Supplemental Data

Dissecting Tunicamycin Biosynthesis by Genome Mining: Cloning and Heterologous Expression of a Minimal Gene Cluster

Filip J. Wyszynski, Andrew R. Hesketh, Mervyn J. Bibb and Benjamin G. Davis

S1 Overview of prior studies into tunicamycin biosynthesis, related to Introduction

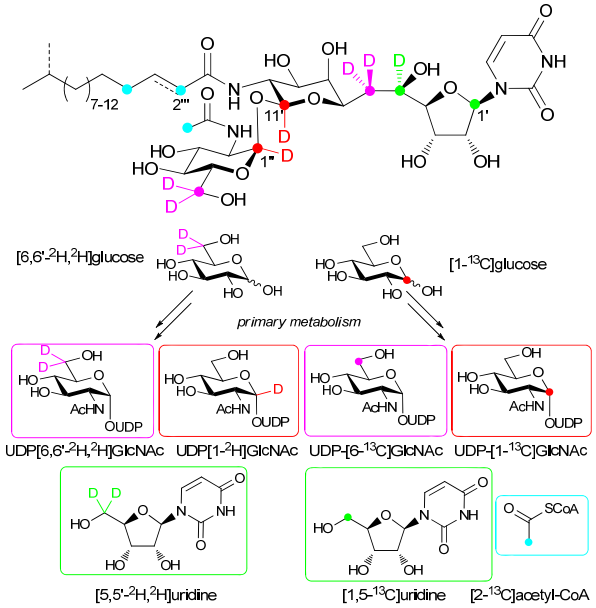

Fig. S1 Prior studies using isotope enrichment in tunicamycin through feeding of labelled precursors

In 2002, Tsvetanova *et al*. performed isotopic enrichment of tunicamycins through metabolic incorporation experiments with radiolabelled precursors (Fig. S1).[1,2] The loci of isotopic enrichment were determined by TLC, MS and NMR characterisation.[3]

^{14}C labelling: Feeding tunicamycin-producer *S. chartreusis* with [2-^{14}C]uridine led to successful incorporation into tunicamycin, suggesting uridine was a biosynthetic precursor. Feeding *S. chartreusis* with D-[1-^{14}C]glucosamine, followed by acid hydrolysis of isolated tunicamycins, indicated this compound had been incorporated into both the 11-carbon tunicaminyl moiety and the terminal GlcNAc residue, suggesting the involvement of a 6-carbon hexose precursor in both tunicamine biosynthesis and terminal glycosylation.

^{13}C labelling: *S. chartreusis* was fed with D-[1-^{13}C]glucose (Fig. S1). The largest enrichments in tunicamycin were detected at both C-11′ and C-1′′, signifying direct incorporation of hexose units in each case from the same metabolic pool of GlcNAc or UDP-GlcNAc. Moderate enrichment was observed at C-6′′, C-6′, C-5′ and C-1′, derived from primary metabolism of D-[1-^{13}C]glucose to D-[6-^{13}C]glucose then UDP-[6-^{13}C]GlcNAc via glycolysis and gluconeogenesis and to [1′,5′-^{13}C]uridine via the pentose phosphate pathway. Enrichment was also observed in the acyl lipid chain at C-2′′ and C-4′′, together with the *N*-acetyl methyl carbon. These incorporations were again derived from primary metabolism of [1-^{13}C]glucose to [2-^{13}C]acetyl-CoA. These results supported the hypothesis that tunicamycin is derived from uridine and GlcNAc building blocks and that its lipid is formed by a fatty acid synthase utilising acetyl-CoA.

^{2}H labelling: *S. chartreusis* was fed with D-[6,6′-^{2}H,^{2}H]glucose and deuterium incorporation monitored by analysing ESI-CDI-MS fragments. A total of seven ^{2}H atoms were incorporated into tunicamycin (Fig. S1), attached to C-5′, C-6′, C-11′, C-1′′ and C-6′′ and showing a similar pattern of labelling to that with D-[1-^{13}C]glucose (Fig. S1). Assignments were supported by competitive metabolic experiments, where *S. chartreusis* cells were co-incubated with D-[6,6′-^{2}H, ^{2}H]glucose and one of a selection of unlabelled metabolites: GlcNAc, uridine, ribose, glycerol or succinate. Competition of these supplements with D-[6,6′-^{2}H, ^{2}H]glucose led to depletion of isotopic enrichment at certain positions in tunicamycin, which allowed the metabolic origin of enriched loci to be ascribed to the previously proposed shunt pathways from primary metabolism.

Together, these experiments allowed a biosynthetic pathway for the tunicamycins to be rationalised.[1,2] Uridine and UDP-GlcNAc are first converted to proposed intermediates uridine-5′-aldehyde and UDP-6-deoxy-4-keto-GlcNAc-5,6-ene. Following the coupling of these intermediates to form the tunicamine core, hydrolysis of UDP was proposed to yield a free anomeric hydroxyl to which GlcNAc is stereospecifically transferred from UDP-GlcNAc. This formed the requisite α,β-1′′,11′ glycosidic linkage. The completed tunicamycin skeleton was then selectively *N*-deacetylated and *N*-acylated with a range of unsaturated fatty acids. An alternative ordering existed whereby the α,β-trehalose linkage was formed first, followed by activation and coupling to uridine-5′-aldehyde. However, this required two different sources of GlcNAc: UDP-GlcNAc and free *N*-acetylglucosamine. The equivalent enrichment of both C-1′′ and C-11′ strongly

suggested the incorporated GlcNAc molecules come from the same metabolic pool, supporting the former pathway, which utilised only UDP-GlcNAc.

S2 Description of *tun* Genes, related to Figure 4

Despite being tightly clustered and likely to utilise a single polycistronic transcript during their expression, the fourteen genes in the *tun* cluster have diverse functions and each control very different aspects of tunicamycin biosynthesis. The genes do not appear to be clustered by function, although this is how they will be treated in the following sections. It should be noted that when referring the closest homologue of a particular *tun* gene in the present discussion, hits from *S. clavuligerus* ATCC27064 and *A. mirum* DSM43827 have been omitted.

Carbohydrate Building Block Biosynthesis

The two suggested coupling partners involved in the formation of the first eleven-carbon dialdose intermediate on the way to tunicamine are uridine 5´-aldehyde and UDP-4-keto-5,6-ene-*N*-acetylglucosamine. The closest homologues of the *tunA* gene product are NAD-dependent epimerase/dehydratases from various bacteria (42% identity and 58% similarity with one from *Streptomyces* sp. Mg1). The protein contains a TGxxGxxG signature at the N-terminus denoting the presence of a nucleotide-binding Rossmann fold,[4] as well as a $Sx_{24}Yx_3K$ catalytic triad.[5] These features are characteristic of short-chain dehydrogenase/reductase superfamily members, which includes the epimerase/dehydratase family of enzymes.[6] Examples sharing these motifs include FlaA1, a UDP-GlcNAc C_6-dehydratase/C_4-oxidoreductase from *Helicobacter pylori* producing UDP-6-deoxy-GlcNAc,[7] and dTDP-glucose 4,6-dehydratase from *Streptomyces venezuelae* yielding dTDP-6-deoxy-4-keto-glucose.[8] We propose that TunA converts UDP-GlcNAc – which has been shown to be a precursor in tunicamycin biosynthesis – to the UDP-4-keto-5,6-ene-GlcNAc intermediate involved in undecose construction, although reduction to the allylic alcohol is possible, resulting in UDP-5,6-ene-GlcNAc.

