EcoFlex - A Multifunctional MoClo Kit for *E. coli* Synthetic Biology

Journal:	*ACS Synthetic Biology*
Manuscript ID	sb-2016-00031f.R1
Manuscript Type:	Letter
Date Submitted by the Author:	02-Apr-2016
Complete List of Authors:	Moore, Simon; Imperial College London, Medicine Lai, Hung-En; Imperial College London, Medicine Kelwick, Richard; Imperial College London, Department of Life Sciences Chee, Soo Mei; Imperial College London, Medicine Bell, David; Imperial College London, Medicine Polizzi, Karen; Imperial College London, Dept of Life Sciences Freemont, Paul; Imperial College London, Medicine

EcoFlex - A Multifunctional MoClo Kit for *E. coli* Synthetic Biology

Simon John Moore[1,3], Hung-En Lai[1,3], Richard Kelwick[1,3], Soo Mei Chee[3], David Bell[3], Karen Marie Polizzi[1,2] and Paul Freemont[1,3]

[1]Centre for Synthetic Biology and Innovation, Imperial College London, South Kensington Campus, Exhibition Road, London, SW7 2AZ, UK
[2]Department of Life Sciences, Imperial College London, South Kensington Campus, Exhibition Road, London, SW7 2AZ, UK
[3]Department of Medicine, Imperial College London, South Kensington Campus, Exhibition Road, London, SW7 2AZ, UK

Abstract

Golden Gate cloning is a prominent DNA assembly tool in synthetic biology for the assembly of plasmid constructs often used in combinatorial pathway optimisation, with a number of assembly kits developed specifically for yeast and plant-based expression. However, its use for synthetic biology in commonly used bacterial systems such as *Escherichia coli*, has surprisingly been overlooked. Here, we introduce EcoFlex a simplified modular package of DNA parts for a variety of applications in *E. coli*, cell-free protein synthesis, protein purification and hierarchical assembly of transcription units based on the MoClo assembly standard. The kit features a library of constitutive promoters, T7 expression, RBS strength variants, synthetic terminators, protein purification tags and fluorescence proteins. We validate EcoFlex by assembling a 68-part containing (20 genes) plasmid (31 kb), characterise *in vivo* and *in vitro* library parts, and perform combinatorial pathway assembly, using pooled libraries of either fluorescent proteins or the biosynthetic genes for the antimicrobial pigment violacein as a proof-of-concept. To minimise pathway screening, we also introduce a secondary module design site to simplify MoClo pathway optimisation. In summary, EcoFlex provides a standardised and multifunctional kit for a variety of applications in *E. coli* synthetic biology.

Keywords

Synthetic biology, Golden Gate, *Escherichia coli*, recombinant protein production, cell-free protein synthesis, violacein

Introduction

A prime target for synthetic biology is the production of natural and commodity chemical products such as drug precursors, fine and specialised chemicals and biofuels using renewable feedstocks [1]. This is driven by the need to transition away from a current reliance on petroleum-derived chemical products and production processes. A more sustainable and environmentally friendly solution to this problem is in the application of synthetic biology to engineer enzyme biosynthetic pathways primarily within microbial cells to produce such chemical products. Essential to the synthetic biology workflow is the use of rapid, efficient and hierarchical DNA assembly techniques and a number of methods have been proposed [2]. One such method, Golden Gate cloning, is a modular, combinatorial technique used for a wide spectrum of DNA assembly applications, from the construction of simple transcription units (TUs) to diverse biochemical pathway libraries originally developed by Marillonnet and colleagues [3–5].

Golden Gate cloning utilises the non-palindromic type IIS restriction endonuclease BsaI, which cuts outside of its recognition sequence (GGTCTCN$^{\downarrow}$NNNN), producing a four base pair overhang. This overhang, also known as a fusion site, can be designed to provide directional DNA part assembly. BsaI sites are placed at both the 5' and 3' ends of a DNA part in inverse orientation, allowing the restriction site to be removed before assembly by the T4 DNA ligase. DNA parts are held within "entry vectors" and can include promoters, protein tags, ribosome-binding sites (RBS), open reading frames (ORF), protein domains, TALEN repeats [6], fusion linkers and terminators. A combined one-pot reaction of "entry vectors" is then digested and ligated into "destination vectors" using BsaI and T4 DNA ligase. Golden Gate assembly has been demonstrated to assemble up to 10 DNA fragments with greater than 95% efficiency [5] and relies solely on the use the BsaI enzyme. Recently, modifications to the original Golden Gate cloning technique have included Golden Braid and MoClo standards [7,8], which rely on the use of alternating antibiotic marker plasmids and dual type IIS endonucleases (BsaI/BsmBI or BsaI/BpiI) to rapidly construct hierarchical multi-gene pathways up to 33 kb in size [9]. In summary, several plasmid kits have been developed including MoClo and GoldenBraid for plants [7,10,11], *Saccharomyces cerevisiae* [12,13] and *Schizosaccharomyces pombe* [14]. For a detailed review of assembly strategies utilising type IIS endonucleases please see Casini *et al* [2]. Focusing on MoClo, this particular assembly scheme allows flexible design of multiple transcription units into modular pathways with each gene under the control of its own promoter, RBS and terminator.

Until the recent release of the CIDAR MoClo kit [15], Golden Gate assembly kits for commonly used bacterial systems such as *E. coli* have surprisingly been overlooked. CIDAR focuses on an automated workflow for MoClo assembly.

However, we wanted to expand this further by modifying and creating a new flexible Golden Gate tool kit specifically for *E. coli* synthetic biology, based on a custom-designed modular library of standard biological parts for hierarchical DNA assembly and optimisation, using the MoClo assembly standard [9]. Herein, we present EcoFlex, a user-friendly, characterised and customisable Golden Gate kit with parts that can be used for example in combinatorial pathway optimisation, cell-free circuit design, recombinant protein purification and in the future genome engineering. We have also included a custom-designed library of constitutive σ^{70} promoters, inducible expression based on the T7 RNA polymerase, a standardised RBS library and a range of synthetic terminators previously characterised [16]. The EcoFlex kit also provides options for recombinant protein production and purification, such as N-terminal linkers for hexahistidine and Strep(II)-tag purification, with either constitutive or T7-inducible expression as single parts or combined with a T7 promoter, RBS, His_6-tag and thrombin cleavage site. We also provide a plasmid cassette system (and individual parts) for inducible T7 RNA polymerase (RNAP) production and purification (Figure S6). We validate T7 RNAP activity using cell-free protein synthesis (CFPS). We have also made two improvements to the MoClo system, firstly by simplifying the storage plasmid for open reading frames (ORFs) and secondly we introduce a secondary site to offer flexibility in pathway design. For example, three TUs can be placed in the secondary module under constant expression, whilst the primary site can be used to assemble and optimise a specific gene, regulatory cascade or biosynthetic pathway.

In synthetic biology, there is an increasing need to provide a repertoire of biological parts (promoters, RBS, terminators) characterised in the context of the chosen DNA assembly technique. For example, DNA parts can behave differently when placed in a different context (known as “context dependency”), which directly challenges the engineering dogma surrounding synthetic biology. A number of investigators have tried to address this problem through the design of neutral insulator sequences [17–19]. To understand such effects in our EcoFlex kit, we have validated our library of biological parts both *in vivo* and *in vitro* (in cell-free reactions). To demonstrate the broad utility of EcoFlex, we also provide exemplars of rapidly assembling and purifying His_6-tagged proteins, rational library design using flow-cytometry and the optimisation of the violacein biosynthetic pathway using a secondary MoClo module to simply pathway design and increase the frequency of successful library assemblies.

Results and Discussion

E. coli is one of the major workhorses in synthetic biology, providing a very well-studied chassis for the manipulation of genetic circuits and implementation of biosynthetic pathways. Using a repertoire of biological parts, modular DNA pathways can be efficiently assembled using a range of techniques such as Golden Gate assembly [3], BglBricks [20], AQUA [21], BASIC [22] and unique nucleotide sequence (UNS)-guided assembly [23]. In addition, other cloning methods include PCR and restriction enzyme-based cloning, Gibson assembly [24] and sequence-ligation independent cloning [25]. Eventually we desire to assemble a universal cloning system based on the MoClo standard for prokaryotic host expression to allow an ORF or DNA part (e.g. promoter, RBS) to be easily portable between multiple systems. To begin building this foundation, we unveil EcoFlex that is designed to provide users a flexible combinatorial library toolbox that can also be customised for variety of synthetic biology applications in *E. coli*.

The EcoFlex toolbox

We have built a versatile Golden Gate toolbox based on the MoClo assembly standard [9] for *E. coli* synthetic biology, which we make available through AddGene (www.addgene.org/cloning/moclo/freemont-ecoflex). To design this kit, we focused on providing flexibility and compatibility with a variety of applications such as pathway assembly, cell-free protein synthesis (CFPS) and recombinant protein purification (Figure 1). Promoters, RBS, tags, ORFs and terminator "parts" were assembled and stored in Level 0 "entry vectors" with chloramphenicol resistance gene and pMB1 origin of replication. A unique feature of our assembly kit is the "entry vector" for storing ORFs (pBP-ORF), whereby to standardise the cloning procedure, each gene is PCR cloned into NdeI and BamHI sites, which are flanked by inverse BsaI sites for Golden Gate assembly. The NdeI (**CATA**TG) site provides the fusion site (in bold) for Golden Gate assembly to the RBS sequence (or N-terminal tag) and the start codon for each gene, so no additional amino acids are added to the encoded protein. Optionally, if NdeI or BamHI sites are located within the gene, a NcoI site is also provided in the vector, whilst BglII, which produces cohesive ends can be used in place of BamHI. ORFs can also be cloned with flanking BsaI restriction sites with compatible fusion sites (5' CATA and 3' TCGA). With NdeI and BamHI cloning, these genes are also suitable for transfer between some of the commercial pET vectors (e.g. pET3a, pET14b, pET15b) for the purpose of recombinant protein production. Alternatively, our kit also provides customized N-terminal fusion tags with either constitutive or T7 RNA polymerase-dependent expression. The "entry vectors" pBP-*lacZ*α and pBP-ORF are customizable for assembly of new promoter parts or ORFs. Further instructions for user modifications are provided in the SI text.

The first tier of assembly (Level 1) requires BsaI to assemble monocistronic TUs, with each ORF under control of its own unique promoter, RBS and transcription terminator (Figure 1A). This is ligated into a "destination vector" backbone (pTU1-A, -B, -C, -D) containing a pMB1 origin and β-lactamase gene. Level 1 fusion sites and assembly options are summarised in Figure 1B. Level 1 TUs contain flanking inverse BsmBI sites, which are subsequently used to direct the MoClo (Level 2) stage of assembly (ABCD). The pTU1-A, -B, -C and -D (Level 1) plasmids allow assembly of 2-4 TUs into the Level 2 "destination vectors" pTU2-a (to accommodate 2 TUs), -b (3 TUs) and pTU2-A, -B, -C and -D (4 TUs). In addition, optional pTU1-D_1 and pTU1-E Level 1 plasmids also allow an additional Golden Gate fusion site for assembly up to 5 TUs in pTU2-A, -B, -C or -D. Level 2 plasmids are available with either *lacZ*α or RFP negative markers, a selection of origins of replication (pMB1, ColE1 and p15A) and a chloramphenicol resistance marker. For larger pathways, up to 20 TUs can be assembled into a Level 3 "destination vector" (pTU3-A for 2 pTU2 plasmids; pTU3-B for 4 pTU2 plasmids).

