Robust platform for *de novo* production of heterologous polyketides and nonribosomal peptides in *Escherichia coli*

Kenji Watanabe*[a] and Hideaki Oikawa*[b]

Received 26th October 2006
First published as an Advance Article on the web 21st December 2006
DOI: 10.1039/b615589h

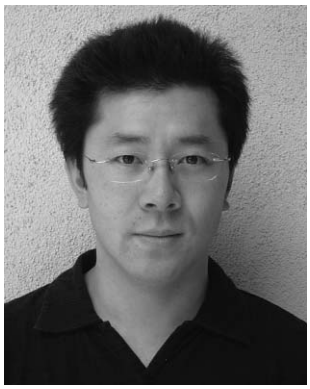
During the past decade, numerous gene clusters responsible for the biosynthesis of important natural products have been identified from a variety of organisms. Heterologous expression utilizing *E. coli* has been employed to provide proteins for mechanistic understanding and structural analyses. It was very recently shown that this system is also capable of *de novo* production of biologically active forms of heterologous nonribosomal peptides, echinomycin and triostin A, through the introduction of genes encoding modules responsible for their assembly into this model bacterial host. The superlative advantage of using *E. coli* as a heterologous host is the availability of a wealth of well-established molecular biological techniques for its genetic and metabolic manipulation. The platform described above which was developed in our laboratory is ideal for use in the production of metabolites found in marine and symbiotic bacteria that are not amenable to artificial cultivation. Development and tailoring of our system will allow for the design of these natural products and ultimately combinatorial yet rational modification of these compounds. This review focuses on the heterologous expression of biosynthetic gene clusters for the assembly of therapeutically potent compounds.

Introduction

E. coli has been used for molecular cloning in genetics, molecular biology and microbiology, and generation of recombinant proteins in biological chemistry and structural biology in quantities sufficient for proteins to be characterized. Scientists of various disciplines have relied on *E. coli* for the progression of their scope by taking advantage of this bacterium's fast life cycle. In addition, knowledge of the complete genome and extensive understanding of many aspects of the metabolic pathways of this organism also greatly reduces the uncertainty in the research process. This organism is a powerful tool that has spread across to other research fields.

[a]*Department of Pharmaceutical Sciences, University of Southern California, 1985 Zonal Ave PSC 718 Los Angeles, California 90033, USA. E-mail: kenjiwat@usc.edu*
[b]*Division of Chemistry, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan. E-mail: hoik@sci.hokudai.ac.jp*

Kenji Watanabe received his BA (1996) and PhD (2000) in biological chemistry from Hokkaido University, Sapporo, Japan. His postdoctoral fellowships included a year at the School of Pharmacy, University of Wisconsin-Madison, and a two-year appointment at Stanford University studying in the departments of chemistry and chemical engineering. In 2003, he returned to Hokkaido University as an assistant in the agriculture department. In 2004, he joined the department of pharmaceutical sciences at the School of Pharmacy, University of Southern California, as a senior research associate. In 2006, he was promoted to research assistant professor (non-tenure track faculty) in the department of pharmaceutical sciences.



Kenji Watanabe

Hideaki Oikawa

Hideaki Oikawa is a Professor at the Division of Chemistry and the Graduate School of Science, Hokkaido University. He received his PhD (1984) in organic chemistry from Hokkaido University. He worked as a postdoctoral fellow with Professor D. E. Cane at Brown University from 1984 and with Dr K. Isono at RIKEN in Japan from 1985. He joined the department of agriculture, Hokkaido University, as an Assistant Professor in 1986 and was promoted to Associate Professor in 1999. He moved to the Division of Chemistry and Graduate School of Science at Hokkaido University as a Professor in 2003.

Over 10% of drugs currently in commercial use are microbial products.[1] Some examples include antibiotics such as streptomycin, kanamycin, chloramphenicol, rifamycin and vancomycin which were isolated from streptomycetes producing these secondary metabolites and bringing notoriety to these bacteria, making them a topic of interest. Katz's and Leadlay's groups in the early 1990s independently isolated the erythromycin biosynthetic gene cluster from *Saccharopolyspora erythraea* (formerly *Streptomyces erythraeus*).[2,3] This was a feat of genetic contribution to understanding polyketide synthase (PKS). It is well known that streptomycetes produce two classes of natural products, polyketides (PKs) and nonribosomal peptides (NRPs) (Fig. 1).[4,5] PKs can be further broken down into multiple subcategories (types I, II, and III) according to chemical structure and method of assembly. The DNA cluster coding for a type I polyketide synthase is a modularly-organized enzyme responsible for PK biosynthesis (Fig. 2a). The second category, NRPs, are synthesized by large and multi-domain complexes and are genetically encoded in a similar manner to PKs (Fig. 2b). Because of the enormous structural diversity and complexity of both PKs and NRPs, many pharmaceutically and agrochemically important pharmacophores have been identified and are in use at the present. High density cell fermentation of the strain producing these natural products is the most popular method of preparing these bioactive molecules for clinical use. The process is economical, requiring mostly inexpensive starting materials and mild conditions that involve minimal energy input for operation. The process is also scalable, which facilitates any need to meet an increase in demand and obtain larger quantities of the product for clinical use. However, many organisms, including the vast majority of soil bacteria, are frequently not amenable to artificial cultivation for isolation of these products.

Today, 54% of clinical drugs on the market were developed from natural origins.[6] To create a successful drug, a lead compound may require some modification in order to improve its potency for treatment and reduce or completely abolish any unexpected toxicities and unfavorable pharmacokinetic properties. Currently, there are three major approaches to preparing these lead compounds and their analogs: (1) fermentation, (2) total chemical synthesis, and (3) semisynthesis. Each of these methods carries its own strengths and weaknesses in terms of its productivity and its ability to generate analogs, but none stand out as an ideal system for facile scalable production and none permits easy access to rational analog design and production. We believe that an *E. coli*-based heterologous host biosynthetic system could combine the advantages of the above-mentioned approaches and would become a valuable alternative for the advancement of the biosynthetic field. We will start with an overview of the heterologous expression of type I PKS and mixed PKS-NRPS biosynthetic gene clusters responsible for producing medicinally important secondary metabolites from bacteria. This review will then survey our efforts towards the isolation of the echinomycin biosynthetic gene cluster and heterologous production of NRP antibiotics in *E. coli* developed in our laboratory.

