Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development

Chris J. Paddon[1] *and Jay D. Keasling*[2,3,4,5]

Abstract | Recent developments in synthetic biology, combined with continued progress in systems biology and metabolic engineering, have enabled the engineering of microorganisms to produce heterologous molecules in a manner that was previously unfeasible. The successful synthesis and recent entry of semi-synthetic artemisinin into commercial production is the first demonstration of the potential of synthetic biology for the development and production of pharmaceutical agents. In this Review, we describe the metabolic engineering and synthetic biology approaches that were used to develop this important antimalarial drug precursor. This not only demonstrates the incredible potential of the available technologies but also illuminates how lessons learned from this work could be applied to the production of other pharmaceutical agents.

[1]*Amyris, Inc., 5885 Hollis Street, Suite 100, Emeryville, California 94608, USA.*
[2]*Department of Chemical and Biomolecular Engineering, University of California, Berkeley, California 94720, USA.*
[3]*Department of Bioengineering, University of California, Berkeley, California 94720, USA.*
[4]*Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.*
[5]*Joint BioEnergy Institute, 5885 Hollis Street, Emeryville, California 94608, USA.*
e-mails: Paddon@amyris.com; keasling@berkeley.edu
doi:10.1038/nrmicro3240
Published online 1 April 2014

Many drugs are derived from, or inspired by, natural products. The most famous example is probably aspirin[1,2], which is a synthetic drug that is derived from salicylic acid — a compound that was originally isolated from willow bark[3] and that has a long history of use in folk medicine. The chemical structure of aspirin is relatively simple, and it was first synthesized in 1897 by chemists at the Bayer Company[1]. However, it is not possible to chemically synthesize all natural products in an economical manner, despite the advances in synthetic organic chemistry during the past century; for example, the potent antimalarial drug precursor artemisinin cannot be economically produced by chemical synthesis alone[4].

Similarly to salicylic acid, artemisinin is naturally produced by a plant, *Artemisia annua*, which has a long history of use in traditional Chinese medicine. The antimalarial properties of *A. annua* extracts were re-discovered by Chinese scientists in the 1970s, and the active compound was identified as artemisinin; this was followed shortly afterwards by the elucidation of its chemical structure[5]. Although artemisinin derivatives were designated first-line antimalarial drugs by the WHO in 2002 (REF. 6), artemisinin availability has been difficult to predict owing to the 18 month lag between initial requisition and eventual supply, which is dependent on the weather and overall harvest. As a result, both the supply and price of the drug have fluctuated substantially over the past decade[7]. Thus, the Semi-synthetic Artemisinin Project was initiated with the aim of providing an alternative source of artemisinin to stabilize the supply and price of antimalarial combination therapies (ACTs) for patients in the developing world[8]. The project involved the metabolic engineering of microorganisms to produce a late-stage chemical precursor of artemisinin (artemisinic acid), followed by synthetic organic chemistry to produce semi-synthetic artemisinin — a product that is indistinguishable from plant-derived artemisinin. The term 'semi-synthetic' indicates that only the final few synthetic steps are accomplished by organic chemistry, whereas the earlier steps are achieved by biological synthesis. The success of the project was made possible by advances both in metabolic engineering and in synthetic biology (BOX 1). During the initial stages of development, the bacterium *Escherichia coli* was used as the producing organism, but as this bacterium is limited in its ability to produce high concentrations of artemisinic acid, *Saccharomyces cerevisiae* (also known as baker's yeast) was later used. The industrial production of semi-synthetic artemisinin by Sanofi has now commenced and the product will be used to supplement the world supply of artemisinin.

In this Review, we describe the development of semi-synthetic artemisinin and highlight the use of synthetic biology and metabolic engineering approaches for this purpose. We consider the burden of malaria in the developing world, the discovery of the antimalarial properties

Box 1 | **Basic principles of metabolic engineering and synthetic biology**

The crux of metabolic engineering is the use of recombinant DNA methods to modify metabolic networks to either increase the production of native metabolites[87,88] or to re-route metabolism for the synthesis of new chemical products that are not naturally produced by the host organism, such as the microbial production of 1,3-propanediol, which is a key component of polymers[89]. Metabolic engineering is akin to rerouting the pipes or unit operations in a chemical factory to produce more of a product that it already produces or to produce an entirely new product. Although it is possible to reconfigure the unit operations of a chemical factory to make it more productive, a computerized control system is necessary to ensure that all unit operations in the factory achieve optimal performance and that each unit is correctly configured to interact with the next.

One of the major challenges in translating metabolic engineering projects from the bench to large-scale production is in the control of the primary metabolic pathways that are directly involved in the production of the desired product (and also all of the other metabolic reactions inside the cell that may affect the primary metabolic pathway). Furthermore, although it might be possible to engineer a microbial cell to produce small amounts of the desired compound, producing that compound in the titres and yields that are necessary for economic viability can be challenging. In the chemical factory example, unit operations can be modelled using sophisticated design software and process control systems, and strategies can be developed to maximize profit, minimize inventories and maximize safety. Analogous systems for designing and building biology are in their infancy[90–93] and belong to the field of synthetic biology. The overall aim of synthetic biology is to simplify biological engineering[94] by applying engineering principles and designs — which emanate from electronic and computer engineering — to biology[95–100]. The development of solid-phase DNA synthesis and assembly of oligonucleotides[101] has enabled the construction of synthetic genes and control elements, such as biological switches[99] and clocks[100]. Recent efforts to standardize and characterize biological components will make the construction of larger genetic circuits easier and more reliable[94,95,102,103,104]. Owing to the inherent complexity of biological systems, it is difficult to accurately design and control them, which has necessitated implementation of the design–build–test–analyse (DBTA) cycle for optimizing microbial production systems[87]. This design strategy can be likened to building millions of variants of a chemical factory, selecting or screening for control system variants that yield the most product and discarding all but one or two of the most productive designs. The knowledge that is gained from the DBTA cycle can be used to improve future designs using computer-aided design software. The better the software we have for building biology (including the ability to predict all interactions in the cell) and the more well-characterized the biological components are, the more the requirement for the DBTA cycle will diminish and biological engineering will move closer to other engineering disciplines.

of artemisinin and its biosynthesis in *A. annua*. This is followed by a discussion of the Semi-synthetic Artemisinin Project, which is divided into three stages: amorphadiene production in *E. coli*, artemisinic acid biosynthesis and the switch to yeast-based production, and the chemical conversion of artemisinic acid to artemisinin. Finally, we provide a brief overview of the main lessons to be learned from the success and shortcomings of the approaches that were used, which should aid the development of other pharmaceutical products.

Malaria and the discovery of artemisinin

Malaria is predominantly a disease of the developing world. Although malaria mortality rates decreased by 26% between 2000 and 2010, an estimated 219 million cases of malaria were reported and 660,000 malaria-associated deaths occurred in 2010. About 90% of all malaria deaths occur in Africa, predominantly among children under 5 years of age[9]. The disease is caused by *Plasmodium* spp. parasites, primarily *Plasmodium falciparum* and *Plasmodium vivax*, but *P. falciparum* causes the most virulent form of malaria. As *P. falciparum* resistance to older, widely used and relatively inexpensive antimalarial drugs (such as chloroquine and sulfadoxine–pyrimethamine)[6] has become widespread, the development of other antimalarial agents is a priority.

The pharmaceutical agents that have succeeded the older antimalarial drugs are artemisinin derivatives. Artemisinin is derived from the sweet wormwood plant *A. annua* (known historically as quinghao), extracts of which were first used in traditional Chinese medicine almost 1,800 years ago[10]. The antimalarial properties of *A. annua* extracts were re-discovered by Chinese scientists during the Cultural Revolution as part of a research programme (known as Project 523) to find a drug that was effective against chloroquine-resistant *P. falciparum*. Within a few years, the active compound had been identified as artemisinin, and its structure was subsequently elucidated[11,12].

Following its recognition by the WHO as an effective antimalarial drug [13], artemisinin derivatives are now primary components of ACTs, which are recommended by the WHO for the treatment of uncomplicated malaria caused by *P. falciparum*, and in some cases, by *P. vivax*[6]. ACTs are composed of an artemisinin derivative (that has improved pharmacokinetic properties compared with artemisinin) in combination with another drug that has an independent mode of action and a longer duration of action. The mode of action of artemisinin is not fully understood, but it is thought that artemisinin-derived free radicals probably damage cellular targets via alkylation[14]. Several different combination drugs are available — for example, the combination of artemether and lumefantrine[15], which is marketed by Novartis. The rationale for combining an artemisinin derivative with another antimalarial drug is twofold; first, the combination is more effective at reducing treatment times and the incidence of recrudescence, and second, the potential for artemisinin resistance to emerge is reduced, as parasites that develop resistance to artemisinin will be killed by the companion drug and the probability of the parasite developing resistance to both drugs is quite low[6]. Although reduced parasite clearance rates have recently been reported in Southeast Asia[6,16–18] (owing to ACT resistance), resistance has not yet been widely observed in Africa, and ACTs remain the mainstay of antimalarial therapies in the developing world. Recent studies of genetic loci that are associated with the delayed clearance of *P. falciparum* in Southeast Asia[19,20] have contributed to the discovery of a molecular marker for artemisinin resistance (known as K13-propellor alleles) which will be useful for surveillance efforts to contain the spread of artemisinin-resistant parasites[21].

