ACS Synthetic Biology

Research Article

pubs.acs.org/synthbio

Yeast Golden Gate (yGG) for the Efficient Assembly of *S. cerevisiae* Transcription Units

Neta Agmon,[†,∥] Leslie A. Mitchell,[†,∥] Yizhi Cai,[‡,§] Shigehito Ikushima,[‡] James Chuang,[‡] Allen Zheng,[‡] Woo-Jin Choi,[‡] J. Andrew Martin,[†] Katrina Caravelli,[‡] Giovanni Stracquadanio,[‡] and Jef D. Boeke*,[†]

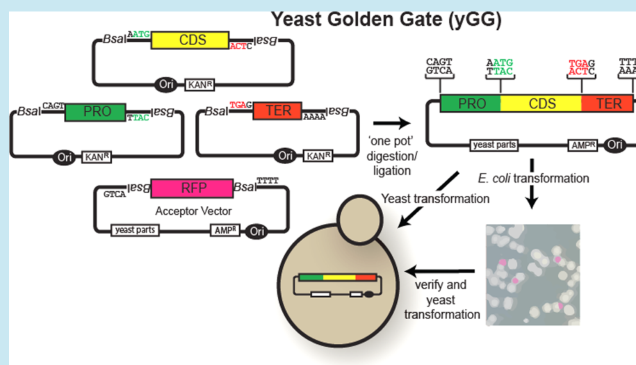
[†]Institute for Systems Genetics, Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, 550 First Avenue, New York, New York 10016, United States

[‡]High Throughput Biology Center, School of Medicine, Johns Hopkins University, Edward D. Miller Research Building, 733 North Broadway, Baltimore, Maryland 21205, United States

[§]School of Biological Sciences, University of Edinburgh, The King's Buildings, Edinburgh, EH9 3JR, United Kingdom

Supporting Information

ABSTRACT: We have adapted the Golden Gate DNA assembly method to the assembly of transcription units (TUs) for the yeast *Saccharomyces cerevisiae*, in a method we call yeast Golden Gate (yGG). yGG allows for the easy assembly of TUs consisting of promoters (PRO), coding sequences (CDS), and terminators (TER). Carefully designed overhangs exposed by digestion with a type IIS restriction enzyme enable virtually seamless assembly of TUs that, in principle, contain all of the information necessary to express a gene of interest in yeast. We also describe a versatile set of yGG acceptor vectors to be used for TU assembly. These vectors can be used for low or high copy expression of assembled TUs or integration into carefully selected innocuous genomic loci. yGG provides synthetic biologists and yeast geneticists with an efficient new means by which to engineer *S. cerevisiae*.



KEYWORDS: *synthetic biology, DNA assembly, S. cerevisiae, transcription unit, yeast Golden Gate*

The yeast *Saccharomyces cerevisiae* is a prominent model organism and a highly valued chassis in the field of synthetic biology. In this space, metabolic engineering is a major focus, as the expression of one or more heterologous enzymes can transform *S. cerevisiae* into a tiny cellular factory. The most well-known example of this to date is the engineering of *S. cerevisiae* to produce commercially relevant concentrations of artemisinic acid, a precursor to the antimalarial drug artemisinin.[1] These metabolic engineering projects require both the introduction of heterologous genes whose expression levels are finely tuned and the redirection of endogenous biosynthetic pathways via modification of native genes. The development of tools to aid in the construction and manipulation of both native and non-native genes for expression in *S. cerevisiae* thus facilitates metabolic engineering and synthetic biology in yeast.

Typical yeast protein coding genes have a relatively simple anatomy, due, in part, to the compact structure of the *S. cerevisiae* genome. Promoters are short, generally extending only ~500 bp upstream of the start codon. Only ~20% of promoters in the yeast genome contain TATA boxes.[2,3] On average, native coding sequences (CDS) are ~1 kb long, and less than 5% contain introns. Sequences associated with 3′-end formation, which typically extend ~200 bp downstream of the stop codon, are usually AT-rich and contain information for both transcriptional termination and 3′-end processing.[4] The simple structure of yeast genes means that expression of non-native proteins in yeast is easily achieved by encoding the CDS of interest between a promoter and terminator that can function in *S. cerevisiae*. Tuning of CDS expression level can then be accomplished by varying the promoter and terminator sequences, changing the gene copy number (e.g., high or low copy plasmid), or altering the genomic locus in which the gene is integrated.

Here, we present a versatile new method that enables modular assembly of the three major classes of yeast gene parts into transcription units (TUs) for expression in yeast. Dubbed yeast Golden Gate, or yGG, this method relies on Golden Gate assembly[5,6] and type IIS restriction enzymes. A major simplifying principle of our strategy is to not decompose 5′ sequences into UAS, promoter, and 5′ UTR subparts and, similarly, to specify single, rather than multiple, 3′-end-forming parts. This is distinct from the standard strategy employed in

Received: December 30, 2014
Published: March 10, 2015

 © 2015 American Chemical Society

DOI: 10.1021/sb500372z
ACS Synth. Biol. 2015, 4, 853−859

Table 1. Standardized Prefix and Suffix Sequences for yGG[a]

