Electronic Supplementary Material (ESI) for Medicinal Chemistry Communications

SUPPORTING INFORMATION

The muraminomicin biosynthetic gene cluster and enzymatic formation of the 2-deoxyaminoribosyl appendage

Xiuling Chi,[a] Satoshi Baba,[b] Nidhi Tibrewal,[a] Masanori Funabashi,[c] Koichi Nonaka,[d] and Steven G. Van Lanen*[a]

[a] Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, Lexington, KY 40536, United States

[b] Biologics Research Laboratories, R&D Division, Daiichi Sankyo Co. Ltd., Tokyo 140-8710, Japan

[c] Natural Products Research Group, Discovery Science and Technology Department, Drug Discovery and Biomedical Technology Unit, Daiichi Sankyo RD Novare Co., Ltd., Tokyo 134-8630, Japan

[d] Biopharmaceutical Research Group I, Biopharmaceutical Technology Research Laboratories, Pharmaceutical Technology Division, Daiichi Sankyo Co., Ltd., Gunma 370-0503, Japan

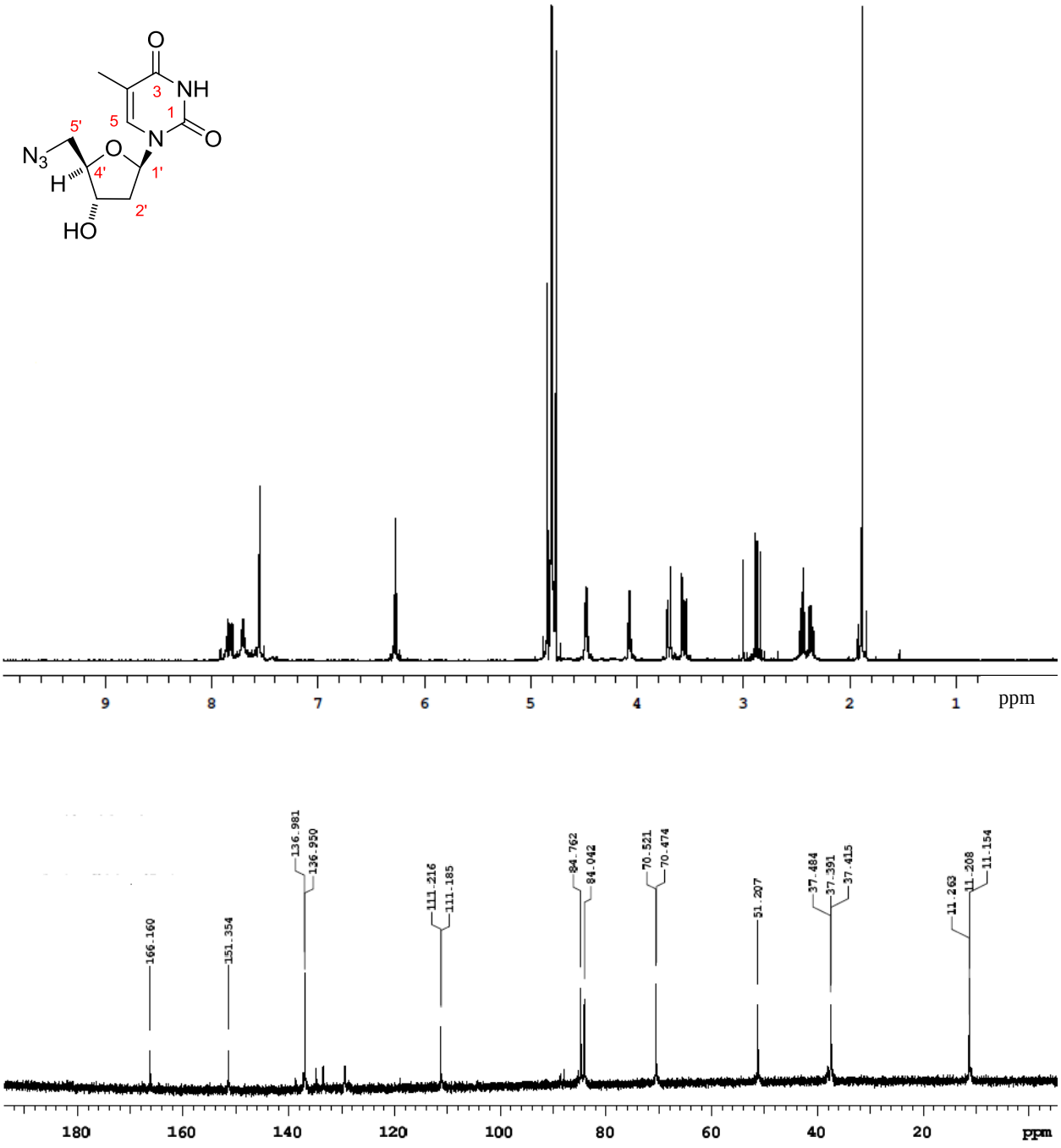
*svanlanen@uky.edu

EXPERIMENTAL PROCEDURES

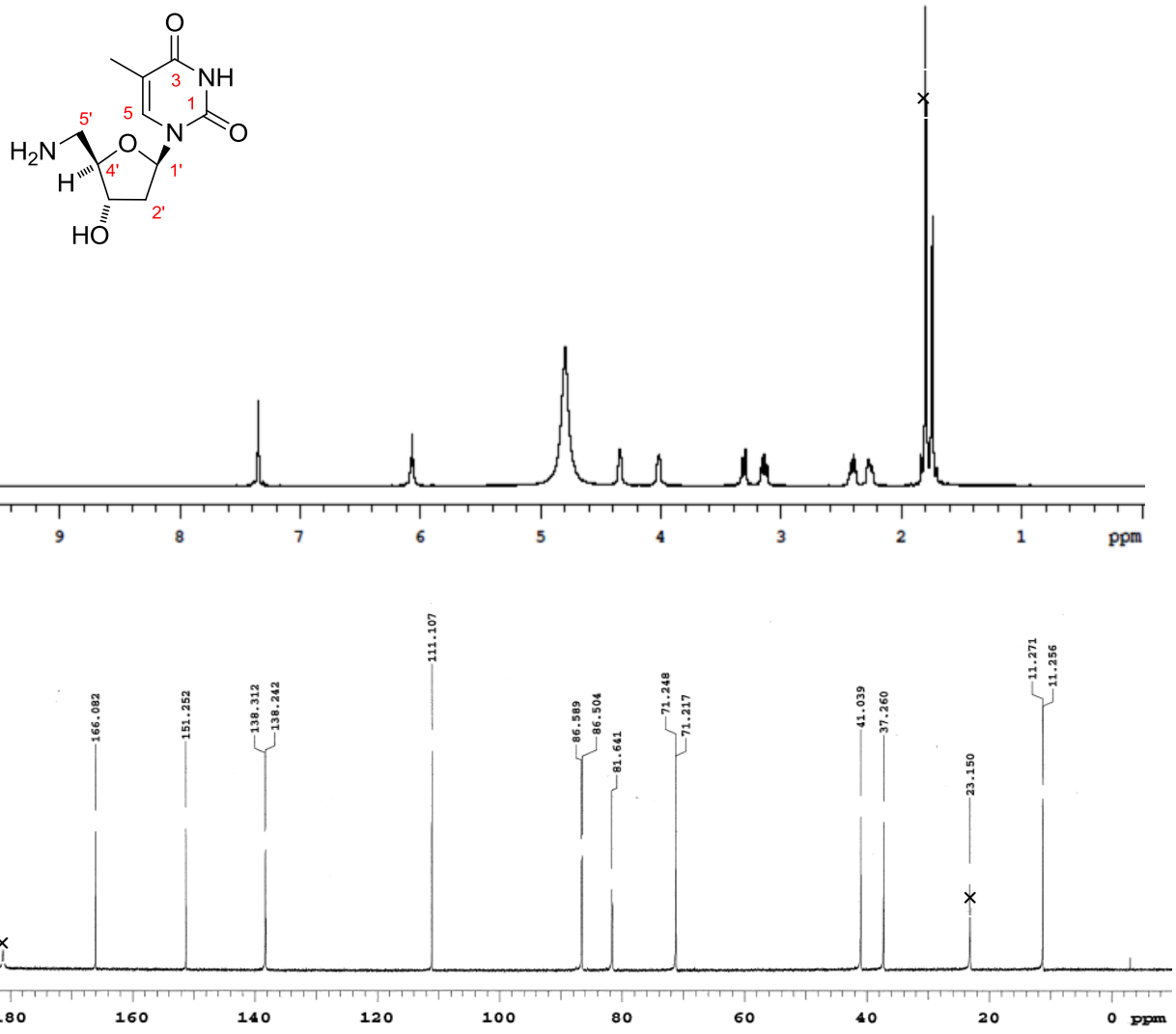