Significantly, we have modified the MoClo system to include a secondary module. This is a unique feature to EcoFlex, whereby separate modules of Level 2 TUs (2-5 TUs) derived from pTU2-a, -b or -A plasmids can be sub-cloned into a secondary site (pTU2^S-A, -a, or -b), using two compatible BpiI sites, which is located upstream of the primary MoClo BsmBI assembly site (Figure 1C). We have introduced this secondary module site as an aid to minimise pathway diversity and increase assembly efficiency during combinatorial optimisation. For example, instead of optimising a six-gene pathway randomly, one module of 2 or 3 TUs can be placed in the secondary module, whilst the remaining 4 or 3 TUs, respectively, are randomised in the primary module. This is important since the screening technique and transformation efficiency are major limiting factors for library screening. An example of its use is demonstrated with the violacein pathway (see below). Finally, all EcoFlex "destination vectors" mentioned (pTU1, pTU2 and pTU3) also contain a monomeric red fluorescence protein (RFP) gene pTU2^S-A, -a and -b due to internal BpiI sites in RFP) under a LacI- and CAP-sensitive promoter and strong RBS, sourced from the iGEM Registry of standard biological parts (BBa_J04450). The RFP transcription unit is inserted between flanking BsaI sites (for pTU1 and pTU3) or BsmBI sites (for pTU2) so that negative colony-forming units (CFUs) can be visibly identified (red) after plate growth.

Promoter, RBS and terminator libraries – *in vivo* and *in vitro* validation

A range of promoter and RBS strength variants were created by randomising the *E. coli* σ^{70} Anderson promoter collection [26] and PET RBS Shine-Dalgarno consensus sequences, respectively, upstream of a eGFP (hereafter referred to as GFP) fluorescence reporter with flanking BsaI sites provided for

downstream Golden Gate Level 1 assembly (see SI text). We utilised the iGEM Anderson promoters J23100, J23108 and J23114 as internal reference standards. For the RBS library a 6 bp Shine-Dalgarno consensus sequence AGGAGG was randomised to RRRRRR (A/G) to vary ribosome binding strength and a range of low-high strengths are provided. The promoter and RBS strengths in Level 1 context were characterised in the destination vector pTU1-A with a constant J23100/SJM901 promoter (RBS library) or PET RBS (promoter library), GFP and BBa_B0015 terminator. This was performed using *E. coli* as a chassis with *in vivo* or *in vitro* (CFPS) fluorescence measurements. Previously we have demonstrated a link between *in vivo* and *in vitro* characterisation of promoter and RBS activity [27]. Herein, we provide further evidence of this relationship by characterising a wider range of promoter (Figure 2A, -E), RBS (Figure 2B) and synthetic terminator parts (Figure 2C, -D).

In summary, most of the parts behaved similarly between the two systems. For example, a strong promoter or terminator *in vivo* was generally strong *in vitro*. However, there are some exceptions to this general rule. For example, *in vivo*, the SJM901 promoter variant showed a moderate GFP synthesis rate of 1.2-fold in comparison to the J23100 promoter standard. Unexpectedly, *in vitro*, this rate increased to 7.8-fold (Figure 2A, -E and –F). Whilst promoter strengths demonstrated a clear link between *in vivo* and *in vitro* activity, we observed some differences between RBS and terminator activity, which was dependent on the strength of promoter used. Unlike *in vivo*, *in vitro*, the signal-noise ratio for the J23100 promoter was low (Figure S4), thus it is unclear to differentiate between weak-strong RBS variants. This could be improved by using the strong SJM901 promoter (Figure S4). However, with the SJM901 promoter the range of weak-strong RBS variants indicated from *in vivo* observations was not as clear *in vitro*, whilst some variants showed enhanced activity (TL2 and TL4) of almost 3-fold over the PET standard. These findings could suggest that the strong activity of the SJM901 promoter saturates ribosome binding *in vitro* therefore narrowing the range of observed activities for apparent weak and strong RBS variants.

For terminator characterisation, a different assay was designed to quantify terminator efficiency using EcoFlex parts. To do this, GFP was assembled into pTU1-A with the J23100 or SJM901 promoter and PET RBS, followed by a variable strength terminator. In pTU1-B, mCherry was assembled with the weak J23114 promoter, PET RBS and BBa_B0015 terminator. These two TUs were then assembled into pTU2-a as a Level 2 assembly (Figure S3). If the variable terminator allowed read-through transcription, this was detected by mCherry fluorescence. In comparison to the original publication describing the synthetic terminator parts [16] used herein, the simple assay we utilise differs in the use of a strong constitutive promoter instead of the arabinose

induction system used previously [16]. We represent our data as a mCherry and GFP rate that is normalised to a control plasmid containing two strong BBa_B0015 terminators for both GFP and mCherry (Figure 2C, -D). Although our data is not a direct comparison to Chen *et al* [16], our *in vivo* measurements suggest that the activity of the synthetic terminators is reproducible in the EcoFlex context with either the SJM901 (Figure 2C) and J23100 promoters (see SI files). It should also be noted that growth rate of the Level 2 plasmids (chloramphenicol) is considerably slower than for the promoter and RBS characterisation data, where Level 1 plasmids (carbenicillin) are used. The standard error of the growth and GFP rates also increases significantly by utilising the stronger SJM901, which is not apparent from the original Level 1 promoter characterisation. Interestingly, by using the J23100 promoter this error decreases, suggesting that the SJM901 promoter causes more stress to the cells in Level 2 plasmids. We utilised both promoters, since for cell-free measurements, only the stronger SJM901 promoter was sufficient to drive production of mCherry.

With the SJM901 promoter *in vivo*, the strongest terminators were L3S2P21 (0.37-fold mCherry rate), L3S1P51 (0.73-fold), L3S1P32 (1.26-fold) and L3S1P11 (1.98-fold), each showing a low rate of mCherry fluorescence in comparison to the strong BBa_B0015 terminator. The remaining library members show a 3 to 10-fold increased mCherry fluorescence rate. The signal intensity of GFP varied between the constructs but remained within a 2.5-fold range of the BBa_B0015 control plasmid. Interestingly, for all of the strong terminators observed *in vivo*, these also behaved similarly *in vitro* (Figure 2D). However, of the weaker terminators only L2U5H1 (2.88) and BBa_B0012 (3.09) showed a clear increase in relative mCherry signal. We did observe a broad range of activities, suggesting that small changes in the terminator structure can have significant effects on either GFP or mCherry production. Mechanisms outlined by Chen *et al* [16] can include changes in mRNA degradation, sequestering of the mCherry RBS site through complex secondary structure or decreased translation coupling. Through continued study and a broader in depth characterisation of DNA parts for EcoFlex, the changes we have observed could be minimised by implementing the method of Mutalik *et al* [18] or introducing neutral insulator DNA sequences [17–19] to minimise complex mRNA secondary structure.

N-terminal affinity tags for recombinant protein purification

ORFs can be fused to a N-terminal fusion tag such as Strep(II) or His_6-tag for recombinant protein production and purification. In addition, a combined T7 promoter, PET RBS, His_6-tag and thrombin cleavage site can be fused to an ORF of interest, terminator and “destination vector” (e.g. pTU1-A-RFP) as a four-part assembly and transformed directly into a lysogenic DE3 strain such as BL21 Gold or KRX for IPTG-inducible T7-expression. Alternatively, lower

strength constitutive promoters can also be used to lower protein levels to prevent aggregation and inclusion body formation, a problem often encountered with T7-based expression [28]. In summary, users can construct TUs rapidly with the desired combinations of promoter strength, RBS variants, N-terminal linkers, ORFs and terminator strength and conveniently identify correctly assembled vectors using blue/white (LacZ) or red/white (RFP) screening. Furthermore, for protein production, an ORF can be assembled with EcoFlex, expressed and purified within 48 hours.

As an example, we sub-cloned CFP and mCherry into the pET15b vector using the NdeI and BamHI restriction sites. As a comparison, the same genes were MoClo assembled into pTU1-A with the combined T7 promoter, His_6-tag and thrombin cleavage site, together with the BBa_B0015 terminator. We then compared protein production and purification, using BL21 (DE3) Gold pLysS as a host strain. In summary, in comparison to the pET15b expression system, our plasmid system produces equivalent levels of the recombinant His_6-tagged CFP and mCherry proteins, with a summary of the SDS-PAGE purification of CFP and mCherry shown in Figure 3. To further demonstrate this system, we also tested a variety of proteins (from our own library collection) with this system showing high levels of production in the *E. coli* KRX auto-induction strain (Figure S5, Table S4).

Pooled library distribution of GFP and mCherry

The assembly of random DNA part libraries for promoter, RBS, ORF and terminator variants is a powerful technique to explore the design space for constructing pathway combinations. Using pooled libraries of promoters and RBS variants for GFP and mCherry in pTU1-A and pTU1-B, respectively, these two TUs were randomly assembled into pTU2-a. Three separate libraries for both GFP and mCherry, with either 3^2, 5^2 or 10^2 variants (for list of parts see Table S4) for coupled promoter and RBS combinations, were built in Level 1 plasmids, creating pTU1-A-GFP and pTU1-B-mCherry. These pooled libraries were sequenced and combinations of pTU1-A-GFP and pTU1-B-mCherry were assembled into pTU2-a, therefore creating theoretical library sizes of 81 (3^2 variants of GFP combined with 3^2 variants of mCherry), 625 (5^2 variants of GFP combined with 5^2 variants of mCherry) or 10,000 (10^2 variants of GFP combined with 10^2 variants of mCherry). The libraries were analysed by flow cytometry to determine the relative distribution of pathway variants expressing GFP and mCherry. The fluorescence of GFP and mCherry Level 1 (Figure S7) and Level 2 libraries (Figure 4) was compared to individual Level 1 reference standards (Figure S7) prepared with the J23100, J23108 and J23114 promoters, strong PET RBS and BBa_B0015 terminator, as well as an empty vector control (pTU1-A-*lacZ*α). The population distributions of the Level 1 and 2 libraries shifted in response to the diversity of the libraries (Table S5) assembled in the MoClo reactions (Table S6). In general, an

increase in the diversity of the library from 3^2, to 5^2 or 10^2, resulted in an increase in diversity for populations that were compared to the expression profiles of the medium (J23108) and strong expression (J23100) profiles of the GFP or mCherry standards. For instance, the population distributions of Level 1 GFP libraries shifted towards an increase in the strong expression profile (J23100-GFP) from 4.41% (3^2) towards 12.3% (5^2) and 11% (10^2) and a marked decrease in the weak expression profile (J23114-GFP) from 70.9% (3^2) decreasing to 64.1% (5^2) and 46.3% (10^2). Additionally, the Level 1 mCherry library profiles displayed an increased representation of the medium (J23108-mCherry) expression from 35.7% (3^2) towards 45.1% (5^2) and 57.2% (10^2), as the diversity of the library was increased.

Analysis of the Level 2 libraries has also shown that as the library diversity was increased from 5^2 to 10^2 there was greater representation of populations in the strong expression profiles of both J23100-GFP (5^2: 0.68% to 10^2: 4.59%) and J23100-mCherry (5^2: 4.14% to 10^2: 16.9%). The bias towards low-expression in the smaller libraries is likely to occur during the overnight growth phase during Level 1 MoClo assembly. *E. coli* transformed with pooled libraries are likely to be subject to selection pressures in which those cells transformed with weaker promoter and/or RBS combinations were less burdened, in terms of their usage of cellular resources directed to GFP and/or mCherry production, and thus grew more quickly than those cells producing higher levels of the reporter proteins [29] . As a result, the cells that had lower rates of fluorescent reporter production were over represented in the final population. We show that as the size of the library is increased there is likely to be more cells transformed with a larger diversity of medium-to-strong promoter and RBS strength combinations. Therefore, there was a greater diversity of cells with different levels of cellular burden and thus a greater representation of *E. coli* producing higher levels of fluorescent reporter proteins in the final population. Whilst these selection biases exist, this type of approach can provide a quality control check of the efficiency for EcoFlex combinatorial assembly, using fluorescent reporter proteins to rationally enrich library design.

Violacein pathway optimisation with the secondary module

Violacein is a violet pigment first isolated from *Chromobacterium violaceum*, and the pathway belongs to the indolocarbazole biosynthetic family that utilises L-tryptophan as a common precursor [30]. We chose the violacein pathway as a model pathway for EcoFlex combinatorial optimisation with the secondary module, as it requires five enzymes (VioA, -B, -C, -D and -E) encoded by the *C. violaceum vioABCDE* operon, with gene sizes ranging from 576 to 2997 bp and a total size of the synthetic operon of 7,242 bp. Herein, we utilise the violacein pathway for optimisation as the violet colour produced is a qualitative indication of pathway flux during library assembly [31].

