

## Computer-assisted $^1\text{H}$ NMR Analysis of the anti-Tuberculosis Drug Lead Ecumicin

Wei Gao,<sup>a,b</sup> José G. Napolitano,<sup>a</sup> David C. Lankin,<sup>a</sup> Jin-Yong Kim,<sup>c,d</sup> Ying-Yu Jing,<sup>c</sup> Hanki Lee,<sup>c,d</sup> Joo-Won Suh,<sup>c,d</sup> Shao-Nong Chena<sup>b</sup> and Guido F. Paulia<sup>b\*</sup>

---

<sup>a</sup> Department of Medicinal Chemistry and Pharmacognosy, College of Pharmacy, University of Illinois at Chicago, Chicago, IL 60612, U.S.A.

<sup>b</sup> Institute for Tuberculosis Research, College of Pharmacy, University of Illinois at Chicago, Chicago, IL 60612, U.S.A.

<sup>c</sup> Division of Bioscience and Bioinformatics, College of Natural Science, Myongji University, Cheoin-gu, Yongin, Gyeonggi-Do, 17058, South Korea

<sup>d</sup> Center for Nutraceutical and Pharmaceutical Materials, Myongji University, Cheoin-gu, Yongin, Gyeonggi-Do, 17058, South Korea



10TrpH3B/	1	3.693816	1*1*1	STAT=Y	PRED= 3.367	RANGE= 0.629	WIDTH(Y)= 2.091	RESP(Y)= 0.2077	HSQC= C66
10Trp_H5/	1	6.698366	1*1*1	STAT=Y	PRED= 6.772	RANGE= 0.276	WIDTH(Y)= 1.646	RESP(Y)= 0.2387	HSQC= C68
10Trp_H9/	1	6.441818	1*1*1	STAT=Y	PRED= 7.089	RANGE= 0.219	WIDTH(Y)= 1.124	RESP(Y)= 0.2512	HSQC= C73
10Trp_H7/	1	6.918046	1*1*1	STAT=Y	PRED= 7.333	RANGE= 0.309	WIDTH(Y)= 0.720	RESP(Y)= 0.2360	HSQC= C75
10TrpH12/	1	3.825512	1*1*3	STAT=Y	PRED= 4.072	RANGE= 0.249	WIDTH(Y)= 1.239	RESP(Y)= 0.2230	HSQC= C77
11Val_H2/	1	4.528574	1*1*1	STAT=Y	PRED= 4.303	RANGE= 1.449	WIDTH(Y)= 1.881	RESP(Y)= 0.2193	HSQC= C80
11Val_H3/	1	2.200073	1*1*1	STAT=Y	PRED= 2.094	RANGE= 0.589	WIDTH(Y)= 1.787	RESP(Y)= 0.2038	HSQC= C81
11Val_H4/	1	1.028232	1*1*3	STAT=Y	PRED= 1.094	RANGE= 0.249	WIDTH(Y)= 2.114	RESP(Y)= 0.2186	HSQC= C82
11Val_H5/	1	0.989110	1*1*3	STAT=Y	PRED= 1.080	RANGE= 0.299	WIDTH(Y)= 1.810	RESP(Y)= 0.1543	HSQC= C83
12Phe_H2/	1	4.853685	1*1*1	STAT=Y	PRED= 4.499	RANGE= 0.989	WIDTH(Y)= 1.248	RESP(Y)= 0.1976	HSQC= C86
12Phe_H3/	1	5.343836	1*1*1	STAT=Y	PRED= 5.208	RANGE= 0.489	WIDTH(Y)= 1.873	RESP(Y)= 0.2247	HSQC= C87
12Phe_H5/	1	7.239580	1*2*1	STAT=Y	PRED= 7.524	RANGE= 0.309	WIDTH(Y)= 2.159	RESP(Y)= 0.2433	HSQC= C89_93
12Phe_H6/	1	7.255476	1*2*1	STAT=Y	PRED= 7.372	RANGE= 0.579	WIDTH(Y)= 1.167	RESP(Y)= 0.2391	HSQC= C90_92
12Phe_H7/	1	7.200628	1*1*1	STAT=Y	PRED= 7.281	RANGE= 0.309	WIDTH(Y)= 1.552	RESP(Y)= 0.2565	HSQC= C91
13Val_H2/	1	4.396075	1*1*1	STAT=Y	PRED= 4.169	RANGE= 0.669	WIDTH(Y)= 1.198	RESP(Y)= 0.2136	HSQC= C97
13Val_H3/	1	1.967770	1*1*1	STAT=Y	PRED= 1.935	RANGE= 0.869	WIDTH(Y)= 1.937	RESP(Y)= 0.2420	HSQC= C99
13Val_H4/	1	0.934957	1*1*3	STAT=Y	PRED= 1.277	RANGE= 0.219	WIDTH(Y)= 2.639	RESP(Y)= 0.2817	HSQC= C100
13Val_H5/	1	0.919351	1*1*3	STAT=Y	PRED= 0.933	RANGE= 0.278	WIDTH(Y)= 2.019	RESP(Y)= 0.2636	HSQC= C101

COUPLING CONSTANTS(HZ):

