African Journal of Biotechnology Vol. 9 (54), pp. 9277-9285, 27 December, 2010
Available online at http://www.academicjournals.org/AJB
ISSN 1684–5315 © 2010 Academic Journals

Review

Progress in strategies for sequence diversity library creation for directed evolution

Zhang Tong, Wu Zhenfang, Chen Hui*, Wu Qi, Tang Zizhong, Gou Junbo Wang Lihua, Hao Weiwei, Wang Chunmei and Li Chunmei

College of Life Science, Sichuan Agricultural University, Ya'an, Sichuan, China 625014.

Accepted 27 December, 2010

Protein engineering has been the most attractive strategy for biologists to redesign enzymes. As the simplest technique of protein engineering, directed evolution has been applied to many fields, such as industry, agriculture and medicine. An experiment of directed evolution comprises mutant libraries creation and screening or selection for enzyme variants with desired properties. Therefore, a successful application of directed evolution depends on whether or not one can generate a quality library and perform effective screening to find the desired properties. Directed evolution is already increasingly used in many laboratories to improve protein stability and activity, alter enzyme substrate specificity, or design new activities. Meanwhile, many more effective novel strategies of mutant library generation and screening or selection have emerged in recent years, and will continue to be developed. Combining computational/rational design with directed evolution has been developed as more available means to redesign enzymes.

Key words: Protein engineering, directed evolution, sequence diversity creation, novel strategy, computational design, rational design.

INTRODUCTION

Enzymes are nature's catalysts, tremendously accelerating the rates of a wide range of biochemical reactions with extreme specificity. Despite its high popularity for chemical synthesis in laboratory, enzymes almost always present too many problems to become suitable for industrial application. Limitations include incompatibility with non natural substrates, low stability, poor activity, and requirements for expensive cofactors (Farinas and Bulter, 2001). Therefore, applying enzymes to the industrial scale requires engineering strategies to improve their stability or activity. Protein engineering has been the most effective strategy for that purpose.

Rational design is reported to be an effective strategy to modify enzymes, including altering the reaction mechanism of the enzyme to catalyse new reactions, switching substrate specificity, expanding substrate specificity, and improving substrate specificity, such as enantioselectivity in kinetic resolutions, and most enzymes have been engineered successfully through this way (Garmaise, 2001; Jermutus et al., 2001; Antikainen, 2005). However, it is not always useful. It reacts only when the knowledge of sequence-structure-function is available, but there are small number of enzymes whose structure or/and function are available. Besides, rational design is complex and multifactorial. Consensus concept has been established to improve thermostability of phytase, however, the complex calculated program limits its extensive

*Corresponding author. E-mail: chenhui@sicau.edu.cn. Tel: +86-835-2886126.

Abbreviations: epPCR, Error-prone PCR; **RID,** random insertion/deletion mutagenesis; **RPR,** random priming recombination; **StEP,** staggered extension process; **ITCHY,** incremental truncation for the creation of hybrid enzyme; **PCR,** polymerase chain reaction; **RACHITT**, random chimeragenesis on transient templates; **SHIPREC,** sequence homology-independent site-directed chimeragenesis; **MEGAWHOP,** megaprimer polymerase chain reaction of whole plasmid; **EP-RCA,** error-prone rolling circle amplification; **SeSaM,** sequence saturation mutagenesis; **RAISE,** random insertion/deletion strand exchange mutagenesis; **RCA,** rolling circle amplification; **YLBS,** Y-ligation-based block shuffling; **MUPREC,** multiplex-polymerase chain reaction-based recombination; **SISDC,** sequence-independent site-directed chimeragenesis; **NRR,** non-homologous random recombination.