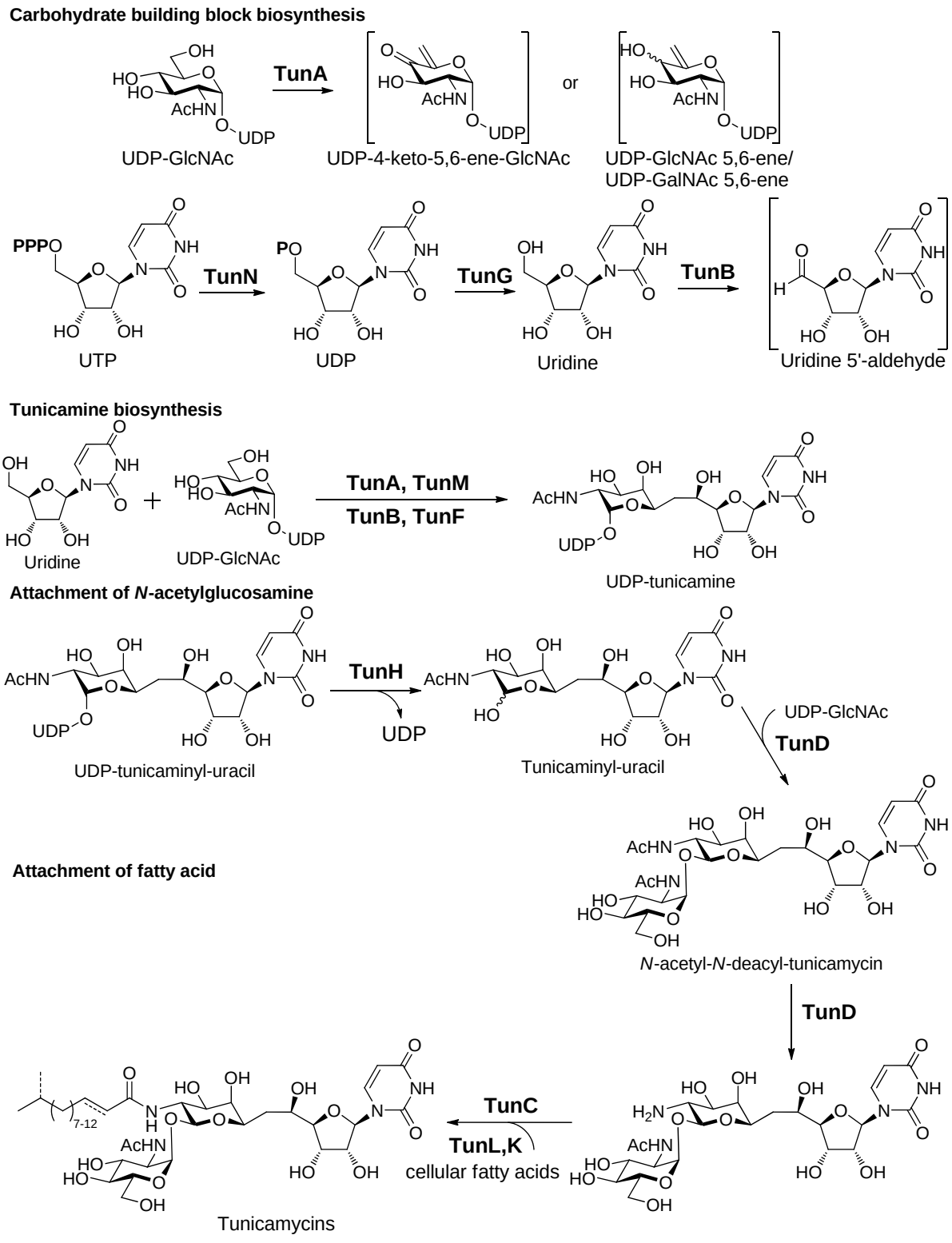

Fig S2. The role of Tun proteins in tunicamycin biosynthesis, related to Figure 4

Formation of the suggested uridine 5′-aldehyde coupling partner begins with the action of TunN. The closest homologues of the *tunN* gene are NUDIX hydrolases from various bacteria (36% identity and 55% similarity with one from *Nakamurella multipartita* DSM 44233). It contains a 23-residue $Gx_5Ex_7RExxEExGV$ motif highly conserved amongst members of the NUDIX-hydrolase superfamily, forming a structural motif involved in Mg^{2+} or Mn^{2+} binding and subsequent catalysis.[9] Substrates of this family of enzymes have the general structure of a nucleotide diphosphate linked to another moiety X (NUDIX), and include nucleoside triphosphates, dinucleotide polyphosphates and nucleotide sugars. Enzymes acting on uridine 5′-triphosphate (UTP) have been identified.[10] In our case, TunN is likely to act as a pyrophosphatase on UTP to give uridine 5′-monophosphate (UMP), which acts as a substrate for further phosphate hydrolysis by TunG. The sequence of *tunG* shows homology to phosphoglycerate mutases (29% identity and 47% similarity with one from *Frankia* sp. CcI3) and α-ribazole phosphatases from a number of sources (27% identity and 42% similarity with one from *Clostridium botulinum* Bf). α-Ribazole 5′-phosphate is an intermediate in cobalamin (vitamin B_{12}) biosynthesis, and its structure is closely related to that of UMP – our proposed substrate for TunG.[11,12] Analysis using the Conserved Domains Database (CDD)[13] showed TunG to belong to the phosphoglycerate mutase-like subgroup of the histidine phosphatase superfamily, to which both of the aforementioned enzyme classes belong.[14] The action of TunN and TunG on UTP is suggested to be similar to the actions of Pur7 and Pur3 in the biosynthesis of nucleoside antibiotic puromycin, where adenosine 5′-triphosphate is the initial substrate.[15,16] The uridine which results from these transformations we suggest is next oxidized to uridine 5′-aldehyde by TunB. All of the closest homologues of *tunB* are radical SAM domain proteins (32% identity and 48% similarity with one from *Haloterrigena turkmenica* DSM 5511) with examples of homology to coenzyme pyrroloquinoline quinone (PQQ) synthesis protein E (*pqqE*) being the most common (33% identity and 51% similarity with one from *Mycobacterium abscessus* ATCC 19977). Members of the radical SAM superfamily all bind a 4Fe4S cluster and S-adenosyl-L-methionine (SAM); indeed TunB contains a characteristic CxxxCxxC motif for binding the iron-sulfur cluster.[17] Enzymes of this class generate an enzyme bound 5′-deoxyadenosyl radical from SAM, a powerful oxidant which catalyses a range of radical-mediated reactions at specific positions of bound substrates.[18] In our case, we propose a hydrogen is

abstracted from C5′ of uridine to generate a substrate radical, which could undergo further oxidation to uridine 5′-aldehyde or feature directly in the subsequent coupling with UDP-4-keto-5,6-ene-GlcNAc.

Tunicamine Biosynthesis

The undecose skeleton of tunicaminyl uracil is proposed to be formed by the tail-to-tail coupling of uridine 5′-aldehyde and UDP-4-keto-5,6-ene-GlcNAc to install the necessary C–C bond. We suggest that these intermediates are the products of UTP and UDP-GlcNAc through the actions of TunN, TunG, TunB and TunA. The gene product TunM shows greatest homology to a methyltransferase family protein from *Saccharomonospora viridis* DSM 43017 (48% identity and 63% similarity). The vast majority of homologues are from this family, an assignment supported by CDD analysis which shows TunM to belong to the *S*-adenosylmethionine dependant methyltransferase superfamily. Since methyl transfer is not envisaged during tunicamycin biosynthesis, the precise role of TunM remains uncertain. Although it contains a SAM-binding motif of ExGxGxG, it lacks any CxxxCxxC motif for the binding of 4Fe4S, thus explaining its classification as a methyltransferase family protein rather than a radical SAM superfamily protein.[19] However, in tunicamycin biosynthetic pathway it likely acts to mediate a radical reaction between the aforementioned coupling partners, via a 5′-deoxyadenosyl radical. One possibility is that TunB and TunM are closely associated and act together to bring both coupling partners in close proximity of an enzyme-bound reactive radical species, which is first created from the SAM cofactor of TunB by the action of its 4Fe4S cluster and then transferred to the SAM cofactor of its mediating partner TunM where the C–C bond forming event takes place. Clearly further work is necessary to obtain direct experimental evidence in support of this hypothesis, but there remains the exciting prospect of deciphering a unique and highly unusual biosynthetic mechanism, possibly uncovering novel chemistry unprecedented for an enzymatic reaction.