Chemicals and instrumentation. Nucleoside bases, nucleosides, and nucleotides were purchased from Sigma or Promega. Buffers, salts, and media components were purchased from Fisher Scientific. UV/Vis spectroscopy was performed with a Bio-Tek μQuant microplate reader using Microtest™ 96-well plates (BD Biosciences) or a Shimadzu UV/Vis-1800 Spectrophotometer. Synthetic oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). DNA sequencing was performed using the BigDye™Terminator version 3.1 Cycle Sequencing kit from Applied Biosystems, Inc. (Foster City, CA) and analyzed at the University of Kentucky Advanced Genetic Technologies Center. HPLC was performed with a Waters Alliance 2695 separation module (Milford, MA) equipped with a Waters 2998 diode array detector and an analytical Apollo C-18 column (250 mm x 4.6 mm, 5 μm) or a semi-preparative Apollo C-18 column (250 mm x 10 mm, 5 μm) purchased from Grace (Deerfield, IL). Electrospray ionization-MS was performed using an Agilent 6120 Quadrupole MSD mass spectrometer (Agilent Technologies, Santa Clara, CA) equipped with an Agilent 1200 Series Quaternary LC system and an Eclipse XDB-C18 column (150mm x 4.6 mm, 5 μm, 80Å). NMR data were collected using a Varian Unity Inova 300, 400 or 500 MHz Spectrometer (Varian, Inc., Palo Alto, CA).