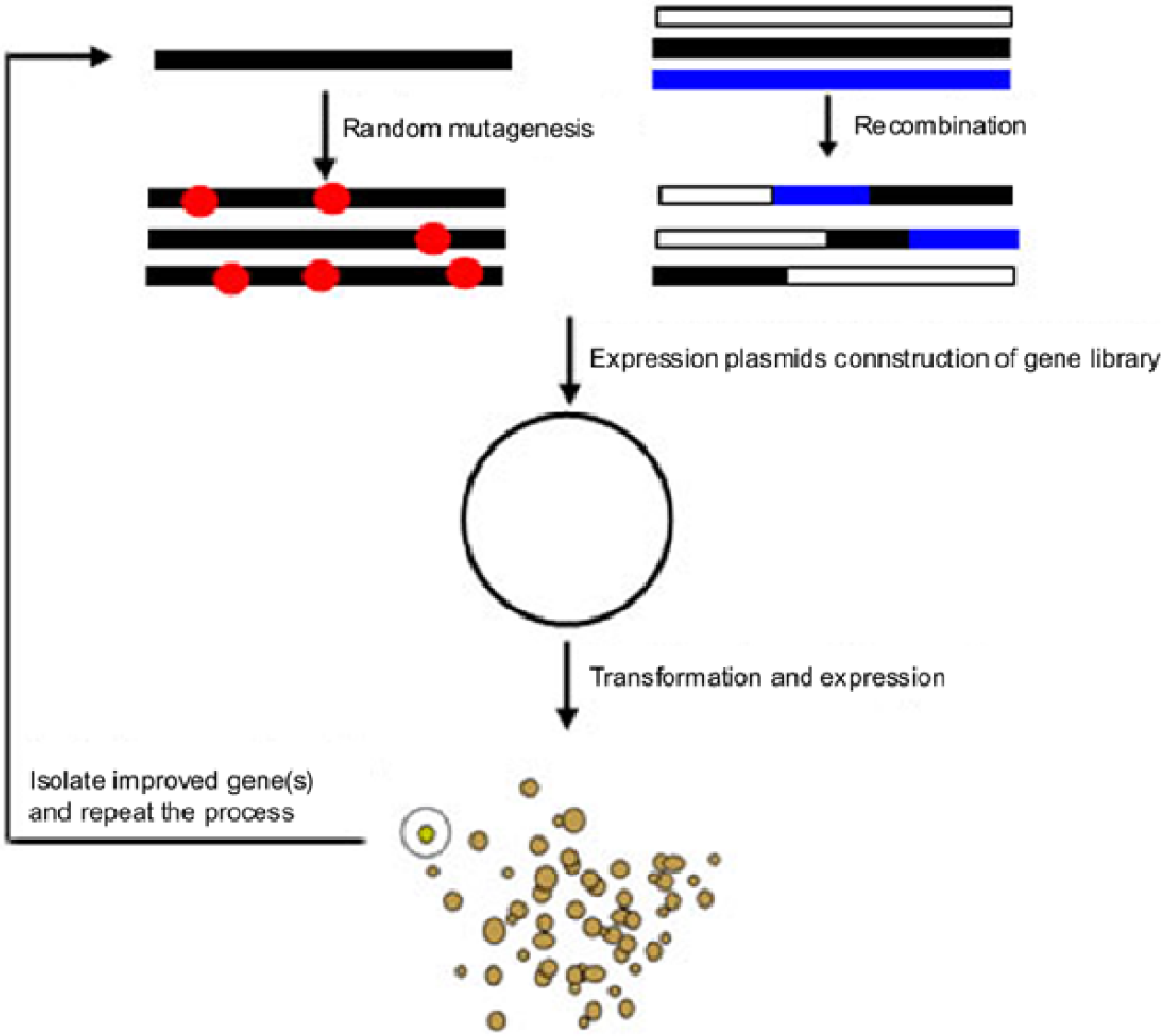

Figure 1. Schematic diagram of directed protein evolution. A typical directed protein evolution experiment was carried out starting at construction of gene mutant library, followed by gene library expression and improved gene(s) isolation.

application (Lehmann et al., 2000).

Directed evolution is a general term used to describe various techniques for generation of protein mutants and screening/selection of desirable functions (Farinas and Bulter., 2001) and it is a powerful complement to 'rational' engineering approaches in the absence of detailed mechanistic understanding (Dougherty et al., 2009). Directed evolution has emerged since the 1970s and it plays an important role in improving, altering enzyme properties and studying sequence-structure-function relationships of enzymes. The laboratory evolution experiment, contrast to Darwiman evolution, often has a defined goal, and the key processes-library creation and selection or screening-are carefully controlled by the experimenter (Figure 1). Comparing to rational design, as an “irrational design” strategy, it does not need to provide insight into sequence-structure-function relationships of proteins. As one of the most effective approaches to engineer enzymes, directed evolution has being used to redesign many enzymes for the requirement of industrial, medical and research application successfully.

A successful application of directed evolution requires one to generate a quality library and perform effective screening to find the desired properties. It is essential to bear a strategy for creating directed evolution sequence libraries that are rich in proteins with the desired enzymatic function, and we will focus on it here. The strategies of selection and screening for enzyme variants with the desired properties will not be discussed here, since you can find them elsewhere. This review is based on an overview of the different strategies available for creating directed evolution sequence libraries, to discuss

the recent advances in the techniques and strategies, including the introduction of computational strategies aimed at improving the quality and potential of mutant libraries.

OVERVIEW OF DIFFERENT STRATEGIES FOR MUTANT LIBRARY CREATION

Mutagenesis and recombination have been considered as two main sorts for the creation of sequence diversity libraries. The techniques of mutagenesis, such as error-prone PCR (epPCR) and random insertion/deletion mutagenesis (RID), generate sequence diversity in the form of point mutations, insertions or deletions. However, recombination techniques do not directly create new sequence diversity but combine existing diversity in new ways. Recombination techniques can be divided into homologous recombination and non-homologous recombination depending on whether genes contain sequence homology. A range of different techniques are available for recombining diverse sequences, such as DNA shuffling, random priming recombination (RPR), the staggered extension process (StEP) for homologous recombination, and incremental truncation for the creation of hybrid enzyme (ITCHY) for non-homologous recombination. The ultimate purpose for all recombination techniques is to recombine DNAs from different sources in new ways to form novel sequences.

Error-prone PCR is a familiar method used to generate variants with random mutations, which is designed to alter and enhance the natural error rate of polymerase on the bases of standard polymerase chain reaction (PCR) methods (Pritcharda and Corneb, 2005). The combination of introducing a small amount of Mn^{2+} and various ratios of nucleotides in the error-prone PCR process are usually effective to create sequence diversity. Higher concentrations of Mg^{2+} are usually used to stabilize non-complementary pairs. Although most enzymes have been designed successfully through this way, it is difficult. The frequency of beneficial mutations is generally low in error-prone PCR reactions. Sequential error-prone PCR is a better choice to generate variants, which performs several times of error-prone PCR to accumulate beneficial mutations (Kong et al., 2001). On the other hand, error-prone PCR produces libraries with biases. Random insertion/deletion mutagenesis offers a general method to reduce these biases as well as access insertions and deletions at random positions within the amino acid sequences (Murakami et al., 2003). However, the procedure of RID is complex, and the experimenter must be sure each step is working correctly.

