

Laboratory-Directed Protein Evolution

Ling Yuan,^{1*} Itzhak Kurek,² James English,²
and Robert Keenan³

Department of Plant and Soil Sciences, and Kentucky Tobacco Research and Development Center, University of Kentucky, Cooper and University Drives, Lexington, Kentucky 40546¹; Verdia Research Campus, Pioneer International, A Dupont Company, 700A Bay Road, Redwood City, California 94063²; and Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637³

INTRODUCTION	374
STRATEGIES FOR DIRECTED EVOLUTION IN PROTEIN DESIGN	374
DNA Shuffling	374
Whole-Genome Shuffling	376
Heteroduplex	376
Random Chimeragenesis on Transient Templates	376
Assembly of Designed Oligonucleotides	376
Mutagenic and Unidirectional Reassembly	377
Exon Shuffling	377
Y-Ligation-Based Block Shuffling	377
Nonhomologous Recombination	377
Combining Rational Design with Directed Evolution	378
APPLICATIONS OF DIRECTED EVOLUTION	378
Directed Evolution of Nucleic-Acid-Modifying Enzymes	378
Polymerases	379
Nucleases	379
Transposase	379
Integrase/recombinase	379
Reporter genes	379
Directed Evolution of Biochemical Catalysts	380
Proteolytic enzymes	380
Cellulolytic enzymes	380
Enzymes for bioremediation	381
Lipases and esterases	381
Cytochrome P450 enzymes	381
Directed Evolution of Metabolic Pathways	382
Directed Evolution of Pharmaceuticals	383
Protein pharmaceuticals	383
Antibodies	383
Vaccines	384
Viruses	384
Therapeutic chemicals	384
Directed Evolution of Agriculturally Important Traits	384
Existing traits	385
(i) Glyphosate tolerance	385
(ii) <i>B. thuringiensis</i> toxin	385
(iii) Golden rice	385
Next-generation traits	385
(i) Chitinase for antifungal properties	385
(ii) Mycotoxin detoxification	385
(iii) Viral vectors	386
CLOSING REMARKS	386
ACKNOWLEDGMENTS	386
REFERENCES	386

* Corresponding author. Mailing address: Department of Plant and Soil Sciences, and Kentucky Tobacco Research and Development Center, University of Kentucky, Cooper and University Drives, Lexington, KY 40546. Phone: (859) 257-4806. Fax: (859) 323-1077. E-mail: lyuan3@uky.edu.

INTRODUCTION

The concept of laboratory-directed protein evolution is not new. Systematic approaches to directed evolution of proteins have been documented since the 1970s (39, 106, 110). One early example is the evolution of the EbgA protein from *Escherichia coli*, an enzyme having almost no β -galactosidase activity. Through intensive selection of a LacZ⁻ deletion strain of *E. coli* for growth on lactose as a sole carbon source, the wild-type EbgA was “evolved” as a β -galactosidase sufficient to replace the *lacZ* gene function (39). Perhaps surprisingly, the evolution of new functions of an enzyme can require few mutations, as was the case for the EbgA protein. EbgA enzyme variants with newly acquired hydrolytic activities toward a variety of β -galactoside sugars contain only one to three mutations (102, 104, 107). The ability to recruit new protein functions was noted by Roy Jenson to arise from the considerable substrate ambiguity of many proteins (136). The substrate ambiguity of a protein can be interpreted as the evolutionary potential that allows a protein to acquire new specificities through mutation or to regain function via mutations that differ from the original protein sequence. All organisms have evolutionarily exploited this substrate ambiguity. When exploited in a laboratory under controlled mutagenesis and selection, it enables a protein to “evolve” in desired directions.

Directed protein evolution is a general term used to describe various techniques for generation of protein mutants (variants) and selection of desirable functions. Over the last three decades, directed protein evolution has emerged as a powerful technology platform in protein engineering. This technology has been advanced considerably by the availability of molecular biology tools and emerging high-throughput screening technologies. These methodologies have simplified the experimental processes and facilitated the identification of mutants with even small improvements in desired function. Advanced recombinant DNA technologies have allowed the transfer of single structural genes or genes for an entire pathway to a suitable surrogate host for rapid propagation and/or high-level protein production. Furthermore, it is now possible to control the rate of mutagenesis in widely applied methods such as error-prone PCR and to modify proteins by systematic insertions or deletions. In addition, site-directed, site-saturation mutagenesis and synthetic oligonucleotides can be used to expand the localized amino acid diversity. While functional complementation of mutant strains is still an excellent choice when possible, the development of sensitive instrumentation and the ability to miniaturize many chemical or biological assays allow the screening of large numbers of samples for selection of desired functions. The ability to rapidly obtain DNA sequence information for gene variants not only provides insight into protein sequence-function relationship but also enhances our ability to select the strategy best suited for the evolution of a particular protein. Thus, directed protein evolution has been expanded from the original *in vivo* approach (e.g., the evolution of EbgA) to include *in vitro* exploration.

One of the most effective strategies in directed protein evolution is to gradually accumulate mutations, either sequentially or by recombination, while applying selective pressure. This is typically achieved by the generation of libraries of mutants followed by efficient screening of these libraries for targeted

functions and subsequent repetition of the process using improved mutants from the previous screening. Many formats of directed protein evolution have been, and continue to be, developed (8, 9).

Here, we review the more recent progress in directed protein evolution (referred as directed evolution hereafter) in a wide range of scientific disciplines and its impacts in chemical, pharmaceutical, and agricultural sciences. Although many strategies for directed evolution are described, we focus on the directed evolution of proteins through gradual accumulation of beneficial mutations, and examples of recombination-based approaches are used primarily to illustrate the power of this technology. The advances in screening technologies for identification of useful functions will not be discussed here, as they have been reviewed elsewhere (8, 184, 207, 273).

STRATEGIES FOR DIRECTED EVOLUTION IN PROTEIN DESIGN

One of the primary goals of protein design is to generate proteins with new or improved properties. In addition to deepening our understanding of the design processes used in nature, the ability to confer a desired activity on a protein or enzyme has considerable practical application in the chemical, agricultural, and pharmaceutical industries. Two strategies are currently being employed towards this goal. The first is directed evolution, in which libraries of variants are searched experimentally for clones possessing the desired properties. The second is rational design, in which proteins are modified based on an understanding of the structural and mechanistic consequences of a particular change or set of changes. While the power of directed evolution is now widely appreciated, our present knowledge of structure-function relationships in proteins is still insufficient to make rational design a robust approach. In this section, we review a few methods and strategies of DNA mutagenesis and recombination for directed evolution, and we discuss ways in which rational design is now being used to facilitate the development of proteins with new and improved properties. Table 1 summarizes some of the methods that have been successfully utilized for directed evolution of a variety of proteins. This is not a complete list, as techniques and strategies of DNA mutagenesis and recombination for directed evolution are constantly arising (54, 148, 149, 150, 224, 236, 344, 352; reviewed by Farinas et al. [84] and by Lutz and Patrick [196]).

DNA Shuffling

The goal of directed evolution is to accumulate improvements in activity through iterations of mutation and screening. The extent to which it succeeds depends critically on the delicate interplay between the quality of biological diversity present in the library, the size of the library, and the ability of an assay to meaningfully detect improvements in the desired activity. The strength of directed evolution lies in the ability of its scoring function (i.e., assay) to mimic the property being evolved, while its weakness lies in the relatively small number of sequences that can be experimentally measured (on the order of 10^3 to 10^6 for high-throughput screening to $>10^{12}$ for display methods [251]).

TABLE 1. Selected methodologies for directed evolution

Method for mutagenesis and recombination	Target gene	Selection	Reference(s)
Single-gene DNA shuffling	β -Lactamase	Cefotaxime degradation	286, 287
Gene family shuffling	Cephalosporinase	Moxalactam degradation	64
Staggered extension process	Subtilisin E	Protease thermostability	351
In vitro-in vivo DNA recombination	GFP	Restore fluorescence	313
Error-prone PCR	Geranylgeranyl diphosphate synthase	Lycopene production	321
Combinatorial libraries enhanced by recombination in yeast	Cytochrome P450	Aromatic hydrocarbon oxidation	1
Compartmentalized self-replication	DNA polymerase	Thermostability, heparin resistance	89
Random chimeragenesis on transient templates	Monoxygenase	Reaction rate, altered substrate specificity	56
Incremental truncation for the creation of hybrid enzymes	N terminus of <i>E. coli</i> glycinamide ribonucleotide formyltransferase (PurN) and C terminus of human glycinamide ribonucleotide formyltransferase (hGART)	PurN-hGART functional hybrid enzymes	195, 225
Sequence homology-independent protein recombination	Cytochrome P450	Deethylation of 7-ethoxyresorufin	276
Sequence-independent site-directed chimeragenesis	β -Lactamases TEM-1 and PSE-4	Functional hybrid enzymes	119
Nonhomologous random recombination	Chorismate mutase and fumarase	Chorismate mutase activity	23
Synthetic shuffling	Subtilisin	Protease thermostability at pH 10	217
Mutagenic and unidirectional reassembly	Phospholipase A ₁	Altered substrate specificity	281
Y-ligation-based block shuffling	GFP	Fluorescence	155
Assembly of designed oligonucleotide	Lipases	Enantioselectivity	343

Library diversity is created through mutagenesis or recombination. Traditionally, libraries have been generated by random point mutagenesis (using, for example, error-prone PCR) or by site-directed mutagenesis of a starting sequence. These libraries are screened (or selected), and the best variant is selected for additional mutagenesis. Because the frequency of beneficial mutations is generally low relative to that of deleterious mutations, only single beneficial mutations are added in each cycle of mutagenesis and screening. Indeed, the probability of improvement decreases rapidly when multiple mutations are made. Thus, iterative, point-mutation-based approaches are generally limited to improvements made in small steps.

DNA shuffling overcomes this limitation by allowing the direct recombination of beneficial mutations from multiple genes. In DNA shuffling a population of DNA sequences are randomly fragmented and then reassembled into full-length, chimeric sequences by PCR (286, 287). In so-called "single-gene" formats, mutations are introduced during the reassembly process by controlling the error rate of DNA polymerase. After screening or application of selective pressure, progeny sequences encoding desirable functions are identified. These clones are then shuffled (bred) iteratively, creating offspring that contain multiple beneficial mutations. Because of this poolwise recombination of beneficial mutations, DNA shuffling gives rise to dramatic increases in the efficiency with which large phenotypic improvements are obtained.

While such methods are relatively efficient when small steps through sequence space are required, the relationship between library diversity, library size, and assay capability dictates that the evolution of phenotypes requiring larger steps through sequence space employ a more efficient search strategy. A simple and powerful way to do this is to use naturally occurring homologous genes as the source of starting diversity (64). In contrast to single-gene shuffling, in which library members are typically 95 to 99% identical, so-called "family shuffling" allows block exchanges of sequences that are typically >60% identical. In part because the sequence diversity comes from related, parental sequences that have survived natural selection ("functional" sequence diversity), much larger numbers of mutations are tolerated in a given sequence without introducing deleterious effects on the structure or function. The increased sequence diversity of these chimeric libraries thus results in sparse sampling of much greater regions of sequence and function space.

Even greater control over the incorporation of sequence diversity can be achieved through "synthetic shuffling." In this approach, no physical starting genes are required. Instead, a series of degenerate oligonucleotides that incorporate all desired diversity (for example, naturally occurring diversity and diversity identified by structural analysis) are used to assemble a library of full-length genes (217). In contrast with fragmentation-based methods, in synthetic shuffling every amino acid from a set of parents is allowed to recombine independently of

every other amino acid. By breaking the linkages between amino acids normally present in parental genes, synthetic shuffling methods access unique regions of sequence space.

All directed evolution experiments must contend with the constraints described above: principally, the type and quality of diversity present in the library, the library size, and the ability of an assay to accurately identify desired clones from that library. To the extent that a desired phenotype is accessible within these constraints, standard DNA shuffling formats and other formats described below provide a rapid and powerful method to optimize activity. For more demanding phenotypes, such as de novo enzyme design, novel substrate specificity, novel enzyme chemistry, etc., there is a need to maximize the information content of a library so that larger steps through vast regions of sequence and function space may be efficiently explored.

Whole-Genome Shuffling

Whole-cell biocatalysts are widely used for industrial applications such as conversion of feedstock to high-value products, production of high-value natural products, and production of protein pharmaceuticals. Fermentation-based bioprocesses are often limited by the sensitivity of microorganisms to temperature, pH, and solvent, resulting in low yield and productivity. Microorganisms represent a delicate and complex system that infrequently can be modified for industrial production by a single gene alteration. Therefore, the ability to evolve an organism at the whole-genome level is highly desirable. A process known as whole-genome shuffling has been developed in an effort to accomplish this objective (347). This approach combines the advantages of family DNA shuffling with the benefits of crossing entire genomes that occur in conventional breeding (347). Traditional breeding is a long, continuous process of genetic recombination of the parental genomes accompanied by phenotypic selections. It is usually limited to two parental genomes per generation and is affected by the genetic compatibility of the parents. On the other hand, manipulation of commercial microorganisms can also be achieved by an asexual process of repeated cycles of random mutagenesis and screening, often referred to as classical strain improvement (CSI) (3). In contrast, the driving force for the accelerated evolution is the recombination of multiple parents in a recursive manner. The advantage of whole-genome shuffling over CSI has been recently demonstrated with *Streptomyces fradiae*, a commonly used strain for commercial production of the complex polyketide antibiotic tylosin (347), and with the industrial strain of *Lactobacillus* for acid tolerance (234). Using a low-production parental strain, two rounds of genome shuffling based on protoplast fusion of mixed populations and screening for tylosin production resulted in mutant strains with productivities similar to that of the commercial strain SF2 (347). However, while it took 20 years and about 1,000,000 assays for the 20 rounds of CSI required to obtain SF2, similar results were produced with 24,000 assays in 1 year of whole-genome shuffling. Patnaik et al. (234) demonstrated the use of whole-genome shuffling for improved acid tolerance in production of lactic acid by lactobacilli. *Lactobacillus* strains with improved low-pH tolerance were first obtained by CSI in order to generate the initial biodiversity pool and then shuffled for

five rounds by protoplast fusion. The improved strains produce threefold more lactic acid than the wild-type strains at pH 4.0.

Whole-genome shuffling is powerful and beneficial in manipulation of organisms (52, 67). It allows the evolution of desired phenotypes by rapid genomic manipulation and stabilization. Directed whole-genome evolution is not limited to microorganisms. By a variety of means, genomes from eukaryotic cells, including regenerable cells from animals and plants, can be recombined recursively for accelerated phenotypic improvement.

Heteroduplex

Recombination in vitro of large genes, such as operons or artificial chromosomes, is difficult to achieve. In an alternative method, in vitro-in vivo DNA recombination takes place with a parental plasmid heteroduplex in an in vivo repair process and has been suggested to be useful for recombination of large genes or entire operons (313). A heteroduplex formed in vitro is used to transform bacterial cells, where repair of regions of nonidentity in the heteroduplex creates a library of new, recombined sequences composed of elements from each parent. However, this method, which is based on the ability of host cells to repair mismatched heteroduplexes, requires high parental gene homology and is limited to two parental genes per event.

Random Chimeragenesis on Transient Templates

Annealing of small fragments as primers, spiking of oligonucleotides as linkers between regions of low homology, and generation of complete synthetic chimera are some of the ways that have been designed to increase frequencies of recombination between low-homology sequences. For example, libraries generated by the random chimeragenesis on transient templates (RACHITT) method showed an average of 14 crossovers per parental gene, a much higher rate than with other reported methods (56). In addition, the RACHITT-derived chimeric genes showed high-frequency recombination at a short region (a few nucleotides). 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, this approach generates high-frequency crossovers. One of the common issues in "family DNA shuffling" is the bias against the incorporation of the less homologous genes in the parental gene pool. By selecting one gene as the sole template, RACHITT is able to force the incorporation of a particular gene even when it shares relatively low homology. In some cases, especially when the background activity of one parent is problematic for library screening, RACHITT allows the selection of this parent as a fragmented donor, thus avoiding the presence of the wild-type gene of this parent in the library.

Assembly of Designed Oligonucleotides

Assembly of designed oligonucleotides (ADO) has been described as a useful technique for gene recombination (343). ADO relies on sequence information on the nonconserved

regions to design a set of synthetic degenerate oligonucleotides. The flanking region of each synthetic fragment contains sequences of conserved regions that can be used as linkers in homologous recombination. PCR assembly of the fragments is then performed in two steps. First, double-stranded DNA is formed by PCR of the single-stranded oligonucleotides in the absence of primers. The double-stranded DNA is then used for PCR amplification of the whole gene, and the full-length gene products are ligated into an expression vector. The two major advantages of the method are that it allows crossing over to occur for low-homology fragments and that self-hybridization of parental genes is minimized or eliminated. High-quality libraries without a parental gene background are essential, especially when high-throughput screening is not available. The limitation due to relatively short lengths of the synthetic oligonucleotides could be overcome by fragment ligation. ADO has been successfully applied to improve the activities of two *Bacillus subtilis* lipases, LipA and LipB (343). One library of 3,000 variants obtained by ADO was sufficient to identify six variants with improved enantioselectivity. The major advantage of this method is the ability to create a large diversity of active variants and to eliminate codon bias in parental genes.

Mutagenic and Unidirectional Reassembly

Song et al. (281) developed mutagenic and unidirectional reassembly (MURA) for the generation of libraries of DNA-shuffled and randomly truncated proteins. In this method, DNA sequences that have been generated by DNA shuffling or by incremental truncation can be simultaneously introduced into a parental gene in a single experiment. The MURA process consists of four steps. First, a random fragmentation of the parental gene is obtained by PCR amplification or restriction digestion. The fragments then are reassembled in the presence of the unidirectional primers that contain a specific restriction site. The DNA fragments are gel purified, subjected to T4 DNA polymerase or S1 nuclease treatments in order to polish both termini, and then digested by a primer-specific restriction enzyme. The MURA method has been used for generation of an N-terminally truncated and DNA-shuffled library of *Serratia* sp. phospholipase A₁ (PlaA) in order to alter the substrate specificity of PlaA from a phospholipase to a lipase (281). The authors isolated nine variants that exhibit lipase and phospholipase activities by high-throughput screening of 2,500 to 3,000 transformants. All these variants showed high lipase activity while retaining their phospholipase activities. All the mutant enzymes possess N-terminal deletions of 61 to 71 amino acids as a result of the MURA process and a relatively small number of amino acid substitutions. The dual activities exhibited by the truncated enzymes suggest that the N-terminal region is critical for the phospholipid substrate interactions.

