Supplementary Information

Title: Complete characterization of the Seventeen Step moenomycin biosynthetic pathway

C. Bohdan Ostash,[‡,Π,1] Emma Doud,[‡,ξ,1] Cecilie Lin,[‡,ξ] Iryna Ostash,[‡,Π] Deborah Perlstein,[‡] Shinichiro Fuse,[ξ,⊥] Manuel Wolpert,[‡] Daniel Kahne,[ξ] Suzanne Walker[‡,*]

[‡]Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA 02115
[ξ]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138
[Π] current address: Department of Genetics and Biotechnology, Ivan Franko National University of L'viv, Grushevskogo st. 4, L'viv 79005, Ukraine
[⊥] current address: Department of Applied Chemistry, Graduate School of Science and Engineering, Tokyo Institute of Technology, 2-12-1, Ookayama, Meguro, Tokyo 152-8552, Japan

[1] E.D. and B.O. contributed equally to this work.

[*]**Corresponding author:**
Suzanne Walker
Department of Microbiology and Molecular Genetics
Harvard Medical School
200 Longwood Avenue
Boston, MA 02115
Phone: 617-432-5488
Fax: 617-738-7664
Email: suzanne_walker@hms.harvard.edu

Table S1. Oligonucleotides Used in This Study

Name	Sequence (5Ō to 3Ō)	Notes
moeGT5up_P1	GCAGTGCGACGCGAGCGCACGAGCAGACGTCGTCATGTGTAG GCTGGAGCTGCTTC	to delete *moeGT5* in cosmid moeno38-1 (l-RED system)
moeGT5rp_P2	TCGGGGTGACCTCGTGTGTCAGCGCCCGGCGGCCGCTCCATA TGAATATCCTCCTTAG	
moeGT2up_P1	CGAGGAGCCCGCCGCGGGAGCGGCCGCCGGGCGCTGACAG TGTAGGCTGGAGCTGCTTC	to delete *moeGT2* in cosmid moeno38-1 (l-RED system)
moeGT2rp_P2	GCCGAGGTGCCGTCCACGCCGTTCCCCCTCCGTCGGCTACAT ATGAATATCCTCCTTAG	
moeGT2up_check	ACGAGGGGGACTTCCGCAG	to verify *moeGT2* deletion
38start_KD4	CGTGCGCAGCGCGGTCTTCGGCTTCGACGGGGTACGGATGAA TATCCTCCTTAGTTC	to delete *moeB5moeA5* in moeno38-1 (l-RED system)
moeA5_P3	AGACGCGCCGGGCGGCCCCCAGTTCGGACCAGATGCCGTAG GCTGGAGCTGCTTCG	
P2_KD4	CATATGAATATCCTCCTTAGTTC	Primer homologous to P2 site of pKD4[3]
alsrev1	AAATCTAGATCAAGAGCGGCCGGGGTC	to probe moeA5B5 deletion
moeF5up_P3	CGGCTCCTCGGTGTCCGTGCCGCGGCTGTAGGCGGCATGTAG GCTGGAGCTGCTTC	to delete *moeF5* in cosmid moeno38-1 (l-RED system)
moeF5rp_P2	TGGACGAGCGGTCGGTCGGGGGCAGCCATGGGTCTCCTACAT ATGAATCTCCTTAGTTC	
F5check_up	GTCTCGGTCAACGAAGTGGTC	to verify *moeF5* deletion
F5_check_rp	CTCTCCAGGGAGATGGTCCG	
moeGT4up_P3	TGCACAGCCTGTACCGGTCGACCTCCAACACCGACCGTGTAG GCTGGAGCTGCTT	to delete *moeGT4* in cosmid moeno38-1 (l-RED system)
moeGT4rp_P2	TCAGCTCTCCTGACGCGTGGGTGAGGACGACGGAGTGAGCAT ATGAATCTCCTTAGTTC	
moeK5-P1	TCCAGAAGCGGGCCGGCGTGCTGCCGCACCTCGGGGCTGTA GGCTGGAGCTGCTTCG	to delete *moeK5* in cosmid moeno38-1 (l-RED system)
moeK5-P2	TGTGCAGGCCGTCCAGCGTGTTGCGCCACTGGCCGGTCATAT GAATATCCTCCTTAG	
GT2con72up	AAAAAGCTTGTTCTGCGACGCGGACGAC	to amplify 0.5 kb internal moeGT3 fragment
GT2con73rp	AAAGAATTCAGGTTCGGAACGTGCAGCA	
moeH5up_P1	AGGCCGCCCTCCAGCCCCTGCTGGACGCCCGATGACGGTGTA GGCTGGAGCTGCTTC	to delete *moeH5* in cosmid moeno38-1 (l-RED system)
moeH5rp_P2	TCTCGTGAAGTGGGGGTCTGCGGCGGTCCGGCCCCGCTACAT ATGAATATCCTCCTTAG	
moeN5up_P1	CCGGCCACGGCCCTGCCGGCGGACTACACGGAGACCATGTAG GCTGGAGCTGCTTCG	to delete *moeN5* in cosmid moeno38-1 (l-RED system)
moeN5rp_P2	GGACGGCCGGCCGGAGACGCTCCGGCCGGCCGTCGGTCATA TGAATATCCTCCTTAG	
MoeE52P1	TTAGCGGACGGCGACGTTTTGGTCACCGGTGCGGCCTGTAGG CTGGAGCTGCTTCG	to delete an internal fragment of *moeE5* in cosmid moeno38-1 (-RED system)
MoeE52P2	TAACGCGGCCAAGRGGAAGACCGTCGAGATGTTTTTCATATGA ATATCCTCCTTAG	
moeGT3intMfeI	AAACAATTGTTCTGCGACGCGGACGAC	to amplify *moeR5* and *moeS5* along with P*moeS5*
orf1intXbaI	AAATCTAGAGGACTCTGCACCCTGAC	
moeR5XbaIup	AAATCTAGACGCGATGAACCGTCACG	to amplify *moeR5*
moeGT3XbaIup	AAATCTAGACGTGCCCTTCGACGACCCG	to amplify *moeGT3* along with its RBS
moeGT3EcoRIrp	AAAGAATTCCCACGCCCTGGTCCTGGAC	
moeN5XbaIup	AAATCTAGACAGGTCACCGAGTACCTCGA	to amplify *moeN5* along with its RBS
moeN5EcoRIrp	AAAGAATTCCGCTGATCAACACGTCGCTC	