Recombination techniques bypass the above limitations by allowing the collection of beneficial mutations from multiple genes. DNA shuffling is a method for *in vitro* recombination of homologous genes invented, which has been the most popular recombination technique in recent years (Stemmer, 1994). In DNA shuffling, a single DNA sequence or a population of DNA sequences are randomly fragmented and then reassembled into full-length, chimeric sequences by PCR, and they are usually called “single-gene DNA shuffling” and “gene family shuffling”, respectively. Many new shuffling strategies have emerged in recent years, such as synthetic shuffling, exon shuffling, whole-genome shuffling and Y-ligation-based block shuffling. Random priming recombination is a simple and efficient method for *in vitro* mutagenesis and recombination of polynucleotide sequences reported by Shao et al. (1998). RPR generally consist of two steps: Generation of a large number of short DNA fragments using random primers and production of a library of full-length sequences through DNA polymerization. The staggered extension process recombination is performed through series of polymerase-catalyzed primer extensions, generating cross hybridization during growing of gene fragments (Zhao et al., 1998). This method commences with priming the template sequence(s) followed by repeated cycles of denaturation, annealing and abbreviated extension. The extending fragments anneal to different templates in each cycle until full-length sequences are obtained (Figure 2). DNA shuffling, RPR and StEP all require the sequences similar enough to be recombined. Random chimeragenesis on transient templates (RACHITT) is a technique that is conceptually similar to StEP and DNA shuffling but is designed to produce chimera with a much larger number of crossovers. It can increase the frequencies of recombination between sequences with low homology, which generated an average of 14 crossovers per parental gene, a much higher rate than with the above recombination methods (Coco, 2003). RACHITT produces a single-stranded, full-length transient template containing uracil and single-stranded partial donor fragments. As one or more parental donor gene fragments can simultaneously anneal to the template, which generates high-frequency crossovers.

Nonhomologous recombination techniques make possible to construct hybrid proteins even when the genes have little or no sequence homology (Lutz et al., 2001). They also allow the efficient creation of new protein folds, while protein variants generated by homologous recombination or mutagenesis are most likely to maintain structural similarity to the parental proteins. Therefore, they enable the generation of protein structural diversity that may not exist in nature. A high-throughput screening is critical following library construction, since large numbers of nonfunctional progeny is present in the library created by non-homologous recombination. There are several non-homologous recombination techniques that are available for directed evolution, which include incremental truncation for the creation of hybrid enzyme, sequence homology-independent site-directed chimeragenesis (SHIPREC) (Udit et al., 2003), sequence homology-independent protein recombination (Sieber et

