Toward Biosynthetic Design and Implementation of *Escherichia coli*-Derived Paclitaxel and Other Heterologous Polyisoprene Compounds

Ming Jiang,[a] Gregory Stephanopoulos,[b] and Blaine A. Pfeifer[a]

Department of Chemical and Biological Engineering, State University of New York at Buffalo, Buffalo, New York, USA,[a] and Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA[b]

Escherichia coli offers unparalleled engineering capacity in the context of heterologous natural product biosynthesis. However, as with other heterologous hosts, cellular metabolism must be designed or redesigned to support final compound formation. This task is at once complicated and aided by the fact that the cell does not natively produce an abundance of natural products. As a result, the metabolic engineer avoids complicated interactions with native pathways closely associated with the outcome of interest, but this convenience is tempered by the need to implement the required metabolism to allow functional biosynthesis. This review focuses on engineering *E. coli* for the purpose of polyisoprene formation, as it is related to isoprenoid compounds currently being pursued through a heterologous approach. In particular, the review features the compound paclitaxel and early efforts to design and overproduce intermediates through *E. coli*.

Isoprenoids are an important class of natural products that have been converted to several therapeutic medicines (2, 41, 42). However, this process can be complicated by the need to harness or reconstruct native biosynthesis. As an example, the well-known isoprenoid natural product paclitaxel is natively produced by the Pacific yew tree (*Taxus brevifolia*), yet utilizing production from the native host proved both uneconomical and environmentally destructive (5, 17, 25). An economical total chemical-synthetic route to the compound is challenging due to the molecule's complex final architecture. This common scenario has driven alternative approaches toward the viable production of this and other therapeutically relevant natural products. One such approach is the use of heterologous biosynthesis, in which the genetic content responsible for a given natural compound is transplanted from the native organism to a surrogate host for the purpose of leveraging the new host's innate biological and engineering capabilities. This review focuses on the heterologous biosynthesis approach, using the compound paclitaxel as a primary example. More specifically, emphasis is placed on the logic and initial steps guiding heterologous biosynthesis, including the choice of host organism, the provision of required intracellular substrates, and the initiation of polyisoprene formation. For more comprehensive or alternative discussions of isoprenoid biosynthesis, readers are directed to several additional review articles (1, 11, 21, 30, 31, 33, 40, 46, 67).

Isoprenoids are often associated with plant natural products, but they are also produced by microorganisms (19, 58). The compounds play several roles in basic cellular function, including light absorption, electron transport, and protein modification. In addition, they are perhaps better known for their roles in less clearly defined chemical communication scenarios. For example, cellular damage and potential foreign-organism invasion both spur isoprenoid biosynthesis in an attempt to protect plant viability. Such a capacity for isoprenoids to serve as chemical cues in response to environmental stimuli, including as repellents to foreign organism invasion, provides one supporting theory as to their medicinal properties, which have been widely recognized (42). In addition to paclitaxel, other isoprenoids with therapeutic value include artemisinin (antimalarial), lycopene (antioxidant), and astaxanthin (antioxidant).

From the standpoint of harnessing this medicinal potential, production processes may then take advantage of the native hosts. From one perspective, this approach simplifies production efforts by utilizing a host that has been evolutionarily optimized for compound formation. Attempts to produce paclitaxel have included semisynthesis in which a key intermediate harvested from the needles of the yew tree is chemically converted to the final active compound (24). More recently, plant cell culture has been used to generate paclitaxel, and there is active research dedicated to characterizing and optimizing this process (53). As mentioned above, harvesting the full compound from the bark of the yew tree proved impractical, spurring the more environmentally friendly and scalable semisynthetic and plant cell culture approaches. However, plant-derived production processes, while taking advantage of native paclitaxel biosynthetic capabilities, are also dictated by the intrinsic biological traits associated with the plant host (17). The resulting cell cycle kinetics associated with plants will clearly fall behind the high growth rates associated with microbial hosts like *Escherichia coli* and *Saccharomyces cerevisiae*. As a result, a classical heterologous production situation is presented: can the biosynthesis of a complex therapeutic natural product be established in a technically superior, yet evolutionarily unoptimized, alternative host?

Within the last 15 years, this question has been asked and pursued with increasing frequency. Driven by the possibility of expediting and simplifying the production of therapeutic isoprenoid compounds, heterologous biosynthesis has also taken advantage of rapid technology development for candidate surrogate hosts. As a result, a series of successful examples have emerged, prompting continued efforts to apply and extend the technology. Each success story, however, must carefully address the biosynthetic needs that are in-built within the original producers. In essence,

heterologous biosynthesis attempts to couple the medicinal value of therapeutic natural products, like isoprenoids, with the engineering possibilities associated with the chosen alternative hosts. This review explains this process, featuring the recent application of the approach toward early-stage paclitaxel intermediates and the emergence of technical advances that have aided and will continue to aid heterologous biosynthetic attempts.

HETEROLOGOUS HOST CHOICE

Generally, the traits that define a promising heterologous host relate to technical convenience and engineering potential. Growth speed and advanced molecular biology protocols are an understated reason why *E. coli* and *S. cerevisiae* are commonly selected as heterologous hosts. In particular, the abundance of experimental protocols associated with these hosts derived from the rapid growth kinetics associated with each. Prior to heterologous biosynthetic applications, many experimental protocols and tools were developed for recombinant protein production. Since then, these same tools have been adopted or modified for heterologous biosynthetic efforts, providing a formidable arsenal of options to aid heterologous reconstitution of desired isoprenoid pathways.

Another criterion for the selection of a heterologous host is the intrinsic ability to support heterologous biosynthesis. This is the main advantage of basing production upon a native host system; the original host will have the innate, evolutionarily optimized ability to support biosynthesis. The selection of a heterologous host may be biased by the potential of the alternative host to contribute to the biosynthetic effort, and this potential must be weighed against growth rate advantages and engineering opportunities.