Synthesis of substrates. The synthesis of 5'-amino-5'-deoxyuridine and 5'-amino-2',5'-dideoxyuridine **2** was previously reported.[13] The identical two-step procedure for the latter was used to synthesize 5'-amino-5'-deoxythymidine **4** (see below).

Step 1: 5'-Azido-5'-deoxythymidine. ^{1}H-NMR (500MHz, D_2O): δ 7.55 (s, 1H), 6.27 (t, 1H, J = 7 Hz), 4.51-4.45 (m, 1H), 4.08-4.04 (m, 1H), 3.72-3.51 (m, 2H), 2.50-2.31 (m, 2H), 1.89-1.82 (s, 3H); ^{13}C-NMR (100MHz, D_2O): δ 166.2 (C), 151.4 (C), 136.97 (CH), 111.2 (CH), 84.8 (CH), 84.1 (CH), 70.5 (CH), 51.2 (CH_2), 37.4 (CH_2), 11.2 (CH_3).



Step 2: 5'-Amino-5'-deoxythymidine, ***2****.* ^{1}H-NMR (500 MHz, D_2O): *δ* 7.36 (s, 1H), 6.07 (t, 1H, *J* = 7 Hz), 4.38-4.30 (m, 1H), 4.08-3.95 (m, 1H), 3.35-3.08 (m, 2H), 2.45-2.22 (m, 2H), 1.78-1.70 (s, 3H); ^{13}C-NMR (125 MHz, D_2O): *δ* 166.1 (C), 151.3 (C), 138.3 (CH), 111.2 (CH), 86.5 (CH), 81.6 (CH), 71.2 (CH), 41.0 (CH_2), 37.3 (CH_2), 11.3 (CH_3). (×: acetate)



Cloning of the muraminomicin gene cluster. *Streptosporangium amethystogenes* sp. SANK 60709 genomic DNA was partially digested with *Sau*3AI to give ~40-kb DNA fragments that were dephosphorylated with bacterial alkaline phosphatase (BAP) and ligated into *Bam*HI-digested cosmid vector SuperCos1-P (Stratagene, Cedar Creek, USA) that contains the loxP site in SuperCos1 and was dephosphorylated by BAP after *Xba*I digestion. The ligation products were packaged with Gigapack III Gold packaging extract as described by the manufacturer (Stratagene), and the resulting recombinant phage was used to transfect *E. coli* XL-1 Blue MR. Approximately 10,000 colonies from the obtained genomic library were screened by colony hybridization using digoxigenin (DIG)-labeled DNA obtained by PCR using degenerate primers as previously described.[15] Hybridization was carried out using DIG easy hyb (Roche, Indianapolis, USA) at 42 °C and the resulting filter was washed under high stringency conditions (0.1 X SSC including 0.1 % SDS, 68 °C). Detection was performed using CDP-Star (Roche) according to the manufacturer's procedures.

Based on restriction digest analysis, four positive cosmids, pMra01-04, were isolated and sequenced using a Roche GS FLX system (Operon Biotechnologies). Potential open reading frames were defined using Frameplot 4.0, and database comparison for sequence homology was performed with BLAST search tools using the National Center for Biotechnology Information (Bethesda, MD). The sequence has been deposited at GenBank under accession no. AB746937.

Cloning of genes for heterologous expression. Genes were amplified by PCR using Expand Long Template PCR System from Roche (Indianapolis, IN) with supplied Buffer 2, 200 μM dNTPs, 5% DMSO, 10 ng DNA pMra02, 5 U DNA polymerase, and 200 nM each of the following primer pairs: *mra20* (forward) 5'- GGTATTGAGGGTCGCATGAACGAGAT -3' / (reverse) 5'- AGAGGAGAGTTAGAGCCTCACCCGG -3' and *mra23* (forward) 5'- GGTATTGAGGGTCGCATGTCGGTCG -3' / (reverse) 5'- AGAGGAGAGTTAGAG CCTCAGCCGA -3'. The PCR program included an initial hold at 94 °C for 2 min, followed by 30 cycles of 94 °C for 10 s, 56 °C for 15 s, and 68 °C for 60 s. The gel-purified PCR product was inserted into pET-30 Xa/LIC using ligation-independent cloning as described in the provided protocol to yield pET30-*mra20* and pET30-*mra23*. The genes were sequenced to confirm PCR fidelity.

