Published in final edited form as:

Metab Eng. 2008 ; 10(3-4): 131–140. doi:10.1016/j.ymben.2008.02.001.

Transfer of the high GC cyclohexane carboxylate degradation pathway from *Rhodopseudomonas palustris* to *Escherichia coli* for production of biotin

Jeffrey R. Bernstein^{1,2}, Thomas Bulter¹, and James C. Liao^{1,2}

¹Department of Chemical and Biomolecular Engineering, University of California, Los Angeles

²Biomedical Engineering Interdepartmental Program, University of California, Los Angeles

Abstract

This work demonstrates the transfer of the five gene cyclohexane carboxylate (CHC) degradation pathway from the high GC alphaproteobacterium *Rhodopseudomonas palustris* to *Escherichia coli*, a gammaproteobacterium. The degradation product of this pathway is pimeloyl-CoA, a key metabolite in *E. coli*'s biotin biosynthetic pathway. This pathway is useful for biotin overproduction in *E. coli*, however, the expression of GC-rich genes is troublesome in this host. When the native *R. palustris* CHC degradation pathway is transferred to a $\Delta bioH$ pimeloyl-CoA auxotroph of *E. coli*, it is unable to complement growth in the presence of CHC. To overcome this expression problem we redesigned the operon with decreased GC content and removed stretches of high GC intergenic DNA which comprise the 5' untranslated region of each gene, replacing these features with shorter low GC sequences. We show this synthetic construct enables growth of the $\Delta bioH$ strain in the presence of CHC. When the synthetic degradation pathway is overexpressed in conjunction with the downstream genes for biotin biosynthesis, we measured significant accumulation of biotin in the growth medium, showing that the pathway transfer is successfully integrated with the host metabolism.

Introduction

The vitamin biotin, a cofactor used in various carboxylation reactions, is an industrially important compound, added to feed stocks and cosmetics, comprising a world market of 10–30 tons per year (Shaw et al., 1999; Streit and Entcheva, 2003). However, the majority of biotin is produced via a chemical process, which has a high environmental burden. The vitamin is synthesized by most bacteria, plants, and fungi, leading to considerable efforts towards discovering a cost-effective bioprocess. In bacteria, most of the genetic analysis of biotin biosynthesis has been performed in *Escherichia coli* and *Bacillus* species. Pimeloyl-CoA is a key metabolite common to all known microbes for the synthesis of biotin in four enzymatic steps (Fig. 1)(Streit and Entcheva, 2003). In gram positive bacteria such as *Bacillus*, the enzymes BioW and BioI are responsible for the conversion of pimelic acid to pimeloyl-CoA. In gram negative bacteria such as *E. coli*, the biochemical route to pimeloyl-CoA is not properly understood although *bioC* and *bioH* are known to play a role in the formation of this compound (Streit and Entcheva, 2003; Tomczyk et al., 2002), which likely occurs through the

Address correspondence to: James C. Liao, 5531 Boelter Hall, 420 Westwood Plaza, Los Angeles, CA 90095. Tel.: 310-825-1656; Fax: 310-206-4107; E-mail: liaoj@ucla.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

condensation of malonyl-CoA and acetate (Sanyal et al., 1994). The *E. coli* $\Delta bioH$ mutant is auxotrophic for pimeloyl-CoA, and has been shown to be complemented by the *bioZ* gene from *Mesorhizobium* (Sullivan et al., 2001). *E. coli* represents an attractive host for the engineering of biotin biosynthesis as there are many available genetic tools which allow for the manipulation of its pathways. Its genes for biotin production, with the exception of *bioH*, are clustered as the *bioBFCD* operon, divergently transcribed from *bioA*. The expression of these genes is tightly repressed in the presence of excess biotin (Weaver et al., 2001). Efforts to design an overproducing strain of this organism have involved de-repressing or overexpressing these genes (Shaw et al., 1999). Here we focus on the integration of the cyclohexane carboxylate (CHC) degradation pathway from *Rhodopseudomonas palustris* (Egland et al., 1997) with *E. coli* metabolism, to provide for a new biochemical route to the pimeloyl-CoA pool incorporated into the vitamin biotin.

CHC is an alicyclic compound produced as an intermediate in the anaerobic degradation of aromatic rings, which are the monomeric constituents of plant lignins, by phototrophic and denitrifying bacteria (Dutton and Evans, 1969; Egland et al., 1997; Kuver et al., 1995). In *R. palustris*, CHC supports relatively rapid growth under both aerobic and anaerobic conditions. The degradation of this compound proceeds through five enzymatic steps, first creating a coenzyme A thioester, which is then metabolized via a series of β -oxidation-like reactions before a ring opening step to yield the metabolite pimeloyl-CoA (Fig. 1) (Samanta and Harwood, 2005) (Pelletier and Harwood, 2000) (Kuver et al., 1995) (Egland et al., 1997). Pimeloyl-CoA is then further degraded enzymatically to acetyl-CoA and CO₂ (Harrison and Harwood, 2005). This pathway offers an attractive route to the pimeloyl-CoA pool destined for biotin synthesis in *E. coli* as it provides a well-characterized alternative to the poorly defined native pathway.