In this regard, it is fortuitous that the most convenient heterologous host options have the native metabolism required to support isoprenoid biosynthesis. *E. coli*, *S. cerevisiae*, and *Bacillus subtilis*, arguably the most commonly used recombinant microbial systems, each have the ability to generate isoprenoid compounds, although none display the health-related properties associated with isoprenoids from plant or more fastidious microbial sources. Nonetheless, this built-in capability can be utilized in the context of heterologous biosynthesis.

Table 1 summarizes the use of *E. coli* and *S. cerevisiae* in the heterologous production of isoprenoid compounds. The literature cited in the table summarizes efforts to establish heterologous biosynthesis according to compound and host organism. There have been several highly visible examples of successful compound formation through both *E. coli* and *S. cerevisiae*, implying that both have the potential to be viable production options. The choice may then be decided by compound-specific requirements, the host-specific potential to meet these requirements, and the experimental preference of the individual researcher. Interestingly, although the genus *Bacillus* has an apparently robust ability to produce isoprene (34), only a few efforts have been made to use the organism for isoprene or polyisoprene formation (68).

SUBSTRATE PROVISION, PATHWAY SELECTION, AND IMPACT ON HETEROLOGOUS BIOSYNTHESIS

There are two distinct metabolic pathways capable of providing the required isopentenyl diphosphate (IPP) starting substrate for isoprenoid biosynthesis. The first route, the mevalonate (MVA) pathway, was identified in the 1950s and completely elucidated shortly thereafter (7, 44). The MVA pathway is primarily found in

eukaryotes and starts with the common intracellular metabolite acetyl coenzyme A (acetyl-CoA), which is converted through a series of six reactions to IPP. As Fig. 1 indicates, multiple routes have been used to provide the acetoacetyl-CoA intermediate, including an acetoacetyl-CoA synthase and an acetoacetyl-CoA ligase (22, 47). With regard to the heterologous hosts previously referenced, the MVA pathway is natively found in *S. cerevisiae*.

E. coli possesses the alternative pathway for IPP formation which begins with a molecule each of pyruvate and glyceraldehyde-3-phosphate. This route is referred to as the 2C-methyl-D-erythritol-4-phosphate (MEP) pathway and produces both IPP and dimethylallyl diphosphate (DMAPP) (20, 35, 55). The MVA pathway includes an essential isomerase to convert between IPP and DMAPP, whereas the MEP pathway isomerase is not required, but such activity is used to alter the ratio between IPP and DMAPP for the purpose of subsequent product formation. The MEP pathway features seven steps, and both the MVA and MEP pathways require energy units (ATP and CTP) and reducing equivalents (NADPH) during substrate generation (Fig. 1).

Table 1 also summarizes the substrate provision pathways used in heterologous production attempts. It is interesting that the MVA pathway has been successfully implemented into *E. coli* despite the existence of the organism's endogenous MEP pathway. A primary reason given for this approach is the unknown regulatory features associated with the native MEP pathway and how these might limit subsequent engineering. As a result, a completely heterologous pathway would then be decoupled from native control and would be easier to alter for the purpose of supporting eventual isoprenoid biosynthesis (39). Alternatively, the MEP pathway has been used extensively and successfully during *E. coli*-based heterologous efforts. More recently, two studies compared both substrate provision pathways during efforts to generate a single isoprenoid compound using *E. coli* (45, 54). In both reports, the MVA pathway supported higher isoprenoid titers under the specific conditions of the experimental systems employed. As is discussed below in the case of *E. coli*-derived paclitaxel intermediate biosynthesis, such head-to-head comparisons of the MVA and MEP pathways can be difficult to interpret without an exhaustive effort to balance and optimize precursor supply toward final isoprenoid biosynthesis.

Use of a heterologous substrate provision pathway has been successful in supporting accompanying biosynthetic efforts and also helps to emphasize general considerations when heterologous biosynthesis is being attempted. The reliance on a nonnative pathway carries an initial requirement for eventual success: functional gene expression. This would appear to be of little concern with a system native to *E. coli*, like the MEP pathway. But a heterologous MVA pathway may pose the common problems that have long plagued the use of *E. coli* for the production of recombinant proteins, namely, that foreign gene expression elements may not transition to the new host. However, thanks to the range of expression and control elements (promoter, operator, ribosomal binding site, and termination sequences) identified and utilized for *E. coli*, challenges in foreign gene expression initiation are becoming more manageable, with a key result being a greater frequency of successful final enzyme activity. Another advance that has aided heterologous gene expression is the use of gene synthesis to address codon bias between organisms. In this way, native codons associated with the heterologous MVA pathway genes can be replaced with those optimal for *E. coli*. Such technical advances have

TABLE 1 Representative examples of heterologously produced isoprenoids, along with hosts, supporting substrate provision pathways, product titers, and engineering steps that influenced subsequent studies or titer improvements[a]

Compound	Heterologous host	Substrate provision pathway	Final titer(s) (intermediate or final compound)[b]	Step to enable or overproduce product	Reference
Paclitaxel	E. coli	MEP	1.3 mg/liter (taxadiene)	Overexpression of *dxs* and *idi*	26
			300 mg/liter (taxadiene), 58 mg/liter (taxadiene-5α-ol)	Overexpression and variation of copy number and promoter of *dxs*, *idi*, *ispD*, and *ispF*	2
	S. cerevisiae	MVA	1.0 mg/liter (taxadiene), 0.025 mg/liter (taxadiene-5α-ol)	None	15
			8.7 mg/liter (taxadiene)	Overexpression of a truncated HMG-CoA reductase; reduction of competing steroid pathway	16
Artemisinin	E. coli	MVA	155 mg/liter (amorphadiene; initial titer)	Introduction and overexpression of MVA pathway; foreign gene codon optimization; promoter variation; plasmid copy number variation; specific overexpression of mevalonate kinase	39
			1,084 mg/liter[c] (amorphadiene; optimized titer)		4
			105 mg/liter (artemisinic acid)	Overexpression of MVA pathway; foreign gene codon optimization; promoter variation	10
	S. cerevisiae	MVA	0.6 mg/liter	None	38
			153 mg/liter (amorphadiene), 32 mg/liter (artemisinic acid)	Overexpression of a truncated HMG-CoA reductase; reduction of competing steroid pathway	52
Carotenoids	E. coli	MEP	6 mg/g DCW (β-carotene)	Overexpression of *dxs*, *ispD*, *ispF*, *ispB*, and *idi*	72
		MVA	22 mg/g DCW (lycopene)	Overexpression of *mvaK1* (*mvk*), *mvaK2* (*pmk*), *mvaD* (*pmd*), and *idi*	71
			12.5 mg/g DCW (lycopene)	Overexpression of MVA pathway with acetoacetyl-CoA ligase introduced	21
	S. cerevisiae	MVA	0.113 mg/g DCW (lycopene)	None	70
			5.9 mg/g DCW (β-carotene)	Overexpression of a truncated HMG-CoA reductase and endogenous geranylgeranyl diphosphate synthase	64

