Chapter 11

Generation of Families of Construct Variants Using Golden Gate Shuffling

Carola Engler and Sylvestre Marillonnet

Abstract

Current standard cloning methods based on the use of restriction enzymes and ligase are very versatile, but are not well suited for high-throughput cloning projects or for assembly of many DNA fragments from several parental plasmids in a single step. We have previously reported the development of an efficient cloning method based on the use of type IIs restriction enzymes and restriction–ligation. Such method allows seamless assembly of multiple fragments from several parental plasmids with high efficiency, and also allows performing DNA shuffling if fragments prepared from several homologous genes are assembled together in a single restriction–ligation. Such protocol, called Golden Gate shuffling, requires performing the following steps: (1) sequences from several homologous genes are aligned, and recombination sites defined on conserved sequences; (2) modules defined by the position of these recombination sites are amplified by PCR with primers designed to equip them with flanking *Bsa*I sites; (3) the amplified fragments are cloned as intermediate constructs and sequenced; and (4) finally, the intermediate modules are assembled together in a compatible recipient vector in a one-pot restriction–ligation. Depending on the needs of the user, and because of the high cloning efficiency, the resulting constructs can either be screened and analyzed individually, or, if required in larger numbers, directly used in functional screens to detect improved protein variants.

Key words: DNA shuffling, High-throughput cloning, Restriction–ligation, Type IIs restriction enzymes, Seamless cloning, Modular cloning, Directed evolution

1. Introduction

The discovery of restriction enzymes almost 40 years ago provided biologists with the ability to create recombinant DNA molecules at will, in practice opening the field of modern molecular biology (1). Since then, cloning with restriction enzymes and ligase has become the base for work in molecular biology and has been hugely successful. However, despite such success, generating constructs is still a relatively slow and tedious process that has several limitations.

Chaofu Lu et al. (eds.), *cDNA Libraries: Methods and Applications*, Methods in Molecular Biology, vol. 729,
DOI 10.1007/978-1-61779-065-2_11,

In particular, each new construct requires the design of a specific cloning strategy, and only one or a few DNA fragments can be ligated together into a recipient vector in one step. Among some of the significant progresses made in cloning in the past few years, recombination-based cloning has been developed to facilitate the construction of plasmids for applications where a gene of interest needs to be transferred into several different acceptor expression vectors (2). Recombination-based cloning allows making constructs in a one-pot reaction with great efficiency. However, such approach still suffers several drawbacks, including the fact that unwanted recombination sites are left at the cloning junctions and that only a limited number of fragments (up to 4) can be cloned in a vector in one step.

We have recently developed a cloning method called Golden Gate cloning that overcomes many of the limitations of current cloning methods (3). This cloning strategy allows nine separate fragments to be cloned in a defined linear order into an acceptor vector in one step, with more than 90% of the colonies obtained containing the desired construct (4). Interestingly, this cloning method is not based on new exotic enzymes or genetic elements, rather it relies on the use of several previously known elements that, when used in concert, provide extremely high efficiency. The most important of these elements consists of the use of type IIs restriction enzymes combined with restriction–ligation, ensuring a very high cloning efficiency, and that only the correct desired construct remains at the end of the restriction–ligation. This cloning method not only has applications for cloning of any construct in general, but is also useful for DNA shuffling. Here, we provide a detailed protocol on how to perform Golden Gate shuffling, but the information provided is by default also useful for making any other type of construct that a user would want to make.

The principle of Golden Gate cloning/shuffling is based on the ability of type IIs enzymes to cleave outside of their recognition site sequence, allowing two DNA fragments flanked by compatible restriction sites to be digested and ligated seamlessly (5–8) (Fig. 1a). Since the ligated product of interest does not contain the original type IIs restriction site, it will not be subject to redigestion in a restriction–ligation. However, all other products that reconstitute the original site will be redigested, allowing their components to be made available for further ligation, leading to the formation of an increasing amount of desired product with increasing time of incubation. Since the sequence of the overhangs at the ends of the digested fragments can be chosen to be any four nucleotide sequence of choice, more than one fragment of interest can be assembled together in a defined linear order in a single restriction–ligation. Moreover, if fragments of several homologous genes are available in the same restriction–ligation mix, DNA shuffling is possible.