Plasmids were introduced into *E. coli* BL 21(DE3) cells, and the transformed strains were grown in LB supplemented with 50 μg/mL kanamycin. Following inoculation of 500 mL of LB with 50 μg/mL kanamycin, the cultures were grown at 18 °C until the cell density reached an OD_{600} ~ 0.5 when expression was induced with 0.1 mM IPTG. Cells were harvested after an overnight incubation at 18 °C and lysed using a French Press with one pass at 15000 psi. Following centrifugation the protein was purified using affinity chromatography with Ni-NTA agarose from Qiagen (Valencia, CA), and the recombinant proteins were desalted into 50 mM Tris-HCl (pH 8), 100 mM NaCl, and 5 % glycerol using a PD-10 desalting column (GE Healthcare). The purified protein was concentrated using an Amicon Ultra 10000 MWCO centrifugal filter (Millipore) and stored as glycerol stocks (40%) at -20 °C. Protein purity was assessed as by 12% acrylamide SDS-PAGE; His_6-tagged proteins were utilized without further modifications.

In vitro characterization of Mra20. Reactions consisted of 25 mM potassium phosphate pH 7.5, 2 mM **2** or analogue, and 100 nM Mra20 at 30 °C, and terminated by the addition of cold TCA to 5% (w/v) or by ultrafiltration using a Microcon YM-3. Following centrifugation to remove protein, the reaction components were analyzed by HPLC using a C-18 reverse-phase column. A series of linear gradients was developed from 40 mM phosphoric acid-triethylamine pH 6.5 (A) to 20% methanol (B) in the following manner (beginning time and ending time with linear

increase to % B): 0-8 min, 0% B; 8-18 min, 60% B; 18-25 min, 95% B; 25-32 min, 95% B; and 32-35 min, 0% B. The flow rate was kept constant at 1.0 mL/min, and elution was monitored at 260 nm.

To determine the kinetic constants with respect to the co-substrate nucleoside, reactions were carried out in 50 mM Tris-HCl pH 9.0 consisting of saturating phosphate (1.5 mM) and variable nucleoside (25 – 5,000 μM), and 100 nM Mra20 at 30 °C under initial velocity conditions. The reactions were terminated with 0.1 M sodium hydroxide, and nucleobase formation was determined by UV/Vis spectroscopy with $\Delta\varepsilon_{300\ nm}$ = 3, 700 M^{-1} at pH 13 for **6** and $\Delta\varepsilon_{290\ nm}$ = 5, 700 M^{-1} at pH 13 for **5**.

In vitro characterization of Mra23. Reactions consisted of 50 mM potassium phosphate pH 7.5, 5 mM $MgCl_2$, 2 mM **3** or analogue, 5 mM uridine 5'-triphosphate (UTP), 5 μM Mra20, and 1 μM Mra23 at 30 °C, and the reaction terminated by the addition of cold TCA to 5% (w/v) or by ultrafiltration using a Microcon YM-3. The activity of ORF39 was tested with sugar-1-phosphates generated in situ from synthetic **2**, **3**, **4** or uridine with co-substrate UTP. Following centrifugation to remove protein, the reaction components were analyzed by HPLC using a phosphoric acid-triethylamine mobile phase as described above. LC-MS was performed using a linear gradient from 0.1% formic acid in water to 0.1% formic acid in acetonitrile over 20 min. The flow rate was kept constant at 0.4 mL/min, and elution was monitored at 254 nm.

SUPPORTING TABLES

Table S1. Annotation of ORFs within the muraminomicin gene cluster.

