Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Supporting Information

Total Syntheses of the Bilirubin Oxidation End Product *Z*-BOX C and its Isomeric Form *Z*-BOX D

Daniel Schulze, Juliane Traber, Marcel Ritter, Helmar Görls, Georg Pohnert, and Matthias Westerhausen

Table of Contents

I. NMR spectra

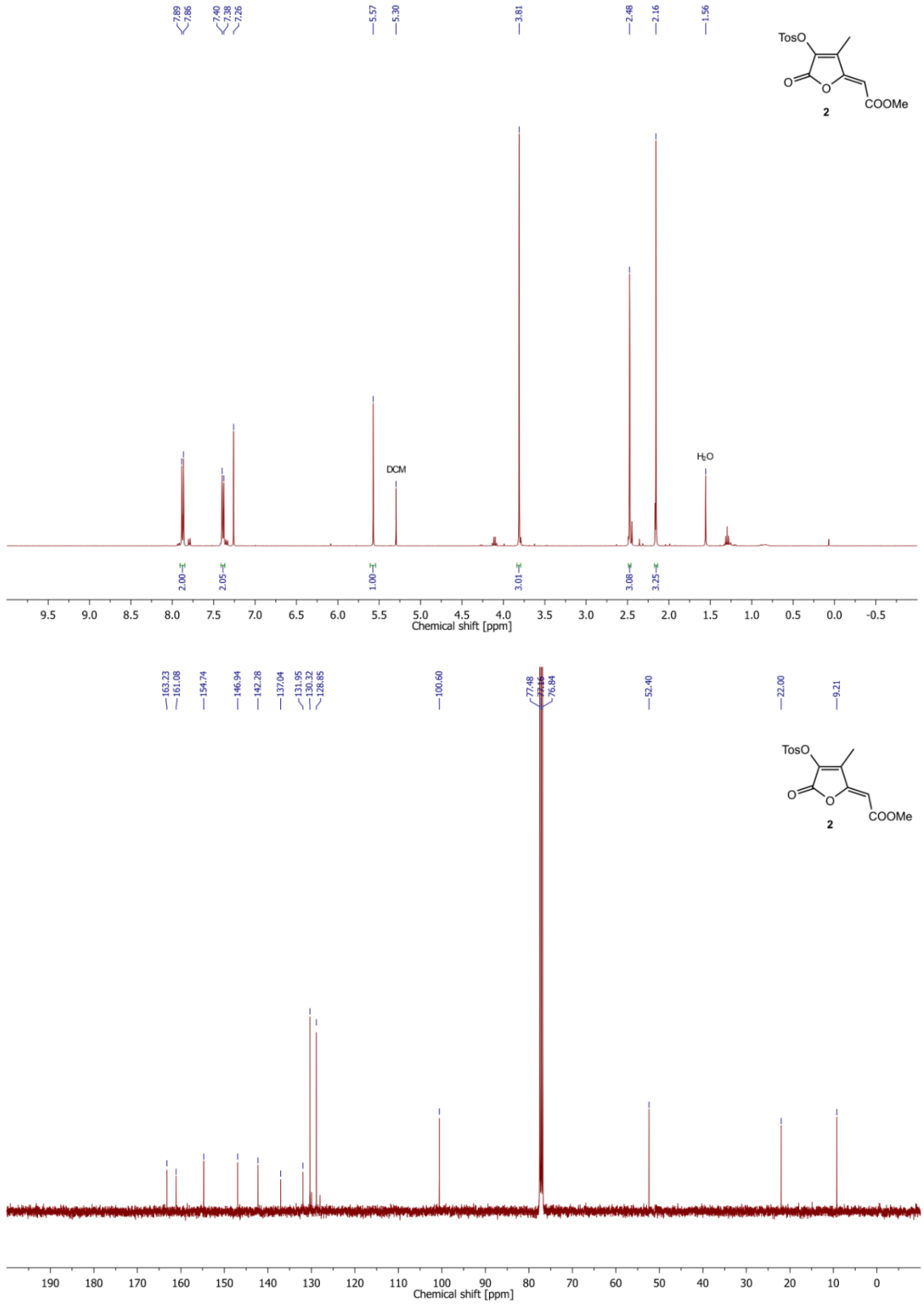

Figure S1. ^{1}H NMR spectrum (400 MHz, $CDCl_3$, 297 K, top) and $^{13}C\{^{1}H\}$ NMR spectrum (101 MHz, $CDCl_3$, 297 K, bottom) of **2**.

