C-nucleoside Formation in the Biosynthesis of the Antifungal Malayamycin A

Hui Hong,[1] Markiyan Samborskyy,[1] Yongjun Zhou,[1,2] and Peter F. Leadlay[1,*]

[1]Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK

[2]Present address: Marine Drugs Research Center, Department of Pharmacy, School of Medicine, Shanghai Jiao Tong University, 200127 Shanghai, P. R. China

*Correspondence and Lead Contact: pfl10@cam.ac.uk (P.F.L.)

SUMMARY

Malayamycin A is an unusual bicyclic *C*-nucleoside, with interesting antiviral, antifungal and anticancer bioactivity. We report here the discovery and characterization of the biosynthetic pathway to malayamycin by using genome mining of near-identical clusters from both the known producer *Streptomyces malaysiensis* and from *Streptomyces chromofuscus*. The key precursor 5'-pseudouridine monophosphate (5'-Ψ-MP) is supplied chiefly through the action of MalD, a TruD-like pseudouridine synthase. In vitro assays showed that MalO is an enoylpyruvyltransferase acting almost exclusively on 5'-Ψ-MP rather than 5'-UMP, while in contrast the counterpart enzyme NikO in the nikkomycin pathway readily accepts either substrate. As a result, deletion of *malD* in *S. chromofuscus* coupled with introduction of the gene for NikO led to production of non-natural *N*-malayamycin, as well as malayamycin A. Conversely, cloning *malO* into the nikkomycin producer *Streptomyces tendae* in place of *nikO* diverted biosynthesis towards *C*-nucleoside formation.

KEYWORDS

pseudouridine synthase, *Streptomyces*, enoylpyruvyltransferase, biosynthetic engineering, nucleoside antibiotic.

INTRODUCTION

Modified nucleosides that mimic UDP-linked metabolites are a prominent group of natural products, produced by *Streptomyces* and allied bacterial genera, that show particularly diverse activities (Winn et al., 2010; Niu et al., 2017; Chen et al., 2017). Malayamycin A (Figure 1, **1**) is a promisingly potent antifungal uracil C-nucleoside, originally isolated from the soil bacterium *Streptomyces malaysiensis* (Benner et al., 2003; Li et al., 2008). Such uracil C-nucleosides are rare, although pseudouridimycin (Figure 1, **2**) (this is a re-assigned structure of the compound previously reported under the name strepturidin (Pesic et al., 2014)) has been recently unveiled as an RNA polymerase inhibitor (Maffioli et al., 2017; Sosio et al., 2018); while pseudouridine itself (Figure 1, **3**) is produced in all living cells through the breakdown of non-coding RNAs containing specific Ψ residues (Spenkuch et al., 2014). The exact fungal target(s) of malayamycin A have not been identified, although it appears particularly to inhibit the sporulation of phytopathogenic fungi (Li et al., 2008). The bicyclic *N*-glycoside ezomycin A_1 **6** (Figure 1), which shares the same *trans*-fused perhydrofuropyran core structure as **1**, likewise shows selective activity against phytopathogenic fungi (Sakata et al., 1974).

In contrast, uracil *N*-nucleosides of the nikkomycin and polyoxin classes (Figure 1) have been shown to act as structural analogues of UDP-*N*-acetylglucosamine and to inhibit chitin synthase (Dähn et al., 1976; Winn et al., 2010; Niu et al., 2017), specifically disrupting fungal and insect cell wall biosynthesis. Analysis of the nikkomycin biosynthetic gene cluster from *Streptomyces tendae* (Bormann et al., 1996) and of the polyoxin gene clusters from *Streptomyces cacaoi* var. *asoensis* and *Streptomyces aureochromogenes* (Chen et al., 2009) has now led to a fairly detailed understanding of the enzymology of the pathway that leads from 5'-uridine monophosphate (UMP) **7** via 3'-enoylpyruvyl-UMP (3'-EPUMP) **8** and the high-carbon (>6Cs) sugar nucleoside octosyl acid **9** to the 5'-amino-5'-carboxy-5'-deoxyhexuronic acid (AHA) **10** (Figure 2), the most advanced common intermediate in formation of uracil-based polyoxins and nikkomycins. However, it is not known to what extent this pathway (Lilla and Yokoyama, 2016; He et al., 2017) is also used in the biosynthesis of uracil C-nucleosides such as malayamycin; nor at which point the Ψ nucleobase is introduced.

SUPPLEMENTARY FIGURES AND TABLES

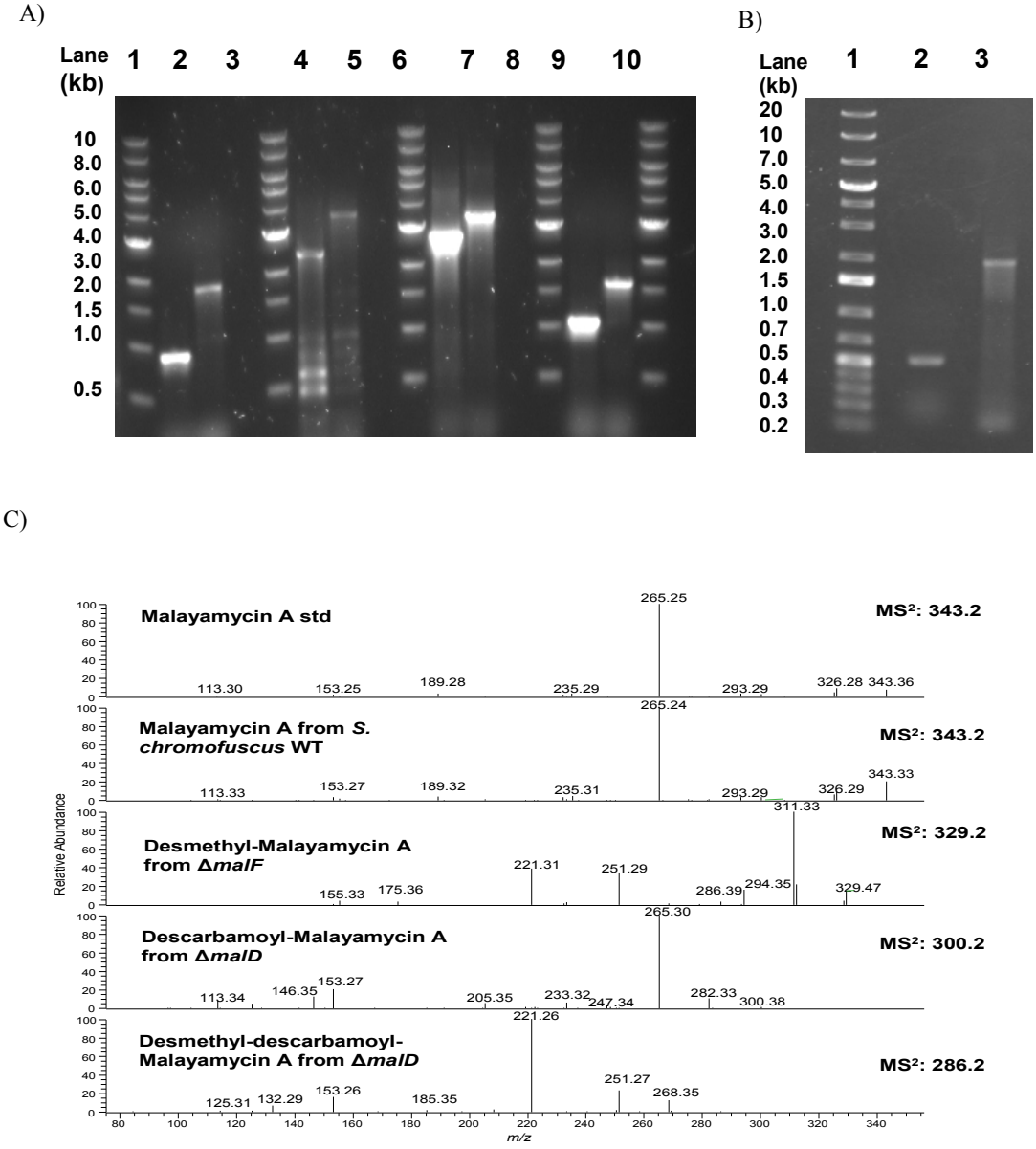
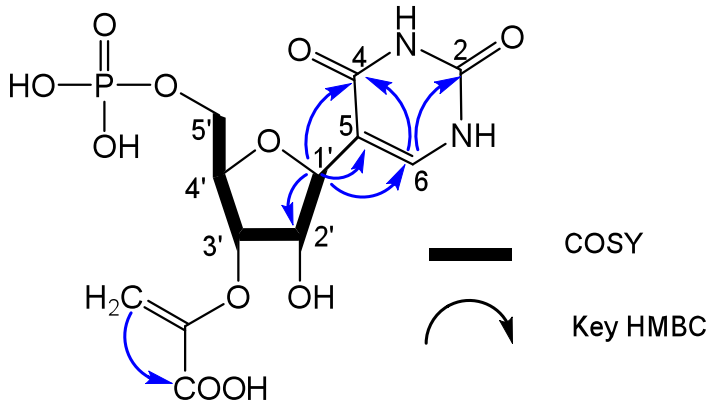
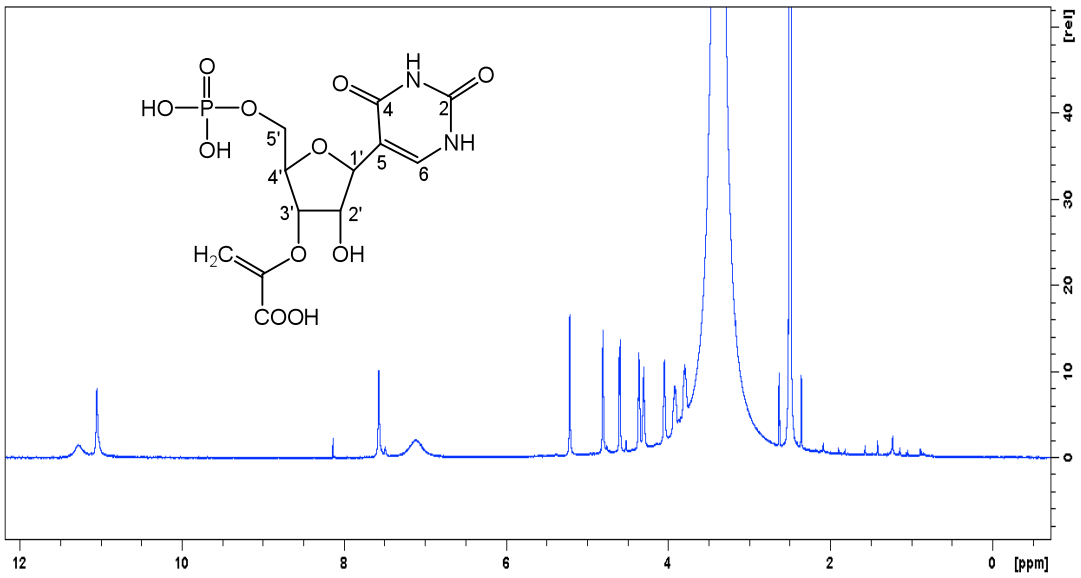


Figure S1. A) In-frame deletion of truD, malD, malF, and coaE genes in *S. chromofuscus*, Related to STAR Method and Figure 3. Lane 1: marker; Lane 2 and 3: PCR product from Δ*truD* (93 bp) and WT (1,897 bp), respectively; Lane 4: marker; Lane 5 and 6: PCR product from Δ*malD* (2,437 bp) and WT (3,703 bp), respectively; Lane 7: marker; Lane 8 and 9: PCR product from Δ*malF* (2,631 bp) and WT (3,315 bp), respectively; Lane10: marker; Lane 11 and 12: PCR product from Δ*coaE* (1,108 bp) and WT (1,675 bp), respectively; Lane 13: marker. **B) In-frame deletion of nikO gene in *S. tendae*, Related to STAR Method and Figure 3.** Lane 1: marker; Lane 2 and 3: PCR product from Δ*nikO* (520 bp) and WT (1,942 bp), respectively. **C) ESI-MS/MS spectra of malayamycin A and its analogous.** Malayamycin Astandard ($[M+H]^+$: 343.2), malayamycin A produced from *S. chromofuscus* WT, desmethyl-malayamycin A from Δ*malF* ($[M+H]^+$: 329.2), descarbamoyl-malayamycin A from Δ*malD* ($[M+H]^+$: 300.2) and descarbamoyl-descarbamoyl-malayamycin A from Δ*malD* ($[M+H]^+$: 286.2).

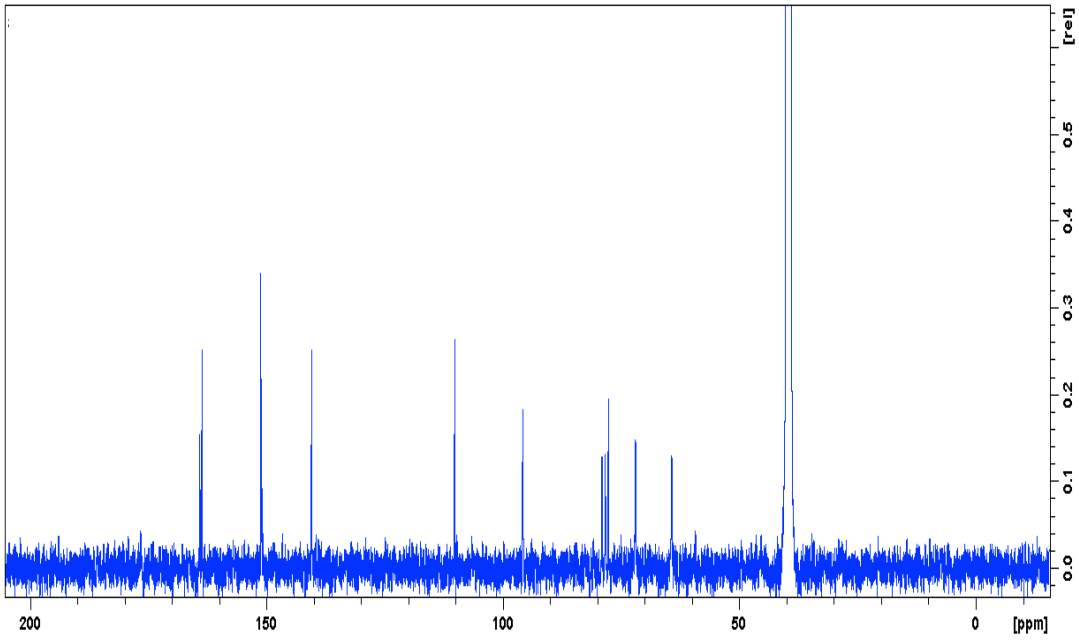
Data S1. 1D and 2D NMR Spectra of 3'-EP-Ψ-MP