Traditionally, heterologous gene expression is accomplished using inducible promoters and optimal ribosome-binding sites (RBS), with single genes added in series to achieve the desired pathway extension and metabolite production (Bulter et al., 2003). However, recent examples have described the functional transfer of entire pathways from one bacterium to another. The phenylacetic acid catabolic pathway from *Pseudomonas putida* U was transferred to an *E. coli* deletion strain whose native pathway for assimilation of this molecule had been removed (Garcia et al., 2004) (Ferrandez et al., 1998). Moreover, a group successfully produced the polyketides epothilone C and D in *E. coli* from a *Sorangium cellulosum* pathway using a combination of DNA synthesis, inducible promoters, chaperone co-expression, and the separation of a large protein into two polypeptides (Mutka et al., 2006). Previously, we assessed the ability of *Escherichia coli* to recognize unaltered foreign DNA from the high GC photosynthetic α -bacterium *R. palustris* (Bernstein et al., 2007). Multiple transcription signals were recognized by the host, but translation of the donor DNA was found to be less efficient which was partially attributed to the high GC content of the 5'-untranslated region (UTR) of transcripts, even in the presence of a strong RBS. Here we show that the introduction of the native pathway under control of either the native or inducible P_{lacO-1} promoter is unable to functionally express in *E. coli*. Through a strategy of lowering the GC content of the *R. palustris* open reading frames, and replacing high GC 5'-UTRs with shorter low GC stretches containing a strong RBS, we integrate this pathway with the *E. coli* biotin biosynthetic genes and measure significant accumulation of this vitamin.

Materials and Methods

Plasmid constructs

pJRB502 is a high copy plasmid derived from pJF118EH (Furste et al., 1986), with the *badH**laliB**abadK* operon under the control of its native promoter as amplified from the *R. palustris* chromosome (Larimer et al., 2004). pJRB510 is a high copy plasmid derived from

pJRB202 (Bernstein et al., 2007). It contains the *badH**aliB**abadK* operon from amplified from the *R. palustris* chromosome, expressed from the P_{lacO-1} promoter.

The protein fusions of genes from the *badH**aliB**abadK* operon were created as described using the high copy pTB120 backbone (Bernstein et al., 2007). Protein fusions to LacZ were made to the ~20th amino acid of each gene in the pathway (Simons et al., 1987), such that the fusion to the final gene in the operon, *badK*, includes the promoter and each of the 4 genes upstream of it.

To generate pJRB620, the individual genes from the *badH**aliB**abadK* operon of *R. palustris* were optimized for *E. coli* codon usage (Grote et al., 2005), which lowered the GC content by 12%. The high GC intergenic DNA comprising the 5'-UTR of each gene was removed and replaced with a shorter low GC sequence containing a strong RBS, and the P_{LtetO-1} promoter was placed in front of the operon. The RBS/UTR for *badH* is derived from pZA31-Luc (Lutz and Bujard, 1997), while the RBS/UTRs for the remaining genes are variants of this sequence, maintaining a minimum of 5 out of 6 bp of Shine Dalgarno consensus sequence 8 bp upstream of the start codon and possessing other AT-rich nucleotides and restriction sites (Supplementary Table 1). A transcription terminator followed by a second P_{LtetO-1} promoter was placed upstream of *aliB*, which is the first gene poorly expressed by *E. coli* in the native, unaltered CHC degradation operon, as measured by LacZ protein fusions. The DNA construct was synthetically produced (Celtek Genes), and was subcloned into the medium copy plasmid pZA31-Luc (Lutz and Bujard, 1997). The second P_{LtetO-1} promoter and T₀ terminator upstream of *aliB* were removed using a duplicated restriction site flanking this region, to generate pJRB621, allowing the five gene operon to be expressed from the single P_{LtetO-1} promoter upstream of *badH*.

pJRB554 is a high copy plasmid derived from pJF118EH (Furste et al., 1986), containing the lacI_q repressor and expressing the biotin biosynthetic genes *bioBAFD* under the control of the P_{lacO-1} promoter, followed by an *rrnB* terminator. The *bioBAFD* genes from the *E. coli* biotin biosynthetic cluster were individually amplified by PCR from the genome using KOD Hot Start Polymerase, and cloned sequentially using restriction enzymes (New England Biolabs) into a multiple cloning site replacing the native RBS with ones closely resembling a consensus Shine-Dalgarno sequence.

Construction of the Δ *bioH* mutant

The *bioH* gene was removed from the chromosome of BW25113 (Lessard et al., 1998) using the λ Red recombination method (Datsenko and Wanner, 2000). The gene deletion was confirmed by PCR and the kanamycin resistance cassette was removed from the chromosome. Finally, the LacR, TetR, Sp^r cassette (Lutz and Bujard, 1997) was moved onto the chromosome from DH5 α Z1 by phage *P1* transduction.

Complementation study

The pimeloyl-CoA auxotrophy of Δ *bioH*-Z1 cells is observed during growth on minimal medium, once cells are starved of biotin. Δ *bioH*-Z1 cells were transformed with either pJRB502 or pJRB510, containing either the *badH**aliB**abadK* operon expressed from the native *R. palustris* or inducible P_{lacO-1} promoter, respectively. Cells were grown for 7 hours in LB plus appropriate antibiotics, which was used for a 0.25% (v/v) inoculation of 2mL of M9 minimal medium plus glucose (0.5%), 0.1 mM isopropyl- β -D-thiogalactopyranoside (IPTG) and appropriate antibiotics. After 20 hours of growth in minimal medium, cells were inoculated (0.25%) into fresh M9 glucose (0.5%) minimal medium with 1mM IPTG and either: a) no CHC; b) 1mM CHC; or c) 1mM CHC plus 30 ng biotin. Cell growth was monitored over 72 hours. The same assay was also performed using pJRB510 in conjunction with pRARE2 (EMD

Biosciences), a plasmid expressing tRNAs for 7 rare codons. Additionally, $\Delta bioH$ -Z1 cells containing either pJRB620 or pJRB621, were cultured using the same assay.

