Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the potent antibiotic taromycin

Kazuya Yamanaka[a,b], Kirk A. Reynolds[a,c], Roland D. Kersten[a], Katherine S. Ryan[a,d], David J. Gonzalez[e], Victor Nizet[e,f], Pieter C. Dorrestein[a,c,f], and Bradley S. Moore[a,f,1]

[a]*Scripps Institution of Oceanography, Departments of* [c]*Chemistry and Biochemistry and* [e]*Pediatrics, and* [f]*Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California at San Diego, La Jolla, CA 92093-0204*

[b]*Yokohama Research Center, JNC Corporation, 5-1 Ookawa, Kanazawa-ku, Yokohama, Kanagawa 236-8605, Japan*

[d]*Department of Chemistry, University of British Columbia, Vancouver, Canada*

[1] ***Corresponding author***: bsmoore@ucsd.edu, Tel.: +1 (858) 822-6650, Fax: +1 (858) 534-1318

1. Experimental methods

Strains and culture conditions.

Strains used in this study are listed in Table S1. A highly transformable *Saccharomyces cerevisiae* strain VL6-48 (MAT alpha, his3-D200, trp1-D1, ura3-52, lys2, ade2-101, met14, psi+cir0) was used as a host for gene cluster direct cloning experiments. This strain was obtained from the American Type Culture Collection (ATCC no. MYA-3666). The yeast cells grown in liquid YPD medium (Yeast extract Peptone Dextrose medium; 2% D-glucose, 1% yeast extract, and 2% peptone) supplemented with 100 mg/l adenine were used for spheroplasting prior to transformation associated recombination (TAR). Yeast transfomants were selected on synthetic tryptophan drop-out agar (SD-Trp agar) consisting of 0.17% yeast nitrogen base without amino acids and ammonium sulfate (Sigma Y1251), 0.19% yeast synthetic drop-out medium supplements without tryptophan (Sigma Y1876), 1 M sorbitol, 2% D-glucose, 0.5% ammonium sulfate, 100 mg/l adenine and 2% agar. *E. coli* Top10 (Invitrogen) was used as a host for propagation of plasmids. *E. coli* ET12567 was used as the non-methylating plasmid donor strain for intergeneric conjugation with *Streptomyces coelicolor* strain M1146 (*Δact Δred Δcpk Δcda*)[1] *S. coelicolor* M1146 was used for heterologous expression experiments and was cultured in MP media containing 1% glucose, 1% soluble starch, 0.4% peptone, 0.3% yeast extract, 0.3% soytone, 0.2% meat extract, 0.2% $CaCO_3$, and 0.5% NaCl. Kanamycin (Kan, 50 μg/ml), apramycin (Apr, 50 μg/ml), chloramphenicol (Chl, 12.5 μg/ml), and nalidixic acid (Nal, 25 μg/ml) were used for selection of *E. coli* and *Streptomyces* recombinant strains.

Construction of the gene cluster capture vector pCAP01.

The yeast-element consisting of ARSH4/CEN6 (replication origin) and TRP1 marker was amplified from a yeast centromeric plasmid, pRS314 (ATCC 77143), with primers yeast-element-F (5'-TCG**ACTAGT**AAC**CTCGAG**ACTTGA**GGTACC**TGTATTTAGAaaaataaacaaataggg-3') and yeast-element-R (5'-TAAACTTGGTCTGACAGTTAGgttcacgtagtgggccatcgcc-3'). Newly introduced restriction sites for SpeI, XhoI, and KpnI are shown in bold, and the capitalized letters represent overlapping regions with flanking elements for the entire capture vector assembly in yeast. *pUC ori* from SuperCos1 (Stratagene) was amplified as the *E. coli*-element with primers Ecoli-element-F (5'-GATGGCCCACTACGTGAACCTAACTGTCAGACCAAGTTTActcatatata-3') and Ecoli-element-R (5'-CCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAaatg-3'). To generate the *Streptomyces* element consisting of the φC31 integrase gene (*int*) and its attachment site (*attP*), origin of DNA transfer (*oriT*), and Kan resistance marker, the *aac(3)IV* gene (apramycin resistance) on the integrative plasmid pSET152 was replaced with the *aph(3)II* gene (Kan/Neo resistance) from pK18mob (accession No. AF012346) as follows. The *aph(3)II* gene with its promoter was PCR amplified with primers aph-sacF (5'-cgataag**GAGCTC**tcacgctgccgcaagcactcag-3') and

aph-sacR (5'-cccagag**GAGCTC**tcagaagaactcgtcaagaaggcga-3'). Bold letters represent SacI restriction sequence. The PCR product was digested with SacI and ligated with the larger fragment of similarly digested pSET152, yielding a Kan resistant variant of pSET152 called pLAE101. The *Streptomyces* elements generated as detailed above were then PCR amplified from pLAE101 with primers Streptomyces-elements-F (5'- TCGAACCCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAAGGcgat-3') and Streptomyces-elements-R (5'- TCTAAATACA**GGTACC**TCAAGT**CTCGAG**GTT**ACTAGT**CGAtctttgtagaaaccatcgg-3').

To assemble all three elements in yeast, fragments having 40 to 41-nt overlaps at both ends were combined and were introduced into the *S. cerevisiae* strain VL6-48 by electroporation. A single transformant colony selected on SD-Trp agar was inoculated into SD-Trp liquid media (agar was omitted from SD-Trp agar) and grown for 36 h at 30°C with shaking. The construct extracted from the cells was re-introduced into *E. coli* Top10 cells by electroporation and then was isolated from an Kan^r clone, yielding the gene cluster capture vector pCAP01 (Figure 1a).

Construction of the taromycin (*tar*) gene cluster specific capture vector.

The *tar* pathway specific capture vector was constructed by introducing into the vector two short homology arms corresponding to flanking regions of the target gene cluster.[2-4] For the left end capture arm of the taromycin (*tar*) gene cluster, a 1.0-kb region corresponding to the *tar1* membrane protein gene was PCR

amplified with tar-leftend-F (5'-ttatatt**actagt**tgtgttctttcctcgccttgaaccgt-3') and tar-leftend-R (5'-AGCGGCATCG**GGATCC**CGCTCATCGCcggggccatg-3'). SpeI and BamHI restriction sites are shown in bold. The other 1.0-kb region, on the right end of the *tar* gene cluster, corresponding to *orf6* was amplified with primers tar-rightend-F (5'-GCGATGAGCG**GGATCC**CGATGCCGCtattaggcggctg-3') and tar-rightend-R (5'-aacctta**ctcgag**tggaccgggtcctgatcagg-3'). BamHI and XhoI restriction sites are shown in bold. The two PCR amplified capture arms were then combined and assembled into single piece (2.0-kb) by PCR with primers tar-leftend-F and tar-rightend-R. The assembled fragment was digested with SpeI and XhoI and introduced into pCAP01, yielding the *tar* gene cluster specific capture vector. Prior to direct TAR cloning, the "circular" construct was digested with BamHI, generating linear capture vector flanked by 1.0-kb capture arms at both ends.

Direct cloning of the *tar* gene cluster using TAR.

Direct TAR cloning of the *tar* gene cluster from genomic DNA was carried out using a previously reported protocol with some modifications.[4] *S. cerevisiae* strain VL6-48 was grown in 100 ml YPD medium supplemented with adenine (100 mg/l) at 30°C with shaking until OD600 of 0.7-1.0 was reached. The cells were harvested and washed with ice-cold water and osmotically stabilized in 1 M sorbitol at 4°C for 20 h prior to spheroplasting. Preparation of spheroplast cells was carried out using a lytic enzyme, Zymolyase-20T (MP Biomedicals), as reported[4] but with extended reaction time (40 min incubation). Spheroplast cells were transformed with 0.5 to 1 μg of genomic DNA fragments and 0.4 μg linearized *tar* gene cluster specific capture vector according to the protocol.[4] The transformed spheroplasts were mixed with 7 ml SD-Trp top agar (SD-Trp containing 3% agar) equilibrated at 55°C and overlaid on SD-Trp agar. The plates were incubated at 30°C for 3 days. One hundred of transformants appeared on the plates were picked with toothpicks and transferred onto new SD-Trp agar plate, and the plate was incubated for 2 days at 30°C.

For screening of the directly cloned *tar* gene cluster, colony PCR was conducted. Cells were spheroplasted using Zymolyase-20T at 30°C for 2 h and then were boiled at 98°C for 5 min. The supernatants were used as a template for PCR screening using primers amplifying a 1-kb region in the middle of the cluster (tarscreeningF = 5'-accgtggtgccggtggaggttc-3', tarscreeningR = 5'-acggcaccgcgctcaccagtc-3'). Plasmids were extracted from identified clones and reintroduced into *E. coli* Top10 cells by electroporation. The plasmids were then purified from Kan^r *E. coli* clones, and direct cloning of the *tar* gene cluster into the capture vector pCAP01 was confirmed by restriction analysis with NcoI or PstI. The yielded construct was designated as pCAP01-*tar*.