	*Bsa*I prefixes	*Bsa*I suffixes	*Bsm*BI prefixes	*Bsm*BI suffixes
PRO	**GGTCTC**ACAGT	AATGC**GAGACC**	**CGTCTC**ACAGT	AATGC**GAGACG**
CDS	**GGTCTC**AAATG	TGAGC**GAGACC**	**CGTCTC**AAATG	TGAGC**GAGACG**
TER	**GGTCTC**ATGAG	TTTTC**GAGACC**	**CGTCTC**ATGAG	TTTTC**GAGACG**

[a]Bold 6 bp sequences are recognition sites; underlined 4 base sequences are overhang sites. All sequences are written 5′ to 3′ on the top strand of the final part.

bacteria, e.g., by the BioFab.[7−9] Using specifically chosen designer overhangs, we generate virtually seamless TUs by yGG. We have built a series of custom yGG acceptor vectors in which TU assembly can be distinguished from unmodified acceptor vector by white/red colony screening. Importantly, the acceptor vectors enable copy number control, can be used for integration into one of several genomic loci, and are entirely compatible with library construction. As a variant of yGG, we also present a way to C-terminally tag CDSs. We also discuss potential future applications of yGG.

RESULTS AND DISCUSSION

Construction of Transcription Units (TUs) for Expression in *S. cerevisiae*. We define a TU to consist of three parts: a UAS/promoter/5′UTR (PRO), a coding sequence (CDS), and a 3′ UTR/polyadenylation signal/terminator (TER). To enable directional assembly of TUs, we assign specific prefix and suffix sequences to each of the three parts that encode inwardly facing Type IIS restriction sites, typically, *Bsa*I or *Bsm*BI, and a 4 bp designer overhang separated from the recognition sequence by a single base to accommodate the offset cutting by the enzyme (Table 1). These Type IIS RE sites are oriented such that they are eliminated upon digestion, exposing designer overhangs as follows: 5′-CAGT-PRO-AATG-3′, 5′-AATG-CDS-TGAG-3′, and 5′-TGAG-TER-TTTT-3′, respectively (Figure 1). (Note that all overhangs are listed here as top-strand sequences for clarity.) The overhangs selected are known to be highly compatible with regulated gene expression and represent the smallest possible scars, as they exploit the natural punctuation marks ATG and TGA. Specifically, the AATG overhang between the PRO and CDS provides the ATG start codon. Importantly, this overhang provides a favorable context for gene expression, as most well-expressed yeast genes have their ATG initiation codons preceded by one or more A's.[10] Additionally, the TGAG overhang at the CDS/TER junction provides a universal TGA stop codon. PRO, CDS, and TER parts flanked by the appropriate prefix and suffix sequences are cloned into kanamycin resistant vectors that do not encode genetic information for replication in yeast. The subsequent yGG assembly of TUs is performed using an ampicillin resistant acceptor vector (Figure 1) in a one-pot digestion-ligation reaction (see Methods). The parental acceptor vector encodes a red fluorescent protein (RFP) gene with *Escherichia coli* promoter and terminator sequences. Thus, following *E. coli* transformation, white/red colony color screening can be used to distinguish clones encoding putative TU assemblies from those containing unmodified parental vector.

Acceptor Vectors. We have constructed a series of acceptor vectors with multiple markers and applications for use in yGG assembly (Table 2). To facilitate TU assembly, yGG acceptor vectors lack *Bsa*I and/or *Bsm*BI restriction sites except for two outwardly facing sites flanking the RFP cassette described above. The overhangs generated following *Bsa*I (or *Bsm*BI) digestion are compatible with receiving the 5′ overhang of the PRO part (CAGT) and the 3′ overhang of the TER part (TTTT). Subsequent to assembly, these vectors permit direct transformation of TUs into yeast cells.

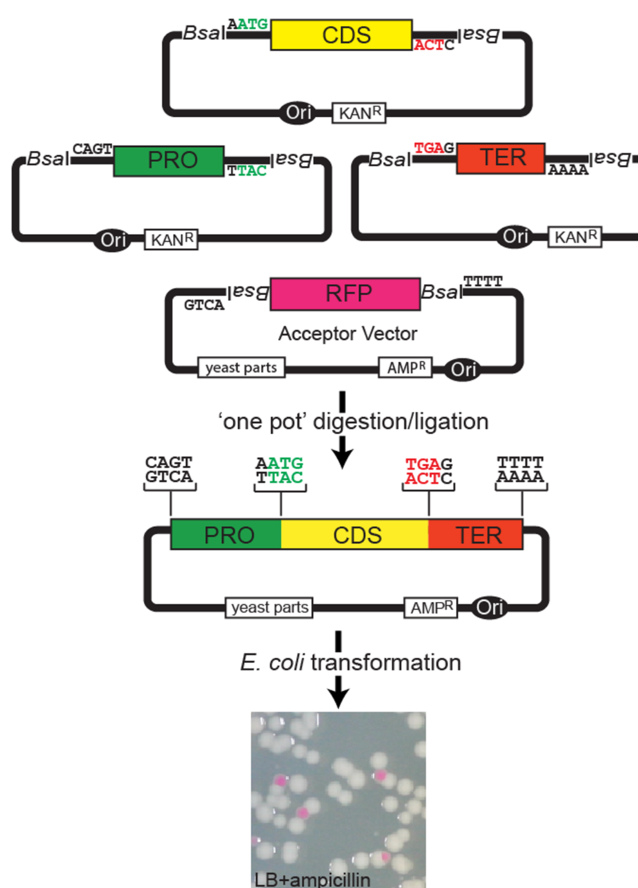


Figure 1. One-pot yGG assembly. PRO, CDS, and TER parts flanked by the appropriate prefix and suffix sequences are cloned into kanamycin resistant vectors. For the one-pot digestion−ligation reaction, cloned parts are mixed in equimolar ratio with ampicillin resistant acceptor vector for subsequent yGG assembly of TUs. The parental acceptor vector encodes a red fluorescent protein (RFP) gene with *E. coli* promoter and terminator sequences. Following *E. coli* transformation, white/red screening can be used to distinguish clones encoding putative TU assemblies as compared to unmodified parental vector.