moeF5XbaIup	AAATCTAGACACCCAGATCGAGTGGACC	to amplify *moeF5* along with its RBS
moeF5EcoRIrp	AAAGAATTCATGGGTCTCCTAGGAGAG	
moeGT4XbaIup	AAATCTAGAGTACCGCTCCTTCTTCATGC	to amplify *moeGT4* along with its RBS
moeGT4EcoRIrp	AAAGAATTCAGTGGAGCGACAGTACCTTC	
moeH5XbaIup	AAATCTAGACTGGACCAGGACGCGGTG	to amplify *moeH5* along with its RBS
moeH5EcoRIrp	AAAGAATTCGCTGATGTCTCGTGAAGTGG	
moeGT5XbaIup	AAATCTAGAGGGACCGGACTCGGACGT	to amplify *moeGT5* along with P*moeGT5*
moeGT5EcoRIrp	AAAGAATTCGGTGACCTCGTGTGTCAGC	
moeGT2XbaIup	AAATCTAGAAGGGCCTGCACTTCACCT	to amplify *moeGT2* along with its RBS
moeGT2EcoRIrp	AAAGAATTCGCCGTCCGGATCGACCA	
moeK5XbaIup	AAATCTAGATCCAGCGTGTTGCGC	to amplify *moeK5* along with its RBS
moeK5EcoRIrp	AAAGAATTCACGAGACATCAGCCG	
moeO5HinDIIIup	AAAAAGCTTCGGGGCGTGCCTTCTTC	to amplify *moeO5* along with P*moeO5*
moeO5XbaIrp	AAATCTAGACCGCCCGCTCCCCGGAC	
moe2HindIII_up	AAAAAGCTTGACGTGAGGCCGTGCAGTTC	to amplify *moe* cluster 2
moe2MfeI_rp	AAACAATTGGCACATCTTGACGTGGACGG	
moeGT1Nde1up	AAACATATGGCTGCCCCCGACCGAC	to amplify *moeGT1*
moeGT1Not1rp	ATAAAGCGGCCGCTCGGGCGTC	
MoeE5BamHI	AAAAAGGATCCGGTGTCGAGCGATACACACGG	to amplify *moeE5*
MoeE5XhoI	AAAAACTCGAGCTAC-AGCCGCGGCACGGAC	

Table S2. Strains and plasmids

Strain or plasmid	Relevant characteristics	Source or reference[a]
Escherichia coli		
DH5α	general cloning host	Novagen
XL1 Blue MR	general cloning host	Stratagene
ET12567 (pUB307)	methylation deficient strain for intergeneric conjugation	C.P. Smith (University of Manchester, U.K.) (Flett)
BW25113 (pIJ790)	strain for λ-RED mediated recombination	John Innes Center (Norwich, U.K.)
Bacillus cereus ATCC19637	MmA sensitive bioassay strain	ATCC
Streptomyces ghanaensis		
ATCC14672	wild type	ATCC
MO12	*moeGT3* disruption in ATCC14672	Ostash, 2007
Streptomyces lividans TK24	wild type	M. Kobayashi (University of Tsukbua, Japan)

Strain or plasmid	Relevant characteristics	Source or reference[a]
Escherichia coli		
DH5α	general cloning host	Novagen
XL1 Blue MR	general cloning host	Stratagene
ET12567 (pUB307)	methylation deficient strain for intergeneric conjugation	C.P. Smith (University of Manchester, U.K.) (Flett)
BW25113 (pIJ790)	strain for λ-RED mediated recombination	John Innes Center (Norwich, U.K.)
Bacillus cereus ATCC19637	MmA sensitive bioassay strain	ATCC
Streptomyces ghanaensis		
ATCC14672	wild type	ATCC
MO12	*moeGT3* disruption in ATCC14672	Ostash, 2007
Streptomyces lividans TK24	wild type	M. Kobayashi (University of Tsukbua, Japan)
Plasmids		
pSET152	Apr^R integrative cloning vector, *aac(3)IV* gene; apramycin resistance gene; int gene; integrase; lacZ	C.P. Smith (Manchester University, U.K.) (Practical Streptomyces Genetics)
pSOK804	$Apra^R$ *Streptomyces* integrative vector	S. Zotchev (Norwegian University of Science and Technology, Trondheim, Norway) (Sekorova)
pKC1139	AprR cloning vector with a temperature-sensitive pSG5 replicon	C.P. Smith (Machester University, U.K.) (Practical Streptomyces Genetics)
pMK19	pKC1139 digested with *HindIII/XbaI* and ligated with the *ermE* promoter	I. Ostast (L'viv University, Ukraine)
pHP45	source of Spectinomycin resistance cassette	J.-L. Pernodet (Université Paris-Sud , France) (BlondeletR)
pKD4	source of Kan^R casette for λ-RED mediated recombination	J. Beckwith (Harvard Medical School, USA) (Datsenko)
pCP20	FLP recombinase plasmid	J. Beckwith (Harvard Medical School, USA) (Datsenko)
pAF/urdR	ori^{pIJ101} *bla tsr*, P*ermE**, 6His tag	A. Bechthold (Freiburg University, Germany)
pTB146	ori^{ColEI} *bla* $lacI^q$ P_{T7}::h-sumo	T. Bernhardt (Harvard Medical School)
pTBsyn05	pTB146 digested with SapI and XmaI and ligated to *synO5*	this study
pED2	pAF/urdR digested with *HindIII/XbaI* and ligated to *moeO5* PCR product amplified from cosmid moeno38-1	this study
pED4	pMK19 digested with *Xba1/EcoRI* and ligated to *moeK5* PCR product amplified from cosmid moeno38-1	this study