J115_217	7.7949	J	10Trp_H8	10Trp_H9	STAT=Y	PRED= 8.100	RANGE= 0.890
J115_218	8.1831	J	10Trp_H8	10Trp_H7	STAT=Y	PRED= 7.850	RANGE= 0.800
J122_123	9.1830	J	01Val_H2	01Val_H3	STAT=Y	PRED= 12.510	RANGE= 4.000
J123_124	6.5874	J	01Val_H3	01Val_H6	STAT=Y	PRED= 6.640	RANGE= 0.200
J123_127	6.6158	J	01Val_H3	01Val_H7	STAT=Y	PRED= 6.640	RANGE= 0.200
J131_132	8.7551	J	02Val_H2	02Val_H3	STAT=Y	PRED= 12.600	RANGE= 6.000
J132_133	6.7883	J	02Val_H3	02Val_H4	STAT=Y	PRED= 6.640	RANGE= 0.200
J132_136	6.7203	J	02Val_H3	02Val_H5	STAT=Y	PRED= 6.640	RANGE= 0.200
J142_143	11.2368	J	03Ile_H2	03Ile_H3	STAT=Y	PRED= 13.150	RANGE= 4.000
J143_144	1.5555	J	03Ile_H3	03IleH4A	STAT=Y	PRED= 1.950	RANGE= 3.000
J143_145	3.3192	J	03Ile_H3	03IleH4B	STAT=Y	PRED= 12.660	RANGE= 2.800
J143_146	6.6322	J	03Ile_H3	03Ile_H6	STAT=Y	PRED= 6.640	RANGE= 0.200
J144_145	-12.9087	J	03IleH4A	03IleH4B	STAT=Y	PRED= -13.400	RANGE= 0.600
J144_149	7.2572	J	03IleH4A	03Ile_H5	STAT=Y	PRED= 7.440	RANGE= 0.150
J145_149	7.6370	J	03IleH4B	03Ile_H5	STAT=Y	PRED= 7.440	RANGE= 0.150
J153_154	2.3441	J	04Thr_H2	04Thr_H3	STAT=Y	PRED= 2.690	RANGE= 6.600
J154_155	6.5142	J	04Thr_H3	04Thr_H4	STAT=Y	PRED= 6.250	RANGE= 0.600
J161_162	3.7286	J	05Thr_H2	05Thr_H3	STAT=Y	PRED= 2.440	RANGE= 4.400
J162_164	6.4627	J	05Thr_H3	05Thr_H4	STAT=Y	PRED= 6.250	RANGE= 0.600
J168_169	8.9241	J	06Val_H2	06Val_H3	STAT=Y	PRED= 2.980	RANGE= 6.600
J169_170	6.6943	J	06Val_H3	06Val_H4	STAT=Y	PRED= 6.640	RANGE= 0.200
J169_173	7.0127	J	06Val_H3	06Val_H5	STAT=Y	PRED= 6.640	RANGE= 0.200
J179_180	8.5788	J	07Leu_H2	07LeuH3A	STAT=Y	PRED= 3.040	RANGE= 6.000
J179_181	6.5941	J	07Leu_H2	07LeuH3B	STAT=Y	PRED= 13.950	RANGE= 4.000
J180_181	-13.5581	J	07LeuH3A	07LeuH3B	STAT=Y	PRED= -14.710	RANGE= 1.600
J180_182	5.4402	J	07LeuH3A	07Leu_H4	STAT=Y	PRED= 12.990	RANGE= 2.560
J181_182	7.9864	J	07LeuH3B	07Leu_H4	STAT=Y	PRED= 2.270	RANGE= 3.000
J182_183	6.5528	J	07Leu_H4	07Leu_H5	STAT=Y	PRED= 6.640	RANGE= 0.200
J182_186	6.5909	J	07Leu_H4	07Leu_H6	STAT=Y	PRED= 6.640	RANGE= 0.200
J190_191	8.8837	J	08Val_H2	08Val_H3	STAT=Y	PRED= 12.880	RANGE= 6.000
J191_192	6.5495	J	08Val_H3	08Val_H4	STAT=Y	PRED= 6.640	RANGE= 0.200
J191_195	6.7835	J	08Val_H3	08Val_H5	STAT=Y	PRED= 6.640	RANGE= 0.200
J201_202	7.6228	J	09Val_H2	09Val_H3	STAT=Y	PRED= 12.960	RANGE= 4.000
J202_203	6.4970	J	09Val_H3	09Val_H4	STAT=Y	PRED= 6.640	RANGE= 0.200
J202_206	6.8287	J	09Val_H3	09Val_H5	STAT=Y	PRED= 6.640	RANGE= 0.200
J212_213	11.1639	J	10Trp_H2	10TrpH3A	STAT=Y	PRED= 12.890	RANGE= 4.000
J212_214	4.7143	J	10Trp_H2	10TrpH3B	STAT=Y	PRED= 1.900	RANGE= 4.400
J213_214	-13.7268	J	10TrpH3A	10TrpH3B	STAT=Y	PRED= -16.760	RANGE= 2.560