[a] Examples are not comprehensive.
[b] DCW, dry cell weight.
[c] Estimated based upon an optical density at 600 nm (OD_{600}) of 3.7 (39); the actual value reported is 293 mg/liter/unit of OD_{600}.

developed in parallel with heterologous biosynthesis, and it is reasonable to conclude that the latter has benefited from the former. Besides gene expression, additional steps have also been useful in ensuring functional enzyme activity, including the adjustment of culture temperature, the use of chaperonin proteins, and the further modification of gene sequences to remove any foreign signaling or localization amino acids (or in some cases, to add such sequences) (2, 39, 73).

The considerations outlined in the paragraph above accompany any heterologous biosynthetic effort. As a result, there is a lessened degree of concern in attempts to achieve substrate provision through the endogenous MEP pathway. There have also been stoichiometric and energetic arguments that the MEP pathway is capable of more efficient substrate conversion (2,

74). However, there is the question of native utility and how the evolutionary control elements associated with endogenous activity will respond to genetic and metabolic engineering. For the purpose of isoprenoid heterologous biosynthesis, either the native or heterologous substrate provision pathways will potentially overlap native regulatory or metabolic networks. This awareness heightens the need for metabolic engineering, which is discussed below.

In the case of attempts to engineer heterologous biosynthesis of paclitaxel in *E. coli*, the MEP pathway was selected for substrate provision (2). This choice was made, in part, due to the previously mentioned potential for improved theoretical yield and as a result of a growing number of examples in which the pathway supported polyisoprene formation (Table 1). However, the native chromo-
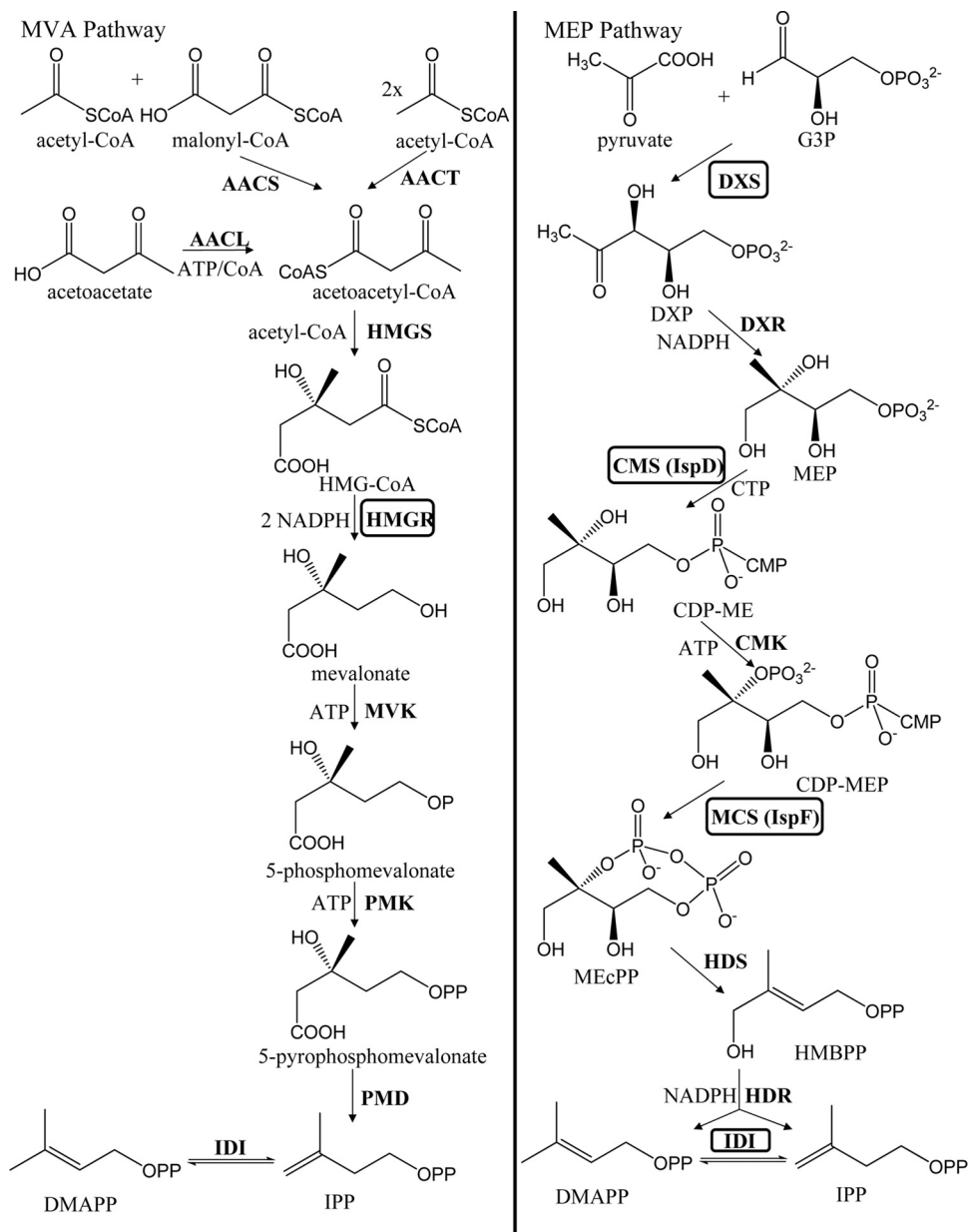
FIG 1 Mevalonate (MVA) and 2C-methyl-D-erythritol-4-phosphate (MEP) polyisoprene substrate provision pathways. Enzymes that have been targeted during the course of metabolic engineering studies for improved paclitaxel intermediate titers are shown in boxes. Specific *E. coli* enzyme designations are in parentheses. AACS, acetoacetyl-CoA synthase; AACT, acetoacetyl-CoA thiolase; AACL, acetoacetyl-CoA ligase; HMGS, hydroxymethylglutaryl (HMG)-CoA synthase; HMGR, HMG-CoA reductase; MVK, mevalonate kinase; PMK, phosphomevalonate kinase; PMD, mevalonate pyrophosphate decarboxylase; IDI, isopentenyl diphosphate isomerase; IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; G3P, glyceraldehyde-3-phosphate; DXS, 1-deoxy-D-xylulose 5-phosphate (DXP) synthase; DXR, DXP reductase; MEP, 2-*C*-methyl-D-erythritol 4-phosphate; CMS, 4-diphosphocytidyl-2-*C*-methyl-D-erythritol (CDP-ME) synthase; CMK, CDP-ME kinase; CDP-MEP, CDP-ME 2-phosphate; MCS, 2C-methyl-D-erythritol 2,4-cyclodiphosphate (MEcPP) synthase; HDS, 1-hydroxy-2-methyl-2-(*E*)-butenyl 4-diphosphate (HMBPP) synthase; HDR, HMBPP reductase.