A major consideration for optimising combinatorial pathways is library size and the limitations of the screening technique. Putting this into perspective, depending on the assembly level and the number of parts included, Golden Gate assemblies typically yield between 10^2 to 10^5 CFU per transformation. Also, unlike the GFP and mCherry libraries described earlier, a significant decrease in transformation efficiency (approximately 100-fold) is observed with violacein pathway assembly. For example, during a five TU Level 2 assembly of the complete *vioABECD* pathway with the low-strength J23114 promoter, only 20-30 purple colonies were obtained from either JM109 or KRX high efficiency (1×10^8 CFU per μg DNA) competent cells. Toxicity was also noted during optimisation of violacein with VEGAS cloning in Brewer's yeast [13]. To tackle this problem and provide an aid to debug pathway diversity towards a more favourable design space, we decided to introduce a secondary module site into the EcoFlex cloning scheme (Figure 1C). This secondary module encodes two BpiI sites (see SI text), which provide compatible overhangs (5' CTAT) and (3' GTAC) for transferring a sub-pathway (or other DNA parts) from the Level 2 plasmids pTU2-A, -a, or -b through BsaI digest, thus allowing entry of 2-5 TUs into the secondary site.

As an example, a constant sub-pathway for the *vioA* and *vioE* genes was assembled into the RFP-containing pTU1-A and -B destination plasmids respectively, with a J23114 promoter, PET RBS and BBa_B0015 terminator. Next, the *vioA* and *vioE* TUs were assembled into pTU2-a, before sub-cloning of this module by BsaI digest into the BpiI-linearized secondary module plasmid pTU2^S-b, creating pTU2^S-*vioA*-*vioE*-b with either pMB1 (high-copy), ColE1 (medium-copy) and p15A (low-copy) origins of replication. Next, pooled promoter, RBS and terminator libraries of *vioB, vioC* and *vioD* were assembled in pTU1-A-*lacZ*α, pTU1-B-*lacZ*α and pTU1-C-*lacZ*α respectively. 50 μL of the transformation mixture was plated onto carbenicillin agar to estimate Level 1 efficiency using blue/white colony screening (>99% efficiency), whilst the remaining mixture (~200 μL) was directly inoculated into 10 mL of LB medium (with carbenicillin) and grown overnight at 30 °C. The library plasmid DNA was isolated, verified for homogeneity by restriction digest and sequenced (Figure 5A). Level 1 promoter-RBS-terminator libraries of *vioB*, *vioC* and *vioD* were then assembled into pTU2^S-*vioA*-*vioE*-b (pMB1, ColE1 and p15A) using BsmBI and T4 DNA ligase. 5 μL of the reaction mixture was then transformed into 50 μL of *E. coli* JM109 competent cells and plated onto chloramphenicol plates and counted for CFUs (Figure 5B and Table 1). We utilised JM109 instead of DH10β as we have observed instability issues with complex biosynthetic pathways with DH10β. For example, transformation of the lycopene or violacein pathway (pTU2-A-*vioABCDE* – positive pathway control) yields a low percentage of white colonies, which

upon plasmid purification yields the same mini-prep restriction digest map as the original plasmid. Whilst we did not pursue sequencing to identify potential mutations, it is known that DH10β has a high basal mutation frequency [32]. With JM109 transformants, we observed a variety of violet, black, brown and white CFU phenotypes (Table 1). Unlike the low-efficiency 5 TU *vioABCDE* assembly, approximately 10^2-10^3 CFUs were obtained with the plasmids containing the pMB1, ColE1 or p15A origins of replication (Table 1).

Initially, we assayed violacein production and sequenced the library parts around each TU (*vioB, vioC* and *vioD*) in a few colonies from the library transformation plate with pMB1 origin, and found that most of them contain a mutated promoter or RBS in at least one of the *vioB*, *vioC* or *vioD* TUs after overnight growth (see SI text and Figure S9-10). It is also interesting to note that some of the clones after overnight liquid or plate growth lose their ability to produce violacein, which is consistent with a mutation or recombination event. To test whether this was linked to plasmid copy number, we also screened colonies from the library transformation using also ColE1 (medium-copy) and p15A (low-copy) origins of replication.

Focusing on the ColE1 and p15A plasmid variants, we re-streaked purple colonies onto fresh plates and single colonies were then grown in biological triplicate to semi-quantify crude violacein, whilst liquid culture aliquots were sent for high-throughput 96-well sequencing (Bugs2Bases, Source Bioscience) of the *vioB*, *vioC* and *vioD* library regions, which requires an additional sub-culturing step before DNA purification and sequencing. For ColE1 plasmid variants, only 5/11 clones were absent of mutations, whilst for p15A clones, this increased to 9/11 clones. However, in the two p15A clones that did contain a mutation (L32 *vioC* RBS and L40 *vioD* promoter), these were single base-pair deletions immediately adjacent to the MoClo fusion site and could be derived from initial Level 1 assembly (Figure S13). In addition the sequencing trace of the L32 *vioD* terminator was weak, despite repeat sequencing. The impact of the violacein pathway as a strong negative-selection marker also appears to amplify the frequency of these mutation events. Golden Gate and MoClo assembly is renowned to be highly efficient and a restriction digest of DNA is generally considered acceptable for quality-control checking [33]. In contrast, the mutations found in pMB1 and ColE1 variants generally represent large-scale deletions (probable recombination events), which form spontaneously during sub-culturing. We speculate that these events are due to repetitive promoter, RBS and terminator sequences [16,34] used during assembly and strong selection pressures caused by violacein production. This is further supported by the fact that no mutations or deletions were detected within the sequenced regions of the *vioB*, *vioC* and *vioD* ORFs, thus mutations only seem to occurs within the promoter, RBS or terminator sequences. This provides a strong focal point for improvement in our future

studies of library design. Please see supporting information for a full summary of these findings.

In general, the library colonies gave a wide range of crude violacein yields, with the highest producing strain L15 giving 66.3 ± 5.6 mg L^{-1} and the lowest strain L21 giving 3.6 ± 1.7 mg L^{-1} (Figure 5D). However, the ColE1 positive control strain (with all violacein genes under J23114 promoter, PET RBS and BBa_B0015 terminator) had the highest crude violacein yield at 69.1 ± 7.4 mg L^{-1}, whereas the equivalent p15A positive control strain produced 9.9 ± 1.3 mg L^{-1}, the lowest among the p15A library strains (Figure 5). From the sequencing information, the ColE1 and p15A libraries are predominated by weak-medium strength promoters and a range of RBS combinations. Interestingly, the strongest promoter library member SJM901 appears in two instances with *vioC*, but it is only found in the p15A variants. It should also be noted that when violacein production was previously optimised with ePathOptimize using five engineered variants of the T7 promoter, all strong producing clones (up to 60-240 mg L^{-1} crude violacein) were predominated by combinations of the weakest T7 promoter variant [35]. Although our study only uses constitutive promoters, the frequent occurrence of weak promoters with strong pathway variants is a shared common theme between these two studies. Fine-tuning of violacein flux was also recently examined by rationally minimising library size of RBS sites in a recent study [36].

For the terminators, there was a strong bias towards the L3S2P21, L2U5H08, L2U2H08 and Bba_B0015 members with only a single occurrence of Bba_B0012. Considering all variants of the EcoFlex library are included at approximately the same concentration at the initial Level 1 assembly, it is unclear why certain terminators become more frequent at the final stage of selection of violacein CFUs.

Future perspectives and conclusions

While the EcoFlex kit is primarily designed for *E. coli*, we are also interested in expanding this toolkit towards a universal cloning system between alternative prokaryotic hosts that offer attractive advantages in biotechnology, such as thermophilic growth, utilisation of alternative carbon feedstocks, resistance to growth-inhibitory metabolites and high titre natural product synthesis. For example, a pathway design could be studied in a non-traditional host and an individual gene from the same pathway could then be rapidly swapped into *E. coli* for enzyme or structural studies. Furthermore, whilst we demonstrate that the secondary module can be used for the optimisation of biosynthetic pathways, this feature could also be used for constructing novel shuttle vectors or genome engineering tools. EcoFlex provides a foundation for the design and implementation of a universal cloning system as we look towards arcane chassis [37], which are becoming increasingly important for future

synthetic biology directions. In summary, the EcoFlex kit provides a simplified cloning scheme that can be used for a variety of synthetic biology applications for *E. coli* and in the future alternative prokaryotic host systems.

Materials and Methods

Bacterial strains and growth

Routine bacterial growth was performed at 37 °C in LB broth, with agar (15 g L^{-1}), carbenicillin (100 µg mL^{-1}), kanamycin (50 µg mL^{-1}) or chloramphenicol (35 µg mL^{-1}) as needed. Chemically competent *E. coli* DH10β (NEB), JM109 (Promega) and KRX (Promega) were used for routine transformations and maintenance of plasmids.

Molecular biology

Restriction enzymes (NEB), T4 polynucleotide kinase (NEB), T4 DNA ligase (Promega), Phusion® (Agilent) and Q5 DNA polymerase® (NEB) were used for routine cloning methods. Molecular biology kits were purchased from Qiagen (QIAprep Miniprep kit and QIAquick gel extraction kit). Primers and the *lacZ*α fragment (see SI text) were synthesised by IDT. DNA parts and primer sequences are provided in the supplementary information. Routine and high-throughput Bugs2Bases DNA sequencing was performed by Source Bioscience, Cambridge.

EcoFlex assembly conditions

For Level 1 and 3 assembly, 100 ng of each DNA part was combined with 50 ng of destination plasmid and incubated with BsaI and T4 ligase for 15-30 cycles of 5 min at 37 °C and 10 min at 16 °C, followed by 5 min at 50 °C and 5 min at 80 °C. For Level 2 assembly, BsmBI was used as the type IIS restriction enzyme and the reaction was incubated at 37 °C overnight. 5 µL of Golden Gate reaction mix was transformed into 50 µL of chemically competent *E. coli* DH10β, JM109 or KRX cells using heat shock transformation. After recovery in 200 µL of SOC medium (NEB), 50-100 µL of cells were plated onto LB antibiotic plates and grown at 37 °C overnight. For assembly of Level 1 libraries, a pooled library of promoters or RBS (100 ng uL^{-1}) variants were assembled in a Level 1 assembly reaction and transformed into DH10β. Instead of incubating on plates, 200 µL of the SOC mixture was directly inoculated into 10 mL of 2YT medium with antibiotic and grown overnight at 30 °C and 160 rpm shaking. High purity library mini-preps was verified by restriction digest and sequencing.

Flow-cytometry

Golden gate combinatorial library reactions were transformed into *E. coli* and grown overnight in 1 ml 2YT and antibiotics at 30 °C. Cell populations were diluted into 1X phosphate buffered saline (2:1000) and loaded for detection into a S3e Cell sorter (Biorad, CA, USA). GFP detection (FL1 - Excitation 488 nm, Emission 525/30 nm) and mCherry detection (FL3 - Excitation 561 nm, Emission 615/25 nm) for approximately 48,000 cells per sample, was carried

out and analysed using FlowJo (v10.1r5) software (FlowJo, LLC, Oregon, USA). In order to analyse the library populations, a gating strategy was used in which reference regions were quantified. To do this, Level 1 plasmids were prepared with either GFP or mCherry assembled with low (J23114), low-medium (J23108) and high (J23100) promoter strengths, the PET RBS and BBa_B0015 terminator. Single plasmid clones were transformed into DH10**β.** The distribution of library populations was then estimated by overlapping with the known reference standards.

Cell-free protein synthesis

Based on the protocol from Sun *et al* [38], a cell-extract was prepared from *E. coli* Rosetta (DE3) pLysS and reactions were monitored in a CLARIOStar© (BMG, Germany) plate reader. For full details, please see SI text and Table S2.