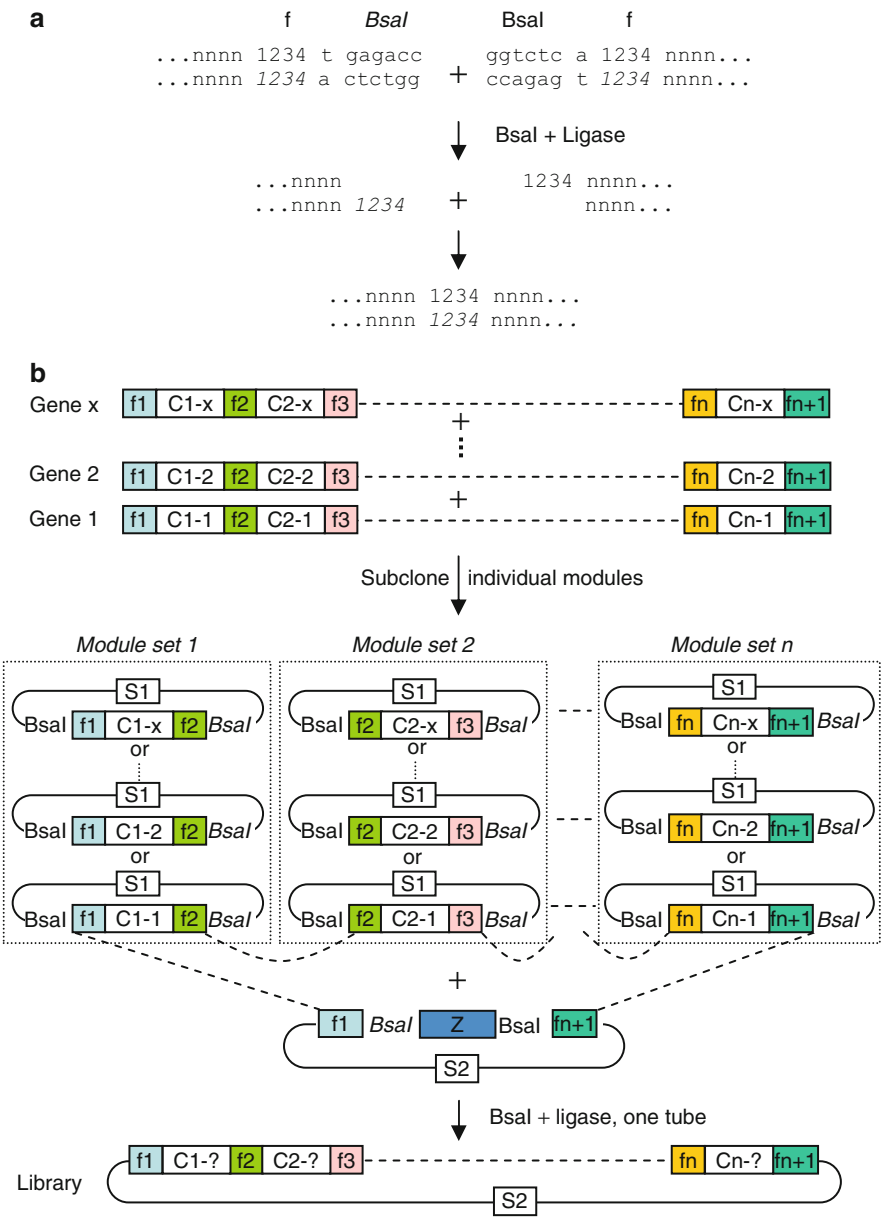

Fig. 1. DNA shuffling strategy. (**a**) Two DNA ends terminated by the same four nucleotides (sequence f, composed of nucleotides 1234, complementary nucleotides noted in *italics*) flanked by a *Bsa*I recognition sequence form two complementary DNA overhangs after digestion with *Bsa*I. (**b**) For shuffling, genes of interest are aligned, and recombination points consisting of four nucleotide sequences (f1 to fn+1) are defined on conserved sequences. Module fragments (core sequence, C1 to Cn, plus flanking four nucleotide sequences) are amplified by PCR and cloned in an intermediate cloning vector. Module fragment plasmids and the acceptor vector are assembled in one restriction–ligation with *Bsa*I and ligase. S1 and S2, two different selectable markers. *Z* lacZ alpha gene fragment.

Basically, using Golden Gate shuffling requires performing the following steps (Fig. 1b): (1) an alignment between the sequences to be shuffled is made and recombination sites are defined on conserved sequences, (2) modules defined by the position of these recombination sites are amplified by PCR with primers designed to equip them with flanking *Bsa*I sites, (3) the amplified fragments are cloned as intermediate constructs and sequenced, and (4) finally, the intermediate modules are assembled together in a compatible recipient vector in a one-pot restriction–ligation. Depending on the needs of the user, and because of the high cloning efficiency, the resulting constructs can either be screened and analyzed individually, or, if required in larger numbers, directly used in functional screens to detect improved protein variants.