pET48b-MoeE5	pET48b digested with *BamHI/XhoI* and ligated to *moeE5* PCR product amplified from cosmid moeno38-1	this study
pET24b-MoeGT1	pET24b digested with *NdeI/NotI* and ligated to *moeGT1* PCR product amplified from cosmid moeno38-1	this study
pMO12	pKC1139 digested with *HindIII/EcoRI* and ligated to an internal fragment of *moeGT3* PCR product amplified from *S. ghanaensis* genomic DNA	this study
pMO13	pMK19 digested with *Xba1/EcoRI* and ligated to *moeGT3* PCR product amplified from cosmid moeno38-1	this study
pMO14	pMO13 digested with *HindIII/EcoRI* to release P*ermE-moeGT3* fusion, blunt ended with Klenow enzyme, and ligated to pOOB40 digested with *EcoRV*	this study
pMO17	pMK19 digested with *Xba1/EcoRI* and ligated to *moeN5* PCR product amplified from cosmid moeno38-1	this study
pOOB40	pSET152 with the central portion of *aac(3)IV* excised with *SacI*, treated with Klenow fragment, and ligated to *EcoRV* digested *hyg* casette	Ostash 2007
pOOB46e	pMK19 digested with *Xba1/EcoRI* and ligated to *moeB4* fragment from pOOB12	this study
pOOB48a	pMK19 digested with *Xba1/EcoRI* and ligated to *moeF5* PCR product amplified from cosmid moeno38-1	this study
pOOB49f	pMK19 digested with *Xba1/EcoRI* and ligated to *moeR5moeS5* PCR product amplified from cosmid moeno38-1	this study
pOOB50	pMK19 digested with *Xba1/EcoRI* and ligated to *moeGT4* PCR product amplified from cosmid moeno38-1	this study
pOOB51	pMK19 digested with *Xba1/EcoRI* and ligated to *moeH5* PCR product amplified from cosmid moeno38-1	this study
pOOB52	pMK19 digested with *Xba1/EcoRI* and ligated to *moeGT5* PCR product amplified from cosmid moeno38-1	this study
pOOB55	pMK19 digested with *Xba1/EcoRI* and ligated to P*moeS5-moeS5* fragment from pOOB49f	this study
pOOB56c	pMK19 digested with *Xba1/EcoRI* and ligated to *moeGT2* PCR product amplified from cosmid moeno38-1	this study
pOOB58	spectinomycin resistance gene *aadA* retrieved as a *DraI* fragment from pHP45 and ligated to pMO14 digested with *XhoI,* treated with Klenow fragment	this study
pOOB59	pMK19 digested with *Xba1/EcoRI* and ligated to *moeR5* PCR product amplified from cosmid moeno38-1	this study

	pOOB63a	pSOK18 digested with *HindIII/EcoRI* and ligated to *moeO5moeN5* PCR product amplified from cosmid moeno38-1	this study
	pOOB64b	pSOK18 digested with *HindIII/Eco*RI and ligated to *moeA4moeB4moeC4* PCR product amplified from *S. ghanaensis* genomic DNA	this study
Cosmids			
	SuperCos1	AmpR *E. coli* cloning vector, *neo* gene, pUC^{ori}	Stratagene
	SuperCos1Hy	SuperCos1 where *neo* was replaced via λ-Red recombination by $Hyg\text{-}oriT^{RK2}\text{-}int^{\cdot C31}$ PCR product amplified from pOOB40	this study, Ostash 2007
	moeno38-1	contains *moe* cluster 1 except *moeR5moeS5*	this study, Ostash 2007
	moeno38-5	moeno38-1 *ΔmoeA4ΔmoeB4,* genes replaced by Kan^R casette	this study
	moeno38-7	moeno38-1 *ΔmoeN5,* gene replaced by Kan^R casette	this study
	moeno38-9	moeno38-1 *ΔmoeGT5,* gene replaced by Kan^R casette	this study
	moeno38-91	moeno38-1 *ΔmoeGT5*	this study
	moeno38-31	moeno38-1 *ΔmoeGT4*	this study
	moeno38-81	moeno38-1 *ΔmoeGT2*	this study
	moeno38-911	moeno38-91 *ΔmoeGT3, moeGT3* replaced with *moeGT3*:*aadA* from pOOB58	this study
	moeno38-41	moeno38-1 *ΔmoeF5*	this study
	moeno38-61	moeno38-1 *ΔmoeH5*	this study
	moeno38-21	moeno38-1 *ΔmoeK5*	this study
	moeno38-412	moeno38-1 Δ*moeE5*	this study

[a]ATCC, American Type Culture Collection, Manassas, VA

1. Construction of *S. ghanaensis* strain MO12

1.1 Disruption of *moeGT3* in *S. ghanaensis* ATCC14672.

An internal fragment of *moeGT3* was amplified with primers GT2con73up and GT2con73rp, digested with *Hin*dIII and *Eco*RI and cloned into the respective sites of pKC1139. The resulting plasmid, pMO12, was used to insertionally inactivate the *moeGT3* gene within the chromosome of *S. ghanaensis* ATCC14672 following a previously described protocol (Fig. S1).(*1*)

The site-specific integration of pMO12 into the *S. ghanaensis* MO12 strain was confirmed via Southern analysis using a DIG-labeled *moeGT3* internal fragment as a probe (Fig. S1). In the wild type strain, the *moeGT3* gene resides in a 4.3-kb *Bam*HI fragment whereas in the MO12 strain the corresponding hybridizing band is absent and a new 11-kb band is present. The latter corresponds to integration of the 7-kb pMO12 plasmid into the 4.5-kb *Bam*HI *moeGT3*-containing fragment of the *S. ghanaensis* chromosome. Similarly, the hybridization pattern of *Xho*I-digests of the wild type and mutant strain confirms the insertional inactivation of *moeGT3* (Fig. S1a, c). Introduction of plasmid pMO14 (P_{ermE}-*moeGT3*) into the MO12 strain restored MmA production.