The product of gene *tunF* shows homology to a large number of UDP-glucose-4-epimerase proteins (46% identity and 63% similarity with one from *Paenibacillus* sp. oral taxon 786 str. D14). It contains an N-terminal Rossmann fold as denoted by CDD analysis and the presence of a characteristic TGxxGxxG nucleotide-binding motif. It also contains the $Sx_{24}Yx_3K$ catalytic triad.[5] Since TunF is not predicted to have 4,6-dehydratase activity, it may serve to function purely as a C-4 epimerase. It is interesting to note

that the dehydration of UDP-GlcNAc and epimerization in tunicamine occurs by a similar mechanism and the overall could in principle by catalysed by one (TunA) or two (TunA and TunF) enzymes. The requirement for an epimerization step may derive from the stereoelectronic preferences of hydride abstraction and return by TunA on its substrate UDP-GlcNAc. After oxidizing C-4 to a ketone and catalyzing dehydration to a 5,6-ene, the abstracted hydride is returned to the substrate from NADH to complete the catalytic cycle; in this case reduction of the 4-keto moiety is likely over that of the 5,6-ene – as observed in 6-deoxysugar biosynthesis[20] – since this alkene moiety is likely to play an important role in the subsequent undecose-forming event. Since the hydride at C-4 was initially abstracted from the axial position it may return from the same face, reducing the 4-keto functionality to an equatorial hydroxyl and yielding a *gluco*- configured product, perhaps requiring epimerization by TunF either before or after undecose formation.

Attachment of *N*-acetylglucosamine

The gene product of *tunH* shows homology to type I phosphodiesterases/nucleotide pyrophosphatases (35% identity and 50% similarity with one from *Burkholderia* sp. 383). CDD analysis reveals N- and C-terminal domains characteristic of the sulfatase superfamily, with particular homology to pfam01663 family members, namely Type I phosphodiesterases/nucleotide pyrophosphatases. The likely function of TunH is in the hydrolysis of UDP-tunicaminyl-uracil, releasing the anomeric 11´-OH for subsequent glycosylation with GlcNAc. Formation of such an unusual α,β-1,1-trehalose linkage is extremely rare and is most probably catalysed by the only putative glycosyltransferase present in the *tun* gene cluster TunD, using UDP-GlcNAc as the glycosyl donor. TunD is homologous to group 1 glycosyl transferases from various bacteria (26% identity and 42% similarity with one from *Magnetococcus* sp. MC-1). CDD analysis predicts TunD is a member of the GT-1 Yqgm-like family, named after glycosyltransferase Yqgm from *Bacillus lichenformis* about which little is known. It contains a motif present in GT-1 family glycosyltransferases (pfam00534) and is part of the larger glycosyltransferase GT-B superfamily possessing the GT-B fold. Through analysis of substrates and products it can be assigned as a retaining glycosyltransferase.[21]

Attachment of Fatty Acid

TunE shows homology to LmbE family proteins from a variety of organisms (43% identity and 57% similarity with one from *Syntrophomonas wolfei* subsp. *wolfei* str. Goettingen). This family is part of the larger PIG-L superfamily and also contains *N*-acetyl-glucosaminyl-phosphatidylinositol de-*N*-acetylases from glycosyl-phosphatidyl-inositol (GPI) anchor biosynthesis[22] and 1-D-*myo*-inosityl-2-acetamido-2-deoxy-α-D-glucopyranoside deacetylases from mycothiol biosynthesis.[23] TunE contains an unusual HPDD zinc-binding motif conserved across both of the latter two enzyme families[24] and CDD analysis also places it in the PIG-L superfamily (COG2120). Based on the similarities between *N*-acetyl-*N*-deacyl-tunicamycin and the substrates of mycothiol and GPI-anchor biosynthesis, we propose that TunE catalyses Zn^{2+}-mediated cleavage of the 10′-acetamido group of the pseudotrisaccharide precursor to reveal the free amine, in preparation for subsequent fatty acid attachment.

Since heterologously produced tunicamycins in *S. coelicolor* were fully acylated although there are no fatty acid synthase genes in the *tun* cluster, the range of fatty acids observed amongst tunicamycin homologues are likely to be derived from the cellular pool of fatty acids; this has previously been observed in teicoplanin biosynthesis.[25] The gene product of *tunL* shows homology to type 2 phosphatidic acid phosphatases (PAP2) from various bacteria (33% identity and 42% similarity with one from *Micromonospora aurantiaca* ATCC 27029). CDD analysis identifies it as a member of the PAP2-like protein subfamily specific to bacteria (cd03392), and its sequence contains a three-part catalytic motif characteristic of this Mg^{2+}-independent phosphatase family – KxxxxxxRP (domain 1), PSGH (domain 2) and SRxxxxxHxxxD (domain 3).[26] The most likely function of TunL is in the regulation of lipid synthesis in the producing bacterium. By down-regulating the levels of cellular phosphatidic acid and up-regulating levels of its cleavage product diacylglycerol, phospholipid biosynthesis is repressed and cellular pools of fatty acids can be diverted for use in tunicamycin biosynthesis via β-oxidative degradation pathways.[27] It is fascinating that tunicamycin-producing organisms appear to have evolved an efficient way of perturbing the complex and subtle signaling pathways of lipid metabolism regulation, allowing increased tunicamycin biosynthesis without negatively affecting vital cellular processes. Intriguingly, the *tun* cluster contains an acyl carrier protein (ACP) in TunK – which exhibits homology to ACPs from numerous bacterial sources (32% identity and 61% similarity with one from *Catenulispora acidiphila* DSM 44928) – but does not

encode a fatty acyl ACP ligase. Presumably, fatty acids are activated as ACP-thioesters for use as substrates in the subsequent *N*-acylation of the tunicamycin skeleton through the action of existing fatty acyl ACP ligases from primary metabolism. The gene product of *tunC* shows homology to GCN5-related *N*-acetyltransferases (GNAT) from a variety of bacteria (33% identity and 50% similarity with one from *Fervidobacterium nodosum* Rt17-B1). Analysis by HHPred shows homology to mycothiol synthase MshD across the full length of TunC (PDB 1P0H), but suggests the presence of two separate GNAT domains making up the N-terminal and C-terminal halves of the protein respectively.[28] Indeed, MshB has been shown to similarly contain two GNAT domains, one of which appears to be catalytically non-functional.[29] We propose that TunC catalyses the acylation of the tunicamycin core skeleton 10′-NH_2, using fatty acids sequestered from primary metabolism and activated as ACP-thioesters. There is, however, a possibility that TunC uses fatty acyl-coenzyme A (acyl-CoA) substrates directly from the β-oxidative degradation pathway as substrates and first catalyses the *S*-acyl transfer from CoA to an associated ACP (in this case TunK), before also catalyzing subsequent *N*-acyl transfer to tunicamycin. This activity has been observed in one of the polyketide synthase modules of curacin A,[30] although further experimental evidence is clearly required to distinguish between these proposals.