Streptomyces- and *E. coli*-based heterologous systems for PK and NRP production

There are four major advantages to using a *Streptomyces*-based expression system for heterologous production of PKs or NRPs found in other *Streptomyces*: (1) ample genetic compilation, (2) innate availability of starting and building units, (3) facile expression of immense gene cluster, and (4) expression of intact biosynthetic pathways devoid of reconstitution. It is believed that *Streptomyces* are more capable of producing natural products in comparison to conventional laboratory strains.[7] The compulsory biosynthetic pathway components were originally uncovered in these strains

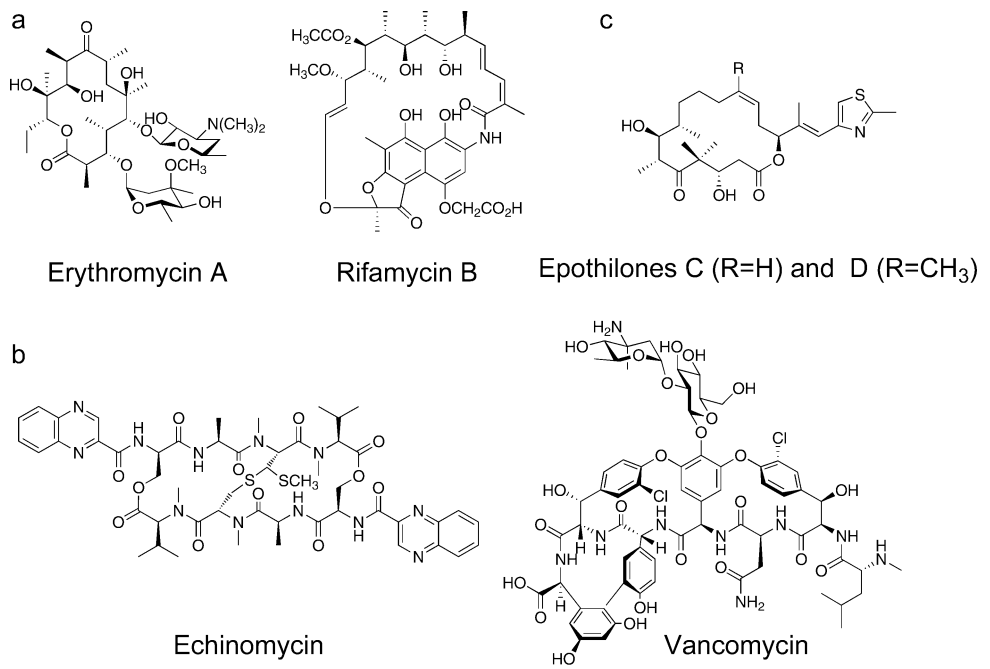


Fig. 1 Chemical structures of (a) PKs, (b) NRPs, and (c) mixed PK-NRP classes of natural products.

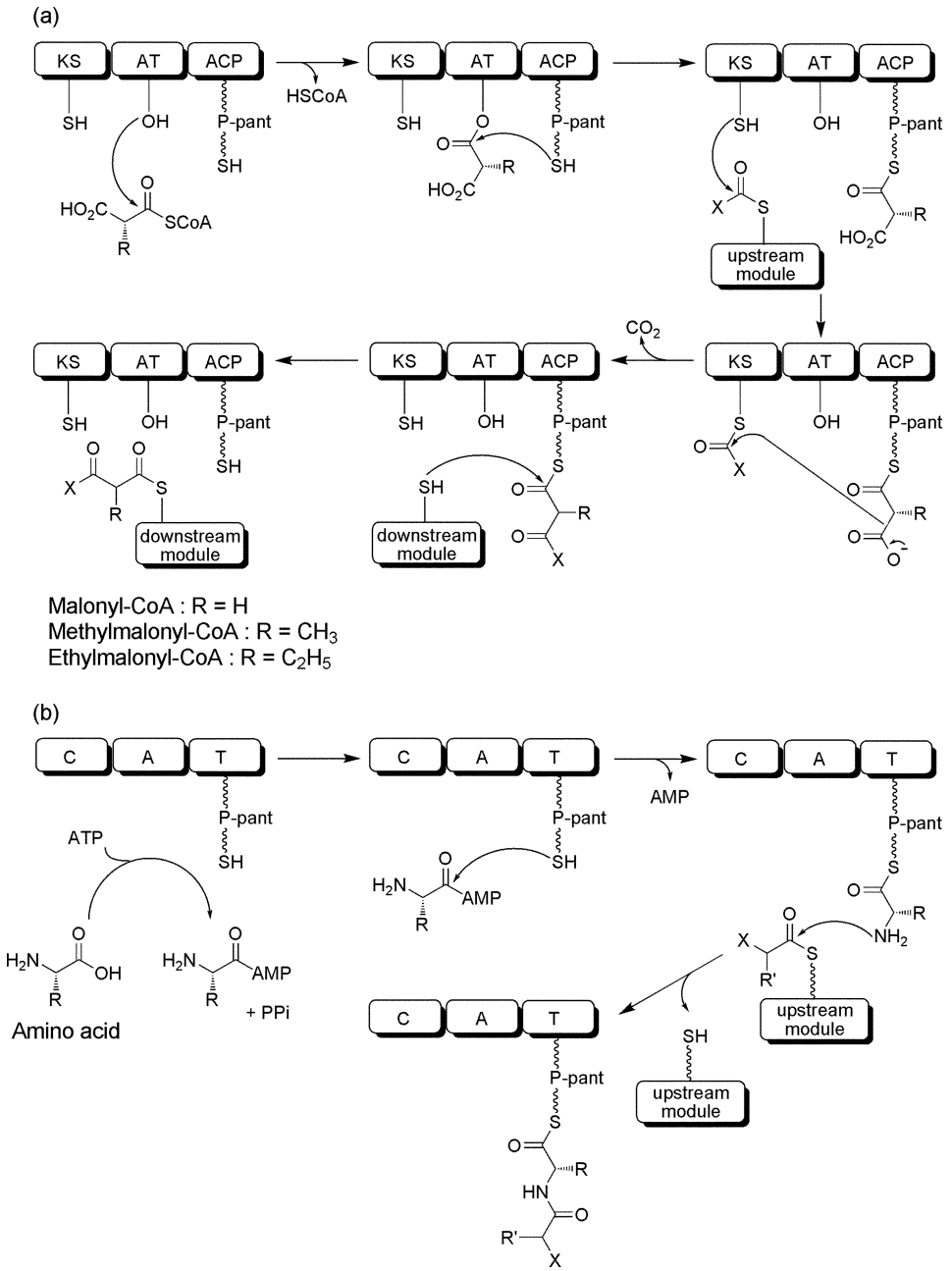


Fig. 2 Modular organization of (a) type I PKS and (b) NRPS. Type I PKS and NRPS modules contain the catalytic domains responsible for acyl and peptide chain elongations, respectively. Hypothetical type I PKS and NRPS modules are shown with core domains. KS, ketosynthase; AT, acyltransferase; ACP, acyl carrier protein; C, condensation; A, adenylation; T, thiolation.

and are known to initiate production of these amazing compounds by manufacturing the starter unit and building blocks, which include propionyl-, methyl- and ethylmolonyl-CoA, necessary to a variety of PKs.[8] Each PKS or NRPS can then be efficiently converted from inactive apo- to phosphopantetheinyl-containing holo-acyl carrier protein (ACP)/peptidyl carrier protein (PCP) forms for elongation of polyketide or polypeptide assembly. Well-characterized *Streptomyces* or model strain, *Streptomyces coelicolor* and *S. lividans*, have been used to produce recombinant polyketides at titers greater than 1 g l^{-1} by means of high density cell fermentation.[9] This was accomplished through the use of a *Streptomyces* expression system where gene clusters responsible for the production of the recombinant PKs were integrated into a chromosome of the host in question. Noteworthy titer may be attributed to the model strain's ability to regulate the production of this secondary metabolite. In recent studies, the complete biosynthetic gene cluster for daptomycin, a lipopeptide antibiotic originally produced by *S. roseosporus*, was successfully inserted into the chromosome of *S. lividans* using a ϕC31 integration system flanked with a phage lambda attL and attR site, a routinely used site for the integration of foreign DNA. In addition, the production of hybrid lipopeptide antibiotics was demonstrated through pathway modification as a method by which to engineer NRP antibiotics.[10,11] Use of *Streptomyces* as a heterologous host

is advantageous, however, there are scores of unidentified genes housed in this bacterium's genome making molecular and cellular manipulation arduous, creating a hurdle for metabolic engineers and slowing advancement in the field of natural products.