***in vivo* growth and fluorescence measurements**

Promoter, RBS and terminator parts were characterised in *E. coli* DH10β. Single colonies were picked in triplicate and grown overnight at 30 °C in a 24-well plate in 2 mL of 2YT medium with antibiotics. Cell density was measured in a CLARIOStar© plate reader at 600 nm and cells were sub-cultured into 100 μL of 2YT with antibiotics in a 96-well Greiner plate to a starting OD_{600} of 0.05. Plates were sealed with a Breathe-Easy® membrane (Sigma) and grown at 30 °C for 6-12 hours at 600 rpm. OD_{600}, GFP and mCherry measurements were recorded every 10 min (see SI text). Data analysis of characterisation data was quantified as described previously with one modification [27]. Rates of GFP and mCherry production were quantified during exponential phase from an average of triplicate data over a 1 hour time period. For cell-free activity measurements only, mCherry fluorescence accumulates significantly later than GFP and was measured at a different time point.

Protein production and purification

Plasmids (pET15b or EcoFlex) were transformed into BL21 Gold (DE3) pLysS. A 10 mL 2YT overnight culture with antibiotics was grown overnight and 0.5 mL was sub-cultured into 100 mL 2YT and grown at 37 °C, 200 rpm until an OD_{600} of 1.0 was reached. IPTG (0.4 mM) was added and the culture was left for 16 hours at 19 °C, 200 rpm. Gravity flow immobilized metal ion chromatography (IMAC) purification was performed. Please see SI text for further details.

Associated content

Supporting Information

Table S1-S14 and Figures S1-S13, as well as oligonucleotide and DNA part sequences. Full plasmid sequences are available on AddGene (www.addgene.org/cloning/moclo/freemont-ecoflex). Excel files of DNA part characterisation and sequencing are also provided. For technical queries or requests please E-mail: simon.moore@imperial.ac.uk or sjmoore505@gmail.com.

Author information

Corresponding authors

For general enquiries, please contact Professor Paul Freemont: *Tel: +44 (0) 2075945327, Fax: +44 (0) 207 594 30, E-mail: p.freemont@imperial.ac.uk.

Author Contributions

S.M., K.P. and P.F. designed the study, analysed data, and wrote the manuscript. S.M. and H.L. assembled EcoFlex. S.M. H.L., R.K., S.C. and D.B. performed the experiments.

Notes

The authors declare no competing financial interests

Acknowledgements

SM and RK are postdoctoral researchers funded by the EPSRC grants EP/K038648/1 (Frontier Engineering) and EP/J02175X/1 (Flowers Consortium), respectively. HL is funded by a studentship award from Imperial College London. We would also like to thank Dr Lisa Goers and Dr James Arpino from Imperial College London for providing plasmids for cloning.

Figure legends

Figure 1. EcoFlex assembly. (A) Level 0 parts are held in pBP (bioparts – promoters, RBS, tags and terminators) and pBP-ORF (open reading frames) plasmids and assembled with BsaI and T4 DNA ligase into pTU1 destination plasmids (A, -B, -C, -D or -D_1, -E) forming Level 1 TUs. Genes are cloned with NdeI (5') and BamHI (3'), with the NdeI site providing an overlap between the start codon and RBS fusion site (CATA). Alternatively, if these sites are located within the gene, custom designed BsaI primers are used. Level 2 modules (2-5 TUs) were assembled with BsmBI and T4 DNA ligase into pTU2 plasmids, which were modified to provide flexibility for assembling 3-5 TUs into a module. In addition, Level 3 pathways (up to 20 TUs) can be assembled from compatible Level 2 modules into pTU3-A or pTU3-B using BsaI and T4 DNA ligase. (B) Assembly options (i) for purification tags and promoter choice, (ii) fusion site sequences and positions for Level 1 assembly of TUs. (C) Secondary module cloning for optimising library screening and minimising pathway variation. Table details compatible fusion sites and assembly options.

Figure 2 – Part characterisation *in vitro* and *in vivo*. A select range of promoters (A), RBS (B) and terminator variants (C) were tested for relative activity in *E. coli* DH10β and cell-free protein synthesis. The J23100 and PET RBS were used for normalisation of promoter and RBS data, respectively. For the terminator assay (Level 2 plasmids), GFP was assembled with a SJM901 promoter, PET RBS and a variable strength terminator. This was followed by a weak J23114 promoter, PET RBS, mCherry and BBa_B0015 terminator. Data is represented as normalised relative activity of GFP and mCherry, respectively. (D) The Sigma70 promoter library characterisation using cell-free protein synthesis. Inset picture shows visual GFP production from cell-free extracts with the SJM901 promoter in comparison to J23100.

Figure 3 – T7 protein production and His_6-tag purification of mCherry and CFP. *E. coli* BL21 Gold (DE3) plasmids strains were grown at 37°C and induced with 0.4 mM IPTG overnight at 19°C. Recombinant proteins were purified using IMAC gravity-flow resin and washed with increasing concentrations of imidazole. Figure abbreviations are as follows, M – protein marker, S – soluble extract, I – insoluble proteins and EF – elution fraction.

Figure 4 – Flow cytometry of GFP and mCherry pooled promoter and RBS libraries. Mixed promoter, RBS libraries were assembled with GFP and mCherry into pTU1-A and pTU1-B, respectively. Instead of picking single colonies, plasmid DNA was prepared directly from liquid culture, checked for purity by restriction digest and sequencing. Level 1 GFP and mCherry libraries were then assembled into pTU2-a and transformed into DH10β. Library cultures of Level 1 (Figure S7) and Level 2 were analysed by flow-cytometry using J23100, J23108 and J23114 promoter standards of GFP (labelled as G100, G108 and G114) and mCherry (labelled as C100, C108 and C114) as a reference of relative strength to estimate population distribution in both 3^2 (9 variants), 5^2 (25 variants) and 10^2 (100 variants) promoter-RBS combinations. These analyses are representative of ~48,000

cells per sample and these data were analysed using FlowJo (v10.1r5) software.

Figure 5 – Secondary module optimisation: Violacein pathway assembly. (A) The *vioA* and *vioE* genes were assembled with the J23114 promoter, PET RBS and BBa_B0015 terminator in pTU1-A and pTU1-B, respectively and then joined together in pTU2-a. The *vioA*-*vioE* fragment was sub-cloned by BsaI into BpiI digested pTU2^{S}-b to form a *vioA*-*vioE* secondary module. Level 1 promoter-RBS-terminator libraries of *vioB* (pTU1-A), *vioC* (pTU1-B) and *vioD* (pTU1-C) were then assembled. Inset sequencing trace demonstrates purity of the library. Library mixtures (pMB1, ColE1 and p15A backbones) were assembled into pTU2^{S}-*vioA*-*vioE*-b to complete the pathway. (B) 5 μL of assembly mixture was then transformed into 50 μL JM109 competent cells then grown overnight at 37 °C on 2YT and chloramphenicol plates. A range of white, brown, light purple and dark purple coloured colonies were obtained (Table 1). Some of the white colonies developed a pale purple shade after a further day of growth at 30°C. (C) A selection ColE1 and p15A variants (in as a triplicate biological repeat) were grown in 5 mL of 2YT liquid culture for high-throughput Bugs2Bases sequencing (Source Bioscience). A number of promoter-RBS regions with insertions are highlighted with an asterisk*. Crude violacein content was estimated by UV-Vis as described [40]. PC, positive control with all violacein genes under J23114 promoter, PET RBS and B0015 terminator. (D) Summary of sequencing results for p15A variants showing a mixture of promoter-RBS and terminator combinations. For classification, promoter, RBS and terminator are colour coded as strong, medium and weak based on GFP and mCherry characterisation data. L32 vioD terminator could not be identified, highlighted as a question mark. Asterisk indicates single base-pair deletions in p15A library. For full sequencing results, please refer to supporting information. Sequencing information is representative of a single culture after additional growth step performed at Source Bioscience.

Table 1. CFUs for violacein pathway optimisation with the secondary module

CFU phenotypes				
Origin of replication	**Violacein (High)**	**Violacein (Low)**	**PDV**	**White**
pMB1 (high)	27	7	18	48
ColE1 (medium)	88	6	57	115
p15A (low)	270	0	118	154

References

(1) Ran, N., Zhao, L., Chen, Z., and Tao, J. (2008) Recent applications of biocatalysis in developing green chemistry for chemical synthesis at the industrial scale. *Green Chem. 10*, 361.

(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. 9*, 1–9.

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

(4) Engler, C., and Marillonnet, S. (2011) Generation of families of construct variants using golden gate shuffling. *Methods Mol. Biol. 729*, 167–81.

(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) Weber, E., Gruetzner, R., Werner, S., Engler, C., and Marillonnet, S. (2011) Assembly of designer TAL effectors by Golden Gate cloning. *PLoS One 6*, e19722.

(7) Sarrion-Perdigones, A., Falconi, E. E., Zandalinas, S. I., Juárez, P., Fernández-del-Carmen, A., Granell, A., and Orzaez, D. (2011) GoldenBraid: An iterative cloning system for standardized assembly of reusable genetic modules. *PLoS One 6*.

(8) 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.

(9) 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.

(10) Engler, C., Youles, M., Gruetzner, R., Ehnert, T.-M., Werner, S., Jones, J. D. G., Patron, N. J., and Marillonnet, S. (2014) A golden gate modular cloning toolbox for plants. *ACS Synth. Biol. 3*, 839–43.

(11) Sarrion-Perdigones, A., Vazquez-Vilar, M., Palací, J., Castelijns, B., Forment, J., Ziarsolo, P., Blanca, J., Granell, A., and Orzaez, D. (2013) GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. *Plant Physiol. 162*, 1618–31.

(12) Agmon, N., Mitchell, L. A., Cai, Y., Ikushima, S., Chuang, J., Zheng, A., Choi, W.-J., Martin, J. A., Caravelli, K., Stracquadanio, G., and Boeke, J. D. (2015) Yeast Golden Gate (yGG) for the Efficient Assembly of *S. cerevisiae* Transcription Units. *ACS Synth. Biol. 4*, 853–9.

(13) Mitchell, L. a., Chuang, J., Agmon, N., Khunsriraksakul, C., Phillips, N. a., Cai, Y., Truong, D. M., Veerakumar, A., Wang, Y., Mayorga, M., Blomquist, P., Sadda, P., Trueheart, J., and Boeke, J. D. (2015) Versatile genetic assembly system (VEGAS) to assemble pathways for expression in *S. cerevisiae. Nucleic Acids Res. 43*, 6620–6630.

(14) Kakui, Y., Sunaga, T., Arai, K., Dodgson, J., Carazo-salas, R., Ji, L., Csika, A., and Sato, M. (2015) Module-based construction of plasmids for chromosomal integration of the fission yeast *Schizosaccharomyces pombe*. *Open Biol. 5, 150054.*

(15) Iverson, S., Haddock, T. L., Beal, J., and Densmore, D. (2015) CIDAR MoClo: Improved MoClo Assembly Standard and New E. coli Part Library Enables Rapid Combinatorial Design for Synthetic and Traditional Biology. *ACS Synth. Biol.*

(16) Chen, Y.-J., Liu, P., Nielsen, A. A. K., Brophy, J. A. N., Clancy, K., Peterson, T., and Voigt, C. A. (2013) Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat. Methods 10*, 659–664.

(17) Casini, A., Christodoulou, G., Freemont, P. S., Baldwin, G. S., Ellis, T., and MacDonald, J. T. (2014) R2oDNA designer: computational design of biologically neutral synthetic DNA sequences. *ACS Synth. Biol. 3*, 525–8.

(18) 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*, 354–60.

(19) Lou, C., Stanton, B., Chen, Y.-J., Munsky, B., and Voigt, C. A. (2012) Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat. Biotechnol. 30*, 1137–42.

(20) Anderson, J. C., Dueber, J. E., Leguia, M., Wu, G. C., Goler, J. A., Arkin, A. P., and Keasling, J. D. (2010) BglBricks: A flexible standard for biological part assembly. *J. Biol. Eng. 4*, 1.

(21) Beyer, H. M., Gonschorek, P., Samodelov, S. L., Meier, M., Weber, W., and Zurbriggen, M. D. (2015) AQUA Cloning: A Versatile and Simple Enzyme-Free Cloning Approach. *PLoS One 10*, e0137652.