Transporter Genes

The gene products of *tunI* and *tunJ* comprise components of the ATP-binding cassette (ABC) transporter systems from a variety of bacteria. TunI is homologous to ABC transporter ATP-binding subunits (41% identity and 61% similarity with one from *Streptomyces scabiei* 87.22) and TunJ is homologous to ABC-2 transporter permeases (32% identity and 51% similarity with one from *Thermobaculum terrenum* ATCC BAA-798). Such transporter systems are well studied and have been observed in many bacteria producing antibiotics, functioning as a self-resistance mechanism by coupling ATP hydrolysis to metabolite efflux through the cell membrane.[31,32] As such, TunI and TunJ together probably provide the important self-resistance mechanism for tunicamycin-producing organisms.

S3 LC/MS Confirmation of Heterologous Tunicamycin Production, related to Figure 3

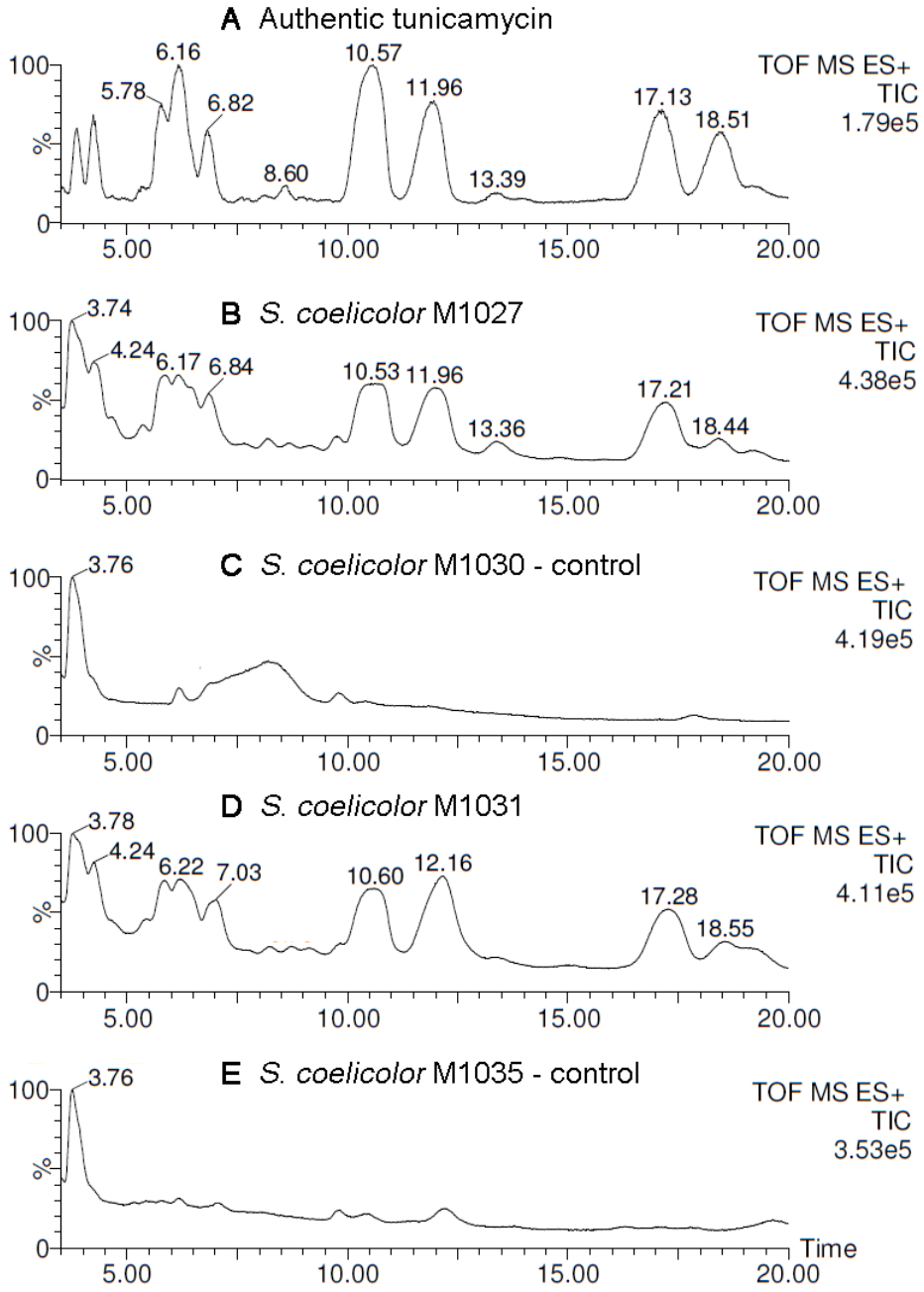

Fig. S3 Total ion chromatograms of extracts from recombinant S. coelicolor strains (for strain descriptions see Tables S1 and S2)

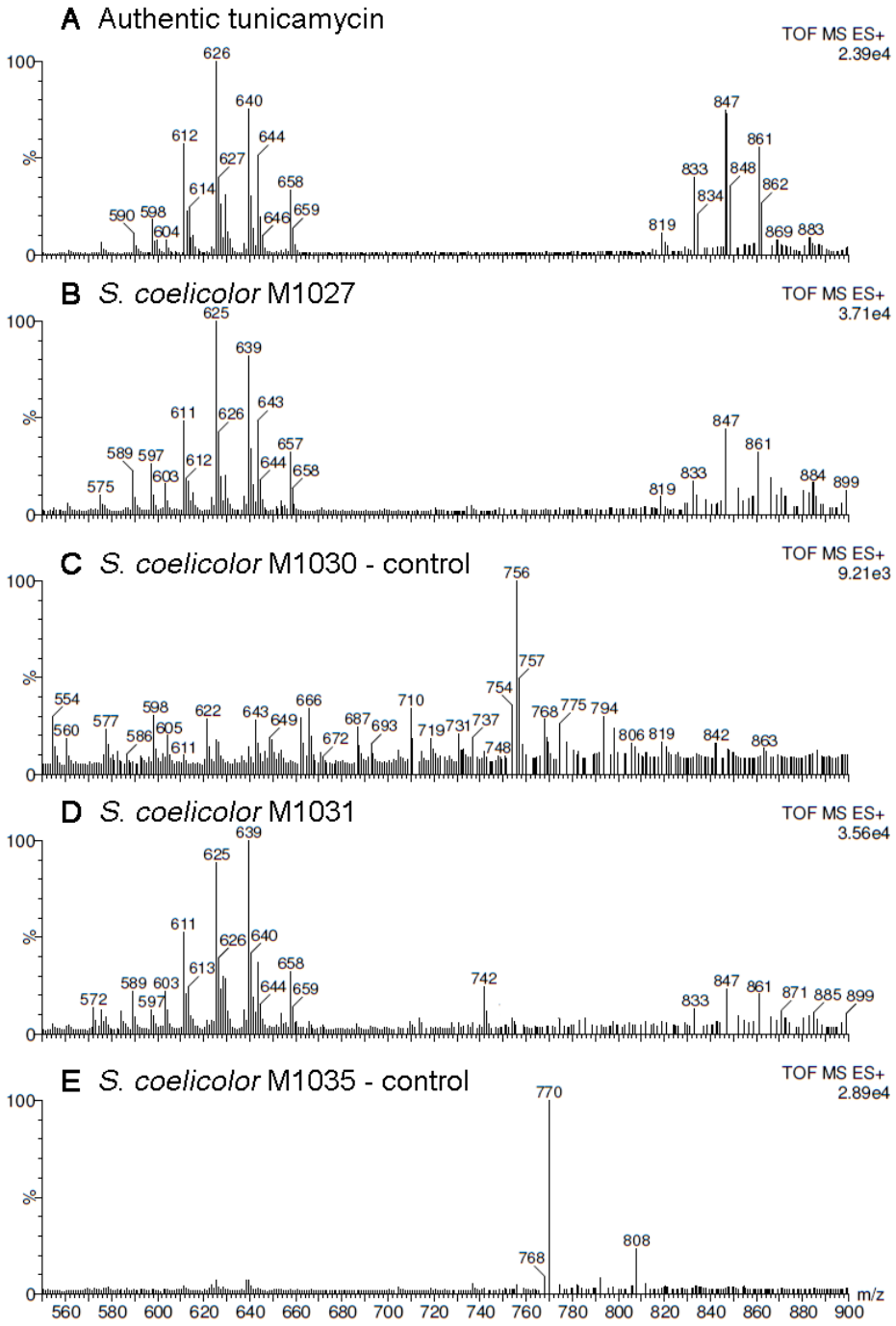

Fig. S4 Mass spectra from LC/MS analysis of extracts from recombinant *S. coelicolor* strains (for strain descriptions see Tables S1 and S2)