J215_218	0.4848	J	10Trp_H5	10Trp_H7	STAT=Y	PRED= -0.300	RANGE= 0.500
J217_218	0.6721	J	10Trp_H9	10Trp_H7	STAT=Y	PRED= 1.020	RANGE= 1.200
J223_224	7.9107	J	11Val_H2	11Val_H3	STAT=Y	PRED= 3.450	RANGE= 9.000
J224_225	6.7500	J	11Val_H3	11Val_H4	STAT=Y	PRED= 6.640	RANGE= 0.200
J224_228	6.8020	J	11Val_H3	11Val_H5	STAT=Y	PRED= 6.640	RANGE= 0.200
J232_233	1.8995	J	12Phe_H2	12Phe_H3	STAT=Y	PRED= 1.900	RANGE= 6.600
J234_238	1.3331	J	12Phe_H5	12Phe_H5	STAT=Y	PRED= 1.980	RANGE= 1.200
J234_235	7.6508	J	12Phe_H5	12Phe_H6	STAT=Y	PRED= 7.660	RANGE= 0.340
J234_237	0.7591	J	12Phe_H6	12Phe_H5	STAT=Y	PRED= 0.550	RANGE= 0.750
J234_236	1.2534	J	12Phe_H5	12Phe_H7	STAT=Y	PRED= 1.260	RANGE= 1.200
J235_237	1.7067	J	12Phe_H6	12Phe_H6	STAT=Y	PRED= 1.480	RANGE= 1.200
J235_236	7.4723	J	12Phe_H6	12Phe_H7	STAT=Y	PRED= 7.400	RANGE= 0.240
J241_242	8.8995	J	13Val_H2	13Val_H3	STAT=Y	PRED= 12.320	RANGE= 6.000
J242_243	6.8280	J	13Val_H3	13Val_H4	STAT=Y	PRED= 6.640	RANGE= 0.200
J242_246	6.4124	J	13Val_H3	13Val_H5	STAT=Y	PRED= 6.640	RANGE= 0.200

CONTROL PARAMETERS:

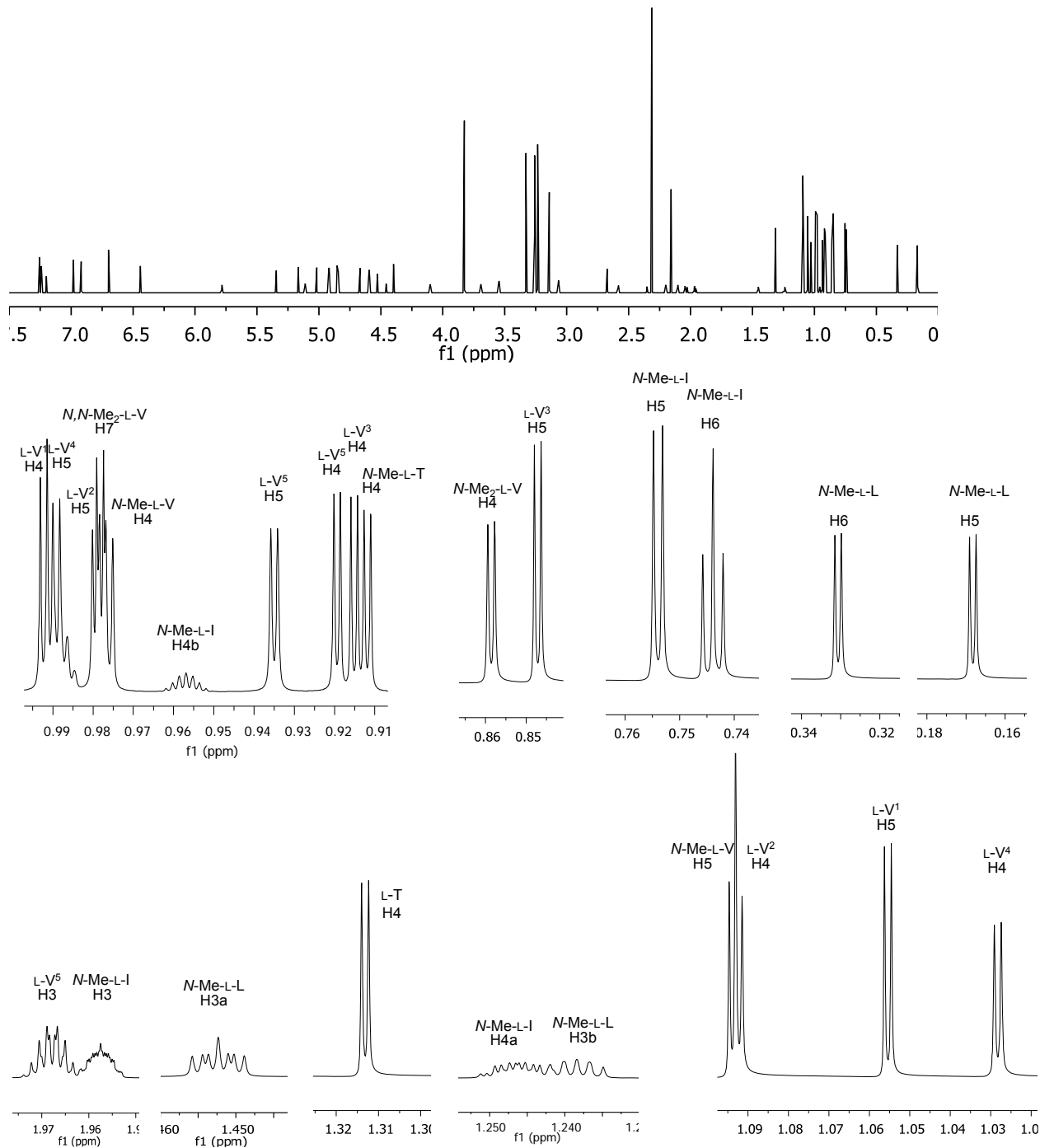
Solvent = none (def. 99% enriched)  
 1.000 = Concentration (vol%, def=1.0%)  
 0.00100000 = Minimum line-intensity  
 0.00100000 = Diagonalization criterium (not in use)  
 11.63298470 = Left frequency (ppm)  
 -4.33373638 = Right frequency (ppm)  
 10.000 = Acquisition time (s, for QMILS)  
 0.000 = Line-width (for modes D, P & T, 0=use defaults)  
 0.103379714 = Data-point resolution (Hz)  
 27.781 = GAUSSIAN (% , 0=use default from INF)  
 -9.448 = Dispersion contribution (% , 0=use default from INF)  
 0.00000000 = Decoupling frequency (for DORES)