1.2 Analysis of *S. ghanaensis* MO12.

The purified cell extracts of MO12 showed strong antibacterial activity, implying that none of the steps essential for MmA pharmacophore formation were affected in the mutant (Fig. S1b). LC-MS analysis revealed that MO12 accumulated the known compound moenomycin C4 (**26**, MmC4, Fig. S2 and S3) as a final product.(*2*) This conclusion was confirmed by high-resolution mass-spectral analysis (calculated $[M-H]^-$ of MmC4: 1418.6007, observed: 1418.6016) (Fig. S3). Furthermore, the strain accumulated the MmC4 precursor lacking the A ring (**27**, Fig. S3, calcd. $[M-H]^-$: 1322.5747; obsvd: 1322.5799). We also observed di- and trisaccharide fragments of MmC4 (Fig. S2) as a result of fragmentation during MS^1 experiments (such a phenomenon has been reported for moenomycins).(*3*)

2. Heterologous *S. lividans* TK24 strains expressing various subsets of *moe* genes

2.1 *S. lividans* moeno38-7 (*ΔmoeN5*)

2.1.A Construction. The kanamycin resistance gene from plasmid pKD4 was amplified with primers moeN5up_P1 and MoeN5rp_P2. The resulting amplicon was used to replace *moeN5* on cosmid moeno38-1 (Fig. 2). We did not evict the *kanR* gene region from moeno38-7 because it did not exert any polar effects on MmA production.

2.1.B Analysis. The strain accumulates two new closely related compounds **5**/**6** (Fig. S4), which were not detected in the extracts of either the empty heterologous host (TK24) or the other recombinant *S. lividans* strains. The dramatically shifted retention time of **5**/**6** (Rt 4.1/4.2 min, respectively) compared to MmA or compound **3** (Rt 9.9 min) is an indication of a shortened lipid chain. The analysis of the Δ*moeN5* extract revealed the presence of precursors to **5**/**6**, namely the compounds **9**, **10**, **11**, **18**, **19**, **29**, and **29c** (Fig. S4). The exact $[M-H]^-$ of all aforementioned compounds coincide with calculated ones (compound **18** – calcd: 943.3694, obsvd: 943.3684; **19** - calcd: 986.3752, obsvd: 943.3732; **5** – calcd: 1365.4861, obsvd: 1365.4867; **6** – calcd: 1364.5021, obsvd: 1364.5023; **29** - calcd: 1189.4546, obsvd: 1189.4503) and point to the presence of a common 15 carbon polyprenyl chain (Fig. S5). Additionally, the presence of a 1072/1073 peak in the MS^2 spectra of **5**/**6** (Fig. S6) confirms that these compounds possess the pentasaccharide-phosphoglyceric acid moiety found in other moenomycins.(*1, 3*) Compounds **5**/**6** show biological activity against the

reporter *B. cereus* strain (Table 2, Fig. S21). Based on bioactivity, exact mass, and similarity to known fragmentation patterns, we proposed the structures of **5**/**6**.

2.1.C Complementation. We failed to restore the production of moenomycins having a C25 isoprenoid chain (e.g. **3**) in the Δ*moeN5* strain using a construct where *moeN5* expression was driven from a constitutive P_{ermE^*} promoter (pMO17). Moreover, we revealed that *moeN5* overexpression even led to a decrease in **5**/**6** production (data not shown), pointing to the existence of an as-yet-unknown regulatory mechanism governing *moeN5* expression under natural conditions.

We constructed plasmid pOOB63a which contains both *moeO5* and *moeN5* genes along with 0.6 kb *moeX5-moeO5* intergenic region. We assumed that the intergenic region contains a promoter responsible for the expression of the *moeO5moeN5* operon. Indeed, introduction of pOOB63a into Δ*moeN5* strain restored the production of **3**. The plasmid pED2, containing only *moeO5* and it's putative promoter under P_{ermE^*}, did not restore production of **3** in the Δ*moeN5* strain, meaning that *only* gene *moeN5* in plasmid pOOB63a is responsible for the restoration of **3** production.

2.2 *S. lividans* TK24 moeno 38-31 *(ΔmoeGT4)*

2.2.A. Construction of the strain. *KanR* was amplified with primers moeGT4up_P3 and moeGT4rp_P2 and used to replace the *moeGT4* gene. The resulting cosmid moeno38-3 was introduced into *E. coli* DH5α (pCP20) to excise

the *kanR* gene in an FLP-mediated reaction.(*4*) The cosmid carrying an 81bp "scar" sequence instead of *moeGT4* was named moeno38-31. The deletion of *moeGT4* in moeno38-31 was confirmed via PCR (see Table S1).

2.2.B Analysis. The LC-MS analysis of extracts from the Δ*moeGT4* strain revealed a mass-peak (Fig. S7) corresponding to one of the proposed monosaccharide precursors to MmA detected in the Δ*moeN5* strain. The exact mass, chromatographic and biological properties of this compound are consistent with the structure of **9** or its equatorial epimer (Fig. 4, Table 1).

2.2.C Complementation. *MoeGT4* was amplified from cosmid moeno38-1 via PCR using the primers listed in Table S1 and cloned into vector pMKI9 to create plasmid pOOB50. The resulting plasmid was transferred to strain *S. lividans* TK24 moeno38-31 via conjugation, and the resulting strain produced the desmethylated pentassacharide compound **4**. Resequencing of *moeK5* revealed that the C-terminus of this gene overlaps with *moeGT4* and was inadvertently deleted in this study.

2.3 *S. lividans* TK24 moeno 38-91 (*ΔmoeGT5*)

2.3.A Construction. Strain moeno38-91 was constructed in the same way as moeno38-31 (Section 2.2.A). The gene deletion was confirmed by PCR (primers listed in Table S1).

2.3.B Analysis. LC-MS analysis indicates that the Δ*moeGT5* strain accumulates the trisaccharide intermediate **20** (Fig. 4, Fig. S8). Unlike the mono- and disaccharide MmA intermediates accumulated by the Δ*moeGT4* and

Δ*moeGT5moeGT3* strains (see above), compound **20** has a full-length polyprenyl chain as judged from both retention time and high resolution MS data (exact $[M-H]^-$ calcd:1122.4998, obsvd: 1122.5005; Fig. S8). The structure of **20** is also supported by MS^2 (Fig. S9) and ^1H NMR experiments presented below.