β -Galactosidase measurements

β -Galactosidase assays for the protein fusions were run in triplicate for two individual colonies of each construct and were performed as described previously (Arvidson et al., 1991; Bernstein et al., 2007).

Time course of the synthesized CHC degradation operon

$\Delta bioH$ -Z1 cells were transformed with pJRB621 and grown for 7 hours in LB medium supplemented with appropriate antibiotics, followed by a 0.25% inoculation of M9 fructose (0.5%) minimal medium, including 100 nM anhydrotetracycline (aTc). After 15 hours of growth, cells were inoculated (0.5%) to 50 mL of fresh M9 fructose medium contained in 250 mL baffled Erlenmeyer flasks supplemented with 100 nM aTc. One culture received 1.5 mM CHC, another received 1.5 mM CHC plus 500 ng D-biotin, another culture did not receive any CHC, and finally one culture was supplemented with only 1 nM aTc and 1.5 mM CHC. Cells were grown at 37 °C and 250 rpm and absorbance was monitored at 600 nm for 30 hours.

Real-time PCR

Duplicate colonies of BW-Z1 cells harboring either pJRB510 or pJRB621 plasmids were grown 16 hours in 3 mL LB medium containing appropriate antibiotics. New LB cultures containing antibiotics and either 1 mM IPTG (for pJRB510) or 100 nM aTc (for pJRB621) were inoculated with 0.5% (v/v) of the overnight culture and grown at 37 °C until an OD₆₀₀ of 0.4 – 0.5. Cells were briefly chilled on ice and a 0.5 mL aliquot was spun down 2 minutes at 14,000 rpm. The supernatant was removed and pellets were frozen at –80 °C until further processing. RNA was isolated and first strand cDNA synthesis was performed as described previously (Bernstein et al., 2007). Gene specific primers for both the native *R. palustris* DNA sequence and also for the *E. coli* codon-optimized sequence were designed using VectorNTI software (Invitrogen). Amplification products ranged between 100–130 bp and were checked by gel electrophoresis for the correct size and absence of non-specific products. Real-time PCR was performed using a Mastercycler EP Gradient cycler with Realplex software (Eppendorf). Reactions were performed in 96-well plates with 25 μ L reaction volumes using the QuantiTect SYBR green PCR kit (Qiagen). A four-step program consisting of denaturation, annealing, extension, and data acquisition was used and melt curves verified that only one species of DNA product was amplified using the primer pairs. Calibration curves with R² values of at least 0.99 were performed for the CHC degradation genes using plasmid DNA titrated over five orders of magnitude, and were normalized to the housekeeping gene *chaA*. Calibration curves for the *chaA* gene were performed using genomic DNA. Primers were also designed for the antibiotic resistance gene present on each plasmid as an additional reference for normalization.

Biotin Production and Bioassay

E. coli BW-Z1 cells were transformed with either pJRB554 or pJRB621, or both plasmids. Duplicate colonies of the above strains and also the wild type cells were inoculated into LB medium containing appropriate antibiotics and grown at 37 °C with shaking. After 7 hours of growth, cells were inoculated (0.25% v/v) into M9 glucose (0.5%) medium containing appropriate antibiotics, 10 mM Na₂HCO₃, Hutner's mineral salts (Rolls and Lindstrom, 1967), and either 0.05 mM IPTG for cells containing pJRB554 or 100 nM aTc for cells containing pJRB621. After 22 hours of growth, cells were inoculated (0.25% v/v), into the same M9 glucose medium described above, using 1mM IPTG or 1 μ M aTc as inducers, where appropriate. In addition to the 2 cultures containing both the pJRB554 and pJRB621 plasmids,

an additional 2 cultures of this strain were set up using the identical medium supplemented with 0.3% L-alanine. After 23 hours of growth, cells were harvested and 1 mL of supernatant was obtained by filtration through a 0.22 μ M membrane (Millipore), which was frozen until later use. Biotin was measured using the auxotrophic bacteria *Lactobacillus plantarum* 8014 (ATCC), according to previously described methods (Waller, 1970), cultured in 1X biotin assay medium (ATCC) for ~40 hr. Standard curves of biotin were performed from 0.05 to 0.5 ng biotin per 5 mL culture tube and were linear with R^2 values of greater than 0.99. Sample measurements were performed by diluting filtered supernatants from 1X to 1,000X in water and adding 50 μ L to the 5 mL culture of *Lactobacillus* in biotin assay medium.