Following the recommendation by the WHO in 2002 that ACTs be used as the first-line treatment for uncomplicated malaria[13], the number of treatments that were procured by the public sector increased from 11 million in 2005 to an estimated 287 million in 2010. The price and availability of artemisinin have fluctuated substantially during this time period, ranging from shortfalls as a result of inconsistencies in the agricultural supply (for example, unfavourable weather reduced crop yield in 2004) to

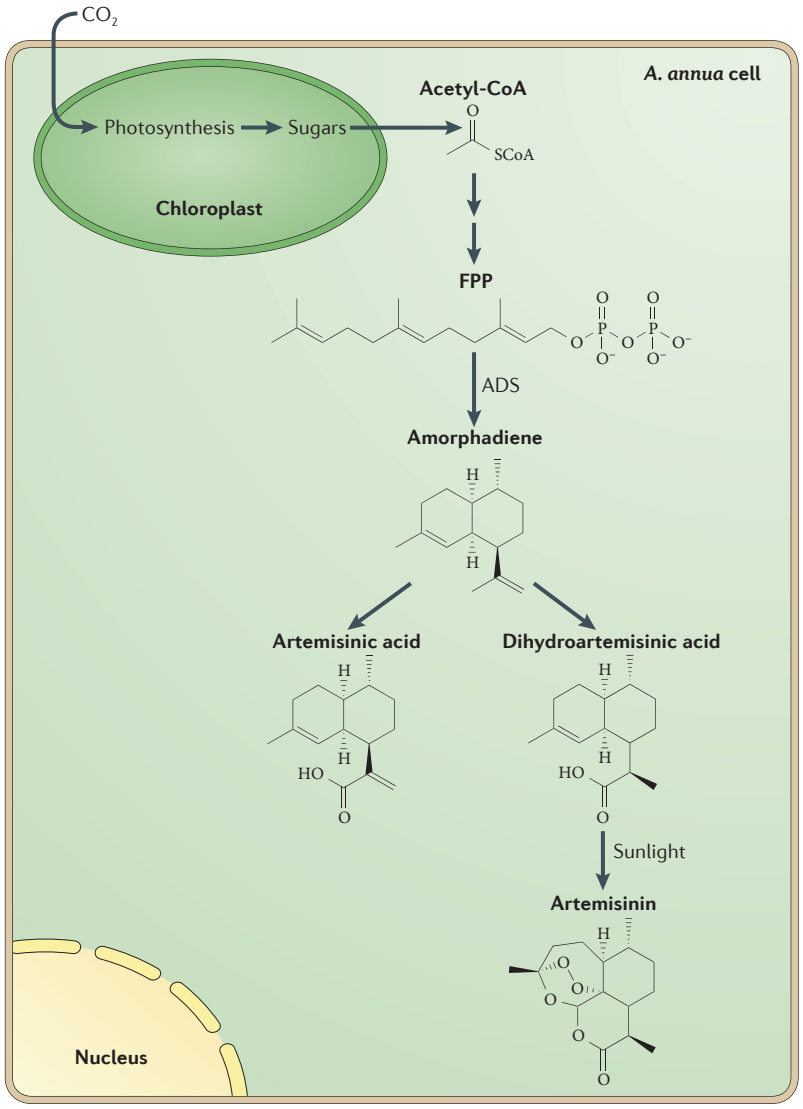


Figure 1 | **Artemisinin biosynthesis pathway in the plant *Artemisia annua*.** Sugars that are produced by photosynthesis in plant chloroplasts are converted to acetyl-CoA in the cytosol. The two-carbon acetyl-CoA feeds into the cytosolic mevalonate pathway, which produces farnesyl diphosphate (FPP). FPP is converted to the 15-carbon sesquiterpene amorphadiene by the enzyme amorphadiene synthase (ADS). Amorphadiene is subsequently enzymatically oxidized to either artemisinic acid or dihydroartemisinic acid; different chemotypes of *A. annua* produce different ratios of these products[30,118]. Dihydroartemisinic acid is the precursor of artemisinin and is thought to be converted to artemisinin in a non-enzymatic photochemical reaction that is stimulated by sunlight[28].

oversupply in 2007, with concomitant price reductions[22]. During times of oversupply, many growers of *A. annua* switched to other crops that have a more dependable price outlook, resulting in shortages in subsequent years. Since 2007, the price of plant-derived artemisinin has continued to fluctuate[23], varying from a recent high of over US$600 per kg in early 2012 to under $300 per kg in late 2013 (REF. 24), and future pricing is unpredictable. Thus, the Semi-synthetic Artemisinin Project was borne out of an urgent need and desire to develop an assured alternative source of affordable artemisinin to stabilize both the cost and the supply of artemisinin derivatives[8].

A major factor to be considered for the development of a drug that is destined for use in the developing world is the price-point requirement. For artemisinin, which is particularly central to healthcare in the developing world, the price point must be comparable to that of naturally produced artemisinin if it is to supplement the plant supply. In addition, the healthcare expenditures of countries in which the drug is in highest demand are considered[25,26]. Many malaria-endemic developing nations have a per capita healthcare expenditure of under $100. By contrast, developed nations have a much higher per capita healthcare budget; for example, the UK has a spend of $3609 and the USA of $8608 (REFS 25,26). Thus, the required price point for semi-synthetic artemisinin was achieved by considering both of these fundamental factors to ensure that it is accessible to those populations that have the greatest demand for the drug.

The Semi-synthetic Artemisinin Project

The project was funded by a grant from the Bill and Melinda Gates Foundation in 2004, and led to a partnership between the University of California (Berkeley, USA), Amyris Inc. (previously known as Amyris Biotechnologies) and the Institute for OneWorld Health (which is a non-profit pharmaceutical company that is now known as PATH Drug Solutions; see further information). The success of the semi-synthetic production of artemisinin is based on a thorough understanding of its production in *A. annua*[27,28] (FIG. 1), the development of rapid methods of microbial genetic manipulation (using metabolic engineering), the ability to rapidly synthesize genes (using synthetic biology) and the application of pharmaceutical chemistry.

Artemisinin is an isoprenoid (also known as a sesquiterpene or simply terpene) molecule that contains 15 carbon atoms (BOX 2) and a lactone endoperoxide[4] (FIG. 1). *A. annua* uses sugars that are generated by photosynthesis to produce the two-carbon intermediate acetyl-CoA, which then enters the mevalonate biosynthetic pathway to produce the 15-carbon intermediate farnesyl diphosphate (FPP). The first committed step of artemisinin biosynthesis is the conversion of FPP to amorphadiene (which is a 15-carbon isoprenoid hydrocarbon) by the terpene synthase enzyme amorphadiene synthase (ADS)[29,30]. At the commencement of the project, it was known that *A. annua* accumulates two oxidation products that were derived from amorphadiene: artemisinic acid and dihydroartemisinic acid[31] (FIG. 1). Dihydroartemisinic acid is the direct precursor of artemisinin and its conversion is thought to occur spontaneously in the presence of sunlight[27,28].

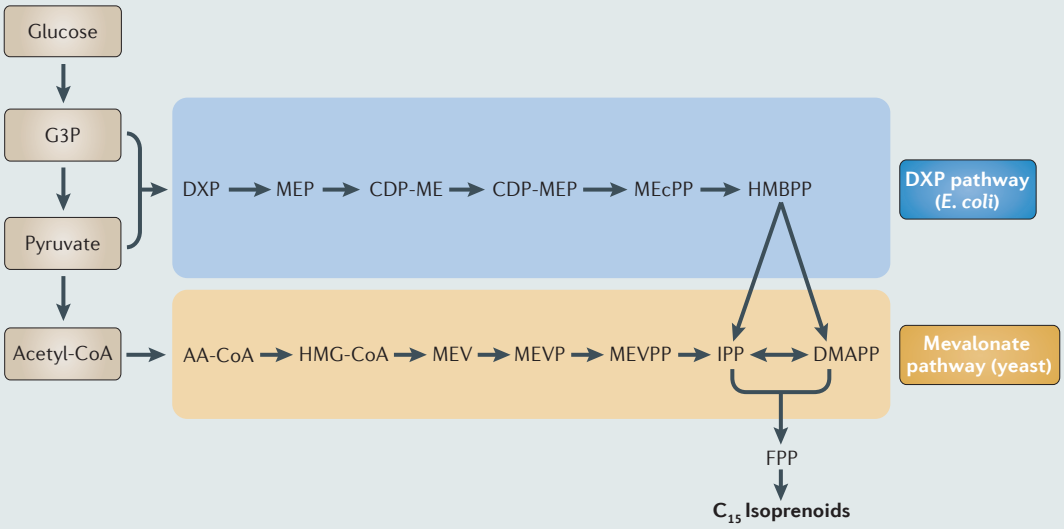
The overall aim of the Semi-synthetic Artemisinin Project was to engineer a microorganism to produce an artemisinin precursor at high titres, rates and yields, followed by chemical conversion to artemisinin (FIG. 2). The decision to produce artemisinic acid rather than dihydroartemisinic acid was made because the enzyme that is responsible for the production of dihydroartemisinic acid was unknown at the time (it was subsequently shown to be DBR2 (REF. 32)), and the chemical conversion

Box 2 | **Isoprenoid production in yeast and *E. coli***

Isoprenoids are a large family of natural products with more than 55,000 members[105]. Family members vary in size, structure and function but they are all derived from two five-carbon precursors: isopentenyl diphosphate (IPP) and its isomer, dimethylallyl diphosphate (DMAPP). Similarly to all eukaryotes, yeast uses the classical mevalonate pathway for the synthesis of isoprenoids[106]. The pathway commences with acetoacetyl-CoA (AA-CoA) that has been produced in the cytosol (not mitochondrial acetyl-CoA)[107], and proceeds through five enzymatic steps via the intermediate mevalonate (MEV) to IPP. IPP is isomerized to DMAPP by IPP isomerase, and the two molecules are combined to form larger isoprenoids, such as farnesyl diphosphate (FPP) (see the figure).

During the 1990s, an alternative pathway for IPP synthesis, known as the 1-deoxy-D-xylulose 5-phosphate (DXP) pathway (also known as the 2-*C*-methyl-D-erythritol phosphate (MEP) pathway, or simply as the non-mevalonate pathway), was discovered [108]. This pathway is used for the production of isoprenoids in *Escherichia coli* and several other bacteria[109]. Plants possess both pathways; they have a classical mevalonate pathway in the cytosol for the production of sterols and sesquiterpenes and a plastid-borne DXP pathway for the biosynthesis of plastidic isoprenoids such as isoprene, monoterpenes and carotenoids[110]. The DXP pathway commences with the glycolytic intermediates glyceraldehyde-3-phosphate (G3P) and pyruvate and proceeds through seven enzymatic steps to produce both IPP and DMAPP[33] (see the figure).