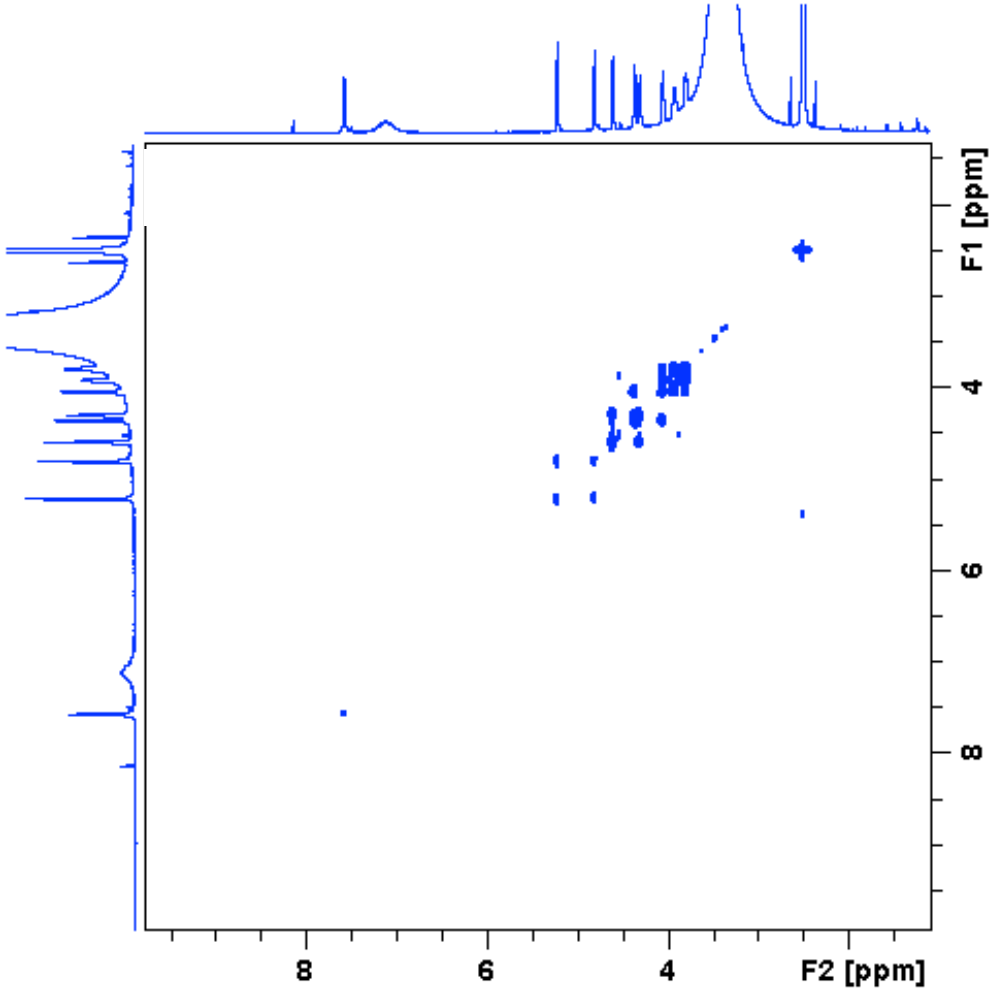
^{1}H NMR spectrum of 3'-EP-Ψ-MP



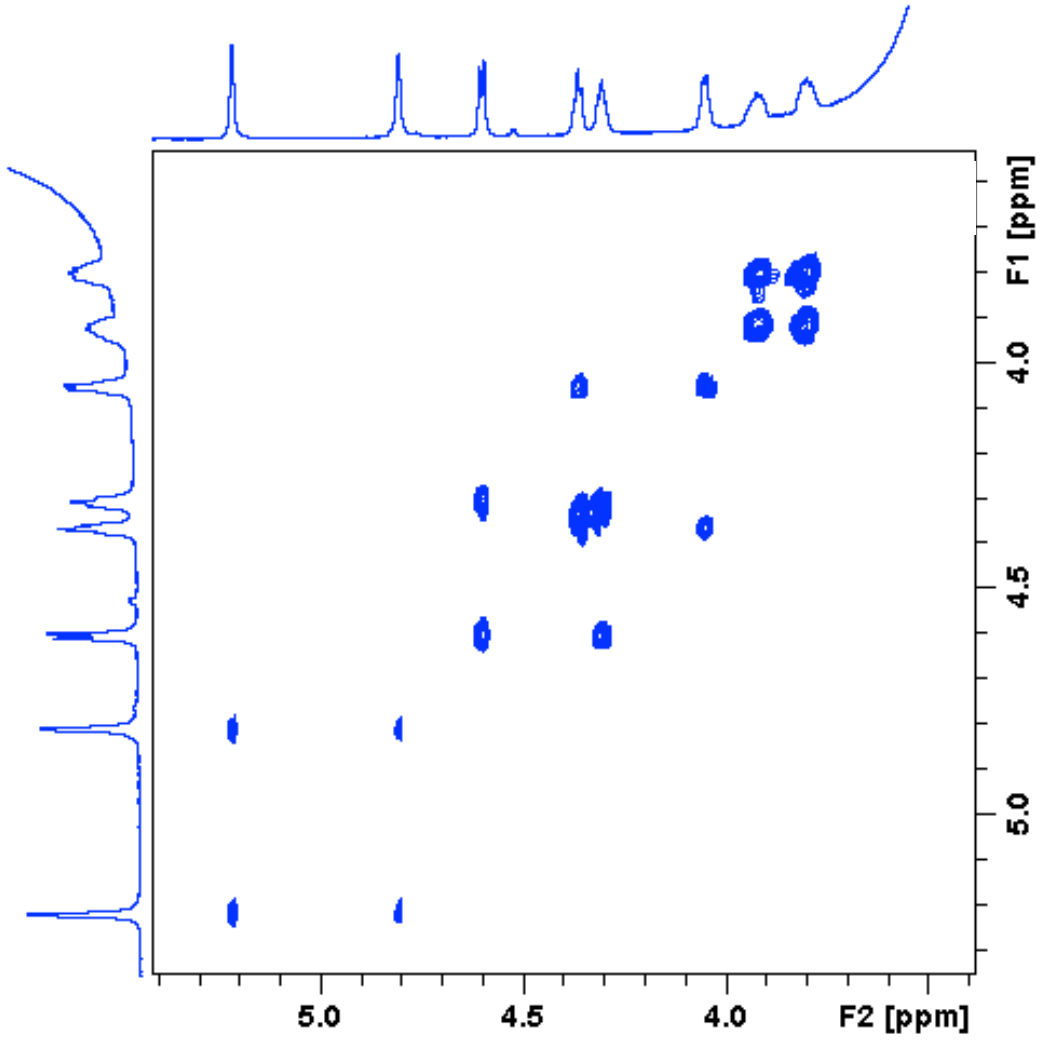
^{13}C NMR spectrum of 3'-EP-Ψ-MP



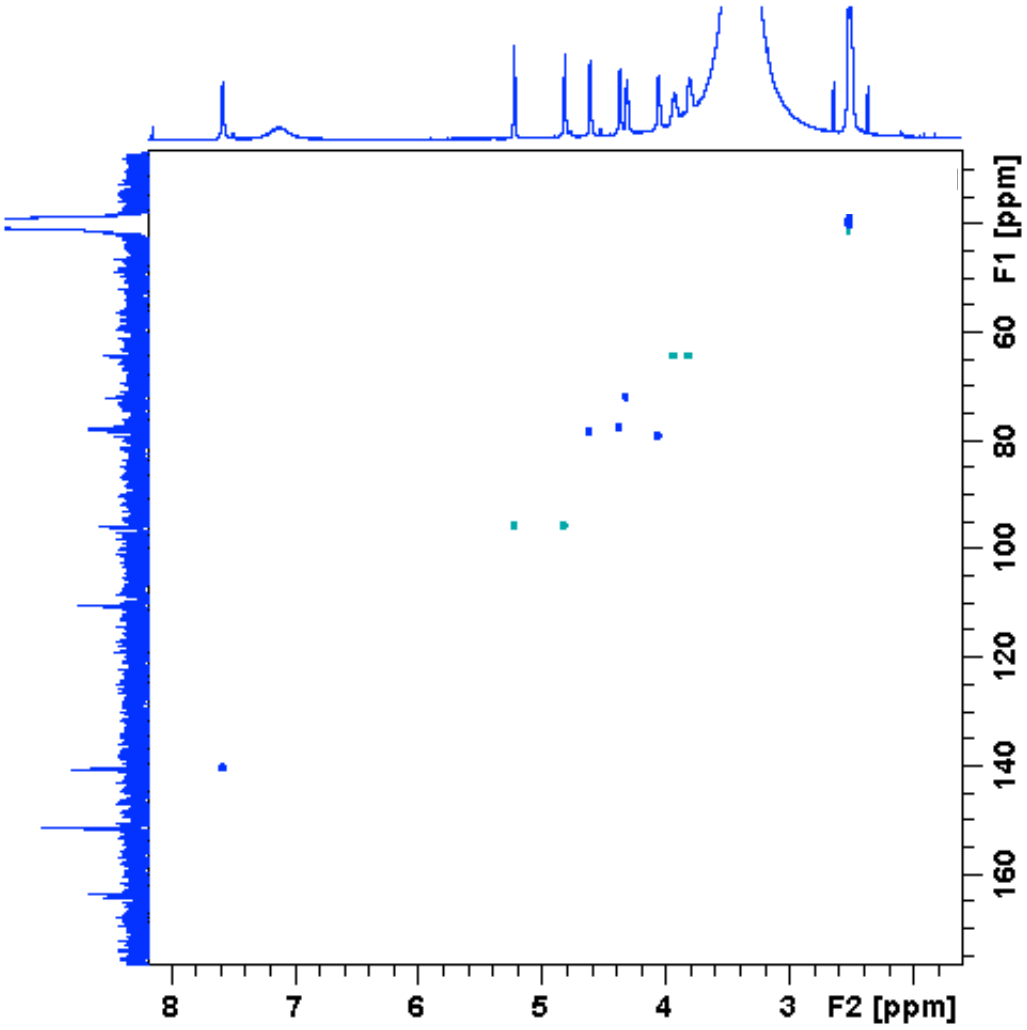
^{1}H-^{1}H COSY spectrum of 3'-EP-Ψ-MP



Zoomed ^{1}H-^{1}H COSY spectrum of 3'-EP-Ψ-MP



HSQC spectrum of 3'-EP-Ψ-MP



HMBC spectrum of 3'-EP-Ψ-MP

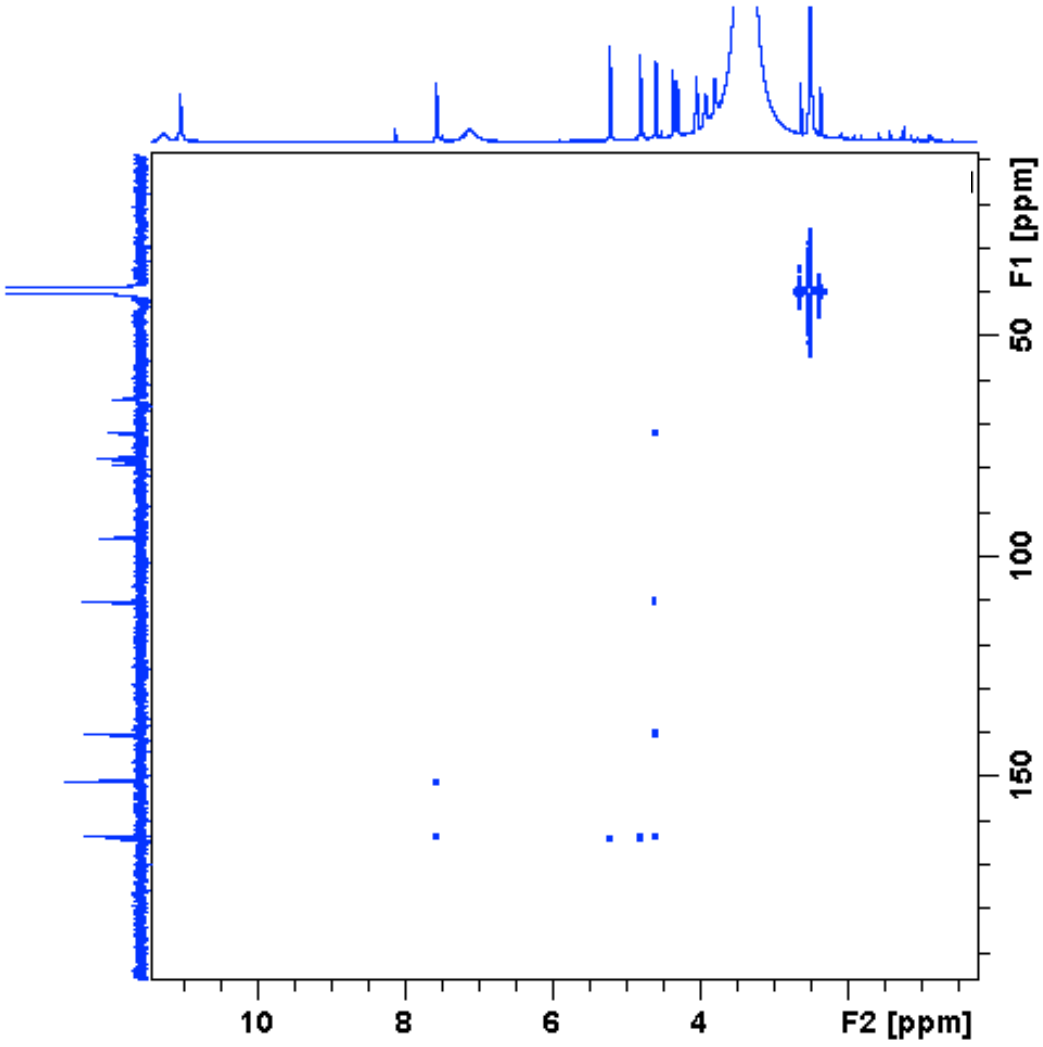


Figure S2. Sequence alignment of TruD family of tRNA pseudouridine synthases, Related to Figure 5. The sequences of TruDs [TruD_SM: TruD from *Streptomyces malaysiensis*; TruD_SC: TruD from *Streptomyces chromofuscus*; TruD_SS: TruD from *Streptomyces* sp. ID38640 (accession no. AVT42379); TruD_HP : TruD from *Helicobacter pylori* (accession no. NP_207718); TruD_AF: TruD from *Archaeoglobus fulgidus* (accession no. O28596); TruD_EC: TruD from *E. coli* (accession no. AQZ29382)] are aligned using Clustal Omega. The conserved sequence motifs are highlighted in boxes. The catalytic aspartate is denoted with an asterisk.