2. Materials

2.1. Polymerase Chain Reaction

1. Novagen KOD Hot Start DNA polymerase (Merck KGaA, Darmstadt), supplied with 10× buffer, 25 mM $MgSO_4$, and 2 mM dNTPs.
2. Custom-made primers can be ordered from any of the many commercial vendors (for example, Invitrogen, Karlsruhe).
3. NucleoSpin® Extract II kit (Macherey Nagel, Dueren), for purification of PCR products.

2.2. Cloning

1. Restriction endonuclease *Sma*I (10 U/μL) (NEB, New England Biolabs Inc., Ipswich, MA, USA), supplied with 10× NEBuffer 4 (200 mM Tris–acetate pH 7.5, 100 mM magnesium acetate, 500 mM potassium acetate, and 10 mM dithiothreitol).
2. Restriction endonuclease *Bsa*I (10 U/μL) (NEB), supplied with 10× NEBuffer 4.
3. T4 DNA ligase 3 U/μL or T4 DNA ligase (HC) 20 U/μL (Promega, Mannheim), both supplied with 10× ligation buffer (300 mM Tris–HCl pH 7.8, 100 mM $MgCl_2$, 100 mM DTT, and 10 mM ATP).
4. For measuring of DNA concentration, we use the NanoDrop ND1000 (Peqlab, Erlangen).
5. Luria-Bertani (LB) medium: 1% bacto-tryptone, 0.5% yeast extract, 1% NaCl in deionized water, and adjusted to pH 7.0 with 5 N NaOH. For plates, 1.5% agar is added.
6. Antibiotics carbenicillin (used instead of ampicillin) and kanamycin: filter-sterilized stocks of 50 mg/mL in H_2O (stored in aliquots at −20°C) are diluted 1:1,000 (final concentration: 50 μg/mL) in an appropriate amount of medium after the medium has been autoclaved and cooled down.

For spectinomycin, a stock of 40 mg/mL is made and is used at a final concentration of 100 μg/mL (dilution 1:400).

7. 5-Bromo-4-chloro-3-indolyl-β-d-galactopyranoside (X-gal): stock solution of 20 mg/mL in dimethylformamide (DMF). For preparation of plates, the stock is diluted 1:500 (final concentration: 40 μg/mL) in an appropriate amount of LB agar after autoclaving/melting and cooling down.

2.3. Preparation of Chemically Competent Cells

1. Solution TFB1: 30 mM potassium acetate, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 100 mM RbCl, and 15% glycerol; adjust to pH 5.8 (with 1 M acetic acid), filter-sterilize, and store at 4°C (ready to use) or at room temperature (cool down before use).
2. Solution TFB2: 100 mM MOPS (or PIPES), 75 mM $CaCl_2$, 10 mM RbCl, and 15% glycerol; adjust to pH 6.5 (with 1 M KOH), filter-sterilize, and store at 4°C (ready to use) or at room temperature (cool down before use).
3. The OD_{600} of bacterial cultures is measured in a SmartSpec™3000 spectrophotometer (Biorad, Muenchen).

2.4. Screening of Colonies

1. DNA minipreps: NucleoSpin® Plasmid Quick Pure (Macherey Nagel, Dueren).
2. Restriction endonucleases (NEB or Fermentas, St. Leon-Rot), all supplied with 10× buffer and if necessary also with 100× BSA (dilute 1:10 and store in aliquots at −20°C).
3. DNA ladder: GeneRuler™ 1-kb DNA Ladder Plus (Fermentas) is used as marker for gel electrophoresis.
4. 50× TAE buffer: 242.0 g Tris, 57.1 mL acetic acid, and 100 mL 0.5 M EDTA, pH 8.0, in 1 L of deionized water.
5. Gels: agarose (0.7–1.5%) in 1× TAE is melted in a microwave oven and one drop of 0.025% ethidium bromide solution (Carl Roth GmbH, Karlsruhe) is added per 100 mL of melted agarose solution.
6. Running buffer of agarose gels is 1× TAE.
7. Gels are checked visually using a Syngene GelVue transilluminator (VWR, Darmstadt), and pictures are taken by using a Quantity one® gel analysis software (Biorad).
8. DNA maps of plasmids are made by using the Vector NTI software (Invitrogen).

2.5. Sequencing

1. DNA/constructs to sequence are sent to an external contractor (GATC Biotech, Konstanz). Sequence data are analyzed using the DNASTAR's Lasergene software.
2. Primers M13RP (CAGGAAACAGCTATGACC) and/or M13FP (TGTAAAACGACGGCCAGT) are used for sequencing of inserts cloned in pUC19-derived vectors.

