GoldBricks: An improved cloning strategy that combines features of Golden Gate and Biobricks for better efficiency and usability

Vishalsingh R Chaudhari, Maureen R Hanson*.

Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853, USA.

Tel: +1 (607)-254-4833   Email: mrh5@cornell.edu

ABSTRACT

With increasing complexity of expression studies and the repertoire of characterized sequences, combinatorial cloning has become a common necessity. Techniques like Biobricks and Golden Gate aim to standardize and speed up the process of cloning large constructs while enabling sharing of resources. The Biobricks format provides a simplified and flexible approach to endless assembly with a compact library and useful intermediates but is a slow process, joining only two parts in a cycle. Golden Gate improves upon the speed with use of TypeIIS enzymes and joins several parts in a cycle but requires a larger library of parts and logistical inefficiencies scale up significantly in the multigene format. We present here a method that provides improvement over these techniques by combining their features. By using Type IIS enzymes in a format like Biobricks, we have enabled a faster and efficient assembly with reduced scarring, which performs at a similarly fast pace as Golden Gate, but significantly reduces library size and user input. Additionally, this method enables faster assembly of operon-style constructs, a feature requiring extensive workaround in Golden Gate. Our format allows such inclusions resulting in faster and more efficient assembly.

KEYWORDS: Biobricks, Golden Gate, Type IIS enzymes, operon cloning, bacteria, chloroplast

1. INTRODUCTION

Synthetic biology, often starting with coding complex designs in the form of DNA constructs, aspires to make or modulate aspects of biology. With increasingly complex schemes in synthetic biology and the growing repertoire of characterized expression elements (commonly referred to as parts), advanced cloning methods frequently become essential for generating combinatorial or multigene constructs.

There are a variety of cloning approaches to choose from that enable speedy and specific assembly of multiple segments of DNA in a short time (1). Some of these strategies also aim to improve efficiency by enabling reutilisation and sharing of the parts across the community, a vision which is supported by iGEM, Biobricks Foundation and several Biofoundries across the globe (2-4). A vetted library of parts is useful not only for quick turnaround times and low cost but also enables community science and labs in developing countries.

Amongst the approaches that enable parts to be reused, Biobricks and Golden Gate are two popular strategies for assembling short DNA fragments into a large construct, with many iterations published (1). The Biobricks cloning strategy was one of the initial attempts to bring standardization to traditional cloning methods (3). It defines a Biobrick part as a DNA sequence flanked by two specific restriction enzymes on each side, one of which produces a compatible overhang with the restriction site on the other end (5). This strategy enables the combining of two parts in the desired order and the product itself becomes a Biobrick part, hence enabling endless assembly with the addition of one part to the front or rear in each cycle. This makes it a fairly simple and flexible approach with a compact library where a sequence of interest (SOI) needs to be represented only once. Although this method provides the advantage of a standard design for endless and specific assembly, it has a few caveats. Notably, each joining event also produces an eight-base pair scar sequence from the disabled mixed restriction sites, which can be problematic for fusion protein assembly (6). And it is a relatively slower assembly method,

restricted to the addition of one part in each cycle. Furthermore, the four restriction enzymes present on the flank cannot be present in the SOI itself, necessitating domestication of such sites if present.

The Golden Gate cloning strategy overcomes the speed limitations of Biobricks by using BsaI, a Type IIS restriction enzyme (7). Type IIS restriction enzymes cut outside the recognition sequence and provide the opportunity to choose the overhang sequence, with the advantage of making it non-palindromic. A non-palindromic sequence prevents self-ligation of the part and can greatly improve the efficiency of end-joining reaction, especially when combining many parts in one cycle. With these two features, Golden Gate strategy enables multi-part assembly, including fusion proteins, while significantly reducing scar sequences (Figure 1). An optimized choice of overhang sequences and ligation protocol has demonstrated the assembly of up to 24 parts in one cycle with greater than 90% accuracy enabling rapid assembly of large constructs (8). However, this strategy does not allow addition of more parts to the final construct, i.e., endless assembly.

To enable extension of assembly, maintain universal compatibility of parts, and sharing across labs, a standard for overhang sequence was formulated for Golden Gate, named MoClo, with the addition of one more Type IIS enzyme BpiI (9, 10). In the MoClo standard, a level 0 part is a basic unit or plasmid that hosts a SOI like a promoter, UTR, CDS, etc. Such level 0 parts are combined in level 1 Assembly, which can take one part of any number of compatible types (11, 12). This assembly frequently represents one transcriptional unit. The level 1 assembly host also determines the position of the transcriptional unit in the next round of super assembly of multiple level 1 transcriptional units into a level 2 multigene construct. This approach has a few limitations. Level 1 assemblies cannot have multiple parts of the same type and hence disallow operons out of the box. Secondly, level 1 parts are preset for their order in level 2 assembly and this order cannot be changed, limiting their reuse or sharing. Finally, it necessitates multiple level 0 parts of the same element for each type of fusion partner. This problem and the previous issue combined usually result in a much larger library of parts

and intermediate plasmids that are not very useful. The GoldenBraid 2.0 standard was introduced for

Golden Gate assembly to reduce the problem of fixed intermediates in MoClo; however, it requires

additional cycles of cloning and 3 restriction enzymes, which leads to a higher domestication load (13).