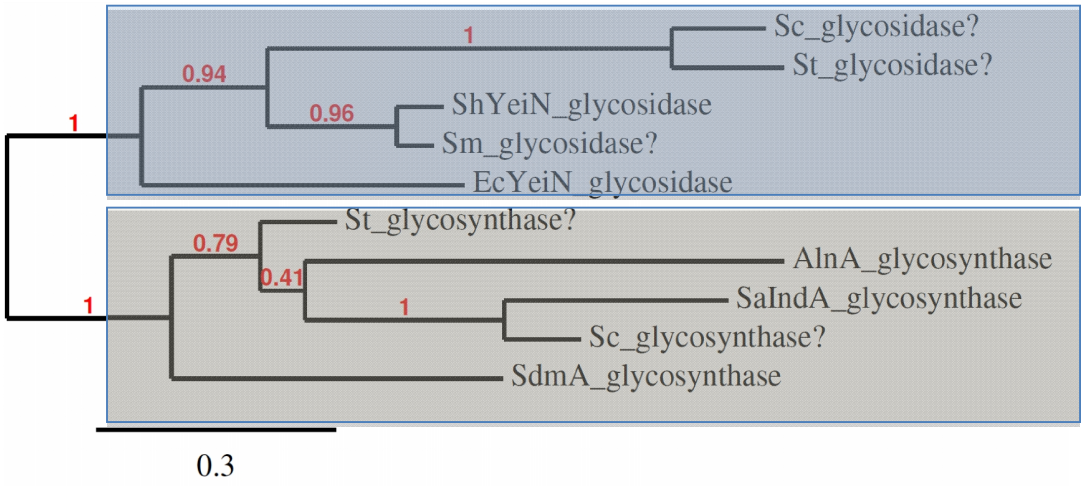


Figure S3. Phylogenetic analysis of YeiN-related glycosidase and putative glycosynthase enzymes, Related to Figure 4. Sc: *Streptomyces chromofuscus*; St: *Streptomyces tendae*; Sm: *Streptomyces malaysiensis*; ShYeiN: YeiN from *Streptomyces himastatinicus* (accession no. WP_009720488); EcYeiN: YeiN from *Escherichia coli* (accession no. AAA60517); AlnA: alnumycin biosynthesis from *Streptomyces* sp. CM020 (accession no. ACI88875); SaInda: indigoidine biosynthesis from *Streptomyces albus* (accession no. AMM12432); Sdma: showdomycin biosynthesis from *Streptomyces showdoensis* ATCC 15227.

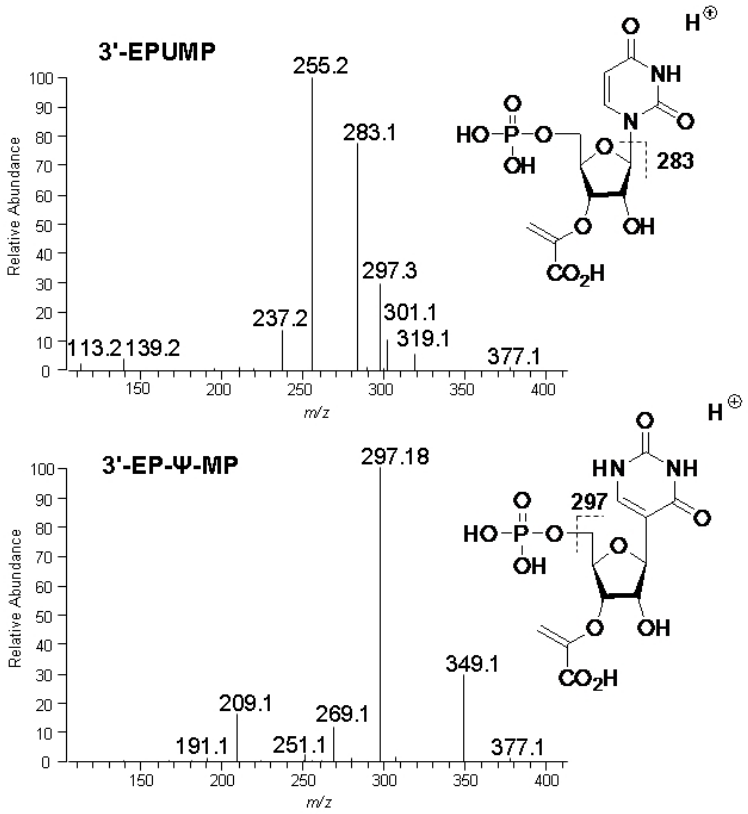
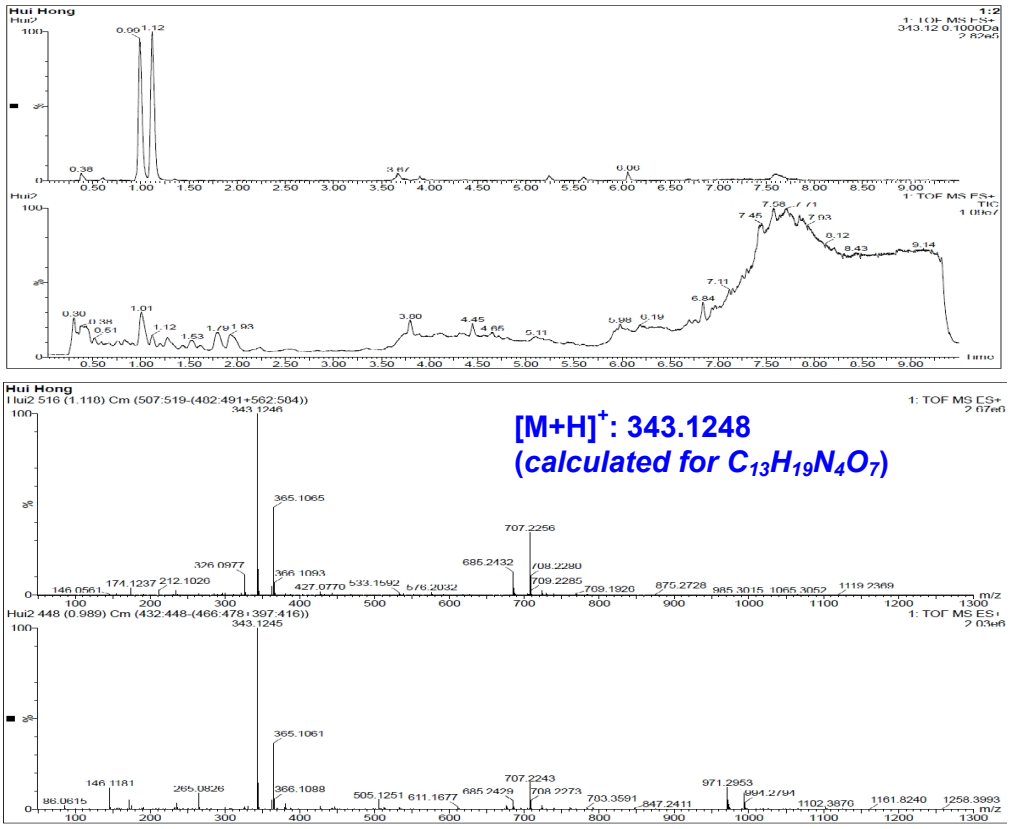


Figure S4. Comparison of ESI-MS/MS spectra of 3'-EPUMP ([M+H]$^+$: 395.2) and 3'-EP-Ψ-MP ([M+H]$^+$: 395.2), Related to Figure 4. Loss of C-N linkage in 3'-EP-Ψ-MP abolished the fragment at *m/z* 283.2, which was a key fragment in the MS/MS of 3'-EPUMP.

A)



B)

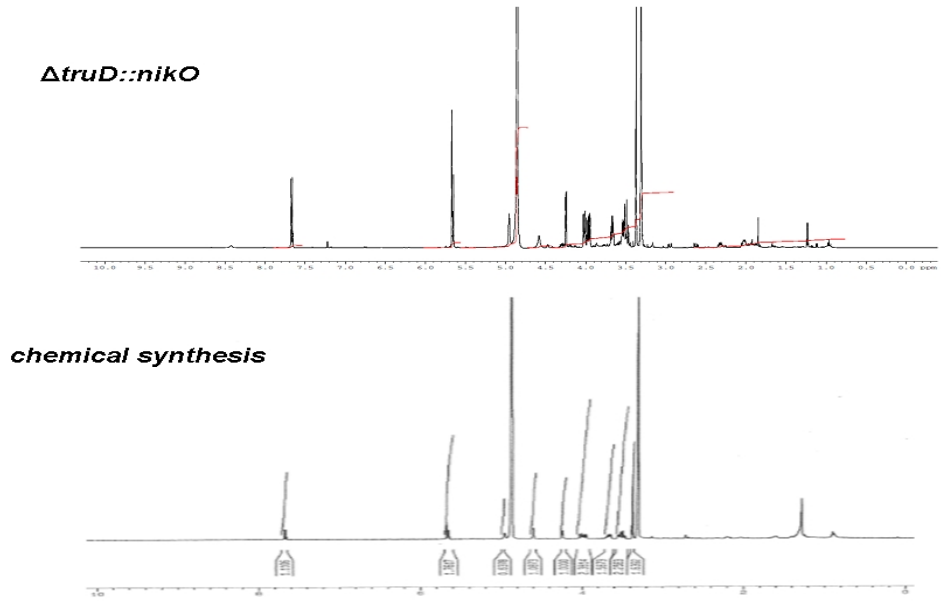


Figure S5. Analysis of *N*-malayamycin A by MS and ^{1}H NMR, Related to Figure 5A. A) High-resolution UPLC-MS analysis of *N*-malayamycin A and malayamycin A from fermentation extract of *S. chromofuscus* Δ*truD::nikO* mutant. **B)** Comparison of ^{1}H spectra of *N*-malayamycin A purified from fermentation of *S. chromofuscus* Δ*truD::nikO* mutant and from chemical synthesis (Hanessian et al., 2005).

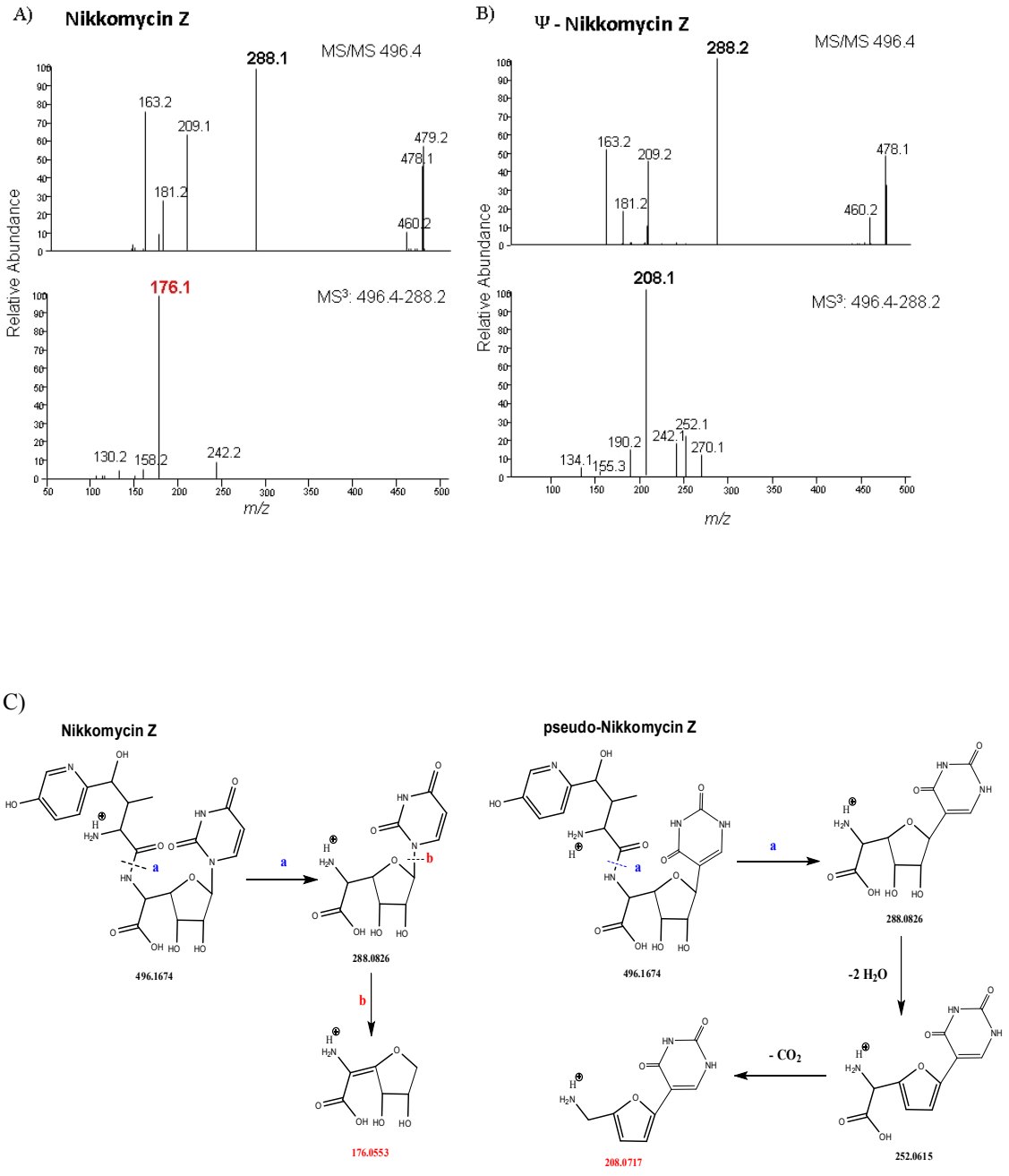


Figure S6. MS/MS and MS^3 analysis of nikkomycin Z and Ψ-nikkomycin Z, Related to Figure 5B. A) MS/MS and MS^3 spectra of nikkomycin Z. **B)** MS/MS and MS^3 spectra of Ψ-nikkomycin Z. **C)** Fragmentation pathways for the generation of *m/z* 176 fragment in nikkomycin Z and of *m/z* 208 fragment in Ψ-nikkomycin Z.