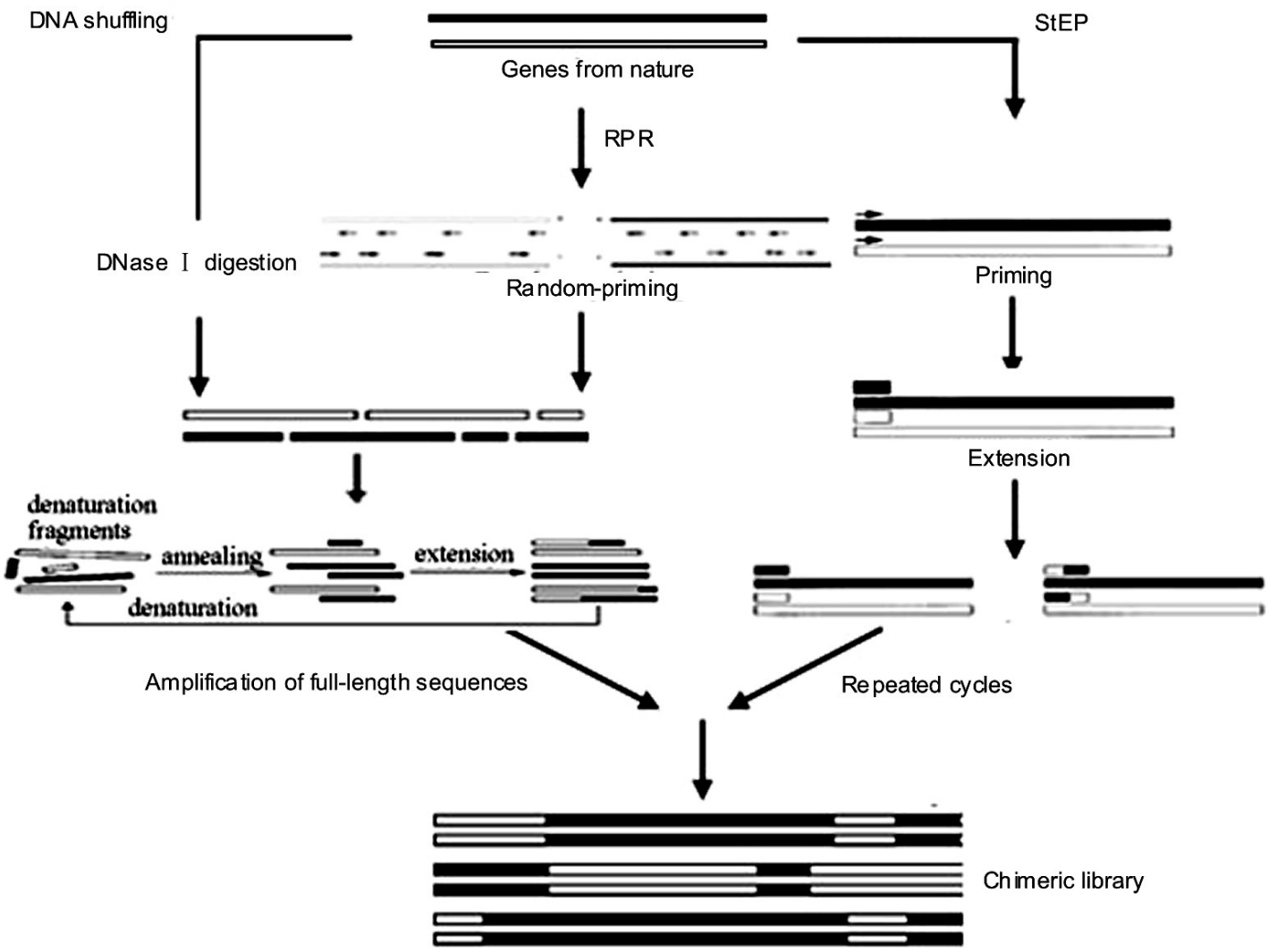

Figure 2. Schematic diagram of three familiar homologous recombination techniques. All methods start at a series of homologous genes. In DNA shuffling and RPR, these DNAs are fragmentized by DNase 1 digestion and random priming, respectively. These fragments are then reassembled into full-length chimerical genes in a self-priming reaction, followed by a standard PCR to amplify the full-length genes. In StEP oligonucleotide primers anneal to the denatured templates, a series of denaturation, annealing, and short extension are repeated until full-length sequences is obtained.

al., 2001) and some new strategies will be introduced below in detail.

NEW STRATEGIES FOR MUTANT LIBRARY CREATION

The above traditional methods for mutant library creation have redesigned many enzymes for the requirements of industrial, medical and research application. However, alternating and improving methods is necessary for the directed evolution of enzymes. In the following, this review will discuss new strategies for mutant library creation on the base of mutangenesis, homologous recombination and non-homologous recombination, respectively.

Mutangenesis

Although error-prone PCR methods remain one of the most popular approaches through mutangenesis to generate libraries for directed evolution, some more samples and effective methods on the basis of mutangenesis have emerged in recent years, such as megaprimer PCR of whole plasmid (MEGAWHOP) (Miyazaki, 2003), error-prone rolling circle amplification (EP-RCA) (Fujii et al., 2004), sequence saturation mutagenesis (SeSaM) (Wong et al., 2004), random insertion/deletion strand exchange mutagenesis (RAISE) (Fujii et al., 2006).

EP-RCA is a very simple random mutagenesis method using rolling circle amplification (RCA) described by Fujii et al. (2004). This method consists of only one RCA step

followed by direct transformation of the host strain. Compare to error-prone PCR, EP-RCA can yield mutants with an adequate mutation frequency for directed evolution experiments but it does not require treatments with any restriction enzymes or DNA ligases. EP-RCA has given rise to broad substitution and some mutants with improved ceftazidime resistance when the entire pUC19 plasmid was used which has TEM-1 β-lactamase gene, to check the effect of this technique.

RAISE method is also described based on gene shuffling (Fujii et al., 2006). The protocol of RAISE consists of only three steps: Firstly, obtain 100 to 300 bp DNA fragments from a DNaseI digestion; secondly, use of terminal deoxynucleotidyl transferase (TdT) to introduce random nucleotides at the 3' terminus of the digested DNA fragments; thirdly, use of self-priming PCR to obtain entire target gene which have introduced various lengths of random insertions, deletions and substitutions (Figure 3). Fujii et al. (2006) obtained some of the deletion mutations of TEM β-lactamase through RAISE that caused higher activities than point mutations, which indicates that RAISE method may be a powerful technique of protein engineering in future.

The SeSaM is a four-step method that can saturate every single nucleotide position of the target sequence with all four standard nucleotides (Wong et al., 2006). SeSaM method is partly similar to RAISE, with the difference in utilization universal base to tail the DNA fragments in SeSaM, rather than random nucleotides in RAISE. Finally, after full-length sequence synthesis, universal base replacement by standard nucleotides is necessary to introduce mutations for SeSaM (Figure 3). SeSaM overcomes limitations caused by biased polymerases, and it is a fast technique of mutagenesis to accomplish a process of directed evolution within 2 to 3 days.

Homologous recombination

Applying recombination for directed evolution is a breakthrough, which allows multiple genes rearrangement to generate improved enzymes. DNA shuffling is the most classic and popular method for homologous recombination, which means multiple genes used to recombine, must possess high similarity with each other. Similar to mutagenesis, many new techniques have emerged for the development of homologous recombination in recent years.

Compared to other recombination techniques, Y-ligation-based block shuffling (YLBS) rearranges DNA blocks and recombines point mutations (Kitamura et al., 2002). In the YLBS method, two types of single-stranded DNAs are used to form a “Y” structure with a stem and two branches, and YLBS performs sequence blocks to rearrange by repeating cycles of Y-ligation. Then YLBS introduces or eliminates sequence blocks efficiently, which can generate a huge diversity of shuffled protein. Therefore, YLBS is a much more effective strategy to obtain some novel proteins that do not exist in nature.

PCR-based recombination is based on homologous recombination of oligonucleotides which have about 20 bp overlap with each other. The following several PCR steps are performed to generate full-length target genes through self-priming of these oligonucleotides or the previous PCR products. PCR-based two-step DNA synthesis (PTDS) is a simple, rapid, high-fidelity and cost-effective technique described by Xiong et al. (2004). The protocol of PTDS only contains two step; synthesis of individual fragments and the entire sequence by PCR. High-fidelity DNA polymerase Pfu is often used to ensure the high-fidelity of PTDS. PTDS shows two prominent advantages compared to other methods: (i) The PTDS method is rapid, as the entire process is about 5 to 7 days; (ii) the PTDS method is suitable for reengineering long genes (5 to 6 kb). Multiplex-PCR-based recombination (MUPREC) is another high-fidelity method based on PCR, described by Eggert et al. (2005). A multiplex-PCR generates gene fragments that contain preformed point mutations, which are subsequently assembled into full-length genes by a recombination-PCR step.