Protein[a]	*Size*[b]	*Proposed function*	*Sequence similarity (protein, accession no., origin)*	*Identity / similarity*[c]	*A-90289*[d]	*Identity %*
ORF-16	258	RNA polymerase, sigma 28 subunit, SigD/FliA/WhiG	SrosDRAFT_47470 (ZP_04474166) *Streptosporangium roseum* DSM 43021	93/97		
ORF-15	123	anti-anti-sigma factor	SrosDRAFT_47460 (ZP_04474165) *Streptosporangium roseum* DSM 43021	89/94		
ORF-14	352	hypothetical protein	SghaA1_010100033943 (ZP_04690229) *Streptomyces ghanaensis* ATCC 14672	62/72		
ORF-13	238	response regulator	SrosDRAFT_47450 (ZP_04474164) *Streptosporangium roseum* DSM 43021	89/94		
ORF-12	563	PAS domain S-box protein	SrosDRAFT_47440 (ZP_04474163) *Streptosporangium roseum* DSM 43021	89/95		
ORF-11	393	glycosyltransferase	SrosDRAFT_47420 (ZP_04474161) *Streptosporangium roseum* DSM 43021	83/92		
ORF-10	329	hypothetical protein	SrosDRAFT_47400 (ZP_04474159) *Streptosporangium roseum* DSM 43021	74/82		
ORF-9	243	hypothetical protein	FRAAL6120 (YP_716259) *Frankia alni* ACN14a	49/65		
ORF-8	361	sugar diacid utilization regulator	SrosDRAFT_47390 (ZP_04474158) *Streptosporangium roseum* DSM 43021	79/84		
ORF-7	308	L-proline dehydrogenase	SrosDRAFT_47380 (ZP_04474157) *Streptosporangium roseum* DSM 43021	83/89		
ORF-6	560	delta-1-pyrroline-5-carboxylate dehydrogenase	SrosDRAFT_47370 (ZP_04474156) *Streptosporangium roseum* DSM 43021	92/95		
ORF-5	183	hypothetical protein	SrosDRAFT_40310 (ZP_04473452) *Streptosporangium roseum* DSM 43021	84/90		
ORF-4	150	DNA-binding protein	SrosDRAFT_40320 (ZP_04473453) *Streptosporangium roseum* DSM 43021	70/79		
ORF-3	293	hypothetical protein	SGR_1524 (YP_001823036) *Streptomyces griseus* NBRC 13350	49/62		
ORF-2	301	NmrA family protein	Franean1_1468 (YP_001505814) *Frankia* sp. EAN1pec	65/76		
ORF-1	199	tetR-family transcriptional regulator	Franean1_1469 (YP_001505815) *Frankia* sp. EAN1pec	72/80		
Mra1	394	rhamnosyltransferase	Cpz31 (ACQ63639) *Streptomyces* sp. MK730-62F2	78/88	LipB1	74
Mra2	395	sugar *O*-methyltransferase	Cpz30 (ACQ63638) *Streptomyces* sp. MK730-62F2	78/82	LipA1	74
Mra3	521	radical SAM domain-containing protein	CalU22 (AAM94801) Micromonospora echinospora	67/81		
Mra4	382	predicted membrane protein	NdasDRAFT_4699 (ZP_04335569) *N. dassonvillei* sub. dassonvillei DSM	55/67		
Mra5	265	sugar *O*-methyltransferase	Cpz29 (ACQ63637) *Streptomyces* sp. MK730-62F2	81/89	LipZ	83
Mra6	214	TmrB-like protein	Cpz27 (ACQ63635) *Streptomyces* sp. MK730-62F2	71/80	LipX	67
Mra7	399	SAM-dependent methyltransferase	Cpz26 (ACQ63634) *Streptomyces* sp. MK730-62F2	74/84	LipW	72
Mra8	322	alcohol dehydrogenase	Cpz25 (ACQ63633) *Streptomyces* sp. MK730-62F2	88/93	LipV	86
Mra9	600	hypothetical protein	Cpz24 (ACQ63632) *Streptomyces* sp. MK730-62F2	69/78	LipU	68
Mra10	344	lipase	Cpz23 (ACQ63631) *Streptomyces* sp. MK730-62F2	78/88	LipT	79
Mra11	1226	ABC multidrug resistance transporter	Cpz22 (ACQ63630) *Streptomyces* sp. MK730-62F2	77/84	LipS	76
Mra12	490	acyltransferase	Cpz21 (ACQ63629) *Streptomyces* sp. MK730-62F2	75/85	LipR	76

Table S1 cont.