There are also several polymerase-based methods, most prominently Gibson assembly and its

variations (1, 14). These techniques, however, do not generate parts that are easily reusable and have

significant one-time use investment of resources. They often require resource investment in designing

and ordering specific primers and subsequently validating constructs for the absence of PCR-based

errors. Additionally, most PCR-based methods have issues with repetitive sequences, very small

sequences, sequences with strong secondary structures, and often non-specific amplification. Some of

these issues are frequently encountered in regulatory sequences.

With advances in synthetic biology, many regulatory elements are characterized, and the

repertoire of SOIs is constantly increasing. Standardized protocols and several efforts to contribute such

validated parts hold promise for improving speed while cutting down significantly on the costs making

such experiments accessible for a larger community (15-18). Our interest in improving multi-gene

cloning is derived from our effort to engineer chloroplasts, which are lucrative sites for synthetic

biology, both for transgene expression and metabolic tweaking, and are now targets for expression of

large pathways (19-23). Current techniques like Biobricks fall short because of their slow speed when it

comes to assembling a large number of genes. Golden Gate's speed is best utilized when making one

transcription unit, but not an operon, from several fragments. In a multigene and combinatorial

scenario, it quickly becomes logistically complex. Considering the limited variety of well-defined

chloroplast regulatory elements across plant species, researchers will frequently have to resort to

optimizing their constructs in operon formats and combinatorial trials for such endeavors. Next-

generation efforts to improve or add novel metabolite pathways linked to food security or biotech

applications are requiring increasingly complex expression systems.

Here we are presenting a strategy, GoldBricks, that strikes a balance between simplicity and flexibility of Biobricks and the speed of Golden Gate approaches by including a Type IIS enzyme in the Biobrick. This strategy greatly improves upon the speed of Biobricks and at the same time has advantages over Golden Gate by enabling operon-based constructs with reduced library size and user input. While this cloning method is an improved approach to do multigene and combinatorial cloning in most common practical scenarios, we believe it would be of exceptional interest in the engineering not only of chloroplast or bacterial operons, but would also be useful for constructing multi-gene assemblies.

2. MATERIALS AND METHODS

2.1 Bacterial strains, growth conditions and reagents

E. coli DH5α cells were used for cloning and transformation. Cells were grown in Luria-Bertani medium at 37˚C with ampicillin 100 μg mL^{-1} or kanamycin 50 μg mL^{-1} for selection. All the fast digest restriction enzymes and Fast alkaline phosphatase (Fast AP cat. #EF0652) used in this study were purchased from ThermoFisher. The molecular weight marker used in gels is Generuler 1 kb plus from ThermoFisher. PCR amplifications were performed using NEB Q5 High fidelity 2x Master mix with the primers listed in the supplementary Genbank files. Plasmid preparations were performed using the Macherey-Nagel Nucleospin plasmid kit. For DNA or PCR product purification, Macherey-Nagel Nucleospin Gel and PCR Clean-up kit were used. For ligation cloning, T4 ligase from Invitrogen was used. All reagents were used according to the manufacturer's protocol. For three-part and five-part cloning, parts were added in approximately equimolar ratios including the destination vector. We have found that an equal amount of plasmids digested provides a good approximation for part sizes smaller than 1kb.

2.2 Construction of parts and assembly

Parts were picked from our Goldbricks library that demonstrated a good range of sizes and had useful restriction sites to visually demonstrate the order of assembly by restriction digestion. These parts belong to our catalog for carboxysome expression in chloroplasts and consist of chloroplast regulatory elements and bacterial coding sequences The full sequence and primers used are provided as Genbank files in supplementary data. Tailed PCR primers were used to generate part amplicons with specific restrictions sites. These amplicons were purified and digested with NotI and MauBI and cloned into pStart-III to make the holding vector of the part. For the 3-parts demonstration of assembly, parts were excised from the holding vectors using restriction enzymes marked in a dashed box in Figure 2 and ligated to pStart-IV linearized with NotI and MauBI without part purification. For the 5-part assembly, parts were excised using the restriction enzymes marked in Figure 4A. For parts with one end phosphorylated, reactions were performed sequentially. First, the end to be dephosphorylated was treated with a corresponding restriction enzyme and Fast AP. After incubation, the enzyme mix was heat-inactivated. In the same mix, the second restriction enzyme was added and incubated. For parts with both ends dephosphorylated, both enzymes and Fast AP were added simultaneously. Digested parts were used without purification and used in a ligation-based assembly in equimolar concentration. The ligation mix was transformed into high-efficiency chemically competent DH5α cells (NEB catalog #C2987H) and Luria-Bertani agar plates with kanamycin (50 µg mL^{-1}) were used for selection.

2.3 Validation of assembly

For 3-part assembly, 8 colonies from the selection plate were randomly picked, cultured and plasmids were extracted. Plasmids obtained were digested using PagI, ScaI and MauBI and run on agarose gel electrophoresis for analysis. For 5-part assembly, colonies from the selection plates were checked by colony PCR using NEB OneTaq quick load 2X master mix and manufacturer's recommendation. The primer pair that was used flanked the NotI and MauBI sites. PCR positive colonies with appropriate size were further verified by restriction digestion to ensure components and their orientation.

