Supplementary Information

Discovery of MRSA active antibiotics using primary sequence from the human microbiome

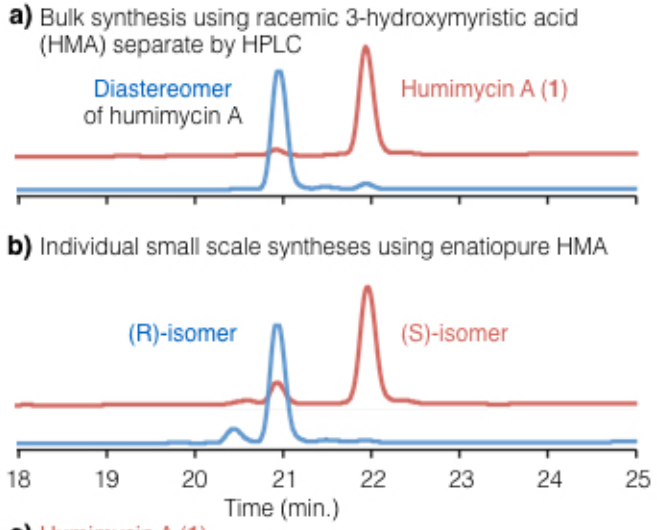
John Chu,[1‡] Xavier Vila-Farres,[1‡] Daigo Inoyama,[3] Melinda Ternei,[1] Louis J. Cohen,[1] Emma A. Gordon,[1] Boojala Vijay B. Reddy,[1] Zachary Charlop-Powers,[1] Henry A. Zebroski,[2] Ricardo Gallardo-Macias,[3] Mark Jaskowski,[3] Shruthi Satish,[3] Steven Park,[4] David S. Perlin,[4] Joel S. Freundlich,[3] and Sean F. Brady[1]*

[1]Laboratory of Genetically Encoded Small Molecules and [2]Proteomics Resource Center, The Rockefeller University, New York, New York, USA. [3]Department of Pharmacology, Physiology, and Neuroscience, Rutgers University – New Jersey Medical School, Newark, New Jersey, USA. [4]Public Health Research Institute, Rutgers University – New Jersey Medical School, Newark, New Jersey, USA.

‡These authors contributed equally to this work. *Corresponding author e-mail: sbrady@rockefeller.edu.

Supplementary Results

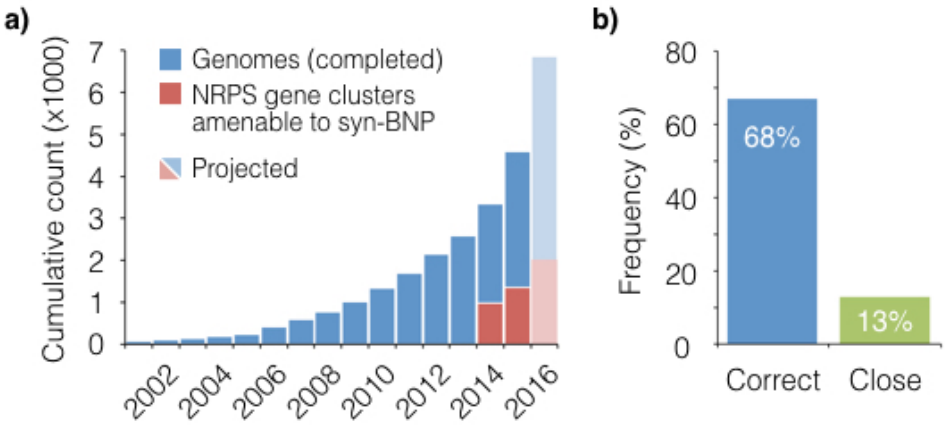
Supplementary Figure 1. Stereochemistry assignment of humimycin A. a) Humimycin A used in this study was synthesized in bulk using a racemic mixture of 3-hydroxymyristic acid (HMA), and the resulting pair of diastereomers were separated by HPLC. **b)** The absolute stereochemistry of humimycin A was assigned by comparison to small scale re-syntheses of humimycins using enantiopure (R)- and (S)-HMA. Humimycin A (**1**) refers to the peptide *N*-acylated with (S)-HMA and shows more potent antibiotic activity (**Supplementary Table 3**). **c)** Structures of humimycin A and its diastereomer. **d)** Solvent gradient for HPLC runs. Characterization and purification were carried out on an Agilent instrument at 4 mL/min using an XBridge BEH C18 column (130 Å, 5 um, 4.6 x 150 mm, Waters Corporation). Solvent A and B are water and acetonitrile with 0.1% TFA, respectively.