Nonhomologous recombination

It is difficult for homologous recombination to recombine genes with little sequence homology, which leads to the development of techniques for non-homologous recombination. This approaches enable the generation of protein structural diversity that may or may not exist in nature, and potentially very useful in evolution of multifunctional proteins. In contrast to DNA shuffling and related methods, ITCHY invented by Ostermeie et al. (1999) does not rely on the parental genes containing regions of DNA sequence homology to create crossovers. There are some other new non-homologous recombination techniques described for recombining sequences without homology.

Sequence-independent site-directed chimeragenesis (SISDC) is a simple and general method that allows crossovers at multiple sites independent on DNA sequence identity (Hiraga and Arnold, 2003). SISDC comprises four steps: Firstly, align the nucleotide sequences of parent genes to determine consensus sequences for targeted sites; secondly, insert marker tags designed with the same restriction endonuclease site at both terminals by experimenter into targeted sites; thirdly, remove the tag regions by digestion with the restriction endonuclease introduced in the previous step; fourthly, mix fragments which can ligate with each other, which generate the final, chimeric library. Hiraga and Arnold (2003) have recombined β-lactamases TEM-1 and PSE-4 which only have 49% nucleotide sequence identity at seven sites to evaluate the SISDC method. More

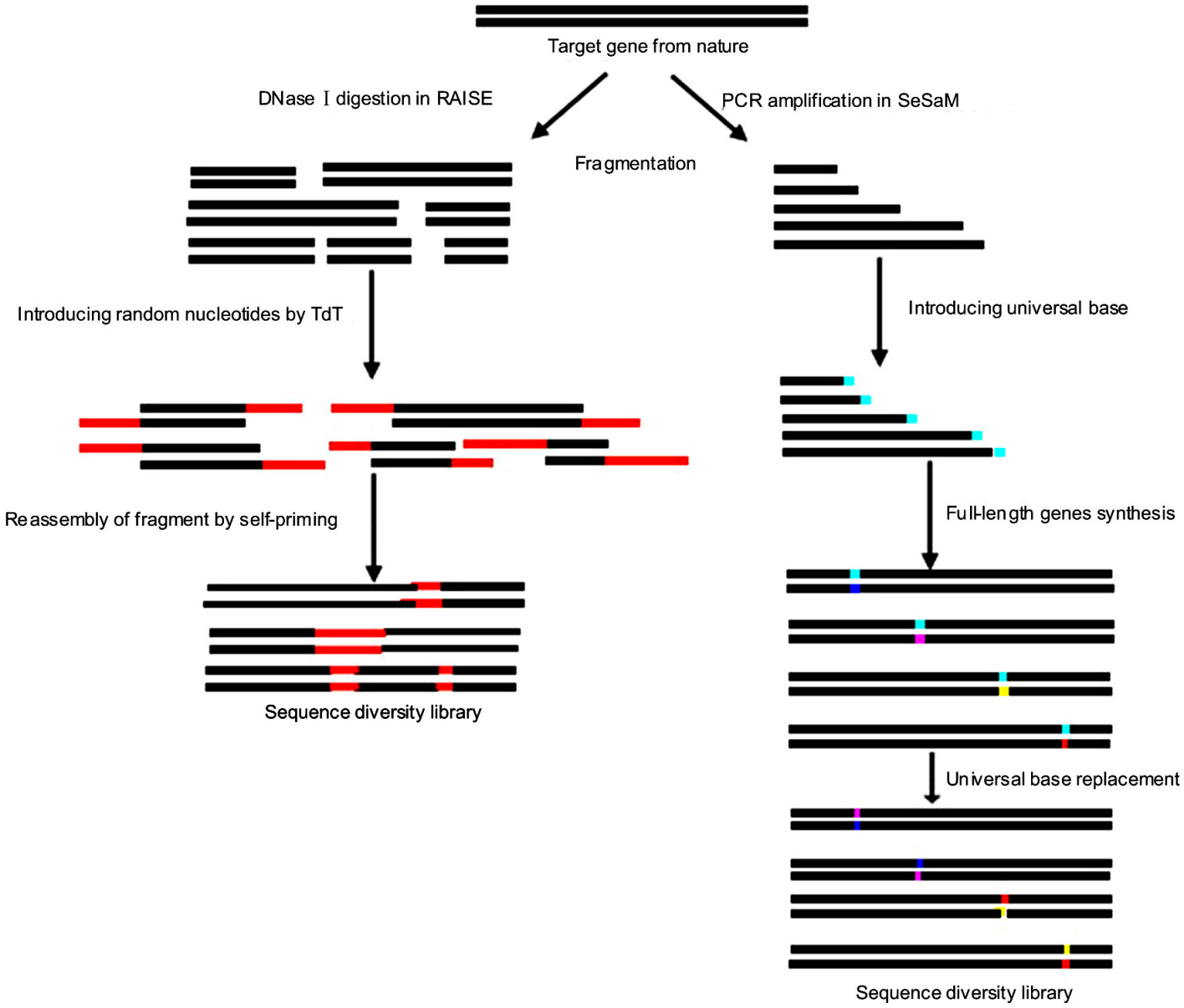

Figure 3. Schematic diagram of RAISE and SeSaM. RAISE and SeSaM are both starting at DNA fragmentation. In RAISE, the fragments are attached with random nucleotides by TdT, and then reassembled by self-priming. In SeSaM the fragments are introduced universally base by terminal transferase at 3′ terminus, followed by full-length synthesis and universal base replacement with standard bases, which are labeled with different color in the figure.

recently, non-homologous random recombination (NRR) method has been described, allowing recombining non-homologous proteins, which is difficult or impossible to accomplish through the previous techniques (Bittker et al., 2004; Doyon and Liu, 2007). NRR method consists of DNaseI fragmentation, blunt-end ligation/extension, and capping using two DNA hairpins to stop the extension. Bittker et al. (2004) applied NRR for recombining chorismate mutase (CM) enzymes, and found that functional CM mutants contained many insertions, deletions and rearrangements.