CONSTRAINTS (in equations X0 = 1.0)...use no empty lines

EQUAL 01Val\_H2 = 13Val\_H2  
 EQUAL 06Val\_H5 = 11Val\_H4  
 EQUAL 07Leu\_H5 = 07Leu\_H6  
 IGNORE(HZ): 940.350 to 936.356  
 IGNORE(HZ): 2062.779 to 2027.293  
 IGNORE(HZ): 1175.285 to 1149.842  
 IGNORE(HZ): 1232.768 to 1209.210  
 IGNORE(HZ): 5188.061 to 5161.182  
 IGNORE(HZ): 4803.141 to 4787.420  
 IGNORE(HZ): 812.060 to 780.231

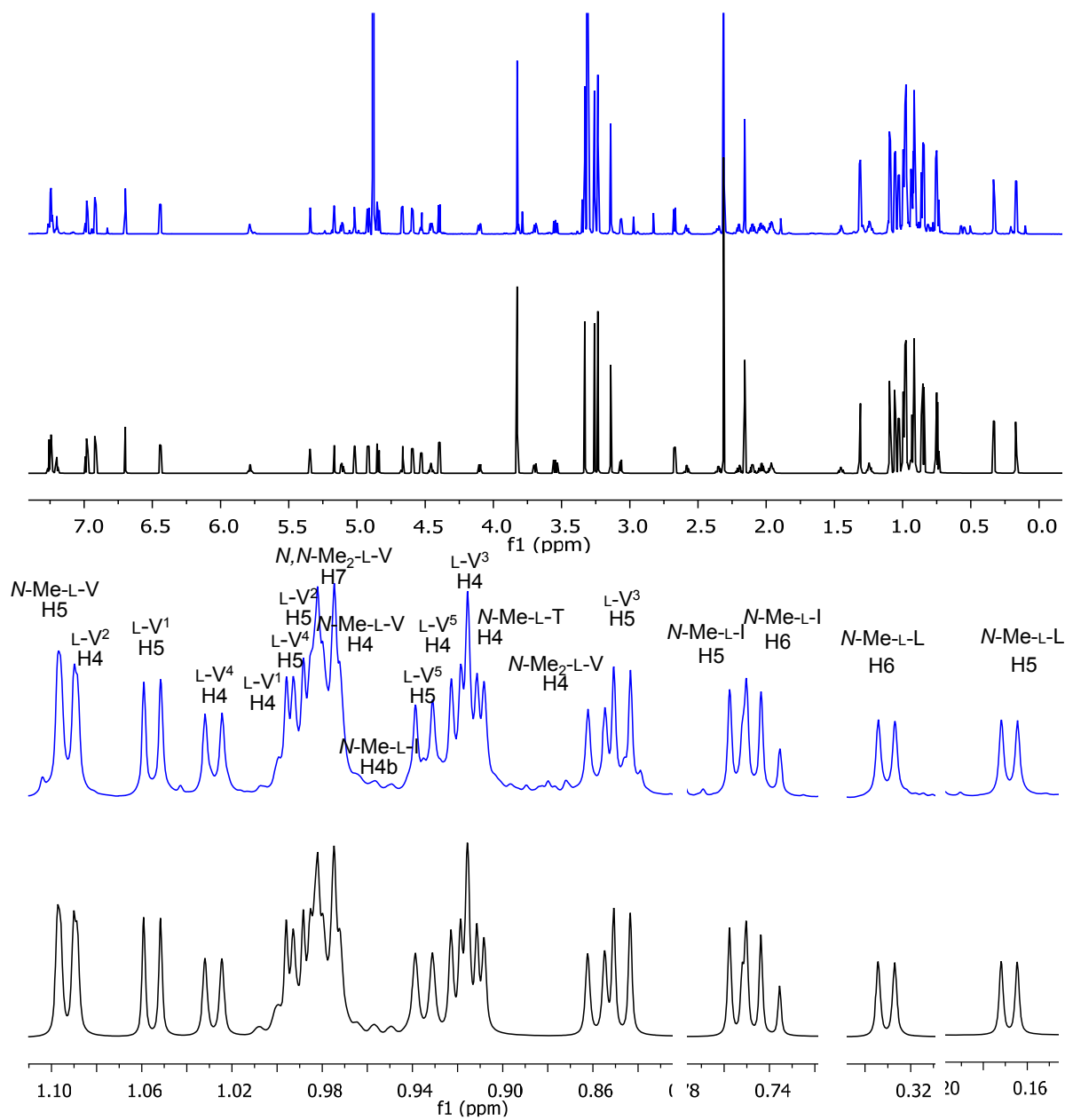
END of FILE

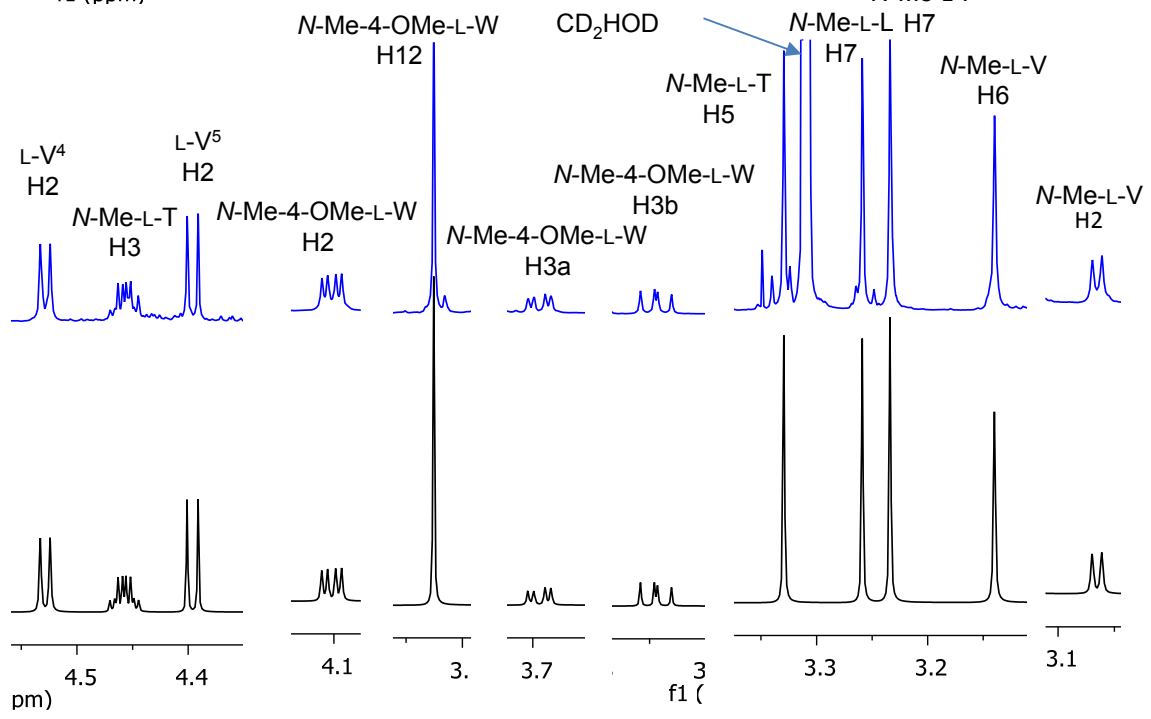
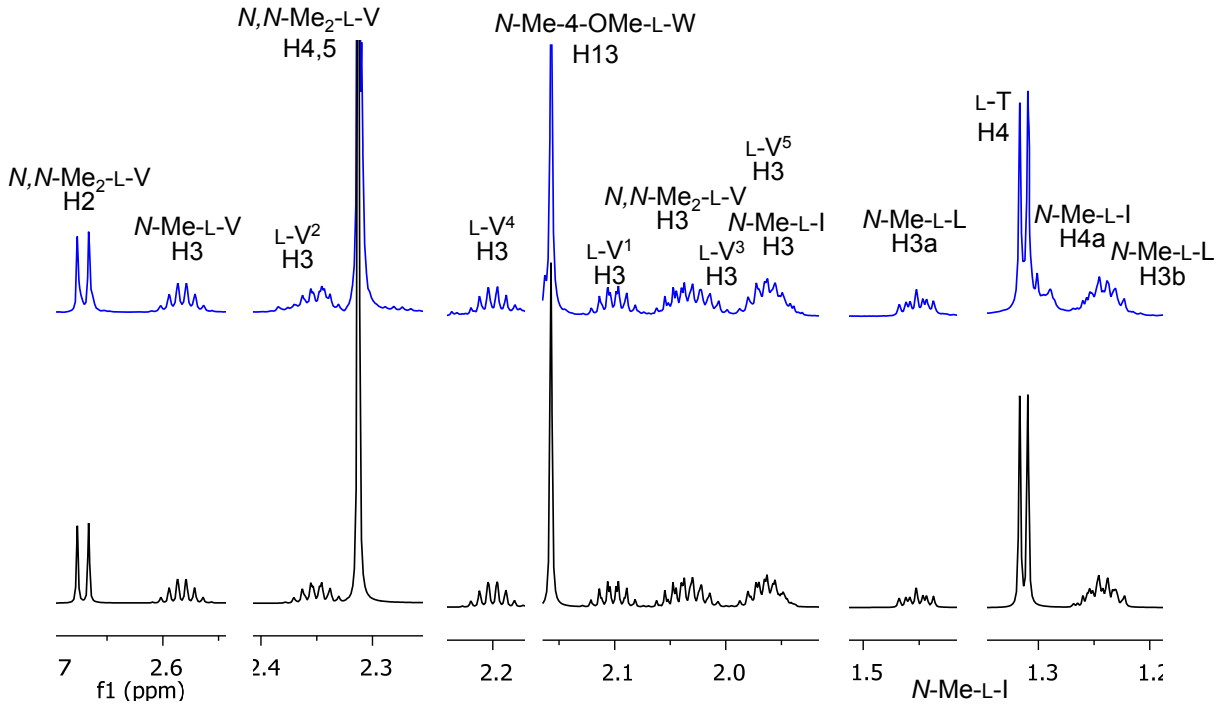
**Supporting Figure 1.** The HiFSA fingerprint of ecumicin (**1**) calculated at 4 GHz demonstrates that presently unattainable NMR magnetic fields will be required to achieve a near first order  $^1\text{H}$  NMR spectrum, suitable for visual analysis of this oligopeptide.



