New insights into the glycosylation steps in the biosynthesis of Sch47554 and Sch47555

Ozkan Fidan†[a], Riming Yan,†[a,b] Gabrielle Gladstone,[a] Tong Zhou,[a] Du Zhu,*[b] Jixun Zhan*[a]

Abstract: Sch47554 and Sch47555 are antifungal compounds from *Streptomyces* sp. SCC-2136. The availability of the biosynthetic gene cluster made it possible to track down the genes encoding biosynthetic enzymes responsible for the structural features of these two angucyclines. Sugar moieties play important roles in the biological activities of many natural products. Investigation of glycosyltransferases (GTs) may potentially help diversify pharmaceutically significant drugs through combinatorial biosynthesis. Sequence analysis indicated that SchS7 is a putative C-GT, while SchS9 and SchS10 were proposed to be O-GTs. In this study, we characterized the roles of these three GTs in the biosynthesis of Sch47554 and Sch47555. Co-expression of the aglycone and sugar biosynthetic genes with *schS7* in *Streptomyces lividans* K4 resulted in the production of C-glycosylated rabelomycin, which revealed that SchS7 attaches a D-amicetose moiety to the aglycone core structure at the C-9 position. Gene inactivation studies revealed that subsequent glycosylation steps take place in a sequential manner in which SchS9 first attaches either a L-aculose or L-amicetose moiety to the 4'-OH of the C-glycosylated aglycone, then SchS10 transfers a L-aculose moiety to the 3-OH of the angucycline core.

Introduction

Angucycline antibiotics are a group of biologically active natural products synthesized by type II polyketide synthase (PKS) complexes through a sequence of reactions.[1,2] Angucyclines exhibit diverse biological activities including antiviral (SM 196 B), antibacterial (marangucycline A, vineomycin A_1), antitumor (landomycin A, urdamycin A, and marangucycline B) and antifungal properties (Sch47554 (**1**) and Sch47555 (**2**), Figure 1).[3–8] Sch47554 and Sch47555 are produced by *Streptomyces* sp. SCC-2136 (ATCC 55186) and exhibit antifungal activity against various yeasts and dermatophytes.[6] The entire biosynthetic gene cluster encoding the biosynthesis of Sch47554 and Sch47555 has been previously reported.[9] The availability of biosynthetic gene cluster (NCBI accession number AJ628018) made it possible to track down the genes encoding biosynthetic enzymes responsible for the structural features of these two angucyclines. The Sch47554 and Sch47555 (*sch*) biosynthetic gene cluster includes enzymes responsible for the biosynthesis of the core polyketide backbone (SchP6-8) and nucleotidyl-activated sugar moieties (SchS1-6). It additionally encodes some tailoring enzymes such as aromatase (SchP4), ketoreductase (SchP5), cyclase (SchP9), oxygenase (SchP10), and three putative glycosyltransferases (GTs) (SchS7, SchS9, and SchS10).[9]

GTs catalyze the transfer of activated sugar moieties to the acceptor molecules and are commonly involved in the biosynthesis and modification of many pharmaceutically significant natural products.[10] Regio- and stereo-specifically attached sugar moieties play important roles in both binding of drugs to biological targets and biological activity of many natural products.[11,12] For instance, landomycin A with a hexasaccharide side chain exhibits much stronger antitumor activity compared to its derivative, landomycin E with a trisaccharide side chain.[13] GTs hold huge potential for helping diversify pharmaceutically important drugs through the combination of chemoenzymatic and *in vivo* methods.[14] For this purpose, heterologous expression and in-frame gene deletion approaches have been implemented in many studies. For example, Rodriguez et al. generated hybrid elloramycin analogs by combinatorial biosynthesis using genes from anthracycline-type and macrolide biosynthetic pathways.[15] Tang and McDaniel also benefited from combinatorial biosynthesis to produce desosaminylated macrolactones and paved the way for the production of 'unnatural' natural product libraries.[16] Trefzer et al. were able to produce novel glycosylated urdamycin derivatives by overexpressing GT genes from landomycin-producing strain of *Streptomyces cyanogenus* S136 in a mutant of urdamycin-producing strain of *Streptomyces fradiae* Tü 2717.[17] With a similar strategy, Ostash et al. generated novel landomycins in a mutant strain.[13] Künzel et al.

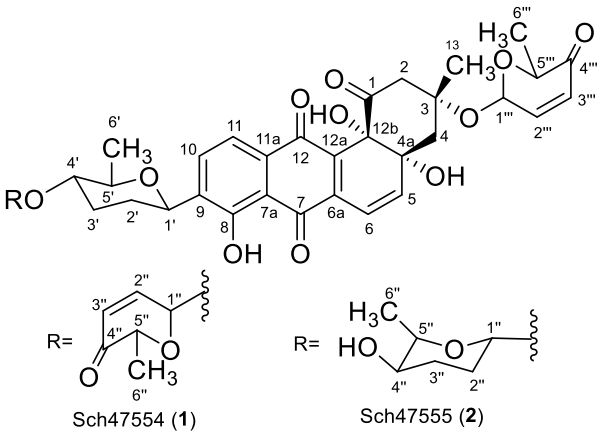


Figure 1. Structures of Sch47554 (**1**) and Sch47555 (**2**).

[a] *O. Fidan, Dr. R. Yan, G. Gladstone, Dr. T. Zhou, Prof. Dr. J. Zhan*
Department of Biological Engineering, Utah State University, 4105 Old Main Hill, Logan, UT 84322, USA
E-mail: jixun.zhan@usu.edu
[b] *Dr. R. Yan, Prof. Dr. D. Zhu*
Key Laboratory of Protection and Utilization of Subtropic Plant Resources of Jiangxi Province, College of Life Science, Jiangxi Normal University, Nanchang, Jiangxi 330022, China
Email: zhudu12@163.com
[†] *These authors contributed equally to this work.*

Supporting information for this article is given via a link at the end of the document.((Please delete this text if not appropriate))

and Luzhetskyy et al. performed targeted gene inactivation including the inactivation of genes encoding GTs, leading to the formation of urdamycins I, J, K and landomycins M and O, respectively.[18,19] All these studies indicate that novel and/or unnatural natural products can be generated by the inactivation of GT genes and combinatorial biosynthesis.

In the present study, we elucidated the roles of three GTs involved in the biosynthesis of **1** and **2**. We have used two aforementioned approaches to functionally characterize GTs in the *sch* biosynthetic gene cluster. The function of *schS7* was characterized through heterologous expression, while we performed targeted gene inactivation for *schS9* and *schS10*. These two approaches not only revealed the function of GTs, but also help to diversify the pharmaceutically relevant molecules through the production of two novel compounds.