Exon Shuffling

Exon shuffling is an evolutionary mechanism in which recombination of nonhomologous genes generates new genes known as mosaic proteins. The natural exon shuffling process has been described for a number of gene families by domain organization and splice frame analysis of the hemostatic proteases and by structural and sequence analysis of SCAN do-

main-containing genes (78). As a result, a new method to evolve proteins by in vitro exon shuffling has been suggested (157). Similar to the natural exon shuffling process, in vitro exon shuffling can be carried out using a mixture of chimeric oligonucleotides that allows the control of which exon or combination of exons is to be spliced. One application of exon shuffling is to develop protein pharmaceuticals based on natural human gene sequences, thus potentially reducing the possibility of immune responses (260). For example, it may be possible to minimize the immunogenicity of therapeutic proteins by constructing high-quality human gene libraries that lack random mutations. To complement the construction of such high-quality libraries, protocols such as that described by Zhao and Arnold (350) can be applied. Inclusion of Mn²⁺ or Mg²⁺ and high-fidelity DNA polymerase during amplification and reassembly can significantly reduce the point mutation rate. Exon-shuffled libraries of unrelated domains that share no sequence or functional homology can potentially generate new "humanized" genes with valuable functions.

Y-Ligation-Based Block Shuffling

While many methods improve functions through creating and recombining point mutations, Y-ligation-based block shuffling (YLBS) is a general methodology that mimics evolution processes such as domain shuffling, exon shuffling, and module shuffling, and it can be used for generating high-diversity libraries (155, 156). YLBS is based on repeated cycles of ligation of sequence blocks with a stem and two branches (Y-ligation) formed by two types of single-stranded DNA. The ability to integrate desired blocks with variable sizes (from several amino acids to a whole domain) into proteins at any site and any frequency will dramatically increase the diversity pool for directed evolution. YLBS can be an efficient technology to introduce or to eliminate (by deletion block or null block) peptides, exons, and domains.

Nonhomologous Recombination

While protein variants generated by homologous recombination or random point mutations are more likely to maintain structural similarity to the parental proteins, nonhomologous recombination allows the efficient creation of new protein folds. This approach enables the generation of protein structural diversity that may or may not exist in nature, and it is potentially very useful in evolution of multifunctional proteins. Several methods for nonhomologous recombination have been described. They include incremental truncation for the creation of hybrid enzyme (ITCHY) (225), sequence-independent site-directed chimeragenesis (119), sequence homology-independent protein recombination (276), and nonhomologous random recombination (NRR) (23). ITCHY libraries are created by cloning two genes (or gene fragments) in tandem in an expression vector containing two unique restriction sites. The linearized vector allows the generation of truncated fragments either by time-dependent exonuclease III enzyme digestion (224) or by the incorporation of α -phosphorothioate deoxynucleoside triphosphates (194). Subsequent blunt-ending and treatment with the second restriction enzyme release truncated fragments in various lengths, and chimeras can then be

generated by ligation to recy­clize the vector. This approach has been combined with an additional recombination step to develop SCRATCHY (193). More recently, the NRR method has been described (23). NRR is based on DNase I fragmentation, blunt-end ligation/extension, and capping using two asymmetrical DNA hairpins to stop the extension. This method potentially provides higher flexibility in modulating fragment size and crossover frequency, as well as in the number of parental genes. The major challenge facing all techniques for sequence-independent recombination of proteins is the presence of large numbers of nonfunctional progeny in the libraries (due to nonsense mutations caused by, for example, frameshifting and/or reversed DNA fragment orientation), thus hindering the search for functional mutants. Therefore, it is critical that a high-throughput screening is in place for the selection; otherwise, a preselection strategy, e.g., downstream fusion of a reporter or selection marker to reduce mutants with internal stop codons, can be applied to generate high-quality libraries.

Combining Rational Design with Directed Evolution

One of the most seductive features of rational/computational approaches to protein design is the ability to access vastly larger regions of sequence space ($>10^{25}$) than can be searched experimentally. The success of such approaches depends on the ability to successfully predict the fitness of a given sequence. For certain properties, such as protein stability, simple “packing” algorithms are capable of predicting sequences with reasonable accuracy. For more complex phenotypes, the successful application of purely rational/computational methods requires sophisticated scoring (energy) functions. The recent *de novo* design of a novel protein fold is a spectacular example of the increasing power of computational design (163).

A powerful application of rational design is using it to focus library diversity for directed evolution experiments. In general, computational analysis of a protein’s structure is first used to generate sequence diversity and to test those sequences for functional properties that can be modeled (scored) *in silico*. Only those variants that pass this prescreen are then synthesized and tested experimentally. In this manner, costly and time-consuming experimental searches are limited to regions of sequence space that are consistent with a protein’s structure.

In an elegant example of structure-based computational design, Dwyer et al. introduce triosephosphate isomerase activity into a catalytically inert protein scaffold, ribose-binding protein (79). The design strategy consists of three stages. First, a chemical and geometric definition of the catalytic machinery was generated. Second, a combinatorial search was performed to identify positions within the active site where the catalytic machinery and substrate could be placed, while simultaneously satisfying the above constraints. Third, the remainder of the active site was optimized to form a stereochemically complementary binding surface. A total of 14 designs were tested, and one of these exhibited a k_{cat}/K_m ratio of 1.5×10^2 for the conversion of dihydroxyacetone phosphate to glyceraldehyde-3-phosphate. This is about 3 orders of magnitude less than the ratio for wild-type triosephosphate isomerase but is nevertheless a rate enhancement of more than 10^5 over that of the

uncatalyzed reaction. Subsequently, the authors use directed evolution to improve the k_{cat}/K_m ratio of the designed enzyme. As is often the case, many of the accumulated changes identified by directed evolution lie in regions distal from the active site, and their effect on activity is therefore difficult to rationalize. A key issue for future design strategies lies in understanding how such mutations, which often contribute cooperatively and over long distances, improve activity (284).

One of the great advantages that emerges from the synthesis of rational design and directed evolution is that once a gene with even low levels of starting activity is obtained through design, it may be rapidly optimized by directed evolution (275). Thus, the goal of rational design becomes detecting even a weak starting activity from a focused library, rather than obtaining an optimized level of activity. The complementary use of rational design with directed evolution is a promising path towards the production of proteins with new and improved properties.

APPLICATIONS OF DIRECTED EVOLUTION

Directed evolution is increasingly used in academic and industrial laboratories to improve protein stability and enhance the activity or overall performance of enzymes and organisms or to alter enzyme substrate specificity and to design new activities. Together with novel techniques for large-scale screening, directed evolution enables the selection of redesigned molecules without the necessity for detailed structural and mechanistic information (reviewed by Arnold [7] and Minshull and Stemmer [209]). In the past years we have seen broad applications of directed evolution in research and product developments of recombinant DNA technologies, biocatalysts, metabolic pathway engineering, pharmaceuticals, and important agricultural traits. Regardless of the research discipline, some common themes or parameters can be observed in the application of directed evolution. For example, directed evolution increasingly appears to be the tool of choice for studying the evolution of and relationship between protein structure and function (2, 114, 138, 192, 226, 259) and for interpretation of the evolutionary significance of biomolecular systems (122, 323). It is also a popular tool for accelerated adaptation of protein functions (e.g., stability, specificity, or affinity) in extreme conditions such as unusual temperatures and organic solvents (198, 204, 221, 222, 327–330), as well as for improvement of recombinant protein biosynthesis (152, 185). Directed evolution has also given rise to altered specificities and activities of enzymes (113–115, 126, 141, 294, 337), enhanced intramolecular interactions (292), modified protein-protein interaction (180), and altered metabolic pathways (263). In the following sections we present some examples of the applications of these technologies.

Directed Evolution of Nucleic-Acid-Modifying Enzymes

An emerging area in biotechnology is the directed evolution of DNA-modifying enzymes. Improving or modifying the site selectivity of restriction endonucleases, recombinases, and other DNA-modifying enzymes (46, 57, 82) can lead to novel applications in genetic engineering, functional genomics, and gene therapy.

Polymerases. Molecular biology technologies such as DNA labeling, PCR, sequencing, site-directed mutagenesis, and some cloning often require DNA polymerases with high activity under suboptimal conditions, such as extreme temperatures and/or in the presence of inhibitors. Compartmentalized self-replication (CSR) is a useful strategy for directed evolution of DNA polymerases or RNA polymerases (89). CSR is based on a feedback loop consisting of a polymerase that replicates only its own encoding gene. Self-replications of polymerase variants generated by error-prone PCR are performed in separated compartments formed by water-in-oil emulsions. Genes encoding improved polymerase under the selection conditions used replicate at higher rates and eventually dominate the mutant population. CSR has been used for evolution of *Taq* polymerase in the presence of increasing amounts of the inhibitor heparin, resulting in the isolation of a variant that exhibits a 130-fold increase in heparin resistance (89).

Directed evolution has been successfully applied to DNA polymerase for enhanced activity (233) and conversion to an efficient RNA polymerase (232, 333). The 2'-*O*-methyl-RNA is more stable and has been produced by chemical synthesis. Chelliserrykattil and Ellington established an efficient screening system for selection of highly active polymerases (47). This system creates a so-called "autogene" by cloning the T7 RNA polymerase under the control of its own promoter. In this system the polymerase variants with higher activity will generate more mRNA and can thus be selectively amplified by a reverse transcription-PCR process. The autogene system has allowed the identification of T7 RNA polymerase variants that can efficiently incorporate various 2'-modified nucleotides with good processivities (47, 48). Mixtures of the polymerase mutants with different specificities have produced transcripts with multiple modified nucleotides. DNA polymerase that is capable of incorporating 2'-*O*-methyl nucleotides has also been created by directed evolution (82).

Nucleases. Nucleases, including restriction endonucleases, are essential enzymes in modern molecular biology and thus are active targets for directed evolution. An intelligently designed selection by compartmentalization of each gene variant in a rabbit reticulocyte transcription/translation system overcomes limitations associated with *in vivo* screening techniques, allowing the efficient screening of restriction endonuclease libraries (74). Novel selection methods have also been developed for selection of restriction enzymes with altered substrate specificities (80, 168, 256, 353). DNA cleavage specificities have been created from the *E. coli* RNase P derivatives (59).

Transposase. Naumann and Reznikoff (216) used directed evolution to generate a mutated Tn5 bacterial transposase that could function on transposons with mutated end binding sequences. The Tn5 transposon encodes a 53-kDa transposase protein (Tnp) that facilitates the movement of the entire transposon by first binding to each of the two 19-bp specific binding sequences (known as outside end [OE]), followed by formation of a nucleoprotein complex, blunt-end cleavage, and then transfer to the target DNA. The transposon also promotes the movement of a single OE by using an additional 19-bp inside end sequence (IE). The wild-type Tn5 Tnp activity is inhibited in *E. coli* as a result of Dam methylation at the IE (IE^{ME}). In order to screen for a transposase mutant that functions with mutated inverted repeats, the IE was modified at position 12

from thymine to adenine (IE12A), which results in loss of recognition by the wild-type transposase. As a consequence, insertion of IE12A in the flanking region of the *lacZ* gene between the transcription and translation start sites results in an inactive transposon. Three rounds of gene shuffling and high-throughput screening for LacZ activity at about 10⁴ colonies per round, followed by analysis of the active variants for activities against OE and IE, has allowed the isolation of a specific hyperactive Tnp variant (TnpsC7). While methylation of IE reduced the wild-type Tnp activity by 100-fold, TnpsC7 activity in the presence of IE^{ME} was markedly higher.

Integrase/recombinase. Improved site specificity for large genome modifications has been recently demonstrated for the wild-type ϕ C31 integrase (265). Scimienti et al. (265) applied two rounds of DNA shuffling in combination with a genetic screen that is capable of identifying improved variants expressing the *lacZ* reporter gene. This improved enzyme possesses strong preference for target-site DNA sequences and has 10- to 20-fold-higher absolute integration frequencies than the wild-type ϕ C31 integrase. In addition to the demonstration of improved site specificity of this integrase, several other groups have successfully altered the site specificity of the Cre/Flp recombinases by directed evolution (35, 36, 252, 258, 314). The Cre recombinase catalyzes the integration, excision, and rearrangement of two 34-bp, double-stranded recombination sites known as *loxP*. Santoro and Schultz (258) designed a fluorescence-activated cell sorting-based screening for recombinases that recognize unnatural recombination sites. The screening system consists of a recombinase variant and a reporter gene plasmid, expressing either enhanced yellow fluorescent protein (YFP) or green fluorescent protein (GFP). Using this high-throughput selection system, the authors isolated recombinase variants that show high specificity for unnatural *loxP* sites and low activity for the wild-type *loxP* site. Site-specific manipulation of genomes by recombinases is a powerful functional genomic tool. Recombinases such as Cre have been widely used to mutagenize and replace genes in mice. Expanding the recombination sequences of recombinases will improve the efficiency and the quality of production of transgenic animals and plants. The ability to evolve proteins that interact with DNA has broad implications. Efforts to evolve other DNA-binding proteins, such as transcription factors, for tailor-made specificities are under way.

Reporter genes. Although by themselves they usually do not modify nucleic acids, in molecular biology, reporter proteins are often closely associated with other proteins that do. Directed evolution has been applied to optimize the physical properties of fluorescent proteins and small-molecule probes for real-time imaging of live cells (21, 40, 142). Fluorescent probes function as "passive" markers that provide high sensitivity for real-time visualization and tracking of cellular events without perturbing the cells. GFP is widely used for tracking protein localization *in vivo* and has been evolved by directed evolution (65). Additional fluorescent variants such as YFP and cyan fluorescent protein have been generated by mutagenesis of the wild-type GFP. These fluorescent variants may be used as companion markers for protein colocalization and for tracking protein-protein interactions by fluorescent resonance energy transfer (FRET). Nguyen and Daugherty (220) addressed the dynamic range and sensitivity limitations associ-

ated with FRET by designing a strategy in which a cyan fluorescent protein-YFP fusion system is used to allow the detection of subtle improvements, enabling gradual optimization of FRET signals. When this system is coupled with random mutagenesis and targeted saturation mutagenesis, substantial enhancement of FRET dynamic range and sensitivity has been achieved. Another example is the engineering of the *Discosoma* red fluorescent protein (DsRed). The wild-type, tetrameric DsRed has poor solubility that can affect the function and localization of the tagged proteins. DsRed is also slow in the chromophore maturation process. By applying seven rounds of site-directed mutagenesis and error-prone PCR followed by high-throughput visual screening for fluorescence in microbial cells, Bevis and Glick (21) isolated soluble DsRed variants that also mature 10 to 15 times faster than the wild-type protein. While the improved DsRed isolated by Bevis and Glick retained its tetrameric state, Campbell et al. (40) evolved DsRed to an active monomeric form that matures 10 times faster than the wild-type protein. Their approach was a step-wise evolution of DsRed first to a dimer and then to a monomer. This sequential improvement of DsRed resulted in an active monomeric protein with improved solubility and shorter maturation time, leading to greater tissue penetration and spectral separation from autofluorescence and other fluorescent probes. The next generation of the monomeric fluorescent proteins have been shown to be more photostable, mature more completely, and be more tolerant to forming fusion proteins (274). The improvement of another well-known reporter protein, beta-glucuronidase, was achieved (200, 202). Further evolution successfully converted this enzyme into a beta-galactosidase (202). Beta-galactosidase activity has also been evolved from a fucosidase (72, 345).

Increasing protein solubility by directed evolution is not limited to reporter proteins. Overexpressed proteins in heterologous systems such as *E. coli* often fail to fold into their native states and are thus accumulated as insoluble inclusion bodies. An efficient method to generate more soluble forms of insoluble proteins is directed evolution. One way to screen for soluble variants is to fuse the variants of an insoluble protein to a reporter for heterologous expression, followed by screening of the reporter protein activity (reviewed by Waldo [317]). Yang et al. (336) utilized a GFP-based screening to evolve the solubility of the *Mycobacterium tuberculosis* Rv2002 gene product. While overexpression of Rv2002 in *E. coli* resulted in inclusion bodies, five soluble mutants were identified after three rounds of error-prone PCR and DNA shuffling. Because the Rv2002 mutants are fused with GFP, the soluble Rv2002-GFP emits brighter fluorescence than the wild-type protein. Enzymatic assays indicated that a soluble mutant Rv2002-M3 protein possesses high catalytic activity as an NADH-dependent $3\alpha,20\beta$ -hydroxysteroid dehydrogenase.

Directed Evolution of Biochemical Catalysts

Since the 1980s, recombinant DNA technologies, and recombinant protein expression technology in particular, have revolutionized the chemical industry. Enzymatic catalysts are superior in many industrial processes because of their high selectivity and minimum energy requirement. However, for the potential of industrial enzymes to be fully exploited, many

challenges remain. In order to be effective and practical, these enzymes need to be consistently available in high quantities and at low cost, and they need to be active and stable under process conditions. In some cases, product inhibitions pose problems. In addition, many enzymes required for specific reactions have yet to be identified and produced. Directed evolution offers viable solutions for enzyme optimization and development of novel specificities. This area of research has been the subject of a number of recent review articles (11, 27–29, 51, 90, 98, 123, 126, 161, 162, 230, 241, 242, 279, 296, 302, 318).

Proteolytic enzymes. The serine endoprotease subtilisin is a commercially important enzyme. With annual sales over \$500 million, the highest among industrial enzymes, subtilisins are widely applied as additives in laundry detergents and other uses. A major challenge in improvement of most industrial enzymes is that the performance is defined not by any single property but by a complex mix of parameters. Although rational design and random mutagenesis have been used to improve single properties such as the thermostability of activity in organic solvents, it is often at the expense of other critical properties. Ness et al. (218) demonstrated multidimensional improvement of subtilisin by DNA shuffling. Twenty-five subtilisin gene fragments obtained from different *Bacillus* isolates were bred together with the full-length gene for a leading commercial protease and screened for thermostability, solvent stability, and pH dependence (at pH 5, pH 7.5, and pH 10). High frequencies of improvements (4 to 12%) in all parameters were achieved using a relatively small library (654 active clones). In addition, the diversity of combinations of properties ranged well beyond that of the properties of the parental enzymes. Sequence analysis of several high performers under each set of conditions revealed that variants with similar properties could be encoded by different sequences. Thermostability, for example, could be conferred by any one of the at least three different genetic elements. Because of the importance of proteolytic enzymes, directed evolution of proteases and peptidases remains one of the most actively pursued research areas (10, 12, 34, 100, 160, 210, 211, 285, 297, 304, 327–329, 349).