(22) Storch, M., Casini, A., Mackrow, B., Fleming, T., Trewhitt, H., Ellis, T., and Baldwin, G. S. (2015) BASIC: A New Biopart Assembly Standard for Idempotent Cloning Provides Accurate, Single-Tier DNA Assembly for Synthetic Biology. *ACS Synth. Biol. 4*, 781–7.

(23) Torella, J. P., Lienert, F., Boehm, C. R., Chen, J.-H., Way, J. C., and Silver, P. A. (2014) Unique nucleotide sequence-guided assembly of repetitive DNA parts for synthetic biology applications. *Nat. Protoc. 9*, 2075–89.

(24) Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. a, and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods 6*, 343–345.

(25) Li, M. Z., and Elledge, S. J. (2012) SLIC: a method for sequence- and ligation-independent cloning. *Methods Mol. Biol. 852*, 51–9.

(26) Kelwick, R., Bowater, L., Yeoman, K. H., and Bowater, R. P. (2015) Promoting microbiology education through the iGEM synthetic biology competition. *FEMS Microbiol. Lett. 362*.

(27) Chappell, J., Jensen, K., and Freemont, P. S. (2013) Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. *Nucleic Acids Res. 41*, 3471–3481.

(28) Jürgen, B., Breitenstein, A., Urlacher, V., Büttner, K., Lin, H., Hecker, M., Schweder, T., and Neubauer, P. (2010) Quality control of inclusion bodies in *Escherichia coli*. *Microb. Cell Fact. 9*, 41.

(29) Ceroni, F., Algar, R., Stan, G.-B., and Ellis, T. (2015) Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods 12*, 415–418.

(30) Knaggs, A. R. (2003) The biosynthesis of shikimate metabolites. *Nat. Prod. Rep. 20*, 119–36.

(31) Balibar, C. J., and Walsh, C. T. (2006) *in vitro* biosynthesis of violacein from L-tryptophan by the enzymes VioA-E from *Chromobacterium violaceum*. *Biochemistry 45*, 15444–15457.

(32) Durfee, T., Nelson, R., Baldwin, S., Plunkett, G., Burland, V., Mau, B., Petrosino, J. F., Qin, X., Muzny, D. M., Ayele, M., Gibbs, R. a., Csörgo, B., Pósfai, G., Weinstock, G. M., and Blattner, F. R. (2008) The complete genome sequence of *Escherichia coli* DH10B: Insights into the biology of a laboratory workhorse. *J. Bacteriol. 190*, 2597–2606.

(33) Lee, M. E., DeLoache, W. C., Cervantes, B., and Dueber, J. E. (2015) A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth. Biol. 4*, 975–86.

(34) Sleight, S. C., Bartley, B. A., Lieviant, J. A., and Sauro, H. M. (2010) Designing and engineering evolutionary robust genetic circuits. *J. Biol. Eng. 4*, 12.

(35) Jones, J. A., Vernacchio, V. R., Lachance, D. M., Lebovich, M., Fu, L., Shirke, A. N., Schultz, V. L., Cress, B., Linhardt, R. J., and Koffas, M. a. G. (2015) ePathOptimize: A Combinatorial Approach for Transcriptional Balancing of Metabolic Pathways. *Sci. Rep. 5*, 11301.

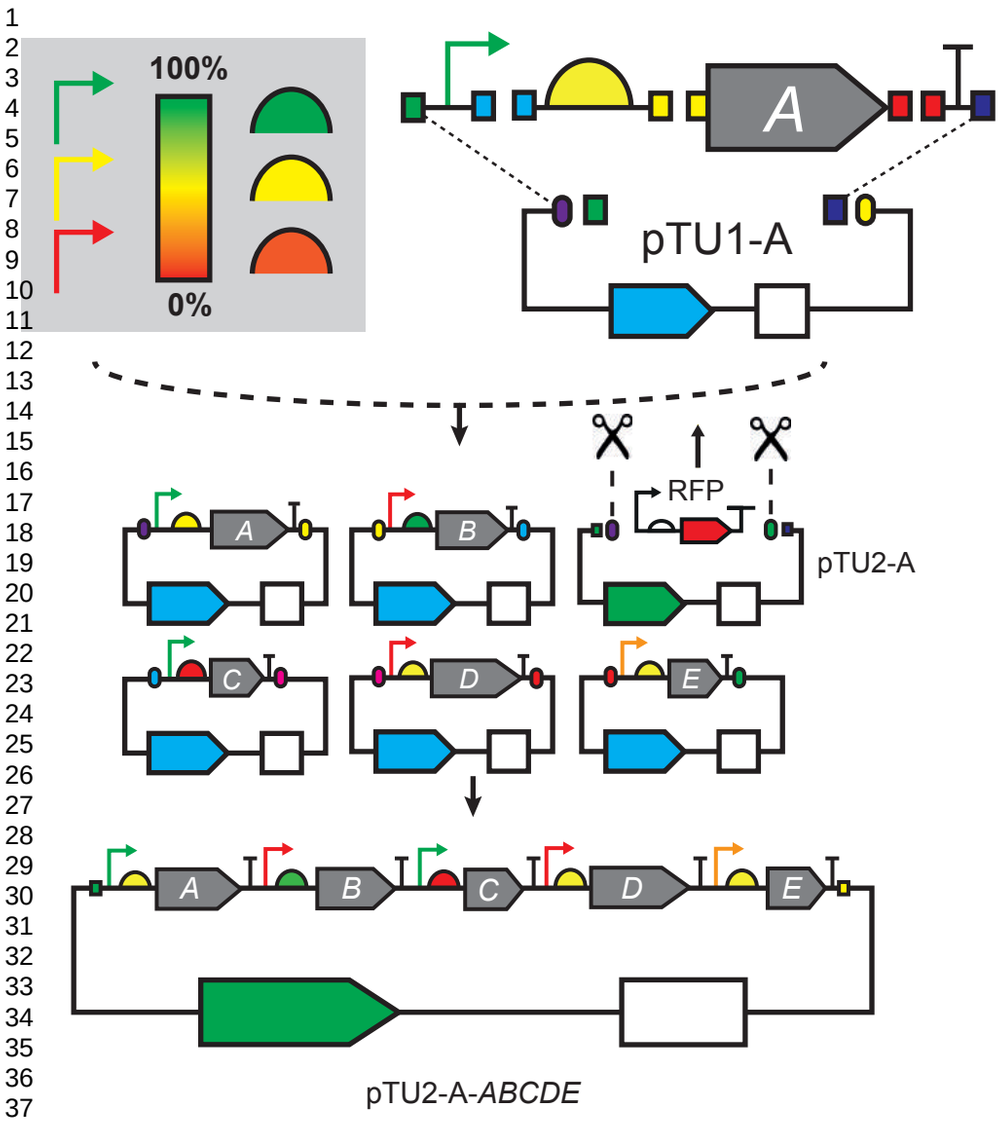
(36) Jeschek, M., Gerngross, D., and Panke, S. (2016) Rationally reduced libraries for combinatorial pathway optimization minimizing experimental effort. *Nat. Commun. 7*, 11163.

(37) Eisenstein, M. (2016) Living factories of the future. *Nature 531*, 401–403.

(38) Sun, Z. Z., Hayes, C. A., Shin, J., Caschera, F., Murray, R. M., and Noireaux, V. (2013) Protocols for implementing an Escherichia coli based TX-TL cell-free expression system for synthetic biology. *J. Vis. Exp. 50762*, e50762.

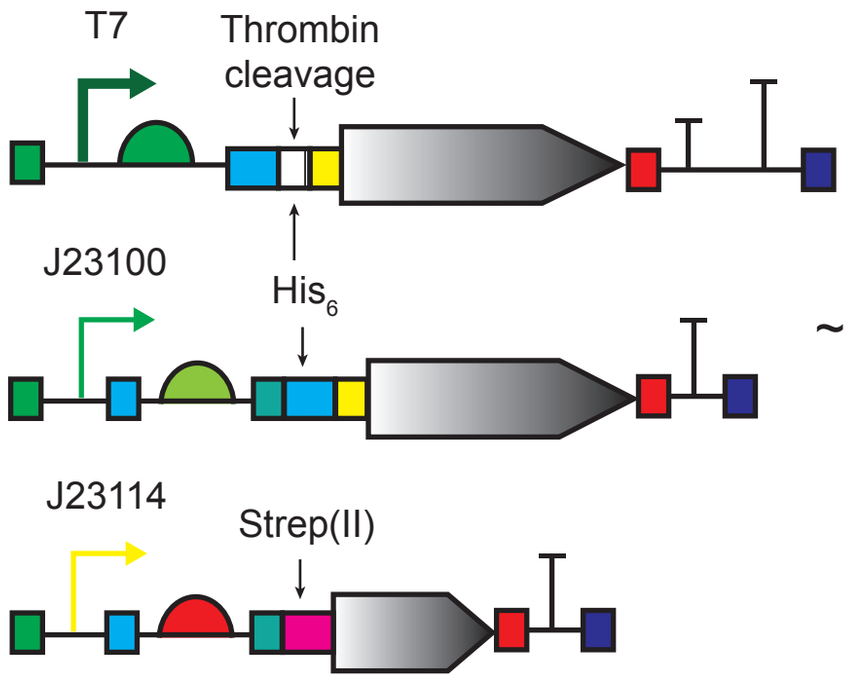
(39) Durán, N., Justo, G. Z., Ferreira, C. V, Melo, P. S., Cordi, L., and Martins, D. (2007) Violacein: properties and biological activities. *Biotechnol. Appl. Biochem. 48*, 127–33.

