ACS SyntheticBiology

Research Article

pubs.acs.org/synthbio

Assembly of Highly Standardized Gene Fragments for High-Level Production of Porphyrins in *E. coli*

Morten T. Nielsen,[†,‡,§] Karina M. Madsen,[†] Susanna Seppälä,[†] Ulla Christensen,[†] Lone Riisberg,[†] Scott J. Harrison,[†] Birger Lindberg Møller,[†,‡,§] and Morten H. H. Nørholm*[,†,‡]

[†]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kogle Allé 6, DK-2970 Hørsholm, Denmark

[‡]Plant Biochemistry Laboratory, Department of Plant and Environmental Sciences, [§]Center for Synthetic Biology: bioSYNergy, University of Copenhagen, Thorvaldsensvej 40, Frederiksberg C, Copenhagen Denmark

Supporting Information

ABSTRACT: Standardization of molecular cloning greatly facilitates advanced DNA engineering, parts sharing, and collaborative efforts such as the iGEM competition. All of these attributes facilitate exploitation of the wealth of genetic information made available by genome and RNA sequencing. Standardization also comes at the cost of reduced flexibility. We addressed this paradox by formulating a set of design principles aimed at maximizing standardization while maintaining high flexibility in choice of cloning technique and minimizing the impact of standard sequences. The design principles were applied to formulate a molecular cloning pipeline and iteratively assemble and optimize a six-gene pathway for protoporphyrin IX synthesis in *Escherichia coli*. State of the art production levels were achieved through two simple cycles of engineering and screening. The principles defined here are generally applicable and simplifies the experimental design of projects aimed at biosynthetic pathway construction or engineering.

KEYWORDS: *molecular cloning, biobricks, synthetic biology, uracil-excision cloning, porphyrin bioengineering*

The ability to recombine DNA with restriction enzymes and ligase moved molecular biology into a new era and was pivotal in the birth of biotechnology. More recently, whole genome sequencing has revolutionized molecular biology and paved the way for fields such as metabolic engineering and synthetic biology. Exploitation of the unprecedented availability of sequence information seems now mainly limited by our ability to build, test, and analyze variants of constructs, genes, and pathways. Accordingly, a range of advanced methods for DNA cloning have recently emerged, addressing the more complex demands.[1−5] Moreover, the need to standardize molecular cloning and share biological parts have been strongly emphasized.[6] A prime example is the BioBrick standard proposed by Tom Knight[7] that was followed by a parts sharing initiative heavily promoted by the iGEM synthetic biology competition. Several alternative BioBrick formats and technologies have subsequently been described (BBF RFC process. http://openwetware.org/wiki/The_BioBricks_Foundation:RFC). However, the history of BioBrick exposes two inherent complications in the pursuit of standardizing molecular cloning: (1) newly developed molecular cloning methodologies are not *per se* backward compatible, which compromises the benefits of part sharing and (2) standardization invariably results in addition of short sequences mediating the standardized DNA assembly. We decided to face these challenges, by formulating a set of design principles to guide the construction of a standardized DNA assembly pipeline. With the proposed design principles, the level of standardization is maximized while maintaining high flexibility in choice of cloning technique and minimizing the impact of standard sequences. The approach incorporates a substantial degree of variability and opportunities for custom design. As test case for our cloning pipeline, we selected production of protoporphyrin IX in *Escherichia coli*. Porphyrins have several properties that make them attractive as a model system: (1) the red color of porphyrins allows for an easy optical assay; (2) the fluorescent properties of porphyrins allow specific and sensitive detection and relative quantitative comparison by HPLC; (3) authentic standards are available for several porphyrins; and (4) the enzymes involved in porphyrin biosynthesis are thoroughly characterized.[8] Protoporphyrin IX is ubiquitously distributed in nature as a precursor in the synthesis of essential prosthetic tetraporphyrin groups such as in b- and c-type cytochromes, peroxidases, chlorophylls, cytochrome P450s, and hemoglobins. Moreover, porphyrins, particularly the manganese containing ones, have application potential within the biomedical field as drug precursors and modulators of cancer cell growth.[9]

Received: January 21, 2014
Published: June 6, 2014

ACS Publications © 2014 American Chemical Society

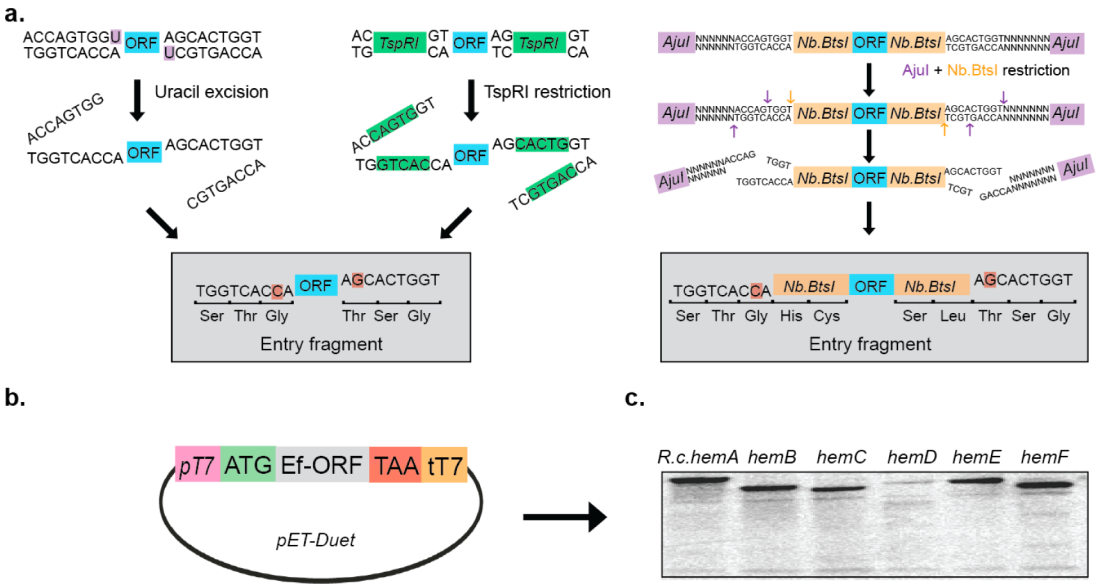

Figure 1. Activation of entry format fragments. The entry format is defined by the 9-bp single stranded overhang (sso) specified in the design principles. The 9-bp can be formed by a number of PCR and restriction enzyme based techniques. (a) Left: PCR based technique where uracils are incorporated in oligo nucleotides and sso's are generated by enzymatic removal of the uracils after the PCR. The uracils mediating sso formation are highlighted in purple. Middle: TspRI restriction. Right: Combined type-II-S restriction endonuclease (here AjuI) and nicking endonuclease (here Nb.BtsI). A consequence of this approach is that the recognition sites for the nicking endonuclease are also translated resulting in addition of two amino acids on both sides of the ORF. The resulting entry fragments with processed sso's are illustrated below. One base pair difference in the upstream and downstream sso's ensures directional cloning (highlighted in red). The amino acids encoded by the 9 bp entry format sso are displayed below the nucleotide sequence. (b) Illustration of the entry vector with entry-ORF inserted. The entry vector adds a T7 promoter, a start codon, a stop codon, and a T7 terminator to the cloned ORF (these elements are highlighted in pink, green, red and orange respectively). (c) Results of T7 dependent ^{35}S-methionine labeling of heme biosynthetic genes *hemA*–*F*.