Results and Discussion

Sequence analysis of the putative GTs in the *sch* biosynthetic gene cluster

The *sch* biosynthetic gene cluster contains three putative GT genes (*schS7*, *schS9*, and *schS10*). Multiple amino acid sequence analysis of SchS7 using Clustal Omega revealed its homology to known *C*-GTs: 72.33% identity with UrdGT2 from *Streptomyces fradiae*, 68.78% identity with SaqGT5 from *Micromonospora* sp. Tü 6368 and 55.97% identity with SimB7 from *Streptomyces antibioticus* Tü6040, as well as to a known *O*-GT, LanGT2 from *Streptomyces cyanogenus* with 58.71% identity [14,20–22]. Multiple amino acid sequence analysis of SchS9 revealed its similarity to several characterized *O*-GTs: 58.91% identity with UrdGT1c from *Streptomyces fradiae*, 58.21% identity with SaqGT4 from *Micromonospora* sp. Tü 6368, and 54.38% identity with LanGT1 from *Streptomyces cyanogenus*.[17,21,23] Similarly, multiple amino acid sequence analysis of SchS10 revealed its homology to a few *O*-GTs: 54.50% identity with SaqGT2 from *Micromonospora* sp. Tü6368, 49.40% identity with LanGT4 from *Streptomyces cyanogenus*, and 48.82% identity with UrdGT1a from *Streptomyces fradiae*.[17,21,23] Clustal Omega multiple sequence alignments for each GTs (Figures S1-S3) indicated that SchS7 is expected to possess the function of *C*-GT, whereas SchS9 and SchS10 are predicted as *O*-GTs.

Functional characterization of *schS7* through heterologous expression

Based on the sequence analysis, SchS7 is predicted to be responsible for the attachment of a sugar moiety to C-9 of the aglycone backbone via the formation of a C-C bond. To verify the function of this putative *C*-GT, we first aimed at the expression of minimal PKS genes of *sch* gene cluster (*schP6-8*) and of *oxy* gene cluster (*oxyABC*) to test the putative functions of SchP6, SchP7, and SchP8, which led to the production of SEK15, same product from the co-expression of OxyABC from the oxytetracycline biosynthetic pathway (data not shown).[24] We then expressed more genes in an attempt to produce a stable aglycone for GT studies. To this end, *schP4-10* genes was ligated into the pRM5 vector, an *Escherichia coli/Streptomyces* shuttle vector to yield pOF10, which was then expressed in *S. lividans* K4. HPLC analysis revealed that a major product **3** was generated by *S. lividans* K4/pOF10. through heterologous expression of *schP4-10* genes (pOF10) in *S. lividans* K4 (Figure 2A). The ESI-MS spectrum of **3** showed the $[M+H]^+$ ion peak at *m/z* 339.0 (Figure S4), indicating that its molecular weight is 338, which is same as that of rabelomycin, a common intermediate in angucycline biosynthesis. The structure of **3** was confirmed to be rabelomycin (Figure 2B) based on a comparison of its ^{1}H NMR data for **3** with previously reported data.[25]

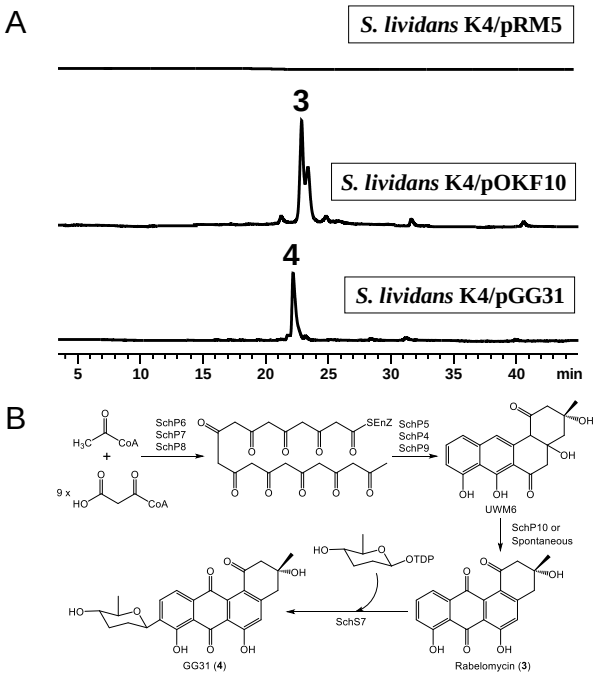


Figure 2. Production of rableomycin (**3**) and 9-D-amicetosyl-rabelomycin (GG31, **4**) in *S. lividans* K4 through heterologous expression. (A) HPLC analysis (420 nm) of the products of *S. lividans* K4/pRM5 (upper), pOF10 (middle) and pGG31 (bottom). **(B)** Biosynthetic pathway of **3** and **4**.

Using the same approach, we introduced more genes including *schS7* and the genes involved in the biosynthesis of activated sugar moieties (*schS1-S6* and *schS8*) to pOF10 to yield pGG31. Expression of this set of genes in *S. lividans* K4 yielded another major product **4**, as shown in Figure 2A. The ESI-MS spectrum of **4** revealed the $[M+H]^+$ ion peak at *m/z* 453.1 (Figure S4), suggesting a molecular weight of 452. **4** is 114 mass units larger than **3**, indicated that **4** is a C-amicetosylated derivative of **3**. To elucidate the structure of **4**, HR-MS analysis was conducted and showed a $[M+H]^+$ ion peak at *m/z* 453.1561, which is consistent with $C_{25}H_{25}O_8$ (cacld. 453.1505) (Figure S5). The formula of **4** revealed that it has 25 carbons, which has 6 more carbons than **3**, further indicating that a D-amicetose moiety has been added. This was supported by 1D and 2D NMR analysis (Table 1 and Figure 3). The NMR data of **4** are similar to those of **3**, except that it has extra signals that belong to the D-amicetose moiety. The ^{13}C NMR spectrum showed 25 carbon signals. The chemical shifts of C-6 (δ 152.3) and C-8 (δ 157.9) indicated that they are hydroxylated carbons on the aromatic rings. Additionally, an oxygenated quaternary carbon signal at δ 71.3 was assigned to

C-3. These signals are consistent with those in rabelomycin. The six carbon signals at δ 73.0, 33.2, 32.0, 79.2, 71.2 and 18.2, and another hydroxylated CH signal belong to the D-amicetose moiety, which further confirmed the presence of this sugar moiety in **4**. The HMBC correlation of the anomeric proton H-1 at δ_H 4.79 to C-9 (δ_C 139.8) confirmed the linkage of the D-amicetose moiety to the C-9 position (Figure 3). Accordingly, the structure of **4** can be identified as 9- D-amicetosyl-rabelomycin and this compound was named as GG31. Therefore, the function of SchS7 was identified as a *C*-GT in the biosynthesis of Sch47554 and Sch47555 (Figure 2B).

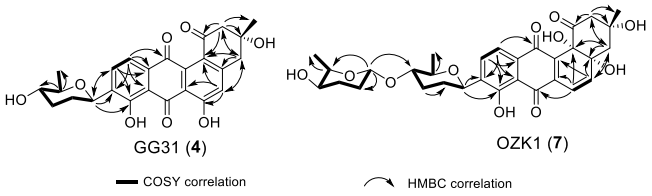


Figure 3. Selected ^{1}H-^{1}H COSY and HMBC correlations for **4** and **7**.