COMBINING RATIONAL/COMPUTATIONAL DESIGN WITH DIRECTED EVOLUTION

Combining rational/computational design with directed evolution has been a revolution of protein engineering in the recent years. Combine computational design pre-screening experimental library construction can diminish mutant library site, which can save a large number of time and money for high-throughput screening (Funke et al., 2005). Computational design is generally based on the structure of proteins, which comprises three stages.

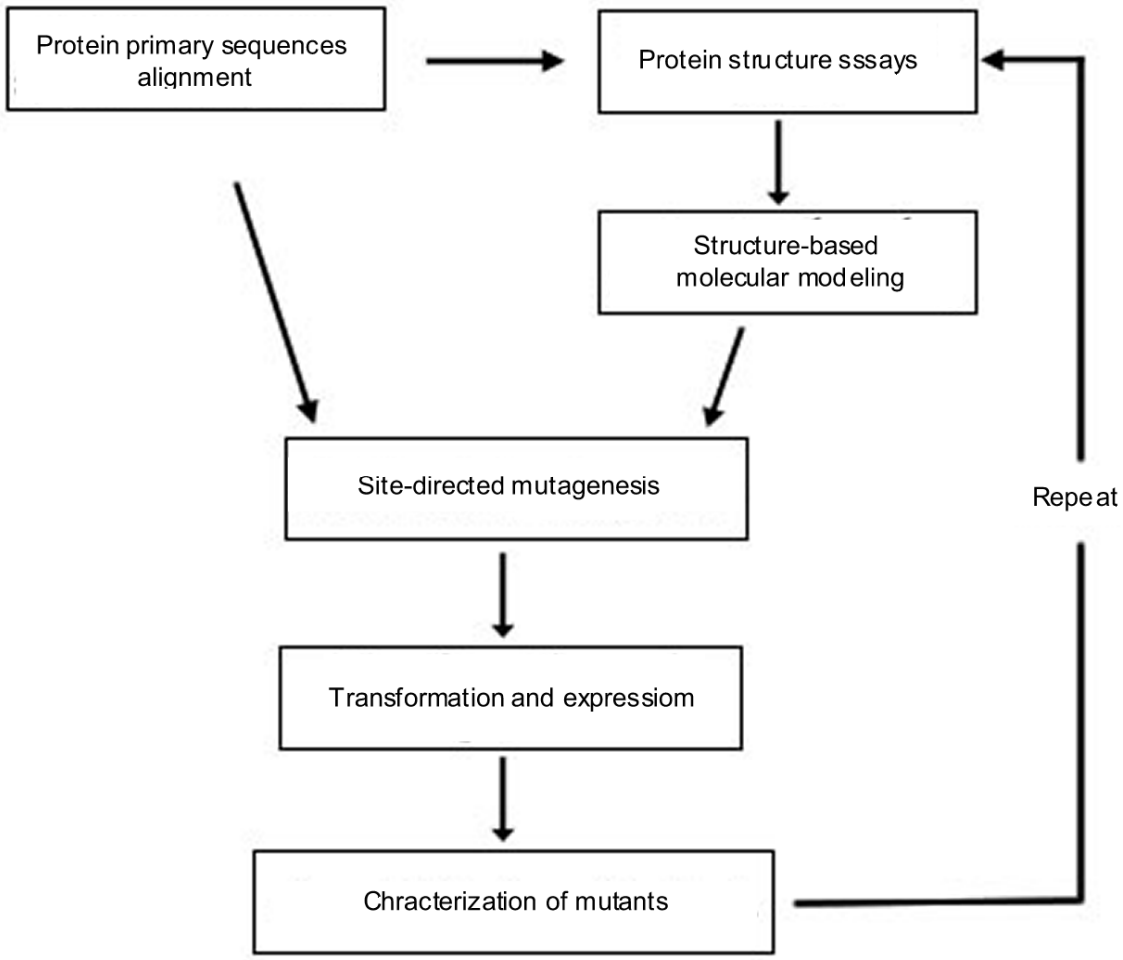

Figure 4. Schematic diagram of rational design. Protein primary sequences alignment and structure-based molecular modeling are usual foundation of rational protein design, and then site-directed mutagenesis is performed to prove the hypotheses. The whole process is often repeated until the improved gene(s) obtained.

Firstly, generate a chemical and geometric definition of the catalytic machinery. Secondly, generate placement of the active site residues through geometric definition. Thirdly, form a stereochemically complementary binding surface by optimizing the remainder of the active site (Dwyer et al., 2004). *De novo* protein design (Kuhlman et al., 2003) and Famclash (Saraf et al., 2004) were effective techniques of computational design which show the increasing power of computational design in the future.