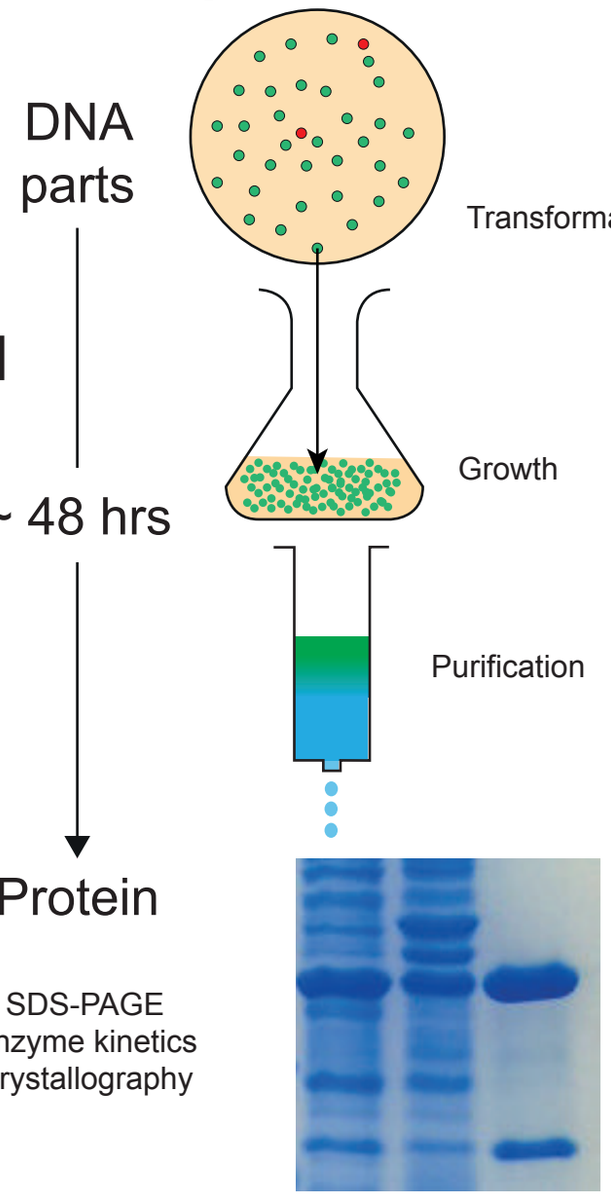
Pathway engineering



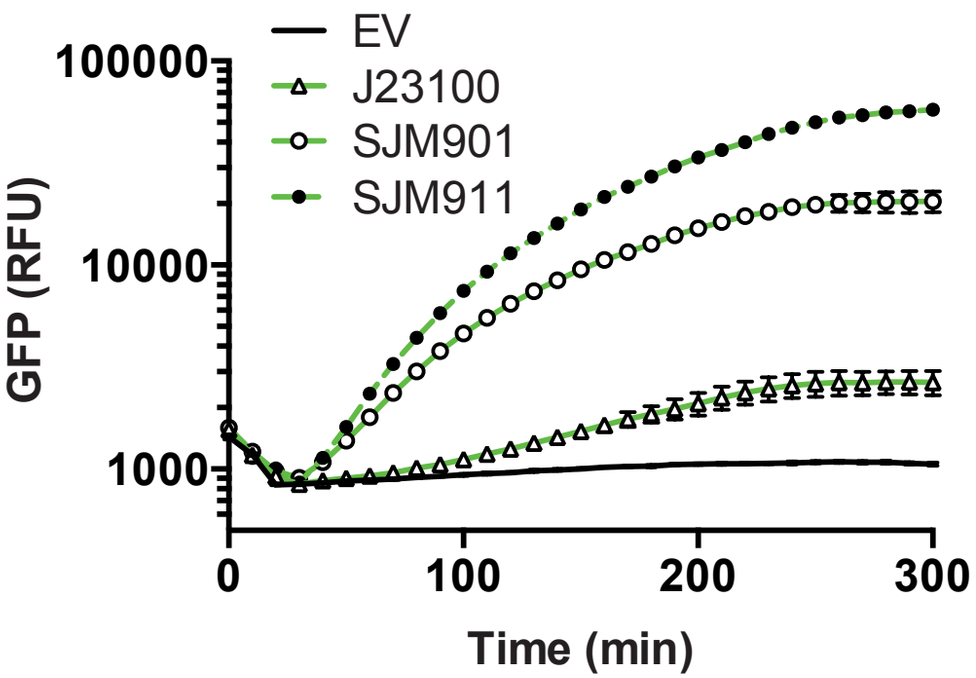
pTU2-A-*ABCDE*

Protein production and purification



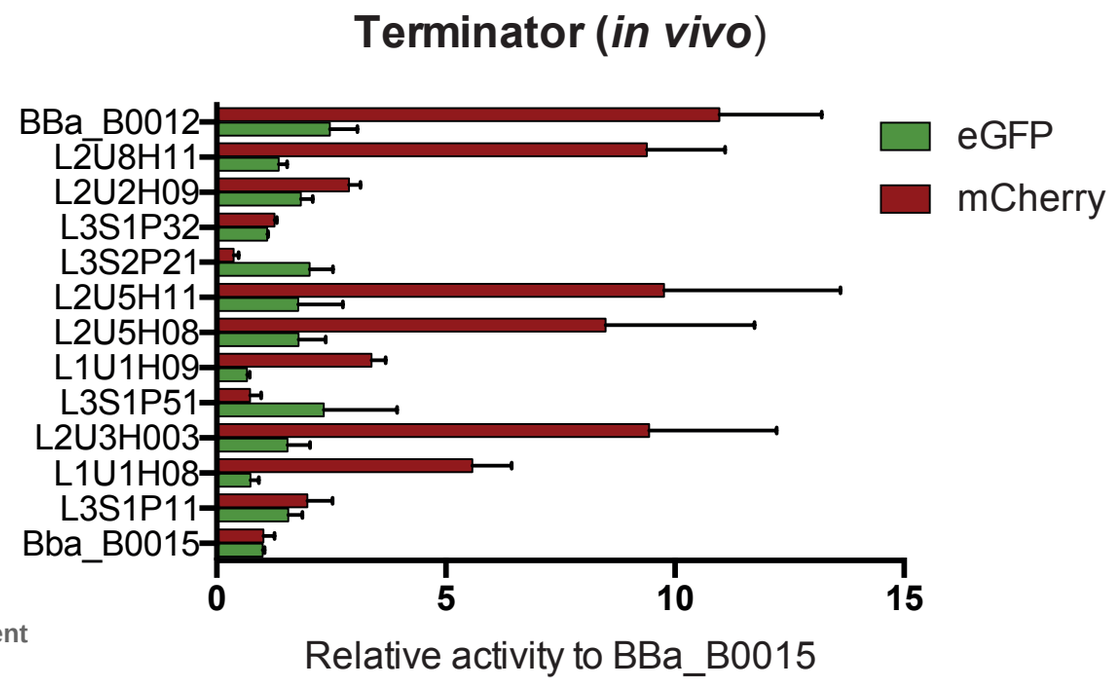


Cell-free transcription-translation



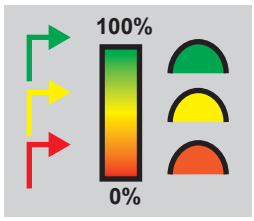
J23100 SJM901

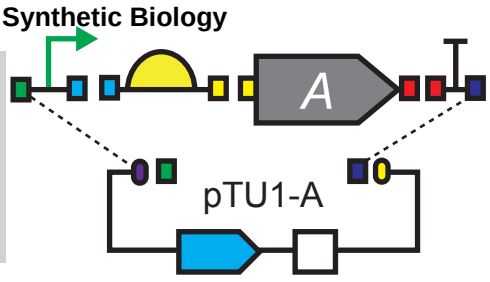
DNA part characterisation



A

Level 0
BioParts

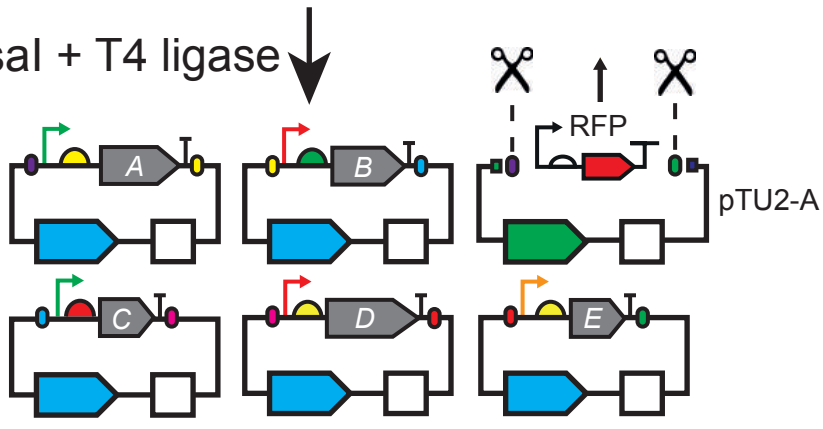




Part storage
pBP
pBP-ORF

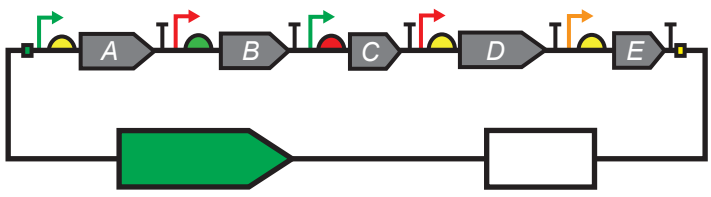
BsaI + T4 ligase

Level 1
Transcription
Units (TUs)

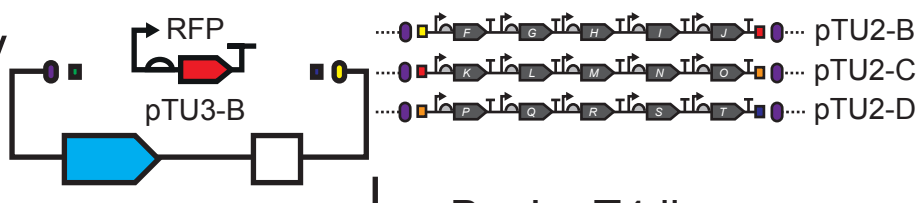


BsmBI + T4 ligase

Level 2
2-5 TUs



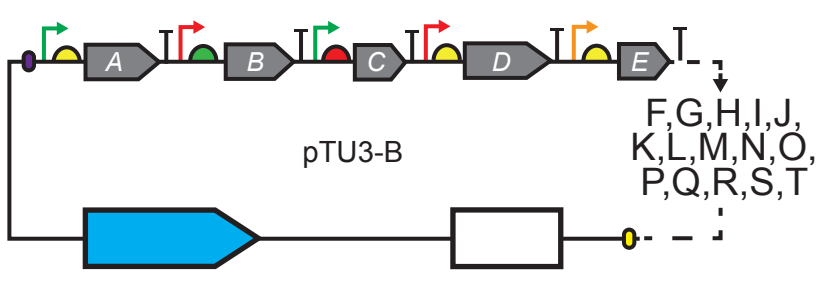
Sub-pathway



BsaI + T4 ligase

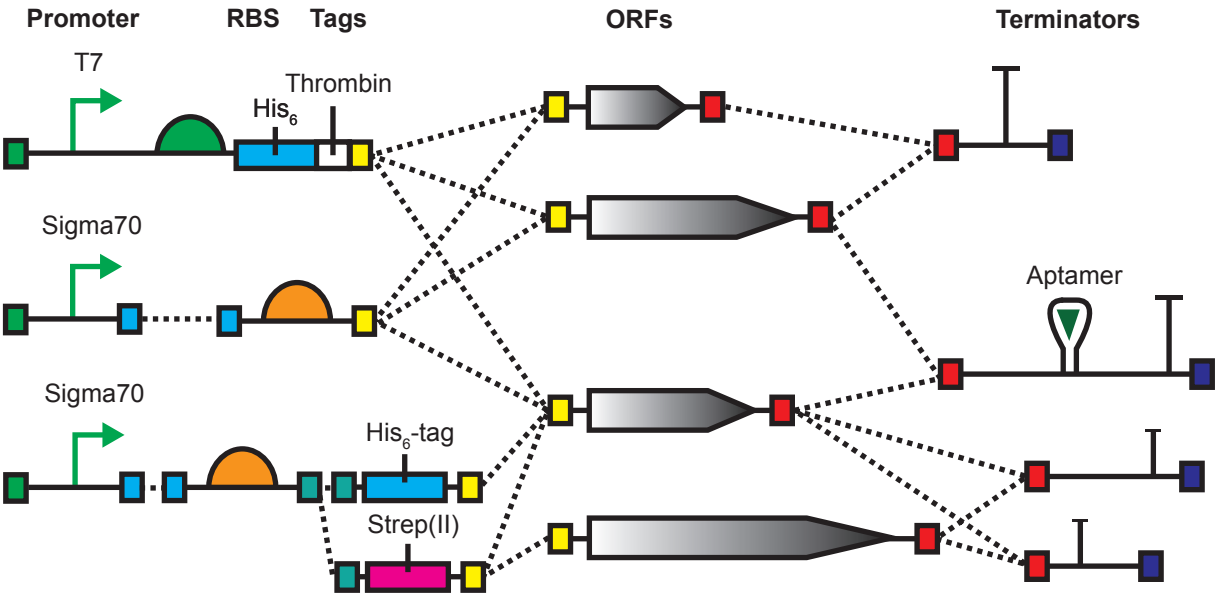
Level 3
6-20 TUs

Pathway

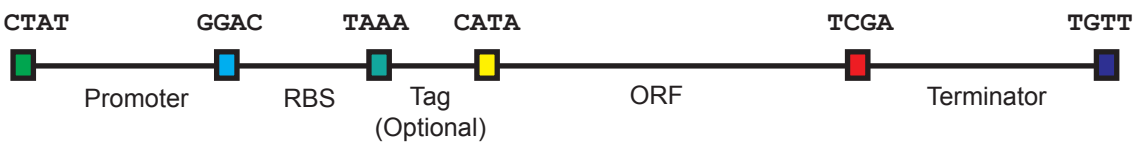


B

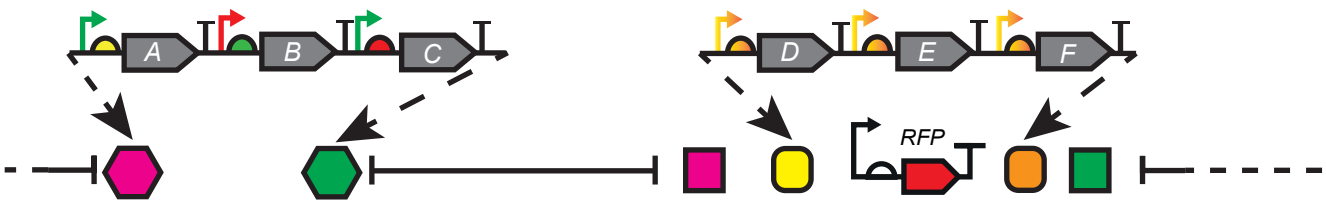
i.



ii.



