HHS Public Access

Author manuscript

Org Lett. Author manuscript; available in PMC 2017 May 31.

Published in final edited form as:

Org Lett. 2014 December 05; 16(23): 6044–6047. doi:10.1021/ol5026603.

Discovery and Characterization of the Tuberculosis Drug Lead Ecumicin

Wei Gao[†,‡], **Jin-Yong Kim**[§,‖], **Shao-Nong Chen**[†], **Sang-Hyun Cho**[‡], **Jongkeun Choi**[⊥], **Birgit U. Jaki**[†,‡], **Ying-Yu Jin**[‖], **David C. Lankin**[†], **Ji-Ean Lee**[‖], **Sun-Young Lee**[§,‖], **James B. McAlpine**[†,‡], **José G. Napolitano**[†], **Scott G. Franzblau**[‡], **Joo-Won Suh**[§,‖,*], and **Guido F. Pauli**[†,‡,*]

[†]Department of Medicinal Chemistry and Pharmacognosy, College of Pharmacy, University of Illinois at Chicago, Chicago, Illinois 60612, United States

[‡]Institute for Tuberculosis Research, College of Pharmacy, University of Illinois at Chicago, 833 South Wood Street, Chicago, Illinois 60612, United States

[§]Division of Bioscience and Bioinformatics, College of Natural Science, Myongji University, Cheoin-gu, Yongin, Gyeonggi-Do 449-728, South Korea

[‖]Center for Nutraceutical and Pharmaceutical Materials, Myongji University, Cheoin-gu, Yongin, Gyeonggi-Do 449-728, South Korea

[⊥]Department of Cosmetic Science, Chungwoon University, Hongseong, Chungnam 350-701, South Korea

Abstract

The new tuberculosis (TB) lead ecumicin (**1**), a cyclic tridecapeptide, was isolated from *Nonomuraea sp.* MJM5123, following a high-throughput campaign for anti-TB activity. The large molecular weight of 1599 amu detected by LC-HR-MS precluded the initial inference of its molecular formula. The individual building blocks were identified by extensive NMR experiments. The resulting two possible planar structures were distinguished by LC-MS^2. Determination of absolute configuration and unambiguous structural confirmation were carried out by X-ray crystallography and Marfey's analysis.

Graphical abstract

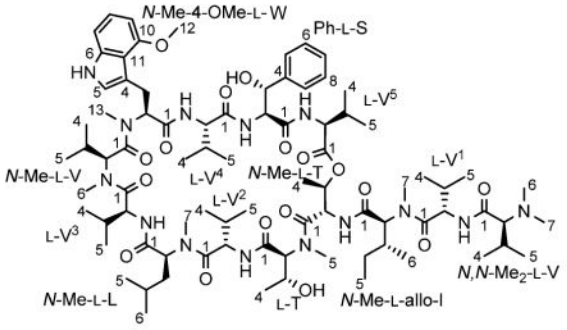
[*]Corresponding Authors: jwsuh@mju.ac.kr. gfp@uic.edu.

Notes

The authors declare no competing financial interest.

Supporting Information

Experimental details, spectral data, crystal structure (PDB), qHNMR (^{1}H) and ^{13}C FIDs (Bruker format, ZIP). This material is available free of charge via the Internet at http://pubs.acs.org.



To date, *Mycobacterium tuberculosis* (*Mtb*) remains the cause of one of the most deadly diseases, tuberculosis (TB). The treatment of multiple and extensively drug-resistant (MDR/XDR) TB and the shortening of the treatment duration remain difficult challenges, making new anti-TB drug discovery imperative.[1] While the intensive search for antibiotics from soil microorganisms in the mid-20th century yielded several clinically useful TB drugs, the pathogenic nature of *Mtb* and its very slow growth rate did not allow classical agar diffusion tests and excluded *Mtb* from the initial target panel. Thus, the discovery of TB drugs has relied upon a broader spectrum of activity. Following the current understanding that *Mtb* is uniquely susceptible to a number of antimicrobial agents, there is the distinct possibility that novel anti-TB agents were not detected during the heyday of actinomycete screening. Therefore, we conducted a high-throughput screening of over 65 000 actinomycete extracts directly against the virulent H_{37}Rv strain. With early stage biological dereplication methods applied,[2] this campaign prioritized 22 strains, among which was *Nonomuraea sp.* MJM5123 whose mycelial methanolic extract contains agents with selective anti-*Mtb* activity.