The first two sets of yGG acceptor vectors are intended for independent replication and segregation once transformed into *S. cerevisiae* and derive from the popular pRS series of yeast shuttle vectors pRS41X and pRS42X.[11] These vectors encode either a centromere/autonomously replicating sequence (CEN/ARS) combination (pRS41X; single copy) or 2 micron parts (pRS42X; high copy), in addition to a selectable marker for yeast, plus selection and replication parts for bacteria (ampicillin resistance and a replication origin; Figure 2).

DOI: 10.1021/sb500372z
ACS Synth. Biol. 2015, 4, 853−859

Table 2. Acceptor Vectors for yGG[a]

plasmid name	yeast marker (organism)	yeast replication parts	yeast integrative locus	*E. coli* marker	2° RE[b]	plasmid no.	Addgene ID
CEN/ARS (Low Copy)							
pAV113	*HIS3* (*Sc*)	CEN/ARS	n/a	Amp	n/a	pLM108	63180
pAV114	*TRP1* (*Sc*)	CEN/ARS	n/a	Amp	n/a	pLM264	63181
pAV115	*LEU2* (*Sc*)	CEN/ARS	n/a	Amp	n/a	pLM109	63182
pAV116	*URA3* (*Sc*)	CEN/ARS	n/a	Amp	n/a	pLM304	63183
pAV11K	*KanMX*	CEN/ARS	n/a	Amp	n/a	pLM200	63184
pAV113.loxP[c]	*HIS3* (*Sc*)	CEN/ARS	n/a	Amp	n/a	pJC081	63186
pAV115.loxP[c]	*LEU2* (*Sc*)	CEN/ARS	n/a	Amp	n/a	pJC082	63187
pAV116.loxP[c]	*URA3* (*Sc*)	CEN/ARS	n/a	Amp	n/a	pJC106	63188
2 μm (μ) (High Copy)							
pAV123	*HIS3* (*Sc*)	2 μ	n/a	Amp	n/a	pAM090	63189
pAV124	*TRP1* (*Sc*)	2 μ	n/a	Amp	n/a	pLM266	63190
pAV125	*LEU2* (*Sc*)	2 μ	n/a	Amp	n/a	pLM270	63191
pAV126	*URA3* (*Sc*)	2 μ	n/a	Amp	n/a	pAM078	63192
Integrative							
pAV103	*HIS3* (*Sc*)	n/a	*HIS3*	Amp	n/a	pLM346	63193
pAV104	*TRP1* (*Sc*)	n/a	*TRP1*	Amp	n/a	pLM262	63194
pAV105	*LEU2* (*Sc*)	n/a	*LEU2*	Amp	n/a	pLM107	63195
pAV106	*URA3* (*Sc*)	n/a	*URA3*	Amp	n/a	pLM302	63196
pAV10.F3	*HIS3* (*Sc*)	n/a	*chrVI*: 97873−98803	Cam	*Not*I or *Bci*VI	pSIB055	63199
pAV10.F3.loxP[c]	*HIS5* (*Sp*)	n/a	*chrVI*: 97873−98803	Amp	*Not*I	pSIB581	63200
pAV10.F6.loxP[c]	*URA3* (*Kl*)	n/a	*chrVI*: 97873−98803	Amp	*Not*I or *Bci*VI	pSIB582	63201
pAV10.K3.loxP[c]	*HIS5* (*Sp*)	n/a	*YKL162C-A*	Amp	*Not*I	pSIB584	63202
pAV10.K6.loxP[c]	*URA3* (*Kl*)	n/a	*YKL162C-A*	Amp	*Not*I or *Bci*VI	pSIB585	63203
pAV10.K5.loxP[c]	*LEU2* (*Sc*)	n/a	*YKL162C-A*	Amp	*Not*I or *Bci*VI	pSIB586	63204
pAV10.HO3.loxP[c]	*HIS5* (*Sp*)	n/a	HO locus	Amp	*Not*I	pSIB587	63205
pAV10.HO5.loxP[c]	*LEU2* (*Sc*)	n/a	HO locus	Amp	*Not*I or *Bci*VI	pSIB589	63206
pAV10.K3	*HIS5* (*Sp*)	n/a	*YKL162C-A*	Amp	*Not*I	pSIB596	63207
pAV10.KH	hygromycin	n/a	*YKL162C-A*	Amp	*Not*I	pSIB599	63208
pAV10.KN	cloNAT	n/a	*YKL162C-A*	Amp	*Not*I	pSIB601	63209
pAV10.K5	*LEU2* (*Sc*)	n/a	*YKL162C-A*	Amp	*Not*I or *Bci*VI	pSIB604	63210
pAV10.HO6	*URA3* (*Kl*)	n/a	HO locus	Amp	*Not*I or *Bci*VI	pSIB843	63211
pAV10.K4	*TRP1* (*Sc*)	n/a	*chrIXR*: 387328−388330	Amp	n/a	pKF091	63212
pAV10	n/a	n/a	n/a	Amp	*Not*I or *Fse*I	pNA0179	63213

[a]*Sc*, *Saccharomyces cerevisiae*; *Sp*, *Schizosaccharomyces pombe*; *Kl*, *Kluyveromyces lactis*; Amp, ampicillin; Cam, chloramphenicol. For additional information on yGG acceptor vector nomenclature, see Figure 2. [b]2° RE refers to the secondary restriction enzyme used to release an assembled TU prior to integrative yeast transformation. [c].loxp indicates the inclusion of loxp sites in the yGG vector. The TU is flanked by two Loxp sites.

