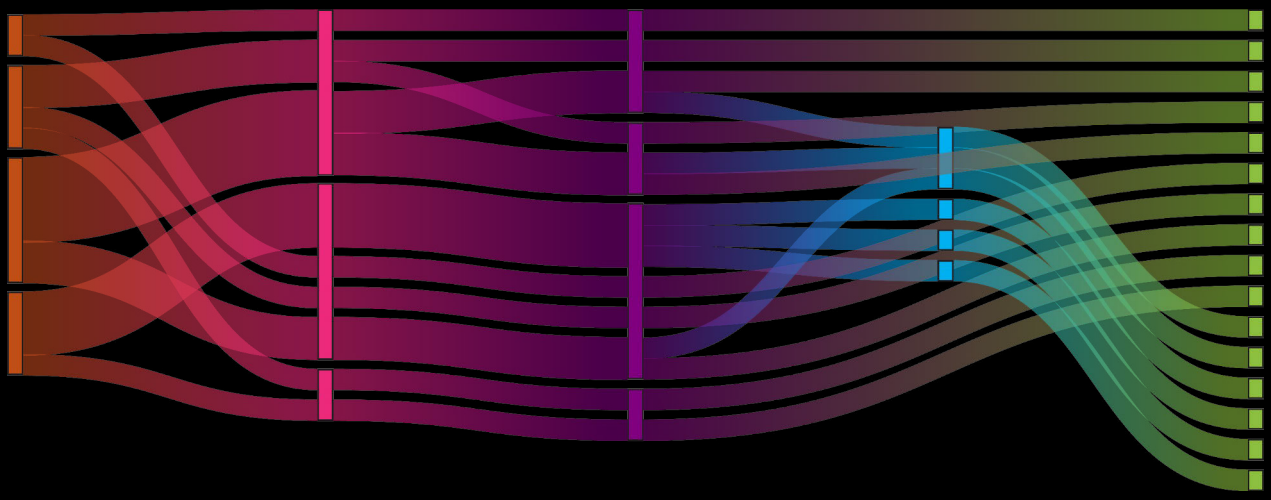
TURUN
YLIOPISTO
UNIVERSITY
OF TURKU



SYNTHETIC BIOLOGY PLATFORM FOR PRODUCTION OF NOVEL ANTHRACYCLINE ANTICANCER AGENTS

Rongbin Wang

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AI OSA – TOM. 743 | ASTRONOMICA – CHEMICA – PHYSICA – MATHEMATICA | TURKU 2025

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

SYNTHETIC BIOLOGY PLATFORM FOR PRODUCTION OF NOVEL ANTHRACYCLINE ANTICANCER AGENTS

Rongbin Wang

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AI OSA – TOM. 743 | ASTRONOMICA – CHEMICA – PHYSICA – MATHEMATICA | TURKU 2025

University of Turku

Faculty of Technology
Department of Life Technologies
Biochemistry
Doctoral Programme in Technology

Supervised by

Professor, Mikko Metsä-Ketelä
Department of Life Technologies
University of Turku

Docent, Jarmo Niemi
Department of Life Technologies
University of Turku

Reviewed by

Research Professor, Yvonne Nygård
VTT Technical Research Centre of Finland Ltd.

Docent, Suvi Santala
Faculty of Engineering and Natural Sciences
Tampere University

Opponent

Professor, Eriko Takano
Department of Chemistry
University of Manchester

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

Cover image: Rongbin Wang

ISBN 978-952-02-0268-2 (PRINT)
ISBN 978-952-02-0269-9 (PDF)
ISSN 0082-7002 (Print)
ISSN 2343-3175 (Online)
Painosalama, Turku, Finland 2025

UNIVERSITY OF TURKU
Faculty of Technology
Department of Life Technologies
Biochemistry
RONGBIN WANG: Synthetic Biology Platform for Production of Novel Anthracycline Anticancer Agents
Doctoral Dissertation, 177 pp.
Doctoral Programme in Technology
April 2025

ABSTRACT

Anthracyclines, microbial natural products primarily produced by soil-dwelling *Streptomyces* bacteria, are potent chemotherapeutic agents used clinically worldwide. However, despite their potent antiproliferative activities, their therapeutic application is limited by irreversible cardiotoxicity. Furthermore, their stereochemical complexity hinders the discovery of improved semi-synthetic analogs. Synthetic biology offers a promising alternative for modification of complex natural products. However, this approach is constrained by the lack of a well-established synthetic biology platform tailored for anthracycline biosynthesis.

To address these issues, this doctoral thesis establishes a complete synthetic biology platform designed for anthracycline engineering, consisting of an amenable *Streptomyces* chassis, modular expression vectors, promoters with tunable strengths, efficient terminators, and a curated library of anthracycline biosynthetic genes. To systematically explore anthracycline production, the biosynthetic pathway was redesigned and assembled into four distinct functional modules: (i) polyketide aglycones, (ii) TDP-carbohydrate, (iii) self-resistance and glycosyltransferases, and (iv) post-PKS tailoring, using a standardized BioBricks approach.

Initially, the platform was employed to achieve the first complete biosynthesis of three clinically relevant anthracyclines—nogalamycin, doxorubicin, and aclacinomycin. Furthermore, the modular design facilitated combinatorial biosynthesis through strategic mixing and matching of biosynthetic modules. This approach led to the production of 16 anthracycline derivatives, 13 of which are entirely new. The streamlined biosynthetic process and balanced expression enabled efficient downstream purification from small-scale fermentations, which facilitated structural characterization by nuclear magnetic resonance and cytotoxicity test on human cell lines, leading to a discovery of six high potency anthracyclines and giving an insight on structure-activity-relationship of anthracyclines.

Overall, this thesis demonstrates the utility of a modular synthetic biology platform for the rational biosynthesis of natural products and the discovery of novel anthracycline analogs, offering a powerful strategy for expanding the structural diversity and therapeutic potential of this important class of anticancer agents.

KEYWORDS: synthetic biology; anticancer agent; anthracycline; BioBricks; complete biosynthesis; combinatorial biosynthesis

TURUN YLIOPISTO
Teknillinen tiedekunta
Bioteknologian laitos
Biokemia
RONGBIN WANG: Synteettisen biologian työkalusarja uusien antrasykliinien syöpälääkkeiden tuotantoon
Väitöskirja, 177 s.
Teknologian tohtoriohjelma
Huhtikuu 2025

TIIVISTELMÄ

Antrasykliinit ovat maaperässä esiintyvien Streptomykeetti-bakteerien tuottamia luonnonyhdisteitä. Antrasykliinejä käytetään syövän hoidossa maailmanlaajuisesti, mutta niiden käytössä on rajoituksia annoksesta riippuvan sydäntoksisuuden takia. Vähemmän toksisia antrasykliinien johdannaisia on vaikea tuottaa synteettisen kemian keinoin antrasykliinien monimutkaisen stereokemian takia. Synteettinen biologia toimii paremmin luonnonyhdisteiden johdannaisten kehittämiseen, mutta antrasykliinien biosynteesin muokkauksen vaatimat strandardoidut työkalut puuttuvat.

Edellämainittujen ongelmien ratkaisemiseksi tässä väitöskirjatutkimuksessa luodaan kokonainen synteettisen biologian työkalusarja antrasykliinien biosynteesin muokkaukseen. Tämä työkalusarja perustuu standardisoituihin BioBrick-osiin ja sisältää tuottoisäntänä toimivan Streptomykeetti-bakteerin modulaariset ekspressiovektorit, promoottorit säädettävällä vahvuudella, tehokkaat terminaattorit sekä tärkeimmät antrasykliinien biosynteettiset geenit. Näitä työkaluja käyttäen antrasykliinien biosynteesi uudelleenjärjesteltiin neljäksi eri moduuliksi: (i) polyketidien aglykonit, (ii) TDP-sokerit, (iii) antibioottiresistenssi ja glykosyylitransferaasit ja (iv) polyketidin synteesin jälkeiset muokkaukset.

Työkalusarjan avulla rakennettiin ensimmäistä kertaa kokonainen biosynteesireitti kolmelle tärkeälle syöpälääkkeelle: nogalamysiinille, doksorubisiinille ja aklasinomysiinille. Tämän lisäksi työkalusarjaa onnistuttiin käyttämään kombinatoriseen biosynteesiin yhdistelemällä eri biosynteettisiä moduuleja toisiinsa. Kombinatorisen biosynteesin tuloksena saatiin luotua 16 antrasykliinien johdannaista, joista 13 ovat uusia yhdisteitä. Näiden yhdisteiden tuotto onnistui tehokkaasti pienen mittakaavan fermentoreissa, mikä mahdollisti niiden kemiallisen rakenteen ratkaisemisen sekä bioaktiivisuuden testaamisen. Kaiken kaikkiaan näistä yhdisteistä kuusi osoittautui tehokkaiksi syöpäsoluja vastaan, ja bioaktiivisuustestien perusteella voitiin myös päätellä minkälaiset kemialliset rakenteet lisäävät yhdisteiden tehoa syöpäsoluja vastaan.

Yhteenvetona väitöskirjatutkimuksessa osoitetaan modulaarisen synteettisen biologian hyöty luonnonyhdisteiden johdannaisten kehittämisessä. Väitöskirjatutkimuksessa kehitetyn synteettisen biologian työkalusarjan avulla onnistuttiin tuottamaan 13 uutta antrasykliinien johdannaista. Tulevaisuudessa tätä

työkalusarjaa voidaan hyödyntää myös muiden luonnonyhdisteiden johdannaisten valmistuksessa.

ASIASANAT: synteettinen biologia, syöpälääke, antrasykliini, BioBricks, kombinatorinen biosynteesi

Table of Contents

Abbreviations

ABC	ATP-Binding Cassette
ACP	Acyl Carrier Protein
BAC	Bacterial Artificial Chromosome
BGC	Biosynthetic Gene Cluster
CoA	Coenzyme A
CRISPR/Cas	Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR Associate Protein
crRNA	CRISPR RNA
DNA	Deoxyribonucleic Acid
DM-DXR	N,N-dimethyl-doxorubicin
EC_{50}	Half Maximal Effective Concentration
FAC	Fungal Artificial Chromosome
G-C	Guanine-Cytosine
GT	Glycosyltransferase
HR-MS	High Resolution-Mass Spectrometry
IR	Integrase-Mediated Recombination
minPKS	Minimum Polyketide Synthase
NAD(P)H	Nicotinamide Adenine Dinucleotide (Phosphate) Hydrogen
NMR	Nuclear Magnetic Resonance
NP	Natural Product
PAC	Phage Artificial Chromosome
PTM	Polycyclic Tetramate Macrolactam
RBS	Ribosome Binding Site
RNA	Ribonucleic Acid
SAR	Structure-Activity Relationship
SARP	*Streptomyces* Antibiotic Regulatory Protein
TAR	Transformation-Associated Recombination

TDP	Thymidine Diphosphate
UHPLC	Ultra High Performance Liquid Chromatography
YAC	Yeast Artificial Chromosome

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

I Wang R[#], Nguyen J[#], Hecht J, Schwartz N, Brown K, Ponomareva L, Niemczura M, van Dissel D, van Wezel G, Thorson J, Metsä-Ketelä M*, Shaaban K*, Nybo S*. A BioBricks metabolic engineering platform for the biosynthesis of anthracyclinones in *Streptomyces coelicolor*. ACS Synthetic Biology, 2022; 11, 12: 4193–4209.

II Wang R[#], Wandi B[#], Schwartz N[#], Hecht J[#], Ponomareva L, Paige K, West A, Desanti K, Nguyen J, Niemi J, Thorson J, Shaaban K*, Metsä-Ketelä M*, Nybo S*. Diverse combinatorial biosynthesis strategies for C–H functionalization of anthracyclinones. ACS Synthetic Biology, 2024; 13, 5: 1523–1536.

III Wang R, Tirkkonen H, van der Zanden S, Ponomareva L, Fagon I, Fève Leslie, Paige K, Brown C, Schwartz N, Hecht J, Thorson J, Neefjes J, Nybo S*, Shaaban K*, and Metsä-Ketelä M*. Complete pathway refactoring and combinatorial biosynthesis of anthracyclines. Unpublished, 2025.

[#]Equal contribution
*Corresponding authors

1 Introduction

1.1 *Streptomyces* and anthracyclines

Micro-organisms are a prolific source of bioactive natural products (NPs), accounting for around two-thirds of antibiotics and one-third of chemotherapeutic agents in clinical use[1]. In particular, soil-dwelling *Streptomyces* bacteria are a rich source of chemical diversity and NPs originating from secondary metabolites[2]. The genus *Streptomyces* is a group of aerobic, Gram-positive, multicellular, and filamentous bacteria with complex life cycles[3]. *Streptomyces* exhibit a large linear chromosome with an average size of 8.4 Mbp and high G-C content DNA[4], where typically more than 20 secondary metabolic pathways are hidden[5,6]. Secondary metabolites are non-essential for microbial growth, development, or reproduction, but provide ecological advantages such as defense, signaling, or competition[7]. A great example of secondary metabolites derived primarily from *Streptomyces* is the clinically used anti-cancer anthracyclines[1]. Anthracyclines are a class of type II aromatic polyketides and chemically consist of a linear tetracyclic 7,8,9,10-tetrahydro-5,12-naphthacenoquinone carbon scaffold with one or more carbohydrate moieties decorated[8]. The first anthracycline, rhodomycin B, was discovered in the 1950s[9], and over five hundred anthracyclines have been isolated to date (**Figure 1**)[10]. Anthracyclines have been prescribed to more than one million patients per year, among which doxorubicin, daunorubicin, and aclacinomycin A (aclarubicin) have been widely used to treat both solid and hematological tumors[11,12].

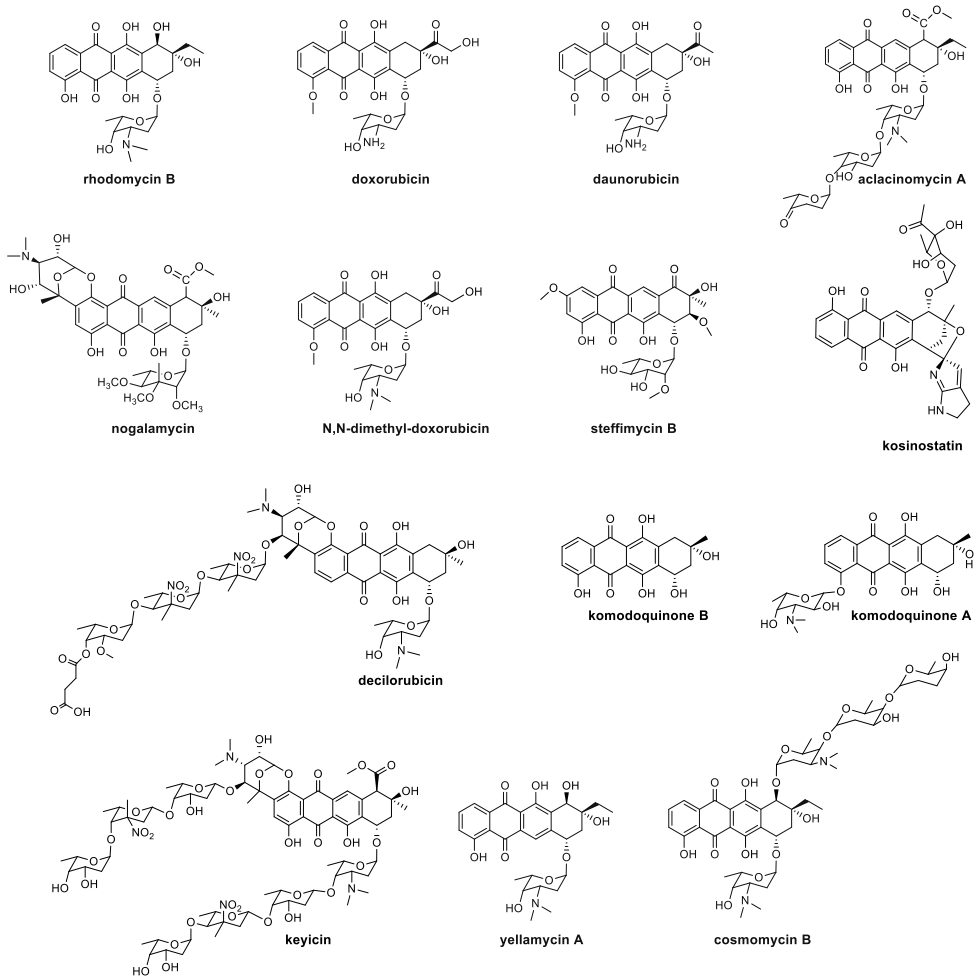


Figure 1. Structures of common anthracyclines.

Anthracyclines are highly efficient anti-proliferative agents, but they can cause accumulative and irreversible cardiotoxicity, which limits the amount that can be safely administered over the lifetime of a patient[13]. Therefore, there is an urgent demand and great interest in searching for novel anthracyclines with improved therapeutic indices[14]. The cardiotoxicity of doxorubicin-like anthracyclines could result from two mechanisms: topoisomerase II–mediated DNA double-strand breaks and chromatin damage via histone eviction at specific genomic loci[15]. Fortunately, it has been demonstrated that anthracycline-induced cardiotoxicity is caused by the combined impact of DNA damage and chromatin disruption[16,17]. Further modifications to anthracyclines have effectively uncoupled the two bioactivities, leading to the discovery of potentially cardiotoxicity-free chemotherapy agents, such

as the semi-synthetic anthracycline N,N-dimethyl-doxorubicin (DM-DXR)[15]. DM-DXR is currently progressing towards clinical trials and is generated by organic semi-synthesis from doxorubicin, which is known to be less efficient and less environmentally friendly compared to biosynthetic manufacturing[15]. To accomplish the biosynthesis of novel anthracyclines, the advantages of synthetic biology and a comprehensive understanding of anthracycline biosynthetic pathways should be leveraged, which will be reviewed in the following sections.

1.2 Biosynthesis of anthracyclines

Anthracycline biosynthesis is directed by biosynthetic gene clusters (BGCs), which are groups of genes encoding the enzymes and proteins required for the synthesis of various molecules[8]. BGCs of anthracyclines usually contain a large number of biosynthetic genes due to the complexity of the chemical structures[8]. These genes can be divided into several groups based on their functions, including biosynthetic genes for polyketide aglycones, carbohydrates, glycosyltransferases, post-PKS tailoring, and self-defense genes (transporters and DNA damage repair genes)[18]. In addition, anthracycline BGCs commonly contain several regulatory genes, which tightly regulate the expression of the pathways[19]. Finally, a few genes of unknown function typically co-exist in anthracycline BGCs[18]. The details of each group will be discussed in the following sections.

1.2.1 Biosynthesis of anthracycline aglycones

Anthracyclines are a group of type II aromatic polyketides that share similar core carbon scaffolds[8]. These scaffolds are initially synthesized by type II minimum polyketide synthase (minPKS) to form enzyme-linked decaketide intermediates[20]. The minPKS usually contains a ketoacyl synthase (KSa), a chain length factor (KSb), and an acyl carrier protein (ACP), which condense an acetyl-CoA or propionyl-CoA starter unit with nine malonyl-CoA units through nine rounds of iterative Claisen condensations[18]. In **Figure 2**, I present the schemes of the two most common anthracyclinone pathways: aklavinone and nogalamycinone. Specifically, the minPKS Snoa123/AknBCDE2F/DpsABCDG complex synthesizes decaketide intermediates, which are inherently unstable and prone to forming various differently folded polyketides in the absence of subsequent biosynthetic enzymes[21,22]. When intermediates undergo reactions with ketoreductases (SnoaD/AknA/DpsE), aromatases (SnoaE/AknE1/DpsF), cyclases (SnoaM/AknW/DpsY), and oxygenases (SnoaB/AknX/DnrG), the first stable tricyclic anthraquinone intermediate, nogalonic acid or aklanonic acid, is generated[18]. The difference between these two molecules originates from the longer propionate starter unit for aklanonic acid. The

fourth ring cyclization is carried out by enantiospecific cyclases, which determine the configuration at the C9 position[23]. For example, SnoaL from the nogalamycin pathway results in a 9S configuration, while AknH from the aclacinomycin pathway and DnrG from the doxorubicin pathway contribute to a 9R configuration[18,24]. The methyltransferases (SnoaC/AknG/DnrG) add the methyl group to form an ester group at C10[18,25]. The final step of aglycone biosynthesis is the reduction of the keto group at C7, catalyzed by an NAD(P)H dependent ketoreductase (SnoaF/AknU/DnrE) to form a 7S configuration for the hydroxyl group[18,26]. The reduction is extremely important as it allows 7-*O*-glycosylation to occur later in the biosynthesis, which is essential for the bioactivity of most anthracyclines[8].

Apart from the most common aglycone biosynthetic pathways mentioned above, there are some exceptions[8]. For instance, in the steffimycin pathway, the aromatization occurs directly after the minPKS reactions, as the BGC does not contain a ketoreductase equivalent to SnoaD[27,28]. This reaction results in a polyketide-derived 2-hydroxyl group.

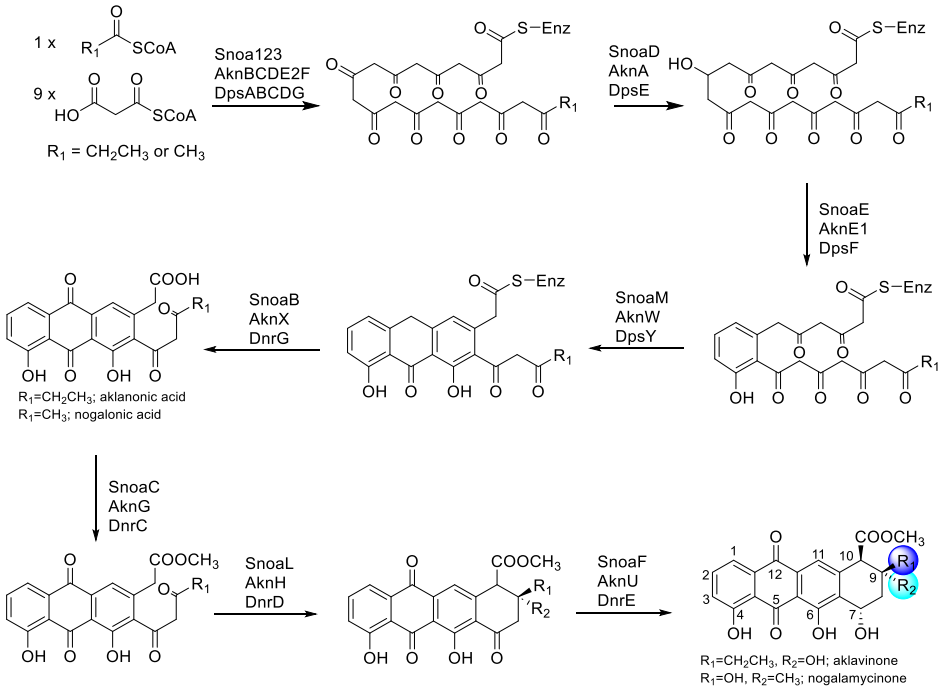


Figure 2. The biosynthetic pathway of polyketide aglycones aklavinone and nogalamycinone. Sno, nogalamycin proteins; Dps/Dnr, daunorubicin proteins; Akn, aclacinomycin proteins.

1.2.2 Biosynthesis of carbohydrates

Anthracyclines are prominently glycosylated molecules bearing diverse sugar moieties, which are crucial for their biological activity, solubility, and pharmacokinetic properties[29]. The most common form of sugar donor for anthracyclines is thymidine diphosphate (TDP) sugars[18]. TDP-sugars are exclusively synthesized from glucose-1-phosphate by glucose-1-phosphate thymidylyltransferases (*e.g.*, DnmL, AknY, SnogJ) to form TDP-D-glucose (**Figure 3**)[18]. TDP-D-glucose is subsequently converted into the key intermediate TDP-4-keto-6-deoxy-D-glucose by the action of TDP-glucose 4,6-dehydratase (*e.g.*, DnmM, AknR, SnogK)[30]. Various modifications, including deoxygenation, epimerization, methylation, transamination, and so on, can occur on TDP-4-keto-6-deoxy-D-glucose, branching the metabolic pathways towards more than 30 different kinds of TDP-sugars[31].

The TDP-sugar moieties on anthracyclines can be divided into neutral and amino sugars, depending on whether a transamination step occurs[18]. The classification is important because amino groups on sugar units drastically alter the physicochemical properties of the molecules, including solubility in water[32]. In case of anthracyclines, amino sugars have been shown to impact histone eviction and double-strand break activities[32,33]. For amino sugars, the most widely used anthracyclines, daunorubicin and doxorubicin, both contain L-daunosamine attached to C7[8]. Aclacinomycin A has L-rhodosamine decorating the aglycone, a sugar presenting two extra methyl groups in comparison to L-daunosamine[18]. During TDP-L-daunosamine biosynthesis, TDP-4-keto-6-deoxy-D-glucose is modified through 2-deoxygenation (DnmT, AclN, SnogH), 3-aminotransfer (DnmJ, AknZ, SnogI), 3,5-epimerization (DnmU, AknL, SnogF), and 4-ketoreduction (DnmV, AknM, SnogG) to form the final product[18,34]. TDP-L-daunosamine further undergoes dual *N*-methylations (AknX2/AclP, SnogAX) to form TDP-L-rhodosamine[35,36].

Many kinds of neutral TDP-sugars are also found in anthracycline pathways[8]. For example, TDP-L-nogalose is present in the nogalamycin pathway[37], while TDP-2-deoxy-L-fucose and TDP-L-rhodinose are seen in aclacinomycin N[38]. Additionally, the oxidoreductase AknOx catalyzes two FAD-dependent consecutive reactions on the TDP-L-rhodinose of aclacinomycin N, leading to the formation of aclacinomycin A and aclacinomycin Y[38]. TDP-L-nogalose is biosynthesized by 3-methylation (SnogG2), 3,5-epimerization (SnogF), 4-ketoreduction (SnogC), and three subsequent *C*-methylations (SnogYLM)[37]. On the other hand, TDP-L-rhodinose is generated through 3-ketoreduction (AknQ), 3-dehydration (AknP), 3,5-epimerization (AknL), and 4-ketoreduction (AknM) of TDP-4-keto-6-deoxy-D-glucose[38]. TDP-2-deoxy-L-fucose only exhibits an extra hydroxyl group at the 3' position compared to TDP-L-rhodinose since AknP is not involved[18,39].

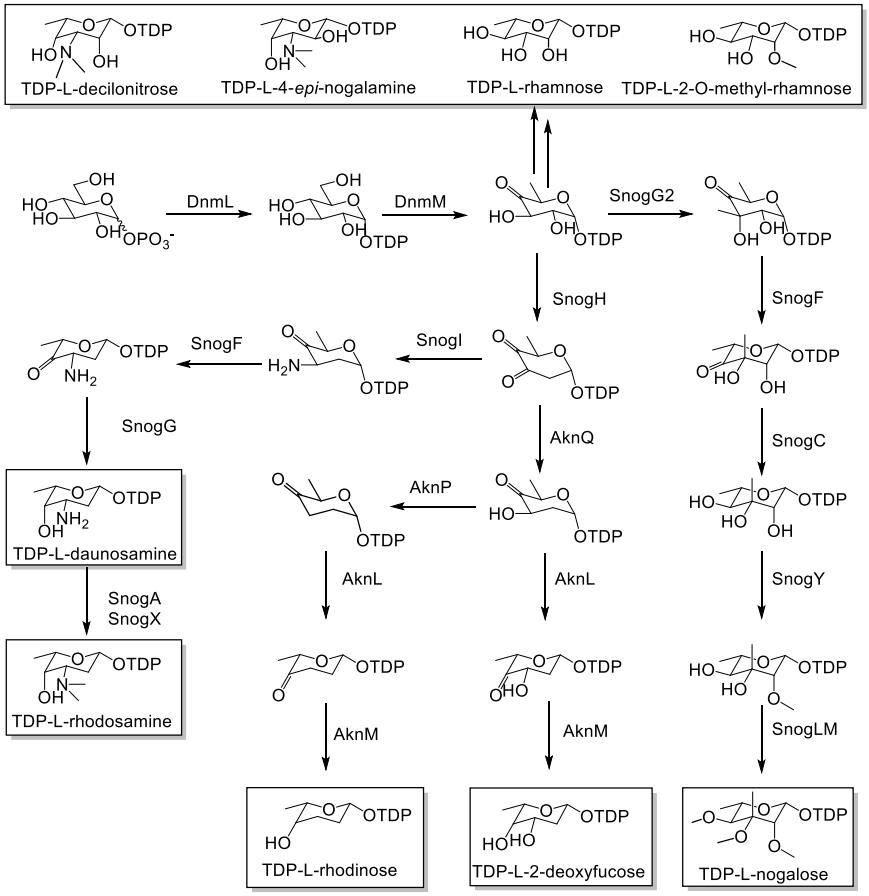


Figure 3. The biosynthetic pathways of common sugar units in anthracyclines. Snog, nogalamycin proteins; Dnm, daunorubicin proteins; Akn, aclacinomycin proteins.

1.2.3 Glycosylation of anthracyclines

Glycosylation of polyketide aglycones is mediated by glycosyltransferases (GTs)[8]. Carbohydrate units can be attached to several positions on the aglycones, including the C1, C4, C7, and C10 positions[20]. Here, the most common 7-*O*-glycosylation is first discussed, which is exclusively catalyzed by B family GTs with additional assistance from cytochrome P450 auxiliary enzymes[40]. Anthracycline 7-*O*-glycosyltransferases possess a GT-B fold, with an N-terminal domain responsible for binding the aglycone acceptor and a C-terminal domain that interacts with the sugar nucleotide donor[41]. It has been proven that P450 enzymes have an activating effect on GTs, possibly by causing a conformational change in GTs, thereby stabilizing the transition state and ultimately enhancing carbohydrate transfer[42]. For example, in the aclacinomycin pathway, the P450 auxiliary protein AknT can

accelerate the turnover rate of the GT AknS for TDP-L-rhodosamine transfer by 200-fold[43]. Although no protein structures of GT and P450 partner pairs for anthracyclines have been solved, the hypothesis was supported by the structure complex of GTs in the erythromycin D pathway[44].

On the other hand, the crystal structure of GT SnogD, responsible for 1-*O*-glycosylation in the nogalamycin pathway, was solved. The structure showed that SnogD contains two Rossman fold domains: one that binds the aglycone acceptor and the other responsible for binding the sugar donor[45]. Evidence showed that SnogD can catalyze the 1-*O*-glycosylation independently[45]. The sugar donor for SnogD is TDP-L-rhodosamine instead of TDP-L-nogalamine, as seen from the nogalamycin structure[36]. Later research clarified that the 2''-hydroxylation by SnoT occurs after the attachment of TDP-L-rhodosamine[46]. C10 is another position where *O*-glycosylation can happen, such as in the rhodomycin and cosmomycin pathways[47,48]. The mechanism of glycosylation is similar to 7-*O*-glycosylation, requiring both GTs and P450 enzymes.

The 4-*O*-glycosylation was found in some anthracyclines, such as komodoquinone A[49]. The rarer *O*-glycosylation can occur through the hydroxyl group on the epoxyoxocin rings of the atypical anthracyclines, such as decilorubicin and keyicin[50,51]. Furthermore, additional sugar molecules can be attached to a previously attached sugar through the formation of glycosidic bonds, creating complex polyglycosylated anthracyclines (**Figure 1**), such as aclacinomycin A and its derivatives[52], cosmomycin B and its derivatives[53], rhodomycin SS-288 derivatives[54], decilorubicin[48], and keyicin[49].

1.2.4 Post-PKS tailoring reactions

Post-PKS tailoring enzymes responsible for modifications on the aglycone core structures are crucial to the chemical diversity of anthracyclines. These modifications can occur at several positions, including C1, C2, C4, C6, C8, C9, C10, and C11 (**Table 1**)[20].