C



Secondary module **Primary module**

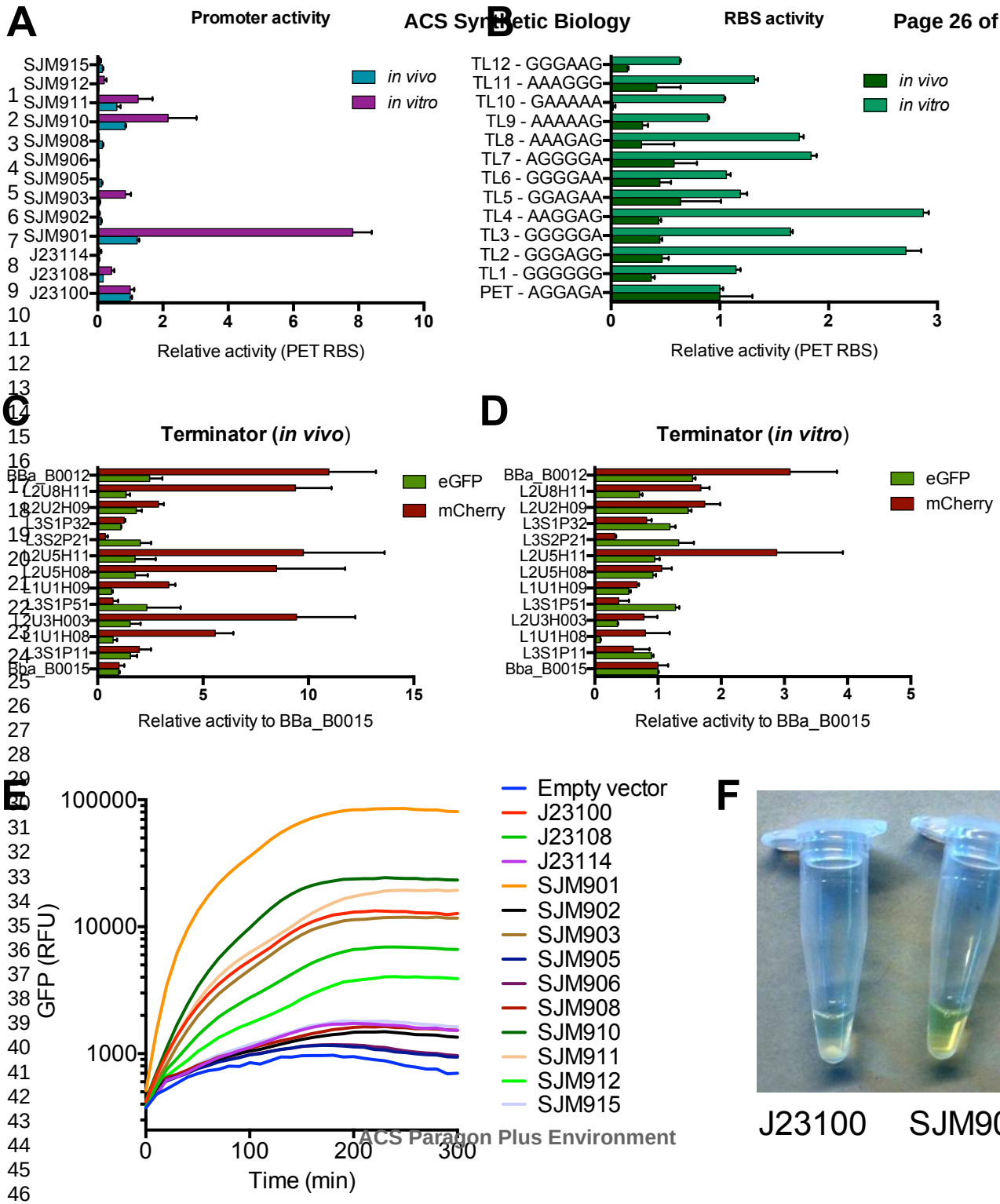
5'BpiI	3'BpiI
CTAT	GTAC
NA	NA
NA	NA
NA	NA
CTAT	GTAC
CTAT	GTAC

Fusion sites

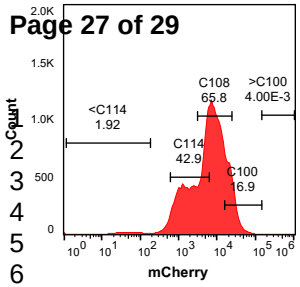
pTU2	TU's	5' BsaI	5' BsmBI	3' BsmBI	3' BsaI
A	4-5	CTAT	ATCT	TTAG	GTAC
B	4-5	GTAC	ATCT	TTAG	GGAC
C	4-5	GGAC	ATCT	TTAG	TCGA
D	4-5	TCGA	ATCT	TTAG	TGTT
a	2	CTAT	ATCT	CCGG	GTAC
b	3	CTAT	ATCT	GAAG	GTAC