Several attempts have been made to engineer both the mevalonate and DXP pathways for high-level production of terpenoids, such as the artemisinin intermediate amorphadiene. Thus far, the mevalonate pathway has been more successful, and the highest reported titres from both *E. coli*[44] (25 g per L amorphadiene) and *S. cerevisiae* (40 g per L amorphadiene)[54] were obtained by overexpression of this pathway. Heterologous production of isoprenoids by the DXP pathway to increase flux through the pathway in *E. coli* has also been achieved using a combinatorial approach, resulting in the production of 1 g per L taxol, which is the first committed precursor of the anticancer drug taxadiene[64]. An attempt to functionally express the *E. coli* DXP pathway in yeast was unsuccessful owing to the lack of activity of the final two enzymes[111]; this is probably due to the fact that these enzymes are iron–sulphur (FeS) enzymes and yeast has a different cytosolic FeS-assembly machinery than bacteria[112]. CDP-ME, methylerythritol cytidyl diphosphate; CDP-MEP, 4-diphosphocytidyl-2-*C*-methyl-D-erythritol-2-phosphate; HMBPP, 4-hydroxy-3-methyl-butenyl-1-diphosphate; HMG-CoA, 3-hydroxy-3-methyl-glutaryl-CoA; MEcPP, 2-*C*-methyl-D-erythritol-2,4-cyclodiphosphate; MEVP, phosphomevalonate; MEVPP, mevalonate diphosphate.



of artemisinic acid to dihydroartemisinic acid is relatively simple and industrially feasible. After synthesis and secretion into the fermentation broth, artemisinic acid is extracted from the fermentation broth and then chemically converted to artemisinin (FIG. 2). As originally envisaged, the project aimed to produce artemisinic acid in *E. coli* by fermentation and then use synthetic organic chemistry to produce artemisinin itself[8]; however, difficulties that were encountered in the oxidation of amorphadiene to artemisinic acid in engineered *E. coli* led to the use of *S. cerevisiae* for the production of artemisinic acid (FIG. 2). The engineered *S. cerevisiae* strain was capable of producing 40 g per L amorphadiene, which was substantially higher than the 25 g per L yield that was possible with *E. coli*. This switch in chassis organism was important as the ability to produce a high concentration of amorphadiene is a prerequisite for the production of high concentrations of artemisinic acid.

Chassis organism
The host microorganism that is used for the production of a desired product; it is typically subjected to genetic manipulation using the tools of synthetic biology.

Stage 1: amorphadiene synthesis in *E. coli*

Isoprenoids are produced in nature by two different pathways: the mevalonate pathway (in eukaryotes and some prokaryotes) and the more recently discovered 1-deoxy-D-xylulose 5-phosphate (DXP) pathway (in bacteria and plant chloroplasts)[33] (BOX 2). To increase the production of heterologous isoprenoids in *E. coli* (C10 and C20 isoprenoids in this case), the rate-limiting enzymes of the native DXP pathway were overexpressed

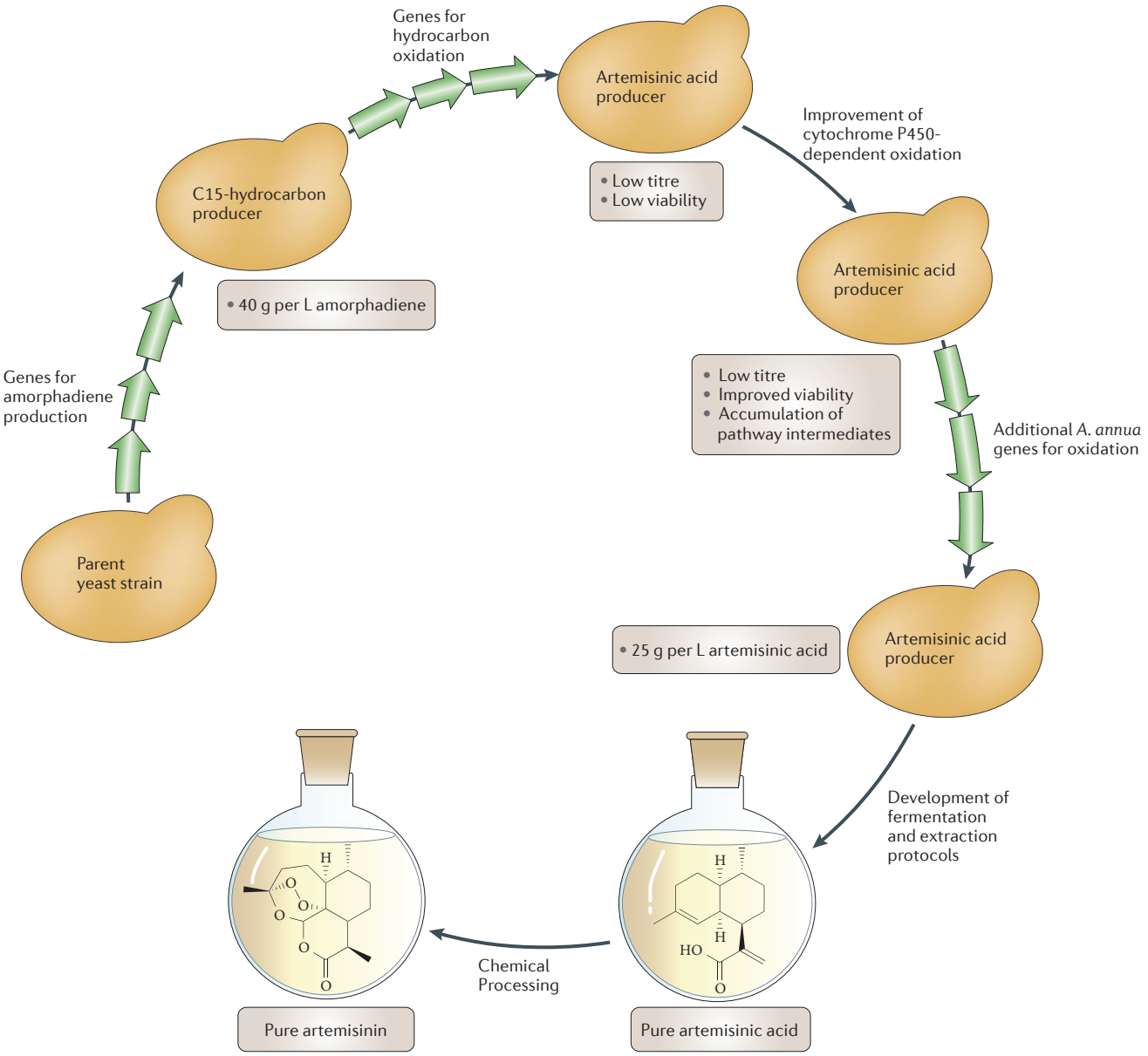
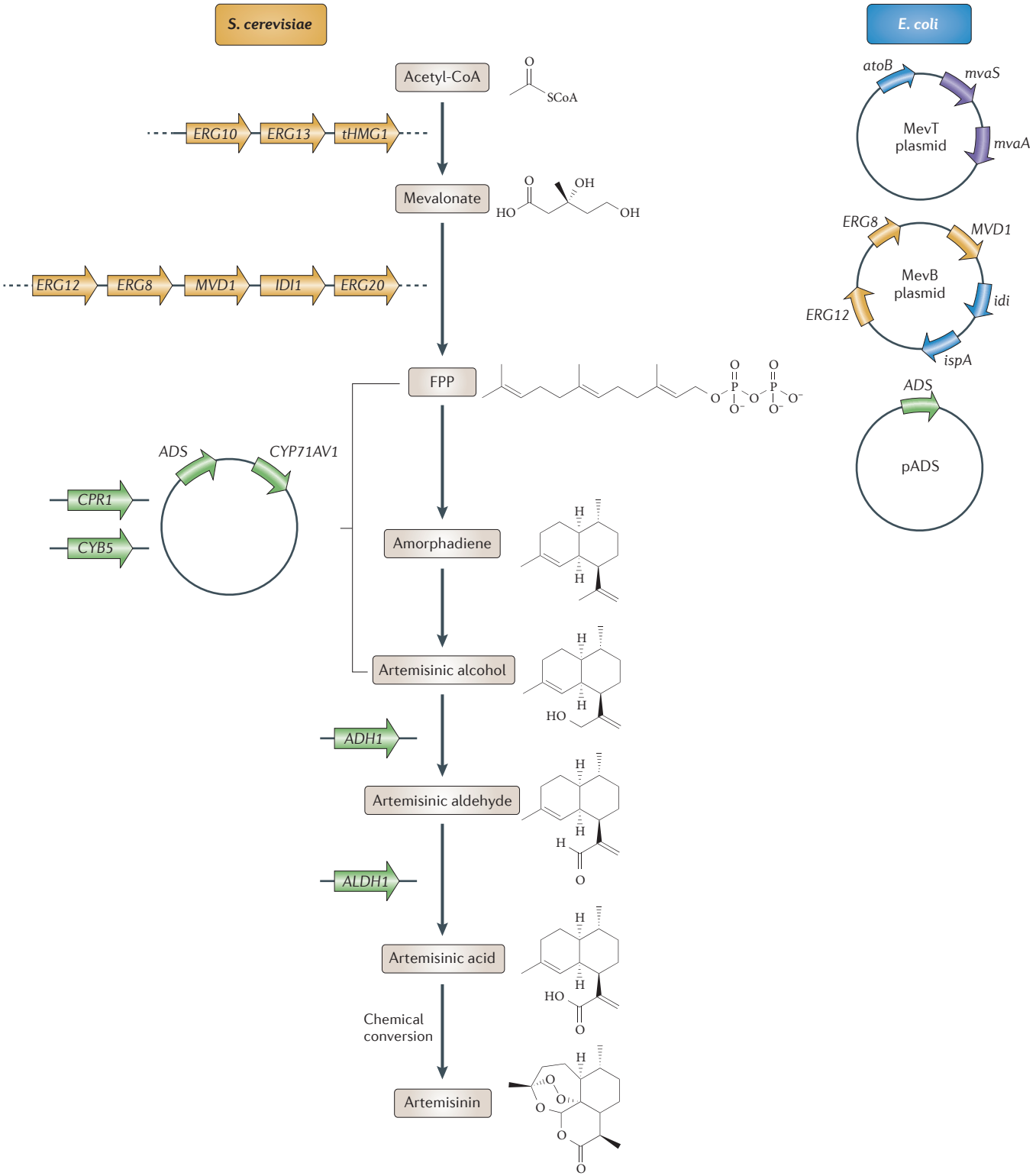


Figure 2 | **The main stages involved in the synthesis of semi-synthetic artemisinin.** The initial stage involved the engineering of the parent *Saccharomyces cerevisiae* strain to produce >25 g per L amorphadiene by overexpression of nine genes of the mevalonate pathway and expression of the *Artemisia annua* amorphadiene synthase. Subsequent engineering steps, involving expression of the *A. annua* cytochrome P450 enzyme, CYP71AV1, its cognate reductase CPR1, cytochrome b_5 and two dehydrogenases enabled the oxidation of amorphadiene to artemisinic acid, which was extracted from fermentation broth and chemically converted to artemisinin. The maximum titre achieved using this procedure was 25 g per L artemisinic acid. Figure is modified, with permission, from REF. 55 © Macmillan Publishers Ltd. All rights reserved.

to increase flux through the pathway, and the genes that encode isoprenoid synthases were also overexpressed. However, this strategy had only limited success and produced terpenes in the low mg per L range[34]. Although an increase in isoprenoid production was observed, the low titres suggested that production was limited by control mechanisms that inhibited synthesis in the native host. However, a dramatic increase in production was observed when the yeast mevalonate pathway was heterologously expressed in *E. coli* along with a synthetic version of the *A. annua* amorphadiene synthase (ADS)[35], which is required for the conversion of endogenously produced FPP to amorphadiene (FIG. 1; FIG. 3).