Rational design is often used to cooperate with library diversity generation by structure analysis of a protein through computational techniques. Primary sequences alignment and secondary and spatial structure modeling are usual strategies for rational design. Diverse sequences will be tested to find out those that have functional properties and suppose why; which can be modeled in silico (Figure 4). Sandgren et al. (2003) compared the Cell2A amino acid sequences of *Hypocrea schweinitzii* and *Hypocrea jecorina* which have different thermostability, and found they show high similarity. Therefore, they substituted the differences to study which are important for thermostability (Sandgren et al., 2003). Combining rational design with directed evolution is semi-rational design of protein engineering, offering a target for directed evolution. This approach can greatly diminish the library size, which can save a great deal of time and money for the following high-throughput screening.

CONCLUSION

Directed evolution mimics the process of Darwiman evolution in a test tube, comprising mutant libraries creation and screening or selection for enzyme variants with the desired properties. The successful application of directed evolution depends on whether or not one can generate a quality library and perform effective screening strategy. The size and sequence diversity of a library are preconditions of screening or selection, so it is essential

Table 1. Different strategies for sequence diversity (mutant library) for directed evolution and their main advantages and limitations.

Category	Technique	Member
Mutagenesis	Generate sequence diversity in the form of point mutations, insertions or deletions. It is usually simple, but difficult to avoid the bias problem.	Error-prone PCR, RID MEGAWHOP,EP-RCA, SeSaM and RAISE
Homologous recombination	Recombine genes to form sequence diversity depending on homology. Aims for high recombination, but difficult to recombine some genes without high identity.	DNA shuffling, Gene family shuffling, RPR, StEP, RACHITT, MUPREC,PTDS and YLBS
Nonhomologous recombination	Recombine genes without sequence homology to form sequence diversity. Be able to obtain novel protein not exist in the nature, however, a high-throughput screening is critical because of the presence of large numbers of nonfunctional progeny in the libraries	ITCHY, SHIPREC, SISDC NRR

to develop more effective strategies for creating directed evolution sequence libraries that are rich in protein with the desired enzymatic function (Yuan et al., 2005; Neylon, 2004).

In this review, we, on the base of an overview of strategies for mutant library generation, have discussed some new techniques in the form of mutagenesis, homlogous recombination, and non-homologous recombination. The advantages and limitations of every strategy have been discussed above. The experimenter can choose the best one suitable for the experiment based on the enzyme used and the condition of the laboratory. Traditional strategies, such as error-prone PCR and DNA shuffling, remain the most popular for directed evolution. The combination of error-prone PCR followed by shuffling of selected mutants with improved function is the most commonly used strategy for directed evolution experiments, such as error-prone PCR followed by DNA shuffling or StEP. Non-homologous recombination techniques are more effective for complex enzymes evolution, and it may obtain novel protein not existing in nature. Therefore, non-homologous recombination techniques will be increasingly popular for directed evolution. Combination of rational/computational design with directed evolution has been developed as a new breakthrough for directed enzyme evolution. The use of computational design has expanded to include the thermostabilization of enzymes (Korkegian et al., 2005) and the redesign of an enzyme active site for improved catalytic activity (Park et al., 2006). However, it is still difficult to apply this approach for some complex enzymes, because of the limitation of protein modeling algorithms (Chica et al., 2005) (Table 1).

An effective screening or selection method is also necessary for a successful directed evolution, and many advanced technologies have emerged in recent years (Matsuura and Yomo, 2006; Lin and Cornish, 2002; Hibbert and Dalby, 2005). The advances in screening technologies and application were not been discussed here, as they could be found elsewhere.

REFERENCES

Antikainen NM (2005). Altering protein specificity: techniques and applications. Bioorg. Med. Chem. 13(8): 2701-2716.

Bittker JA, Le BV, Liu JM, Liu DR (2004). Directed evolution of protein enzymes using nonhomologous random recombination. Proc. Natl. Acad. Sci. USA, 101(18): 7011-7016.

Chica RA, Doucet N, Pelletier JN (2005). Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. Curr. Opin. Biotechnol. 16(4): 378-384.

Coco WM (2003). RACHITT: Gene family shuffling by random chimeragenesis on transient templates. Methods Mol. Biol. 231: 111-127.

Dougherty MJ, Arnold FH (2009). Directed evolution: new parts and optimized function. Curr. Opin. Biotechnol. 20(4): 486-491

Doyon JB, Liu DR (2007). Identification of eukaryotic promoter regulatory elements using nonhomologous random recombination. Nucleic Acids Res. 35(17): 5851-5860.

Dwyer MA, Looger LL, Hellinga HW (2004). Computational design of a biologically active enzyme. Science, 304(5679): 1967-1971.

Eggert T, Funke SA, Rao NM, Acharya P, Krumm H, Reetz MT et al., (Provide Complete Name) (2005). Multiplex-PCR-based recombination as a novel high-fidelity method for directed evolution. Chembiochem. 6(6): 1062-1067.

Farinas ET, Bulter T (2001). Directed enzyme evolution. Curr. Opin. Biotechnol. 12(6): 545-551.

Fujii R, Kitaoka M, Hayashi K (2004). Error-prone rolling circle amplification: the simplest random mutagenesis protocol. Nat. Protoc. 1(5): 2493-2497.

Fujii R, Kitaoka M, Hayashi K (2006). RAISE: a simple and novel method of generating random insertion and deletion mutations. Nucleic Acids Res. 34(4): e30.

Funke SA, Otte N, Eggert T, Bocola M, Jaeger KE, Thiel W (2005). Combination of computational prescreening and experimental library construction can accelerate enzyme optimization by directed evolution. Protein Eng. Des. Sel. 18(11): 509-514.

Garmaise M (2001). Rational beliefs and security design. Rev. Finance Stud. 14: 1183-1213.