3. RESULTS AND DISCUSSION

3.1 GoldBricks merges BioBricks and Golden Gate

The Goldbrick strategy uses a four-restriction enzyme scheme for defining a part. Like Biobricks there are two restriction sites on each side, where the inner flank is made up of TypeIIS sites and the outer flank is made of TypeII sites. Unlike BioBricks, where there are no part types, Goldbricks typifies parts based on overhangs and only compatible parts can be ligated. The use of Type IIS restriction enzymes provides an opportunity to use a non-palindromic overhang sequence, which prevents self-ligation of parts and improves efficiency and fidelity of end-joining reaction (Figure 1). This allows more parts to be included in the assembly. Our approach aims to reduce library size by avoiding redundant cloning of a SOI into different part types by using common overhangs whenever possible (Supplementary Figure 1A).

The required specificity in such a scenario is provided by TypeII sites and occasionally by end-specific phosphorylation, which is possible by the use of two TypeIIS enzymes. Furthermore, specificity of the order of parts and endless extensions are provided by the outer flank of TypeII enzymes. Biobricks assembly also produces 8bp scars at junctions, which are unsuitable for certain purposes such as fusion proteins or when sequence sanctity is mandatory. Our strategy reduces scar sequences and is on par with Golden Gate. Further, our method avoids the use of linkers, and associated scars, for inserting parts in most scenarios due to greater cross-compatibility of parts.

3.2 Format of GoldBricks

An entry vector with a part cloned in it forms the holding vector that is the basic unit of the library or collection of parts. The entry vector, pStart-III, is defined by NotI and MauBI sites and these two sites are the external flanking sites in all the part types (Figure 2A). Each part is flanked on the inner side by the two TypeIIS enzymes, Esp3I and Eco31I, which define the compatibility of ligation of each part. TypeIIS enzymes are oriented to cut towards the SOI (Figure 2B). A mix of universal overhangs and

compatible overhangs define the order of ligation. For example, a promoter with a 5'UTR will have a

Type IIS Eco31I site in its suffix that will generate a 5' TTAC overhang for complementarity to A*ATG* start

codon. To match, a coding sequence (CDS) part will have a Esp3I site in the prefix generating a 5'A*ATG*

overhang. The CDS part's Eco31I site in the suffix produces 5'AAGC, which is complementary to the

Terminator prefix Esp3I overhang, and so on. The full list of overhang compatibility of parts is listed in

Supplementary Figure 1A. Some less common part types (for eg. N-tag or C-tag) have overhangs that can

oligomerize; however, this potential problem can be easily remedied by reducing the number of parts in

a ligation cycle or by end specific phosphorylation.

3.3 Three-part assembly

Parts are excised from the holding vectors using a combination of two enzymes. As shown in Figure 2,

for an assembly of three parts, boxed sites are used to eject each part. Ligation has only one possible

transformable outcome and hence a high-fidelity assembly is achieved. This scheme further reduces

user effort by following optional recommendations. First, selection marker can be shuttled between

holding vectors and destination vector. When parts held in pStart-III with the ampicillin marker (blue

backbone in Figure 2) are assembled in destination vector of a different selection marker, pStart-IV with

a kanamycin marker (red backbone in Figure 2), the user can avoid purification of parts for ligation. This

is especially helpful with parts that are of very small size (<200bp). Alternatively, when antibiotic

shuttling is not possible due to mixed Level 0 and level 1 parts, the backbone can be degraded by MbiI or

other blunt cutters. Second, all restriction enzymes are heat-inactivable and heat-inactivated reaction

mixtures are directly used in the ligation reaction, simplifying and reducing user effort.

For the demonstration of three-part assembly, two 5'UTR parts and a CDS were ligated

according to the scheme (Figure 3A), a combination not possible in Golden Gate or Biobricks in one

cycle. This reaction used antibiotic shuttling and hence purification of digested parts was not necessary.

From the antibiotic selection plate, eight colonies were randomly inoculated and DNA isolated and digested to verify insertion. All the colonies showed the restriction pattern expected of correct assembly (Figure 3B). In our experience, this assembly has matched theoretical expectations and yields almost all colonies with valid constructs. As the resultant product is also a Goldbrick part, it can be utilized further for higher-level or multigene assembly by extending on either side, unlike MoClo. As MoClo level 1 assembly is defined for order and number of partners in the next assembly, it is not useful as a part but only as an intermediate. Goldbricks' intermediate assembly is as flexible and hence useful as the parts that made them, a feature akin to Biobricks.

3.4 Five-part assembly

A modified approach allows the assembly of a greater number of parts while including parts of the same type in one assembly. As parts are excised by using two different restriction enzymes, it is possible to specifically remove the phosphate group from one end. All the restriction enzymes and alkaline phosphatase use a common reaction buffer and can be heat-inactivated. Such digestion and phosphatase reactions can be achieved within a tube without the need for additional purification. The scheme for specific assembly of five parts is shown in Figure 4A, which marks the restriction enzyme used to eject parts from holding vectors and phosphatase-treated ends. Two parts require end-specific phosphorylation. Such a scenario will have four legitimate Type IIS ligations and two illegitimate ligations. The illegitimate ligations, where #3 prefixes #2 or #4 prefixes #3, result in untransformable outcomes or require at least 7 parts to circularize, making the likelihood extremely rare over 5-part assembly. Resultant colonies from the ligation were screened by colony PCR for insert sizes. Colonies with inserts almost always had the correct size (Supplementary Figure 2). Plasmids were isolated from the PCR-positive colonies and were further checked by restriction digestion for order of the assembly, which is always found to be correct (Figure 4B).