An earlier attempt to produce sesquiterpenes in *E. coli* using native plant synthases had only succeeded in producing a low titre of these products, which was assumed to be due to poor expression of the plant synthases in *E. coli*[36]. Thus, a synthetic version of ADS that was codon-optimized for *E. coli* was engineered and proved to be successful, which is one of the earliest examples of a synthetic biology application. The mevalonate pathway was expressed in *E. coli* on two plasmids: one encoding



the *mevT* operon (known as the 'top pathway'), which comprises three genes (*atoB*, *ERG13* and *tHMG1*) that are needed for the conversion of acetyl-CoA to mevalonate, and the second plasmid encoding the *mevB* operon (known as the 'bottom pathway'), which comprises five genes (*idi*, *ispA*, *MVD1*, *ERG8* and *ERG12*) for the conversion of mevalonate to FPP (FIG. 3). Cultures of *E. coli* strains that expressed the MevT and MevB plasmids and the codon-optimized ADS produced 112 mg per L amorphadiene. Optimization of the fermentation conditions further increased production to 0.5 g per L (REF. 37), although this was still 50-fold lower than the 25 g per L estimate concentration that is required for the economical production of semi-synthetic artemisinin.

◀ Figure 3 | **Engineering approaches required for pathway construction and semi-synthetic artemisinin production.** *Escherichia coli* was initially used as the chassis organism for the optimization of the mevalonate expression pathway, but steps subsequent to amorphadiene synthesis were difficult to recapitulate in *E. coli*, and *Saccharomyces cerevisiae* was therefore used as the chassis organism for engineering the entire synthesis pathway. The mevalonate pathway was heterologously expressed in *E. coli* on two plasmids (MevB and MevT), along with a plasmid encoding the *Artemisia annua* amorphadiene synthase (ADS), which enabled the conversion of acetyl-CoA to amorphadiene. The yeast strain was engineered to overexpress the mevalonate pathway, and all genes were integrated into the genome, with the exception of *ADS* and *CYP71AV1*, which were plasmid-borne, and were required for the conversion of amorphadeine to artemisinic alcohol. Chromosomal insertion of two more genes (*ADH1* and *ALDH1*) was required for conversion to artemisinic acid. The final step of the pathway, which is the conversion of artemisinic acid to artemisinin, required chemical conversion. The genes expressed encode the following enzymes: *ADH1*, artemisinic alcohol dehydrogenase; *ADS*, amorphadiene synthase; *ALDH1*, artemisinic aldehyde dehydrogenase; *atoB and ERG10*, Acetoacetyl-CoA thiolase; *CPR1*, cytochrome P450 reductase; CYB5, cytochrome b_5; *CYP71AV1*, cytochrome P450 enzyme that converts amorphadiene to artemisinic alcohol; *ERG8*, phosphomevalonate kinase; *ERG12*, mevalonate kinase; *ERG13* and *mvaS*, HMG-CoA synthase; *idi* and *IDI1*, isopentenyl diphosphate isomerase; *ispA* and *ERG20*, farnesy diphosphate (FPP) synthase; *mvaA*, HMG-CoA reductase; *MVD1*, mevalonate diphosphate decarboxylase; *tHMG1*, truncated HMG-CoA reductase (tHMGR). Genes coloured blue are derived from *E. coli*, yellow genes are derived from *S. cerevisiae*, purple genes are derived from *Staphylococcus aureus* and green genes are derived from *A. annua*. All *A. annua* genes were synthetic and codon-optimized for the chassis organism.

Shortcomings of the* E. coli *strain. Despite the production of amorphadiene, there were obvious shortcomings with the strain; the most notable being that increased expression of the *mevT* operon inhibited growth[38]. To eliminate this problem, a synthetic combinatorial approach was used[39]. It was reasoned that expression of the three enzymes encoded by the *mevT* operon was imbalanced, leading to the accumulation of an unknown intermediate that inhibited growth. In the absence of a priori knowledge of optimal enzyme expression levels, an approach was developed using libraries of tunable intergenic regions (TIGRs) followed by screening for increased mevalonate production using a bioassay[39,40]. This approach revealed that reduced expression of 3-hydroxy-3-methyl-glutaryl-coenzyme A synthase (HMGS) and a truncated HMG-CoA reductase (tHMGR; BOX 2) led to a sevenfold increase in mevalonate production[39]. Systematic analysis of gene expression and metabolite build-up subsequently showed that accumulation of the pathway intermediate HMG-CoA caused the observed growth inhibition and that an increase in tHMGR activity (as a result of overexpressing the enzyme) resulted in higher titres of mevalonate[38]. Finally, transcriptomic and metabolomic analyses of strains that accumulated HMG-CoA showed that build-up of HMG-CoA inhibits fatty acid biosynthesis. This was prevented by the addition of palmitic acid to the medium[41], which illustrates the contribution of systems biology to strain engineering.

Considering that the expression of yeast HMGR (encoded by *tHMG1*), and possibly of HMG-CoA synthase (encoded by *ERG13*), was imbalanced or limiting, further effort was directed towards improving the production of amorphadiene by concentrating on this node. An earlier study showed that high titres of mevalonate could be produced in *E. coli* by expressing a type II HMGR from *Enterococcus faecalis*[42]; thus, the expression of type II HMGRs[43] was investigated as part of an integrated strategy that also included the development of the fermentation step[44]. Replacement of the yeast HMGR (encoded by *tHMG1*) with type II HMGR from either *E. faecalis* or *Staphylococcus aureus* was tested, and the *S. aureus* HMGR (encoded by *mvaA*) was found to produce the highest concentrations of amorphadiene; subsequent replacement of the yeast HMGS (encoded by *ERG13*) with *S. aureus* HMGS (encoded by *mvaS*) further increased production. Together with the development of a novel nitrogen- and carbon-limited fed-batch fermentation process, this led to an average titre of more than 25 g per L amorphadiene[44]. Interestingly, although the cell lysate from the *E. coli* strain expressing *E. faecalis* HMGR had a higher HMGR specific activity than the strain expressing the yeast enzyme, it produced lower concentrations of amorphadiene. This observation, along with the earlier unexpected observation from the unbiased combinatorial screen, that reduced HMGS and HMGR expression increased mevalonate production[39], illustrates the limitations of rational engineering (such that an increase in enzyme activity does not necessarily result in more product), and that strain engineering requires the exploration of many permutations. This is an example of the design–build–test–analyse (DBTA) cycle that is described in BOX 1.

Stage 2: the switch to yeast-based production

Although the production of 25 g per L amorphadiene in *E. coli* was a key step forwards, the oxidized derivative of amorphadiene (artemisinic acid) is the preferred substrate for chemical conversion to semi-synthetic artemisinin[45,46]. Cell-free assays indicated that the oxidation of amorphadiene is accomplished by a cytochrome P450 enzyme that is located in the trichomes of *A. annua*[27], and a strategy was developed to identify and functionally express this enzyme in *S. cerevisiae*. In order to test the functionality of the identified P450 (CYP71AV1) in *S. cerevisiae*, it was expressed along with its cognate reductase (CPR1) in an amorphadiene-producing yeast strain, which enabled the production of 100 mg per L artemisinic acid[47].

The identification of CYP71AV1 and the ability to produce artemisinic acid in *S. cerevisiae* placed the project team in a quandary, given that the concentration of amorphadiene that was produced in yeast (150 mg per L) was much lower than that in *E. coli* (25 g per L). However, *E. coli* is typically unsuitable for the expression of eukaryotic P450 enzymes and it was unclear whether high titres of artemisinic acid were obtainable. After considering a number of possibilities (for example, extraction of amorphadiene from *E. coli*, followed by the use of *S. cerevisiae* for bioconversion to artemisinic acid), it was concluded that a single fermentation step would be needed to produce artemisinic acid at the price point that is required for the developing world (the per capita expenditure on healthcare in the developing world is only 1–10% of that in the developed world[25]). Two project teams (one of which worked with *E. coli* and one of which worked with *S. cerevisiae*) examined which organism was optimal and found that

Tunable intergenic regions (TIGRs). Stretches of DNA that are located between genes and that can be modified (for example, by the insertion of hairpins and cleavage sites) to enable each gene in an operon to be varied independently of all others in a random manner.

Fed-batch fermentation process
A batch culture that is fed continuously with nutrient medium. It differs from continuous culture owing to the variation in culture volume from beginning to end.

Cytochrome P450 enzyme
A member of a superfamily of monooxygenase enzymes that catalyse the oxidation of organic substrates. Plant microsomal cytochrome P450 enzymes require a reductase enzyme (CPR) to transfer electrons from NADPH and may also require cytochrome b_5 to supply electrons.

Trichomes
Hair-like or glandular structures on the surface of plants. Glandular trichomes are the major biosynthetic site of many natural plant products, including artemisinin.

S. cerevisiae was the superior organism. Although *E. coli* could produce oxidized intermediates of amorphadiene (for example, artemisinic alcohol) following the expression of CYP71AV1, it could only produce more than 1 g per L artemisinic acid when grown at 20 °C — an unsustainably low temperature for industrial fermentation[48] (D. Pitera, unpublished observations). By optimizing the production pathway in yeast, synthesis of 2.5 g per L artemisinic acid was achieved[49], and further improvements in titre were possible using an alternative yeast strain.

Use of an alternative yeast strain. The yeast strain that was initially used for the production of artemisinic acid was derived from *S. cerevisiae* strain S288C — the workhorse of yeast genetics[50] and the first eukaryote to have its whole genome sequenced[51]. However, because S288C sporulates poorly[52], strain construction using Mendelian genetics is difficult and little information is available on its performance in industrial fermentations. By contrast, *S. cerevisiae* strain CEN.PK2 was shown to have characteristics that are desirable for industrial fermentation, and is capable of sporulating sufficiently[53]. Initially, strain CEN.PK2 was engineered to have the same properties as the S288C-derived strain[47], and comparable artemisinic acid production was obtained in small-scale cultures. However, growth of the engineered CEN.PK2 strain in bioreactors led to the accumulation of acetate, which resulted in growth arrest. This was found to be due to the overexpression of a single gene (*UPC2-1*; C. J. P, unpublished observations). Thus, S288C and CEN.PK2 clearly differ as chassis organisms for the production of artemisinic acid in ways that were unanticipated. Rather than attempting to identify the underlying reason for the detrimental effect of *UPC2-1* expression in CEN.PK2, every enzyme of the mevalonate pathway up to and including *ERG20* (the final step for the production of FPP; BOX 2) was overexpressed in CEN.PK2 in an effort to increase the production of amorphadiene[54]. In addition, glucose was used as the carbon source instead of galactose, as it is approximately 100-fold cheaper than galactose and the use of galactose would have made the fermentation process prohibitively expensive for a pharmaceutical product that was destined for the developing world. The engineered CEN.PK2 strain produced fivefold higher concentrations of amorphadiene compared with the original S288C-derived strain, and, following further improvements in the fermentation process, the strain produced 40 g per L amorphadiene[54]. However, disappointingly, the increase in amorphadiene production did not result in a corresponding increase in artemisinic acid production when *CYP71AV1* was expressed, and amorphadiene production was tenfold higher than artemisinic acid production.