Genetic manipulation of the *tar* regulatory genes.

Regulatory gene manipulations were carried out using yeast *in vivo* homologous recombination-mediated PCR targeting.[5] The *tar19* and/or *tar20* regulatory genes located at boundary of the cluster and all unrelated genes beyond the regulator genes were eliminated from pCAP01-*tar* by gene replacement with PCR amplified URA3 auxotrophic marker gene flanked by 39 nucleotide homology arms, which are identical to upstream or downstream region of the target region, as follows. To eliminate the region from the *tar20* to *orf6,* the URA3 gene including its promoter was PCR amplified from pGAL-MF plasmid with primers tarluxdel-F (5'-AGCCTCAACAGTTCGCGGCCGGCGTCCGCTGGTGCAGTGtctagacagattgtactgagagtgca-3') and tarrightldel-R (5'-TGGACCGGGTCCTGATCAGGCGGTCTCCCGGTCACGTTCtctagaatttgtgagtttagtatacatgc-3'). The capitalized letters represent homology arms for recombination. *S. cerevisiae* VL6-48 carrying pCAP01-*tar* was then transformed with the URA3 cassette (1.1-kb) by electroporation. Resultant plasmid from a single transformant selected on uracil deficient SD agar was reintroduced into *E. coli* Top10 cells by electroporation. The plasmid was purified from Kan^r clone, and elimination of the target region was confirmed by restriction analysis. The resulting construct, in which the *tar20* gene and all the downstream genes were eliminated, was designated as pCAP01-*tar*M2. For elimination of the region from the *tar19* to *orf6,* the URA3 cassette was prepared by PCR with primers tarsarpdel-F (5'-GCAAGTGTGATGAATGTTTCGTGACAACGCACTAGGCTCtctagacagattgtactgagagtgca-3') and tarrightldel-R. *S. cerevisiae* VL6-48 carrying pCAP01-*tar* was transformed with the cassette, and the mutant plasmid was then prepared in the same way, yielding pCAP01-*tar*M1, in which both the *tar19* and *20* genes were eliminated.

Heterologous expression of the *tar* gene cluster.

The pCAP01-*tar* and its derivatives that have the φC31 integrase (*int*) gene with its attachment site (*attP*) were introduced into *E. coli* ET1256[6] and transferred to *S. coelicolor* M1146[1] by triparental intergeneric conjugation facilitated by *E. coli* ET12567/pUB307.[7] Kan^r exconjugants were selected on MS agar containing Nal and Kan. Heterologous *S. coelicolor* strains were routinely pre-cultivated in Trypticase Soy Broth (TSB) containing Kan and Nal at 30 °C for 3 days. A portion (0.5 ml) of the precultures was used to inoculate 50 ml modified MP production medium[8] consisting of 1% glucose, 1% soluble starch, 0.2% peptone, 0.3% yeast extract, 0.3% soytone, 0.2% malt extract, 25 mM TES (pH7.2), and 0.5% NaCl, which was grown for 4 to 5 days at 30°C in a 250 ml flask with rotary shaking. Solid phase extraction of the culture supernatants was performed with Amberlite XAD resin (SIGMA) as described previously.[9] The methanol extracts from the resin were analyzed by reversed phase HPLC and also HPLC-MS. The analytical conditions for HPLC were as follows: Luna 5 μ C18 column (100 x 4.6 mm; Phenomenex Inc., USA) was used at room temperature with gradient flow of aqueous acetonitrile buffered with 0.1% TFA (20 – 60% acetonitrile over 30 min at a flow rate of 1.0 ml/min); detection was at 254 nm.

Genetic complementation of the *tar* regulatory genes.

To reintroduce the eliminated regulatory genes back into the regulatory gene deficient mutants, a chromosomal integrative expression vector having φBT1 phage integrase (*int*) gene with its attachment site (*attP*) and strong constitutive *ermE** promoter, pKY01, was used.[8] The vector has the *aac(3)IV* gene providing Apr^r in both *E. coli* and actinobacteria as a selective marker. The *tar19* sarp regulator gene was PCR amplified with primers tarSARP-ndeF (5'-aggtccc**catatg**gcggtcctgggacagttg-3') and tarSARP-hindR(5'-ctctagt**aagctt**tcagggcttcctccgcagcg-3'), and the resultant product was inserted into pKY01 after restriction digestion, generating pKY01-*tar19*. For the *tar20* luxR-type regulator expression construct, pKY01-*tar20*, a primer set tarLuxR-ndeF (5'-aggtccc**catatg**gagttagacccaaaggtaccgatc-3') and tarLuxR-hindR (5'-ctctagt**aagctt**tcaacagttcgcggccggcgt-3') was used. NdeI and HindIII restriction sites are shown in bold. The resultant two expression constructs were then transferred to two regulatory gene deficient mutants, *S. coelicolor* M1146 / pCAP01-*tar*M1 and *S. coelicolor* M1146 / pCAP01-*tar*M2, by conjugal transfer as mentioned above. Four mutants were selected against Apr and Nal and subjected to expression.

Overexpression of the codon redressed *tar20* regulatory genes.

The TTA codon coding for Leu is known to be very rare in *Streptomyces* strains. To redress all five TTA codons found in the *tar20* gene, PCR-mediated mutagenesis exchanging codons was carried out as follows. The *tar20* gene was PCR amplified as five overlapping fragments divided at TTA codons with the following primer pairs:

tar20S1-ndeF (5'-aggtctc**catatg**gagctggacccaaaggtacc-3')
tar20S1-R (5'-TCAGCAGCTCCGTCTTTCCGctggcgacactgcccgtga-3')

tar20S2-F (5'-CGGAAAGACGGAGCTGCTGAacactttcgccgatacggcct-3')
tar20S2-R (5'-AGCAGCCGCTGAACATTTCGgggcgcctcggccgtcgatg-3')

tar20S3-F (5'-CGAAACGTTCAGCGGCTGCTcgaggaaccgggcactgacg-3')
tar20S3-R (5'-TCCCGCAGCGCCGTCCTGAGgccgttgatccgcaggtcgg-3')

tar20S4-F (5'-CTCAGGACGGCGCTGCGGGAcgggcacctgagcgcacac-3')
tar20S4-R (5'-GGGCAGCGCGCCGTCACCGAccacatcggagggaacgtcgg-3')

tar20S5-F (5'-TCGGTGACGGCGCGCTGCCCccggcggaggaaccgtggc-3')
tar20S5-hindR (5'-ctctagt**aagctt**tcaacagttcgcggccggcgt -3').

The capitalized letters represent overlapping regions for following PCR assembling and NdeI and HindIII restriction sites are shown in bold. In the above shown primers, TTA codons were redressed to CTG to give the same amino acid, Leu, as shown in underlined letters. Five overlapping PCR products were then assembled and amplified as a single piece by PCR using primers tar20S1-ndeF and tar20S5-hindR. The resultant codon redressed *tar20* gene was cloned into the NdeI-HindIII site of pKY01. The construct, in which correct codon replacement was confirmed by sequencing, was then transferred to two regulatory gene deficient mutants, *S. coelicolor* M1146 / pCAP01-*tar*M1 and *S. coelicolor* M1146 / pCAP01-*tar*M2, by conjugal transfer in the same way. Resultant mutants selected against Apr and Nal were subjected to expression.

Isolation of taromycins

A methanol extract from a 1 L fermentation produced ~150 mg of crude material. The crude methanol extract was redissolved in H_2O (1 mL) and purified by preparative HPLC (C18, 15 mL/min, 20% CH_3CN/H_2O+0.1% TFA to 80% CH_3CN/H_2O+0.1% TFA over 30 min). Fractions containing taromycin, as indicated by MS-MS fragmentation, were combined and dried under vacuum to produce a crude taromycin mixture (15 mg). The crude taromycin mixture was then purified by semi-preparative HPLC (C18, 15 mL/min, 20% CH_3CN+0.1%/H_2O+0.1% TFA to 80% CH_3CN+0.1%/H_2O+0.1% TFA over 30 min) to yield taromycin A (1 mg).

Structural elucidation of taromycins.