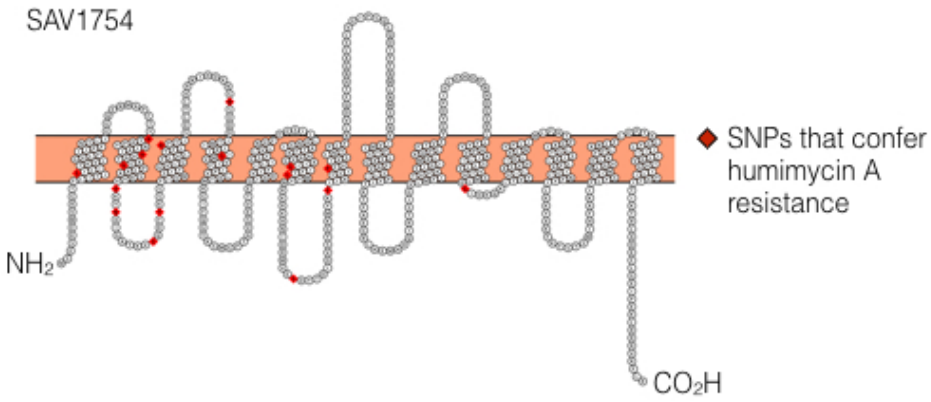
d) HPLC gradient

Time (min.)	0	1.5	5	35	38
Solvent B (%)	5	5	35	95	95

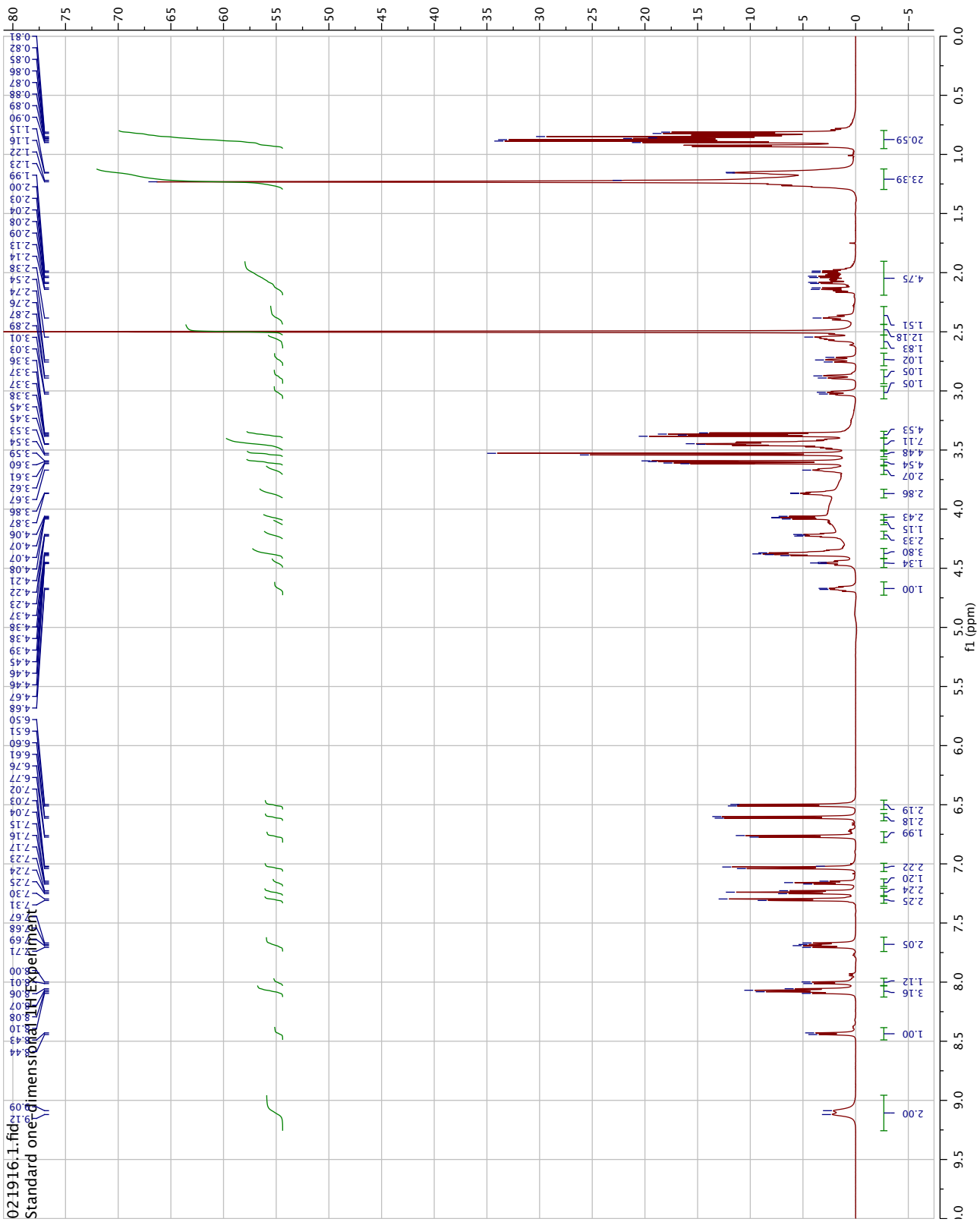
Supplementary Figure 2. Scope of the syn-BNP approach. **a)** Cumulative counts of completed bacterial genomes in GenBank (blue bars) and NRPS gene clusters amenable to the syn-BNP approach (red bars). **b)** Our analysis of Minimum Information about a Biosynthetic Gene Cluster (MIBiG, http://mibig.secondarymetabolites.org) indicates that the existing bioinformatic tools correctly predicted the core peptide encoded by the vast majority of well studies NRPS gene clusters. Overall 303 adenylation domains from 42 NRPS gene clusters (containing ≥3 modules, full information available) were analyzed. "Correct" predictions include those identical to the natural product as well as the following nearly exact matches: Val/Leu/Ile, Dab/Orn/Lys, Asp/Glu, Asn/Gln. Predictions were deemed "close" in case of "correct" predictions that have additional modifications, e.g., β-hydroxylation, β-methylation, etc.