somal version of the MEP pathway was incapable of supporting paclitaxel overproduction. Instead, based upon information regarding MEP pathway bottleneck points (72), additional copies of the *dxs*, *ispD*, *ispF*, and *idi* genes were introduced at a range of gene dosage levels. This particular strategy was adopted because of wide experimental variation in resulting taxadiene titers (the first dedicated intermediate in the paclitaxel pathway) without a systematic approach to expressing either the genes associated with substrate provision or those needed for eventual taxadiene formation.

The results suggested the potential for high-titer biosynthesis, but a careful experimental design was required to assess and establish the correct metabolic environment for consistent overproduction. This was accomplished by balancing the flux of carbon through both substrate provision (referred to as the upstream pathway) and subsequent taxadiene formation (the downstream pathway). There may very well be remaining inefficiencies with this system and the MEP pathway, but this result demonstrated the potential of the MEP pathway to be used for high-level iso-
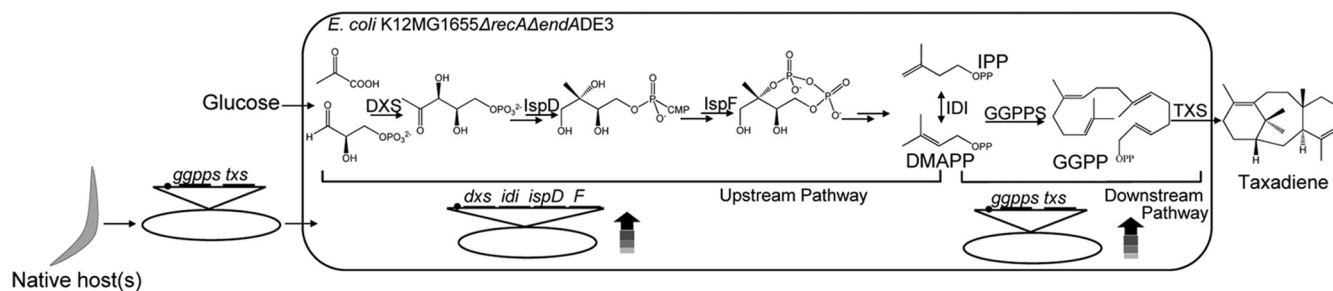
FIG 2 Heterologous taxadiene design, implementation, and optimization. Strain MG1655 was engineered for plasmid stabilization (ΔrecA ΔendA) and to support the T7 promoter (DE3). Having been previously identified and sequenced, the heterologous geranylgeranyl diphosphate synthase (*ggpps*) and taxadiene synthase (*txs*) genes were synthesized, codon optimized, and truncated to remove suspected signaling regions before being introduced via plasmid. Both native bottleneck MEP genes and the heterologous genes were designed for a range of expression levels through modifications to promoter (black circle preceding operon) strength (*trc*, T5, and T7) and gene dosage. Upstream and downstream pathway balancing allowed optimized taxadiene titers.

prenoid biosynthesis in the context of heterologous taxadiene production.

FORMATION OF POLYISOPRENE UNITS

Prior to taxadiene biosynthesis, which is the result of a dedicated terpene synthase, the polyisoprene unit geranylgeranyl diphosphate (GGPP) must be formed. From IPP and DMAPP, polyisoprene units are generated through the consecutive condensation of IPP to allylic diphosphate compounds (typically beginning with a DMAPP unit). The resulting C_{10} (geranyl diphosphate [GPP]), C_{15} (farnesyl diphosphate [FPP]), and C_{20} (geranylgeranyl diphosphate) chains are the basis for conversion by additional enzymes into isoprenoid classes of monoterpenes, sesquiterpenes, and diterpenes, respectively (57). *E. coli* has the native capability to produce FPP as a result of the FPP synthase IspA (18). However, the cell accumulates only limited amounts of GGPP (6, 28).

As a result, a GGPP synthase (GGPPS) from *Taxus canadensis* was used to allow the production of paclitaxel intermediates from *E. coli* (2, 23). Subject to the same design parameters introduced above for heterologously expressed genes, the GGPP synthase gene was synthesized, with care being taken to optimize codon usage for *E. coli* and to remove any unneeded native plastidial insertion regions. This gene, in conjunction with a taxadiene synthase and bottleneck MEP genes, was then tested over a wide expression range. Chromosomal, low, medium, and higher gene dosages were implemented through chromosomal integration or the use of plasmids with well-established copy numbers. Gene dosage was coupled to a range of promoters also varying in strength. As a result, a comprehensive landscape of metabolic control was mapped to taxadiene production titers, and a maximum was identified for this particular system. The resulting balance of upstream substrate provision and downstream biosynthetic pathways resulted in a taxadiene titer of ~300 mg/liter. Figure 2 depicts the steps involved in heterologous isoprenoid biosynthesis featuring the paclitaxel case.