Porphyrins are synthesized from glycine and succinyl-CoA through the Shemin pathway[10,11] or from tRNA loaded with glutamate through the C5-pathway.[12−14] δ-Aminolevulinic acid (ALA) is the first common intermediate in the two pathways. *E. coli* naturally utilize the C5-pathway[15] and introduction of the Shemin pathway by heterologous expression of δ-amino-levulinic acid synthase (*hemA*)[16,17] has previously been exploited to obtain high ALA titers.[18] In this manuscript, we describe a standardized yet generic pipeline for DNA cloning and demonstrate its feasibility for pathway engineering by constructing a pair of plasmids that enable high production of protoporphyrin IX (80 μM). The plasmids are selected through two engineering cycles based on screening of pathway variants, balancing expression of the six biosynthetic genes with semirandomized Shine-Dalgarno (SD) sequences.

RESULTS AND DISCUSSION

Design Principles for Cloning Pipeline. The overall purpose of our cloning pipeline is to achieve streamlined, robust, high-throughput cloning with a minimal requirement for *de novo* design. In addition, the pipeline design, and underlying principles, should be sufficiently generic to be compatible with a range of advanced cloning techniques. The major principle that we arrive at to fulfill these criteria is to standardize at two levels: (1) define an entry format and corresponding entry vector to achieve fully standardized incorporation of new DNA-parts and (2) design a general strategy for downstream parts assembly. The overall design principles that we arrive at for the entry format is that all entry-parts will be flanked by defined 9-bp sequences that facilitate directional ligase independent cloning with both PCR and restriction enzyme based technologies (Figure 1a). Specifically, the 9-bp flanking sequences are designed to contain recognition sequences for TspRI. TspRI is a unique restriction enzyme that generates a 9-bp single stranded overhang, thus providing a one step, error free method for activation prior to cloning (Figure 1a). The major advantage of this approach is that synthetic genes, free from internal TspRI sites, can be subcloned in a PCR-error-free manner. The flexible entry format can also be created with simple PCR-based methods such as uracil-excision,[19,20] enzyme free cloning,[21] and PliCing.[5] These methodologies will likely be useful when genes are cloned from the native source DNA. As a third alternative, compatible entry fragments can be created using a methodology that resembles type-II-S restriction enzyme-based technologies such as MoClo[22] and Golden Gate.[23] This third alternative relies on addition of internal 6-bp *Nb.BtsI* nicking enzyme recognition sites and is therefore ligase independent. Type-II-S restriction enzyme-based technologies may be particular useful in robotics-based systems. Homology/recombination-based cloning methodologies such as isothermal *in vitro* recombination,[1] ligase independent cloning,[24] or ligase cycling reaction[25] could also be use the generate entry clones because these are highly flexible and sequence independent technologies. The second principle is that protein coding sequences are designed without start and stop codons. This ensures flexibility in terms of modifications such as addition of tags, linkers, or fused gene constructs. A consequence of this is that the 9-bp sequences that facilitate cloning will be translated and could potentially interfere with protein function. To minimize this risk, the TspRI-containing 9-bp sequences are designed to encode only serine, threonine, and glycine. Serine and glycine are the most commonly used residues in protein linker sequences,[26] and threonine only differs from serine by one methyl group. Finally, we decided to

dx.doi.org/10.1021/sb500055u | *ACS Synth. Biol.* 2015, 4, 274−282

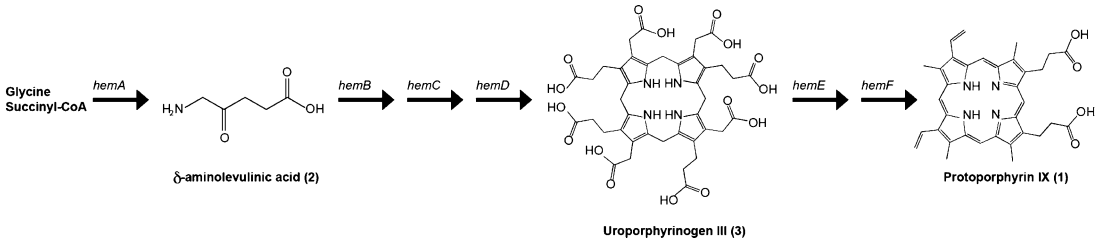

Figure 2. Biosynthetic route for protoporphyrin IX by the Shemin pathway. Chemical structures of the expected products of *hemD*, uroporphyrinogen III and *hemF*, protoporphyrin IX as well as the first dedicated intermediate δ-aminolevulinic acid are displayed.

utilize a T7-based expression vector as entry plasmid backbone. This introduces a simple quality checkpoint for proper translation of protein encoding entry fragments based on an assay for specific T7-dependent expression in the presence of rifampicin.[27] In our case, we utilized a pET-DUET-1 backbone as entry vector. The backbone was modified to position a start codon in optimal distance from a consensus SD sequence and immediately upstream of the 9-bp entry sequence to ensure that no unwanted amino acids were included in the resulting protein. In principle any backbone can be prepared for hosting BioBricks in this format by a plethora of techniques including PCR, TspRI or type-II-S restriction enzymes (Supporting Information Figure S1).

Application of the Standardized Cloning Pipeline to Protoporphyrin IX Production. With the basic principles established we decided to utilize the cloning pipeline to explore aspects of pathway engineering in *E. coli* using protoporphyrin IX as test case. We introduced the Shemin pathway in *E. coli* by heterologous expression of the model ALA synthase, *hemA*, from *Rhodobacter capsulatus*, as previously described.[18,28] The remaining enzymes required for protoporphyrin IX biosynthesis, encoded by *hemB*, *hemC*, *hemD*, *hemE*, and *hemF*,[8] were acquired from *E. coli*. A schematic overview of the biosynthetic pathway is presented in Figure 2. All six genes were inserted in the pET-Duet-1 based entry vector as entry fragments by uracil-excision and transformed in *E. coli* BL21 DE3 for verification of transcription and translation by ^{35}S-methionine labeling (Figure 1b). Indeed, translation of all six genes was confirmed. *hemD* appeared to be translated at a lower level indicating a possible bottleneck in pathway engineering (Figure 1c).