3.5 Further manipulations

As these assemblies preserve the terminal restriction sites, it is possible to extend the construct on either side. We combined 3 transcription units in a cycle using the scheme shown in Figure 5A. This reaction imposes a strict ligation condition. The digested parts were used without purification in the ligation reaction to extend 5P-A assembly. Since the 5P-B backbone also had the kanamycin resistance, it was degraded using MbiI. MbiI is a blunt end cutter that cuts at two sites and hence there is a very low probability of recircularizing. Six colonies were randomly selected without screening, DNA prepared and tested by restriction digestion (Figure 5B). Similar to three-part assembly, all were found to be correct assemblies. A few common scenarios and extensions of intermediate assemblies are described in Supplementary Figure 3. Transcription units can be fused in any desired order. This also allows parallel assembly of parts, which can be fused in a second round for reduced time of assembly. Even though parts are typified, a GoldBrick part can bind to more part types than Golden Gate (Figure 6, Supplementary Figure 3A); for example, the CDS can be prefixed by a 5'UTR or an N-tag and can be suffixed by a C-tag or a 3'UTR/Terminator. For certain non-standard scenarios, the cloning can still be achieved by using compatible restriction schemes and without linkers or recloning of SOI, for example, a fusion protein made of several pieces (Supplementary Figure 3F). Such scenarios are, however, presumed to be rare and can be easily accommodated most often without making new parts.

3.6 Comparison of cloning methods

We have demonstrated that a relatively simple and Biobrick-like scheme can effectively ligate multiple parts. Our method aims to ligate up to 5 parts in a cycle, which we believe is good enough to form a transcriptional unit or more under most common scenarios. We forfeit the ability to ligate a large number of parts in one cycle to allow compaction of the library, improve the flexibility of intermediates and reduce user effort. We would like to draw a comparison using hypothetical scenarios and show how

our technique occupies an optimal position for most practical aspects. First, consider the library size requirement to construct multigene or combinatorial constructs. Biobricks does not typify parts and hence each SOI needs to be cloned only once. This keeps the cost and size of a library at a minimum. As shown in Supplementary Figure 1, GoldBricks combines the part definitions to reduce the number of part types compared to the MoClo standard (9). A MoClo scheme of 15 part types can be reduced to 7 types in GoldBricks format, reducing redundant cloning of SOIs. Additionally, compared to the MoClo specification of 9 overhang sequences, Goldbricks specifies only three overhangs, showing the increased cross-compatibility of parts. An additional layer of traditional enzymes and capping the maximum number of parts in a reaction to five provides markedly improved flexibility and reduced library size. For a hypothetical scenario of combinatorial cloning of 4 constructs using 6 SOIs, Biobrick will need 6 parts, GoldBricks will need 8 parts and MoClo will need 12 parts (Supplementary Figure 4A). GoldBricks reduces the recloning of SOI in multiple types compared to MoClo, which in turn reduces the logistical requirements and complexity of the library of parts. GoldBricks enables operon-style constructs or concatemerization of parts (e.g., multimeric fusion proteins) without requiring linkers or unnecessary recloning of SOI.

Second, we need to consider the speed of the cloning in an operon format. We compare a hypothetical scenario of constructing a four-gene operon using GoldBricks and MoClo in Supplementary Figure 4B. As MoClo does not allow two parts of the same type in a cycle, such an assembly will need to be split into 4 level 1 assemblies, which are then combined in the next cycle to get the final construct (9, 11). On the other hand, Goldbricks is limited only by the number of parts in a cycle and similar constructs are achieved in the same amount of time but with less user input. Additionally, the GoldBricks intermediates have higher flexibility for reuse and do not have adapter/linker scars. Biobricks will take more time and more user input than either technique.