The third type of acceptor vector is intended for integration into a specific locus in the yeast genome and therefore lacks genetic parts that enable independent replication in yeast (e.g., CEN/ARS or 2 micron sequences). To this end, as with the other two sets of pRS vectors, we have converted the pRS40X series into yGG TU acceptor vectors for integration into the *URA3*, *LEU2*, *TRP1*, and *HIS3* loci[12] (Figure 2). Furthermore, we have designed and built a series of custom integrative acceptor vectors (Figure 2). Here, a yeast selectable marker is encoded on one side of the *Bsa*I-RFP-*Bsa*I cassette, and together these parts are flanked by ∼500 bp regions targeting an innocuous site in the yeast genome. These sites include the *HO* locus on chromosome 4, intergenic regions on the left arm of chromosome 6 and right arm of chromosome 9, and a dubious ORF on chromosome 11 (*YKL162C-A*) (Table 2). To facilitate integration, on either side of the targeting sequences each custom integrating acceptor vector encodes rare cutting restriction sites such as *Bci*VI and/or *Not*I (Table 2). Digestion with the second enzyme can excise the entire integration cassette, generating a substrate for recombination with the corresponding endogenous locus in the yeast chromosome.

Although our yGG acceptor vector set is extensive, it was constructed with our own needs in mind. Any other vector could, in principle, be manipulated to serve as an acceptor vector for yGG (details on yGG acceptor vector construction can be found in ref 12).

Designing and Constructing PRO, CDS, and TER Parts. The boundaries of PRO, CDS, and TER parts are determined using a simple set of rules that enable the automated extraction of their sequences from the yeast genome (or elsewhere). Because the start and stop codons of a CDS are encoded by the designer overhangs as part of prefix and suffix sequences, a CDS part is defined to extend from the second codon of the open reading frame of a gene through the last sense codon. For PRO and TER parts extracted from the *S. cerevisiae* genome, we have developed boundary definition rules based on commonly accepted, average sequence length for these two types of genetic elements, fully aware that such rules are imperfect. Specifically, we somewhat arbitrarily define yeast PRO segments as the DNA extending 5′ of the ATG codon of the gene of interest for either (i) 500 bp or (ii) the nearest upstream gene boundary, whichever is shorter. TER sequences are defined as the sequence 3′ of the CDS that extend either (i)

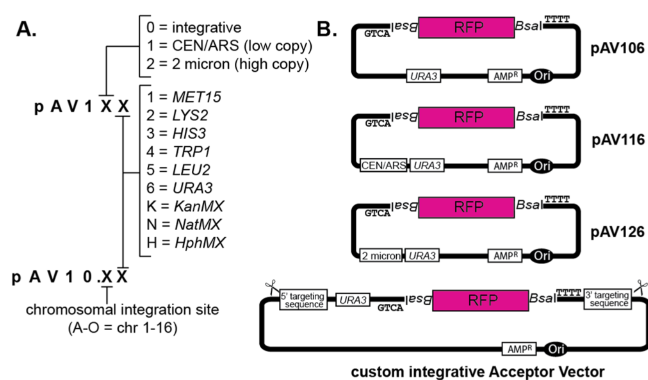


Figure 2. yGG acceptor vectors. (A) Schematic representation of acceptor vector nomenclature. (B) Schematic of representative acceptor vectors. All yGG acceptor vectors (AVs) encode resistance to ampicillin (AMP^R) or chloramphenicol (not pictured) to permit construction of TUs in a one-pot reaction with PRO, CDS, and TER parts that are cloned into kanamycin resistant vectors. Following transformation of yGG reaction products into *E. coli*, white/red screening can be used to identify clones encoding assembled constructs.

200 bp or (ii) the nearest downstream gene boundary, whichever is shorter.

Prefix and suffix sequences can be appended to parts in at least three ways: (i) The appropriate overhang can be encoded by primers such that the resulting PCR product encodes the appropriate sequences; this is typically done for PRO and TER sequences cloned out of *S. cerevisiae*. (ii) The prefix and suffix can be built into the design of parts to be made by polymerase chain assembly[13] or other means of DNA synthesis; this is typically done for CDSs derived from other organisms as we first recode the CDS to optimize codon usage for expression in *S. cerevisiae* using GeneDesign.[14] (iii) The prefix and suffix could be ligated to a pre-existing part as adapter or linker sequences.

In cases where a forbidden site exists internally to a part, there are multiple ways to eliminate the site. Most directly, after subcloning, the forbidden site can be changed using site-directed mutagenesis. Alternatively, one can design a modified version of the part to be synthesized. Finally, the forbidden type IIS restriction site can be eliminated by constructing a pair of subparts that can be used together in yGG reactions (Supporting Information Figure S1). For the most part, changing one base in a PRO or TER part is unlikely to alter the function of the part, and recoding forbidden sites internal to CDS parts can also be carried out using GeneDesign[14] and similar software.