A large-scale mycelial methanolic extract of strain MJM5123 was obtained from a 20 L fermenter culture. The extract underwent a chemical fractionation process in parallel with biological characterization using Microplate Alamar Blue[3] and Vero cell toxicity assays[4,5] to monitor the anti-TB activity and selectivity index (SI = IC_{50}/MIC), respectively.[6] Isolation of active constituents involved three sequential steps of orthogonal separation methods: adsorption, gel permeation, and countercurrent chromatography. Vacuum liquid chromatography of the extract (128.7 g) on POLYGOPREP 100–50 C_{18} silica gel using a H_2O/MeOH gradient in 20% steps yielded seven chemically distinct fractions, VC1–7. Fractions VC6 and 7 eluting with neat MeOH and $CHCl_3$ exhibited MICs of <0.76 and 0.74 μg/mL, and SIs of >66 and 68, respectively. After recombination (8.47 g), gel permeation on a Sephadex LH-20 open column with 100% MeOH as eluent resulted in 81 fractions, which upon TLC monitoring yielded a panel of 11 recombined fractions, S1 to 11. The subfractions S2 and S3 both showed MICs <0.21 μg/mL and SIs >238. Using high speed countercurrent chromatography (HSCCC), S2 (177 mg) was separated into 140 fractions with HEMWat +2, which is a mixture of hexanes, EtOAc, MeOH, and H_2O in the ratio 3:7:5:5 (v/v). Selection of HEMWat +2 as the most suitable two-phase solvent system was based on S2 having a partition coefficient (w/w) close to 1 (0.926). Upon TLC and ^{1}H NMR analysis, the 140 fractions were recombined to a panel of 26 fractions, H1 to H26. The MICs of subfractions H3, H5, H7, and H9 to H23 were <0.50 μg/mL, and their SIs >100. Using the same solvent system, S3 (374 mg) was also separated by HSCCC into 110 fractions and

recombined to a panel of 11 fractions, H′1 toH′11. The MICs were <0.39 μg/mL and the SIs were >128 for fractions H′4, H′6, H′8, and H′9.

Among the most active fractions, H14 was the purest (91%) by qHNMR (100% method)[7] and, therefore, selected for structural elucidation. This led to the discovery of a new macrocyclic tridecapeptide, designated as ecumicin (**1**). Retesting confirmed that ecumicin has an MIC against replicating *Mtb* of 0.26 μg/mL and a selectivity index of >320. **1** is equally active against isogenic *Mtb* strains that are resistant to single actinomycete-derived anti-TB drugs, including rifampin, streptomycin, and cycloserine.[2] The biological profile summarized in Table 1 indicated that **1** was highly likely to be a new anti-TB agent.

High resolution ESI-MS produced pseudomolecular ions at m/z 1600.0189 $[M + H]^+$ and 1621.9982 $[M + Na]^+$, which indicated an exact mass of 1599.0111. Such a large exact mass precluded a reliable inference of the molecular formula. The IR spectrum suggested the presence of multiple amide moieties (1632 cm^{-1}), and the UV spectrum indicated that **1** is being a peptide with aromatic residues, having absorption bands at 211, 219, 263, 281, and 291 nm. Despite the relatively large molecular weight, which in turn indicates large numbers of protons and carbons, the NMR spectra of **1** in CD_3OD showed relatively good signal dispersion and, therefore, were rich in structural information. The ^1H NMR spectrum in CD_3OH displayed eight exchangeable proton signals between δ_H 10.30 and 7.72 that vanished upon exchange in fully deuterated CD_3OD. In addition, signals occurred in the aromatic region coding for one phenyl group and another aromatic system. Moreover, doublets and triplets were found between δ_H 0.16 to 1.33 integrating for 20 aliphatic methyl groups. Surprisingly for the case of a peptide, singlets integrating for 3H at δ_H 2.16, 3.14, 3.23, 3.26, 3.33, and 3.82, and 6H at δ_H 2.31, appeared, indicating eight *N*-methyl and/or methoxyl groups. The signal distribution with doublets in the amide region and signals covering the entire range between δ_H 0.16 and 5.4 was a clear indication that **1** is a peptide indeed.