High resolution mass spectrometry analysis of taromycin A showed an isotopic distribution consistent with two chlorine atoms. FT-MS analysis of taromycin A showed the amino acid sequence identical to daptomycin with the exception of the serine residue being substituted with an additional alanine (Figure S9). The mass fragment of the kynurenine and tryptophan residues indicated a single chlorine atom on each residue. The locations of the chlorine atoms on the tryptophan and kynurenine residues were identified by comparative HSQC and HMBC NMR analysis against daptomycin (Tables S3, S4 and Figures S10, S11). Taromycin A (200 µg) was hydrolyzed (6 M HCl, 160°C, 5 min) and dried down under a stream of nitrogen. The resulting solid was redissolved in 1 M sodium bicarbonate (200 µl) and to this was added 1 mL of 1-fluoro-2,4-dinitrophenyl-D-alanine amide (D-FDAA) in acetone (1 mg/ml). The reaction was stirred at 45 °C for 1 h, quenched with 1 M HCl (200 µL), and then dried down under a stream of nitrogen. The resulting solid was then redissolved in 1:1 H_2O:CH_3CN (200 µl) and filtered into an LC-MS vial. The resulting solution was analyzed (15 µL) by LC-MS under the conditions described under Table S5. A second batch of taromycin A (200 µg) was hydrolyzed (6 M HCl containing 5% phenol, 160°C, 3 min)[12] and treated with D-FDAA as reported above. Daptomycin was treated identically to taromycin A to yield additional retention times recorded in Table S5. Amino acid standards (50 µl of 50 mM in 1 M $NaHCO_3$) were reacted with D-FDAA at 45°C for 1 h. Following concentration by N_2 and filtration, each amino acid standard was injected according to identical conditions used for taromycin A and daptomycin. Retention times are tabulated in Table S3.

Functional validation of pCAP01 system.

Function of the newly developed pCAP01 was validated by the direct cloning of the 30-kb marinopyrrole (*mpy*) biosynthetic gene cluster from *Streptomyces* sp. CNQ-418 that we previously reported (accession no. JX157625) and its heterologous expression in *Streptomyces coelicolor* M512 prior to direct cloning of the *tar* gene cluster as follows. For construction of the *mpy* cluster specific capture vector, a 1-kb region containing the *mpy1* gene was PCR amplified as left side capture arm with primers mpy-leftendF (5'-aagatcg**actagt**agagcgccaacagggccg-3') and mpy-leftendR (5'-TACCGGGAAAGCGCC**GGATCC**TCAGATGGTGAAGGTGCGAggaaa-3'). SpeI and BamHI restriction sites are shown in bold. The right side capture arm (1-kb) containing the *mpy16* gene was amplified by PCR with primers mpy-rightendF (3'-TCGCACCTTCACCATCTGA**GGATCC**GGCGCTTTCCCGGTActggagacc-5') and mpy-rightendR (3'-cggaacc**ctcgag**ctcgctcttcctggctg-5'). BamHI and XhoI restriction sites are shown in bold. For each primer set used to amplify the 1 kb capture arms, the capitalized letters represent overlapping regions of homology between the two 1-kb fragments for PCR assembly of a 2-kb insert. The 2-kb insert was assembled from the PCR products by PCR using primers mpy-leftendF and mpy-rightendR. The assembled fragment was digested with SpeI and XhoI and introduced into pCAP01, yielding the *mpy* gene cluster

specific capture vector. Prior to direct TAR cloning, the “circular” construct was digested with BamHI, generating linear capture vector flanked by 1.0-kb capture arms at both ends.

Spheroplast cells of *S. cerevisiae* strain VL6-48 were transformed with HindIII digested genomic DNA from *S.* sp. CNQ418 and the linearized (BamHI digested) *mpy* pathway specific capture vector as described above. Colonies grown on SD-Trp agar were subjected to PCR screening with primers mpy-ckF (5’-tccacgaccgtagaggccagtg-3’) and mpy-ckR (5’-gactcgctgacttcggagagacac-3’), and the desired plasmid was extracted from an identified clone. Restriction analysis of the plasmid propagated in *E. coli* confirmed direct cloning of the *mpy* cluster, yielding pCAP01-*mpy*.

pCAP01-*mpy* was then transferred to *S. coelicolor* M512 by triparental intergeneric conjugation facilitated by *E. coli* ET12567/pUB307.[7] A resultant Kan^r clone was subjected to heterologous expression experiment as previously reported,[8] and successful production of marinopyrrole A was confirmed (Figure S12).

2. Strains and plasmids used in this study

Table S1. Strains and plasmids used in this study.

strain / plasmid	description	source
plasmids		
pRS314	source of ARSH4/CEN6 and TRP1: ARSH4/CEN6, TRP1, *pMB1 ori*, *bla.*	ATCC 77143
SuperCos-1	source of *pUC ori*: *pUC ori, bla, neo, cos.*	Stratagene
pLAE101	pSET152 derivative, source of *aph(3)II,* φC31 *int-attP, and oriT* (RP4): *pMB1 ori, aph(3)II,* φC31 *int-attP, oriT* (RP4).	this study
pGAL-MF	source of URA3 gene: *pBS ori, bla,* 2micron ori, URA3.	Dualsystems
pCAP01	gene cluster capture vector: ARSH4/CEN6*, pUC ori, aph(3)II, φC31 int-attP, oriT* (RP4).	this study
pUB307	self-transmissible plasmid that mobilizes other plasmids *in trans* for DNA transfer into hosts: RP4, *neo*	7
pKY01	pMS82 derivative for gene expression in *Streptomyces*: *pMB1 ori,* φ*BT1 int-attP, oriT* (RP4), *aac(3)IV, ermEp*.*	8
pK18mob	source of *aph(3)II* gene: *aph(3)II, pMB1 ori, oriT*(RP4)	10
pCAP01-*tar*	pCAP01 derivative that carries 73-kb genomic region containing the entire *tar* gene cluster (*tar1-20*).	this study
pCAP01-*tar*M1	pCAP01-*tar* derivative (*Δtar19-20*): *tar1-18,* URA3.	this study
pCAP01-*tar*M2	pCAP01-*tar* derivative (*Δtar20*): *tar1-19*, URA3.	this study
strains		
S. cerevisiae		
VL6-48	host strain for *in vivo* homologous recombination: MAT alpha, his3-D200, trp1-D1, ura3-52, lys2, ade2-101, met14, psi+cir0.	ATCC MYA-3666
E. coli		
Top10	host strain for routine cloning: F-, *Δ(araA-leu)7697*, *[araD139]$_{B/r}$*, *Δ(codB-lacI)3*, φ80d*lacZ58*(M15), *galK0*, *mcrA0*, *galU-*, *recA1*, *endA1*, *nupG-*, *rpsL-* (strR), *Δ(mcrC-mrr)715*	Invitrogen
ET12567	DNA methylation deficient donor strain for conjugation: F2 *dam13*::Tn*9, dcm,6 hsdM, hsdR, recF143, zjj-202*::Tn*10, galK2, galT22 ara-14, lacY1, xyl-5, leuB6, thi-1, tonA31, rpsL136, hisG4, tsx-78, mtl-1, glnV44*	6
S. coelicolor		
M512	host strain for heterologous expression: *ΔactII-orf4*, Δ*redD*, SCP1-, SCP2-	11
M1146	host strain for heterologous expression derived from *S. coelicolor* M145: *Δact, Δred, Δcpk, Δcda.*	1

3. Deduced functions and sequence comparison of the *tar* genes

Table S2. Deduced functions of ORFs in the *tar* locus in *Saccharomonospora* sp. CNQ490. Gene organization is shown in Figure 2.