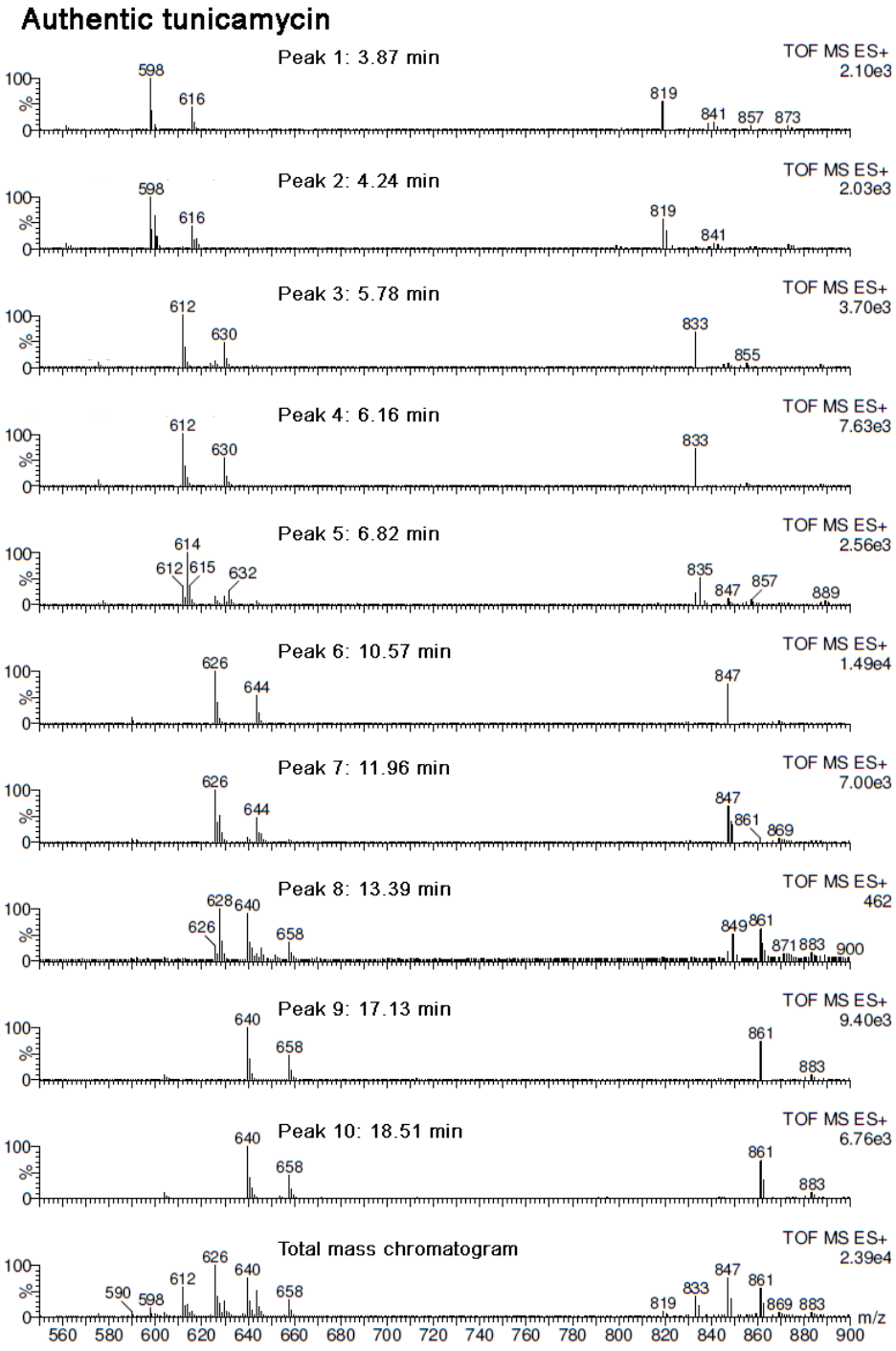

Fig. S5 Extracted mass spectra of individual peaks from LC/MS analysis of authentic tunicamycin

