Anthracycline biosynthesis in Streptomyces: engineering, resistance and antimicrobial activity

Hulst, M.B.

Anthracyclines: biosynthesis, engineering and clinical applications

Mandy B. Hulst*, Thadee Grocholski*, Jacques J.C. Neefjes, Gilles P. van Wezel† and Mikko Metsä-Ketelä†

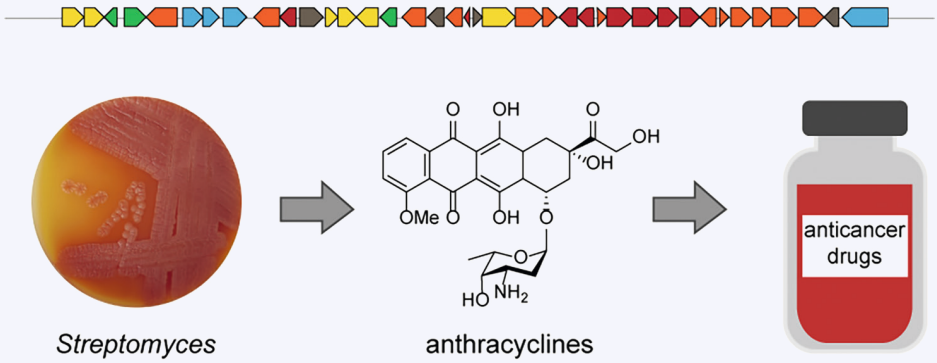
* These authors contributed equally.
† Corresponding authors.

Nat. Prod. Rep. 2022; **39**, 814–841.

Abstract

Covering: January 1995 to June 2021

Anthracyclines are glycosylated microbial natural products that harbor potent antiproliferative activities. Doxorubicin has been widely used as an anticancer agent in the clinic for several decades, but its use is restricted due to severe side effects such as cardiotoxicity. Recent studies into the mode-of-action of anthracyclines have revealed that effective cardiotoxicity-free anthracyclines can be generated by focusing on histone eviction activity, instead of canonical topoisomerase II poisoning leading to double-strand breaks in DNA. These developments have coincided with an increased understanding of the biosynthesis of anthracyclines, which has allowed generation of novel compound libraries by metabolic engineering and combinatorial biosynthesis. Coupled to the continued discovery of new congeners from rare Actinobacteria, a better understanding of the biology of *Streptomyces* and improved production methodologies, the stage is set for the development of novel anthracyclines that can finally surpass doxorubicin at the forefront of cancer chemotherapy.



Introduction

Natural products derived from the secondary metabolism of bacteria and fungi are an important source of antibiotics and other drugs. About 70% of the approved drugs are natural products or derived from a natural product[20]. The majority of the bioactive secondary metabolites are produced by Actinobacteria and especially by members of the genus *Streptomyces*[2,4,21,22]. There are many classes of secondary metabolites, and these compounds may have a range of bioactivities, including antibacterial, anticancer, antifungal, antiviral, anthelmintic, herbicidal, or immunosuppressive activity[3,20]. The compounds can be divided into different structural families, with polyketides, non-ribosomal peptides, ribosomally synthesized post-translationally modified peptides (RiPPs), terpenoids, and alkaloids as major examples.

Subject of this review are the anthracyclines, which are aromatic type II polyketides assembled by sequential condensation of acyl-CoA units. Aromatic polyketides are classified based on the carbon scaffold as anthracyclines, angucyclines, aureolic acids, tetracyclines, tetracenomycins, pentangular polyphenols, and benzoisochromanequinones[23]. The polyphenolic aglycone units are typically decorated by a number of tailoring reactions such as methylation, amination, oxidation and glycosylation to generate further structural diversity[24]. The best-known aromatic polyketides are the tetracyclines, which are active against a wide range of both Gram-positive and Gram-negative bacteria. Chlortetracycline or aureomycin was isolated from *Streptomyces aureofaciens* in 1947[25]. Tetracyclines are protein synthesis inhibitors that target the bacterial ribosome[26,27]. The first anthracyclines that were discovered were the rhodomycins (e.g. rhodomycin B, **1**) (Figure 1), which displayed strong antibiotic activity against *Staphylococcus aureus*[28]. However, the anthracyclines are particularly known for their potent anticancer activity[11,29]. Brockmann later defined anthracyclines as "yellow-red or red, optically active dyes" that consist of a linear tetracyclic 7,8,9,10-tetrahydro-5,12-naphtacenoquinone scaffold, which is decorated with one or more sugar moieties[30].

The best-known anthracyclines are daunorubicin (**2**) and doxorubicin (**3**) (Figure 1). In the early 1960s, **2** was isolated from *Streptomyces peucetius*[9,10]. The molecule demonstrated exceptionally strong activity against acute leukemia[12]. Random mutagenesis of the daunorubicin producing strain resulted in the even more potent derivative **3**[8]. Despite serious dose-limiting adverse effects, the two anthracyclines continue to be used in first-line chemotherapy for the treatment of solid and hematological tumors[31]. More anthracyclines were isolated from *Streptomyces* species in the following years, including nogalamycin (**4**) from *Streptomyces nogalater*[32], aclacinomycin A (**5**) from *Streptomyces galilaeus*[33], and steffimycin B (**6**) from *Streptomyces steffisburgensis*[34] (Figure 1). The structurally unique **4** demonstrated high activity against Gram-positive bacteria and several cancer lines, but severe toxicity prevented its progress in clinical trials[35]. The anthracycline **5** demonstrated potent antileukemia activity and low cardiotoxicity and is currently prescribed exclusively in Japan and China[13]. Many attempts, particularly in the 1980s, were made to expand the chemical space of existing anthracyclines by semi-synthesis in search for analogues with reduced cardiotoxicity[36]. Notably, these failed to yield clinically relevant derivatives with increased activity spectrum and reduced cardiotoxicity[13].

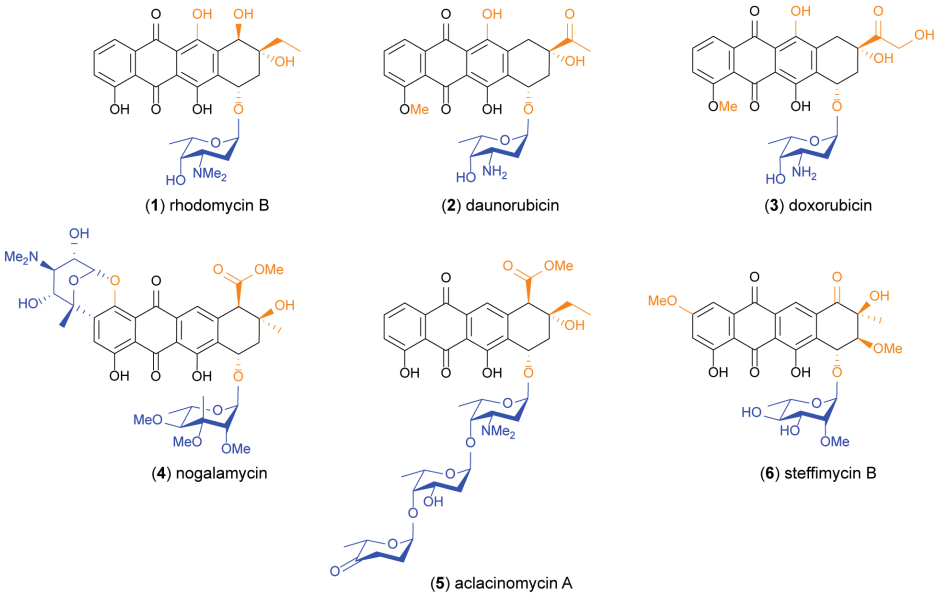


Figure 1. Chemical structures of classical anthracyclines. Structures of rhodomycin B (**1**), daunorubicin (**2**), doxorubicin (**3**), nogalamycin (**4**), aclacinomycin A (**5**) and steffimycin B (**6**), which are produced by *S. purpurascens*, *S. peucetius* (both **2** and **3**), *S. nogalater, S. galilaeus* and *S. steffisburgensis*, respectively. The common linear tetracyclic 7,8,9,10-tetrahydro-5,12-naphtacenoquinone scaffold is present in all structures. The chemical diversity of these compounds occurs either from their tailoring enzymes (orange) or their sugar moieties (blue).

Recent new insights into the mechanism of action and biosynthetic pathways of anthracyclines suggest that it is finally feasible to find "a better doxorubicin"[13]. The aim of this review is to provide a comprehensive overview of the current literature on biosynthesis, mode-of-action and medical application of anthracyclines, and provide pointers towards further exploration of the chemical space of these important tetracyclic polyketides. First, we will discuss the production of anthracyclines by Actinobacteria and their biological role. Second, we will review the application of anthracyclines as anticancer compounds. Third, recent advances in the understanding of the biosynthetic pathways of anthracyclines will be discussed. Finally, we evaluate approaches that may be employed to find new anthracyclines as well as strategies to optimize the production of anthracyclines by genetic engineering and optimization of the fermentation process.

Actinobacteria as anthracycline producers

Biology and secondary metabolism of Actinobacteria

Actinobacteria are a diverse phylum of Gram-positive bacteria with high G+C DNA content, which are widely distributed in both aquatic and terrestrial ecosystems[37]. The bacteria have a multicellular lifestyle[38]. The best characterized genus of Actinobacteria is *Streptomyces*. The

bacteria play a key ecological role in their environment due to their ability to break down a wide range of macromolecules and scavenge nutrients. Streptomycetes are mycelial organisms that reproduce by sporulation, with a lifecycle that is similar to that of filamentous fungi[38]. A single uninucleoid spore germinates and then grows out by a combination of hyphal tip extension and branching into a complex mycelial network of large multinucleoid cells[39,40]. During this process, enzymes are released extracellularly to break down natural polymers in the environment to provide nutrients for the vegetative mycelium[41]. When conditions become limiting, a complex regulatory network initiates the production of aerial hyphae. The nutrients required for this process are supplied by autolytical degradation of the old vegetative mycelium, also known as substrate mycelium[42,43]. This involves a process of programmed cell death (PCD), which is a hallmark of all multicellular bacteria[38]. Eventually, the reproductive aerial hyphae differentiate into chains of uninucleoid spores, which are dispersed into the environment in search of new nutrients[44].

This process of morphological differentiation is closely connected to chemical differentiation, namely the onset of secondary metabolism and antibiotic production, which correlates temporally with development[5,6,45]. The genes involved in the production of secondary metabolites are typically clustered on the genome in so-called biosynthetic gene clusters (BGCs). Whole genome sequencing revealed that streptomycetes generally contain a multitude of BGCs, and far more than originally anticipated[46]. However, many compounds originating from these BGCs have escaped screening, because the BGCs may be poorly expressed under laboratory conditions, or because the production levels of the metabolites are too low to detect[47]. The environmental cues that activate the expression of many BGCs in their natural environment are likely missing in the laboratory[48–51]. Such 'cryptic' BGCs potentially specify a rich variety of bioactive molecules and are thus attractive targets for new studies. Developments in bioinformatics, genetic and analytical tools may lead to the discovery of new natural products[48,49,52].

Most antibiotics are produced by filamentous microorganisms. A likely explanation is that antibiotics play a role in their defense to protect the biochemical building blocks that are released by the autolytic degradation of their mycelium in a nutrient-deprived environment, which will attract motile competitors[21,38,53]. Anthracyclines generally display antibiotic activity against Gram-positive bacteria, although synergistic effects together with the antibiotic rifampicin have also been observed against Gram-negative bacteria[54]. Moreover, as anthracyclines intercalate into DNA, they are effective antiproliferative agents against eukaryotic cell lines. Recently, the DNA-targeting activity of anticancer compounds was also shown to limit infection by DNA phages, which may have been an evolutionary driver for streptomycetes to produce anthracyclines[55]. Interestingly, anticancer compounds may play an important physiological role in the bacterial life cycle that is complementary to that of antibiotics, by actually driving the PCD process. The accumulation of DNA damaging compounds right at the onset of development may promote PCD by destroying a network of cells within the substrate mycelium, thus accelerating the provision of nutrients to feed the build-up of the aerial mycelium[56,57]. Besides a role in chemical warfare, secondary metabolites

may play a role in interspecies and intraspecies communication[58]. These diverse roles of secondary metabolites have provided a significant selective advantage for the producing organisms and have promoted the evolution of the tremendous chemodiversity of microbial natural products[59,60].

Self-resistance to anthracyclines

Anticancer agents that affect DNA structure and/or inhibit DNA functions are also cytotoxic to their producers[57]. Therefore, the production of these compounds is tightly regulated and usually coincides with expression of resistance genes[4,7]. Different self-resistance mechanisms against anticancer compounds include efflux, sequestration, modification, self-sacrifice, and metabolic dormancy[57].

Anthracyclines are usually exported by ABC-transporters in the producer strains of *Streptomyces*. This limits the toxicity for the producer strains, analogous as human or mouse tumor cells can protect themselves (through overexpression of ABC transporter MDR1 (ABCB1))[61]. As additional protection mechanism some anthracycline BGCs encode UvrA-like proteins that are involved in the nucleotide excision repair of DNA (Figure 2). The daunorubicin BGC includes four genes involved in self-resistance[62,63]. The first genes, *drrA* and *drrB*, encode an ABC transporter that supports the efflux of **2** and **3**[64,65]. Directly downstream of these genes lies *drrD*, which is involved in self-resistance as well. A *drrD* deletion mutant showed reduced self-resistance, but the mechanism has not yet been elucidated[66]. The gene product of *drrC* is homologous to UvrA-like proteins and is involved in the recognition and repair of DNA damage[67–69]. Recently, another mechanism of self-resistance was identified in *S. peucetius*. A yet unidentified secreted protein binds to extracellular **2**, thereby controlling extracellular drug concentrations and preventing uptake by the bacterium[70]. The resistance and biosynthetic genes are regulated by a cascade of pathway-specific regulators. The mechanism is similar to the regulation of the undecylprodigiosin cluster of the model organism *Streptomyces coelicolor*, which we will discuss first.

Transcriptional control of biosynthetic gene clusters

Much of our knowledge on antibiotic production and regulation originates from the model organism *S. coelicolor*[71]. Before the publication of its complete genome sequence, which was the first *Streptomyces* genome to be published[46], it was recognized that this strain produces the blue-pigmented actinorhodin (Act), the red-pigmented undecylprodigiosin (Red), the calcium-dependent antibiotic (CDA), and the plasmid-encoded methylenomycin (Mmy)[72,73]. The gene clusters of the genome-encoded antibiotics are controlled by the cluster-situated regulators (CSRs) ActII-ORF4, RedD and CdaR, respectively, which all belong to the family of *Streptomyces* antibiotic regulatory proteins (SARPs)[74]. The genes for these pivotal regulators are all subject to translational control by the BldA tRNA, which is required for the translation of the rare UUA codon for leucine. This tRNA is also required for the proper translation of many developmental genes and is therefore a key 'regulator' in the control of development and antibiotic production[75,76].