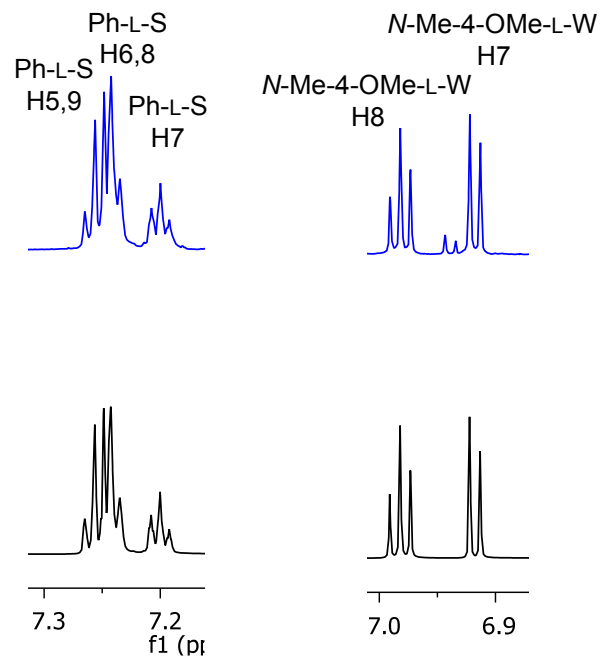
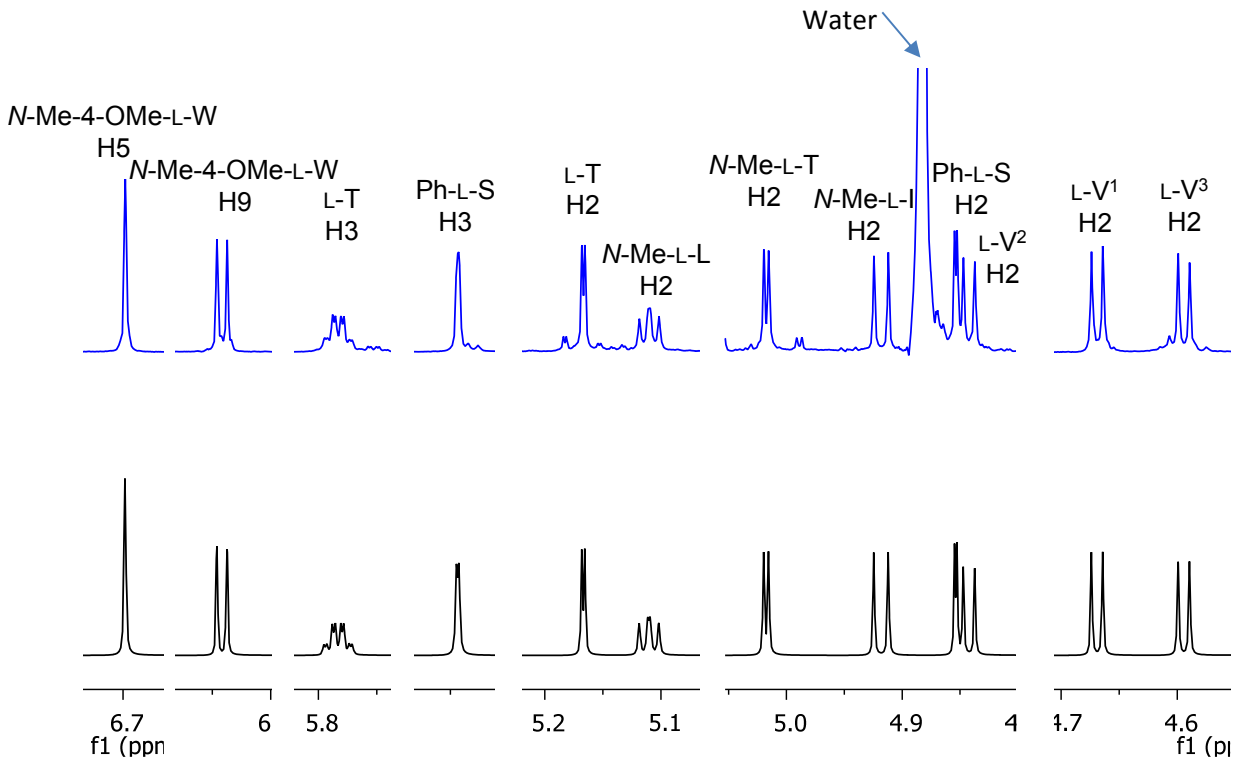


**Supporting Figure 2.** Comparison of the HiFSA fingerprint of ecumicin (**1**) calculated at 900 MHz (black) to the 900MHz experimental spectrum (blue) with expansions. RMS = 0.104%.