orf	product (aa size)	closest homolog [source]	homologous Dpt enzymes	identity/ similarity (%)	accession no.	proposed function
orf-3	294	sugar kinase [*Saccharomonospora cyanea* NA-134]	-	87/93	ZP_09746218.1	-
orf-2	268	gluconolaconase/LRE domain-containing protein *[Streptomyces* sp. SirexAA-E]	-	63/72	YP_004806341.1	-
orf-1	308	putative oxidoreductase [*Gordonia soli* NBRC 108243]	-	75/81	ZP_21227872.1	-
tar1	230	membrane-associated protein [*Saccharomonospora viridis* DSM 43017]		79/87	YP_003133683.1	transporter
			DptP	47/63	AAX31552.1	
tar2	282	ABC-type multidrug transport system [*Saccharomonospora viridis* DSM 43017]		86/90	YP_003133683.1	transporter
			DptN	62/75	AAX31554.1	
tar3	295	ABC transporter ATP-binding subunit [*Saccharomonospora viridis* DSM 43017]		86/92	YP_003133683.1	transporter
			DptM	62/74	AAX31553.1	
tar4	579	acyl-CoA synthetase [*Saccharomonospora viridis* DSM 43017]		74/82	YP_003133686.1	acyl-CoA synthetase
			DptE	47/61	AAX31555.1	
tar5	572	acyl-CoA dehydrogenase [*Saccharomonospora viridis* DSM 43017]	-	78/84	YP_003133687.1	acyl-CoA dehydrogenase
tar6	549	acyl-CoA dehydrogenase [*Saccharomonospora viridis* DSM 43017]	-	78/85	YP_003133688.1	acyl-CoA dehydrogenase
tar7	87	acyl carrier protein [*Saccharomonospora viridis* DSM 43017]		76/89	YP_003133689.1	acyl carrier protein
			DptF	38/60	AAX31556.1	
tar8	5773	non-ribosomal peptide synthase [*Saccharomonospora viridis* DSM 43017]		81/87	YP_003133690.1	non-ribosomal peptide synthatase
			DptA	51/62	AAX31557.1	
tar9	7305	non-ribosomal peptide synthase [*Saccharomonospora viridis* DSM 43017]		82/87	YP_003133691.1	non-ribosomal peptide synthatase
			DptBC	51/62	AAX31558.1	
tar10	2385	non-ribosomal peptide synthase [*Saccharomonospora viridis* DSM 43017]		82/89	YP_003133692.1	non-ribosomal peptide synthatase
			DptD	55/67	AAX31559.1	
tar11	71	hypothetical protein [*Saccharomonospora viridis* DSM 43017]		90/95	YP_003133693.1	MbtH-like
			DptG	70/79	AAX31560.1	
tar12	345	methyltransferase [*Saccharomonospora viridis* DSM 43017]		82/89	YP_003133694.1	glutamate 3-methyl-transferase
			DptI	55/66	AAX31562.1	

tar13	236	tryptophan 2,3-dioxygenase [*Saccharomonospora viridis* DSM 43017]	DptJ	73/81 52/68	YP_003133695.1 AAX31562.1	tryptophan 2,3-dioxygenase
tar14	534	tryptophan halogenase [*Streptomyces viridochromogenes* Tue57]	-	79/86	ZP_21114879.1	FAD-dependent tryptophan halogenase
tar15	178	flavin reductase [*Kutzneria* sp. 744]	-	63/70	ABV56599.1	flavin reductase
tar16	288	hypothetical protein [*Saccharomonospora viridis* DSM 43017]	-	78/86	YP_003133696.1	kynurenine formamidase
tar17	247	thioesterase [*Saccharomonospora viridis* DSM 43017]	-	70/83	YP_003133697.1	thioesterase
tar18	85	-	-	-	-	unknown
tar19	980	SARP family transcriptional activator [*Saccharomonospora xinjiangensis* XJ-54]	-	75/82	ZP_09983150.1	Transcriptional activator
tar20	982	luxR family transcriptional regulator [*Saccharomonospora xinjiangensis* XJ-54]	-	73/81	ZP_09983147.1	transcriptional regulator
orf1	222	*O*-methyltransferase [*Saccharomonospora xinjiangensis* XJ-54]	-	83/90	ZP_09983146.1	-
orf2	366	methoxymalonate biosynthesis protein *[Saccharomonospora xinjiangensis* XJ-54]	-	86/91	ZP_09983145.1	-
orf3	370	acyl-CoA dehydrogenase [*Saccharomonospora xinjiangensis* XJ-54]	-	86/91	ZP_09983144.1	-
orf4	94	acyl carrier protein [*Saccharomonospora xinjiangensis* XJ-54]	-	82/89	ZP_09983143.1	-
orf5	287	3-hydroxyacyl-CoA dehydrogenase [*Saccharomonospora xinjiangensis* XJ-54]	-	85/92	ZP_09983142.1	-
orf6	298	hypothetical protein [*Saccharomonospora xinjiangensis* XJ-54]	-	74/84	ZP_09983141.1	-
orf7	274	enoyl-CoA hydratase/carnithine racemase [*Saccharomonospora xinjiangensis* XJ-54]	-	87/93	ZP_09983140.1	-
orf8	297 (truncated)	enoyl-CoA hydratase/carnithine racemase [*Saccharomonospora xinjiangensis* XJ-54]	-	85/91	ZP_09983138.1	-
orf9	116 (truncated)	enoyl-CoA hydratase/carnithine racemase [*Saccharomonospora xinjiangensis* XJ-54]	-	57/67	ZP_09983138.1	-

4. Schematic diagram of the *tar* gene cluster direct cloning using TAR.

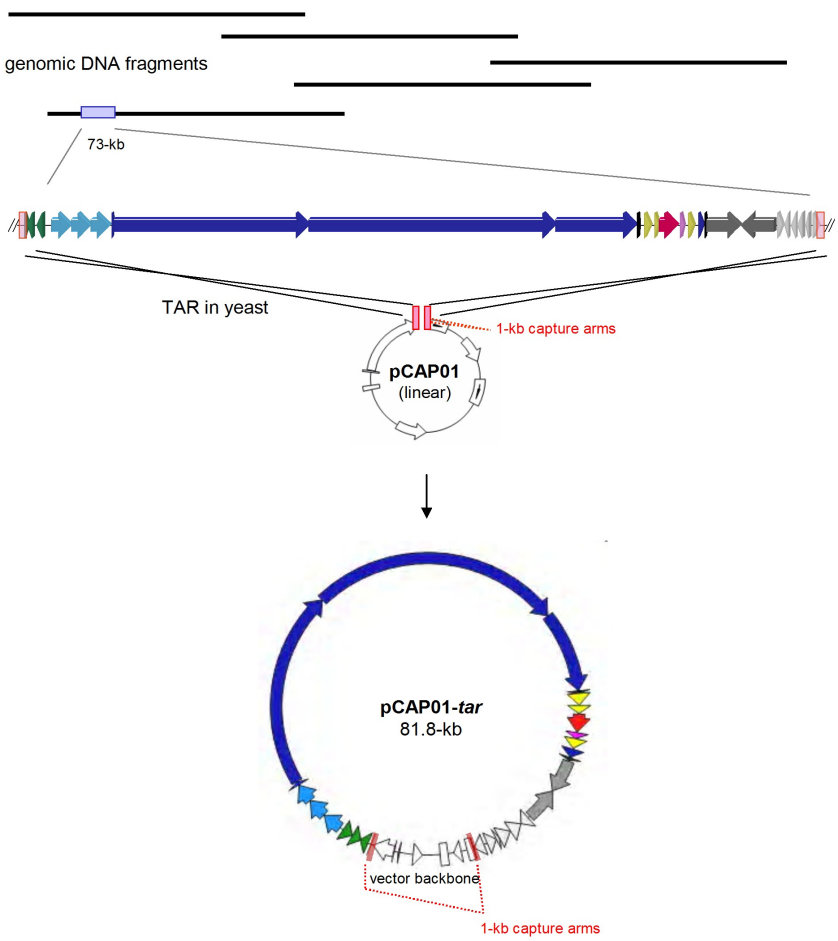

Figure S1
TAR cloning of the taromycin biosynthetic gene cluster (*tar*). Spheroplast cells of *S. cerevisiae* VL6-48 were transformed with DNA mixture consisting of enzymatically digested genomic DNA fragments of *S.* sp CNQ-490 and the linearized tar pathway specific capture vector. In yeast cells, homologous recombination occurred between homology arms pre-installed in the vector and homologous genomic DNA fragment, and thus 73-kb genomic region containing the *tar* loci was directly captured into the vector, yielding replicable circular construct pCAP01-*tar*. Yeast clones carrying the *tar* pathway were selected on agar plate lacking Trp. In the case when there is no homology between the capture vector and co-introduced genomic DNA fragment, the capture vector stays linear, resulting in no colony formation on selective agar plate.