Further spectroscopic evidence supporting the classification of **1** as a peptide was derived from the ^{13}C NMR spectrum, which revealed a total of 12 (amide-) carbonyl signals integrating for 13 carbons in the range between δ_C 170.92 and 175.14. By using DEPT-135 and HSQC experiments, 34 methine, 3 methylene, and 28 methyl signals were identified. These signals, along with five quaternary carbon signals at δ_C 155.23, 142.68, 140.04, 118.36, and 112.47 added up to a total of 83 carbons present in **1**. Extensive analysis of the 2D NMR spectra, particularly COSY, TOCSY, HSQC, and HMBC, resulted in the elucidation of 15 discrete ^1H,^1H spin systems belonging to 13 amino acid residues: seven valines, two threonines, one leucine, one isoleucine, one 4-methoxy-tryptophan, and one phenylserine.

As conventional HSQC and/or HMBC experiments cannot resolve the crowded aliphatic methyl region of **1** with 20 ^{13}C signals presented in a $\Delta\delta_C$ 1.17 window, as well as the carbonyl region with 12 ^{13}C signals presented in a $\Delta\delta_C$ 4.22 window, semiselective HMBC was employed, which yielded high-resolution in the indirect ^{13}C dimension by suppressing homonuclear proton coupling modulations.[8] The ^1H signals of aliphatic methyl groups were attached to corresponding ^{13}C signals via HSQC-type cross peaks (^1J~126 Hz) observed on

one semiselective HMBC spectrum. The individual amino acid residues were linked sequentially via correlations between the carbonyl carbon and the following protons: α-protons (^2J), β-protons (^3J), neighboring α-protons (^3J), and/or neighboring *N*-methyl protons (^3J). The observation of HMBC correlation between the β-proton of a threonine residue and the carbonyl carbon of a valine residue suggested **1** to be a depsipeptide cyclized between the C-terminal carboxyl and the side chain hydroxyl group of the threonine. As two carbonyl carbons resonated at precisely the same frequency, the NMRexperiments left two structural possibilities (Figure S1, Supporting Information (SI)), both containing five valines ($V^{1–5}$), one *N*-methyl-valine (*N*-Me-V), one *N,N*-dimethyl-valine (*N,N*-Me_2-V), one threonine (T), one *N*-methyl-threonine (*N*-Me-T), one *N*-methyl-leucine (*N*-Me-L), one *N*-methyl-isoleucine (*N*-Me-I), one *N*-methyl-4-methoxy-tryptophan (*N*-Me-4-OMe-W), and one phenylserine (Ph-S). Both possible structures exhibit the same molecular formula, $C_{83}H_{134}N_{14}O_{17}$, which calculates for an exact mass of 1599.0051 amu and is consistent with the observed value. The key COSY and HMBC correlations are presented in Figure 1, and the detailed ^1H and ^{13}C assignments are given in Table 2.

The remaining ambiguity in the planar structure of **1** was resolved by the following fragments in the tandem mass experiment (Figures S2 and S16, SI), which supported the structure of **1** shown below: *m*/*z* (rel. Int.) 1600.23 $[M + H]^+$ (8), 1246.95 [M − *N,N*-Me_2-V − V^1 − *N*-Me-I + 2H]$^+$ (48), 990.86 [M − *N,N*-Me_2-V − V^1 − *N*-Me-I − T − V^5 + 2H]$^+$ (7), 800.60 $[M + 2H]^{2+}$ (50), 610.41 [*N,N*-Me_2-V + V^1 + *N*-Me-I + T + Val^5]$^+$ (36), 354.32 [*N,N*-Me_2-V + Val^1 + *N*-Me-I]$^+$ (87).

Colorless needle-like crystals of **1** were obtained by slow evaporation from a solvent composed of MeOH, MeCN, and H_2O (2:2:1, v/v). X-ray crystallography with these crystals added unambiguous support to the proposed structure. As shown in Figure S18, **1** possesses a twisted β-hairpin structure stabilized by five intermolecular hydrogen bonds. This is congruent with the result drawn from the CD spectrum, which also suggested **1** to be a peptide with an antiparallel β-sheet structure. As the hydrophobic side chains of alternate residues point away from the β-hairpin backbone, the surface of ecumicin is hydrophobic and neutral. A plot of the *B*-factors shows that the tail residues are more flexible than those located in the cyclic core. In addition, n $\rightarrow$ π^* interactions, an electronic delocalization effect analogous to the hydrogen bond,[9] further stabilizes the secondary structure of **1**. Two observed n $\rightarrow$ π^* interactions (Table 2) reside in the turning site of the β-sheet, contributing to the stabilization of the turn. The *B*-factor putty tube representation of **1** (Figure S19) shows that its structure is mostly ordered, with only the two *N*-terminal residues showing overall conformational flexibility. Cross-ring ROESY correlations confirmed the macrocyclic nature of ecumicin and support that the two *N*-terminal amino acids are located in the most flexible portion of the molecule.