**Supporting Table 1.** Comparison of the final HiFSA-based  $^1\text{H}$  chemical shifts ( $\delta_{\text{H}}$ , in ppm) and multiplicities (Mult.) of ecumicin (1), the chemical shifts of previously reported free amino acids (FAA)<sup>[7]d</sup>, and the PERCH predicted values.

Amino Acid	Position	$\delta_{\text{H}}$ (ppm)			Amino Acid	Position	$\delta_{\text{H}}$ (ppm)			
		HiFSA	FAA	Predicted			HiFSA	FAA	Predicted	
<i>N,N</i> -Me <sub>2</sub> -L-V	2	2.672	3.945	4.169	L-V <sup>3</sup>	2	4.594	3.945	4.489	
	3 <sup>a</sup>	2.042	2.358	1.952		3 <sup>a</sup>	2.027	2.358	1.995	
	4	0.847	1.041	0.509		4	0.915	1.066	1.060	
	5	0.978	1.066	0.818		5	0.859	1.041	0.999	
	6,7 ( <i>N,N</i> -Me <sub>2</sub> )		2.313			2.370				
L-V <sup>1</sup>	2	4.669	3.945	4.508	<i>N</i> -Me-L-V	2	3.065	3.945	5.153	
	3 <sup>a</sup>	2.101	2.358	2.155		3 <sup>a</sup>	2.583	2.358	20524	
	4	0.992	1.041	1.038		4	1.094	1.066	1.319	
	5	1.055	1.066	0.943		5	0.976	1.041	1.075	
						6 ( <i>N</i> -Me)	3.140		3.160	
<i>N</i> -Me-L- <i>allo</i> -I	2	4.918	4.022	4.617	<i>N</i> -Me-4-OMe-L-W	2	4.102	4.384	5.043	
	3 <sup>a</sup>	1.957	2.065	2.032		3a	3.545	3.460	3.613	
	4a <sup>a</sup>	0.989	1.337	1.465		3b	3.694	3.535	3.367	
	4b <sup>a</sup>	1.246	1.508	1.175		5 <sup>b</sup>	6.698	7.337	6.772	
	5	0.754	0.948	0.723		7 <sup>c</sup>	6.918	7.698	7.333	
	6	0.744	0.948	0.880		8	6.984	7.192	7.318	
	7 ( <i>N</i> -Me)		3.234			3.196	9	6.442	7.278	7.089
L-T	2	5.167	3.989	4.668	12 ( <i>O</i> -Me)	3.826		4.072		
	3	5.783	4.422	4.010	13 ( <i>N</i> -Me)	2.157		2.936		
	4	1.313	1.351	1.270						
<i>N</i> -Me-L-T	2	5.018	3.989	4.099	L-V <sup>4</sup>	2	4.529	3.945	4.303	
	3	4.457	4.422	4.345		3 <sup>a</sup>	2.200	2.358	2.094	
	4	0.912	1.351	1.252		4	1.028	1.066	1.094	
	5 ( <i>N</i> -Me)		3.330			3.257	5	0.989	1.041	1.080
L-V <sup>2</sup>	2	4.842	3.945	4.525	Ph-L- <i>threo</i> -S	5 (o) <sup>b</sup>	7.240	7.332	7.524	
	3 <sup>a</sup>	2.350	2.358	1.975		6 (m) <sup>b</sup>	7.255	7.429	7.372	
	4	1.092	1.066	0.938		7 (p) <sup>b</sup>	7.201	7.338	7.281	
	5	0.979	1.041	1.094		8 (m) <sup>b</sup>	7.255	7.429	7.372	
						9 (o) <sup>b</sup>	7.240	7.332	7.526	
<i>N</i> -Me-L- <i>allo</i> -L	2	5.110	4.061	5.342	L-V <sup>5</sup>	2	4.396	3.945	4.169	
	3a <sup>a</sup>	1.453	1.857	1.444		3 <sup>a</sup>	1.968	2.358	1.935	
	3b <sup>a</sup>	1.239	1.744	1.914		4	0.935	1.066	1.277	
	4 <sup>a</sup>	0.957	1.766	1.630		5	0.919	1.041	0.933	
	5	0.168	0.958	0.972						
	6	0.331	0.970	0.972						
	7 ( <i>N</i> -Me)		3.259			3.110				

<sup>a</sup> Defined multiplicities are assigned to signals previously reported as multiplets (m).<sup>[4]</sup>

<sup>b</sup> Defined multiplicities are assigned to signals previously reported as singlets (s) or broad singlets (br s).<sup>[4]</sup>

<sup>c</sup> HiFSA reveals the signal previously assigned as a doublet of doublets (dd)<sup>[4]</sup> to be a ddd.

<sup>d</sup> The reported values are based on basic free amino acids with no structural modification, e.g. *N*-methylation.

**Supporting Table 2.** HiFSA-based  $^1\text{H}$ ,  $^1\text{H}$  spin-spin coupling constants ( $J_{\text{HH}}$ , in Hz) of ecumicin (1) compared with the chemical shifts of previously reported free amino acids (FAA)<sup>[7]b</sup>, and the PERCH predicted values.