3. Methods

3.1. Selection of Recombination Points

The first step for performing DNA shuffling consists of selecting "recombination sites" within several parental sequences. We use the term "recombination sites" in functional analogy to the recombination sites used in recombination systems such as the phage P1 Cre-loxP recombination system. However, for the purpose described here, "recombination sites" can be any three or four nucleotide sequences of choice defined as such (with the limitations described below), and will serve as the sequence where restriction enzyme digestion and ligation will take place; no real recombinase will in fact be used. We call the sequence between recombination sites as "modules" (more generally, plasmids containing such sequences are also referred to as modules). One "set" of modules consists of all equivalent modules prepared from homologous sequences and flanked by the same two recombination sites (Fig. 1b).

1. A first and obvious requirement for choosing recombination sites is to select them within sequences conserved between all parental sequences. This requirement is easy to satisfy since only four nucleotides need to be conserved for each recombination site (one amino acid) if an enzyme such as *Bsa*I is used for shuffling (digestion with *Bsa*I results in a four-nucleotide overhang), and only three nucleotides if an enzyme such as *Sap*I is used (see Note 1). Therefore, the first step in gene shuffling consists of performing an alignment of the amino acid sequence, and then of the nucleotide sequence (we use the vector NTI program, but any other program will also be suitable). Performing a sequence alignment can, however, be omitted if nonhomologous parental sequences are used (see Note 2).
2. A second requirement for recombination sites is to avoid selecting the same sequence twice, as this would lead to illegitimate recombination and deletion of the sequences between the two sites. It is also important to make sure that the sequence of any site does not match the sequence of any of the other chosen sites, both on the same and on the complement strand. For example, choice of the sequence ATTC will preclude the choice of the sequence GAAT for any of the other recombination sites used for this shuffling experiment. Use of two such sites would sometimes lead to ligation of two inappropriate fragments, one in the opposite orientation. This would lead to the formation of molecules that will not be able to form circular plasmids, but that would continue to ligate to further modules and form long linear multimeric concatemers.

3. A third requirement is to avoid the 16 palindromic sequences, since any palindromic DNA end can be ligated to another copy of the same DNA fragment in opposite orientation, and lead to the same problem as described above. For enzymes leading to a four-nucleotide extension, 240 different sequences are, therefore, available.
4. Finally, a fourth but optional requirement can be defined to maximize the efficiency of DNA shuffling. We have observed that inappropriate ligation of fragments can occur between ends with four-nucleotide overhangs that match for three of four consecutive nucleotides, for example, as in sequences GGTG and AGTG, or GGTG and CACT. Therefore, combination of two such sites should be avoided if possible (see Note 3).

Other than the minimal requirements defined above, the number of recombination sites, as well as their position within the gene to shuffle, is chosen depending on the needs and the goal of the user for each specific protein. Therefore, the size (see Note 4) and number of fragments to shuffle will vary for each gene and each experiment. We have tested up to eight recombination points within a gene (nine module sets), but a higher number should be possible as well, although probably with reduced efficiency.

The following steps consist of amplifying the defined modules by PCR, cloning them in intermediate vectors, and sequencing them. Alternatively, these steps can be replaced by simply ordering these modules from a gene synthesis company (see Note 5).

3.2. PCR Amplification of the Modules

Modules defined by the position of the recombination sites need to be amplified by PCR using primers designed to add two *Bsa*I sites flanking each module. Primers are designed such that the overhangs created by digestion of the amplified products with *Bsa*I (or any other type IIs enzyme chosen) correspond to the sequence of the chosen recombination sites. Therefore, the sequence ttg gtctca is added to each primer sequence, for example, ttggtctca CAGG nnnnn (CAGG being the recombination site, followed by 16–20 nucleotides of target sequence). For nine modules prepared from three homologs, 54 primers need to be made.

Moreover, a requirement for Golden Gate shuffling is to not have any internal *Bsa*I sites present within any of the DNA fragments used for shuffling. Indeed, the presence of a *Bsa*I site within one of the modules would lead to redigestion of the shuffled DNA molecules containing such fragment at the end of the assembly step. These linear molecules will not transform *Escherichia coli*. Therefore, any such site needs to be removed before this step, and doing so at the time of generation of the entry clones is appropriate. We have previously described a method to remove internal *Bsa*I sites upon cloning in the entry

vector, which requires using a specific cloning vector corresponding to the combination of recombination sites flanking the given fragment (3). However, this method is not useful for DNA shuffling since new and different recombination sites are chosen for each new shuffling experiment, and therefore, specific cloning vectors are not available. Instead, removal of internal *Bsa*I sites from PCR fragments can be done easily using gene SOEing (9) (see Note 6).

PCR is usually performed using plasmid DNA, but genomic DNA or cDNAs can also be used as starting material, depending on the source available to the user; for small modules, template DNA is not even necessary (see Note 7). For amplification, we use the enzyme KOD Hot Start DNA polymerase since it has a very low error rate and, unlike Taq polymerase or many enzyme mixes, produces DNA products with blunt ends (other thermostable polymerases that a user would prefer to use are of course perfectly acceptable as well). Blunt ends are advantageous, as the products can be easily cloned in any standard vector such as pUC19 by blunt-end cloning (see below), and therefore, cloning them does not require the purchase of a kit.

1. The PCR mix is set up following the manufacturer's instructions, for example, using KOD polymerase, with the following conditions: 1 μL of plasmid DNA (5–20 ng/μL), 5 μL of 10× buffer, 3 μL of 25 mM $MgSO_4$, 5 μL of 2 mM dNTPs, 1.5 μL each of 10 μM sense and antisense primers, and 1 μL of KOD Hot Start DNA polymerase (10 U/μL, final concentration 0.02 U/μL) in a total reaction volume of 50 μL.
2. PCR is performed using the following cycling conditions: (1) incubation at 95°C for 2 min for polymerase activation, (2) denaturation at 95°C for 20 s, (3) annealing at 58°C for 10 s – the temperature for the annealing step can be adjusted for specific primers, but the temperature of 58°C usually works well for primers designed as described above, (4) extension at 70°C, the duration depends on the length of the expected fragment (from 10 s/kb for fragments smaller than 500 bp up to 25 s/kb for fragments larger than 3 kb, see manufacturer's instructions); steps 2–4 are repeated 35 times and are followed by a final extension step at 70°C for 20 s to 2 min (depending on fragment length). The reaction is then incubated at 12°C until taken out of the thermocycler.
3. Of the PCR product obtained, 2 μL is then analyzed by gel electrophoresis to make sure that a product of the correct size has been amplified.
4. The amplified fragment is purified from remaining primers, potential primer dimers, and remaining polymerase enzyme by using the NucleoSpin® Extract II kit and following the kit protocol. DNA is eluted from the column with 30–50 μL of

elution buffer (5 mM Tris–HCl, pH 8.5). In case several bands were amplified rather than only the expected fragment, the same kit can also be used to cut and extract the appropriate DNA fragment from an agarose gel.

3.3. Blunt-End Cloning of the Modules

Many commercial kits are available for cloning PCR products, including the pGEM-T kit (Promega), pJET (Fermentas), and TOPO® TA kit (Invitrogen). However, PCR products can also be cloned very efficiently without purchasing a kit, by using blunt-end cloning performed with a restriction–ligation (10, 11). This method is very efficient and has the advantage that the DNA fragment of interest to be cloned does not need to be flanked by any specific sequence, and, therefore, restrictions on primer design are minimal (see Note 8). Another advantage of this cloning method is that any plasmid of choice by a user can be used as long as it contains a unique blunt site, preferably in a reporter gene such as *LacZ*. This is useful since cloning vectors for generating entry clones for shuffling need to fulfill preferentially two requirements: (1) they should preferably not contain any restriction site for the type IIs enzyme chosen for shuffling (see Note 9) and (2) the antibiotic resistance gene of the entry vector should preferably be different from the one in the destination vector. Since several commercial cloning vectors have a *Bsa*I restriction site in the ampicillin resistance gene (for example, pGEM-T or pJET), we have made our own entry cloning vectors that simply consist of pUC19 lacking a *Bsa*I restriction site (see Note 10).

1. Add 0.5 μL of vector (50 ng), 1 μL of PCR product (50–100 ng), 2 μL of 10× ligation buffer (Promega), 1 μL of *Sma*I enzyme (10 U; NEB), 1 μL of ligase (3 U; Promega), and 14.5 μL of water (total volume of 20 μL) into a tube. The reaction mix is incubated for 1–2 h at room temperature or in a 25°C incubator, if one is available.
2. The entire ligation mix is transformed to DH10B chemically competent cells and plated on LB plates with X-gal and the appropriate antibiotic (the transformation protocol is described below in paragraph Subheading 3.6).
3. White colonies (or sometimes pale blue when small inserts are cloned) are picked and inoculated in 5 mL of LB medium containing the appropriate antibiotic.
4. Plasmid DNA is extracted using the NucleoSpin® Plasmid Quick Pure kit from Macherey Nagel following the manufacturer's instructions.
5. Plasmid DNA can be checked by restriction enzyme digestion using *Bsa*I and analysis of the digested DNA by agarose gel electrophoresis. A fragment of the size of the expected module should be visible.