AFTER BIOSYNTHESIS: PROTEIN, METABOLIC, AND PROCESS ENGINEERING

A heterologous biosynthetic attempt can be classified as a success if any level of production is accomplished. However, success involving only low levels of production will not serve the larger objective of viable mass production. Here, the choice of heterologous biosynthesis offers a second level of engineering to improve upon initial titers, yields, and productivities. Similar to the advan-

tages afforded by the heterologous host during attempts at pathway reconstitution, the engineering tools available with hosts like *E. coli* and *S. cerevisiae* offer a number of engineering opportunities across scales.

At the protein level, specific enzymes within the precursor or polyisoprene formation stages of biosynthesis have been targeted for improvements in final titers. The GGPPS responsible for initiating the downstream isoprenoid formation process has been a frequent target for protein engineering strategies. Both direct and random mutations have been used to alter or improve the capabilities of this enzyme (37, 50, 65). In these studies (and numerous others), the engineering schemes were aided or enabled by the colorimetric properties of carotenoid compounds. Other groups have used dedicated foreign GGPPSs to aid in a particular isoprenoid biosynthetic effort (similar to the paclitaxel case). This is done to channel substrate flow to the correct-chain-length polyisoprene unit prior to subsequent isoprenoid formation reactions (16, 66).

Metabolic engineering in the context of isoprenoid heterologous biosynthesis has primarily been dedicated to the MEP and MVA pathways. Numerous studies have focused on individual components of these pathways to identify and address potential bottlenecks (a comprehensive assessment of these studies is presented in the reviews cited at the beginning of this article). With regard to the paclitaxel case, several groups have optimized production of related compounds through chromosomal modification of the MEP pathway (12, 72). This approach appears to be particularly well suited to the native *E. coli* MEP pathway and has the potential to both minimize metabolic burden and eliminate the need for plasmids, as well as to maintain the dynamic range of gene expression through variation in promoter strength. The well-curated metabolic backgrounds of hosts like *E. coli* and *S. cerevisiae* also offer the application of computational modeling to predict and implement beneficial metabolic changes. This has been demonstrated in a number of cases using stoichiometric modeling of background metabolism to identify genes to be either deleted or overexpressed for the purpose of improving final titers (3, 13, 61, 62). The advantage of such models is the ability to identify alternative native metabolism contributing to or detracting from heterologous production. In a recent effort focused on identifying gene overexpression targets capable of improving *E. coli*-derived taxadiene, the *ppk*, *sthA*, *purN*, and *folD* targets were computationally and experimentally overexpressed, with three of four tar-

gets resulting in improved taxadiene levels (8). Finally, given the native and heterologous components associated with or influencing total biosynthesis, studies have begun to use various -omics technologies as another route to characterize and potentially engineer the biosynthetic process (9, 32, 49, 51).

Hosts like *E. coli* also have well-developed bioreactor protocols available to boost cell density and volumetric productivity (36, 60, 69). These strategies were applied in the paclitaxel case to build upon the metabolic engineering of upstream and downstream pathway balances to push final taxadiene titers to >1 g/liter. Other efforts in isoprenoid heterologous biosynthesis have shown similarly impressive final titers (52, 63). In summary, *E. coli* and other well-characterized heterologous host systems offer more than excellent molecular biology tools available to facilitate pathway transfer and reconstitution. The new hosts have been the primary platforms for bioproduct protein, metabolic, and process engineering for the last 3 decades, and this knowledge base is supportive of the product development strategies that are now being applied to isoprenoid heterologous biosynthesis.

COMPLETE ISOPRENOID BIOSYNTHESIS AND HETEROLOGOUS HOST BIAS

This review focuses primarily on the logic and steps needed to establish polyisoprene (i.e., GGPP) formation without providing the detailed reactions that support full isoprenoid compound (i.e., paclitaxel) formation. These reactions are unique for each individual isoprenoid compound, but biosynthetic similarities include multiple reaction steps that modify a polyisoprene frame. These reactions often involve dedicated cytochrome P450 hydroxylases. The paclitaxel compound is a good example, in which 17 steps are predicted to convert taxadiene to the final compound, with the majority of these transformations being catalyzed by P450 enzymes (14, 27).

Establishing full heterologous paclitaxel biosynthesis will require the same dedication that has led to the success of high-level *E. coli* taxadiene production. Specifically, a combination of heterologous and pathway metabolic engineering will be required to first establish paclitaxel intermediate and full compound formation, followed by efforts to maximize specific and volumetric productivities. As before, the transfer of genes from plant sources will require special attention to establish active gene expression, both from a mechanistic standpoint (through gene synthesis and expression technology) and in terms of postexpression functionality (as a by-product of protein localization, folding, and activity). As a key example, in the same study in which substrate and biosynthetic requirements were balanced for taxadiene overproduction, the resulting *E. coli* strain was engineered to generate the subsequent paclitaxel pathway intermediate taxadien-5α-ol, catalyzed by a cytochrome P450 enzyme (2). As for the foreign GGPPS and taxadiene synthase, the P450 gene and that for an accompanying NADPH reductase were synthesized and altered to maximize gene expression, cellular protein localization (through removing unwanted and introducing beneficial localization and fusion sequences), and final enzyme activity. Even so, the 58-mg/liter titer for the taxadien-5α-ol product was significantly below the optimized value obtained for taxadiene. Such a result indicates the need for renewed efforts in heterologous and pathway metabolic engineering to boost yields. Furthermore, this process must take place repeatedly to reach the paclitaxel endpoint or a pathway intermediate capable of being converted semisynthetically to the