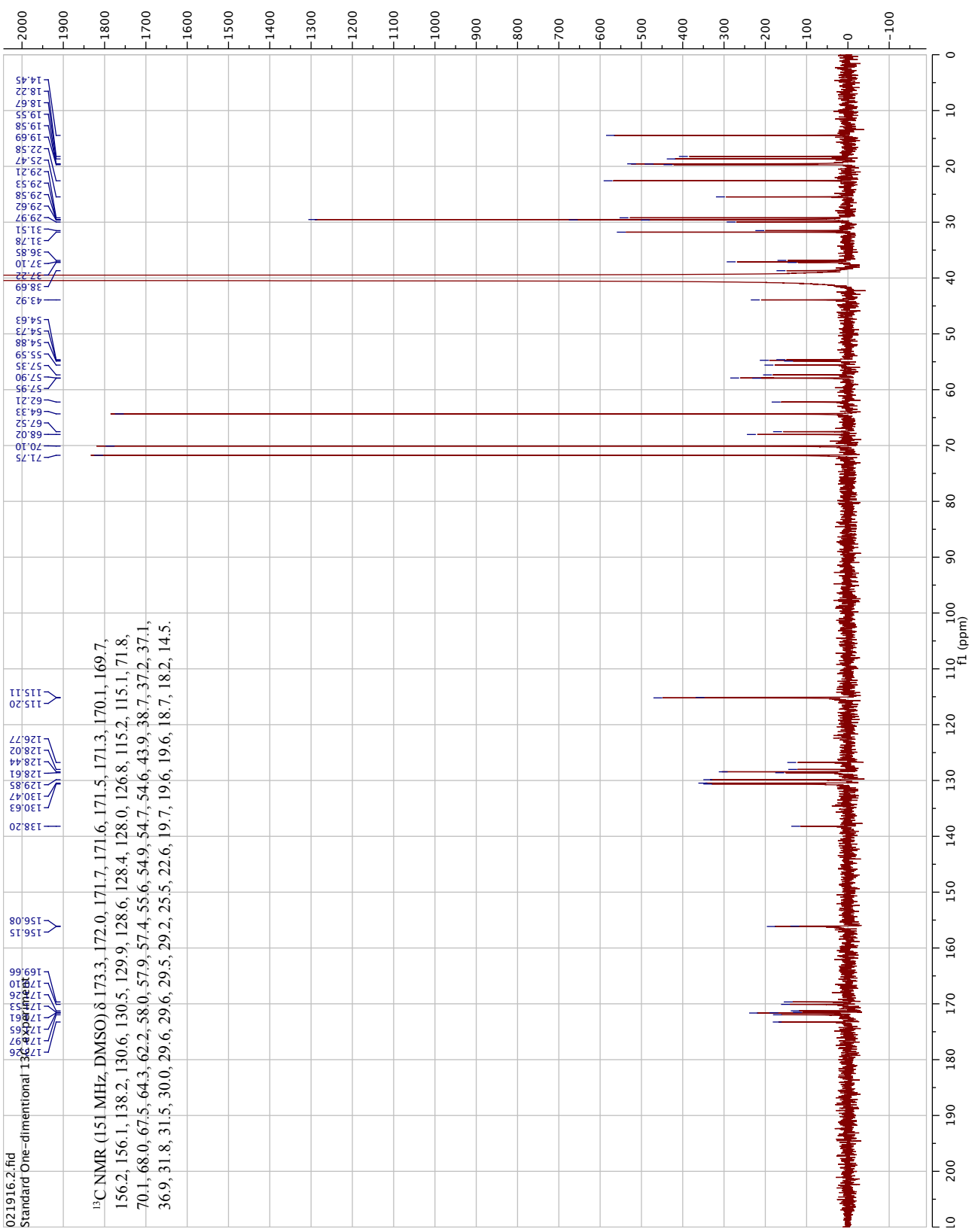
Supplementary Figure 3. SNPs in SAV1754 that confer humimycin A resistance. We selected *S. aureus* USA300 mutants that could survive on 2.5 times the MIC (20 µg/mL) and sequenced the genome of 23 resistant mutants. Upon comparison to the parent strain, we found that all 23 mutants contained one non-synonymous mutation in SAV1754. Sixteen of the 23 are unique mutations and displayed below.