For instance, hydroxylation at C1 is observed in several anthracycline pathways, such as nogalamycin, kosinostatin, and decilorubicin[55]. The modification is necessary for subsequent 1-*O*-glycosylation in the nogalamycin and decilorubicin pathways, playing an important role in bioactivity[56]. The reaction is catalyzed by an atypical two-component mono-oxygenase system composed of a cyclase-like protein and a short-chain alcohol dehydrogenase/reductase (SDR)[55]. Several enzymatic pairs from different pathways and the mechanism for C1-hydroxylation have been reported recently[55]. The C4 position in anthracyclines is typically a polyketide-derived hydroxy group, but dehydration or methylation has been noted on selected pathways. For example, in the kosinostatin pathway, a four-enzyme system

(KstA15161011) catalyzes the 4-hydroxyl regioisomerization that removes a hydrogen group from the C4 position[57]. *O*-methylation modification is much more common, as demonstrated in the doxorubicin pathway, where the multi-functional enzyme DnrK is responsible for the reaction[58].

The carbon length and regiochemistry at the C9 group is determined by the starter unit and the fourth ring cyclase, respectively[56]. However, in the doxorubicin pathway, the cytochrome P450 enzyme DoxA is responsible for three-step oxidation events on the ethyl side chain at C9 to generate compounds daunorubicin and doxorubicin[59]. The C10 position remains a methyl ester group after aglycone synthesis due to the presence of oxygenase and methylase in most anthracycline pathways[18]. Additional modifications, such as demethylation, decarboxylation, and hydroxylation, are found in many pathways[20]. Demethylation, catalyzed by DnrP, RdmC, and EamC, is found in the doxorubicin, rhodomycin, and komodoquinone B pathways, respectively[60]. The pathways diverge in the next step, where DnrK in the doxorubicin pathway and EamK in the komodoquinone B pathway decarboxylate the C10 position, while RdmB in the rhodomycin pathway catalyzes hydroxylation[60,61]. The hydroxyl group at the C10 position is essential for atypical 10-*O*-glycosylation in the rhodomycin pathway[47]. Hydroxylation at the C11 position is observed in the doxorubicin, decilorubicin, and rhodomycin pathways, and the reaction is catalyzed by classical flavoenzymes (*e.g.*, RdmE)[50,62,63].

There are some uncommon post-PKS tailoring modifications in certain anthracycline pathways. For instance, in the steffimycin pathway, the C8 group can be further modified to a methoxy group[28]. In addition, steffimycin is a 2-hydroxylated anthracycline and this position can be further methylated to a methoxy group[27,28]. The hydroxyl group at C6 is derived from the polyketide backbone, but the substituent is removed in very rare anthracyclines such as yellamycin[64].

Table 1. The common post modifications on natural anthracycline aglycones.

Anthracycline	C1	C2	C4	C6	C8	C9[#]	C10	C11
aklavinone	H	H	OH	OH	H	Et, OH	COOMe	H
nogalamycin	OH*	H*	OH	OH	H	Me, OH	COOMe	H
kosinostatin	OH	H	H	OH	H	Me, OH[#]	OH*	H
decilorubicin	OH*	H*	H	OH	H	Me, OH	H	OH
doxorubicin	H	H	OMe	OH	H	Et, OH	H	OH
daunorubicin	H	H	OMe	OH	H	Et, OH	H	OH
rhodomycin B	H	H	OH	OH	H	Et, OH	OH*	OH
steffimycin B	H	OMe	OH	OH	OMe	Me, OH	O	H
yellamycin A	H	H	OH	H	H	Et, OH	OH*	OH

*Further modifications can occur in this position. [#]An atypical modification happens later. Et, ethyl group; Me, methyl group.

1.2.5 Transporters and self-resistance to anthracyclines

Anthracyclines can alter DNA structure and/or inhibit its functions, thereby exerting cytotoxic effects on their producers[14]. To ensure the survival of natural producers, the BGCs usually contain genes for transporters and self-resistance[18]. Anthracyclines are mostly exported to the extracellular environment by ATP-binding cassette (ABC) transporters in natural *Streptomyces* producers[65,66]. The most well-studied case involves the transporters DrrAB from the doxorubicin pathway (**Figure 4**)[65]. DrrAB are among the simplest ABC drug transporters. DrrA is a peripheral membrane protein, while DrrB is an integral membrane protein containing seven transmembrane domains[43]. When both DrrA and DrrB are present, their interaction stabilizes DrrA in the membrane, enabling it to adopt an active conformation. DrrA can subsequently bind ATP in a doxorubicin-dependent manner, functioning as an ATP-driven pump to efflux doxorubicin and daunorubicin. Evidence suggests that the DrrAB function as a multidrug transporter system with broad drug specificity[67].

The repair of damaged DNA by UvrA-like proteins is an additional protection mechanism. In the doxorubicin pathway, the UvrA-like protein DrrC is vital for resistance and the *drrC* mutant is less resistant to daunorubicin[68,69]. DrrC could scan the DNA and dislodge the intercalated daunorubicin/doxorubicin, thereby allowing free DNA for replication[68]. A similar damage repair gene, *snorO*, also exists in the nogalamycin pathway[37].

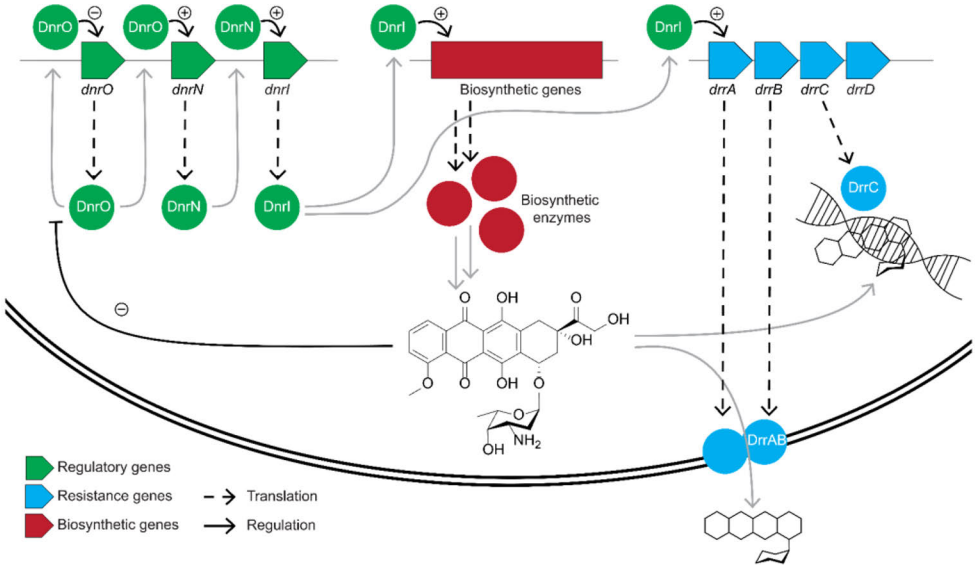


Figure 4. Regulation of doxorubicin biosynthesis and resistance mechanism. Reprinted from the reference with permission granted[8].

1.2.6 Regulatory genes

Anthracycline pathways are mostly derived from secondary metabolism, therefore, the expression of these BGCs is tightly regulated[19]. The regulation is mediated by a variety of global and pathway-specific regulators[19]. Understanding the mechanism of these regulators is crucial, as they can be targets for engineering strains for enhanced production of NPs or the activation of silent BGCs[70]. Here, I would like to discuss more on the pathway-specific regulators.

For instance, the regulation of the doxorubicin BGC involves three pathway-specific regulators (**Figure 4**). DnrO is a TetR-family transcriptional regulator that can inhibit its own transcription by binding to its promoter region. Meanwhile, DnrO can bind to all glycosylated products of the pathway, which inhibits DNA-*dnrO* binding and leads to increased *dnrO* expression[71]. DnrO can also activate the transcription of *dnrN* by binding to its promoter region[72]. Therefore, the translation of DnrN is upregulated, which subsequently activates *dnrI*. The increased DnrI then directly regulates the expression of the biosynthetic and self-defense genes in the BGC, leading to the activation of doxorubicin production[73].

1.3 Synthetic biology tools in *Streptomyces*

Synthetic biology is an interdisciplinary field focused on creating new biological components, devices, and systems, or redesigning existing biological systems using

engineering principles[74]. Since the beginning of this century, synthetic biology has developed significantly, thanks to recent advances in bioinformatics, recombinant DNA techniques and synthetic DNA techniques, especially in the model microorganisms like *Escherichia coli* (*E. coli*) and yeast[75–77]. Although the development of synthetic biology in *Streptomyces* is constrained by their complex life cycle and high G-C ratio of DNA content, there have been many successful cases of developing various synthetic biology tools in *Streptomyces*[78–80]. These include new recombinant DNA tools, well-established *Streptomyces* chassis, and other essential DNA elements for engineering, which will be discussed in the following sections.

1.3.1 *Streptomyces* chassis development

As a fundamental step in the development of synthetic biology in *Streptomyces*, several heterologous expression hosts have been built for discovering and characterizing novel metabolites[79]. An ideal heterologous host should be well-characterized and possess essential characteristics, including fast and dispersed growth, a clear genetic background, a clean metabolic background, amenability to genetic engineering, and enhanced secondary metabolite production[79]. To this end, several engineered *Streptomyces* strain chassis have been created and been widely used for heterologous expression of BGCs, including *Streptomyces coelicolor* (*S. coelicolor*), *S. venezuelae*, *S. lividans*, *S. albus* and *S. avermitilis*.

Among these, *S. coelicolor* is the best-known species of *the Streptomyces* genus[79]. *S. coelicolor* A3 (2) was the first *Streptomyces* species to have its genome fully sequenced, showing an 8.7 Mb linear chromosome and harboring over 20 BGCs[81]. Many attempts have been made to construct a better *S. coelicolor* chassis[79]. For example, four endogenous BGCs were knocked out to generate *S. coelicolor* M1146 with a clean production background[82]. In addition, point mutations in *rpoB* encoding RNA polymerase β-subunit (*S. coelicolor* M1152) and *rpsL* encoding ribosomal protein S12 (*S. coelicolor* M1152) enhanced the overall BGC production without growth impairment[82]. *S. lividans* is genetically close to *S. coelicolor* and popular for lacking an endonuclease restriction system, which makes it accept methylated foreign DNA and results in high transformation efficiency[83]. The derivative *S. lividans* TK24 has been widely used as heterologous hosts because of clean background resulting from knock-out of endogenous BGCs[84,85]. *S. albus* is another versatile species with a small genome size (~6.8Mb). *S. albus* J1074, defective in the SalI restriction-modification system[86], allows efficient transformation of heterologous BGCs and has shown excellent performance in numerous heterologous expression studies[28,87,88]. More details of widely used chassis

regarding their genotypes, characteristics, and applications are listed in **Table 2** below.

Table 2. Examples of widely used *Streptomyces* chassis

Chassis	Species advantages	Genotype/characteristics	Product	Reference
S. coelicolor M145	Genetically best- studied species; Various genetic tools.	ΔSCP1 ΔSCP2	Chloramphenicol Congocidine	82
S. coelicolor M1146		Δ*act*, Δ*red*, Δ*cpk*, and Δ*cda* compared to *S. coelicolor* M145	Cypemycin Formicamycin	89,90
S. coelicolor M1152		Point mutations in *rpoB* [S433L] in *S. coelicolor* M1146	Chloramphenicol Tatiomicin	91,92
S. coelicolor M1154		Point mutations in *rpsL* [K88E] in *S. coelicolor* M1152	FK506 Clorobiocin	93,94
S. lividans TK24	Low endogenous protease activity; Accept methylated DNA; High transformation efficiency; Fast growth.	Δ*act*, Δ*red*, and Δ*cda*	mithramycin A Syringolin	84,85
S. lividans TK23		One copy integration of AfsRS increased plasmid uptake efficiency	JBIR-159 Kinanthraquinone	95,96
S. lividans K4-114		Δ*act* in *S. lividans* TK24	Platencin Salinomycin	97,98
S. venezuelae DHS2001	Fast growth; Sporulates in liquid; Easy manipulation; High conjugation efficiency.	Δ*pikA*	Phenylpropanoids Tylactone	99,100
S. venezuelae YJ028		Δ*pikA* and Δ*des*	Doxorubicin derivatives Flaviolin	101,102
S. albus J1074	Small genome; Efficient genetic toolkits.	Δ*pfk* and Δ*wblA*; Overexpression of *cpk*; Deletion of paulomycin BGC	Steffimycin Thaxtomins	28,88
S. albus Del14		Deletion of 15 BGCs in *S. albus* J1074	Tunicamycin B2 Moenomycin M	103
S. avermitilis SUKA5	Stable genome; Efficient precursor supply.	Deletion of 1.51 Mb left arm; Deleted oligomycin BGC;	Pladienolide Streptomycin	6
S. avermitilis SUKA22		Deletion of three BGCs; Prevent undesired recombination	Resistomycin Bafilomycin	104

1.3.2 Vector development

Vectors play a crucial role in identifying and manipulating heterologous genes, facilitating the heterologous expression of BGCs in *Streptomyces*[79,105]. The vectors developed for *Streptomyces* can be divided into three groups: cosmids, artificial chromosomes, and standardized and orthogonal vectors.

1.3.2.1 Cosmid

Cosmid vectors are designed to accommodate large DNA fragments, up to 45 kb in size, making them ideal for cloning small to medium-sized BGCs[106]. They combine features of plasmids and λ phage[107]. First, they harbor essential plasmid elements, including the origin of replication, multiple cloning sites, and selective markers. Additionally, cosmids encode the *cos* sequences of bacteriophages, allowing them

to be packaged *in vitro* into phage capsids[107]. After packaging, they can be transduced into a target host, such as *E. coli*, where they replicate as plasmids. Cosmids are usually designed as multi-copy plasmids in *E. coli* to benefit DNA isolation and further manipulation. In contrast, fosmids, which are low-copy-number cosmids with the *E. coli* F factor replicon, were constructed to improve structural stability of large DNA inserts[108].

1.3.2.2 Artificial chromosome vector

Cosmids are designed for medium-sized DNA[105]. However, natural product BGCs can reach up to 200 kb[109]. Therefore, vectors based on artificial chromosomes were designed, which have the capacity to carry 100–350 kb of DNA[110]. Artificial chromosome vectors can be grouped based on the origin of the replicons. The most common categories are phage artificial chromosome (PAC)[111], bacterial artificial chromosome (BAC)[112], yeast artificial chromosome (YAC)[113], and fungal artificial chromosome (FAC), among which PACs and BACs have been more frequently utilized in *Streptomyces*[114].

The first PAC vector, pCYPAC-1, was designed by merging the features of P1-phage and F factor[111]. It can accommodate DNA inserts ranging from 100 to 300 kb in bacterial cells. The recombinant PAC vector can remain as a single copy and propagate stably, while the inserted DNA can exist without issues of chimerism or rearrangements. More derivatives of pCYPAC-1 were developed later, such as pESAC, to include essential elements for gene manipulations in *Streptomyces*[93,115].

On the other hand, the first BAC vector, pBAC108L[112], was developed by taking advantage of the well-studied *E. coli* single-copy plasmid F factor. It contained the *oriS* and *repE* genes mediating replication. Meanwhile, *parA* and *parB* were also retained to control the copy number at a level of one or two in *E. coli*, which helped maintain large intact DNA. Additionally, a chloramphenicol resistance marker and a cloning segment were introduced to benefit the cloning process. Thanks to this design, pBAC108L was able to stably carry a 300 kb size of human DNA. Similarly to PAC vectors, more BAC vectors were designed later for *Streptomyces*, including pStreptoBAC[116] and pSBAC[117], which included more user-friendly elements for genetic engineering in *Streptomyces*.

1.3.2.3 Standardized and orthogonal vectors

With the rapid advancements in synthetic biology, a wide range of standardized and orthogonal vectors has been designed for *Streptomyces*, including both replicative and integrating plasmids[79]. The replicative vectors usually contain replicons originating from natural *Streptomyces* plasmids, such as SCP2 from *S. coelicolor*.

Meanwhile, the integrative vectors are based on integration mediated by phage integrases. Specifically, integrases are a group of phage-encoded, site-specific recombinases that facilitate a single crossover event between the *attP* site on the phage and the *attB* site on the bacterial chromosome. Therefore, designing the integrase coding genes and *attP* site on a vector allows integration at the *attB* sites that naturally exist in *Streptomyces* chromosomes, thereby obtaining stable expression[118]. The common replicative and integrative plasmids for *Streptomyces* and their characteristics are listed in **Table 3**.

Table 3. Examples of common plasmids for *Streptomyces*

Vector Name	Type	Characteristics	Reference
pWHM3	replicative	amp^R, tsr^R, *oriT*, pIJ101 *ori*	[119]
pUWL201	replicative	amp^R, tsr^R, pIJ101 *ori*	[120]
pSET152	integrative	amp^R, apr^R, *oriT*, *attP*-ΦC31	[121]
pAV	integrative	apr^R, RP4, *attP*-VWB	[122]
pSC	integrative	spc^R, RP4, *attP*-ΦC31	[122]
pTB	integrative	tsr^R, RP4, *attP*-ΦBT1	[122]
pOSV801–812	integrative	apr^R/hyg^R/kan^R, *oriT*, *attP*-ΦBT1/ΦC31/pSAM2/VWB	[123]
pENSV1	integrative	amp^R, spc^R, *oriT*, *attP*-SV1	[124]
pENTG1	integrative	amp^R, vio^R, *oriT*, *attP*-TG1	[124]

amp^R, ampicillin^R; tsr^R, thiostrepton^R; apr^R, apramycin^R; hyg^R, hygromycin^R; kan^R, kanamycin^R; spc^R, spectinomycin; vio^R, viomycin^R; *oriT* and RP4, origin of transfer enabling conjugative transfer from *E. coli* into *Streptomyces*; pIJ101 *ori*, origin of replication in *Streptomyces*; *attP*, attachment sites of the recombinases.

1.3.3 Promoter engineering

Promoters are fundamental genetic elements that control gene transcriptions. They can be categorized into two primary categories: constitutive and inducible promoters. Constitutive promoters, usually derived from housekeeping genes, ensure stable expression levels across various growth conditions. In contrast, inducible promoters, often originating from genes that are naturally activated under specific conditions, enable significant changes in expression levels in response to environmental stimuli[125]. Numerous natural and synthetic promoters have been identified to manipulate gene expression in *Streptomyces*, encompassing both constitutive and

inducible types[78]. The most widely used promoters along with their origins and characteristics are listed in **Table 4** below.

Table 4. Examples of commonly used promoters in *Streptomyces*

Promoters	Origin and characteristics	Reference
Constitutive promoters		
*kasOp**	Derived from the promoter of *Streptomyces* antibiotic regulatory protein (SARP) family regulator in *S. coelicolor* A3 Removing the binding sites of ScbR and ScbR2 regulators, then obtained by random mutation	126–128
*ermEp**	Derived from the promoter of erythromycin resistance gene in *Streptomyces erythraeus*; trinucleotide deletion in the ermEp1 region of *ermEp*	129–131
gapdhp(EL)	Promoter region of glyceraldehyde-3-phosphate dehydrogenase in *Eggerthella lenta*	132–134
rpsLp(CF)	Promoter region of 30S ribosomal protein in *Cellulomonas flavigena*	132,133
SF14	Derived from phage 119 isolated from *Streptomyces ghanaensis*	135,136
p21	Screened from the synthetic promoter library built based on the consensus -35 and -10 sequences of the *ermEp1* promoter	137–139
p15	RNA sequencing screening from *S. albus* J1074	140
sp1–*sp44*	Flow cytometry screening from mutated *kasOp** promoters induced by osmotic stress	131,132,141
Inducible promoters		
tipAp	Thiostrepton inducible promoter from *S. lividans*	142–145
xylAp	Xylose inducible promoter	146
nitAp	ε-caprolactam and transcription factor NitR induced promoter from the nitrilase gene *nitA*	147
tcp830	Synthetic tetracycline-inducible promoters	148–150
P21-cmt-CymR	Synthetic cumate inducible promoter	138,151

1.3.4 Terminator engineering

Terminators are essential genetic elements that stop transcription, insulating and stabilizing the transcripts. Many natural terminators have been screened and characterized in *Streptomyces* (**Table 5**). Furthermore, many well-characterized terminators from other microorganisms have also been found to function in

Streptomyces, owing to the shared fundamental principle of intrinsic termination. Specifically, their palindromic sequences fold into stable hairpin structures, which disrupt interactions between RNA polymerase and the nascent RNA, leading to transcription termination[152].

Table 5. Examples of terminators for *Streptomyces*

Terminators	Origin and characteristics	Reference
fd	An intrinsic transcription terminator from fd bacteriophage	153
B0015	Combined ribosomal RNA operon B (rrnB) T1 and T2 terminators	154
B0017	Terminator with two copies of rrnB T1	154
tt-sbi-A	Synthetic bidirectional terminator	155
10 synthetic terminators	Synthetic terminators tested in *S. venezuelae*	151,156

1.3.5 DNA assembly tools for *Streptomyces*

1.3.5.1 Traditional cloning tools for natural product BGC

The sizes of NP BGCs are typically large, ranging from 10 to 200 kb, making their cloning challenging and time-consuming[105,109]. To address this issue, various methods have been developed over the last several decades[114]. One of the earlier BGC cloning methods is based on cosmid library construction[107]. In this method, chromosome DNA containing the target BGCs is extracted and partially digested with restriction enzymes or random shearing. The resulting fragments are then ligated into cosmid vectors with compatible restriction sites. The DNA library is packaged into phages, which are used to infect *E. coli*. However, the subsequent PCR-based screening is relatively tedious, requiring the screening of thousands of colonies[106]. In addition, cosmid library cloning relies on the random digestion of the genomic DNA and is limited by the capacity of the cosmid, often resulting in incomplete BGCs[37].

Due to the limitations of cosmid library-based BGC cloning, more efficient methods were developed, such as transformation-associated recombination (TAR) system, Red/ET recombineering, integrase-mediated recombination (IR) system, and plasmid *Streptomyces* bacterial artificial chromosome (pSBAC) system[114]. The TAR system leverages homologous recombination in *Saccharomyces cerevisiae* (**Figure 5A**)[157,158]. The key aspect of this method is the construction of a capture plasmid with two homology arms of the target BGCs[158]. Enzymes are required to digest intact

BGCs from genomic DNA[114]. The linear capture plasmid and the digested BGCs are then co-transformed into yeast to allow *in vivo* homologous recombination. A well-designed capture plasmid typically contains elements for replication in yeast, *E. coli*, and actinobacteria, enabling the recombinant plasmid to shuttle between these organisms[158]. For *Streptomyces*, the *oriT* sequence is usually included in the capture plasmid to facilitate intergeneric conjugation from *E. coli*[159]. The *attP-int* sequence and integrase coding sequence are often introduced for stable expression of the BGCs by integrating them into the *Streptomyces* chromosome, with the TAR plasmid pCAP01 being a typical example[159].

Red/ET recombineering (**Figure 5B**) is another BGC cloning tool developed in *E. coli*, utilizing phage-recombinases, either Redα/Redβ from the λ phage or RecE/RecT from the Rac prophage[160]. RecE/Redα are 5'→3' ATP-independent exonucleases that create single-stranded overhangs, whereas RecT/Redβ are DNA annealing proteins that promote the annealing of complementary single-stranded DNA during recombination[160]. The procedure for this method is similar to the TAR system, where digested genomic DNA with BGCs is co-transformed into *E. coli* with a linear vector containing homology arms of the BGC[161]. Homologous recombination subsequently occurs *in vivo* by the recombinases to generate the recombinant vector, which can be shuttled and further expressed in *Streptomyces*[160].

The IR system (**Figure 5C**) was developed to directly clone large DNA fragments from bacterial genomes through site-specific recombination[114]. This method utilizes phage integrases for *Streptomyces*. For example, the ΦBT1 integrase system allows the direct cloning of BGCs from genomic DNA[162]. The system consists of three different plasmids: a pUC119-based suicide vector (pSV) containing a mutated *attP* site of the ΦBT1 integrase and a homologous region to the 5' end of the target BGC, a pKC1139 with an *attB* site and a region homologous to the 3' end of the target BGC, and an integrative plasmid pIJ10500 that introduces the ΦBT1 integrase. The two plasmids can be inserted adjacent to the BGC by single crossover, and the subsequent introduction of ΦBT1 integrase enables the excision of the pKC1139 harboring the target BGC. The pKC1139 can then be extracted and transferred for recovery or heterologous expression.

The pSBAC system employs the pSBAC vector for capturing NP BGCs (**Figure 5D**)[163]. This system leverages the ability of pSBAC vector to carry large BGCs, its replicons for propagation in *E. coli*, and its *oriT* and ΦC31 *attP-int* for BGC integration in *Streptomyces*[163]. This method depends on the presence or insertion of unique restriction enzymes on both sides of the BGC. The pSBAC vector needs to be inserted between the unique restriction enzyme sites in the chromosome through homologous recombination. The genomic DNA is subsequently extracted and digested by the unique restriction enzyme, allowing straightforward capture by self-ligation.

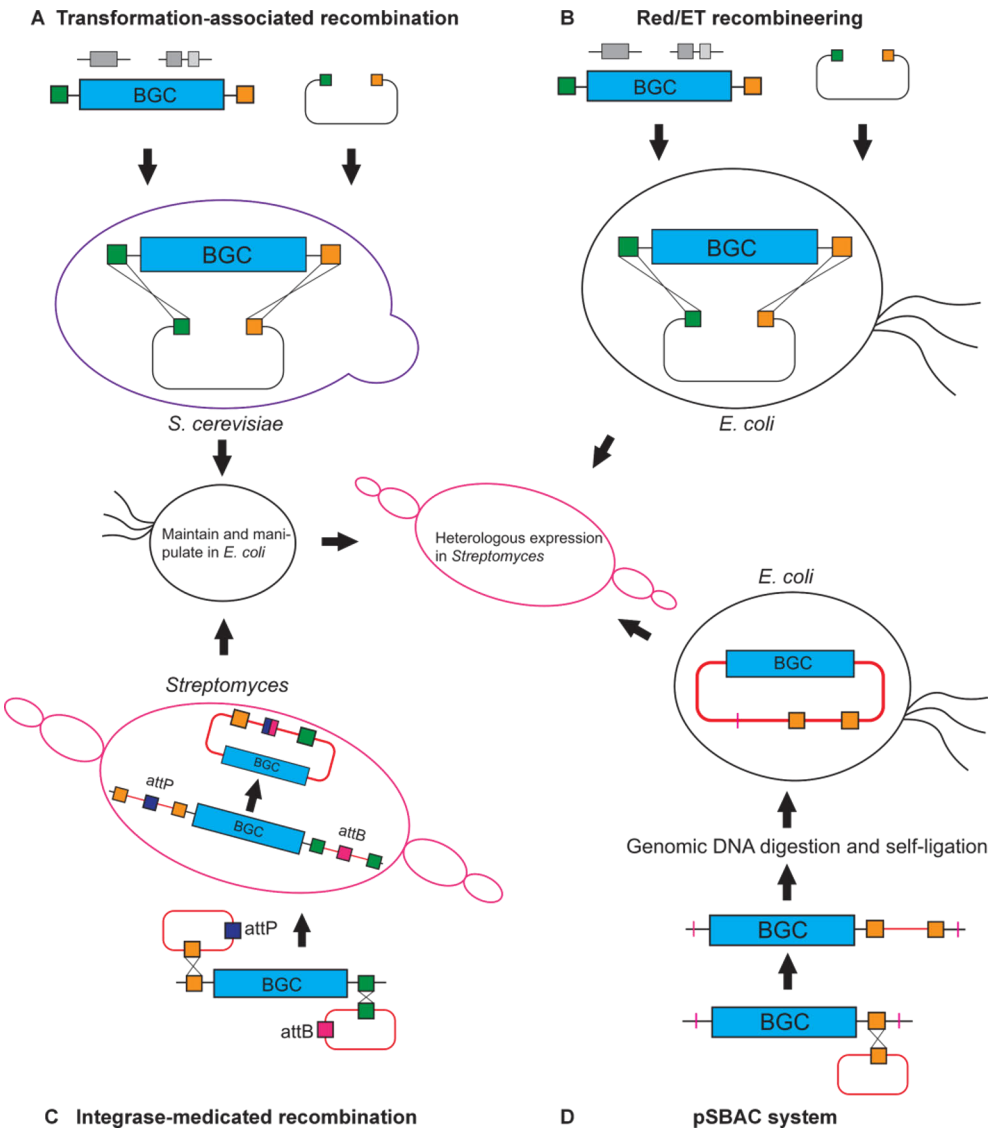


Figure 5. Cloning tools for direct capture large BGCs. A, Digested genomic DNA containing the target BGC is co-transformed into *S. cerevisiae* with a capture vector bearing homology arms. Homologous recombination assembles the BGC into the vector. **B**, Red/ET recombineering, similar to TAR, enables *in vivo* assembly of the BGC in *E. coli* using phage-derived recombinases (Redα/Redβ or RecE/RecT). **C**, A suicide vector with an *attP* site is integrated at the 5' end of the BGC, and a second vector with an *attB* site at the 3' end. An integrase expressed from a third plasmid excises the BGC onto the second plasmid. **D**, The pSBAC system inserts a capture vector near the BGC. After genomic digestion with restriction enzymes, self-ligation recovers the BGC on the vector. All captured BGCs can be maintained in *E. coli* for modification and transferred to *Streptomyces* for heterologous expression.

1.3.5.2 Advanced tools for pathway assembly

Recombinant DNA techniques are crucial for constructing or redesigning microbial cell factories[105]. In addition to traditional techniques like restriction enzyme assembly, several advanced recombinant DNA techniques have emerged, including Gibson assembly, Golden Gate assembly, and BioBricks assembly (**Figure 6**).

Gibson assembly is a seamless technique that relies on the enzyme activities of exonuclease, DNA polymerase, and DNA ligase[164]. The assembly of multiple fragments using Gibson assembly involves the following steps: 20–40 bp overlaps are first created between adjacent DNA fragments. The 5'-exonuclease generates long overhangs, which then anneal at the complementary region. The polymerase fills in the gaps in the single-stranded regions, while DNA ligase repairs the nicks. The overlaps guide the DNA fragments to anneal in a designated order, but this also means each DNA fragment contains unique overlapping DNA sequences, making them non-reusable in other applications. PCR is typically used to create overlaps, but PCR amplification can be challenging in *Streptomyces* due to the high G-C content of its DNA[165]. Although synthetic DNA fragments can be used, it will make this approach less cost-effective. Additionally, it has been shown that the high G-C level of the overlapping DNA significantly decreases the accuracy of this method[166].

Golden Gate assembly is another method for assembling multiple fragments. It utilizes type IIS restriction enzymes, which are a class of enzymes that recognize a specific DNA sequence, but cleave the DNA at a distance from the recognition site, usually outside the recognition sequence itself[167,168]. This allows a single type IIS enzyme to create several designed DNA ends to seamlessly join multiple DNA fragments. However, like Gibson assembly, each fragment has two ends for ligation that only match with those on neighboring fragments, limiting their reuse in other applications. PCR is preferred for designing or adding type IIS recognition sites to gene fragments, but it poses challenges, especially when amplifying the high G-C content DNA sequencing for *Streptomyces*. Conversely, synthetic DNA, while overcoming PCR limitations, is hindered by high cost and non-reusable drawbacks.

BioBricks is a standardized DNA assembly method based on the recycling of restriction sites with compatible sticky ends. After ligation, the ends generate a scar sequence that is not recognized by either enzyme, allowing new genes to be added indefinitely by repeating the process[169]. In the BioBricks [RFC-10] standard, gene fragments are flanked by four BioBricks restriction enzymes EcoRI, XbaI, SpeI, and PstI. XbaI and SpeI form compatible sticky end 5'-CTAG-3' after digestion[124]. A typical BioBricks assembly round involves digesting the upstream and downstream inserts from plasmids using EcoRI/SpeI and XbaI/PstI, respectively. Simultaneously, a destination plasmid is digested with EcoRI/PstI to carry the inserts. The advantage of Biobricks assembly is that it provides standard and interchangeable biological parts, making them reusable for various purposes, even across different laboratories.

For *Streptomyces*, an efficient way to build a BioBricks library is to use synthetic DNA for all genetic elements, such as genes, promoters, and terminators, overcoming PCR limitations. Due to its high reusability, BioBricks assembly is more cost-efficient compared to Gibson and Golden Gate assembly. Nevertheless, BioBricks assembly has several disadvantages: it is not a seamless technique and introduces a sequence scar with each ligation event[169]. Additionally, it does not allow multiple-fragment cloning and requires many rounds when handling multiple fragments.