5. Restriction mapping of pCAP01-*tar* and its derivatives.

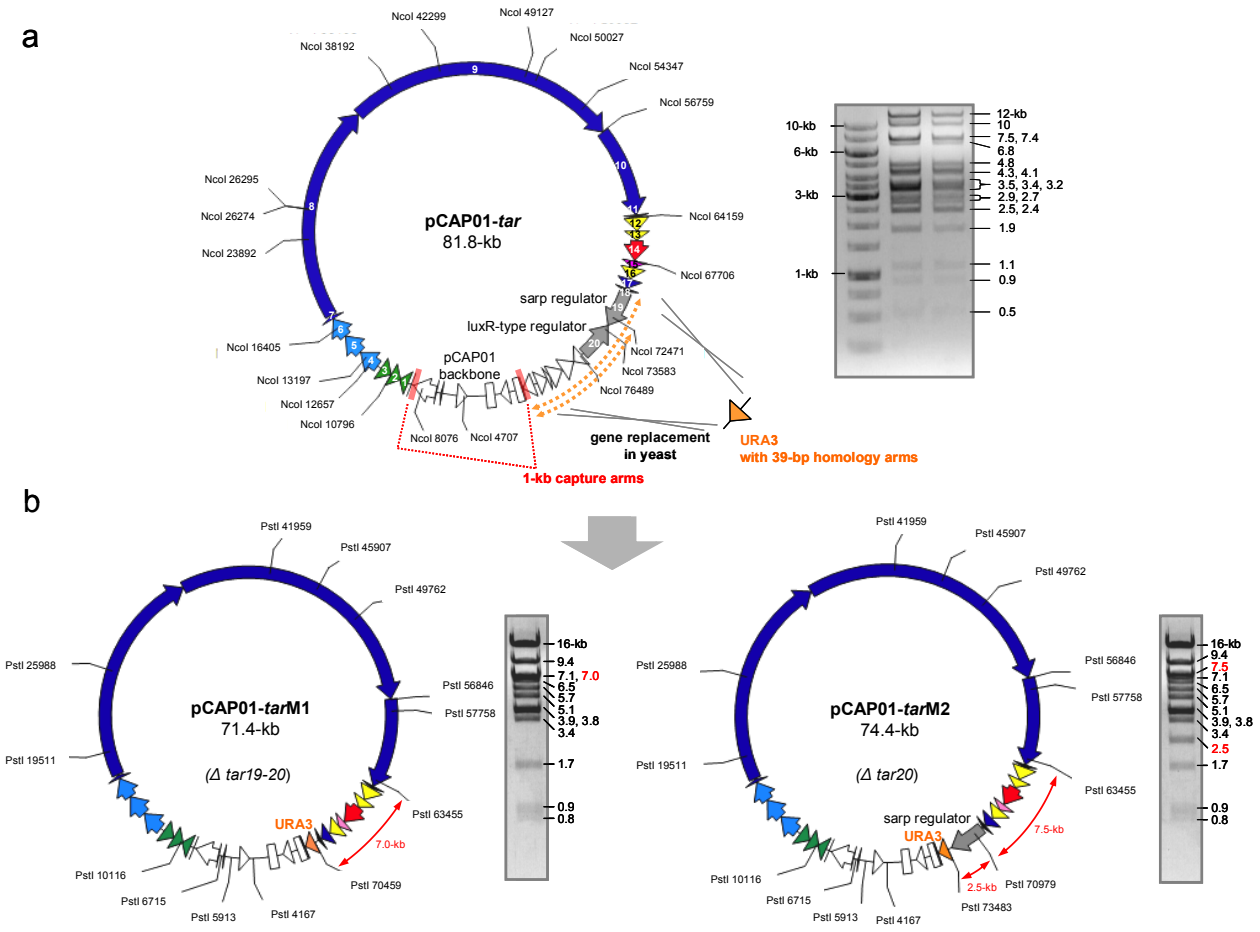

Figure S2
Restriction maps of the TAR-cloned *tar* gene clusters. (a) NcoI restriction mapping of pCAP01-*tar* propagated in two different *E. coli* clones confirmed that the 73-kb genomic region containing *tar* was correctly cloned and that the resultant 81.8-kb construct had been stably carried in *E. coli*. (b) The *tar* regulatory genes *tar19/tar20* and unrelated genes beyond the cluster border on pCAP01-*tar* [shown as dotted arrows in (a)] were replaced with the URA3 auxotrophic marker gene in yeast, generating regulatory gene deficient constructs pCAP01-*tar*M1 (Δ*tar19-20*) and pCAP01-*tar*M2 (Δ*tar20*). Successful gene replacements were confirmed by PstI restriction mapping. All three constructs were introduced into the φC31 attachment site on the chromosome of the heterologous host *S. coelicolor* M1146.

6. Result of heterologous expression of the *tar* gene cluster in *S. coelicolor* M1146

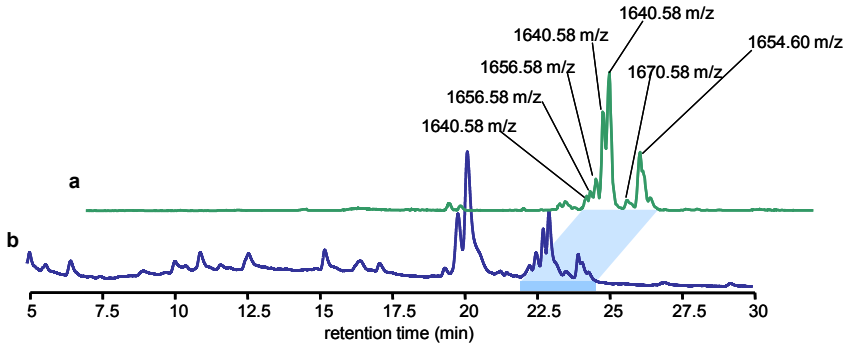

Figure S3
HPLC-MS analysis of the taromycins produced by *S. coelicolor* M1146 / pCAP01-*tar*M2 (Δ*tar20 luxR*). (a) extracted ion chromatogram (1500 – 1700 m/z) and (b) UV chromatogram monitored at 254 nm are shown. Observed ions from each peak are also indicated.

7. Result of genetic complementation of the *tar* regulatory genes.

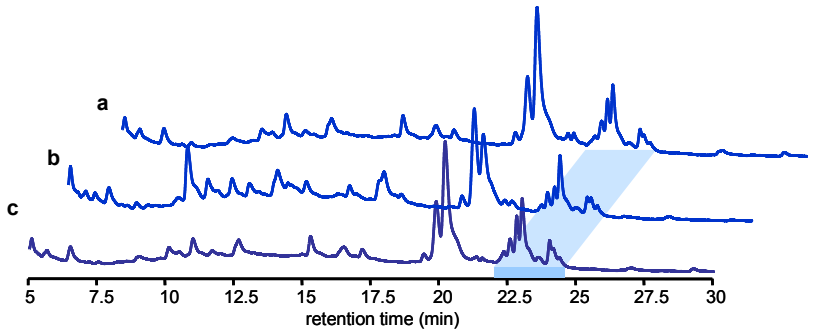

Figure S4
Effect of genetic complementation with *tar* regulatory genes in *S. coelicolor* M1146 / pCAP01-*tar*M1 (Δ*tar19 sarp*, Δ*tar20 luxR*). Extracts from (a) *S. coelicolor* M1146 / pCAP01-*tar*M1 (Δ*tar19 sarp*, Δ*tar20 luxR*) complemented with the *tar20 luxR* gene via pKY01-*tar20*, (b) *S. coelicolor* M1146 / pCAP01-*tar*M1 (Δ*tar19 sarp*, Δ*tar20 luxR*) complemented with the *tar19 sarp* gene via pKY01-*tar19*, and (c) *S. coelicolor* M1146 / pCAP01-*tar*M1 (Δ*tar19 sarp*, Δ*tar20 luxR*) were analyzed by C18 reversed phase HPLC and monitored at 254 nm.

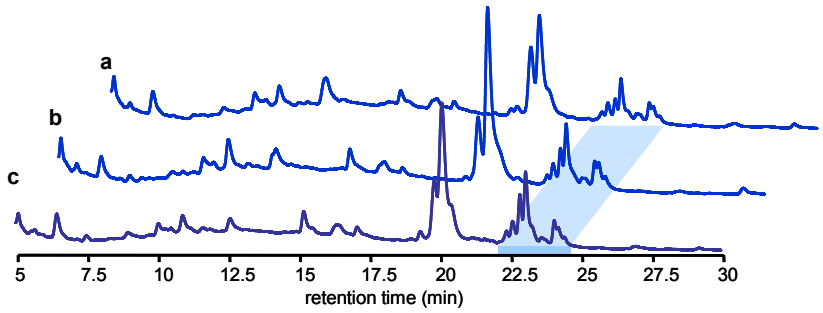

Figure S5
Effect of genetic complementation of the *tar* regulatory genes in *S. coelicolor* M1146 / pCAP01-*tar*M2 (Δ*tar20 luxR*). Extracts from (a) *S. coelicolor* M1146 / pCAP01-*tar*M2 (Δ*tar20 luxR*) complemented with the *tar20 luxR* gene via pKY01-*tar20*, (b) *S. coelicolor* M1146 / pCAP01-*tar*M2 (Δ*tar20 luxR*) complemented with the *tar19 sarp* gene via pKY01-*tar19*, and (c) *S. coelicolor* M1146 / pCAP01-*tar*M2 (Δ*tar20 luxR*) were analyzed by C18 reversed phase HPLC and monitored at 254 nm.