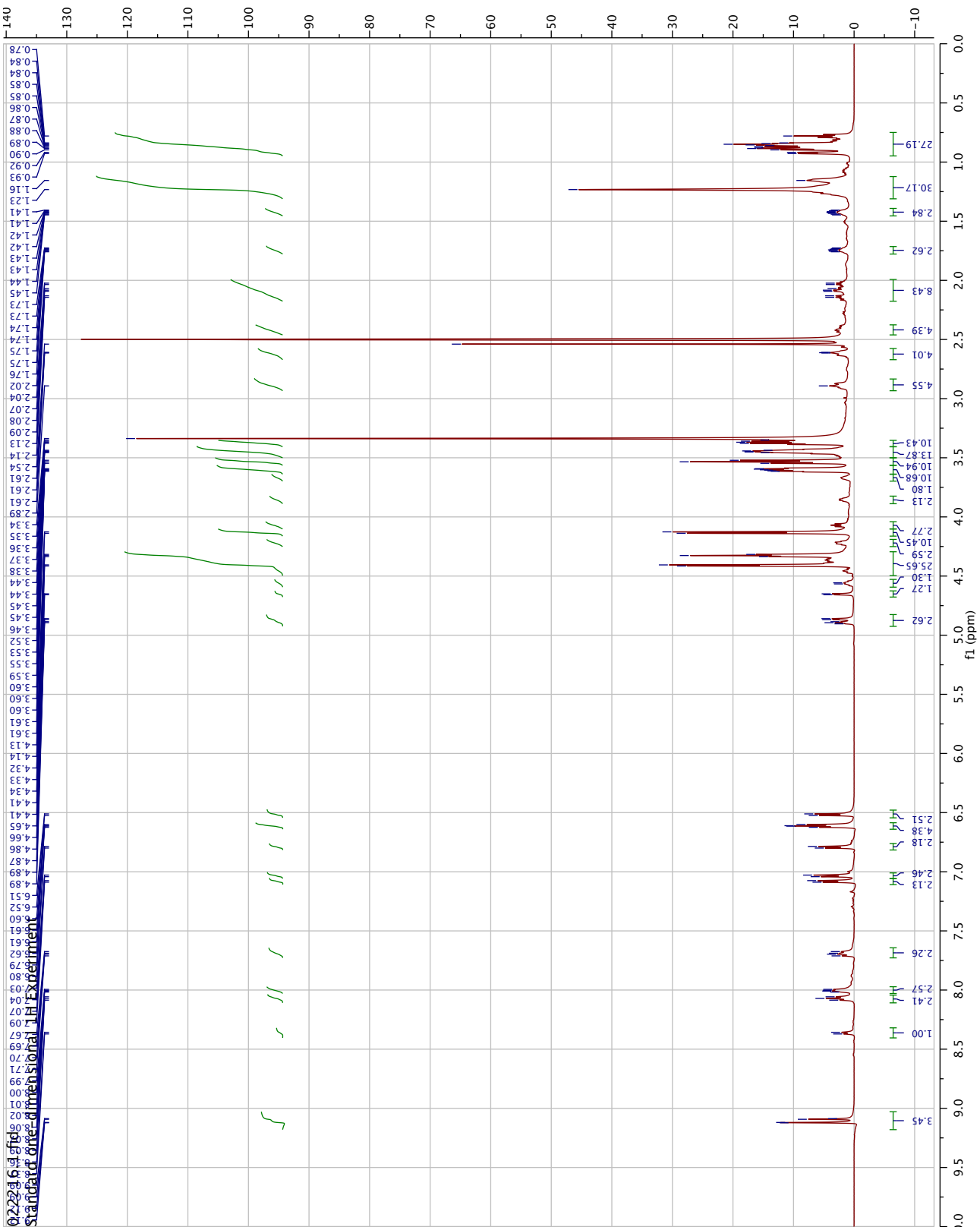
Supplementary Figure 4. ^{1}H NMR spectrum of humimycin A



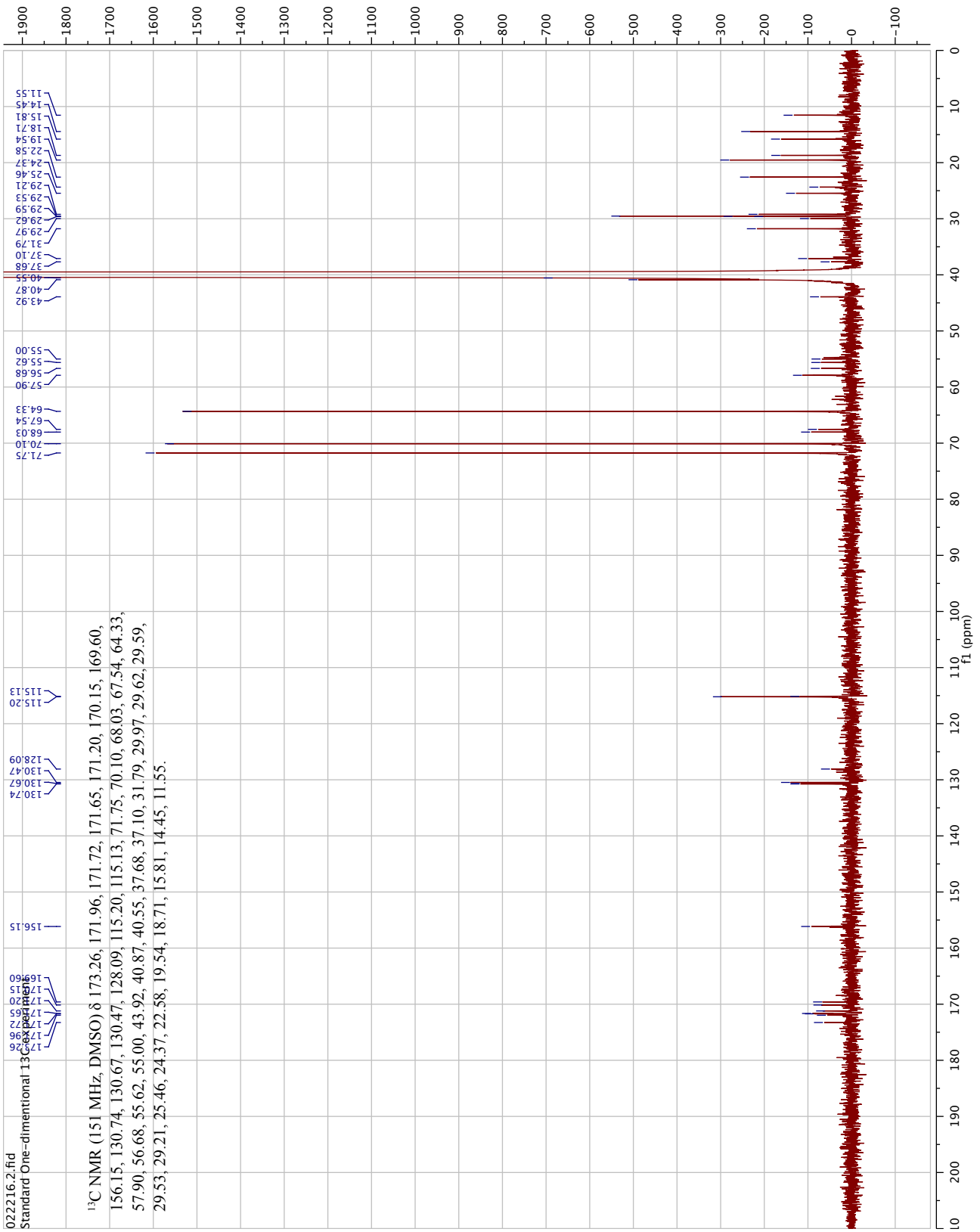
Supplementary Figure 5. ^{13}C NMR spectrum of humimycin A