In lieu of changing forbidden sites within part sequences, the yGG reaction conditions can also be modified to skip the 5 min incubation at 50 °C, the second to last step. Eliminating the type IIS restriction digest in this step increases the background of intact parental vector, but it leaves some proportion of correct assemblies with ligated internal sites. Although one would expect both a lower yield of correct assemblies as well a higher background of intact parental vector to transform *E. coli*, the white/red selection system built into our yGG workflow makes it easy to distinguish clones with assembled constructs.

Efficiency of yGG Assembly. The most basic yGG reaction consists of four parts for assembly: a PRO, CDS, TER, and an acceptor vector. In some cases, however, the number of parts can increase, for example, if a single CDS is composed of multiple subparts or when generating a TU with a C-terminal fusion tag (see below). To examine how the number of parts affects assembly efficiency, we compared four- and eight-part yGG reactions (Supporting Information Data) using white/red screening as the output (see Methods). Initially we followed the protocol described by Engler et al.,[5] which specified stock *Bsa*I at 10 U/μL. Here, for the four-part assembly, 92% of recovered colonies were white, and for the eight-part assembly, 72% of colonies were white (Figure 3).

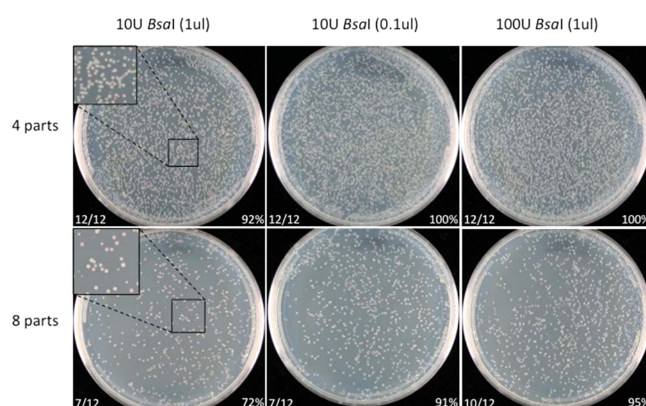


Figure 3. Efficiency of yGG with different numbers of parts. Four- and eight-part yGG assembly was performed in the presence of 10, 10, and 100 U of *Bsa*I with a volume of 1, 0.1, and 1 μL of enzyme, respectively. The yGG reaction products were transformed into bacteria and plated on LB-Carb plates. Pictures were taken after a 1 day incubation at 37 °C. White and red colonies were counted; white colonies percentage is indicated on the lower right of the picture. The fraction on the left-hand side is the number of correct assembly clones, as measured by plasmid prep and digest of 12 white colonies.

The observation of red transformants suggested to us that the final yGG reaction product contained undigested parental acceptor vector encoding the RFP cassette. We hypothesized that an insufficient active *Bsa*I might underlie this result. To test this, we obtained a concentrated stock of *Bsa*I (100 U/μL) from New England BioLabs to circumvent the problem that addition of extra *Bsa*I at the standard concentration (10 U/μL) yielded a prohibitive glycerol concentration in the final reaction mixture. Using 100 U of *Bsa*I per reaction (1uL of 100 U/μL), we recovered 95% white colonies in the eight-part assembly reaction. Moreover, we discovered this result could be recapitulated for the eight-part assembly (91% white colonies) using only 10 U of the 100 U/μL *Bsa*I stock (0.1 μL). This result indicates that reduced glycerol concentration underlies the improved *Bsa*I digestion efficiency in yGG reactions.

To test whether white-colored transformants encoded correctly assembled TU constructs, we picked 12 colonies from each reaction condition, prepped the plasmids, and digested with an appropriate restriction enzyme to test the assembly structure. For the four-part assembly, in each of the three experimental conditions, 100% of the selected white colonies yielded the expected digestion pattern. However, in the eight-part assembly, the 10, 100, and 10 U low-glycerol reactions yielded only 7/12, 10/12, and 7/12 correct assemblies, respectively. Four independent incorrect digestion patterns were observed, and a single clone representing each class was sequenced to investigate the cause of each misassembly. In two cases, an internal CDS overhang (TGGT or GTTG) misassembled with a designer overhang

DOI: 10.1021/sb500372z
ACS Synth. Biol. 2015, 4, 853−859

in which 2 bp were mismatched to a designer overhang (CAGT or TTTT). The third misassembly occurred between an internal CDS overhang (ACGG) with 3 bp mismatches to the designer overhang (CAGT). In the final misassembled clone we analyzed, the sequencing reaction failed, possibly due to a large deletion or a plasmid contamination. Thus, it is possible that the overhangs internal to the CDS part assembly may not be optimal for yGG assembly and should be evaluated on a case-by-case basis.

In summary, we conclude that yGG is most efficiently performed using 10 U of highly concentrated restriction enzyme to minimize the concentration of glycerol in the reaction. This is particularly important when assembling TUs with more than four parts. Moreover, our results suggest that the faithful assembly of parts in yGG reactions requires the use of maximally different overhang sequences when possible. These observations should be broadly applicable to Golden Gate reactions in general.