Cellulolytic enzymes. Enzymes that hydrolyze carbohydrates are also active targets for directed evolution. Up to sevenfold enhancement of the thermostability of the endoglucanase EngB has been achieved by introducing sequence diversities from a partially homologous endoglucanase, EngD (213, 214). A library was constructed using genes encoding the cellulosomal endoglucanase EngB and noncellulosomal cellulase EngD from *Clostridium cellulovorans*. The more thermostable cellulosomal endoglucanases are of high industrial relevance. Cellulosomes from clostridia are efficient at hydrolyzing microcrystalline cellulose. The relatively high efficiency has been attributed to (i) the correct ratio between catalytic domains, which optimizes synergism between them; (ii) appropriate spacing between the individual components to further promote synergism; and (iii) the presence of different enzymatic activities (cellulolytic or hemicellulolytic) in the cellulosome, which can remove other polysaccharides in heterogeneous cell wall materials.

Applications of cell wall-loosening enzymes can be found in a variety of industrial processes. In the pulp and paper industry, enzymatic degradation of the hemicellulose-lignin com-

plexes present in pulps preserves intact cellulose fibers and strongly reduces the amount of bleaching chemicals required. The enzyme laccase is of interest for biobleaching and has been improved in industrially relevant parameters by directed evolution (38). Other applications in which cellulosic hydrolases are used include improvement of dough quality in the baking industry, increasing the feed conversion efficiency of animal feed, clarifying juices, and producing xylose, xylobiose, and xylo-oligomers. In addition, cellulosic hydrolases are important in biomass conversion for novel biofuel and other valuable chemicals. In a broader aspect, directed evolution has been successfully applied to improve many enzymes involved in carbohydrate biosynthesis, modification, and degradation. Examples include ADP-glucose pyrophosphorylase (254), amylosucrase (310), aldolase (86, 326), sugar kinase (120), cellulase (153), amylases (19, 20, 154, 312), xylanases (49, 129, 203), glucose dehydrogenase (14), and beta-glucosidase (13).

Enzymes for bioremediation. Enzymes that cleave carbon-halogen bonds are being studied not only because of the important chemical reactions they catalyze but also for potential use in environmental sciences. Haloalkane dehalogenase converts alkylhalide functionality to an alcohol group with broad substrate specificity. This enzyme has been subjected to directed evolution for improved function in detoxification of halogenated compounds (30, 38, 95, 96, 240, 348). Organophosphate-degrading enzymes have been evolved and selected for broadened substrate specificity (53, 335). Broadened substrate specificity of a biphenyl dioxygenase has also been achieved (33, 87, 164, 291). Efforts in cleaning underground water contamination prompted the evolution of an enzyme for chlorinated ethene degradation (41).

Lipases and esterases. Lipases, which comprise another class of hydrolases, have broad industrial applications. Lipases catalyze the hydrolysis and synthesis of long-chain acylglycerols from triglycerides. For production of biofuel, a single transesterification reaction using lipases in organic solvents can convert vegetable oil to methyl- or other short-chain alcohol esters. Biodegradable biopolymers such as polyphenols, polysaccharides, and polyesters show a considerable degree of diversity and complexity. Lipases and esterases are used as catalysts for polymeric synthesis (e.g., stereoselectivity, regioselectivity, and chemoselectivity) under mild reaction conditions. Lipases are also used in synthesis of fine chemicals, agrochemicals, and pharmaceuticals.

Directed evolution of industrially important lipases has been extensively reviewed (131–134, 247–249). The enantioselectivity of lipases is of biochemical interest. The ability to engineer lipases with high enantioselectivities allows the production of desired enantiopure compounds. A *Pseudomonas aeruginosa* lipase has been evolved to increase enantioselectivity towards the chiral substrate 2-methyldecanoic acid *p*-nitrophenyl ester. A few rounds of directed evolution produced greater than 25-fold improvement of the enantioselectivity. It is interesting that the best variants contain five amino acid changes and most of them are located in the flexible loop regions (183, 249). Using the ADO approach, increased enantioselectivities of two *B. subtilis* lipases have been identified by screening of a small number of variants (343). The substrate specificity and stability of lipases can also be modified by directed evolution (147, 282). The lipase from *Bacillus thermocatenuatus* BTL2 exhibits low phos-

pholipase activity. A single round of random mutagenesis followed by screening of 6,000 variants generated progeny with more than a 10-fold increase in phospholipase activities (147). Most of the variants show reduced activities towards medium- and long-chain fatty acyl methyl esters compared to the wild-type enzyme. Moreover, in combination with structure-guided site-directed mutagenesis, further improvement of the phospholipase activity has been achieved. The best variant, which exhibits 17-fold improvement in phospholipase selectivity, has 1.5- to 4-fold-higher activity towards long-chain fatty acyl substrates. In an effort to achieve the opposite goal, the phospholipase A of *Serratia* has been converted to a lipase by using a combination of DNA shuffling and N-terminal truncations (281).

By sequential generation of random mutagenesis and screening, Moore and Arnold (212) evolved an esterase for deprotection of an antibiotic *p*-nitrobenzyl ester in aqueous organic solvents. A variant has been found to perform as well in 30% dimethylformamide as the wild-type enzyme in water, a 16-fold improvement in esterase activity. As in many other directed evolution experiments, the successful outcome of this work relied on the establishment of a high-throughput screening assay, this time using the *p*-nitrophenyl ester. In recent years, a great deal of effort has been devoted to design of screening tools for improvement of lipases and esterases (91, 97). Droge et al. (77) reported the binding of a phosphonate suicide inhibitor to lipase A that is presented by phage display. The specific interaction with the suicide inhibitor provides a fast and reproducible method for selection lipases with novel substrate specificities. Two new triglyceride analogue biotinylated suicide inhibitors have been designed, synthesized, and applied in directed evolution of phage-displayed lipolytic enzymes (70, 71).

Cytochrome P450 enzymes. The cytochrome P450 superfamily is a highly diversified set of heme-containing proteins, and members serve a wide spectrum of functions. In addition to the most common function of catalyzing hydroxylation, P450 proteins perform a variety of reactions, including N oxidation; sulfoxidation; epoxidation; N, S, and O dealkylation; peroxidation; deamination; desulfuration; and dehalogenation. In mammals they are critical for drug metabolism, blood homeostasis, cholesterol biosynthesis, and steroidogenesis. In plants they are involved in plant hormone synthesis, phytoalexin synthesis, flower petal pigment biosynthesis, and most likely hundreds of additional, unknown functions. In fungi they make ergosterol and are involved in pathogenesis by detoxification of host plant defenses. Bacterial P450s are key players in antibiotic synthesis. More recently, cytochrome P450 enzymes have shown promise in industrial applications as new methods for high-level production and high-throughput assays have been developed (4, 18, 306).

A number of cytochrome P450 enzymes have been the targets of directed evolution (50, 54, 83, 250, 255, 306, 307, 331, 332). Cytochrome P450 enzymes are often found to be poorly active, with narrow substrate specificity. The wild-type P450 BM-3, which is specific for long-chain fatty acids, was a target for rational design and directed evolution (181). Based on the crystal structure, eight amino acids were identified for creation of libraries by site-specific randomization mutagenesis of each residue. The libraries were screened by a spectroscopic assay

using omega-*p*-nitrophenoxycarboxylic acids as substrates. By sequential evolution, variants showing specificity towards medium-chain substrates were identified. In a subsequent study (182), one of the variants was found to be able to efficiently hydroxylate indole, resulting in the formation of indigo and indirubin. Further characterization of this mutant revealed that it is capable of hydroxylating several alkanes and alicyclic, aromatic, and heterocyclic compounds, all of which are non-natural substrates for the wild-type enzyme (6). Many cytochrome P450 monooxygenases are multimeric and membrane associated, with low catalytic efficiencies. Glieder et al. (92) evolved the *Bacillus megaterium* cytochrome P450 BM-3, which is specific for C₁₂ to C₁₈ fatty acids, to efficiently catalyze the conversion of C3 to C8 alkanes to alcohols. In this case the evolved enzyme exhibits a broad range of substrate specificities, including the gaseous alkane propane, as well as improved activity towards the natural fatty acid substrates. BM-3 has also been engineered to be significantly more tolerant to several cosolvents, including the organic cosolvents dimethyl sulfoxide and tetrahydrofuran (332). Furthermore, the regioselectivity and enantioselectivity of BM-3 have been engineered through a combination of in vitro evolution, and the selectivity appears to be retained in vivo with *E. coli* cells (238).

Successful evolution of cytochrome P450 requires efficient high-throughput screens that are sensitive to the activities of interest. Horseradish peroxidase couples the phenolic products of hydroxylation of aromatic substrates to generate colored or fluorescent compounds that are easily detectable in high-throughput formats. Joo et al. (139) have taken advantage of this system by coexpressing the coupling enzymes with functional mono- and dioxygenases. Using fluorescent digital imaging, they screened libraries of cytochrome P450cam from *Pseudomonas putida* for novel activity of chlorobenzene hydroxylation. Joo et al. (140) also utilized this so-called "peroxide shunt" pathway to identify variants showing significantly improved activity for naphthalene hydroxylation in the absence of the NADPH cofactor. Interestingly, the P450 enzyme has recently been used as a model for computational structure-guided evolution (227).

Directed Evolution of Metabolic Pathways

The evolution of whole metabolic pathways is a particularly attractive concept, because most natural and novel compounds are produced by pathways rather than by single enzymes. Genetically up-regulating one enzyme activity in a pathway does not always guarantee an increase in the final product. Therefore, metabolic pathway engineering usually requires the coordinated manipulation of all enzymes in the pathway. The potential for evolving a pathway in the laboratory has long been recognized. For instance, using the *ebg* operon of *E. coli* as a model, it has been demonstrated that a pathway can be redirected and that such evolution requires a series of mutations in several structural and regulatory genes (103, 109, 111). However, instead of operons, genes of a pathway are often located in different positions in the genome, making such coordinated engineering difficult. Several strategies can be applied to the directed evolution of metabolic pathways, as follows.

(i) Whole genomes are shuffled (see above) and selected for

desired phenotypes or products (239). The successful engineering of polyketide and lactic acid production in *Lactobacillus* (234, 347) has demonstrated that whole-genome shuffling is one of the most powerful tools in directed evolution of pathways. It is particularly useful when a pathway is not well characterized and key enzymes or genes have not yet been identified or cloned. Phenotypic improvement by whole-genome shuffling is an important milestone for bioprocess optimization. Together with novel techniques for cultivating and identifying previously unrecognized microorganisms (342) and information on biodiversity in terms of species, distribution, and ecosystem function (reviewed by Bull et al. [37]), whole-genome shuffling will continue to expand its impact to the production of high-value biomolecules.

(ii) The genes encoding key enzymes are heterologously expressed to alter an existing pathway. Introduction of an enzyme with novel specificity can redirect the metabolic flux in a host and result in production of new products (261, 321). These recombinant enzymes can be obtained from other organisms known to produce the compounds (299) or by directed evolution to create the desired specificity from an enzyme that normally catalyzes other reactions (144, 315). For instance, under anaerobic conditions yeast does not efficiently produce ethanol by using xylose. By heterologous expression of a xylose isomerase from the fungus *Piromyces* and selection of yeast transformants on xylose, Kuyper et al. (166) have isolated a mutant strain that exhibits a sixfold increase in the anaerobic growth rate on xylose and higher yields of ethanol. Pathway engineering often requires alteration of the substrate pools for the key steps. Thus, directly targeting enzymes responsible for the production of these substrates can enhance or even redirect biosynthetic pathways (177). To engineer a multienzyme pathway for novel carotenoid production in *E. coli*, Schmidt-Dannert and colleagues first introduced two genes to produce the precursor phytoene. Subsequently, a library of two shuffled desaturase genes from *Erwinia* was introduced for the desaturation of phytoene. Divergent lycopene-like compounds with different degrees and positions of desaturation were identified. The pathway of a chosen mutant was further modified by introducing a library of shuffled cyclase genes. The engineering of the carotenoid pathway represents a fine example of how directed evolution can be used to redesign a complex pathway (68, 147, 167, 175, 176, 178, 205, 206, 257, 262, 263, 305, 320, 324).

(iii) In nature, many pathway genes are organized in gene clusters or operons (171, 172). Well-known examples include pathways for polyketide biosynthesis (125) and biosynthesis of certain secondary metabolites (190). Early work using the *ebg* operon presented convincing arguments for directed evolution of an operon as an effective approach in pathway engineering (103, 105, 108, 109, 111). Directed evolution of naturally existing operons and, in some cases, artificially assembled operons offers a unique and coordinated approach to engineer novel functions. Another demonstration of this approach is the manipulation of an arsenate detoxification pathway by DNA shuffling (63). A plasmid containing the operon of four *ars* genes was shuffled and selected for increased resistance to arsenic. While the native operon does not confer *E. coli* resistance to arsenic, several rounds of selection resulted in cell growth in media where the arsenate concentration reached the

solubility limit. In another example, the trehalose-6-phosphate synthase/phosphatase operon was evolved to achieve greater trehalose production in *E. coli* (159, 160). In *E. coli*, trehalose-6-phosphate synthase and trehalose-6-phosphate phosphatase are encoded by the *otsBA* operon. Directed evolution of the *otsBA* operon and screening for trehalose synthesis resulted in 15 positive clones and 12-fold improvement in trehalose production compared to that with the wild-type strain. The same strategy can be applied to artificial operons similar to that constructed for the production of the biopolymer poly(3-hydroxybutyrate-co-3-hydroxyhexanoate) (231). In another example, a metabolically engineered *E. coli* strain for astaxanthin production has been generated by overexpression of three metabolic enzymes from different origins: the *E. coli* isopentenyl diphosphate isomerase, the *Archaeoglobus fulgidus* geranylgeranyl diphosphate synthase (GPS), and the *Agrobacterium aurantiacum* astaxanthin biosynthesis enzymes (*crtWZYIB* gene products) (322). In a subsequent effort, repeated cycles of error-prone PCR, which employs a low-fidelity replication step to introduce random point mutations at each round of amplification, were used to evolve one of these key enzymes, GPS (321). A 100% improvement in lycopene production has been detected by screening for deeper orange color in 3,500 colonies. It is tempting to speculate that the application of directed evolution to the synthetic operon that contains isopentenyl diphosphate isomerase, GPS, and *crtWZYIB* might result in larger amounts of astaxanthin than the levels observed by single-gene evolution.

(iv) The characteristics of a metabolic pathway are a result of the dynamic interaction between its structural genes and the gene regulatory apparatus. Therefore, directed pathway evolution can be achieved by engineering of gene regulation factors that control these pathways (61). The recent exciting progress in engineering of artificial transcription factors has shown that this approach is not only feasible but also advantageous in certain areas of metabolic engineering. Notable advances have mainly been in the generation of artificial zinc finger transcription factors (17, 25, 75, 76, 127, 128, 135, 146, 174, 186, 187, 215, 266–271, 300). Chimeric proteins containing novel DNA-binding domains (such as polydactyl zinc fingers) have shown promise in high-throughput ligand-binding screens, genome-wide gene activation/repression, targeted DNA cleavage, DNA/chromatin modification, and site-specific integration (135). This strategy is particularly powerful when dealing with pathways that are undefined or normally inactive without induction. Engineered transcription factors can also be used to target a known gene regulatory region(s). For example, they can be evolved to bind specific promoter sequences proximal to the binding sites of known and natural transcription factors (94). Transcription factors and their target genes comprise the basic unit in the complex transcriptional regulatory network. Network-wide engineering must deal with higher levels of complexity. The ability to evolve the transcriptional network, however, represents a new possibility in pathway engineering. Yokobayashi et al. proposed the construction of an artificial transcriptional control network and provided examples of how such a genetic circuit can be optimized by a combination of rational design and directed evolution (338, 339). Metabolic pathways often respond to cell-cell communications. An elegantly designed “population control” system was con-

structed based on a quorum-sensing system, allowing a synthetic bacterial ecosystem to be controlled by cell-cell communication (340). Directed evolution of the major component of this system, the LuxR-type transcriptional regulators, revealed the evolutionary plasticity of the quorum-sensing mechanism (60). Another challenge in pathway engineering is to control the timing of gene expression. Inducible gene regulation systems such as the tetracycline/Tet receptor can be used to switch pathways on and off. Evolving these systems to recognize novel inducers has tremendous practical implications in pathway engineering (264, 280).

Directed Evolution of Pharmaceuticals

Protein pharmaceuticals. Directed evolution has revolutionized the development of novel therapeutic proteins (5, 93, 118, 145, 157, 165, 173, 235, 253). DNA family shuffling of more than 20 human alpha interferon genes, followed by selection of antiviral and antiproliferation activities in murine cells, resulted in greater than 250,000-fold improvement (44). Interestingly, no random mutation occurred in the highly improved proteins; i.e., the novel chimeras were created from the genetic diversity within the parental gene family, a result with intriguing implications for gene evolution. Homologous recombination approaches have also been successfully applied to improvement of the human p53 protein, a tumor suppressor (201, 334). The human prolyl endopeptidase is important in activation of the melphalan prodrug, but the wild-type enzyme is thermolabile. Robotic-assisted directed evolution has significantly improved the thermostability of the enzyme (117). By combining receptor structure-based engineering and directed evolution, an amphioxus insulin-like peptide was converted to mammalian insulin (99). Another exciting area to explore functional diversities is the evolution of hormones and hormone receptors (55, 69, 293). Directed evolution has led to the increase of peroxidase activity of horse heart myoglobin (319). Therapeutic proteases and protease inhibitors are also active targets for directed evolution (191, 288–290). The macromolecular protease inhibitor ecotin is of therapeutic value. By combining directed evolution and stepwise engineering, Stoop and Craik (288) generated ecotin libraries that contain variants with significantly enhanced selectivity towards plasma kallikrein.