Amino acid	Coupling	$J_{\text{HH}}$ (Hz)			Amino acid	Coupling	$J_{\text{HH}}$ (Hz)			
		HiFSA	FAA	Predicted			HiFSA	FAA	Predicted	
<i>N,N</i> -Me <sub>2</sub> -L-V	$^3J$ (2,3)	9.18	4.44	12.51	L-V <sup>3</sup>	$^3J$ (2,3)	8.88	4.44	12.88	
	$^3J$ (3,4)	6.59	7.01	6.64		$^3J$ (3,4)	6.55	7.01	6.64	
	$^3J$ (3,5)	6.62	7.02	6.64		$^3J$ (3,5)	6.78	7.02	6.64	
L-V <sup>1</sup>	$^3J$ (2,3)	8.75	4.44	12.6	<i>N</i> -Me-L-V	$^3J$ (2,3)	7.62	4.44	12.96	
	$^3J$ (3,4)	6.79	7.01	6.64		$^3J$ (3,4)	6.50	7.01	6.64	
	$^3J$ (3,5)	6.72	7.02	6.64		$^3J$ (3,5)	6.82	7.02	6.64	
<i>N</i> -Me-L- <i>allo</i> -I	$^3J$ (2,3)	11.23	3.92	13.15	<i>N</i> -Me-4-OMe-L-W	$^3J$ (2,3a)	11.16	7.33	12.89	
	$^3J$ (3,4a) <sup>a</sup>	1.56	7.00	1.95		$^3J$ (2,3b)	4.71	5.30	1.90	
	$^3J$ (3,4b) <sup>a</sup>	3.32	9.00	12.66		$^2J$ (3a,3b)	-13.73	-15.40	-16.76	
	$^3J$ (3,6)	6.63	5.27	6.64		$^5J$ (5,7) <sup>a</sup>	0.48	-	-0.3	
	$^2J$ (4a,4b) <sup>a</sup>	-12.91	-13.68	-13.40		$^4J$ (7,9) <sup>a</sup>	0.67	1.10	1.02	
	$^3J$ (4a,5)	7.26	7.40	7.44		$^3J$ (8,9)	7.79	7.02	8.10	
	$^3J$ (4b,5) <sup>a</sup>	7.64	7.47	7.44		$^3J$ (7,8)	8.18	8.05	7.85	
L-T	$^3J$ (2,3)	2.34	3.88	2.69	L-V <sup>4</sup>	$^3J$ (2,3)	7.91	4.44	3.45	
	$^3J$ (3,4)	6.51	6.64	6.25		$^3J$ (3,4)	6.75	7.01	6.64	
						$^3J$ (3,5)	6.81	7.02	6.64	
<i>N</i> -Me-L-T	$^3J$ (2,3)	3.73	3.88	2.44		$^3J$ (2,3)	1.90		1.90	
	$^3J$ (3,4)	6.46	6.64	6.25		$^4J$ [5(o),9(o)] <sup>a</sup>	1.33	2.06	1.98	
L-V <sup>2</sup>	$^3J$ (2,3)	8.92	4.44	2.98	<i>Ph</i> -L- <i>threo</i> -S	$^3J$ [5(o),6(m)] <sup>a</sup>	7.65	7.69	7.66	
	$^3J$ (3,4)	6.69	7.01	6.64		$^3J$ [8(m),9(o)] <sup>a</sup>	7.65	7.69	7.66	
	$^3J$ (3,5)	7.01	7.02	6.64		$^5J$ [5(o),8(m)] <sup>a</sup>	0.76	0.59	0.55	
<i>N</i> -Me-L-L	$^3J$ (2,3a)	8.58	8.36	3.04		$^5J$ [6(m),9(o)] <sup>a</sup>	0.76	0.59	0.55	
	$^3J$ (2,3b)	6.59	5.84	13.95		$^4J$ [5(o),7(p)] <sup>a</sup>	1.25	1.25	1.26	
	$^2J$ (3a,3b) <sup>a</sup>	-13.56	-14.40	-14.71		$^4J$ [7(p),9(o)] <sup>a</sup>	1.25	1.25	1.26	
	$^3J$ (3a,4) <sup>a</sup>	5.44	6.31	12.99		$^4J$ [6(m),8(m)] <sup>a</sup>	1.71	1.41	1.48	
	$^3J$ (3b,4) <sup>a</sup>	7.99	8.27	2.27		$^3J$ [6(m),7(p)] <sup>a</sup>	7.47	7.524	7.40	
	$^3J$ (4,5)	6.55	6.50	6.64		$^3J$ [7(p),8(m)] <sup>a</sup>	7.47	7.524	7.40	
	$^3J$ (4,6)	6.59	6.62	6.64		L-V <sup>5</sup>	$^3J$ (2,3)	8.90	4.44	12.32
							$^3J$ (3,4)	6.83	7.01	6.64
					$^3J$ (3,5)	6.41	7.02	6.64		

<sup>a</sup> Scalar coupling constants that were not reported previously due to extensive resonance overlap.<sup>[4]</sup>

<sup>b</sup> The reported values are based on basic free amino acids with no structural modification, e.g. *N*-methylation.