C-Terminally Tagging TUs Generated by yGG. It is often useful to express a tagged version of a protein for fluorescence microscopy, immunopurification, expression level analysis, etc. To this end, we have devised a yGG-compatible strategy to generate TUs encoding C-terminal fusion tags. Here, we assign a special suffix to the CDS part that permits its assembly into a TU in either the untagged or tagged format. Our design utilizes the enzyme *Bce*AI in combination with either *Bsa*I or *Bsm*BI (Figure 4). *Bce*AI is a long-reach type IIS restriction enzyme that cuts 12 and 14 bp from the recognition, leaving a 2 base overhang. Although the *Bce*AI recognition sequence is 5 bp in length, it contains the very rarely encountered CpG dinucleotide and is thus underrepresented in the yeast genome relative to other sequences of this length. By embedding a *Bsa*I (or *Bsm*BI) site inside the *Bce*AI site (**ACGGC**ATA**GGTCTC**GCTCA), it is possible to generate one of two different overhangs; *Bsa*I (or *Bsm*BI) digestion in a standard yGG reaction generates a standard 3′ CDS overhang of TGAG, whereas *Bce*AI digestion leaves a 2 base overhang consisting of only the AC of the complementary strand to the TGA stop codon, allowing read-through to occur. Due to the moderately unreliable digestion pattern of long-reach restriction enzymes like *Bce*AI and to ensure assembly with the digested fragments, we use annealed oligonucleotides in combination with a standard acceptor vector. Those oligos can contain either a short tag (e.g., flag, V5, HA, etc.) or a linker to ensure the C-terminus of the CDS is in frame with sequences of longer tags (e.g., GFP, mCherry, TAP, GST). Longer tags can be provided as yGG-compatible subcloned constructs to which we assign the 3′ overhang sequence GGAT. In contrast to the untagged yGG, which may be performed as a one-pot reaction, a tagging yGG reaction requires predigestion of the CDS construct with *Bce*AI and gel purification prior to the yGG reaction.

To test the efficiency of C-terminal tagging by yGG, we built a CDS construct with the appropriate sequences flanking *HTP1*, whose protein product functions in the purine salvage pathway (Supporting Information Data). The *HTP1* CDS part was assembled into a TU by yGG with its native promoter and terminator along with a V5 tag plus a fluorescent protein tag (either mCherry or GFP). The V5 sequence was provided as annealed oligos and served as a linker to put the fluorescent protein sequence in frame. Assembly efficiency, assessed by PCR with primers spanning the GFP tag, revealed that 10 out of 13 white colonies produced amplicons consistent with