Targeted gene inactivation of *schS9* and *schS10*

Sequence analysis of SchS9 and SchS10 suggested that they are responsible for *O*-glycosylations at 3-OH and 4'-OH of **1** and **2**. We inactivated both genes through homologous recombination to generate mutants lacking corresponding *O*-GTs (Figure 4A). Apramycin resistant exconjugants were obtained from intergenic conjugation between *Streptomyces* sp. SCC-2136 and *E. coli* ET12567 respectively harboring pRY9 and pRY10. The confirmation of the single crossover events was performed by PCR using vector- (RV-M and M13-47) and genome-specific primers (Figure 4B). As illustrated in Figure 4A and 4B, the 2.2 kb and 2.1 kb PCR products were amplified from *Streptomyces* sp. SCC-2136/Δ*schS9* genome, whereas *Streptomyces* sp. SCC-2136 wild type genome did not yield these fragments. Similarly, we amplified the 2.6 kb and 2.8 kb fragments from the genome of *Streptomyces* sp. SCC-2136/Δ*schS10*, while we did not obtain these expected fragments from the genome of wild type strain. Thus, these PCR results confirmed the targeted gene inactivation for *schS9* and *schS10*.

Upon confirmation of gene inactivation, the positive exconjugants are grown on YM plates supplemented with apramycin for product analysis. HPLC analysis of the extracts of the cultures of *Streptomyces* sp. SCC-2136/Δ*schS9* and *Streptomyces* sp. SCC-2136/Δ*schS10* at 420 nm showed two same products (**5** and **6**) at 19.6 min and 26.5 min in both extracts. *Streptomyces* sp. SCC-2136 Δ*schS10* produced two additional compounds (**8** and **7**) at 26.1 and 32.5 min (Figure 4C). The ESI-MS spectra of **5**, **6**, **7**, and **8** showed the $[M-H]^-$ ion peaks at *m/z* 373.0, 469.1, 583.1, and 579.1, respectively (Figure S6), indicating that their molecular weights are 374, 470, 584, and 580, respectively. We purified all these compounds for chemical structure elucidation, except **8** due to the low titer. We confirmed the chemical structures of **5** and **6** based on the ^{1}H NMR data as listed in Table1. We, additionally, ran ^{1}H, ^{13}C, HSQC, and HMBC NMR as well as HR-MS for OZK1

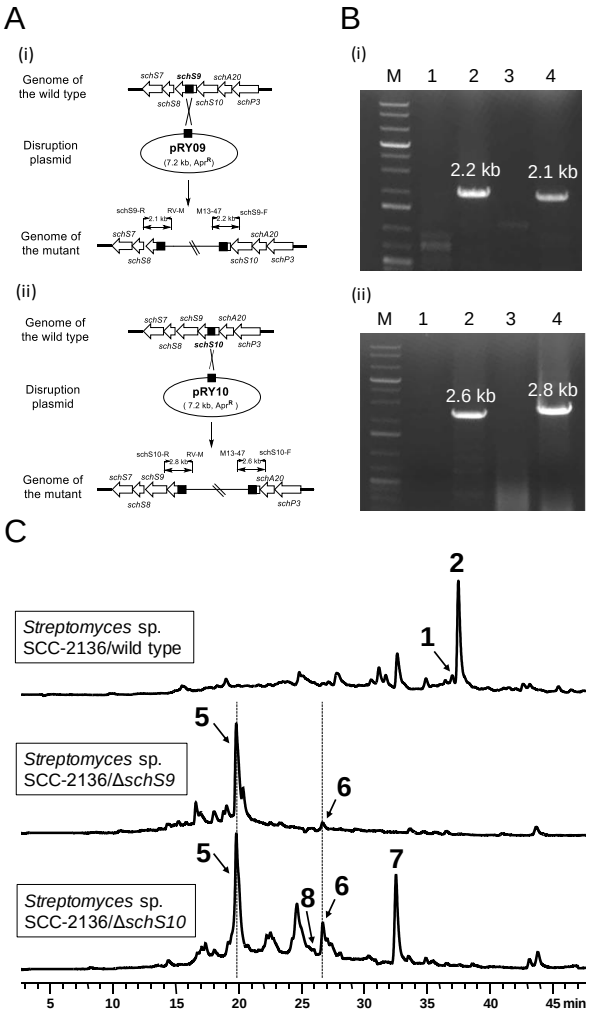


Figure 4. Gene disruption of *schS9* and *SchS10*. (A) Single crossover recombination strategy for disruption of *schS9* (i) and *SchS10* (ii) in *Streptomyces* sp. SCC-2136. (B) PCR confirmation of the *Streptomyces* sp. SCC-2136/Δ*schS9* (i) and *Streptomyces* sp. SCC-2136/Δ*schS1*0 (ii) single crossover mutants. (i) Confirmation of the insertion of pRY09 into the genome of *Streptomyces* sp. SCC-2136. M: 1 kb plus DNA ladder; 1: primers sch9-F/M13-47 with the wild type as the template; 2: primers sch9-F/M13-47 with the Δ*schS9* mutant as the template; 3: primers sch9-R/RV-M PCR with the wild type as the template; 4: primers sch9-R/RV-M PCR product with the Δ*schS9* mutant as the template. (ii) Confirmation of the insertion of pRY10 into the genome of *Streptomyces* sp. SCC-2136. M: 1 kb plus DNA ladder; 1: primers sch10-F/M13-47 PCR with the wild type as the template; 2: primers sch10-F/M13-47 with the Δ*schS10* mutant as the template; 3: primers sch10-R/RV-M with the wild type as the template; 4: primers sch10-R/RV-M with the Δ*schS10* mutant as the template. (C) HPLC traces (at 420 nm) for the cultures *Streptomyces* sp. SCC-2136/wild-type (**upper**), Δ*schS9* (**middle**) and Δ*schS10* (**bottom**).

(**7**) (Table 1 and Figure 3) to elucidate its chemical structure of this new compound. NMR data analysis for **5** and **6** confirmed the structure of previously reported compounds of urdamycin X and 3'-deoxyaquayamycin, respectively.[6,26] Urdamycin X was reported to be a surprising shunt product in *Streptomyces fradie*/Δ*urdQ/R*, while 3'-deoxyaquayamycin is a common *C*-glycosylated intermediate compound in angucycline biosynthesis.

Compounds **7** and **8**, on the other hand, are *O*-glycosylated intermediate compounds. HR-MS for **7** indicated a $[M-H]^-$ ion peak at 583.2559, which is consistent with the calculated exact mass of 583.2224 (Figure S7), indicating that the molecular formula of **7** is $C_{31}H_{36}O_{11}$. This was supported by the NMR data (Table 1) containing 31 carbon signals. The chemical shift of C-8 (δ 159.3) indicates a hydroxyl group attached to the aromatic ring. In addition, similar to **6**, **7** contains the tertiary hydroxyl groups at the C-4a (δ 132.1) and C-12b (δ 133.9) positions and an anomeric hydroxyl group at C-4' position (δ 72.14). As the molecular weight of **7** suggests, we found an additional anomeric proton signal (CH-1'') at δ_C 98.4 and δ_H 4.91, indicating that **7** has one more

Table 1. ^{1}H NMR data for **3** (in $CDCl_3$), **4** (in $(CD_3)_2CO$), **5** (in CD_3OD), **6** (in DMSO-d_6), and **7** (in $CDCl_3$) and ^{13}C NMR data for **4** and **7**.