To circumvent such a stumbling block, metabolic engineers have resorted to using *E. coli* as a heterologous host, taking advantage of the availability of well-established molecular biological techniques for its genetic and metabolic manipulation. In addition, the ease of fermentation for *E. coli*, makes this organism particularly suitable and attractive for overproduction of natural products. Its robust tolerance of exogenous proteins and fast life cycle will also promote our engineering efforts. The means for large-scale protein production will also facilitate detailed biochemical characterization of the biosynthetic pathway, and conceive a scaffold for a rational approach to optimizing the pathway and modifying the production of exogenous natural products and analog biosynthesis. Khosla and Cane, pioneers of this field, have successfully shown the heterologous production of a complex polyketide natural product 6-deoxyerythronolide B (6dEB), an erythromycin aglycon, in *E. coli* (Fig. 3).[12] Recent studies performed at KOSAN Biosciences, a biotech company have shown optimization and improvement to the already impressive yield of this compound to over one gram per liter using *E. coli* as a heterologous host.[13,14] Armed with the successful heterologous expression of a complex polyketide natural product in *E. coli*, under the merit of utilizing the bacterium as a heterologous host, we will summarize our recent experimental results when the NRP echinomycin gene cluster is transplanted from its natural host into *E. coli* allowing for the development of an *E. coli*-based host system for the unmitigated engineered biosynthesis of natural products.[15]

Erythromycin production in *E. coli*

The erythromycin biosynthetic study was the first example of the heterologous production of complex PKs in *E. coli* and has paved the way for the field of natural products engineering in *E. coli*.[12] These studies have helped scientists decipher the PK biosynthetic pathway by providing crucial pieces to this exceedingly complex puzzle. Heterologous production of 6dEB in *E. coli* entailed the exogenous feeding of propionic acid as a source for propionyl-CoA, the starting unit, which is also converted into (2*S*)-methylmalonyl-CoA, the building unit for the PK core structure. The starting unit, propionyl-CoA, is effectively supplied from propionic acid undergoing a CoA-ligation catalyzed by propionyl-CoA synthetase (PrpE). Catalyzed by propionyl-CoA carboxylase (PCC), propionyl-CoA is carboxylated to afford the building unit, (2*S*)-methylmalonyl-CoA. Genes encoding PrpE and PCC were installed and integrated into *E. coli's* chromosome for expression under control of the T7 promoter.[13] Additionally, the *sfp* gene which codes for a phosphopantetheinyl transferase from *Bacillus subtilis*,[16] responsible for converting the apo- to phosphopantetheinyl-containing holo-ACP conformer for elongation of the polyketide assembly was also included in this chromosome.[12] Furthermore, to avoid squandering these acyl-CoAs by unexpected metabolism for the biosynthesis of undesirable metabolites, the *prp* operon thought to be responsible for propionic acid catabolism in *E. coli* was shut down and the *ygfG* gene coding a methylmalonyl-CoA decarboxylase was deleted from the chromosome of the expression host.[13]

To produce 6dEB in *E. coli*, plasmids carrying the PKS genes responsible for the production of the macrocyclic scaffold were transformed into the engineered bacterium. *E. coli's* ability to tolerate 6dEB's massive biosynthetic gene cluster and capability of producing the heterogeneous metabolite was demonstrated when the resultant strain was furnished with its indispensable substrates. After further optimization, the titer of 6dEB was improved to 1.1 g per liter of fermentation culture.[14] In recent studies, the co-expression of two biosynthetic gene clusters responsible for the production of deoxysugars, TDP-L-mycarose and TDP-D-desosamine, reconstituted operons with the above pathways into *E. coli* provide evidence that a potent antibiotic, erythromycin C is prepared at 0.4 mg l^{-1} of culture.[17] These results have provided

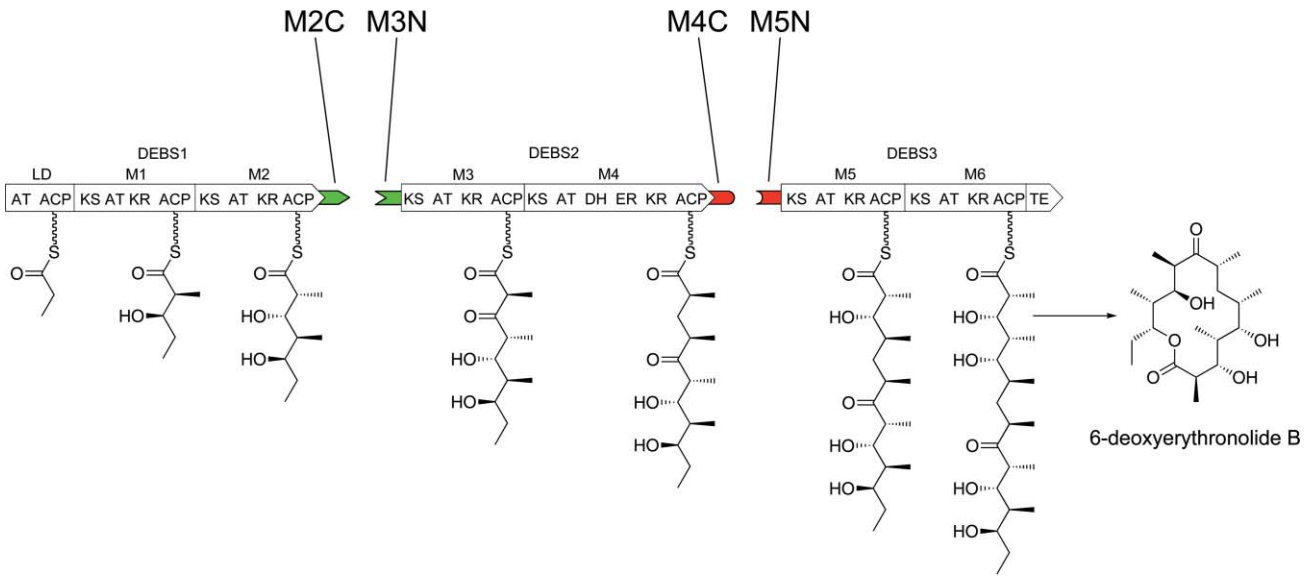


Fig. 3 6-Deoxyerythronolide B biosynthesis. M, module; LD, loading domain; KR, ketoreductase; DH, dehydratase; ER, enoylreductase; TE, thioesterase.

metabolic engineers with an additional arsenal for studying natural products through the expression of complete biosynthetic gene clusters in *E. coli* in place of *Streptomyces*. Moreover, the novel expression system made the detailed mechanism analysis of biosynthetic reactions possible by providing a substantial quantity of biosynthetic enzymes for any, if not all, of the desired mechanistic steps.