The X-ray crystallographic data also suggest that, except for *N,N*-Me_2-V, all amino acid residues exhibit the same configuration. As the crystallographic data were unable to distinguish between right-handed and left-handed structures, all-L- and all-D-isomers were both possible. The absolute configuration of *N,N*-Me_2-V could not be determined at this stage, as the distinction was complicated by the close similarity of the electron density maps of the corresponding dimethyl amino and isopropyl groups. However, a Marfey

experiment[10] identified distinctive L-V and *N*-Me-L-V residues. Combining information from both X-ray and Marfey's analysis, it was determine that **1** contains at least 12 L-amino acids. Only the chirality of *N,N*-Me_2-V remained unclear, as this amino acid does not react with Marfey's reagent. However, the lack of an epimerization domain in the nonribosomal peptide synthetase of **1** (detailed data will be published separately) suggests that indeed all amino acid residues, including *N,N*-Me_2-V, possess L-configuration.

The discovery of **1** demonstrates that new potent anti-TB secondary metabolites may be uncovered from actinomycetes by direct whole-cell based high-throughput screening. An orthogonal set of fractionation methods and an extensive panel of analytical methods led to the structural characterization of this relatively high-MW natural product. The ability of **1** to maintain its high potency against *Mtb* H_{37}Rv-isogenic strains that are monoresistant to rifampin, streptomycin and cycloserine (Table 1), suggests that **1** has a different molecular target from these drugs. Genome mining of lab-generated, spontaneous ecumicin-resistant *Mtb* identified the ClpC1 ATPase complex as the putative target, and this was confirmed by a drug affinity response test.[2] Complete inhibition of *Mtb* growth by **1** was achieved in a mouse model of acute TB infection, making **1** the only TB drug lead among three known compounds targeting *Mtb* ClpC1.[2]

While the present report was in preparation, Um et al. noted the structural similarities between **1**[11] and two cyclic dodecapeptides, ohmyungsamycins A/B,[1213] which were obtained from a *Streptomyces* species and studied from a chemical perspective, including the determination of absolute configuration. Their biological characterization as antimicrobial and cytotoxic agents, albeit limited to primary bioassays, indicates a potential overlap with **1** as a TB drug lead. However, as high anti-TB selectivity and no general toxicity have been observed for **1**, it is likely that the structural differences between **1** and the ohmyungsamycins are sufficient to produce diverging bioactivities.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. B. Ramirez, UIC Center for Structural Biology, for NMR support, and Drs. María F. Rodríguez Brasco, Chang-Hwa Hwang, and Larry L. Klein, all from the University of Illinois at Chicago, for helpful discussions. This work was partially supported by Grant No. PJ009643 to J.-W. Suh from the Next-Generation BioGreen 21 Program, Rural Development Administration, Republic of Korea. Funding for the 900 MHz NMR spectrometer was provided by NIH (Grant P41 GM068944) awarded to Dr. P. G. W. Gettins by NIGMS and is gratefully acknowledged.

References

1. Lienhardt C, Raviglione M, Spigelman M, Hafner R, Jaramillo E, Hoelscher M, Zumla A, Gheuens J. J Infect Dis. 2012; 205(Suppl 2):S241–S249. [PubMed: 22448022]
2. Gao W, Kim JY, Anderson JR, Akopian T, Hong S, Jin JY, Kandror O, Kim JW, Lee IA, Lee SY, McAlpine JB, Mulugeta S, Sunoqrot S, Wang Y, Yang SH, Yoon TM, Goldberg AL, Pauli GF, Suh JW, Franzblau SG, Cho S. Antimicrob Agents Chemother. 2014; doi: 10.1128/AAC.04054-14
3. Collins L, Franzblau S. Antimicrob Agents Chemother. 1997; 41:1004–1009. [PubMed: 9145860]