2.3.B.1 Proton NMR analysis of compound 20. Full NMR assignments for moenomycin and related compounds have been reported(*5*) and all resonances for compound **20** could be assigned based on reported data. Multiplicities are reported by using the following abbreviations : s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad, dt = doublet of triplets, *J* = coupling constant in Hertz. δ_H(500 MHz, D_2O) d 5.79 (dd, F-1, $^3J_{1,P}$ = 5.8 Hz, $J_{1,2}$ = 3.9 Hz, 1H), 5.45 (d, I-7, $J_{6,7}$ = 14.4 Hz, 1H), 5.42 (t, I-2, $J_{1,2}$ = 5.3 Hz, 1H), 5.36 (dt, I-6, $J_{5,6}$ = 5.9 Hz, $J_{6,7}$ = 14.4 Hz, 1H), 5.23 (t, I-13, $J_{12,13}$ = 7.3 Hz, 1H), 5.13-5.21 (m, I-17, 1H), 5.04 (d, F-3, $J_{2,3}$ = 10.8 Hz, 1H), 4.76 (s, I-22, 2H), 4.66 (d, E-1, $J_{1,2}$ = 7.8 Hz, 1H), 4.56 (d, D-1, $J_{1,2}$ = 7.8 Hz, 1H), 4.44 (s, F-5, 1H), 3.92-4.27 (m, E-6, H-1, H-2, H-2', I-1, I-1', 6H), 3.91 (dd, D-6, $J_{5,6'}$ = 2.6 Hz, $J_{6,6'}$ = 12.7 Hz, 1H), 3.42-3.92 (m, D-3, D-5, D-6', E-2, E-3, E-4, E-5, E-6', F-2, F-4, 10H), 3.38 (dd, D-4, $J_{4,5}$ = 9.3 Hz, 1H), 3.30 (dd, D-2, $J_{2,3}$ = 8.3 Hz, 1H), 2.77 (brd, I-12, $J_{12,13}$ = 6.9 Hz, 1H), 1.97-2.22 (m, I-4, I-4', I-5, I-5', I-15, I-15', I-16, I-16', 8H), 2.04 (s, E-2-NAc, 3H), 1.87-1.97 (m, I-10, I-10', 2H), 1.76 (s, I-25, 3H), 1.69 (s, I-19, 3H), 1.65 (s, I-21, 3H), 1.63 (s, I-20, 3H), 1.36-1.42 (I-9, m, 2H), 1.23 (s, F-4-Me, 3H), 0.97 (s, I-23, I-24, 6H).

2.3.C Complementation. *MoeGT5* was amplified from cosmid moeno38-1 via PCR using the primers listed in Table S1 and cloned into vector pMKI9 to create

plasmid pOOB52. pOOB52 was transferred to strain *S. lividans* TK24 moeno38-91 via conjugation and restored production of compound **3**.

2.4 *S. lividans* TK24 moeno38-911 (*ΔmoeGT5ΔmoeGT3)*

2.4.A Construction. The cosmid moeno38-911 was constructed by disrupting *moeGT3* in cosmid moeno38-91. A 3.3kb *Xba*I-*Eco*RI fragment containing the *moeGT3*::*aadA* allele was retrieved from pOOB58 and used to replace *moeGT3* in cosmid moeno38-91. The replacement of *moeGT3* with *moeGT3*::*aadA* in moeno38-911 was verified by PCR (primers moeGT3XbaIup and moeGT3EcoRIrp).

2.4.B Analysis. The Δ*moeGT5*Δ*moeGT3* strain accumulated the disaccharide compound **11** (exact $[M-H]^-$ calcd: 781.3160, obsvd: 781.3147; Fig. S10, Fig. 2). A precursor to **11**, compound **9**, was also detected. Both compounds are present in moenomycin extracts from the Δ*moeN5* strain (see above).

2.4.C Complementation. *MoeGT3* was amplified from cosmid moeno38-1 via PCR using the primers listed in Table S1 and cloned into vector pMKI9 to create plasmid pMO13. pMO13 was transferred to strain *S. lividans* TK24 moeno38-911 via conjugation and restored production of compound **20** (see Section 2.3).

2.5 *S. lividans* TK24 moeno38-81 (*ΔmoeGT2)*

2.5.A Construction. Strain moeno38-81 was constructed in the same way as moeno38-31 (Section 2.2.A). The gene deletion was checked by PCR (primers listed in Table S1).

2.5.B Analysis. The Δ*moeGT2* strain accumulates the tetrasaccharide MmA intermediate **22** (exact $[M-H]^-$ calcd: 1325.5792, obsvd: 1325.5786) (Fig. 2, Fig. S11).

2.5.C Complementation. *MoeGT2* was amplified from cosmid moeno38-1 via PCR using the primers listed in Table S1 and cloned into vector pMKI9 to create plasmid pOOB56c. pOOB56c was transferred to strain *S. lividans* TK24 moeno38-81 via conjugation and restored production of compound **3**.

2.6 *S. lividans* TK24 moeno38-61 (*ΔmoeH5*)

2.6.A Construction. Strain moeno38-61 was constructed in the same way as moeno38-31 (Section 2.2.A). The gene deletion was checked by PCR (primers listed in Table S1).

2.6.B Analysis. The LC-MS of the Δ*moeH5* strain extracts revealed a new mass-peak corresponding to pentasaccharide **24** (exact $[M-H]^-$ calcd: 1501.6113, obsvd: 1501.6113) (Fig. 2, Fig. S12). The strain also accumulates significant quantities of intermediates including tri- and tetrasaccharide **20** and **22**, respectively (data not shown). The pattern of accumulated intermediates, MS^2 data (Fig. S13), exact mass of **24**, as well as its biological and chromatographic properties allowed us to propose the structure of **24** as shown in Figure 2. The distinctive feature of **24** is the acid form of unit B (galactopyranuronic acid) instead of galactopyranuronamide in **3**/**25**.(*1*) In addition, the conversion of **24** into the known compound **23**(*3*) via *moeR5* expression in Δ*moeH5* strain also supports the structure of **24** (see below).