Rifamycin intermediate P8/1-OG production in *E. coli*

Rifamycins and other ansamycins are a PK class of natural products containing a benzoic or naphthalenic chromophore bridged by an aliphatic PK chain, which terminates at the chromophore by an amide linkage.[18–21] The biosynthetic pathway for the ansamycin type PK enlists 3-amino-5-hydroxybenzoic acid (AHBA) to serve as a starting unit for the assembly of these molecules.[22–25] In order to study and appreciate the assembly of intermediary rifamycin, 2,6-dimethyl-3,5,7-trihydroxy-7-(3′-amino-5′-hydroxyphenyl)-2,4-heptadienoic acid (P8/1-OG) (Fig. 4), Khosla and co-workers introduced two powerful technologies to express multiple open reading frames (ORFs) in the expansive biosynthetic gene cluster coding megasynthase documented to biosynthesize the PK backbone structure.

To facilitate the reconstitution of a biosynthetic operon for expression in *E. coli*, seven biosynthetic genes responsible for the production of AHBA must be reconstituted for the bacterium capable of governing the expression of this gene cluster.[26] Each ORF calls for a ribosome binding site (rbs) for successful expression in an *E. coli*-translation system and the reconstituted operon requires an available promoter for transcription in *E. coli*. Seven genes from the rifamycin gene cluster were cloned into the pET28 expression vector. By taking advantage of the compatibility between restriction endonucleases *Xba*I and *Spe*I sticky ends to construct a single vector, individual ORF cassettes with rbs (*Xba*I–*Spe*I) were assembled as a polycistronic expression plasmid.

In addition, significant protein expression levels of these hefty modular genes in *E. coli* require that the *rifA*, which codes for the first polypeptide of rifamycin PKS, be stretched across 15 kilobases (kb) of DNA and the molecular weight of posttranslational product is estimated at 0.53 megadalton (MDa). To accommodate expression of the considerable ORF in *E. coli*, engineering *rifA*, the loading domain (LD) and module1 (M1), and M2 and M3, necessitates that they be expressed separately as two polypeptides. The 6dEB synthase (DEBS) was demonstrated to contain two pairs of intermodular linkers: *C*-terminal linker of DEBS M2 (M2C)/*N*-terminal linker of DEBS M3 (M3N) and *C*-terminal linker of DEBS M4 (M4C)/*N*-terminal linker of DEBS M5 (M5N) (Fig. 3).[27–30] These linkers, M2C/M3N and M4C/M5N, are believed to act as a "glue" to augment binding affinity between M2 and M3, and M4 and M5, respectively. It has also been recognized that antiparallel helical interactions between intermodular linkers play a part in the chain transfer of a linear PK from the polypeptide to a downstream polypeptide.[31,32] With regard to the functional reconstitution of RifA protein, a pair of linker M2C and M3N were fused to the C terminus of Rif M1 and the N terminus of Rif M2, respectively, which were then cloned into a single plasmid.

The two plasmids were introduced into an *E. coli* strain, BAP1 which was engineered by Blaine A. Pfeifer for heterologous PK production, described in the erythromycin section.[12] Accordingly, the genetically engineered bacterium was shown to produce the expected intermediate P8/1-OG at a final yield of 2.5 mg l^{-1} of culture. The realization of an *E. coli*-based expression system by the plasmid-borne operon of the complex biosynthetic pathway and functional expression of the mega-gene responsible for producing PK and NRP core structures was confirmed by this reported achievement.

Epothilone production in *E. coli*

Today, epothilone is one of the most prominent natural products as an anticancer agent.[33] In 2000, it was reported that the biosynthetic gene cluster was uncovered by KOSAN Biosciences and successfully expressed by a *Streptomyces*-based system (Fig. 5). The *Streptomyces*-based system was also able to produce the recombinant epothilones C and D at 0.05–0.1 mg l^{-1} of culture.[34] Subsequently, KOSAN Biosciences made an attempt to produce epothilone using a heterologous *E. coli*-based expression system. At that time, two interesting ideas were tested for the expression of the biosynthetic gene cluster in *E. coli*.[35]

To express the biosynthetic gene cluster encoding epothilone's PKS and NRPS, and produce the epothilone core structure, these genes were obtained by the use of chemical synthesis[36]

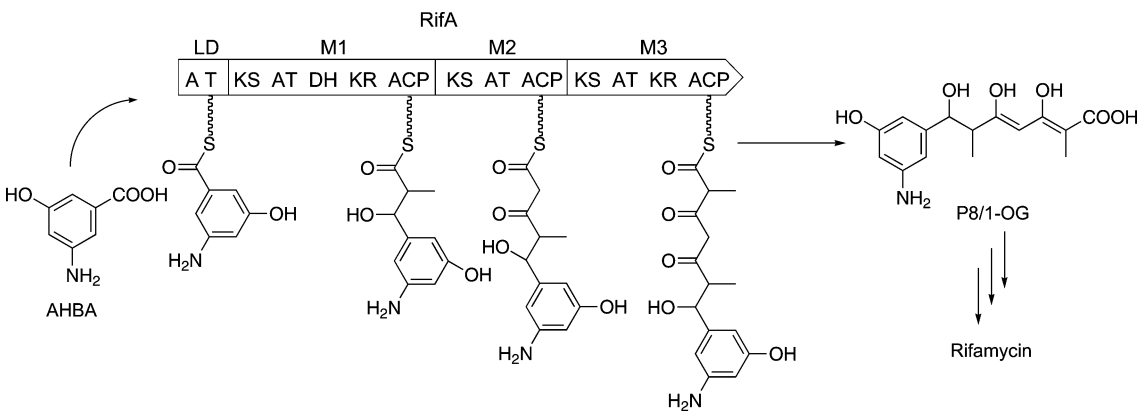


Fig. 4 P8/1-OG biosynthesis.