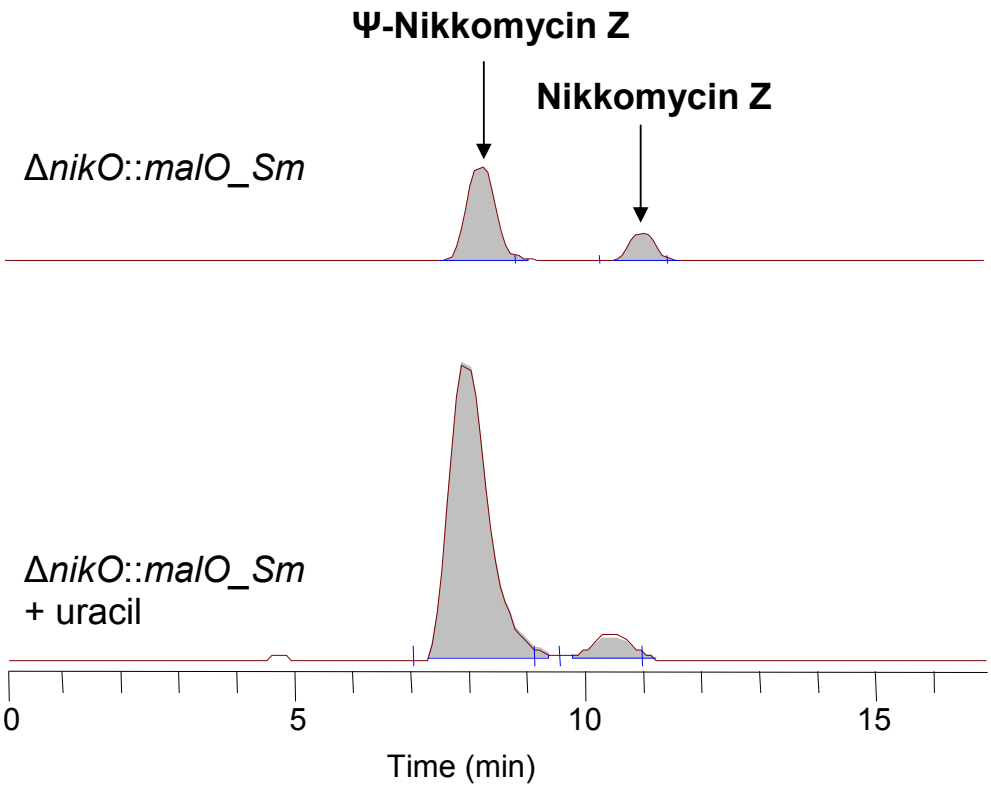


Figure S7. Comparison of Ψ–nikkomycin Z production with and without uracil feeding to the Δ*nikO*::*malO_Sm* mutant, Related to Figure 5B. When uracil (0.4%) was fed to the Δ*nikO*::*malO_Sm* mutant, the production of Ψ-nikkomycin Z was increased 5-fold**.**

Table S1. Deduced Functions of ORFs in the *mal* Biosynthetic Gene Cluster of *Streptomyces malaysiensis* DSM 14702, Related to Table 1.

ORF[a]	Size[b]	Protein Homolog (Accession No.), Origin	I/S[c] (%)	Proposed Function
0224c	246	ribonuclease PH (WP_047176847) *Streptomyces* sp. MNU77	93/95	unknown function
0225 *coaE*	206	dephospho-CoA kinase (ACU97583) *Saccharomonospora viridis*	40/50	unknown function
0226c *truD*	353	tRNA ψ-uridine synthase D (AGZ78434) *Streptomyces chromofuscus*	98/99	provision of ψ-UMP
0227c *malA*	513	major facilitator transporter (AEZ64581) *Streptomyces chromofuscus*	97/98	transporter
0228c *malM*	245	putative hydroxylase (AEZ64582) *Streptomyces chromofuscus*	76/81	hydroxylase
0229c *malO*	479	UDP-GlcNac 1-carboxy-vinyl transferase (AEZ64583) *Streptomyces chromofuscus*	82/87	enoylpyruvyl-UMP synthase
0230c *malI*	208	nikkomycin biosynthesis protein NikI (AEZ64584) *Streptomyces chromofuscus*	83/90	hydroxylase
0231c *malB*	263	short chain reductase (AEZ64585) *Streptomyces chromofuscus*	83/86	unknown function
0232c *malL*	246	protein tyrosine/serine phosphatase (ABV97039) *Streptomyces chromofuscus*	72/81	octosyl acid phosphate hydrolase
0234c *malC*	257	no hit		unknown function
0237 *malJ*	454	radical SAM protein (ABX24496) *Streptomyces chromofuscus*	83/87	octosyl acid phosphate synthase
0238 *malK*	395	aminotransferase NikK (AEZ64589) *Streptomyces chromofuscus*	84/89	aminotransferase in AHA synthesis
0239 *malD*	490	carbamoyltransferase (AEZ64590) *Streptomyces chromofuscus*	92/95	carbamoyl transferase
0240	276	DNA topology modulation protein (AEZ64591) *Streptomyces chromofuscus*	64/69	unknown function
0241 *malE*	370	*O*-methyltransferase (AEZ64592) *Streptomyces chromofuscus*	81/87	methyltransferase
0242 *malR1*	308	transcriptional regulator ArsR (NP_033310326) *Streptomyces iakyrus*	45/57	regulator
0243	473	peptidase M20 (ABL83816) *Nocardiodes* sp. JS614	76/84	unknown function

[a]Suffix c denotes complementary strand
[b]Amino acids
[c]Identity/similarity

Table S2. Plasmids used in this work, Related to STAR Methods

Plasmid	Genotype/Characteristics	Reference
pYH7	*E.coli-Streptomyces* shuttle vector	Sun et al., 2010
pYH7-*truD*	*truD* gene (from *S. chromofuscus*) disruption construct in which a 966 bp internal fragment of *truD* was deleted in-frame	this work
pYH7-*malF*	*malF* gene (from *S. chromofuscus*) disruption construct in which a 684 bp internal fragment of *malF* was deleted in-frame	this work
pYH7-*malD*	*malD* gene (from *S. chromofuscus*) disruption construct in which a 1266 bp internal fragment of *malD* was deleted in-frame	this work
pYH7-*coaE*	*coaE* gene (from *S. chromofuscus*) disruption construct in which a 567 bp internal fragment of *coaE* was deleted in-frame	this work
pYH7-*nikO*	*nikO* gene (from *S. tendae*) disruption construct in which a 1422 bp internal fragment of *nikO* was deleted in-frame	this work
pGP9	*E.coli-Streptomyces* shuttle vector, *attP* (ΦBT1), *int*, PactI	Gregory et al., 2003
pGP9-*truD*	pseudouridine synthase *truD* (from *S. chromofuscus*) complementation plasmid	this work
pGP9-*nikO*	enoylpyruvyl-UMP synthase *nikO* (from *S. tendae*) complementation plasmid	this work
pGP9-*malD*	carbamoyltransferase *malD* (from *S. chromofuscus*) complementation plasmid	this work
pIB139	*E.coli-Streptomyces* shuttle vector, *attP* (ΦC31), *int*, PermE*	
pIB139-*truD_Sm*	pseudouridine synthase *truD* (from *S. malaysiensis*) complementation plasmid	this work
pIB139-*nikO*	enoylpyruvyl-UMP synthase *nikO* (from *S. tendae*) complementation plasmid	this work
pIB139-*malO_Sm*	enoylpyruvyl-UMP synthase *malO* (from *S. malaysiensis*) complementation plasmid	this work
pIB139-*malO_Sm*-*truD_Sm*	enoylpyruvyl-UMP synthase *malO* (from *S. malaysiensis*) together with pseudouridine synthase *truD* from (*S. malaysiensis*) complementation plasmid	this work
pET28a(+)	*E. coli* protein expression vector	Invitrogen
pET28a-*truD_Sm*	truD (from *S. malaysinesis*) protein expression construct with N-terminal His-tag based on pET28a(+)	this work
pET28a-*YeiN*	*E. coli* pseudouridine-5′-phosphate glycosidase yeiN protein expression construct with N-terminal His-tag based on pET28a(+)	this work
pET28a-*malO_Sm*	malO (from *S. malaysinesis*) protein expression construct with N-terminal His-tag based on pET28a(+)	this work
pET28a-*nikO*	nikO (from *S. tendae*) protein expression construct with N-terminal His-tag based on pET28a(+)	this work

Table S3. Oligonucleotide primers used in this work, Related to STAR Methods

Primer	Nucleotide sequence (5' to 3')
primers for protein expression	
TruD_Sm-p28_Fd	CTGGTGCCGCGCGGCAGCCATATGGAGTACCCCGTCCTCAAGGCCCAC
TruD_Sm-p28_Rv	TCCACCAGTCATGCTAGCCATATGCTAGGCGTTGGCGTTCCG GGGTGC
MalO_Sm-p28_Fd	CTGGTGCCGCGCGGCAGCCATATGGAACCTCGGCTGACCCCCTCCGTG
MalO_Sm-p28_Rv	TCCACCAGTCATGCTAGCCATATGTTACTCCTGGGA GCCCGG TGGGTC
NikO-p28_Fd	CTGGTGCCGCGCGGCAGCCATATGCAGGACCGTTGGCGAGAACCCGCT
NikO-p28_Rv	TCCACCAGTCATGCTAGCCATATGTCAGGCCGGGTCCACACCGGCCCC
YeiN-p28_Fd	CTGGTGCCGCGCGGCAGCCATATGTCTGAATTAAAAATTTCCCCT
YeiN-p28_Rv	TCCACCAGTCATGCTAGCCATATGTTAACCCACGAGACGCTGATATTC
primers for genes in-frame deletion based on pYH7 vector	
truD-L1	TGATCAAGGCGAATACTTCATATGTCTGGTACTTCAACTCCGAGTGGC
truD-L2	GGGCAGGAAGAGGGTCACTCCGTTCCTTTCGTCCTG
truD-R1	AACGGAGTGACCCTCTTCCTGCCCTCCGGCTGCTAC
truD-R2	CCGCGCGGTCGATCCCCGCATATGAGATCGCGAGCAGCAGACAGACGA
coaE-L1	TGATCAAGGCGAATACTTCATATGAGCGCATGAGGGTGTTGAAGCTGT
coaE-L2	CATCGTGGCGGAGTCCTTCGGCATCCGGCGATTCTA
coaE-R1	ATGCCGAAGGACTCCGCCACGATGACGGTCAGGACC
coaE-R2	CCGCGCGGTCGATCCCCGCATATGATCCTCTGGTGGACCCGTGGCTCT
malD_L1	GTCCACCGGGACTGATCAAGGCGAATACTTCATTCGAGTACGTCCACGAACTC
malD_L2	GAGCGGTACAGCAGACGCAAGAACAACACCAGTCTCAACATCAAGG
malD_R1	CCTTGATGTTGAGACTGGTGTTGTTCTTGCGTCTGCTGTACCGCTC
malD_R2	CTTCCCCGTCCGGGACCCGCGCGGTCGATCCCCGCAGTGGAAATGTGCGAC
malF_L1	GTCCACCGGGACTGATCAAGGCGAATACTTCAACGAGCTGAGGAAGAAGAAG
malF_L2	CTCACCTGCCGGGACATCTTCACCTCGCACGATCTGCGCATGCTG
malF_R1	CAGCATGCGCAGATCGTGCGAGGTGAAGATGTCCCGGCAGGTGAG
malF_R2	CTTCCCCGTCCGGGACCCGCGCGGTCGATCCCCGCAATCTGGTCTTCTC
nikO_L1	TGATCAAGGCGAATACTTCATATGACCCACCTTTTCTACTTCAACGAC
nikO_L2	AGGCCCAGGACGGGTGGTACCTCCTGAGCTGAGTGG
nikO_R1	GGAGGTACCACCCGTCCTGGGCCTCGACCGCGACGA
nikO_R2	CCGCGCGGTCGATCCCCGCATATGCGGTTGTATTCGTCCTGCACCACT
primers for PCR screening of deletion mutants	
truD-CP1	TGTCATACTTCTCCGTCGTCCTG
truD-CP2	ACCTACGCCGATGTCAAGGATCC
coaE-CP1	TACTGGACCGCAGTGCCTACATC
coaE-CP2	TGTTCTTCGTCAACTACTACGAC
malD-CP1	TTGATGGCAGAAGGCGGTCTTG
malD-CP2	TTCTACGCCAAGGACGAGCAC
malF-CP1	TTTCGTGAGGACGACCTGGATCATG
malF-CP2	ACGTCCCTATGGAGGAGCACATGAC
nikO-CP1	ACACGTTCGTCTGGCTCGACCTGC
nikO-CP2	TGCGCGGTCGATGAGCTGGTCGTA
primers for genes complementation using pGP9 vector	
truD-p9_Fd	GATTAATTAAGGAGGACA CATATG GAGTACCCCATCCTCAAGGCTCAC
truD-p9_Rv	GGCTGGGCAGAATAGGGC CATATG TCAGCGCTCCGACCCGGCCCAGGC
malD-p9_Fd	GATTAATTAAGGAGGACA CATATG GCGCACATCCTCGGCCTGTCCTGC
malD-p9_Rv	GGCTGGGCAGAATAGGGC CATATG TCACAGGACGGGGATGCCATGGAG
nikO-p9_Fd	GATTAATTAAGGAGGACA CATATG CAGGACCGTTGGCGAGAACCCGCT
nikO-p9_Rv	GGCTGGGCAGAATAGGGC CATATG TCAGGCCGGGTCCACACCGGCCCC
primers for genes complementation using pIB139 vector	
truD_Sm-p139_Fd	GCCGGT TGGTAG GATCCA CATATG GAGTACCCCGTCCTCAAGGCCCAC

truD_Sm-p139_Rv	CCTCTA GAGGAT CCCCAA CATATG CTAGGCGTTGGCGTTCCGGGGTGC
nikO-p139_Fd	GCCGGT TGGTAG GATCCA CATATG CAGGACCGTTGGCGAGAACCCGCT
nikO-p139_Rv	CCTCTA GAGGAT CCCCAA CATATG TCAGGCCGGGTCCACACCGGCCCC
malO_Sm-p139_Fd	GCCGGT TGGTAG GATCCA CATATG GAACCTCGGCTGACCCCCTCCGTG
malO_Sm-p139_Rv	CCTCTA GAGGAT CCCCAA CATATG TTACTCCTGGGAGCCCGGTGGGTC
truD_Sm-XbaI_Fd	GAGTAACATATGTTGGGGATCCTCTAGAgaaggaacggagtgaacATGGAGTACCC
truD_Sm-XbaI_Rv	GCGCGCGGCCGCGGATCCTCTAGACTAGGCGTTGGCGTTCCGGGGTGC

Table S4. NMR spectral data of 3'-enolpyruvyl-pseudouridine monophosphate (3'-EP-Ψ-MP) and 3'-enolpyruvyl-uridine monophosphate (3'-EPUMP) in [$^2H_6$]DMSO, Related to Figure 4.