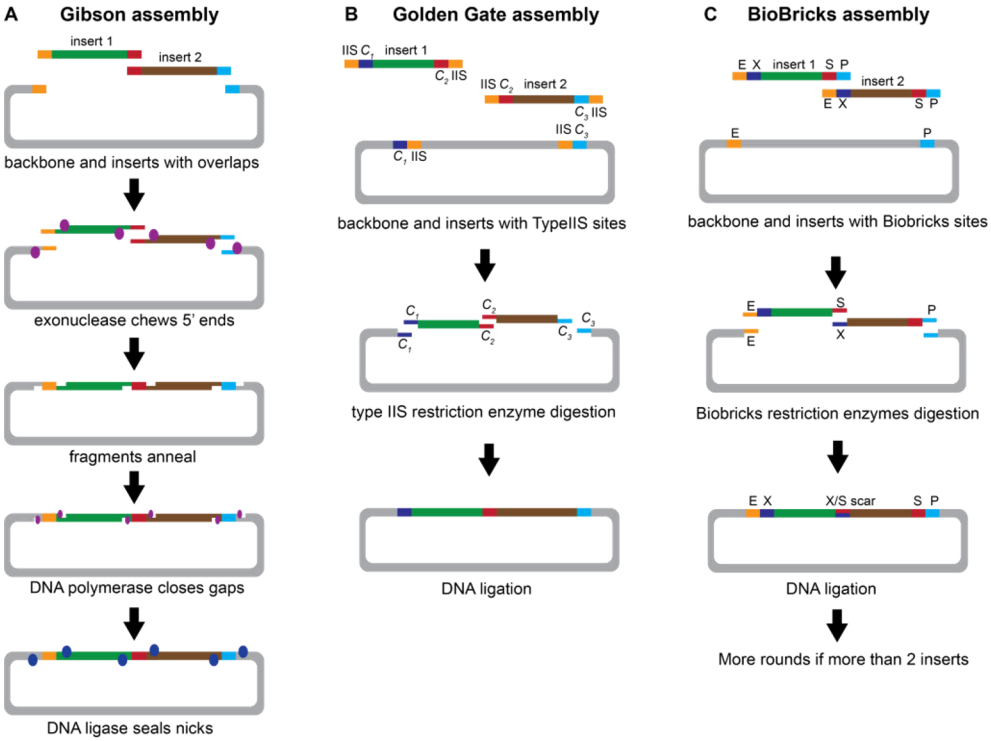


Figure 6. Three advanced assembly tools for pathway assembly. A, In Gibson assembly, DNA fragments with overlapping ends are combined in a one-pot reaction where a 5' exonuclease creates single-stranded overhangs, fragments anneal, DNA polymerase fills in gaps, and DNA ligase seals the nicks, resulting in seamless assembly. **B**, In Golden Gate assembly, DNA parts flanked by Type IIS restriction sites are digested to release fragments with user-defined overhangs, allowing simultaneous and ordered ligation of multiple parts in a scarless manner. **C**, In BioBricks assembly, DNA parts with standardized restriction sites (e.g., EcoRI, XbaI, SpeI, PstI) are sequentially digested and ligated; each assembly step typically results in a fixed scar sequence and requires multiple rounds for combining more than two fragments. Legend: IIS, Type IIS restriction enzyme recognition sites; C1–C3, Type IIS restriction enzyme cutting sites; E, EcoRI; X, XbaI; S, SpeI; P, PstI.

1.3.6 Genome editing tools for *Streptomyces*

Genome editing has become a cornerstone of metabolic engineering and synthetic biology, enabling the deliberate redesign of microbial pathways for the production of valuable chemicals[170]. In *Streptomyces*, precise and efficient genome manipulation is especially important for harnessing their vast biosynthetic capabilities, which allows scientists to activate silent BGCs, eliminate competing metabolic pathways, reengineer native regulatory systems, and incorporate synthetic components to enhance metabolite production[171].

Several genome editing tools have been developed in *Streptomyces*[172]. The most conventional method is homologous recombination, which uses a gene replacement cassette flanked by homology arms[173]. However, it is limited by the low recombination efficiency in *Streptomyces*, especially for deleting large genome regions. To overcome these limitations, alternative methods have been developed, such as the PCR-targeting system[174], the Cre-loxP recombination system[6] and the FLP/FRT recombination system[175]. The PCR-targeting system, adapted from traditional homologous recombination, requires a cosmid library[174,176]. Target genes on the cosmids are deleted in *E. coli* using the λ Red system, and the modified cosmids are transferred to *Streptomyces* for promoted gene replacement via long homology regions[174]. The Cre-loxP and FLP/FRT systems employ site-specific recombinases (Cre and Flp) to excise DNA flanked by two loxP/FRT sites[177]. To delete the target genes, the loxP/FRT sites must be introduced on both sides of the gene, and the recombinase expressed triggers deletion through two single-crossover events[172,178]. However, these methods are time-consuming and non-seamless techniques, requiring more efficient and scarless genome editing strategies.

On the other hand, more advanced clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein (Cas) systems enable the precision and efficiency of genome editing in *Streptomyces*[172,179]. CRISPR/Cas system is an adaptive immune system against invading foreign DNA or RNA, which widely exist in archaea and bacteria[180]. A CRISPR array consists of short direct repeats separated by short variable DNA sequences (‘spacers’), which is flanked by diverse *cas* genes[181]. The CRISPR array is transcribed to CRISPR RNAs (crRNAs) after processing and maturement, which can guide Cas proteins to cleave the targeting nucleic acids[180]. Researchers adapted CRISPR/Cas system as genome editing techniques by introducing *cas* gene cassettes in the host and designing the artificial crRNA array targeting gene interests[182]. Genome editing is accomplished during the repair of DNA double-stranded breaks through endogenous error-prone nonhomologous end-joining (NHEJ) or homology-directed repair (HDR) assisted by the introduced homology templates[182].

The most widely used CRISPR/Cas system in *Streptomyces* is class 2 type II Cas9 originating from *Streptococcus pyogenes*[80,172,183]. Cas9 requires the presence of

a protospacer adjacent motif (PAM) sequence (5'-NGG-3') for target recognition and cleavage[184]. To address the limitations in non-high G-C regions, the class 2 type V Cas12a/Cpf1 system, which recognizes a different PAM (5'-TTTV-3'), was developed as a complementary tool[133]. However, both Cas9 and Cas12a are heterologous to *Streptomyces* and can exhibit cytotoxicity, posing challenges for their application in these organisms[184,185]. Recently, a native type I-E CRISPR system has been reported as an alternative, offering reduced toxicity and improved compatibility for genome engineering in *Streptomyces*[186].

However, it is worth noting that the efficiency of these editing methods in *Streptomyces* is strain-specific and even sequence-specific[133,185,187], which could be related to the presence of restriction-modification system and tolerance level of toxic proteins like Cas nucleases[184,185]. Therefore, there is no one-size-fits-all genome editing tool across the *Streptomyces* genus and the tool selection is case-specific.

1.4 Heterologous expression of secondary metabolite biosynthetic pathways in *Streptomyces*

1.4.1 Traditional heterologous expression of BGCs in *Streptomyces*

1.4.1.1 Significance of heterologous expression of BGC

Heterologous expression of secondary metabolite pathways is an effective strategy widely used for discovering new NPs, owing to its advantages over expression in native hosts[79,114]. First, the expression of secondary metabolite pathways is usually tightly regulated and often silent because of the regulatory networks from the native hosts or the lack of environmental triggers[188]. Heterologous expressions can overcome these regulatory constraints, allowing the production of silent NPs[189]. Second, model microorganisms are typically selected for heterologous expression due to their beneficial properties: easy culturing, clear genetic background, clean production background, and advanced synthetic biology toolbox[81]. Therefore, heterologous expression in model hosts facilitates easier manufacturing and downstream purification processes. In addition, the available synthetic biology toolboxes in model chassis allow genetic manipulation of biosynthetic pathways, leading to improved production or new metabolite derivatives.

1.4.1.2 Examples of traditional heterologous expression of BGCs in *Streptomyces*

To date, more than 100 NP BGCs have been successfully expressed in *Streptomyces* heterologous hosts using various strategies[79,105,114,190]. Most BGCs were isolated and captured by traditional cosmid library screening. For instance, the partial 19.7 kb cluster of anthracycline steffimycin was captured by cosmid library construction using the cosmid pKC505[28]. A 15 kb cluster was recovered by restriction digestion, with 11 kb involved in steffimycin biosynthesis. Assembling the two partial gene clusters into an expression plasmid and transforming it into the heterologous host *S. albus* lead to steffimycin production[28]. However, this method is laborious, time-consuming, and often results in incomplete pathways when dealing with large BGCs (>20 kb).

More efficient methods have been used to simplify the process of acquiring target BGCs. One such method is homologous recombination inspired strategies. For example, the TAR system was used to successfully capture a 54 kb BGC encoding anthracycline cosmomycin from *Streptomyces* sp. CNT-302 strain[191]. The recombinant pCAP01 vector with the captured cluster was further expressed in the heterologous host *S. coelicolor* M512, leading to the production of cosmomycin C and the identification of a new cosmomycin analog[191]. The cryptic type II PKS gene cluster (*skt*) from *Streptomyces* sp. Tü 6314 was identified through Red/ET recombineering[192]. In brief, *Nhe*I-digested genomic DNA with the intact target BGC was co-transformed into *E. coli* with a linear capture vector pSET152 containing 50 bp homology arms of the cluster. The expression in heterologous host *S. coelicolor* M1152/M1154 lead to the production of six aromatic polyketides, including three new streptoketides.

Site-specific recombination-based methods have also been developed for this purpose. For instance, the ΦBT1 integrase-mediated IR system have been proven successful for cloning many gene clusters, including the *act* cluster for typical type II polyketide actinorhodin from *S. coelicolor* M145[162]. The excised 25 kb cluster was introduced into *S. coelicolor* M1146, restoring the production of the blue pigment actinorhodin. Excision of the *act* cluster to a multi copy plasmid increased the gene cluster copy number in the host, contributing to higher production yields. The pSBAC system leverages the versatile pSBAC vector for capturing BGCs. For example, the BGC encoding pikromycin, a type I polyketide macrolide antibiotic, was captured via the pSBAC system[193]. Briefly, a unique HindIII site was inserted into the left border of the BGC by PCR-targeted gene insertion, while the pSBAC vector was inserted into the right border through a single cross, harboring the homology arms and another HindIII site[193]. The chromosomal DNA was digested with HindIII, and self-ligated to form a recombinant pSBAC vector carrying the pikromycin cluster. The recombinant vector was further introduced into the

heterologous host *S. lividans* and *S. coelicolor*, leading to the production of a major pikromycin derivative, 10-deoxymethynolide[193].

1.4.2 NP pathway refactoring in *Streptomyces*

1.4.2.1 Significance of pathway refactoring

Traditional heterologous expression strategies have several limitations[216]. For example, native promoters from BGCs may be incompatible with the heterologous host, leading to low or no transcription of essential biosynthetic genes. Additionally, biosynthetic genes may not be properly translated due to unrecognized ribosome binding site (RBS) or codon usage differences. Furthermore, traditional heterologous expression of BGCs cannot bypass regulation by pathway-specific regulators within the BGCs[194].

To address these issues, the concept of pathway refactoring has been introduced and applied to NP heterologous expression, thanks to the advancements of synthetic biology, bioinformatics and our understanding of BGCs. Borrowed from software development, "refactoring" describes the process of rewriting the underlying code of a program to reduce the complexity without changing its functionality[195]. The first example in genetic engineering was presented in 2005, demonstrating the simplification of a phage genome through the redesign of known genetic elements to allow individual modifications[196]. Pathway refactoring involves systematic redesigning and optimization of pathways without altering their functionality. This method aims to: 1) minimize biosynthetic pathways, 2) bypass native regulatory systems, and 3) achieve controllable expression[197].

By leveraging pathway refactoring, NP production can escape regulation by both global and pathway-specific regulators. In addition, fine-tuned expression can be accomplished in industrial strains using synthetic and controlled genetic elements, enabling industrial and pharmaceutical applications. Moreover, pathway refactoring provides more opportunities to activate cryptic NP pathways beyond traditional heterologous expression, facilitating novel compound discovery and the study of biosynthetic steps[194].

1.4.2.2 Strategies and examples for pathway refactoring in *Streptomyces*

To achieve these goals, several strategies for NP pathway refactoring have been developed in *Streptomyces*[189,194,198]. First, the expression of biosynthetic genes can be regulated or limited by regulatory elements like promoters, terminators, and RBS in BGCs. Therefore, using synthetic and/or characterized regulatory elements could

release these regulations and achieve tuned enzyme levels, resulting in high yields of target products with clean production profiles (**Figure 7**)[199]. For example, the Tam gene cluster for the type II aromatic polyketide tetarimycin was initially captured from environmental DNA and activated through heterologous expression in *S. albus* with the up-regulated SARP family positive regulator *tamI* (**Figure 7A**)[200]. The cluster was predicted to include six gene operons controlled by four promoters. To refactor this pathway, all four promoter regions were substituted with synthetic promoters via the TAR method. The heterologous expression of the refactored cluster in *S. albus* demonstrated levels identical to the activated expression, clearly showing the advantage of pathway refactoring in activating silent gene clusters and releasing native regulation. In another case, constitutive promoters, including *ermE**p, *kasO*p*, and *sp44*, were utilized to replace the native promoter regions of four biosynthetic genes in the BGC for rapamycin, a type I polyketide compound used to treat many diseases[184,201]. The engineered strains showed enhanced rapamycin production, with yields up to 87.7 mg/L, demonstrating the ability of pathway refactoring to improve NP production.

To fully accomplish the goals of pathway refactoring of systematic and rational redesign of NP pathways, bottom-up reconstruction of the entire NP pathway can be carried out, rewriting the DNA sequence of BGCs[197,197]. This approach heavily relies on the understanding of the BGC and the advances of synthetic biology[189]. The first step involves eliminating all noncoding DNA, non-essential genes, and regulatory genes from the BGCs. Then, each essential gene is recoded by selecting codons to eliminate internal regulation. Protein engineering or ortholog swapping might be performed when native enzymes are less efficient. The recoded genes are arranged into artificial operons, where synthetic or characterized regulatory elements, like promoters, terminators, and RBSs, are utilized to control the expressing dynamics and conditions. Bottom-up refactoring aims to reconstruct all genetic elements, allowing the refactored pathway to bypass all native regulations. This reconstruction requires efficient DNA assembly methods for multiple fragment assembly[184,198,202].

Several publications report successful cases of bottom-up pathway refactoring. For instance, a gene cluster from *Streptomyces griseus* encoding hybrid PKS-NRPS polycyclic tetramate macrolactams (PTMs) was refactored using a plug-and-play strategy (**Figure 7B**)[203]. The gene functions in the cluster were predicted, and six characterized constitutive promoters were selected and inserted upstream of each essential coding gene. Meanwhile, the native cluster and an upregulated gene cluster with a single *ermE**p added upstream of the whole gene cluster were used as controls. All three clusters were integrated into *S. lividans* for heterologous expression. Real-time PCR analysis showed that all genes in the reconstructed operons were transcribed at significantly higher levels compared to those in the native producer and the other two engineered strains, suggesting that the transcription of a silent gene

cluster can be significantly enhanced through gene cluster reconstruction. As a result, the strain with the refactored pathway produced six potential PTMs, two of which were further confirmed by structure elucidation. In addition, bottom-up reconstruction facilitated the deletion of specific genes, allowing elucidation of the biosynthetic steps. In another case, the nitrophenyl-substituted polyketide spectinabilin, known for its antimalarial and antiviral activities, was traditionally heterologously expressed in *S. lividans*. However, no production was observed, even after knocking out the repressor[204]. To activate the gene cluster, bottom-up refactoring was carried out by using nine characterized promoters for each gene. The transcription levels of the refactored spectinabilin pathway in *S. lividans* were much higher than those of the native cluster, and unlike *S. lividans* with the native cluster, the strain with the refactored pathway clearly produced the target product.

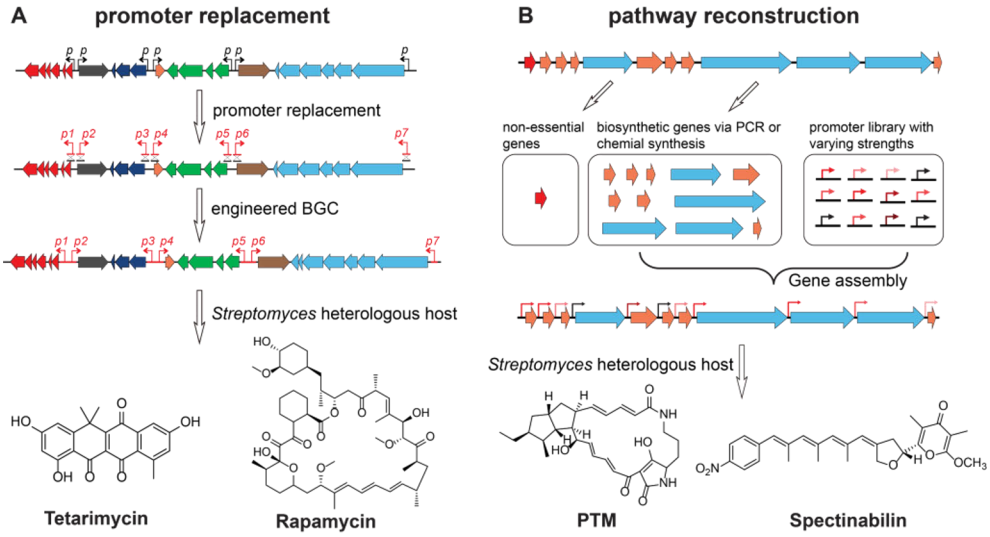


Figure 7. Strategies and examples of pathway refactoring. A, Native promoter regions on the BGCs are replaced with a set of well-characterized promoters by homologous recombination. The promoter-refactored BGCs are then transferred into a *Streptomyces* heterologous host for refactoring. **B**, The essential genes of the BGC were PCR-amplified or synthesized, after which they are re-assembled with promoters with varying strengths to create a fully refactored pathway. The synthetic BGC is subsequently expressed in heterologous host.

1.5 Combinatorial biosynthesis in *Streptomyces*

1.5.1 Principles of combinatorial biosynthesis

NPs have long been a key source of pharmacotherapy. However, their applications are often limited by suboptimal properties or the development of drug resistance[205].

Chemical modifications have been attempted to improve their biological properties, but the structural complexity of NPs makes this challenging. In this context, the concept of combinatorial biosynthesis was proposed to extend the chemistry of nature[206,207]. It aims to produce "unnatural" NPs by manipulating biosynthetic enzymes to modify functional groups, regio-chemistry and scaffold backbones. The development of combinatorial biosynthesis aligns with advancements in genetic engineering. Initially, combinatorial biosynthesis was primarily conducted through precursor-directed biosynthesis before genetic engineering became feasible[208]. Genetic engineering was first applied to this field in 1985[209]. Today, with emerging synthetic biology techniques, we can easily manipulate NP pathways and generate a large library of diverse compounds or pathways, providing more opportunities to explore chemical diversity and discover novel NPs with high potential biological activities[207].

There are several advantages of combinatorial biosynthesis[207]. First, it aims to enrich the novelty and diversity of NPs, potentially enhancing their biological features. Second, it offers a more sustainable way to produce improved NPs by exploiting their analogs compared to the traditional chemical modifications. Third, combinatorial biosynthesis is usually investigated in well-studied heterologous hosts to increase the compound titer, ultimately resulting in less expensive compounds. Additionally, structure-activity relationships can be explored through the large compound analog library created by combinatorial biosynthesis, which is crucial for drug development[206].

1.5.2 Strategies and examples of combinatorial biosynthesis in *Streptomyces*

Several strategies have been developed to enable combinatorial biosynthesis. It initially began with precursor-directed biosynthesis, leveraging the substrate promiscuity of enzymes involved in biosynthetic pathways. For instance, the polyether antibiotic monensin is encoded by type I polyketide synthases in *Streptomyces cinnamonensis*[210]. Type I polyketide synthases are composed of sequentially organized modules, where acyltransferases in each module serve as gatekeepers for accepting the substrate for the next round of chain extension. It was found that the acyltransferase domain in the fifth module of the monensin PKS can accept nonnatural malonic acid derivatives, leading to the discovery of new premonensin derivatives (**Figure 8A**).

Genetic engineering has made it possible to introduce or swap exogenous enzymes, domains, modules, and subunits from different NP pathways to achieve combinatorial biosynthesis, which has been a major strategy for many years. There are numerous examples of discovering new anthracycline analogs through this

method (**Figure 8B**). For example, gene segments from *Streptomyces purpurascens* were cloned into *S. galilaeus* ATCC 31615, the aclacinomycin producer, resulting in new derivatives of glycosylated ε-rhodomycinone[211]. Later, one of the genes was cloned into a *S. galilaeus* mutant strain producing aclacinomycin T, and the reengineered strain produced two hybrid anthracyclines[212]. Similarly, the aklavinone-11-hydroxylase DnrF was introduced into *S. galilaeus* ATCC 31133, leading to four 11-hydroxylated aclacinomycin derivatives[213]. A plasmid carrying the polyketide synthases for nogalamycin aglycone intermediates was transformed into the aclacinomycin non-producing strain *S. galilaeus* H038[214]. The fermentation of the engineered strains produced a mixture of hybrid anthracyclines, two of which were nogalamycin based compounds containing aclacinomycin sugar moieties. Epirubicin (4'-epidoxorubicin) is considered to show less cardiotoxicity than doxorubicin[215]. The production of epirubicin relied on chemical synthesis, which was inefficient and not environmentally friendly. By replacing the 4'-ketoreductase *dnmV* with *avrE* from the avermectin cluster or *eryBIV* gene from the erythromycin cluster, the production of 4'-epidaunorubicin and epirubicin was observed[216]. Recently, researchers focused on producing N,N-dimethylated anthracyclines, potential cardiotoxicity-free agents, by combining biosynthetic genes from the doxorubicin and aclacinomycin pathways[35]. They introduced the N-methyltransferases AclP and AknX2 for the methylation of TDP-L-daunosamine in the industrial strain *S. peucetius* G001. To allow the attachment of TDP-L-rhodosamine sugar moiety, they replaced the native glycosyltransferase DnrS with AknST. Although the target production of DM-DXR was not observed because of the limitation of DoxA, the engineered strain was able to produce the intermediate N,N-dimethyl-13-deoxydaunorubicin as the main product, with low production of N,N-dimethyl-daunorubicin.

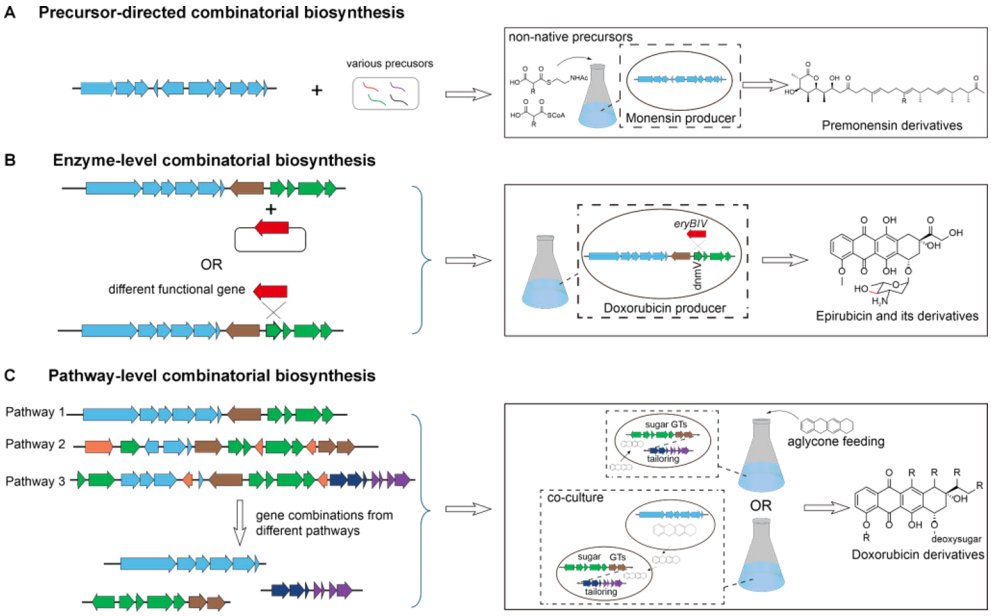


Figure 8. Strategies and examples of combinatorial biosynthesis. A, Non-native precursors are supplemented into the culture of native producers to exploit the substrate promiscuity of biosynthetic enzymes, enabling the generation of novel derivative compounds. **B**, Heterologous gene cassettes encoding modifying enzymes can be introduced to modify existing products or replace native biosynthetic genes, resulting in the production of structurally diverse analogs. **C**, Genes from different biosynthetic pathways can be assembled and expressed in a heterologous host to create new combinatorial pathways. When full pathway reconstruction is challenging, precursor feeding can be used as an alternative. In cases where it is difficult to express the complete pathway in a single host, distribution of pathway modules across multiple strains followed by co-culture offers a viable strategy.

Advancements in molecular and synthetic biology have facilitated the transfer of biosynthetic genes from various species for heterologous expression in well-characterized hosts, creating more opportunities for combinatorial biosynthesis to produce a vast library of compounds[217,218]. For instance, a combinatorial biosynthetic platform for generating doxorubicin derivatives was reported (**Figure 8C**)[101]. Briefly, a *S. venezuelae* strain resistant to doxorubicin was constructed by introducing three doxorubicin resistance genes, *drrABC*. Different deoxysugar synthesis genes from several pathways were introduced to synthesize a variety of TDP-deoxysugars. Subsequently, different GTs and their auxiliary helper proteins were screened and expressed to bind the deoxysugars to the exogenously fed aglycone ε-rhodomycinone (**Figure 8C**). Furthermore, post-PKS tailoring genes were added to some specific combinations to create doxorubicin derivatives. In total, the combinatorial biosynthesis system resulted in 17 glycosylated doxorubicin precursors, as well as doxorubicin and its two intermediates, among which seven

glycosylated derivatives of rhodomycin D are novel anthracyclines. However, this system relied on available feeding compounds. The ideal strategy is to coexpress biosynthetic genes for polyketide aglycones along with all other necessary genes into the heterologous host, which is challenging[207]. To address this issue, a one-pot combinatorial biosynthesis system for glycosylated anthracyclines was developed (**Figure 8C**)[219]. They constructed two aglycones (aklavinone and ε-rhodomycinone) cassettes and eight deoxysugars cassettes with their GTs. The cassettes were transformed into *S. venezuelae* mutant strains separately, and co-cultivation of the engineered strains accomplished the combinatorial biosynthesis, leading to 16 glycosylated anthracyclines, seven of which are new. These cases demonstrate the potential of generating novel anthracycline analogs with improved anticancer agents through combinatorial biosynthesis.

2 Aims of the Study

The thesis aims to build up a modular synthetic biology platform for refactoring of natural anthracycline pathways and discovering novel compounds through combinatorial biosynthesis.

Specifically, the aims are the following:

I Build up a synthetic biology platform consisting of all necessary elements for metabolic engineering (Study I-III).

II Re-design anthracycline pathways in a modular manner and build four BioBricks libraries based on gene functions (Study I-III).

III Utilize the BioBricks libraries for complete refactoring of natural anthracycline pathways (Study III).

IV Produce novel anthracyclines through combinatorial biosynthesis (Study II and III).

3 Materials and Methods

3.1 Strains and culture conditions

The *E. coli* TOP10 strain was used for gene manipulation and protein expression. *E. coli* ET12567/pUZ8002 and *E. coli* ET12567/pR9406 were used for intergeneric conjugation with *Streptomyces*[220]. *S. coelicolor* M1146, M1152, M1152*ΔmatAB*, and *S. lividans* K4−114 were used as the heterogeneous expression hosts. *S. albus* was used for glycosyltransferase complementation experiments.

E. coli strains were cultivated at 37°C in Lysogeny broth (LB) or LB agar plates for selection and at room temperature in 2 × TY medium for protein expression. *Streptomyces* strains were grown at 30°C in liquid media, including Tryptic soy broth (TSB), SG-TES (Soytone 10 g/L, Glucose 20 g/L, Yeast extract 5 g/L, TES free acid 5.73 g/L, $CoCl_2$ 1 mg/L), Mannitol Soya flour (MS), and E1 for strain cultivation and compound production. For solid cultivation, MS, R5 and Inorganic Salt Starch Agar No.4 (ISP4) were applied[173]. When screening the exconjugants from intergeneric conjugation, 10 mM $MgCl_2$ was added to the plates. *Streptomyces* fermentation processes for compound production are as follows: *Streptomyces* strains from glycerol stocks were revived on the solid plates for 3–4 days. Then, the *Streptomyces* cells were pre-cultured in 50 mL of TSB or SG-TES medium for 3–4 days and finally inoculated in fresh SG-TES or E1 producing media for 5–7 days in 250–2000 mL Erlenmeyer flasks.

When appropriate, chloramphenicol (25 μg/mL), ampicillin (100 μg/mL), kanamycin (50 μg/mL), apramycin (50 μg/mL), spectinomycin (100 μg/mL), viomycin (20 μg/mL), hygromycin (50 μg/mL), and nalidixic acid (25 μg/mL) were supplemented to media for the strain selection.

3.2 Feeding experiment

For the feeding experiment, around 200 μg of aklavinone or nogalamycinone was added to the cultures of engineered strains. For comparison, two control groups were included: one with the engineered strains grown under normal conditions and another with the same amount of compounds added to an empty culture.

3.3 Inhibitory test

Streptomyces pre-culture was plated on the MS plates and incubated at 30°C for 6 hours. The sterilized metal cylinders were then put on the plates. 30 µL of various concentrations of anthracyclines (0.01 g/L to 1 g/L) were added in the metal cylinders. The plates were finally incubated at 30°C for 3–4 days to see the zones of inhibition.

3.4 Plasmid and strain construction

BioBricks assembly was used for gene construction, following the 3A assembly method as previously described[124]. Briefly, the pSB1X3 (X = C, K, A, or T) series served as destination plasmids for cloning, each containing a constitutively active Red Fluorescent Protein (RFP) construct (J04450) to facilitate red-white screening. A pSB1X3 plasmid with an antibiotic resistance marker distinct from those of both insert-carrying plasmids was selected and digested with EcoRI and PstI, followed by treatment with alkaline phosphatase to prevent self-ligation. The upstream insert was digested with EcoRI and SpeI, while the downstream insert was digested with XbaI and PstI. Instead of gel purification, all digested products were purified using a PCR cleanup kit to remove residual restriction enzymes. Following all rounds of BioBricks assembly, the final constructs were cloned into BioBricks-compatible *Streptomyces* plasmids via the appropriate restriction sites. The plasmids were verified by restriction enzyme digestions and full plasmid sequencing.

For protein expression plasmid construction, enzyme coding sequences were codon-optimized based on *E. coli* codon usage and tagged with an N-terminal polyhistidine (His) tag. The genes were then cloned into the pBAD/HisB backbone using the designated restriction enzyme sites[55].

Intergeneric conjugation was applied to transform plasmids into *Streptomyces* using the standard *Streptomyces* techniques as described previously[173]. The recombinant strains were screened by corresponding antibiotic stress. A standard technique of heat shock transformation was applied to *E. coli*.

3.5 *Streptomyces* integrating plasmids

The *Streptomyces* integrating plasmids used in this study (**Table 6**) were collected and modified by our collaborator Prof. S. Eric Nybo.

Table 6. Integrating plasmids used in the synthetic biology platform

Vector name	Resistance in *Streptomyces*	Integrase	Note	Reference
pSET154BB	Apramycin	ΦC31	pSET152BB with 5'-*tt-sbi-A* and 3'-*fd* phage terminators	124,153,155
pOSV821	Apramycin	ΦC31	pOSV801 with 5′-*tt-sbi-A* and 3′-*fd* phage terminators	123,153,155
pOSV808	Hygromycin	VWB		123
pOSV820	Kanamycin	pSAM2	pOSV811 with 5′-ECK120033736 and 3′-ECK120033737 terminators	123,151,156
pENTG3	Viomycin	ΦTG1	pENTG1 with 5'-L3S2P21 and 3'-L3S3P41 terminators	124,151,156
pENSV3	Spectinomycin	SV1	pENSV1 with 5'-ECK120010818 and 3'-ECK120029600 terminators	124,151,156

3.6 Compound extraction

Two methods were utilized for extracting compounds from the cultural supernatant. For anthracycline aglycones, the cultural supernatant was added with the same volume of ethyl-acetate with 1% acetic acid in centrifuge tubes. The phases were separated by 5 min of centrifugation at 4,000 g after 10 min of vortexing. The ethyl-acetate phase was collected and dried with vacuum. Methanol was added to suspend the extracts for analysis.

For glycosylated anthracycline compounds, the metabolites were collected from the cultural supernatant and recovered by overnight binding to 20 g/L of LXA-1180 resin at 8°C with gentle shaking. The resin was collected by discarding the medium and washed with Milli-Q water. The absorbed compounds were extracted with methanol. For large scale purification, methanol was evaporated using a rotary evaporator and the dried compounds were resuspended in minimal volume of methanol.

3.7 Analytical HPLC method

The extracted compounds were analyzed by UHPLC (Shimadzu Nexera LC-40 system with a diode array detector) using a Phenomenex Kinetex C18 column (2.6 μm, 100 Å, 4.6×100 mm) or HPLC (Agilent 1260 Infinity II) using a Poroshell 120 Phenyl-Hexyl Column (2.7 μm, 4.6 mm×100 mm). Method A: Solvent A: 15% CH_3CN/0.1% HCOOH; solvent B: CH_3CN; flow rate: 0.3 mL/min; 0–2 min, 0% B; 2–20 min, 0–40% B; 20–24 min, 100% B; 24–29 min, 0% B. Method B: Solvent A: 0.1% HCOOH; solvent B: CH_3CN/0.1% HCOOH; flow rate: 0.5 mL/min; 0–10 min, 5–95% B; 10–13 min, 95% B; 13.1–15.1 min, 95–5% B.

3.8 Compound purification and structure elucidation

Methanol extracts were first purified using silica chromatography with high-purity grade silica (pore size 60 Å, 230–400 mesh particle size), and, if needed, further purified with Sephadex LH-20 (MeOH; 2.5 × 65 cm). Fractions containing the compounds of interest were pooled, evaporated using a rotary evaporator, and then re-suspended in a minimal volume of methanol for semi-preparative HPLC.

Semi-preparative HPLC was carried out using a Shimadzu LC-20AP/CBM-20A system with a diode array detector and a Phenyl-Hexyl (5 µm, 100 Å, 250 × 21.2 mm) Kinetex column, along with an Agilent 1260 Infinity II Prep HPLC system equipped with a diode array detector and a Gemini C18 column (5 µm, 110 Å, 250 × 10 mm). Fractions containing pure compounds were evaporated using a rotary evaporator and desiccator, then re-suspended in deuterated solvents for NMR analysis. Method A: Solvent A: H_2O/0.1% HCOOH; Solvent B: CH_3CN; flow rate: 20 mL/min; 0–2 min, 0% B; 2–22 min, 0–100% B; 22–24 min, 100% B; 24–29 min, 0% B. More methods can be found in original publication III.

NMR spectra were recorded with a 600 MHz Bruker AVANCE-III system with liquid nitrogen cooled Prodigy TCI cryoprobe, a 500 MHz Bruker AVANCE-III system with liquid nitrogen cooled Prodigy BBO cryoprobe, Bruker Avance NEO 400 MHz (Bruker BioSpin Corporation, Billerica, MA), Varian 500 MHz (Agilent, Santa Clara, CA, USA) spectrometers, and/or the Bruker Avance NEO 600 MHz NMR spectrometer with triple channel TCI 5-mm cryoprobe. All NMR spectra were processed in Bruker TopSpin 4.1.3 version, and the signals were internally referenced to the solvent signals or tetramethylsilane.

High resolution electrospray ionization mass spectra were recorded using Waters Acquity RDa Detector or AB SCIEX Triple TOF® 5600 system.

3.9 Protein purification and enzymatic reactions

Protein purification was performed using polyHis-tag affinity chromatography with TALON SuperFlow resin, followed by desalting with PD-10 columns. Enzymes were concentrated using Pierce™ Protein Concentrators PES (10K MWCO), and the purity of enzymes were assessed by SDS-PAGE.

Enzymatic reactions were conducted as previously described[60]. Each reaction was set up in a 1.5 mL Eppendorf tube containing 200 µL of reaction buffer (50 mM phosphate, 50 mM NaCl, pH 7.5). Substrates were added to a final concentration of 10–50 µM, followed by enzyme addition at a final concentration of 5–10 µM. For reactions involving DnrF, RdmE, KstA1516, and KstA1011, NADPH was included at a final concentration of 0.5 mM. When RdmB was present, 10 µM DTT and 400 µM SAM were also added. Reactions were incubated at 30°C for 1–4 hours. After

incubation, reaction products were extracted with $CHCl_3$ by vigorous vortexing. The organic phase was dried under vacuum, resuspended in MeOH, and analyzed by HPLC.

3.10 Cancer cell line viability assay

Cytotoxicity assays were performed at the University of Kentucky in triplicate using human cancer cell lines A549 (non-small cell lung cancer), PC3 (prostate cancer), and HCT116 (colorectal cancer), as well as Merkel cell carcinoma lines MKL1 and MCC26, following previously described protocols[221]. Briefly, an initial viability screening was conducted at 80 μM to identify active compounds. Compounds that reduced cancer cell viability to below 50% were then evaluated at 10 μM, and their EC_{50} values were determined. Doxorubicin and nogalamycin served as positive controls, while vehicle (DMSO) was used as a negative control.

3.11 Artificial intelligence tool usage

ChatGPT and Microsoft copilot were used to correct grammar and polish sentences in accordance with the guideline on the use of artificial intelligence in research at the University of Turku.

4 Results and Discussion

4.1 Establishment of a synthetic biology platform for anthracycline biosynthesis (Study I-III)

As fertile producers of numerous NPs, *Streptomyces* remain the optimal cell factories for genetic engineering to produce value-added compounds[79]. Although many genetic tools have been developed for *Streptomyces*, it still lacks a comprehensive synthetic biology platform for biosynthesis of desired metabolites. To address this issue, I aim to build a synthetic biology platform by testing and evaluating all the necessary genetic tools for metabolic engineering, including *Streptomyces* chassis, vectors, promoters, terminators, RBS, and biosynthetic genes.

4.1.1 *Streptomyces* chassis evaluation (Study I)

An optimal *Streptomyces* chassis is fundamental for the heterologous expression of NPs, making its evaluation a priority. In this thesis, two well-studied *Streptomyces* hosts, *S. coelicolor* and *S. lividans*, were initially selected and evaluated by our collaborators through the production of minPKS *snoa123* from the nogalamycin pathway[79]. Two engineered *S. coelicolor* strains, M1146 and M1152, were studied due to their: 1) clear genetic background, 2) removal of interfering endogenous gene clusters, and 3) enhanced secondary metabolite production (M1152). *S. lividans* is closely related to *S. coelicolor*, but has the advantage of accepting methylated DNA, facilitating easier genetic manipulation. The modified version K4−114 was characterized here because the endogenous actinorhodin cluster was deleted for a clean production background, while retaining the enhanced secondary metabolite properties of the parental strain TK24[83].

First, the production of SEK15 from codon-optimized s*noa123* in three engineered strains was evaluated on R5 solid plates. All three strains produced similar levels of SEK15 at 30–40 mg/L. However, *S. lividans* K4−114 also produced large quantities of undesired prodiginines, leading to its exclusion from further experimentation[83]. Given that *S. coelicolor* M1152 contains a mutation regulating RNA polymerase B that is absent in *S. coelicolor* M1146, *S. coelicolor* M1152 was selected for further evaluation (**Figure 9B**). Knockout of *matAB* has been noted to

eliminate the mycelial aggregation phenotype and improve biomass accumulation. Therefore, *S. coelicolor* M1152Δ*matAB* was constructed and compared with *S. coelicolor* M1152 by expressing wild-type *snoa123*. As shown in **Figure 9C**, the knockout of *matAB* led to a threefold improvement to 40 mg/L in SG-TES medium. Based on these results, *S. coelicolor* M1152Δ*matAB* was selected as the chassis for further heterologous expression.

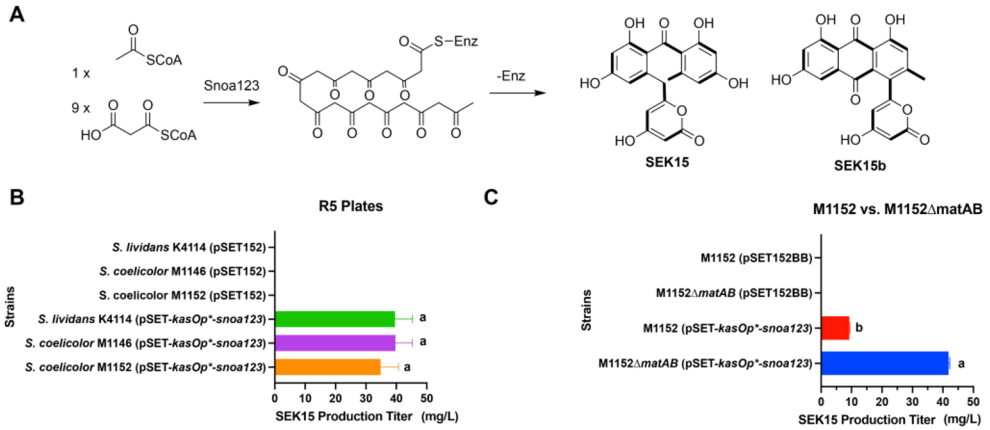


Figure 9. ***Streptomyces*** **chassis evaluation by C-20 minPKS engineering**. **A**, biosynthetic pathway scheme depicting the synthesis of SEK15 and SEK15b from the minPKS Snoa123. **B**, SEK15 titers on R5 solid agar plates from codon-optimized *snoa123* in three heterologous expression hosts. **C**, SEK15 titers in SG-TES liquid medium from wild-type *snoa123* expressed in two *S. coelicolor* engineered hosts. ANOVA analysis was performed to assess statistical significance between strains. Different letters next to the columns indicate statistically significant difference (p<0.05).

4.1.2 Genetic elements in the platform (Study I-III)

BioBricks assembly was employed to leverage the construction of standardized genetic parts. Unlike Golden Gate and Gibson assembly, BioBricks genetic parts are reusable in various applications and are well-suited for the parallel assembly of multiple branching pathways. Additionally, BioBricks design addresses the challenges associated with multi-gene assembly techniques (e.g. Gibson assembly), especially involving low recombination efficiency of multiple high G-C content *Streptomyces* DNA[166].

To implement BioBricks assembly in our synthetic biology platform, all genetic elements were designed and synthesized according to the BioBricks [RFC-10] standard[124]. The plasmids, promoters, terminators, and biosynthetic genes were synthesized or cloned to lack internal EcoRI, XbaI, SpeI, and PstI restriction sites, with silent mutations introduced at these restriction sites if necessary. Furthermore, all genetic elements were provided the BioBricks prefix (5'-GAATTCGCGGCCGC

TTCTAGAG-3') and suffix (5'-TACTAGTAGCGGCCGCTGCAG-3') sequences to enable isocaudomer cloning.

The anthracycline pathways are complex and may encompass more than 30 genes[18]. Consequently, to accommodate all necessary biosynthetic genes, a modular design was employed to dissect the anthracycline pathways into smaller sections, facilitating plasmid manipulation. To ensure stable production in the heterologous host, integrating BioBricks compatible plasmids were selected for the platform (**Table 6**), which can be easily transformed into hosts via intergeneric conjugation. Plasmids with different integrases and resistance markers are compatible within the same host.

Promoters and terminators are critical for controlling transcription. Therefore, a well-established platform should contain characterized promoters with varying strengths. Constitutive promoters are selected to release pathway control from native regulations. Specifically, four natural promoters, *rpsLp*(CF)[132], *gapdhp*(EL)[132], *SF14*[135], and *p15*[140], screened from different species were selected. Additionally, mutated versions of commonly used natural promoters for *Streptomyces* were also selected, including *kasOp**[126], *ermEp**[129], *p21*[137], *sp41-sp44*[132]. The relative strength of the promoters has been summarized in the reference recently[222]. For terminators, rho-independent intrinsic terminators, including *fd*[153], *tt-sbi-A*[155], B0015[154], and B0017[154], were chosen to insulate the operons.

Anthracycline biosynthetic genes were codon-optimized based on the native codon preference for *S. coelicolor*. The well-studied BBa_B0034 RBS (5'-AAAGAGGAGAAA-3') was included during gene synthesis. In case of failure in the expression of codon-optimized genes, wild-type sequences with upstream native RBS regions were synthesized according to the BioBricks standard.

4.2 *De novo* anthracycline biosynthesis (Study I-III)

4.2.1 Modular design for anthracycline biosynthesis (Study I-III)

NPs like anthracyclines are tightly regulated by the regulatory genes within their BGCs, which often limits their heterologous expression[216]. To address this issue, I discarded the regulatory genes and utilized strong constitutive promoters to drive the expression of biosynthetic genes, thereby releasing them from regulatory network constraints. In addition, I modularly redesigned anthracycline pathways by arranging the biosynthetic genes into four BioBricks-compatible *Streptomyces* plasmids according to their gene functions. Specifically, the ΦC31 integrase plasmid pSET154BB carries genes for polyketide aglycones; glycosyltransferases, transporters, and resistance genes are assembled into the VWB integrase plasmid

pOSV808; pOSV820 with pSAM2 integrase is responsible for TDP-carbohydrates; and potential post-PKS tailoring genes are cloned into pENTG3 with ΦTG1 integrase. All four plasmids are compatible within a single *Streptomyces* strain.

4.2.1.1 Polyketide aglycone biosynthesis (Study I-III)

4.2.1.1.1 minPKS engineering

For C-20 aglycones, the minPKS *snoa123* operon has been studied in section **4.1.1**. Our collaborators then optimized production of C-21 polyketides. The wild-type gene cassettes minPKS *aknBCDE2F* from *S. galilaeus* and *dpsABCDG* from *S. peucetius* were synthesized as a single BioBricks part due to unsuccessful attempts at codon optimization and decoupling translational coupling. Then, they were flanked with either the medium strength promoter *ermE*p* or the strong promoter *kasOp** and transformed into *S. coelicolor* M1152Δ*matAB* for C-21 minPKS engineering (**Figure 10B**). The synthesis of the intermediate UWM7 from AknBCDE2F was significantly better than that from DpsABCDG. Additionally, the strong promoter *kasOp** enhanced the expression of the minPKS, resulting in a titer of 7.5 mg/L in SG-TES medium.

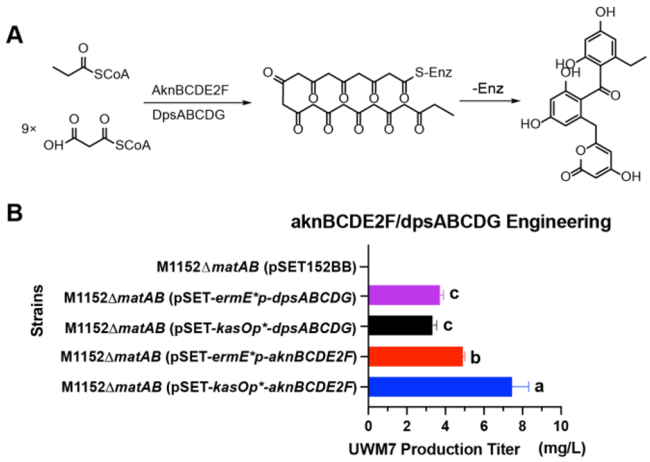


Figure 10. C-21 minPKS engineering. A, Biosynthetic pathway scheme illustrating the synthesis of UWM7 from the AknBCDE2F/DpsABCDG minPKS. **B**, UWM7 titers in SG-TES liquid medium from wild-type *dpsABCDG*/*aknBCDE2F* expressed under different promoters in *S. coelicolor* M1152Δ*matAB* engineered hosts. ANOVA statistical analysis was performed as described in **Figure 9**.

4.2.1.1.2 Nogalonic acid and aklanonic acid engineering

After optimizing the minPKS operons, the subsequent operons carrying ketoreductase, aromatase, cyclase, and oxygenase were cloned and assembled with the minPKS operons to create plasmids for producing stable tricyclic anthracyclinone intermediates, nogalonic acid and aklanonic acid (**Figure 11A**). For nogalonic acid biosynthesis, the strong promoter *p15* was used for genes from the aclacinomycin and doxorubicin pathways. The results showed that the oxygenases SnoaB/AknX are necessary, although a small amount of the target compounds was detected from self-deoxygenation. The combination of *snoaDEMB* was found to be optimal among the four options (**Figure 11B**). It is worth mentioning that the first round of nogalonic acid production failed, which was later found to be caused by inaccuracies in the sequencing data for *snoaD*, *snoaE*, *snoaM*, and *snoaB*. Re-sequencing of the nogalamycin BGC allowed correction of the sequencing errors and cloning of functional gene constructs. For aklanonic acid, combinations of *aknAE1WX* and *dpsEFYdnrG* were fused with the *kasOp** promoter and assembled with *aknBCDE2F* operons. Among these, the *kasOp*-aknBCDE2F+kasOp*-aknAE1WX* combination produced the highest yield of aklanonic acid equivalents at 30 mg/L (**Figure 11C**).

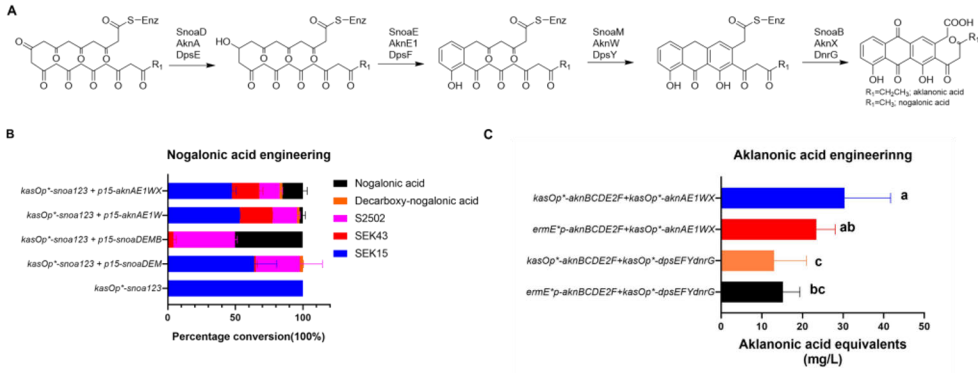


Figure 11. Nogalonic and aklanonic acid engineering. A, Biosynthetic pathways of aklanonic acid and nogalonic acid. **B**, Percentages of nogalonic acid and its intermediates from different combinations of promoters and biosynthetic genes. Percent conversion was calculated based on the integration of peak areas at 290 nm. **C**, Yields of aklanonic acid-derived shunt products from various combinations of promoters and biosynthetic genes. ANOVA was conducted as previously described.

4.2.1.1.3 Anthracyclinone pathway engineering and optimization

To obtain full pathways for producing polyketide aglycones for anthracyclines, different combinations of *O*-methyltransferase genes, fourth-ring cyclases, and 7-

ketoreductases were fused with the stronger *sp44* promoter and cloned into the integrating plasmid pENTG1. These expressions were then evaluated in optimal nogalonic acid and aklanonic acid-producing strains. The genes were selected from the doxorubicin pathway (*dnrCDE*), aclacinomycin pathway (*aknGHU*), and nogalamycin pathway (*snoaLCF*). In addition, the putative fourth-ring cyclase *kyc34* from the keycin pathway was tested in combination with *snoaLFkyc34*. The configuration of the C9 position is determined by fourth-ring cyclases, where DnrD and AknH contribute to the 9R configuration, while Kyc34 and SnoaL produce 9S-configured anthracyclinones.

After combining these genes with nogalonic and aklanonic acid strains, four different aglycones were synthesized (**Figure 12A-C**), including three natural aglycones: aklavinone (**1**), nogalamycinone (**2**), and auramycinone (**3**), and an unnatural aglycone, 9-*epi*-aklavinone (**4**). Among them, the optimal combinations, marked with star symbols, were assembled into a single plasmid pSET154BB to characterize the yields in two production media, SG-TES and E1 (**Figure 12D–F**). The yields were quantified based on the standard curve of **2**. The results clearly showed that the yields in SG-TES were higher than those in E1 medium. However, the SG-TES medium also led to the production of large amounts of 7-deoxy compounds, which are typical anthracycline degradation products[223]. The structures of the intermediates were elucidated by HR-MS and NMR (data not shown; more details in original publication I). The production yields varied from 1.5–4.5 mg/L in SG-TES, while the tricyclic intermediates presented titers of over 20 mg/L when incorporating two plasmids for the biosynthesis of aglycones. This indicated that the expression cassettes for the single plasmids were limited. It was presumed that, because of no internal terminators for operon insulation, the single-plasmid constructs could cause long transcripts, which were unstable and easy to be degraded, hereby limiting the transcription level of the genes.

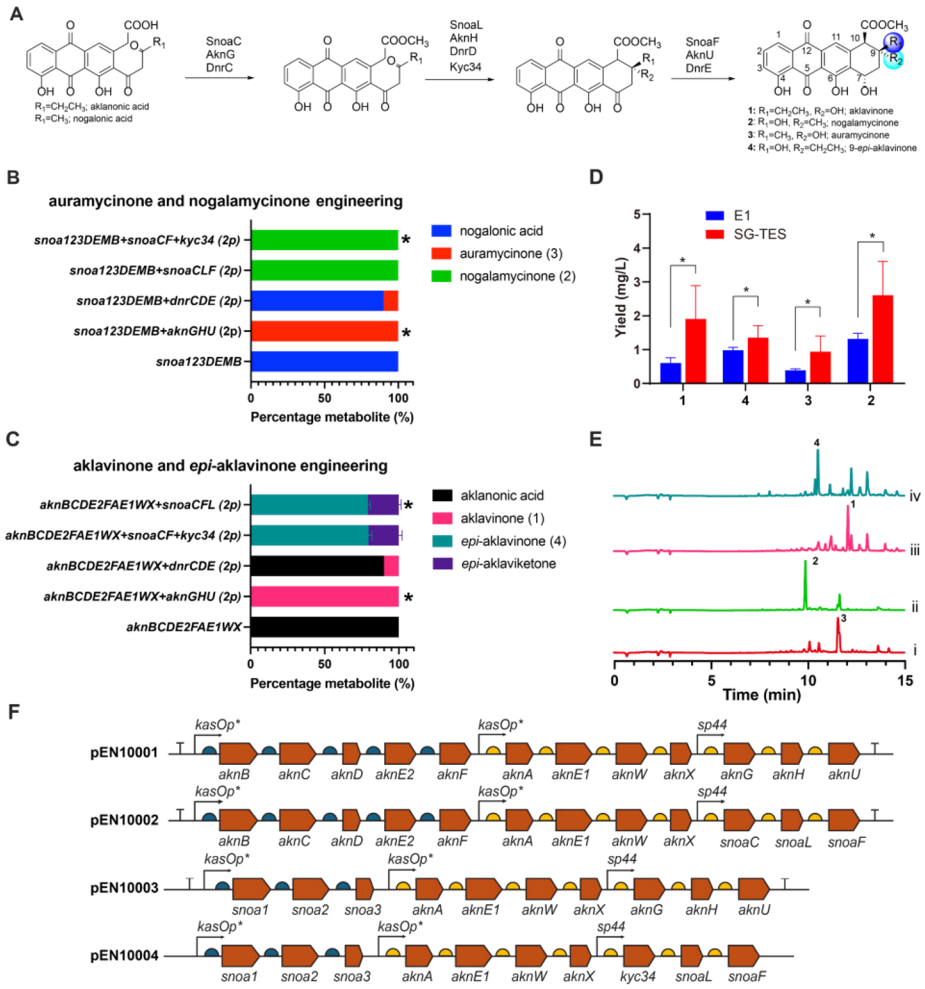


Figure 12. Full aglycone pathway engineering. **A**, Late steps for aglycone biosynthesis. **B-C**, Percentage conversion of four aglycones and their intermediates. The best-performing combinations were marked with a star symbol. **D**, Production yields of four anthracyclinones from one single plasmid in SG-TES and E1 media. The t-test was performed for the statistical analysis. *p<0.05. **E**, Representative HPLC traces illustrating the production of four anthracyclinones, labeled with star symbols. i, auramycinone; ii, nogalamycinone; iii, aklavinone; iv, 9-*epi*-aklavinone. **F**, SBOL diagrams depicting gene cassettes for aglycone biosynthesis.

Therefore, several strategies were applied to further improve yields of aglycones (**Figure 13**). Terminators and ribozyme-based insulators were used between the operons. Ribozyme-based insulators operate by cleaving the 5'-untranslated region of the mRNA, creating a hairpin loop that stabilizes the transcript[224]. Different promoters were also utilized to tune expressions, and the sequences and orders of the biosynthetic genes were evaluated.

To optimize the nogalamycinone plasmid, the sequences of minPKS *snoa123* were codon-optimized according to the codon usage of *S. coelicolor*, and translation coupling was released (**Figure 13A**). Large amounts of SEK15 were detected, indicating that the expression of *snoa123* was too high in comparison to genes downstream of the pathway. Therefore, stronger promoters were used for the two subsequent operons. Additionally, the fourth ring cyclase *snoaL* in the third operon was placed immediately after the promoter to further enhance production. The *fd* terminator and ribozyme-based insulator *riboJ* were employed to insulate the operons. These combined strategies led to a twelve-fold increase in production, reaching 38 mg/L in E1 medium.

For 9-*epi*-aklavinone optimization (**Figure 13B**), the *kasOp** promoter for the first two operons was replaced by the stronger promoters *sp41* and *sp42*, respectively. In addition, three ribozyme-based insulators were placed after each promoter to insulate and stabilize the transcripts. The expression of *snoaL* was enhanced using the same strategy as in nogalamycinone optimization. As a result, the production of 9-*epi*-aklavinone reached 4.5 mg/L in E1 medium.

Aklavinone is the core structure of many clinically used anthracyclines. Therefore, improving its production yield is crucial for enhancing the subsequent biosynthesis of glycosylated anthracyclines. I focused more on the first and third operons, as the second operon has been functioning well in the nogalamycinone plasmids (**Figure 13C**). A stronger promoter *sp41* was used to replace the *kasOp** promoter. Similarly, the expression of cyclase *aknH* was improved by changing its order in the operon. Furthermore, two extra terminators were inserted between the operons. These improvements resulted in a 1.5-fold increase, yielding 1.6 mg/L in E1 medium.

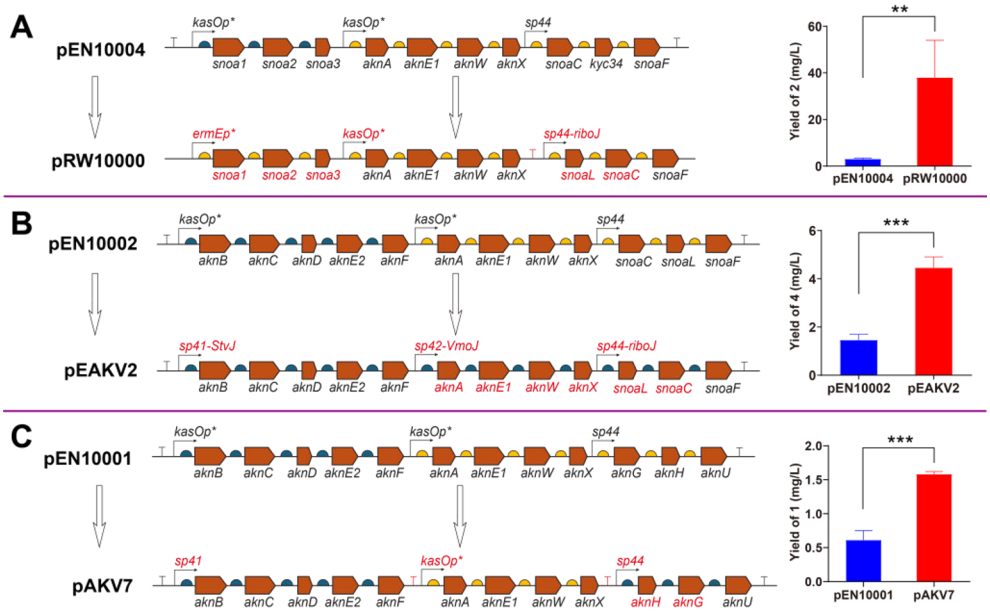


Figure 13. Further optimization for aglycone production. A-C, The SBOL diagrams illustrate the optimization strategies for aglycone biosynthesis and bar charts show the yield improvements. The changes are marked in red. The t-test was performed for the statistical analysis. *p<0.05; **p<0.01; ***p<0.001.

Prior to this study, there were only a few reports on the heterologous expression of anthracyclinone aglycones[23,219,225–227]. In the cases of auramycinone and nogalamycinone, the biosynthetic genes were obtained through restriction enzyme digestion of the native clusters[23,225]. As a result, the reconstructed pathways retained additional open reading frames and native regulatory elements, which limited compound production due to uncontrolled or suboptimal gene expression. A more recent study reported the heterologous expression of aklavinone, in which biosynthetic genes were amplified by PCR and placed under *ermE*p* promoter to enhance expression[219]. Nevertheless, native regulatory sequences were still retained during PCR amplification, which may have affected pathway efficiency. Although the activator gene *dnrI* was included, the reported yield of aklavinone was only 0.5 mg/L, considerably lower than the 1.6 mg/L achieved in this study. Notably, there had been no prior reports of heterologous expression of the unnatural compound 9-*epi*-aklavinone before this work. This compound was identified upon the introduction of a gene cassette containing *snoaL* into a *S. peucetius* mutant strain[225].

In this study, all four core anthracycline aglycones were successfully expressed in a heterologous host, establishing a foundational framework for the complete biosynthesis of anthracyclines. Despite improvements made to the expression system and fermentation process, the production yield of aklavinone—a key precursor for

many anthracycline—remains low and warrants further optimization to facilitate its practical application in engineered biosynthetic pathways.

4.2.1.2 Resistance and glycosyltransferase (Study III)

Anthracyclines that disrupt DNA structure and/or inhibit DNA function also exhibit cytotoxicity towards their producers[14]. To ensure the survival of the chassis host during anthracycline production, genes for transporters and self-resistance are needed. To construct a robust *Streptomyces* strain with broad resistance to anthracyclines, the ABC-transporters *drrAB* from the doxorubicin BGC were introduced. Additionally, the UvrA-like proteins *drrC* from the doxorubicin pathway and *snorO* from the nogalamycin pathway were included to repair potential DNA damage in the producer host strains. The expression of resistance genes was enhanced with the strong constitutive promoter *kasOp** to enable high production of anthracyclines (**Figure 14A**). An inhibitory test was conducted to assess the functionality of the resistance genes. As shown in **Figure 14B**, the starting strain *S. coelicolor* M1152Δ*matAB* was susceptible to 1 and 0.5 g/L daunorubicin and DM-DXR, while no inhibition was observed in strains with *drrABC* and *snorO* under the same conditions. Nogalamycin showed less toxicity to *S. coelicolor* strains even without the presence of *drrABC* and *snorO* genes. The inhibitory results gave confidence that the selected resistance genes could provide sufficient self-defense and drug efflux against various anthracycline analogs, thereby obtaining a robust host with broad anthracycline resistance for heterologous expression.

To accomplish the prominent 7-*O*-glycosylation, the transferase module was built based on well-studied anthracycline pathways, with the strong promoter *gapdhp*(EL) applied to ensure sufficient expression. To test the functionality of GTs and P450 enzymes, complementation experiments and compound feeding were conducted. For instance, the codon-optimized GT *snogE* and P450 enzyme *snogN* from the nogalamycin pathway with synthetic RBS BBa_0034 were complemented in the knock-out strains *S. albus snogE*Δ and *snogN*Δ, respectively. However, no expected complementation was observed. Then, we went back to their wild-type sequences, restoring the production of the expected compound nogalamycin F (**5, Figure 14D**). In addition, a feeding experiment was conducted for strains harboring operons for wild-type GT *aknS* and P450 enzyme *aknT* from the aclacinomycin pathway. The strain also carried the operon for producing TDP-L-rhodosamine. After feeding with compound **1**, the engineered strain produced the monoglycosylated compound aclacinomycin T (**6**, **Figure 14C**), proving the functionality of both GT and TDP-L-rhodosamine operons.

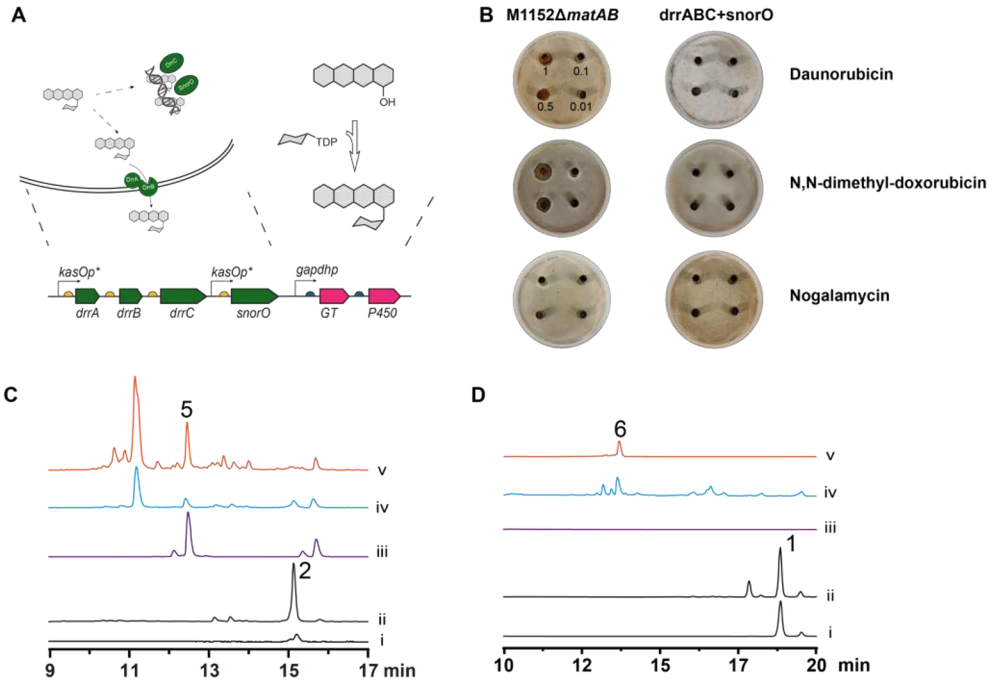


Figure 14. Construction and functional test of resistance and GT modules. A, Assembly strategy for the resistance and GT module. **B**, Functionality test of resistance genes using an inhibitory test. Left panel, plates of the parental strain *S. coelicolor* M1152Δ*matAB*; right panel, plates of the strains harboring resistance genes *drrABC* and *snorO*. The three anthracyclines were added to the metal cylinders on each plate at concentrations of 1, 0.5, 0.1, and 0.01 g/L. **C**, Complementation experiment for *snogE* and *snogN*. i, *snogN*Δ strain; ii, *snogE*Δ strain; iii, authentic standard of **5**; iv, *snogN*Δ complemented with *snogN*; v, *snogE*Δ complemented with *snogE*. **D**, Functionality of GT *aknST* cassette assessed by a feeding experiment. i, authentic standard of **1**; ii, feeding of **1** to an empty culture; iii, strain harboring operons for GT *aknST* and the corresponding sugar unit; iv, feeding of **1** to the strain harboring operons for GT *aknST* and the sugar unit; v, authentic standard of **6**.

4.2.1.3 TDP-carbohydrate biosynthesis (Study III)

Anthracyclines are decorated with various 6-deoxysugar units. I aim to build a BioBricks library consisting of plasmids to produce numerous TDP-carbohydrates. Previously, this possibility had been verified by constructing plasmids producing several neutral deoxysugars[228]. However, intermediates were found in some strains, possibly due to the use of a single promoter for the entire operon. In my design, DesIII and DesIV were chosen to obtain the key intermediate TDP-4-keto-6-deoxy-D-glucose, which is common to all amino- and deoxysugar pathways (**Figure 3**), as they functioned well in previous research[228]. I then advanced the pathways by adding another strong promoter, *p21*, when branching the sugar pathways.

First, I cloned gene cassettes for TDP-aminosugars, including the core construct for TDP-L-daunosamine (*desIII+IV*, *snogHI, snogFG*), with a third promoter, *gaphdp*(EL), inserted to ensure sufficient expression. TDP-L-rhodosamine was obtained by including the *O*-methyltransferases (*snogAX*) in the TDP-L-daunosamine construct, while TDP-L-acosamine production was achieved by changing the 4'-ketoreductase in the TDP-L-daunosamine construct (*snogG* to *aveBIV*). I also built gene sets for several neutral TDP-deoxysugars, such as TDP-3',4'-demethyl-L-nogalose (*snogG2FC, snogY*) from the nogalamycin pathway, which lacks methylations at the C3' and C4' positions compared to TDP-L-nogalose. Similarly, the third promoter *gapdhp*(EL) was introduced for *snogY*. The full biosynthesis of TDP-L-nogalose was achieved by including the remaining methylase genes *snogLM*. In addition, TDP-2-deoxy-L-fucose (*aknLMQ*) and TDP-L-rhodinose (addition of *aknP*) from the aclacinomycin pathway were biosynthesized (**Figure 15A**). The third strong promoter *sp44* was used to enhance the transcription of *aknPQ*, as their expressions were limited without the additional promoter (data not shown).

However, since the direct identification of TDP-carbohydrates from engineering is challenging, I proceeded to detect formation of glycosylated compounds instead. I conducted aglycone feeding experiments as described earlier. The biosynthesis of TDP-L-rhodosamine was confirmed by feeding compound **1** to strains harboring both GT and carbohydrate modules (**Figure 14D**). A similar feeding experiment was conducted for TDP-L-3',4'-demethyl-nogalose biosynthesis. After adding compound **2** to the strains with GT and carbohydrate modules, the production of L-3',4'-demethyl-nogalose-nogalamycinone (**7**) was observed and confirmed (**Figure 15B**). It is worth mentioning that although there are several papers proposing the biosynthetic pathways of these TDP-deoxysugar pathways previously[36,37,227], this is the first time these pathways have been experimentally verified *in vivo*, thus completing the understanding of the last remaining steps in the biosynthesis of aclacinomycin and nogalamycin.

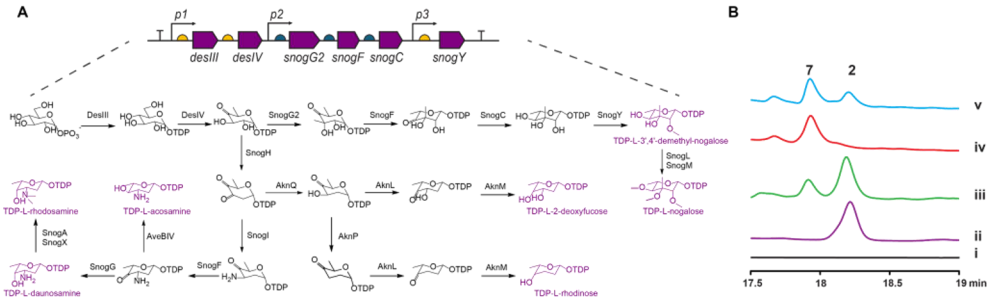


Figure 15. Construction and functionality of TDP-carbohydrate units. A, Assembly strategy for sugar plasmids, with the construction of TDP-L-3',4'-demethyl-nogalose demonstrated as an example. **B**, Feeding experiment to verify the functionality of biosynthetic genes for TDP-L-3',4'-demethyl-nogalose. i, strain harboring GT SnogEN and resistance genes; ii, feeding of **2** to an empty culture; iii, *snoaW*Δ strain; iv, feeding of **2** to a strain containing GT and TDP-L-3',4'-demethyl-nogalose module; v, co-elution of iv and purified compound **2**.

4.2.1.4 Post-PKS tailoring reactions (Study II)

Many anthracycline pathways harbor post-PKS tailoring genes for modifications on the core structure. Therefore, a genetic module containing diverse post-PKS tailoring genes was constructed to fulfill the biosynthesis of anthracyclines. The tailoring enzymes involved in modifications of the most common positions on the aglycones were selected from well-characterized anthracycline pathways. For example, FAD-dependent monooxygenase DnrF from the doxorubicin pathway and RdmE from the rhodomycin pathway were recruited for C11-hydroxylation[62,63]. The 15-methylesterase RdmC and the methyltransferase-like RdmB were utilized for C10-demethylation and C10-hydroxylation, respectively[60,61]. In addition, the isozyme of RdmC, EamC from the komodoquinone pathway, was included for C10-demethylation, while the methyltransferase-like EamK from the same pathway was incorporated to achieve C10-decarboxylation. Furthermore, the short-chain aldol reductase KstA16 and the cyclase-like KstA15 from the kostinostin pathway were chosen to jointly catalyze C1-hydroxylation[57]. The reaction cascade was further extended to 4-hydroxyl regioisomerization by the NmrA-like short-chain dehydrogenase/reductase enzymes KstA11 and KstA10[57].

Post-PKS tailoring module was constructed by assembling the gene cassettes on the pENTG3 plasmid with *gapdhp*(EL) fused for sufficient expression levels. The constructs were tested in all four aglycone producers. For DnrF reactions (**Figure 16**), it was revealed that DnrF could catalyze C11-hydroxylation for all four aglycones, but the enzyme displayed a preference for 9R configuration (**1** and **3**) instead of 9S-configured anthracyclines (**2** and **4**). The corresponding enzymatic reactions were carried out *in vitro* to probe the substrate promiscuity, yielding consistent results (data not shown; more details in original publication II). The

production of both *in vivo* and *in vitro* samples was confirmed by MS analysis. A similar strategy was applied to the remaining catalytic reactions. I next focused on the EamCRdmB and EamCK combinations for C10-hydroxylation and C10-decarboxylation, respectively (**Figure 16**). Both *in vivo* and *in vitro* results demonstrated that all four aglycones were quantitatively converted by EamCK into 10-decarboxylated products. On the other hand, only **1** and **3** were converted into 10-hydroxylated products, indicating RdmB exhibits a preference toward 9R-configured substrates. The unreacted EamC products were eventually degraded into 10-decarboxylated products, as previously reported.

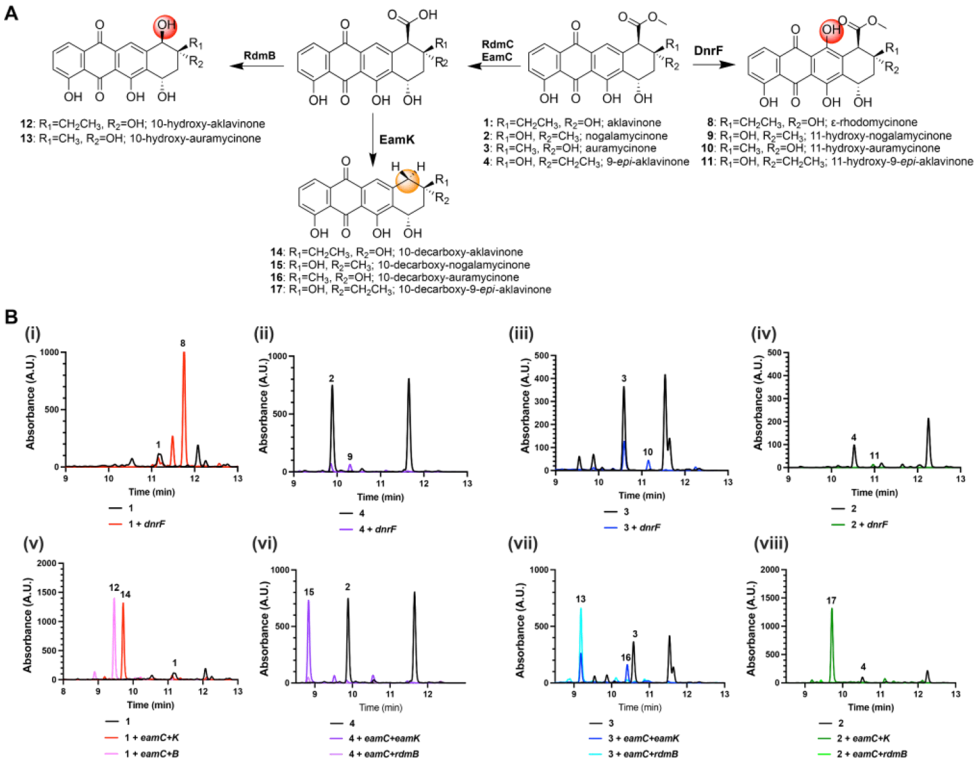


Figure 16. Metabolic engineering of 11-hydroxylated, 10-hydroxylated and 10-decarboxylated anthracyclinones. A, Post-PKS tailoring pathways for C11-hydroxylation, C10-hydroxylation and C10-decarboxylation for anthracyclinones. **B**, HPLC traces of strains engineered with expression constructs encoding **1**−**4** and their corresponding tailoring genes, or control strains producing only **1**−**4**; i-iv, traces recorded at 490 nm; v-viii, traces recorded at 430 nm.

The RdmECB combination was also investigated to achieve 10,11-hydroxylation. Since DnrF, RdmE and RdmB are more selective for 9R-configured substrates, the production profiles of 9S-configured anthracyclinones (**2** and **4**) were

identical to those in EamC and RdmB reactions. In contrast, 10,11-hydroxylated anthracyclinones were observed in **1** and **3** strains (**Figure 17A-B**). Finally, I studied C1-hydroxylation by KstA1516, and the expected 1-hydroxylated products were detected as the main products in all strains (**Figure 17**C-D). Furthermore, C4-hydroxyl regioisomerization reactions were achieved in all strains by adding *kstA10* and *kstA11* genes.

The previous studies on post-PKS tailoring enzymes either focus on one or specific configured anthracyclinone substrate or completely focus on *in vitro* data[60–63]. By contrast, my study presents a systematic evaluation of the substrate promiscuity of tailoring enzymes on all four possible aglycones both *in vivo* and *in vitro*. In addition, incorporating tailoring genes from unnatural pathways provides great opportunities for the chemodiversification of anthracyclines and novel drug discovery. During the study, I observed that certain enzymatic reactions follow a specific sequence. For example, DnrF or RdmE are inactive if RdmC or EamC act first (data not shown). However, only a limited number of reaction combinations were explored in this work. Further investigation of additional reaction orders and combinations will be important to better understand the pathway dynamics.

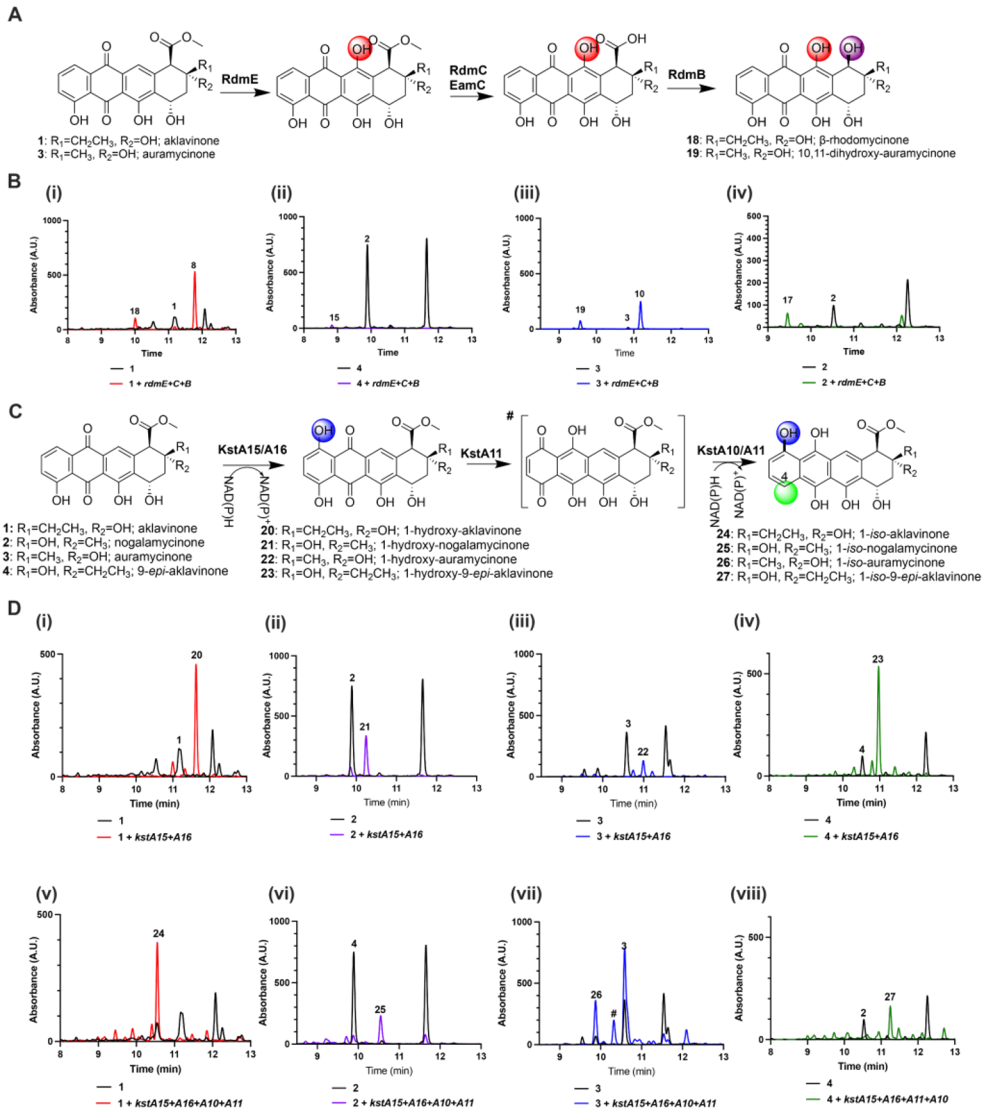


Figure 17. Metabolic engineering of 1-hydroxylated and 4-regioisomerized anthracyclinones. **A,** Post-PKS tailoring pathways for C10,11-hydroxylation of anthracyclinones. **B**, HPLC traces at 490 nm of strains engineered with expression constructs encoding **1**−**4** and *rdmECB* genes, or control strains producing only **1**−**4**; **C,** Post-PKS tailoring pathways for C1-hydroxylation and C4-regioisomerization of anthracyclinones. **D**, i-iv, HPLC traces at 490 nm of strains engineered with expression constructs encoding **1**−**4** and *kstA1516* genes, or control lines producing only **1**−**4**; v-viii, HPLC traces at 430 nm of strains engineered with expression constructs encoding **1**−**4** and *kstA15161011* genes, or control lines producing only **1**−**4**.

4.2.2 Complete biosynthesis of natural anthracyclines (Study III)

After completing the individual modules, I focused on the complete biosynthesis of three important natural anthracyclines, including nogalamycin, doxorubicin, and aclacinomycin.

4.2.2.1 Nogalamycin pathway refactoring

To accomplish the complete refactoring of the nogalamycin pathway, I utilized all the necessary genes from the native pathway or homologs from other anthracycline pathways (**Figure 18**). The nogalamycin pathway was constructed based on the nogalamycinone plasmid pRW10000 described earlier. After introducing two additional plasmids carrying the resistance and GTs, and TDP-carbohydrate modules, the production of glycosylated compound **7** was observed by UHPLC analysis, compared to the authentic standard, and further confirmed by MS analysis. Thanks to the strong expression of GTs and sufficient self-resistance, the complete metabolic flux to compound **6** was achieved.

Due to the lack of genes for C1-hydroxylation, the attachment of the second sugar TDP-L-rhodosamine at this position did not occur. Therefore, I introduced a fourth plasmid carrying genes responsible for C1-hydroxylation (*snoaL2W*)[229] and tailoring genes for 2”-hydroxylation (*snoT*), epoxyoxocin ring formation (*snoK*), 4”-epimerization (*snoN*), and *O*-methylations (*snogLM*), resulting in the clean production of nogalamycin (**28**) in E1 medium[230,231]. The production was identified by UHPLC analysis by comparing to authentic standard and LC-(-)-ESI-HR-MS analysis (calculated: 786.2973 $[M-H]^-$; found: 786.2990 $[M-H]^-$). The production yield was quantified using the standard curve and reached 33.6 mg/L after 7 days of cultivation of. Approximately 72% of the compound was present in the cultural supernatant, indicating efficient efflux by the non-native transporters.

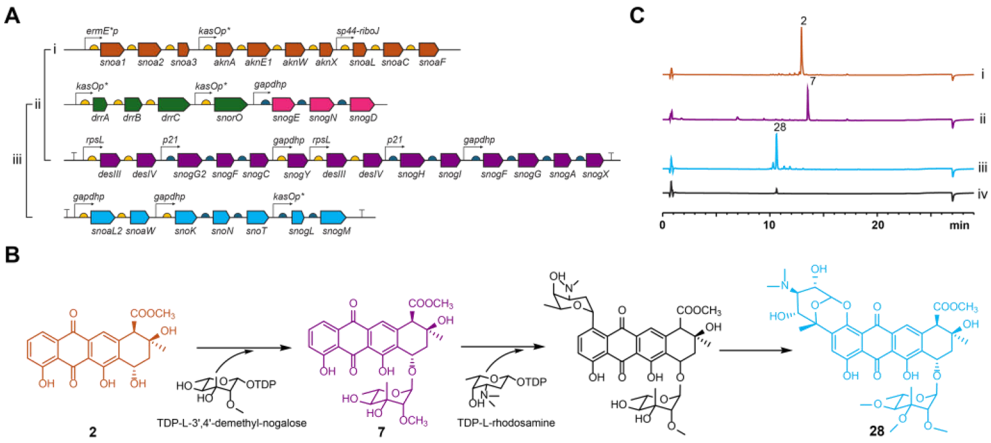


Figure 18. Complete refactoring of the nogalamycin pathway. The gene cassettes, HPLC traces and pathway schemes are presented, with matching colors for biosynthetic genes, HPLC traces, and compound structures. Left brackets with Roman numerals indicate the plasmids included in the strains of each UHPLC trace. The UHPLC traces of authentic standards and the structures of pathway intermediates are shown in black. **A**, SBOL diagrams illustrating the gene cassettes of each plasmid and the redesign strategy for the nogalamycin pathway refactoring. **B**, Simplified pathway scheme of nogalamycin biosynthesis. **C**, UHPLC traces showing the production profiles of representative strains. i, strain NOG at 430 nm; ii, strain RW-30 at 430 nm; iii, strain RW-48 at 470 nm; iv, authentic standard of **28** at 470 nm.

On the other hand, production in SG-TES was also tested due to its potential for higher yields. However, UHPLC analysis revealed only intermediate peaks, suggesting unexpected reactions or degradations induced by the medium (data not shown).

4.2.2.2 Doxorubicin pathway refactoring

A similar strategy was applied to the doxorubicin pathway. Specifically, M1152Δ*matAB*::pEN10001 (AKV1) was chosen as the initial chassis. The second and third expression cassettes for resistance and GTs DnrQS, and carbohydrate TDP-L-daunosamine were introduced, leading to the clean production of the glycosylated compound 11-deoxy-rhodomycin D (**29**). Production was identified by comparison to an authentic standard and MS analysis (data not shown; more details in original publication III).

However, when the remaining tailoring genes (*dnrFPKV*, *doxA*) were added, only a tiny amount of doxorubicin was produced by the recombinant strain. Stepwise development cycles were subsequently carried out for troubleshooting. Introduction of the 11-hydroxylation (*dnrF*)[232] and 15-demethylation (*dnrP*)[233] activities led to the full conversion of the substrate **29** to 15-demethoxy-rhodomycin D (**30**), but only 18.7% was further converted to 13-deoxydaunomycin (**31**) by *dnrK*[234], indicating

that 4-*O*-methylation by *dnrK* is a bottleneck step, even though another promoter *ermEp** was added (**Figure 19B**). The natural operon structure of P450 monooxygenase (*doxA)* and the auxiliary gene (*dnrV*) was preserved to ensure functional expression. Additionally, native ferredoxin reductase *fdr3* and ferredoxin *fdx4*, responsible for transferring electrons to DoxA, were introduced[35]. As a result, doxorubicin (**35**) was produced as the main product in the engineered strain, but an equivalent level of the early intermediate **30** was also observed. To address this issue, another copy of *dnrK* driven by the strong promoter *kasOp** was added, resulting in a clean production profile of **35** in strain RW-42 with a yield of 1.5 mg/L (**Figure 19B-C**). However, trace amounts of daunorubicin (**34**) were still detected in the strain without the ferredoxin pair.

I next switched to the aklavinone plasmid pAKV7 which was shown to result in higher yields. The results revealed that pAKV7 could supply more substrates, and the expression of sugar moieties was also sufficient, as all aglycones were converted to glycosides. Although the yield of **35** reached 2.85 mg/L in RW-44 strain, the intracellular DoxA level became limiting, preventing full conversion to **35** and resulting in the production of mostly 13-dihydrodaunorubicin (**33**) and **34** produced. On the other hand, a high producing strain of **34**, RW-43, with a yield of 14.6 mg/L in E1 medium, was achieved.

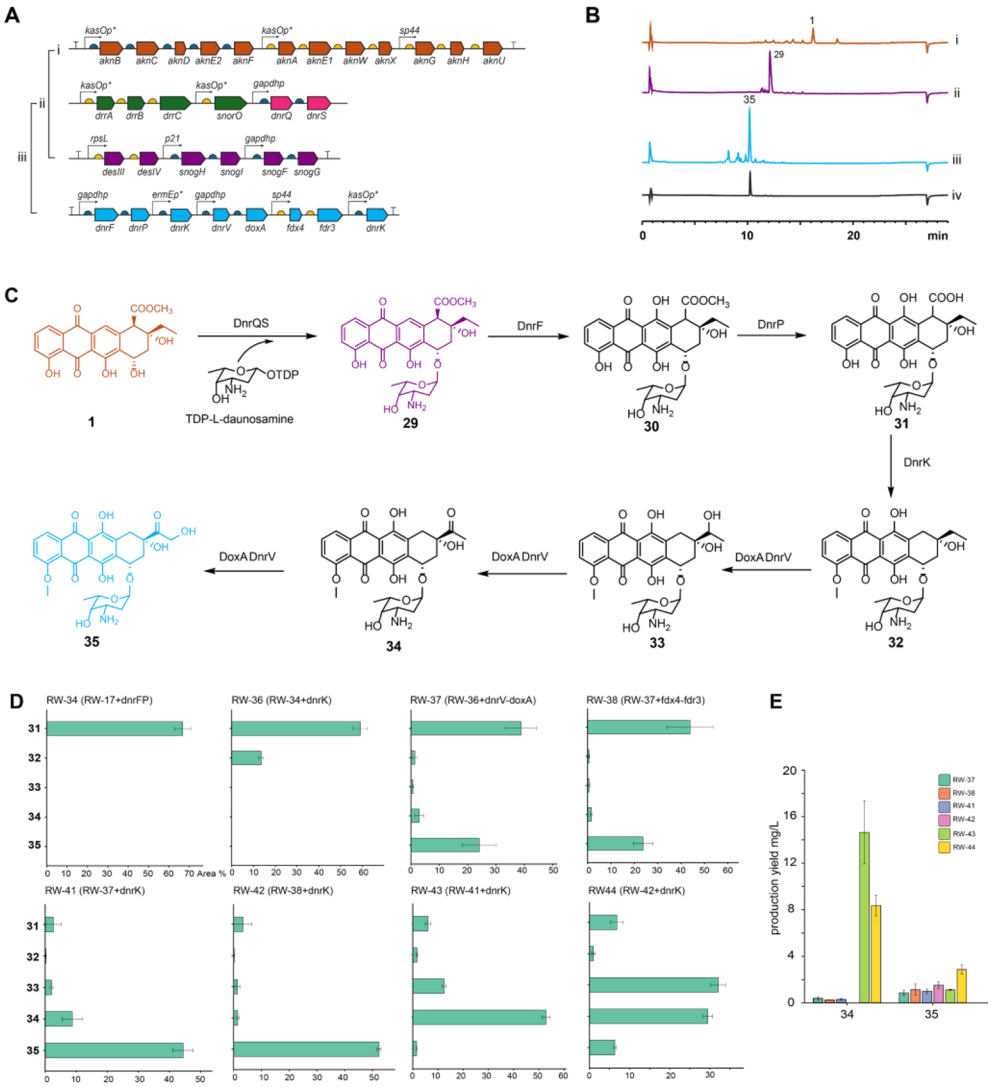


Figure 19. Complete refactoring of the doxorubicin pathway. A, Construction strategy for the redesigned doxorubicin pathway. **B**, UHPLC traces presenting the production profiles of representative strains. i, strain AKV1 at 430 nm; ii, strain RW-17 at 430 nm; iii, strain RW-38 at 490 nm; iv, authentic standard of **35** at 490 nm. **C**, Late biosynthetic pathway of doxorubicin. Panels **A**, **B**, and **C** are presented in the same manner as in **Figure 18**. **D**, Percentages of doxorubicin derivatives produced in the engineered strains. **E**, Compound **34** and **35** production titers of the engineered strains.

4.2.2.3 Aclacinomycin pathway refactoring

Aclacinomycin A is an anthracycline that demonstrates potential cardiotoxicity-free anticancer activity. Unlike nogalamycin or doxorubicin, aclacinomycins lack core scaffold modifications, but possess multiple sugar moieties, significantly influencing

their cytotoxicity and pharmacological properties. In this report, I aimed to refactor the aclacinomycin N pathway (*e.g.*, L-rhodosaminyl-2-deoxy-L-fucosyl-L-rhodinosyl aklavinone) because the final oxidation steps by AknOx create a mixture of aclacinomycin derivatives, complicating downstream purifications[38].

Since aclacinomycin N and doxorubicin biosynthesis proceed through the same core scaffold, the aklavinone-producing strain AKV1 was transformed with the resistance and *S. galilaeus* GT cassettes, and TDP-deoxysugar modules. In the transferase module, the GT AknS and P450 enzyme AknT are responsible for attaching the first sugar unit, TDP-L-rhodosamine[227]. The bifunctional GT AknK appends TDP-2-deoxy-L-fucose and TDP-L-rhodinose to the growing sugar chain to afford aclacinomycin N (**36**)[39]. The gene cassette for TDP-L-rhodosamine was sourced from the nogalamycin refactoring, and additional genes (*aknLMPQ*) from the aclacinomycin pathway were included for the two neutral sugar units (**Figure 20A**). The production of **36** was detected in SG-TES medium, identified by comparison to an authentic standard and by HR-MS analysis (calculated: 814.3645 $[M+H]^+$; found: 814.3639 $[M+H]^+$). However, the extract also contained many intermediates (data not shown; more details in original publication III).

To improve production, I incorporated the enhanced aklavinone cassette, pAKV7, and tested it in E1 medium, which typically supports better production profiles. This resulted in a cleaner and increased production of **36**, with a yield of 2.4 mg/L (**Figure 20B**). However, the accumulation of aglycone **1** was still detected, highlighting the challenges in balancing the expression of aglycone, tandemly-acting GTs, and multiple TDP-sugar donors.

To the best of our knowledge, this study represents the first successful refactoring of complete anthracycline biosynthetic pathways. Previous efforts in the heterologous expression of anthracyclines have relied solely on the direct capture and expression of native BGCs, which remain subject to native regulatory constraints[37,225]. Moreover, such approaches do not readily support combinatorial biosynthesis for generating novel derivatives. In contrast, our work involved the cloning, introduction, and fine-tuning of up to 38 genes from multiple pathways into a heterologous *Streptomyces* host (more than 63,000 bp of designed DNA sequence was integrated to the chromosome), enabling the modular reconstruction of functional anthracycline pathways. This achievement marks a significant milestone in synthetic biology within the *Streptomyces* genus.

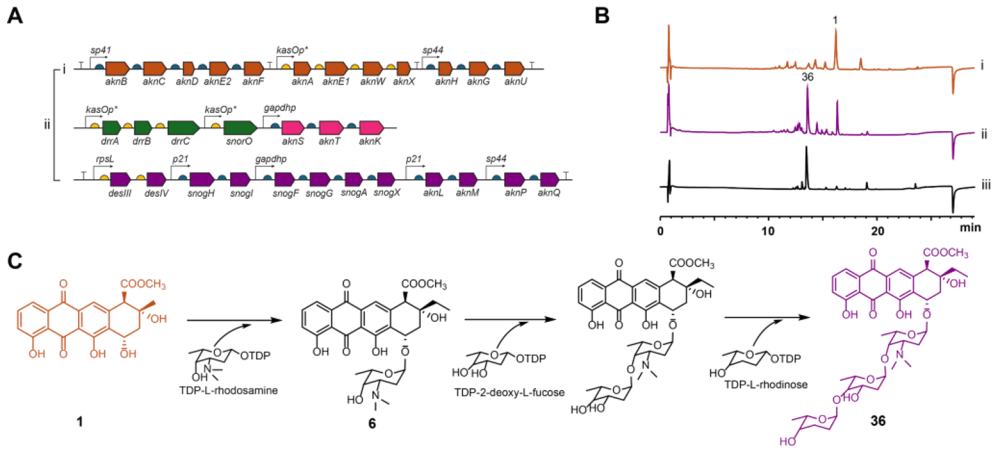


Figure 20. Complete refactoring of the aclacinomycin pathway. A, SBOL diagrams illustrate the construction strategy for the redesigned aclacinocmyin pathway. **B**, UHPLC traces at 430 nm showing the production profiles of representative strains. i, strain AKV7; ii, strain RW-21; iii, authentic standard of **36**. **C**, Simplified biosynthetic pathway of the aclacinomycin pathway. The colors and Roman numerals are presented similarly to those in **Figure 18**.

4.2.3 Combinatorial biosynthesis for novel anthracyclines (Study III)

Anthracyclines share similar core structures yet display remarkable chemodiversity due to variations in sugar units and multifaceted tailoring modifications. Although genetic permutations in nature have led to the discovery of more than 500 anthracyclines, many are not clinically suitable due to drug resistance, side effects, and unfavorable pharmacokinetic profiles[8]. One attractive methodology for generating "unnatural" NPs is combinatorial biosynthesis. The current synthetic biology platform provides the requisite tools to accomplish combinatorial biosynthesis of anthracyclines: an amenable genetic host, modular pathway construction, and BioBricks libraries of different gene sets, facilitating the rapid generation of recombinant hosts.

Therefore, I leveraged the platform to generate new anthracycline analogs by modifications of the core scaffold and attachment of various carbohydrate moieties. Prior to this story, the specificity of GTs towards altered substrates was mainly unknown. However, we hypothesized that since the four aglycone acceptors have similar structures, the native GTs would exhibit relaxed specificity towards the aglycone and a stricter preference towards different sugar nucleotide donors. Therefore, I incorporated the assembled GT pairs (AknST, DnrQS, SnogEN) with their natural sugar nucleotide donor cassettes to improve the likelihood of obtaining novel anthracyclines by exploring non-cognate aglycones (**Figure 21**). I first utilized the doxorubicin GT DnrQS to combine the natural TDP-L-daunosamine with

atypical aglycones, **3**, **2**, and **4**, which resulted in production of **37**, **38**, and **39**, respectively. The DnrQS pair has been demonstrated to attach TDP-L-acosamine, the 4'-epimer of TDP-L-daunosamine, for the biosynthesis of epirubicin[216]. Therefore, I transferred TDP-L-acosamine to **2** and **4** to form **40** and **41**, respectively. As for the GT AknST from the aclacinomycin pathway, I took advantage of it for decorating the natural TDP-L-rhodosamine to all three unnatural aglycones **3**, **2**, and **4** leading to **42**, **43**, and **44**, respectively. Finally, I switched to GT SnogEN and its natural sugar unit, TDP-L-3',4'-demethyl-nogalose, enabling the glycosylation of non-native C-21 carbon aglycones **1** and **4** to generate **45** and **46**, respectively.

Next, the chemical diversity was further enriched by incorporating additional tailoring functionalities. The two-component monooxygenase system KstA1516 from the kosinostatin pathway was recruited for C1-hydroxylation on three nogalamycinone-based compounds, resulting in new anthracyclines, **47**, **48**, and **49**[57]. The strain producing compound **6** was used to further demonstrate regiospecific hydroxylation. By using C11-hydroxylation by DnrF, the production of the known anthracycline rhodomycin T (**50**) was observed. The genes from the rhodomycin pathway were employed for C11-hydroxylation (*rdmE*), C10-demethylation (*rdmC*), and C10-hydroxylation (*rdmB*), resulting in another known anthracycline, rhodomycin B (**51**)[62].

DM-DXR is a promising cardiotoxicity-free anthracycline due to its unique mechanism of action for evicting histones from open chromatin regions. In contrast to **1**, DM-DXR has an L-rhodosamine substitution at the 7-position. However, the daunorubicin GT DnrS does not accept the dimethylamino sugar nucleotide[35]. Therefore, AknS from the aclacinomycin pathway is a more suitable GT for TDP-L-rhodosamine. Starting with the strain producing **6**, the first TDP-L-rhodosamine unit and GT pair *aknST* were incorporated, followed by the introduction of *dnrFKPV*, *doxA*, *fdx4*, and *fdr3*. However, only the penultimate intermediate N,N-dimethyl-13-deoxydaunorubicin (**52**) was produced, indicating that DoxA does not accept the modified substrate[35]. Additional protein engineering work is required to increase the substrate flexibility of DoxA.

The efficient and balanced expression of the required modules afforded quantitative and complete conversion to the intended anthracyclines in most cases, as determined by UHPLC analysis, facilitating downstream purification. Specifically, the fermentation of the recombinant strains was scaled up to 1-5 L of SG-TES or E1 culture. The purification process involved the silica column chromatography or size-exclusive chromatography, followed by further purification of preparative HPLC. The structures of the novel compounds were rigorously determined via NMR spectroscopy and HR-MS analysis, while the three known molecules were identified by comparison to authentic standards (More details can be found in original publication III). Using the comprehensive combinatorial

biosynthesis platform, in total 16 anthracycline analogs were obtained by combining four genetic cassettes, where 13 anthracycline analogs are entirely novel.

Prior to this study, numerous efforts have been made to achieve the combinatorial biosynthesis of anthracyclines[211–215]. However, most previous approaches focused on enzyme-level modifications, in which a limited number of heterologous genes were introduced into native producers to modify or replace specific biosynthetic steps. These strategies only allowed for minimal structural diversification of natural anthracyclines. To address these limitations, some researchers attempted to reconstruct entire anthracycline biosynthetic pathways to enable more extensive combinatorial biosynthesis[101,219]. However, the cloning and co-expression of complete anthracycline gene clusters in a single heterologous host remained challenging due to the lack of a comprehensive synthetic biology platform. As a result, alternative strategies such as bioconversion and co-cultivation were employed[101,219]. Despite their utility, these methods were not readily scalable due to limitations in substrate availability and the complexity of fermentation processes.

In contrast, our study successfully overcomes these challenges by enabling the modular combination of all key components of anthracycline biosynthesis—including aglycones, TDP-sugars, GTs, and tailoring enzymes—within a unified synthetic biology platform. This strategy resulted in the production of a wide variety of anthracycline analogs with enhanced diversity. Furthermore, whereas earlier reconstituted pathways typically relied on a single promoter driving long gene cassettes—often resulting in inefficient expression and accumulation of intermediates—our system utilizes well-characterized promoters with varying strengths to independently fine-tune the expression of each biosynthetic module[101,219]. Consequently, the target compounds were produced with high efficiency and purity, as demonstrated by the dominant peaks observed in HPLC analysis. Therefore, this synthetic biology platform will act a powerful and versatile tool for the rational design of improved anthracycline drugs via combinatorial biosynthesis in the future.

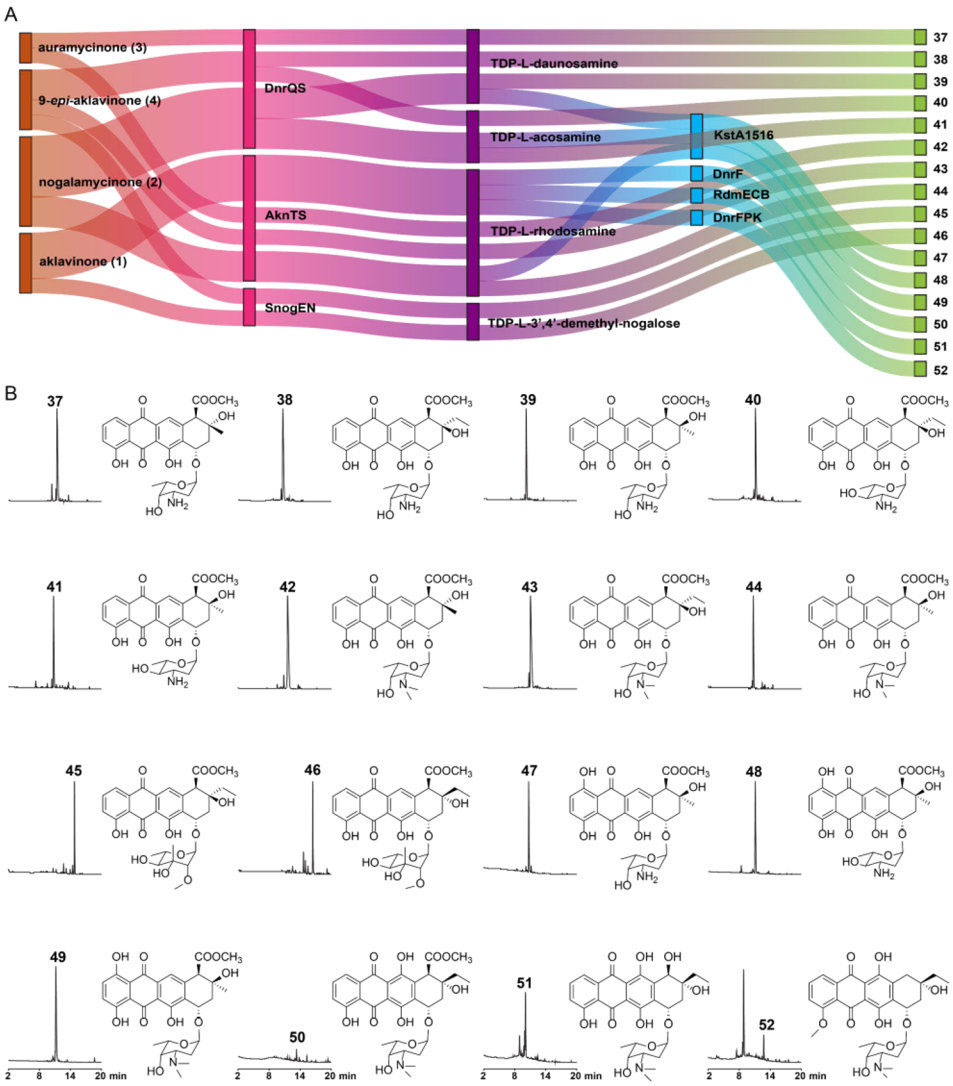


Figure 21. Combinatorial biosynthesis for designer anthracyclines. A, Streamlines illustrating the modular combinations to produce designer anthracyclines. **B**, UHPLC traces showing the production profiles of each strain, along with the structures of the desired compounds. The production profiles of compounds **37**-**46** are presented by visible absorbance at 430 nm, while the production of compounds **47**-**52** is shown by absorbance at 490 nm.

4.3 Human cancer cell line viability (Study III)

To better understand the structure-activity relationship (SAR) of the novel anthracyclines, their cytotoxic potency was evaluated at the University of Kentucky. All newly synthesized compounds (**37**–**49**), along with control compounds, were assessed for cytotoxic activity against five human cancer cell lines at a concentration

of 80 µM. These cell lines included A549 (non-small cell lung cancer), PC3 (prostate cancer), and HCT116 (colorectal cancer), as well as Merkel cell carcinoma lines MKL1 and MCC26 (**Table 7**; further details in original publication III). Consistent with previous reports, compounds **45** and **46**, which are decorated with neutral sugars, along with aglycones **1** and **2** that lack sugar moieties, exhibited no cytotoxic effects on any of the five cancer cell lines[235]. In contrast, other compounds bearing amino sugars demonstrated significant cytotoxic activity, underscoring the crucial role of amino sugars in the cytotoxicity of anthracyclines.

As a result, compounds **45** and **46** were excluded from further bioactivity testing. Subsequently, a second assay was conducted on the remaining compounds at a lower concentration (10 µM) to better differentiate their cytotoxic effects, as most compounds exhibited similar toxicity at 80 µM. As shown in **Figure 22**, the novel anthracyclines displayed varying degrees of cytotoxicity across different cancer cell lines at 10 µM. Notably, the sugar unit, aglycone core, and post-PKS tailoring modifications all influenced cytotoxicity. Compounds **41**–**43**, which contain rhodosamine, exhibited higher toxicity compared to compounds **37**–**41** that possess daunosamine or acosamine, emphasizing the impact of 3'-*N*-methylation of sugars on bioactivity. Interestingly, auramycinone-based compounds demonstrated the lowest cytotoxic activity among all tested compounds, whereas nogalamycinone- and 9-*epi*-aklavinone-based derivatives exhibited similar toxicity levels (*e.g.*, **37** vs. **42**; **38** vs. **40** vs. **43**; **39** vs. **41** vs. **44**). These findings suggest that carbon chain length and stereochemistry at the C9 position may influence the DNA-binding ability of anthracyclines, thereby affecting their cytotoxic properties. Additionally, C1-hydroxylation significantly enhanced cytotoxicity compared to the non-hydroxylated analogs.

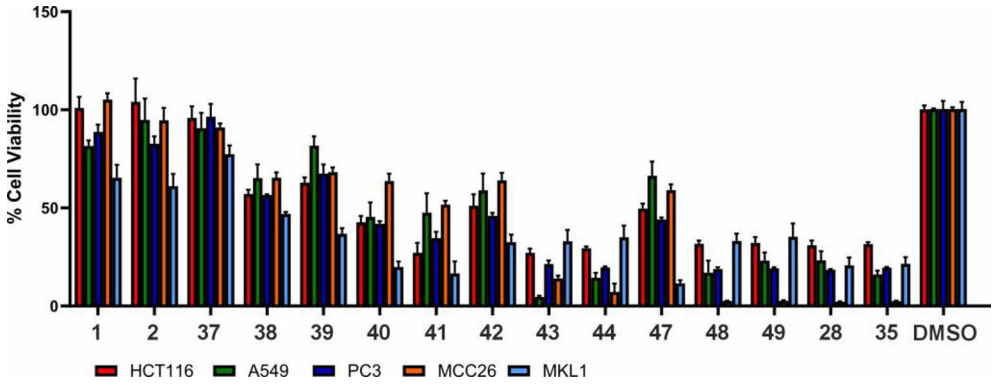


Figure 22. Viability of five human cancer lines against novel anthracycline compounds at 10 µM. Cell viability data were normalized relative to the DMSO-treated control group.

Finally, a dose-response experiment was conducted on compounds **37**–**44** and **47**–**49**, starting at 10 μM, to determine their EC_{50} values (**Table 7**). This experiment enabled a precise comparison of the cytotoxic potency of the novel anthracyclines. The results revealed that six compounds (**40**, **41**, **43**, **44**, **48**, and **49**) exhibited EC_{50} values below 10 μM against cancer cell lines. Among them, compounds **43**, **44**, and **49** were cytotoxic to PC3, HCT116, MKL1, and MCC26 cells. Interestingly, A549 cells exhibited tolerance to all novel compounds except compound **49**, which displayed mild toxicity. These findings further highlight the importance of rhodosamine and C1-hydroxylation in enhancing cytotoxicity. Although positive control compounds **28** and **35** exhibited significantly stronger cytotoxicity, with IC_{50} values below 1 μM, the results provide valuable insights into the SAR of anthracyclines and will contribute to the rational engineering of anthracycline derivatives for future drug development.

Table 7. Cytotoxic activities of novel anthracycline compounds.

Compounds	EC_{50} μM			EC_{50} μM	
	A549	PC3	HCT116	MKL1	MCC26
Doxorubicin (**28**)	0.77±0.05	0.33±0.01	0.17±0.03	0.044±0.000	0.22±0.02
Nogalamycin (**35**)	0.103±0.001	0.072±0.003	0.023±0.004	0.078±0.002	0.061±0.005
Aklavinone (**1**)	>10	>10	>10	>10	>10
Nogalamycinone (**2**)	>10	>10	>10	>10	>10
7-*O*-*α*-L-Daunosaminyl-auramycinone (**37**)	>10	>10	>10	>10	>10
7-*O*-*α*-L-Daunosaminyl-9-*epi*-aklavinone (**38**)	>10	>10	>10	>10	>10
7-*O*-*α*-L-Daunosaminyl-nogalamycinone (**39**)	>10	>10	>10	>10	>10
7-*O*-*α*-L-Acosaminyl-9-*epi*-aklavinone (**40**)	>10	>10	5.80±1.75	9.38±0.63	>10
7-*O*-*α*-L-Acosaminyl-nogalamycinone (**41**)	>10	>10	>10	4.50±1.91	7.17±0.86
7-*O*-*α*-L-Rhodosaminyl-auramycinone (**42**)	>10	>10	>10	>10	>10
7-*O*-*α*-L-Rhodosaminyl-9-*epi*-aklavinone (**43**)	>10	4.86±0.07	1.48±0.50	2.38±0.14	8.08±1.22
7-*O*-*α*-L-Rhodosaminyl-nogalamycinone (**44**)	>10	7.40±1.94	1.56±0.12	1.75±0.08	8.40±0.21
7-*O*-*α*-L-3′,4′-Demethyl-nogalosyl-9-*epi*-akalvinone (**45**)	>80	>80	>80	>80	>80
7-*O*-*α*-L-3′,4′-Demethyl-nogalosyl-akalvinone (**46**)	>80	>80	>80	>80	>80
1-Hydroxy-7-*O*-*α*-L-daunosaminyl-nogalamycinone (**47**)	>10	>10	>10	>10	>10
1-Hydroxy-7-*O*-*α*-L-acosaminyl-nogalamycinone (**48**)	>10	5.67±0.29	2.59±0.41	1.83±0.04	>10
1-Hydroxy-7-*O*-*α*-L-rhodosaminyl-nogalamycinone (**49**)	9.35±1.86	4.49±0.82	1.16±0.03	1.59±0.21	8.31±1.27

5 Conclusions and Future Perspectives

In this doctoral thesis, I established a comprehensive synthetic biology platform that includes an amenable *Streptomyces* host, well-characterized promoters of tunable strengths, efficient terminators, user-friendly vectors, and an extensive collection of biosynthetic genes for anthracyclines. This platform enables the streamlined assembly of anthracycline biosynthetic pathways with minimal effort, facilitating the complete biosynthesis of anthracyclines. Moreover, it provides a powerful framework for combinatorial biosynthesis, unlocking new opportunities for the discovery of structurally diverse and potential anticancer agents.

A critical prerequisite for this advancement was ensuring the functionality of all essential components within the platform. To achieve this, I constructed an engineered *S. coelicolor* M1152Δ*matAB* as a chassis optimized for the heterologous expression. The platform includes more than ten well-characterized constitutive promoters and over ten efficient intrinsic terminators to regulate gene expression. Five integrative *Streptomyces* plasmids, designed to be compatible within a single strain, were constructed to accommodate all biosynthetic genes and enable modular pathway assembly. A curated library of over 100 biosynthetic genes from various anthracycline pathways was organized into distinct functional modules, including polyketide aglycones, GTs, TDP-carbohydrates, self-resistance, and post-PKS tailoring. Their enzymatic functions and substrate promiscuity were systematically evaluated through gene complementation, exhaustive precursor feeding, and *in vivo* metabolite analysis. The efficient construction and assembly of all genetic elements were facilitated by a standardized BioBricks design.

By leveraging the synthetic biology platform, I accomplished the complete biosynthesis of three clinically relevant anthracyclines–nogalamycin, doxorubicin, and aclacinomycin–for the first time. This groundbreaking achievement required the coordinated heterologous expression of 38, 31, and 31 genes, respectively, underscoring the complexity of these biosynthetic pathways and the innovative strategies we employed to successfully reconstitute them in a heterologous host. Specifically, I first built and optimized the expressing plasmids for the core structures, anthracyclinones, resulting in yields of 1.6-40 mg/L in E1 production

medium. The subsequent optimization and integration of GTs, self-defense, TDP-carbohydrates, and post-PKS tailoring modules led to the clean production of target anthracyclines, with yields ranging from 1.5 to 33.6 mg/L. Notably, this work also elucidated the final unsolved steps in the sugar biosynthesis of nogalamycin and aclacinomycin, further advancing our understanding of the biosynthetic pathways.

Beyond the biosynthesis of natural anthracyclines, the modular nature of the platform facilitated the generation of 16 anthracycline analogs, 13 of which represent entirely novel compounds. The high-yield and clean production of these analogs enabled efficient purification and structural characterization via NMR. Structural variations among these analogs arose from differences in the carbon chain length and stereochemistry of aglycones, TDP-carbohydrate structures, and post-PKS tailoring modifications. The following cytotoxicity test led to the discovery of six high potency anthracyclines and provided valuable insights into SAR studies, supporting the rational design of improved anticancer agents in the future.

The development of synthetic biology in *Streptomyces* has long been hindered by the absence of a comprehensive genetic toolbox. The platform established in this thesis addresses this limitation, providing a robust foundation for the complete biosynthesis of complex natural products in *Streptomyces*. Furthermore, this work demonstrates a scalable and systematic approach for the rapid exploration of functional gene combinations, facilitating the discovery of novel bioactive compounds.

However, there are still several limitations in this study that require further investigation. For instance, although the clean production of three natural anthracyclines was achieved, the yields of aklavinone-derived aclacinomycin N and doxorubicin remain far below industrial requirements. This low production is primarily due to the limited yield of aklavinone, which in turn results from the insufficient expression of the minPKS AknBCDE2F. To address this bottleneck, more efforts should focus on enhancing minPKS expression through strategies such as decoupling translational coupling, optimizing codon usage, and alleviating endogenous regulatory constraints by using fully artificial genetic elements.

Second, verifying the functional expression of individual genes in *Streptomyces* remains challenging due to the long developmental cycle of this organism. In many cases, gene functionality is inferred from the accumulation of desired metabolites. However, such metabolites may be undetectable or only appear after the inclusion of several genes. This issue is compounded by the nature of the BioBricks assembly method, since individual genes cannot be removed from the constructs. The troubleshooting is laborious even if a single gene in a multi-gene cassette is non-functional. To overcome this bottleneck, efficient gene editing systems such as CRISPR/Cas could be employed to replace specific gene elements within BioBricks constructs, enabling more flexible and rapid pathway optimization.

There are several promising avenues for further development based on this platform. The biological evaluations of the newly generated anthracycline analogs that were conducted in this study were limited to cytotoxicity assays in cancer cell lines. A more comprehensive assessment, including studies on DNA damage and chromatin disruption, is warranted—particularly given their relevance to cardiotoxicity, a key concern for anthracycline therapeutics. With improved SAR understanding, further engineering of *Streptomyces* strains could combine beneficial modifications to produce analogs with enhanced therapeutic potential. In addition, while this study successfully achieved the biosynthesis of aclacinomycin N, the clinically used compound is aclacinomycin A[33], which requires an additional oxidative modification catalyzed by AknOx. Although AknOx can introduce multiple oxidative products, the previous study has demonstrated that specific mutations in AknOx can restrict its activity to yield primarily aclacinomycin A[38]. Therefore, introducing a rationally engineered AknOx variant into the platform could enable targeted production of aclacinomycin A. Finally, although this platform was originally developed for anthracyclines, its modular architecture is broadly applicable and could be adapted for the biosynthesis of other valuable natural product classes, particularly glycosylated aromatic polyketides.

Acknowledgements

My doctoral thesis work was conducted in the Antibiotic Biosynthesis Engineering (ABE) laboratory at the Department of Life Technologies, Faculty of Technology, University of Turku, from 2020 to 2025. I would like to thank all the personnel from our faculty for creating such a lovely working and research environment. I also extend my thanks to all the teachers from my courses for their excellent instruction. The contributions made by the administrative staff to the Doctoral Programme of Technology are greatly appreciated. The personal funding from Turku University Foundation, China Scholarship Council, and the Doctoral Programme of Technology are sincerely acknowledged. Additional funding support from Research Council of Finland is much appreciated. I am thankful for the travel grant that supported my attendance at the scientific conferences provided by the Doctoral Programme of Technology, Turku University Foundation, Microbes and Immunity Research Program, and Computational and Molecular Methodologies for Life Sciences research program.

I sincerely thank Professor Yvonne Nygård and Docent Suvi Santala for the pre-examination of my doctoral thesis and I am honored to have Professor Eriko Takano as the opponent for my public defense.

I would like to express my deepest appreciation to my principal supervisor, Professor Mikko Metsä-Ketelä. I am sincerely grateful for his extraordinary supervision, valuable advice, and patience. I also thank the support, encouragement, and freedom he gave me in my research. My gratitude also extends to my co-supervisor, Docent Jarmo Niemi, for his guidance and valuable suggestions on my doctoral project. He has always been an outstanding lecturer and scientist. I am grateful to our external collaborators, especially to Associate Professor S. Eric Nybo from Ferris State University and Associate Professor Khaled A. Shaaban from the University of Kentucky.

Special thanks to all ABE members, both past and present, for their peer support, knowledge sharing, and companionship during coffee break and lunch. They are not only my colleagues but also my friends who have supported me and brought happiness to my life. In particular, I would like to thank Heli Tirkkonen, Arina Koroleva, Magdalena Niemczura, Keith Yamada, Benjamin Nji Wandi, Aleksi Nuutila for their friendship and efforts in our shared projects.

Last but not least, I would like to express my profound gratitude to my family for their support and love in my life. I am especially grateful to my parents and sister for their unconditional love and constant encouragement. I also thank my beloved wife, Yuqing, for her love, encouragement, and the happiness she brings to my life.

Turku, April 2025
Rongbin

List of References

1. Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83**, 770–803 (2020).
2. Alam, K. *et al. Streptomyces*: The biofactory of secondary metabolites. *Front. Microbiol.* **13**, 968053 (2022).
3. Flärdh, K. & Buttner, M. J. *Streptomyces* morphogenetics: dissecting differentiation in a filamentous bacterium. *Nat. Rev. Microbiol.* **7**, 36–49 (2009).
4. Lee, N. *et al.* Systems and synthetic biology to elucidate secondary metabolite biosynthetic gene clusters encoded in *Streptomyces* genomes. *Nat. Prod. Rep.* **38**, 1330–1361 (2021).
5. Lee, Y. *et al.* CRISPR-aided genome engineering for secondary metabolite biosynthesis in *Streptomyces*. *J. Ind. Microbiol. Biotechnol.* **51**, kuae009 (2024).
6. Komatsu, M., Uchiyama, T., Ōmura, S., Cane, D. E. & Ikeda, H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc. Natl. Acad. Sci.* **107**, 2646–2651 (2010).
7. Demain, A. L. & Fang, A. The Natural Functions of Secondary Metabolites. in *History of Modern Biotechnology I* (ed. Fiechter, A.) vol. 69 1–39 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2000).
8. Hulst, M. B., Grocholski, T., Neefjes, J. J. C., Van Wezel, G. P. & Metsä-Ketelä, M. Anthracyclines: Biosynthesis, engineering and clinical applications. *Nat. Prod. Rep.* **39**, 814–841 (2022).
9. Brockmann, H. & Bauer, K. Rhodomycin, ein rotes antibioticum aus actinomyceten. *Naturwissenschaften* **37**, 492–493 (1950).
10. Elshahawi, S. I., Shaaban, K. A., Kharel, M. K. & Thorson, J. S. A comprehensive review of glycosylated bacterial natural products. *Chem. Soc. Rev.* **44**, 7591–7697 (2015).
11. Arcamone, F. *et al.* Adriamycin, 14-hydroxydaunomycin, a new antitumor antibiotic from S. peucetius var. caesius. *Biotechnol. Bioeng.* **11**, 1101–1110 (1969).
12. Murzyn, A. *et al.* Aclarubicin: Contemporary insights into its mechanism of action, toxicity, pharmacokinetics, and clinical standing. *Cancer Chemother. Pharmacol.* **94**, 123–139 (2024).
13. Zhang, S. *et al.* Identification of the molecular basis of doxorubicin-induced cardiotoxicity. *Nat. Med.* **18**, 1639–1642 (2012).
14. Martins-Teixeira, M. B. & Carvalho, I. Antitumour anthracyclines: Progress and perspectives. *ChemMedChem* **15**, 933–948 (2020).
15. Qiao, X. *et al.* Uncoupling DNA damage from chromatin damage to detoxify doxorubicin. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 15182–15192 (2020).
16. Pang, B. *et al.* Drug-induced histone eviction from open chromatin contributes to the chemotherapeutic effects of doxorubicin. *Nat. Commun.* **4**, 1908 (2013).
17. Van Der Zanden, S. Y., Qiao, X. & Neefjes, J. New insights into the activities and toxicities of the old anticancer drug doxorubicin. *FEBS J.* **288**, 6095–6111 (2021).
18. Metsä-Ketelä, M., Niemi, J., Mäntsälä, P. & Schneider, G. Anthracycline Biosynthesis: Genes, Enzymes and Mechanisms. in *Anthracycline Chemistry and Biology I* (ed. Krohn, K.) vol. 282 101–140 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2007).
19. Xu, Z. *et al.* Metabolic engineering of *Streptomyces* to enhance the synthesis of valuable natural products. *Eng. Microbiol.* **2**, 100022 (2022).

20. Niemi, J., Metsä-Ketelä, M., Schneider, G. & Mäntsälä, P. Biosynthetic Anthracycline Variants. in *Anthracycline Chemistry and Biology I* (ed. Krohn, K.) vol. 282 75–99 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2008).
21. Fu, H., Ebert-Khosla, S., Hopwood, D. A. & Khosla, C. Engineered biosynthesis of novel polyketides: Dissection of the catalytic specificity of the act ketoreductase. *J. Am. Chem. Soc.* **116**, 4166–4170 (1994).
22. Zhou, H., Li, Y. & Tang, Y. Cyclization of aromatic polyketides from bacteria and fungi. *Nat. Prod. Rep.* **27**, 839 (2010).
23. Torkkell, S. *et al.* Identification of a cyclase gene dictating the C-9 stereochemistry of anthracyclines from *Streptomyces nogalater*. *Antimicrob. Agents Chemother.* **44**, 396–399 (2000).
24. Kallio, P., Sultana, A., Niemi, J., Mäntsälä, P. & Schneider, G. Crystal structure of the polyketide cyclase AknH with bound substrate and product analogue: Implications for catalytic mechanism and product stereoselectivity. *J. Mol. Biol.* **357**, 210–220 (2006).
25. Dickens, M. L., Ye, J. & Strohl, W. R. Analysis of clustered genes encoding both early and late steps in daunomycin biosynthesis by *Streptomyces* sp. strain C5. *J. Bacteriol.* **177**, 536–543 (1995).
26. Dickens, M. L., Ye, J. & Strohl, W. R. Cloning, sequencing, and analysis of aklaviketone reductase from *Streptomyces* sp. strain C5. *J. Bacteriol.* **178**, 3384–3388 (1996).
27. Wang, G., Chen, J., Zhu, H. & Rohr, J. One-pot enzymatic total synthesis of presteffimycinone, an early intermediate of the anthracycline antibiotic steffimycin biosynthesis. *Org. Lett.* **19**, 540–543 (2017).
28. Gullón, S. *et al.* Isolation, characterization, and heterologous expression of the biosynthesis gene cluster for the antitumor anthracycline steffimycin. *Appl. Environ. Microbiol.* **72**, 4172–4183 (2006).
29. Zunino, F., Pratesi, G. & Perego, P. Role of the sugar moiety in the pharmacological activity of anthracyclines: Development of a novel series of disaccharide analogs. *Biochem. Pharmacol.* **61**, 933–938 (2001).
30. Thompson, M. W., Strohl, W. R. & Floss, H. G. Purification and characterization of TDP-D-glucose 4,6-dehydratase from anthracycline-producing Streptomycetes. *J. Gen. Microbiol.* **138**, 779–786 (1992).
31. Thibodeaux, C. J., Melançon, C. E. & Liu, H. Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew. Chem. Int. Ed.* **47**, 9814–9859 (2008).
32. Brown, K. V., Wandi, B. N., Metsä-Ketelä, M. & Nybo, S. E. Pathway engineering of anthracyclines: Blazing trails in natural product glycodiversification. *J. Org. Chem.* **85**, 12012–12023 (2020).
33. Wander, D. P. A. *et al.* Doxorubicin and aclarubicin: Shuffling anthracycline glycans for improved anticancer agents. *J. Med. Chem.* **63**, 12814–12829 (2020).
34. Mohideen, F. I., Nguyen, L. H., Richard, J. D., Ouadhi, S. & Kwan, D. H. *In vitro* reconstitution of the dTDP-L-daunosamine biosynthetic pathway provides insights into anthracycline glycosylation. *ACS Chem. Biol.* **17**, 3331–3340 (2022).
35. Hulst, M. B. *et al.* Metabolic engineering of *Streptomyces peucetius* for biosynthesis of N,N-dimethylated anthracyclines. *Front. Bioeng. Biotechnol.* **12**, 1363803 (2024).
36. Siitonen, V., Nji Wandi, B., Törmänen, A.-P. & Metsä-Ketelä, M. Enzymatic synthesis of the C-glycosidic moiety of nogalamycin R. *ACS Chem. Biol.* **13**, 2433–2437 (2018).
37. Siitonen, V. *et al.* Identification of late-stage glycosylation steps in the biosynthetic pathway of the anthracycline nogalamycin. *ChemBioChem* **13**, 120–128 (2012).
38. Alexeev, I., Sultana, A., Mäntsälä, P., Niemi, J. & Schneider, G. Aclacinomycin oxidoreductase (AknOx) from the biosynthetic pathway of the antibiotic aclacinomycin is an unusual flavoenzyme with a dual active site. *Proc. Natl. Acad. Sci.* **104**, 6170–6175 (2007).
39. Lu, W., Leimkuhler, C., Oberthür, M., Kahne, D. & Walsh, C. T. AknK is an L-2-deoxyfucosyltransferase in the biosynthesis of the anthracycline aclacinomycin a. *Biochemistry* **43**, 4548–4558 (2004).

40. Rini, J. M., Moremen, K. W., Davis, B. G. & Esko, J. D. Glycosyltransferases and Glycan-Processing Enzymes. in *Essentials of Glycobiology* (eds. Varki, A. et al.) (Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), 2022).
41. Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: Structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77**, 521–555 (2008).
42. Borisova, S. A. & Liu, H. Characterization of glycosyltransferase DesVII and its auxiliary partner protein DesVIII in the methymycin/pikromycin biosynthetic pathway. *Biochemistry* **49**, 8071–8084 (2010).
43. Kaur, P. & Russell, J. Biochemical coupling between the DrrA and DrrB proteins of the doxorubicin efflux pump of *Streptomyces peucetius*. *J. Biol. Chem.* **273**, 17933–17939 (1998).
44. Moncrieffe, M. C. *et al.* Structure of the glycosyltransferase EryCIII in complex with its activating P450 homologue EryCII. *J. Mol. Biol.* **415**, 92–101 (2012).
45. Claesson, M., Siitonen, V., Dobritzsch, D., Metsä -Ketelä, M. & Schneider, G. Crystal structure of the glycosyltransferase SnogD from the biosynthetic pathway of nogalamycin in *Streptomyces* nogalater. *FEBS J.* **279**, (2012).
46. Nji Wandi, B. *et al.* Evolution-guided engineering of non-heme iron enzymes involved in nogalamycin biosynthesis. *FEBS J.* **287**, 2998–3011 (2020).
47. Miyamoto, Y., Johdo, O., Nagamatsu, Y. & Yoshimoto, A. Cloning and characterization of a glycosyltransferase gene involved in the biosynthesis of anthracycline antibiotic Î²-rhodomycin from *Streptomyces violaceus*. *FEMS Microbiol. Lett.* **206**, 163–168 (2002).
48. Garrido, L. M. *et al.* Insights in the glycosylation steps during biosynthesis of the antitumor anthracycline cosmomycin: Characterization of two glycosyltransferase genes. *Appl. Microbiol. Biotechnol.* **73**, 122–131 (2006).
49. Itoh, T., Kinoshita, M., Wei, H. & Kobayashi, M. Stereostructure of komodoquinone A, a neuritogenic anthracycline, from marine *Streptomyces* sp. KS3. *Chem. Pharm. Bull. (Tokyo)* **51**, 1402–1404 (2003).
50. Ishii, K. *et al.* Decilorubicin,a new anthracycline antibiotic. *J. Antibiot. (Tokyo)* **36**, 451–453 (1983).
51. Acharya, D. *et al.* Omics technologies to understand activation of a biosynthetic gene cluster in *Micromonospora* sp. WMMB235: Deciphering keyicin biosynthesis. *ACS Chem. Biol.* **14**, 1260–1270 (2019).
52. Oki, T. *et al.* Antitumor anthracycline antibiotics, aclacinomycin A and analogues. I. Taxonomy, production, isolation and physicochemical properties. *J. Antibiot. (Tokyo)* **32**, 791–800 (1979).
53. Morioka, H. *et al.* Production and isolation of cosmomycins A, B, C and D: New differentiation inducers of friend cell F5-5. *Agric. Biol. Chem.* **49**, 1951–1958 (1985).
54. Saito, S., Katsuda, Y., Johdo, O. & Yoshimoto, A. New rhodomycin analogs, SS-288A and SS-288B, produced by a *Streptomyces violaceus* A262 mutant. *Biosci. Biotechnol. Biochem.* **59**, 135–137 (1995).
55. Wandi, B. N. *et al.* Mechanism of Two-component mono-oxygenases involved in anthracycline 1-hydroxylation. *ACS Catal.* **14**, 12359–12371 (2024).
56. Wang, R. *et al.* A BioBricks metabolic engineering platform for the biosynthesis of anthracyclinones in *Streptomyces coelicolor*. *ACS Synth. Biol.* **11**, 4193–4209 (2022).
57. Zhang, Z. *et al.* Hydroxyl regioisomerization of anthracycline catalyzed by a four-enzyme cascade. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 1554–1559 (2017).
58. Jalali, E. *et al.* Biochemical and structural studies of the carminomycin 4-*O*-methyltransferase DnrK. *J. Nat. Prod.* **87**, 798–809 (2024).
59. Yang, L. *et al.* Functional expression and purification of DoxA, a key cytochrome P450 from *Streptomyces peucetius* ATCC 27952. *PeerJ* **10**, e14373 (2022).
60. Dinis, P. *et al.* Evolution-inspired engineering of anthracycline methyltransferases. *PNAS Nexus* **2**, 1–10 (2023).
61. Grocholski, T., Dinis, P., Niiranen, L., Niemi, J. & Metsä-Ketelä, M. Divergent evolution of an atypical S-adenosyl-L-methionine-dependent monooxygenase involved in anthracycline biosynthesis. *Proc. Natl. Acad. Sci.* **112**, 9866–9871 (2015).

62. Wang, Y., Niemi, J. & Mäntsälä, P. Modification of aklavinone and aclacinomycins in vitro and in vivo by rhodomycin biosynthesis gene products. *FEMS Microbiol. Lett.* **208**, 117–122 (2002).
63. Hwang, C. K. *et al.* Expression of *Streptomyces peucetius* genes for doxorubicin resistance and aklavinone 11-hydroxylase in *Streptomyces* galilaeus ATCC 31133 and production of a hybrid aclacinomycin. *Antimicrob. Agents Chemother.* **39**, 1616–1620 (1995).
64. Johdo, O. *et al.* Anthracycline metabolites from *Streptomyces violaceus* A262. IV. New anthracycline yellamycins produced by a variant strain SC-7. *J. Antibiot. (Tokyo)* **44**, 1155–1159 (1991).
65. Guilfoile, P. G. & Hutchinson, C. R. A bacterial analog of the mdr gene of mammalian tumor cells is present in *Streptomyces peucetius*, the producer of daunorubicin and doxorubicin. *Proc Natl Acad Sci USA* **88**, 8553–8557 (1991).
66. Tenconi, E. & Rigali, S. Self-resistance mechanisms to DNA-damaging antitumor antibiotics in Actinobacteria. *Curr Opin Microbiol* **45**, 100–108 (2018).
67. Li, W., Sharma, M. & Kaur, P. The DrrAB efflux system of *Streptomyces peucetius* is a multidrug transporter of broad substrate specificity. *J. Biol. Chem.* **289**, 12633–12646 (2014).
68. Furuya, K. & Richard Hutchinson, C. The DrrC protein of *Streptomyces peucetius*, a UvrA-like protein, is a DNA-binding protein whose gene is induced by daunorubicin. *FEMS Microbiol. Lett.* **168**, 243–249 (1998).
69. Lomovskaya, N. *et al.* The *Streptomyces peucetius drrC* gene encodes a UvrA-like protein involved in daunorubicin resistance and production. *J. Bacteriol.* **178**, 3238–3245 (1996).
70. Wang, B., Guo, F., Dong, S.-H. & Zhao, H. Activation of silent biosynthetic gene clusters using transcription factor decoys. *Nat. Chem. Biol.* **15**, 111–114 (2019).
71. Furuya, K. & Hutchinson, C. R. The DnrN protein of *Streptomyces peucetius*, a pseudo-response regulator, is a DNA-binding protein involved in the regulation of daunorubicin biosynthesis. *J. Bacteriol.* **178**, 6310–6318 (1996).
72. Jiang, H. & Hutchinson, C. R. Feedback regulation of doxorubicin biosynthesis in *Streptomyces peucetius*. *Res. Microbiol.* **157**, 666–674 (2006).
73. Prija, F., Srinivasan, P., Das, S., Kattusamy, K. & Prasad, R. DnrI of *Streptomyces peucetius* binds to the resistance genes, *drrAB* and *drrC* but is activated by daunorubicin. *J. Basic Microbiol.* **57**, 862–872 (2017).
74. Cameron, D. E., Bashor, C. J. & Collins, J. J. A brief history of synthetic biology. *Nat. Rev. Microbiol.* **12**, 381–390 (2014).
75. Dai, Z., Liu, Y., Guo, J., Huang, L. & Zhang, X. Yeast synthetic biology for high-value metabolites. *FEMS Yeast Res.* n/a-n/a (2014).
76. Karim, A. S. *et al.* Deconstructing synthetic biology across scales: A conceptual approach for training synthetic biologists. *Nat. Commun.* **15**, 5425 (2024).
77. Khalil, A. S. & Collins, J. J. Synthetic biology: Applications come of age. *Nat. Rev. Genet.* **11**, 367–379 (2010).
78. Lee, N. *et al.* Synthetic biology tools for novel secondary metabolite discovery in *Streptomyces*. *J. Microbiol. Biotechnol.* **29**, 667–686 (2019).
79. Hwang, S. *et al. Streptomyces* as microbial chassis for heterologous protein expression. *Front. Bioeng. Biotechnol.* **9**, 804295 (2021).
80. Ye, S., Enghiad, B., Zhao, H. & Takano, E. Fine-tuning the regulation of Cas9 expression levels for efficient CRISPR-Cas9 mediated recombination in *Streptomyces*. *J. Ind. Microbiol. Biotechnol.* **47**, 413–423 (2020).
81. Bentley, S. D. *et al.* Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141–147 (2002).
82. Gomez-Escribano, J. P. & Bibb, M. J. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb. Biotechnol.* **4**, 207–215 (2011).
83. Ahmed, Y. *et al.* Engineering of *Streptomyces lividans for heterologous* expression of secondary metabolite gene clusters. *Microb. Cell Factories* **19**, 5 (2020).
84. Novakova, R. *et al.* Increased heterologous production of the antitumoral polyketide mithramycin A by engineered *Streptomyces lividans* TK24 strains. *Appl. Microbiol. Biotechnol.* **102**, 857–869 (2018).

85. Huang, F. *et al.* Heterologous expression and antitumor activity analysis of syringolin from *Pseudomonas syringae* pv. *syringae* B728a. *Microb. Cell Factories* **17**, 31 (2018).
86. Rodicio, M. R. & Chater, K. F. The Sa1I (SalGI) restriction-modification system of *Streptomyces albus* G. *Gene* **74**, 39–42 (1988).
87. Zaburannyi, N., Rabyk, M., Ostash, B., Fedorenko, V. & Luzhetskyy, A. Insights into naturally minimised *Streptomyces albus* J1074 genome. *BMC Genomics* **15**, 97 (2014).
88. Jiang, G. *et al.* High-yield production of herbicidal thaxtomins and thaxtomin analogs in a nonpathogenic *Streptomyces* strain. *Appl. Environ. Microbiol.* **84**, e00164-18 (2018).
89. Claesen, J. & Bibb, M. Genome mining and genetic analysis of cypemycin biosynthesis reveal an unusual class of posttranslationally modified peptides. *Proc. Natl. Acad. Sci.* **107**, 16297–16302 (2010).
90. McDonald, H. P. *et al.* Heterologous expression of the formicamycin biosynthetic gene cluster unveils glycosylated fasamycin congeners. *J. Nat. Prod.* **86**, 1677–1689 (2023).
91. Kumar, K. & Bruheim, P. A comparative study at bioprocess and metabolite levels of superhost strain *Streptomyces coelicolor* M1152 and its derivative M1581 heterologously expressing chloramphenicol biosynthetic gene cluster. *Biotechnol. Bioeng.* **119**, 145–161 (2022).
92. Kepplinger, B. *et al.* Discovery, isolation, heterologous expression and mode-of-action studies of the antibiotic polyketide tatiomicin from *Amycolatopsis* sp. DEM30355. *Sci. Rep.* **12**, 15579 (2022).
93. Jones, A. C. *et al.* Phage P1-derived artificial chromosomes facilitate heterologous expression of the FK506 gene cluster. *PLoS ONE* **8**, e69319 (2013).
94. Flinspach, K. *et al.* Heterologous expression of the biosynthetic gene clusters of coumermycin A_1, clorobiocin and caprazamycins in genetically modified *Streptomyces coelicolor* strains. *Biopolymers* **93**, 823–832 (2010).
95. Hashimoto, T. *et al.* A novel oxazole-containing tetraene compound, JBIR-159, produced by heterologous expression of the cryptic trans-AT type polyketide synthase biosynthetic gene cluster. *J. Antibiot. (Tokyo)* **74**, 354–358 (2021).
96. Sakai, K. *et al.* Production of kinanthraquinone D with antimalarial activity by heterologous gene expression and biotransformation in *Streptomyces lividans* TK23. *J. Nat. Prod.* **87**, 855–860 (2024).
97. Smanski, M. J. *et al.* Expression of the Platencin biosynthetic gene cluster in heterologous hosts yielding new platencin congeners. *J. Nat. Prod.* **75**, 2158–2167 (2012).
98. Jiang, C. *et al.* Establishing an efficient salinomycin biosynthetic pathway in three heterologous *Streptomyces* hosts by constructing a 106-kb multioperon artificial gene cluster. *Biotechnol. Bioeng.* **118**, 4668–4677 (2021).
99. Park, S. R. *et al.* Engineering of plant-specific phenylpropanoids biosynthesis in *Streptomyces venezuelae*. *J. Biotechnol.* **141**, 181–188 (2009).
100. Jung, W. S. *et al.* Heterologous expression of tylosin polyketide synthase and production of a hybrid bioactive macrolide in *Streptomyces venezuelae*. *Appl. Microbiol. Biotechnol.* **72**, 763–769 (2006).
101. Han, A. R. *et al.* Development of a *Streptomyces venezuelae*-based combinatorial biosynthetic system for the production of glycosylated derivatives of doxorubicin and its biosynthetic intermediates. *Appl. Environ. Microbiol.* **77**, 4912–4923 (2011).
102. Maharjan, S., Park, J. W., Yoon, Y. J., Lee, H. C. & Sohng, J. K. Metabolic engineering of *Streptomyces venezuelae* for malonyl-CoA biosynthesis to enhance heterologous production of polyketides. *Biotechnol. Lett.* **32**, 277–282 (2010).
103. Myronovskyi, M. *et al.* Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. *Metab. Eng.* **49**, 316–324 (2018).
104. Komatsu, M. *et al.* Engineered *Streptomyces avermitilis* host for heterologous expression of biosynthetic gene cluster for secondary metabolites. *ACS Synth. Biol.* **2**, 384–396 (2013).
105. Wang, W., Zheng, G. & Lu, Y. Recent advances in strategies for the cloning of natural product biosynthetic gene clusters. *Front. Bioeng. Biotechnol.* **9**, 692797 (2021).

106. Saito, I. & Stark, G. R. Charomids: Cosmid vectors for efficient cloning and mapping of large or small restriction fragments. *Proc. Natl. Acad. Sci.* **83**, 8664–8668 (1986).
107. Collins, J. & Hohn, B. Cosmids: A type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage lambda heads. *Proc. Natl. Acad. Sci.* **75**, 4242–4246 (1978).
108. Kim, U.-J., Shizuya, H., De Jong, P. J., Birren, B. & Simon, M. I. Stable propagation of cosmid sized human DNA inserts in an F factor based vector. *Nucleic Acids Res.* **20**, 1083–1085 (1992).
109. Meng, J. *et al.* Improvement of pristinamycin I (PI) production in *Streptomyces pristinaespiralis* by metabolic engineering approaches. *Synth. Syst. Biotechnol.* **2**, 130–136 (2017).
110. Bajpai, B. High Capacity Vectors. in *Advances in Biotechnology* (eds. Ravi, I., Baunthiyal, M. & Saxena, J.) 1–10 (Springer India, New Delhi, 2014).
111. Loannou, P. A. *et al.* A new bacteriophage P1–derived vector for the propagation of large human DNA fragments. *Nat. Genet.* **6**, 84–89 (1994).
112. Shizuya, H. *et al.* Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. *Proc. Natl. Acad. Sci.* **89**, 8794–8797 (1992).
113. Murray, A. W. & Szostak, J. W. Construction of artificial chromosomes in yeast. *Nature* **305**, 189–193 (1983).
114. Nah, H.-J., Pyeon, H.-R., Kang, S.-H., Choi, S.-S. & Kim, E.-S. Cloning and heterologous expression of a large-sized natural product biosynthetic gene cluster in *Streptomyces* species. *Front. Microbiol.* **8**, (2017).
115. Dracopoli, N. C., Haines, J. L. & Korf, B. R. Current protocols in human genetics. (1994).
116. Miao, V. *et al.* Daptomycin biosynthesis in *Streptomyces roseosporus*: Cloning and analysis of the gene cluster and revision of peptide stereochemistry. *Microbiology* **151**, 1507–1523 (2005).
117. Sosio, M. *et al.* Artificial chromosomes for antibiotic-producing actinomycetes. *Nat. Biotechnol.* **18**, 343–345 (2000).
118. Olorunniji, F. J. *et al.* Gated rotation mechanism of site-specific recombination by ϕC31 integrase. *Proc. Natl. Acad. Sci.* **109**, 19661–19666 (2012).
119. Vara, J., Lewandowska-Skarbek, M., Wang, Y. G., Donadio, S. & Hutchinson, C. R. Cloning of genes governing the deoxysugar portion of the erythromycin biosynthesis pathway in S*accharopolyspora erythraea* (*Streptomyces erythreus*). *J. Bacteriol.* **171**, 5872–5881 (1989).
120. Brown, K. V. & Nybo, S. E. Complete sequences of pIJ101-based *Streptomyces-Escherichia coli* shuttle vectors. *Access Microbiol.* **6**, (2024).
121. Bierman, M. *et al.* Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to *Streptomyces* spp. *Gene* **116**, 43–49 (1992).
122. Phelan, R. M. *et al.* Development of next generation synthetic biology tools for use in *Streptomyces venezuelae*. *ACS Synth. Biol.* **6**, 159–166 (2017).
123. Aubry, C., Pernodet, J.-L. & Lautru, S. Modular and integrative vectors for synthetic biology applications in *Streptomyces* spp. *Appl. Environ. Microbiol.* **85**, e00485-19 (2019).
124. Nguyen, J. T., Riebschleger, K. K., Brown, K. V., Gorgijevska, N. M. & Nybo, S. E. A BioBricks toolbox for metabolic engineering of the tetracenomycin pathway. *Biotechnol. J.* **17**, 2100371 (2022).
125. Peng, B., Williams, T. C., Henry, M., Nielsen, L. K. & Vickers, C. E. Controlling heterologous gene expression in yeast cell factories on different carbon substrates and across the diauxic shift: A comparison of yeast promoter activities. *Microb. Cell Factories* **14**, 91 (2015).
126. Wang, W. *et al.* An engineered strong promoter for Streptomycetes. *Appl. Environ. Microbiol.* **79**, 4484–4492 (2013).
127. Wang, W. *et al.* Development of a synthetic oxytetracycline-inducible expression system for Streptomycetes using *de novo* characterized genetic parts. *ACS Synth. Biol.* **5**, 765–773 (2016).
128. Li, X., Sang, Z., Zhao, X. & Wen, Y. Metabolic engineering of *Streptomyces roseosporus* for increased production of clinically important antibiotic daptomycin. *Microb. Biotechnol.* **17**, e70038 (2024).
129. Bibb, M. J., Janssen, G. R. & Ward, J. M. Cloning and analysis of the promoter region of the erythromycin resistance gene (ermE) of *Streptomyces* erythraeus. *Gene* **38**, 215–226 (1985).

130. Wang, X., Deng, Z. & Liu, T. Marker-free system using ribosomal promoters enhanced xylose/glucose isomerase production in *Streptomyces rubiginosus*. *Biotechnol. J.* **14**, 1900114 (2019).
131. Guo, W. *et al.* Identification and characterization of a strong constitutive promoter stnYp for activating biosynthetic genes and producing natural products in *Streptomyces*. *Microb. Cell Factories* **22**, 127 (2023).
132. Bai, C. *et al.* Exploiting a precise design of universal synthetic modular regulatory elements to unlock the microbial natural products in *Streptomyces*. *Proc. Natl. Acad. Sci.* **112**, 12181–12186 (2015).
133. Zhang, J. *et al.* Efficient multiplex genome editing in *streptomyces* via engineered CRISPR-Cas12a systems. *Front. Bioeng. Biotechnol.* **8**, 726 (2020).
134. Glöckle, A., Schuler, S., Einsiedler, M. & Gulder, T. A. M. A plug-and-play system for polycyclic tetramate macrolactam production and functionalization. *Microb. Cell Factories* **24**, 13 (2025).
135. Labes, G., Bibb, M. & Wohlleben, W. Isolation and characterization of a strong promoter element from the *Streptomyces ghanaensis* phage I19 using the gentamicin resistance gene (aacC1) of Tn1696 as reporter. *Microbiology* **143**, 1503–1512 (1997).
136. Yin, S. *et al.* Identification of a cluster-situated activator of oxytetracycline biosynthesis and manipulation of its expression for improved oxytetracycline production in *Streptomyces rimosus*. *Microb. Cell Factories* **14**, 46 (2015).
137. Siegl, T., Tokovenko, B., Myronovskyi, M. & Luzhetskyy, A. Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes. *Metab. Eng.* **19**, 98–106 (2013).
138. Horbal, L., Fedorenko, V. & Luzhetskyy, A. Novel and tightly regulated resorcinol and cumate-inducible expression systems for *Streptomyces* and other actinobacteria. *Appl. Microbiol. Biotechnol.* **98**, 8641–8655 (2014).
139. Bauman, K. D. *et al.* Refactoring the cryptic streptophenazine biosynthetic gene cluster unites phenazine, polyketide, and nonribosomal peptide biochemistry. *Cell Chem. Biol.* **26**, 724-736.e7 (2019).
140. Luo, Y., Zhang, L., Barton, K. W. & Zhao, H. Systematic identification of a panel of strong constitutive promoters from *Streptomyces albus*. *ACS Synth. Biol.* **4**, 1001–1010 (2015).
141. Li, Z. *et al.* Positive regulation of a LuxR family protein, MilO, in mildiomycin biosynthesis. *Appl. Environ. Microbiol.* **91**, e01654-24 (2025).
142. Chiu, M. L. *et al.* Broad spectrum thiopeptide recognition specificity of the *Streptomyces lividans* TipAL protein and its role in regulating gene expression. *J. Biol. Chem.* **274**, 20578–20586 (1999).
143. Takano, E., White, J., Thompson, C. J. & Bibb, M. J. Construction of thiostrepton-inducible, high-copy-number expression vectors for use in *Streptomyces* spp. *Gene* **166**, 133–137 (1995).
144. Holgado, G. G. *et al.* Radamycin, a novel thiopeptide produced by *Streptomyces* sp. RSP9. I. taxonomy, fermentation, isolation and biological activities. *J. Antibiot. (Tokyo)* **55**, 383–390 (2002).
145. Miranzo, D., Seco, E. M., Cuesta, T. & Malpartida, F. Isolation and characterization of *pcsB*, the gene for a polyene carboxamide synthase that tailors pimaricin into AB-400. *Appl. Microbiol. Biotechnol.* **85**, 1809–1819 (2010).
146. Noguchi, Y. *et al.* Development of a strictly regulated xylose-induced expression system in *Streptomyces*. *Microb. Cell Factories* **17**, 151 (2018).
147. Komeda, H., Hori, Y., Kobayashi, M. & Shimizu, S. Transcriptional regulation of the *Rhodococcus rhodochrous* J1 *nitA* gene encoding a nitrilase. *Proc. Natl. Acad. Sci.* **93**, 10572–10577 (1996).
148. Rodríguez-García, A., Combes, P., Pérez-Redondo, R., Smith, M. C. A. & Smith, M. C. M. Natural and synthetic tetracycline-inducible promoters for use in the antibiotic-producing bacteria *Streptomyces*. *Nucleic Acids Res.* **33**, e87–e87 (2005).

149. Dangel, V., Westrich, L., Smith, M. C. M., Heide, L. & Gust, B. Use of an inducible promoter for antibiotic production in a heterologous host. *Appl. Microbiol. Biotechnol.* **87**, 261–269 (2010).
150. Hernandez-Abanto, S. M., Woolwine, S. C., Jain, S. K. & Bishai, W. R. Tetracycline-inducible gene expression in mycobacteria within an animal host using modified *Streptomyces* tcp830 regulatory elements. *Arch. Microbiol.* **186**, 459–464 (2006).
151. Bai, C. & Van Wezel, G. P. CUBIC: A versatile cumate-based inducible CRISPRi system in *Streptomyces*. *ACS Synth. Biol.* **12**, 3143–3147 (2023).
152. You, L. *et al.* Structural basis for intrinsic transcription termination. *Nature* **613**, 783–789 (2023).
153. Otsuka, J. & Kunisawa, T. Characteristic base sequence patterns of promoter and terminator sites in φX174 and fd phage DNAs. *J. Theor. Biol.* **97**, 415–436 (1982).
154. Brown, J. The iGEM competition: Building with biology. *IET Synth. Biol.* **1**, 3–6 (2007).
155. Huff, J., Czyz, A., Landick, R. & Niederweis, M. Taking phage integration to the next level as a genetic tool for mycobacteria. *Gene* **468**, 8–19 (2010).
156. Chen, Y.-J. *et al.* Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat. Methods* **10**, 659–664 (2013).
157. Larionov, V., Kouprina, N., Solomon, G., Barrett, J. C. & Resnick, M. A. Direct isolation of human *BRCA2* gene by transformation-associated recombination in yeast. *Proc. Natl. Acad. Sci.* **94**, 7384–7387 (1997).
158. Kouprina, N. & Larionov, V. TAR cloning: Insights into gene function, long-range haplotypes and genome structure and evolution. *Nat. Rev. Genet.* **7**, 805–812 (2006).
159. Yamanaka, K. *et al.* Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proc. Natl. Acad. Sci.* **111**, 1957–1962 (2014).
160. Fu, J. *et al.* Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat. Biotechnol.* **30**, 440–446 (2012).
161. Yin, S. *et al.* Heterologous expression of oxytetracycline biosynthetic gene cluster in *Streptomyces venezuelae* WVR2006 to improve production level and to alter fermentation process. *Appl. Microbiol. Biotechnol.* **100**, 10563–10572 (2016).
162. Du, D. *et al.* Genome engineering and direct cloning of antibiotic gene clusters via phage ϕBT1 integrase-mediated site-specific recombination in *Streptomyces*. *Sci. Rep.* **5**, 8740 (2015).
163. Nah, H.-J., Woo, M.-W., Choi, S.-S. & Kim, E.-S. Precise cloning and tandem integration of large polyketide biosynthetic gene cluster using *Streptomyces* artificial chromosome system. *Microb. Cell Factories* **14**, 140 (2015).
164. Avilan, L. Assembling Multiple Fragments: The Gibson Assembly. in *DNA Manipulation and Analysis* (ed. Scarlett, G.) vol. 2633 45–53 (Springer US, New York, NY, 2023).
165. Li, L., Jiang, W. & Lu, Y. A Modified Gibson Assembly Method for Cloning Large DNA Fragments with High GC Contents. in *Synthetic Metabolic Pathways: Methods and Protocols* (eds. Jensen, M. K. & Keasling, J. D.) 203–209 (Springer, New York, NY, 2018).
166. Casini, A. *et al.* One-pot DNA construction for synthetic biology: The Modular Overlap-Directed Assembly with Linkers (MODAL) strategy. *Nucleic Acids Res.* **42**, e7–e7 (2014).
167. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE* **3**, e3647 (2008).
168. Marillonnet, S. & Grützner, R. Synthetic DNA assembly using golden gate cloning and the hierarchical modular cloning pipeline. *Curr. Protoc. Mol. Biol.* **130**, e115 (2020).
169. Shetty, R. P., Endy, D. & Knight, T. F. Engineering BioBrick vectors from BioBrick parts. *J. Biol. Eng.* **2**, 5 (2008).
170. Zeng, Y. *et al.* Advanced genome-editing technologies enable rapid and large-scale generation of genetic variants for strain engineering and synthetic biology. *Curr. Opin. Microbiol.* **69**, 102175 (2022).
171. Krysenko, S. Current approaches for genetic manipulation of *Streptomyces* spp.—key bacteria for biotechnology and environment. *BioTech* **14**, 3 (2025).
172. Zhao, Y., Li, G., Chen, Y. & Lu, Y. Challenges and advances in genome editing technologies in *Streptomyces*. *Biomolecules* **10**, 734 (2020).