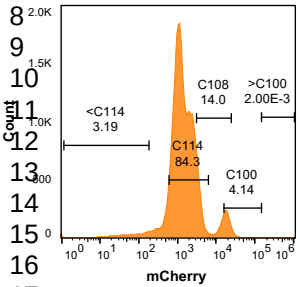
Fusion sites

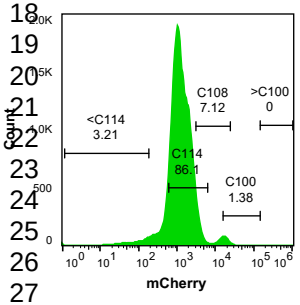
Restriction enzymes: BsaI BsmBI BpiI

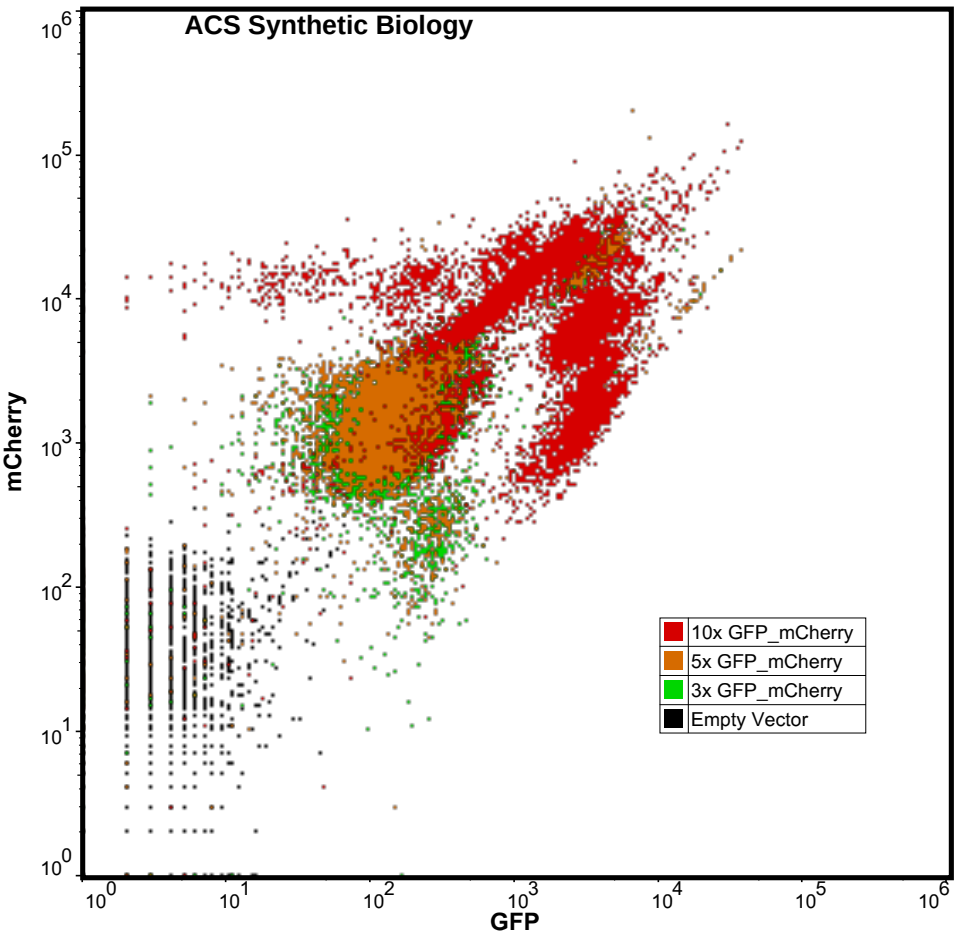


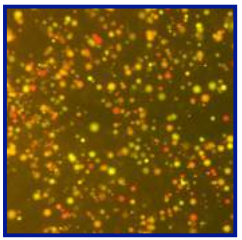
ACS Synthetic Biology

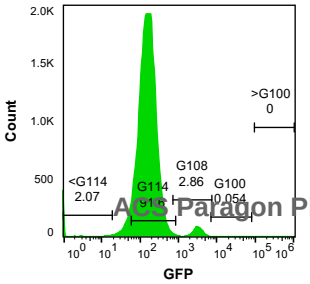


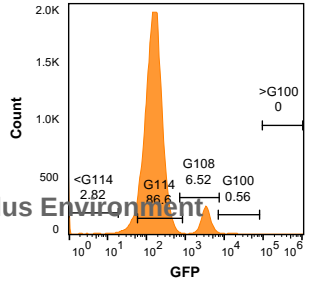


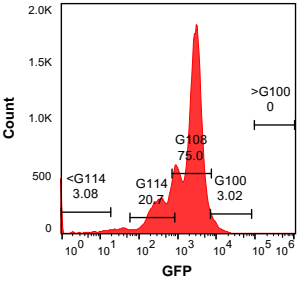


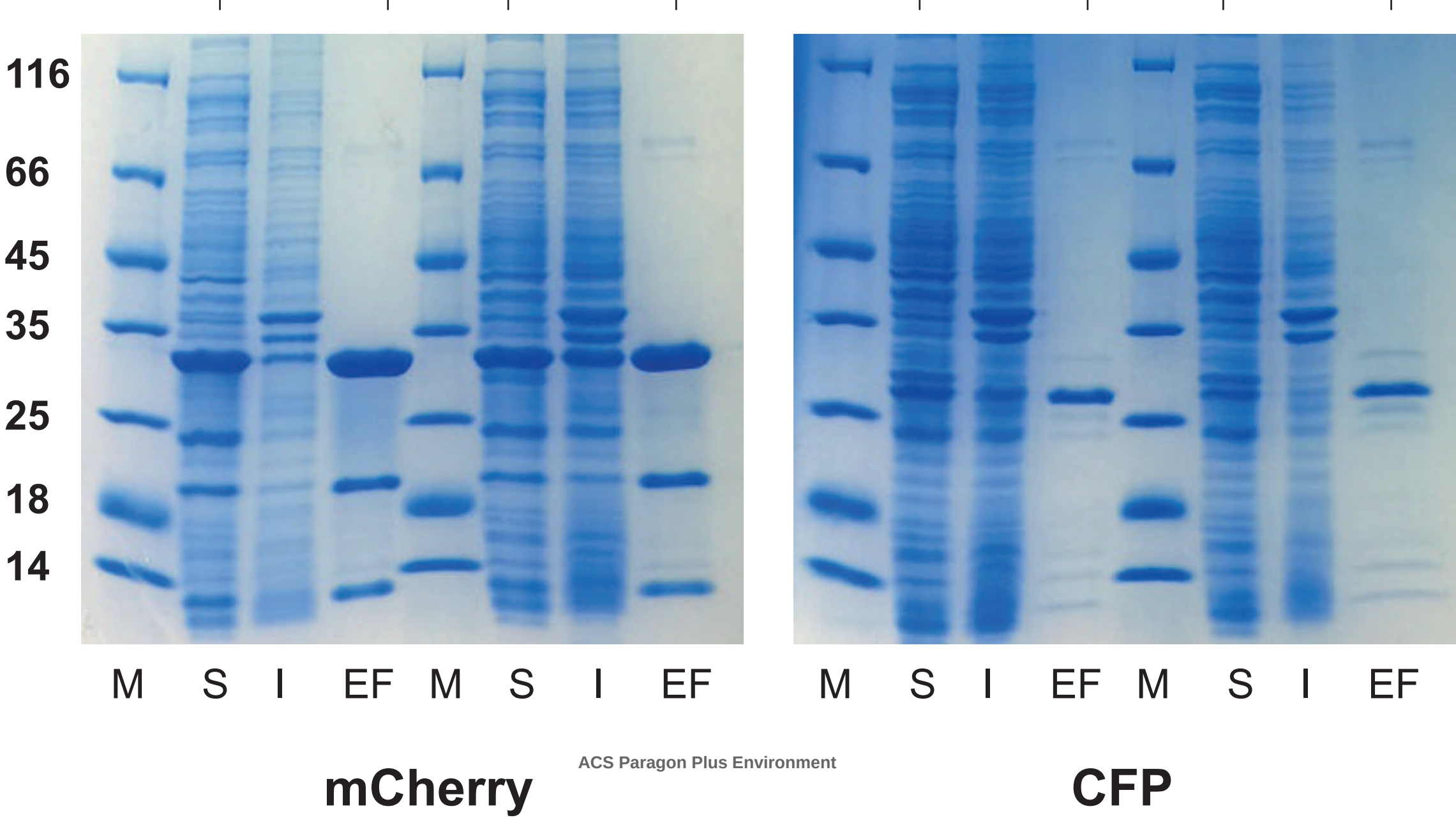


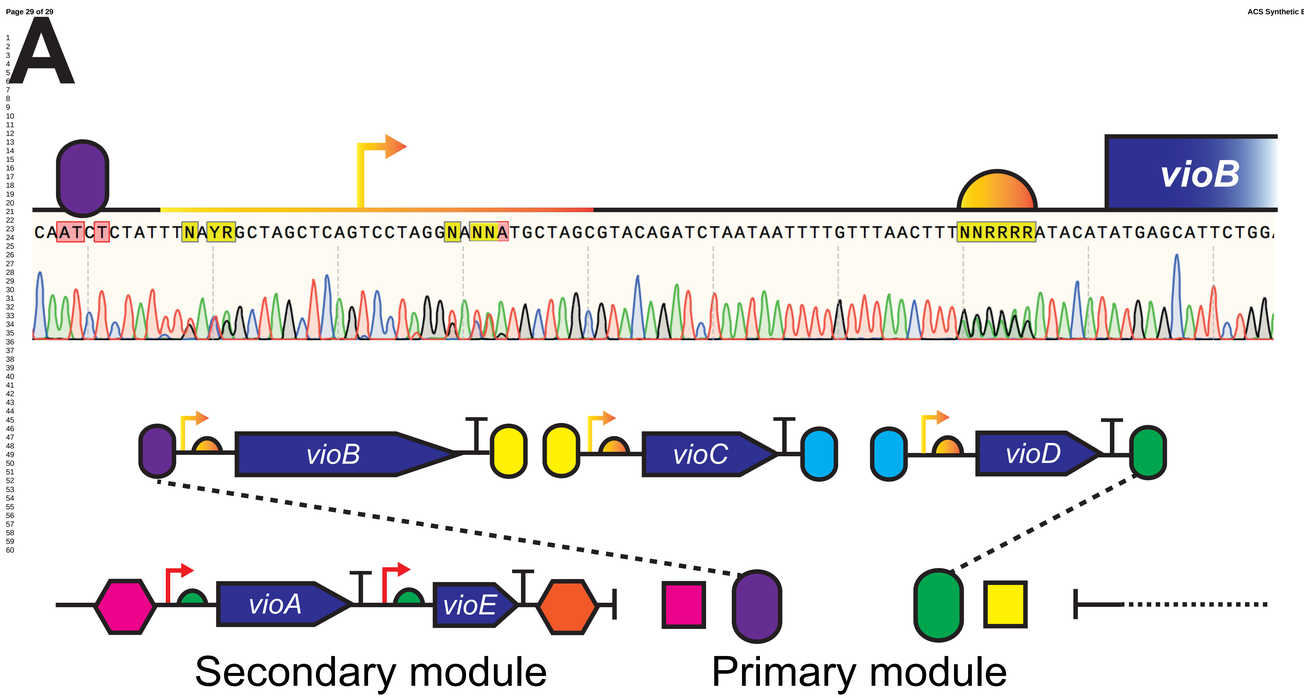


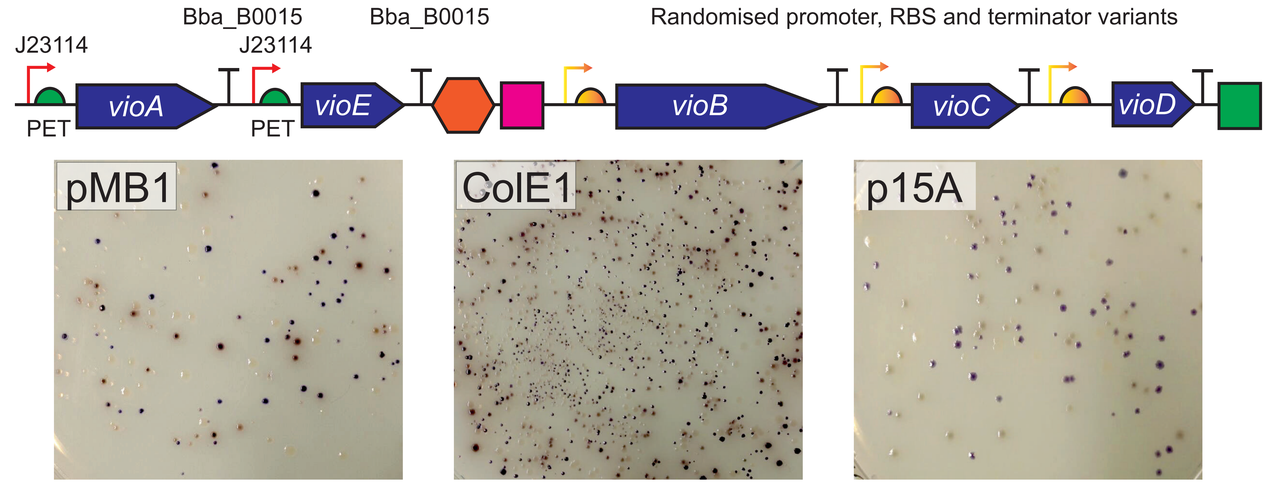


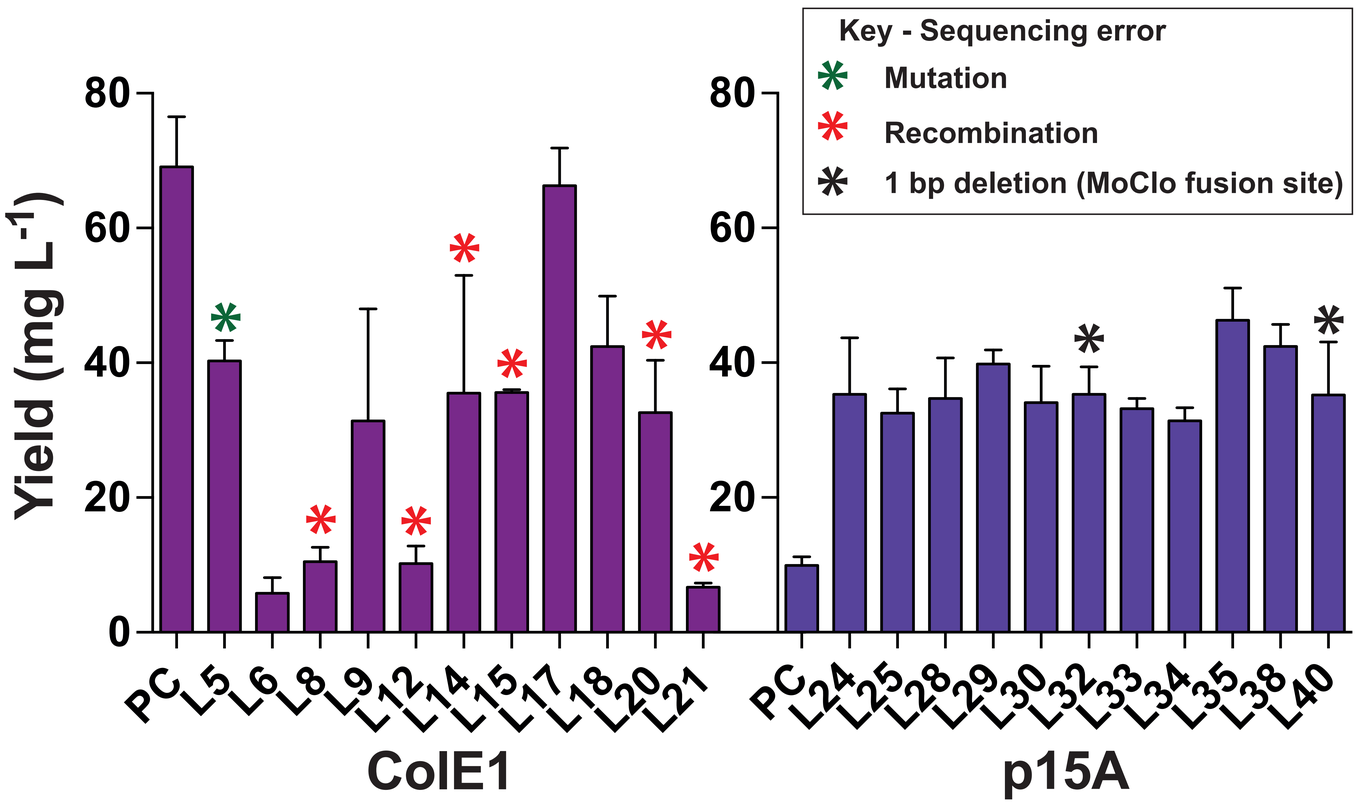





pET15b
EcoFlex
pET15b
EcoFlex
116
66
45
35
25
18
14
M S I EF M S I EF
M S I EF M S I EF
mCherry
CFP







D

p15A sequencing summary

ID	vioB			vioC			vioD		
	Promoter	RBS	Terminator	Promoter	RBS	Terminator	Promoter	RBS	Terminator
L24	902	8	L3S2P21	912	9	L3S2P21	902	8	B0015
L25	108	7	L3S2P21	912	9	L2U5H08	114	3	L2U5H08
L28	114	3	L2U5H08	114	1	L2U5H08	905	9	L2U5H08
L29	908	6	L3S2P21	906	4	L3S2P21	908	10	L2U5H08
L30	114	10	L2S2P21	901	10	L2U5H08	902	10	L2S2P21
L32	906	6	B0015	906	4*	L3S2P21	108	6	?
L33	906	10	L2U5H08	901	9	L2U5H08	908	1	L3S2P21
L34	114	6	L2U2H09	114	4	L2U5H08	114	2	L3S2P21
L35	902	6	L3S2P21	906	7	L2U5H08	915	9	L3S2P21
L38	915	1	L3S2P21	114	4	L2U2H09	902	10	L3S2P21
L40	114	3	L3S2P21	114	4	L2U2H09	905*	8	L2U5H08

Promoter/RBS Strength
Weak
Medium
Strong

Terminator Strength
Weak
Medium
Strong