The production of secondary metabolites by *Streptomyces* is tightly regulated *via* a complex network of signaling cascades and feedback loops, which involves both translational and transcriptional regulation[5,6]. The most abundant pleotropic regulators in *Streptomyces* are two-component systems, which allow for a rapid response to altering environments. The pleotropic regulators can act directly by binding to cluster-situated regulators (CSRs) or indirectly by influencing other regulatory pathways. Subsequently, the CSR—or a hierarchical cascade of CSRs within the BGC – modulates the expression of the genes in the BGC[45].

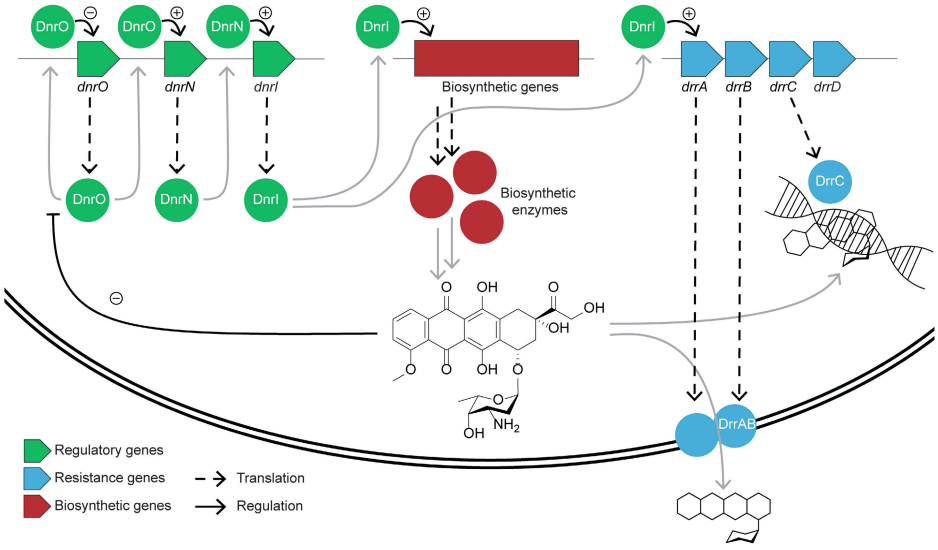


Figure 2. Regulation of the daunorubicin biosynthetic gene cluster and resistance mechanisms. The doxorubicin BGC is regulated by a cascade of three regulators (green). *DnrI* directly controls the expression of the biosynthetic (red) and resistance (blue) genes by binding to the promoter regions within the cluster. *DnrN* is also a DNA-binding protein that binds to the *dnrI* promoter to activate transcription. *DnrO* activates the transcription of *dnrN* and represses its own transcription. The glycosylated products of the biosynthetic pathway inhibit the binding of *dnrO* to the *dnrO* promoter, which leads to the activation of *dnrN* and hence *dnrI*. The cluster has four resistance genes (blue) that are activated by DnrI. The two first genes, *drrA* and *drrB*, together encode an ABC transporter that supports the efflux of daunorubicin (**2**) and doxorubicin (**3**). The gene product of *drrC* is homologous to UvrA-like proteins and is involved in the recognition and repair of DNA damage caused by the products of the pathway. The precise function of *drrD* has not been elucidated. The biosynthetic genes are visualized as a single box and include around 30 genes that together encode the minimal PKS for the tetracyclic skeleton, the post-PKS tailoring enzymes, and those for glycosylation.

The regulation of the BGC for Red is one of the best studied examples and may partly serve as an example for the regulation of the daunorubicin BGC. The production of Red is under positive control of the CSR RedD[77,78]. The transcription of *redD* is in turn regulated by RedZ, which is a pseudo-response regulatory protein[79]. The regulation of the daunorubicin BGC (Figure 2) is similar to that of the Red BGC, whereby *dnrI* and *dnrN* are homologues of *redD* and *redZ*, respectively. DnrI directly controls the transcription of the biosynthetic and resistance genes by binding to the promoter regions within the cluster[80]. DnrN is also a DNA-binding protein that binds to the

dnrI promoter to activate transcription[81]. However, regulation of the daunorubicin BGC is a three-tier transcriptional activator system, and also involves the TetR-family transcriptional regulator *dnrO*, which activates the transcription of *dnrN*[82]. *DnrO* also represses its own transcription.

Interestingly, anticancer compounds may play a direct role in the transcriptional control of their BGCs by virtue of their DNA-intercalating activities. Jadomycin is an angucycline-type polyketide with anticancer activity and is produced by *Streptomyces venezuelae* ISP5230. The jadomycin BGC encodes five regulators (JadR1-R2 and JadW1-W3)[83,84]. JadR1 is the pathway-specific activator of jadomycin biosynthesis and indispensable for its production[85]. The DNA binding activity of the atypical response regulators JadR1 (OmpR family) and RedZ (NadR family) is controlled by the intracellular concentration of jadomycin and Red, respectively. The compounds bind to these transcription factors to inhibit transcription[86]. Similarly, in *S. peucetius*, drug-mediated control is mediated *via* inhibition of the DNA binding capacity of *dnrO* by **2**[82]. In fact, all of the glycosylated products in the pathway inhibit the DNA–*dnrO* binding and *dnrO* expression is increased, which leads to the activation of *dnrN* and hence *dnrI* (Figure 2). Thus, the transcriptional activation of the enzymes required for antibiotic production go hand-in-hand with the activation of the resistance machinery.

Understanding of the regulatory networks that control the anthracycline production, and of the mechanisms that provide self-resistance, is of critical importance for the development of successful metabolic engineering and strain engineering approaches. Understanding of regulation will allow for increased production levels and understanding of resistance mechanisms will facilitate engineering strategies to reduce toxicity issues.

Anthracyclines as anticancer drugs

Many anthracyclines, including **2**, were originally isolated as potential new antibiotics[10]. However, the observation of the antiproliferative activity *in vitro* and in mice initiated the development of **2** as an anticancer drug. The success of the molecule in the clinic led to the generation of hundreds of anthracycline analogues by modified bacteria and half-chemistry[15]. Currently, six different semi-synthetic variants of **2** (Figure 1) are used in the clinic in cancer treatment: **3** (Figure 1), epirubicin (**7**), idarubicin (**8**), pirarubicin (**9**) and valrubicin (**10**) (Figure 3, Table 1)[87]. In addition, the totally synthetic amrubicin (**11**), which contains a minimalistic version of the daunosamine sugar, is used in lung cancer chemotherapy in Japan (Figure 3, Table 1)[88]. These anthracycline drugs are a cornerstone in oncology treatments and more than one million cancer patients are treated with these drugs annually.

While these drugs are effective in the treatment of a series of solid and blood cancer types, their application is limited by life-threatening toxicity, especially cardiotoxicity. Cardiotoxicity resulting in heart failure may be acute, but in general is increasing with accumulated dose. For these reasons, anthracycline application is usually limited to six courses, because further

accumulation of the drugs is considered an unacceptable risk of heart failure. Juvenile cancer patients may get higher doses (because 'their heart can handle it'), while senior patients (usually over 70–75 years) and people with a poor heart function are excluded from anthracycline-based therapy[15]. Since cancer is often a disease of elderly people, a major patient group is excluded from this often effective treatment or simply treated at a non-effective low dose. Many young cancer survivors suffer heart problems later in life due to the cardiotoxic effects of anthracyclines. Interestingly, the triglycosylated **5**, which is the only anthracycline used in the clinic that is not based on **2**, is not significantly cardiotoxic in animal models, but is mainly used in the treatment of hematological malignancies such as acute myeloid leukemia (AML)[89–91]. **11** is also a variant with limited cardiotoxicity and is used in Japan for the treatment of lung cancer. Most tumor types are treated with anthracycline variants that show dose-dependent cardiotoxicity that limits treatment. In fact, cancer treatment is limited by the risk of cardiotoxicity rather than by the anticancer activity.

Table 1. Activity of classical and semi-synthetic anthracyclines. The application of anthracyclines as anticancer drugs is limited by their severe side effects, including cardiotoxicity resulting in heart failure. Anthracyclines may have two activities: poisoning of topoisomerase II resulting in DNA breaks, and the eviction of histones from the DNA resulting in the loss of epigenetic information. Not all anthracyclines have been tested for their activities and side effects. Notably, cardiotoxicity does not occur for anthracyclines that exhibit either DNA breaks or histone eviction activity (**5**, **11** and **12**).

Name	Cytotoxicity	Cardiotoxicity	DNA breaks	Histone eviction
Rhodomycin B (**1**)	+	?	?	?
Daunorubicin (**2**)	+	+	+	+
Doxorubicin (**3**)	+	+	+	+
Nogalamycin (**4**)	+	+	+*	?
Aclacinomycin A (**5**)	+	-	-	+
Steffimycin B (**6**)	+	?	?	?
Epirubicin (**7**)	+	+	+	+
Idarubicin (**8**)	+	+	+	+
Pirarubicin (**9**)	+	+	+	?
Valrubicin (**10**)	+	+	+	?
Amrubicin (**11**)	+	-	+	-
N,*N*-dimethyldoxorubicin (**12**)	+	-	-	+

* Topoisomerase I poison.

An antibiotic turned into an anticancer drug

Modern anticancer drugs are designed to inhibit proteins typically involved in the epigenetic code, DNA repair, apoptosis, or epigenetics. Anthracyclines were identified in screens for compounds that inhibit cell growth under tissue culture conditions. They would not easily have been identified in more target-directed screens, as is the standard in these times. Anthracyclines consist of a tetracyclic aglycone associated to an amino sugar. The tetracyclic moiety ensures intercalation into the DNA double helix, while the amino sugar positions into

the DNA minor groove (Figure 4A)[92]. This results in the poisoning of the enzyme topo-isomerase II (Topo II) and traps this enzyme on the DNA[93]. The function of Topo II is to alter the topology of DNA and relax supercoiled DNA by cutting the DNA double helix and passing it through the enzyme for re-ligation[94]. The molecular interaction of anthracyclines with Topo II is unclear, but another intercalating agent etoposide has been shown to stabilize the cleavage complex of Topo II, which prevents re-ligation and leads to DNA double-strand breaks (Figure 4B)[95]. The treatment of patients with anthracyclines results in cells containing many DNA breaks[15]. DNA breaks—if not swiftly repaired—will activate ATM1 and TP53 to initiate apoptosis. The cellular response to DNA breaks can easily be detected by probing Western blots for histone H2AX phosphorylated by ATM1 (also called γH2AX). DNA breaks can be directly detected by constant field electrophoresis[16]. Since fast-growing cells (like cancer cells) are more sensitive to the effects of DNA breaks, they die more efficiently thus providing a therapeutic window for Topo II poisons. However, other fast-growing cells such as hair cells and immune cells are also affected, which explains some of the direct reversible side effects of these drugs.

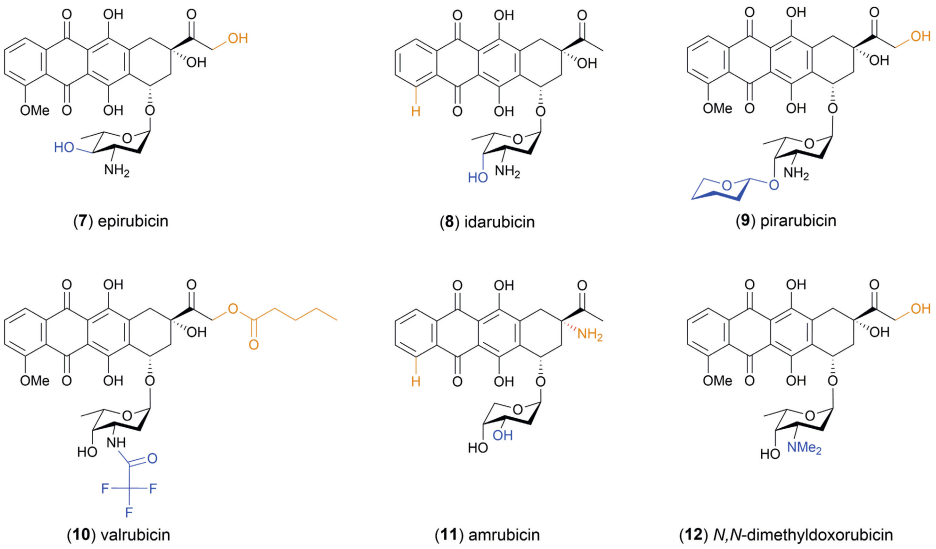


Figure 3. Chemical structures of semi-synthetic anthracyclines as anticancer drugs. Semi-synthetic anthracyclines epirubicin (**7**), idarubicin (**8**), pirarubicin (**9**) and valrubicin (**10**) are in clinical use. The fully synthetic amrubicin (**11**) is also used in anticancer chemotherapy. *N*,*N*-dimethyldoxorubicin (**12**) is a promising drug under development. The chemical differences in the aglycone (orange) or the sugar moiety (blue) are highlighted.

The mechanism of action of **5** is, however, an interesting exception. This drug does not cause any DNA breaks because it traps and poisons Topo II at an earlier step in the DNA break-and-repair cycle before the DNA break is initiated. Yet, this drug is highly effective in the treatment of AML tumors[14,96]. Further studies revealed a new activity of anthracycline drugs: the removal of nucleosomes from defined areas in the genome as detected by life imaging of photoactivate PAGFP-histones, magnetic tweezers on chromatin and FAIR-Seq (Figure 4C)[97]. Histones are the

major proteins of chromatin and facilitate packaging of DNA within cells[14,16,97]. Post-translational modifications of histone tails are important determinants for compactization of chromatin and regulation of gene expression, and mark the epigenetic code. Since histones released from DNA are degraded and replaced by nascent histones, the anthracyclines are effectively altering the epigenetic code and could thus be considered epigenetic drugs as well with specificity for the regions where histone eviction occurs[15].