Third, we should evaluate the speed of cloning in the non-operon format. We compare a hypothetical scenario of combinatorial assembly of 7 genes in form of 7 transcriptional units (TU), where each TU is made of 4 parts (Supplementary Figure 4C). This represents a common cloning scenario for multigene nuclear expression constructs. In this scenario, MoClo is the quickest at accomplishing the given configuration in 4 days. GoldBricks takes 6 days and Biobricks takes 10 days. However, GoldBricks requires the least number of reactions to be set up, 10 vs. 50 for MoClo and vs. 27 in BioBricks. The intermediates generated here are capable of giving any combination of 7 genes in any order required. The number of intermediates is only 7 (1 per TU) for Biobricks and GoldBricks but is very high at 49 (7 per TU) for Golden Gate. Hence, GoldBricks presents a path of least effort and complexity at a slightly elevated requirement of time by generating more versatile intermediates. GoldenBraid 2.0, MOBIUS and LOOP methods have tried to address the inflexibility of level 1 assembly to organize into higher assembly by introducing variations in the collection of Level 1 holding plasmids (13, 24, 25). GoldenBraid 2.0 increases the reusability of level 1 assembly at cost of speed and increased domestication load (Table1). GoldenBraid drops down to combining two TUs in a cycle and with the requirement of occasional swapping of host-vector, the average is lower than 2 per cycle. MOBIUS ligates 4 TUs, partly relieves the domestication issue of GoldenBraid and adds visual indicators for tracking cloning, but it brings back rigid intermediates, linking sequence scars, and speeds slower than MoClo. LOOP also uses a high number of intermediates and its advantages become apparent only when combining several TUs. While these methods provide tradeoffs suitable for certain applications, for these iterations to be backward compatible, they have not resolved the issues inherent to the Golden Gate scheme, primarily lack of out-of-the-box compatibility for operons and a larger library due to repetitive cloning of SOI. Our method does not need specific Level 1 plasmids for organizing higher assemblies. Any vector with NotI/MauBI site can host at any stage.

GoldBricks uses a scheme of 4 enzymes, similar to BioBricks. However, the choice of enzymes does not cause a substantial change in the domestication requirement when compared with BioBricks and is significantly lower than Golden Gate versions that enable endless assembly (Table 1).

In summary, we present here a cloning strategy that is a hybrid approach of Biobricks and Golden Gate and brings in their best attributes. The GoldBricks strategy significantly improves the assembly speed compared to Biobricks. Like Biobricks, our strategy preserves the flexibility and idempotent nature of assemblies constructed. It also allows multiple parts of the same type in an assembly, hence significantly reducing the library size for constructing multigene operons compared to Golden Gate. By allowing a flexible partner choice, the parts and intermediates generated are more versatile leading to reduced library size and user input. These hybrid advantages are achieved by sacrificing the type-less structure of Biobricks and a very high number of unique part types that can be assembled in Golden Gate. Advantages gained by losing these features outweigh the disadvantages in most circumstances for general cloning and are especially useful for making multigene operons. Multigene operons are useful strategies for expressing a large number of genes in a compact format, especially for plastid gene expression studies, where the repertoire of well-characterized regulatory elements is small and reusing the same elements multiple times is prohibited by homologous recombination.

The major caveat of our strategy is that it is not backward compatible with the existing library of parts that are designed for Biobricks or Golden Gate, which is a common problem found amongst earlier iterations of those methods as well. Nevertheless, we believe the approach we present strikes a balance between speed and flexibility in cloning, especially for multi-gene operon kind of constructs.

Supplementary Data

Supplementary Data are available at SYNBIO Online.

14

Material Availability Statement

Unique materials presented in the manuscript will be available from the authors upon reasonable request and through a materials transfer agreement.

Data Availability Statement

DNA sequences for all the vectors, parts, primers and assemblies used in the paper are available at SYNBio online.

References

1. Chao,R., Yuan,Y. and Zhao,H., (2015) Recent advances in DNA assembly technologies. *FEMS Yeast Res* 15,1-9.

2. Casini,A., Storch,M., Baldwin,G.S. and Ellis,T., (2015) Bricks and blueprints: methods and standards for DNA assembly. *Nat Rev Mol Cell Biol* 16,568-576.

3. Smolke,C.D., (2009) Building outside of the box: iGEM and the BioBricks Foundation. *Nat Biotechnol* 27,1099-1102.

4.   Hillson,N. *et al.,* (2019) Building a global alliance of biofoundries. *Nat Commun* 10,2040.

5.   Knight,T., (2007) The BioBricks Foundation:BBFRFC10.

6.   Cai,Y., Wilson,M.L. and Peccoud,J., (2010) GenoCAD for iGEM: a grammatical approach to the design of standard-compliant constructs. *Nucleic Acids Res* 38,2637-2644.

7.   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.

8.   Potapov,V. *et al.,* (2018) Comprehensive profiling of four base overhang ligation fidelity by T4 DNA ligase and application to DNA assembly. *ACS Synth Biol* 7,2665-2674.

9.   Marillonnet,S. and Werner,S., (2020) Assembly of multigene constructs using the modular cloning system MoClo. *Methods Mol Biol* 2205,125-141.

10.   Werner,S., Engler,C., Weber,E., Gruetzner,R. and Marillonnet,S., (2012) Fast track assembly of multigene constructs using Golden Gate cloning and the MoClo system. *Bioeng Bugs* 3,38-43.

11.   Grutzner,R. and Marillonnet,S., (2020) Generation of MoClo standard parts using Golden Gate cloning. *Methods Mol Biol* 2205,107-123.

12.   Weber,E., Engler,C., Gruetzner,R., Werner,S. and Marillonnet,S., (2011) A modular cloning system for standardized assembly of multigene constructs. *PloS one* 6,e16765.

13.   Sarrion-Perdigones,A. *et al.,* (2013) GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. *Plant Physiol* 162,1618-1631.

14.   Gibson,D.G. *et al.,* (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6,343-345.

15.   Vafaee,Y., Staniek,A., Mancheno-Solano,M. and Warzecha,H., (2014) A modular cloning toolbox for the generation of chloroplast transformation vectors. *PloS one* 9,e110222.