Supplementary Figure 6. ^{1}H NMR spectrum of humimycin B



Supplementary Figure 7. ^{13}C NMR spectrum of humimycin B



Supplementary Table 1. List of all syn-BNPs for this study

No.	Name	N-Cap	1	2	3	4	5	6	7	8	9	10	11	12	13	MW	IsoMS	Obsd MS	Ion	Organism of origin	Genome	Cl.
1	Human.1	β-Hydroxymyristic acid	D-Tyr	Ser	D-Tyr	Tyr	D-Thr	Ile	Val	← Humimycin B						1134.4	1133.6	1156.6	[M+Na]	*Rhodococcus erythropolis* SK121	gi_229493436	2
2	Human.2	none (NH_2)	D-Tyr	Ser	Ser	Ser	Thr	Gly	Ser							687.7	687.2	688.3		*Rhodococcus erythropolis* SK121	gi_229491953	1
3	Human.3	β-Hydroxymyristic acid	Tyr	Ser	D-Tyr	Phe	D-Thr	Val	Val	← Humimycin A						1104.4	1103.6	1104.2		*Rhodococcus equi* ATCC 33707	gi_326328562	2
4	Human.4	none (NH_2)	Gly	Orn	D-Ser	Ser	Gly	Phe	Val							893.1	892.5	893.55		*Paenibacillus* sp. HGF7	gi_334134556	1
5	Human.5v1	β-Hydroxymyristic acid	D-Thr	Ser	Ala	Ser	Leu	D-Gln	Orn							946.2	945.6	946.6		*Bacillus* sp. 7_6_55CFAA_CT2	gi_365164252	2
6	Human.6	none (NH_2)	Leu	D-Ser	D-Arg	Lys	Orn	Thr								717.9	717.5	718.45		*Pseudomonas* sp. 2_1_26	gi_355654810	2
7	Human.7	none (NH_2)	Ala	Phe	D-Ala	Ala	D-Ala	Phe								596.7	596.3	597.35		*Mycobacterium parascrofulaceum* ATCC BAA-614	gi_296172066	2
8	Human.8v1	β-Hydroxymyristic acid	Tyr	Glu	D-Asp	D-Glu	D-Gln									909.0	908.4	909.5		*Rhodococcus equi* ATCC 33707	gi_326328562	3
9	Human.8v2	β-Hydroxymyristic acid	Tyr	Lys	D-Lys	D-Glu										921.5	922.1	922.55		*Rhodococcus equi* ATCC 33707	gi_326328562	3
10	Oral.1	β-Hydroxymyristic acid	Glu	D-Orn	Tyr	D-Thr	Glu	D-Val	Pro	Glu	D-Tyr	Ile				1482.7	1481.8	1481.55	[M-H]	*Bacillus subtilis* BSn5	gi_320017650	1
11	Oral.2	none (NH_2)	D-Ser	Lys	Gly	Orn	D-Ser	Ser	Gly	Lys	D-Orn	Ser				965.1	964.5	965.55		*Pseudomonas fluorescens* SS101	gi_387997393	2
12	Oral.3	none (NH_2)	Asp	D-Orn	Dab	Thr	Ala	D-Ala	D-Orn	Dab						804.9	804.5	805.5		*Pseudomonas protegens* Pf-5	gi_68342549	3
13	Oral.4v1	none (NH_2)	D-Asp	Orn	D-Phe	Gly	D-Glu	D-Glu	Gly	Glu						895.9	895.4	897.4		*Bradyrhizobium elkanii* WSM1741	gi_548696015	1
14	Oral.4v2	none (NH_2)	D-Asp	Orn	D-Phe	Gly	D-Orn	D-Glu	Gly	Orn						865.9	865.4	866.45		*Bradyrhizobium elkanii* WSM1741	gi_548696015	1
15	Oral.5	β-Hydroxymyristic acid	Phe	Glu	Asp	Ile	Thr	Leu	Ala	Pro						1131.4	1130.7	1129.5	[M-H]	*Uncultured bacterium* ACD_77C00477	gi_406883084	1
16	Oral.6	none (NH_2)	D-Leu	D-Thr	D-Ile	D-Leu	D-Ser	Leu	D-Ser	Ile						859.1	858.5	859.55		*Pseudomonas fluorescens* SS101	gi_387997393	1
17	Oral.7	β-Hydroxymyristic acid	Asn	D-Tyr	D-Asn	Gln	Pro	D-Ser	Asn							1062.2	1061.5	1084.5	[M+Na]	*Bacillus subtilis* subsp. spizizenii str. W23	gi_301410941	4
18	Oral.8	none (NH_2)	D-Ser	Arg	D-Ser	Orn	Lys	Orn	Thr	Thr						907.0	906.5	907.5		*Pseudomonas aeruginosa* UCBPP-PA14	gi_115583796	4
19	Oral.9v1	none (NH_2)	D-Ser	D-Orn	Orn	Gly	D-Thr	Ser	Orn							692.8	692.4	693.4		*Pseudomonas aeruginosa* PA7	gi_150958624	3
20	Oral.9v2	none (NH_2)	D-Ser	D-Orn	Glu	Gly	D-Thr	Ser	Glu							722.7	722.3	723.35		*Pseudomonas aeruginosa* PA7	gi_150958624	3
21	Oral.10v1	none (NH_2)	Glu	Orn	Asp	Dab	Asp	D-Orn	Gly							762.8	762.4	763.3		*Ralstonia pickettii* DTP0602	gi_546340292	1
22	Oral.10v2	none (NH_2)	Asp	Dab	Glu	Asp	D-Orn	Orn	Gly							762.8	762.4	763.35		*Ralstonia pickettii* DTP0602	gi_546340292	1
23	Oral.11	β-Hydroxymyristic acid	Leu	D-Ser	Pro	Val	Cys	Gly								801.1	800.5	801.4		*Streptococcus mutans* TCI-145	gi_349996068	1
24	Oral.12	none (NH_2)	Ser	D-Asp	Ser	Orn	Ser									508.5	508.2	509.25		*Achromobacter xylosoxidans* A8	gi_310757913	1
25	Oral.13	none (NH_2)	Ala	Leu	MeOrn	Ser	Thr									532.6	532.3	533.35		*Mycobacterium neoaurum* ATCC 25795	gi_635284330	1
No.	Name	N-Cap	1	2	3	4	5	6	7	8	9	10	11	12	13	MW	IsoMS	Obsd MS	Ion	Organism of origin	Genome	Cl.
Failed peptide syntheses																						
	Human.5v2	β-Hydroxymyristic acid	D-Thr	Gln	Ala	Ser	Leu	D-Glu	Glu							1003.2	1002.6			*Bacillus* sp. 7_6_55CFAA_CT2	gi_365164252	2
	Human.9	β-Hydroxymyristic acid	Ser	Glu	Phe	Asp	Ser	Ser	Phe	Ser	Phe	Ser	Leu	Ser		1652.8	1651.8			*Rhodococcus equi* ATCC 33707	gi_326328562	7
	Human.10	β-Hydroxymyristic acid	D-Ala	Thr	D-Ala	Leu	D-Thr	Leu	Phe							962.2	961.6			*Rhodococcus equi* ATCC 33707	gi_326328562	5
	Human.11	β-Hydroxymyristic acid	Phe	Leu	Thr	Phe	Tyr	Tyr								1079.4	1078.6			*Rhodococcus equi* ATCC 33707	gi_326328562	6
	Oral.14	β-Hydroxymyristic acid	D-Leu	D-Asp	D-Thr	D-Ile	D-Leu	D-Ser	Leu	Leu	D-Ser	Ile				1313.7	1312.9			*Pseudomonas protegens* Pf-5	gi_68342549	1