Mra13	354	acyl-CoA synthase	Cpz20 (ACQ63628) *Streptomyces* sp. MK730-62F2	87/90	LipQ	84
Mra14	422	SHMT-like (L-Thr:uridine-5'-aldehyde transaldolase)	Cpz14 (ACQ63622) *Streptomyces* sp. MK730-62F2	81/87	LipK	83
Mra15	441	aminotransferase	Cpz13 (ACQ63621) *Streptomyces* sp. MK730-62F2	77/86	LipJ	76
Mra16	189	TmrB-like protein	Cpz12 (ACQ63620) *Streptomyces* sp. MK730-62F2	72/81	LipI	71
Mra17	213	SAM-dependent methyltransferase	Cpz11 (ACQ63619) *Streptomyces* sp. MK730-62F2	75/86	LipH	78
Mra18	173	beta-hydroxylase	Cpz10 (ACQ63618) *Streptomyces* sp. MK730-62F2	84/91	LipG	80
Mra19	329	AraC family transcriptional regulator	Cpz9 (ACQ63617) *Streptomyces* sp. MK730-62F2	65/72	LipF	61
Mra20	455	pyrimidine-nucleoside phosphorylase	DealDRAFT_0710 (ZP_03728855) *Dethiobacter alkaliphilus* AHT 1	46/63	LipP	42
Mra21	433	aminotransferase	CetH (ACH85568) *Actinomyces* sp. Lu 9419	52/64	LipO	41
Mra22	367	glycosyltransferase	Orf13 (BAI23321) *Streptomyces griseus*	39/57	LipM	36
Mra23	232	nucleotidylyltransferase	Amir_3895 (YP_003101615) *Actinosynnema mirum* DSM 43827	40/51	LipM	33
Mra24	217	dioxygenase	Orf7 (BAI23315) *Streptomyces griseus*	41/56	LipL	40
ORF1	328	DNA polymerase beta domain protein region	Mpop_3269 (YP_001925955) *Methylobacterium populi* BJ001	29/43		
ORF2	615	glucose-methanol-choline oxidoreductase	SACE_4301 (YP_001106495) *Saccharopolyspora erythraea* NRRL	46/58		
ORF3	161	hypothetical protein	N9414_23023 (ZP_01630686) *Nodularia spumigena* CCY9414	33/48		
ORF4	192	hypothetical protein	SrosDRAFT_27630 (ZP_04472191) *Streptosporangium roseum* DSM 43021	74/83		
ORF5	501	hydrolase	Svir_18520 (YP_003133702) *Saccharomonospora viridis* DSM 43017	62/72		
ORF6	181	NADPH-dependent FMN reductase	Pat9bDRAFT_2775 (ZP_05729459) *Pantoea* sp. At-9b	43/65		
ORF7	3307	amino acid adenylation domain-containing protein	Sare_4562 (YP_001539321) *Salinispora arenicola* CNS-205	49/62		
ORF8	396	acyl-CoA dehydrogenase	TcurDRAFT_15750 (ZP_04030556) *Thermomonospora curvata* DSM 43183	73/84		
ORF9	286	hypothetical protein	hyg24 (ABC42561) *Streptomyces hygroscopicus*	43/56		
ORF10	300	Short-chain dehydrogenase/reductase	Bxe_B2739 (YP_552606) *Burkholderia xenovorans* LB400	57/71		
ORF11	384	hypothetical protein	Plav_2009 (YP_001413280) *Parvibaculum lavamentivorans* DS-1	63/78		
ORF12	140	hypothetical protein	Veis_2934 (YP_997688) *Verminephrobacter eiseniae* EF01-2	57/73		
ORF13	386	acyl-CoA dehydrogenase	TcurDRAFT_39460 (ZP_04032895) *Thermomonospora curvata* DSM 43183	50/65		

[a] Sequences are deposited at NCBI (accession no. AB746937).
[b] Numbers are in amino acids.
[c] % sequence identity and similarity for the entire length of the proteins.
[d] Sequences are deposited at NCBI (accession no. AB530986).

SUPPORTING FIGURES

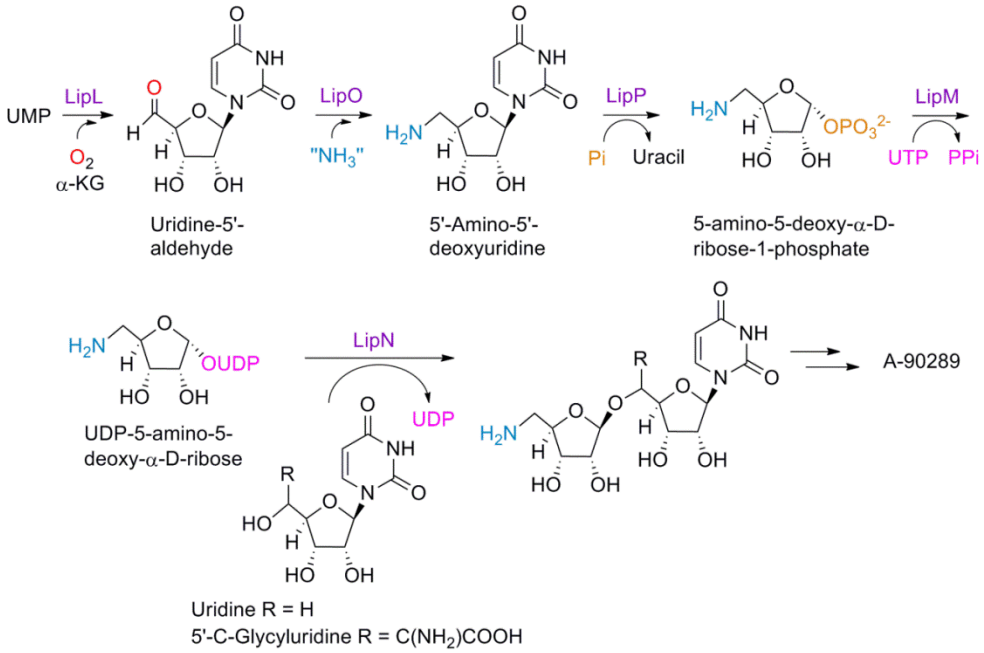


Figure S1. Biosynthetic pathway leading to the assembly of the β(1→5) disaccharide core of the lipopeptidyl nucleoside antibiotics.

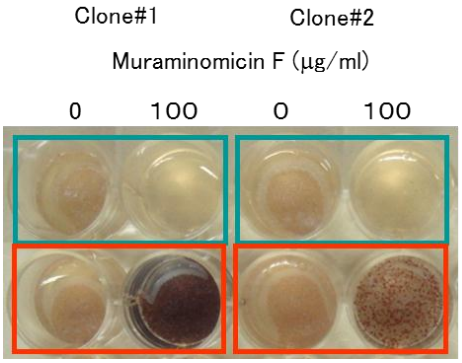


Figure S2. Resistance conferred by pMra02 upon heterologous expression in *Streptomyces lividans* TK21.