16.   Lampropoulos,A. *et al.,* (2013) GreenGate---a novel,versatile,and efficient cloning system for plant transgenesis. *PloS one* 8,e83043.

Synthetic Biology

17.   Engler,C. *et al.,* (2014) A Golden Gate modular cloning toolbox for plants. *ACS Synth Biol* 3,839-843.

18.   Occhialini,A. *et al.,* (2019) MoChlo: A versatile,modular cloning toolbox for chloroplast biotechnology. *Plant Physiol* 179,943-957.

19.   Bock,R., (2015) Engineering plastid genomes: methods, tools, and applications in basic research and biotechnology. *Annual review of plant biology* 66,211-241.

20.   Lin,M.T. *et al.,* (2014) Beta-Carboxysomal proteins assemble into highly organized structures in Nicotiana chloroplasts. *Plant J* 79,1-12.

21.   Hanson,M.R., Lin,M.T., Carmo-Silva,A.E. and Parry,M.A., (2016) Towards engineering carboxysomes into C3 plants. *Plant J* 87,38-50.

22.   Rosenblueth,M. *et al.,* (2018) Nitrogen fixation in cereals. *Front Microbiol* 9,1794.

23.   Ivleva,N.B., Groat,J., Staub,J.M. and Stephens,M., (2016) Expression of active subunit of nitrogenase via integration into plant organelle genome. *PloS one* 11,e0160951.

24.   Andreou,A.I. and Nakayama,N., (2018) Mobius Assembly: A versatile Golden-Gate framework towards universal DNA assembly. *PloS one* 13,e0189892.

25.   Pollak,B. *et al.,* (2019) Loop assembly: a simple and open system for recursive fabrication of DNA circuits. *New Phytol* 222,628-640.

Figure Legends

Figure 1. Advantage of TypeIIS ligation: Sticky ends of TypeIIS enzymes can be designed to be non-palindromic, which prevents the issue of self-ligation responsible for several unintended byproducts.

Figure 2. A basic scheme for defining a GoldBrick part and a three-part assembly: (A) Here each part has a 4-enzyme flank and part is typified by the overhang sequence of the TypeIIS site. Each part has specific compatibility for the suffix and the prefix part. For an example of the 3-part assembly, the parts can be excised using sites marked in the box and subsequently ligated to give a strictly ordered product. If the destination vector has different antibiotic marker, parts do not need to be purified for assembly. (B) sequence format to show directionality of TypeIIS enzymes.

Figure 3. Demonstration of three-part assembly: (A) A Three-part assembly was performed using parts U1, CDS1 and U2 that were 62bp, 309 bp and 112 bp long, respectively. After transformation, 8 colonies were selected randomly and DNA was prepared. (B) DNA was digested using PagI, ScaI and MauBI, which would give bands of the mentioned sizes from successful assemblies. The agarose gel of the digested DNA of empty pStart-IV (lane C) and eight randomly selected colonies (lane 1-8) demonstrates that all clones harbored the correct assembly. Specific fragment sizes are marked in blue.

Figure 4. Demonstration of five-part assembly: (A) Schematic for two assemblies 5P-A and 5P-B is shown. Each part was cut in its holding vector using the enzyme shown in the figure (N= NotI, Ec=Eco31I, Es=Esp3I and M=MauBI). Part ends that were dephosphorylated are marked by *. These 5 parts were then ligated into pStart-IV and selected on kanamycin. (B) 6 colonies that were PCR positive for inserts were randomly picked and DNA prepared. These were validated for correct assembly using restriction digestion according to the schematic. All colonies showed the correct insertion. Specific fragment sizes are marked in blue.

Figure 5. Assembly extension: Level 1 assemblies 5P-A, 5P-B and a tobramycin resistance cassette were combined using the restriction enzyme scheme shown in (A). Sites marked with '*' were dephosphorylated. 6 random clones of Super Assembly A (SA-A) were checked using restriction digestion according to schematic (B) and all were found to have correct assemblies. Specific fragment sizes are marked in blue.

Table 1. Domestication requirement of GoldBricks and comparison with other techniques (A) Table of the 5 restriction sites used in the current strategy and the length of the restriction site recognition sequence. (B) The combined frequency of all the restriction sites used in a particular strategy in a sample genome of the tobacco chloroplast and *E. coli*. Note *: Of the four cloning methods listed here, Golden Gate cloning is not an endless cloning method.
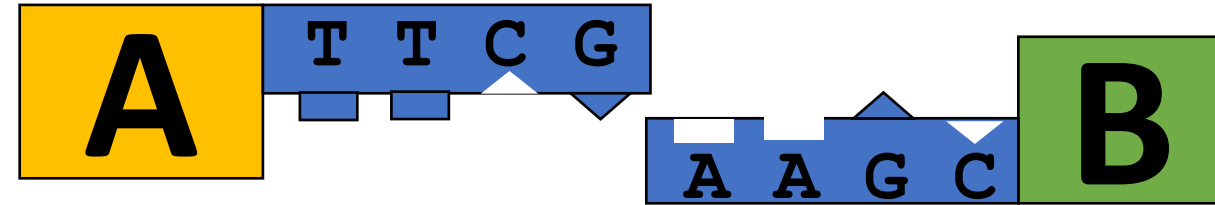
TypeII ligation