Position	Rabelomycin (**3**)	GG31 (**4**)	Urdamycin X (**5**)	3'-Deoxyaquayamycin (**6**)	OZK1 (**7**)	GG31 (**4**)[a]	OZK1 (**7**)[a]
1	-	-	-	-	-	195.1	174.4
2	2.99 (d, 1H, 15.1) 3.07 (d, 1H, 15.1)	2.80 (d, 1H, 14.6) 2.95 (d, 1H, 14.6)	1.95 (dd, 1H, 13.8, 3.0) 2.33 (d, 1H, 13.8)	2.35 (d, 2H, 15.0)	2.62 (d, 2H, 14.8)	53.4	44.8
3	-	-	-	-	-	71.3	72.16
4	3.12 (s, 2H)	3.14 (d, 2H, 11.9)	2.20 (d, 1H, 10.5) 2.03 (dd, 1H, 10.5, 3.3)	2.90 (d, 1H, 13.0) 3.02 (d, 1H, 13.0)	3.08 (m, 2H)	44.0	41.2
4a	-	-	-	-	-	129.7	132.1
5	7.04 (s, 1H)	7.08 (s, 1H)	1.67 (m, 1H) 2.11 (m, 1H)	7.81 (d, 1H, 7.8)	7.64 (d, 1H, 7.7)	121.5	139.5
6	-	-	2.61 (td, 1H, 12.1, 3.4) 2.85 (td, 1H, 11.9, 5.6)	7.72 (d, 1H, 7.7)	7.82 (d, 1H, 7.7)	152.3	119.1
6a	-	-	-	-	-	116.8	115.8
7	-	-	-	-	-	193.1	188.3
7a	-	-	-	-	-	114.8	115.3
8	-	-	-	-	-	157.9	159.3
9	7.30 (d, 1H, 2.6)	-	7.25 (dd, 1H, 6.3, 2.7)	-	-	134.5	139.8
10	7.72 (d, 1H, 4.7)	7.94 (d, 1H, 7.7)	7.59 (d, 1H, 7.7)	7.86 (d, 1H, 7.9)	7.92 (d, 1H,7.9)	134.1	133.4
11	7.70 (d, 1H, 4.8)	7.56 (d, 1H, 7.6)	7.58 (d, 1H, 7.6)	7.79 (d, 1H, 6.3)	7.85 (d, 1H, 7.9)	118.2	119.6
11a	-	-	-	-	-	138.1	131.5
12	-	-	-	-	-	182.9	188.0
12a	-	-	-	-	-	138.2	161.0
12b	-	-	-	-	-	128.6	133.9
13	1.52 (s, 3H)	1.47 (s, 3H)	1.54 (s, 3H)	1.16 (s, 3H)	1.34 (s, 3H)	29.6	27.2
1'	-	4.79 (d, 1H, 10.8)	-	4.70 (d, 1H, 10.6)	4.85 (d, 1H, 10.1)	73.0	73.2
2'	-	1.61 (m, 2H)	-	1.39 (m, 1H) 2.09 (m, 1H)	1.46 (m, 2H)	33.2	31.7
3'	-	2.14 (m, 2H)	-	1.55 (m, 1H) 2.02 (m, 1H)	2.27 (m, 2H)	32.0	31.8
4'	-	3.23 (m, 1H)	-	3.12 (m, 1H)	3.32 (m, 1H)	79.2	79.7
5'	-	3.40 (m, 1H)	-	3.33 (m, 1H)	3.55 (m, 1H)	71.16	77.6
6'	-	1.33 (d, 3H, 6.0)	-	1.26 (d, 3H, 6.0)	1.35 (d, 3H, 6.1)	18.2	18.5
1''	-	-	-	-	4.91 (s, 1H)	-	98.4
2''	-	-	-	-	1.90 (m, 2H)	-	27.7
3''	-	-	-	-	1.80 (m, 2H)	-	30.5
4''	-	-	-	-	3.31 (m, 1H)	-	72.14
5''	-	-	-	-	3.76 (m, 1H)	-	69.9
6''	-	-	-	-	1.27 (d, 3H, 6.2)	-	17.8

[a] ^{13}C NMR data for **4** and **7**.

amicetose sugar moiety. HBMC correlation of H-1'' to C-4' revealed the linkage of the second amicetose moiety to 4'-OH of **6**. Thus, **7** was characterized as L-amicetosyl-3'-deoxyaquayamycin. All the signals are assigned according to the 1D and 2D NMR spectra (Table 1). Based on our findings, we proposed the glycosylation steps in the biosynthesis of **1** and **2** as illustrated in Figure 5. Glycosylation occurs in a sequential manner in which SchS7 first *C*-glycosylates the polyketide aglycone, followed by the *O*-glycosylation of 3'-deoxyaquayamycin by SchS9. Lastly, SchS10 attaches the third sugar moiety to the C-3 position, leading to the production of **1** and **2**.