Supplementary Table 2. Bacteria strain information

a)	Name of Bacteria	Strain	Media	Resistance
ESKAPE pathogens	*Enterococcus faecium*	Com15	LB	
	Staphylococcus aureus	USA300	LB	Methicillin
	Klebsiella pneumonia	ATCC 10031	LB	
	Acinetobacter baumannii	ATCC 17978	LB	
	Pseudomonas aeruginosa	PAO1	LB	
	Enterobacter cloacae	ATCC 14037	LB	
Various MRSA strains	*Staphylococcus aureus*	ATCC BAA-42	LB	Methicillin
	Staphylococcus aureus	NRS100	LB	Methicillin, tetracycline
	Staphylococcus aureus	NRS108	LB	Gentamicin
	Staphylococcus aureus	ATCC BAA-1721	LB	Hyper-virulent
	Staphylococcus aureus	NRS281	LB	Erythromycin
	Staphylococcus aureus	NRS22	LB	Vancomycin
	Staphylococcus aureus	USA800	LB	Methicillin
	Staphylococcus aureus	USA500	LB	Methicillin
	Staphylococcus aureus	USA300	LB	Methicillin
	Staphylococcus aureus	USA200	LB	Methicillin
	Staphylococcus aureus	USA100	LB	Methicillin
Staphylo-cocci	*Staphylococcus intermedius*	NCTC 11048	LB	
	Staphylococcus delphini	8086	LB	
	Staphylococcus pseudintermedius	ED99	LB	

b)	Name of Bacteria	Strain	Media	Resistance
Actino-bacteria	*Bifidobacterium adolescentis* *	L2-32	LYBHI	
	Bifidobacterium longum *	CCUG 52486	LYBHI	
	Corynebacterium amycolactum §	SK46	LYBHI	
	Rhodococcus erythropolis	ATCC 11048	BHI	
	Rothia dentocariosa	M567	LYBHI	
	Rothia mucilaginosa §	CC87LB	LYBHI	
Bacter-oidetes	*Bacteroides dorei* *	5_1_36/D4	LYBHI	
	Bacteroides finegoldii *	CL09T03C10	LYBHI	
	Parabacteroides merdae *	CL03T12C32	LYBHI	
	Prevotella melaninogenica *	D10	LYBHI	
Firmicutes	*Bacillus subtilis*	168 1A1	LYBHI	
	Eubacterium sp. *	3_1_31	LYBHI	
	Staphylococcus aureus	RN4220	LB/BHI	
	Staphylococcus aureus	USA300	LB	Methicillin
	Staphylococcus epidermidis	RP62A	LB	Methicillin
	Streptococcus mitis	B6	LYBHI	
	Streptococcus mutans	UA159	LYBHI	
	Streptococcus pneumoniae	TCH8431/19A	BHI	
	Streptococcus sanguinis	SK36	BHI	
Proteo-bacteria	*Escherichia coli*	LF82	LB	
	Neisseria mucosa	ATCC 25996	LYBHI	
	Salmonella enterica	subsp. *enterica*	LYBHI	