Results

The native DNA encoding the *R. palustris* CHC degradation pathway is unable to complement a Δ *bioH* *E. coli* strain

We chose to transfer the CHC degradation pathway from *R. palustris* to *E. coli* in order to develop a new metabolic route for pimeloyl-CoA synthesis in the production of biotin. If the pathway functioned properly in its host, then CHC would be degraded into pimeloyl-CoA, complementing the *E. coli* Δ *bioH* strain auxotrophic for this molecule. First we constructed pJRB502, a high copy plasmid containing the *badH**AliB**AbadK* operon from *R. palustris* under the control of its native promoter and transformed the Δ *bioH* strain. When grown in minimal medium lacking biotin, but containing CHC, the strain was unable to grow, however growth occurred in the presence of biotin, indicating that pimeloyl-CoA could not be generated from CHC to complement the biotin biosynthetic pathway (Fig 2a and Table 2). To evaluate whether the lack of complementation in this strain was due to poor recognition of the *R. palustris* promoter, we created pJRB510, a high copy plasmid expressing the native *R. palustris* *badH**AliB**AbadK* operon under the control of the P_{lacO-1} promoter and transformed the Δ *bioH* strain. When grown in minimal medium lacking biotin and containing CHC, this strain was also unable to grow, indicating that even in the presence of a strong, induced promoter, the pathway was not functionally expressed (Fig. 2a and Table 2). To evaluate whether the lack of complementation in this strain was due to the presence of rare codons encoded by the foreign DNA, we transformed pJRB510 into the Δ *bioH* strain in the presence of pRARE2, a plasmid expressing 7 rare tRNAs in *E. coli*. When the two plasmids were co-expressed in the pimeloyl-CoA auxotroph, the strain was still unable to grow in the presence of CHC, but lacking biotin, indicating that the pathway was not functionally expressed. Previously, we have noted that although native promoters from *R. palustris* are often recognized and transcribed by *E. coli*, translation of this DNA is severely limited by sequence elements in the 5'-UTRs of mRNAs (Bernstein et al., 2007).

The final three genes of the CHC degradation operon are poorly expressed in *E. coli*

To further understand *E. coli*'s expression of genes from the CHC degradation operon, we created protein fusions to LacZ of BadH, BadI and AliA in addition to the previously studied fusions to both AliB and BadK (Fig. 2b)(Bernstein et al., 2007). The expression of LacZ from the fusions with BadH or BadI were both higher than those of other *R. palustris* protein fusions (BadA, BadD, BadR, CrtI) previously tested using the same vector and strain of *E. coli* (Fig. 2c)(Bernstein et al., 2007). However, there was a precipitous drop in measured β -galactosidase activity from protein fusions to the final 3 genes of the operon, a difference of more than 2 orders of magnitude between the fusions to BadI and BadK, suggesting that the inability to degrade CHC to pimeloyl-CoA may result from incomplete expression of the genes from this pathway. The poor expression of the last 3 genes of the native operon, as measured by protein fusions to LacZ, was also corroborated by measuring mRNA transcript abundance, with decreasing amounts of transcript observed for each successive gene in the operon (see Fig. 4).

A synthetic CHC degradation operon with lower GC content is able to degrade CHC to pimeloyl-CoA in *E. coli*

Previously, we showed that the high GC content of the *R. palustris badD* UTR disabled *E. coli*'s ability to express a LacZ protein fusion construct transcribed from the fully induced P_{lacO-1} promoter (Bernstein et al., 2007). Replacement of the 14 bp RBS of *R. palustris* with an *E. coli* RBS enabled minimal expression from LacZ, indicating that altering the Shine-Dalgarno sequence is insufficient for robust expression of these high GC foreign genes. Replacement of the *R. palustris* UTR with an *E. coli* optimized UTR enhanced expression of this construct ~60-fold, while changing the *R. palustris* codons to the lower GC codons favored by *E. coli*, doubled expression from this promoter. We considered these results to create a synthetic CHC degradation operon optimized to *E. coli*. The codon usage of the open reading frames was altered to the lower GC codons favored by *E. coli* (Grote et al., 2005), and shorter, lower GC 5'-UTRs containing a strong RBS for each gene were employed (Fig. 3a, Table 3 and Supplementary Table 1). When the synthetic CHC degradation operon containing either the duplicated (pJRB620) or the single $P_{LtetO-1}$ promoter (pJRB621) was transformed to the $\Delta bioH$ strain of *E. coli* and grown in minimal medium, growth was dependent on the presence of CHC, complementing the strain's pimeloyl-CoA auxotrophy (Table 2). Growth of the $\Delta bioH$ strain containing pJRB621 was followed over time (Fig. 3b). The culture supplemented with biotin had the shortest lag phase and displayed an exponential doubling time of 95.3 minutes. The fully-induced culture in the presence of 1.5 mM CHC, but without biotin, had a slightly longer lag phase, while displaying a similar exponential doubling time of 91.6 minutes. The culture which contained 100-fold less inducer grew linearly throughout the time course, with a doubling time of approximately 200 minutes. Finally, when CHC was omitted from the medium, but the synthetic operon was fully induced, no growth was measured.

The optimized CHC degradation pathway maintains higher mRNA levels across the five gene operon than the native DNA sequence expressed from an inducible promoter

Once we determined that the optimized CHC pathway (pJRB621) was functionally expressed in *E. coli*, we compared its mRNA expression level using RT-PCR to the construct with native DNA sequence expressed from the P_{lacO-1} promoter (pJRB510). The two constructs behaved differently in *E. coli* with the native DNA experiencing a ~33-fold drop in mRNA abundance from *badH* to *badK* (Fig. 4), while the optimized construct experienced only a 3-fold drop in measured transcript abundance across these same genes. This phenomenon was observed when normalizing to either *chaA*, a housekeeping gene on the genome or when using the antibiotic resistance marker from the respective expression plasmids as a reference gene.