	3'-EP-Ψ-MP[a]		3'-EPUMP[b]	
Position	δ_C	δ_H (multi, *J* in Hz)	δ_C	δ_H
C-2	151.3		151.7	
C-4	163.7		164.0	
C-5	**110.3**		**103.2**	**5.67**
C-6	140.5	7.57 (1H, s)		7.74 (1H, d)
C-1'	**78.4**	**4.60 (1H, d, 5.35)**	**88.5**	**5.88**
C-2'	72.0	4.30 (1H, m)	72.0	4.42
C-3'	77.8	4.36 (1H, m)	77.1	4.5
C-4'	79.1	4.05 (1H, m)	80.9	4.20
C-5'a	64.3	3.92 (1H, m)	65.7	4.06
C-5'b		3.80 (1H, m)		3.99
$=CH_2$a	95.9	5.21 (1H, s)	97.5	5.33
$=CH_2$b		4.81 (1H, s)		4.88
C=	151.0		150.9	
COOH	164.2		164.6	

[a]Measured at 500 MHz (^{1}H) and 125.7 MHz (^{13}C);

[b]Data from (Ginj et al., 2005). Measured at 500 MHz (^{1}H) and 125.9 MHz (^{13}C).

We present here a genomics-led approach to deciphering the biosynthetic origin of the Ψ C-nucleoside linkage in malayamycin A; and to identifying the factors that direct the biosynthesis of either N- or C-nucleosides. This has enabled engineering of the malayamycin A biosynthetic pathway to produce non-natural *N*-malayamycin, and the engineering of nikkomycin biosynthesis in *Streptomyces tendae* to increase production of the C-glycoside Ψ-nikkomycin Z. The origin of the Ψ moiety appears to be the unexpected repurposing of an enzyme catalyzing a universal editing step in RNA metabolism, in which specific uridine residues are isomerized into pseudouridine. These insights open up additional possibilities for the engineered biosynthesis of new nucleosides (Niu et al., 2017; Chen et al., 2017).

RESULTS AND DISCUSSION

Identification of the Malayamycin A, *mal*, Gene Cluster

Whole-genome next-generation de novo sequencing of the known malayamycin producer *Streptomyces malaysiensis* DSM14702 was carried out using the Illumina platform and the reads were assembled into two chromosomal scaffolds, one of 8,908,535 bp (SC1) and one of 1,727,442 bp (SC2). Bio-informatic analysis using the AntiSMASH program (Blin et al., 2016) revealed the presence of at least 39 biosynthetic gene clusters, among which was one predicted to encode biosynthesis of a polyketide nucleoside. The structure of malayamycin A indicates that a carbamoyltransferase should be required for the biosynthesis, and so in our initial search for the *mal* locus we used the sequence of an authentic carbamoyltransferase gene from the geldanamycin biosynthetic gene cluster (Genbank accession number ATL81095.1) in a BLAST (Altschul et al., 1990) search of the genome sequence. This identified a single carbamoyltransferase gene homolog, *malD*, in the same cluster as highlighted by AntiSMASH. Detailed inspection of this region revealed two adjacent clusters, one the gene cluster for a macrocyclic polyketide and the other a strong candidate to be the malayamycin biosynthetic gene cluster, comprising 20 predicted open reading frames (ORFs). These included several homologs of genes in the nikkomycin gene cluster whose functions have been previously elucidated (Figure 3A), including genes for enoylpyruvyl-UMP transferase (*malO*), for octosyl acid synthesis (*malJ*, *malL*), and for further processing of octosyl acid (*malI*, *malK*, *malM*).

The presence of the predicted carbamoyltransferase gene *malD* and methyltransferase gene *malF* are also fully consistent with the production of malayamycin. Strikingly, the candidate *mal* locus was also found to contain a gene with significant sequence similarity to the tRNA Ψ synthase gene *truD* (Charette and Gray, 2000; Hamma and Ferré-D'Amaré, 2006), which immediately suggested a possible origin of the *ψ*-uridine moiety in malayamycin (Figure 3A, see Table S1 for a complete list of predicted gene functions for the *mal* cluster of *S. malaysiensis*).

Gene sequences from the candidate malayamycin cluster were then used in BLAST searches against a library of >50 actinomycete genomes fully sequenced in-house, and a very significant match was found with an orphan gene cluster in *Streptomyces chromofuscus* ATCC 49982 (Genbank accession number JN671974) (See Table 1 for the complete list of predicted gene functions). Fermentation of *S. chromofuscus* revealed that it is indeed also a malayamycin A producer (Figure 3B). This strain proved more amenable to genetic manipulation than *S. malaysiensis*, so it was used in further genetic experiments.

Gene Deletions of *malF* and *malD*: *O*-Methylation and *N*-Carbamoylation are Late Steps in Malayamycin A biosynthesis

Methyltransferase gene *malF*, carbamoyltransferase gene *malD*, dephospho-CoA kinase gene *coaE*, and tRNA pseudouridine synthase gene *truD* were individually deleted (Figure S1). The in-frame deletion of the *coaE* gene did not alter malayamycin A production, suggesting either that the *coaE* gene lies outside the cluster or that it is involved but not essential. The in-frame deletion of the *malF* methyltransferase gene of *S. chromofuscus* led to the complete loss of malayamycin A production ($[M+H]^+$: *m/z* 343.2) and to the appearance of a peak ($[M+H]^+$: *m/z* 329.2) corresponding to the known compound desmethylmalayamycin A, a minor component of the wild type fermentation of *S. malaysiensis* (Benner et al., 2003) (Figure 3B). Similarly, in-frame deletion of the *malD* carbamoyltransferase gene led to the replacement of malayamycin A by a peak ($[M+H]^+$: *m/z* 300.2) corresponding to the new analog descarbamoyl-malayamycin A (Figure 3B). The identity of this metabolite was confirmed by MS/MS fragmentation (Figure S1). As expected, complementation of the Δ*malD* mutant with a wild type copy of *malD* restored malayamycin A production (Figure 3B). Only the Δ*malD* mutant revealed an

additional peak ($[M+H]^+$: *m/z* 286.2) that would correspond to descarbamoyl-desmethylmalayamycin A. The identity of this metabolite was confirmed by MS/MS fragmentation (Figure S1). These data taken together fully confirm the identity of the *mal* cluster, and suggest that, although the methyltransferase prefers carbamoylated substrate, there is no obligatory order in the timing of these two late-stage modifications.

Pseudouridine-5'-phosphate (5'-Ψ-MP) as a Precursor for Malayamycin Biosynthesis: Repurposing of a Universal RNA Editing Mechanism?

In principle, the Ψ nucleobase in malayamycin might be installed at either an early or a late stage in the biosynthetic pathway, but the finding of a gene in the *mal* biosynthetic gene cluster with significant sequence similarity to authentic TruD pseudouridine synthases (Ψ synthases) offered a substantial clue that 5'-Ψ-MP might be the key precursor. Across all kingdoms of life, cofactor-independent Ψ synthases belong to one of six different enzyme families named respectively after the *E. coli* enzymes TruA, TruB, RsuA, RluA and TruD and the archaeal enzyme Pus10 (Spenkuch et al., 2014). These enzymes catalyze the isomerization of specific U residues within particular RNA species into Ψ residues (Hamma and Ferré-D'Amaré, 2006). Although the biological functions of Ψ remain unclear, the additional H-bonding available to the free N1-H in Ψ residues is proposed to enhance base stacking and to increase the rigidity of the sugar phosphate backbone (Hamma and Ferré-D'Amaré, 2006; Spenkuch et al., 2014). TruD was originally discovered as the enzyme uniquely capable of catalyzing the isomerization of uridine-13 in $tRNA^{Glu}$ (Kaya and Ofengand, 2003). The TruD homologs encoded in the *mal* clusters of *S. malaysiensis* and *S. chromofuscus* both possess the conserved active site motifs of authentic TruD enzymes (Figure S2), consistent with their being active catalysts for this isomerization reaction. Subsequent turnover of tRNA to mononucleotides, catalyzed by cellular ribonucleases, would then generate 5'-Ψ-MP for incorporation into malayamycin. In *E. coli*, tRNA has been shown to become significantly unstable as a stress response (Sørensen et al., 2017), so it is plausible that the entry of *Streptomyces* into stationary phase, which triggers malayamycin biosynthesis, might analogously accelerate tRNA breakdown. Alternatively, the TruD homologs in the *mal* clusters, and the recently-reported similar TruD homolog in the pseudouridimycin

(PUM) biosynthetic gene cluster (Sosio et al., 2018) may have diverged to be active against a non-RNA uridine-based mononucleotide substrate. Sosio and colleagues suggest that either 5'-UMP or 5'-UDP is phosphorylated at the 3' position, catalyzed by PumH, a putative kinase annotated as an adenylate kinase. The hypothetical product is proposed to mimic an RNA-embedded uridine residue sufficiently for the TruD homolog PumJ to isomerize it, after which a putative phosphatase PumD is proposed to remove both 3'- and 5'-phospho groups to generate the Ψ needed for PUM biosynthesis. However, no genetic or biochemical evidence was adduced in support of this proposal. The malayamycin clusters do house an adenylate kinase, but there is no gene with significant similarity to PumD in either cluster (or indeed elsewhere in the genome).

To examine the role of TruD in malayamycin biosynthesis, the *truD* gene was specifically deleted from *S. chromofuscus*. Analysis of extracts from fermentation of the Δ*truD* mutant showed that malayamycin production levels were reduced to only 5-10% of wild-type levels (Figure 5A). The specificity of this effect was confirmed by complementation of the mutant *in trans* by a copy of the *truD* gene, which restored malayamycin production to wild-type levels (Figure 5A). We ascribe the residual level of malayamycin production in the Δ*truD* mutant to the background production of 5'-Ψ-MP, either from breakdown of RNA edited by other cellular Ψ synthases, or another as yet unidentified pathway. We also obtained TruD_Sm from *S. malaysiensis* as a purified recombinant protein in *E. coli*, and assayed it directly in vitro for its putative ability to convert 5'-UMP into 5'-Ψ-MP. No such activity was detectable. It has been argued (Hamma and Ferré-D'Amaré, 2006) that such an enzymatic activity would anyway be deleterious to the cell, both by depleting the pool of precursors for DNA synthesis; and by risking the random misincorporation of Ψ into RNA. The exquisite specificity of all known Ψ synthases, imposed by the RNA context, would neatly circumvent such potential problems, as would the phosphorylation-dephosphorylation pathway suggested by Sosio and colleagues (Sosio et al., 2018), particularly if the timing of expression of the genes were under tight control. Further work will be needed to clarify the substrate specificity of these TruD homologs.

There is also a widely-distributed salvage pathway present in bacterial cells for the degradation of 5'-Ψ-MP to uracil and ribose 5-phosphate, catalyzed by pseudouridine-5′-phosphate glycosidases such as YeiN from *E. coli* (Preumont et al.,

2008; Thapa et al., 2014), which might normally operate to keep levels of 5'-Ψ-MP low. In malayamycin-producing strains the glycosidase reaction might conceivably operate in the reverse direction as a glycosynthase, as has been recently shown for distinctive YeiN-like enzymes involved in biosynthesis of the blue pigment indigoidine (Takahashi et al., 2007), of the unusual polyketide alnumycin (Oja et al., 2012) and of the C-nucleoside showdomycin (Palmu et al., 2017) (Figure S3).

The malayamycin enoylpyruvyltransferase MalO, unlike its nikkomycin counterpart NikO, specifically accepts C-nucleoside 5'-Ψ-MP as substrate rather than 5'-UMP

The gene *nikO* encoding the nikkomycin 3'-enoylpyruvyl-transferase was cloned from *S. tendae* (Ginj et al., 2005) and NikO was expressed and purified as a recombinant protein from *E. coli*, as described in Experimental Procedures. Likewise, the *malO_Sm* gene from the malayamycin biosynthetic gene cluster of *S. malaysiensis* was cloned, and MalO_Sm was obtained as a purified protein from *E. coli*. Each enzyme was incubated with PEP and either 5'-UMP or 5'-Ψ-MP as substrate (Experimental Procedures) and the reaction mixtures were analyzed by HPLC. The analysis showed that 5'-UMP was as previously demonstrated (Ginj et al., 2005) an excellent substrate for NikO to yield 3'-EPUMP **8**, while in contrast no detectable reaction was found with MalO_Sm (Figure 4A).

The alternative substrate 5'-Ψ-MP was generated in situ through the reverse action of recombinant monophosphate glycosidase YeiN (Preumont et al., 2008) on uracil and ribose 5-phosphate (R5P). When incubated in the presence of either MalO_Sm or NikO, this led to the appearance of 3'-EP-Ψ-MP **11** (Figure 4B). The identity of **11** was confirmed by MS/MS (Figure S4) and NMR analysis of purified compound from a large scale enzymatic reaction (Experimental Procedures) (Table S4). Comparing the ^{1}H and ^{13}C NMR data of 3'-EPUMP (Ginj et al., 2005), the H-5 signal disappeared, the H-6 signal changed from a doublet peak to a singlet peak, and the chemical shift of H-1' shifted from 5.88 to 4.60, and the C-1' shifted from 88.5 to 78.4, suggesting a C-C link between C-5 and C-1'. Furthermore, the HMBC spectrum showed correlations of H-1' to C-5, C-4 and C-6, confirming the C-1' and C-5 linkage. The in vitro results strongly suggest that the pathway to the C-nucleoside malayamycin is as shown in Figure 6, in which the key precursor is 5'-Ψ-MP rather

peter leadlay 2/12/18 15:37
Deleted: '

than 5'-UMP. The adenosylcobalamin-dependent radical SAM enzyme MalJ is proposed to catalyze C-C bond formation as previously demonstrated for nikkomycin/polyoxin biosynthesis (Lilla and Yokoyama, 2016; He et al., 2017) to form Ψ-octosyl acid phosphate **12**, which is then acted upon by phosphatase MalL to give Ψ-octosyl acid **13**. Decarboxylation and amination, probably catalyzed by the combination of putative 2-oxoglutarate-Fe(II)-dependent oxygenases MalM and MalI, together with aminotransferase MalK, would give descarbamoyl-desmethylmalayamycin A **14**. The final steps involve the action of carbamoyltransferase MalF and O-methyltransferase MalD to provide malayamycin A 1 (Figure 6). The observed tolerance of NikO towards 5'-Ψ-MP as substrate in vitro also helps to account for the previous isolation of two pseudo-nikkomycins (differing from known nikkomycins only in the C-C rather than C-N attachment of the nucleobase) from fermentation extracts of mutant *S. tendae* (Heitsch et al., 1988; Decker et al., 1989).

Engineered biosynthesis of *N*-malayamycin and of *C*-nikkomycins

We reasoned that the different substrate specificity of MalO and NikO might be exploited to divert biosynthesis either towards N- or C-nucleosides. The *S. chromofuscus* Δ*truD* strain, which produces greatly reduced levels of malayamycin A (Figure 5A), was used as the starting point of an attempt to produce *N*-malayamycin by direct fermentation. This compound has only been obtained previously by an 18-step stereocontrolled synthetic route (Hanessian et al., 2005). The *nikO* gene from *S. tendae* was cloned on expression plasmid pGP9-*nikO* and used for conjugation into the *S. chromofuscus* Δ*truD* mutant (Experimental Procedures). LC-MS analysis of extracts from the fermentation of this strain (Figure 5A) revealed two peaks with the same $[M+H]^+$ of 343.2: the peak with a retention time of 13.7 minutes represents malayamycin A, and the second peak at 16 minutes is *N*-malayamycin A, as confirmed by high-resolution MS and NMR analysis (Figure S5). This result shows that all the enzymes of the malayamycin A pathway are indeed able to act on 3'-EPUMP and confirms that *N*-malayamycin can be obtained by fermentation. The switch to *N*-nucleoside production is however only partial, because 5'-Ψ-MP is still available in the Δ*truD* mutant and is also processed by MalO and NikO.