1. Superscripts denote special growth conditions, i.e., anaerobic atmosphere (*) and 30°C incubation(§).
2. The LYBHI medium contains (per liter) brain heart infusion (BHI, 37 g), yeast extract (5 g), maltose (1 g), cellobiose (1 g), cysteine (0.5 g), and hemin (5 mg).

Supplementary Table 3. MICs of humimycin A and its alanine scan analogs

		MIC (μg/mL) [ab]	
Name [c]	**Sequence** [de]	**(R) diastereomer** [f]	**(S) diastereomer** [f]
Original syn-BNP	HMA–YSyFtVV	32	8 (humimycin A, **1**)
Ala1	HMA–**A**SyFtVV	>	>
Ala2	HMA–Y**A**yFtVV	>	>
Ala3	HMA–YS**a**FtVV	>	>
Ala4	HMA–YSy**A**tVV	>	>
Ala5	HMA–YSyF**a**VV	>	>
Ala6	HMA–YSyFt**A**V	>	32
Ala7	HMA–YSyFtV**A**	>	>

[a] Minimal inhibitory concentration (MIC) values are expressed in units of μg/mL.
[b] ">" denotes MIC values greater than 128 μg/mL (highest concentration tested).
[c] Alanine substitution analogs were synthesized on 2-chlorotrityl resins. The first amino acid (1.2 equiv.) was loaded in the presence of DIEA (3 equiv.) and capped by DCM/MeOH/DIEA (80:15:5) treatment for 30 min. The rest of the SPPS steps were identical to those described in *Peptide Synthesis* section of the *Online Methods*.
[d] HMA denotes *N*-terminal modification with 3-hydroxymyristic acid (MHMA).
[e] Amino acids are abbreviated using standard 1-letter codes; upper and lower case letters denote amino acids in L- and D- forms, respectively.
[e] Humimycin A (**1**) refers to the syn-BNP *N*-acylated with the (S)-HMA (**Supplementary Figure S1**).

Supplementary Table 4. MICs of humimycin A against various clinical isolates of MRSA

Strain	MIC (μg/mL)
USA300	8
BAA-42	8
NRS108	8
NRS22	8
USA100	8
USA200	16
USA800	16
BAA-1721	32
NRS100	128

Supplementary Table 5. List of SNPs in SAV1754 that confer humimycin A resistance

Summary

The following table lists all SNPs acquired through one round of resistance selection.
Among 23 fully sequenced mutants, 22 acquired a single point mutation on protein SAV1754 and 1 acquired two mutations
A number of these mutants (9) acquired mutations, in addition to SAV1754, located elsewhere in the genome