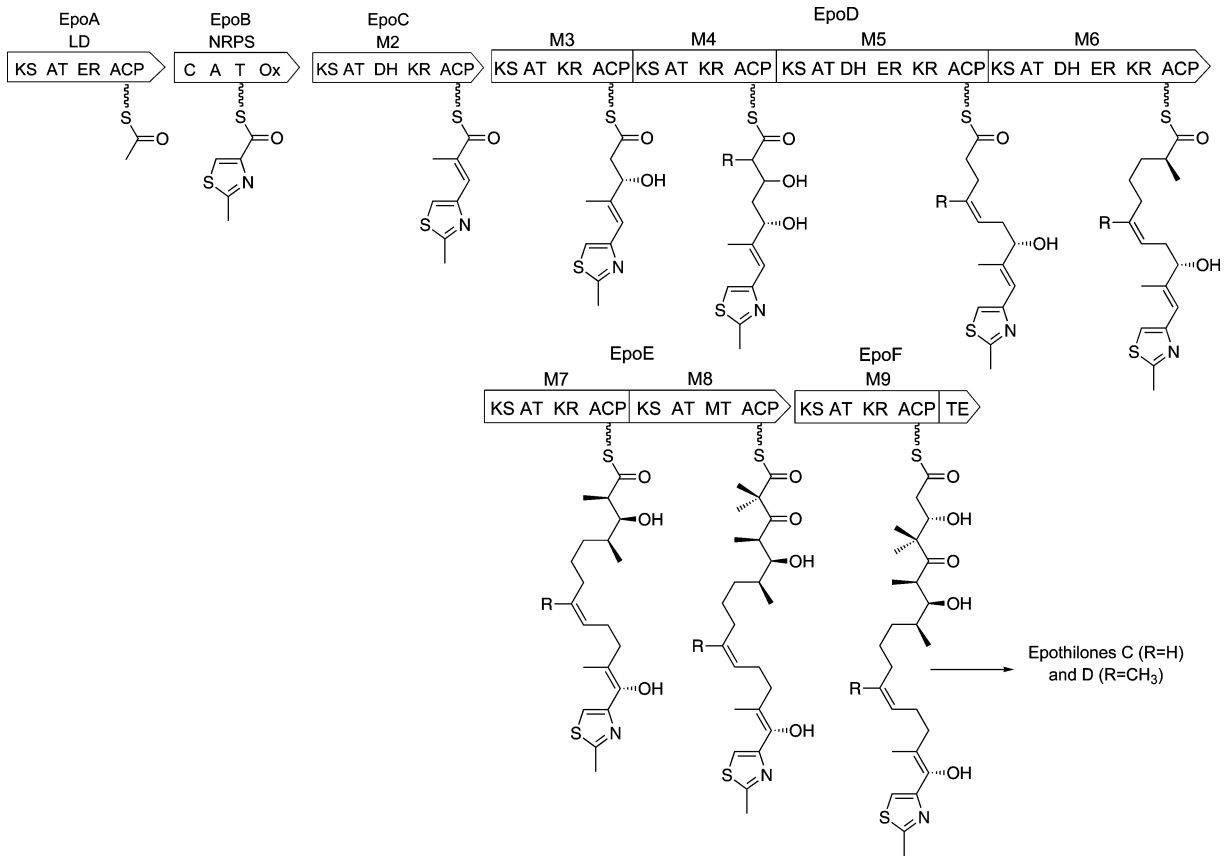


Fig. 5 Epothilone biosynthesis. Ox, oxidase.

and replacement of restriction endonuclease sites in the synthetic genes, which were conveniently cloned into expression plasmids. Low level protein expression for the epothilone biosynthetic genes was observed in *E. coli.* To increase the expression level of these genes, codon usage in the synthetic genes was optimized for *E. coli*. A criterion for optimization of the *E. coli* expression system mandates that these genes code the same amino acids reported for the natural proteins.

An additional approach taken to resolve the low level expression of functional proteins was to express it in concert with chaperones, which could aid more accurate and efficient protein folding and prevent protein precipitation. Because *epoD* is stretched across 22 kb of DNA and the molecular weight of this tetramodular PKS is estimated at 0.76 MDa, *epoD* was expressed as two separated polypeptides, M3/M4 and M5/M6, with appropriate linker regions mirroring the strategy reported for P8/1-OG assembly in *E. coli*.

To produce epothilones C and D in *E. coli*, plasmids carrying the NRPS and PKS genes responsible for the production of the macrocyclic scaffold were transformed into an engineered strain of *E. coli* as described in the erythromycin section. It was observed and reported that the transformant strain was capable of producing epothilones C and D at estimated titers of less than 1 μg l^{-1} of fermentation culture. These results compelled us to further investigate the *de novo* production of biologically active compounds while using *E. coli* as a vehicle.

Proposed mechanism for echinomycin biosynthesis

Echinomycin is an NRP isolated from an assortment of bacteria including *S. lasaliensis* and belongs to the large family of quinomycin antibiotics.[37] As the name implies, this class of natural products possesses a bicyclic aromatic quinoxaline or quinoline chromophore attached to the dimerized cyclic peptide core structure. To identify and isolate the echinomycin biosynthetic gene cluster, a cosmid library was constructed using *S. lasaliensis* total DNA. A sequence of the echinomycin biosynthetic pathway involved in quinoxaline-2-carboxylic acid (QC) and peptide backbone biosyntheses is shown in Fig. 6. Based on past findings, L-tryptophan has been identified as the precursor to QC[38] and on the predicted functions of the proteins encoded within the gene cluster we can hypothesize and construct the QC biosynthetic pathway as described in Fig. 6a. Moreover, we have performed feeding studies by synthesizing (2*S*,3*S*)- and (2*S*,3*R*)-diastereomers of the proposed substrate for Ecm11 of the QC biosynthesis and found that only one substrate, (2*S*,3*S*)-ß-hydroxytryptophan, resulted in QC formation when fed to *S. lasaliensis*.[39] Thus, we were able to establish the absolute configuration of the key intermediate for the biosynthesis of QC and plans to decipher this pathway in its entirety are ongoing.

A glimpse into echinomycin's peptide core revealed that the cluster contains four intact NRPS modules, which are responsible for the assembly of echinomycin. Analysis of the deduced