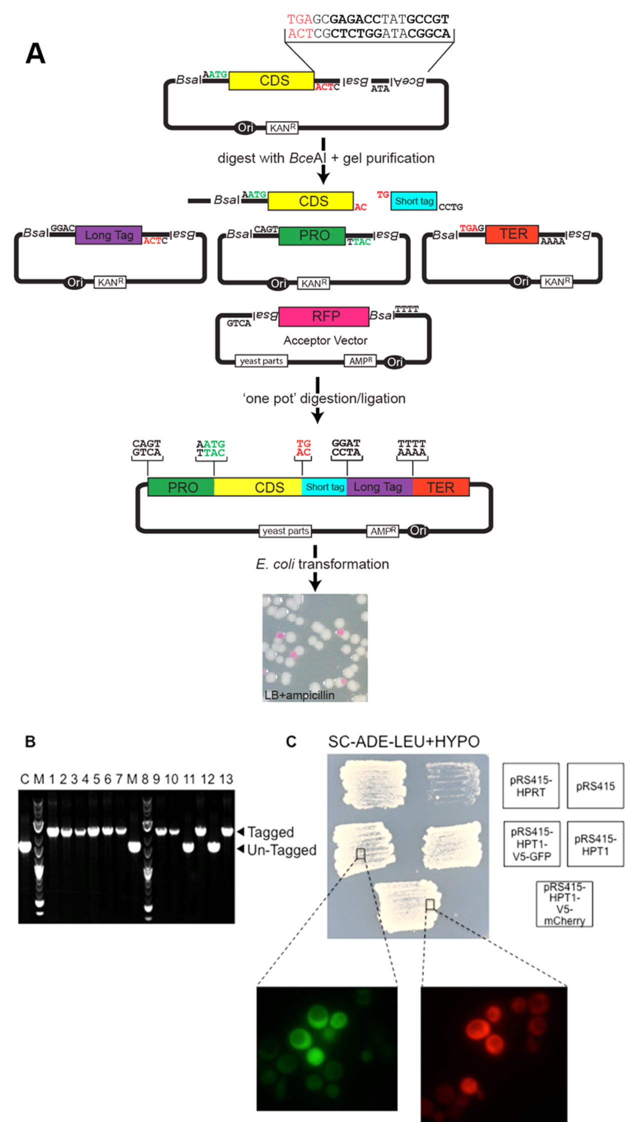


Figure 4. yGG to construct a TU encoding a C-terminally tagged CDS. (A) The CDS is part flanked by the appropriate prefix and suffix sequences and cloned into a kanamycin resistant vector (Supporting Information Data). Prior to TU assembly by yGG, the CDS construct is digested with *Bce*AI and subsequently gel purified. The *Bce*AI digested CDS fragment is mixed with PRO, TER, Long Tag (e.g., GFP, mCherry, TAP, GST) constructs, each flanked by the appropriate prefix and suffix sequences, plus Short Tag (e.g., flag, V5, HA etc.) or linker annealed oligos and desired yGG acceptor vector. The mixture is then subjected to a one-pot digestion−ligation reaction with the appropriate enzymes to assemble the tagged TU. Following *E. coli* transformation, white/red screening can be used to distinguish clones encoding putative TU assemblies as compared to unmodified parental vector. (B) Colony PCR was performed on 13 white colonies from yGG assembly carried out as described in (A). Primers amplified a region around the C-terminus of the V5-GFP tagged protein to differentiate tagged and untagged clones. Ten of 13 amplicons are consistent with the predicted size for the tagged construct. C, untagged construct; M, DNA ladder. (C) Functional validation of C-terminally tagged *HTP1*. *HTP1* C-terminally tagged with GFP or mCherry is functional; thus, cells grow on medium containing hypoxanthine as the sole purine source. HPRT, the human ortholog of HPT1 that is known to functionally complement, is a positive control. Fluorescence microscopy reveals expression of both mCherry and GFP in these cells.

DOI: 10.1021/sb500372z
ACS Synth. Biol. 2015, 4, 853−859

correct assembly. Similar results were obtained with the mCherry tag (data not shown). One of each GFP or mCherry tagged *HTP1* TU constructs was then subjected to two functional assays. To determine whether *HTP1* was expressed, we transformed the constructs into a yeast strain in which *ADE4* and *HTP1* had been deleted from their native genomic loci. In the absence of *HTP1* expression, this strain cannot grow on medium containing hypoxanthine as the sole purine source; however, both the mCherry and GFP tagged *HTP1* TUs fully complemented the growth on this medium (Figure 4B). Expression of both mCherry and GFP was also confirmed by fluorescence microscopy (Figure 4C).

Together, this work demonstrates successful construction of C-terminally tagged TUs by yGG. The reduced efficiency of assembly here as compared to the untagged assembly likely lies in the digestion and gel purification step. Specifically, any undigested CDS carried through the gel extraction step can lead to untagged TU assembly during the yGG reaction. Additionally, long reach IIS enzymes typically cut with less precision than short reach IIS enzymes like *Bsa*I and *Bsm*BI, as the sequence composition between the recognition and cut sites can impact DNA movement and stretching.[15]

Perspective for Further Applications Using yGG. One of the biggest advantages of our yGG method is the use of the bacterial RFP to select against unmodified parental acceptor vector. However, in some cases, there is an obvious selection in yeast that can differentiate between correct and incorrect clones, such as with assembly of an essential yeast transcription unit. In this case, we can bypass the bacterial step and transform the yGG product straight into the yeast cells. A modification that could be made to the yGG acceptor vector in this case is to express a yeast marker between the *Bsa*I sites.

Although we propose using yGG to assemble yeast transcription units for expression in yeast, there are other useful applications for this method. For example, we have used the PRO and TER sequences to serve as homologous sequences for targeted deletion of a specific yeast gene.[16] Using yGG, a selectable marker gene (*URA3*, *KanMX*, etc.) can be assembled between the PRO and TER of the gene to be deleted. For this, we built a specific acceptor vector (pAV10) without a yeast selectable marker and lacking a yeast replication origin. Additionally, we included rare restriction enzyme (*Not*I and *Fse*I) recognition sites flanking the TU assembly site. Thus, following assembly using yGG, the fragment containing the PRO, marker, and TER can be digested and transformed into yeast for targeted deletion of the gene of interest.

yGG can also be used for expression of non-native genes in yeast by assembling a heterologous CDS with a yeast promoter and terminator. Importantly, to enable optimal expression in yeast, the gene sequence should be first codon optimized for *S. cerevisiae*, keeping in mind the forbidden sites to ensure efficient assembly.

In addition, the concepts of yGG can easily be adapted to mammalian or plant cells. Expression in mammalian cells may require larger, more complex promoters, but the same yGG concepts can be used once these are defined. Similarly, there is strong evidence that encoding an intron in mammalian expression constructs has a positive influence on expression. An intron could be contained within the PRO segment, or a separate intron segment could be interposed between the PRO and CDS segments or between multiple CDS parts, allowing the evaluation of large numbers of different introns on gene expression, for example.

Thus, use of the assembly strategy in yGG can be expanded for easy cloning in a variety of uses and organisms. Future work could expand the use of this method, creating an extensive library of parts complying with yGG design standards, allowing the combinatorial assembly of transcription units from different organisms for expression in yeast and other organisms.

METHODS

One-Pot yGG Assembly. One-pot Golden Gate assembly reagents used for yGG reactions are described in detail.[5,6] Briefly, TU parts (PRO, CDS and TER), each subclone cloned into a kanamycin resistant vector (pCR Blunt II TOPO, Invitrogen/Life technologies, Carlsbad, CA, 450245), were combined in equimolar amounts and mixed with reaction master mix [1.5 μL 10× T4 DNA ligase reaction buffer (New England BioLabs, M0202), 0.15 μL 100× bovine serum albumin (BSA, New England BioLabs), 600 U T4 DNA ligase (rapid) (Enzymatics, Beverly, MA, L6030-HC-L), and 10 or 100 U of *Bsa*I or/and *Bsm*BI (New England BioLabs, Beverly, MA, R0535 or R0580, respectively)] to a final volume of 15 μL. The high-concentration *Bsa*I was a custom order from New England BioLabs. One-pot digestion−ligation assembly was performed in a thermocycler as follows: 25 cycles of 37 °C for 3 min and 16 °C for 4 min, followed by 50 °C for 5 min and 80 °C for 5 min. Five microliters of each assembly reaction was transformed into 50 μL of competent DH5α *E. coli* cells and plated on the appropriate selection media. For C-terminal tagging yGG assembly reactions, before the one-pot yGG assembly, 1 μg of cloned CDS was digested with *Bce*AI and loaded on a gel, and the appropriate band was extracted (Zymo Research, Irvine, CA).