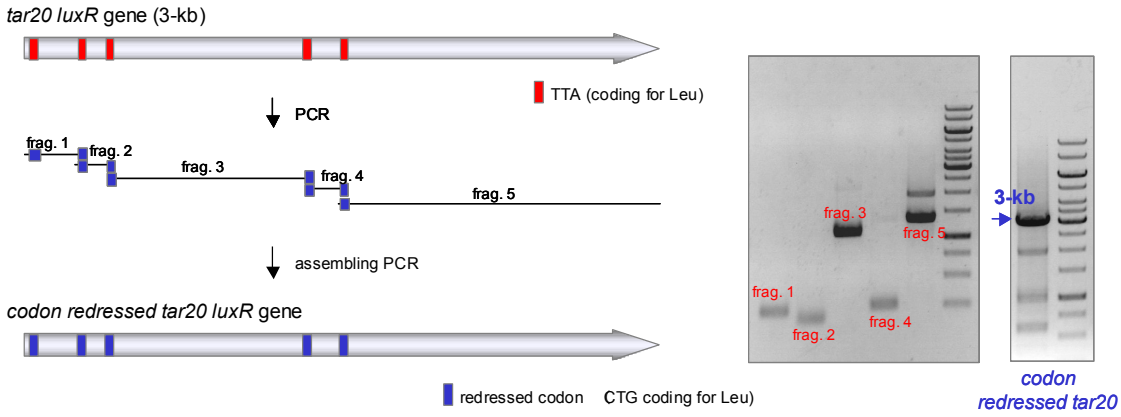

Figure S6
Schematic diagram for construction of the codon redressed *tar20 luxR* regulatory gene. Five TTA codons coding for Leu observed in the native *tar20* gene (shown as red boxes) were replaced with the alternate codon CTG by PCR-mediated mutagenesis. The *tar20* gene was divided into five pieces and PCR amplified with mutagenesis primers (gel picture on left). The five fragments having replaced codon at ends (shown as blue boxes) were then reassembled into a single fragment, generating the codon redressed *tar20* gene where all five TTA codons were replaced with CTG (gel picture on right). The resultant gene was cloned into pKY01 to give pKY01-*tar20m* and used for genetic complementation.

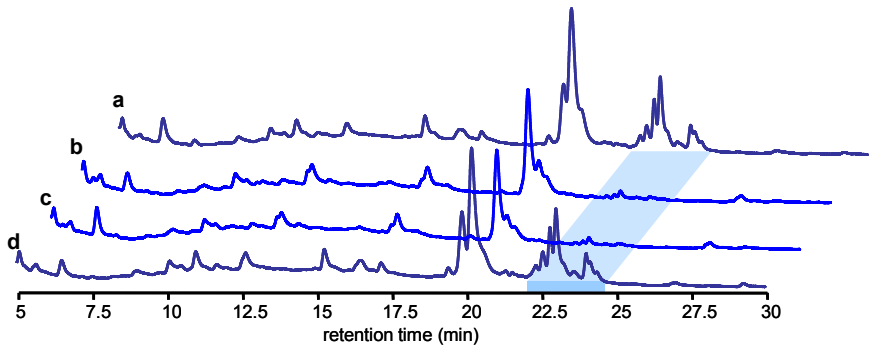

Figure S7
Effect of genetic complementation of the codon redressed *tar20 luxR* regulatory gene. Extracts from (a) *S. coelicolor* M1146 / pCAP01-*tar*M2 (Δ*tar20 luxR*), (b) *S. coelicolor* M1146 / pCAP01-*tar*M2 (Δ*tar20 luxR*) complemented with the codon redressed *tar20 luxR* gene via pKY01-*tar20m*, (c) *S. coelicolor* M1146 / pCAP01-*tar*M1 (Δ*tar19 sarp*, Δ*tar20 luxR*) complemented with pKY01-*tar20m*, and (d) *S. coelicolor* M1146 / pCAP01-*tar*M1 (Δ*tar19 sarp*, Δ*tar20 luxR*) were analyzed by HPLC and monitored at 254 nm.

8. NMR and Marfey analysis tables

Table S3. NMR correlations of the daptomycin Kyn and Trp residues (d_6-DMSO, 600 MHz)

Residue	Position	δH, mult, *J* (Hz)	δC, type	HMBC	COSY
Kyn	1				
	2				
	3				
	4				
	5		118.6, C		
	6		151.3, C		
	7	6.74, d (6.0)	116.5, CH	113.9	7.23
	8	7.23, t (6.0)	133.9, CH	130.6, 151.3	6.53, 6.74
	9	6.53, t (6.0)	113.9, CH	116.5	7.23, 7.72
	10	7.72, br s	130.66, CH	118.6	6.53
Trp	52				
	53	4.45, s	53.1, CH		2.88, 3.05
	54a	2.88, t (9.0)	26.8, CH_2	110.1, 123.2, 127.3	3.05, 4.45
	54b	3.05, s		110.1	2.88, 4.45
	55		110.1, C		
	56	7.13, s	123.3, CH	110.1, 127.3, 136.0	10.77
	NH_{indole}	10.77, s		110.1, 123.3, 127.3, 136.0	7.13
	57		136.0, C		
	58	7.30, d (6)	110.7	117.5, 127.3	7.04
	59	7.04, t (7.2)	120.4	117.8, 136.0	7.30, 6.96
	60	6.96, t (7.8)	117.5	110.7	7.04, 7.56
	61	7.56, d (8.4)	117.8	110.1, 120.4, 127.3, 136.0	6.96
	62		127.3, C		

Table S4. NMR correlations of the taromycin A 4-Cl-Kyn and 6-Cl-Trp residues (d_6-DMSO, 600 MHz)

Residue	Position	δH, mult, *J* (Hz)	δC, type	HMBC	COSY
4-Cl-Kyn	1				
	2				
	3				
	4				
	5				
	6				
	7	6.82, d (6.0)	115, CH		
	8				
	9	6.53, m	113.9, CH		7.78
	10	7.78, m	132.8 CH		
6-Cl-Trp	52				
	53	4.45, s	53.1, CH		

	54a	2.88, t (7.8)	26.8, CH_2		
	54b	3.05, s		53.1	
	55		110.3, C		
	56	7.19, s	124.4, CH	110.3, 126.3, 136.4	
	NH_{indole}	10.97, s			
	57		136.4, C		
	58	7.35, s	110.3, CH	118.5, 125.3, 126.3	
	59				
	60	6.96, d (6.0)	118.1, CH	110.9, 126.3	7.60
	61	7.60, d (6.0)	119.4, CH	125.3, 136.4	6.96
	62		126.3, C		

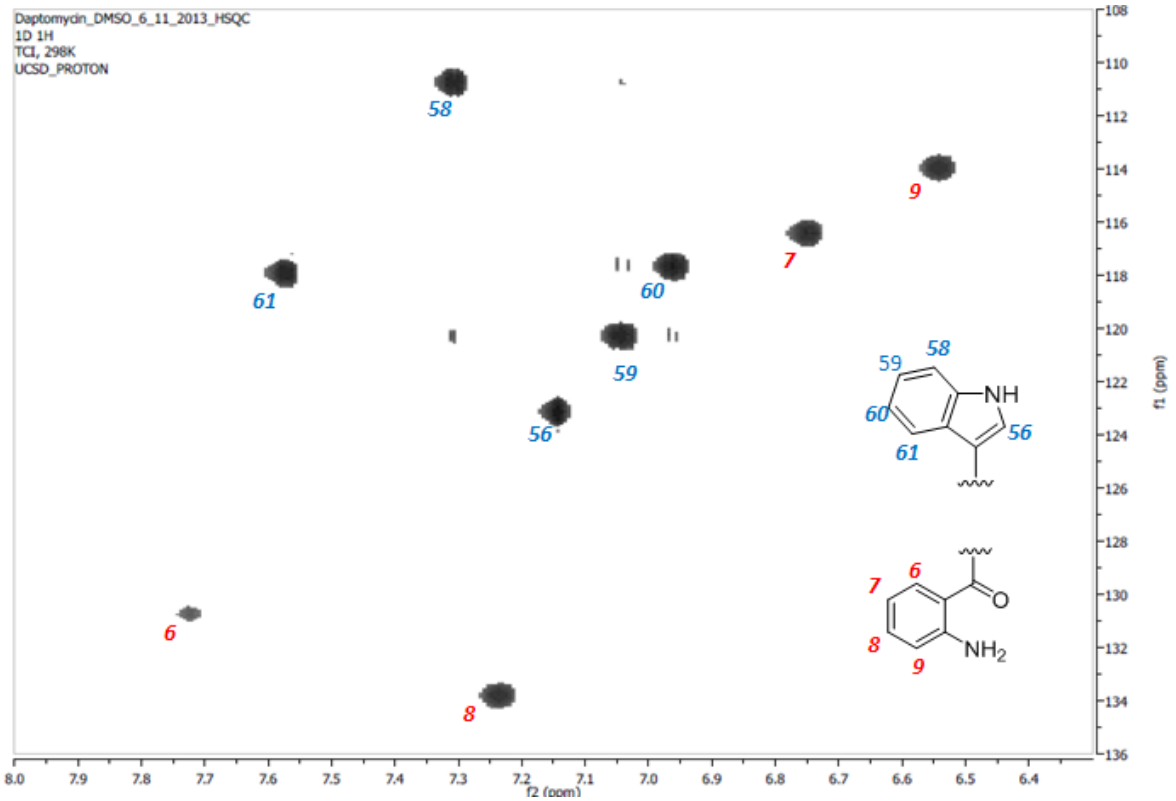

Figure S8
HSQC spectra of daptomycin aromatic region (d_6-DMSO, 600 MHz)