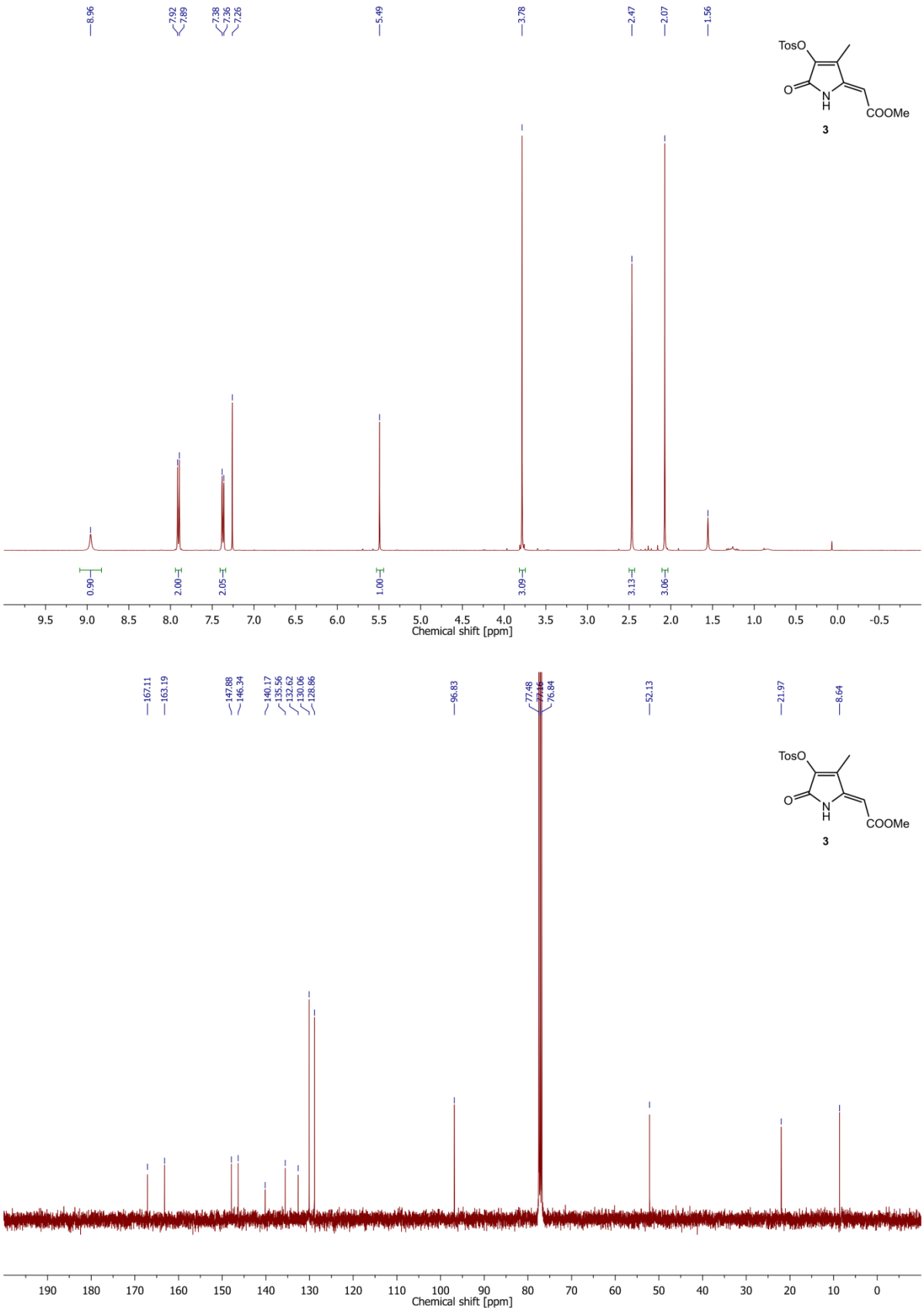

Figure S2. ^{1}H NMR spectrum (400 MHz, $CDCl_3$, 297 K, top) and $^{13}C\{^{1}H\}$ NMR spectrum (101 MHz, $CDCl_3$, 297 K, bottom) of **3**.