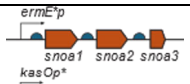
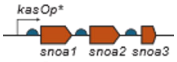
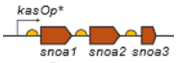
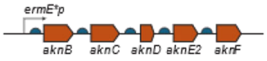
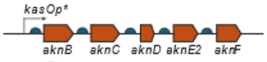
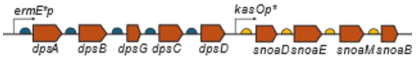
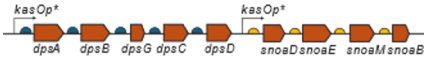
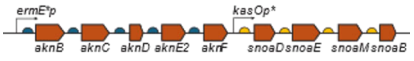
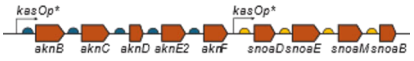
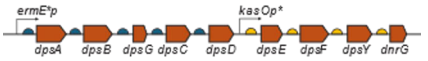
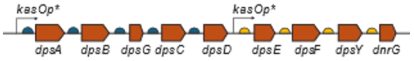
173. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. *Practical Streptomyces Genetics*. vol. 291 (John Innes Foundation Norwich, 2000).
174. Gust, B., Challis, G. L., Fowler, K., Kieser, T. & Chater, K. F. PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl. Acad. Sci.* **100**, 1541–1546 (2003).
175. Schweizer, H. P. Applications of the *Saccharomyces cerevisiae* Flp-*FRT* System in Bacterial Genetics. *Microb. Physiol.* **5**, 67–77 (2003).
176. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97**, 6640–6645 (2000).
177. Siegl, T. & Luzhetskyy, A. Actinomycetes genome engineering approaches. *Antonie Van Leeuwenhoek* **102**, 503–516 (2012).
178. Herrmann, S. *et al.* Site-specific recombination strategies for engineering actinomycete genomes. *Appl. Environ. Microbiol.* **78**, 1804–1812 (2012).
179. Tao, W., Yang, A., Deng, Z. & Sun, Y. CRISPR/Cas9-based editing of *Streptomyces* for discovery, characterization, and production of natural products. *Front. Microbiol.* **9**, 1660 (2018).
180. Makarova, K. S. *et al.* An updated evolutionary classification of CRISPR–Cas systems. *Nat. Rev. Microbiol.* **13**, 722–736 (2015).
181. Wang, J. Y. & Doudna, J. A. CRISPR technology: A decade of genome editing is only the beginning. *Science* **379**, eadd8643 (2023).
182. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
183. Alberti, F. & Corre, C. Editing streptomycete genomes in the CRISPR/Cas9 age. *Nat. Prod. Rep.* **36**, 1237–1248 (2019).
184. Kim, D. G. *et al.* Engineered CRISPR-Cas9 for *Streptomyces* sp. genome editing to improve specialized metabolite production. *Nat. Commun.* **16**, 874 (2025).
185. Li, L. *et al.* CRISPR-Cpf1-assisted multiplex genome editing and transcriptional repression in *Streptomyces*. *Appl. Environ. Microbiol.* **84**, e00827-18 (2018).
186. Zhou, Q. *et al.* Repurposing endogenous type I-E CRISPR-Cas systems for natural product discovery in *Streptomyces*. *Nat. Commun.* **15**, 9833 (2024).
187. Yeo, W. L. *et al.* Characterization of Cas proteins for CRISPR-Cas editing in Streptomycetes. *Biotechnol. Bioeng.* **116**, 2330–2338 (2019).
188. Martín, J.-F. & Liras, P. Engineering of regulatory cascades and networks controlling antibiotic biosynthesis in *Streptomyces*. *Curr. Opin. Microbiol.* **13**, 263–273 (2010).
189. Medema, M. H., Breitling, R., Bovenberg, R. & Takano, E. Exploiting plug-and-play synthetic biology for drug discovery and production in microorganisms. *Nat. Rev. Microbiol.* **9**, 131–137 (2011).
190. Kang, H.-S. & Kim, E.-S. Recent advances in heterologous expression of natural product biosynthetic gene clusters in *Streptomyces* hosts. *Curr. Opin. Biotechnol.* **69**, 118–127 (2021).
191. Larson, C. B., Crüsemann, M. & Moore, B. S. PCR-independent method of transformation-associated recombination reveals the cosmomycin biosynthetic gene cluster in an ocean Streptomycete. *J. Nat. Prod.* **80**, 1200–1204 (2017).
192. Qian, Z. *et al.* Discovery of the streptoketides by direct cloning and rapid heterologous expression of a cryptic PKS II gene cluster from *Streptomyces* sp. Tü 6314. *J. Org. Chem.* **85**, 664–673 (2020).
193. Pyeon, H., Nah, H.-J., Kang, S.-H., Choi, S.-S. & Kim, E.-S. Heterologous expression of pikromycin biosynthetic gene cluster using *Streptomyces* artificial chromosome system. *Microb. Cell Factories* **16**, 96 (2017).
194. Tan, G.-Y. & Liu, T. Rational synthetic pathway refactoring of natural products biosynthesis in actinobacteria. *Metab. Eng.* **39**, 228–236 (2017).
195. Fowler, M. & Beck, K. *Refactoring: Improving the Design of Existing Code*. (Addison-Wesley, Boston Columbus New York San Francisco Amsterdam Cape Town Dubai London Munich, 2019).

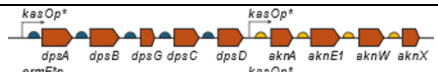
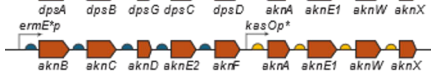
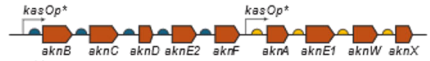
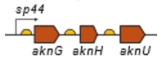
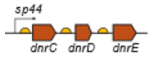
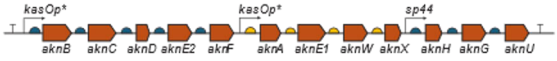
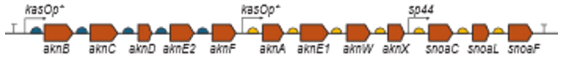
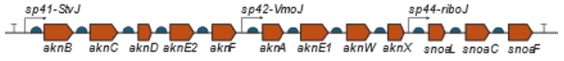
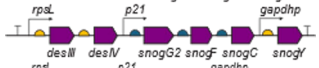
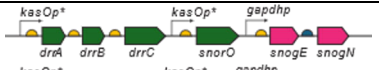
196. Chan, L. Y., Kosuri, S. & Endy, D. Refactoring bacteriophage T7. *Mol. Syst. Biol.* **1**, 2005.0018 (2005).
197. Temme, K., Zhao, D. & Voigt, C. A. Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*. *Proc. Natl. Acad. Sci.* **109**, 7085–7090 (2012).
198. Li, L., Maclntyre, L. W. & Brady, S. F. Refactoring biosynthetic gene clusters for heterologous production of microbial natural products. *Curr. Opin. Biotechnol.* **69**, 145–152 (2021).
199. Olano, C. *et al.* Activation and identification of five clusters for secondary metabolites in *Streptomyces albus* J 1074. *Microb. Biotechnol.* **7**, 242–256 (2014).
200. Montiel, D., Kang, H.-S., Chang, F.-Y., Charlop-Powers, Z. & Brady, S. F. Yeast homologous recombination-based promoter engineering for the activation of silent natural product biosynthetic gene clusters. *Proc. Natl. Acad. Sci.* **112**, 8953–8958 (2015).
201. Li, J., Kim, S. G. & Blenis, J. Rapamycin: One drug, many effects. *Cell Metab.* **19**, 373–379 (2014).
202. He, F. *et al.* CRISETR: An efficient technology for multiplexed refactoring of biosynthetic gene clusters. *Nucleic Acids Res.* **52**, 11378–11393 (2024).
203. Luo, Y. *et al.* Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster. *Nat. Commun.* **4**, 2894 (2013).
204. Shao, Z. *et al.* Refactoring the silent spectinabilin gene cluster using a plug-and-play scaffold. *ACS Synth. Biol.* **2**, 662–669 (2013).
205. Rutledge, P. J. & Challis, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat Rev Microbiol* **13**, 509–523 (2015).
206. Ang, E. L., Sun, H., Liu, Z. & Zhao, H. Recent advances in combinatorial biosynthesis for drug discovery. *Drug Des. Devel. Ther.* 823 (2015).
207. Kim, E., Moore, B. S. & Yoon, Y. J. Reinvigorating natural product combinatorial biosynthesis with synthetic biology. *Nat. Chem. Biol.* **11**, 649–659 (2015).
208. McCormick, J. R. D., Johnson, Sylvia. & Sjolander, N. O. Biosynthesis of the tetracyclines. V. naphthacenic precursors. *J. Am. Chem. Soc.* **85**, 1692–1694 (1963).
209. Hopwood, D. A. *et al.* Production of 'hybrid' antibiotics by genetic engineering. *Nature* **314**, 642–644 (1985).
210. Bravo-Rodriguez, K. *et al.* Predicted incorporation of non-native substrates by a polyketide synthase yields bioactive natural product derivatives. *ChemBioChem* **15**, 1991–1997 (2014).
211. Niemi, J. *et al.* Hybrid anthracycline antibiotics: Production of new anthracyclines by cloned genes from *Streptomyces purpurascens* in *Streptomyces galilaeus*. *Microbiology* **140**, 1351–1358 (1994).
212. Kunnari, T., Niemi, J., Ylihonko, K., Mäntsälä, P. & Hakala, J. Hybrid anthracyclines by a genetically engineered *Streptomyces galilaeus* mutant. *Bioorg. Med. Chem. Lett.* **7**, 725–726 (1997).
213. Kim, H. S., Hong, Y. S., Kim, Y. H., Yoo, O. J. & Lee, J. J. New anthracycline metabolites produced by the aklavinone 11-hydroxylase gene in *Streptomyces galilaeus* ATCC 31133. *J. Antibiot. (Tokyo)* **49**, 355–360 (1996).
214. Kunnari, T. J., Ylihonko, K. P. J., Klika, K. D., Mäntsälä, P. I. & Hakala, J. M. L. Hybrid anthracyclines from a genetically engineered *Streptomyces galilaeus* mutant. *J. Org. Chem.* **65**, 2851–2855 (2000).
215. Jain, K. K. *et al.* A prospective randomized comparison of epirubicin and doxorubicin in patients with advanced breast cancer. *J. Clin. Oncol.* **3**, 818–826 (1985).
216. Madduri, K. *et al.* Production of the antitumor drug epirubicin (4′-epidoxorubicin) and its precursor by a genetically engineered strain of *Streptomyces peucetius*. *Nat. Biotechnol.* **16**, 69–74 (1998).
217. Olano, C. *et al.* Glycosylated derivatives of steffimycin: Insights into the role of the sugar moieties for the biological activity. *ChemBioChem* **9**, 624–633 (2008).
218. Luzhetskyy, A. *et al.* Aranciamycin analogs generated by combinatorial biosynthesis show improved antitumor activity. *Appl. Microbiol. Biotechnol.* **80**, 15–19 (2008).

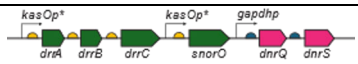
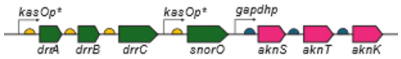
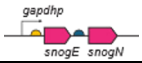
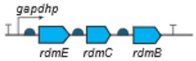
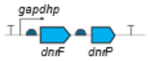
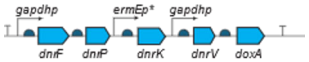
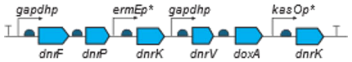
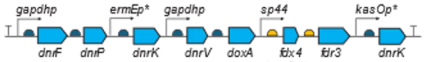
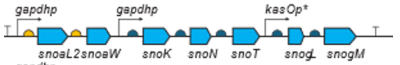
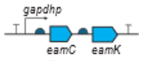
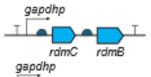
219. Kim, E. *et al.* One-pot combinatorial biosynthesis of glycosylated anthracyclines by cocultivation of *Streptomyces* strains producing aglycones and nucleotide deoxysugars. *ACS Comb. Sci.* **19**, 262–270 (2017).
220. Castro, J. F. *et al.* Identification and heterologous expression of the chaxamycin biosynthesis gene cluster from *Streptomyces leeuwenhoekii*. *Appl. Environ. Microbiol.* **81**, 5820–5831 (2015).
221. Shaaban, K. A. *et al.* Herbimycins D–F, ansamycin analogues from *Streptomyces* sp. RM-7-15. *J. Nat. Prod.* **76**, 1619–1626 (2013).
222. Myronovskyi, M. & Luzhetskyy, A. Native and engineered promoters in natural product discovery. *Nat. Prod. Rep.* **33**, 1006–1019 (2016).
223. Eckardt, K., Schumann, G., Tresselt, D. & Ihn, W. Biosynthesis of anthracyclinones. Isolation of a new early cyclization product aklaviketone. *J. Antibiot. (Tokyo)* **41**, 788–793 (1988).
224. Lou, C., Stanton, B., Chen, Y.-J., Munsky, B. & Voigt, C. A. Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat. Biotechnol.* **30**, 1137–1142 (2012).
225. Torkkell, S. *et al.* The entire nogalamycin biosynthetic gene cluster of *Streptomyces* nogalater: Characterization of a 20-kb DNA region and generation of hybrid structures. *Mol. Genet. Genomics* **266**, 276–288 (2001).
226. Kantola, J. *et al.* Elucidation of anthracyclinone biosynthesis by stepwise cloning of genes for anthracyclines from three different *Streptomyces* spp. *Microbiology* **146**, 155–163 (2000).
227. Räty, K. *et al.* Cloning and characterization of *Streptomyces galilaeus* aclacinomycins polyketide synthase (PKS) cluster. *Gene* **293**, 115–122 (2002).
228. Tirkkonen, H. *et al.* Engineering BioBricks for deoxysugar biosynthesis and generation of new tetracenomycins. *ACS Omega* **8**, 21237–21253 (2023).
229. Siitonen, V., Blauenburg, B., Kallio, P., Mäntsälä, P. & Metsä-Ketelä, M. Discovery of a two-component monooxygenase SnoaW/SnoaL2 involved in nogalamycin biosynthesis. *Chem. Biol.* **19**, 638–646 (2012).
230. Siitonen, V. *et al.* Divergent non-heme iron enzymes in the nogalamycin biosynthetic pathway. *Proc. Natl. Acad. Sci.* **113**, 5251–5256 (2016).
231. Nji Wandi, B., Siitonen, V., Palmu, K. & Metsä-Ketelä, M. The rieske oxygenase SnoT catalyzes 2″-hydroxylation of L-rhodosamine in nogalamycin biosynthesis. *ChemBioChem* 1–6 (2020).
232. Niemi, J. *et al.* Characterization of aklavinone-11-hydroxylase from *Streptomyces purpurascens*. *Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol.* **1430**, 57–64 (1999).
233. Madduri, K. & Hutchinson, C. R. Functional characterization and transcriptional analysis of a gene cluster governing early and late steps in daunorubicin biosynthesis in *Streptomyces peucetius*. *J. Bacteriol.* **177**, 3879–3884 (1995).
234. Jansson, A., Koskiniemi, H., Mäntsälä, P., Niemi, J. & Schneider, G. Crystal structure of a ternary complex of DnrK, a methyltransferase in daunorubicin biosynthesis, with bound products*. *J. Biol. Chem.* **279**, 41149–41156 (2004).
235. van Gelder, M. A. *et al.* Re-exploring the anthracycline chemical space for better anti-cancer compounds. *J. Med. Chem.* **66**, 11390–11398 (2023).

Appendix Tables

Appendix table 1. List of plasmids constructed in this study

No.	Plasmid name	Backbone	SBOL diagram	Reference
Aglycone				
1	pSET-S1	pSET152BB		This study
2	pSET-S2	pSET152BB		This study
3	pSET-S3	pSET152BB		This study
4	pSET-A1	pSET152BB		This study
5	pSET-A2	pSET152BB		This study
6	pSET-D1	pSET152BB		This study
7	pSET-D2	pSET152BB		This study
8	pSET-S2S5	pSET152BB		This study
9	pSET-D1S5	pSET152BB		This study
10	pSET-D2S5	pSET152BB		This study
11	pSET-A1S5	pSET152BB		This study
12	pSET-A2S5	pSET152BB		This study
13	pSET-D1D5	pSET152BB		This study
14	pSET-D2D5	pSET152BB		This study
15	pSET-A1D5	pSET152BB		This study
16	pSET-A2D5	pSET152BB		This study
17	pSET-D1A5	pSET152BB		This study

No.	Plasmid name	Backbone	SBOL diagram	Reference
18	pSET-D2A5	pSET152BB		This study
19	pSET-A1A5	pSET152BB		This study
20	pSET-A2A5	pSET152BB		This study
21	pTG-S6	pENTG1		This study
22	pTG-S7	pENTG1		This study
23	pTG-A6	pENTG1		This study
24	pTG-D6	pENTG1		This study
25	pEN10001	pSET154BB		This study
26	pEN10002	pSET154BB		This study
27	pEN10003	pSET154BB		This study
28	pEN10004	pSET154BB		This study
29	pRW10000	pSET154BB		This study
30	pAKV7	pSET154BB		This study
31	pEAKV2	pOSV821		This study
TDP-carbohydrate				
32	pDAU	pOSV820		This study
33	pRHO	pOSV820		This study
34	pACO	pOSV820		This study
35	pDENOG	pOSV820		This study
36	pRHO_DEFUC_RHODI	pOSV820		This study
37	pRHO_DENOG	pOSV820		This study
Resistance and/or GT				
38	pR_SnogEN	pOSV808		This study
39	pR_SnogEND	pOSV808		This study

No.	Plasmid name	Backbone	SBOL diagram	Reference
40	pR_DnrQS	pOSV808		This study
41	pR_AknST	pOSV808		This study
42	pR_AknSTK	pOSV808		This study
43	pSnogEN	pOSV808		This study
Post-PKS tailoring				
44	pF	pENTG3		This study
45	pECB	pENTG3		This study
46	p15_16	pENTG3		This study
47	pFP	pENSV3		This study
48	pFPK	pENSV3		This study
49	pFP_K	pENSV3		This study
50	pFP_K_VA	pENTG3		This study
51	pFP_K_VA_x4_r3	pENTG3		This study
52	pFP_K_VA_K	pENTG3		This study
53	pFP_K_VA_x4_r3_K	pENTG3		This study
54	pL2W_KNT_LM	pENTG3		This study
55	pSV-dnrF	pENSV3		This study
56	pSV-eamCK	pENSV3		This study
57	pSV-rdmCB	pENSV3		This study
58	pSV-rdmECB	pENSV3		This study
59	pSV-kstA1516	pENSV3		This study
60	pSV-kstA15161011	pENSV3		This study

Appendix table 2. List of *Streptomyces* strains constructed in this study.

Strain	Parental strain	Genotype and relevant characteristics	Note	Reference
M1152Δ*matAB*	*S. coelicolor* M1152	SCP1- SCP2- Δ*act* Δ*red* Δ*cpk* Δ*cda rpoB*(C1298T) *Δsco2961−2962*	chassis	This study
Δ*snogE*+*snogE*	*S.albus*/Δ*snogE*	*S.albus*/Δ*snogE*::pSnogEN	snogE complementation	This study
Δ*snogN*+*snogN*	*S.albus*/Δ*snogN*	*S.albus*/Δ*snogN*::pSnogEN	snogN complementation	This study
AKV1	M1152Δ*matAB*	M1152Δ*matAB*::pEN10001	Producer of **1**	This study
AKV7	M1152Δ*matAB*	M1152Δ*matAB*::pAKV7	Producer of **1**	This study
EAKV2	M1152Δ*matAB*	M1152Δ*matAB*::pEAKV2	Producer of **4**	This study
AURA2	M1152Δ*matAB*	M1152Δ*matAB*::pAURA2	Producer of **3**	This study
NOG	M1152Δ*matAB*	M1152Δ*matAB*::pRW10000	Producer of **2**	This study
NOG2	M1152Δ*matAB*	M1152Δ*matAB*::pNOG2	Producer of **2**	This study
S1	M1152Δ*matAB*	M1152Δ*matAB*::pSET-S1	C-20 minPKS	This study
S2	M1152Δ*matAB*	M1152Δ*matAB*::pSET-S2	C-20 minPKS	This study
S3	M1152Δ*matAB*	M1152Δ*matAB*::pSET-S3	C-20 minPKS	This study
A1	M1152Δ*matAB*	M1152Δ*matAB*::pSET-A1	C-21 minPKS	This study
A2	M1152Δ*matAB*	M1152Δ*matAB*::pSET-A2	C-21 minPKS	This study
D1	M1152Δ*matAB*	M1152Δ*matAB*::pSET-D1	C-21 minPKS	This study
D2	M1152Δ*matAB*	M1152Δ*matAB*::pSET-D2	C-21 minPKS	This study
S2S5	M1152Δ*matAB*	M1152Δ*matAB*::pSET-S2S5	Nogalonic acid	This study
D1S5	M1152Δ*matAB*	M1152Δ*matAB*::pSET-D1S5	Aklanonic acid	This study
D2S5	M1152Δ*matAB*	M1152Δ*matAB*::pSET-D2S5	Aklanonic acid	This study
A1S5	M1152Δ*matAB*	M1152Δ*matAB*::pSET-A1S5	Aklanonic acid	This study
A2S5	M1152Δ*matAB*	M1152Δ*matAB*::pSET-A2S5	Aklanonic acid	This study
D1D5	M1152Δ*matAB*	M1152Δ*matAB*::pSET-D1D2	Aklanonic acid	This study
D2D5	M1152Δ*matAB*	M1152Δ*matAB*::pSET-D2D5	Aklanonic acid	This study
A1D5	M1152Δ*matAB*	M1152Δ*matAB*::pSET-A1D5	Aklanonic acid	This study
A2D5	M1152Δ*matAB*	M1152Δ*matAB*::pSET-A2D5	Aklanonic acid	This study
D1A5	M1152Δ*matAB*	M1152Δ*matAB*::pSET-D1A5	Aklanonic acid	This study
D2A5	M1152Δ*matAB*	M1152Δ*matAB*::pSET-D2A5	Aklanonic acid	This study
A1A5	M1152Δ*matAB*	M1152Δ*matAB*::pSET-A1A5	Aklanonic acid	This study
A2A5	M1152Δ*matAB*	M1152Δ*matAB*::pSET-A2A5	Aklanonic acid	This study
S2S5S6	M1152Δ*matAB*	M1152Δ*matAB*::pSET-S2S5::pTG-S6	Producer of **2**	This study
S2S5S7	M1152Δ*matAB*	M1152Δ*matAB*::pSET-S2S5::pTG-S7	Producer of **2**	This study
S2S5D6	M1152Δ*matAB*	M1152Δ*matAB*::pSET-S2S5::pTG-D6	Producer of **3**	This study
S2S5A6	M1152Δ*matAB*	M1152Δ*matAB*::pSET-S2S5::pTG-A6	Producer of **3**	This study
A2A5S7	M1152Δ*matAB*	M1152Δ*matAB*::pSET-A2A5::pTG-S7	Producer of **4**	This study
A2A5D6	M1152Δ*matAB*	M1152Δ*matAB*::pSET-A2A5::pTG-D6	Producer of **1**	This study
A2A5A6	M1152Δ*matAB*	M1152Δ*matAB*::pSET-A2A5::pTG-A6	Producer of **1**	This study
AKV_F	AKV1	M1152Δ*matAB*::pEN10001::pSV-dnrF	Producer of **8**	This study
AKV_CK	AKV1	M1152Δ*matAB*::pEN10001::pSV-eamCK	Producer of **14**	This study
AKV_CB	AKV1	M1152Δ*matAB*::pEN10001::pSV-rdmCB	Producer of **12**	This study
AKV_ECB	AKV1	M1152Δ*matAB*::pEN10001::pSV-rdmECB	Producer of **18**	This study
AKV_1516	AKV1	M1152Δ*matAB*::pEN10001::pSV-kstA1516	Producer of **20**	This study

Strain	Parental strain	Genotype and relevant characteristics	Note	Reference
AKV_15161011	AKV1	M1152Δ*matAB*::pEN10001::pSV-kstA15161011	Producer of **24**	This study
EAKV_F	EAKV2	M1152Δ*matAB*::pEAKV2::pSV-dnrF	Producer of **9**	This study
EAKV_CK	EAKV2	M1152Δ*matAB*::pEAKV2::pSV-eamCK	Producer of **15**	This study
EAKV_CB	EAKV2	M1152Δ*matAB*::pEAKV2::pSV-rdmCB	Production of **15**	This study
EAKV_ECB	EAKV2	M1152Δ*matAB*::pEAKV2::pSV-rdmECB	Production of **15**	This study
EAKV_1516	EAKV2	M1152Δ*matAB*::pEAKV2::pSV-kstA1516	Producer of **21**	This study
EAKV_15161011	EAKV2	M1152Δ*matAB*::pEAKV2::pSV-kstA15161011	Producer of **25**	This study
AURA_F	AURA2	M1152Δ*matAB*::pAURA2::pSV-dnrF	Producer of **10**	This study
AURA_CK	AURA2	M1152Δ*matAB*::pAURA2::pSV-eamCK	Producer of **16**	This study
AURA_CB	AURA2	M1152Δ*matAB*::pAURA2::pSV-rdmCB	Producer of **13**	This study
AURA_ECB	AURA2	M1152Δ*matAB*::pAURA2::pSV-rdmECB	Producer of **19**	This study
AURA_1516	AURA2	M1152Δ*matAB*::pAURA2::pSV-kstA1516	Producer of **22**	This study
AURA_15161011	AURA2	M1152Δ*matAB*::pAURA2::pSV-kstA15161011	Producer of **26**	This study
NOG_F	NOG2	M1152Δ*matAB*::pNOG2::pSV-dnrF	Producer of **11**	This study
NOG_CK	NOG2	M1152Δ*matAB*::pNOG2::pSV-eamCK	Producer of **17**	This study
NOG_CB	NOG2	M1152Δ*matAB*::pNOG2::pSV-rdmCB	Production of **17**	This study
NOG_ECB	NOG2	M1152Δ*matAB*::pNOG2::pSV-rdmECB	Production of **17**	This study
NOG_1516	NOG2	M1152Δ*matAB*::pNOG2::pSV-kstA1516	Producer of **23**	This study
NOG_15161011	NOG2	M1152Δ*matAB*::pNOG2::pSV-kstA15161011	Producer of **27**	This study
RW-1	AKV1	M1152Δ*matAB*::pEN10001::pR_DnrQS	**1**+R+GT(DnrQS)	This study
RW-2	AKV7	M1152Δ*matAB*::pAKV7::pR_DnrQS	**1**+R+GT(DnrQS)	This study
RW-3	AKV1	M1152Δ*matAB*::pEN10001::pR_AknST	**1**+R+GT(AknST)	This study
RW-4	AKV7	M1152Δ*matAB*::pAKV7::pR_AknST	**1**+R+GT(AknST)	This study
RW-5	AKV1	M1152Δ*matAB*::pEN10001::pR_AknSTK	**1**+R+GT(AknSTK)	This study
RW-6	AKV7	M1152Δ*matAB*::pAKV7::pR_AknSTK	**1**+R+GT(AknSTK)	This study
RW-7	AKV1	M1152Δ*matAB*::pEN10001::pR_SnogEN	**1**+R+GT(SnogEN)	This study
RW-8	EAKV2	M1152Δ*matAB*::pEAKV2::pR_DnrQS	**4**+R+GT(DnrQS)	This study
RW-9	EAKV2	M1152Δ*matAB*::pEAKV2::pR_AknST	**4**+R+GT(AknST)	This study
RW-10	EAKV2	M1152Δ*matAB*::pEAKV2::pR_SnogEN	**4**+R+GT(SnogEN)	This study
RW-11	AURA	M1152Δ*matAB*::pAURA2::pR_DnrQS	**3**+R+GT(DnrQS)	This study
RW-12	AURA	M1152Δ*matAB*::pAURA2::pR_AknST	**3**+R+GT(AknST)	This study
RW-13	pNOG	M1152Δ*matAB*::pRW10000::pR_SnogEN	**2**+R+GT(SnogEN)	This study
RW-14	pNOG	M1152Δ*matAB*::pRW10000::pR_SnogEND	**2**+R+GT(SnogEND)	This study
RW-15	pNOG	M1152Δ*matAB*::pRW10000::pR_AknST	**2**+R+GT(AknST)	This study
RW-16	pNOG	M1152Δ*matAB*::pRW10000::pR_DnrQS	**2**+R+GT(DnrQS)	This study
RW-17	RW-1	M1152Δ*matAB*::pEN10001::pR_DnrQS::pDAU	Producer of **29**	This study
RW-18	RW-2	M1152Δ*matAB*::pAKV7::pR_DnrQS::pDAU	Producer of **29**	This study
RW-19	RW-3	M1152Δ*matAB*::pEN10001::pR_AknST::pRHO	Producer of **6**	This study
RW-20	RW-5	M1152Δ*matAB*::pEN10001::pR_AknSTK::pRHO_DEFUC_RHODI	Producer of **36**	This study
RW-21	RW-6	M1152Δ*matAB*::pAKV7::pR_AknSTK::pRHO_DEFUC_RHODI	Producer of **36**	This study
RW-22	RW-7	M1152Δ*matAB*::pEN10001::pR_SnogEN::pDENOG	Producer of **46**	This study
RW-23	RW-8	M1152Δ*matAB*::pEAKV2::pR_DnrQS::pDAU	Producer of **38**	This study
RW-24	RW-8	M1152Δ*matAB*::pEAKV2::pR_DnrQS::pACO	Producer of **40**	This study
RW-25	RW-9	M1152Δ*matAB*::pEAKV2::pR_AknST::pRHO	Producer of **43**	This study

Strain	Parental strain	Genotype and relevant characteristics	Note	Reference
RW-26	RW-10	M1152Δ*matAB*::pEAKV2::pR_SnogEN::pDENOG	Producer of **45**	This study
RW-27	RW-11	M1152Δ*matAB*::pAURA2::pR_DnrQS::pDAU	Producer of **37**	This study
RW-28	RW-12	M1152Δ*matAB*::pAURA2::pR_AknST::pRHO	Producer of **42**	This study
RW-29	RW-13	M1152Δ*matAB*::pRW10000::pR_SnogEN::pDENOG	Producer of **7**	This study
RW-30	RW-14	M1152Δ*matAB*::pRW10000::pR_SnogEND::pRHO_DENOG	Producer of **28**	This study
RW-31	RW-15	M1152Δ*matAB*::pRW10000::pR_AknST::pRHO	Producer of **44**	This study
RW-32	RW-16	M1152Δ*matAB*::pRW10000::pR_DnrQS::pDAU	Producer of **39**	This study
RW-33	RW-16	M1152Δ*matAB*::pRW10000::pR_DnrQS::pACO	Producer of **41**	This study
RW-34	RW-17	M1152Δ*matAB*::pEN10001::pR_DnrQS::pDAU::pFP	Producer of **31**	This study
RW-35	RW-17	M1152Δ*matAB*::pEN10001::pR_DnrQS::pDAU::pFPK	Producer of **32**	This study
RW-36	RW-17	M1152Δ*matAB*::pEN10001::pR_DnrQS::pDAU::pFP_K	Producer of **32**	This study
RW-37	RW-17	M1152Δ*matAB*::pEN10001::pR_DnrQS::pDAU::pFP_K_VA	Producer of **35**	This study
RW-38	RW-17	M1152Δ*matAB*::pEN10001::pR_DnrQS::pDAU::pFP_K_VA_x4_r3	Producer of **35**	This study
RW-39	RW-18	M1152Δ*matAB*::pAKV7::pR_DnrQS::pDAU::pFP_K_VA	Producer of **34**	This study
RW-40	RW-18	M1152Δ*matAB*::pAKV7::pR_DnrQS::pDAU::pFP_K_VA_x4_r3	Producer of **34**, **35**	This study
RW-41	RW-17	M1152Δ*matAB*::pEN10001::pR_DnrQS::pDAU::pFP_K_VA_K	Producer of **35**	This study
RW-42	RW-17	M1152Δ*matAB*::pEN10001::pR_DnrQS::pDAU::pFP_K_VA_x4_r3_K	Producer of **35**	This study
RW-43	RW-18	M1152Δ*matAB*::pAKV7::pR_DnrQS::pDAU::pFP_K_VA_K	Producer of **34**	This study
RW-44	RW-18	M1152Δ*matAB*::pAKV7::pR_DnrQS::pDAU::pFP_K_VA_x4_r3_K	Producer of **35**	This study
RW-45	RW-19	M1152Δ*matAB*::pEN10001::pR_AknST::pRHO::pF	Producer of **50**	This study
RW-46	RW-19	M1152Δ*matAB*::pEN10001::pR_AknST::pRHO::pECB	Producer of **51**	This study
RW-47	RW-19	M1152Δ*matAB*::pEN10001::pR_AknST::pRHO pFP_K_VA_x4_r3	Production of **52**	This study
RW-48	RW-29	M1152Δ*matAB*::pRW10000::pR_SnogEND::pDENOG::pL2W_KNT_LM	Producer of **28**	This study
RW-49	RW-31	M1152Δ*matAB*::pRW10000::pR_AknST::pRHO::p15_16	Producer of **49**	This study
RW-50	RW-32	M1152Δ*matAB*::pRW10000::pR_DnrQS::pDAU::p15_16	Producer of **47**	This study
RW-51	RW-33	M1152Δ*matAB*::pRW10000::pR_DnrQS::pACO::p15_16	Producer of **48**	This study

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Painosalama, Turku, Finland 2025

ISBN 978-952-02-0268-2 (PRINT)
ISBN 978-952-02-0269-9 (PDF)
ISSN 0082-7002 (Print)
ISSN 2343-3175 (Online)