High-level artemisinic acid production. Further investigation of artemisinic acid production by the engineered CEN.PK2 strain expressing CYP71AV1 and its cognate reductase showed that artemisinic acid production resulted in a severe decrease in viability, which did not occur in amorphadiene-producing strains that lacked CYP71AV1 (REF. 55). It was initially suggested that yeast cells were unable to cope with the export of such a high flux of artemisinic acid, as it had previously been shown that production of artemisinic acid in the original S288C-derived strain caused the high-level induction of transporter gene expression[56]. However, transcript analysis indicated that cells that express CYP71AV1 experience severe oxidative stress, which was subsequently confirmed by use of the redox dye dihydrorhodamine 123 (K. Benjamin, personal communication). It was noted that poor coupling between cytochrome P450 enzymes and their cognate reductases can result in the release of reactive oxygen species[57] and that cytochrome P450 enzymes are normally present in molar excess compared to their cognate reductases[58]. CYP71AV1 and its reductase (CPR1) were expressed in the yeast strain from strong galactose-regulated promoters on a multicopy plasmid, and it was therefore assumed that the proteins were expressed at similarly high concentrations. The expression of CPR1 was consequently reduced, which led to improved viability, albeit with reduced production of artemisinic acid. It was also noted that the reaction rate of some cytochrome P450 enzymes increases following interaction with cytochrome b_5 (REFS 59,60). Thus, a cytochrome b_5 from *A. annua* was expressed in the CEN.PK2 production strain, which resulted in increased production of artemisinic acid, but the concentration achieved was still considerably lower than the target concentration of 25 g per L (REF. 55). Worryingly, this strain also produced a high concentration of artemisinic aldehyde, which is an oxidation intermediate that is highly reactive and is presumed to be toxic.

Owing to the difficulties of producing the required high concentrations of artemisinic acid in *S. cerevisiae*, the *A. annua* artemisinic aldehyde dehydrogenase[61] (*ALDH1*) was expressed, and this resulted in a substantial increase in the production of artemisinic acid. A gene that encodes a putative alcohol dehydrogenase was also found in the *A. annua* trichome-specific expressed sequence tag (EST) library. The gene was isolated and the encoded alcohol dehydrogenase was subsequently characterized, which revealed that it was an NAD-dependent artemisinic alcohol dehydrogenase (ADH1). The combined expression of *ADH1* and *ALDH1* produced the highest concentration of artemisinic acid that had been achieved so far[55] (FIGS 2,3). Finally, to ensure constitutive expression of the overexpressed mevalonate pathway enzymes and the *A. annua*-derived genes, the *GAL80* gene was deleted, which also made production independent of galactose (galactose had previously been required as a non-metabolized inducer). Coupled with development of the fermentation process (see below), this strain was capable of attaining the target titre of 25 g per L artemisinic acid that had been set at the beginning of the project[55].

It had been observed with the initial S288C-derived strain that artemisinic acid could be efficiently extracted from the cell pellet, presumably as a result of efficient transport out of yeast cells and subsequent extracellular

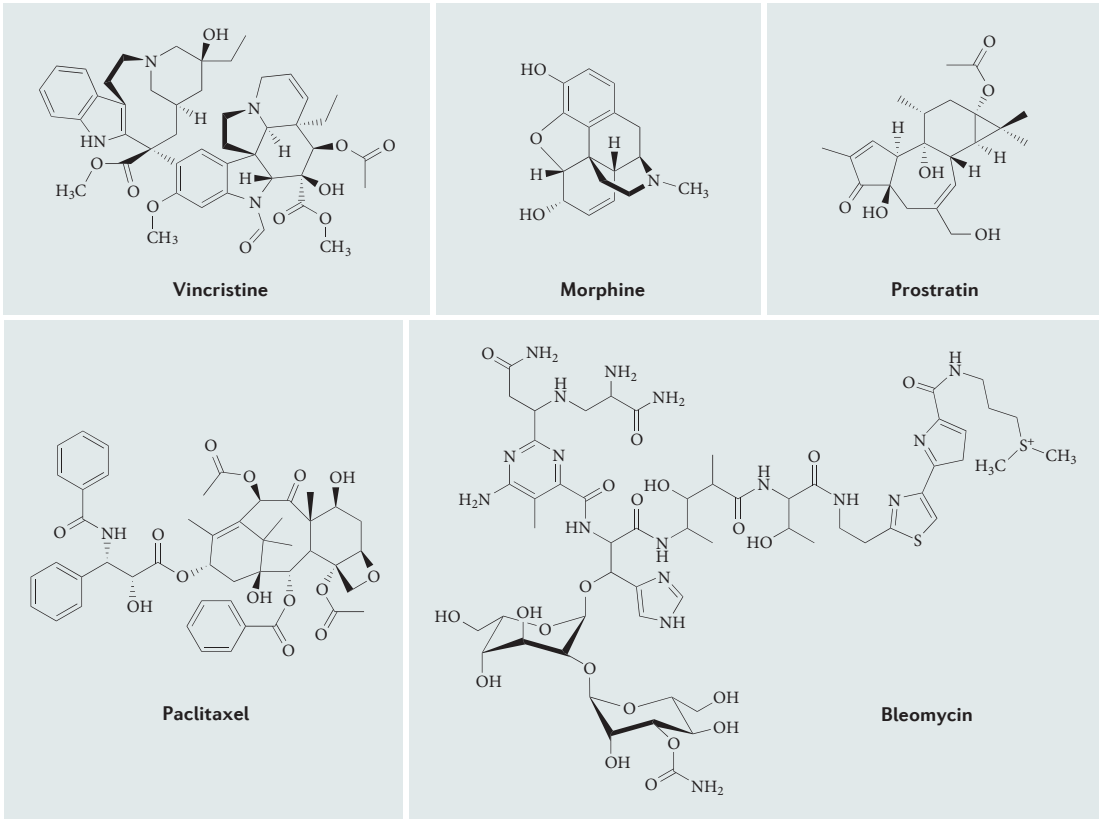


Figure 4 | Molecules amenable to synthetic production using similar principles to those described for semi-synthetic artemisinin. These molecules include vincristine, a vinca alkaloid; morphine, a benzyl isoquinoline alkaloid; prostratin, an isoprenoid; paclitaxel, an isoprenoid and bleomycin, a glycopeptide. Vincristine and paclitaxel are anticancer drugs, morphine is an analgesic, prostratin has anti-HIV properties and bleomycin is an antibiotic.

accumulation[47]. Following the increased production owing to *ALDH1* and *ADH1* expression, artemisinic acid was found to precipitate out of the growth medium and was actually visible to the naked eye. It was subsequently found that solubilization of artemisinic acid in the medium by the addition of isopropyl myristate (a hydrophobic overlay) enabled the extraction of higher concentrations (25 g per L) of artemisinic acid at reasonably high purity.

Stage 3: chemical conversion to artemisinin

For the final step, a novel chemistry process was developed for the conversion of the purified artemisinic acid to artemisinin[55]. The chemical process that was initially used[55] involves a four-step conversion: reduction of artemisinic acid to dihydroartemisinic acid, esterification of the carboxylic acid moiety on dihydroartemisinic acid (to block the subsequent formation of side products), generation of singlet oxygen to produce the 3-hydroperoxide and acid-catalysed Hock fragmentation and rearrangement, which leads to the generation of artemisinin in the presence of molecular oxygen. It is possible to scale-up this chemical process using batch reactors in organic chemical manufacturing sites worldwide, which does not require large amounts of capital or large-scale photoreactors that are expensive to operate. However, it should be noted that the commercial process that is used by Sanofi involves an alternative high-efficiency photochemical conversion process[62], which requires the use of specialized large-scale photoreactors.

Outlook: production of other pharmaceuticals

The production of semi-synthetic artemisinin is the first success story for the combined use of metabolic engineering and synthetic biology in the production of a pharmaceutical agent at industrial scale. As semi-synthetic artemisinin is functionally equivalent to the plant-derived drug[55], it has now been approved by the WHO for the preparation of approved artemisinin derivatives (such as artesunate) for incorporation into ACTs[63].

In the future, we predict that it will be possible to produce many different pharmaceutical products using similar techniques to those that were developed in the Semi-synthetic Artemisinin Project (FIG. 4). For example, biosynthesis of the potent anticancer drug Taxol has already been attempted in both *E. coli* and yeast[64,65]. This pathway contains many P450 enzymes and other enzymes that decorate the diterpene backbone of the molecule, which complicates its production in microorganisms. The candidate HIV drug prostratin might be more amenable to synthetic production, as the biosynthetic pathway that is involved seems to be less complex. Furthermore, precursors to several useful alkaloids

Hock fragmentation
A reaction of hydroperoxides connected to an unsaturated system, leading to cleavage of the C–C bond and the formation of two carbonyl compounds.

Photochemical conversion
A chemical reaction that is initiated by the absorption of energy in the form of visible, ultraviolet or infrared light.

Alkaloids
Nitrogen-containing natural products, many of which have pharmacological properties. Examples include cocaine, caffeine and the antimalarial drug quinine.

Box 3 | **Overcoming developmental obstacles**

A number of obstacles were overcome during the course of semi-synthetic artemisinin development to ensure that the process would deliver the final product at an affordable price for the developing world. For example, to divert carbon flux to artemisinic acid production in *Saccharomyces cerevisiae*, the expression of squalene synthase was reduced. In both the original S288C-derived strain and the production strain CEN.PK2, the methionine-repressible *S. cerevisiae MET3* promoter was used to inhibit the production of squalene synthase and channel flux into amorphadiene production[113,114]. This strategy required the addition of 1 mM methionine to the growth medium, which added to the cost of the process. Thus, for the production of artemisinic acid, the *MET3* promoter was replaced with the *CTR3* promoter, which is repressed by high copper concentrations. This modification enables the repression of squalene synthase by simply adding copper sulphate ($CuSO_4$) to the growth medium, which is a much cheaper alternative to methionine[55].