Antibodies. Therapeutic antibodies represent the fastest growing area in pharmaceutical development. Considering that in nature the combinatorial antibody diversity is a result of somatic recombination, it is not surprising that directed evolution can be a powerful and practical tool for the creation of high-affinity antibodies in vitro. Techniques such as surface display facilitate high-throughput selection for desired activity (32, 62, 85, 124, 143, 295, 308). Recombination of phage-displayed, low-affinity immunoglobulin M antibodies resulted in variants with increased affinity of several orders of magnitude in just two rounds of evolution (85). The same strategy has yielded stable disulfide bond-free antibody single-chain fragments (244). The requirement for disulfide bond formation has hindered antibody production in systems such as *E. coli*, and disulfide bond-free antibodies not only potentially simplify production but also provide insight into antibody protein folding. Additional research has aimed at engineering antibodies

to achieve extremely high affinities (15, 26, 66, 112, 137, 246). The gene for the llama heavy chain antibody fragment was evolved and selected for improvement in production (309). Antibody variants were identified that exhibited two- to four-fold increases in production while retaining their antigen specificity (341). Crystallographic analysis of one of the evolved antibodies revealed that the mutations conferring significant improvement in affinity do not directly contact the antigen, suggesting that it would be difficult to obtain such results via rational design. Nonetheless, the strategy of combining rational design and directed evolution should accelerate antibody engineering more rapidly than using either approach alone.

Catalytic antibodies are also of interest for directed evolution (298, 301). Superior catalysts for aryl phosphate were generated from synthetic human antibody libraries (43). Antibodies have also been engineered for diagnostic purposes (161).

Vaccines. Directed evolution has played and continues to play an important role in the development of new vaccines (58, 188, 189, 197, 235, 245, 325). To boost immunity, directed evolution can be used to generate improved proteous antigens or other immunomodulatory molecules, DNA vaccines, and whole viruses (see below). On the other hand, certain cytokines and allergens can be bred for down-regulation of allergic immune responses. Recursive library construction and selection allowed the isolation of high-affinity, protective mimotopes against *Cryptococcus neoformans* (16). Highly immunogenic mimotopes of the hepatitis C virus hypervariable regions have been selected by a combination of DNA shuffling and phage display-based screening (346). A DNA vaccine of the E7 oncogene has been developed and shown to provide protection against tumor cells (223). This strategy of rearranging oncogene sequences presents an advantage over wild-type oncogene-derived DNA vaccines, which carry a risk of de novo tumor induction. Toxic side effects have been associated with the direct administration of recombinant antitumor interleukin-12 protein. A DNA vaccine based on the interleukin-12 gene has been shown to reduce adverse side effects, while its potency and effectiveness have been further improved by directed evolution (179). In addition, high-affinity T-cell receptor variants can be generated and used for detecting peptide-major histocompatibility complex complexes on antigen-presenting cells (121).

Viruses. Breeding of viruses has tremendous practical implications in gene therapy and vaccine development (283). The feasibility was demonstrated using the murine leukemia viruses (MLV). Family shuffling of six MLV produced variants with novel tropism (283). The MLV envelope protein consists of two subunits, SU and TM, associated by a labile disulfide bond. This complex, which interacts with a cellular receptor and mediates fusion with the plasma membrane, is highly sensitive to physical forces during the manufacturing process. As a result, the concentration procedure commonly used for retrovirus vectors is ineffective for manufacturing stocks of high titer. To improve the resistance of the MLV envelope protein to the process of concentration by ultracentrifugation, the envelope regions of six ecotropic strains were shuffled (243). Screening for survival after three consecutive concentration steps resulted in 30- to 100-fold-improved stability compared to the parental viruses. In an effort to establish a pig-tailed macaque

model for human immunodeficiency virus (HIV) infection, Pekrun et al. evolved a HIV type 1 variant with a substantially enhanced replication rate (237). In an interesting attempt to control the risks associated with pathogenic phenotypes of high-replicating viral vaccines, a tetracycline-inducible system was introduced to control the HIV gene replication (199). By application of directed evolution, highly infectious viral variants have been isolated; however, the viral replication is strictly controlled by a doxycycline-dependent switching system. An alternate strategy to control viral replication by using the bacteriophage T7 polymerase has also been developed (31).

Therapeutic chemicals. The role of biocatalysis in pharmaceutical production has been rapidly expanding since the establishment of recombinant DNA technology (45, 123). The involvement of enzyme and metabolic pathway engineering in therapeutic chemical production is moving towards the mainstream in the industry, and directed evolution technologies are leading the advance. Applications of directed evolution in development of anti-infection agents were among the early examples demonstrating the power and effectiveness of the technologies. Evolution of polyketide synthases to generate novel antibiotic activities demonstrated that novel compounds can be identified even in small libraries (123). The modular nature of the polyketide synthetic pathway allows an efficient way to create large numbers of polyketide variants by replacing individual modules with a shuffled library (151). Directed evolution of a toluene-xylene monooxygenase resulted in variants that catalyze the synthesis of various valuable fine chemicals, such as catechol (311). The substrate specificity of the cephalosporin acylase has been altered for the improvement of cephalosporin and penicillin production (229, 278). Directed evolution has allowed the identification of "hot spots," in this case, a single amino acid residue crucial for substrate specificity. When this hot spot was subjected to saturation mutagenesis, variants with further improvement or novel specificity were identified (228). Protein engineering using site-directed and/or saturation mutagenesis, guided by information generated from directed evolution, can be an extremely powerful approach to create novel functionalities (73, 88, 208, 316).

Directed Evolution of Agriculturally Important Traits

Agricultural biotechnology offers tremendous promise. Possibilities exist for improvement of crop yields through resistance to pests, including weeds, insects, and disease, as well as tolerance to environmental stresses such as cold and drought. Other areas which may affect eventual yield include postharvest characteristics such as ripening control and prevention of potato sweetening.

In the 20 years since it has been possible to introduce transgenes into plants, many novel strategies have been devised to improve the quality of crops. Many strategies for pest control, cold tolerance, disease control, and other areas of improvement have had positive initial results in laboratory settings; however, the genes have not provided sufficient efficacy to produce commercially viable genetically modified (GM) products. In retrospect this makes sense, since many transgenes that were used in these experiments clearly had not been optimized for use in GM crop plants.

Directed evolution can be used to improve existing traits

such as glyphosate resistance and *Bacillus thuringiensis* toxin expression in commercial crops. It can also be used to develop traits from programs in which initial leads (genes) provided insufficient efficacy. Furthermore, directed evolution can be applied to develop desirable gene functions from gene targets that have low or no activity, resulting in novel traits that would otherwise not have been possible (169).

Existing traits. (i) Glyphosate tolerance. Existing glyphosate resistance traits in corn, cotton, and soybean, based on expression of a microbial enopyruvylshikimate-3-phosphate synthase that is not affected by the herbicide, are effective. However, there is clearly room for improvement. He et al. (116) bred *E. coli* and *Salmonella enterica* serovar Typhimurium enopyruvylshikimate-3-phosphate synthases (the enzyme which, when carrying a specific mutation, confers tolerance to the herbicide) to develop variants with superior properties. Several gene variants from a single round of directed evolution resulted in enzymes simultaneously improved over the best parent in multiple kinetic parameters, including a twofold-improved specific activity, a fivefold-improved K_m for phosphoenolpyruvate, and a fivefold decrease in sensitivity to glyphosate. Interestingly, the mutations identified in that study do not coincide with the mutations identified previously by other researchers in their efforts to improve the properties of this enzyme. These results demonstrate that directed evolution can provide novel solutions to improving protein function even for proteins that have undergone extensive improvement through random mutagenesis and/or structure-based protein design.

Recently, Castle et al. reported the development of an alternative method for producing glyphosate-tolerant crop plants (42). First, the researchers searched for an enzyme that would detoxify glyphosate. After growing several hundred strains of common microbes, they determined that the most effective was the soil microbe *Bacillus licheniformis*. The researchers identified three related genes encoding an enzyme, glyphosate *N*-acetyltransferase (GAT), from a microbial diversity collection consisting of predominantly *Bacillus* species. The starting genes, identified from *B. licheniformis*, encoded GAT enzymes which acetylated glyphosate, albeit very poorly. After 11 iterations of DNA shuffling, the enzyme activity was improved nearly 10,000-fold. To test its potential, corn plants were transformed with improved GAT gene variants. The transgenic plants tolerated six times the concentration of glyphosate that farmers normally apply, with no apparent effect on health or reproduction.

(ii) *B. thuringiensis* toxin. Plants expressing *B. thuringiensis* toxin genes are the second most widely grown transgenic crops. This trait has been widely used by corn and cotton farmers. Currently there are two limitations of *B. thuringiensis* Cry proteins that can be addressed by directed evolution. First, the spectrum of insects controlled by any given *B. thuringiensis* Cry protein is relatively narrow. *B. thuringiensis* Cry proteins with broadened specificity have the potential to further reduce the use of synthetic pesticides in commercial agriculture. Second, it is relatively difficult to express *B. thuringiensis* Cry proteins in transgenic plants at sufficiently high levels to control many insect pests. *B. thuringiensis* Cry proteins exhibiting increased specific activity against current insect targets could reduce the effort required to generate a commercially useful level of insect

resistance. Directed evolution has been successfully used to address both of these issues (170).

(iii) Golden rice. Golden rice is a rice variety developed to express elevated levels of β -carotene (a precursor of vitamin A) in the grain (22). Vitamin A deficiency afflicts more than 100 million people in at least 26 developing countries, including highly populated areas of Asia, Africa, and Latin America. Every year 1 to 2 million people die because of infectious and other diseases as a consequence of weakened immune systems caused by this deficiency. In addition, hundreds of thousands go permanently blind due to vitamin A deficiency. Many of the victims are children. Rice, therefore, is an important target for enhanced nutritional qualities, as it is a staple in the diets of a majority of the world's population. Golden rice has been touted as a breakthrough GM product which could supplement vitamin A deficiencies in the diets of millions of people around the world. Currently developed golden rice varieties represent a good start toward this goal. However, it is unlikely that the amount of vitamin A precursor produced in current golden rice varieties is enough to have a significant impact (219).

This is undoubtedly an application in which directed evolution could be of great benefit. The metabolic pathway engineered into golden rice requires the coordinated expression of multiple transgenes (see "Directed evolution of metabolic pathways" above). By evolving these genes toward higher overall activity and better synergistic behavior, there is the potential to significantly boost the amount of β -carotene produced in next-generation golden rice varieties.

Next-generation traits. As mentioned above, traits that have already found their way to the marketplace have room for improvement that may be effectively addressed by directed evolution. Over the past 20 years, there have been numerous traits for which promising results were seen in laboratories but which did not translate into commercially viable products. There have also been concepts which showed initial promise but which did not show efficacy even in a laboratory setting because the starting genes did not function in the required plant cell environment. Directed evolution may open the door for opportunities for turning these concepts into reality.

(i) Chitinase for antifungal properties. The antifungal properties of plant-expressed chitinases have been known for more than 10 years (130). Still, there are no commercial crop plant products based on expression of these enzymes. This is a prime example of promising results being seen in the laboratory which did not translate to a commercial product. Through the application of directed evolution to dramatically increase the activity of antifungal chitinases expressed in transgenic crop plants, there is the potential of controlling fungal diseases.

(ii) Mycotoxin detoxification. Mycotoxin production is a toxic side effect of fungal infection of crop plants. *Fusarium moniliforme* infection of maize can result in contamination by mycotoxins, the most prominent of which is fumonisin. A transgenic approach to reducing fumonisin contamination was based on amine oxidase enzymes isolated from black yeasts found on *Fusarium*-infected ears by cultures of *Exophiala spinifera* (24). However, the starting enzymes had no activity in the extracellular space where they were required to work. Five rounds of DNA shuffling and screening were performed using surrogate hosts, including a plant screening system. Significant improvements were generated in enzyme activity at the low pH

environment of the apoplast as well as in efficiency of protein secretion. Functional assays showed significant improvement of in planta fumonisin detoxification (J. English and J. Duvick, unpublished observations).

(iii) **Viral vectors.** Viral vectors offer the possibility of very high-level expression of valuable compounds in a relatively short time frame. However, current tobacco mosaic virus-based vectors are in need of improvements in order for this to be a viable process. Scientists at the Scottish Crop Research Institute in collaboration with Large Scale Biology Corporation used random mutagenesis coupled with recombination to improve the performance of their vectors in planta. The mutagenized tobacco mosaic virus variants were subjected to gene shuffling and screened for faster movement around the plant as well as higher transgene expression. Variants that moved significantly faster throughout the plant were recovered (303).

CLOSING REMARKS

Using keywords in this review to search major scientific databases can result in hundreds or even more than a thousand hits. It is difficult, if not impossible, to cover all literature on laboratory-directed protein evolution. Directed evolution is a system that compares and utilizes the mounting genetic information generated in this era of genomics. It is also a mechanism to expand the genetic diversity in our search for novel functions. Its power as a postgenomics technology platform is being increasingly recognized.

For a newcomer to the field of laboratory-directed evolution, the two volumes of *Methods in Molecular Biology* (8, 9) edited by Arnold and Georgiou are a good place to start. The 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. A quick assessment on the possibility of handling large numbers of variants is typically the first consideration for the feasibility of starting a directed evolution experiment. However, the technology is rapidly moving in more sophisticated directions. Efforts are being made to model and validate the minimum sampling numbers; i.e., what is the minimum number of screenings necessary in order to obtain measurable improvements? In some practices, assaying pooled samples instead of single samples is one of the effective ways to drastically reduce the number of experiments. Directed evolution is a process where progressive partial change built upon previous partial changes. It is possible, or even preferable, to accelerate the improvement by performing multiple rounds of evolution in which mutants with small but measurable degrees of enhancement are identified by a limited number of assays and then using these mutants as parents for the next round of evolution. In most cases, greater improvements can be achieved by successive rounds of evolution than by screening a larger number of mutants in one experiment. Furthermore, taking advantage of the tremendous computing power in the genomics era, directed evolution is also being carried out in silico (101). Computer-assisted analysis can significantly reduce the demand for labor- and cost-intensive wet-lab experiments (81, 277). Finally, the ever-increasing information on protein structure-function relationships and gene sequence-function relation-

ships not only can provide insights into the impacts of mutations but also can refine the targets for directed evolution.

We have just begun to see the impacts of directed evolution on biological sciences and biotechnology. Future reviews of this subject will no doubt describe further levels of complication and sophistication in the application of directed evolution technologies.

ACKNOWLEDGMENTS

We thank K. Shen and M. Lassner for insightful comments.

This work is supported in part by a grant (to L.Y.) from the Kentucky Tobacco Research and Development Center, University of Kentucky.

REFERENCES

1. Abecassis, V., D. Pompon, and G. Truan. 2000. High efficiency family shuffling based on multi-step PCR and in vivo DNA recombination in yeast: statistical and functional analysis of a combinatorial library between human cytochrome P450 1A1 and 1A2. *Nucleic Acids Res.* **28**:E88.
2. Agrawal, V., and R. K. Kishan. 2001. Functional evolution of two subtly different (similar) folds. *BMC Struct. Biol.* **1**:5.
3. Aharonowitz, Y., and G. Cohen. 1981. The microbiological production of pharmaceuticals. *Sci. Am.* **245**:140–152.
4. Alcalde, M., E. T. Farinas, and F. H. Arnold. 2004. Colorimetric high-throughput assay for alkene epoxidation catalyzed by cytochrome P450 BM-3 variant 139-3. *J. Biomol. Screen.* **9**:141–146.
5. Anthony-Cahill, S. J., and T. J. Magliery. 2002. Expanding the natural repertoire of protein structure and function. *Curr. Pharm. Biotechnol.* **3**:299–315.
6. Appel, D., S. Lutz-Wahl, P. Fischer, U. Schwaneberg, and R. D. Schmid. 2001. A P450 BM-3 mutant hydroxylates alkanes, cycloalkanes, arenes and heteroarenes. *J. Biotechnol.* **88**:167–171.
7. Arnold, F. H. 1998. Design by directed evolution. *Acc. Chem. Res.* **31**:125–131.
8. Arnold, F. H., and G. Georgiou (ed.). 2003. Directed enzyme evolution: screening and selection methods. *Methods in molecular biology*, vol. 230. Humana Press, Totowa, N.J.
9. Arnold, F. H., and G. Georgiou (ed.). 2003. Directed evolution library creation: methods and protocols. *Methods in molecular biology*, vol. 231. Humana Press, Totowa, N.J.
10. Arnold, F. H., L. Giver, A. Gershenson, H. Zhao, and K. Miyazaki. 1999. Directed evolution of mesophilic enzymes into their thermophilic counterparts. *Ann. N. Y. Acad. Sci.* **870**:400–403.
11. Arnold, F. H., and J. C. Moore. 1997. Optimizing industrial enzymes by directed evolution. *Adv. Biochem. Eng. Biotechnol.* **58**:1–14.
12. Arnold, F. H., P. L. Wintrode, K. Miyazaki, and A. Gershenson. 2001. How enzymes adapt: lessons from directed evolution. *Trends Biochem. Sci.* **26**:100–106.
13. Arrizubieta, M. J., and J. Polaina. 2000. Increased thermal resistance and modification of the catalytic properties of a beta-glucosidase by random mutagenesis and in vitro recombination. *J. Biol. Chem.* **275**:28843–28848.
14. Baik, S. H., T. Ide, H. Yoshida, O. Kagami, and S. Harayama. 2003. Significantly enhanced stability of glucose dehydrogenase by directed evolution. *Appl. Microbiol. Biotechnol.* **61**:329–335.
15. Barbas, C. F., III, D. Hu, N. Dunlop, L. Sawyer, D. Cababa, R. M. Hendry, P. L. Nara, and D. R. Burton. 1994. In vitro evolution of a neutralizing human antibody to human immunodeficiency virus type 1 to enhance affinity and broaden strain cross-reactivity. *Proc. Natl. Acad. Sci. USA* **91**:3809–3813.
16. Beenhouwer, D. O., R. J. May, P. Valadon, and M. D. Scharff. 2002. High affinity mimotope of the polysaccharide capsule of *Cryptococcus neoformans* identified from an evolutionary phage peptide library. *J. Immunol.* **169**:6992–6999.
17. Beerli, R. R., D. J. Segal, B. Dreier, and C. F. Barbas III. 1998. Toward controlling gene expression at will: specific regulation of the *erbB-2/HER-2* promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. USA* **95**:14628–14633.
18. Bell, S. G., C. F. Harford-Cross, and L. L. Wong. 2001. Engineering the CYP101 system for in vivo oxidation of unnatural substrates. *Protein Eng.* **14**:797–802.
19. Berk, H., and R. J. Lebbink. 2003. High-throughput screening of mutant alpha-amylase libraries for increased activity at 129 degrees C. *Methods Mol. Biol.* **230**:127–135.
20. Bessler, C., J. Schmitt, K. H. Maurer, and R. D. Schmid. 2003. Directed evolution of a bacterial alpha-amylase: toward enhanced pH-performance and higher specific activity. *Protein Sci.* **12**:2141–2149.
21. Bevis, B. J., and B. S. Glick. 2002. Rapidly maturing variants of the *Drosophila* red fluorescent protein (DsRed). *Nat. Biotechnol.* **20**:83–87.