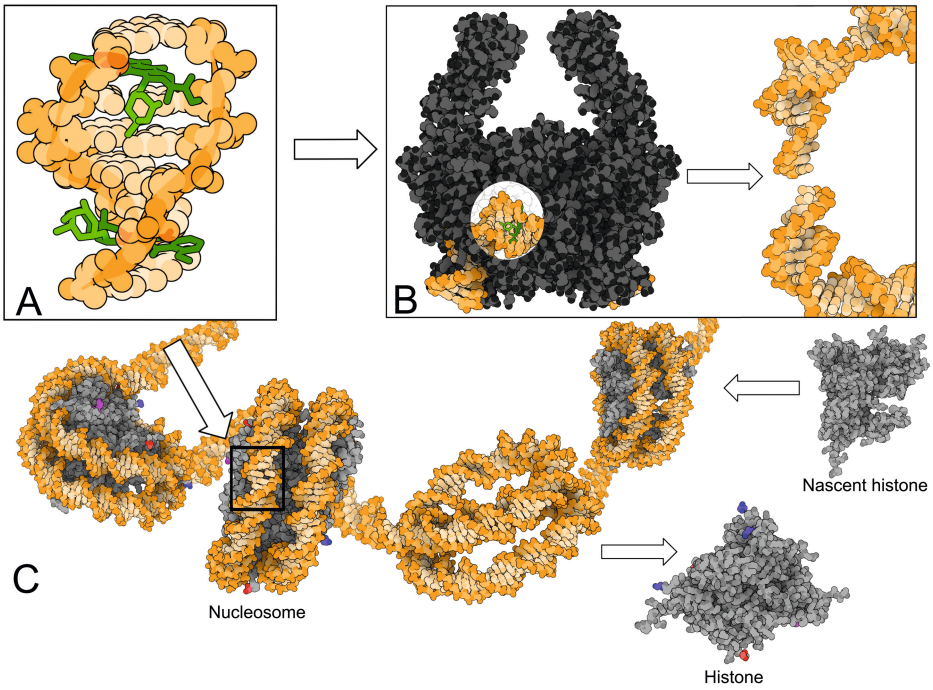


Figure 4. Mechanism of action of anthracyclines. (A) The doxorubicin (**3**) aglycone (dark green) intercalates into the DNA double helix (orange), while the amino sugar (light green) is positioned in the DNA minor groove. **(B)** The anthracycline–DNA complex poisons topoisomerase II (black) so that the enzyme is trapped on the DNA. In case of **3** (green), the poisoning leads to DNA double-strand breaks, whereas other anthracyclines such as **5** and **12** do not induce DNA breaks. **(C)** Anthracyclines act as epigenetic drugs *via* release of histones from the nucleosome. Histones (grey) are commonly modified by methylation (red), acetylation (purple), ubiquitination (blue) or phosphorylation (not shown). Histone eviction alters the accessibility of DNA and influences gene transcription. The introduction of nascent histones (grey) leads to loss of epigenetic information encoded in the histone tails.

The question then is which of these activities is more important for the anticancer effects of anthracyclines. Structure/function relationships revealed that **11** (where the amino group is located on the tetracyclic aglycone instead of the sugar moiety) showed only DNA break activity, while semi-synthetic *N*,*N*-dimethyldoxorubicin (**12**) (Figure 3, Table 1) solely displayed nucleosome eviction activity[16]. Recent studies have shown that **12** has high cytotoxic activity similar to **3**, which is in contrast to **11** that has only mild cytotoxicity. This suggests that histone eviction could be the most relevant bioactivity in anthracycline chemotherapy[15,16,89–91].

While anthracycline drugs have been used in the clinic for almost 50 years for the treatment of cancer patients, the molecular basis of the anticancer effects is only partially understood. For example, why nucleosome eviction causes cytotoxicity is unclear, as is the reason for the relative tumor selectivity of these drugs.

The side effects

The major and treatment-limiting side effect of anthracyclines is accumulated cardiotoxicity. In other words, the treatment is not halted because of the failure to eliminate cancer cells, but due to the side effects. In the 1980s, scientists of the National Cancer Institute performed an extensive synthesis of analogues of **3** with the aim to identify compounds lacking the cardiotoxic side effects[98]. And failed. Later studies revealed that **7** and other newer anthracyclines are somewhat less cardiotoxic. This variant was tested in mice, rats, and rabbits, but was not further developed for unknown reasons. Out of all of the doxorubicin derivatives in clinical use, only **11** can be considered significantly less cardiotoxic[99], but the compound has minor effects on tumors and is difficult to obtain. For these reasons, **11** is hardly used in cancer therapy.

Both **5** and **12** have a *N*,*N*-dimethyl group on the sugar moiety and both fail to induce DNA breaks, but still promote nucleosome eviction. Both compounds trap Topo II on DNA but poison the enzyme before the DNA breaks are made. The reduced (cardio)toxic effects of cancer treatment with **5** have been observed by oncologists, but never systematically studied. Since the prospects of AML patients are very poor, cardiotoxicity is possibly a minor issue in the treatment of these patients and therefore largely ignored. When tested in mice as well as human cardiomyocyte microcultures, **5** and **12** did not show any cardiotoxicity[16,18]. However, as mentioned above, **11**—the anthracycline drug that only induces DNA breaks—also lacks cardiotoxicity. This suggests that cardiotoxicity of **2**, **3**, **7** and **8** is caused by the combination of the nucleosome eviction and DNA breaks, whereas drugs with one of the two activities do not induce cardiotoxicity[16,18]. Consequently, chemical separation of these activities may lead to drug variants lacking cardiotoxic activity.

Anthracyclines have additional activities that have been considered to be the source of cardiotoxicity. Especially the formation of reactive oxygen species (ROS) has gained wide attention[100]. Anthracyclines may poison Topo IIβ in mitochondria, which affects transcription and mitochondria replication and results in the induction of ROS. Since heart muscle cells have many mitochondria for the production of ATP, cardiomyocytes may be particularly hit by anthracyclines. For these reason, ROS inhibitors have been generated that, however, have only limited effects on the cardiotoxicity of anthracyclines[101]. Besides that, **12** is equally efficient in producing radicals as **3**, but is not cardiotoxic. This suggests that radical formation is not directly related to cardiotoxicity.

Anthracyclines are not only toxic to the heart. As can be expected, they eliminate not only fast-growing tumor cells, but also immune cells, often resulting in reversible neutropenia. This implies that patients are initially more susceptible for infections after treatment with these drugs.

These drugs also affect spermatogenesis (albeit this was not observed for **5** in mice)[102]. Sperm cells from juvenile cancer patients are often isolated and stored for future family planning.

All anthracyclines used in the clinic except **5** induce DNA breaks. Consequently, errors can be introduced in the subsequent DNA repair process. A critical step in DNA repair is the phosphorylation of histone H2AX by ATM kinases. Because histone H2AX is also removed by anthracyclines, DNA repair is strongly attenuated when compared to Topo II poisons that do not evict histones (such as etoposide), which may further increase the error rate of DNA repair[14]. In almost one percent of patients treated with **2** or **3**, so-called treatment induced tumors emerge, usually an Acute Myeloid Leukemia (AML) or Acute Lymphocytic Leukemia (ALL) within two years after completion of the anticancer treatment[103]. Anthracycline variants lacking DNA break activity will not induce these 'secondary tumors', as shown for **5** in TP53 heterozygous mice[16]. Anthracycline variants *N*,*N*-dimethylated amino sugar all showed a loss in DNA break inducing activity, while retaining histone eviction and cytotoxicity[18,19]. Rhodomycin B could also be a member of this subgroup of anthracyclines, acting mainly *via* histone eviction.

The anthracyclines used in the clinic are effective as anticancer drugs for a variety of tumor types. The drugs also have major side effects of which cardiotoxicity is most acute and treatment limiting. Controlling the side effects of **5** and **12** would strongly improve the application of this cornerstone in cancer treatment.

Biosynthesis

The biosynthesis of anthracyclines in Actinobacteria can be divided into two stages, where conserved biosynthetic pathways are utilized for formation of the anthracyclinone carbon scaffold, followed by generation of structural diversity in later stages in so-called tailoring steps[104]. Particularly significant are glycosylation steps, where anthracyclinones are decorated with diverse carbohydrate moieties originating from D-glucose-6-phosphate[105]. Other important tailoring steps include methylations and various redox reactions, which further modify the metabolites. Anthracyclines are thus formed by complex multi-enzyme pathways, which explain that they cannot easily be made by full chemistry.

Biosynthesis of anthracycline aglycones *via* canonical polyketide pathways

The formation of anthracyclines has been intensely investigated due to their medical importance and most biosynthetic steps have been elucidated at the molecular level. Anthracyclines belong to type II polyketides and thus share common ancestry for carbon chain formation with bacterial fatty acid synthase machineries[106]. The starter unit may be either acetyl-CoA or propionyl-CoA, where the latter is generated by dedicated cluster-situated acyltransferases and propionyl-CoA synthases[107–109]. Nine rounds of iterative Claisen condensations with malonyl-CoA by the so-called minimal polyketide synthase (minPKS) leads to formation of the initial anthracycline decaketide (Figure 5A)[110–112].

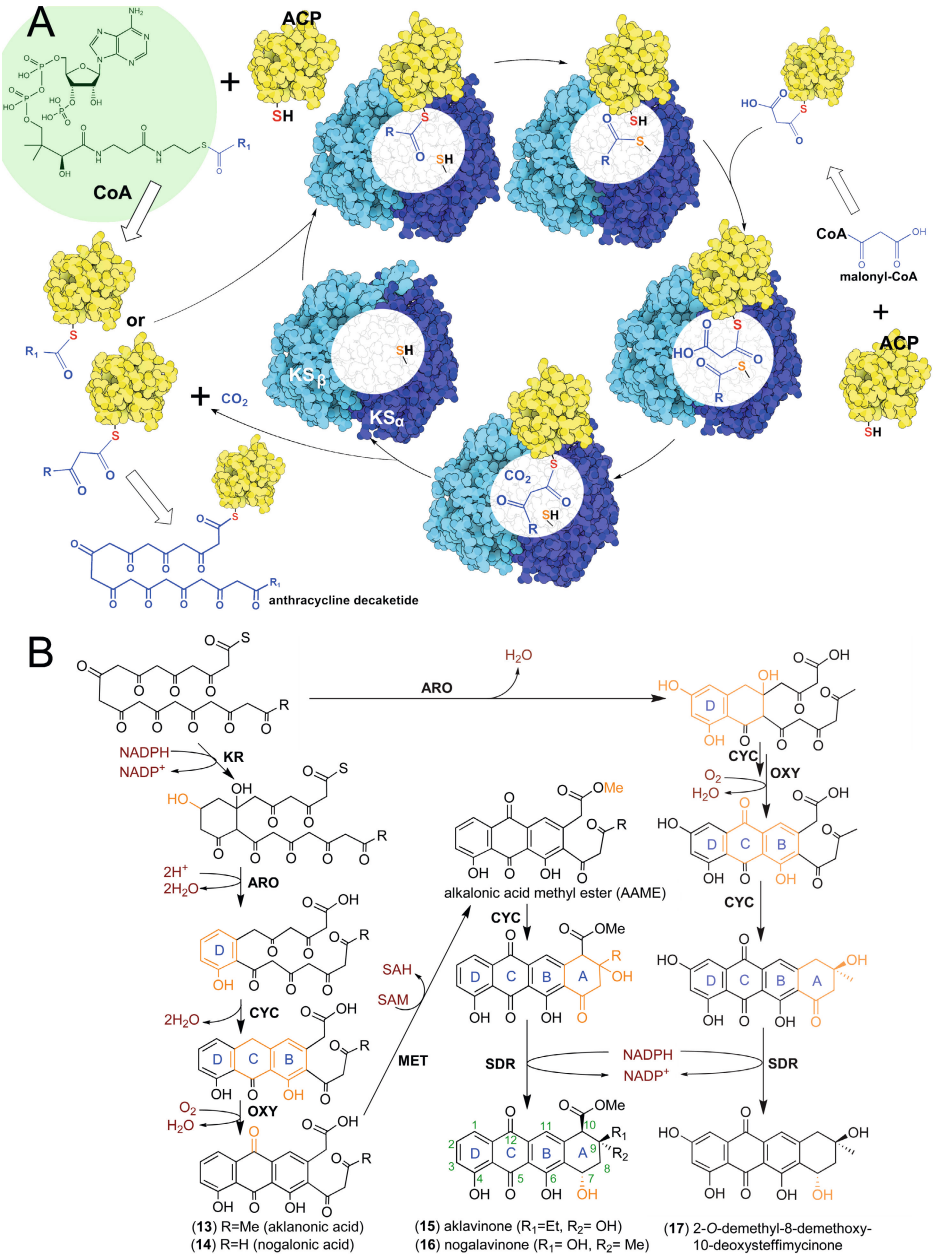


Figure 5. Biosynthesis of anthracycline aglycones. (A) The biosynthesis of type II polyketides, including anthracyclines, involves an iterative cycle of reactions involving three proteins that form the minimal polyketide synthase (minPKS). The acyl carrier protein (ACP; in yellow) contains a phosphopantetheine prosthetic group derived from CoA onto which the polyketide is assembled. The ketosynthase (KS) consists of two subunits, namely the catalytic KS_{α} (dark blue) and chain length determinant KS_{β} (light blue). Polyketide biosynthesis starts with coupling of an acetyl or propionyl (R_1) unit to the ACP. During the catalytic cycle, the growing polyketide is transferred from the ACP to an active site cysteine of KS_{α}.

After initial transfer to the KS_α, the ACP also supplies the ketosynthase with malonyl units. Incorporation of each malonyl unit in a Claisen condensation reaction catalyzed by KS_β extends the polyketide by two carbons. The polyketide is simultaneously transferred back to the ACP. Nine extension rounds are required for formation of the anthracycline decaketide structure. **(B)** After formation of the anthracycline decaketide, the polyketide undergoes a series of subsequent reactions catalyzed by ketoreductases (KR), aromatases (ARO), cyclases (CYC), oxygenases (OXY), methyltransferases (MET) as well as short-chain dehydrogenase/reductases (SDR), leading to the stable aglycone intermediates (**15**, **16**, **17**). Depending on the cyclase that performs closure of the A-ring, the stereochemistry of C*9* may differ, such as *9R** for (**15**) and *9S** for (**16**). Steffimycin biosynthesis branches off at an earlier stage and includes unique features such as an unusual 2-hydroxyl group due to lack of a KR enzyme and decarboxylative cyclization of the A-ring. The relevant positions of reaction products in each step are illustrated in orange.

The minPKS is composed of a heterodimeric ketosynthase (KS_α/KS_β) responsible for the chain elongation reactions and an acyl carrier protein (ACP), where the highly reactive poly-β-keto intermediate is covalently tethered[113,114].