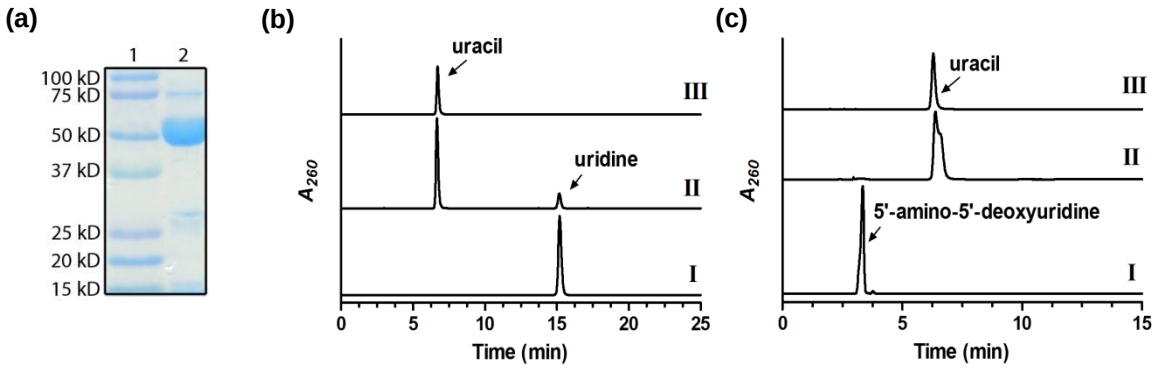


Figure S3. *In vitro* characterization of Mra20 (a) SDS-PAGE analysis of partially purified His_6-Mra20 (expected MW of 52.5 kD). (b) HPLC analysis using uridine after 1 h without Mra20 (I), 1 h reaction (II) and authentic uracil (III). (c) HPLC analysis using 5'-amino-5'-deoxyuridine after 1 h without Mra20 (I), 1 h reaction (II) and authentic uracil (III). $A_{260,}$ absorbance at 260 nm.

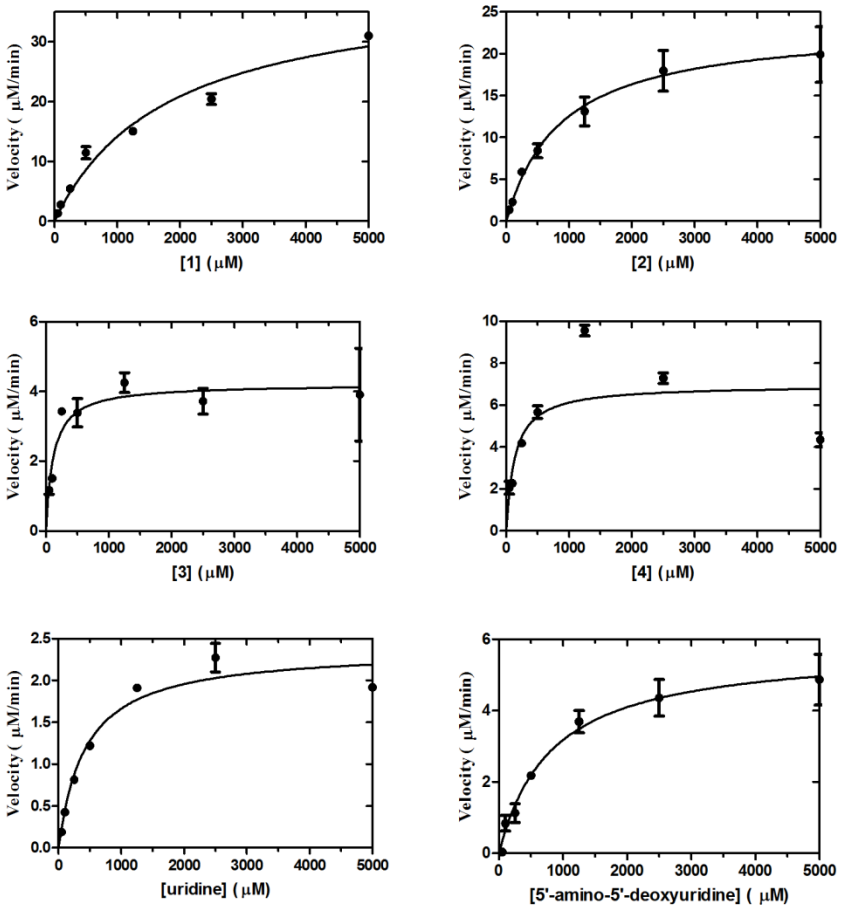


Figure S4. Single-substrate kinetic analysis of Mra20**.** Data were fitted to the Michaelis-Menten equation and extracted kinetic constants are given in Table 1.