22. Beyer, P., S. Al-Babili, X. D. Ye, P. Lucca, P. Schaub, R. Welsch, and I. Potrykus. 2002. Golden rice: introducing the beta-carotene biosynthesis pathway into rice endosperm by genetic engineering to defeat vitamin A deficiency. *J. Nutr.* **132**:506s–510s.
23. Bittker, J. A., B. V. Le, J. M. Liu, and D. R. Liu. 2004. Directed evolution of protein enzymes using nonhomologous random recombination. *Proc. Natl. Acad. Sci. USA* **101**:7011–7016.
24. Blackwell, B. A., J. T. Gilliam, M. E. Savard, J. David Miller, and J. P. Duvick. 1999. Oxidative deamination of hydrolyzed fumonisins B(1) (AP(1)) by cultures of *Exophiala spinifera*. *Nat. Toxins* **7**:31–38.
25. Blancafort, P., L. Magnenat, and C. F. Barbas III. 2003. Scanning the human genome with combinatorial transcription factor libraries. *Nat. Biotechnol.* **21**:269–274.
26. Boder, E. T., K. S. Midelfort, and K. D. Wittrup. 2000. Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity. *Proc. Natl. Acad. Sci. USA* **97**:10701–10705.
27. Bornscheuer, U. T. 2000. Directed evolution of enzymes. *Chim. Oggi-Chem. Today* **18**:65–67.
28. Bornscheuer, U. T. 2002. Methods to increase enantioselectivity of lipases and esterases. *Curr. Opin. Biotechnol.* **13**:543–547.
29. Bornscheuer, U. T., and M. Pohl. 2001. Improved biocatalysts by directed evolution and rational protein design. *Curr. Opin. Chem. Biol.* **5**:137–143.
30. Bosma, T., J. Damborsky, G. Stucki, and D. B. Janssen. 2002. Biodegradation of 1,2,3-trichloropropane through directed evolution and heterologous expression of a haloalkane dehalogenase gene. *Appl. Environ. Microbiol.* **68**:3582–3587.
31. Brakmann, S., and S. Grzeszik. 2001. An error-prone T7 RNA polymerase mutant generated by directed evolution. *Chembiochem* **2**:212–219.
32. Brophy, S. E., P. D. Holler, and D. M. Kranz. 2003. A yeast display system for engineering functional peptide-MHC complexes. *J. Immunol. Methods* **272**:235–246.
33. Bruhlmann, F., and W. Chen. 1999. Tuning biphenyl dioxygenase for extended substrate specificity. *Biotechnol. Bioeng.* **63**:544–551.
34. Bryan, P. N. 2000. Protein engineering of subtilisin. *Biochim. Biophys. Acta* **1543**:203–222.
35. Buchholz, F., P. O. Angrand, and A. F. Stewart. 1998. Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nat. Biotechnol.* **16**:657–662.
36. Buchholz, F., and A. F. Stewart. 2001. Alteration of Cre recombinase site specificity by substrate-linked protein evolution. *Nat. Biotechnol.* **19**:1047–1052.
37. Bull, A. T., A. C. Ward, and M. Goodfellow. 2000. Search and discovery strategies for biotechnology: the paradigm shift. *Microbiol. Mol. Biol. Rev.* **64**:573–606.
38. Bulter, T., M. Alcalde, V. Sieber, P. Meinhold, C. Schlachtbauer, and F. H. Arnold. 2003. Functional expression of a fungal laccase in *Saccharomyces cerevisiae* by directed evolution. *Appl. Environ. Microbiol.* **69**:987–995.
39. Campbell, J. H., J. A. Lengyel, and J. Langridge. 1973. Evolution of a second gene for beta-galactosidase in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **70**:1841–1845.
40. Campbell, R. E., O. Tour, A. E. Palmer, P. A. Steinbach, G. S. Baird, D. A. Zacharias, and R. Y. Tsien. 2002. A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. USA* **99**:7877–7882.
41. Canada, K. A., S. Iwashita, H. Shim, and T. K. Wood. 2002. Directed evolution of toluene *ortho*-monooxygenase for enhanced 1-naphthol synthesis and chlorinated ethene degradation. *J. Bacteriol.* **184**:344–349.
42. Castle, L. A., D. L. Siehl, R. Gorton, P. A. Patten, Y. H. Chen, S. Bertain, H. J. Cho, N. Duck, J. Wong, D. Liu, and M. W. Lassner. 2004. Discovery and directed evolution of a glyphosate tolerance gene. *Science* **304**:1151–1154.
43. Cesaro-Tadic, S., D. Lagos, A. Honegger, J. H. Rickard, L. J. Partridge, G. M. Blackburn, and A. Pluckthun. 2003. Turnover-based in vitro selection and evolution of biocatalysts from a fully synthetic antibody library. *Nat. Biotechnol.* **21**:679–685.
44. Chang, C. C., T. T. Chen, B. W. Cox, G. N. Dawes, W. P. Stemmer, J. Punnonen, and P. A. Patten. 1999. Evolution of a cytokine using DNA family shuffling. *Nat. Biotechnol.* **17**:793–797.
45. Chartrain, M., P. M. Salmon, D. K. Robinson, and B. C. Buckland. 2000. Metabolic engineering and directed evolution for the production of pharmaceuticals. *Curr. Opin. Biotechnol.* **11**:209–214.
46. Chelliserrykattil, J., G. Cai, and A. D. Ellington. 2001. A combined in vitro/in vivo selection for polymerases with novel promoter specificities. *BMC Biotechnol.* **1**:13.
47. Chelliserrykattil, J., and A. D. Ellington. 2003. Autogene selections. *Methods Mol. Biol.* **230**:27–43.
48. Chelliserrykattil, J., and A. D. Ellington. 2004. Evolution of a T7 RNA polymerase variant that transcribes 2'-O-methyl RNA. *Nat. Biotechnol.* **22**:1155–1160.
49. Chen, Y. L., T. Y. Tang, and K. J. Cheng. 2001. Directed evolution to produce an alkalophilic variant from a *Neocallimastix patriciarum* xylanase. *Can. J. Microbiol.* **47**:1088–1094.
50. Cherry, J. R. 2000. Directed evolution of microbial oxidative enzymes. *Curr. Opin. Biotechnol.* **11**:250–254.
51. Cherry, J. R., and A. L. Fidantsef. 2003. Directed evolution of industrial enzymes: an update. *Curr. Opin. Biotechnol.* **14**:438–443.
52. Chiang, S. J. 2004. Strain improvement for fermentation and biocatalysis processes by genetic engineering technology. *J. Ind. Microbiol. Biotechnol.* **31**:99–108.
53. Cho, C. M., A. Mulchandani, and W. Chen. 2002. Bacterial cell surface display of organophosphorus hydrolase for selective screening of improved hydrolysis of organophosphate nerve agents. *Appl. Environ. Microbiol.* **68**:2026–2030.
54. Chopra, S., and A. Ranganathan. 2003. Protein evolution by “codon shuffling”: a novel method for generating highly variant mutant libraries by assembly of hexamer DNA duplexes. *Chem. Biol.* **10**:917–926.
55. Coco, W. M., L. P. Encell, W. E. Levinson, M. J. Crist, A. K. Loomis, L. L. Licato, J. J. Arensdorf, N. Sica, P. T. Pienkos, and D. J. Monticello. 2002. Growth factor engineering by degenerate homoduplex gene family recombination. *Nat. Biotechnol.* **20**:1246–1250.
56. Coco, W. M., W. E. Levinson, M. J. Crist, H. J. Hektor, A. Darzins, P. T. Pienkos, C. H. Squires, and D. J. Monticello. 2001. DNA shuffling method for generating highly recombined genes and evolved enzymes. *Nat. Biotechnol.* **19**:354–359.
57. Cohen, H. M., D. S. Tawfik, and A. D. Griffiths. 2004. Altering the sequence specificity of HaeIII methyltransferase by directed evolution using in vitro compartmentalization. *Protein Eng. Des. Sel.* **17**:3–11.
58. Cohen, J. 2001. ‘Breeding’ antigens for new vaccines. *Science* **293**:236–238.
59. Cole, K. B., and R. L. Dorit. 2001. Protein cofactor-dependent acquisition of novel catalytic activity by the RNase P ribonucleoprotein of *E. coli*. *J. Mol. Biol.* **307**:1181–1212.
60. Collins, C. H., F. H. Arnold, and J. R. Leadbetter. 2005. Directed evolution of *Vibrio fischeri* LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones. *Mol. Microbiol.* **55**:712–723.
61. Collins, C. H., Y. Yokobayashi, D. Umeno, and F. H. Arnold. 2003. Engineering proteins that bind, move, make and break DNA. *Curr. Opin. Biotechnol.* **14**:665.
62. Crameri, A., S. Cwirla, and W. P. Stemmer. 1996. Construction and evolution of antibody-phage libraries by DNA shuffling. *Nat. Med.* **2**:100–102.
63. Crameri, A., G. Dawes, E. Rodriguez, Jr., S. Silver, and W. P. Stemmer. 1997. Molecular evolution of an arsenate detoxification pathway by DNA shuffling. *Nat. Biotechnol.* **15**:436–438.
64. Crameri, A., S. A. Raillard, E. Bermudez, and W. P. Stemmer. 1998. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**:288–291.
65. Crameri, A., E. A. Whitehorn, E. Tate, and W. P. Stemmer. 1996. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat. Biotechnol.* **14**:315–319.
66. Cumbers, S. J., G. T. Williams, S. L. Davies, R. L. Grenfell, S. Takeda, F. D. Batista, J. E. Sale, and M. S. Neuberger. 2002. Generation and iterative affinity maturation of antibodies in vitro using hypermutating B-cell lines. *Nat. Biotechnol.* **20**:1129–1134.
67. Dai, M., and S. D. Copley. 2004. Genome shuffling improves degradation of the anthropogenic pesticide pentachlorophenol by *Sphingobium chlorophenolicum* ATCC 39723. *Appl. Environ. Microbiol.* **70**:2391–2397.
68. de Boer, A. L., and C. Schmidt-Dannert. 2003. Recent efforts in engineering microbial cells to produce new chemical compounds. *Curr. Opin. Chem. Biol.* **7**:273–278.
69. Dell’Era, P., R. Ronca, L. Coco, S. Nicoli, M. Metra, and M. Presta. 2003. Fibroblast growth factor receptor-1 is essential for in vitro cardiomyocyte development. *Circ. Res.* **93**:414–420.
70. Deussen, H. J., S. Danielsen, J. Breinholt, and T. V. Borchert. 2000. Design and synthesis of triglyceride analogue biotinylated suicide inhibitors for directed molecular evolution of lipolytic enzymes. *Bioorg. Med. Chem. Lett.* **10**:2027–2031.
71. Deussen, H. J., S. Danielsen, J. Breinholt, and T. V. Borchert. 2000. A novel biotinylated suicide inhibitor for directed molecular evolution of lipolytic enzymes. *Bioorg. Med. Chem.* **8**:507–513.
72. Dion, M., A. Nisole, P. Spangenberg, C. Andre, A. Glottin-Fleury, R. Mattes, C. Tellier, and C. Rabiller. 2001. Modulation of the regioselectivity of a *Bacillus* alpha-galactosidase by directed evolution. *Glycoconjugate J.* **18**:215–223.
73. Dion, M., G. Osanjo, C. Andre, P. Spangenberg, C. Rabiller, and C. Tellier. 2001. Identification by saturation mutagenesis of a single residue involved in the alpha-galactosidase AgaB regioselectivity. *Glycoconj. J.* **18**:457–464.
74. Doi, N., S. Kumadaki, Y. Oishi, N. Matsumura, and H. Yanagawa. 2004. In vitro selection of restriction endonucleases by in vitro compartmentalization. *Nucleic Acids Res.* **32**:e95.
75. Dreier, B., R. R. Beerli, D. J. Segal, J. D. Flippin, and C. F. Barbas III. 2001. Development of zinc finger domains for recognition of the 5'-ANN-3' family of DNA sequences and their use in the construction of artificial transcription factors. *J. Biol. Chem.* **276**:29466–29478.
76. Dreier, B., D. J. Segal, and C. F. Barbas III. 2000. Insights into the molec-