A complementary experiment was carried out on the biosynthetic pathway to nikkomycin Z **4** in *S. tendae*, to attempt diversion of biosynthesis towards Ψ-nikkomycin Z (Heitsch et al., 1988). First, the *nikO* gene of *S. tendae* was specifically deleted in-frame. This abolished the production of nikkomycin Z (Figure 5B). Complementation of the *S. tendae* Δ*nikO* mutant with a wild-type copy of *nikO* in trans successfully restored nikkomycin Z production. However, when the *truD_Sm* gene of *S. malaysiensis* was expressed in the *S. tendae* Δ*nikO* mutant strain, as expected it did not lead to production of Ψ-nikkomycin Z (Figure 5B). Encouragingly, when the enoylpyruvyl-transferase gene *malO_Sm* from *S. malaysiensis* was used to complement *S. tendae* Δ*nikO*, LC-MS analysis revealed a new, earlier-eluting peak with the same mass as nikkomycin Z, but almost no peak at the position expected for nikkomycin Z. It has been reported that Ψ-nikkomycin Z formation is boosted by feeding a high concentration (0.4%) of uracil to the wild-type *S. tendae* fermentation (Decker et al., 1988). Indeed, when uracil was fed to the wild-type culture, the same earlier-eluting peak as observed in the Δ*nikO::malO_Sm* fermentation appeared with the same *m/z* at 496.4 as nikkomycin Z (Figure 5B). The MS/MS spectrum of this species was almost identical to that of nikkomycin Z (Figure S6). However, MS^3 on *m/z* 288.2 of the two showed significant differences. In the MS^3 of nikkomycin Z, a fragment at *m/z* 176.2 was generated due to C-N cleavage between uracil and ribose moiety. MS^3 spectrum of the earlier eluted peak did not show this characteristic fragment, instead a fragment at *m/z* 208.2 was generated, which was in agreement with the C-C glycosidic linkage in Ψ-nikkomycin Z (Figure S6). These results were in full agreement with the in vitro specificity we have determined for MalO_Sm. Expression of both *truD_Sm* and *malO_Sm* in *S. tendae* Δ*nikO* did not further alter the amount of Ψ-nikkomycin Z formed (Figure 5B), so a step other than that catalyzed by TruD_Sm must be rate-limiting. A YeiN-like monophosphate glycosidase is encoded in the *S. tendae* genome (Figure S3) and this enzyme may be responsible for the stimulation of Ψ-nikkomycin Z production by uracil. When uracil was fed to the Δ*nikO*::*malO_Sm* mutant, the production of Ψ-nikkomycin Z was increased 5-fold (Figure S7).

SIGNIFICANCE

There is an urgent need for new antifungal compounds for both plant protection and clinical use. The uridine C-nucleoside malayamycin is a promising lead compound, but exploration of analogs and their activity has so far been limited by the complexity of the synthetic route. We report here the discovery of the biosynthetic gene cluster (*mal*) for malayamycin A from two different *Streptomyces* strains, which has revealed the molecular basis for C-nucleoside formation. The pathway parallels the early steps of biosynthesis of known uridine-based *N*-nucleosides, except that the key nucleotide building block is 5'-pseudouridine monophosphate (5'-Ψ-MP) instead of 5'-UMP. The key determinant of C-nucleoside formation is the malayamycin 3'-enoylpyruvyltransferase MalO, which acts specifically on 5'-Ψ-MP and not on 5'-UMP. We show that the supply of 5'-Ψ-MP for malayamycin biosynthesis may depend either upon the repurposing of a universal RNA editing mechanism, in which a specific U residue in RNA is converted to Ψ by the action of a TruD-like enzyme encoded in the malayamycin biosynthetic gene cluster; or alternatively on the evolution of such an enzyme to isomerize a uridine mononucleotide substrate. These insights have allowed engineering of the malayamycin pathway to create analogs by fermentation, and suggest the presence of a TruD gene is sufficient to identify an orphan actinomycete gene cluster as leading to a pseudouridine-based antibiotic.

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and four tables and can be found with this article online at XXXXX.

ACKNOWLEDGMENTS

This work was co-funded by BBSRC and Syngenta Ltd via a strategic Long and Large (sLoLa) project grant (BB/K002341/1) awarded jointly to the Universities of Warwick, Manchester, Bristol and Cambridge (P.F.L.). We thank Syngenta for kind provision of the malayamycin analytical reference standard, and for advice on extraction and purification of malayamycin; and S. Dickens, N. Scott and R. Deglau of the Department of Biochemistry Sequencing Facility for genome sequencing,

AUTHOR CONTRIBUTIONS

H.H. and P.F.L. designed the experiments and wrote the manuscript. M.S. sequenced, assembled and annotated genomes. H.H. and Y.Z. conducted the experiments.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Benner, J. P., Boehlendorf, B. G. H., Kipps, M. R., Lambert, N. E. P., Luck, R., Molleyres, L.-P., Neff, S., Schuez, T. C., and Stanley, P. D. (2003). WO 03/062242 [Chem. Abstr. *139*, 132519]. Biocidal compounds and their preparation.

Blin, K., Wolf, T., Chevrette, M. G., Lu, X., Schwalen, C. J., Kautsar, S. A., Suarez Duran, H. G., de Los Santos, E. L. C., Kim, H. U., Nave, et al. (2017). antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. Nucl. Acids Res. *45*, W36-W41.

Bormann, C. Möhrle, V., and Brundtner, C. (1996). Cloning and heterologous expression of the entire set of structural genes for nikkomycin biosynthesis from *Streptomyces tendae* Tü901 in *Streptomyces lividans*. J. Bacteriol. *178*, 1216–1218.

Charette, M., and Gray, M.W. (2000). Pseudouridine in RNA: What, where, how, and why. IUBMB Life *49*, 341–351.

Chen, S., Kinney, W. A., and Van Lanen, S. (2017). World J. Microbiol. Biotechnol. *33*, 66.

Chen, W., Huang, T., He, X., Meng, Q., You, D., Bai, L., Li, J., Wu, M., Li, R., Xie, Z., et al. (2009). Characterization of the polyoxin biosynthetic gene cluster from *Streptomyces cacaoi* and engineered production of polyoxin H. J. Biol. Chem. *284*, 10627–10638.

Dähn, U Hagenmeier, H, Höhne, H., König, W. A., Wolf, G. and Zähner, H. (1976) . Stoffwechselprodukte von mikroorganismen. 154. Arch. Microbiol. *107*, 143–160.

Decker, H., Bormann, C., Fiedler, H.-P., and Zähner, H. (1988). Metabolic products of microorganisms. 252. Isolation of new nikkomycins from *Streptomyces tendae*. J. Antibiot. *42*, 230–235.

peter leadlay 4/12/18 12:37
Deleted: -

peter leadlay 4/12/18 12:37
Formatted: Font:Italic

peter leadlay 4/12/18 12:37
Deleted: :

Ericsson, U. B., Nordlund, P., and Hallberg, B. M. (2004). X-ray structure of tRNA pseudouridine synthase TruD reveals an inserted domain with a novel fold. FEBS Lett. *565*, 59–64.

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchinson III, C. A., and Smith, H. O. (2009). Nat. Methods *6*, 343–345.

Ginj, C., Rüegger, H., Amrhein, N., and Macheroux, P. (2005). 3'-enolpyruvoyl-UMP, a novel and unexpected metabolite in nikkomycin biosynthesis. ChemBioChem. *6*, 1974–1976.

Gregory, M. A., Till, R., and Smith, M. C. (2003). J. Bacteriol. *185*, 5320–5323.

Hamma, T., and Ferré-D'Amaré, A. R. (2006). Pseudouridine synthases. Chem. Biol. *13*, 1125–1135.

Hanessian, S., Huang, G., Chenel, C., Machaalani, R., and Loiseleur, O. (2005). Total synthesis of N-malayamycin A and related bicyclic purine and pyrimidine nucleosides. J. Org. Chem. *70*, 6721–6734.

He, N., Wu, P., Lei, Y., Xu, B., Zhu, X., Xu, G., Gao, Y., Qi, J., Deng, Z., Tang, G., et al. (2017). Construction of an octosyl acid backbone catalyzed by a radical S-adenosylmethionine enzyme and a phosphatase in the biosynthesis of high-carbon sugar nucleoside antibiotics. Chem. Sci. *8*, 444–451.

Heitsch, H., König, W. A., Decker, H., Bormann, C., Fiedler, H.-P., and Zähner, H. (1988). Metabolic products of microorganisms. 254. Structure of the new nikkomycins pseudo-X and pseudo-J. J. Antibiot. *42*, 711-717.

Hoang, C., and Ferre-D'Amaré, A. R. (2004). Crystal structure of the highly divergent pseudouridine synthase TruD reveals a circular permutation of a conserved fold. RNA *10*, 1026–1033.

Kaya, Y., and Ofengand, J. (2003). A novel unanticipated type of pseudouridine synthase with homologs in bacteria, archaea, and eukarya. RNA *9*, 711–721.

Kaya, Y., Del Campo, M., Ofengand, J., and Malhotra, A. (2004). Crystal structure of TruD, a novel pseudouridine synthase with a new protein fold. J. Biol. Chem. *279*, 18107–18110.

Kieser, T., Bibb, M., Buttner, M., Chater, K. F., and Hopwood, D. A. (2001). Practical *Streptomyces Genetics*. The John Innes Foundation, Norwich, UK.

Li, W., Csukai, M., Corran, A., Crowley, P., Solomon, P. S., and Oliver, R. P. (2008). Malayamycin, a new streptomycete antifungal compound, specifically inhibits

sporulation of *Stagonospora nodorum* (Berk) Castell and Germano, the cause of wheat glume blotch disease. Pest Manag. Sci. *64*, 1294–1302.

Lilla, E. A., and Yokoyama, K. (2016). Carbon extension in peptidylnucleoside biosynthesis by radical SAM enzymes. Nat. Chem. Biol. *12*, 905–907.

MacNeil, D. G., Gewain, K. M., Ruby, C. L., Dezeny, G., Gibbons, P. H., and MacNeil, T. (1992). Gene *111*, 61–68.

Maffioli, S. I., Zhang, Y., Degen, D., Carzaniga, T., Del Gatto, G., Serina, S., Monciardini, P., Mazzetti, C., Guglierame, P., Candiani, G., et al. (2017). Antibacterial nucleoside-analog inhibitor of bacterial RNA polymerase. Cell *169*, 1240–1248.

Niu, G., Zheng, J., and Tan, H. (2017). Biosynthetic and combinatorial biosynthesis of antifungal nucleoside antibiotics. Science China Life Sciences *60*, 939–947.

Oja, T., Klika, K. D., Appassamy, L., Sinkkonen, J., Mäntsälä, P., Niemi, J., and Metsä-Ketelä, M. (2012). Biosynthetic pathway toward carbohydrate-like moieties of alnumycins contains unusual steps for C-C bond formation and cleavage. Proc. Natl. Acad. Sci. USA *109*, 6024–6029.

Palmu, K., Rosenqvist, P., Thapa, K., Ilina, Y., Siitonen, V1., Baral, B., Mäkinen, J., Belogurov, G., Virta, P., Niemi, J., et al. (2017). Discovery of the showdomycin gene cluster from *Streptomyces showdoensis* ATCC 15227 yields insight into the biosynthetic Logic of C-nucleoside antibiotics. ACS Chem. Biol. *12*, 1472–1477.

Pesic, A., Steinhaus, B., Kemper, S., Nachtigall, J., Kutzner, H. J., Höfle, G., and Süssmuth RD. (2014). Isolation and structure elucidation of the nucleoside antibiotic strepturidin from *Streptomyces albus* DSM 40763. J. Antibiot. *67*, 471–477.

Preumont, A., Snoussi, K., Stroobant, V., Collet, J. F., and Van Schaftingen, E. (2008). Molecular identification of pseudouridine-metabolizing enzymes. J. Biol. Chem. *283*, 25238–25246.

Sakata, K., Sakurai, A., and Tamura, S. (1974). Structures of ezomycins A_1 and A_2. Tetrahedron Lett. *49*, 4327–4330.

Sambrook, J., and Russell, D. W. (2001). Molecular Cloning: A Laboratory Manual, 3rd edn. Cold Spring Harbor Laboratory Press, New York, USA.

Sørensen, M. A., Fehler, A. O., and Svenningsen, S. L. (2018). Transfer RNA instability as a stress response in *Escherichia coli*: rapid dynamics of the tRNA pool as a function of demand. RNA Biol. *15*, 586–593.

peter leadlay 4/12/18 12:35
Deleted: 7

peter leadlay 4/12/18 12:36
Formatted: Font:Italic

Sosio, M., Gaspari, E., Iorio, M., Pessina, S., Medema, M. H., Bernasconi, A., Simone, M., Maffioli, S. I., Ebright, R. H., and Donadio, S. (2018). Analysis of the pseudouridimycin biosynthetic pathway provides insights into the formation of C-nucleoside antibiotics. Cell Chem. Biol. *25*, 540–549.

Spenkuch, F., Motorin, Y., and Helm, M. (2014). Pseudouridine: still mysterious, but never a fake (uridine)! RNA Biol. *11*, 1540–1554.

Sun, Y., Hahn, F., Demydchuk, Y., Chettle, J., Tosin, M., Osada, H., and Leadlay, P. F. (2010). Nat. Chem. Biol. *6*, 99–101.

Svenningsen, S. L., Kongstad, M., Stenum, T. S., Muñoz-Gómez, A. J., and Sørensen, M. A. (2017). Transfer RNA is highly unstable during early amino acid starvation in *Escherichia coli*. Nucl. Acids Res. *45*, 793–804.

Takahashi, H., Kumagai, T., Kitani, K., Mori, M., Matoba, Y., and Sugiyama, M. (2007). Cloning and characterization of a *Streptomyces* single module type non-ribosomal peptide synthetase catalyzing a blue pigment synthesis. J. Biol. Chem. *282*, 9073–9081.

Thapa, K., Oja, T., and Metsä-Ketela, M. (2014). Molecular evolution of the bacterial pseudouridine-5'-phosphate glycosidase protein family. FEBS J. *281*, 4439–4449.