The CHC pathway in the presence of overexpressed biotin biosynthetic genes *bioBAFD* allows for the accumulation of biotin in minimal medium

After demonstrating the complementation of the $\Delta bioH$ phenotype of *E. coli*, we sought to determine the ability of the CHC degradation pathway to supply pimeloyl-CoA in the synthesis of the vitamin biotin. A high copy plasmid was designed to overexpress the *E. coli bioBAFD* genes under the control of the inducible P_{lacO-1} promoter. When this plasmid was transformed into wild type *E. coli*, no detectable biotin was measured (Table 4). However, in the presence of the CHC degradation operon and 1mM CHC, about 44 $\mu\text{g/L}$ of biotin was excreted into the medium, approximately 1,000 fold higher than the reported 15 ng/L of biotin excreted by wild type *E. coli* (Pai, 1972). When pJRB620, the synthetic CHC pathway containing the second $P_{LtetO-1}$ promoter was expressed in conjunction with the biotin production plasmid, similar amounts of biotin were measured in the medium compared with fermentations containing the single $P_{LtetO-1}$ promoter of pJRB621 (data not shown). The first enzymatic reaction in biotin synthesis subsequent to the production of pimeloyl-CoA involves the decarboxylative condensation of L-alanine by BioF (Fig. 1) (Streit and Entcheva, 2003). When L-alanine was

supplemented in the growth medium, production of biotin more than doubled to ~106 µg/L, indicating that integration of the two pathways can be further enhanced by supplying co-reactants.

Discussion

Recently, a number of studies have aimed to transfer large pieces of unaltered DNA from one bacterium to another such as the fusing of two genomes (Itaya et al., 2005), or the replacement of one genome with another (Lartigue et al., 2007). However, with these advances comes a new hurdle which must be addressed, namely the recognition of foreign or donor DNA by a host's transcription and translation systems such that it may be integrated with the underlying biological system to provide new hybrid functionality within the cell. Here we show the successful incorporation of the CHC degradation pathway from the high GC bacterium *R. palustris* to *E. coli*, allowing for a novel route to the pimeloyl-CoA pool destined for biotin production in the host. At first we cloned the native *R. palustris* operon, consisting of 5 genes over 5.3 kb, into a high copy plasmid expressed from either the native or inducible P_{lacO-1} promoter. The fact that these constructs were unable to complement a pimeloyl-CoA auxotroph in the presence of CHC, regardless of the inclusion of a plasmid expressing rare *E. coli* tRNAs, indicates there are significant challenges to expressing high GC DNA in this host. To alleviate this problem we synthesized a construct with lower GC content, by altering codon usage and replacing native 5'-UTRs with lower GC sequences containing a strong RBS. The synthetic operon had a final GC content of 54.3% which is significantly lower than the native DNA sequence at 66.3%. The drop in GC content, coupled with the inclusion of low GC 5'-UTRs, enables the functional expression of the CHC degradation operon from a single promoter. The largest fold change in β -galactosidase activity from the native, unaltered DNA pathway, as measured by protein fusion (Fig. 2c), occurs at *aliB*, the 3rd position of the operon, which possesses both the longest and highest GC 5'-UTR (Table 3). Additionally, the mRNA abundance decreased sequentially for each gene of the native operon, while the synthesized operon did not experience the same drop in transcript levels (Fig. 4). The fact that the *E. coli*-optimized operon was functionally expressed, while the native *R. palustris* pathway was not expressed in *E. coli*, is partially supported by a study linking *E. coli* translation initiation efficiency *in vitro* to lower GC content (Voges et al., 2004). The authors tested 756 constructs which differed in a 39 bp insert downstream of the start codon, fused to GFP. Low GC coding regions were more likely to achieve high expression, while high GC DNA in this position was more likely to cause low or no expression.

The creation of a new metabolic pathway to pimeloyl-CoA production in *E. coli*, allows for the replacement of the native pathway which has poorly defined biochemistry. Although *bioH* and *bioC* are both known to play a role in pimeloyl-CoA formation, without knowing the mechanism or role of other genes in this process, it is difficult to engineer the pathway. Indeed one group has reported that the overexpression of *bioH* in the presence of other biotin biosynthesis genes significantly decreased biotin productivity (Koga et al., 1996). Other groups have also studied biotin overproduction in *E. coli*, but did not address engineering the pimeloyl-CoA pool (Shaw et al., 1999). Their approach involved reorganizing biotin biosynthetic genes into a single operon, *bioBFCD*, expressed from the P_{tac} promoter, and further optimization of this operon, by removing a stem loop after *bioD* and changing the RBS for *bioB*, which catalyzes the final step in biotin production. The chemistry of CHC degradation is well studied (Egland et al., 1997) (Kuver et al., 1995) (Pelletier and Harwood, 2000), occurring within a cluster of genes for the anaerobic degradation of benzoate and 4-hydroxybenzoate (Egland and Harwood, 2000), offering further opportunities to expand the range of substrates which can be incorporated into useful metabolites in *E. coli*. As the cost of DNA synthesis lowers, the ability to synthesize large pathways will not only become affordable, but will offer unique advantages over traditional cloning. Not only will optimization of codons, GC content, promoters and 5'-

UTRs be facilitated, but the inclusion of unique restriction sites flanking each genetic building block will allow for considerable flexibility in further optimization.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This research was supported by the office of Science (BER), U.S. Department of Energy, Grant No. DE-FG02-07ER64490.