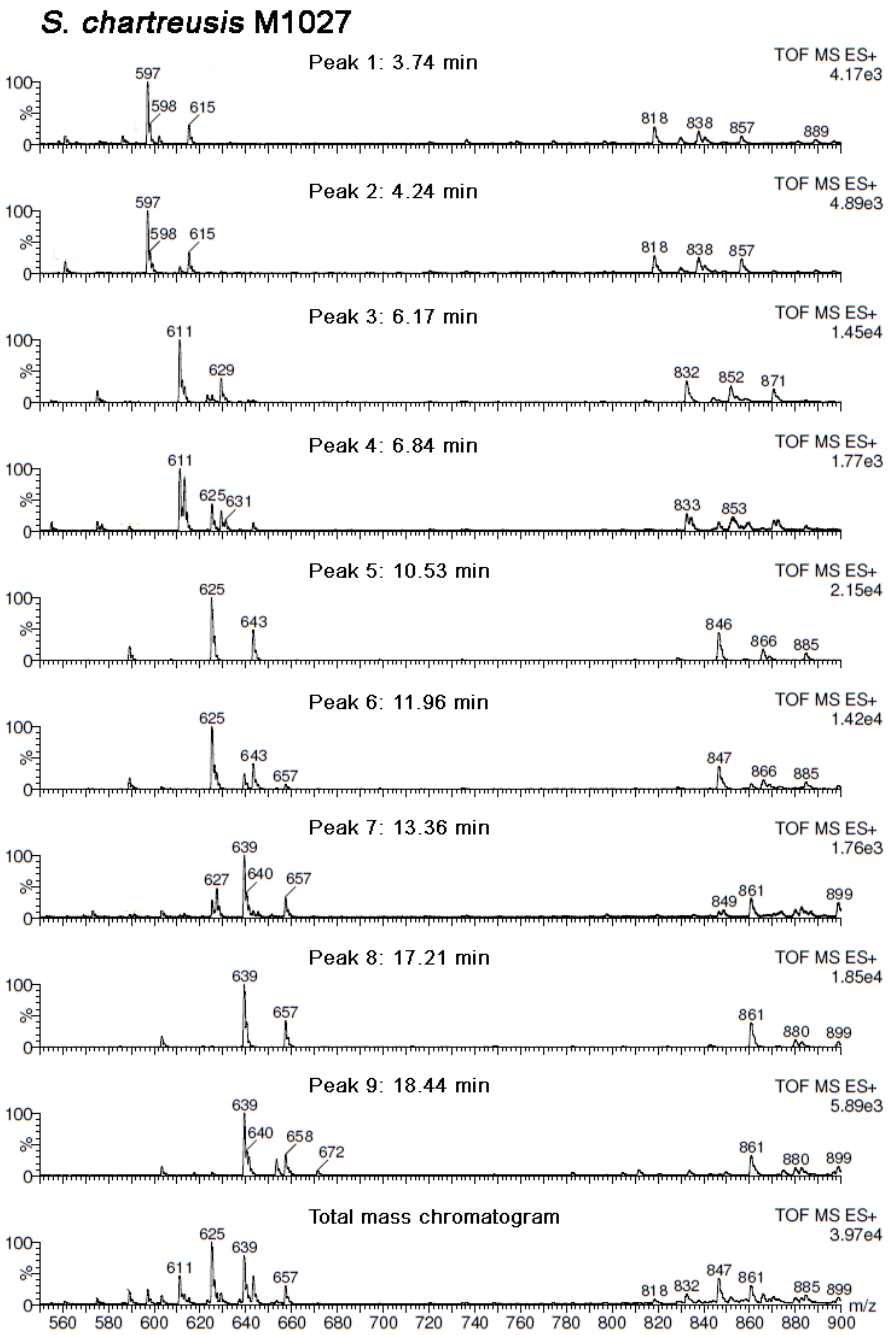

Fig. S6 Extracted mass spectra of individual peaks from LC/MS analysis of extracts from *S. coelicolor* M1027 (for strain descriptions see Tables S1 and S2)

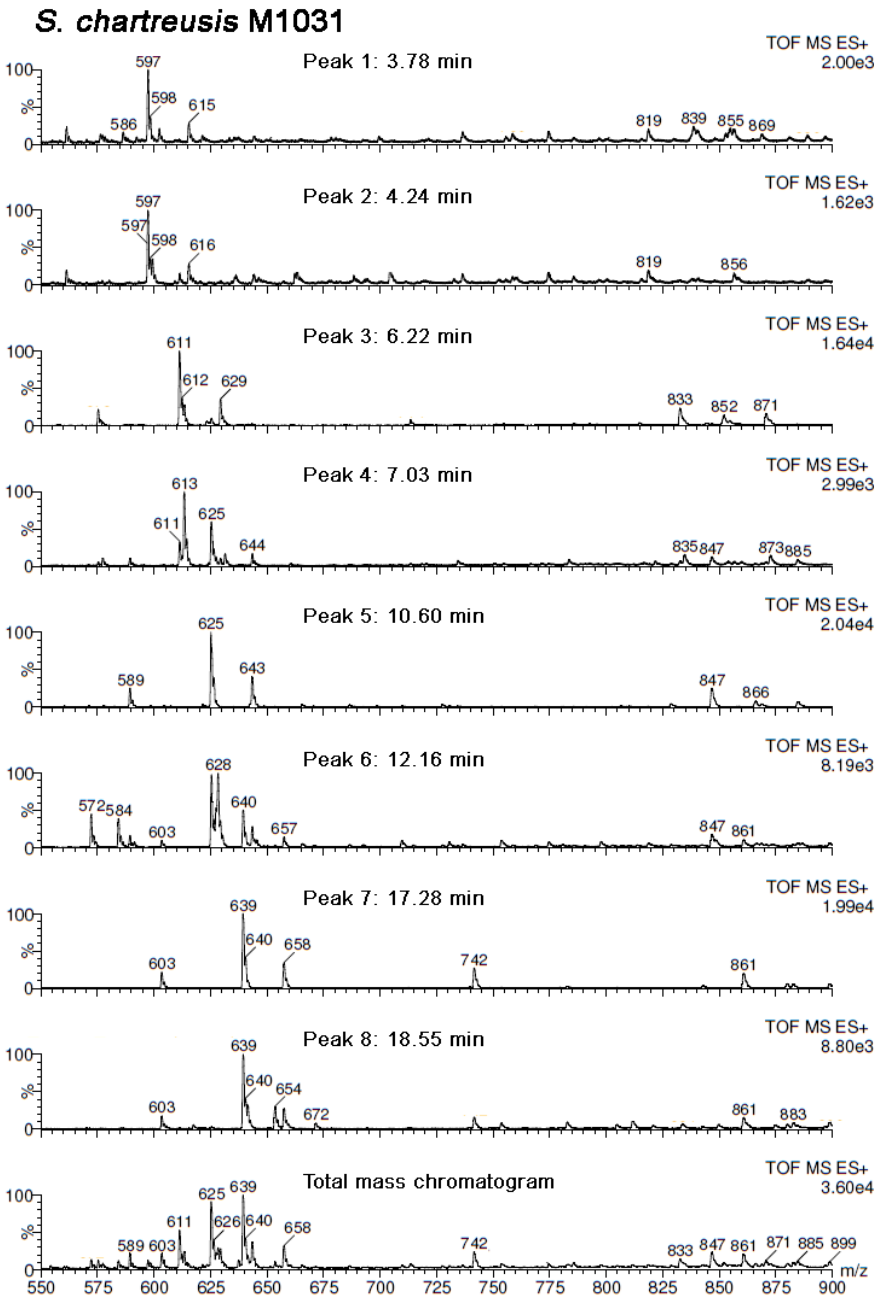

Fig. S7 Extracted mass spectra of individual peaks from LC/MS analysis of extracts from *S. coelicolor* M1031 (for strain descriptions see Tables S1 and S2)

For both authentic and heterologously produced tunicamycin samples, lipid length increases with retention time. The lipid distribution of tunicamycins in the authentic sample is very similar to that of both heterologously produced samples. The mass spectra of individual LC/MS peaks – corresponding to separated tunicamycin homologues – are consistent with those obtained by Tsvetanova *et al*.[3] Based on the masses observed in our mass spectra and on comparison with published analyses, the tunicamycin homologues present in each LC/MS peak can be estimated as follows: Peaks 1 and 2 (Tun 14:1$_{A/B}$), Peak 3

(Tun 15:$1_{A/B}$), Peak 4 (Tun 15:$1_{A/B}$, Tun15:0), Peaks 5 and 6 (Tun 16:$1_{A/B}$), Peak 7 (Tun 16:0, Tun17:$1_{A/B}$/Tun17:$0_{A/B/C}$), Peaks 8 and 9 (Tun 17:$1_{A/B}$/Tun17:$0_{A/B/C}$). Tunicamycin homologues have been named according to the nomenclature proposed by Tsvetanova *et al*.; for a particular homologue Tunx:y_N, x refers to the number of carbons in the acyl chain (including C=O), y refers to the degree of unsaturation in the chain and N refers to the order of elution of regioisomers.[3]

S4 ^{1}H NMR Confirmation of Heterologous Tunicamycin Production, related to Figure 3