Winn, M., Goss, R. J., Kimura, K., and Bugg, T. D. (2010). Antimicrobial nucleoside antibiotics targetting cell wall assembly. Nat. Prod. Rep. *27*, 279–304.

peter leadlay 4/12/18 12:35
Deleted: Oct 12:0. doi: 10.1080/15476286.2017.1391440. [Epub ahead of print].

peter leadlay 4/12/18 12:36
Deleted: -

peter leadlay 4/12/18 12:33
Formatted: Font:Cambria

peter leadlay 4/12/18 12:33
Formatted: Font:Cambria

Main Figure titles and legends

Figure 1. Structures of uracil-containing nucleoside natural products

Figure 2. The biosynthetic pathway to nikkomycin and polyoxin N-nucleosides

Figure 3. Production of malayamycin from *Streptomyces chromofuscus* and mutants

(A) Malayamycin biosynthetic gene cluster in *Streptomyces chromofuscus* and *Streptomyces malaysiensis*.

(B) LC-MS analysis of malayamycin A standard ($[M+H]^+$: 343.2) and its analogs produced by *S. chromofuscus* wild-type (WT_Sc) and mutant strains: the methyltransferase (*malF*) deletion mutant (Δ*malF*) produced desmethylmalayamycin A ($[M+H]^+$: 329.2); the carbamoyltransferase (*malD*) deletion mutant (Δ*malD*) produced both descarbamoylmalayamycin A ($[M+H]^+$: 300.2) and the descarbamoyl-desmethyl-malayamycin A ($[M+H]^+$: 286.2). Complementation of the Δ*malD* mutant with carbamoyltransferase (Δ*malD*::*malD*) restored malayamycin A production. See also Figure S1.

peter leadlay 2/12/18 15:38
Deleted: s

Figure 4: HPLC Analysis of the Reactions Catalyzed by NikO and MalO_Sm monitored at UV of 262 nm
(A) with UMP as substrate **I**: uridine monophosphate (UMP) standard; **II**: formation of **8** (3'-EPUMP) from UMP in a reaction containing PEP, UMP and NikO; **III**: almost no conversion of UMP in the presence of PEP and MalO_Sm.
(B) with 5'-Ψ-MP as substrate **I**: 5'-pseudouridine monophosphate (5'-Ψ-MP) was generated by uracil, R5P and YeiN; **II**: formation of **11** (3'-EP-Ψ-MP) in a reaction containing uracil, R5P and YeiN, PEP and NikO; **III**: formation of **11** (3'-EP-Ψ-MP) in a reaction containing uracil, R5P and YeiN, PEP and MalO_Sm. See also Figures S3, S4, Table S4 and Data S1.

peter leadlay 4/12/18 12:01
Deleted: '

peter leadlay 4/12/18 12:01
Deleted: '

peter leadlay 4/12/18 11:58
Deleted: and

peter leadlay 4/12/18 11:58
Deleted: .

Figure 5. LC-MS analysis of nucleosides produced by *S. chromofuscus*, *S. tendae* and mutants
(A) LC-MS ($[M+H]^+$: 343.2) analysis of malayamycin A and *N*-malayamycin A. When *truD* was deleted (Δ*truD*), production of malayamycin A was reduced to 5-10% of the amount from wild-type *S. chromofuscus* (WT_Sc). Complementation of *truD* to the Δ*truD* mutant (Δ*truD*::*truD*) rescued malayamycin A production. When the *nikO* gene of nikkomycin cluster was complemented into the Δ*truD* mutant (Δ*truD*::*nikO*). *N*-linked malayamycin (*N*-malayamycin A) was produced along with malayamycin A.
(B) LC-MS ($[M+H]^+$: 496.4) analysis of nikkomycin Z and pseudo-nikkomycin Z production from *S. tendae* and its mutant strains. *nikO* deletion mutant (Δ*nikO*) abolished nikkomycin Z production. Using *nikO* for complementation of the Δ*nikO* mutant (Δ*nikO*::*nikO*) rescued nikkomycin Z production. When the *malO* gene of malayamycin cluster from *S. malaysiensis* was used to complement the Δ*nikO* mutant (Δ*nikO*::*malO_Sm*). *C*-linked nikkomycin Z (Ψ-nikkomycin Z) was almost exclusively produced. There was no significant increase in the amount of Ψ-nikkomycin Z production when *malO* and *truD* were together used to complement the Δ*nikO* mutant (Δ*nikO*::*malO_Sm*). As reported in the literature (Decker et al., 1988), Ψ-nikkomycin Z production was boosted when high concentration of uracil (0.4%) was fed to the wild-type *S. tendae*. See also Figures S2, S5 and S6.

peter leadlay 2/12/18 19:05
Deleted: S3

Figure 6. The proposed biosynthetic pathway to malayamycin A

Main Tables

Table 1. Deduced Functions of ORFs in the *mal* Biosynthetic Gene Cluster of *Streptomyces chromofuscus* ATCC 49982

STAR Methods text

Contact for Reagent and Resource Sharing

"Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Peter F Leadlay (pfl10@cam.ac.uk)"

Experimental Model and Subject Details

Bacterial Strains and Culture Conditions

Streptomyces malaysiensis DSM 14702 (malayamycin A-producing strain) was obtained from the Leibniz Institute DSMZ, the German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany. *Streptomyces chromofuscus* ATCC 49982 and *Streptomyces tendae* ATCC 31160 were obtained from the American Type Culture Collection (ATCC) via LGC Standards, Teddington, U.K. All strains were maintained on SFM agar (2% soya flour (AYKASOY), 2% D-mannitol, 2% agar) at 30°C. *E. coli* strains were grown in Luria-Bertani (LB) broth (10% tryptone, 5% yeast extract, 10% NaCl) or agar (10% tryptone, 5% yeast extract, 10% NaCl, 2% agar) at 37°C with appropriate antibiotic selection (kanamycin, at 50 μg ml^{-1}).

Method Details

DNA Isolation and Manipulations

Plasmids and oligonucleotides (Sigma-Aldrich) used in this work are summarized in Supplementary Tables S3 and S4 respectively. Restriction endonuclease digestion (NewEngland Biolabs Inc), PCR amplifications using Phusion® High-Fidelity PCR Master Mix (NewEngland Biolabs Inc) and ligation using Gibson Assembly® Master Mix (NewEngland Biolabs Inc) were carried out according to the manufacturers' protocols. Liquid cultures for isolation of genomic DNA were grown in tryptone soya broth (Difco). DNA isolation and manipulation in *Streptomyces*, and *E. coli* were carried out using standard protocols (Sambrook and Russell, 2001; Kieser et al., 2001).

Metabolite analysis

For small-scale analysis, Streptomyces chromofuscus and Streptomyces tendae were grown in liquid SFM medium (2% soya flour (ARKASOY), 2% d-mannitol) at 30°C and 200 rpm in a rotary incubator for 4 days. 0.5 mL samples of culture broth were centrifuged at 20,000 x *g* for 15 min. 100 μL of clear supernatant was analyzed by LC-MS. LC-MS analyses were performed on a HPLC (Agilent Technologies 1200 series) coupled to a Thermo Fisher LTQ mass spectrometer fitted with an electrospray

ionization (ESI) source. A Prodigy 5µ C18 column (4.6 x 250 mm, Phenomenex) was used, and the samples were eluted using MilliQ-deionised/distilled water (MQ) containing 0.1% trifluoroacetic acid (A) and methanol (B) at a flow rate of 0.7 ml min^{-1}. The linear elution gradient for analysis of malayamycin A and its analogues was 2% B for 8 min, 2% to 12% B over 10 min, 12% to 98% B over 1 min, 98% B for 10 min, 98% to 2% B over 1 min. The elution gradient for nikkomycin analysis was 2% B for 5 min, 2% to 15% B over 13 min, 15% to 98% B over 1 min, 98% B for 10 min, 98% to 2% B over 1 min. The mass spectrometer was run in positive ionization mode, scanning from *m/z* 200 to 2000 in full scan mode. MS/MS analysis were performed on $[M+H]^+$ ions with a normalized collision energy of 10% to 15%. High-resolution UPLC-MS was carried out on a Waters Xevo G2-S Q-TOF for the semi-purified mixture of malayamycin A and *N*-malayamycin A. A Waters Acquity instrument fitted with a Acquity UPLC BEH 1.7µ C18 column (2.1 x 50 mm) was used, and the sample was eluted using MQ containing 0.1% formic acid (A) and acetonitrile containing 0.1% formic acid (B) at a flow rate of 0.21 ml min^{-1}. The linear elution gradient was 2% B for 4.4 min, 2% to 6% B over 5 min, 6% to 100% B over 1.6 min, 100% to 2% B over 1 min. The ESI was operated in positive mode.

Construction of gene knock-out plasmids

Recombinant plasmids based on the pYH7 vector were constructed by ligating DNA fragments (about 2 kb) PCR-amplified from the upstream and downstream flanks of the target gene into cloning vector pYH7, which was digested with *NdeI*, treated with shrimp alkaline phosphatase (SAP) and gel purified. To ligate the fragments corresponding to the left and right flanking regions of the target gene into pYH7, the isothermal assembly method was used as described (Gibson et al., 2009). The mixture was incubated at 50°C for 60 min, and was used to transform *E. coli* DH10B. The integrity of all recombinant plasmids was checked by restriction digestion and sequencing.

truD*, *malF*, *mal D* and *coaE* gene knock-out in *S. chromofuscus

The constructs used in this study for gene knock-out are summarized in Supplementary Table S2. The pYH7 based constructs were transformed into *E. coli* ET12657/pUZ8002 and then introduced by conjugation into mycelia of the

Streptomyces chromofuscus WT strain. Conjugations were carried out on 20 ml of SFM plates (2% mannitol, 2% soya flour, 2% agar). After incubating at 30°C for around 17 hours, exconjugants were selected with 10 μg ml^{-1} apramycin and 25 μg ml^{-1} nalidixic acid. Single colonies from this plate were transferred to a SFM plate containing 50 μg ml^{-1} apramycin to check for antibiotic resistance. Mutant screening was carried out by streaking transformants on SFM agar medium for non-selective growth, then patching single colonies in parallel onto SFM agar and SFM agar containing apramycin (50 μg ml^{-1}). Candidate colonies with the correct phenotype (Apr^S) were selected for further screening by PCR using screening primers listed in Table S3 to identify double cross-over mutants. The PCR fragments from the double cross-over mutants were further verified by sequencing.

nikO* gene knock-out in *S. tendae

The pYH7-*nikO* plasmid was transformed into *E. coli* ET12657/pUZ8002 and then introduced by conjugation into spores of the *Streptomyces tendae* WT strain. Spores from SFM plates were collected, washed twice with 2TY and heat-shocked at 50°C for 10 min before conjugation. Conjugations were carried out on 20 ml of SFM plates. After incubation at 30°C for 17 hours, exconjugants were selected with 25 μg ml^{-1} apramycin and 25 μg ml^{-1} nalidixic acid. Single colonies from this plate were transferred to a SFM plate containing 50 μg ml^{-1} apramycin to double check for antibiotic resistance. Mutants screening were carried out by streaking transformants on SFM agar medium for non-selective growth, then patching single colonies onto both SFM agar and SFM agar containing apramycin (50 μg ml^{-1}) in parallel. Candidate colonies with the correct phenotype (Apr^S) were selected for further screening by PCR using screening primers listed in Supplementary Table S3 to identify double cross-over mutants. The PCR fragments from the double cross-over mutants were further verified by sequencing.

Complementation of *truD*, *nikO* and *malD* genes into *S. chromofuscus* deletion mutants

The tRNA pseudouridine synthase *truD* and carbamoyltransferase *malD* were PCR amplified from genomic DNA of *S. chromofuscus*, using primer pairs truD-p9_Fd/Rv and malD-p9_Fd/Rev, respectively. The enoylpyruvyl-UMP synthase *nikO* was PCR amplified from genomic DNA of *S. tendae*, using primer pair nikO-p9_Fd/Rv. The

cloning vector pGP9 was digested with *NdeI* and gel purified. The *truD*, *malD* and *nikO* PCR fragments were ligated by the isothermal assembly method with the digested pGP9 plasmid, to yield plasmids pGP9-*truD*, pGP9-*malD* and pGP9-*nikO*. The constructs were then introduced by conjugation into Δ*truD*, Δ*malD* and Δ*truD* strains, respectively, to generate Δ*truD*::*truD*, Δ*malD*::*malD and* Δ*truD*::*nikO* complementation strains. The conjugation procedure was the same as described in 1.4. After incubating at 30°C for around 17 hours, exconjugants were selected with 10 μg ml^{-1} apramycin and 25 μg ml^{-1} nalidixic acid. Single colonies from this plate were transferred to a SFM plate containing 50 μg ml^{-1} apramycin to double check for antibiotic resistance. The patch from the confirmation plate was then inoculated into SFM liquid culture containing 50 μg ml^{-1} apramycin for metabolites production.

Complementation of *nikO, malO_Sm, truD_Sm* and *malO_Sm-truD_Sm* genes into the *S. tendae nikO* deletion mutant

The *S. tendae nikO*-deletion mutant (Δ*nikO*) was complemented with native enoylpyruvyl-UMP synthase *nikO*, as well as enoylpyruvyl-UMP synthase *malO_Sm*, tRNA pseudouridine synthase *truD_Sm* and *malO_Sm-truD_Sm* from *S. malaysinensis*. Gene *nikO* was PCR amplified from genomic DNA of *S. tendae*, using primer pair nikO-p139_Fd/Rv. Genes *truD_Sm* and *malO_Sm* were PCR amplified from genomic DNA of *S. malaysiensis*, using primer pairs truD_Sm-p139_Fd/Rv and malO_Sm-p139_Fd/Rv, respectively. The cloning vector pIB139 was digested with *NdeI* and gel purified. The *nikO*, *truD_Sm* and *malO_Sm* PCR fragments were ligated by the isothermal assembly method with the digested pIB139 vector, to yield plasmids pIB139-*nikO*, pIB139-*truD_Sm* and pIB139-*malO_Sm*, respectively. To create plasmid pIB139-*malO_Sm-truD_Sm*, the plasmid pIB139-*malO_Sm* was digested with *XbaI*, and purified by gel. The *truD_Sm* including its RBS was PCR amplified from genomic DNA of *S. malaysinensis*, using primer pair truD_Sm-XbaI_Fd/Rv. This *truD_Sm* PCR fragment was then ligated by the isothermal assembly method with the *XbaI*-digested pIB139-*malO_Sm* to yield pIB139-*malO_Sm-truD_Sm* plasmid.

The constructs were then introduced by conjugation into Δ*nikO* mutant, to generate Δ*nikO*::*nikO*, Δ*nikO*::*truD_Sm*, Δ*nikO*::*malO_Sm* and Δ*nikO*::*malO_Sm-truD_Sm*. The conjugation procedure was the same as described in 1.5. After incubating at 30°C

for around 17 hours, exconjugants were selected with 25 μg ml^{-1} apramycin and 25 μg ml^{-1} nalidixic acid. Single colonies from this plate were transferred to a SFM plate containing 50 μg ml^{-1} apramycin to double check for antibiotic resistance. The patch from the confirmation plate was then inoculated into SFM liquid culture containing 50 μg ml^{-1} apramycin for metabolites production.