2.6.C Complementation. *MoeH5* was amplified from cosmid moeno38-1 via PCR using the primers listed in Table S1 and cloned into vector pMKI9 to create plasmid pOOB51. pOOB51 was transferred to strain *S. lividans* TK24 moeno38-61 via conjugation and restored production of compound **3**.

2.7 *S. lividans* TK24 moeno38-61moeR5^{+}

2.7.A Construction. *MoeR5* was amplified from moeno38-1 with primers moeR5XbaIup and moeGT3intMfeI and cloned into *Xba*I-*Eco*RI digested pMKI9 to yield pOOB59 which was transferred to *S. lividans* TK24 moeno38-61 via conjugation.

2.7.B Analysis. The strain Δ*moeH5* expressing plasmid pOOB59 (*moeR5* under P*ermE* transcriptional control) possesses an altered profile of produced moenomycins compared to the initial strain. By shortening the fermentation time from 6 days to 4 days, we were able to detect in the extracts of moeR5^{+}Δ*moeH5* strain the initial compound **24** and its derivative **23**(*3*) (Fig. 4, Fig. S14) resulting from MoeR5 activity. The 16 Da difference between the masses of **24** and **23** (exact $[M-H]^{-}$ calcd: 1485.6164, obsvd: 1485.6199) agrees with the proposed role of the *moeR5moeS5* gene pair in conversion of UDP-GlcNAc into its C6-deoxy derivative, UDP-chinovosamine.

2.8 *S. lividans* TK24 moeno38-41 (*ΔmoeF5*)

2.8.A Construction. Strain moeno38-41 was constructed in the same way as moeno38-31 (Section 2.2.A). The gene deletion was checked by PCR (primers listed in Table S1).

2.8.B Analysis. The strain accumulates inactive MmA intermediate compound **8** (exact $[M-H]^-$ compound **8**: calcd. 565.2050, obsvd. 565.2055; compound, Fig. 4, Fig. S15). Its mass is 1 Da more than that of the monosaccharide MmA intermediate **9** accumulated by the Δ*moeGT4* strain as would be expected from the F ring C4 acid versus the amide.

2.8.C Complementation. *MoeF5* was amplified from cosmid moeno38-1 via PCR using the primers listed in Table S1 and cloned into vector pMKI9 to create plasmid pOOB48a. pOOB48a was transferred to strain *S. lividans* TK24 moeno38-41 via conjugation and restored production of compound **3**.

2.9 *S. lividans* TK24 moeno38-21 (*ΔmoeK5)*

2.9.A Construction. Strain moeno38-21 was constructed in the same way as moeno38-31 (Section 2.2.A). The gene deletion was checked by PCR (primers listed in Table S1).

2.9.B Analysis. The strain accumulates new compound **4** (exact $[M-H]^-$ calcd: 1486.6116, obsvd: 1486.6118) (Fig. 4, Fig. S16) with a mass and biological properties which are consistent with the loss of the methyl group from unit F of compound **3**. MS^2 analysis of **4** (Fig. S17) shows a fragmentation pattern consistent with the proposed structure. The stereochemistry of the F ring C4 hydroxyl group is assumed to be axial based on full characterization of

desmethyl moenomycins isolated from fermentation broths,(*5-8*) as well as our biochemical evidence that UDP-GalUA (having an axial configuration at C4) is the natural sugar donor used in F ring biosynthesis (see Section 4.3).

2.9.C Complementation. *MoeK5* was amplified from cosmid moeno38-1 via PCR using the primers listed in Table S1 and cloned into vector pMKI9 to create plasmid pED4. Although this plasmid did not complement the strain, the accumulation of the expected desmethyl pentasaccharide **4** in strain *S. lividans* TK24 moeno38-21 rules out the possibility of any polar effect. The lack of complementation is most likely due to incorrect sequencing of the original *moeK5* since subsequent studies (data not shown) have indicated that the C-terminus of *moeK5* in fact overlaps with *moeGT4*.

2.10 *S. lividans* TK24 moeno38-5 (*ΔmoeA5ΔmoeB5*)

2.10.A Construction. The kanamycin resistance gene from plasmid pKD4 was amplified with primers 38start-KD4 and moeA4-P3. The resulting amplicon was used to replace the *moeA4moeB5* gene pair as well as the entire nonessential "left arm" of moeno38-1 (Fig. 2). Our previous studies showed that deletion of this arm did not alter MmA production (data not shown). We did not evict the *kanR* gene region from moeno38-5 because it did not exert any negative effects on MmA production. The replacement of the *moeA5* and *moeB5* genes with *kanR* in moeno38-5 was confirmed via diagnostic PCR (primers P2_KD4 and alsrev1).

2.10.B Analysis. The moenomycins produced by the Δ*moeA5*Δ*moeB5* strain were the same as from the parental moeno38-1^{+} strain,(*1*) implying that the moeA5 and moeB5 genes are not needed for lipid-pentasaccharide (**3**) biosynthesis (Fig. S18).

2.11 *S. lividans* TK24 moeno38-5 moeR5^{+}

2.11.A Construction. pOOB59 (see section 2.7.A) was transferred to *S. lividans* TK24 moeno38-5 via conjugation.

2.11.B Analysis. Expression of *moeR5* in the Δ*moeA5*Δ*moeB5* strain leads to the production of known MmA intermediate **25** (exact $[M-H]^{-}$ calcd: 1484.6324, obsvd: 1484.6328) (Fig. 4, Fig. S19) which has been reported by others(*3, 9*) and our laboratory.(*1*) A comparison of the spectra of moenomycins produced by this strain (Fig. S19), the parental strain (Section 4.11), and the *moeR5*$^{+}$Δ*moeH5* strain (Section 4.7) indicates that *moeR5* participates in the conversion of UDP-GlcNAc into its C6-deoxy derivative, UDP-chinovosamine.(*1*) Knowledge of the genotypes of the strains listed above combined with exact masses of the compounds **23**, **24**, **25**, and **3** (Table 1) support the inferred structures (Fig. 4).