- ular recognition of the 5'-GNN-3' family of DNA sequences by zinc finger domains. *J. Mol. Biol.* **303**:489–502.
77. Droge, M. J., C. J. Rugeberg, A. M. van der Sloot, J. Schimmel, D. S. Dijkstra, R. M. Verhaert, M. T. Reetz, and W. J. Quax. 2003. Binding of phage displayed *Bacillus subtilis* lipase A to a phosphonate suicide inhibitor. *J. Biotechnol.* **101**:19–28.
 78. Dupuy, D., V. G. Duperat, and B. Arveiler. 2002. SCAN domain-containing 2 gene (SCAND2) is a novel nuclear protein derived from the zinc finger family by exon shuffling. *Gene* **289**:1–6.
 79. Dwyer, M. A., L. L. Looger, and H. W. Hellinga. 2004. Computational design of a biologically active enzyme. *Science* **304**:1967–1971.
 80. Eisenschmidt, K., T. Lanio, A. Jeltsch, and A. Pingoud. 2002. A fluorimetric assay for on-line detection of DNA cleavage by restriction endonucleases. *J. Biotechnol.* **96**:185–191.
 81. Endelman, J. B., J. J. Silberg, Z. G. Wang, and F. H. Arnold. 2004. Site-directed protein recombination as a shortest-path problem. *Protein Eng. Des. Sel.* **17**:589–594.
 82. Fa, M., A. Radeghieri, A. A. Henry, and F. E. Romesberg. 2004. Expanding the substrate repertoire of a DNA polymerase by directed evolution. *J. Am. Chem. Soc.* **126**:1748–1754.
 83. Farinas, E. T., M. Alcalde, and F. Arnold. 2004. Alkene epoxidation catalyzed by cytochrome P450BM-3 139–3. *Tetrahedron* **60**:525–528.
 84. Farinas, E. T., T. Bulter, and F. H. Arnold. 2001. Directed enzyme evolution. *Curr. Opin. Biotechnol.* **12**:545–551.
 85. Fermer, C., I. Andersson, K. Nilsson, and O. Nilsson. 2004. Specificity rescue and affinity maturation of a low-affinity IgM antibody against progastrin-releasing peptide using phage display and DNA shuffling. *Tumour Biol.* **25**:7–13.
 86. Fong, S., T. D. Machajewski, C. C. Mak, and C. Wong. 2000. Directed evolution of D-2-keto-3-deoxy-6-phosphogluconate aldolase to new variants for the efficient synthesis of D- and L-sugars. *Chem. Biol.* **7**:873–883.
 87. Furukawa, K. 2000. Engineering dioxygenases for efficient degradation of environmental pollutants. *Curr. Opin. Biotechnol.* **11**:244–249.
 88. Geddie, M. L., and I. Matsumura. 2004. Rapid evolution of beta-glucuronidase specificity by saturation mutagenesis of an active site loop. *J. Biol. Chem.* **279**:26462–26468.
 89. Ghadessy, F. J., J. L. Ong, and P. Holliger. 2001. Directed evolution of polymerase function by compartmentalized self-replication. *Proc. Natl. Acad. Sci. USA* **98**:4552–4557.
 90. Giver, L., and F. H. Arnold. 1998. Combinatorial protein design by in vitro recombination. *Curr. Opin. Chem. Biol.* **2**:335–338.
 91. Giver, L., A. Gershenson, P. O. Freskgard, and F. H. Arnold. 1998. Directed evolution of a thermostable esterase. *Proc. Natl. Acad. Sci. USA* **95**:12809–12813.
 92. Glieder, A., E. T. Farinas, and F. H. Arnold. 2002. Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. *Nat. Biotechnol.* **20**:1135–1139.
 93. Graddis, T. J., R. L. Remmele, Jr., and J. T. McGrew. 2002. Designing proteins that work using recombinant technologies. *Curr. Pharm. Biotechnol.* **3**:285–297.
 94. Graslund, T., X. Li, L. Magnenat, M. Popkov, and C. F. Barbas III. 2005. Exploring strategies for the design of artificial transcription factors: targeting sites proximal to known regulatory regions for the induction of gamma-globin expression and the treatment of sickle cell disease. *J. Biol. Chem.* **280**:3707–3714.
 95. Gray, K. A., T. H. Richardson, K. Kretz, J. M. Short, F. Bartnek, R. Knowles, L. Kan, P. E. Swanson, and D. E. Robertson. 2001. Rapid evolution of reversible denaturation and elevated melting temperature in a microbial haloalkane dehalogenase. *Adv. Synth. Catal.* **343**:607–617.
 96. Gray, K. A., T. H. Richardson, D. E. Robertson, P. E. Swanson, and M. V. Subramanian. 2003. Soil-based gene discovery: a new technology to accelerate and broaden biocatalytic applications. *Adv. Appl. Microbiol.* **52**:1–27.
 97. Griffiths, A. D., and D. S. Tawfik. 2003. Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization. *EMBO J.* **22**:24–35.
 98. Griffiths, A. D., and D. S. Tawfik. 2000. Man-made enzymes—from design to in vitro compartmentalization. *Curr. Opin. Biotechnol.* **11**:338–353.
 99. Guo, Z. Y., L. Shen, W. Gu, A. Z. Wu, J. G. Ma, and Y. M. Feng. 2002. In vitro evolution of amphioxus insulin-like peptide to mammalian insulin. *Biochemistry* **41**:10603–10607.
 100. Gupta, R., Q. K. Beg, and P. Lorenz. 2002. Bacterial alkaline proteases: molecular approaches and industrial applications. *Appl. Microbiol. Biotechnol.* **59**:15–32.
 101. Gustafsson, C., S. Govindarajan, and R. Emig. 2001. Exploration of sequence space for protein engineering. *J. Mol. Recognit.* **14**:308–314.
 102. Hall, B. G. 1981. Changes in the substrate specificities of an enzyme during directed evolution of new functions. *Biochemistry* **20**:4042–4049.
 103. Hall, B. G. 1990. Directed evolution of a bacterial operon. *Bioessays* **12**:551–558.
 104. Hall, B. G. 2003. The EBG system of *E. coli*: origin and evolution of a novel beta-galactosidase for the metabolism of lactose. *Genetica* **118**:143–156.
 105. Hall, B. G. 1982. Evolution of a regulated operon in the laboratory. *Genetics* **101**:335–344.
 106. Hall, B. G. 1973. In vivo complementation between wild-type and mutant beta-galactosidase in *Escherichia coli*. *J. Bacteriol.* **114**:448–450.
 107. Hall, B. G. 2002. Predicting evolution by in vitro evolution requires determining evolutionary pathways. *Antimicrob. Agents Chemother.* **46**:3035–3038.
 108. Hall, B. G. 1978. Regulation of newly evolved enzymes. IV. Directed evolution of Ebg repressor. *Genetics* **90**:673–681.
 109. Hall, B. G. 1989. Selection, adaptation, and bacterial operons. *Genome* **31**:265–271.
 110. Hall, B. G., and D. L. Hartl. 1974. Regulation of newly evolved enzymes. I. Selection of a novel lactase regulated by lactose in *Escherichia coli*. *Genetics* **76**:391–400.
 111. Hall, B. G., and L. Xu. 1992. Nucleotide sequence, function, activation, and evolution of the cryptic asc operon of *Escherichia coli* K12. *Mol. Biol. Evol.* **9**:688–706.
 112. Hanes, J., C. Schaffitzel, A. Knappik, and A. Pluckthun. 2000. Picomolar affinity antibodies from a fully synthetic naive library selected and evolved by ribosome display. *Nat. Biotechnol.* **18**:1287–1292.
 113. Hansson, L. O., R. Bolton-Grob, T. Massoud, and B. Mannervik. 1999. Evolution of differential substrate specificities in Mu class glutathione transferases probed by DNA shuffling. *J. Mol. Biol.* **287**:265–276.
 114. Hansson, L. O., and B. Mannervik. 2000. Use of chimeras generated by DNA shuffling: probing structure-function relationships among glutathione transferases. *Methods Enzymol.* **328**:463–477.
 115. Hansson, L. O., M. Widersten, and B. Mannervik. 1999. An approach to optimizing the active site in a glutathione transferase by evolution in vitro. *Biochem. J.* **344**:93–100.
 116. He, M., Z. Y. Yang, Y. F. Nie, J. Wang, and P. Xu. 2001. A new type of class I bacterial 5-enopyruvylshikimate-3-phosphate synthase mutants with enhanced tolerance to glyphosate. *Biochim. Biophys. Acta* **1568**:1–6.
 117. Heinis, C., P. Alessi, and D. Neri. 2004. Engineering a thermostable human prolyl endopeptidase for antibody-directed enzyme prodrug therapy. *Biochemistry* **43**:6293–6303.
 118. Heyd, B., F. Pecorari, B. Collinet, E. Adjadj, M. Desmadril, and P. Minard. 2003. In vitro evolution of the binding specificity of neocarzinostatin, an edinyne-binding chromoprotein. *Biochemistry* **42**:5674–5683.
 119. Hiraga, K., and F. H. Arnold. 2003. General method for sequence-independent site-directed chimeragenesis. *J. Mol. Biol.* **330**:287–296.
 120. Hoffmeister, D., J. Yang, L. Liu, and J. S. Thorson. 2003. Creation of the first anomeric D/L-sugar kinase by means of directed evolution. *Proc. Natl. Acad. Sci. USA* **100**:13184–13189.
 121. Holler, P. D., P. O. Holman, E. V. Shusta, S. O'Herrin, K. D. Wittrup, and D. M. Kranz. 2000. In vitro evolution of a T cell receptor with high affinity for peptide/MHC. *Proc. Natl. Acad. Sci. USA* **97**:5387–5392.
 122. Hong, K. W., M. Ibba, and D. Soll. 1998. Retracing the evolution of amino acid specificity in glutamyl-tRNA synthetase. *FEBS Lett.* **434**:149–154.
 123. Huisman, G. W., and D. Gray. 2002. Towards novel processes for the fine-chemical and pharmaceutical industries. *Curr. Opin. Biotechnol.* **13**:352–358.
 124. Huls, G., D. Gestel, J. van der Linden, E. Moret, and T. Logtenberg. 2001. Tumor cell killing by in vitro affinity-matured recombinant human monoclonal antibodies. *Cancer Immunol. Immunother.* **50**:163–171.
 125. Hutchinson, C. R., H. Decker, K. Madduri, S. L. Otten, and L. Tang. 1993. Genetic control of polyketide biosynthesis in the genus *Streptomyces*. *Antonie Leeuwenhoek* **64**:165–176.
 126. Iffland, A., S. Gendrezig, P. Tafelmeyer, and K. Johnsson. 2001. Changing the substrate specificity of cytochrome c peroxidase using directed evolution. *Biochem. Biophys. Res. Commun.* **286**:126–132.
 127. Imanishi, M., Y. Hori, M. Nagaoka, and Y. Sugiura. 2001. Design of novel zinc finger proteins: towards artificial control of specific gene expression. *Eur. J. Pharm. Sci.* **13**:91–97.
 128. Imanishi, M., Y. Hori, M. Nagaoka, and Y. Sugiura. 2000. DNA-bending finger: artificial design of 6-zinc finger peptides with polyglycine linker and induction of DNA bending. *Biochemistry* **39**:4383–4390.
 129. Inami, M., C. Morokuma, A. Sugio, H. Tamanoi, R. Yatsunami, and S. Nakamura. 2003. Directed evolution of xylanase J from alkaliphilic *Bacillus* sp. strain 41M-1: restore of alkaliphily of a mutant with an acidic pH optimum. *Nucleic Acids Res. Suppl.* **2003**:315–316.
 130. Jach, G., B. Gornhardt, J. Mundy, J. Logemann, P. Pinsdorf, R. Leah, J. Schell, and C. Maas. 1995. Enhanced quantitative resistance against fungal disease by combinatorial expression of different barley antifungal proteins in transgenic tobacco. *Plant J.* **8**:97–109.
 131. Jaeger, K. E., and T. Eggert. 2002. Lipases for biotechnology. *Curr. Opin. Biotechnol.* **13**:390–397.
 132. Jaeger, K. E., T. Eggert, A. Eipper, and M. T. Reetz. 2001. Directed evolution and the creation of enantioselective biocatalysts. *Appl. Microbiol. Biotechnol.* **55**:519–530.
 133. Jaeger, K. E., and M. T. Reetz. 2000. Directed evolution of enantioselective enzymes for organic chemistry. *Curr. Opin. Chem. Biol.* **4**:68–73.
 134. Jaeger, L., M. C. Wright, and G. F. Joyce. 1999. A complex ligase ribozyme

- evolved in vitro from a group I ribozyme domain. *Proc. Natl. Acad. Sci. USA* **96**:14712–14717.
135. **Jamieson, A. C., J. C. Miller, and C. O. Pabo.** 2003. Drug discovery with engineered zinc-finger proteins. *Nat. Rev. Drug Discov.* **2**:361–368.
 136. **Jensen, R. A.** 1976. Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* **30**:409–425.
 137. **Jermutus, L., A. Honegger, F. Schwesinger, J. Hanes, and A. Pluckthun.** 2001. Tailoring in vitro evolution for protein affinity or stability. *Proc. Natl. Acad. Sci. USA* **98**:75–80.
 138. **Joerger, A. C., S. Mayer, and A. R. Fersht.** 2003. Mimicking natural evolution in vitro: an N-acetylneuraminidase mutant with an increased dihydrodipicolinate synthase activity. *Proc. Natl. Acad. Sci. USA* **100**:5694–5699.
 139. **Joo, H., A. Arisawa, Z. Lin, and F. H. Arnold.** 1999. A high-throughput digital imaging screen for the discovery and directed evolution of oxygenases. *Chem. Biol.* **6**:699–706.
 140. **Joo, H., Z. Lin, and F. H. Arnold.** 1999. Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation. *Nature* **399**:670–673.
 141. **Joyce, G. F.** 2004. Directed evolution of nucleic acid enzymes. *Annu. Rev. Biochem.* **73**:791–836.
 142. **Juillerat, A., T. Gronemeyer, A. Keppler, S. Gendrezig, H. Pick, H. Vogel, and K. Johnsson.** 2003. Directed evolution of O(6)-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo. *Chem. Biol.* **10**:313–317.
 143. **Jung, S., A. Honegger, and A. Pluckthun.** 1999. Selection for improved protein stability by phage display. *J. Mol. Biol.* **294**:163–180.
 144. **Jurgens, C., A. Strom, D. Wegener, S. Hettwer, M. Wilmanns, and R. Sterner.** 2000. Directed evolution of a (beta alpha)8-barrel enzyme to catalyze related reactions in two different metabolic pathways. *Proc. Natl. Acad. Sci. USA* **97**:9925–9930.
 145. **Kallen, K. J., J. Grotzinger, and S. Rose-John.** 2000. New perspectives on the design of cytokines and growth factors. *Trends Biotechnol.* **18**:455–461.
 146. **Kamiuchi, T., E. Abe, M. Imanishi, T. Kajii, M. Nagaoka, and Y. Sugiura.** 1998. Artificial nine zinc-finger peptide with 30 base pair binding sites. *Biochemistry* **37**:13827–13834.
 147. **Kauffmann, I., and C. Schmidt-Dannert.** 2001. Conversion of *Bacillus thuringiensis* lipase into an efficient phospholipase with increased activity towards long-chain fatty acyl substrates by directed evolution and rational design. *Protein Eng.* **14**:919–928.
 148. **Kawarasaki, Y., K. E. Griswold, J. D. Stevenson, T. Selzer, S. J. Benkovic, B. L. Iverson, and G. Georgiou.** 2003. Enhanced crossover SCRATCHY: construction and high-throughput screening of a combinatorial library containing multiple non-homologous crossovers. *Nucleic Acids Res.* **31**:e126.
 149. **Kikuchi, M., and S. Harayama.** 2002. DNA shuffling and family shuffling in vitro gene evolution. *Methods Mol. Biol.* **182**:243–257.
 150. **Kikuchi, M., K. Ohnishi, and S. Harayama.** 2000. An effective family shuffling method using single-stranded DNA. *Gene* **243**:133–137.
 151. **Kim, B. S., D. H. Sherman, and K. A. Reynolds.** 2004. An efficient method for creation and functional analysis of libraries of hybrid type I polyketide synthases. *Protein Eng. Des. Sel.* **17**:277–284.
 152. **Kim, G. J., Y. H. Cheon, and H. S. Kim.** 2000. Directed evolution of a novel N-carbamylase/D-hydantoinase fusion enzyme for functional expression with enhanced stability. *Biotechnol. Bioeng.* **68**:211–217.
 153. **Kim, Y. S., H. C. Jung, and J. G. Pan.** 2000. Bacterial cell surface display of an enzyme library for selective screening of improved cellulase variants. *Appl. Environ. Microbiol.* **66**:788–793.
 154. **Kim, Y. W., J. H. Choi, J. Y. Kim, C. Park, H. Cha, S. B. Lee, B. H. Oh, T. W. Moon, and K. H. Park.** 2003. Directed evolution of *Thermus* maltogenic amylase toward enhanced thermal resistance. *Appl. Environ. Microbiol.* **69**:4866–4874.
 155. **Kitamura, K., Y. Kinoshita, S. Narasaki, N. Nemoto, Y. Husimi, and K. Nishigaki.** 2002. Construction of block-shuffled libraries of DNA for evolutionary protein engineering: Y-ligation-based block shuffling. *Protein Eng.* **15**:843–853.
 156. **Kitamura, K., C. Yoshida, and K. Nishigaki.** 2003. GFPs of insertion mutation generated by molecular size-altering block shuffling. *FEBS Lett.* **555**:483–488.
 157. **Kolkman, J. A., and W. P. Stemmer.** 2001. Directed evolution of proteins by exon shuffling. *Nat. Biotechnol.* **19**:423–428.
 158. Reference deleted.
 159. **Kong, X. D., Y. Liu, X. J. Gou, H. Y. Zhang, X. P. Wang, and J. Zhang.** 2001. Directed evolution of operon of trehalose-6-phosphate synthase/phosphatase from *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **280**:396–400.
 160. **Kong, X. D., Y. M. Liu, X. J. Gou, S. H. Zhu, H. Y. Zhang, X. P. Wang, and J. Zhang.** 2001. Directed evolution of alpha-aspartyl dipeptidase from *Salmonella typhimurium*. *Biochem. Biophys. Res. Commun.* **289**:137–142.
 161. **Korpimaki, T., J. Rosenberg, P. Virtanen, U. Lamminmaki, M. Tuomola, and P. Saviranta.** 2003. Further improvement of broad specificity hapten recognition with protein engineering. *Protein Eng.* **16**:37–46.
 162. **Kuchner, O., and F. H. Arnold.** 1997. Directed evolution of enzyme catalysts. *Trends Biotechnol.* **15**:523–530.
 163. **Kuhlman, B., G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, and D. Baker.** 2003. Design of a novel globular protein fold with atomic-level accuracy. *Science* **302**:1364–1368.
 164. **Kumamaru, T., H. Suenaga, M. Mitsuoka, T. Watanabe, and K. Furukawa.** 1998. Enhanced degradation of polychlorinated biphenyls by directed evolution of biphenyl dioxygenase. *Nat. Biotechnol.* **16**:663–666.
 165. **Kurtzman, A. L., S. Govindarajan, K. Vahle, J. T. Jones, V. Heinrichs, and P. A. Patten.** 2001. Advances in directed protein evolution by recursive genetic recombination: applications to therapeutic proteins. *Curr. Opin. Biotechnol.* **12**:361–370.
 166. **Kuyper, M., A. A. Winkler, J. P. van Dijken, and J. T. Pronk.** 2004. Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: a proof of principle. *FEMS Yeast Res.* **4**:655–664.
 167. **Kwon, S. J., A. L. de Boer, R. Petri, and C. Schmidt-Dannert.** 2003. High-level production of porphyrins in metabolically engineered *Escherichia coli*: systematic extension of a pathway assembled from overexpressed genes involved in heme biosynthesis. *Appl. Environ. Microbiol.* **69**:4875–4883.
 168. **Lanio, T., A. Jeltsch, and A. Pingoud.** 1998. Towards the design of rare cutting restriction endonucleases: using directed evolution to generate variants of EcoRV differing in their substrate specificity by two orders of magnitude. *J. Mol. Biol.* **283**:59–69.
 169. **Lassner, M., and J. Bedbrook.** 2001. Directed molecular evolution in plant improvement. *Curr. Opin. Plant Biol.* **4**:152–156.
 170. **Lassner, M. W., and D. McElroy.** 2002. Directed molecular evolution: bridging the gap between genomics leads and commercial products. *Omic* **6**:153–162.
 171. **Lawrence, J. G.** 2003. Gene organization: selection, selfishness, and serendipity. *Annu. Rev. Microbiol.* **57**:419–440.
 172. **Lawrence, J. G.** 2002. Shared strategies in gene organization among prokaryotes and eukaryotes. *Cell* **110**:407–413.
 173. **Lazetic, S., S. R. Leong, J. C. Chang, R. Ong, G. Dawes, and J. Punnonen.** 2002. Chimeric co-stimulatory molecules that selectively act through CD28 or CTLA-4 on human T cells. *J. Biol. Chem.* **277**:38660–38668.
 174. **Lee, D. K., W. Seol, and J. S. Kim.** 2003. Custom DNA-binding proteins and artificial transcription factors. *Curr. Top. Med. Chem.* **3**:645–657.
 175. **Lee, P. C., B. N. Mijts, and C. Schmidt-Dannert.** 2004. Investigation of factors influencing production of the monocyclic carotenoid torulene in metabolically engineered *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **65**:538–546.
 176. **Lee, P. C., A. Z. Momen, B. N. Mijts, and C. Schmidt-Dannert.** 2003. Biosynthesis of structurally novel carotenoids in *Escherichia coli*. *Chem. Biol.* **10**:453–462.
 177. **Lee, P. C., R. Petri, B. N. Mijts, K. T. Watts, and C. Schmidt-Dannert.** 2005. Directed evolution of *Escherichia coli* farnesyl diphosphate synthase (IspA) reveals novel structural determinants of chain length specificity. *Metab. Eng.* **7**:18–26.
 178. **Lee, P. C., and C. Schmidt-Dannert.** 2002. Metabolic engineering towards biotechnological production of carotenoids in microorganisms. *Appl. Microbiol. Biotechnol.* **60**:1–11.
 179. **Leong, S. R., J. C. Chang, R. Ong, G. Dawes, W. P. Stemmer, and J. Punnonen.** 2003. Optimized expression and specific activity of IL-12 by directed molecular evolution. *Proc. Natl. Acad. Sci. USA* **100**:1163–1168.
 180. **Levine, K., L. Kiang, M. D. Jacobson, R. P. Fisher, and F. R. Cross.** 1999. Directed evolution to bypass cyclin requirements for the Cdc28p cyclin-dependent kinase. *Mol. Cell* **4**:353–363.
 181. **Li, Q. S., U. Schwaneberg, M. Fischer, J. Schmitt, J. Pleiss, S. Lutz-Wahl, and R. D. Schmid.** 2001. Rational evolution of a medium chain-specific cytochrome P-450 BM-3 variant. *Biochim. Biophys. Acta* **1545**:114–121.
 182. **Li, Q. S., U. Schwaneberg, P. Fischer, and R. D. Schmid.** 2000. Directed evolution of the fatty-acid hydroxylase P450 BM-3 into an indole-hydroxylating catalyst. *Chemistry* **6**:1531–1536.
 183. **Liebeton, K., A. Zonta, K. Schimossek, M. Nardini, D. Lang, B. W. Dijkstra, M. T. Reetz, and K. E. Jaeger.** 2000. Directed evolution of an enantioselective lipase. *Chem. Biol.* **7**:709–718.
 184. **Lin, H., and V. W. Cornish.** 2002. Screening and selection methods for large-scale analysis of protein function. *Angew. Chem.* **41**:4402–4425.
 185. **Lin, Z., T. Thorsen, and F. H. Arnold.** 1999. Functional expression of horseradish peroxidase in *E. coli* by directed evolution. *Biotechnol. Prog.* **15**:467–471.
 186. **Liu, N., G. Caderas, C. Deillon, S. Hoffmann, S. Klauser, T. Cui, and B. Gutte.** 2001. Fusion proteins from artificial and natural structural modules. *Curr. Protein Pept. Sci.* **2**:107–121.
 187. **Liu, Q., D. J. Segal, J. B. Ghiara, and C. F. Barbas III.** 1997. Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *Proc. Natl. Acad. Sci. USA* **94**:5525–5530.
 188. **Locher, C. P., V. Heinrichs, D. Apt, and R. G. Whalen.** 2004. Overcoming antigenic diversity and improving vaccines using DNA shuffling and screening technologies. *Expert Opin. Biol. Ther.* **4**:589–597.
 189. **Locher, C. P., N. W. Soong, R. G. Whalen, and J. Punnonen.** 2004. Development of novel vaccines using DNA shuffling and screening strategies. *Curr. Opin. Mol. Ther.* **6**:34–39.
 190. **Lopez, J. V.** 2003. Naturally mosaic operons for secondary metabolite biosynthesis: variability and putative horizontal transfer of discrete catalytic

- domains of the epothilone polyketide synthase locus. *Mol. Genet. Genomics* **270**:420–431.
191. **Lorentsen, R. H., C. H. Moller, M. Etzerodt, H. C. Thogersen, and T. L. Holtet.** 2003. Substrate turnover and inhibitor binding as selection parameters in directed evolution of blood coagulation factor Xa. *Org. Biomol. Chem.* **1**:1657–1663.
 192. **Lo Surdo, P., M. A. Walsh, and M. Sollazzo.** 2004. A novel ADP- and zinc-binding fold from function-directed in vitro evolution. *Nat. Struct. Mol. Biol.* **11**:382–383.
 193. **Lutz, S., and M. Ostermeier.** 2003. Preparation of SCRATCHY hybrid protein libraries: size- and in-frame selection of nucleic acid sequences. *Methods Mol. Biol.* **231**:143–151.
 194. **Lutz, S., M. Ostermeier, and S. J. Benkovic.** 2001. Rapid generation of incremental truncation libraries for protein engineering using alpha-phosphothioate nucleotides. *Nucleic Acids Res.* **29**:E16.
 195. **Lutz, S., M. Ostermeier, G. L. Moore, C. D. Maranas, and S. J. Benkovic.** 2001. Creating multiple-crossover DNA libraries independent of sequence identity. *Proc. Natl. Acad. Sci. USA* **98**:11248–11253.
 196. **Lutz, S., and W. M. Patrick.** 2004. Novel methods for directed evolution of enzymes: quality, not quantity. *Curr. Opin. Biotechnol.* **15**:291–297.
 197. **Marshall, S. H.** 2002. DNA shuffling: induced molecular breeding to produce new generation long-lasting vaccines. *Biotechnol. Adv.* **20**:229–238.
 198. **Martin, A., V. Sieber, and F. X. Schmid.** 2001. In-vitro selection of highly stabilized protein variants with optimized surface. *J. Mol. Biol.* **309**:717–726.
 199. **Marzio, G., K. Verhoef, M. Vink, and B. Berkhout.** 2001. In vitro evolution of a highly replicating, doxycycline-dependent HIV for applications in vaccine studies. *Proc. Natl. Acad. Sci. USA* **98**:6342–6347.
 200. **Matsumura, I., and A. D. Ellington.** 2001. In vitro evolution of beta-glucuronidase into a beta-galactosidase proceeds through non-specific intermediates. *J. Mol. Biol.* **305**:331–339.
 201. **Matsumura, I., and A. D. Ellington.** 1999. In vitro evolution of thermostable p53 variants. *Protein Sci.* **8**:731–740.
 202. **Matsumura, I., and A. D. Ellington.** 2002. Mutagenic polymerase chain reaction of protein-coding genes for in vitro evolution. *Methods Mol. Biol.* **182**:259–267.
 203. **McCarthy, J. K., A. Uzelac, D. F. Davis, and D. E. Eveleigh.** 2004. Improved catalytic efficiency and active site modification of 1,4-beta-D-glucan glucosylase A from *Thermotoga neapolitana* by directed evolution. *J. Biol. Chem.* **279**:11495–11502.
 204. **Merz, A., M. C. Yee, H. Szadkowski, G. Pappenberger, A. Cramer, W. P. Stemmer, C. Yanofsky, and K. Kirschner.** 2000. Improving the catalytic activity of a thermophilic enzyme at low temperatures. *Biochemistry* **39**:880–889.
 205. **Mijts, B. N., P. C. Lee, and C. Schmidt-Dannert.** 2004. Engineering carotenoid biosynthetic pathways. *Methods Enzymol.* **388**:315–329.
 206. **Mijts, B. N., and C. Schmidt-Dannert.** 2003. Engineering of secondary metabolite pathways. *Curr. Opin. Biotechnol.* **14**:597–602.
 207. **Miller, L. W., and V. W. Cornish.** 2005. Selective chemical labeling of proteins in living cells. *Curr. Opin. Chem. Biol.* **9**:56–61.
 208. **Minagawa, H., J. Shimada, and H. Kaneko.** 2003. Effect of mutations at Glu160 and Val198 on the thermostability of lactate oxidase. *Eur. J. Biochem.* **270**:3628–3633.
 209. **Minshull, J., and W. P. Stemmer.** 1999. Protein evolution by molecular breeding. *Curr. Opin. Chem. Biol.* **3**:284–290.
 210. **Miyazaki, K., and F. H. Arnold.** 1999. Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function. *J. Mol. Evol.* **49**:716–720.
 211. **Miyazaki, K., P. L. Wintrod, R. A. Grayling, D. N. Rubingh, and F. H. Arnold.** 2000. Directed evolution study of temperature adaptation in a psychrophilic enzyme. *J. Mol. Biol.* **297**:1015–1026.
 212. **Moore, J. C., and F. H. Arnold.** 1996. Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. *Nat. Biotechnol.* **14**:458–467.
 213. **Murashima, K., and R. H. Doi.** 2003. Selection of heat-stable *Clostridium cellulovorans* cellulases after in vitro recombination. *Methods Mol. Biol.* **230**:231–237.
 214. **Murashima, K., A. Kosugi, and R. H. Doi.** 2002. Thermostabilization of cellulosomal endoglucanase EngB from *Clostridium cellulovorans* by in vitro DNA recombination with non-cellulosomal endoglucanase EngD. *Mol. Microbiol.* **45**:617–626.
 215. **Nagaoka, M., and Y. Sugiura.** 2000. Artificial zinc finger peptides: creation, DNA recognition, and gene regulation. *J. Inorg. Biochem.* **82**:57–63.
 216. **Naumann, T. A., and W. S. Reznikoff.** 2002. Tn5 transposase with an altered specificity for transposon ends. *J. Bacteriol.* **184**:233–240.
 217. **Ness, J. E., S. Kim, A. Gottman, R. Pak, A. Krebber, T. V. Borchert, S. Govindarajan, E. C. Mundorff, and J. Minshull.** 2002. Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently. *Nat. Biotechnol.* **20**:1251–1255.
 218. **Ness, J. E., M. Welch, L. Giver, M. Bueno, J. R. Cherry, T. V. Borchert, W. P. Stemmer, and J. Minshull.** 1999. DNA shuffling of subgenomic sequences of subtilisin. *Nat. Biotechnol.* **17**:893–896.
 219. **Nestle, M.** 2001. Genetically engineered “golden” rice unlikely to overcome vitamin A deficiency. *J. Am. Diet. Assoc.* **101**:289–290.
 220. **Nguyen, A. W., and P. S. Daugherty.** 2005. Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nat. Biotechnol.* **23**:355–360.
 221. **Oh, K. H., S. H. Nam, and H. S. Kim.** 2002. Directed evolution of N-carbamyl-D-amino acid amidohydrolase for simultaneous improvement of oxidative and thermal stability. *Biotechnol. Prog.* **18**:413–417.
 222. **Oh, K. H., S. H. Nam, and H. S. Kim.** 2002. Improvement of oxidative and thermostability of N-carbamyl-D-amino acid amidohydrolase by directed evolution. *Protein Eng.* **15**:689–695.
 223. **Osen, W., T. Peiler, P. Ohlschlager, S. Caldeira, S. Faath, N. Michel, M. Muller, M. Tommasino, I. Jochmus, and L. Gissmann.** 2001. A DNA vaccine based on a shuffled E7 oncogene of the human papillomavirus type 16 (HPV 16) induces E7-specific cytotoxic T cells but lacks transforming activity. *Vaccine* **19**:4276–4286.
 224. **Ostermeier, M., A. E. Nixon, and S. J. Benkovic.** 1999. Incremental truncation as a strategy in the engineering of novel biocatalysts. *Bioorg. Med. Chem.* **7**:2139–2144.
 225. **Ostermeier, M., J. H. Shim, and S. J. Benkovic.** 1999. A combinatorial approach to hybrid enzymes independent of DNA homology. *Nat. Biotechnol.* **17**:1205–1209.
 226. **Osuna, J., A. Perez-Blancas, and X. Soberon.** 2002. Improving a circularly permuted TEM-1 beta-lactamase by directed evolution. *Protein Eng.* **15**:463–470.
 227. **Otey, C. R., J. J. Silberg, C. A. Voigt, J. B. Endelman, G. Bandara, and F. H. Arnold.** 2004. Functional evolution and structural conservation in chimeric cytochromes P450: calibrating a structure-guided approach. *Chem. Biol.* **11**:309–318.
 228. **Otten, L. G., C. F. Sio, A. M. van der Sloot, R. H. Cool, and W. J. Quax.** 2004. Mutational analysis of a key residue in the substrate specificity of a cephalosporin acylase. *Chembiochem* **5**:820–825.
 229. **Otten, L. G., C. F. Sio, J. Vrielink, R. H. Cool, and W. J. Quax.** 2002. Altering the substrate specificity of cephalosporin acylase by directed evolution of the beta-subunit. *J. Biol. Chem.* **277**:42121–42127.
 230. **Panke, S., and M. G. Wubbolts.** 2002. Enzyme technology and bioprocess engineering. *Curr. Opin. Biotechnol.* **13**:111–116.
 231. **Park, S. J., W. S. Ahn, P. R. Green, and S. Y. Lee.** 2001. Production of poly(3-hydroxybutyrate-co-3-hydroxyhexanoate) by metabolically engineered *Escherichia coli* strains. *Biomacromolecules* **2**:248–254.
 232. **Patel, P. H., H. Kawate, E. Adman, M. Ashbach, and L. A. Loeb.** 2001. A single highly mutable catalytic site amino acid is critical for DNA polymerase fidelity. *J. Biol. Chem.* **276**:5044–5051.
 233. **Patel, P. H., and L. A. Loeb.** 2000. DNA polymerase active site is highly mutable: evolutionary consequences. *Proc. Natl. Acad. Sci. USA* **97**:5095–5100.
 234. **Patnaik, R., S. Louie, V. Gavrilovic, K. Perry, W. P. Stemmer, C. M. Ryan, and S. del Cardayre.** 2002. Genome shuffling of *Lactobacillus* for improved acid tolerance. *Nat. Biotechnol.* **20**:707–712.
 235. **Patten, P. A., R. J. Howard, and W. P. Stemmer.** 1997. Applications of DNA shuffling to pharmaceuticals and vaccines. *Curr. Opin. Biotechnol.* **8**:724–733.
 236. **Pedersen, H., S. Holder, D. P. Sutherlin, U. Schwitter, D. S. King, and P. G. Schultz.** 1998. A method for directed evolution and functional cloning of enzymes. *Proc. Natl. Acad. Sci. USA* **95**:10523–10528.
 237. **Pekrun, K., R. Shibata, T. Igarashi, M. Reed, L. Sheppard, P. A. Patten, W. P. Stemmer, M. A. Martin, and N. W. Soong.** 2002. Evolution of a human immunodeficiency virus type 1 variant with enhanced replication in pig-tailed macaque cells by DNA shuffling. *J. Virol.* **76**:2924–2935.
 238. **Peters, M. W., P. Meinhold, A. Glieder, and F. H. Arnold.** 2003. Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3. *J. Am. Chem. Soc.* **125**:13442–13450.
 239. **Petri, R., and C. Schmidt-Dannert.** 2004. Dealing with complexity: evolutionary engineering and genome shuffling. *Curr. Opin. Biotechnol.* **15**:298–304.
 240. **Pikkemaat, M. G., and D. B. Janssen.** 2002. Generating segmental mutations in haloalkane dehalogenase: a novel part in the directed evolution toolbox. *Nucleic Acids Res.* **30**:E35.
 241. **Pluckthun, A., C. Schaffitzel, J. Hanes, and L. Jermutus.** 2000. In vitro selection and evolution of proteins. *Adv. Protein Chem.* **55**:367–403.
 242. **Powell, K. A., S. W. Ramer, S. B. del Cardayre, W. P. C. Stemmer, M. B. Tobin, P. F. Longchamp, and G. W. Huisman.** 2001. Directed evolution and biocatalysis. *Angew. Chem.* **40**:3948–3959.
 243. **Powell, S. K., M. A. Kaloss, A. Pinkstaff, R. McKee, I. Burimski, M. Pensiero, E. Otto, W. P. Stemmer, and N. W. Soong.** 2000. Breeding of retroviruses by DNA shuffling for improved stability and processing yields. *Nat. Biotechnol.* **18**:1279–1282.
 244. **Proba, K., A. Worn, A. Honegger, and A. Pluckthun.** 1998. Antibody scFv fragments without disulfide bonds made by molecular evolution. *J. Mol. Biol.* **275**:245–253.
 245. **Punnonen, J.** 2000. Molecular breeding of allergy vaccines and antiallergic cytokines. *Int. Arch. Allergy Immunol.* **121**:173–182.
 246. **Rader, C., M. Popkov, J. A. Neves, and C. F. Barbas III.** 2002. Integrin