TABLE 2 Literature analysis of isoprenoid compounds generated heterologously[a]

Compound	Heterologous host	No. of articles	Yr of first report
Paclitaxel (taxadiene)	*E. coli*	11	2001
	S. cerevisiae	9	2006
Artemisinin (amorphadiene)	*E. coli*	24	2003
	S. cerevisiae	18	2006
Carotenoid (lycopene)	*E. coli*	158	1986
	S. cerevisiae	9	1994
Carotenoid (zeaxanthin)	*E. coli*	64	1990
	S. cerevisiae	2	2009
Carotenoid (astaxanthin)	*E. coli*	35	1995
	S. cerevisiae	4	2009

[a] Results of a search conducted (26 October 2011) using the PubMed database with either "coli" or "cerevisiae" in conjunction with the molecules in parentheses. The numbers include both primary research and review articles and may include articles not directly related to heterologous biosynthesis. However, the results illustrate the differences in compound-specific host choice and the associated volume of research since heterologous biosynthesis was initiated.

final product. An additional complication is that not all of the pathway enzymes have been identified; thus, additional research into pathway elucidation or substitution must occur prior to next-step pathway heterologous production. The positive viewpoint is that there is now precedence to support future attempts at heterologously generating and overproducing downstream paclitaxel pathway intermediates. However, this venture will no doubt test the current limits of metabolic engineering and, most likely, will require advances in current technology to obtain ultimate success. Still, considering the advances that have occurred to this point and the implications in overall success, there is reason for optimism in pursuing such a challenging objective.

Finally, Table 2 presents a literature analysis of the heterologous production of recent and common isoprenoid compounds. Two points can be made about the results. First, there has clearly been significantly more research dedicated to the production of carotenoid compounds. This is due in part to the earlier sequence availability of such pathways, which then spurred the first heterologous attempts (29, 43, 48, 56, 59). Since the first publications on heterologous production, the majority of the remaining studies have been dedicated to metabolic engineering and production optimization. As mentioned above, for the carotenoid compounds, engineering studies are aided greatly by the colorimetric nature of the final compounds, most likely bolstering the associated number of research efforts. There is a drop in research citations with increasing complexity of the carotenoid compound (from lycopene to zeaxanthin to astaxanthin). This again may reflect the order or availability associated with pathway knowledge, but it may also reflect the difficulty associated with introducing downstream steps dedicated to polyisoprene adornment. It should also be stated that the above carotenoid trend holds for *E. coli* but not for *S. cerevisiae*. It is interesting to observe such a disparity between these two technically convenient hosts within the carotenoid family of isoprenoids. Our speculation is that the biosynthetic requirements did not favor the use of either *E. coli* or *S. cerevisiae*, and as a result, the research community then relied on

E. coli, which would generally be considered more technically convenient than *S. cerevisiae*.

Second, the more recent trends associated with paclitaxel and artemisinin intermediates show a more even distribution of the *E. coli* and *S. cerevisiae* heterologous hosts. Unlike the carotenoid compounds, these studies began in earnest only during the last 5 to 10 years. But the more prevalent use of *S. cerevisiae* may point to the additional complexity associated with both final compounds. The yeast cellular background obviously has more commonalities with the plant cell structure than the *E. coli* background. This may greatly facilitate the introduction of the downstream pathways required for final compound formation. Essentially, the nature and complexity of the pathway create a bias which favors the use of a yeast system, and this may reflect the early trends in heterologous host choice. Time will tell if the final artemisinin or paclitaxel (or other complex isoprenoid) compounds can be completely synthesized heterologously and which host will serve this purpose.