6. Two minipreps are sent for sequencing using primers M13RP and/or M13FP.
7. When the correct sequence has been verified, DNA concentration of the plasmid prep is measured using the NanoDrop ND1000 (Peqlab).

3.4. Construction of the Destination Vector

A destination vector compatible with the entry modules needs to be made. There are many ways of making a destination vector, but what is important is that the final vector should respect the following criteria. First, it should contain two *Bsa*I sites (or any other type IIs enzyme chosen) with cleavage sites compatible with the beginning of the first and the end of the last entry module sets. The vector backbone should not contain any other *Bsa*I restriction site. And finally, it should have an antibiotic selectable marker different from the one used in the entry clones. No description is provided here for making such a vector, since each vector will require a specific construction strategy.

3.5. DNA Shuffling

Once entry constructs and the recipient vector are made and sequenced, performing DNA shuffling only requires pipetting all components into a reaction mix, incubating the mix in a thermocycler, and transforming it into competent cells. An important factor is to add an equimolar amount of DNA for each of the module sets and the destination vector. Since a module set usually contains several modules, the amount of DNA for each individual module of a set containing x different modules (x alternative homologous sequences) should contain only $1/x$ the amount of DNA compared to the vector; for example, each module from a set containing three modules should have a third of the amount of DNA compared with the recipient vector.

1. A restriction–ligation is set up by pipetting 40 fmol (or 100 ng, see Note 11) of each module set and of the vector, 2 μL of 10× ligation buffer, 10 U (1 μL) of *Bsa*I, and either 3 U (1 μL) of ligase for assembly of 2–4 module sets or 20 U (1 μL) of HC ligase for assembly of more than four module sets, in a total volume of 20 μL into a tube.
2. The restriction–ligation mix is incubated in a thermocycler. For assembly of 2–4 module sets, incubation for 60–120 min at 37°C is sufficient. If more module sets are ligated together, the incubation time is increased to 6 h, or cycling is used as following: 2 min for 37°C followed by 3 min for 16°C, both repeated 50 times (see Note 12).
3. Restriction–ligation is followed by a digestion step (5 min at 50°C) and then by heat inactivation for 5 min at 80°C. The final incubation step at 80°C is very important and should not be omitted. Its purpose is to inactivate the ligase at the end of the restriction–ligation. Omitting this step would lead

to religation of some of the insert and plasmid backbone fragments still present in the mix, when it is taken out of the thermocycler before transformation. Such unwanted products might be ligated more efficiently than they are redigested by the type IIs enzyme at room temperature. Therefore, a larger percentage of colonies would contain such type of undesired ligation products.

3.6. Transformation of the Library in Competent Cells

The entire ligation is transformed into chemically competent DH10B cells (see Note 13).

1. Frozen chemically competent cells (100 μL per tube) are thawed on ice.
2. The entire ligation is added to the cells, and the mix incubated on ice for 30 min.
3. The cells and DNA mix is heat shocked for 90 s at 42°C in a water bath.
4. The cells are allowed to recover on ice for 5 min.
5. To the cells, add 1 mL of LB medium, and incubate the tube at 37°C in a shaker-incubator (150 rpm) for 45 min to 1 h.
6. After incubation, 25–100 μL of the transformation are plated on LB agar plates containing antibiotic and X-gal. Plating of an aliquot of the transformation is necessary to estimate the number of independent constructs that will be obtained. The remainder of the transformation can be inoculated into 5 mL of liquid LB with the appropriate antibiotic if users want to grow the entire library.
7. The plates and liquid culture are incubated overnight at 37°C.
8. Many white and very few blue colonies should be obtained on the plate. A few white colonies from the plate can be picked for preparation of miniprep DNA. Plasmid DNA can be analyzed by restriction digestion and sequencing to estimate the number of correct clones.
9. Miniprep DNA is also prepared from the liquid culture. This DNA prep should represent a library of constructs containing shuffled DNA. Depending on the specific goal of the shuffling experiment, clones can either be functionally screened individually or as a library. The shuffled plasmid library may be transformed in any target organism of choice for functional screening.

3.7. Preparation of Chemically Competent DH10B Cells

Chemically competent or electrocompetent *E. coli* cells can either be purchased from a commercial vendor or made in the laboratory. The protocol that we use is as follows:

1. *E. coli* strain DH10B is inoculated from a glycerol stock onto an LB plate; the inoculum is streaked on the plate using a loop

so as to obtain individual colonies. The plate is incubated overnight at 37°C.