4. Falzari K, Zhou Z, Pan D, Liu H, Hongmanee P, Franzblau SG. Antimicrob Agents Chemother. 2005; 49:1447–1454. [PubMed: 15793125]
5. Mangalindan GC, Talaue MT, Cruz LJ, Franzblau SG, Adams LB, Richardson AD, Ireland CM, Concepcion GP. Planta Med. 2000; 66:364–365. [PubMed: 10865457]
6. Cantrell CL, Lu T, Fronczek FR, Fischer NH, Adams LB, Franzblau SG. J Nat Prod. 1996; 59:1131–1136. [PubMed: 8988597]
7. Pauli GF, Jaki BU, Gödecke T, Lankin DC. J Nat Prod. 2012; 75:834–851. [PubMed: 22482996]
8. Claridge TDW, Perez-Victoria I. Org Biomol Chem. 2003; 1:3632–3634. [PubMed: 14649890]
9. Bartlett GJ, Choudhary A, Raines RT, Woolfson DN. Nat Chem Biol. 2010; 6:615–620. [PubMed: 20622857]
10. Bhushan R, Bruckner H. J Chromatogr B. 2011; 879:3148–3161.
11. Cho, SH., Choi, JK., Franzblau, SG., Friesen, JB., Gao, W., Jaki, BU., Jin, YY., Kim, JY., Kim, JW., Lankin, DC. Cyclic peptide from *Nonomuraea sp.*, process for the production thereof, and pharmaceutical composition for the prevention or treatment of mycobacteria related disease comprising the same. PCT Patent. WO 2012/144790.
12. Um S, Choi TJ, Kim H, Kim BY, Kim SH, Lee SK, Oh KB, Shin J, Oh DC. J Org Chem. 2013; 78:12321–12329. [PubMed: 24266328]
13. Crystallographic data for ecumicin reported in this paper have been deposited at the Cambridge crystallographic Data Centre (CCDC) under the deposition number 940680. These data can be obtained free of charge via http://www.ccdc.cam.ac.uk/data_request/cif, or by e-mailing data_request@ccdc.cam.ac.uk, or by CCDC, 12 Union Road, Cambridge CB21EZ, UK, fax: +44(0)-1233-336033.

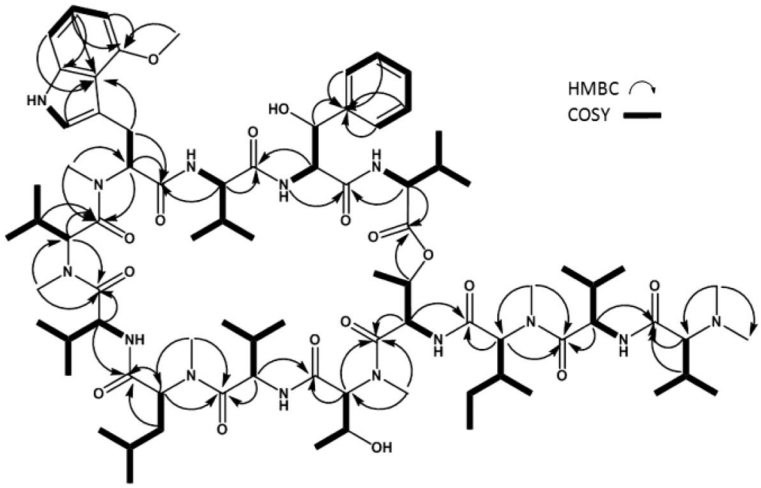


Figure 1.
Key COSY and HMBC correlations for ecumicin.

Table 1

Ecumicin Is Toxic for *Mtb*, but Nontoxic for a Mammalian Cell Line (VERO Cells)[2]

			ecumicin (1)
MIC (μg/mL) vs *Mtb*	H37Rv resistant to:		0.26
		SM[a]	<0.20
		RMP[b]	0.30
		CS[c]	<0.20
IC50 (μg/mL) vs VERO cells			>50

[a]SM, streptomycin.

[b]RMP, rifampin.

[c]CS, cycloserine.