ACKNOWLEDGMENTS

REFERENCES

1. **Ajikumar PK, et al.** 2008. Terpenoids: opportunities for biosynthesis of natural product drugs using engineered microorganisms. Mol. Pharm. **5**:167–190.
2. **Ajikumar PK, et al.** 2010. Isoprenoid pathway optimization for paclitaxel precursor overproduction in *Escherichia coli.* Science **330**:70–74.
3. **Alper H, Jin YS, Moxley JF, Stephanopoulos G.** 2005. Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli.* Metab. Eng. **7**:155–164.
4. **Anthony JR, et al.** 2009. Optimization of the mevalonate-based isoprenoid biosynthetic pathway in *Escherichia coli* for production of the anti-malarial drug precursor amorpha-4,11-diene. Metab. Eng. **11**:13–19.
5. **Arbuck SG, Blaylock BA.** 1995. Paclitaxel: clinical results and current issues in development. CRC Press, Boca Raton, FL.
6. **Asai K, et al.** 1994. The identification of *Escherichia coli ispB* (CEL) gene encoding the octaprenyl diphosphate synthase. Biochem. Biophys. Res. Commun. **202**:340–345.
7. **Bloch K.** 1992. Sterol molecule: structure, biosynthesis, and function. Steroids **57**:378–383.
8. **Boghigian BA, Armando J, Salas D, Pfeifer BA.** Computational identification of gene over-expression targets for metabolic engineering of taxadiene production. Appl. Microbiol. Biotechnol., in press.
9. **Boghigian BA, Salas D, Ajikumar PK, Stephanopoulos G, Pfeifer BA.** 2011. Analysis of heterologous taxadiene production in K- and B-derived *Escherichia coli.* Appl. Microbiol. Biotechnol. **93**:1651–1661.
10. **Chang MC, Eachus RA, Trieu W, Ro DK, Keasling JD.** 2007. Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. Nat. Chem. Biol. **3**:274–277.
11. **Chang MC, Keasling JD.** 2006. Production of isoprenoid pharmaceuticals by engineered microbes. Nat. Chem. Biol. **2**:674–681.
12. **Chiang CJ, Chen PT, Chao YP.** 2008. Replicon-free and markerless methods for genomic insertion of DNAs in phage attachment sites and controlled expression of chromosomal genes in *Escherichia coli.* Biotechnol. Bioeng. **101**:985–995.
13. **Choi HS, Lee SY, Kim TY, Woo HM.** 2010. In silico identification of gene amplification targets for improvement of lycopene production. Appl. Environ. Microbiol. **76**:3097–3105.
14. **Croteau R, Ketchum RE, Long RM, Kaspera R, Wildung MR.** 2006. paclitaxel biosynthesis and molecular genetics. Phytochem. Rev. **5**:75–97.
15. **Dejong JM, et al.** 2006. Genetic engineering of taxol biosynthetic genes in *Saccharomyces cerevisiae.* Biotechnol. Bioeng. **93**:212–224.
16. **Engels B, Dahm P, Jennewein S.** 2008. Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards paclitaxel production. Metab. Eng. **10**:201–206.
17. **Frense D.** 2007. Taxanes: perspectives for biotechnological production. Appl. Microbiol. Biotechnol. **73**:1233–1240.
18. **Fujisaki S, Hara H, Nishimura Y, Horiuchi K, Nishino T.** 1990. Cloning and nucleotide sequence of the *ispA* gene responsible for farnesyl diphosphate synthase activity in *Escherichia coli.* J. Biochem. **108**:995–1000.
19. **Gershenzon J, Dudareva N.** 2007. The function of terpene natural products in the natural world. Nat. Chem. Biol. **3**:408–414.
20. **Grawert T, Groll M, Rohdich F, Bacher A, Eisenreich W.** 2011. Biochemistry of the non-mevalonate isoprenoid pathway. Cell. Mol. Life Sci. **68**:3797–3814.
21. **Harada H, Misawa N.** 2009. Novel approaches and achievements in biosynthesis of functional isoprenoids in *Escherichia coli.* Appl. Microbiol. Biotechnol. **84**:1021–1031.
22. **Harada H, et al.** 2009. Efficient synthesis of functional isoprenoids from acetoacetate through metabolic pathway-engineered *Escherichia coli.* Appl. Microbiol. Biotechnol. **81**:915–925.
23. **Hefner J, Ketchum RE, Croteau R.** 1998. Cloning and functional expression of a cDNA encoding geranylgeranyl diphosphate synthase from *Taxus canadensis* and assessment of the role of this prenyltransferase in cells induced for taxol production. Arch. Biochem. Biophys. **360**:62–74.
24. **Holton RA, Biediger RJ, Boatman PD.** 1995. Semisynthesis of taxol and taxotere. CRC Press, Boca Raton, FL.
25. **Horwitz SB.** 1994. How to make taxol from scratch. Nature **367**:593–594.
26. **Huang Q, Roessner CA, Croteau R, Scott AI.** 2001. Engineering *Escherichia coli* for the synthesis of taxadiene, a key intermediate in the biosynthesis of taxol. Bioorg. Med. Chem. **9**:2237–2242.
27. **Jennewein S, Wildung MR, Chau M, Walker K, Croteau R.** 2004. Random sequencing of an induced Taxus cell cDNA library for identification of clones involved in paclitaxel biosynthesis. Proc. Natl. Acad. Sci. U. S. A. **101**:9149–9154.
28. **Kainou T, et al.** 2001. Dimer formation of octaprenyl-diphosphate synthase (IspB) is essential for chain length determination of ubiquinone. J. Biol. Chem. **276**:7876–7883.
29. **Kajiwara S, et al.** 1995. Isolation and functional identification of a novel cDNA for astaxanthin biosynthesis from *Haematococcus pluvialis*, and astaxanthin synthesis in *Escherichia coli.* Plant Mol. Biol. **29**:343–352.
30. **Kirby J, Keasling JD.** 2009. Biosynthesis of plant isoprenoids: perspectives for microbial engineering. Annu. Rev. Plant Biol. **60**:335–355.
31. **Kirby J, Keasling JD.** 2008. Metabolic engineering of microorganisms for isoprenoid production. Nat. Prod. Rep. **25**:656–661.
32. **Kizer L, Pitera DJ, Pfleger BF, Keasling JD.** 2008. Application of functional genomics to pathway optimization for increased isoprenoid production. Appl. Environ. Microbiol. **74**:3229–3241.
33. **Klein-Marcuschamer D, Ajikumar PK, Stephanopoulos G.** 2007. Engineering microbial cell factories for biosynthesis of isoprenoid molecules: beyond lycopene. Trends Biotechnol. **25**:417–424.
34. **Kuzma J, Nemecek-Marshall M, Pollock WH, Fall R.** 1995. Bacteria produce the volatile hydrocarbon isoprene. Curr. Microbiol. **30**:97–103.
35. **Kuzuyama T, Seto H.** 2003. Diversity of the biosynthesis of the isoprene units. Nat. Prod. Rep. **20**:171–183.
36. **Lee SY.** 1996. High cell-density culture of *Escherichia coli.* Trends Biotechnol. **14**:98–105.
37. **Leonard E, et al.** 2010. Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. Proc. Natl. Acad. Sci. U. S. A. **107**:13654–13659.
38. **Lindahl AL, et al.** 2006. Production of the artemisinin precursor amorpha-4,11-diene by engineered *Saccharomyces cerevisiae.* Biotechnol. Lett. **28**:571–580.
39. **Martin VJ, Pitera DJ, Withers ST, Newman JD, Keasling JD.** 2003. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. Nat. Biotechnol. **21**:796–802.
40. **Maury J, Asadollahi MA, Moller K, Clark A, Nielsen J.** 2005. Microbial isoprenoid production: an example of green chemistry through metabolic engineering. Adv. Biochem. Eng. Biotechnol. **100**:19–51.
41. **McCaskill D, Croteau R.** 1997. Prospects for the bioengineering of isoprenoid biosynthesis. Adv. Biochem. Eng. Biotechnol. **55**:107–146.
42. **McGarvey DJ, Croteau R.** 1995. Terpenoid metabolism. Plant Cell **7**:1015–1026.
43. **Misawa N, et al.** 1990. Elucidation of the *Erwinia uredovora* carotenoid biosynthetic pathway by functional analysis of gene products expressed in *Escherichia coli.* J. Bacteriol. **172**:6704–6712.
44. **Miziorko HM.** 2011. Enzymes of the mevalonate pathway of isoprenoid biosynthesis. Arch. Biochem. Biophys. **505**:131–143.