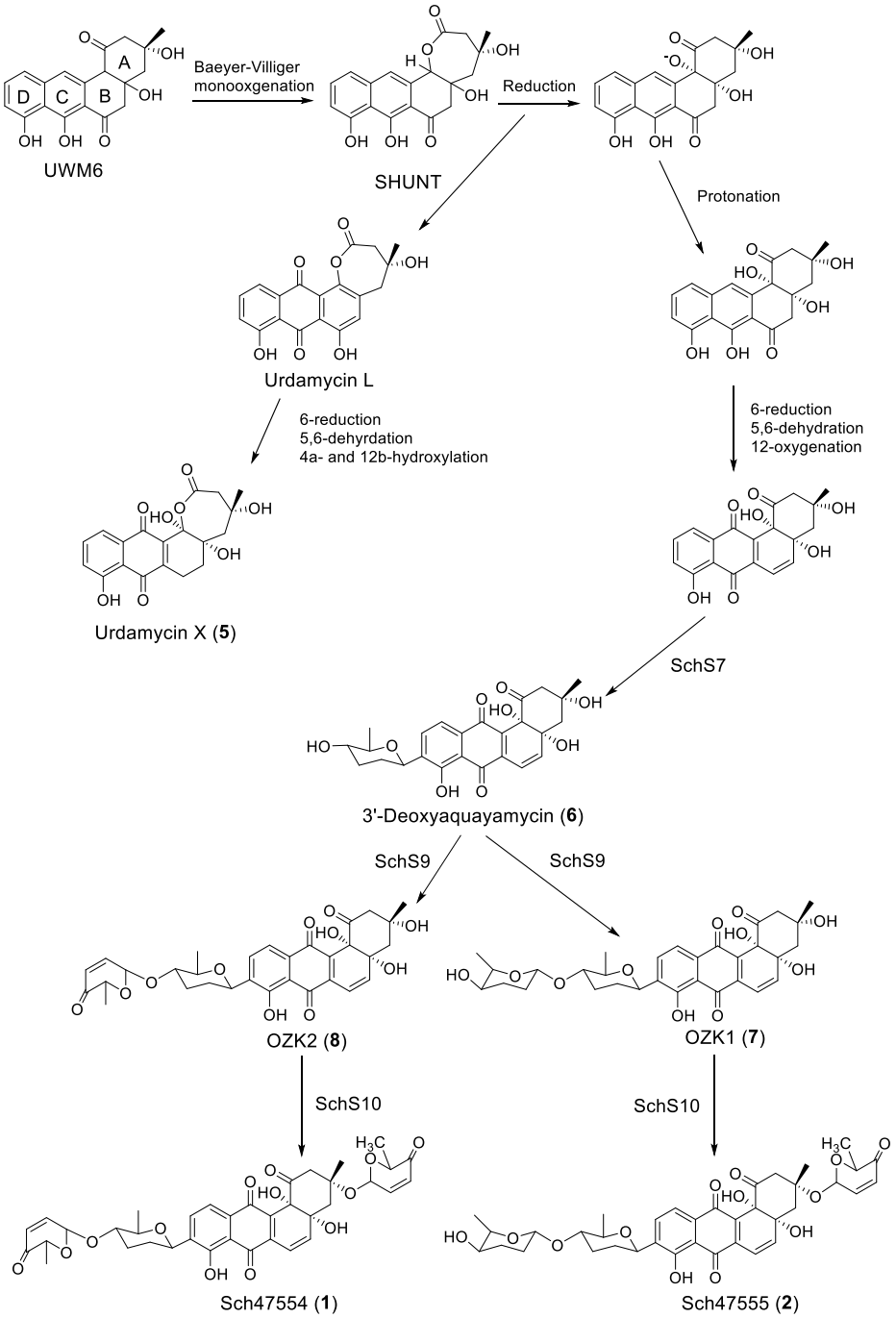


Figure 5. Proposed glycosylation steps in the biosynthesis of **1** and **2**.

Discussion

Many natural products are produced through the assembly of their aglycone core structure followed by the decoration with sugar moieties. The biological activities of many natural products including medicinally significant antibiotics and anticancer drugs depends on the regio- and stereo-specifically attached sugar moieties.[11,27] Chemical synthesis of sugar moieties is excessively difficult to perform and impractical to establish in large scale.[14] Thus, GTs are useful alternatives tools for chemoenzymatic and *in vivo* approaches to develop glycosylated natural products.[27] Many GT genes have been isolated from antibiotic-producing microorganisms and functionally characterized mostly by gene inactivation experiments and in a few cases by heterologous expression. Among these GTs are *lanGT1*, *lanGT2*, *lanGT3* and *lanGT4* in the landomycin gene cluster,[14,17,28,29] *urdGT1a*, *urdGT1b*, *urdGT1c* and *urdGT2* in the urdamycin gene cluster,[20,23,30,31] *saqGT1*, *saqGT2*, *saqGT3*, *saqGT4*, *saqGT5*, *saqGT6* in the saquayamycin Z gene cluster,[21] *mtmGI*, *mtmGII*, *mtmGIII* and *mtmGIV* in the mithraycin gene cluster,[32,33] *oleG1* and *oleG2* in the oleandomycin gene cluster,[34,35] and *dnrS* in the daunosamine gene cluster.[36]

1 and **2** are two antifungal angucyclines that possess three sugar moieties. However, how these sugars are added to the structures remains unknown. The *sch* gene cluster contains three putative GT genes (*schS7*, *schS9*, *schS10*).[9] Based on multiple sequence comparison of the deduced amino acid sequences of GTs in the *sch* biosynthetic gene cluster, SchS7 showed high homology to several known *C*-GTs including UrdGT2, SaqGT5, and SimB7. SaqGT5 attaches an activated D-olivose moiety directly to the C-9 position of angucycline in the case of saquayamycin Z and of tetracenequinone in the case of galtamycin B.[21] SimB7 and UrdGT2 are also responsible for the attachment of an activated D-olivose to C-9 of polyketide aglycone.[22,31] In particular, UrdGT2 attaches an activated D-olivose to C-9 of 2-hydro-3-hydroxy-prejadomycin (UWM6) or terangulol.[14,31] The examples of SaqGT5 and UrdGT2 show the substrate flexibility that will facilitate the synthesis of derivative compounds using those *C*-GTs. UrdGT2 also functions as *O*-GT and transfers the activated D-olivose sugar moiety to 1,2-dihydroxyanthraquinone.[20] Surprisingly, multiple sequence comparison also revealed high homology of SchS7 to a known *O*-GT, LanGT2 from *Streptomyces cyanogenus* which catalyzes the first glycosylation step by attaching activated D-olivose to 8-OH of the aglycone.[14] Interestingly, an engineered version of LanGT2 was able to act as *C*-GT, showing activity for both *C*- and *O*-glycosylation similar to UrdGT2.[20,37] Therefore, we expected SchS7 to be responsible for the attachment of D-amicetose to C-9 of the polyketide aglycone. To characterize its function, we have used a heterologous expression approach to create two expression plasmids, pOF10 and pGG31. pOF10 contains the minimal PKS genes (*schP6*, *schP7*, and *schP8*) and aromatase, ketoreductase, cyclase, as well as oxygenase genes (*schP4*, *schP5*, *schP9*, and *schP10*), while pGG31 additionally includes *schS7* and genes required for the biosynthesis of activated sugar moieties (*schS1*, *schS2*, *schS3*, *schS4*, *schS5*, *schS6*, and *schS8*). *S. lividans* K4 harboring pOF10 and pGG31 respectively produced two major peaks of rabelomycin (**3**) and a novel compound, D-amicetosyl-rabelomycin (**4**), proving that SchS7 is responsible for the attachment of a D-amicetose sugar moiety to the C-9 position of **3**.