Much effort was devoted to optimizing the oxidation of amorphadiene to artemisinic acid, as it was unclear whether CYP71AV1 was capable of providing the flux that is required for the commercial production of semi-synthetic artemisinin. Thus, an alternative metabolic approach was investigated. Eukaryotic P450s, such as CYP71AV1, are anchored on the outer face of the endoplasmic reticulum, as are the cognate reductase (CPR1) and cytochrome b_5. However, cytochrome P450 enzymes from prokaryotes are soluble enzymes, and $P450_{BM3}$ from *Bacillus megaterium* has the highest turnover rate of any known P450 protein[115]. In addition, this enzyme was previously engineered to function on alternative substrates[116]. Thus, $P450_{BM3}$ was mutagenized to find a variant of the enzyme that was capable of oxidizing amorphadiene to artemisinic epoxide in *Escherichia coli*, followed by selective chemical transformation of the epoxide to dihydroartemisinic acid[117]. However, owing to concerns about the stability of the epoxide, combined with further characterization of CYP71AV1, it was decided that CYP71AV1 would be used for the oxidation of amorphadiene.

Recent work has characterized two different alleles of the gene that encodes CYP71AV1 (REF. 118). There are two chemotypes of *Artemisia annua*: high artemisinin producers (HAP chemotype; with low artemisinic acid production) and low artemisinin producers (LAP chemotype; with high artemisinic acid production). The LAP allele (which has a seven-amino-acid amino-terminal extension, unlike the HAP allele) that was used in the project produces substantially more oxidized product than the HAP allele when transiently expressed in the plant *Nicotiana benthamiana*. The allelic differences were unknown at the time that the strains were engineered, but, if the HAP allele of *CYP71AV1* had been used, it is likely that substantially less product would have been produced. The fact that two alleles of a pathway gene could potentially have such an effect on production illustrate that both a thorough understanding of the native pathway and a means to rapidly test different alleles is needed. Strain engineers are frequently confronted by multiple alleles of the same gene, with little knowledge about whether variants show differences in activity. Thus, the availability of a fast, reliable and cheap method of strain engineering, implicit in the vision of synthetic biology, to test multiple engineered variants would be very useful.

(such as analgesics, muscle relaxants and antioxidants) have been produced in *E. coli* and yeast[66,67]. As many of these molecules are naturally produced in plants, it might soon be possible to engineer plants to produce these molecules in higher yields[68]. Plant production might offer more advantages, particularly for metabolic pathways that have many cytochrome P450 enzymes or in which intermediates of the metabolic pathway might be toxic (such that it would be an advantage to sequester the intermediates in a cellular compartment).

So, what lessons can be taken from the project that could be applied to other metabolic engineering or synthetic biology projects? As outlined in this Review, it is crucial to thoroughly understand the production pathway in the native organism; however, the degree of knowledge that is required varies according to the particular goal of the project. For example, the mevalonate pathway is relatively well understood, so if the final desired product had been amorphadiene (which is the hydrocarbon precursor to artemisinin) it would have been possible to engineer the production organism using pre-existing knowledge; this is the approach that was taken for the initial work on amorphadiene production in *E. coli*[44]. However, the oxidation of amorphadiene to artemisinic acid was not well understood (see BOX 3 for an alternative approach, along with descriptions of other developmental obstacles that were overcome to enable the production of semi-synthetic artemisinin at the required price point). Although CYP71AV1 is capable of the three-step oxidation reaction that is required for the conversion of amorphadiene to artemisinic acid[47], the essential role of the two dehydrogenases (ADH1 and ALDH1) was not appreciated until later in the project. The expression of these two enzymes in the production yeast proved to be vital for obtaining the titres and yields that are required to meet the economics of production for a drug for the developing world[55].

The production process that is used can either be completely biosynthetic (such as conversion of a simple carbon substrate to the final product in a living organism), semi-synthetic (such as the biosynthetic production of artemisinic acid, followed by chemical conversion to artemisinin[55]) or even a bioconversion of one intermediate into another (in those situtaions in which a chemical conversion pathway is not possible, bioconversion is preceded by, and frequently followed by, organic chemical synthesis). The production of benzylisoquinoline alkaloids, many of which are pharmacologically active (such as morphine) in either *E. coli*[67] or yeast[66] is another example of semi-synthetic production; however, in this case, the synthesis pathway relies on commercially available intermediates, as many of the plant enzymes that catalyse early biosynthetic steps have not yet been isolated and cloned.

It is also important that the chassis organism can be engineered to produce the required product under manufacturing conditions. Some biosynthetic enzymes in the native host might have reduced activity in the desired chassis organism, which complicates optimization of the production process. For example, in the Semi-synthetic Artemisinin Project, the plant cytochrome P450 enzyme, CYP71AV1, could not be expressed in *E. coli* at 30 °C (REF. 48), which necessitated the switch to production in yeast. It is possible to bypass the need for such enzymes; for example, to produce the benzylisoquinoline alkaloid precursor (S)-reticuline (which is a precursor of many pharmaceutically active alkaloids, such as morphine and codeine) in *E. coli,* a modified biosynthetic pathway was constructed to eliminate the need to express a cytochrome P450 enzyme[69]. However, production of the bioactive alkaloid magnoflorine requires expression of the cytochrome P450 enzyme CYP80G2, which was expressed in yeast cells that were co-cultured with (S)-reticuline-producing *E. coli*[67]. A similar co-culturing approach for the semi-synthetic production of artemisinin was considered, but was not feasible on an industrial scale to meet the required price point. Nonetheless, this study[67] provides proof-of-concept for the value of co-culturing approaches.

Chemotypes
Plants that are morphologically similar or identical but that are distinguished by differences in the production of secondary metabolites.

Yeast homologous recombination
A method to assemble DNA fragments with overlapping homology regions by transforming yeast and using its innate homologous recombination ability to correctly assemble the fragments.

Ligase chain reaction
(LCR). A method to assemble DNA fragments using bridging oligonucleotides, a thermostable DNA polymerase and multiple denaturation–annealing–ligation temperature cycles.

Circular polymerase extension cloning
(CPEC). A polymerase chain reaction (PCR)-based method for the assembly of multiple DNA fragments.

Gibson isothermal assembly
A single-temperature enzymatic method of assembling multiple DNA fragments.

The Semi-synthetic Artemisinin Project used the best understood and most genetically tractable chassis microorganisms — *E. coli* and *S. cerevisiae* — as the basis for all biological engineering. The ability to engineer these organisms using the tools of synthetic biology, and thus rapidly progress through the DBTA engineering cycle (BOX 1), makes them attractive hosts. Progress in synthetic biology since the completion of the Semi-synthetic Artemisinin Project, such as the discovery and development of the CRISPR–Cas system for genome engineering of yeast[70] and bacteria[71] and the development of multiplex genome engineering technologies for both *E. coli*[72] and yeast[73], only add to the allure of these organisms. However, some pharmaceutical agents, especially those of microbial origin, might be more amenable to production in the native organism. Whole-genome sequencing has shown that many fungi encode biosynthetic modules for many more natural products than are currently known — for example, the polyketides (which are a class of complex secondary metabolites, many of which are pharmacologically active, such as the antibiotic erythromycin) and meroterpenoids (which are molecules produced from both polyketide and terpenoid precursors, some of which are known to be pharmacologically active[74]) that are produced by *Aspergillus nidulans*[74,75]. The development of efficient gene targeting for the replacement of the native promoters of polyketide synthase genes with controllable, regulated promoters has enabled the production of novel polyketides in *A. nidulans*[75]. A similar synthetic biology approach (known as promoter replacement), assembly of the reconstituted gene cluster by yeast homologous recombination[76–78] and integration into the genome of a tractable heterologous host has enabled the activation and characterization of a novel macrolactam gene cluster in *Streptomyces griseus*[79]. For large-scale production of such natural products, it will be necessary to assess whether to construct the biosynthetic pathways in the native host or to heterologously engineer production in a more amenable chassis organism, such as yeast or *E. coli*.

Irrespective of the chassis organism that is chosen or the pharmaceutical product that is under development, the ability to rapidly engineer the organism and manipulate the pathways involved (for optimization purposes) are fundamental requirements of synthetic biology. A standardized system of DNA assembly is needed to implement this vision. Several methods have been developed for restriction enzyme-independent assembly of large DNA modules, including yeast homologous recombination[76–78,80], ligase chain reaction (LCR)[81], circular polymerase extension cloning (CPEC)[82] and Gibson isothermal assembly[83]. A recent comparison of these assembly methods indicates that yeast homologous recombination and LCR enable the most reliable assembly of up to 12 DNA components (for example, promoters, genes and markers), but suggests that LCR provides the fastest and easiest workflow[84]. The implementation of automated DNA-assembly methods, preferably designed and implemented using a standardized, roboticized, computer aided design–computer aided manufacturing (CAD–CAM) system[85] (BOX 1) will hasten the DBTA cycle[86] and accelerate the pace of pharmaceutical drug development using synthetic biology.