45. **Morrone D, et al.** 2010. Increasing diterpene yield with a modular metabolic engineering system in *E. coli*: comparison of MEV and MEP isoprenoid precursor pathway engineering. Appl. Microbiol. Biotechnol. **85:** 1893–1906.

46. **Muntendam R, Melillo E, Ryden A, Kayser O.** 2009. Perspectives and limits of engineering the isoprenoid metabolism in heterologous hosts. Appl. Microbiol. Biotechnol. **84:**1003–1019.

47. **Okamura E, Tomita T, Sawa R, Nishiyama M, Kuzuyama T.** 2010. Unprecedented acetoacetyl-coenzyme A synthesizing enzyme of the thiolase superfamily involved in the mevalonate pathway. Proc. Natl. Acad. Sci. U. S. A. **107:**11265–11270.

48. **Perry KL, Simonitch TA, Harrison-Lavoie KJ, Liu ST.** 1986. Cloning and regulation of *Erwinia herbicola* pigment genes. J. Bacteriol. **168:**607–612.

49. **Redding-Johanson AM, et al.** 2011. Targeted proteomics for metabolic pathway optimization: application to terpene production. Metab. Eng. **13:**194–203.

50. **Reiling KK, et al.** 2004. Mono and diterpene production in *Escherichia coli*. Biotechnol. Bioeng. **87:**200–212.

51. **Ro DK, et al.** 2008. Induction of multiple pleiotropic drug resistance genes in yeast engineered to produce an increased level of anti-malarial drug precursor, artemisinic acid. BMC Biotechnol. **8:**83.

52. **Ro DK, et al.** 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature **440:**940–943.

53. **Roberts SC.** 2007. Production and engineering of terpenoids in plant cell culture. Nat. Chem. Biol. **3:**387–395.

54. **Rodriguez-Villalon A, Perez-Gil J, Rodriguez-Concepcion M.** 2008. Carotenoid accumulation in bacteria with enhanced supply of isoprenoid precursors by upregulation of exogenous or endogenous pathways. J. Biotechnol. **135:**78–84.

55. **Rohmer M.** 1999. The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. Nat. Prod. Rep. **16:**565–574.

56. **Ruther A, Misawa N, Boger P, Sandmann G.** 1997. Production of zeaxanthin in *Escherichia coli* transformed with different carotenogenic plasmids. Appl. Microbiol. Biotechnol. **48:**162–167.

57. **Ruzicka L.** 1953. The isoprene rule and the biogenesis of terpenic compounds. Experientia **9:**357–367.

58. **Sacchettini JC, Poulter CD.** 1997. Creating isoprenoid diversity. Science **277:**1788–1789.

59. **Sandmann G, Woods WS, Tuveson RW.** 1990. Identification of carotenoids in *Erwinia herbicola* and in a transformed *Escherichia coli* strain. FEMS Microbiol. Lett. **59:**77–82.

60. **Shiloach J, Fass R.** 2005. Growing *E. coli* to high cell density—a historical perspective on method development. Biotechnol. Adv. **23:**345–357.

61. **Suthers PF, et al.** 2007. Metabolic flux elucidation for large-scale models using ^{13}C labeled isotopes. Metab. Eng. **9:**387–405.

62. **Suthers PF, Chang YJ, Maranas CD.** 2010. Improved computational performance of MFA using elementary metabolite units and flux coupling. Metab. Eng. **12:**123–128.

63. **Tsuruta H, et al.** 2009. High-level production of amorpha-4,11-diene, a precursor of the antimalarial agent artemisinin, in *Escherichia coli*. PLoS One **4:**e4489.

64. **Verwaal R, et al.** 2007. High-level production of beta-carotene in *Saccharomyces cerevisiae* by successive transformation with carotenogenic genes from *Xanthophyllomyces dendrorhous*. Appl. Environ. Microbiol. **73:** 4342–4350.

65. **Wang C, Oh MK, Liao JC.** 2000. Directed evolution of metabolically engineered *Escherichia coli* for carotenoid production. Biotechnol. Prog. **16:**922–926.

66. **Wang CW, Oh MK, Liao JC.** 1999. Engineered isoprenoid pathway enhances astaxanthin production in *Escherichia coli*. Biotechnol. Bioeng. **62:**235–241.

67. **Withers ST, Keasling JD.** 2007. Biosynthesis and engineering of isoprenoid small molecules. Appl. Microbiol. Biotechnol. **73:**980–990.

68. **Xue J, Ahring BK.** 2011. Enhancing isoprene production by genetic modification of the 1-deoxy-D-xylulose-5-phosphate pathway in *Bacillus subtilis*. Appl. Environ. Microbiol. **77:**2399–2405.

69. **Yamane T, Shimizu S.** 1984. Fed-batch techniques in microbial processes. Adv. Biochem. Eng. **30:**147–194.

70. **Yamano S, Ishii T, Nakagawa M, Ikenaga H, Misawa N.** 1994. Metabolic engineering for production of beta-carotene and lycopene in *Saccharomyces cerevisiae*. Biosci. Biotechnol. Biochem. **58:**1112–1114.

71. **Yoon SH, et al.** 2006. Enhanced lycopene production in *Escherichia coli* engineered to synthesize isopentenyl diphosphate and dimethylallyl diphosphate from mevalonate. Biotechnol. Bioeng. **94:**1025–1032.

72. **Yuan LZ, Rouviere PE, Larossa RA, Suh W.** 2006. Chromosomal promoter replacement of the isoprenoid pathway for enhancing carotenoid production in *E. coli*. Metab. Eng. **8:**79–90.

73. **Zhang H, Boghigian BA, Armando J, Pfeifer BA.** 2011. Methods and options for the heterologous production of complex natural products. Nat. Prod Rep. **28:**125–151.

74. **Zhao Y, et al.** 2011. Biosynthesis of isoprene in *Escherichia coli* via methylerythritol phosphate (MEP) pathway. Appl. Microbiol. Biotechnol. **90:**1915–1922.