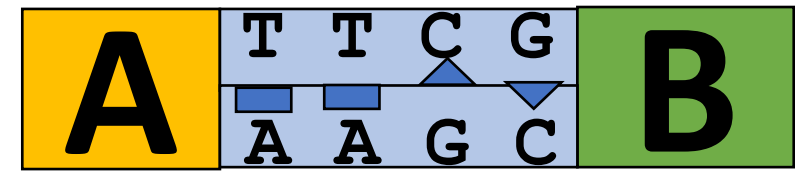# TypeII-S ligation

Palindromic overhangs
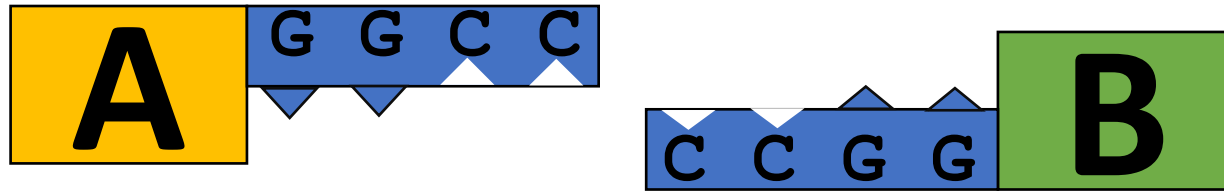
Non-palindromic overhangs
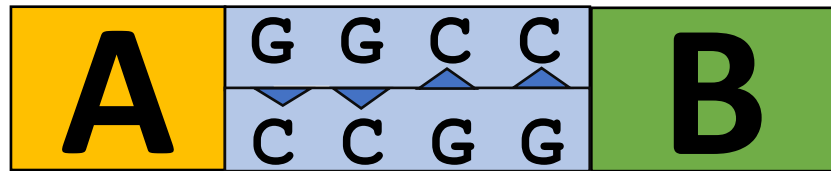
Ligation

Desired product

Self-ligated side products

No self ligated side products

GoldBricks Assembly scheme

A

NotI   MauBI

Destination Vector

Esp3I -(TypeIIS)- Eco31I
NotI   MauBI

GoldBrick Part 1

3 parts Assembly using Boxed sites

Esp3I
NotI

Eco31I
MauBI

GoldBrick assembly

Esp3I -(TypeIIS)- Eco31I
NotI   MauBI

GoldBrick Part 2

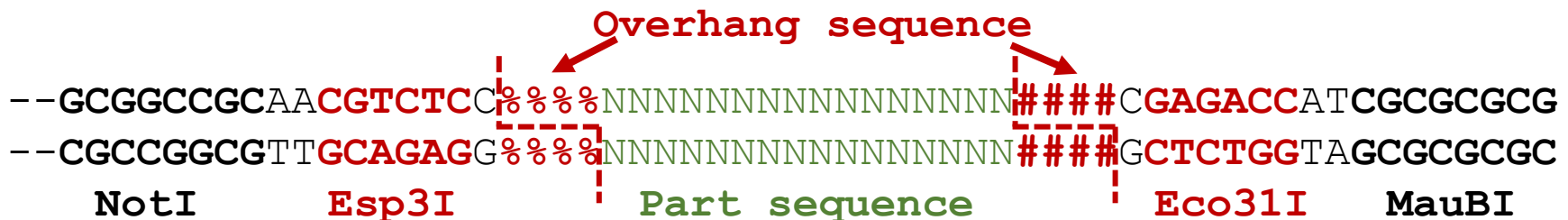Esp3I ---(TypeIIS)--- Eco31I
NotI   MauBI

Goldbrick Part 3

Library of parts

B

Overhang sequence

```
--GCGGCCGCAACGTCTCC%%%%NNNNNNNNNNNNNNNNNNNN###CGAGACCATCGCGCGCG
--CGCCGGCGTTGCAGAGG%%%%NNNNNNNNNNNNNNNNNNNN####GCTCTGGTAGCGCGCGC
  NotI        Esp3I       Part sequence         Eco31I      MauBI
```

A

Destination Vector   **Parts from Library**

Digest part plasmids, inactivate restriction enzymes and ligate. **No purification necessary.**