Standardized Assembly of Multiple Fragments. The next step in our pipeline construction was to standardize assembly of several fragments. Minimal *de novo* design, to avoid DNA assembly being the experimental bottleneck, and compatibility with multiple techniques and flexibility in terms of number of assembled fragments were the drivers for our design principles. At this stage, we focused on PCR based techniques because they allow highly flexible assembly of multiple fragments in a single cloning reaction.[1,5,20] We arrived at the following simple design rules: (1) entry fragments should be PCR amplified exclusively by generic oligo nucleotides, which is possible due to the standardized entry formats allowing generic oligo nucleotides annealing in the common Gly-, Ser-, and Thr-encoding regions, and (2) linking regions that define specificity in multifragment assemblies should be designed with maximal sequence diversity and confined length (8−10bp) (Table 1). These principles allow generation of a very large number of assemblies using only a small set of benchmarked oligo nucleotides, thus minimizing *de novo* optimization and failed assemblies due to erroneous oligo nucleotides. The approach ensures backward compatibility as long as the linkers are maintained, making it straightforward to expand the assembly at a later stage. Moreover, the short overlap sequences allow for incorporation of control elements (e.g., required for transcription and translation) while keeping the oligo nucleotides relatively short. This means that the gene expression levels can be tailored using permutations of the generic oligo nucleotides with, for example, variable promoters, SD-sequences, leader sequences, or peptide tags. In summary, generic oligo nucleotides contain three defined regions: An annealing part, a control element, and an assembly linker (Figure 3a). In contrast to the entry clone strategy, the described second-level multigene assembly strategy comes at the expense of flexibility in terms of cloning technology.

Table 1. Generic Linkers Used for Multi-Fragment Assembly

linker ID	sequence (sense)	complementary sequence (antisense)	length (bp)
entry fragment upstream	ACCAGTGGT	ACCACTGGT	9
entry fragment downstream	AGCACTGGT	ACCAGTGCT	9
generic linker A	AGTCGGTGT	ACACCGACT	9
generic linker B	AAGCAGCGT	ACGCTGCTT	9
generic linker C	AGACGTCAT	ATGACGTCT	9
generic linker D	AGGTCTGAGT	ACTCAGACCT	10

Assembly of Tetrapyrrole Core Biosynthesis Genes. The proof of concept pathway engineering was conducted as a screening study with constitutive promoters (p*trc*) and diversity at the translational level mediated by semirandomized SD-sequences (NNNNNGGAN). The core GGA motif in the SD was fixed to reduce the occurrence of nonfunctional sequences as previously described.[29] The semirandom SD was included as a control element in the upstream- and p*trc* in the downstream-oligo nucleotide for the entry clone, while a p*trc* and a transcriptional terminator was included as control elements in the backbone oligo nucleotides (Figure 3a). Entry clones of *R. capsulatus hemA*, *hemB*, *hemC*, and *hemD* were PCR amplified individually with the linker pairs A−B, B−C, C−D, and D−Ec and assembled with a PCR amplified *pCDF* backbone harboring the Ec−A linkers (Figure 3b) and transformed into *E. coli* DH5*α*. The initial screen for porphyrin producing transformants was conducted directly on the LB-plate used for transformation. Simple optical inspection revealed highly diverse levels of porphyrin production among the transformants (Figure 4a). Approximately 10% of the colonies were intensely red-colored, 10% were white and the remaining had an intermediate color phenotype. Twenty-four *E. coli* colonies spanning the range of porphyrin producers were selected for further analysis. The *pCDF-hemAD* plasmids were purified and sequenced to establish the efficiency of the uracil-excision cloning and map the variance in SD-sequences. The success

dx.doi.org/10.1021/sb500055u | *ACS Synth. Biol.* 2015, 4, 274−282

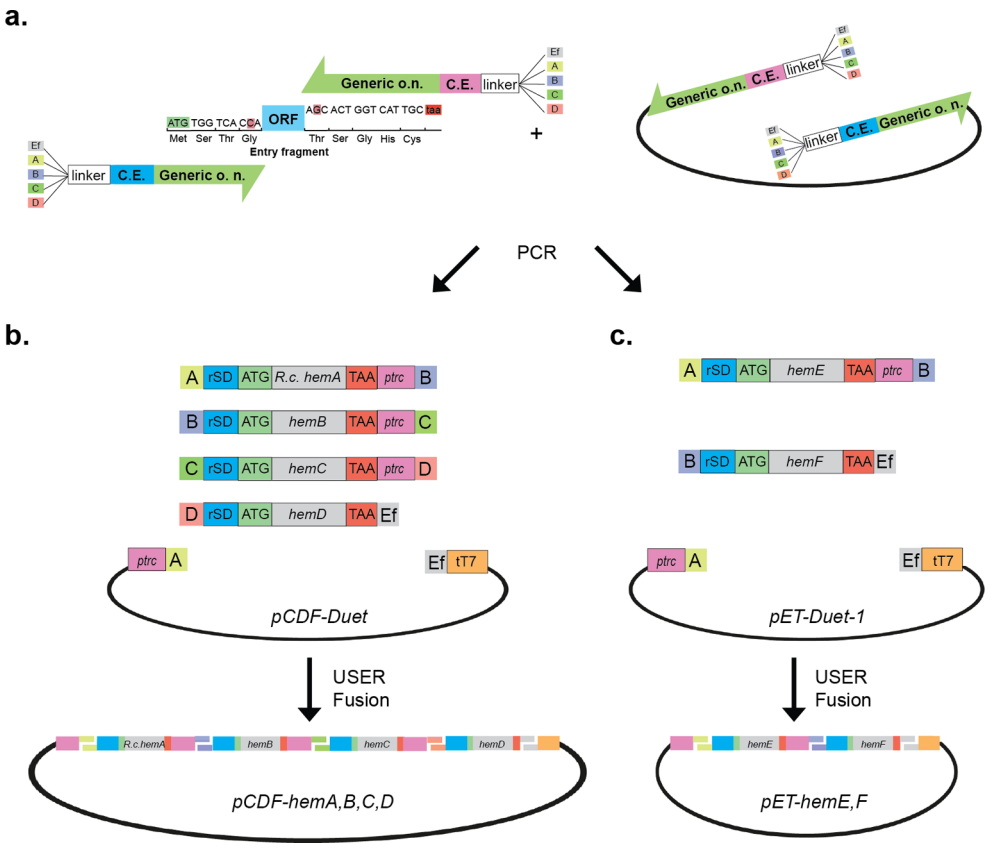

Figure 3. Schematic overview of multifragment assembly from entry clones. (a) Illustration of entry clone (left) and backbone (right) amplification by generic oligo nucleotides. The three elements of the generic oligo nucleotides are depicted as follows: annealing part (green), control elements (blue upstream, purple downstream), and linker (white box). The five variants of the linker (A–D, Entry fragment (Ef)) are shown in extension of the oligo nucleotide. Sequences of the linkers are given in Table 1. (b, c) Schematic overview of assembly process for four (b) and two (c) entry fragments. Entry fragments *R. capsulatus hemA* (*R.c.hemA*), *hemB*, *hemC*, and *hemD* were assembled into amplified *pCDF-Duet*, whereas entry clones *hemE* and *hemF* were assembled into *pET-Duet*. Complementary linkers are illustrated with identical colors and letter codes. The semirandomized SD sequence (rSD) was included in the upstream control element (C.E.) and the *ptrc* promoter downstream in all entry clones. Amplified backbones harbor a *ptrc* upstream and a T7 transcriptional terminator (tT7) downstream. The schematic overviews are not drawn to scale.

rate of the cloning, defined as clones containing all four genes in correct order, was 79% (19 out of 24 clones as verified by sequencing). The five remaining clones contained only one gene, which was either *hemA* (3 clones) or *hemD* (2 clones). This was not surprising, as these genes contain one linker with perfect match to the vector, but the effect could also be caused by instability due to the repetitive sequences used. The predicted strength of the 24 different combinations of SD-sequences were calculated using the Ribosome Binding Site Calculator[30,31] (Table 2), but no statistically significant correlations between predicted RBS strength and porphyrin levels were observed.