2.12 *S. lividans* TK24 moeno38-5 expressing *moe* cluster 2

2.12.A Construction. Genes *moeA4, moeB4,* and *moeC4* (*moe* cluster 2) were amplified with primers moe2HindIII and moe2MfeI. The resulting PCR product was cloned into pSOK804 to give pOOB64b. This plasmid was coexpressed with various moeno38-5 derivatives to study A ring biosynthesis.

2.12.B Analysis. The co-expression of *moe* clusters 1 and 2 in the absence of the *moeR5moeS5* gene pair leads to the production of the expected A ring containing compound, pholipomycin(*2*) (**2**, exact $[M-H]^-$ calcd: 1596.6490, obsvd: 1596.6439; Fig. 4) as judged by MS analysis (Fig. S20). The UV absorption spectrum of **2** has peak at 258nm (absent in MmA derivatives lacking chromophore A ring), which further supports the structure **2**.

2.13. *S. lividans* TK24 moeno38-412 (Δ*moeE5)*

2.13 Construction. *KanR* was amplified and used to replace an internal fragment of *moeE5*. The resulting cosmid was introduced into *E. coli* DH5α (pCP20) to excise the *kanR* gene in an FLP-mediated reaction.(*4*) The cosmid carrying an 81bp "scar" sequence instead of *moeE5* was named moeno38-412. The deletion of *moE5* in moeno38-31 was confirmed via PCR (see Table S1).

2.13.B Analysis. This strain did not accumulate any detectable MmA intermediates, and we decided to investigate its function utilizing *in vitro* biochemistry (see below).

3. MoeO5 reconstituted from *S. lividans*

3.1 MoeO5 cloning

MoeO5 and its promoter (P_{MoeO5}) were PCR amplified from cosmid moeno38-1 using primers moeO5HindIIIup (AAAAAGCTTCGGGGCGTGCCTTCTTC) and MoeO5Xba1rp (AAATCTAGACCGCCCGCTCCCCGGAC). These primers introduced the respective restriction enzyme sites HindIII and XbaI (underlined). PCR

amplification was performed with KOD Hot Start polymerase (Novagen). The PCR product was gel purified, digested with HindIII and XbaI, and ligated using T4 DNA ligase (NEB) into the linearized constitutive expression plasmid pAF1 to give the N-terminal His6 tagged pED2.

3.2 MoeO5 expression and purification

The plasmid pED2 was transferred to *S. lividans* TK24 via protoplast transformation. A single colony was inoculated into 30 mL TSB Thio[50] media and grown at 30 °C in a baffled flask. After 2 days, all 30 mL were used to inoculate 4 L of R5 media in a fermenter (New England Center for Research Excellence, NERCE). The cells were grown at 30 °C for 4 days before harvesting by centrifugation (30 min at 5000g) and storage at -80 °C.

pED2+ mycelia were resuspended in 3 mL/g Buffer A (NaH_2PO_4 pH=8.0, 200 mM NaCl, 5 mM $MgCl_2$) and incubated at room temperature with rlysozyme, benzonase, and protease inhibitor complexes (Novagen) for 2 h. Cells were the lysed by sonication and the cell debris was removed by centrifugation (45 min at 14000g). The supernatant was rocked with TALON® resin (Clontech) for 1 h at 4 ºC. The resin was collected and washed with 20 vol Buffer A by rocking for 30 minutes at 4 ºC, packed into a column, and eluted with 20-200 mM imidazole in buffer A. Fractions containing the target protein (identified by SDS–PAGE) were pooled, desalted, concentrated, flash frozen in liquid nitrogen, and stored at –80 ºC. The protein concentration was determined using the Dc Protein Assay (Biorad) using BSA as the standard. The yield of purified MoeO5 was

approximately 1 mg/L with a purity of 50% (data not shown). The protein sequence was confirmed by mass spectrometry.

3.3 MoeO5 reconstitution

All reactions were carried out as described in Materials and Methods. MoeO5 purified from *S. lividans* and SUMO- MoeO5 from *E. coli* act identically. Mock purification of an empty vector control (pAF1 in *S. lividans*) yielded no band in the region of MoeO5, and the eluent fraction contained no prenyltransferase activity.

4. Sequence of *synO5*

GTTAACGCTAGTCCTCAACTGGATCATCACACGGAGCTGCACGCGGCACCAC
CTCTGTGGCGTCCGGGTCGTGTTCTGGCTCGTCTGCGTGAACATCAACCTGGT
CCTGTCCATATCATCGATCCGTTCAAAGTTCCGGTTACCGAAGCAGTGGAAA
AAGCGGCGGAGCTGACTCGCCTGGGTTTTGCTGCTGTCCTGCTGGCTTCTACC
GATTATGAGTCCTTCGAATCCCACATGGAACCGTATGTTGCTGCCGTGAAAG
CTGCTACTCCTCTGCCTGTTGTTCTGCACTTTCCTCCACGTCCTGGAGCTGGAT
TTCCTGTCGTTCGTGGCGCTGATGCTCTGCTGCTGCCGGCCCTGCTGGGTTCT
GGTGACGATTATTTTGTGTGGAAAAGTTTCCTGGAGACACTGGCGGCATTTCC
GGGGCGTATTCCTCGTGAAGAGTGGCCTGAGCTGCTGCTGACTGTTGCTCTAC
CTTTGGCGAAGATCCTCGTACAGGTGATCTGCTGGGCACAGTGCCTGTTAGT
ACCGCTAGTACCGAAGAGATTGACCGCTATCTGCATGTCGCTCGTGCCTTTGG
TTTTCACATGGTGTATCTGTATAGCCGTAATGAGCACGTTCCTCCGGAGGTCG

TCCGTCATTTTCGTAAAGGGCTGGGACCTGACCAAGTCCTGTTTGTGAGTGGC

AACGTTCGTAGTGGCCGTCAAGTGACCGAATATCTGGATAGTGGCGCTGACT

ATGTTGGTTTTGCCGGGGCTCTGGAACAGCCTGATTGGCGTAGCGCTCTGGCT

GAAATTGCTGGTCGTCGTCCTGCTGCCCCTGCTCGTCCTGGTTCTGGCCGTTG

A

5. Preparation of moenocinyl pyrophosphate

A mixture of moenocinyl bromide (9.8 mg, 23.3 μmol, 1.0 eq.) and Tris(tetrabutylammonium) hydrogen pyrophosphate (53.5 mg, 59.3 μmol, 2.5 eq.) in dry MeCN (0.20 mL) was stirred at room temperature for 1.5 h. The resulting solution was directly loaded onto a normal phase silica column and eluted with *i*-PrOH/20%NH_3 aq. = 8/3 to give mononocinyl pyrophosphate (8.5 mg, 16.4 mmol, 70%). Multiplicities are reported by using the following abbreviations : s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad, dt = doublet of triplets, *J* = coupling constant in Hertz. δ_H (500 MHz, D_2O/CD_3OD = 2/1) 5.45 (t, I-2, $J_{1,2}$ = 6.4 Hz, 1H), 5.39 (d, I-6, $J_{5,6}$ = 15.5 Hz, 1H), 5.33 (dt, I-6, $J_{5,6}$ = 15.5 Hz, $J_{6,7}$ = 5.5 Hz, 1H), 5.14 (t, I-13, $J_{12,13}$ = 7.3 Hz, 1H), 5.10 (t, I-17, $J_{16,17}$ = 8.2 Hz, 1H), 4.66 (br-s, I-22, 1H), 4.66 4.66 (br-s, I-22, 1H), 4.53 (t, I-1, $J_{1,2}$ = 6.4 Hz, 1H), 3.24 (t, NBu-1, $J_{1,2}$ = 8.7 Hz, 24H), 2.69 (d, I-12, $J_{12,13}$ = 7.3 Hz, 2H), 1.99-2.20 (m, I-4, I-4', I-5, I-5', I-15, I-15', I-16, I-16', 8H), 1.91 (m, I-10, 2H), 1.60-1.74 (m, NBu-2, 24H), 1.73 (s, I-25, 3H), 1.67 (s, I-19, 3H), 1.61 (s, I-21, 3H), 1.60 (s, I-20, 3H), 1.42 (tt, NBu-3, $J_{2,3}$ = 7.3 Hz, $J_{3,4}$ = 7.3 Hz, 24H), 1.38 (m, I-9, 2H), 1.03 (t, NBu-4, $J_{3,4}$ = 7.3 Hz, 36H), 0.97 (s, I-23, I-24, 6H); ^{13}C NMR (100 MHz,

D_2O/CD_3OD = 2/1) : δ 150.08 (I-11), 141.01 (I-3 or I-7) 140.82 (I-3 or I-7), 136.80 (I-14), 131.68 (I-18), 126.93 (I-6), 125.30 (I-17), 123.18 (I-2), 122.98 (I-13), 109.62 (I-22), 63.37 (I-1), 59.11 (NBu-1), 42.70 (I-9), 40.77 (I-15), 36.28 (I-8), 35.77 (I-12), 33.02 (I-4), 32.69 (I-5), 32.45 (I-10), 28.07 (I-23), 28.07 (I-24), 27.63 (I-16), 26.49 (I-19), 24.24 (NBu-2), 24.23 (I-25), 20.23 (NBu-3), 18.39 (I-20), 16.65 (I-21), 13.89 (NBu-4); ^{31}P NMR (162 MHz, CD_3OD) : δ -8.88 (d, $J_{1,2}$ = 20.5 Hz, 1P), -8.40 (d, $J_{1,2}$ = 20.5 Hz, 1P), LRMS (ESI) calcd for $C_{25}H_{43}N_5O_7P_2$ $[M-H]^-$ 517.3, found 517.2.

Supplementary References

(1) Ostash, B., Saghatelian, A., and Walker, S. (2007) A streamlined metabolic pathway for the biosynthesis of moenomycin A. *Chem. Biol. 14*, 257-267.

(2) Welzel, P. (2005) Syntheses around the transglycosylation step in peptidoglycan biosynthesis. *Chem. Rev. 105*, 4610-4660.

(3) Zehl, M., Pittenauer, E., Rizzi, A., and Allmaier, G. (2006) Characterization of moenomycin antibiotic complex by multistage MALDI-IT/RTOF-MS and ESI-IT-MS. *J. Am. Soc. Mass. Spectrom. 17*, 1081-1090.

(4) Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A. 97*, 6640-6645.

(5) Hennig, L., Findeisen, M., Welzel, P., and Haessner, R. (1998) H-1 NMR spectroscopic studies of the moenomycins. *Magn. Reson. Chem. 36*, 615-620.

(6) Schuricht, U., Endler, K., Hennig, L., Findeisen, M., and Welzel, P. (2000) Studies on the biosynthesis of the antibiotic moenomycin A. *J. Prakt. Chem. 342*, 761-772.

(7) Ritzeler, O., Hennig, L., Findeisen, M., Welzel, P., and Muller, D. (1997) Search for new moenomycin structure-activity relationships. Synthesis of a trisaccharide precursor of a moenomycin analogue. *Tetrahedron 53*, 5357-5357.

(8) El-Abadla, N., Lampilas, M., Hennig, L., Findeisen, M., Welzel, P., Muller, D., Markus, A., and van Heijenoort, J. (1999) Moenomycin A: The role of the methyl group in the moenuronamide unit and a general discussion of structure-activity relationships. *Tetrahedron 55*, 699-722.

(9) Liu, H., Ritter, T. K., Sadamoto, R., Sears, P. S., Wu, M., and Wong, C. H. (2003) Acceptor specificity and inhibition of the bacterial cell-wall glycosyltransferase MurG. *Chembiochem 4*, 603-609.