Experiments on the daunorubicin and nogalamycin pathways have clarified how the first stable intermediates, aklanonic acid (**13**) and nogalonic acid (**14**), respectively, are folded in a controlled manner (Figure 5B)[111,115]. The process is initiated by 9-ketoreduction (KR) and aromatization (ARO) of the D-ring, which is followed by cyclization of C- and B-rings by a dedicated cyclase (CYC)[116,117]. The precise timing for release of the polyketide from the carrier protein is unknown, but the quinone forming 12-oxygenases (OXY) are able to utilize substrates that are not bound to ACP[118–120]. Next, the terminal carboxyl group is *O*-methylated by *S*-adenosyl-L-methionine (SAM)-dependent methyltransferases (MET), an essential step for cyclization of the A-ring (Figure 5B)[121]. Closure of the fourth ring occurs *via* Aldol condensation and these cyclases (CYC) determine the *9R* or *9S* stereochemistry of anthracyclines[115,122,123]. The final step in the biosynthesis of the key intermediates aklavinone (**15**) and nogalavinone (**16**) is ketoreduction by short-chain dehydrogenase/reductases (SDR) that retain exclusively the *7S* stereochemistry (Figure 5B)[117,124].

Deviation from this classical biosynthetic model transpires on the steffimycin-subclass of anthracyclines, where the lack of a 9-ketoreductase on the pathway leads to the archetypical 2-hydroxyl group[125]. Additionally, the fourth ring cyclase StfX, which is unrelated to SnoaL and AknH, acts on an ACP bound thioester and catalyzes decarboxylative cyclization to form 2-*O*-demethyl-8-demethoxy-10-deoxysteffimycinone (**17**) (Figure 5B)[126].

Biosynthesis of diverse carbohydrates

Anthracyclines contain a broad selection of 6-deoxysugar units that are typically associated with microbial natural products[127,128]. These include many 2,6-dideoxy- and 2,3,6-trideoxysugars[105], but the two most common substrates used in glycosylation of the 7-position are the amino sugars TDP-L-daunosamine (**18**)[129] and TDP-L-rhodosamine (**19**)[130], which have been enzymatically synthesized using proteins from the nogalamycin pathway (Figure 6)[131]. Rare carbohydrates attached to anthracyclines include brasiliose[132], L-decilonitrose[133], a γ-branched octose[134], and the 4-ketosugars cinerulose and L-aculose[130].

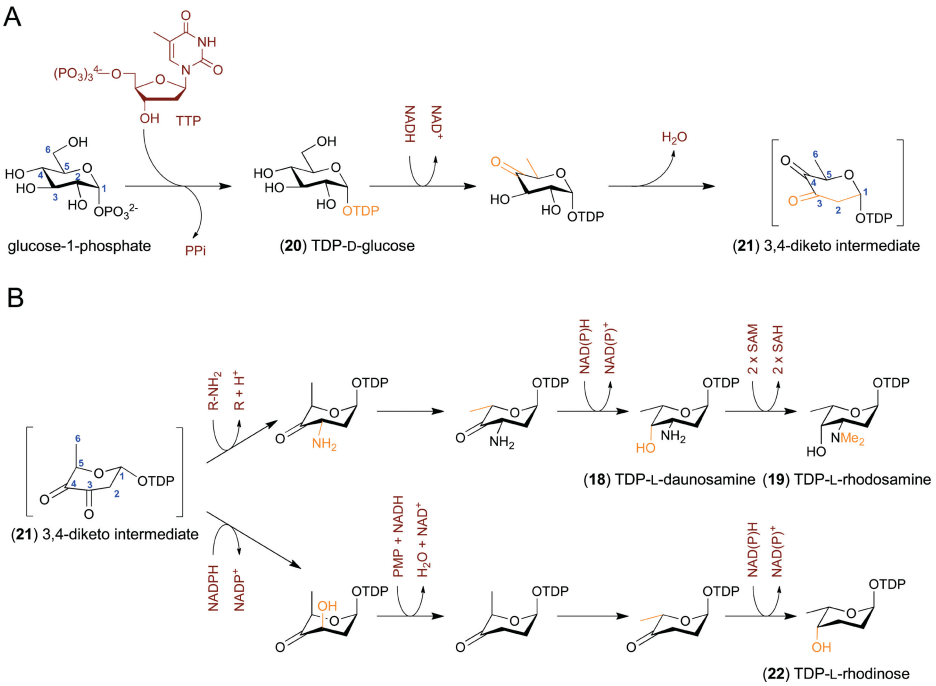


Figure 6. Biosynthesis of diverse carbohydrates. (A) The biosynthesis of 6-deoxysugars starts with the attachment of thymidine monophosphate to glucose-1-phosphate resulting in TDP-D-glucose (**20**). This is followed by two consecutive dehydration reactions by SDR enzymes resulting the formation of a common unstable 3,4-diketo intermediate (**21**). **(B)** The biosynthesis of amino and deoxysugars branch off from the intermediate **21**. On pathways towards amino sugars (**18**, **19**), the pyridoxal-5'-phosphate dependent enzymes (PLP) catalyze a 3-transamination reaction followed by a 5-epimerization reaction. Subsequently, **18** is formed through a NADPH stereospecific 4-ketoreductase. A SAM-dependent *N*-dimethylation of **18** leads to **19**. In contrast, the biosynthesis of 2,3,6-trideoxysugars such as **22** proceeds *via* 3-ketoreduction and removal of the 3-hydroxyl group. Subsequent 5-epimerization and 4-ketoreduction finalize formation of **22**.

In general, the early biosynthetic steps for all 6-deoxysugars are conserved and also found in primary metabolic pathways such as cell-wall synthesis[127]. The first step is the attachment of thymidine monophosphate to glucose-1-phosphate to generate TDP-D-glucose (**20**) with concomitant loss of pyrophosphate by nucleotidyltransferases (Figure 6)[135,136]. Two consecutive dehydration reactions, 4,6-dehydration by SDR enzymes[137], and 2,3-dehydration by Nudix hydrolase superfamily proteins[138], lead to formation of an unstable 3,4-diketo intermediate (**21**). Activation of the TDP-carbohydrate in such a manner allows diverse modification reactions to occur and leads to branching of the biosynthetic pathways towards their end products (Figure 6)[128].

On pathways leading to amino sugars, the next step is 3-transamination by pyridoxal-5'-phosphate (PLP) dependent enzymes using amino acid nitrogen group donors (Figure 6)[139]. The importance of the 4-keto group is further highlighted in the 5-epimerization reaction[140], which is facilitated by an enolate anion intermediate at C4[104]. Formation of **18** is completed

through stereospecific 4-ketoreduction using NADPH as the reducing agent, while biosynthesis of **19** continues through SAM-dependent *N*-dimethylation (Figure 6)[131].

In turn, TDP-L-rhodinose (**22**) can be considered an example of a typical anthracycline 2,3,6-trideoxysugar (Figure 6). The biosynthesis continues from the common 3,4-diketo intermediate through 3-ketoreduction. Subsequent removal of the 3-hydroxyl groups is mechanistically challenging due to the proximity of the 4-keto group[141] and requires enzymes containing an [2S–2Fe] iron sulfur cluster, a pyridoxamine-5'-phosphate cofactor and external electron donors such as ferredoxin/ferredoxin reductase[142]. Similarly to amino sugar biosynthesis, the final steps include 5-epimerization and 4-ketoreduction[104].

Glycosylation of anthracyclines

The carbohydrate units are transferred to aglycones by glycosyl transferases (GTs) that typically belong to the B family (GT-B)[143]. The structure of the L-rhodosamine transferase SnogD[144] responsible for 1-*O*-glycosylation on the nogalamycin pathway has demonstrated how the two Rossman fold domains, one for the aglycone acceptor and one for the sugar donor, provide a framework for transfer of diverse carbohydrates to a variety of positions in anthracyclinones[141,143].

In all known instances, the glycosylation at either the 7-position[131,145–147] or the 10-position[145,148] requires an additional P450-enzyme to increase the activity of the *bona fide* GT. The likely role of this P450-enzyme is to act as an allosteric activator[147], a proposal that has been supported by the crystallization of the α2β2 heterodimeric complex from the erythromycin pathway[149]. Strong non-covalent interactions between the partners have been proposed to induce conformational shifts both in the N-terminal aglycone acceptor and C-terminal sugar donor domains of the GTs to aid catalysis[149].

Other tailoring reactions

Modification of the anthracyclinone core in late stage tailoring steps is critical for the chemical diversity of these natural products. The doxorubicin pathway harbors the flavin adenine dinucleotide (FAD) dependent mono-oxygenase *dnrF* which catalyzes a 11-hydroxylation reaction (Figure 7)[150]. The biosynthesis continues through 15-demethylation by the *dnrP* esterase[151,152] and classical 4-*O*-methylation by *dnrK*[121] to generate 13-deoxydaunorubicin (**23**). The heme-containing cytochrome P450-protein DoxA drives the three consecutive 13- and 14-hydroxylation reactions that define the doxorubicin-subclass of anthracyclines (Figure 7)[153]. The daunomycin BGC also encodes additional enzymes that are responsible for formation of baumycin acetals (**24**), which are formed *via* oxidative carbon–carbon bond cleavage of a second nitrososugar (Figure 7)[154].

Rhodomycin biosynthesis proceeds similarly to daunorubicin with homologous proteins for 11-hydroxylation by RdmE[155,156] and 15-demethylation by RdmC[157,158]. The subsequent 10-hydroxylation step is catalyzed by SAM-dependent methyltransferase-like RdmB[158,159], which is interestingly homologous to *dnrK*, to complete formation of **1** (Figure 7).

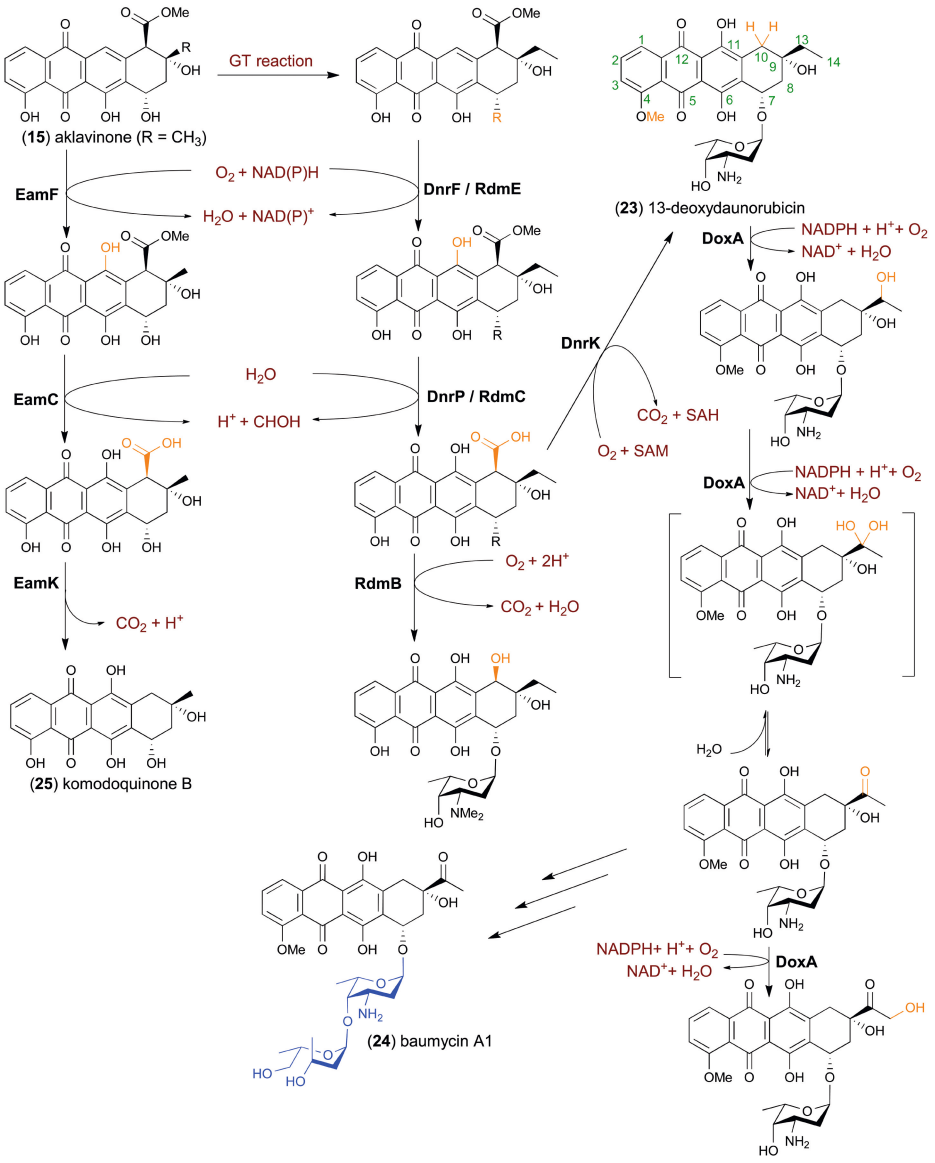


Figure 7. Chemical diversification of rhodomycin, doxorubicin and komodoquinone by tailoring enzymes. Various enzymes introduce chemical diversity (in orange) in the anthracycline backbone in the biosynthesis of **1**, **3** and **25**. All of these pathways share FAD dependent C-11 hydroxylases (DnrF/RdmE/EamF) and C-15 methylesterases (DnrP/RdmC/EamC). The next enzymes on the pathways are SAM-dependent methyltransferase-like proteins (DnrK/RdmB/EamK), which catalyze distinct chemistry. DnrK catalyzes C4-*O*-methylation with C-10 decarboxylated moonlighting activity, RdmB catalyzes C-10 hydroxylation, and EamK catalyzes solely C-10 decarboxylation. The biosynthesis of **3** continues through the action of the heme-dependent cytochrome P450 enzyme DoxA, which catalyzes three successive hydroxylation reactions. The daunomycin BGC also encodes additional enzymes that are responsible for formation of higher glycosides such as baumycin A1 (**24**).