2. Inoculate 5 mL of LB from a single colony and incubate the flask overnight in a shaker-incubator (37°C, shaking 150 rpm).
3. The following day, transfer 2 mL of this culture to a flask containing 200 mL of LB and incubate for around 2 h until OD_{600} reaches 0.6.
4. Cool down the cells on ice for 10 min. The cells are pelleted in a centrifuge for 5 min at 4,500 rpm (4,000 × *g*) at 4°C. The cells are resuspended in 0.4 volume of ice-cold TFB1.
5. Repeat the centrifugation. Resuspend the pellet in 1/25 volume of ice-cold TFB2.
6. The cells are aliquoted 100 μL per tube and shock-frozen in liquid nitrogen. The aliquots are stored at −80°C.

4. Notes

1. Several different type IIs enzymes can be used for construct assembly. We have, for example, tested the enzymes *Bsa*I, *Bpi*I, and *Esp*3I. For all three, restriction–ligation can be performed efficiently in ligase buffer from Promega. All three have a 6-bp recognition sequence and have a four-nucleotide cleavage site located one (*Bsa*I and *Esp*3I) or two nucleotides (*Bpi*I) away from the recognition sequence. Enzymes of the type of *Sap*I such as *Lgu*I can also be used efficiently in a restriction–ligation (12). These enzymes have a 7-bp recognition sequence, meaning that it occurs more rarely than 6-bp cutters, and therefore, fewer sites will have to be removed from sequences of interest to clone (discussed in paragraph Subheading 3.2). However, these enzymes have a three-nucleotide cleavage site, meaning that only 64 different sequences are available to choose from for use as recombination sites.
2. Shuffling does not necessarily require making an alignment of several sequences. A user might want to shuffle sequences with no homology at all; for example, test a range of different promoters and terminators for optimal expression of a coding sequence. In such a case, two recombination sites would be chosen, one between the promoter and the coding sequence, and one between the coding sequence and the terminator. DNA shuffling would require the following three sets of entry modules: (1) the first set containing as many promoter modules as desired, (2) the second set would contain only

one module consisting of the coding sequence, and (3) the third set would contain again several terminator modules.

3. Such requirement can easily be fulfilled if only few recombination sites are necessary within a gene. However, it becomes more difficult to fulfill when a large number of sites are used in one cloning experiment. However, if the choice of such sites cannot be avoided, shuffling will still be possible with nevertheless reasonable efficiency.
4. The minimal size of modules that we have tested is 38 bp (including the recombination sites, but excluding the flanking *Bsa*I recognition sites). In theory, a module needs to be long enough for the two strands to remain annealed under restriction–ligation conditions, in practice 37°C. This means that even smaller modules could be made, which would be useful if a user wants to focus his efforts on a very small region of a protein of interest.
5. For generation of entry clones, the steps consisting of PCR amplification, cloning, and sequencing can be avoided if the entry modules are simply ordered from a gene synthesis company. This can be a useful option if the user does not have access to a DNA sequence of interest, for example, a sequence from sequence databases obtained from a metagenomics project. If a sequence is ordered from a gene synthesis company, the fragment to be synthesized should be ordered directly with the appropriate type IIs restriction sites flanking the sequence of interest. Also, it is useful to make sure upon ordering that the synthesized fragment does not contain any internal *Bsa*I sites, and that the cloning vector in which the ordered DNA fragment will be cloned does not also contain additional sites for the type IIs enzyme chosen. Finally, it is also useful to make sure that the antibiotic resistance gene from the vector in which the synthesized fragment is cloned is different from the one from the vector that will be used for assembly of the shuffled library.
6. Basically, two primers overlapping an internal *Bsa*I site are made, one in each orientation, with a mismatch designed to introduce a silent mutation in the type IIs restriction site. Two separate PCRs are performed with primers designed to amplify the two halfs of the module. The PCR products are purified on a column, and a mix of both is used as a template for a second PCR performed using both flanking primers only (the two primers flanking the given module). This PCR is purified on a column, cloned, and sequenced.
7. For small modules of up to 80 nucleotides, PCR amplification does not necessarily require a DNA template. For example, two complementary primers can be ordered covering the entire sequence of the module (including the flanking type

IIs restriction sites). Both primers are annealed in water and directly used for blunt-end cloning in the cloning vector. For larger but still small modules, two overlapping primers can be ordered that are complementary at their 3′ end on a length of 20–25 nucleotides. A double-stranded DNA fragment can be obtained by performing a PCR with both primers without a template. In theory, one single PCR cycle should be sufficient, but using 35 cycles as for normal PCR also works well.