Sequence analysis of SchS9 and SchS10 revealed that are homologous to known *O*-GTs such as UrdGT1a, UrdGT1c, SaqGT2, SaqGT4, LanGT1, and LanGT4. In particular, SchS9 is highly homologous to SaqGT4 and UrdGT1c which attach activated L-rhodinose to the first sugar moiety in the biosynthesis of saquayamycin Z and urdamycin A, respectively. UrdGT1c transfers a L-rhodinose sugar moiety to urdamycinone B, 100-2, and aquayamycin at 3'-OH, while SaqGT4 attaches the third and fifth sugar moieties at 3-OH of the saccharide chain.[21,23] This indicates the substrate flexibility of these enzymes and SchS9 might potentially recognize different substrates. Additionally, SchS9 has high homology to LanGT1 from the landomycin biosynthetic gene cluster that is responsible for the attachment of D-olivose to 4'-OH of the first sugar moiety.[17] For SchS10, multiple sequence alignment showed that it is highly homologous to SaqGT2 and UrdGT1a that are responsible for the attachment of L-rhodinose at 12b-OH and 3-OH, respectively.[21,23] The amino acid sequence of SchS10 is also similar to that of LanGT4 which, in contrast to SaqGT2 and UrdGT1a, attaches the third sugar moiety (L-rhodinose) to the saccharide chain at C-9.[17] Therefore, both SchS9 and SchS10 are expected to catalyze the *O*-glycosylation steps in Sch47554 and Sch47555 biosynthesis. To confirm their function, we disrupted the corresponding genes through targeted gene deletion. HPLC analysis of the extracts of *Streptomyces* sp. SCC-2136/Δ*schS9* and *Streptomyces* sp. SCC-2136/Δ*schS10* mutants revealed that both mutants produced **5** and **6** in common. *Streptomyces* sp. SCC-2136/Δ*schS10* produced two additional compounds (**7** and **8**) (Figure 4C). Except **8** which was only characterized with LC-MS due to the very low titer, all compounds were purified and characterized by LC-MS and NMR. Production of these compounds in the mutant strains indicates that *O*-glycosylation takes place in a sequential manner in which SchS9 first attaches either L-amicetose or L-aculose moiety to 4'-OH of **6**, then SchS10 transfers a L-aculose moiety to 3-OH of **7** or **8** to yield **1** and **2**. SchS9 has relatively relaxed substrate specificity toward sugar donors as shown by its ability to attach two different sugar moieties, while SchS10 is flexible with sugar acceptors and can recognize both **7** and **8** as substrate. **8** might be a more favorable substrate for SchS10 than **7** due to a higher titer of **1** compared to **2** in the wild type, whereas more efficient production of **7** than **8** in *Streptomyces* sp. SCC-2136/Δ*schS10* mutant indicates that SchS9 favorably attaches an activated L-amicetose moiety to 4'-OH of **6**. Additionally, having **6** in both *Streptomyces sp.* SCC-2136/Δ*schS9* and Δ*schS10* extracts suggests that SchS7 is the dedicated *C*-GT that catalyzes the first glycosylation with 9-deolivosyl aquayamycin as the sugar acceptor substrate. **3** and **4** were not detected in the wild type and mutant strains of *Streptomyces sp.* SCC-2136. The enzymes that are involved in the 6-reduction, 5,6-dehydration and 12-oxygenation may not recognize a glycoside substrate such as **4**. Production of **4** through heterologous expression in *S. lividans* indicated that SchS7 has relaxed specificity towards the sugar acceptors.

In addition, both mutant strains produced urdamycin X (**5**) which was reported as a shunt product. Similar to **5**, urdamycin L was detected in a Δ*urdM* mutant of *S. fradie* during Baeyer-Villiger monooxygenation process in urdamycin A biosynthesis.[25] **5** is different from urdamycin L due to the almost saturated ring B

(additional OH groups at C-4a and C-12b), whereas urdamycin L possesses a phenolic arene ring B with an OH group at C-6 (Figure 5).[26] However, both compounds are related to each other by having an unusual ε-lactone as ring A. Urdamycin L is derived from Baeyer-Villiger monooxygenation reaction through spontaneous oxidation, while **5** is an over-oxidized analogue of urdamycin L.[25,26] In the *sch* biosynthetic gene cluster, there exist oxygenases such as *schP3* and *schP10* with high homology to *urdM* (64%) and *urdE* (84%), respectively. These enzymes might be in the oxygenation of the angucycline ring. It is surprising that **5**, as a shunt product, was produced in larger quantities than the biosynthetic intermediate **6** in the *schS9* and *schS10* disruption mutants (Figure 4C). Because *schS7* is located downstream of *schS9* and *schS10*, there is a possibility that a polar effect occurred when the disruption plasmids were inserted into the genome. This polar effect may cause a low level expression of the *C*-GT SchS7 and thus the overall attenuation of the biosynthetic route of **6**, which in turn led to the increased accumulation of **5**.

Conclusions

We used two different approaches to characterize the function of three GTs in *sch* biosynthetic gene cluster; heterologous expression for SchS7 and targeted gene disruption for SchS9 and SchS10. As expected from multiple sequence analysis of these GTs with their homologues, we found that the glycosylation process in the biosynthesis of Sch47554 and Sch47555 occurs in a sequential manner. First, SchS7, which is C-GT, attaches a D-amicetose moiety to polyketide aglycone followed by the attachment of the second sugar moiety to 4'-OH of the first sugar moiety by SchS9. In the last step, SchS10 transfers an L-aculose to C-3 of glycosylated polyketide aglycone, leading to the biosynthesis of major products of Sch47554 and Sch47555. We isolated two novel compounds, GG31 and OZK1 from these two different approaches, respectively, diversifying the pharmaceutically relevant compounds. With elucidating the function of three GTs, we may combine them with other GTs to change the saccharide chain that possibly enhances the antifungal activity as well as diversify further the pharmaceutically relevant antifungal compounds.

Experimental Section

Bacterial strains, media and culture conditions

Streptomyces lividans K4 strains harboring pOF10 and pGG31 were maintained on R5 agar plates supplemented with thiostrepton at 28 °C for 10 days. *Streptomyces* sp. SCC-2136/Δ*schS9* and *Streptomyces* sp. SCC-2136/Δ*schS10* mutants were grown on YM agar plates supplemented with apramycin at 28 °C for 10 days. *Streptomyces* sp. SCC-2136 was grown in YM broth for 5 days for genomic DNA extraction. *E. coli* XL1-Blue and ET12567, which were grown in LB media at 37 °C overnight, were used for general genetic manipulations and intergenic conjugation, respectively. MS and ISP4 media were used for conjugation and single crossover homologous recombination, respectively. Apramycin (50 μg/mL), thiostrepton (50 μg/mL), chloramphenicol (25 μg/mL), ampicillin (50 μg/mL) and nalidixic acid (25 μg/mL) were supplemented when appropriate.

General genetic manipulation and PCR

Standard molecular biology protocols were performed as previously described.[38] Genomic DNA of *Streptomyces* sp. SCC-2136 and its mutants were extracted by ZR Fungal/Bacterial DNA Miniprep Kit (Irvine, CA, USA). Plasmid DNA extraction from *E. coli* cells was performed using Thermo Scientific GeneJET Plasmid Miniprep Kit (Logan, UT, USA). PCR reactions were performed with Arktik™ Thermal Cycler (Thermo Scientific) using Phusion DNA polymerase (Thermo Scientific, Logan, UT, USA). Oligonucleotide primers were ordered from Sigma-Aldrich and dissolved in TE buffer to the concentration of 100 ng/mL.