Chemical Characterization of Stage-One Clones. All 24 selected *E. coli* colonies were subjected to detailed analyses by chromatography-based methods. HPLC coupled fluorescence detection (FLD) using excitation and emission wavelength selective for porphyrins (410 nm excitation/620 nm emission) revealed the presence of several signals (Figure 4b). Identification of the fluorescent compounds was attempted with Fourier transform mass spectrometry (FT-MS) analysis. The late eluding compound, **1**, was identified as protoporphyrin IX based on both the retention time (1.6 min) and comparison with the mass spectra of an authentic standard (Figure 4c). The major signal obtained, however, turned out to contain two related coeluding compounds **4** and **5** sharing an ion of 1169 m/z and with dominant ions of 577 and 631 m/z, respectively (Supporting Information Figure S2). Neither of these ions matches those derived from expected intermediates. The FT-MS analyses support that both **4** and **5** contain nitrogen and are capable of binding iron in the high molecular mass form (1169 m/z) and the FLD signal indicate a tetrapyrrole structure. Taken together with the exclusive co-occurrence of protoporphyrin IX with elevated levels of **4** and **5** (Figure 4d), this evidence suggests that **4** and **5** are indeed porphyrins. A metabolite with similar mass as **1** (557 m/z) has previously been observed in co-occurrence with protoporhyrin IX in *E. coli* constitutively expressing *R. capsulatus hemA*.[32] The metabolite was identified as heme based on proposed signature ions, however, the accurate mass and isotope patterns available to us demonstrate that iron is not bound in the low molecular form (557 m/z), but exclusively in the high molecular form (1167 m/z) which rules heme out as a target. Previous studies identified uroporphyrin I and III as the dominant species of porphyrins upon constitutive expression of *R. capsulatus hemA* with *E. coli hemB–D*.[28] However, in our experiments, neither uroporphyrin I nor III was detectable in any of the analyzed clones (data not shown). Absence of uroporphyrin III could be explained by insufficient *hemD* activity as the *in vitro* translation data suggested that *hemD* was not as efficiently expressed as the remaining genes. The presence of unknown tetrapyrrols are

dx.doi.org/10.1021/sb500055u | *ACS Synth. Biol.* 2015, 4, 274–282

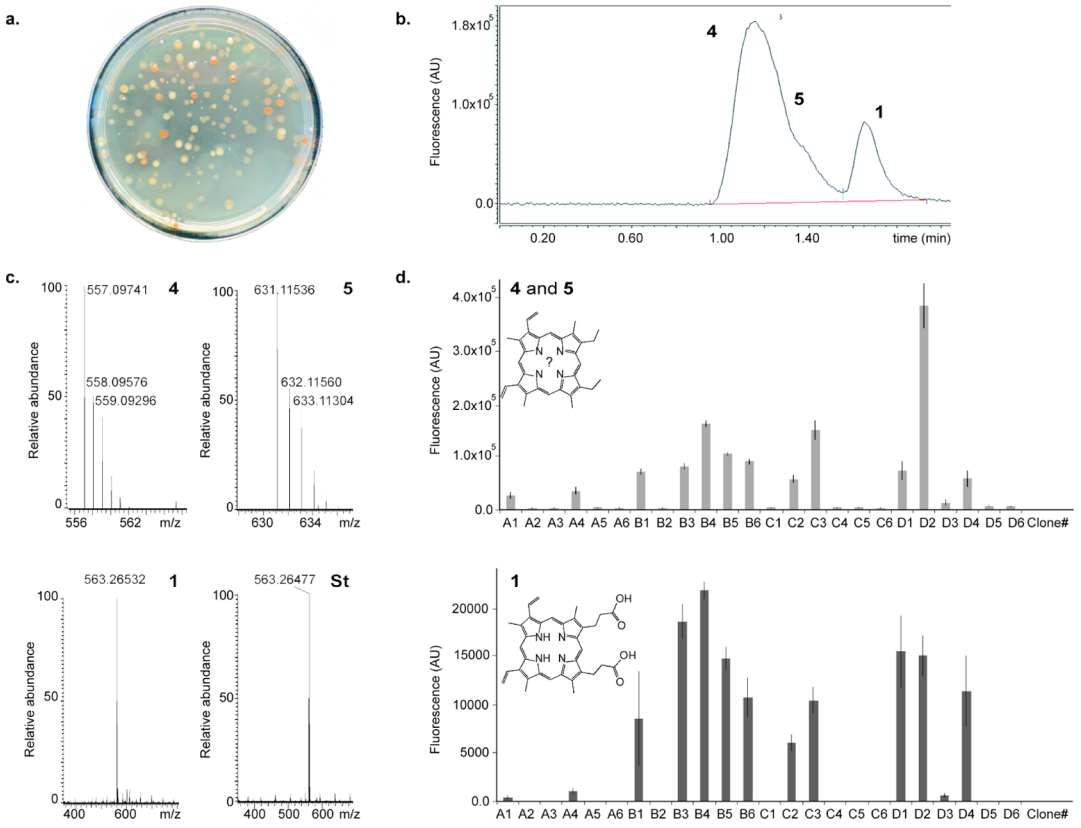

Figure 4. Porphyrin production in stage-one (hemA−D) clones. (a) Colonies were screened by optical inspection directly on plates 72 h after transformation. Porphyrin producing clones were identified based on color intensity and 24 clones spanning the observed color intensity range were selected for analysis. (b) Representative HPLC-FLD chromatogram of porphyrin producing clone (D1). The observed compounds are numbered **1**, **4**, and **5**. (c) Representative mass spectra of porphyrin producing clone (D2). The panels show the dominant ions in each of the compounds **1**, **4**, and **5** as well as an authentic standard (St) of protoporphyrin IX. (d) Quantitative analysis of porphyrin production in 24 stage-one clones based on HPLC-FLD data. Each clone was analyzed in seven replicates and the graphs depict the average signal intensities with standard deviations. The upper panel summarizes the levels of **4** and **5** combined, as they could not be separated with the HPLC parameters used in this study. A general tetrapyrrole structure is embedded to illustrate that the exact structure of the compounds is unknown, but the fluorescence and mass spectrometrical data indicate a porphyrin class molecule. The lower panel depicts the protoporhyrin IX levels and the chemical structure of protoporphyrin IX (Pubchem) is embedded.

therefore best explained by conversion of porphyrin intermediates into unknown derivatives and protoporphyrin IX by endogenous *E. coli* enzymes.