Table 2

^{1}H (900 MHz) and ^{13}C (100 MHz) NMR Data for Ecumicin in CD_3OD (CD_3OH for NH Protons)

no.	δ_H, mult (*J* in Hz)	δ_C	no.	δ_H, mult (*J* in Hz)	δ_C	no.	δ_H, mult (*J* in Hz)	δ_C	no.	δ_H, mult (*J* in Hz)	δ_C
N,N-Me_2-L-V			3	5.783 (dd, 2.3, 6.5)	69.82	L-V[3]			12	3.826(s)	55.59
1		173.28	4	1.313 (d, 6.5)	16.76	1		173.39	13	2.157 (s)	40.90
2	2.672(d, 9.17)	75.61	NH	8.50 (d, 9.1)		2	4.594 (d, 8.9)	55.34	NH	10.22 (d, 2.0)	
3	2.042 (m)	28.68				3	2.027 (m)	32.88			
4	0.847 (d, 6.6)	19.35	*N*-Me-L-T			4	0.915 (d, 6.6)	19.66	L-V[4]		
5	0.978 (d, 6.6)	20.05	1		171.21	5	0.859 (d, 6.8)	19.21	1		174.17
6, 7	2.313 (s)	42.26	2	5.018 (d, 3.7)	63.00	NH	9.07 (d, 9.5)		2	4.529 (d, 7.9)	59.21
			3	4.457 (dd, 3.7, 6.5)	66.93				3	2.200 (m)	33.70
L-V[1]			4	0.912 (d, 6.5)	19.91	*N*-Me-L-V			4	1.028 (d, 6.8)	19.44
1		174.99	5	3.330 (s)	34.30	1		170.92	5	0.989 (d, 6.8)	19.97
2	4.669 (d, 8.8)	55.91				2	3.065 (d, 7.6)	71.37	NH	7.85 (d, 9.6)	
3	2.101 (m)	31.62	L-V[2]			3	2.583 (m)	29.94			
4	0.992 (d, 6.8)	19.62	1		174.73	4	0.976 (d, 6.8)	19.76	Ph-L-*threo*-S		
5	1.055 (d, 6.7)	19.36	2	4.842 (d, 8.9)	56.77	5	1.094 (d, 6.5)	22.02	1		173.76
NH	8.17 (d, 8.2)		3	2.350 (m)	31.76	6	3.140 (s)	40.37	2	4.854 (d, 1.9)	59.78
			4	1.092 (d, 6.7)	19.64				3	5.344 (d, 1.9)	72.97
N-Me-L-*allo*-I			5	0.979 (d, 7.0)	19.70	*N*-Me-4-OMe-L-W			4		142.68
1		172.60	NH	7.73 (d, 8.6)		1		171.61	5, 9	7.240 (br s)	127.06
2	4.918 (d, 11.2)	61.76				2	4.102 (dd, 11.1, 4.7)	71.11	6, 8	7.255 (br s)	129.28
3	1.957 (m)	34.33	*N*-Me-L-L			3	3.545 (dd, 11.1, −13.7)	26.97	7	7.201 (br s)	128.25
4	0.989 (m)	26.61	1		173.28		3.694 (dd, 4.7, −13.7)		NH	8.30 (d, 7.8)	
	1.255 (m)		2	5.110 (dd, 6.2, 8.4)	55.59	4		112.47			
5	0.753 (dd, 7.3, 7.6)	11.50	3	1.239 (m)	39.37	5	6.698 (s)	124.00	L-V[5]		
6	0.743 (d, 6.6)	15.10		1.452 (m)		6		140.04	1		175.13
7	3.233 (s)	31.40	4	0.957 (m)	25.64	7	6.918 (d, 8.18)	106.01	2	4.396 (d, 8.9)	59.26
			5	0.168 (d, 6.6)	21.67	8	6.982 (dd, 8.18, 7.80)	123.30	3	1.968 (m,)	33.19
L-T			6	0.331 (d, 6.6)	23.47	9	6.442 (d, 7.80)	99.77	4	0.919 (d, 6.4)	19.10

no.	δ_H, mult (*J* in Hz)	δ_C	no.	δ_H, mult (*J* in Hz)	δ_C	no.	δ_H, mult (*J* in Hz)	δ_C	no.	δ_H, mult (*J* in Hz)	δ_C
1		171.95	7	3.259 (s)	31.63	10		155.23	5	0.935 (d, 6.8)	19.44
2	5.167 (d, 2.3)	53.39				11		118.36	NH	9.02 (d, 9.9)	