ASSOCIATED CONTENT

Supporting Information

Figure S1: Elimination of a *Bsa*I site using yGG. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: 1-646-501-0503. E-mail: jef.boeke@nyumc.org.

Author Contributions

[‖]N.A. and L.A.M. contributed equally to this work. N.A., L.A.M., and J.D.B. wrote the manuscript. N.A. and L.A.M. created the figures. N.A., L.A.M., Y.C., S.I., J.C., A.Z., W.J.C., J.A.M., K.C., and G.S. contributed experimental data. The work was performed in the lab of J.D.B.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported in part by National Science Foundation grant MCB-0718846 and Defense Advanced Research Projects Agency contract N66001-12-C-4020 (to J.D.B.). L.A.M. was funded by a postdoctoral fellowship from the Natural Sciences and Engineering Research Council of Canada.

ABBREVIATIONS

yGG, yeast golden gate; PRO, promoter; CDS, coding sequence; TER, terminator; TU, transcription unit; DNA, DNA; bp, base pair

DOI: 10.1021/sb500372z
ACS Synth. Biol. 2015, 4, 853−859

REFERENCES

(1) Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M. D., Tai, A., Main, A., Eng, D., Polichuk, D. R., Teoh, K. H., Reed, D. W., Treynor, T., Lenihan, J., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K. W., Fickes, S., Galazzo, J., Gaucher, S. P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Jiang, H., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., Westblade, L. F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievense, J., Covello, P. S., Keasling, J. D., Reiling, K. K., Renninger, N. S., and Newman, J. D. (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature 496*, 528−532.

(2) Basehoar, A. D., Zanton, S. J., and Pugh, B. F. (2004) Identification and distinct regulation of yeast TATA box-containing genes. *Cell 116*, 699−709.

(3) Blazeck, J., and Alper, H. S. (2013) Promoter engineering: recent advances in controlling transcription at the most fundamental level. *Biotechnol. J. 8*, 46−58.

(4) Mischo, H. E., and Proudfoot, N. J. (2013) Disengaging polymerase: terminating RNA polymerase II transcription in budding yeast. *Biochim. Biophys. Acta 1829*, 174−185.

(5) 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.

(6) Engler, C., Kandzia, R., and Marillonnet, S. (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One 3*, e3647.

(7) Cambray, G., Guimaraes, J. C., Mutalik, V. K., Lam, C., Mai, Q. A., Thimmaiah, T., Carothers, J. M., Arkin, A. P., and Endy, D. (2013) Measurement and modeling of intrinsic transcription terminators. *Nucleic Acids Res. 41*, 5139−5148.

(8) Mutalik, V. K., Guimaraes, J. C., Cambray, G., Lam, C., Christoffersen, M. J., Mai, Q. A., Tran, A. B., Paull, M., Keasling, J. D., Arkin, A. P., and Endy, D. (2013) Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods 10*, 354−360.

(9) Mutalik, V. K., Guimaraes, J. C., Cambray, G., Mai, Q. A., Christoffersen, M. J., Martin, L., Yu, A., Lam, C., Rodriguez, C., Bennett, G., Keasling, J. D., Endy, D., and Arkin, A. P. (2013) Quantitative estimation of activity and quality for collections of functional genetic elements. *Nat. Methods 10*, 347−353.

(10) Dvir, S., Velten, L., Sharon, E., Zeevi, D., Carey, L. B., Weinberger, A., and Segal, E. (2013) Deciphering the rules by which 5′-UTR sequences affect protein expression in yeast. *Proc. Natl. Acad. Sci. U.S.A. 110*, E2792−E2801.

(11) Sikorski, R. S., and Hieter, P. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics 122*, 19−27.

(12) Mitchell, L. A., Cai, Y., Taylor, M., Noronha, A. M., Chuang, J., Dai, L., and Boeke, J. D. (2013) Multichange isothermal mutagenesis: a new strategy for multiple site-directed mutations in plasmid DNA. *ACS Synth. Biol. 2*, 473−477.

(13) Stemmer, W. P., Crameri, A., Ha, K. D., Brennan, T. M., and Heyneker, H. L. (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene 164*, 49−53.

(14) Richardson, S. M., Nunley, P. W., Yarrington, R. M., Boeke, J. D., and Bader, J. S. (2010) GeneDesign 3.0 is an updated synthetic biology toolkit. *Nucleic Acids Res. 38*, 2603−2606.

(15) Pingoud, A., Fuxreiter, M., Pingoud, V., and Wende, W. (2005) Type II restriction endonucleases: structure and mechanism. *Cell. Mol. Life Sci. 62*, 685−707.

(16) Terfruchte, M., Joehnk, B., Fajardo-Somera, R., Braus, G. H., Riquelme, M., Schipper, K., and Feldbrugge, M. (2014) Establishing a versatile Golden Gate cloning system for genetic engineering in fungi. *Fungal Genet. Biol. 62*, 1−10.

DOI: 10.1021/sb500372z
ACS Synth. Biol. 2015, 4, 853−859