References

- Arvidson DN, et al. Automated kinetic assay of beta-galactosidase activity. *Biotechniques* 1991;11:733–734. 736, 738. [PubMed: 1809325]
- Bernstein JR, et al. Directed Evolution of Ribosomal Protein S1 for Enhanced Translational Efficiency of High GC *Rhodopseudomonas palustris* DNA in *Escherichia coli*. *J Biol Chem* 2007;282:18929–18936. [PubMed: 17412688]
- Bulter T, et al. A perspective of metabolic engineering strategies: moving up the systems hierarchy. *Biotechnol Bioeng* 2003;84:815–821. [PubMed: 14708122]
- Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 2000;97:6640–6645. [PubMed: 10829079]
- Dutton PL, Evans WC. The metabolism of aromatic compounds by *Rhodopseudomonas palustris*. A new, reductive, method of aromatic ring metabolism. *Biochem J* 1969;113:525–536. [PubMed: 5807211]
- Egland PG, Harwood CS. HbaR, a 4-hydroxybenzoate sensor and FNR-CRP superfamily member, regulates anaerobic 4-hydroxybenzoate degradation by *Rhodopseudomonas palustris*. *J Bacteriol* 2000;182:100–106. [PubMed: 10613868]
- Egland PG, et al. A cluster of bacterial genes for anaerobic benzene ring biodegradation. *Proc Natl Acad Sci U S A* 1997;94:6484–6489. [PubMed: 9177244]
- Ferrandez A, et al. Catabolism of phenylacetic acid in *Escherichia coli*. Characterization of a new aerobic hybrid pathway. *J Biol Chem* 1998;273:25974–25986. [PubMed: 9748275]
- Furste JP, et al. Molecular cloning of the plasmid RP4 primase region in a multi-host-range tacP expression vector. *Gene* 1986;48:119–131. [PubMed: 3549457]
- Garcia B, et al. Strategy for cloning large gene assemblages as illustrated using the phenylacetate and polyhydroxyalkanoate gene clusters. *Appl Environ Microbiol* 2004;70:5019–5025. [PubMed: 15294843]
- Grote A, et al. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res* 2005;33:W526–W531. [PubMed: 15980527]
- Harrison FH, Harwood CS. The pimFABCDE operon from *Rhodopseudomonas palustris* mediates dicarboxylic acid degradation and participates in anaerobic benzoate degradation. *Microbiology* 2005;151:727–736. [PubMed: 15758219]
- Itaya M, et al. Combining two genomes in one cell: stable cloning of the *Synechocystis* PCC6803 genome in the *Bacillus subtilis* 168 genome. *Proc Natl Acad Sci U S A* 2005;102:15971–15976. [PubMed: 16236728]
- Koga N, et al. Analysis of the bioH gene of *Escherichia coli* and its effect on biotin productivity. *Journal of Fermentation and Bioengineering* 1996;81:482–487.
- Kuver J, et al. Metabolism of cyclohexane carboxylic acid by the photosynthetic bacterium *Rhodopseudomonas palustris*. *Arch Microbiol* 1995;164:337–345. [PubMed: 8572887]
- Larimer FW, et al. Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris*. *Nat Biotechnol* 2004;22:55–61. [PubMed: 14704707]
- Lartigue C, et al. Genome transplantation in bacteria: changing one species to another. *Science* 2007;317:632–638. [PubMed: 17600181]

- Lessard IA, et al. Homologs of the vancomycin resistance D-Ala-D-Ala dipeptidase VanX in *Streptomyces toyocaensis*, *Escherichia coli* and *Synechocystis*: attributes of catalytic efficiency, stereoselectivity and regulation with implications for function. *Chem Biol* 1998;5:489–504. [PubMed: 9751644]
- Lutz R, Bujard H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* 1997;25:1203–1210. [PubMed: 9092630]
- Marquet A, et al. Biosynthesis of biotin and lipoic acid. *Vitam Horm* 2001;61:51–101. [PubMed: 11153271]
- Mutka SC, et al. Heterologous production of epothilone C and D in *Escherichia coli*. *Biochemistry* 2006;45:1321–1330. [PubMed: 16430229]
- Pai CH. Mutant of *Escherichia coli* with derepressed levels of the biotin biosynthetic enzymes. *J Bacteriol* 1972;112:1280–1287. [PubMed: 4565539]
- Pelletier DA, Harwood CS. 2-Hydroxycyclohexanecarboxyl coenzyme A dehydrogenase, an enzyme characteristic of the anaerobic benzoate degradation pathway used by *Rhodopseudomonas palustris*. *J Bacteriol* 2000;182:2753–2760. [PubMed: 10781543]
- Rolls JP, Lindstrom ES. Effect of thiosulfate on the photosynthetic growth of *Rhodopseudomonas palustris*. *J Bacteriol* 1967;94:860–869. [PubMed: 6051358]
- Samanta SK, Harwood CS. Use of the *Rhodopseudomonas palustris* genome sequence to identify a single amino acid that contributes to the activity of a coenzyme A ligase with chlorinated substrates. *Mol Microbiol* 2005;55:1151–1159. [PubMed: 15686561]
- Sanyal I, et al. Biosynthesis of pimeloyl-CoA, a biotin precursor in *Escherichia coli*, follows a modified fatty acid synthesis pathway: ¹³C-labeling studies. *J. Am. Chem. Soc* 1994;Vol. 116:2637–2638.
- Shaw N, et al. Biotin production under limiting growth conditions by *Agrobacterium/Rhizobium* HK4 transformed with a modified *Escherichia coli* bio operon. *J Ind Microbiol Biotechnol* 1999;22:590–599. [PubMed: 10455485]
- Simons RW, et al. Improved single and multicopy lac-based cloning vectors for protein and operon fusions. *Gene* 1987;53:85–96. [PubMed: 3596251]
- Streit WR, Entcheva P. Biotin in microbes, the genes involved in its biosynthesis, its biochemical role and perspectives for biotechnological production. *Appl Microbiol Biotechnol* 2003;61:21–31. [PubMed: 12658511]
- Sullivan JT, et al. The bio operon on the acquired symbiosis island of *Mesorhizobium* sp. strain R7A includes a novel gene involved in pimeloyl-CoA synthesis. *Microbiology* 2001;147:1315–1322. [PubMed: 11320134]
- Tomczyk NH, et al. Purification and characterisation of the BIOH protein from the biotin biosynthetic pathway. *FEBS Lett* 2002;513:299–304. [PubMed: 11904168]
- Voges D, et al. Analyzing and enhancing mRNA translational efficiency in an *Escherichia coli* in vitro expression system. *Biochem Biophys Res Commun* 2004;318:601–614. [PubMed: 15120642]
- Waller JR. Increased sensitivity of the microbiological assay for biotin by *Lactobacillus plantarum*. *Appl Microbiol* 1970;20:485–491. [PubMed: 5485730]
- Weaver LH, et al. Competing protein:protein interactions are proposed to control the biological switch of the *E coli* biotin repressor. *Protein Sci* 2001;10:2618–2622. [PubMed: 11714930]