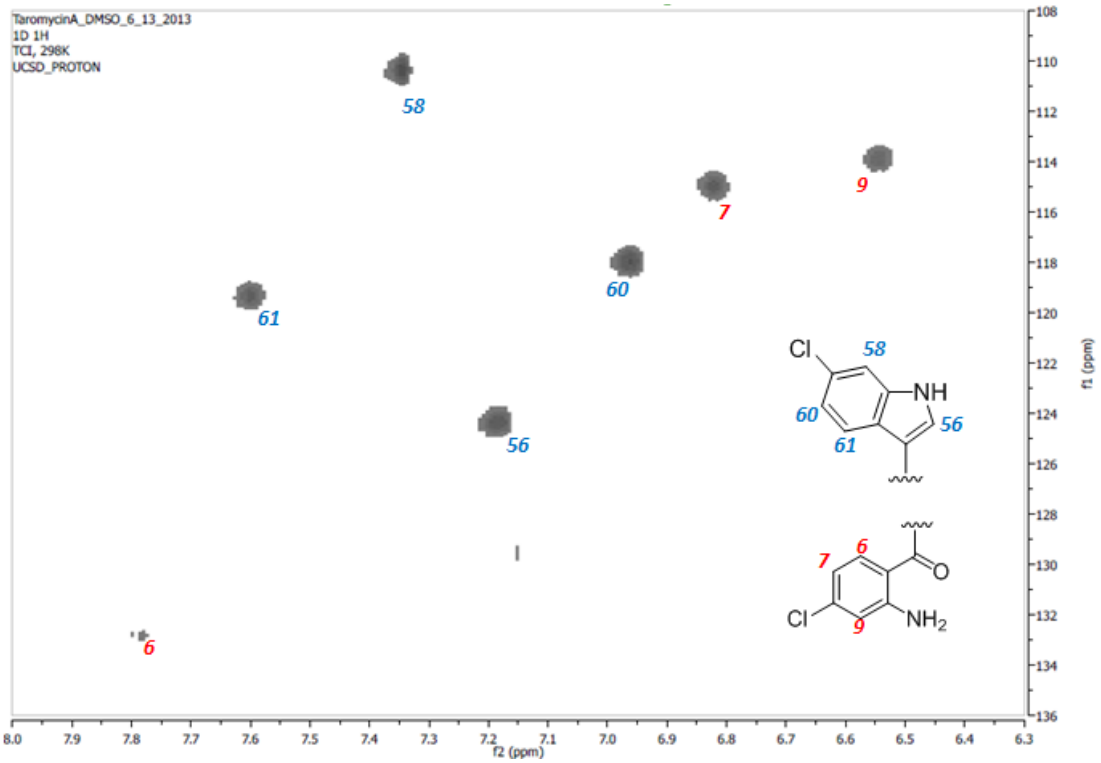

Figure S9
HSQC spectra of taromycin A aromatic region (d_6-DMSO, 600 MHz)

Table S5. Retention times (in min) of taromycin A and daptomycin constituent amino acids derivatized with D-FDAA, as well as standard D-FDAA amino acids.*

*Conditions: Phenomenex Luna 5 µm C_{18} column (4.6 x 250 mm), 1.0% MeCN/H_2O + 0.1% formic acid (FA) for 5 min followed by a gradient to 40% MeCN/H_2O + 0.1% FA over 70 min, followed by another gradient to 65% MeCN/H_2O + 0.1% FA over 15 min and held for 5 min at a flow of 0.4 mL/min, monitoring from 200 to 600 nm. ** D-FDAA-D-Orn retention time was deduced by reacting L-FDAA with L-Orn. ***Asparagine is converted to aspartic acid during hydrolysis giving 1:3 mixture of D:L-Asp. #It is well known that L configured amino acids retain longer than D when derivatized with D-FDAA, and we confirmed this trend with unchlorinated Trp and Kyn residues.[13] This observation led us to assign the earlier retention time as the D stereoisomer and the latter as the L stereoisomer for the chlorinated Trp and Kyn residues.

Amino Acid Residue	D-FDAA-L-Amino Acid	D-FDAA-D-Amino Acid	D-FDAA-Daptomycin	D-FDAA-Taromycin A
Ala	82.36	77.15	77.04	77.31
Orn**	90.76	93.81*	90.85	91.06
Thr	74.92	68.82	74.78	74.73
allo-Thr	71.83	68.91	--	--
β-methyl-Glu	--	--	78.46	78.46
Asp	72.20	70.43	72.14	72.41
Asn***	--	--	69.50	70.44
4-Cl-Kyn#	83.57	78.98	--	83.73
6-Cl-Trp#	98.73	96.73	--	99.19

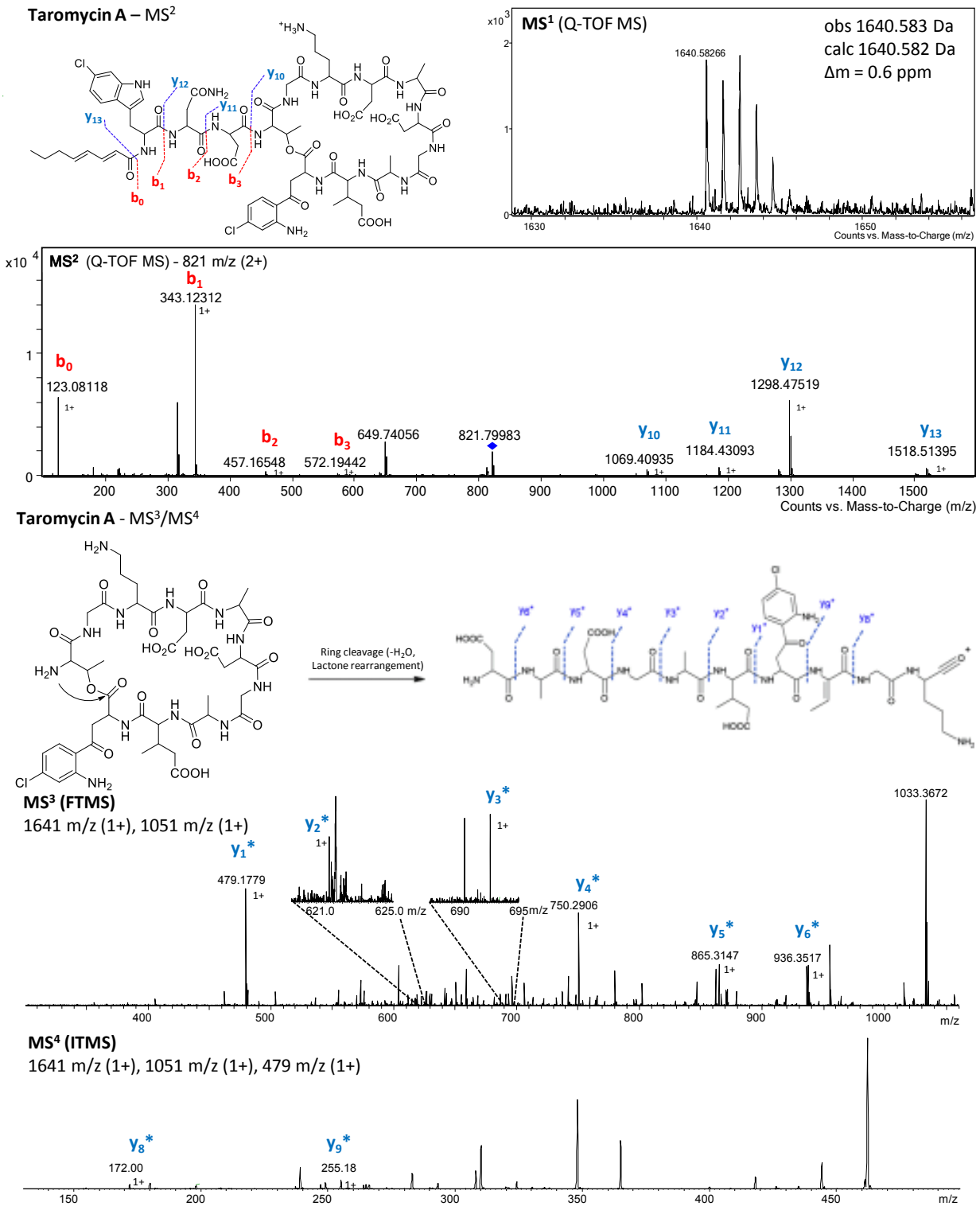

Figure S10

MS and MS^n analysis of taromycin A. MS^2 analysis was performed by ESI-Q-TOF MS, MS^3 analysis was performed by FT-ICR MS, MS^4 analysis was performed by LTQ(IT) MS. * - indicates y-ion series-like fragments based on taromycin sequence with a putative b-ion like C-terminus.