Quantitative Analysis of Stage-One Clones. Based on the data from the HPLC-FLD analyses, we performed a relative quantification of the porphyrin producing capabilities of the clones in our collection (Figure 4d). Approximately half of the clones produced trace amounts of **4** and **5** with no detectable protoporphyrin IX and are thus considered the background level. The remaining clones produced 4−53 fold more **4** and **5** than the background level while the protoporphyrin IX levels were fairly constant. The three clones containing only *hemA* (B3, D1, and D4) all produced protoporphyrin IX as well as decent amounts of **4** and **5** (Figure 4d), thus supporting previous observations that the initial step of heme biosynthesis indeed is limiting in *E. coli*.[18] From a pathway engineering perspective, the clone D2 was by far the most promising as its ability to produce **4** and **5** was superior thus possibly offering a substantial gain in product formation upon re-engineering with the remaining pathway genes (Figure 4d). To verify that the porphyrin producing phenotype of clone D2 is indeed facilitated by the *pCDF-hemA,B,C,D* plasmid, the purified plasmid was retransformed into DH5α cells. Six clones were randomly chosen for analysis, as described above. All six produced both porphyrins in similar levels as the original D2 clone thus confirming that the *pCDF-hemA,B,C,D-D2* plasmid is responsible for this phenotype (data not shown).

Building on Stage-One Clones for Increased Porphyrin IX Production. In a second engineering cycle, entry fragments of *hemE* and *hemF* were amplified with matching generic oligo nucleotides (A−B, B−Ec), assembled with a pET-DUET backbone (Ec−A) and transformed into DH5α harboring *pCDF-hemA,B,C,D-D2* (Figure 3c). Again, transformants were selected based on color intensity and two superior clones c15 and c16 emerged from the subsequent HPLC-FLD screen. Both clones displayed a dramatic increase in protoporphyrin IX production accompanied by an equally dramatic decrease in the accumulation of **4** and **5**, thus further supporting that **4** and **5** are intermediates in or derivatives of intermediates in protoporphyrin IX synthesis (Figure 5a). Absolute quantification of protoporphyrin IX by LC-FTMS revealed production capabilities ranging from 10 to 90uM corresponding to 6 to 50 mg/L (Figure 5b). Thus, the best producing clones obtained (*pCDF-hemAD-D2+pET-hemEF-c15* or *C16)* are within the same range as the highest producers previously reported (84 ± 8 μM).[28] To further investigate the impact of the *pET-hemE,F* plasmids from c15 and c16 on protoporphyrin production, the plasmids were isolated,

dx.doi.org/10.1021/sb500055u | *ACS Synth. Biol.* 2015, 4, 274−282

Table 2. Shine−Dalgarno Sequences from Analyzed *pCDF-hemA−D* Clones, Their Predicted RBS Strengths, and Porphyrin Production Levels

clone ID	pCDF-hemA−D plasmids				predicted Shine−Dalgarno strength				porphyrin (sum of compds. 4 and 5)
	hemA SD	*hemB* SD	*hemC* SD	*hemD* SD	hemA SD	hemB SD	hemC SD	hemD SD	
A1	GCATTGGAG	TGTACGGAA	TCTTTGGAT	CTCCGGGAG	535	238	489	312	26 906
A2	GGTATGGAT	GGTCCGGAT	CCGGAGGAT	TCATTGGAG	228	139	8514	586	3478
A3	AATTCGGAG	ATTTCGGAA	n.d.	CAATGGGAA	1507	586	125	511	3547
A4	TGCCCGGAG	GTGAGGGAT	GCGTGGGAT	CCGCCGGAT	641	229	2065	120	35 769
A5	GACATGGAT	CGCGCGGAC	ATTTTGGAG	ACGTTGGAA	260	326	9452	51	4079
A6	AAGGGGGAT	AAGATGGAT	GTGGTGGAT	CCCTAGGAT	2526	103	1843	208	3447
B1	GTGATGGAG	GAGATGGAA	ACCGCGGAG	CGTTGGGAG	878	278	767	373	71 167
B2	TGGTCGGAC	GGACTGGAG	AGTACGGAT	TTTTTGGAC	97	373	1440	373	3873
B3	AAGAGGGAT				518				81 363
B4	ACCATGGAG	GACTTGGAT	ATGTTGGAT	CGTTAGGAG	767	85	802	1440	160 586
B5	TCCTTGGAG	GTTCGGGAT	GTTCCGGAG	TGAGTGGAA	1725	174	4600	670	104 191
B6	ATTGTGGAG	TGGCAGGAG	CGCTAGGAG	GTTGGGGAT	918	560	2705	97	90 191
C1	GCTCAGGAT	TGTTGGGAA	CTGGAGGAT	AGCCTGGAA	52	447	5429	228	4196
C2	TTCGTGGAG	CGTTGGGAG	TCCTTGGAG	CCCTAGGAG	1259	670	2160	2586	57 429
C3	CTCCAGGAT	ACGTTGGAG	TTGTAGGAT	TGTGGGGAT	139	139	2065	116	148 578
C4	GGATTGGAT	TCGGAGGAG	GTGGTGGAG	GATGTGGAT	238	1440	4642	14	4028
C5	TTTTCGGAT	TCATTGGAG	__TCAGGAG	TTGAGGGAT	260	373	6304	470	4009
C6	CGAGCGGAT	_CGCGGGAG	ATCACGGAG	TCCTTGGAT	50	918	2065	133	3191
D1	GTTTTGGAG				1051				73 002
D2	TATTGGGAG	AGACGGGAT	GTGGGGGAT	GAGTTGGAG	1772	285	13 354	560	380 934
D3	GTAGGGGAA	AGGGGGGAT	TTTTTGGAT	CCCGTGGAT	429	939	1576	139	13 363
D4	GTAGAGGAT				802				58 055
D5				TAAGTGGAC				312	7656
D6				GTCTCGGAC				68	6492