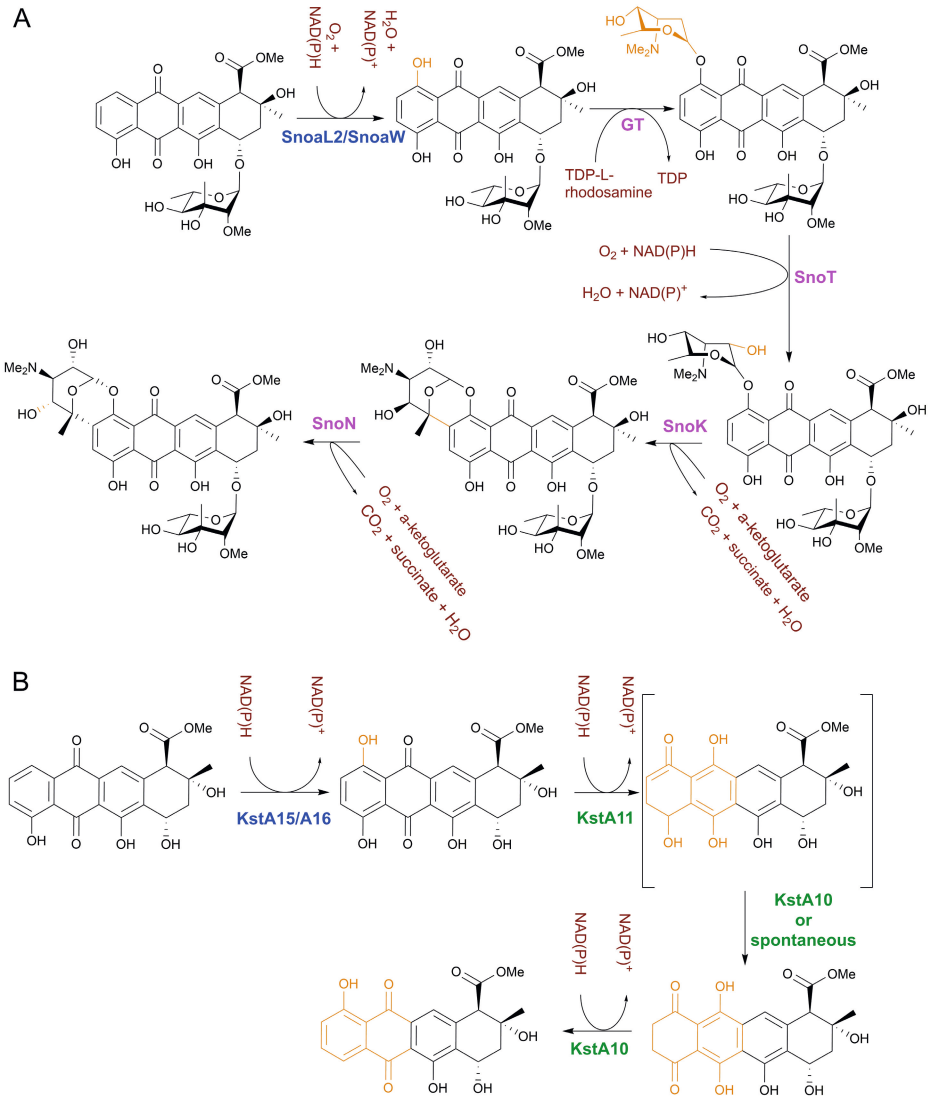


Figure 8. Parallels between tailoring steps in nogalamycin and kosinostatin biosynthesis. (A) The biosynthesis of **4** involves several atypical proteins. Formation of the archetypical epoxyoxocin ring system is initiated by the two-component 1-hydroxylase system SnoaW and SnoaL2. Attachment of the L-rhodosamine sugar is catalyzed by the SnogD glycosyltransferase. The reaction product is 2"-hydroxylated by the Rieske enzyme SnoT. The dual linkage of the amino sugar is completed by SnoK, a non-heme iron and α-ketoglutarate dependent enzyme. The final 4"-epimerization is catalyzed by SnoN. **(B)** The pathway towards **26** also includes a two component 1-hydroxylase system KstA15 and KstA16. In order to remove the 4-hydroxy group, KstA11 catalyzes an asymmetric dearomatization that allows 4-ketoreduction and 3,4-dehydration by KstA10.

Insertion of a single amino acid to *dnrK* suffices to convert 4-*O*-methylation activity to 10-hydroxylation[160], which demonstrates how small changes in biosynthetic enzymes allow the generation of structural diversity in microbial natural products[59]. This is further illustrated by an example from the komodoquinone B (**25**) pathway, where another SAM-dependent methyltransferase-like protein, EamK, catalyzes 10-decarboxylation (Figure 7)[151].

The late stages in the biosynthesis of **4** involve several atypical proteins (Figure 8A). Formation of this nogalamycin-subtype defining epoxyoxocin ring system is initiated by a two-component 1-hydroxylase system[161] that is composed of the SDR enzyme SnoaW and SnoaL2, which is related to polyketide cyclases such as SnoaL[123,162]. A similar system exists on the kosinostatin pathway, where the recruitment of two additional Rossmann-fold proteins KstA10 and KstA11 complete regioisomerization of the phenolic hydroxyl group (Figure 8B)[163]. In nogalamycin biosynthesis, the Rieske enzyme SnoT performs 2"-hydroxylation of the L-nogalamine sugar using a [2S–2Fe] iron sulfur cluster after the second glycosylation event[164]. The dual attachment of the amino sugar is completed by the non-heme iron and α-ketoglutarate dependent SnoK (Figure 8A)[165]. The final step in the biosynthesis of nogalamycin is 4"-epimerization to generate the L-nogalamine carbohydrate by SnoN (Figure 8A), which surprisingly shares significant sequence similarity with SnoK despite the difference in enzymatic activity[165,166].

Other unique post-PKS tailoring steps include oxidation and methylation reactions by P450 and SAM-dependent proteins, respectively, on the steffimycin pathway[125]. An atypical oxidoreductase AknOx catalyzes two subsequent oxidations of the terminal L-rhodosamine sugar to L-aculose supported by a bicovalently bound FAD on the aclacinomycin pathway[167]. Finally, kosinostatin harbors a bicovalently attached pyrrolopyrrole moiety, which is formed in a multi-enzyme cascade including non-ribosomal peptide assembly steps[134].

Discovery

Naturally occurring anthracyclines

The distribution of different classes of microbial natural products is uneven, with some compounds highly specific and only produced by a small number of species, whereas others appear to be 'household' compounds that are produced by almost all species[168]. Examples of household compounds are two compounds produced by nearly all streptomycetes, namely the siderophore desferrioxamine that is required for iron acquisition[169] and the volatile organic compound geosmin, which lends the typical earthy smell to soil[170]. Streptothricin and streptomycin are examples of ubiquitous antibiotics, and during screening at Eli Lilly these compounds were detected in 10% and 1% of the extracts of the *Streptomyces* cultures, respectively, while daptomycin was found in only about 1 in 10^7 extracts[171]. The distribution of genes for biosynthesis of anticancer compounds is similarly uneven[168]. For example, the BGC specifying the macrolide FD-891 is present in almost all publicly available *Streptomyces* genomes[168], which may be due to the fact that apart from cytotoxic activity, FD-891 also acts as a plant

phytotoxin. The distribution of anthracycline BGCs seems to be moderate, with doxorubicin and aclacinomycin pathways identified in 37 and 32 out of 1110 genomes, respectively, while nogalamycin BGC was found only in two genomes[168].

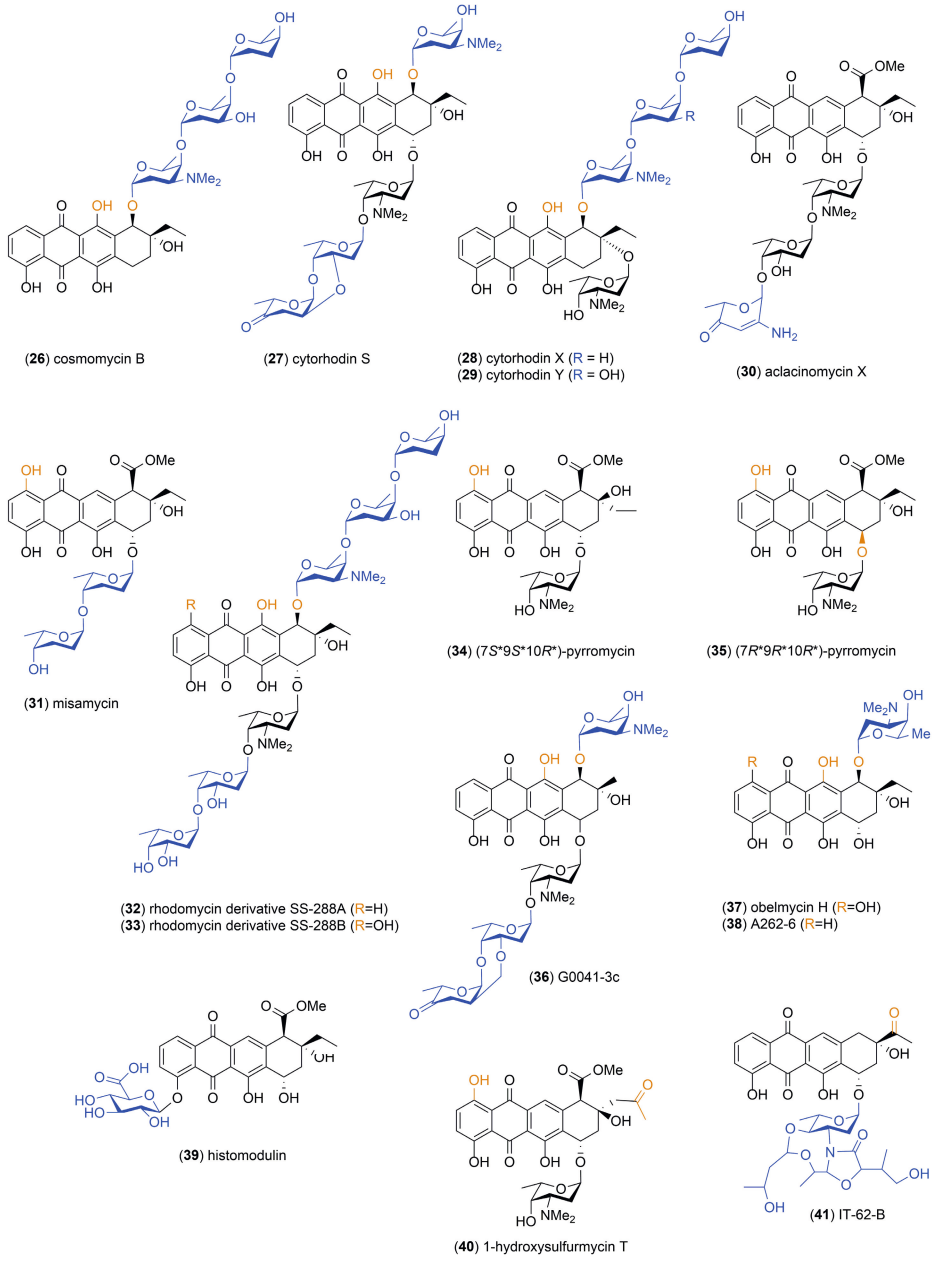


Figure 9. Structures of recently discovered anthracyclines based on the aclacinomycin scaffold. Differences in glycosylation and tailoring are colored in blue and orange, respectively.

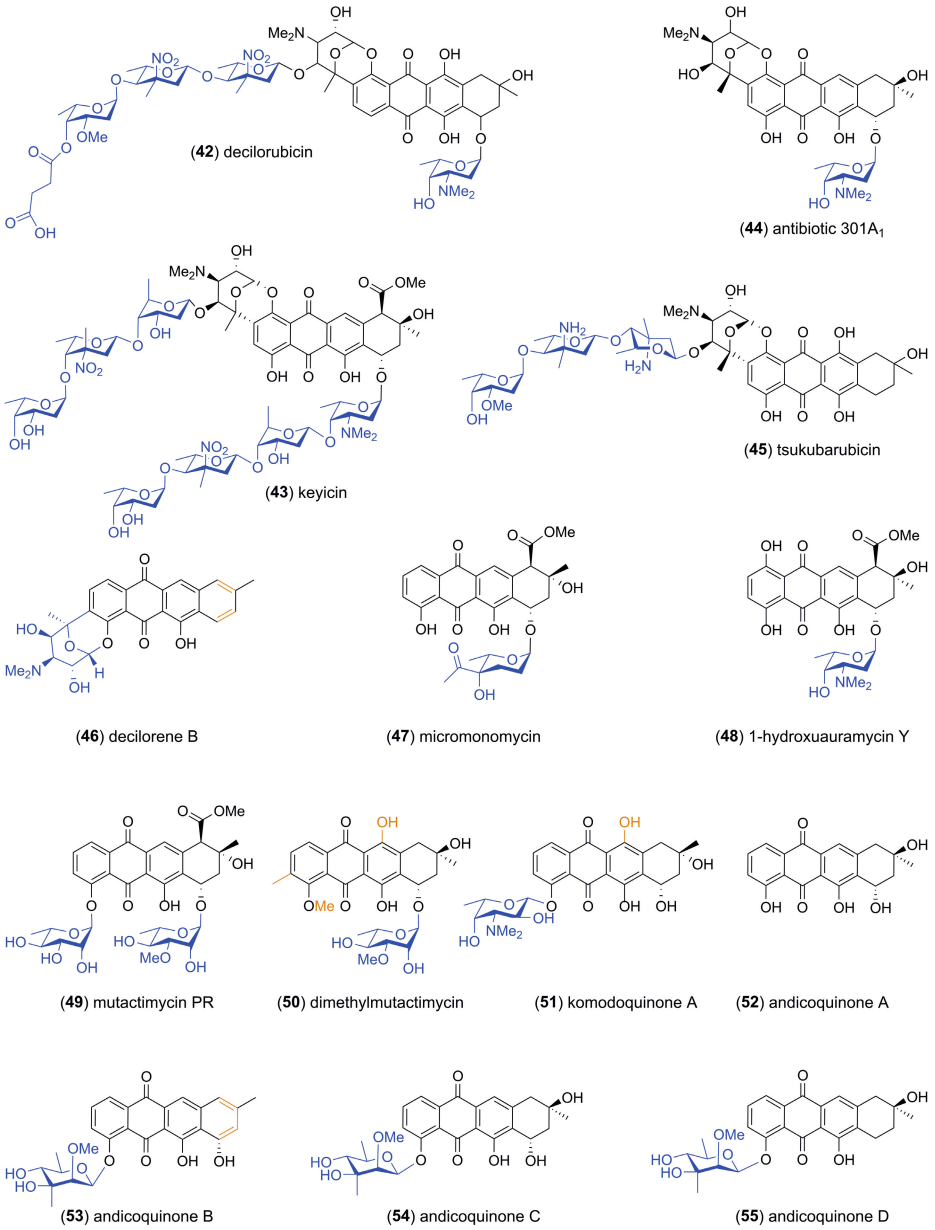


Figure 10. Structures of nogalamycin-type anthracyclines with epoxyoxocin ring systems. Differences in glycosylation and tailoring are colored in blue and orange, respectively. C9 carbon stereochemistry for **42** and **45** was not reported.

