The Logic, Experimental Steps, and Potential of Heterologous Natural Product Biosynthesis Featuring the Complex Antibiotic Erythromycin A Produced Through E. coli.

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
The Logic, Experimental Steps, and Potential of Heterologous Natural Product Biosynthesis Featuring the Complex Antibiotic Erythromycin A Produced Through E. coli

Ming Jiang1, Haoran Zhang2, Blaine A. Pfeifer1

1Chemical and Biological Engineering Department, State University of New York at Buffalo
2Chemical Engineering Department, Massachusetts Institute of Technology

Correspondence to: Blaine A. Pfeifer at blainepf@buffalo.edu

URL: http://www.jove.com/video/4346
DOI: doi:10.3791/4346

Keywords: Biomedical Engineering, Issue 71, Chemical Engineering, Bioengineering, Molecular Biology, Cellular Biology, Microbiology, Basic Protocols, Biochemistry, Biotechnology, Heterologous biosynthesis, natural products, antibiotics, erythromycin A, metabolic engineering, E. coli

Date Published: 1/13/2013


Abstract

The heterologous production of complex natural products is an approach designed to address current limitations and future possibilities. It is particularly useful for those compounds which possess therapeutic value but cannot be sufficiently produced or would benefit from an improved form of production. The experimental procedures involved can be subdivided into three components: 1) genetic transfer; 2) heterologous reconstitution; and 3) product analysis. Each experimental component is under continual optimization to meet the challenges and anticipate the opportunities associated with this emerging approach.

Heterologous biosynthesis begins with the identification of a genetic sequence responsible for a valuable natural product. Transferring this sequence to a heterologous host is complicated by the biosynthetic pathway complexity responsible for product formation. The antibiotic erythromycin A is a good example. Twenty genes (totaling >50 kb) are required for eventual biosynthesis. In addition, three of these genes encode megasynthases, multi-domain enzymes each ~300 kDa in size. This genetic material must be designed and transferred to E. coli for reconstituted biosynthesis. The use of PCR isolation, operon construction, multi-cystronic plasmids, and electro-transformation will be described in transferring the erythromycin A genetic cluster to E. coli.

Once transferred, the E. coli cell must support eventual biosynthesis. This process is also challenging given the substantial differences between E. coli and most original hosts responsible for complex natural product formation. The cell must provide necessary substrates to support biosynthesis and coordinately express the transferred genetic cluster to produce active enzymes. In the case of erythromycin A, the E. coli cell had to be engineered to provide the two precursors (propionyl-CoA and (2S)-methylmalonyl-CoA) required for biosynthesis. In addition, gene sequence modifications, plasmid copy number, chaperonin co-expression, post-translational enzymatic modification, and process temperature were also required to allow final erythromycin A formation.

Finally, successful production must be assessed. For the erythromycin A case, we will present two methods. The first is liquid chromatography-mass spectrometry (LC-MS) to confirm and quantify production. The bioactivity of erythromycin A will also be confirmed through use of a bioassay in which the antibiotic activity is tested against Bacillus subtilis. The assessment assays establish erythromycin A biosynthesis from E. coli and set the stage for future engineering efforts to improve or diversify production and for the production of new complex natural compounds using this approach.

Introduction

Erythromycin A is a polyketide antibiotic produced by the Gram-positive soil bacterium Saccharopolyspora erythraea, and current production has been incrementally improved to ~10 g/L through decades of traditional mutagenesis and screening protocols and more recently through process optimization schemes. Mutagenesis and screening strategies are common in antibiotic natural product development as a result of difficulties in cultivating and/or genetically manipulating native production hosts and because of the readily available antibiotic activity or improved growth phenotypes to aid selection. In the case of erythromycin A, S. erythraea is limited by a slow growth profile and the lack of more direct genetic manipulation techniques (relative to organisms like E. coli), thus, hampering rapid improvements in production and the biosynthesis of new derivatives. Having recognized the production issues and unlocked diversification possibilities with compounds like erythromycin A, the research community began to pursue the idea of heterologous biosynthesis. These efforts coincided with available sequence information for the erythromycin A gene cluster. It should be emphasized that the number of sequenced complex natural product gene clusters has greatly expanded, providing the impetus for continued efforts in heterologous biosynthesis to access encoded medicinal potential. To
do so, heterologous reconstitution requires that the new host meet the needs of the specific biosynthetic pathway. *E. coli* provides technical convenience, a wide-spanning set of molecular biology techniques, and metabolic and process engineering strategies for product development. Yet, when compared to native production hosts, *E. coli* does not exhibit the same level of complex natural product production. It was therefore unknown whether *E. coli* could serve as a viable heterologous option for complex natural product biosynthesis. However, it was assumed that *E. coli* would be an ideal host organism if heterologous biosynthesis could be accomplished.

With this goal in mind, initial efforts began to produce the polyketide aglycone 6-deoxyerythronolide B (6dEB) through *E. coli*. However, native *E. coli* metabolism could not provide appreciable levels of the propionyl-CoA and (2S)-methylmalonyl-CoA precursors needed to support 6dEB biosynthesis nor could the new host post-translationally modify the deoxyerythronolide B synthase (DEBS) enzymes. To remedy these issues, a metabolic pathway composed of native and heterologous enzymes was built into *E. coli* such that exogenously fed propionate was converted intracellularly to propionyl-CoA and then (2S)-methylmalonyl-CoA; during the engineering to complete this pathway, an *sfp* gene was placed into the chromosome of *E. coli* BL21(DE3) to produce a new strain termed BAP1. The *Sfp* enzyme is a phosphopantetheinytransferase capable of attaching the 4'-phosphopantetheinyl co-factor to the DEBS enzymes. The three DEBS genes (each ~10 kb) were then placed on two separately selectable expression vectors containing inducible T7 promoters. After a key adjustment of post-induction temperature (to 22 °C), the DEBS genes were coordinately expressed within BAP1 in an active state capable of generating 6dEB.

The pursuit of full erythromycin A biosynthesis then began using an analogous gene cluster from *Micromonospora megalomicea* or a hybrid pathway composed of genes from *S. erythraea*, *S. fradiae*, and *S. venezuelae* which produced the intermediates erythromycin C and 6-deoxyerythromycin D, respectively. Recently, our group has extended these efforts by producing erythromycin A (the most clinically-relevant form of erythromycin) through *E. coli*. In contrast to previous work, our strategy coordinately expressed the 20 original *S. erythraea* genes needed for polyketide biosynthesis, deoxysugar biosynthesis and attachment, additional tailoring, and self-resistance (Figure 2). In total, 26 (native and heterologous) genes were engineered to allow *E. coli* to produce erythromycin A at 4 mg/L. This result established complete production of a complex polyketide natural product using *E. coli* and serves as a basis to leverage this new production option or pursue new ones.

**Protocol**

The text below is specific for the erythromycin A antibiotic, but the steps are designed to be generally applicable to other natural products as candidates for heterologous biosynthesis.

1. **Erythromycin A Genetic Cluster Transfer**

1. Design PCR primers to amplify all of the genes associated with the erythromycin A cluster within the *S. erythraea* chromosome. This step is specific to those natural product candidates whose genetic sequences have been determined. Alternatively, the required genes may be synthesized (through several commercial vendor options) with the advantage that the newly synthesized genes would be designed for optimal codon usage within the new *E. coli* host. If DNA synthesis is pursued, a current challenge is generating megasynthesis genes which may exceed 10 kb in length. Technically, synthesis can be accomplished, but it is challenging and costly. It is anticipated that continual improvements in gene synthesis technology will address current drawbacks and further validate this approach for providing heterologous genetic material.

2. Perform PCR using freshly prepared *S. erythraea* genomic DNA as a template. The chromosomal DNA may be prepared from cultures of *S. erythraea* or from frozen stocks of the bacterium (Figure 3). PCR reactions may benefit from the addition of DMSO (4 μl in 50 μl reactions) to account for high G+C content in the target *S. erythraea* genes. Each amplified gene should be sequenced to confirm that the full gene has been retrieved and that no mutations have been introduced during the PCR process.

3. Through restriction digest and ligation, insert each of the isolated genes into expression vectors designed for *E. coli* (Figure 3). This is initiated by designing restriction sites within the PCR primers to allow digestion and ligation into suitable expression vectors (an example has been included in Figure 3). We have used both pET21 and pET28 vectors. The pET28 vector allows the inclusion of a 5′ leader sequence which aided gene expression in the erythromycin A case. Confirm individual gene expression through RT-PCR, SDS-PAGE analysis (using soluble protein fractions), or convenient phenotypic assays (Figure 4).

4. Construct operons based upon genes successfully expressed from individual expression plasmids (Figure 3). We have traditionally used combinations of compatible-cosehesive restriction enzymes (such as XbaI and SpeI) to sequentially convert genes to operons; however, the restriction sites chosen for operon construction must not reside within (or must be removed from) the isolated gene sequences (an issue that can be avoided if gene synthesis is used). This step consolidates the 20 erythromycin A genes prior to *E. coli* transformation. Currently, we have used this method to include up to 20 kb of foreign DNA into a standard pET vector (two genes of ~10 kb each). Furthermore, we have used this approach to generate operons containing nine genes. It is unknown exactly how much DNA or how many genes may be integrated into pET-type plasmids; however, the metrics cited above provide a reference for complicated systems like the erythromycin biosynthetic pathway. Finally, standard *in vitro* ligation steps and subsequent successful transformation steps become increasingly difficult as operon and plasmid size increases. Ligation temperature (12-22 °C) and time (3-24 hr) are varied and electro-transformation is used to aid the process of producing final plasmid constructs.

5. Transform the newly designed biosynthetic plasmids into *E. coli* strain BAP1 using electro-transformation (Figure 5). The procedure may be attempted with plasmids introduced either sequentially or in combination. When transforming size-able or numerous plasmids, we have traditionally used electro-transformation (as opposed to chemical transformation). Once transformed, plate the cells on solid LB medium containing tetracycline (5 mg/l), carbenicillin (100 mg/l), kanamycin (50 mg/l), streptomycin (50 mg/l), and apramycin (50 mg/l) to select for plasmid maintenance. It is highly unusual to use six different selection antibiotics post-transformation of an *E. coli* cell, but this simply reflects the total number of plasmids needed to allow final erythromycin A biosynthesis. While possible, the approach will limit the production capabilities of the compound once culturing has begun (please see below).
2. **E. coli** Biosynthetic Reconstitution

1. Strain BAP1 has been designed to provide the required substrates for biosynthesis and to post-translationally modify the deoxyerythronolide B megasynthase. Individual genes of the erythromycin A pathway have been tested for gene expression and consolidated as operons into expression plasmids. This section is dedicated to culture conditions to test for concerted activity of the full erythromycin A pathway.

2. Take 3 separate colonies of the freshly transformed BAP1 and inoculate 1.5 ml LB cultures containing the required plasmid selection antibiotics (at the same concentrations indicated above to select for transformants on solid medium), Isopropyl β-D-1-thiogalactopyranoside (IPTG; 100 μM) and arabinose (2 mg/ml) to induce gene expression, and 20 mM sodium propionate to support intracellular biosynthetic precursor formation. Once production has been confirmed, larger cultures may be conducted in an attempt to scale and maximize product titers.

3. Culture the cells at 22 °C and 250 rpm for 24-48 hr. In this case, antibiotic selection ensures that the required plasmids are maintained in the early stages of cellular culture. However, plasmids may be easily lost as selection antibiotic levels change over the course of the culturing period. This may lead to plasmid instability (i.e., the loss of one or more plasmids during culture), which would then limit production capability. The problem is exacerbated when incompatible plasmids are used to introduce a complete heterologous gene cluster. Such is the case for the erythromycin A compound (Figure 5) and would have to be remedied to truly overproduce the compound from **E. coli**. However, for the purpose of demonstrating initial production, such design flaws are tolerated for the sake of expeditiously establishing the feasibility of the heterologous approach.

4. Clarify the cultures by centrifugation at 10,000 rpm in an Eppendorf microfuge. Remove the supernatant and store at 4 °C for analysis.

5. In the case of erythromycin A, lack of compound production was addressed by the inclusion of the GroES/EL chaperonin (using pGro7 specifically to address lack of activity of the enzymes post-expression) and the inclusion of additional gene copies (for those genes suspected of weak expression/activity). Similar strategies may be needed with other heterologous natural product production efforts.

3. **Product Analysis**

1. To confirm and quantify erythromycin A formation, inject the culture supernatant onto an LC-MS system (Figure 6). Commercially available erythromycin A was used to confirm production and as a standard during quantification. For cases without commercially available standards, LC-MS analysis would require additional chemical characterization by NMR and high resolution mass spectrometry.

2. Compare production between the 3 colonies/cultures from the BAP1 transformation. Choose the highest producer and prepare a glycerol stock from the colony source. This step can be done in parallel with the production cultures of Protocol Section 2 by using a small portion of the cultures to prepare glycerol stocks. Store the glycerol stock in a -80 °C freezer.

3. Using the confirmed, highest producing cell stock, begin a separate production culture and extract the final supernatant with 2x volume ethyl acetate. Use a vacuum evaporative system to dry the extract and resuspend in 100 μl methanol. Transfer the extract to a sterile filter disk (~1/4 inch diameter) and allow the disk to dry. Separately, culture **B. subtilis** in LB liquid overnight. Add 20 μl of the **B. subtilis** culture to 20 ml of liquid LB agar at 45 °C. Plate the agar and allow to solidify. Place the filter disk on the plate and incubate at 37 °C (Figure 7). Alternatively, the extract can be added to liquid cultures within a 96-well plate to quantify the activity using a plate reader (Figure 7).

### Representative Results

The desired outcome of this approach is the production of a fully bioactive natural product from the **E. coli** heterologous host. This is best represented by the LC-MS results used to confirm and quantify production (Figure 6) and the antibacterial bioassay used to confirm final activity (Figure 7). In the overall scheme of heterologous biosynthesis, this result defines success. Once accomplished, research efforts then turn to optimization (both at the cellular and process scales) and molecular engineering. The final objectives include economical production processes for the original compound and control over the heterologous system such that rational modifications to the process can be made to produce variants of the original compound with the potential for a heightened or broadened activity spectrum. Failure to produce the desired compound then triggers numerous contingency plans to assess which portion of the approach is prohibiting successful biosynthesis. Many of these trouble-shooting mechanisms are in-built within the procedure outlined above. In our experience, it is important, at a minimum, to confirm successful gene expression (SDS-PAGE analysis of soluble biosynthetic enzymes) of the newly introduced gene cluster prior to attempting to assess biosynthesis of the formed natural product. Post expression, as indicated above, the use of protein folding chaperones, enhanced gene copy number, and process temperature have been effective means of allowing eventual biosynthesis.

![Native Producer](Saccharopolyspora erythraea) → **Heterologous Host** → **Natural Product**

*Figure 1.* An overall depiction of the heterologous natural product biosynthetic scheme featuring the erythromycin A compound. The experimental steps required include 1) design and transfer of the erythromycin gene cluster from **S. erythraea** to **E. coli**; 2) reconstitution of biosynthesis within **E. coli**; and 3) product analysis of erythromycin A.
Figure 2. **A.** The erythromycin gene cluster as organized in the native *S. erythraea* chromosome. **B.** Erythromycin A biosynthesis. Within *E. coli*, the Sfp enzyme is required for posttranslational modification of the DEBS1, 2, and 3 polyketide synthase enzymes (each >300 kDa); this is denoted by the connecting arm associated with each ACP domain. The DEBS enzymes form an $\alpha_2\beta_2\gamma_2$ complex subdivided into module units that catalyze the consecutive condensations of one starter propionyl-CoA unit and six (2S)-methylmalonyl-CoA extender units to form the polyketide 6-deoxyerythronolide B (6dEB). KS-ketosynthase; AT-acyl transferase; ACP-acyl carrier protein; KR-ketoreductase (with the inactive KR noted in Module 3); DH-dehydratase; ER-enoyl reductase; TE-thioesterase. A PrpE (propionyl-CoA synthetase) and PCC (propionyl-CoA carboxylase, composed of two protein subunits) are required to convert exogenous propionate to the starting substrates propionyl-CoA and (2S)-methylmalonyl-CoA. The genes for Sfp and PrpE were previously engineered within the BAP1 *E. coli* chromosome \(^\text{19}\). Conversion of the 6dEB molecule to erythromycin is accomplished by the deoxysugar biosynthesis and attachment, additional tailoring, and self-resistance enzymes. **Click here to view larger figure.**
Figure 3. A. Isolation of individual genes required for erythromycin A biosynthesis (excluding the three DEBS genes). Individual plasmid design including the PCR oligonucleotide primers used for isolating \textit{eryCI}; primer restriction sites have been indicated in bold and complementary sequence has been underlined. Also included is the restriction analysis of the resulting pET21 and pET28 vectors containing individual genes and operons. B. i. Consolidation of the biosynthetic genes needed for polyketide tailoring as operons for introduction to \textit{E. coli}. ii. Restriction site and other expression elements involved with operon construction and design; \textit{X/S} represents the compatible-cohesive ligation of \textit{XbaI} and \textit{SpeI} sites which results in a new sequence unrecognized by either restriction enzyme. \textit{Click here to view figure 3.}
Figure 4. A. SDS-PAGE analysis of the individual genes presented in Figure 3. The asterisks denote those genes expressed with 5' leader sequences resulting from the pET28 expression vector. B. Phenotypic assessment of ErmE (erythromycin A resistance). Erythromycin A within the solid medium used to test *E. coli* strains harboring plasmids with or without the *ermE* erythromycin A resistance gene. Cell growth is rescued with *ermE* expression.
Figure 5. A. Schematic of the full erythromycin A pathway introduced to E. coli, including a plasmid for the GroEL/ES chaperonin (pGro7) and a plasmid containing an extra copy of eryK. B. Transformation comparison between chemical and electro methods using both combination (i) and sequential (ii) introduction of plasmids. C. Plasmid stability of transformed strain. Click here to view larger figure.
Critical steps in heterologous biosynthesis are encountered at each of the three procedural points in the process: 1) genetic transfer; 2) biosynthetic reconstitution; and 3) product analysis. A problem at any stage will derail the ultimate objective of establishing heterologous biosynthesis. Perhaps the most challenging aspect of the process is establishing reconstituted biosynthesis, since this is absolutely required to allow successful analysis. However, reconstitution is dependent upon careful design and transfer of the genetic material responsible for eventual biosynthesis. Likewise, established, sensitive analytical methods (like LC-MS) can detect small titer levels that otherwise may be interpreted as lack of production. It is the culmination of these situations that makes heterologous biosynthesis challenging.

With regards to the erythromycin A example, confirmed gene sequences isolated by PCR were individually tested for expression prior to consolidating the genes for coordinated expression and natural product formation. During initial attempts at product analysis, intermediate compounds were first identified only when a chaperonin (GroEL/ES) was co-expressed with the required gene cluster. The final product was only produced when an additional copy of the gene responsible for the last step in biosynthesis (eryK) was included. Both of these measures benefitted from the sensitive LC-MS analytical method used to initially assess production.

Now that production has been confirmed, the E. coli cellular background offers numerous engineering opportunities. The first of which is to optimize production of the native erythromycin A compound. This would offer an alternative production route with potential savings in process...
time and cost. Furthermore, the erythromycin A pathway may be manipulated for the purpose of diversifying product formation. New analogs offer the potential of new antibiotic (or other therapeutic) activity. These same opportunities exist for general applications of heterologous biosynthesis and provide the primary motivation for the approach, further supported by the challenges encountered with many original producing organisms. Continual applications and improvements of the approach described in this article will increase the frequency of future heterologous biosynthetic success.

Disclosures

No conflicts of interest declared.

Acknowledgements

The authors thank the NIH (AI074224 and GM085323) and NSF (0712019 and 0924699) for funding to support projects dedicated to heterologous biosynthesis.

References