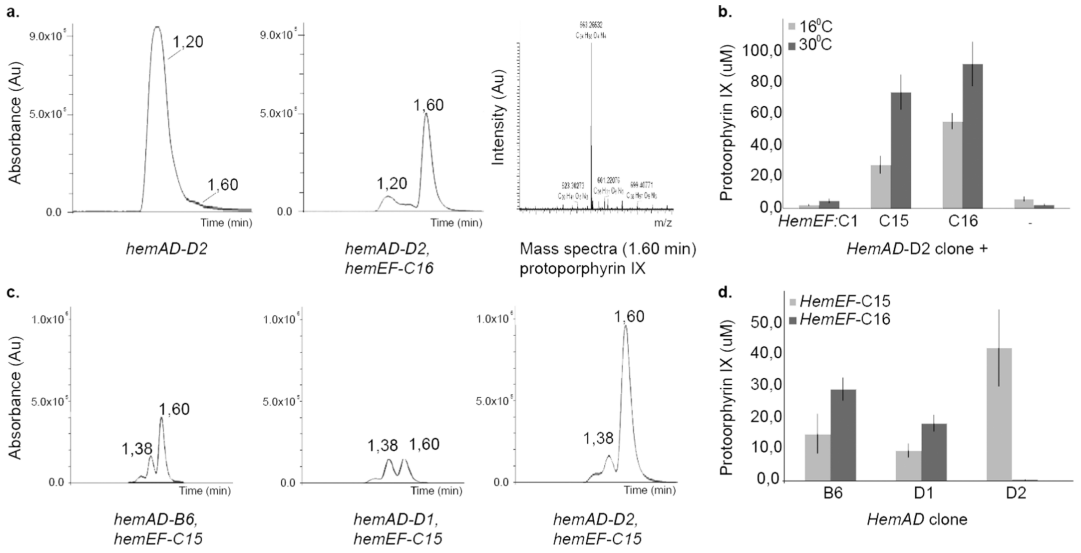

Figure 5. Protoporphyrin IX production in stage-two (*hemAD+hemEF*) clones. (a) HPLC-FLD chromatograms of the parental stage-one clone (*hemAD-D2*, left) and the highest producer among the stage-two clones, *hemAD-D2, hemEF-c16* (middle) as well as mass spectrum of protoporphyrin IX produced in *hemAD-D2, hemEF-c16* (right). (b) Absolute quantitation of protoporphyrin IX produced by stage-two clones at 16 °C (light gray) or 30 °C (dark gray). Bars display the average of three replicates and error bars denote the standard deviation. (c) HPLC-FLD chromatograms from three stage-one clones retransformed with the high producing stage-two plasmid *pET-hemEF-c15*. (d) Absolute quantification of protoporphyrin IX produced by three stage-one clones retransformed with either *pET-hemEF-c15* (light gray) or *pET-hemEF-c16* (dark gray) cultivated at 30 °C. Bars display the average of eight replicates and error bars denote the standard deviation.

purified, sequenced, and retransformed into three stage-one clones: an average producer (B6), a single gene *hemA* clone (D1), and the high producer (D2). In five of six combinations, transformation with *pET-hemE,F* resulted in increased levels of protoporphyrin IX and reduced accumulation of **4** and **5** (Figure 5c, d). Curiously, the last combination (*pCDF-hemAD-D2, pET-hemEF c16)* resulted in clones that did not produce detectable levels of porphyrins as they did in the initial screen and in the identical replicate. To investigate whether this phenotype was caused by alterations at the DNA level, *pCDF-*

dx.doi.org/10.1021/sb500055u | *ACS Synth. Biol.* 2015, 4, 274−282

hemAD-D2 and *pET-hemEF-c16* were isolated individually from a transformant randomly chosen among the analyzed colonies and compared to the corresponding plasmid stocks by restriction analysis. Indeed, the *pCDF-hemAD-D2* plasmid displayed an altered DNA fragment pattern, consistent with loss of *R. capsulatus hemA,* (Supporting Information Figure S3), and upon re-examination of the 24 stage-one clones, we found several cases of instability over the prolonged incubation time of the assay. In the light of the induced plasmid instability observed in this study, and the previously obtained titers in the same range,[28] it is tempting to speculate that production of porphyrins in the observed range is near the maximum that *E. coli* can tolerate.

Concluding Remarks. In summary, we have demonstrated that by utilizing the cloning pipeline and design principles proposed in this manuscript, state of the art levels of protoporphyrin IX production (50 mg/L) can be achieved in *E. coli* through two simple design-build-test cycles. Moreover, the design allows not only maximizing production but also generation of intermediary producers by alternative plasmid combinations. Although we do not obtain higher titers than previously shown,[28] our work demonstrates that standardized genes and proteins (all six enzymes in the pathway contained the same linker sequences) can perform as well as the native versions. In addition to the example described here, we have recently successfully added the same peptide linkers to seven genes and proteins in a heterologous pathway for high-level production of a diterpene in *E. coli* (Nielsen, MT and Nørholm, MHH, unpublished). The incorporated flexibility makes this approach useful for optimization of production scenarios as well as for rapid generation and iterative expansion of collections of clones with varying degrees of a defined phenotype. In the simplified design used for our proof of principle studies, we introduced a risk for genetic instability by repeated use of the p*trc* promoter. This may explain the observed occurrence of instable plasmids. For construction of stable production strains, this risk can trivially be eliminated by introducing different promoters as control elements in each set of generic oligo nucleotides. As the assembly of fragments relies exclusively on generic oligo nucleotides, part amplification can be heavily benchmarked and requires little *de novo* optimization. When combined with a robust cloning technique, such as the uracil excision cloning used in this manuscript (79% success rate with five DNA fragments), screening based approaches become feasible and are readily accommodated in automated processes. Taken together, the simple design process, flexibility for fine-tuning by tailoring control elements in generic primers, reliable amplification and robust cloning, combined with the potential for automation, makes the pipeline a very useful tool for combinatorial assembly and optimization of biosynthetic pathways.

METHODS

Strains and Media. *Escherichia coli* strain DH5α was used throughout this study unless otherwise stated. BL-21(DE3) was used for radioactive labeling studies. Bacteria were propagated on Luria−Bertoni (LB) agar plates supplemented with ampicillin (100 mg/mL) or spectinomycin (50 mg/mL) where required. For liquid cultures, bacteria were propagated in 2xYT for plasmid purification, PASM media for ^{35}S-methionine labeling and LB broth for porphyrin production assays.

PCR and Uracil-Excision Cloning. Uracil-excision compatible PCR products were amplified with 28 cycles in 50 μL reaction mixtures using proof-reading PfuX7 polymerase.[33] USER Fusion was performed as previously described[20] with minor modifications. Uracil-excision cloning cassettes were designed following the principles described by ref 3. Uracil-excision compatible vectors were generated by PCR as described by ref 34 using the settings outlined above or by restriction digest using the enzymes Fast Digest AjuI (Thermo Scientific) and Nb.BtsI (New England Biolabs). Purified PCR products of inserts and vector were mixed in a total insert to vector ratio of 6:1 (volume). When more than one insert is cloned simultaneously, equimolar amounts of each insert was maintained. The reaction mixture was buffered with 5× Phusion HF buffer (Invitrogen) and 1 U of USER enzyme mix (New England Biolabs) was added. The reaction mixture was incubated for 20 min at 37 °C, followed by 20 min at 25 °C before transformation of chemically competent *E. coli* cells. All primers used in this study are summarized in Supporting Information Table S1.

Plasmid Construction. Compatible plasmid backbones were generated by inserting an NheI/PacI-digested AjuI-Nb.BsrDI cassette (GBlock, IDT-DNA, Supporting Information Table S1) into multiple cloning site-1 of pET-DUET (Novagen) using the XbaI and PacI restriction sites, generating pET-MTSG-Aju,BsrDI. The start codon for translation is contained within the 5′Nb.BsrDI site, while a stop codon is positioned downstream of the 3′Nb.BsrDI site. Entry clones of *hemB, E* and *F* were PCR amplified from *E. coli* K-12 genomic DNA with primer sets designed to eliminate internal TspRI sites. *Rhodobactor capsulatus hemA* and *E. coli hemC* and *D* were synthesized as TspRI-free genes flanked by the entry sequence and activated by TspRI-digest. All entry clones were introduced into a PCR amplified pET-MTSG-AjuI-BsrDI vector by uracil excision cloning as previously described.[3] Vector backbones for multigene assemblies were based on pCDF-DUET or pET-DUET (Novagen) and PCR amplified with uracil excision compatible primers. The primers were designed to introduce a *trc*-promoter and the A-linker upstream of the first insert as well as a T7-transcriptional terminator and the Db-linker downstream of the last insert. Entry clones for multigene assemblies were amplified with generic primers introducing a SD-sequence upstream and *ptrc* downstream. USER fusion assembly of entry clone inserts and backbones was essentially performed as previously described.[20] Linkers were selected to match the backbone. All plasmids were validated by restriction digest patterns and sequencing. For a complete list of plasmids generated in this study consult Supporting Information Table S2.

^{35}S-Labeling Methionine Labeling of Proteins. Transcription and translation of porphyrin biosynthetic enzymes flanked by the 9-bp entry sequence was confirmed by using the rifampicin blocking technique and ^{35}S-labeling.[35] Precultures of BL21(DE3) cells carrying relevant plasmids were subcultured (1:50) into 1 mL chemically defined rich medium lacking methionine.[36] The cultures were grown at 37 °C to an optical density of 0.3−0.5, when expression was induced with 0.5 mM isopropyl-β-D-1-thiogalactopyranoside. After 10 min of induction, the cells were treated with 0.25 mg/mL rifampicin for 10 min, labeled with 15 μCi/mL ^{35}S-methionine for 3 min, and harvested. The pellets were solubilized in 50 μL lysis buffer (CellLytic B, 1 mM EDTA, 2 mM $MgCl_2$, lysozyme, benzonase, and complete protease inhibitor cocktail) and after 15 min

dx.doi.org/10.1021/sb500055u | *ACS Synth. Biol.* 2015, 4, 274−282

incubation one volume of SDS-PAGE sample buffer was added (125 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 100 mM DTT, bromphenol blue). Samples were applied on precast 4−20% SDS-PAGE gels (Expedeon). Upon completion of the electrophoresis, the gels were dried and the labeled bands were visualized using a Cyclone scanner (PerkinElmer).

Screening for Porphyrin Production. Initial screening for porphyrin production was conducted directly from the plate used for transformation. Plates were incubated for 16 h at 37 °C after transformation to allow colony formation and then incubated for 72 h at 30 °C. At this stage, porphyrin production could be estimated by simple visual inspection due to red pigmentation accumulating in the colonies and the media.

Assay Conditions and Extraction Procedure. For chromatographic assays, colonies were inoculated from plates into precultures and grown 16 h at 37 °C and 300 rpm shaking. Assay cultures of 500 μL fresh LB was inoculated with 10 μL preculture and grown in 2 mL microtiter plates for 5 d at 30 °C, 300 rpm or 4 h at 37 °C, 300 rpm followed by 5 d at 16 °C, 300 rpm shaking. All screening assays were made in seven replicates. Extracts for screening were processed as follows: cells were precipitated and supernatant discarded, pellet was resuspended in 200 μL H_2O and metabolites were extracted with 600 μL methanol at room temperature for 1 h. Cell debris was precipitated by centrifugation and cleared extract (600 μL) was acidified with 150 μL methanol pH 3.0, prior to analysis. For absolute quantification and compound identification, we followed a procedure described by Kwon *et al.*[28] with minor modifications. Absolute quantifications were made in 3 or 8 replicates. Cells were grown in 2 mL LB in 5 mL microtiter plates for 5 d at 30 °C, 300 rpm. Porphyrins exported to the media were precipitated with 4 mg/mL DEAE Sephadex A-25. Following precipitation, pellets were extracted twice with 2 mL acetone/acetonitrile/acetic acid (10:1:0.1) at 40 °C with sonication. Extracts were cleared by centrifugation, evaporated to dryness and resuspended in 1 mL methanol.

Analytical Method. Chromatographic separation of porhyrins was achieved using a 5 min isocratic method. Methanol adjusted to pH 3.0 with acetic acid was used as mobile phase and the column was an Ascentis Express Phenyl-Hexyl HPLC column (Supelco). Chromatographic separation was achieved on a Dionex UltiMate 3000 UHPLC. Flow 0.4 mL/min, column oven at 30 °C. Fluorescence detection was made with a FLD-3000 Fluorescence Detector (Thermo Fisher Scientific, San Jose, CA) and the data processed by the Chromeleon 7.1 Chromatography Data System (Thermo Fisher Scientific). Mass spectra were recorded by an Orbitrap Fusion Mass Spectrometer (Thermo Fisher Scientific, San Jose, CA) and data were processed with Xcalibur 3.0. Separation was achieved using the following parameters: Fluorescence detection parameters were 410 nm for excitation and 620 nm emission. Mass spectrometry was performed in positive mode using electrospray ionization and full scan mode (350−2000 m/z). The injection volume was 20 μL for HPLC-FLD analysis, 10 μL for HPLC-FT-MS.

■ ASSOCIATED CONTENT

Supporting Information

Schematic examples of different backbone preparations, Fourier transform mass spectrometry (FT-MS) analysis of porphyrins, and restriction analysis of selected stable and unstable clones. This material is available free of charge *via* the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +45 217-99184 Fax: +45-353-33300. Email: morno@biosustain.dtu.dk.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was funded by the Novo Nordisk Foundation. S.S. is the recipient of Villum Foundation's Young Investigator Programme grant VKR023128.

■ REFERENCES

(1) 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.

(2) Walhout, A. J. M., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S., and Vidal, M. (1999) GATEWAY recombinational cloning: Application to the cloning of large numbers of open reading frames or ORFeomes. *Methods Enzymol. 328*, 575−592.

(3) Nour-Eldin, H. H., Hansen, B. G., Nørholm, M. H. H., Jensen, J. K., and Halkier, B. A. (2006) Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Res. 34*, e122−e122.

(4) Engler, C., Gruetzner, R., Kandzia, R., and Marillonnet, S. (2008) Golden Gate Shuffling: A one-pot DNA Shuffling method based on type IIs restriction enzymes. *PLoS One 4*, e5553.

(5) Blanusa, M., Schenk, A., Sadeghi, H., Marienhagen, J., and Schwaneberg, U. (2010) Phosphorothioate-based ligase-independent gene cloning (PLICing): An enzyme-free and sequence-independent cloning method. *Anal. Biochem. 406*, 141−146.

(6) Nielsen, J., and Keasling, J. D. (2011) Synergies between synthetic biology and metabolic engineering. *Nat. Biotechnol. 29*, 693−695.

(7) Shetty, R., Lizarazo, M., Rettberg, R., and Knight, T. F. (2011) Chapter 13: Assembly of BioBrick standard biological parts using three antibiotic assembly. *Methods in Enzymology, Synthetic Biology Part B*, 1st ed., pp 311−326, Elsevier, Inc., San Diego, CA.

(8) Heinemann, I. U., Jahn, M., and Jahn, D. (2008) The biochemistry of heme biosynthesis. *Arch. Biochem. Biophys. 474*, 238−251.

(9) Tovmasyan, A., Sheng, H., Weitner, T., Arulpragasam, A., Lu, M., Warner, D. S., Vujaskovic, Z., Spasojevic, I., and Batinic-Haberle, I. (2013) Design, mechanism of action, bioavailability and therapeutic effects of Mn porphyrin-based redox modulators. *Med. Princ. Pract. 22*, 103−130.

(10) Neuberger, A., and Scott, J. J. (1952) Aminolevulinic acid and porphyrin biosynthesis. *Nature 172*, 1093−1094.

(11) Shemin, D., and Russell, C. S. (1953) δ-Aminolevulinic acid, its role in the biosynthesis of porphyrins and purines. *J. Am. Chem. Soc. 75*, 4873−4874.

(12) Beale, S. I., Gough, S. P., and Granick, S. (1974) Biosynthesis of δ-aminolevulinic acid from the intact carbon skeleton of glutamic acid in greening barley. *Proc. Natl. Acad. Sci. U.S.A. 72*, 2719−2723.

(13) Beale, S. I., and Castelfranco, P. A. (1972) 14C incorporation from exogenous compounds into δ-amino levulinic acid by greening cucumber cotyledons. *Biochem. Biophys. Res. Commun. 52*, 143−149.

(14) Kannangara, C. G., Gough, S. P., Oliver, R. P., and Rasmussen, S. K. (1984) Biosynthesis of δ-aminolevulinate in greening barley leaves VI. Activation of glutamate by ligation to RNA. *Carlsberg Res. Commun. 49*, 417−437.

(15) Jahn, D., Michelsen, U., and Söll, D. (1990) Two glutamyl-tRNA reductase activities in *Escherichia coli*. *J. Biol. Chem. 266*, 2542−2548.

(16) Neidle, E. L., and S, K. (1992) Expression of the *Rhodobacter sphaeroides* hemA and hemT genes, encoding two 5-aminolevulinic acid synthase isozymes. *J. Bacteriol. 175*, 2292−2303.
(17) van der Werf, M. J., and Zeicus, J. G. (1995) 5-Aminolevulinate production by *Escherichia coli* containing the *Rhodobacter sphaeroides* hemA gene. *Appl. Environ. Microbiol. 62*, 3560−3566.
(18) Kang, Z., Zhang, J., Zhou, J., Qi, Q., Du, G., and Chen, J. (2012) Recent advances in microbial production of δ-aminolevulinic acid and vitamin B12. *Biotechnol. Adv. 30*, 1533−1542.
(19) Bitinaite, J., Nichols, N. M., Bitinaite, J., and Nichols, N. M. (2009) DNA cloning and engineering by uracil excision. *Curr. Protoc. Mol. Biol.*, DOI: 10.1002/0471142727.mb0321s86.
(20) Geu-Flores, F., Nour-Eldin, H. H., Nielsen, M. T., and Halkier, B. A. (2006) USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res. 35*, e55.
(21) Tillett, D., and Neilan, B. A. (1999) Enzyme-free cloning: a rapid method to clone PCR products independent of vector restriction enzyme sites. *Nucleic Acids Res. 27*, e26.
(22) Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2010) A Modular Cloning System for Standardized Assembly of Multigene Constructs. *PLoS ONE 6*, e16765.
(23) Engler, C., Kandzia, R., and Marillonnet, S. (2008) A One Pot, One Step, Precision Cloning Method with High Throughput Capability. *PLoS ONE 3*, e3647.
(24) Aslanidis, C., and de Jong, P. J. (1990) Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res. 18*, 6069−6074.
(25) de Kok, S., Stanton, L. H., Slaby, T., Durot, M., Holmes, V. F., Patel, K. G., Platt, D., Shapland, E. B., Serber, Z., Dean, J., Newman, J. D., and Chandran, S. S. (2014) Rapid and reliable DNA Assembly via ligase cycling reaction. *ACS Synth. Biol. 3*, 97−106.
(26) Anderson, J. C., Dueber, J. E., Leguia, M., Wu, G. C., Goler, J. A., Arkin, A. P., and Keasling, J. D. (2009) BglBricks: A flexible standard for biological part assembly. *J. Biol. Eng. 4*, 1−1.
(27) Nevin, D. E., and Pratt, J. M. (1990) A coupled *in vitro* transcription−translation system for the exclusive synthesis of polypeptides from the T7-promoter. *FEBS Lett. 291*, 259−263.
(28) Kwon, S. J. S., de Boer, A. L. A., and Schmidt-Dannert, C. C., 4. (2003) High-level production of porphyrins in metabolically engineered *Escherichia coli*: Systematic extension of a pathway assembled from overexpressed genes involved in heme biosynthesis. *Appl. Environ. Microbiol. 69*, 4875−4883.
(29) Zelcbuch, L., Antonovsky, N., Bar-Even, A., Levin-Karp, A., Barenholz, U., Dayagi, M., Liebermeister, W., Flamholz, A., Noor, E., Amram, S., Brandis, A., Bareia, T., Yofe, I., Jubran, H., and Milo, R. (2012) Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Res. 41*, e98−e98.
(30) Salis, H. M., Mirsky, E. A., and Voigt, C. A. (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol. 27*, 946−U112.
(31) Salis, H. M. (2011) The ribosome binding site calculator. *Meth. Enzymol. 498*, 19−42.
(32) Lee, M. J. (2012) Porphyrin derivatives from a recombinant *Escherichia coli* grown on chemically defined medium. *J. Microbiol. Biotechnol. 22*, 1653−1658.
(33) Nørholm, M. H. H. (2010) A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. *BMC Biotechnol. 10*, 21.
(34) Hansen, B. G., Mnich, E., Nielsen, K. F., Nielsen, J. B., Nielsen, M. T., Mortensen, U. H., Larsen, T. O., and Patil, K. R. (2012) Involvement of a natural fusion of a cytochrome P450 and a hydrolase in mycophenolic acid biosynthesis. *Appl. Environ. Microbiol. 78*, 4908−4913.
(35) Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1989) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol. 185*, 60−89.
(36) Ukkonen, K., Vasala, A., Ojamo, H., and Neubauer, P. (2011) High-yield production of biologically active recombinant protein in shake flask culture by combination of enzyme-based glucose delivery and increased oxygen transfer. *Microb. Cell Fact. 10*, 107.