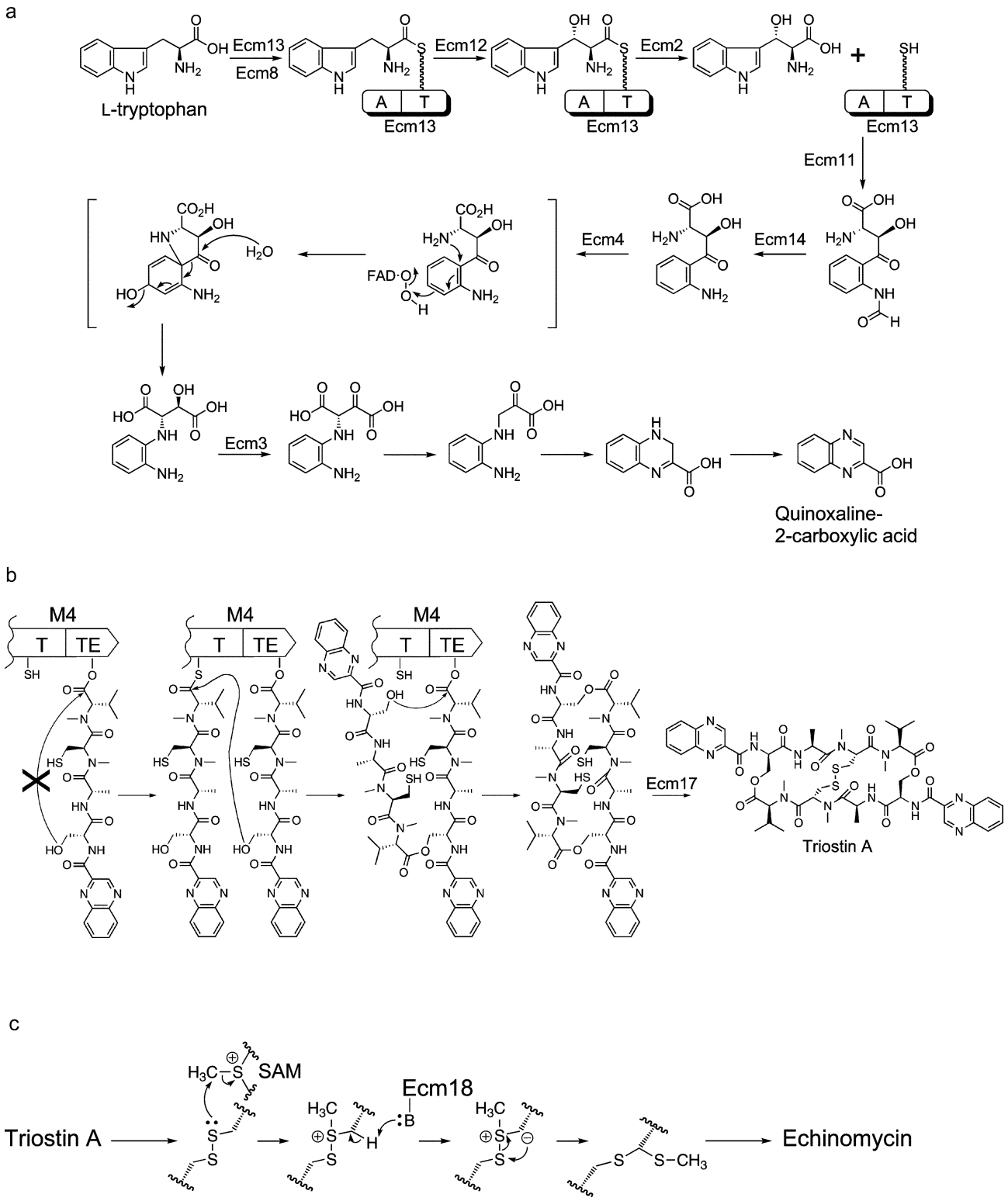


Fig. 6 Proposed pathway for echinomycin biosynthesis and mechanism of thioacetal formation in quinomycin antibiotics. (a) Proposed pathway for quinoxaline-2-carboxylic acid (QC) biosynthesis. (b) Proposed pathway for octadepsipeptide core structure biosynthesis, and proposed mechanism for the homodimerization and cyclorelease of the peptide chain. (c) Proposed mechanism for the thioacetal bridge formation. FabC, fatty acid synthase acyl carrier protein; E, epimerization; M, methyltransferase.

adenylation (A) domain from the nucleotide sequence encoding the NRPS with the aid of a web-based protocol helped to establish that the first, third and fourth A domains are capable of recognizing and ultimately activating L-serine, L-cysteine and L-valine for consecutive addition downstream, respectively. In addition, the epimerization (E) domain in module 1 of the four modules required for the conversion of L- to D-serine, and the methyltransferase (M) domain for *N*-methylcysteine synthesis in module 3 and *N*-methylvaline synthesis in module 4, were also identified and their position in the assembly line deduced. Furthermore, analysis of reported multimeric NRP systems, such as gramicidin S and enterobactin suggest that the echinomycin-synthesizing NRPS system is comprised of four modules and capped off with a thioesterase domain capable of peptide chain homodimerization and cyclorelease (Fig. 6b). Further evaluation of the gene cluster failed to reveal an independent acyl carrier protein in the entire gene cluster. While waiting for the completion and refinement of our sequence, Keller *et al.* presented an article in *J. Biol. Chem.* where the biosynthesis of echinomycin was examined through direct purification of the protein isolated from a different echinomycin producing strain *S. echinatus.*[40] It was demonstrated that the echinomycin biosynthetic pathway recruits the fatty acid ACP for its biosynthesis as a substitute for the putative acyl carrier domain. This result parallels our independent sequencing efforts, confirming the absence of the ACP domain in the cluster.[40] Following results published by Keller's group, the fatty acid ACP from *S. lasaliensis* was isolated and incorporated into our biosynthesis model.

It was proposed over 20 years ago by Waring *et al.* that triostin A is a direct precursor to echinomycin. Undergoing a radical producing reaction, triostin A is interconverted to echinomycin through formation of the unique thioacetal bridge, characteristic of echinomycin.[41] DNA sequence data obtained for the gene cluster resolved *ecm18*, which is highly homologous to a known *S*-adenosyl-L-methionine (SAM)-dependent methyltransferase, which has been thought to convert the disulfide bridge in triostin A to the thioacetal bridge in echinomycin *via* a sulfoniumylide intermediate (Fig. 6c). To verify this, we demonstrated *in vitro* that purified Ecm18 catalyzed the conversion of triostin A to echinomycin in the presence of SAM. This innovative discovery, where a single methyltransferase is capable of biotransforming a disulfide bridge into a thioacetal moiety, is significant.

de novo Biosynthesis of quinomycin antibiotics in *E. coli*

Our main objective here was to engineer *E. coli* for successful heterologous NRP biosynthesis using only glucose and simple salts. This achievement will serve as an important tool and allow for important advancements in the field of natural products. A system such as the one described will establish an *E. coli*-based system for the facile production of essential natural products and their analogs using very simple and economical carbon and nitrogen sources as ingredients.

For *E. coli* production of echinomycin, the feasibility for expression of each of the fifteen *S. lasaliensis* genes in *E. coli* was confirmed. Once confirmed, we then assembled the fifteen genes into three separate plasmids allowing each gene to carry its own T7 promoter and ribosome-binding site. We created two restriction endonuclease sites (*Xba* I and *Spe* I) in a pET vector system placing the *Xba* I site on the 5′ side of the promoter and the *Spe* I site on the 3′ side of the terminator. Both *Xba* I and *Spe* I produce compatible overhangs, hence, a "promoter-gene-terminator" cassette that can ligate into a vector carrying another "promoter-gene-terminator" cassette (s) can be prepared by *Xba* I-*Spe* I double digestion of the cassette and either *Xba* I or *Spe* I digest of the vector (Fig. 7). To ensure the stable retention of all three plasmids in *E. coli*, we decided to use the orthogonal

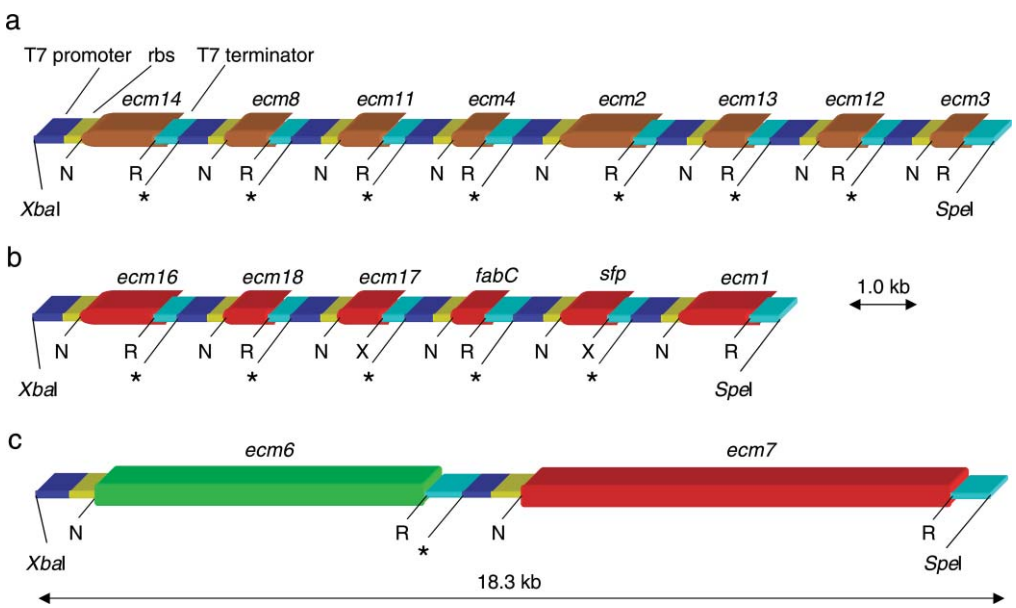


Fig. 7 Organization of the plasmid-borne echinomycin biosynthetic gene cluster. (a) Map of plasmid for the production of the QC biosynthetic proteins. (b) Map of plasmid for the production of NRPSs. (c) Map of plasmid for the production of the auxiliary proteins. N, *Nde* I; E, *Eco*R I; X, *Xho* I; *, ligation product between cohesive *Xba* I and *Spe* I recognition sites; rbs, ribosome binding site.