Hibbert EG, Dalby PA (2005). Directed evolution strategies for improved enzymatic performance. Microb. Cell Fact. 4: p. 29.

Hiraga K, Arnold FH (2003). General method for sequence-independent site-directed chimeragenesis. J. Mol. Biol. 330(2): 287-296.

Huisman GW, Gray D (2002). Towards novel processes for the fine-chemical and pharmacentical industries. Curr. Opin. Biotechnol. 13: 352-358.

Jermutus L, Tessier M, Pasamontes L, van Loon AP, Lehmann M (2001). Structure-based chimeric enzymes as an alternative to directed enzyme evolution: phytase as a test case. J. Biotehnol. 85(1): 15-24.
Kitamura K, Kinoshita Y, Narasaki S, Nemoto N, Husimi Y, Nishigaki K (2002). Construction of block-shuffled libraries of DNA for evolutionary protein engineering: Y-ligation-based block shuffling. Protein Eng. 15(10): 843-853.
Kong X, Liu Y, Gou X, Zhu S, Zhang H, Wang X, et al., (Provide Complete Name) (2001). Directed evolution of alpha-aspartyl dipeptidase from *Salmonella typhimurium*. Biochem. Biophys. Res. Commun. 289(1): 137-142.
Korkegian A, Black ME, Baker D, Stoddard BL (2005). Computational thermostabilization of an enzyme. Science, 308(5723): 857-860.
Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, Baker D (2003). Design of a novel globular protein fold with atomic-level accuracy. Science, 302(5649): 1364-1368.
Lassner MW, Bedbrook J (2001). Directed molecular evolution in plant improvement. Curr. Opin. Plant Biol. 4: 152-156.
Lehmann M, Pasamontes L, Lassen SF, Wyss M (2000). The consensus concept for thermostability engineering of proteins. Biochim. Biophys. Acta. 1543(2): 408-415.
Lin H, Cornish VW (2002). Screening and selection methods for large-scale analysis of protein function. Angew Chem. 41: 4402-4425.
Lutz S, Ostermeier M, Benkovic SJ (2001). Rapid generation of incremental truncation libraries for protein engineering using alpha-phosphothioate nucleotides. Nucleic Acids Res. 29(4): e16.
Matsuura T, Yomo T (2006). *In vitro* evolution of proteins. J. Biosci. Bioeng. 101(6): 449-456.
Miyazaki K (2003). Creating random mutagenesis libraries by megaprimer PCR of whole plasmid (MEGAWHOP). Methods Mol. Biol. 231: 23-28.
Murakami H, Hohsaka T, Sisido M (2003). Random insertion and deletion mutagenesis. Methods Mol. Biol. 231: 53-64.
Neylon C (2004). Chemical and biochemical strategies for the randomization of protein encoding DNA sequence: library construction for directed evolution. Nucleic Acids Res. 32(4): 1448-1459.
Ostermeie M, Shim JH, Benkovic SJ (1999). A combinatorial approach to hybrid enzymes independent of DNA homology. Nat. Biotechnol. 17: 1205-1209.
Park HS, Nam SH, Lee JK, Yoon CN, Mannervik B, Benkovic SJ, et al., (Provide Complete Name) (2006). Design and evolution of new catalytic activity with an existing protein scaffold. Science, 311(5760): 535-538.
Pritcharda L, Corneb D (2005). (Provide Tittle). J. Theor. Biol. 234: 497-509.
Sandgren M, Gualfetti PJ, Shaw A, Gross LG, Saldajeno M, Day AG, et al., (Provide Complete Name) (2003). Comparison of family 12 glycoside hydrolases and recruited substitutions important for thermal stability. Protein Sci. 12: 848-860.
Saraf MC, Horswill AR, Benkoic SJ, Maranas CD (2004). FamClash: a method for ranking the activity of engineered enzymes. Proc. Natl. Acad. Sci. USA, 101(12): 4142-4147.
Shanklin J (2000). Exploring the possibilities presented by protein engineering. Curr. Opin. Plant Biol. 3(3): 243-248.
Shao Z, Zhao H, Giver L, Arnold FH (1998). Random-priming *in vitro* recombination: an effective tool for directed evolution. Nucleic Acids Res. 26: 681-685.
Sieber V, Martinez CA, Arnold FH (2001). Libraries of hybrid proteins from distantly related sequences. Nat. Biotechnol. 19: 456-460.
Stemmer WPC (1994). DNA shuffling by random fragmentation and reassembly *in vitro* recombination for molecular evolution. Proc. Natl. Acad. Sci. USA, 91: 10747-10751.
Udit AK, Silberg JJ, Sieber V (2003). Sequence homology-independent protein recombination (SHIPREC). Methods Mol. Biol. 231: 153-163.
Wong TS, Roccatano D, Zacharias M, Schwaneberg U (2006). A statistical analysis of current random mutagenesis methods for directed protein evolution. J. Mol. Biol. 355: 858-871
Xiong AS, Yao QH, Peng RH, Li X, Fan HQ, Cheng ZM et al., (Provide Complete Name) (2004). A simple, rapid, high-fidelity and cost-effective PCR-based two-step DNA synthesis method for long gene sequences. Nucleic Acids Res. 32(12): e98.
Yuan L, Kurek I, English J, Keenan R (2005). Laboratory-directed protein evolution. Microbiol. Mol. Biol. Rev. 69(3): 373-392.
Zhao H, Giver L, Shao Z, Affholter JA, Arnold FH (1998). Molecular evolution by staggered extension process (StEP) *in vitro* recombination. Nat. Biotech. 16: 258-261.