Abbreviations

CHC, cyclohexane carboxylate; RBS, ribosome-binding sites; UTR, untranslated region; IPTG, isopropyl- β -D-thiogalactopyranoside; aTc, anhydrotetracycline.

Figure 1.

(A) Degradation of cyclohexanecarboxylic acid to pimeloyl-CoA as performed by *R. palustris*. (B) Pimeloyl-CoA is the common intermediate in microbial biotin biosynthesis catalyzed by four enzymes. The final step of biotin synthesis catalyzed by biotin synthase (BioB) is not fully understood, but inserts a sulfur into dethiobiotin, forming a tetrahydrothiophene ring (Marquet et al., 2001; Streit and Entcheva, 2003). The mechanism of pimeloyl-CoA synthesis in *E. coli* has not been determined, but requires *bioC* and *bioH*.

Figure 2.

(A) Constructs expressing the native *R. palustris* cyclohexane carboxylate degradation operon from either the native or inducible P_{lacO-1} promoter. Growth of a $\Delta bioH$ pimeloyl-CoA auxotroph of *E. coli* did not occur with or without 1mM CHC when transformed with these constructs, but was able to grow when supplemented with biotin. Gray is used for the *R. palustris* sequence and white is used for *E. coli* DNA. (B) Protein fusions of the native *R. palustris* DNA were made to the ~20th amino acid of each protein from the operon to a LacZ variant missing its first 9 amino acids. Constructs were expressed in *E. coli* from the native *R. palustris* promoter. Gray is used for the *R. palustris* sequence and white is used for *E. coli* DNA. (C) β -galactosidase activity as measured from the 5 protein fusions of the native CHC degradation pathway.

Figure 3.

(A) A DNA-synthesized CHC degradation pathway was created with lower GC content by changing codon usage and replacing the native high GC 5'-UTRs with shorter low GC sequences, expressed using a repeat of the $P_{LtetO-1}$ promoter (pJRB620). Unique restriction sites flank each gene, while the second $P_{LtetO-1}$ promoter and T_0 transcription terminator are flanked by a duplicated restriction site, allowing for the removal of the second promoter and transcription terminator, creating a construct (pJRB621) where the pathway is expressed from the single $P_{LtetO-1}$ promoter upstream of *badH*. White coloring denotes *E. coli*-optimized DNA sequence. (B) CHC-dependent growth of the $\Delta bioH$ pimeloyl-CoA auxotroph strain of *E. coli* containing pJRB621 in minimal medium. Open squares, 1mM CHC plus biotin; closed diamonds, 1mM CHC only; closed triangles, 100-fold less inducer (aTc) plus 1mM CHC; open circles, without CHC.

Figure 4.

Normalized mRNA abundance as measured by RT-PCR of genes from the CHC degradation operon with either the native *R. palustris* DNA sequence (pJRB510, Fig. 2a), expressed from the P_{lacO-1} promoter, or the *E. coli*-optimized sequence (pJRB621, Fig. 3a), expressed from $P_{LtetO-1}$. Transcripts are normalized to the same housekeeping gene, *chaA*.