origins of replication and three antibiotic resistance genes. Our approach requires multiple monocistronic gene cassettes, rather than a single polycistronic gene cassette for the assembly of the multi-gene plasmid constructs. This approach simplifies plasmid construction and re-construction because it eliminates the need to clone every gene in the construct in the same transcriptional direction. Furthermore, our approach of transcribing each gene independently has helped to achieve similar protein production levels despite the many genes that were incorporated into our plasmids for heterologous expression in *E. coli*. We attribute these desirable findings to the abundance of mRNA transcribed from the polycistronic gene cassette leading to the incomplete synthesis of mRNA and/or unstable mRNA. Also, we believe that our approach has improved our chances and ability to incorporate an even greater number of heterologous genes into *E. coli* in a simple and reliable fashion.

Subjecting *E. coli* strain BL21 (DE3), carrying our three plasmids, to an 8-day-long fed-batch fermentation in M9 minimal medium, we were able to purify echinomycin from the culture extract through a series of chromatographic steps to obtain a final echinomycin yield of 0.3 mg l^{-1} of culture. Furthermore, to demonstrate the ease and effectiveness of modifying the *E. coli*-based heterologous biosynthetic system, we chose conversion of the echinomycin biosynthetic pathway into a triostin A biosynthetic pathway by simply modifying our plasmids and removing *ecm18*. The further modified strain produced the expected compound, triostin A at a yield of 0.6 mg l^{-1} of culture. Unlike the chemical synthesis, our system permits one-pot, hands-free, and easily scalable bioproduction of complex natural products and their analogs in less than two weeks, omitting the need for highly trained organic chemists possessing sophisticated laboratory expertise. Relative to previous attempts at developing an *E. coli*-based heterologous host system for biosynthesis of natural products, our results are comparable if not superior to published results. In fact, we believe that our result was significant for the production of such complex compounds using fermentation conditions that have not yet been optimized.

Based on our past observations, we expect the efficiency of such a demanding transcription and the stability of a lengthy mRNA to be rather low when a polycistronic system is used for ancillary PKS and NRPS systems. In fact, the disproportionate expression level of each of the genes was observed as expected. This may be accredited to the fact that genes located further downstream from its promoter typically reduce expression levels. Aside from transcriptional and translational efficiency issues, the construction of such plasmids is very demanding and requires technical know how. On top of the hurdles of generating plasmids carrying multi-gene inserts, if one is to create a polycistronic cassette, one needs to worry about the orientation of the insert, which can be difficult to control. The efficiency of procuring constructs with unidirectional inserts in the same reading-frame decreases rapidly as the number of genes to be inserted increases. Use of multi-monocistronic cassettes can eliminate the need to control the gene's orientation and significantly improve the efficiency of the plasmid construction. Our results confirmed that replacement of the polycistronic cassette with a multi-monocistronic cassette can consistently exhibit high-level expression of every gene in our constructs and ameliorate plasmid construction. We believe these improvements have made the *de novo* biosynthesis of complex natural products possible.

When introducing any exogenous biosynthetic pathway into *E. coli*, the toxicity of the biosynthetic product can impair the host. This problem can be circumvented by introducing a self-resistance mechanism conferring resistance without degrading our target compound. While considering the echinomycin biosynthetic pathway, the homology of Ecm16 and daunorubicin resistance-conferring factor DrrC,[42,43] and the similarity between their mode of action for echinomycin and daunorubicin, suggests that Ecm16 bequeaths *S. lasaliensis* with a non-destructive resistance against echinomycin. We were able to demonstrate that *ecm16* conferred echinomycin resistance to BL21 (DE3). Also, omission of *ecm16* hampered the growth of the host, suggesting that sufficient amounts of echinomycin and triostin A would have been unattainable without the aid of *ecm16*, an echinomycin-resistance mechanism.

Conclusion and future directions

Unparalleled by any current reports, we have successfully developed a heterologous host, which is capable of biosynthesizing a variety of complex bioactive NRPs using only simple carbon and nitrogen sources. We anticipate our system will pave the way for the development of a general yet economical platform for the one-pot mass production of therapeutically beneficial natural products and their analogs. In addition, our system has unlocked the possibility of the expedient and trivial introduction of even larger and more complex biosynthetic pathways into *E. coli*. Our plasmid-based pathway is more manageable and readily modifiable for engineering purposes, as demonstrated in our report.

Future directions will include the use of a chromosome-based system in *E. coli*. This strategy will require us to integrate our gene cluster of interest minus the genes or a part of the cluster responsible for scaffold or post assembly line modification for producing rationally designed analogs.[44] Establishing a chromosome expression system will do away with the maximum number of genes that is acceptable for heterologous introduction, concerns over the stability of these genes, and other minute barriers commonly associated with using a plasmid expression system. The proposed direction is realistic as it was tried to an extent with the construction of an *E. coli* strain housing genes essential for the biosynthesis of 6-deoxyerythronolide B, at production levels of 1 g l^{-1} of culture in the absence of any genetic engineering. However, integration of the biosynthetic pathway in its entirety *via* a chromosome obscures and retards any attempts at engineering the pathway.

Natural products of interest developed to target life threatening bacteria and malignant tumors will serve as valuable agents to treat these infections and diseases once they are subjected to scrutiny and meet all the criteria set forth by the Food and Drug Administration. However, because the biosynthetic pathways for these compounds are commonly foreign metabolites to *E. coli*, certain genetic implementations are necessary to make the heterologous host an efficient secondary metabolite-producing organism. Principally, care must be taken to introduce self-resistance mechanisms into the bacterium if the natural product being produced is likely to be lethal for *E. coli*. This toxicity can be avoided by identifying the mechanism of self-resistance in the original producing host and supplementing the same or a similar system into *E. coli*. Marahiel and co-workers have reported

the successful engineering of an ABC transporter system into *Bacilllus subtilis* for the safe heterologous expression of the NRPS antibiotic bacitracin.[45] However, we anticipate that use of Gram-positive *Streptomyces* ABC transporters would prove futile in Gram-negative *E. coli* due to the presence of an additional outer membrane. To circumvent this problem, we will take advantage of the ABC transporter system endogenous to *E. coli*, the AcrAB-TolC efflux system.[46] This AcrAB system has provided *E. coli* with the most effective resistance against a broad range of organic compounds. Use of this efflux pumping system could increase or broaden the spectrum of compounds *E. coli* is capable of producing and sustain the bacterium's viability thereby bettering its titer to levels that can benefit and lower production costs for biotech companies.