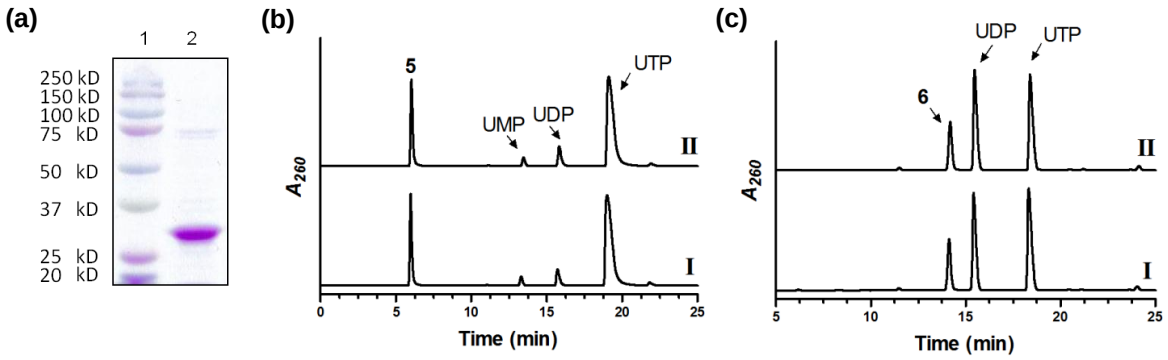


Figure S5. *In vitro* characterization of Mra23. (a) SDS-PAGE analysis of purified His_6-Mra23 (expected MW of 29.6 kD). (b) HPLC analysis of the reaction starting with uridine after 3 h without Mra23 (I) and 3 h reaction (II). (c) HPLC analysis of the reaction starting with **3** after 3 h without Mra23 (I) and 3 h reaction (II). $A_{260,}$ absorbance at 260 nm.

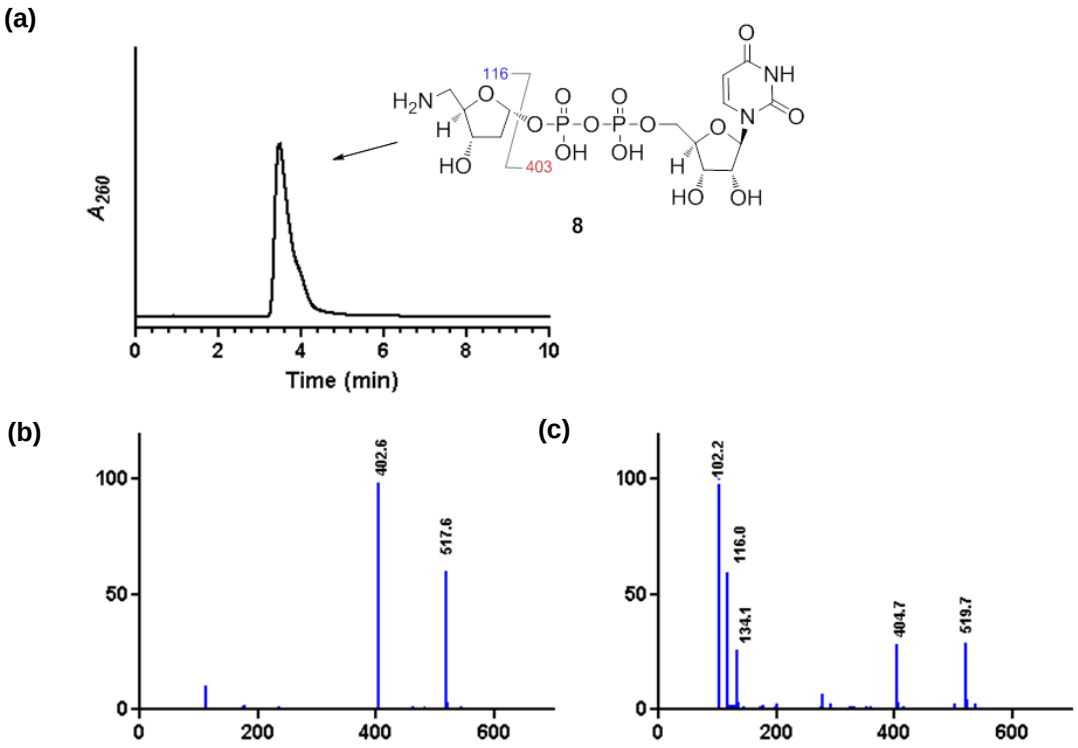


Figure S6. LC-MS of the Mra23 product (**7**) generated from substrate **4**. (a) LC analysis of **7** following purification. (b) Negative ion mass spectrum for the peak at elution time t = 3.5 min. (c) Positive ion mass spectrum for the peak at elution time t = 3.5min. $A_{260,}$ absorbance at 260 nm.