1. Rainsford, K. D. *Aspirin and Related Drugs* (CRC Press, 2004).
2. Garavito, R. M. Working Knowledge: aspirin. *Sci. Am.* **280**, 108 (1999).
3. Gerhardt, C. Untersuchungen über die wasserfreien organischen Säuren. *Justus Liebigs Annalen Chemie* **87**, 57–84 (in German) (1853).
4. White, N. J. Qinghaosu (artemisinin): the price of success. *Science* **320**, 330–334 (2008).
A review of the history and properties of artemisinin.
5. Li, Y. & Wu, Y. L. How Chinese scientists discovered qinghaosu (artemisinin) and developed its derivatives? What are the future perspectives? *Med. Trop.* **58**, 9–12 (1998).
6. World Health Organisation. *Guidelines for the treatment of malaria.* 2nd edn (WHO, 2010).
7. World Health Organisation. *WHO informal consultation with manufacturers of artemisinin-based pharmaceutical products in use for the treatment of malaria* (WHO, 2007).
8. Hale, V., Keasling, J. D., Renninger, N. & Diagana, T. T. Microbially derived artemisinin: a biotechnology solution to the global problem of access to affordable antimalarial drugs. *Am. J. Trop. Med. Hyg.* **77**, 198–202 (2007).
9. World Health Organisation. *World Malaria Report 2012* (WHO, 2012).
10. Hsu, E. Reflections on the 'discovery' of the antimalarial qinghao. *Br. J. Clin. Pharmacol.* **61**, 666–670 (2006).
11. Qinghaosu Antimalarial Coordinating Research Group. Antimalaria Studies on Qinghaosu. *Chin. Med. J.***12**, 811–816 (1979).
12. Klayman, D. L. Qinghaosu (artemisinin): an antimalarial drug from China. *Science* **228**, 1049–1055 (1985).
13. Dalrymple, D. *Artemisia Annua*, artemisinin, ACTs & malaria control in Africa. Tradition, science and public policy [online], http://www.mmv.org/newsroom/publications/artemisia-annua-artemisinin-acts-and-malaria-control-africa-interplay-traditio (2012).
14. Cui, L. & Su, X. Z. Discovery, mechanisms of action and combination therapy of artemisinin. *Expert Rev. Anti-Infective Ther.* **7**, 999–1013 (2009).
15. Barnes, K. I., Chanda, P. & Ab Barnabas, G. Impact of the large-scale deployment of artemether/lumefantrine on the malaria disease burden in Africa: case studies of South Africa, Zambia and Ethiopia. *Malar J.* **8**, S8 (2009).
16. Cheeseman, I. H. *et al.* A major genome region underlying artemisinin resistance in Malaria. *Science* **336**, 79–82 (2012).
17. Dondorp, A. M. *et al.* Artemisinin resistance: current status and scenarios for containment. *Nature Rev. Microbiol.* **8**, 272–280 (2010).
18. Fairhurst, R. M. *et al.* Artemisinin-resistant Malaria: research challenges, opportunities, and public health implications. *Am. J. Trop. Med. Hyg.* **87**, 231–241 (2012).
19. Takala-Harrison, S. *et al.* Genetic loci associated with delayed clearance of *Plasmodium falciparum* following artemisinin treatment in Southeast Asia. *Proc. Natl Acad. Sci.* **110**, 240–245 (2013).
20. Miotto, O. *et al.* Multiple populations of artemisinin-resistant *Plasmodium falciparum* in Cambodia. *Nature Genet.* **45**, 648–655 (2013).
21. Ariey, F. *et al.* A molecular marker of artemisinin-resistant *Plasmodium falciparum* malaria. *Nature* **505**, 50–55 (2014).
This paper reports the identification of a molecular marker for artemisinin-resistant malaria, which will be important for the tracking and elimination of artemisinin-resistant parasites.
22. Shretta, R. & Yadav, P. Stabilizing supply of artemisinin and artemisinin-based combination therapy in an era of wide-spread scale-up. *Malar J.* **11**, 399 (2012).
23. Dahlberg Global Development Advisors. Independent Mid-Term review of the assured Artemisinin Supply System (A2S2) Project. Geneva: UNITAID [online], http://www.unitaid.eu/images/projects/malaria/110406_A2S2_Final_Report.pdf (2012).
24. Assured Artemisinin Supply System. Artemisinin imports into India [online], http://www.a2s2.org/market-data/artemisinin-imports-into-india.html (updated 13 january 2014).
25. WHO Health Financing. Per capita total expenditure on health at average exchange rate (US $) [online], http://gamapserver.who.int/gho/interactive_charts/health_financing/atlas.html?indicator=i3&date=2011 (2011).
26. The World Bank. Health expenditure per capita (current US$). [online] http://Data.Worldbank.org/Indicator/SH.XPD.PCAP (2013).
27. Bertea, C. M. *et al.* Identification of intermediates and enzymes involved in the early steps of artemisinin biosynthesis in *Artemisia annua*. *Planta Med.* **71**, 40–47 (2005).
28. Brown, G. D. The biosynthesis of artemisinin (Qinghaosu) and the phytochemistry of *Artemisia annua* L. (Qinghao). *Molecules* **15**, 7603–7698 (2010).
29. Bouwmeester, H. J. *et al.* Amorpha-4,11-diene synthase catalyses the first probable step in artemisinin biosynthesis. *Phytochemistry* **52**, 843–854 (1999).
30. Covello, P. S., Teoh, K. H., Polichuk, D. R., Reed, D. W. & Nowak, G. Functional genomics and the biosynthesis of artemisinin. *Phytochemistry* **68**, 1864–1871 (2007).
31. Paddon, C. *et al.* In *Isoprenoid Synthesis in Plants and Microorganisms* (eds Bach, T. J. & Rohmer, M.) 91–106 (Springer, 2013).
32. Zhang, Y. *et al.* The molecular cloning of artemisinic aldehyde Δ11(13) reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. *J. Biol. Chem.* **283**, 21501–21508 (2008).
33. Zhao, L., Chang, W. C., Xiao, Y., Liu, H. W. & Liu, P. Methylerythritol phosphate pathway of isoprenoid biosynthesis. *Annu. Rev. Biochem.* **82**, 497–530 (2013).
34. Reiling, K. K. *et al.* Mono and diterpene production in *Escherichia coli*. *Biotechnol. Bioeng.* **87**, 200–212 (2004).
35. Martin, V. J., Pitera, D. J., Withers, S. T., Newman, J. D. & Keasling, J. D. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nature Biotech.* **21**, 796–802 (2003).
This paper provides a description of the initial amorphadiene production pathway in *E. coli*.