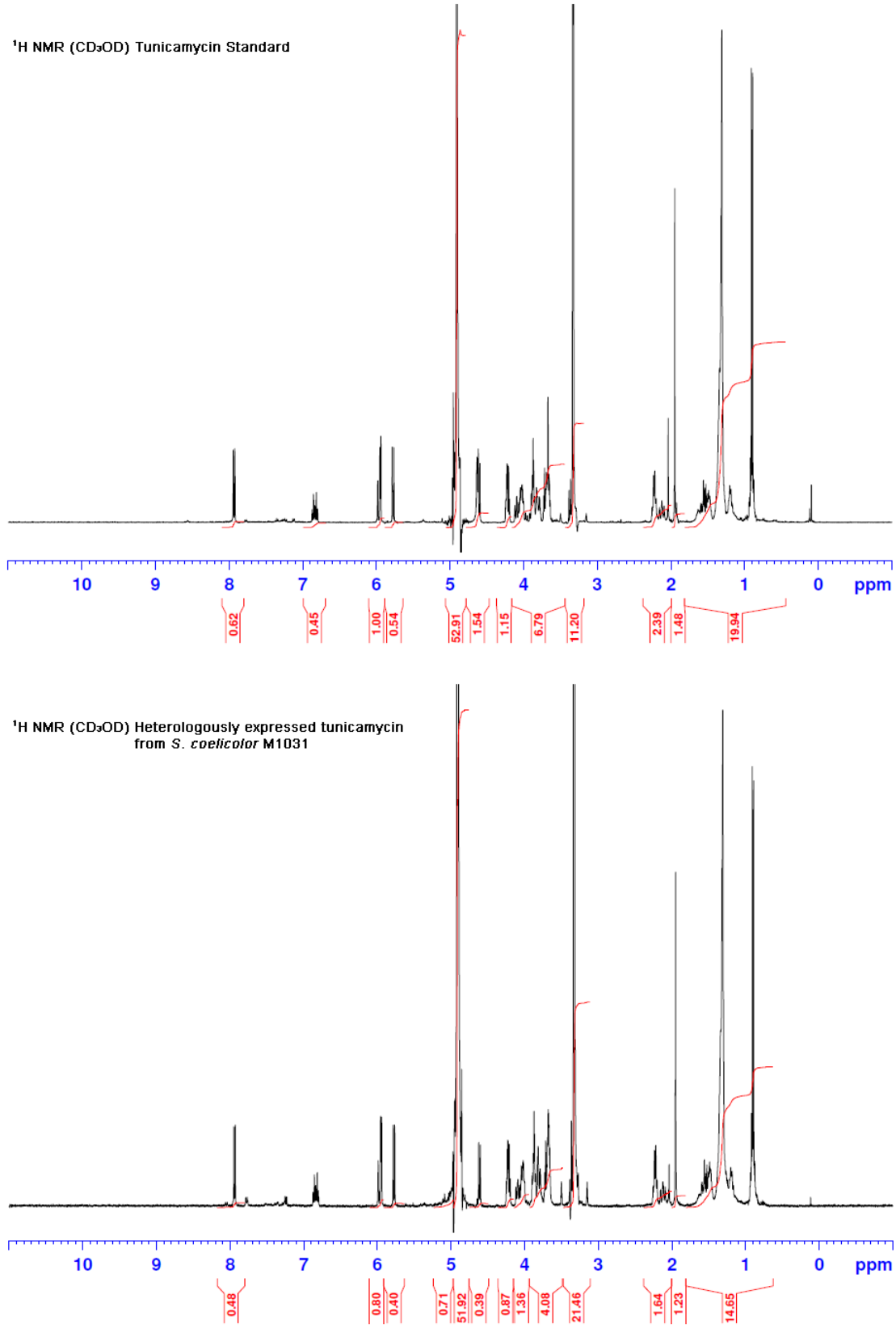

Fig. S8 ^{1}H NMR Confirmation of Heterologous Tunicamycin Production, related to Figure 3

S5 Bacterial Strains, Plasmids and Primers Used, related to Experimental Procedures

Table S1. Bacterial Strains Used in This Study, related to Experimental Procedures

Strain	Relevant genotype	Source or reference
Streptomyces		
S. chartreusis NRRL3882	Tunicamycin producing strain	DSMZ culture collection
S. lysosuperificus ATCC31396	Tunicamycin producing strain	ATCC culture collection
S. coelicolor M1146	*S. coelicolor* M145[33] with deleted *act*, *red*, *cda* and *cpk* gene clusters	Gomez-Escribano J.P. personal communication
S. coelicolor M1152	*S. coelicolor* M1146 carrying a point mutation in *rpsL* for enhanced expression of antibiotic gene clusters	Gomez-Escribano J.P. personal communication
S. coelicolor M1025	*S. coelicolor* M1152 carrying pIJ12315	This study
S. coelicolor M1026	*S. coelicolor* M1152 carrying pIJ12316	This study
S. coelicolor M1027	*S. coelicolor* M1152 carrying pIJ12317	This study
S. coelicolor M1028	*S. coelicolor* M1152 carrying pIJ12318	This study
S. coelicolor M1030	*S. coelicolor* M1152 carrying pMJCOS1 (vector-only control)	This study
S. coelicolor M1031	*S. coelicolor* M1146 carrying pIJ12003a	This study
S. coelicolor M1035	*S. coelicolor* M1146 carrying pRT802 (vector-only control)	This study
Escherichia coli		
E. coli XL1 Blue MR	Restriction-deficient cloning strain for cosmid-based cloning, eg. construction of genomic library	Stratagene
E. coli DH5α	Cloning strain for routine subcloning and high-quality plasmid preparations	Invitrogen
E. coli BW25113	Host strain for λ Red recombination	[34]
E. coli ET12567	Methylation-deficient strain used for conjugation with *Streptomyces*	[35]
E. coli S17-1	Donor strain for conjugative transfer of mobilisable plasmids to *Streptomyces*	[36]
Other bacteria		
Bacillus subtilis EC1524	Tunicamycin-sensitive reporter strain	John Innes Centre, Norwich

Table S2. Plasmids Used in This Study, related to Experimental Procedures

Plasmid	Details	Source or reference
SuperCos1	Cosmid vector used in constructing *S.chartreusis* genomic library	Stratagene
pMJCOS1	Conjugative derivative of Supercos1 that integrates in a single copy at the ϕC31 attachment site in the *Streptomyces* host	[37]
pRT802	Conjugative vector that integrates in a single copy at the ϕBT1 attachment site in the *Streptomyces* host	[38]
pIJ12315	Cosmid clone 4H8 made integrative and conjugative by targeting with the SspI fragment from pMJCOS1	This study
pIJ12316	Cosmid clone 5K7 made integrative and conjugative by targeting with the SspI fragment from pMJCOS1	This study
pIJ12317	Cosmid clone 6N9 made integrative and conjugative by targeting with the SspI fragment from pMJCOS1	This study
pIJ12318	Cosmid clone 7C3 made integrative and conjugative by targeting with the SspI fragment from pMJCOS1	This study
pIJ12003a	A 12.9 Kbp SacI fragment from genomic library cosmid 4H8 containing the minimal *tun* cluster cloned into the SacI site of pRT802	This study

Table S3. Primers Used in This Study, related to Experimental Procedures

Name[a]	Sequence	Use
tunA for	5´-AACACCGGATTCTCATCACC-3´	To generate a probe for *tunA*
tunA rev	5´-TTGAAGGAGCGAGGAACACT-3´	
tunN for	5´-GGTGGGTGTCTACCAGCATC-3´	To generate a probe for *tunN*
tunN rev	5´-GTGGCCTCGAAGAGTGAGAG-3´	
tunD for	5´-ACCGCTACATGCACGACAT-3´	To span the gap between two contigs containing putative *tun* cluster
tunD rev	5´-CTCGTCCTCGACGAACTCC-3´	

[a] All primers were purchased from Sigma-Aldrich

S6 Media Recipes, related to Experimental Procedures

Unless otherwise stated, all media, buffers and solutions used with *Streptomyces* are listed in [39].