Production, purification and NMR analysis of *N*-malayamycin

For the production of *N*-malayamycin, 100 of 20 ml SFM agar plates containing 50 μg ml^{-1} apramycin were set up. Each plate was innoculated with 0.3 ml of 1-day TSBY culture of *S. chromofuscus* Δ*truD*::*nikO* mutant, in which enolpyruvyltransferase gene (*nikO*) from nikkomycin cluster was complemented into the tRNA pseudouridine synthase gene (*truD*) deletion mutant of *S. chromofuscus*. The plates were incubated at 30 °C for 12 days. After the incubation, plates were extracted with MQ water twice. The aqueous solution was then lyophilized. The dried residue was dissolved in methanol, and spun thoroughly to remove insoluble debris. The methanol sample was applied onto a preparative HPLC (Agilent 1200) fitted with a Luna 10μ C18 column (100Å, 21.20 x 250 mm, Phenomenex). Compounds were eluted with MQ (A) and acetonitrile containing 0.1% formic acid (B) at a flow rate of 20 ml min^{-1}. The linear gradient was 0% B for 5 min, 0% to 10% B over 10 min, 10% to 100% B over 5 min, 100% B for 4 min, 100% to 0% B over 1 min. HPLC elution was monitored at 262 nm. Fractions were further checked by LC-MS analysis. Fractions containing *N*-malayamycin were combined. Acetonitrile was removed under reduced pressure, and sample was lyophilized. The dried sample was dissolved in a minimal amount of methanol, and further purified on a ThermoHypersil BDS C8 analytical column (5μ, 4.6 x 250 mm), eluting with MQ (A) and acetonitrile (B) at a flow rate of 1 ml min^{-1}. The gradient was 0% B for 5 min, 0% to 6% B over 10 min, 6% to 0% B over 1 min. Fractions containing *N*-malayamycin were combined, and acetonitrile was removed under reduced pressure, and the sample was lyophilized. The dried residue was dissolved in [$^2H_4$]-methanol, and analysed on a Brucker-500 NMR spectrometer.

Protein expression and purification

The tRNA pseudouridine synthase *truD_Sm* and enolpyruvyl transferase *malO_Sm* genes were PCR amplified from genomic DNA of *Streptomyces malaysiensis*, using primer pairs TruD_Sm-p28_Fd/Rv and MalO_Sm-p28_Fd/Rv, respectively. The *nikO* gene was PCR amplified from genomic DNA of *S. tendae*, using primer pair NikO-p28_Fd/Rv. The *E. coli* pseudouridine-5′-phosphate glycosidase *yeiN* gene was PCR amplified from genomic DNA of *E. coli* K12, using primer pair YeiN-p28_Fd/Rv. The cloning vector pET28a(+) was digested with *NdeI* and gel purified. The *nikO*, *yeiN, truD_Sm* and *malO_Sm* PCR fragments were ligated by the isothermal assembly method with the digested pET28a(+) vector, to yield plasmids pET28a-*nikO*, pET28a-*yeiN*, pET28a-*truD_Sm* and pET28a-*malO_Sm*, respectively.

The pET28a-*nikO*, pET28a-*yeiN*, pET28a-*truD_Sm* and pET28a-*malO_Sm* plasmids were then used to transform *E. coli* BL21(DE3) for protein expression. A single colony was inoculated into 10 mL of LB medium containing 50 μg ml^{-1} kanamycin and grown overnight at 37°C, 250 rpm. The overnight culture was used to inoculate 1 L LB medium containing 50 μg ml^{-1} kanamycin and incubated at 37°C, 200 rpm until A_{600} reached 0.6 before addition of isopropyl-β-d-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM, and incubation at 22°C overnight to induce protein expression. Cells were harvested by centrifugation at 4,000 rpm for 10 min, resuspended in lysis buffer (20 mM Tris-HCl, pH 7.8, 0.5 M NaCl, 10 mM imidazole) and lysed by sonication. The total lysate was centrifuged at 14,000 x *g* for 40 min, and the supernatant was loaded onto a His-Bind column, which had been pre-charged with nickel ions and equilibrated with lysis buffer. The column was washed with 10 column volumes of lysis buffer. Bound proteins were then eluted with a step gradient of increasing imidazole concentration (40, 80, 100, 150, 200, 250 and 500 mM in binding buffer). The fractions containing the expected protein were pooled, concentrated and buffer exchanged to 50 mM Tris-HCl, 0.15 M NaCl, pH 7.7. The purified proteins were aliquoted and stored at -70°C.

***In vitro* activity assays of TruD_Sm, NikO and MalO_Sm**

To test the activity of TruD_Sm, several candidate substrates were tried, including uridine, UMP, UDP and UTP. Each reaction mixture (50 μl) contained 5 μM purified TruD_Sm, and 1 mM substrate, in 50 mM Tris-HCl buffer pH 7.7. Incubations were carried out at 30°C for 0.5 hr, 1 hr, 2 hr, 6 hr and overnight. After the incubation, chloroform was added to the reaction mixtures to precipitate the protein. 10 ul of the

resultant samples were analysed by HPLC-UV-MS (Agilent HP1200 coupled to Thermo fisher LTQ mass spectrometer) on a Prodigy 5µ C18 column (4.6 x 250 mm, Phenomenex), eluting with MQ containing 0.1% trifluoroacetic acid (A) and acetonitrile containing 0.1% trifluoroacetic acid (B) at a flow rate of 0.7 ml min^{-1}. The linear gradient was 0% B for 5 min, 0% to 5% B over 10 min, 5% to 100% B over 10 min, 100% B for 4 min, 100% to 0% B over 1 min. The mass spectrometer was set up in positive electrospray ionisation (ESI) mode, scanning from *m/z* 200 to 1000 in full scan mode. MS/MS analysis were performed on $[M+H]^+$ ions with a normalized collision energy of 35%. Elution was also monitored by a photodiode array (PDA) detector.

To test the enolpyruvyl transfer activity of NikO and MalO_Sm towards uridine 5'-monophosphate (UMP) as substrate, the enzymatic reactions were set up containing 1 mM UMP, 1 mM phosphoenolpyruvate (PEP) and 5 µM purified NikO or MalO_Sm in 50 mM Tris-HCl buffer pH 7.7. Reactions were incubated at 30°C for 2 hr. After incubations, samples were treated with chloroform and 10 µl of the resultant samples were analysed by HPLC-UV-MS as described above.

To test the enolpyruvyl transfer activity of NikO and MalO_Sm towards pseudouridine-5'-monophosphate (Ψ-MP) as substrate, Ψ-MP was generated *in situ* using *E. coli* pseudouridine-5′-phosphate glycosidase YeiN. To verify the activity of the purified YeiN, reaction was set up containing 0.5 mM $MnCl_2$, 1 mM uracil, 1 mM ribose 5-phosphate (R5P), 5 µM YeiN, in 50 mM Tris-HCl buffer pH 7.7. The reaction mixture was incubated at 30°C for 2 hr. The formation of Ψ-MP was checked by HPLC-UV-MS analysis of 5 µl reaction mixture as described above. A one-pot reaction was set up for the activity assay of NikO and MalO_Sm using Ψ-MP as substrate. The reaction contained 0.5 mM $MnCl_2$, 1 mM uracil, 1 mM R5P, 1 mM PEP, 5 µM YeiN and 5 µM NikO or MalO_Sm, in 50 mM Tris-HCl buffer pH 7.7. The reaction mixture was incubated at 30°C for 2 hr. After treatment with chloroform, 10 µl of the resultant samples were analysed by HPLC-UV-MS as described above.

Large-scale production and purification of 3'-EP-ΨMP and NMR analysis

For the production of 3'-EP-ΨMP, 40 ml of reaction containing 1 mM $MnCl_2$, 2 mM uracil, 2 mM R5P, 2 mM PEP, 5 µM YeiN and 6 µM NikO, in 50 mM Tris-HCl buffer pH 7.7. The reaction mixture was incubated at 30°C for 3 hr. After incubation,

chloroform was added to the reaction mixture to precipitate the proteins. The clear supernatant was then transferred and lyophilized. To purify the reaction product, the lyophilized crude mixture was dissolved in water, and applied to a semi-prep C18 column (Gemini 5µ C18, 10.00 x 250 mm, Phenomenex) on an Agilent 1260 Infinity II LC. MQ containing 0.05% TFA (A) and acetonitrile (B) were used to elute the compound at a flow rate of 3 ml min^{-1}. The linear gradient was 0% B for 5 min, 0% to 5% B over 10 min, 5% to 100% B over 10 min, 100% B for 4 min, 100% to 0% B over 1 min. Elution was monitored at 262 nm. Fractions containing 3'-EP-ΨMP were combined. Acetonitrile was removed under reduced pressure with the rotary evaporator. The remaining aqueous solution was lyophilized. The dried residue was dissolved in [$^2H_6$]DMSO, and analysed on a Brucker-500 NMR spectrometer.

Quantification and Statistical Analysis

See individual sections above for details on the statistics used for analysis.

Data and Software Availability

The software used in this study is listed in the Key Resources Table.

The sequence of the *mal* cluster has been deposited in the NCBI under the accession number GenBank MH537786.

Supplemental item titles and legends

Figure S1. A) In-frame deletion of *truD*, *malD*, *malF*, and *coaE* genes in *S. chromofuscus*, Related to STAR Method and Figure 3. Lane 1: marker; Lane 2 and 3: PCR product from Δ*truD* (93 bp) and WT (1,897 bp), respectively; Lane 4: marker; Lane 5 and 6: PCR product from Δ*malD* (2,437 bp) and WT (3,703 bp), respectively; Lane 7: marker; Lane 8 and 9: PCR product from Δ*malF* (2,631 bp) and WT (3,315 bp), respectively; Lane10: marker; Lane 11 and 12: PCR product from Δ*coaE* (1,108 bp) and WT (1,675 bp), respectively; Lane 13: marker. **B) In-frame deletion of *nikO* gene in *S. tendae*, Related to STAR Method and Figure 3.** Lane 1: marker; Lane 2

and 3: PCR product from Δ*nikO* (520 bp) and WT (1,942 bp), respectively. **C) ESI-MS/MS spectra of malayamycin A and its analogous.** Malayamycin A standard ($[M+H]^+$: 343.2), malayamycin A produced from *S. chromofuscus* WT, desmethyl-malayamycin A from Δ*malF* ($[M+H]^+$: 329.2), descarbamoyl-malayamycin A from Δ*malD* ($[M+H]^+$: 300.2) and descarbamoyl-descarbamoyl-malayamycin A from Δ*malD* ($[M+H]^+$: 286.2).

Figure S2. Sequence alignment of TruD family of tRNA pseudouridine synthases, Related to Figure 5. The sequences of TruDs [TruD_SM: TruD from *Streptomyces malaysiensis*; TruD_SC: TruD from *Streptomyces chromofuscus*; TruD_SS: TruD from *Streptomyces* sp. ID38640 (accession no. AVT42379); TruD_HP : TruD from *Helicobacter pylori* (accession no. NP_207718); TruD_AF: TruD from *Archaeoglobus fulgidus* (accession no. O28596); TruD_EC: TruD from *E. coli* (accession no. AQZ29382)] are aligned using Clustal Omega. The conserved sequence motifs are highlighted in boxes. The catalytic aspartate is denoted with an asterisk.

Figure S3. Phylogenetic analysis of YeiN-related glycosidase and putative glycosynthase enzymes, Related to Figure 4. Sc: *Streptomyces chromofuscus*; St: *Streptomyces tendae*; Sm: *Streptomyces malaysiensis*; ShYeiN: YeiN from *Streptomyces himastatinicus* (accession no. WP_009720488); EcYeiN: YeiN from *Escherichia coli* (accession no. AAA60517); AlnA: alnumycin biosynthesis from *Streptomyces* sp. CM020 (accession no. ACI88875); SaInda: indigoidine biosynthesis from *Streptomyces albus* (accession no. AMM12432); Sdma: showdomycin biosynthesis from *Streptomyces showdoensis* ATCC 15227.

Figure S4. Comparison of ESI-MS/MS spectra of 3'-EPUMP ($[M+H]^+$: 395.2) and 3'-EP-Ψ-MP ($[M+H]^+$: 395.2), Related to Figure 4. Loss of C-N linkage in 3'-EP-Ψ-MP abolished the fragment at *m/z* 283.2, which was a key fragment in the MS/MS of 3'-EPUMP.

Figure S5. Analysis of *N*-malayamycin A by MS and ^{1}H NMR, Related to Figure 5A. A) High-resolution UPLC-MS analysis of *N*-malayamycin A and malayamycin A

from fermentation extract of *S. chromofuscus* Δ*truD::nikO* mutant. **B)** Comparison of ^{1}H spectra of *N*-malayamycin A purified from fermentation of *S. chromofuscus* Δ*truD::nikO* mutant and from chemical synthesis (Hanessian et al., 2005).

Figure S6. MS/MS and MS^{3} analysis of nikkomycin Z and Ψ-nikkomycin Z, Related to Figure 5B. A) MS/MS and MS^{3} spectra of nikkomycin Z. **B)** MS/MS and MS^{3} spectra of Ψ-nikkomycin Z. **C)** Fragmentation pathways for the generation of *m/z* 176 fragment in nikkomycin Z and of *m/z* 208 fragment in Ψ-nikkomycin Z.

Figure S7. Comparison of Ψ–nikkomycin Z production with and without uracil feeding to the Δ*nikO*::*malO_Sm* mutant, Related to Figure 5B. When uracil (0.4%) was fed to the Δ*nikO*::*malO_Sm* mutant, the production of Ψ-nikkomycin Z was increased 5-fold**.**

Data S1. NMR Spectra of 3'-EP-Ψ-MP. Related to Figure 4. ^{1}H- and ^{13}C- 1D NMR spectra, and ^{1}H-^{1}H COSY, HSQC and HMBC 2D NMR spectra, of the product of the reaction catalyzed by MalO.

peter leadlay 4/12/18 12:01
Formatted: Font:Not Bold, Superscript

peter leadlay 4/12/18 12:01
Formatted: Font:Not Bold

peter leadlay 4/12/18 12:01
Formatted: Font:Not Bold, Superscript

peter leadlay 4/12/18 12:01
Formatted: Font:Not Bold

peter leadlay 4/12/18 12:01
Formatted: Font:Not Bold

peter leadlay 4/12/18 12:01
Formatted: Font:Not Bold

peter leadlay 4/12/18 12:01
Formatted: Font:Not Bold

peter leadlay 4/12/18 12:01
Formatted: Font:Not Bold

peter leadlay 4/12/18 11:55
Deleted: