Directed evolution mimics natural evolution through screening or selection with the aim of generating genes that encode for proteins with new or improved functional traits (1). Experiments in directed evolution involve iterative cycles of diversification on the gene level (2) followed by screening or selection for the best variant on the protein level (3). Many different methods have been developed to generate rational guided or random mutant libraries (2,4–7), usually targeting only a single gene. DNA recombination strategies differ from directed protein evolution of single genes by generating chimeras through rational or random recombination of at least two genes. On the protein level, secondary structure elements, motifs, or domains preselected for function are mixed and used for the generation of new proteins to screen or select from (8).

Many methods for DNA recombination have been developed and applied in protein engineering campaigns (9). These methods can be grouped into homologous and nonhomologous strategies (10): homology-dependent recombination strategies such as DNA-shuffling (11), staggered extension process (StEP) (12), random priming (13), or random chimeragenesis of transient templates (RACHITT) (14) rely on the principle of hybridization and extension of homologous DNA fragments during PCRs. Hence, methods for homologous recombination require at least 70% sequence identity on the gene level. Homology-independent recombination strategies generate hybrid proteins of more distant related genes by using truncation-based methods such as incremental truncation for the creation of hybrid enzymes [ITCHY (15), SCRATCHY (16)] and sequence homology-independent protein recombination (SHIPREC) (17). Two genes were recombined in a single experiment as proof of concept in the latter methods. The sequence-independent site-directed chimeragenesis (SISDC) (18) method was reported to be able to recombine more than two genes, relying on the site-directed incorporation of marker tags that later have to be removed by Bael-restriction. Degenerate oligonucleotide gene shuffling (DOGS) (19) makes use of degenerate primers, which allow control over relative levels of recombination between the genes that are recombined, but requires endonucleases for cloning of chimeric libraries. Due to its simplicity, overlap extension PCR (OE-PCR) (20,21) is widely used and requires at least 15 overlapping nucleotides.

**Reports**

**Phosphorothioate-based DNA recombination: an enzyme-free method for the combinatorial assembly of multiple DNA fragments**

Jan Marienhagen, Alexander Dennig, and Ulrich Schwaneberg

*Department of Biotechnology, RWTH Aachen University, Aachen, Germany*

Rational guided generation of protein chimeras has developed into an attractive approach in protein engineering for tailoring catalytic and biophysical properties of enzymes. Combinatorial recombination of structural elements or whole protein domains is still technically challenging due to sequence dependent biases diminishing the overall quality of resulting chimeric libraries. Since methods for generating such libraries are often limited by a low frequency of crossover points and suffer from challenges in handling, we developed the phosphorothioate-based DNA recombination method (PTRec). PTRec is an enzyme-free method and only requires a short stretch of four amino acids that is identical among the proteins to be recombined in order to define a single crossover point. In a PTRec-generated chimeric library that shuffled five domains of phytase using genes from three different species, 88% of 42 randomly picked and sequenced genes were efficiently recombined. Furthermore, PTRec is a technically simple, fast, and reliable method that can be used for domain- and exon-shuffling or recombination of DNA fragments in general.

**Keywords:** recombination; directed evolution; protein engineering; phosphorothioate; PTRec

**Method summary:**

Phosphorothioate-based DNA recombination method (PTRec) is a ligase- and restriction site-independent approach for recombining secondary structure elements, motifs or domains of proteins. PTRec comprises only four simple steps including a highly efficient chemical cleavage of phosphorothioated nucleotides to generate 12-nucleotide long overhangs in double-stranded DNA with subsequent hybridization in a one-pot assembly. As proof of principle, five domains of three phytases (sequence identity: 45%–53%; 15 fragments) were recombined in a combinatorial manner with four crossover points. PTRec is "almost" homology-independent, not limited by the number of genes to recombine, and enables immediate and efficient cloning of the resulting chimeras.
between individual DNA-fragments that have to be assembled in a final PCR step. Recently, USER friendly DNA recombination (USERec) (10) has been introduced claiming “unique advantages in comparison to alternative recombination strategies,” for instance by overcoming sequence constraints with the usage of a 5′-ANk-3′-ANk motif at crossover sites. USERec relies on the activity of multiple enzymes: uracil DNA glycosylase, endonuclease VIII (together termed “USER enzyme”) (22), T4 DNA ligase, and requires a final PCR amplification step to generate large quantities of the recombined library. Homology-independent methods for DNA recombination have in common that they require enzymatic activities after DNA amplification/purification to generate the desired libraries. Additionally, a final PCR amplification step is often needed to amplify the chimeric library before transformation.

Here we report the phosphorothioate-based DNA recombination method (PTRec) as a ligase- and restriction site-independent method for recombining secondary structure elements, motifs or domains of proteins. PTRec is based on phosphorothioate chemistry, which allows site-specific cleavage of phosphorothioester bonds in phosphorothioate oligonucleotides in the presence of ethanol and iodine in an alkaline solution (23). PTRec starts with PCR amplification of the target DNA fragments and vector backbone by PCR using primers with complementary phosphorothioate nucleotides at their 5′-end. The PCR products are cleaved in an iodine/ethanol solution at elevated temperatures producing single-stranded overhangs. Subsequently, these complementary ends hybridize at room temperature, and the resulting DNA constructs can be directly transformed into competent host cells. As proof of principle, PTRec was used to recombine five domains of three different phytases (amino acid sequence identity: 45%–53%) by using four crossover points in the targeted proteins. Only a stretch of four amino acids that is identical among the proteins is required to define a crossover point, making PTRec almost independent from sequence constraints. PTRec is not limited by the number of genes to be recombined and enables immediate and efficient cloning of the resulting chimeric library. PTRec was developed after extensive optimization of DNA cleavage and hybridization conditions using the recently published phosphorothioate-based ligase-independent gene cloning (PLICing) method (24) and the OmniChange method (25), which were developed as a DNA fusion technology for cloning random mutant libraries of single genes and a method for the simultaneous site saturation mutagenesis of five codons in a single protein, respectively.

Materials and methods
All chemicals were analytical grade and purchased from Sigma-Aldrich (Steinheim, Germany), Serva (Heidelberg, Germany), or AppliChem (Darmstadt, Germany). Enzymes were obtained from Fermentas (St. Leon-Rot, Germany), and all oligonucleotides were purchased from Eurofins MWG Operon (Ebersberg, Germany). Oligonucleotides were diluted in Milli-Q water (Millipore, Billerica, MA, USA) to a final concentration of 100 µM. All oligonucleotides used in this study are summarized in Supplementary Table S1. The genes of the three phytases—appA of Escherichia coli (GenBank, AF537219), ympH of Verrassia mollarellii (GenBank, JF911533), and phyX of Haefnia alvei (Patent, U.S. no. 2008/0263688 A1)—were used for recombination as proof of concept for the PTRec method. The expression vector pET22b(+) was obtained from Novagen (Merck KGaA, Darmstadt, Germany), and E. coli BL21-Gold (DE3) was purchased from Invitrogen (Karlsruhe, Germany). Chemically competent E. coli BL21-Gold (DE3) cells (transformation efficiency: 1.4 × 10⁷ cfu/µg pUC19) were prepared according to a published procedure (26). PCR amplification of phytase genes, individual phytase domains, and vector pET22b(+) DNA isolations were performed using QIApREP Spin Miniprep kit (Qiagen, Hilden, Germany). PTRec requires only a short stretch of four amino acids (AS) that is identical among the proteins to be recombined in order to define a single crossover point for recombination. In the three model phytases, four crossover points were identified by multiple protein sequence alignments, giving rise to five different protein domains varying in size (domain A, ~50 amino acids; domain B, ~105 amino acids; domain C, ~155 amino acids; domain D, ~50 amino acids; and domain E, ~55 amino acids). As a hybridization method, PTRec requires sequence identity on the nucleotide level at the respective crossover points (see Supplementary Table S1). Despite identity on the protein level, the short crossover point stretches showed diversity on the nucleotide level due to the degeneracy of the genetic code. Therefore, 35 silent mutations in total were introduced into 19 out of 30 primers during oligonucleotide design. PCRs were performed in a final volume of 50 µL in a Gradient Cycler (Eppendorf, Darmstadt, Germany) using thin-wall PCR tubes (Sarstedt, Nuemrecht, Germany). PCRs for subcloning purposes were composed of 1× PfuS buffer, 0.2 mM dNTP mix, 400 nM forward and reverse primer (appA,
PCR Methods

P1 and P2; ympH, P3 and P4; phyX, P5 and P6; see Supplementary Table S1), 5 U PfuS DNA polymerase, and 20 ng template DNA. PCR cycling conditions: initial denaturation at 94°C for 2 min, 35 cycles [94°C for 30 s, 55°C for 30 s, 68°C for 50 s (phytases) or 4 min (pET22b(+)), and one fill up cycle (68°C for 10 min). Before iodine cleavage, the PCR products were DpnI-digested to remove the template DNA, column-purified using a PCR purification kit (Macherey-Nagel, Düren, Germany), and quantified using a NanoDrop 1000 UV spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). For amplification of individual phytase domains, PCRs contained 1× Taq buffer, 0.2 mM dNTP mix, 400 nM forward and reverse primers (see Supplementary Table S1), 5 U Taq DNA polymerase, and 40 ng of the respective pET22b(+) vector templates. PCR cycling conditions: initial denaturation at 94°C for 2 min, 35 cycles [94°C for 30 s, 55°C for 30 s, 72°C for 7 min (domains A1, B1, and C1, which include the pET22b(+) vector sequence; see Figure 1], or 1 min (all domains 2–5; see Figure 1) and one fill-up cycle (72°C for 5 min). Following completion of the PCRs, DpnI was added to remove template DNA. Subsequently, PCR products were purified and quantified (Nanodrop). Agarose-TAE gel electrophoresis was performed according to a standard protocol (27) to confirm presence and size of amplified genes, individual domains, or vectors. Colony PCRs were carried out before sequencing, following the subcloning PCR protocol with only slight modifications: a master mix was dispensed into 20-µL aliquots in PCR tubes. Colonies were directly transferred into PCR tubes with a sterile pipet tip. Initial denaturation time at 94°C was increased to 5 min to ensure cell lysis, and 5 µL the resulting PCR products were analyzed by agarose gel electrophoresis.

Iodine treatment, DNA-fragment hybridization, and transformation

For recombination purposes, iodine treatment and DNA fragment hybridization procedures were optimized to recombine multiple DNA fragments using the phosphorothioate cleavage principle reported by Blanusa et al. (24). Finally, a mixture of 4 µL DNA (0.03–0.6 pmol), 0.5 µL cleavage buffer (0.5 M Tris-HCl, pH 9.0), 0.3 µL iodine stock solution (100 mM iodine in 99% ethanol), and 0.2 µL Milli-Q water proved to be best for DNA cleavage. Samples were incubated in PCR tubes (70°C for 5 min; Eppendorf Mastercycler) and kept on ice until further use. Generated DNA fragments were subsequently hybridized without any purification by carefully mixing with a pipet. For phytase subcloning, 0.15 pmol cleaved pET22b(+) and 0.2 pmol cleaved appA, 0.22 pmol ympH, or 0.25 pmol phyX were combined. Resulting mixtures were incubated at 20°C for 5 min prior to transformation.

For hybridization during recombination experiments, a constant molar ratio of 1:4 [fragments A1, B1, and C1; each 0.09 pmol and each ~5.4 kb in size, including the linearized pET22b(+) vector, to 12 fragments: A2–A5, B2–B5, and C2–C5; each 0.36 pmol and each 0.17–0.45 kb] proved to be best to ensure hybridization (Figure 1). The molar ratio of 0.09 pmol to 0.36 pmol also resulted in good recombination and transformation results (see the Results and discussion section). After iodine cleavage, all 15 DNA fragments (five domains of each of the three phytase genes) were pooled on ice (Figure 1) and subsequently heated (70°C for 3 min; Eppendorf Mastercycler) to avoid preferential hybridization due to the order in pipeting. After

![Figure 2. Analysis of 42 clones recombined by PTRec.](image-url)
incubation at room temperature (20°C for 5 min), 3 μL hybridization reaction were transformed without further purification or ligation into chemically competent E. coli BL21 Gold (DE3) cells (50 μL) following a standard protocol (28). Transformation mixtures were spread on agar plates with Luria Bertani (LB) medium supplemented with 100 mg/mL ampicillin and incubated overnight at 37°C.

Colonies carrying the pET22b(+) vector with shuffled phytase domains were directly transferred into 96-well master plates, grown overnight in 100 μL LB medium with 100 mg/mL ampicillin at 37°C, and stored at -80°C after the addition of 100 μL (50% vol) glycerol.

Results and discussion

Three phytases, appF from Escherichia coli, ympH from Yersinia mollaretii, and phyX from Hafnia alvei were chosen to serve as model genes for development and validation of the PTRec method. The number of crossover points was set to four within the targeted genes, since most proof of concept demonstrations of gene shuffling methods have been performed with one to four crossover points. PTRec comprises four steps: (1) DNA fragment amplification by PCR using primers with complementary phosphorothioate nucleotides at the 5'-end; (2) chemical cleavage of the PCR products in an iodine/ethanol solution for generating single-stranded overhangs (12 nucleotides); (3) random DNA fragment assembly to full-length chimeric genes through hybridization at room temperature; and (4) transformation into competent host cells without further DNA purification or enzymatic DNA-ligation (Figure 1).

As a preparatory step, all phytases were subcloned for domain recombination experiments into the pET22b(+) vector by PLICing. The vector backbone together with the N-terminal portion of the genes was regarded as domain A1, B1, and C1, respectively, for the three phytase genes (Figure 1). A protein sequence alignment of the three phytases was used to select four crossover sites for recombination (Supplementary Figure S1). Identical stretches of four amino acids at potential crossover sites are necessary to design gene-specific complementary oligonucleotides with 12 phosphorothioate nucleotides at the 5’-end (Supplementary Table S1 and Supplementary Figure S1). One additional crossover site near the stop codon was chosen to recombine five protein domains of the three phytases in total (Figure 1). The 15 DNA fragments for domain shuffling were generated and purified as described in the Material and methods section (Figure 1, step 1). PTRec relies on the specific hybridization of single stranded 5'-ends of multiple DNA fragments and a vector backbone as described for the PLICing method (24) and the OmniChange method (25), which were developed as a cloning technology and a method for simultaneous site-saturation of five codons, respectively. In PTRec, overhangs are generated by chemical cleavage of the phosphorothioate bonds with iodine in an alkaline ethanol solution (Figure 1, step 2) (23). Phosphorothioate bonds were incorporated into the DNA fragments during PCR amplification using primers with 12 consecutive phosphorothioated nucleotides at the 5’-end (Supplementary Table S1 and Supplementary Figure S1). Critical for random but oriented assembly of the five fragments was a heat incubation protocol, which enables oriented hybridization of all DNA fragments to full plasmid size (Figure 1, step 3). The protocol ensures an efficient and stable assembly of phytase chimeras consisting of five fragments and 10 nicked DNA positions, even under transformation conditions. The latter could be achieved by heat incubation at 70°C for 3 min in the PCR-Cycler, cooling to 20°C for 5 min in the PCR-Cycler, and subsequent incubation on ice until transformation, which yielded 1100 clones (transformation efficiency: 5.1 × 10⁷ cfu/µg DNA).

As a first screen, 25 clones were randomly picked, and colony PCRs were performed.
In all 25 cases, full-length genes in the expected size range from 1250–1300 bp were obtained as PCR products. The chimeric genes differed slightly in length, since individual domains, depending on their respective origin, also varied in size. Prior to the development of PTRec, we expected that the repair of 10 DNA nicks by E. coli after transformation might become a critical step for PTRec, since hybridization and subsequent repair of 12 nicks proved to be challenging in the past (29). We assume that enzymatic fragment generation might not have been complete prior to hybridization and transformation, thereby causing reduced efficiencies in the past. In PTRec, the chemically generated 12-nucleotide overlaps were stable to transformation conditions, and the E. coli host BL21(DE3) closed the 10 nicked DNA positions without causing frame-shifts or mutations at the crossovers points. Encouraged by these results, 42 clones were randomly picked and fully sequenced. We consider this number of clones to be large enough to determine the efficacy of PTRec for DNA recombination. Figure 2 shows the diverse pattern of chimeric phytase genes. All 42 genes contained the five expected domains in the correct order, demonstrating an efficient assembly through oriented hybridization. No attempt was made to screen the generated chimeric library for activity, because the study was intended to develop PTRec and to evaluate its performance as an universal method for recombining DNA fragments of distinctly related genes. In order to benchmark PTRec, the number of fragments that can be assembled, the number of crossover points, and number of genes successfully recombined were determined. The distribution of individual fragments in the 42 clones was analyzed in detail. In none of the sequenced clones was a domain missing or occurring more than once. No preferential occurrence of any DNA fragment of the three phytase genes could be detected during analyses of the domains 1, 2, 4, and 5 (domain 1: appA, 18; ympH, 12; phyX, 12; domain 2: appA, 19; ympH, 13; phyX, 10; domain 4: appA, 14; ympH, 14; phyX, 14; domain 5: appA, 14; ympH, 15; phyX, 13). Only phyX domain 3 can be regarded as being underrepresented (domain 3: appA, 21; ympH, 14; phyX, 7) in the library, since this domain was identified only seven times. From 42 phytase genes sequenced, only five were not recombined (3× appA and 2× ympH). This can likely be attributed to incomplete DpnI digestion of template DNA prior to chemical cleavage. In addition, two of the chimeric phytase genes occurred twice. The theoretical size of the chimeric phytase library comprises 243 combinations (5 domains of 3 genes: 3^5), and thus screening of 729 clones (three times oversampling) is already sufficient to explore 95% of all possible combinations (30). Interestingly, no mutations were observed around the five crossover sites (+/-50 nucleotides), which would reduce the overall quality of the chimeric library. Altogether three transition mutations could be identified within the chimeric genes, which corresponds to the error-rate of 2.2 x 10^-5 of the used Tag DNA polymerase.

DNA recombination is a powerful tool for the laboratory evolution of proteins, and various methods were developed for generating such genetically diverse libraries (9). Table 1 summarizes the most advanced methods for nonhomologous gene recombination. Approximately one decade ago, the methods ITCHY (15), SCRATCHY (16), and SHIPREC (17) allowing random recombination of two genes with one to three crossover points were developed. The limitation regarding the number of crossover points shifted the attention to the development of methods with rationally guided selection of crossover points. Application of SISDC (18), OE-PCR (20,21), or USERec (10) allows recombination with up to nine crossover points (Table 1). However, enzymatic steps involved in fragment generation, ligation, and final PCR assembly limit the efficiency of nonhomologous recombination methods.

Compared with the above mentioned nonhomologous recombination methods, PTRec is unique, since enzymatic steps during fragment generation are avoided, and neither a ligation step nor a final assembly PCR is required. Key to DNA fragment assembly in PTRec is the efficient and specific chemical cleavage of multiple phosphorothioate diester bonds generating single-stranded 12-nucleotide overlaps for hybridization. Any DNA purification prior to hybridization and transformation can be skipped, since iodine and ethanol do not interfere with hybridization of complementary single-stranded DNA (24). Therefore, the time requirement for recombination experiments with PTRec is significantly reduced compared with all other nonhomologous recombination methods. In addition, avoiding inefficient enzymatic steps during fragment generation and eliminating ligation and PCR assembly steps may reduce the possibility of problems during library generation (24). A further advantage of PTRec is the minimal sequence requirement at each crossover point; a stretch of four amino acids that is identical among the proteins is sufficient to define a crossover point (Supplementary Figure S1). Differences in gene sequences can therefore be compensated by incorporation of silent mutations in the primers used in step 1, which lowers sequence requirements at individual crossover points on the nucleotide level even more. In 19 cases, incorporation of silent mutations in the PCR primers was necessary to ensure complementary single-stranded 5’-overhangs for efficient hybridization (Supplementary Table S1). Due to these minimal requirements, PTRec should enable researchers to use this method to recombine even more distantly related or even unrelated proteins, possibly leading to proteins with new functions. Moreover, due to the modular nature of the PCR-based fragment generation in step 1, PTRec allows generation of chimeric sublibraries; the number of crossover points can be adjusted by simply changing primer combinations. For example, if it turned out that it was better for domain 2 and domain 3 to originate from the same gene in order to yield functional proteins in the previous recombination experiments, the forward primer of domain 2 and the reverse primer of domain 3 are used during fragment generation in step 1 to omit the crossover point between these two domains. In summary, due to its simplicity in handling, high recombination efficiency, and flexibility, the PTRec method represents an innovative and powerful technique to generate high-quality chimeric libraries in a modular manner.

Acknowledgments
We thank Dr. Stefan Haefner (BASF SE) for fruitful discussions. This work was financially supported by BASF Societas Europaea (BASF SE).

Competing interests
The authors declare no competing interests.

References
otide exchanges that complements the bias of error-prone PCR. Biotechnol. J. 3:74-82.

Received 12 December 2011; accepted 9 April 2012.

Address correspondence to Ulrich Schwaneberg, RWTH Aachen University, Department of Biotechnology, Worringer Weg 1, 52074 Aachen, Germany. e-mail: u.schwaneberg@biotec.rwth-aachen.de

Supplementary material for this article is available at www.BioTechniques.com/article/113865.

To purchase reprints of this article, contact: biotechniques@fosterprinting.com