Table S4. Recipes for Media Used in This Study, related to Experimental Procedures

MYM agar	OB Agar
4 g Maltose	40 g Porridge Oats (Tesco, UK)
4 g Yeast Extract	20 g Agar
10 g Malt Extract	1 L Tap water
20 g Agar	
1 L Deionised water/tap water (1:1)	Autoclave twice
After autoclaving add:	
2 mL R2 Trace Elements solution	

S7 References

1. Tsvetanova BC, Kiemle DJ, Price NPJ (2002): Biosynthesis of tunicamycin and metabolic origin of the 11-carbon dialdose sugar, tunicamine. J. Biol. Chem. 277: 35289-35296.
2. Price NPJ, Tsvetanova B (2007): Biosynthesis of the tunicamycins: A review. J. Antibiot. 60: 485-491.
3. Tsvetanova BC, Price NPJ (2001): Liquid Chromatography-Electrospray Mass Spectrometry of Tunicamycin-Type Antibiotics. Anal. Biochem. 289: 147-156.
4. Rossmann MG, Moras D, Olsen KW (1974): Chemical and Biological Evolution of a Nucleotide-Binding Protein. Nature 250: 194-199.
5. Filling C, Berndt KD, Benach J, Knapp S, Prozorovski T, Nordling E, Ladenstein R, Jornvall H, Oppermann U (2002): Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. J. Biol. Chem. 277: 25677-25684.
6. Oppermann U, Filling C, Hult M, Shafqat N, Wu XQ, Lindh M, Shafqat J, Nordling E, Kallberg Y, Persson B, et al. (2003): Short-chain dehydrogenases/reductases (SDR): the 2002 update. Chem.-Biol. Interact. 143: 247-253.
7. Creuzenet C, Schur MJ, Lin JJ, Wakarchuk WW, Lam JS (2000): FlaA1, a new bifunctional UDP-GlcNAc C-6 dehydratase C-4 reductase from Helicobacter pylori. J. Biol. Chem. 275: 34873-34880.
8. Allard STM, Cleland WW, Holden HM (2004): High resolution x-ray structure of dTDP-glucose 4,6-dehydratase from Streptomyces venezuelae. J. Biol. Chem. 279: 2211-2220.
9. McLennan AG (2006): The Nudix hydrolase superfamily. Cell. Mol. Life Sci. 63: 123-143.
10. Xu WL, Shen JY, Dunn CA, Bessman MJ (2003): A new subfamily of the nudix hydrolase superfamily active on 5-methyl-UTP (ribo-TTP) and UTP. J. Biol. Chem. 278: 37492-37496.
11. Roth JR, Lawrence JG, Rubenfield M, Kiefferhiggins S, Church GM (1993): Characterization of the Cobalamin (Vitamin-B12) Biosynthetic Genes of Salmonella typhimurium. J. Bacteriol. 175: 3303-3316.
12. Otoole GA, Trzebiatowski JR, Escalantesemerena JC (1994): The cobC Gene of Salmonella typhimurium Codes for a Novel Phosphatase Involved in the Assembly of the Nucleotide Loop of Cobalamin. J. Biol. Chem. 269: 26503-26511.

13. Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, et al. (2009): CDD: specific functional annotation with the Conserved Domain Database. Nucl. Acids Res. 37: D205-D210.

14. Rigden DJ (2008): The histidine phosphatase superfamily: structure and function. Biochem. J. 409: 333-348.

15. Sanchez MB, Barrado P, Jimenez A, Lobato MF (2006): The pur3 gene from the pur cluster encodes a monophosphatase essential for puromycin biosynthesis in Streptomyces. FEBS Lett. 580: 1807-1811.

16. Espinosa JC, Tercero JA, Rubio MA, Jimenez A (1999): The pur7 gene from the puromycin biosynthetic pur cluster of Streptomyces alboniger encodes a nudix hydrolase. J. Bacteriol. 181: 4914-4918.

17. Wang SC, Frey PA (2007): S-adenosylmethionine as an oxidant: the radical SAM superfamily. Trends Biochem. Sci. 32: 101-110.

18. Jarrett JT (2003): The generation of 5'-deoxyadenosyl radicals by adenosylmethionine-dependent radical enzymes. Curr. Opin. Chem. Biol. 7: 174-182.

19. Martin JL, McMillan FM (2002): SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. Curr. Opin. Struct. Biol. 12: 783-793.

20. Trefzer A, Salas JA, Bechthold A (1999): Genes and enzymes involved in deoxysugar biosynthesis in bacteria. Nat. Prod. Rep. 16: 283-299.

21. Lairson LL, Henrissat B, Davies GJ, Withers SG (2008): Glycosyltransferases: Structures, functions, and mechanisms. Annu. Rev. Biochem. 77: 521-555.

22. Nakamura N, Inoue N, Watanabe R, Takahaski M, Takeda J, Stevens VL, Kinoshita T (1997): Expression cloning of PIG-L, a candidate N-acetylglucosaminylphosphatidylinositol deacetylase. J. Biol. Chem. 272: 15834-15840.

23. Newton GL, Av-Gay Y, Fahey RC (2000): N-acetyl-1-D-myo-inosityl-2-amino-2-deoxy-alpha-D-glucopyranoside deacetylase (MshB) is a key enzyme in mycothiol biosynthesis. J. Bacteriol. 182: 6958-6963.

24. Urbaniak MD, Crossman A, Chang TH, Smith TK, van Aalten DMF, Ferguson MAJ (2005): The N-acetyl-D-glucosaminylphosphatidylinositol De-N-acetylase of glycosylphosphatidylinositol biosynthesis is a zinc metalloenzyme. J. Biol. Chem. 280: 22831-22838.

25. Borghi A, Edwards D, Zerilli LF, Lancini GC (1991): Factors affecting the normal and branched-chain acyl moieties of teicoplanin components produced by Actinoplanes teichomyceticus. J. Gen. Microbiol. 137: 587-592.

26. Carman GM, Han GS (2006): Roles of phosphatidate phosphatase enzymes in lipid metabolism. Trends Biochem. Sci. 31: 694-699.

27. Carman GM, Han GS (2009): Phosphatidic Acid Phosphatase, a Key Enzyme in the Regulation of Lipid Synthesis. J. Biol. Chem. 284: 2593-2597.

28. Soding J (2005): Protein homology detection by HMM-HMM comparison. Bioinformatics 21: 951-960.

29. Vetting MW, de Carvalho LPS, Yu M, Hegde SS, Magnet S, Roderick SL, Blanchard JS (2005): Structure and functions of the GNAT superfamily of acetyltransferases. Arch. Biochem. Biophys. 433: 212-226.

30. Gu LC, Geders TW, Wang B, Gerwick WH, Hakansson K, Smith JL, Sherman DH (2007): GNAT-like strategy for polyketide chain initiation. Science 318: 970-974.

31. Davidson AL, Dassa E, Orelle C, Chen J (2008): Structure, function, and evolution of bacterial ATP-binding cassette systems. Microbiol. Mol. Biol. Rev. 72: 317-364.

32. Mendez C, Salas JA (2001): The role of ABC transporters in antibiotic-producing organisms: drug secretion and resistance mechanisms. Res. Microbiol. 152: 341-350.

33. Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, et al. (2002): Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). Nature 417: 141-147.

34. Datsenko KA, Wanner BL (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.

35. Macneil DJ, Gewain KM, Ruby CL, Dezeny G, Gibbons PH, Macneil T (1992): Analysis of Streptomyces avermitilis Genes Required for Avermectin Biosynthesis Utilizing a Novel Integration Vector. Gene 111: 61-68.

36. Simon R, Priefer U, Puhler A (1983): A Broad Host Range Mobilization System for In vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. Nat. Biotechnol. 1: 784-791.

37. Yanai K, Murakami T, Bibb MJ (2006): Amplification of the entire kanamycin biosynthetic gene cluster during empirical strain improvement of Streptomyces kanamyceticus. Proc. Natl. Acad. Sci. U.S.A. 103: 9661-9666.

38. Gregory MA, Till R, Smith MCM (2003): Integration site for Streptomyces phage fBT1 and development of site-specific integrating vectors. J. Bacteriol. 185: 5320-5323.

39. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000): Practical Streptomyces Genetics. John Innes Foundation: Norwich.