Table 1

Bacterial strains, plasmids and primers

strain/plasmid/primer	relevant features	reference/source
<i>R. palustris</i> CGA010	Wild type strain used for cloning <i>badH</i> and <i>badK</i> genes and promoter	(Larimer et al., 2004)
<i>E. coli</i> BW25113	Strain used for LacZ expression studies	(Lessard et al., 1998), (Datsenko and Wanner, 2000)
BW25113-Z1	BW25113 <i>E. coli</i> strain with LacR, TetR, Sp ^r moved from DH5αZ1 by phage P1 transduction	(Lutz and Bujard, 1997), (Bernstein et al., 2007)
Δ <i>bioH</i> -Z1	BW25113 <i>E. coli</i> strain with knockout of the Δ <i>bioH</i> gene. LacR, TetR, Sp ^r moved from DH5αZ1 by phage P1 transduction	this study
<i>Lactobacillus plantarum</i> 8014	Strain used in biotin bioassay	ATCC
pTB120	Promoter-less <i>lacZ</i> , preceded and followed by <i>rrnB</i> terminator; used for protein fusion; high copy; Ap ^r	(Bernstein et al., 2007)
pJRB502	P _{badH} :: <i>badH</i> and <i>badK</i> ; native <i>R. palustris</i> operon and promoter; Ap ^r	this study
pJRB510	P _{lacO-1} :: <i>badH</i> and <i>badK</i> ; native <i>R. palustris</i> operon; Ap ^r	this study
pJRB620	P _{LietO-1} :: <i>badH</i> and P _{LietO-1} :: <i>aliB</i> and <i>badK</i> ; <i>E. coli</i> codon-optimized pathway; Cm ^r	this study
pJRB621	P _{LietO-1} :: <i>badH</i> and <i>badK</i> ; <i>E. coli</i> codon-optimized pathway; Cm ^r	this study
pJRB554	P _{lacO-1} :: <i>bioBAFD</i> ; <i>E. coli</i> biotin biosynthetic genes; Ap ^r	this study
primers		
badHFor	AATAGAATTCACCGTGGGTCGAGACAGCGC	this study
badHrev	AATTGGGATCCTCGCGCAACCGGCGG	this study
badIrev	ATAGGATCCGACGATTGATGATGATCCACG	this study
aliBrev	ATAGGATCCCGGTCTCGCTCCTGAAATCC	this study
aliArev	ATAGGATCCACGCAGGCGTCGAGATCGT	this study
badKrev	ATAGGATCCCGGTTCAGCGTGATGATGCC	this study
bioBFor(eco)	CGGAATTCATTAAGAGGAGAAATACCATGGCTACCGCCCA	this study
	CGCTG	
bioBRev(xba)	TGCTCTAGATTCTCCTAATCATAATGCTGCCGCTTGTAATAT	this study
	TCG	
bioAFor(xba)	TGCTCTAGATGACAACGGACGATCTTGCCCTTTG	this study
bioARev(cla)	ACCATCGATTCTCCTTATTGGCAAAAAATGTTTCATCCTGTGA	this study
	CC	
bioFFor(claI)	ACCATCGATATGAGCTGGCAGGAGAAAAATCAACG	this study
bioFRev(spe)	CATAAGCTTCAGAATGGCTACAACAAGGCAAGGTT	this study
bioDFor(spe)	GGACTAGTATGAGTAAACGTTATTTTGTACCGGAACG	this study
bioDRev(hind)	CATAAGCTTCAGAATGGCTACAACAAGGCAAGGTT	this study
badH+rbsFor(eco)	CGGAATTCGAAATCAAAGGGGAGCGAGCGAG	this study
badH-Krev(hind)	CATAAGCTTGCACGCTAGCGGTGGGAGAA	this study
badH+pr(eco)	AATAGAATTCACCGTGGGTCGAGACAGCGC	this study
P1 (Δ <i>bioH</i>)	CTACACCTCTGCTTCAACGCCACCAGCAGGTGACAAAACCTC	this study
	GGCCGTGTAGGCTGGAGCTGCTTC	
P2 (Δ <i>bioH</i>)	ATGAATAACATCTGGTGGCAGACCAAGGTCAGGGGAATGTT	this study
	CATCCATATGAATATCCTCCTTAG	

Table 2

Growth of a $\Delta bioH$ strain of *E. coli* in minimal medium lacking biotin transformed with various CHC degradation constructs.

Construct	(-) CHC	(+) CHC	(+) CHC, (+) Biotin
pJRB502 (native <i>R. palustris</i> pathway and promoter)	–	–	+
pJRB510 (native <i>R. palustris</i> pathway, P _{lacO-1} promoter)	–	–	+
pJRB510 + pRARE2 (rare tRNAs)	–	–	+
pJRB620 (<i>E. coli</i> -optimized pathway, duplicated P _{LtetO-1} promoter)	–	+	+
pJRB621 (<i>E. coli</i> -optimized pathway, single P _{LtetO-1} promoter)	–	+	+

Table 3

GC content and length of 5'-UTR for the genes of the CHC degradation pathway (native DNA and *E. coli*-optimized constructs).

Gene	% GC content in 5'-UTR (<i>R. palustris</i> / <i>E. coli</i>)	Length of 5' UTR (Base pairs) (<i>R. palustris</i> / <i>E. coli</i>)
<i>badH</i>	58.3 / 38.1	24 / 21
<i>badI</i>	53.3 / 41.7	30 / 12
<i>aliB</i>	68.8 / 42.1	64 / 19
<i>aliA</i>	60.7 / 42.1	56 / 19
<i>badK</i>	63.0 / 47.4	27 / 19

Table 4

Biotin production of *E. coli* cells in minimal medium with 1mM CHC, expressing variations of the *R. palustris* cyclohexane carboxylate degradation pathway and the *E. coli* *bioBAFD* biotin biosynthetic genes.

Construct / condition	Biotin production (µg/L)
BW-Z1 (wild type)	ND
621 (CHC operon)	ND
554 (<i>bio</i> operon)	ND
621 / 554	43.7 ± 3.2
621 / 554 + 0.3% L-Alanine	106.1 ± 3.2

Data are the average of two independent colonies grown in M9 glucose (0.5%) minimal medium, supplemented with 1mM CHC. ND denotes biotin concentrations which were not detected by the lower limit of the bioassay.