutation may alter this pK_a ensuring the protonated state.

We believe that three factors play important roles in the successful identification of these spectra-shift mutants. First, the highly stabilized template protein sfGFP (Pedelacq *et al.*, 2006) provides extra stability space for beneficial mutations to appear (Bloom *et al.*, 2006). Second, the NDT degenerate codon (12 codons/12 aa) increases the likelihood of a positive hit in screening by reducing the number of degenerated codons (as compared with the conventional NNK codon, 32 codons/20 aa, in which N = A, T, G or C and K = G or T). Third and most importantly, our approach allows the direct construction of mutant libraries with high amino acid diversity, bypassing the limitation of codon degeneracy and the

Fig. 5. Normalized excitation and emission spectra of S205D. For the emission spectra, the original sfGFP and S205D are excited at 395 and 485 nm. When using 395 nm excitation wavelength, the emission spectra of sfGFP and S205D mutant overlay perfectly. The S205D mutant shows no emission when excited at 485 nm.

rare chance of consecutive nucleotide mutations that usually compromise the diversity of libraries constructed by traditional random mutagenesis methods.

In summary, we have developed a mutagenesis method for generating up to three consecutive, random in-frame mutations. Characterization of the constructed libraries suggests that this new method has sufficient mutagenesis efficiency for building large-scale libraries. The ability to control the reading frame enables the delivery of up to three NDT or other desired codons into a protein sequence. As mutant libraries with 'codon mutations' are highly diversified, beneficial mutants may be identified by screening a very small fraction of mutants from the library.

Supplementary data

Supplementary data are available at *PEDS* online.

Acknowledgements

We thank Professor Vincent Lee (University of Maryland) and Professor Jody Turner (Virginia Commonwealth University) for providing the access to a fluorescence plate reader.

Funding

This work was supported by the National Institutes of Health [Grant number R01GM084396], and Virginia Commonwealth University.

References

- Baldwin,A.J., Busse,K., Simm,A.M. and Jones,D.D. (2008) *Nucleic Acids Res.*, **36**, e77.
- Bloom,J.D., Labthavikul,S.T., Otey,C.R. and Arnold,F.H. (2006) *Proc. Natl Acad. Sci.*, **103**, 5869–5874.
- Brejck,K., Sixma,T.K., Kitts,P.A., Kain,S.R., Tsien,R.Y., Ormo,M. and Remington,S.J. (1997) *Proc. Natl Acad. Sci.*, **94**, 2306–2311.
- Cao,Y., Hallet,B., Sherratt,D.J. and Hayes,F. (1997) *J. Mol. Biol.*, **274**, 39–53.
- Cirino,P.C., Mayer,K.M. and Umeno,D. (2003) *Methods Mol. Biol.*, **231**, 3–9.
- Cunningham,B.C. and Wells,J.A. (1989) *Science*, **244**, 1081–1085.
- Daggett,K.A., Layer,M. and Cropp,T.A. (2009) *ACS Chem. Biol.*, **4**, 109–113.
- Edwards,W.R., Busse,K., Allemann,R.K. and Jones,D.D. (2008) *Nucleic Acids Res.*, **36**, e78.
- Gerth,M.L., Patrick,W.M. and Lutz,S. (2004) *Protein Eng. Des. Sel.*, **17**, 595–602.
- Haapa,S., Taira,S., Heikkinen,E. and Savilahti,H. (1999) *Nucleic Acids Res.*, **27**, 2777–2784.
- Hallet,B., Sherratt,D.J. and Hayes,F. (1997) *Nucleic Acids Res.*, **25**, 1866–1867.
- Hayes,F., Hallet,B. and Cao,Y. (1997) *J. Biol. Chem.*, **272**, 28833–28836.
- Jones,D.D. (2005) *Nucleic Acids Res.*, **33**, e80.
- Liu,J., Castaneda,C.A., Wilkins,B.J., Fushman,D. and Cropp,T.A. (2010) *Bioorg. Med. Chem. Lett.*, **20**, 5613–5616.
- Murakami,H., Hohsaka,T. and Sisido,M. (2002) *Nat. Biotechnol.*, **20**, 76–81.
- Ochman,H., Gerber,A.S. and Hartl,D.L. (1988) *Genetics*, **120**, 621–623.
- Ormo,M., Cubitt,A.B., Kallio,K., Gross,L.A., Tsien,R.Y. and Remington,S.J. (1996) *Science*, **273**, 1392–1395.
- Pedelacq,J.D., Cabantous,S., Tran,T., Terwilliger,T.C. and Waldo,G.S. (2006) *Nat. Biotechnol.*, **24**, 79–88.
- Poussu,E., Vihinen,M., Paulin,L. and Savilahti,H. (2004) *Proteins*, **54**, 681–692.
- Reetz,M.T., Kahakeaw,D. and Lohmer,R. (2008) *ChemBiochem*, **9**, 1797–1804.
- Wong,T.S., Roccatano,D., Zacharias,M. and Schwaneberg,U. (2006) *J. Mol. Biol.*, **355**, 858–871.