Instrument	Species	Precursor [m/z]	Observed mass [Da]	Calculated mass [Da]	Difference [Da]	Δm [ppm]
Q-TOF-MS	y13	821 (2+)	1518.514	1518.509	0.005	3
Q-TOF-MS	y12	821 (2+)	1298.475	1298.469	0.006	5
Q-TOF-MS	y11	821 (2+)	1184.431	1184.426	0.005	4
Q-TOF-MS	y10	821 (2+)	1069.409	1069.399	0.01	9
Q-TOF-MS	y10(-H_2O)	821 (2+)	1051.389	1051.388	0.001	1
Q-TOF-MS	b3	821 (2+)	572.194	572.191	0.003	5
Q-TOF-MS	b2	821 (2+)	457.166	457.164	0.002	4
Q-TOF-MS	b1	821 (2+)	343.121	343.123	-0.002	6
Q-TOF-MS	b0	821 (2+)	123.081	123.08	0.001	8
FTMS	y6(-H_2O)*	1641 (1+), 1051 (1+)	936.352	936.361	-0.009	10
FTMS	y5(-H_2O)*	1641 (1+), 1051 (1+)	865.315	865.324	-0.009	10
FTMS	y4(-H_2O)*	1641 (1+), 1051 (1+)	750.291	750.297	-0.006	8
FTMS	y3(-H_2O)*	1641 (1+), 1051 (1+)	693.27	693.276	-0.006	9
FTMS	y2(-H_2O)*	1641 (1+), 1051 (1+)	622.234	622.239	-0.005	8
FTMS	y1(-H_2O)*	1641 (1+), 1051 (1+)	479.178	479.18	-0.002	4
ITMS	y9(-H_2O)*	1641 (1+), 1051 (1+), 479 (1+)	255.18	255.145	0.035	
ITMS	y8(-H_2O)*	1641 (1+), 1051 (1+), 479 (1+)	172.108	172	0.108	

Figure S11

MS and MS^n analysis of taromycin A (continued). MS^2 analysis was performed by ESI-Q-TOF MS, MS^3 analysis was performed by FT-ICR MS, MS^4 analysis was performed by LTQ(IT) MS. * - indicates y-ion series-like fragments based on taromycin sequence with a putative b-ion like C-terminus.

9 Bioactivity of taromycin A compared to daptomycin

The taromycin A stock was prepared by the addition of 100 µl of DMSO to 1000 µg pure material. In a separate microcentrifuge tube, 10 µl of taromycin A stock at 10 mg/ml was added to 490 µl cation-adjusted Mueller Hinton broth (CA-MHB) to prepare the working taromycin A stock. Bacterial strains tested were prepared by removing log-phase cultures from the incubator shaking at 37°C and transferring to a sterile 15 ml conical tube and then centrifuged at 4000 rpm for 10 min. The supernatant was removed and the resulting bacterial pellet was resuspended in 600 µl sterile 1x PBS. The bacterial suspensions in 1x PBS were then diluted to contain $5x10^5$ cfu/ml and added to the test plate. Next, taromycin A at 100 µg/ml was added to the test plate followed by serial dilutions. After all components were added, the test plates were sealed and the bacteria were allowed to propagate at 37°C for 16-18 h. Thereafter, the OD600 for each well was acquired and recorded in a plate reader. The minimal inhibitory concentration for each sample tested was defined as OD600 reading less the 0.06, which was comparable to the observed bacteria alone absorbance. A DMSO, bacteria alone and media blank control were also tested within the same plate. In addition, each plate contained duplicate samples and was run in triplicate to test reproducibility (Table S6).

Table S6. Taromycin A and daptomycin minimum inhibitory concentration (50 µg/ml $CaCl_2$)*

Strain	Details	Taromycin A MIC	Daptomycin MIC
MRSA 0325	Dap S	12	>2
E. faecalis 613	Dap S	6	2
E. faecalis 613D	Dap R	50	>2
E. faecalis 447	Dap S	6	2
E. faecium 447D	Dap R	50	>2
VRE 5938	Dap R	100	>2

*Dap S: daptomycin susceptible; Dap R: daptomycin resistant. Concentrations are reported in µg/mL. OD600<0.06 considered minimum inhibitory concentration.

10 Verified functionality of pCAP01 by direct cloning and heterologous expression of the marinopyrrole biosynthesis (*mpy)* gene cluster

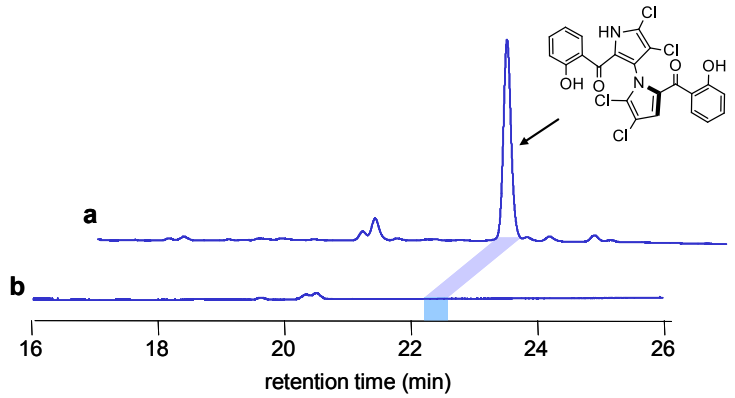

Figure S12
HPLC analysis of the marinopyrrole A produced by *S. coelicolor* mutants. Extracts from (a) *S. coelicolor* M512 / pCAP01-*mpy* and (b) *S. coelicolor* M512 / pCAP01 (empty vector) were analyzed by C18 reversed phase HPLC. UV absorption was monitored at 210 nm.

11 Reference list for supporting information

1. Gomez-Escribano JP, Bibb MJ (2011) Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol* **4**:207-215.
2. Larionov V, Kouprina N, Solomon G, Barrett JC, Resnick MA (1997) Direct isolation of human BRCA2 gene by transformation-associated recombination in yeast. *Proc Natl Acad Sci U S A* **94**:7384-7387.
3. Kouprina N, Annab L, Graves J, Afshari C, Barrett JC, Resnick MA, Larionov V (1998) Functional copies of a human gene can be directly isolated by transformation-associated recombination cloning with a small 3' end target sequence. *Proc Natl Acad Sci U S A* **95**:4469-4474.
4. Kouprina N, Larionov V (2008) Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast *Saccharomyces cerevisiae*. *Nat Protoc* **3**:371-377.
5. Oldenburg KR, Vo KT, Michaelis S, Paddon C (1997) Recombination-mediated PCR-directed plasmid construction *in vivo* in yeast. *Nucleic Acids Res* **25**:451-452.
6. 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.
7. Flett F, Mersinias V, Smith CP (1997) High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting *Streptomycetes*. *FEMS Microbiol Lett* **155**:223-229.
8. Yamanaka K, Ryan KS, Gulder TA, Hughes CC, Moore BS (2012) Flavoenzyme-catalyzed atropo-selective n,c-bipyrrole homocoupling in marinopyrrole biosynthesis. *J Am Chem Soc* **134**:12434-12437.
9. Hughes CC, Prieto-Davo A, Jensen PR, Fenical W (2008) The marinopyrroles, antibiotics of an unprecedented structure class from a marine *Streptomyces* sp. *Org Lett* **10**:629-631.
10. Schafer A, Tauch A, Jager W, Kalinowski J, Thierbach G, Puhler A (1994) Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* **145**:69-73.
11. Floriano B, Bibb M (1996) afsR is a pleiotropic but conditionally required regulatory gene for antibiotic production in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **21**:385-396.
12. Muramoto K, Kamiya, H (1990) Recovery of tryptophan in peptides and proteins by high-temperature

and short-term acid hydrolysis in the presence of phenol. *Anal Biochem* **189**: 223-230.
13. Bhushan R, Bruckner H (2011) Use of Marfey's reagent and analogs for chiral amino acid analysis: Assessment and applications to natural products and biological systems. *J Chromatogr B* **879**: 3148-3161.