Construction of plasmids for heterologous expression and gene disruption

For the heterologous expression approach, *schP4*, *schP5*, *schP6*, *schP7*, *schP8*, *schP9* and *schP10* genes were PCR-amplified from *Streptomyces* sp. SCC-2136 genomic DNA using oligonucleotide primers listed in Table S1. Specific forward and reverse primers contain compatible restriction enzymes such as XbaI, SpeI, and NheI. The PCR products were first ligated into the pJET1.2 cloning vector for DNA sequencing using the Sanger method (Eton Bioscience, San Diego, CA, USA). Each combination of genes was then constructed by the subsequent addition of individual genes into the cloning vector and the correct ligation was confirmed by restriction enzyme digestions. Additionally, each group of genes was ligated into a pRM5-derived vector. The gene cassette containing *schP6*, *schP7,* and *schP8* was excised from pOF1 by using PacI and NheI and ligated into pRM5 between the same sites to yield pOF3. Subsequently, *schP5* gene was excised from pOF2 using NheI and SpeI and ligated into pOF3 between the same sites to yield pOF4. *schP4*, *schP9* and *schP10* were also excised from pOF5-7 and inserted in the same sequential manner to yield pOF8-10, respectively. Similarly, *schS1*, *schS2*, *schS3*, *schS4*, *schS5*, *schS6*, *schS7* and *schS8* were PCR-amplified from the genomic DNA of *Streptomyces* sp. SCC-2136 and subsequently inserted into pJET1.2 to yield pKN57. The gene cassette containing these genes was excised from pKN57 using the NheI and SpeI restriction sites and ligated into pOF10 digested with the same restriction enzymes to yield pGG31. pOF10 and pGG31 plasmids were introduced into *Streptomyces lividans* K4 through PEG-mediated protoplast transformation for product analysis.

For the targeted gene disruption approach, a 672 bp fragment of *schS9* and a 763 bp fragment of *schS10* were PCR-amplified from the genomic DNA of *Streptomyces* sp. SCC-2136 using the primers listed in Table S1. The PCR products for *schS9* and *schS10* were first ligated into the pJET1.2 cloning vector using T4 DNA ligase (New England Biolabs) to yield pRY7 and pRY8, respectively. These plasmids were sequenced to ensure that the genes were correct and free of mutations. The inserts were excised from pRY7 and pRY8 with XbaI and HindIII (New England Biolabs) and ligated to the thermal sensitive plasmid pKC1139 between the same sites to yield pRY9 and pRY10, respectively, which were subsequently used for gene inactivation of *schS9* and *schS10* in *Streptomyces* sp. SCC-2136 through intergenic conjugation using *E. coli* ET12567. Positive exconjugants for *Streptomyces* sp. SCC-2136/Δ*schS9* and Δ*schS10* mutants were transferred to ISP4 plates supplemented with apramycin and NC, and incubated at 37 °C for about 10 days to allow the plasmid to integrate into the genome and yield the single crossover mutants. The positive mutants were shuttled back and forth to YM and ISP4 plates supplemented with apramycin for two generations to guarantee the positive mutants. Positive exconjugants after genome integration were subjected to product analysis and PCR confirmation of gene knockouts using primers listed in Table S1. The details of PEG-mediated protoplast transformation and intergenic conjugation protocols are described in literature.[39]

Production, extraction, purification and characterization of compounds

S. lividans K4 harboring pOF10 and pGG31, and *Streptomyces* sp. SCC-2136/Δ*schS9* and *Streptomyces* sp. SCC-2136/Δ*schS10* mutants were grown for 10 days at 28 °C on R5 and YM agar plates supplemented with appropriate antibiotics, respectively. The cultures were chopped and extracted with a mixture of solvent consisting of 89% ethyl acetate, 10% methanol, and 1% acetic acid (v/v). The resulting extracts were dried *in vacuo*, and the residues were redissolved in methanol for LC-MS analysis. All crude extracts were separated on a silica gel 60 column using different ratios of chloroform–methanol. The fractions with target compounds were further separated by HPLC with an Agilent Eclipse XDB-C18 column (5 µm, 250 mm × 4.6 mm). Low-resolution ESI-MS spectra were obtained on an Agilent 6130 LC-MS. Waters GCT high resolution mass spectrometer, with EI/CI and LIFDI capabilities (University of California, Riverside) was used to obtain the accurate mass spectra for the new compounds. 1D and 2D NMR spectra were recorded on a JEOL ECX-300 or a Bruker AvanceIII HD Ascend-500 NMR instrument. The chemical shift (δ) values and coupling constants (*J* values) are reported in parts per million (ppm) and hertz (Hz), respectively.

Acknowledgements

This work was supported by the National Institutes of Health Grant AI065357 RM DP 008. The Bruker AvanceIII HD Ascend-500 NMR instrument used in this work was funded by the National Science Foundation Award CHE-1429195. R. Y. was supported by the Young Teacher Development Plan Visiting Scholar Special Funds of Jiangxi Province of China.

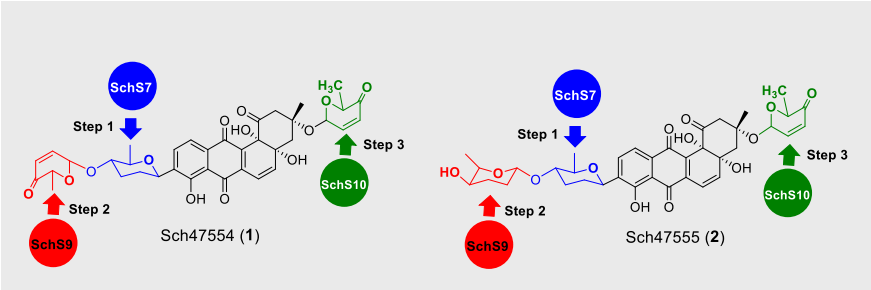
Keywords: Glycosylation • Biosynthesis • Sch47554 • Sch47555 • *Streptomyces* sp. SCC-2136