8. One restriction for blunt-end cloning is that the ends of the primer should not recreate a *Sma*I site (or any other blunt-end restriction site used for cloning) after ligation of the PCR product (i.e., the DNA fragment to subclone should not start with the sequence GGG or finish with CCC). A second restriction is that the fragment to be cloned should not contain an internal restriction site for the enzyme used for cloning. If this is the case, another enzyme should be chosen for cloning, for example, *Eco*RV (and a cloning vector containing a unique *Eco*RV site in the polylinker should be used).
9. The presence of a *Bsa*I site in the vector backbone of the entry modules does not prevent from using them for performing Golden Gate shuffling, since plasmids containing the final shuffled sequences should not contain this vector backbone. However, the presence of such a site in all entry constructs would lead to continuous ligation and redigestion at this site, which would unnecessarily consume some ATP from the ligation mix, at the expense of the desired ligation events.
10. The widely used pUC19 vector also contains a *Bsa*I site in the ampicillin-resistance gene. A simple strategy, enzymatic inverse PCR (13), can be used to eliminate the internal *Bsa*I site in pUC19. The entire plasmid can be amplified with two primers overlapping with the *Bsa*I site: primers bsarem1 (ttt ggtctc a ggtt ctcgcggtatcattgcagc) and bsarem2 (ttt ggtctc a aacc acgctcaccggctccag). These primers are designed to introduce a single silent nucleotide mutation in the *Bsa*I recognition site in the vector. The primers are themselves flanked by two *Bsa*I restriction sites that form two compatible overhangs after *Bsa*I enzyme digestion. After amplification of the entire plasmid with both primers, the PCR is purified with a column (to remove remaining polymerase and nucleotides). The linear fragment is subjected to restriction–ligation using *Bsa*I and ligase, and transformed in *E. coli*.
11. In practice, if all module plasmids and the vector have approximately the same size (4–5 kb), simply adding 100 ng of DNA of each module set and of the vector will work relatively well. However, if plasmids with widely different sizes are used, calculating an equimolar amount should provide a higher cloning efficiency. The following formula (from the NEB catalog) can

be used: 1 μg of a 1,000-bp DNA fragment corresponds to 1.52 pmol.

12. We have found that both types of programs work well when high concentration ligase is used, but both programs can be tested in parallel by the users to optimize ligation efficiency.
13. Any other *E. coli* strain can also be used. If higher transformation efficiency is required, the restriction–ligation mix can be transformed in electrocompetent *E. coli* cells. In this case, DNA from the restriction–ligation mix should first be ethanol-precipitated and resuspended in 10 μL of water.

Acknowledgments

The authors would like to thank Dr. Stefan Werner for critical reading of this manuscript.

References

1. Roberts, R. J. (2005) How restriction enzymes became the workhorses of molecular biology. *Proc. Natl. Acad. Sci. USA* **102**, 5905–5908.
2. Katzen, F. (2007) Gateway® recombinational cloning: a biological operating system. *Expert Opin. Drug Discov.* **2**, 571–589.
3. Engler, C., Kandzia, R., and Marillonnet, S. (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, e347.
4. Engler, C., Gruetzner, R., Kandzia, R., and Marillonnet, S. (2009) Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One* **4**, e5553.
5. Lebedenko, E. N., Birikh, K. R., Plutalov, O. V., and Berlin, Y. A. (1991) Method of artificial DNA splicing by directed ligation (SDL). *Nucleic Acids Res.* **19**, 6757–6761.
6. Szybalski, W., Kim, S. C., Hasan, N., and Podhajska, A. J. (1991) Class-IIS restriction enzymes – a review. *Gene* **100**, 13–26.
7. Berlin, Y. A. (1999) DNA splicing by directed ligation (SDL). *Curr. Issues Mol. Biol.* **1**, 21–30.
8. Lu, Q. (2005) Seamless cloning and gene fusion. *Trends Biotechnol.* **23**, 199–207.
9. Horton, R. M., Ho, S. N., Pullen, J. K., Hunt, H. D., Cai, Z., and Pease, L. R. (1990) Gene splicing by overlap extension. *Biotechniques* **8**, 528–535.
10. Bolchi, A., Ottonello, S., and Petrucco, S. (2005) A general one-step method for the cloning of PCR products. *Biotechnol. Appl. Biochem.* **42**, 205–209.
11. Liu, Z. G., and Schwartz, L. M. (1992) An efficient method for blunt-end ligation of PCR products. *Biotechniques* **12**, 28–30.
12. Kotera, I., and Nagai, T. (2008) A high-throughput and single-tube recombination of crude PCR products using a DNA polymerase inhibitor and type IIS restriction enzyme. *J. Biotechnol.* **137**, 1–7.
13. Stemmer, W. P., and Morris, S. K. (1992) Enzymatic inverse PCR: a restriction site independent, single-fragment method for high-efficiency, site-directed mutagenesis. *Biotechniques* **13**, 214–220.