Despite the infrequent occurrence of anthracycline BGCs, the medical importance of these compounds has led to the isolation of more than 500 naturally occurring anthracyclines to date[172,173]. The majority of anthracyclines have been discovered from terrestrial *Streptomyces*

species, but the number of compounds obtained from marine and rare Actinobacteria have significantly increased in recent years due to increased bioprospecting of exotic locations. A comprehensive survey of all anthracyclines is beyond the scope of this review, and we will rather focus on compounds isolated during the 'genomics era' in the last 25 years.

Several anthracyclines related to the classical compounds such as **1**, **5**, cosmomycin B (**26**), cytorhodins (**27–29**) (Figure 9), but which harbor differences in the glycosylation and tailoring patterns have been reported. These include aclacinomycin X (**30**)[174], misamycin (**31**)[175], rhodomycin derivatives SS-288A (**32**) and SS-288B (**33**)[176], and two pyrromycins (**34**, **35**) (Figure 9)[177]. G0041-3c (**36**) contains an additional carbohydrate at C10[178], while obelmycin H (**37**) and its congener A262-6 (**38**) are glycosylated only at this position[179]. In turn, histomodulin (**39**) is a rhodomycin derivative containing a glucopyranuronosyl unit at C4[180]. Two doxorubicin-type anthracyclines with a C13 carbonyl group, 1-hydroxysulfurmycin (**40**)[177] and IT-62-B (**41**)[181] have been identified (Figure 9). All these metabolites have been isolated either from terrestrial, marine-derived, or endophytic *Streptomyces* sp.

A few novel nogalamycin-type anthracyclines with the archetypical epoxyoxocin ring system have also been identified. These include derivatives of classical decilorubicin (**42**)[182] such as keyicin (**43**)[133] and antibiotic 301A$_1$ (**44**) (Figure 10), which surprisingly displayed synergistic effects with other antibiotics against Gram-negative *E. coli*[54]. Tsukubarubicin (**45**) lacks carbohydrate units at C7, but contains unusual avidinosamine sugar units appended to the dually attached rhodosamine (Figure 10)[183]. Decilorene B (**46**) harbors the epoxyoxocin ring system in a different regiochemistry appended at C4[184]. Other acetate primed anthracyclines include micromonomycin (**47**), which contains a 2,3,6-trideoxy sugar with an unusual acyl substitution at C4'[185], and 1-hydroxyauramycin Y (**48**) (Figure 10)[177]. In addition, anthracyclines such as mutactimycin PR (**49**)[186], demethyl mutactimycin (**50**)[187], komodoquinone A (**51**)[188] and andicoquinones A–D (**52**–**55**)[184] have been isolated (Figure 10). Noteworthy, several of these anthracyclines contain a more uncommon 4-*O*-glycosylation pattern.

In recent years, several steffimycin-type metabolites have been discovered. These include 11-hydroxy-steffimycin (**56**)[189], steffimycin E (**57**)[190] and the more unusual steffimycin F (**58**)[191], which contains a benzyl group appended to the quinone ring (Figure 11). Novel aranciamycins I (**59**) and J (**60**) have been obtained from a marine-derived *Streptomyces*[192], while aranciamycin anhydride (**62**) was discovered from a strain isolated from the rhizosphere of Norway spruce (Figure 11)[193]. Tetracenoquinocin (**63**), with a fully aromatic A-ring, and 5-iminoaranciamycin (**64**) have been isolated from a sponge-associated streptomycete (Figure 11)[194]. *Amycolatopsis* sp. was found to produce mutactimycin E (**65**), with an atypical C3 methyl group[195]. Nocardicyclins A–B (**66**, **67**) isolated from *Nocardia pseudobrasiliensis*[196] and arimetamycins A–C (**68**–**70**) obtained *via* metagenomics approaches[132] harbor rare brasiliose sugar moieties (Figure 11).

Anthracycline scaffolds with atypical modifications have also been isolated in recent years. A novel cyano-substituted anthracycline (**71**)[197] and a fully aromatized anthracycline that is

O-glycosylated at C4 (**72**)[198] has been characterized from terrestrial *Streptomyces* sp. (Figure 12). Kosinostatin (**73**), a stereoisomer of the classical isoquinocycline B (**74**), contains an unusual pyrrolopyrrole unit[199,200]. In addition, several non-glycosylated kosinostatin/isoquinocycline B pathway intermediates (**75**–**80**) have been isolated from cultures of *Micromonospora* sp. (Figure 12)[199–201].

Finally, numerous new anthracycline aglycones without appended carbohydrates have been characterized recently. Aranciamycin K (**81**) from a marine-derived *Streptomyces* sp.[202] and four anthracyclinones (**82**–**85**) from *Micromonospora* sp.[203] originate from a longer four-carbon butyryl-CoA starter unit (Figure 13). Boshramycinones A–B (**86**, **87**)[204], SS-228*R* (**88**) and related sharkquinone (**89**)[205] from marine-derived streptomycetes harbor fully aromatic A-rings.

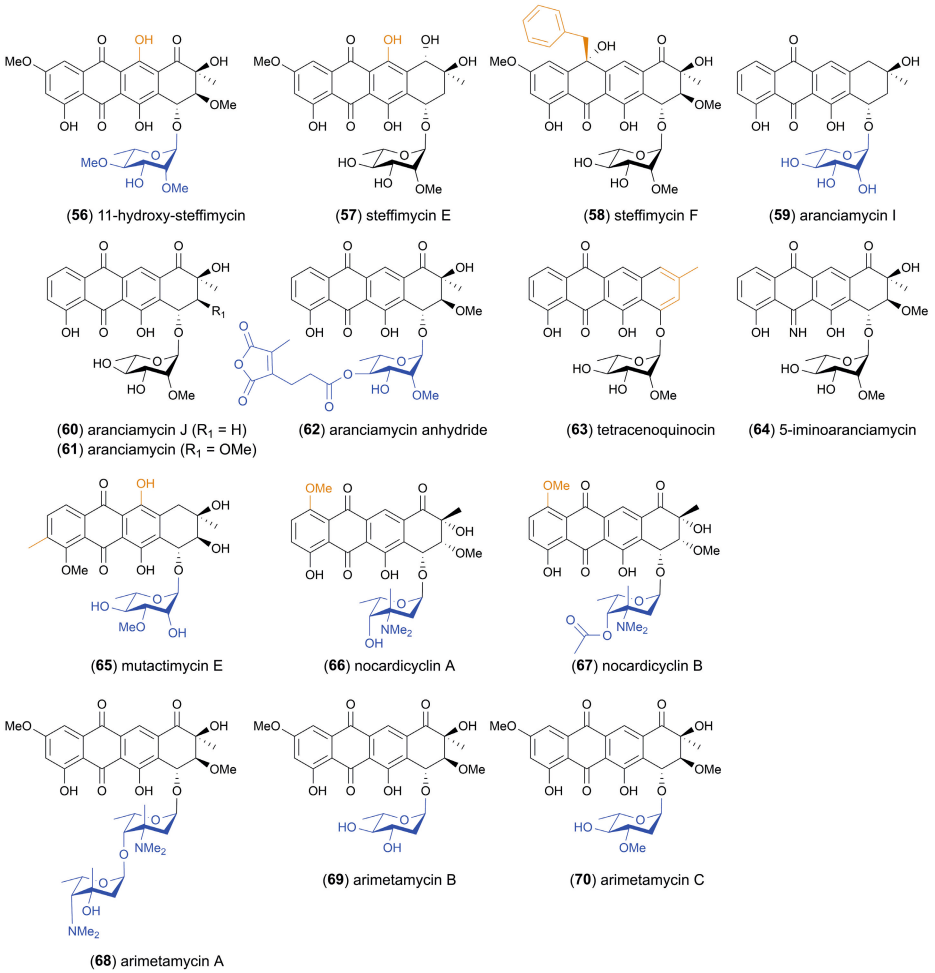


Figure 11. Structures of steffimycin-type anthracyclines. Differences in glycosylation and tailoring are colored in blue and orange, respectively.

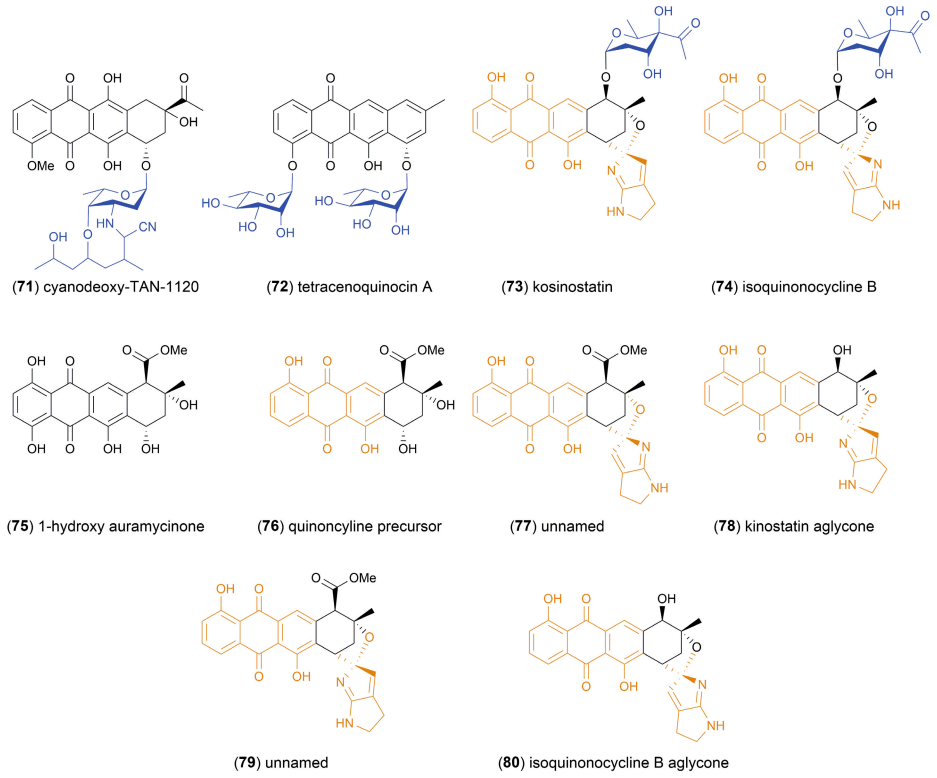


Figure 12. Anthracycline scaffolds with atypical modifications. Differences in glycosylation and tailoring are colored in blue and orange, respectively.

These anthracyclinones are derived from an acetyl-CoA starter unit, as is the related boshramycinone C (**90**) (Figure 13)[204]. The rare actinomycete *Nonomuraea rhodomycinica* NR4-ASC07 was found to produce 3-hydroxy substituted anthracyclinones (**91**)[206], while 2-chloro substituted celastramycin B (**92**) was identified from an endophytic *Streptomyces*[207]. Resomycins A–C (**93**–**95**), with modifications in the A-ring, have been isolated from a chartreusin producing terrestrial *Streptomyces* (Figure 13)[208].

Anthracycline biosynthetic gene clusters

To date 13 BGCs that have been experimentally verified to be involved in anthracycline biosynthesis. The medical importance of **5**[109,209] and **2**[110,210] is reflected in the fact that two BGCs for each have been identified from different *Streptomyces* strains, whereas single BGCs have been confirmed for rhodomycin (**1**)[156], nogalamycin (**4**)[112], cosmomycin (**26**)[148], cytorhodin (**27**)[145], kosinostatin (**73**)[134], aranciamycin (**59**)[211], arimetamycin (**68**)[132], keyicin (**43**)[133], tsukubarubicin (**45**)[183], komodoquinone (**25**)[151], and steffimycin (**6**)[125] pathways (Figure 14).

The majority of BGCs share homologous genes for the assembly of the aglycone units, with the exception of steffimycin, aranciamycin and arimetamycin BGCs, where the biosynthesis proceeds

via a decarboxylative cyclization mechanism[126]. The steffimycin and aranciamycin clusters are exceptional also in the sense that they do not encode any deoxysugar biosynthesis genes and the TDP-L-rhamnose building block has been proposed to be acquired from cell wall metabolism[125,148]. Genes required for formation of amino sugars are widely spread, and include TDP-L-daunosamine, TDP-L-rhodosamine and TDP-brasiliose in the case or arimetamycin[132]. It is noteworthy that also the komodoquinone BGCs harbors deoxysugar biosynthesis genes even though the strain only produced the aglycone komodoquinone B (**25**)[151]. However, glycosylated komodoquinone A (**51**) with a L-rhodosamine moiety has been isolated from another *Streptomyces* strain[188].

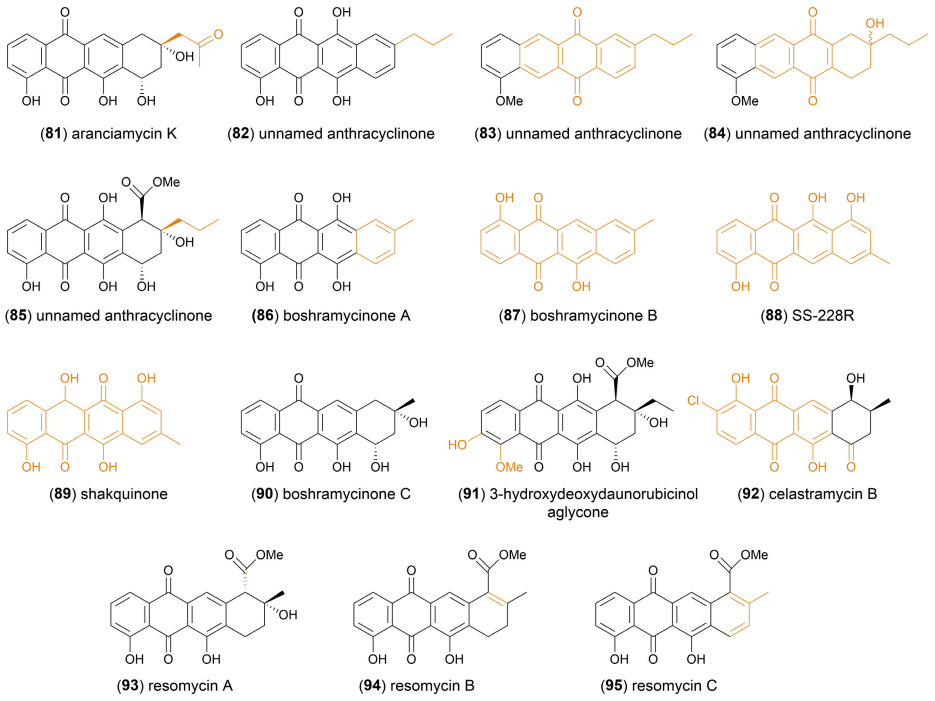


Figure 13. Recently discovered novel anthracyclinones. Unusual features are colored in orange.