36. Martin, V. J., Yoshikuni, Y. & Keasling, J. D. The *in vivo* synthesis of plant sesquiterpenes by *Escherichia coli*. *Biotechnol. Bioeng.* **75**, 497–503 (2001).
37. Newman, J. D. *et al.* High-level production of amorpha-4,11-diene in a two-phase partitioning bioreactor of metabolically engineered *Escherichia coli*. *Biotechnol. Bioeng.* **95**, 684–691 (2006).
38. Pitera, D. J., Paddon, C. J., Newman, J. D. & Keasling, J. D. Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. *Metab. Eng.* **9**, 193–207 (2007).
39. Pfleger, B. F., Pitera, D. J., Smolke, C. D. & Keasling, J. D. Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nature Biotech.* **24**, 1027–1032 (2006).
40. Pfleger, B. F., Pitera, D. J., Newman, J. D., Martin, V. J. & Keasling, J. D. Microbial sensors for small molecules: development of a mevalonate biosensor. *Metab. Eng.* **9**, 30–38 (2007).
41. Kizer, L., Pitera, D. J., Pfleger, B. F. & Keasling, J. D. Application of functional genomics to pathway optimization for increased isoprenoid production. *Appl. Environ. Microbiol.* **74**, 3229–3241 (2008).
42. Tabata, K. & Hashimoto, S. Production of mevalonate by a metabolically-engineered *Escherichia coli*. *Biotechnol. Lett.* **26**, 1487–1491 (2004).
43. Hedl, M., Tabernero, L., Stauffacher, C. V. & Rodwell, V. W. Class II 3-hydroxy-3-methylglutaryl coenzyme A reductases. *J. Bacteriol.* **186**, 1927–1932 (2004).
44. Tsuruta, H. *et al.* High-level production of amorpha-4,11-diene, a precursor of the antimalarial agent artemisinin, in *Escherichia coli*. *PLoS ONE* **4**, e4489 (2009).
This study reports strain engineering and fermentation development, which culminated in the production of 25 g per L amorphadiene in *E. coli*.
45. Roth, R. J. & Acton, N. A simple conversion of artemisinic acid into artemisinin. *J. Nature Prod.* **52**, 1183–1185 (1989).
46. Roth, R. J. & Roth, N. A. Simple conversion of artemisinic acid into artemisinin. US Patent 4992561 (1991).
47. Ro, D. K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943 (2006).
This paper describes the isolation of CYP71AV1 — the cytochrome P450 enzyme that oxidizes amorphadiene — and provides the first demonstration of artemisinic acid production in *S. cerevisiae*.
48. Chang, M. C., Eachus, R. A., Trieu, W., Ro, D. K. & Keasling, J. D. Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. *Nature Chem. Biol.* **3**, 274–277 (2007).
49. Lenihan, J. R., Tsuruta, H., Diola, D., Renninger, N. S. & Regentin, R. Developing an industrial artemisinic acid fermentation process to support the cost-effective production of antimalarial artemisinin-based combination therapies. *Biotechnol. Prog.* **24**, 1026–1032 (2008).
50. Mortimer, R. K. & Johnston, J. R. Genealogy of principal strains of the yeast genetic stock center. *Genetics* **113**, 35–43 (1986).
51. Goffeau, A. *et al.* Life with 6000 genes. *Science* **274**, 563–567 (1996).
52. Ben-Ari, G. *et al.* Four linked genes participate in controlling sporulation efficiency in budding yeast. *PLoS Genet.* **2**, e195 (2006).
53. van Dijken, J. P. *et al.* An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. *Enzyme Microb. Technol.* **26**, 706–714 (2000).
54. Westfall, P. J. *et al.* Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc. Natl Acad. Sci. USA* **109**, E111–E118 (2012).
This paper describes strain engineering and fermentation development, which enabled the production of 40 g per L amorphadiene in *S. cerevisiae*, followed by the chemical conversion of amorphadiene to dihydroartemisinic acid.
55. Paddon, C. J. *et al.* High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**, 528–532 (2013).
This paper describes strain engineering and fermentation development — which enabled the production of 25 g per L artemisinic acid in *S. cerevisiae* — and an efficient, non-photochemical conversion to artemisinin.
56. Ro, D. K. *et al.* 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 (2008).
57. Zangar, R. C., Davydov, D. R. & Verma, S. Mechanisms that regulate production of reactive oxygen species by cytochrome P450. *Toxicol. Appl. Pharmacol.* **199**, 316–331 (2004).
58. Peterson, J. A., Ebel, R. E., O'Keeffe, D. H., Matsubara, T. & Estabrook, R. W. Temperature dependence of cytochrome P-450 reduction. A model for NADPH-cytochrome P-450 reductase:cytochrome P-450 interaction. *J. Biol. Chem.* **251**, 4010–4016 (1976).
59. Schenkman, J. B. & Jansson, I. The many roles of cytochrome b_5. *Pharmacol. Ther.* **97**, 139–152 (2003).
60. Zhang, H., Im, S. C. & Waskell, L. Cytochrome b_5 increases the rate of product formation by cytochrome P450 2B4 and competes with cytochrome P450 reductase for a binding site on cytochrome P450 2B4. *J. Biol. Chem.* **282**, 29766–29776 (2007).
61. Teoh, K. H., Polichuk, D. R., Reed, D. W. & Covello, P. S. Molecular cloning of an aldehyde dehydrogenase implicated in artemisinin biosynthesis in *Artemisia annua*. *Botany* **87**, 635–642 (2009).
This study reports the identification of the *A. annua* aldehyde dehydrogenase, which is required for high-level production of artemisinic acid in *S. cerevisiae*.
62. Sanofi. Prix Potier 2012 (in French) [online], http://www.sanofi.fr/l/fr/fr/layout.jsp?cnt=F4BEF56E-3727-40C9-A6A2-29894774AD12 (2012).
63. WHO Prequalificaion of Medicines Programme. Acceptance of non-plant-derived-artemisinin offers potential to increase access to malaria treatment [online], http://apps.who.int/prequal/info_press/documents/PQ_non-plant_derived_artemisinin_1.pdf (2013).
64. Ajikumar, P. K. *et al.* Isoprenoid pathway optimization for taxol precursor overproduction in *E. coli*. *Science* **330**, 70–74 (2010).
65. Engels, B., Dahm, P. & Jennewein, S. Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. *Metab. Eng.* **10**, 201–206 (2008).
66. Hawkins, K. M. & Smolke, C. D. Production of benzylisoquinoline alkaloids in *Saccharomyces cerevisiae*. *Nature Chem. Biol.* **4**, 564–573 (2008).
67. Minami, H. *et al.* Microbial production of plant benzylisoquinoline alkaloids. *Proc. Natl Acad. Sci. USA* **105**, 7393–7398 (2008).
68. Glenn, W., Runguphan, W. & O'Connor, S. Recent progress in the metabolic engineering of alkaloids in plant systems. *Curr. Opin. Biotechnol.* **24**, 354–365 (2013).
69. Nakagawa, A. *et al.* A bacterial platform for fermentative production of plant alkaloids. *Nature Commun.* **2**, 326 (2011).
70. DiCarlo, J. E. *et al.* Genome engineering in *Saccharomyces cerevisiae* using CRISPR–Cas systems. *Nucleic Acids Res.* **41**, 4336–4343 (2013).
This paper reports the use of the CRISPR–Cas system in yeast, which has the potential to markedly advance genome engineering in this organism.
71. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR–Cas systems. *Nature Biotech.* **31**, 233–239 (2013).
72. Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894–898 (2009).
73. Dicarlo, J. E. *et al.* Yeast oligo-mediated genome engineering (YOGE). *ACS Synth. Biol.* **2**, 741–749 (2013).
74. Lo, H. C. *et al.* Two separate gene clusters encode the biosynthetic pathway for the meroterpenoids austinol and dehydroaustinol in *Aspergillus nidulans*. *J. Am. Chem. Soc.* **134**, 4709–4720 (2012).
75. Ahuja, M. *et al.* Illuminating the diversity of aromatic polyketide synthases in *Aspergillus nidulans*. *J. Am. Chem. Soc.* **134**, 8212–8221 (2012).
76. Ma, H., Kunes, S., Schatz, P. J. & Botstein, D. Plasmid construction by homologous recombination in yeast. *Gene* **58**, 201–216 (1987).
77. Oldenburg, K. R., Vo, K. T., Michaelis, S. & Paddon, C. Recombination-mediated PCR-directed plasmid construction *in vivo* in yeast. *Nucleic Acids Res.* **25**, 451–452 (1997).
78. Shao, Z. & Zhao, H. DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res.* **37**, e16 (2009).
79. Luo, Y. *et al.* Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster. *Nature Commun.* **4**, 2894 (2013).
80. Kuijpers, N. *et al.* A versatile, efficient strategy for assembly of multi-fragment expression vectors in *Saccharomyces cerevisiae* using 60-bp synthetic recombination sequences. *Microb. Cell Factories* **12**, 47 (2013).
81. Pachuk, C. J. *et al.* Chain reaction cloning: a one-step method for directional ligation of multiple DNA fragments. *Gene* **243**, 19–25 (2000).
82. Quan, J. & Tian, J. Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS ONE* **4**, e6441 (2009).
83. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**, 343–345 (2009).
84. de Kok, S. *et al.* Rapid and reliable DNA assembly via ligase cycling reaction. *ACS Synthet. Biol.* **3**, 97–106 (2014).
85. Leguia, M., Brophy, J., Densmore, D. & Anderson, J. C. Automated assembly of standard biological parts. *Methods Enzymol.* **498**, 363–397 (2011).
86. Nielsen, J. & Keasling, J. D. Synergies between synthetic biology and metabolic engineering. *Nature Biotech.* **29**, 693–695 (2011).
87. Bailey, J. E. Toward a science of metabolic engineering. *Science* **252**, 1668–1675 (1991).
This paper provides an early description and vision of metabolic engineering.
88. Woolston, B. M., Edgar, S. & Stephanopoulos, G. Metabolic engineering: past and future. *Annu. Rev. Chem. Biomol. Eng.* **4**, 259–288 (2013).
89. Nakamura, C. E. & Whited, G. M. Metabolic engineering for the microbial production of 1,3-propanediol. *Curr. Opin. Biotechnol.* **14**, 454–459 (2003).
90. Chen, J., Densmore, D., Ham, T. S., Keasling, J. D. & Hillson, N. J. DeviceEditor visual biological CAD canvas. *J. Biol. Eng.* **6**, 1 (2012).
91. Hillson, N. J., Rosengarten, R. D. & Keasling, J. D. j5 DNA assembly design automation software. *ACS Synth. Biol.* **1**, 14–21 (2012).
92. Ham, T. S. *et al.* Design, implementation and practice of JBEI-ICE: an open source biological part registry platform and tools. *Nucleic Acids Res.* **40**, e141 (2012).
93. Linshiz, G. *et al.* PaR–PaR laboratory automation platform. *ACS Synth. Biol.* **2**, 216–222 (2013).
94. Mutalik, V. K. *et al.* Quantitative estimation of activity and quality for collections of functional genetic elements. *Nature Methods* **10**, 347–353 (2013).
95. Endy, D. Foundations for engineering biology. *Nature* **438**, 449–453 (2005).
96. Gardner, T. S. Synthetic biology: from hype to impact. *Trends Biotechnol.* **31**, 123–125 (2013).
97. Andrianantoandro, E., Basu, S., Karig, D. K. & Weiss, R. Synthetic biology: new engineering rules for an emerging discipline. *Mol. Syst. Biol.* **2**, 2006.0028 (2006).
98. Stephanopoulos, G. Synthetic biology and metabolic engineering. *ACS Synth. Biol.* **1**, 514–525 (2012).
99. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
100. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).
References 99 and 100 are two of the earliest publications on synthetic biology and demonstrate the potential to apply engineering principles to biology.
101. Stemmer, W. P. C., Crameri, A., Ha, K. D., Brennan, T. M. & Heyneker, H. L. Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene* **164**, 49–53 (1995).
102. Mutalik, V. K. *et al.* Precise and reliable gene expression via standard transcription and translation initiation elements. *Nature Methods* **10**, 354–360 (2013).
103. Bonnet, J., Subsoontorn, P. & Endy, D. Rewritable digital data storage in live cells via engineered control of recombination directionality. *Proc. Natl Acad. Sci.* **109**, 8884–8889 (2012).
104. Bonnet, J., Yin, P., Ortiz, M. E., Subsoontorn, P. & Endy, D. Amplifying genetic logic gates. *Science* **340**, 599–603 (2013).
105. Ajikumar, P. K. *et al.* Terpenoids: opportunities for biosynthesis of natural product drugs using engineered microorganisms. *Mol. Pharm.* **5**, 167–190 (2008).
106. Paltauf, F., Kohlwein, S. D. & Henry, S. A. in *The Molecular and Cellular Biology of the Yeast* Saccharomyces (eds Jones, E. W., Pringle, J. R. & Broach, J. R.) 425–500 (Cold Spring Harbor Laboratory Press, 1992).

107. Shiba, Y., Paradise, E. M., Kirby, J., Ro, D. K. & Keasling, J. D. Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metab. Eng.* **9**, 160–168 (2006).
108. Eisenreich, W., Bacher, A., Arigoni, D. & Rohdich, F. Biosynthesis of isoprenoids via the non-mevalonate pathway. *Cell. Mol. Life Sci.* **61**, 1401–1426 (2004).
109. Putra, S. R., Disch, A., Bravo, J. M. & Rohmer, M. Distribution of mevalonate and glyceraldehyde 3-phosphate/pyruvate routes for isoprenoid biosynthesis in some Gram-negative bacteria and mycobacteria. *FEMS Microbiol. Lett.* **164**, 169–175 (1998).
110. Lichtenthaler, H. K. The 1-deoxy-d-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 47–65 (1999).
111. Carlsen, S. *et al.* Heterologous expression and characterization of bacterial 2-C-methyl-D-erythritol-4-phosphate pathway in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **97**, 5753–5769 (2013).
112. Lill, R. & Muhlenhoff, U. Maturation of iron–sulfur proteins in eukaryotes: mechanisms, connected processes, and diseases. *Annu. Rev. Biochem.* **77**, 669–700 (2008).
113. Paradise, E. M., Kirby, J., Chan, R. & Keasling, J. D. Redirection of flux through the FPP branch-point in *Saccharomyces cerevisiae* by down-regulating squalene synthase. *Biotechnol. Bioeng.* **100**, 371–378 (2008).
114. Asadollahi, M. A. *et al.* Production of plant sesquiterpenes in *Saccharomyces cerevisiae*: effect of ERG9 repression on sesquiterpene biosynthesis. *Biotechnol. Bioeng.* **99**, 666–677 (2008).
115. Noble, M. A. *et al.* Roles of key active-site residues in flavocytochrome P450 BM3. *Biochem. J.* **339**, 371–379 (1999).
116. Glieder, A., Farinas, E. T. & Arnold, F. H. Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. *Nature Biotech.* **20**, 1135–1139 (2002).
117. Dietrich, J. A. *et al.* A novel semi-biosynthetic route for artemisinin production using engineered substrate-promiscuous P450(BM3). *ACS Chem. Biol.* **4**, 261–267 (2009).
This paper describes the adaptation of a high-activity bacterial cytochrome P450 enzyme for the production of an oxidized artemisinin precursor.
118. Ting, H. M. *et al.* The metabolite chemotype of *Nicotiana benthamiana* transiently expressing artemisinin biosynthetic pathway genes is a function of CYP71AV1 type and relative gene dosage. *New Phytol.* **199**, 352–366 (2013).

Acknowledgements
The authors thank the Bill and Melinda Gates Foundation for their generous support of this project. They also thank their many colleagues at the University of California Berkeley, Amyris, Sanofi, National Research Council of Canada (NRC) and PATH Drug Solutions for their unstinting efforts that made this work a success.

Competing interests statement
The authors declare competing interests: see Web version for details.

FURTHER INFORMATION
PATH Drug Solutions: http://sites.path.org/drugdevelopment
***Nature Reviews Microbiology* Focus on Synthetic Biology:** http://www.nature.com/nrmicro/focus/synbio

ALL LINKS ARE ACTIVE IN THE ONLINE PDF