[1] J. Rohr, R. Thiericke, *Nat. Prod. Rep.* **1992**, *9*, 103–137.
[2] B. J. Rawlings, *Nat. Prod. Rep.* **1999**, *16*, 425–484.
[3] T. Henkel, T. Ciesiolka, J. Rohr, A. Zeeck, *J. Antibiot.* **1989**, *42*, 299–311.
[4] T. Henkel, J. Rohr, J. M. Beale, L. Schwenen, *J. Antibiot.* **1990**, *43*, 492–503.
[5] S. Grabley, P. Hammann, K. Hutter, H. Kluge, R. Thiericke, J. Wink, A. Zeeck, *J. Antibiot.* **1991**, *44*, 670–673.
[6] M. Chu, R. Yarborough, J. Schwartz, M. G. Patel, A. C. Horan, V. P. Gullo, P. R. Das, M. S. Puar, *J. Antibiot.* **1993**, *46*, 861–865.
[7] Y. Song, G. Liu, J. Li, H. Huang, X. Zhang, H. Zhang, J. Ju, *Mar. Drugs* **2015**, *13*, 1304–1316.
[8] Z. Hu, L. Qin, Q. Wang, W. Ding, Z. Chen, Z. Ma, *Nat. Prod. Res.* **2016**, 1–8.
[9] D. B. Basnet, T.-J. Oh, T. T. H. Vu, B. Sthapit, K. Liou, H. C. Lee, J.-C. Yoo, J. K. Sohng, *Mol. Cells* **2006**, *22*, 154–162.
[10] T. Vogt, P. Jones, *Trends Plant Sci.* **2000**, *5*, 380–386.
[11] V. Kren, L. Martinkova, *Curr. Med. Chem.* **2001**, *8*, 1303–1328.
[12] A. C. Weymouth-Wilson, *Nat. Prod. Rep.* **1997**, *14*, 99.
[13] B. Ostash, U. Rix, L. L. R. Rix, T. Liu, F. Lombo, A. Luzhetskyy, O. Gromyko, C. Wang, A. F. Braña, C. Méndez, et al., *Chem. Biol.* **2004**, *11*, 547–555.
[14] A. Luzhetskyy, T. Taguchi, M. Fedoryshyn, C. Dürr, S.-E. Wohlert, V. Novikov, A. Bechthold, *ChemBioChem* **2005**, *6*, 1406–1410.
[15] L. Rodriguez, C. Oelkers, I. Aguirrezabalaga, A. F. Braña, J. Rohr, C. Méndez, J. A. Salas, *J. Mol. Microbiol. Biotechnol.* **2000**, *2*, 271–6.
[16] L. Tang, R. McDaniel, *Chem. Biol.* **2001**, *8*, 547–555.
[17] A. Trefzer, C. Fischer, S. Stockert, L. Westrich, E. Künzel, U. Girreser, J. Rohr, A. Bechthold, *Chem. Biol.* **2001**, *8*, 1239–1252.
[18] E. Künzel, B. Faust, C. Oelkers, U. Weissbach, Daniel W. Bearden, G. Weitnauer, L. Westrich, A. Andreas Bechthold,, Jürgen Rohr, *J. Am. Chem. Soc.* **1999**, *121*, 11058–11062.
[19] A. Luzhetskyy, L. Zhu, M. Gibson, M. Fedoryshyn, C. Dürr, C. Hofmann, D. Hoffmeister, B. Ostash, C. Mattingly, V. Adams, et al., *ChemBioChem* **2005**, *6*, 675–678.
[20] C. Dürr, D. Hoffmeister, S.-E. Wohlert, K. Ichinose, M. Weber, U. von Mulert, J. S. Thorson, A. Bechthold, *Angew. Chemie Int. Ed.* **2004**, *43*, 2962–2965.
[21] A. Erb, A. Luzhetskyy, U. Hardter, A. Bechthold, *ChemBioChem* **2009**, *10*, 1392–1401.
[22] A. Trefzer, S. Pelzer, J. Schimana, S. Stockert, C. Bihlmaier, H.-P. Fiedler, K. Welzel, A. Vente, A. Bechthold, *Antimicrob. Agents Chemother.* **2002**, *46*, 1174–82.
[23] A. Trefzer, D. Hoffmeister, E. Künzel, S. Stockert, G. Weitnauer, L. Westrich, U. Rix, J. Fuchser, K. Bindseil, J. Rohr, et al., *Chem. Biol.* **2000**, *7*, 133–142.
[24] W. Zhang, B. D. Ames, S.-C. Tsai, Y. Tang, *Appl. Environ. Microbiol.* **2006**, *72*, 2573–80.
[25] U. Rix, L. L. Remsing, D. Hoffmeister, A. Bechthold, J. Rohr, *ChemBioChem* **2003**, *4*, 109–111.
[26] M. Fedoryshyn, M. Nur-e-Alam, L. Zhu, A. Luzhetskyy, J. Rohr, A. Bechthold, *J. Biotechnol.* **2007**, *130*, 32–8.
[27] J. Thorson, T. Hosted Jr., J. Jiang, J. Biggins, J. Ahlert, *Curr. Org. Chem.* **2001**, *5*, 139–167.
[28] L. Westrich, S. Domann, B. Faust, D. Bedford, D. A. Hopwood, A. Bechthold, *FEMS Microbiol. Lett.* **1999**, *170*, 381–387.
[29] A. Luzhetskyy, T. Liu, M. Fedoryshyn, B. Ostash, V. Fedorenko, J. Rohr, A. Bechthold, *ChemBioChem* **2004**, *5*, 1567–1570.
[30] E. Künzel, G. Weitnauer, H. Decker, A. Bechthold, D. Hoffmeister, S. Haag, L. Westrich, P. Schneider, B. Faust, J. Rohr, *Microbiology* **2000**, *146*, 147–154.
[31] I. Baig, M. Kharel, A. Kobylyanskyy, L. Zhu, Y. Rebets, B. Ostash, A. Luzhetskyy, A. Bechthold, V. A. Fedorenko, J. Rohr, *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 7842–6.
[32] E. Fernández, U. Weissbach, C. Sánchez Reillo, A. F. Braña, C. Méndez, J. Rohr, J. A. Salas, *J. Bacteriol.* **1998**, *180*, 4929–37.
[33] G. Blanco, E. Fernández, M. J. Fernández, A. F. Braña, U. Weissbach, E. Künzel, J. Rohr, C. Méndez, J. A. Salas, *MGG - Mol. Gen. Genet.* **2000**, *262*, 991–1000.
[34] C. Olano, A. M. Rodriguez, J.-M. Michel, C. Méndez, M.-C. Raynal, J. A. Salas, *Mol. Gen. Genet. MGG* **1998**, *259*, 299–308.
[35] M. Doumith, R. Legrand, C. Lang, J. A. Salas, M.-C. Raynal, *Mol. Microbiol.* **1999**, *34*, 1039–1048.
[36] S. L. Otten, X. Liu, J. Ferguson, C. R. Hutchinson, *J. Bacteriol.* **1995**, *177*, 6688–92.
[37] H. K. Tam, J. Härle, S. Gerhardt, J. Rohr, G. Wang, J. S. Thorson, A. Bigot, M. Lutterbeck, W. Seiche, B. Breit, et al., *Angew. Chemie Int. Ed.* **2015**, *54*, 2811–2815.
[38] S. Joseph, W. R. David, *Cold Spring Harb. Lab. Press. Cold Spring Harb. New York* **2001**, *2*.
[39] T. Kieser, M. J. Bibb, K. F. Chater, D. A. Hopwood, *Practical Streptomyces Genetics*, John Innes Foundation, Norwich, **2000**.

Entry for the Table of Contents (Please choose one layout)

Layout 2:

FULL PAPER



Ozkan Fidan Riming Yan, Gabrielle Gladstone, Tong Zhou, Du Zhu, Jixun Zhan**

Page No. – Page No.

New insights into the glycosylation steps in the biosynthesis of Sch47554 and Sch47555

The roles of three glycosyltransferases in the biosynthesis of Sch47554 and Sch47555 in *Streptomyces* sp. SCC-2136 were characterized. SchS7 is a C-glycosyltransferase that attaches a D-amicetose moiety to the aglycone core structure at the C-9 position. SchS9 then attaches either a L-aculose or L-amicetose moiety to the 4'-OH of the C-glycosylated aglycone, and SchS10 finally transfers a L-aculose moiety to the 3-OH of the angucycline core.