Acknowledgements

This work was funded by Grant-in-Aids for Scientific Research from the Japan Society for the Promotion of Science (A) 17208010 (H. O.). K. W. is the recipient of an award in Japan Synthetic Organic Chemistry (Daiichi Pharmaceutical Co. Ltd.) and from the Japan Society for Bioscience, Biotechnology, and Agrochemistry Research Foundation.

References

1 G. M. Cragg, D. J. Newman and K. M. Snader, *J. Nat. Prod.*, 1997, **60**, 52.
2 S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson and L. Katz, *Science*, 1991, **252**, 675.
3 J. Cortes, S. F. Haydock, G. A. Roberts, D. J. Bevitt and P. F. Leadlay, *Nature*, 1990, **348**, 176.
4 D. E. Cane, C. T. Walsh and C. Khosla, *Science*, 1998, **282**, 63.
5 C. T. Walsh, *Science*, 2004, **303**, 1805.
6 D. J. Newman, G. M. Cragg and K. M. Snader, *J. Nat. Prod.*, 2003, **66**, 1022.
7 T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Chater and D. A. Hopwood, *Practical Streptomyces Genetics*, The John Innes Foundation: Norwich, England, 2000.
8 C. R. Hutchinson, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 3010.
9 R. P. Desai, T. Leaf, E. Woo and P. Licari, *J. Ind. Microbiol. Biotechnol.*, 2002, **5**, 297.
10 P. Julia, X. Li, A. Whiting, M. Latif, T. Gibson, C. J. Silva, P. Brain, J. Davies, V. Miao, S. K. Wrigley and R. H. Baltz, *J. Ind. Microbiol. Biotechnol.*, 2006, **33**, 121.
11 V. Miao, M. C. Gal, K. Nguyen, P. Brian, J. Penn, A. Whiting, J. Steele, D. Kau, S. Martin, R. Ford, T. Gibson, M. Richard and H. Baltz, *Chem. Biol.*, 2006, **13**, 269.
12 B. A. Pfeifer, S. J. Admiraal, H. Gramajo, D. E. Cane and C. Khosla, *Science*, 2001, **291**, 1790.
13 S. Murli, J. Kennedy, L. C. Dayem, J. R. Carney and J. T. Kealey, *J. Ind. Microbiol. Biotechnol.*, 2003, **30**, 500.
14 J. Lau, C. Tran, P. Licari and J. Galazzo, *J. Biotechnol.*, 2004, **110**, 95.
15 K. Watanabe, K. Hotta, A. P. Praseuth, K. Koketsu, A. Migita, C. N. Boddy, C. C. C. Wang, H. Oguri and H. Oikawa, *Nat. Chem. Biol.*, 2006, **2**, 423.
16 R. H. Lambalot, A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, C. Khosla and C. T. Walsh, *Chem. Biol.*, 1996, **3**, 923.
17 S. Peiru, H. G. Menzella, E. Rodriguez, J. Carney and H. Gramajo, *Appl. Environ. Microbiol.*, 2005, **71**, 2539.
18 M. G. Anderson, J. J. Kibby, R. W. Rickards and J. M. Rothchild, *J. Chem. Soc., Chem. Commun.*, 1980, 1277.
19 O. Ghisalba and J. Nuesch, *J. Antibiot.*, 1981, **34**, 64.
20 A. M. Becker, A. J. Herlt, G. L. Hilton, J. J. Kibby and R. W. Rickards, *J. Antibiot.*, 1983, **36**, 1323.
21 P. Traxler and O. Ghisalba, *J. Antibiot.*, 1982, **35**, 1361.
22 O. Ghisalba, *Chimia*, 1985, **39**, 79.
23 H. G. Floss and J. M. Beale, *Angew. Chem., Int. Ed. Engl.*, 1989, **28**, 146.
24 H. G. Floss, *Nat. Prod. Rep.*, 1997, **14**, 433.
25 D. Hunziker, T. W. Yu, C. R. Hutchinson, H. G. Floss and C. Khosla, *J. Am. Chem. Soc.*, 1998, **120**, 1092.
26 K. Watanabe, M. A. Rude, C. T. Walsh and C. Khosla, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 9774.
27 R. S. Gokhale, S. Y. Tsuji, D. E. Cane and C. Khosla, *Science*, 1999, **284**, 482.
28 S. Y. Tsuji, D. E. Cane and C. Khosla, *Biochemistry*, 2001, **40**, 2326.
29 N. Wu, S. Y. Tsuji, D. E. Cane and C. Khosla, *J. Am. Chem. Soc.*, 2001, **123**, 6465.
30 N. Wu, D. E. Cane and C. Khosla, *Biochemistry*, 2002, **41**, 5056.
31 P. Kumar, Q. Li, D. E. Cane and C. Khosla, *J. Am. Chem. Soc.*, 2003, **125**, 4097.
32 Y. Tang, C. Y. Kim, I. I. Mathews, D. E. Cane and C. Khosla, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 11124.
33 K. Gerth, N. Bedorf, G. Höfle, H. Irschik and H. Reichenbach, *J. Antibiot.*, 1996, **49**, 560.
34 L. Tang, S. Shah, L. Chung, J. Carney, L. Katz, C. Khosla and B. Julien, *Science*, 2000, **287**, 640.
35 S. C. Mutka, J. R. Carney, Y. Liu and J. Kennedy, *Biochemistry*, 2006, **45**, 1321.
36 S. J. Kodumal, K. G. Patel, R. Reid, H. G. Menzella, M. Welch and D. V. Santi, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 15573.
37 N. Steinerova, H. Lipavska, K. Stajner, J. Caslavska, M. Blumauerova, J. Cudlin and Z. Vanek, *Folia Microbiol. (Praha)*, 1987, **32**, 1.
38 D. G. Reid, D. M. Doddrell, D. H. Williams and K. R. Fox, *Biochim. Biophys. Acta*, 1984, **798**, 111.
39 K. Koketsu, H. Oguri, K. Watanabe and H. Oikawa, *Org. Lett.*, 2006, **8**, 4719.
40 G. Schmoock, F. Pfenning, J. Jewiarz, W. Schlumbohm, W. Laubinger, F. Schauwecker and U. Keller, *J. Biol. Chem.*, 2005, **280**, 4339.
41 A. Cornish, M. J. Waring and R. D. Nolan, *J. Antibiot. (Tokyo)*, 1983, **36**, 1664.
42 N. Lomovskaya, S. K. Hong, S. U. Kim, L. Fonstein, K. Furuya and C. R. Hutchinson, *J. Bacteriol.*, 1996, **178**, 3238.
43 K. Furuya and C. R. Hutchinson, *FEMS Microbiol. Lett.*, 1998, **168**, 243.
44 H. G. Menzella, R. Reid, J. R. Carney, S. S. Chandran, S. J. Reisinger, K. G. Patel, D. A. Hopwood and D. V. Santi, *Nat. Biotechnol.*, 2005, **23**, 1171.
45 K. Eppelmann, S. Doekel and M. A. Marahiel, *J. Biol. Chem.*, 2001, **276**, 34824.
46 H. Nikaido, *J. Bacteriol.*, 1996, **20**, 5853.