Transformation and miniprep
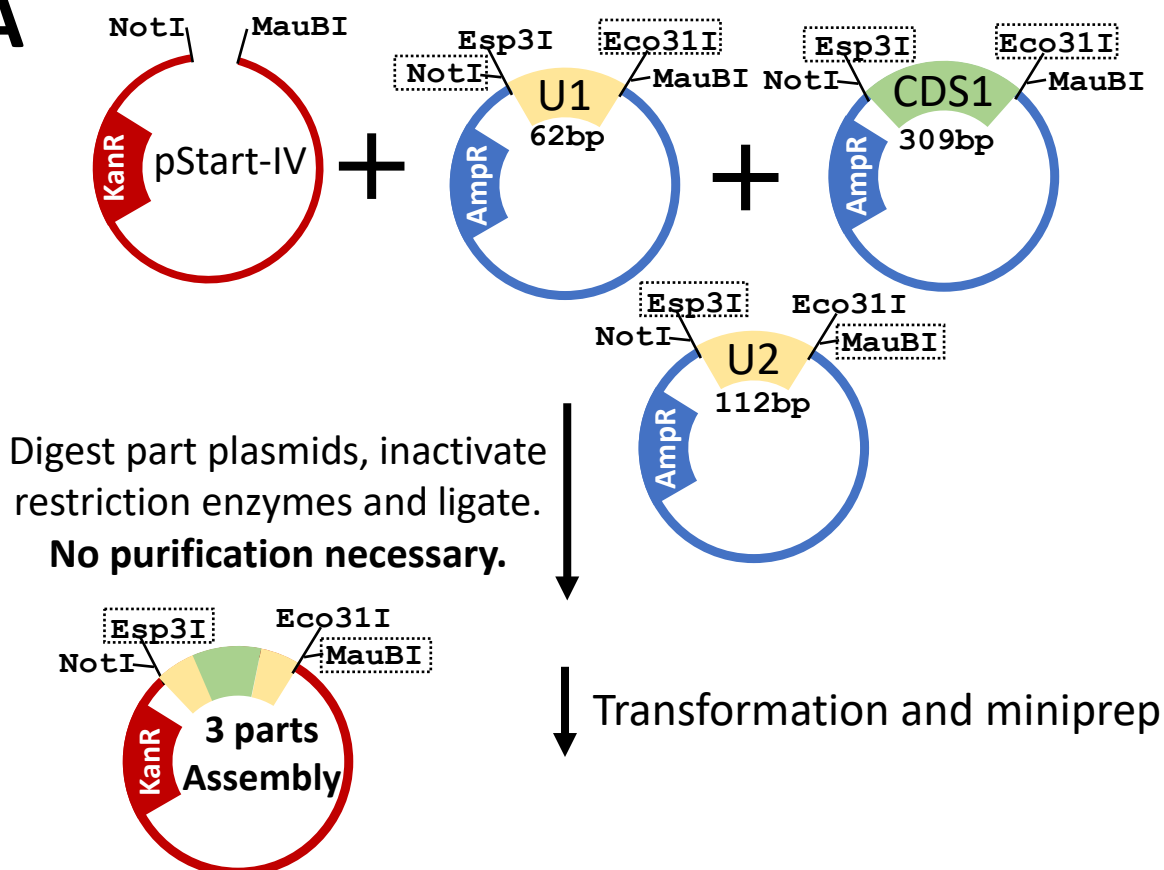
B

Testing of 8 random clones by restriction digestion

1 Kb

0.4Kb

0.2Kb

Synthetic Biology

A

Assembly 5P-A

N — N* — Ec* | UTR1 | Es — Ec* | CDS1 | Es — Ec | UTR2 | Es* — Ec | CDS2 | Es* — M* — M | TER1 |

Part order #1 #2 #3 #4 #5

Assembly 5P-B

N — N* — Ec* | PRO | Es — Ec* | CDS3 | Es — Ec | UTR1 | Es* — Ec | CDS4 | Es* — M* — M | TER2 |

Part order #1 #2 #3 #4 #5

B

5P-A

PagI — PagI — PagI | KanR | UTR1 | CDS1 | UTR2 | CDS2 | NdeI | TER1 | MauBI

839b — 624b — 318b — 533b — 1173b

5P-B

PagI — PagI — PagI | KanR | PRO | EcoRI | CDS3 | BamHI | UTR1 | CDS4 | BamHI | TER1 | PagI

891b — 624b — 318b — 249b — 1540b — 533b — 185b

5P-A digested with PagI, NdeI, MauBI

1 2 3 M 4 5 6

1.5Kb
0.5Kb

5P-B digested with PagI, EcoRI, BamHI

1 2 3 M 4 5 6

1.5Kb
0.5Kb
0.1Kb

A

NotI — Esp3I   Eco31I* / MauBI

5P-A   KanR

+

MbiI — NotI   Esp3I   Eco31I / MauBI — MbiI

KanR   **5P-B**

Backbone degraded by blunt cutter

+

Esp3I* / NotI   Eco31I / MauBI

AmpR   **Tob^R**

→ **Selection on Kanamycin** →

NotI — Esp3I   Eco31I / MauBI

KanR   **Super Assembly A**

Synthetic Biology

B

PagI   PagI   PagI   NdeI   PagI   PagI

KanR   5P-A   5P-B   Tobramycin^R

←624b→ ←318b→ ←533b→ ←——3557b——→ ←——1822b——→ ←1124b→

SA-A digested with PagI, NdeI

1   2   3   M   4   5   6

5Kb

1.5Kb

0.5Kb

Domestication requirements of the restriction enzyme scheme

A

Enzyme	Type	Restriction site length
NotI	TypeII	8 bp
MauBI	TypeII	8 bp
Esp3I	TypeII-S	6 bp
Eco31I	TypeII-S	6 bp

B

Strategy	Frequency of restriction sites in NtCp genome (150kb)	Frequency of restriction sites in *E.coli* genome
BioBricks	208	1448
Goldbricks	53	1581
GoldenBraid 2.0	75 (100 for MoClo)	6798 (3136 for MoClo)
Golden Gate*	29	261