The key differences in the BGCs are related to the number of glycosyl transferases, gene sets for the biosynthesis of diverse deoxysugars and the arsenal of tailoring genes present. Different combinations of these elements allow the rationalization for the diversification of anthracyclines. The aclacinomycin BGC does not contain any tailoring genes that modify the aglycone and therefore **5** can be considered as an ancient anthracycline. A noteworthy feature is that some of the more complex pathways appear to have been built on top of ancestral core pathways. For instance, the genomic organization of the arimetamycin BGC suggests that a TDP-brasiliose gene cassette may have merged with an ancestral steffimycin BGC in response to multidrug resistance efflux pumps that recognize the parental compound[132]. Further examples include expansion of the presumably parental nogalamycin and aclacinomycin BGCs through the inclusion of additional glycosyltransferases on the keyicin/tsukubarubicin and cosmomycin/cytorhodin pathways, respectively.

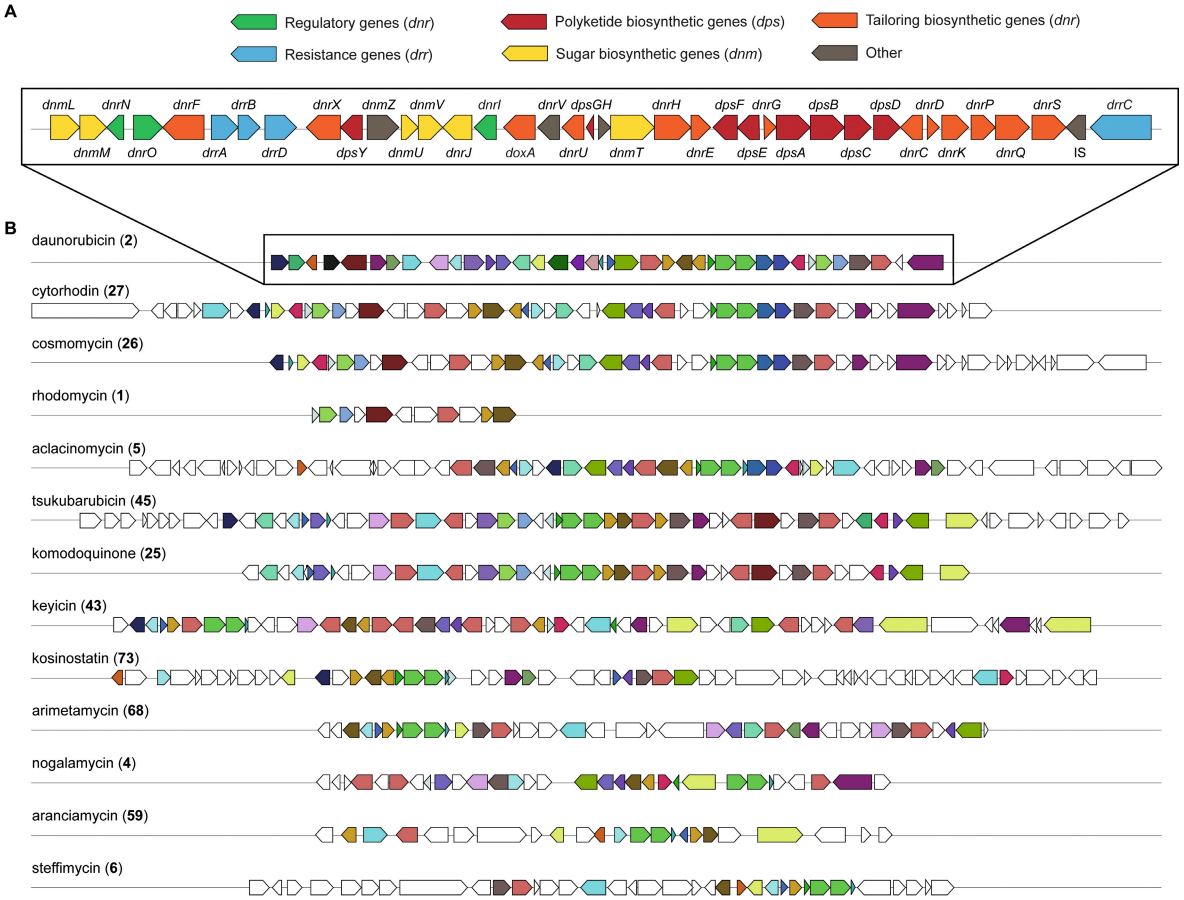


Figure 14. Currently known anthracycline biosynthetic gene clusters. (A) Schematic representation of the daunorubicin BGC. The color scheme represents the global gene function, classified as regulatory genes (green), resistance genes (blue), polyketide biosynthetic genes (red), sugar biosynthetic genes (yellow), tailoring genes (orange), or other (grey). **(B)** Gene cluster homology search of the daunorubicin BGC against twelve other experimentally verified anthracycline BGCs using MultiGeneBlast (v1.1.13) with default parameters[212]. The daunorubicin BGC is visualized in a color scheme that represents homology between genes of different clusters. For most BGCs the total extent of the clusters is unclear. Note that the aranciamycin and steffimycin BGCs lack sugar biosynthesis genes. The rhodomycin BGC has been sequenced only partially.

Genome mining strategies for discovery of novel BGGs

Actinobacteria harbor numerous cryptic or silent BGCs and activation of these pathways has become an increasingly important tool in the discovery of novel secondary metabolites.[40] The correct bioinformatic identification of anthracycline metabolic pathways is an essential prerequisite and few strategies, each with its own benefits and caveats, are available.

Identification of the full complement of cyclase genes is probably the most robust method to classify aromatic polyketides[213], but this is still relatively time consuming and manual curation of sequences is required. However, the fourth ring cyclases such as SnoaL[123] and AknH[122] appear to be highly specific for anthracyclines and can be used as a guide to drive the genome mining efforts. The added benefit is that phylogenetic analysis of the cyclase sequences allows prediction of the stereochemical configuration at C9. The limitation of the method is that steffimycin-type compounds are not detected as they harbor fourth ring cyclases such as StfX[197] that do not share sequence similarities to the canonical anthracycline cyclases.

Another approach is to focus on phylogenetic analysis of ketosynthase sequences, which reflect structures of the end products[214]. The analysis is complicated by the fact that anthracyclines appear to form several distinct clades[215]. Ketosynthase sequence analysis was utilized in the discovery of the arimetamycin pathway from a metagenomics eDNA library[132].

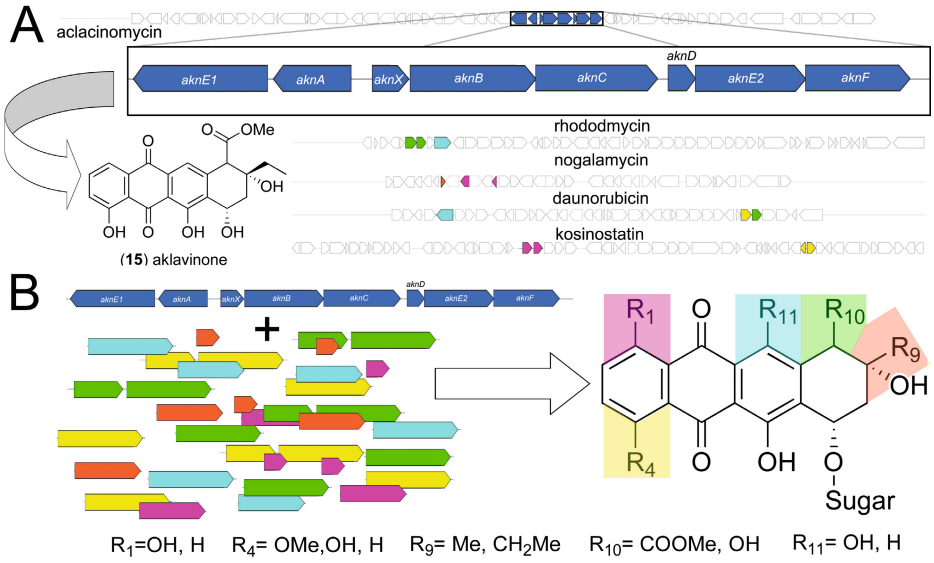


Figure 15. Generation of hybrid anthracyclines in *S. galilaeus* producing aclacinomycins. (A) The aclacinomycin BGC results in the formation of aklavinone (**15**). This aglycone is not modified by native tailoring enzymes, and therefore cloning of tailoring genes from the rhodomycin, nogalamycin, daunorubicin and kosinostatin pathways has led to formation of novel hybrid compounds. **(B)** Libraries of anthracyclines have been generated through modification of the aglycone at R_1 (magenta), R_4 (yellow), R_9 (orange), R_{10} (green), and R_{11} (cyan).

Metabolic engineering and combinatorial biosynthesis

Anthracycline pathways appear to be particularly amenable for pathway engineering efforts to further increase chemical diversity. The metabolic engineering efforts have focused particularly on expression of tailoring genes from related pathways[216] and glycodiversification[105].

S. galilaeus producing **5** and other aclacinomycins has been an ideal host for the generation of hybrid anthracyclines, since the aclacinomycin BGC does not harbor any native aglycone modification genes[216]. This has allowed modification of substituents at R_1, R_4, R_9, R_{10} and R_{11} of the aklavinone aglycone by cloning different combinations of tailoring genes from the rhodomycin[217,218], nogalamycin[115,219,220], daunorubicin[221–223] and kosinostatin[224] pathways (Figure 15). The efficacy of such pathway engineering experiments was demonstrated by expression of the minPKS from the nogalamycin pathway together with the 10-methylesterase, 10-hydroxylase and 11-hydroxylase from the rhodomycin pathway in a *S. galilaeus* mutant HO75, which produces a mixture of differently glycosylated aclacinomycins due to a defective dTDP-hexose-3-dehydratase[225], that resulted in the production of a library of 60 anthracyclines[226]. Hybrid anthracyclines have also been generated by incorporating additional foreign genes to the daunorubicin[219], epelmycin[223] and steffimycin[227] pathways.

Modification of the carbohydrate units of anthracyclines affects the bioactivity of the compounds and therefore the engineering of deoxysugar pathways has generated wide interest. The pioneering example is the production of biosynthetic **7** in *S. peucetius*, which was achieved through inactivation of the natural TDP-4'-keto-2,3,6-trideoxyhexulose reductase gene *dnmV* and expression of 4-ketoreductases with opposite stereoselectivity[228].

The success of these approaches relies on the ability of the glycosyltransferases to recognize and append unnatural TDP-carbohydrates to anthracyclinones. The L-rhamnose glycosyltransferases StfG and AraGT from the steffimycin and aranciamycin pathways appear to be particularly promiscuous towards foreign TDP-carbohydrates. Expression of the aranciamycin BGC in two heterologous hosts led to the isolation of eight differently glycosylated derivatives (Figure 16)[213,229]. In turn, co-expression of the steffimycin BGC together with gene cassettes for synthesis of various TDP-carbohydrates in *S. albus* led to the formation of twelve derivatives (Figure 16)[230].

The heterologous host *S. venezuelae* has also been engineered to produce numerous TDP-carbohydrates and to express the two-component aclacinomycin *aknS* and *aknT* glycosyltransferase system[231]. Bioconversion experiments with extracellularly provided ε-rhodomycinone led to attachment of TDP-L-daunosamine, TDP-L-rhodosamine, TDP-L-ristosamine, TDP-L-vancosamine, TDP-D-digitoxose, TDP-L-digitoxose and TDP-L-rhamnose to the aglycone, and demonstrated the substrate flexibility of glycosyltransferases with the formation of twenty anthracyclines (Figure 17)[231]. The method was expanded to a one-pot combinatorial biosynthesis system, where distinct anthracyclinones and carbohydrates are produced in separate *Streptomyces* strains, but glycosylated anthracyclines are obtained upon co-cultivation. The experiments led to formation of 16 anthracyclines, including seven novel metabolites (Figure 17)[232].

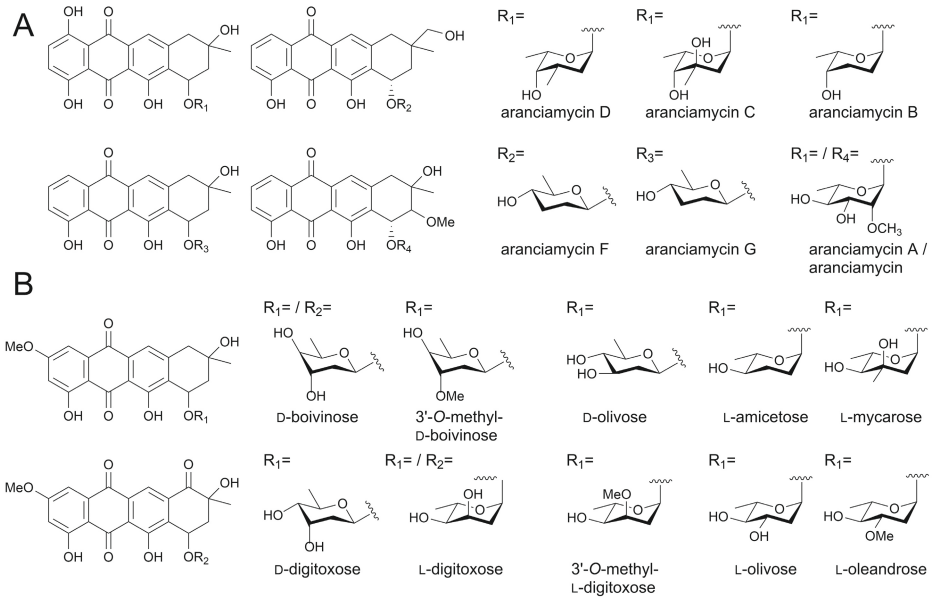


Figure 16. Combinatorial biosynthesis of aranciamycin and steffimycin in heterologous hosts. **(A)** Aranciamycin analogs have been generated by expression of the aranciamycin BGC from *S. echinatus* in *S. albus*, *S. fradiae* Tü2717, *S. fradiae* A0, and *S. diastatochromogenes* Tü6028. Novel glycosides are formed due to the availability of different TDP-carbohydrates in the heterologous hosts and the promiscuity of the glycosyltransferases. **(B)** Heterologous expression of the steffimycin BGC together with various TDP-carbohydrate pathways in *S. albus* led to the generation of twelve glycosylated derivatives of steffimycin.