	Mutations within SAV1754													Non-SAV1754 mutations					
MUTANT	**CHROM**	**POS**	**TYPE**	**REF**	**ALT**	**EVIDENCE**	**STRAND**	**NT_POS**	**AA_POS**	**EFFECT**	**LOCUS_TAG**	**GENE PRODUCT**		**MUTANT**	**CHROM**	**POS**	**TYPE**	**REF**	**ALT**	**EVIDENCE**	**STRAND**
1	NC_007793	1875948	SNP	C	T	T:65 C:0	-	169/1662	57/553	missense_variant c.169G>A p.Ala57Thr	RS09285	polysaccharide biosynthesis protein		**2**	NC_007793	813007	SNP	A	G	G:93 A:0	-
2	NC_007793	1874892	SNP	C	T	T:50 C:0	-	1225/1662	409/553	missense_variant c.1225G>A p.Gly409Ser	RS09285	polysaccharide biosynthesis protein		**5**	NC_007791	2378	SNP	G	A	A:6692 G:0	+
3	NC_007793	1875468	SNP	G	A	A:144 G:0	-	649/1662	217/553	missense_variant c.649C>T p.Leu217Phe	RS09285	polysaccharide biosynthesis protein		**8**	NC_007791	2378	SNP	G	A	A:6862 G:8	+
4	NC_007793	1876068	SNP	T	C	C:43 T:0	-	49/1662	17/553	missense_variant c.49A>G p.Ser17Gly	RS09285	polysaccharide biosynthesis protein		**10**	NC_007793	2363468	SNP	G	T	T:234 G:1	-
5	NC_007793	1874892	SNP	C	T	T:34 C:0	-	1225/1662	409/553	missense_variant c.1225G>A p.Gly409Ser	RS09285	polysaccharide biosynthesis protein		**12**	NC_007793	626985	SNP	C	T	T:60 C:0	-
6	NC_007793	1875894	SNP	T	A	A:69 T:0	-	223/1662	75/553	missense_variant c.223A>T p.Asn75Tyr	RS09285	polysaccharide biosynthesis protein		**12**	NC_007793	123206	SNP	G	C	C:100 G:0	-
7	NC_007793	1875791	SNP	G	A	A:69 G:0	-	326/1662	109/553	missense_variant c.326C>T p.Ala109Val	RS09285	polysaccharide biosynthesis protein		**14**	NC_007793	2433481	SNP	A	C	C:30 A:0	-
8	NC_007793	1875948	SNP	C	T	T:57 C:0	-	169/1662	57/553	missense_variant c.169G>A p.Ala57Thr	RS09285	polysaccharide biosynthesis protein		**17**	NC_007793	2350655	SNP	G	A	A:47 G:0	-
9	NC_007793	1875903	SNP	C	T	T:66 C:0	-	214/1662	72/553	missense_variant c.214G>A p.Ala72Thr	RS09285	polysaccharide biosynthesis protein		**18**	NC_007791	2378	SNP	G	A	A:9202 G:6	+
10	NC_007793	1875416	SNP	G	A	A:120 G:0	-	701/1662	234/553	missense_variant c.701C>T p.Ser234Phe	RS09285	polysaccharide biosynthesis protein		**22**	NC_007791	2378	SNP	G	A	A:8436 G:102	+
11	NC_007793	1875350	SNP	G	T	T:41 G:1	-	767/1662	256/553	missense_variant c.767C>A p.Pro256Gln	RS09285	polysaccharide biosynthesis protein									
12	NC_007793	1875737	SNP	C	A	A:32 C:0	-	380/1662	127/553	missense_variant c.380G>T p.Trp127Leu	RS09285	polysaccharide biosynthesis protein									
13	NC_007793	1875974	SNP	G	A	A:43 G:0	-	143/1662	48/553	missense_variant c.143C>T p.Ala48Val	RS09285	polysaccharide biosynthesis protein									
14	NC_007793	1875736	SNP	C	G	G:28 C:0	-	381/1662	127/553	missense_variant c.381G>C p.Trp127Cys	RS09285	polysaccharide biosynthesis protein									
15	NC_007793	1875869	SNP	C	T	T:15 C:0	-	248/1662	83/553	missense_variant c.248G>A p.Ser83Asn	RS09285	polysaccharide biosynthesis protein									
16	NC_007793	1875372	SNP	C	G	G:99 C:0	-	745/1662	249/553	missense_variant c.745G>C p.Glu249Gln	RS09285	polysaccharide biosynthesis protein									
	NC_007793	1875792	SNP	C	T	T:44 C:0	-	325/1662	109/553	missense_variant c.325G>A p.Ala109Thr	RS09285	polysaccharide biosynthesis protein									
17	NC_007793	1875902	SNP	G	A	A:46 G:0	-	215/1662	72/553	missense_variant c.215C>T p.Ala72Val	RS09285	polysaccharide biosynthesis protein									
18	NC_007793	1875930	SNP	C	A	A:85 C:2	-	187/1662	63/553	missense_variant c.187G>T p.Val63Leu	RS09285	polysaccharide biosynthesis protein									
19	NC_007793	1875698	SNP	C	G	G:67 C:0	-	419/1662	140/553	missense_variant c.419G>C p.Ser140Thr	RS09285	polysaccharide biosynthesis protein									
20	NC_007793	1875480	SNP	C	T	T:206 C:0	-	637/1662	213/553	missense_variant c.637G>A p.Gly213Ser	RS09285	polysaccharide biosynthesis protein									
21	NC_007793	1875857	SNP	T	A	A:102 T:0	-	260/1662	87/553	missense_variant c.260A>T p.Tyr87Phe	RS09285	polysaccharide biosynthesis protein									
22	NC_007793	1874892	SNP	C	T	T:56 C:1	-	1225/1662	409/553	missense_variant c.1225G>A p.Gly409Ser	RS09285	polysaccharide biosynthesis protein									
23	NC_007793	1875870	SNP	T	A	A:30 T:0	-	247/1662	83/553	missense_variant c.247A>T p.Ser83Cys	RS09285	polysaccharide biosynthesis protein									

Supplementary Table 6. Overexpression of SAV1754 confers humimycin A resistance.

a) Primers for the amplification of SAV1754 [a]

Primer F	gcgcgagctcagaggaggactctct**ATG**AGTGAAAGTAAAGAAATGGTGC
Primer R	gcgcggtacc**TCA**TCGTAAAAACCTAACTCTACGTC

b) PCR protocol

Step 1	98°C for 8 min
Step 2	98°C for 30 sec
Step 3	62°C for 30 sec
Step 4	72°C for 2 min
Step 5	Go to Step 2 for 32 times
Step 6	72°C for 8 min

c) Susceptibility to humimycin A of *S. aureus* hosting various plasmids [bc]

Insert	Vector	MIC (μg/mL) [d]
Empty vector	pRMC2	8
SAV1754 (WT)	pRMC2	>128
SAV1754 (mutant no.8)	pRMC2	>128

[a] Start and stop codons of the SAV1754 gene (upper case) are in bold; restriction sites for cloning are underlined (*SacI*/*KpnI*); ribosome binding sequence was built into the forward primer (lower case).
[b] The SAV1754 gene was placed under the control of the tetracycline-inducible promoter $P_{xyl/tetO}$.
[c] Empty and recombinant plasmids were transformed into *S. aureus* RN4220.
[d] Minimal inhibitor concentrations (MIC) were tested in the presence of 5 ng/mL of anhydrotetracycline and recorded after 48 h of growth; ">128" denotes MIC values greater than 128 μg/mL (highest concentration tested).