- alpha(v)beta3 targeted therapy for Kaposi's sarcoma with an in vitro evolved antibody. *FASEB J.* **16**:2000–2002.
247. **Reetz, M. T.** 2004. Controlling the enantioselectivity of enzymes by directed evolution: practical and theoretical ramifications. *Proc. Natl. Acad. Sci. USA* **101**:5716–5722.
 248. **Reetz, M. T.** 2002. Lipases as practical biocatalysts. *Curr. Opin. Chem. Biol.* **6**:145–150.
 249. **Reetz, M. T., R. Wenkel, and D. Avnir.** 2000. Entrapment of lipases in hydrophobic sol-gel-materials: efficient heterogeneous biocatalysts in aqueous medium. *Synth. Stuttgart* **2000**:781–783.
 250. **Roberts, G. C. K.** 1999. The power of evolution: accessing the synthetic potential of P450s. *Chem. Biol.* **6**:R269–R272.
 251. **Roberts, R. W., and J. W. Szostak.** 1997. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl. Acad. Sci. USA* **94**:12297–12302.
 252. **Rufer, A. W., and B. Sauer.** 2002. Non-contact positions impose site selectivity on Cre recombinase. *Nucleic Acids Res.* **30**:2764–2771.
 253. **Ryu, D. D., and D. H. Nam.** 2000. Recent progress in biomolecular engineering. *Biotechnol. Prog.* **16**:2–16.
 254. **Salamone, P. R., I. H. Kavakli, C. J. Slattery, and T. W. Okita.** 2002. Directed molecular evolution of ADP-glucose pyrophosphorylase. *Proc. Natl. Acad. Sci. USA* **99**:1070–1075.
 255. **Salazar, O., P. C. Cirino, and F. H. Arnold.** 2003. Thermostabilization of a cytochrome P450 peroxigenase. *ChemBiochem* **4**:891–893.
 256. **Samuelson, J. C., and S. Y. Xu.** 2002. Directed evolution of restriction endonuclease BstYI to achieve increased substrate specificity. *J. Mol. Biol.* **319**:673–683.
 257. **Sandmann, G.** 2003. Novel carotenoids genetically engineered in a heterologous host. *Chem. Biol.* **10**:478–479.
 258. **Santoro, S. W., and P. G. Schultz.** 2002. Directed evolution of the site specificity of Cre recombinase. *Proc. Natl. Acad. Sci. USA* **99**:4185–4190.
 259. **Schaffitzel, C., and A. Pluckthun.** 2001. Protein-fold evolution in the test tube. *Trends Biochem. Sci.* **26**:577–579.
 260. **Schellekens, H.** 2002. Immunogenicity of therapeutic proteins: clinical implications and future prospects. *Clin. Ther.* **24**:1720–1740.
 261. **Schmidt-Dannert, C.** 2001. Directed evolution of single proteins, metabolic pathways, and viruses. *Biochemistry* **40**:13125–13136.
 262. **Schmidt-Dannert, C.** 2000. Engineering novel carotenoids in microorganisms. *Curr. Opin. Biotechnol.* **11**:255–261.
 263. **Schmidt-Dannert, C., D. Umeno, and F. H. Arnold.** 2000. Molecular breeding of carotenoid biosynthetic pathways. *Nat. Biotechnol.* **18**:750–753.
 264. **Scholz, O., M. Kostner, M. Reich, S. Gastiger, and W. Hillen.** 2003. Teaching TetR to recognize a new inducer. *J. Mol. Biol.* **329**:217–227.
 265. **Sclimenti, C. R., B. Thyagarajan, and M. P. Calos.** 2001. Directed evolution of a recombinase for improved genomic integration at a native human sequence. *Nucleic Acids Res.* **29**:5044–5051.
 266. **Segal, D. J.** 2002. The use of zinc finger peptides to study the role of specific factor binding sites in the chromatin environment. *Methods* **26**:76–83.
 267. **Segal, D. J., and C. F. Barbas III.** 2001. Custom DNA-binding proteins come of age: polydactyl zinc-finger proteins. *Curr. Opin. Biotechnol.* **12**:632–637.
 268. **Segal, D. J., R. R. Beerli, P. Blancafort, B. Dreier, K. Effertz, A. Huber, B. Koksich, C. V. Lund, L. Magnenat, D. Valente, and C. F. Barbas III.** 2003. Evaluation of a modular strategy for the construction of novel polydactyl zinc finger DNA-binding proteins. *Biochemistry* **42**:2137–2148.
 269. **Segal, D. J., B. Dreier, R. R. Beerli, and C. F. Barbas III.** 1999. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. *Proc. Natl. Acad. Sci. USA* **96**:2758–2763.
 270. **Segal, D. J., J. Goncalves, S. Eberhardy, C. H. Swan, B. E. Torbett, X. Li, and C. F. Barbas III.** 2004. Attenuation of HIV-1 replication in primary human cells with a designed zinc finger transcription factor. *J. Biol. Chem.* **279**:14509–14519.
 271. **Segal, D. J., J. T. Stege, and C. F. Barbas III.** 2003. Zinc fingers and a green thumb: manipulating gene expression in plants. *Curr. Opin. Plant Biol.* **6**:163–168.
 272. Reference deleted.
 273. **Sengupta, D., H. Lin, S. D. Goldberg, J. J. Mahal, and V. W. Cornish.** 2004. Correlation between catalytic efficiency and the transcription read-out in chemical complementation: a general assay for enzyme catalysis. *Biochemistry* **43**:3570–3581.
 274. **Shaner, N. C., R. E. Campbell, P. A. Steinbach, B. N. Giepmans, A. E. Palmer, and R. Y. Tsien.** 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* **22**:1567–1572.
 275. **Shanklin, J.** 2000. Exploring the possibilities presented by protein engineering. *Curr. Opin. Plant Biol.* **3**:243–248.
 276. **Sieber, V., C. A. Martinez, and F. H. Arnold.** 2001. Libraries of hybrid proteins from distantly related sequences. *Nat. Biotechnol.* **19**:456–460.
 277. **Silberg, J. J., J. B. Endelman, and F. H. Arnold.** 2004. SCHEMA-guided protein recombination. *Methods Enzymol.* **388**:35–42.
 278. **Sio, C. F., L. G. Otten, R. H. Cool, and W. J. Quax.** 2003. Analysis of a substrate specificity switch residue of cephalosporin acylase. *Biochem. Biophys. Res. Commun.* **312**:755–760.
 279. **Skandalis, A., L. P. Encell, and L. A. Loeb.** 1997. Creating novel enzymes by applied molecular evolution. *Chem. Biol.* **4**:889–898.
 280. **Skarftad, E., E. O'Neill, J. Garmendia, and V. Shingler.** 2000. Identification of an effector specificity subregion within the aromatic-responsive regulators DmpR and XylR by DNA shuffling. *J. Bacteriol.* **182**:3008–3016.
 281. **Song, J. K., B. Chung, Y. H. Oh, and J. S. Rhee.** 2002. Construction of DNA-shuffled and incrementally truncated libraries by a mutagenic and unidirectional reassembly method: changing from a substrate specificity of phospholipase to that of lipase. *Appl. Environ. Microbiol.* **68**:6146–6151.
 282. **Song, J. K., and J. S. Rhee.** 2001. Enhancement of stability and activity of phospholipase A(1) in organic solvents by directed evolution. *Biochim. Biophys. Acta* **1547**:370–378.
 283. **Soong, N. W., L. Nomura, K. Pekrun, M. Reed, L. Sheppard, G. Dawes, and W. P. Stemmer.** 2000. Molecular breeding of viruses. *Nat. Genet.* **25**:436–439.
 284. **Spiller, B., A. Gershenson, F. H. Arnold, and R. C. Stevens.** 1999. A structural view of evolutionary divergence. *Proc. Natl. Acad. Sci. USA* **96**:12305–12310.
 285. **Sroga, G. E., and J. S. Dordick.** 2001. Generation of a broad esterolytic subtilisin using combined molecular evolution and periplasmic expression. *Protein Eng.* **14**:929–937.
 286. **Stemmer, W. P.** 1994. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* **91**:10747–10751.
 287. **Stemmer, W. P.** 1994. Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**:389–391.
 288. **Stoop, A. A., and C. S. Craik.** 2003. Engineering of a macromolecular scaffold to develop specific protease inhibitors. *Nat. Biotechnol.* **21**:1063–1068.
 289. **Stoop, A. A., E. Eldering, T. R. Dafforn, R. J. Read, and H. Pannekoek.** 2001. Different structural requirements for plasminogen activator inhibitor 1 (PAI-1) during latency transition and proteinase inhibition as evidenced by phage-displayed hypermutated PAI-1 libraries. *J. Mol. Biol.* **305**:773–783.
 290. **Stoop, A. A., L. Jespers, I. Lasters, E. Eldering, and H. Pannekoek.** 2000. High-density mutagenesis by combined DNA shuffling and phage display to assign essential amino acid residues in protein-protein interactions: application to study structure-function of plasminogen activation inhibitor 1 (PAI-I). *J. Mol. Biol.* **301**:1135–1147.
 291. **Suenaga, H., M. Mitsuoka, Y. Ura, T. Watanabe, and K. Furukawa.** 2001. Directed evolution of biphenyl dioxygenase: emergence of enhanced degradation capacity for benzene, toluene, and alkylbenzenes. *J. Bacteriol.* **183**:5441–5444.
 292. **Sun, J., J. Baudry, J. A. Katzenellenbogen, and B. S. Katzenellenbogen.** 2003. Molecular basis for the subtype discrimination of the estrogen receptor-beta-selective ligand, diarylpropionitrile. *Mol. Endocrinol.* **17**:247–258.
 293. **Sun, J., J. A. Katzenellenbogen, H. Zhao, and B. S. Katzenellenbogen.** 2003. DNA shuffling method for generating estrogen receptor alpha and beta chimeras in yeast. *BioTechniques* **34**:278–280.
 294. **Sun, L., T. Bulter, M. Alcalde, I. P. Petrounia, and F. H. Arnold.** 2002. Modification of galactose oxidase to introduce glucose 6-oxidase activity. *ChemBiochem* **3**:781–783.
 295. **Swers, J. S., B. A. Kellogg, and K. D. Wittrup.** 2004. Shuffled antibody libraries created by in vivo homologous recombination and yeast surface display. *Nucleic Acids Res.* **32**:e36.
 296. **Szostak, J. W.** 1997. In vitro selection and directed evolution. *Harvey Lect.* **93**:95–118.
 297. **Taguchi, S., A. Ozaki, T. Nonaka, Y. Mitsui, and H. Momose.** 1999. A cold-adapted protease engineered by experimental evolution system. *J. Biochem. (Tokyo)* **126**:689–693.
 298. **Takahashi, N., H. Kakinuma, L. Liu, Y. Nishi, and I. Fujii.** 2001. In vitro abzyme evolution to optimize antibody recognition for catalysis. *Nat. Biotechnol.* **19**:563–567.
 299. **Takase, K., S. Taguchi, and Y. Doi.** 2003. Enhanced synthesis of poly(3-hydroxybutyrate) in recombinant *Escherichia coli* by means of error-prone PCR mutagenesis, saturation mutagenesis, and in vitro recombination of the type II polyhydroxyalkanoate synthase gene. *J. Biochem. (Tokyo)* **133**:139–145.
 300. **Tan, W., K. Zhu, D. J. Segal, C. F. Barbas III, and S. A. Chow.** 2004. Fusion proteins consisting of human immunodeficiency virus type 1 integrase and the designed polydactyl zinc finger protein E2C direct integration of viral DNA into specific sites. *J. Virol.* **78**:1301–1313.
 301. **Tanaka, F., R. Fuller, H. Shim, R. A. Lerner, and C. F. Barbas III.** 2004. Evolution of aldolase antibodies in vitro: correlation of catalytic activity and reaction-based selection. *J. Mol. Biol.* **335**:1007–1018.
 302. **Tobin, M. B., C. Gustafsson, and G. W. Huisman.** 2000. Directed evolution: the 'rational' basis for 'irrational' design. *Curr. Opin. Struct. Biol.* **10**:421–427.
 303. **Toth, R. L., G. P. Pogue, and S. Chapman.** 2002. Improvement of the

- movement and host range properties of a plant virus vector through DNA shuffling. *Plant J.* **30**:593–600.
304. Uchiyama, H., T. Inaoka, T. Ohkuma-Soyejima, H. Togame, Y. Shibana, T. Yoshimoto, and T. Kokubo. 2000. Directed evolution to improve the thermostability of prolyl endopeptidase. *J. Biochem. (Tokyo)* **128**:441–447.
 305. Umeno, D., and F. H. Arnold. 2003. A C-35 carotenoid biosynthetic pathway. *Appl. Environ. Microbiol.* **69**:3573–3579.
 306. Urlacher, V., and R. D. Schmid. 2002. Biotransformations using prokaryotic P450 monooxygenases. *Curr. Opin. Biotechnol.* **13**:557–564.
 307. Urlacher, V. B., S. Lutz-Wahl, and R. D. Schmid. 2004. Microbial P450 enzymes in biotechnology. *Appl. Microbiol. Biotechnol.* **64**:317–325.
 308. van den Beucken, T., N. van Neer, E. Sablon, J. Desmet, L. Celis, H. R. Hoogenboom, and S. E. Hufton. 2001. Building novel binding ligands to B7.1 and B7.2 based on human antibody single variable light chain domains. *J. Mol. Biol.* **310**:591–601.
 309. van der Linden, R. H., B. de Geus, G. J. Frenken, H. Peters, and C. T. Verrips. 2000. Improved production and function of llama heavy chain antibody fragments by molecular evolution. *J. Biotechnol.* **80**:261–270.
 310. van der Veen, B. A., G. Potocki-Veronese, C. Albenne, G. Joucla, P. Monsan, and M. Remaud-Simeon. 2004. Combinatorial engineering to enhance amylosucrase performance: construction, selection, and screening of variant libraries for increased activity. *FEBS Lett.* **560**:91–97.
 311. Vardar, G., and T. K. Wood. 2004. Protein engineering of toluene-*o*-xylene monooxygenase from *Pseudomonas stutzeri* OX1 for synthesizing 4-methylresorcinol, methylhydroquinone, and pyrogallol. *Appl. Environ. Microbiol.* **70**:3253–3262.
 312. Verhaert, R. M. D., J. Beekwilder, R. Olsthoorn, J. van Duin, and W. J. Quax. 2002. Phage display selects for amylases with improved low pH starch-binding. *J. Biotechnol.* **96**:103–118.
 313. Volkov, A. A., Z. Shao, and F. H. Arnold. 1999. Recombination and chimeragenesis by in vitro heteroduplex formation and in vivo repair. *Nucleic Acids Res.* **27**:e18.
 314. Voziyanov, Y., J. H. Konieczka, A. F. Stewart, and M. Jayaram. 2003. Stepwise manipulation of DNA specificity in F1p recombinase: progressively adapting F1p to individual and combinatorial mutations in its target site. *J. Mol. Biol.* **326**:65–76.
 315. Wackett, L. P. 1998. Directed evolution of new enzymes and pathways for environmental biocatalysis. *Ann. N. Y. Acad. Sci.* **864**:142–152.
 316. Wada, M., C. C. Hsu, D. Franke, M. Mitchell, A. Heine, I. Wilson, and C. H. Wong. 2003. Directed evolution of N-acetylneuraminic acid aldolase to catalyze enantiomeric aldol reactions. *Bioorg. Med. Chem.* **11**:2091–2098.
 317. Waldo, G. S. 2003. Genetic screens and directed evolution for protein solubility. *Curr. Opin. Chem. Biol.* **7**:33–38.
 318. Walsh, C. 2001. Enabling the chemistry of life. *Nature* **409**:226–231.
 319. Wan, L., M. B. Twitchett, L. D. Eltis, A. G. Mauk, and M. Smith. 1998. In vitro evolution of horse heart myoglobin to increase peroxidase activity. *Proc. Natl. Acad. Sci. USA* **95**:12825–12831.
 320. Wang, C. W., and J. C. Liao. 2001. Alteration of product specificity of *Rhodospirillum rubrum* phytoene desaturase by directed evolution. *J. Biol. Chem.* **276**:41161–41164.
 321. Wang, C. W., M. K. Oh, and J. C. Liao. 2000. Directed evolution of metabolically engineered *Escherichia coli* for carotenoid production. *Biotechnol. Prog.* **16**:922–926.
 322. Wang, C. W., M. K. Oh, and J. C. Liao. 1999. Engineered isoprenoid pathway enhances astaxanthin production in *Escherichia coli*. *Biotechnol. Bioeng.* **62**:235–241.
 323. Wang, J. D., C. Herman, K. A. Tipton, C. A. Gross, and J. S. Weissman. 2002. Directed evolution of substrate-optimized GroEL/S chaperonins. *Cell* **111**:1027–1039.
 324. Watts, K. T., P. C. Lee, and C. Schmidt-Dannert. 2004. Exploring recombinant flavonoid biosynthesis in metabolically engineered *Escherichia coli*. *Chembiochem* **5**:500–507.
 325. Whalen, R. G., R. Kaiwar, N. W. Soong, and J. Punnonen. 2001. DNA shuffling and vaccines. *Curr. Opin. Mol. Ther.* **3**:31–36.
 326. Williams, G. J., S. Domann, A. Nelson, and A. Berry. 2003. Modifying the stereochemistry of an enzyme-catalyzed reaction by directed evolution. *Proc. Natl. Acad. Sci. USA* **100**:3143–3148.
 327. Winthrope, P. L., and F. H. Arnold. 2000. Temperature adaptation of enzymes: lessons from laboratory evolution. *Adv. Protein Chem.* **55**:161–225.
 328. Winthrope, P. L., K. Miyazaki, and F. H. Arnold. 2000. Cold adaptation of a mesophilic subtilisin-like protease by laboratory evolution. *J. Biol. Chem.* **275**:31635–31640.
 329. Winthrope, P. L., K. Miyazaki, and F. H. Arnold. 2001. Patterns of adaptation in a laboratory evolved thermophilic enzyme. *Biochim. Biophys. Acta* **1549**:1–8.
 330. Winthrope, P. L., D. Zhang, N. Vaidehi, F. H. Arnold, and W. A. Goddard III. 2003. Protein dynamics in a family of laboratory evolved thermophilic enzymes. *J. Mol. Biol.* **327**:745–757.
 331. Wiseman, A. 2003. Novel cytochromes P450 applications arising from the directed-evolution of recombinant micro-organisms. *Letts. Appl. Microbiol.* **37**:264–267.
 332. Wong, T. S., F. H. Arnold, and U. Schwaneberg. 2004. Laboratory evolution of cytochrome P450BM-3 monooxygenase for organic cosolvents. *Biotechnol. Bioeng.* **85**:351–358.
 333. Xia, G., L. Chen, T. Sera, M. Fa, P. G. Schultz, and F. E. Romesberg. 2002. Directed evolution of novel polymerase activities: mutation of a DNA polymerase into an efficient RNA polymerase. *Proc. Natl. Acad. Sci. USA* **99**:6597–6602.
 334. Xiroidimas, D. P., and D. P. Lane. 1999. Molecular evolution of the thermosensitive Pab1620 epitope of human p53 by DNA shuffling. *J. Biol. Chem.* **274**:28042–28049.
 335. Yang, H., P. D. Carr, S. Y. McLoughlin, J. W. Liu, I. Horne, X. Qiu, C. M. J. Jeffries, R. J. Russell, J. G. Oakshott, and D. L. Ollis. 2003. Evolution of an organophosphate-degrading enzyme: a comparison of natural and directed evolution. *Protein Eng.* **16**:241.
 336. Yang, J. K., M. S. Park, G. S. Waldo, and S. W. Suh. 2003. Directed evolution approach to a structural genomics project: Rv2002 from *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA* **100**:455–460.
 337. Yano, T., S. Oue, and H. Kagamiyama. 1998. Directed evolution of an aspartate aminotransferase with new substrate specificities. *Proc. Natl. Acad. Sci. USA* **95**:5511–5515.
 338. Yokobayashi, Y., C. H. Collins, J. R. Leadbetter, F. H. Arnold, and R. Weiss. 2003. Evolutionary design of genetic circuits and cell-cell communications. *Adv. Complex Syst.* **6**:37–45.
 339. Yokobayashi, Y., R. Weiss, and F. H. Arnold. 2002. Directed evolution of a genetic circuit. *Proc. Natl. Acad. Sci. USA* **99**:16587–16591.
 340. You, L., R. S. Cox III, R. Weiss, and F. H. Arnold. 2004. Programmed population control by cell-cell communication and regulated killing. *Nature* **428**:868–871.
 341. Zahnd, C., S. Spinelli, B. Luginbuhl, P. Amstutz, C. Cambillau, and A. Pluckthun. 2004. Directed in vitro evolution and crystallographic analysis of a peptide-binding single chain antibody fragment (scFv) with low picomolar affinity. *J. Biol. Chem.* **279**:18870–18877.
 342. Zengler, K., G. Toledo, M. Rappe, J. Elkins, E. J. Mathur, J. M. Short, and M. Keller. 2002. Cultivating the uncultured. *Proc. Natl. Acad. Sci. USA* **99**:15681–15686.
 343. Zha, D., A. Eipper, and M. T. Reetz. 2003. Assembly of designed oligonucleotides as an efficient method for gene recombination: a new tool in directed evolution. *Chembiochem* **4**:34–39.
 344. Zha, W., T. Zhu, and H. Zhao. 2003. Family shuffling with single-stranded DNA. *Methods Mol. Biol.* **231**:91–97.
 345. Zhang, J. H., G. Dawes, and W. P. Stemmer. 1997. Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc. Natl. Acad. Sci. USA* **94**:4504–4509.
 346. Zhang, X. X., Q. Deng, S. Y. Zhang, J. Liu, Q. Cai, Z. M. Lu, and Y. Wang. 2003. Broadly cross-reactive mimotope of hypervariable region 1 of hepatitis C virus derived from DNA shuffling and screened by phage display library. *J. Med. Virol.* **71**:511–517.
 347. Zhang, Y. X., K. Perry, V. A. Vinci, K. Powell, W. P. Stemmer, and S. B. del Cardayre. 2002. Genome shuffling leads to rapid phenotypic improvement in bacteria. *Nature* **415**:644–646.
 348. Zhao, H. 2003. A pH-indicator-based screen for hydrolytic haloalkane dehalogenase. *Methods Mol. Biol.* **230**:213–221.
 349. Zhao, H., and F. H. Arnold. 1999. Directed evolution converts subtilisin E into a functional equivalent of thermitase. *Protein Eng.* **12**:47–53.
 350. Zhao, H., and F. H. Arnold. 1997. Optimization of DNA shuffling for high fidelity recombination. *Nucleic Acids Res.* **25**:1307–1308.
 351. Zhao, H., L. Giver, Z. Shao, J. A. Affholter, and F. H. Arnold. 1998. Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol.* **16**:258–261.
 352. Zhao, H. M., K. Chockalingam, and Z. L. Chen. 2002. Directed evolution of enzymes and pathways for industrial biocatalysis. *Curr. Opin. Biotechnol.* **13**:104–110.
 353. Zhu, Z., J. C. Samuelson, J. Zhou, A. Dore, and S. Y. Xu. 2004. Engineering strand-specific DNA nicking enzymes from the type IIS restriction endonucleases BsaI, BsmBI, and BsmAI. *J. Mol. Biol.* **337**:573–583.