Another means for generating novel anthracyclinones has been the assembly of polyketide biosynthesis genes in a heterologous host. A number of *Streptomyces* strains have been optimized for the production of heterologous metabolites[233,234]. An example is *S. coelicolor* M1152, which was obtained by deletion of the four main antibiotics clusters and ribosome engineering[235]. *S. venezuelae* is especially well-suited for the heterologous expression of polyketides due to the presence of a large number of innate phosphopantetheinyl transferase genes, which are required for the activity of the PKS machinery[236]. The anthracyclines aklavinone and ε-rhodomycinone have been produced in this host[232]. In addition, **15** and auravinone, which is identical to **16** except for opposite stereochemistry at C9, have been produced in *S. lividans* TK24 and *S. coelicolor* CH999 through assembly of biosynthetic genes from three different anthracycline pathways[109,117].

Improved production of anthracyclines

The secondary metabolites produced by streptomycetes are generally produced in low concentrations (μg·L^{-1} to mg·L^{-1}). *Streptomyces peucetius* ATCC 27952 produces about 1–20 mg·L^{-1} of **3** in production media[63,237]. However, the production titers need to be much higher for industrial-scale production (g·L^{-1}). Traditionally, strains with improved productivity were

obtained by random mutagenesis and screening, also called classical strain improvement[236]. This method has significant drawbacks as it is time-consuming and there is a risk of accumulation of unwanted mutations that affect other properties of the strain[238]. Developments in the fields of genetic engineering and systems biology have resulted in rational strain improvement methods[239,240]. These approaches focus on modification of regulatory networks, increasing precursor availability, increasing expression of rate-limiting biosynthetic enzymes, deleting genes for competing pathways, increasing product tolerance, combinatorial biosynthesis, and heterologous hosts.

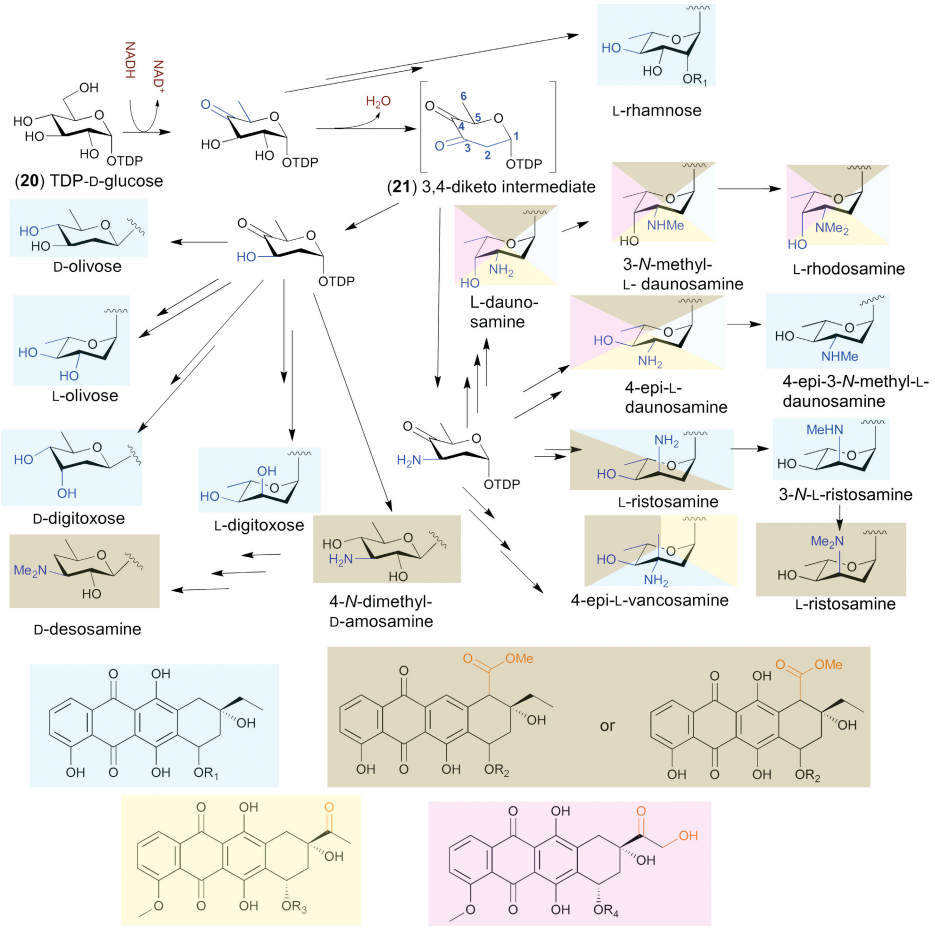


Figure 17. Production of differently glycosylated novel rhodomycins and daunorubicins in *S. venezuelae*. A series of plasmids encoding TDP-carbohydrate pathways (highlighted in cyan, yellow, or magenta) were assembled in *S. venezuelae*. Novel glycosylated anthracyclines were generated by feeding ε-rhodomycinone to the cultures for bioconversion reactions or by co-cultivation (highlighted in brown) with engineered *S. venezuelae* strains expressing core anthracycline pathways. Incorporation of the glycoside to the aglycone was possible due to the promiscuity of the AknS/AknT glycosyltransferase.

Engineering regulatory networks

The production of secondary metabolites is tightly regulated by global and pathway-specific control. Pathway-specific regulators may be positive or negative regulators and are typically regulated by environmental sensors and global regulators. The regulatory network is a key target for strain improvement, for example by overexpression of positive regulators and deletion of negative regulators.

The doxorubicin BGC contains a negative regulator (*dnrO*), and two positive regulators (*dnrN* and *dnrI*) (Figure 2)[82]. Overexpression of *dnrN* and *dnrI* in *S. peucetius* ATCC 29050 resulted in a two-fold (about 200 mg·mL^{-1}) overproduction of **2**[241]. Overexpression of the global regulatory gene *afsR* in *S. peucetius* ATCC 27950 resulted in 4-fold (about 80 mg·mL^{-1}) overproduction of **3**[242]. Overexpression of *dnrN* and *dnrI* under the *ermE** promoter in combination with *afsR* and SAM synthase genes to improve co-substrate availability resulted in up to 4.3-fold (about 7.5 mg·mL^{-1}) overproduction[65]. Other global regulatory proteins that have been targeted to increase the production of **3** are *wblA*[243] and *doxR*[244].

Metabolic engineering for yield improvement

The building blocks of secondary metabolites originate from primary metabolism[238]. The aglycone backbone of anthracyclines is typically constructed from acetyl-CoA or propionyl-CoA and is extended with several units of malonyl-CoA. The sugar moieties generally originate from glucose-6-phosphate. Hence, the anthracycline precursors are generated in glycolysis, the tricarboxylic acid cycle, and the pentose-phosphate pathway. The production of anthracyclines may be improved by increasing the availability of these precursors. Acetyl-CoA carboxylase (Aac) catalyzes the conversion of acetyl-CoA to malonyl-CoA. The biosynthesis of glucose-6-phosphate depends on the activity of phosphoglucomutase (Pgm). Overexpression of *acc* and *pgm* resulted in increased production of anthracycline(-like) compounds[245,246]. The production of **4** by *S. nogalater* NRRL3035, **6** by *S. steffisburgensis* NRRL 3193, and elloramycin by *Streptomyces olivaceus* Tü2353 was increased 1.5 to 2-fold[246]. The availability of acetyl-CoA depends on the triacylglycerol pool. The acetyl-CoA synthase is responsible for the activation of fatty acids with coenzyme A (CoA). Inducible expression of the *S. coelicolor* acetyl-CoA synthase SCO6196 resulted in increased production of actinorhodin (1.9-fold; 216.1 mg·L^{-1}), jadomycin B (1.7-fold; 133.0 mg·L^{-1}), oxytetracycline (4.7-fold; 4.54 g·L^{-1}), and avermectin B_{1a} (0.5-fold; 9.31 g·L^{-1}) in respectively, *Streptomyces coelicolor* M145, *Streptomyces venezuelae* ISP5230, *Streptomyces rimosus* M4018, and *Streptomyces avermitilis* A56[247]. This strategy has not yet been applied for the production of anthracyclines. However, overexpression of two genes involved in CoA biosynthesis (*coaA* and *coaE*) resulted in 2.1-fold (3.34 mg·mL^{-1}) increased production of **3** in *S. peucetius* ATCC 27952[248]. Precursor supply of polyketides has recently been reviewed elsewhere[249].

The overexpression of biosynthetic genes has also been widely applied to improve production levels, with a special focus on rate-limiting enzymes. Heterologous enzymes from the sugar pathway of *S. venezuelae* ATCC 15439 were expressed in *S. peucetius* ATCC 27952, which

resulted in an increased production of **3** by 2.6-fold (about 3.5 mg·mL^{-1})[237]. The activity of the glycosyltransferase DnrS is rate-limiting, resulting in an accumulation of ε-rhodomycin. Combined overexpression of the *S. venezuelae* sugar genes and the native glycosyltransferase pair (*dnrS*/*dnrQ*) resulted in 5.6-fold (about 9.5 mg·mL^{-1}) increased production[237]. The final enzyme in the doxorubicin pathway (DoxA) is responsible for the last three steps towards **3**. Especially the last step from **2** to **3** is 170-fold less efficient than the previous step in the pathway[153]. Overexpression of *doxA* is not sufficient to significantly improve production of **3**, which is still an interesting target for engineering[250].

Another strategy may be to express multiple copies of the BGC. During classical strain improvement, amplification of BGCs has been reported as a cause for yield improvement, for example for penicillin[251] and kanamycin[252] production. Currently, several methods like the TAR system, the IR system and the BAC system are available for the cloning of entire BGCs, which can be used for heterologous expression or duplication of the BGC in the native host[253].

Secondary metabolites are often toxic to their producers. Consequently, increasing product tolerance by overexpression of resistance genes is another widely applied strategy. The DNA damaging activity of anthracyclines may also affect the producing strain. Therefore, the BGCs generally include exporters and DNA repair mechanisms. Overexpression of the doxorubicin resistance genes *drrA* and *drrB* (Figure 2) in *S. peucetius* ATCC 27952 resulted in a 2.2-fold (about 3 mg·mL^{-1}) increased production of **3**, and overexpression of *drrC* (Figure 2) in a 5.1-fold (about 7.0 mg·mL^{-1}) increase[254]. In addition, improved resistance enhances the stability of the production strain, which is an important consideration for industrial production.

Another approach for improving production is the inactivation of competing pathways. Two examples can be taken from the doxorubicin pathway. First, the availability of the sugar moiety daunosamine can be increased by inactivation of the rhamnose pathway that competes for intermediates. Inactivation of the final enzyme in the biosynthetic pathway of rhamnose (RmdD) may result in an increased daunosamine availability, which could result in increased production of **3**[65,255]. Besides that, the production levels of **2** and **3** are decreased due their conversion into baumycin-like higher glycosides. The enzymes responsible for these conversions are *dnrH*, *dnrU*, and *dnrX*. Deletion of *dnrH* resulted in increased production of **2** in *S. peucetius* ATCC 27952 of 8.5-fold (about 15 mg·mL^{-1}). Deletion of *dnrH* and *dnrX* resulted in a 3-fold (14 mg·mL^{-1}) increased production of **2**[256,257].

Fermentation and media

The production of secondary metabolites is linked to changes in the environment, and therefore optimizing media and fermentation methods is crucial to increase production levels. Carbon catabolite repression plays a role in most streptomycetes, resulting in reduced production of secondary metabolites in the presence of a high concentration of carbon source, mainly glucose[258]. However, high growth rates are often obtained with glucose as carbon source. Consequently, medium design is always a trade-off between growth rates and production

rates[238]. The production of **3** is reduced in high glucose conditions[259], and high levels of inorganic phosphate also repress production of anthracyclines by *S. peucetius*[260]. Phosphate levels dictate the transition from primary to secondary metabolism[261]. Recently, the production of **3** was improved to 1.1 $g \cdot L^{-1}$ by a combination of classical strain improvement and medium optimization[262].

Perspectives

Back in 1992, Weiss wondered whether 'We will ever find a better doxorubicin' after the generation of hundreds of semi-synthetic derivatives that did not significantly improve the outcome of anticancer chemotherapy or limit the side effects[13]. Yet, recent studies with these old drugs using new technologies have identified unrecognized activities that are unique to anthracycline drugs: nucleosome eviction. The complex bioactivities of anthracyclines suggest that additional relevant activities may be uncovered in the future. While the anticancer activities of anthracyclines can sometimes be improved as exemplified by the recently discovered semi-synthetic utorubicin[263], the major limitation in the use of these drugs are the toxic side effects. The strongly reduced toxicity profile of **12**, without compromising anticancer efficacy, shows that overcoming toxicity issues is possible. Further points of interest are modifications affecting tissue distribution and/or reducing the removal of drugs from cells by ABC transporters such as MDR1, which may contribute to better anthracycline drugs.

These developments have coincided with an increased understanding of how *Streptomyces* bacteria are able to produce numerous anthracyclines with relative ease. The accumulation of next-generation sequencing data and modern synthetic biology techniques will offer unprecedented possibilities for manipulation of anthracycline biosynthetic pathways to further increase the chemical diversity within this family of compounds. Discovery of novel anthracyclines continues to thrive also by traditional means, by focusing on culture extracts from isolates of rare Actinobacteria.

One critical issue for future development of anthracycline drugs is the need for coordination between natural product discovery and bioactivity profiling. Most of the anthracyclines described in this review have never been tested, for instance, towards histone eviction activity or cardiotoxicity. The primary reason is the shift in anthracycline research back to academia due to current lack of interest from pharmaceutical companies. An international multidisciplinary community effort to routinely screen all newly discovered compounds for bioactivities beyond basic cytotoxicity parameters is gravely needed. If successful, the combination of metabolic engineering, semi-synthesis and modern bioactivity profiling will allow re-evaluation of the chemical space around anthracyclines and discovery of improved drugs. For these reasons, the future of anthracycline drugs looks bright and, yes, we believe that a better doxorubicin can indeed be found.