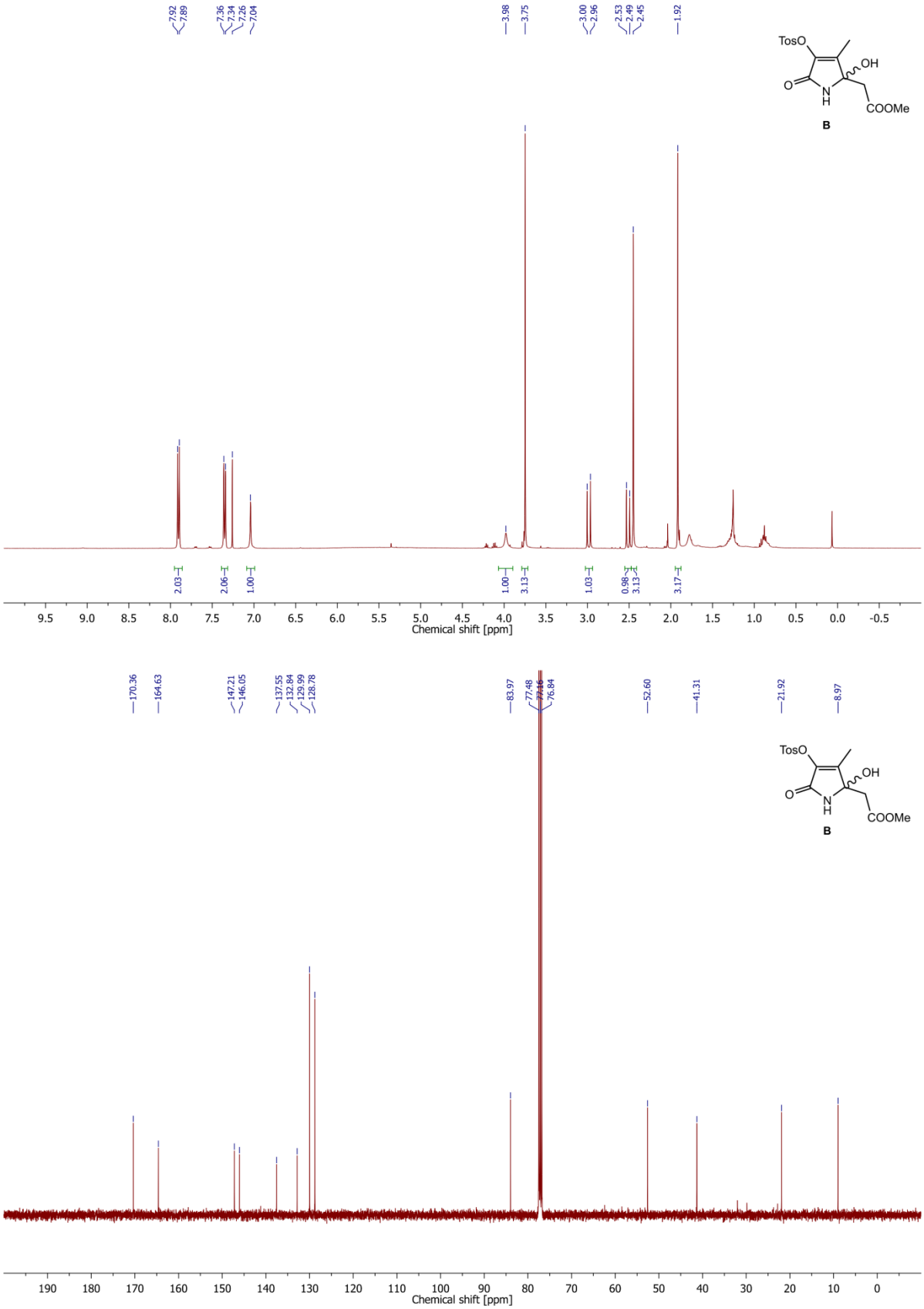

Figure S3. ^{1}H NMR spectrum (400 MHz, $CDCl_3$, 297 K, top) and $^{13}C\{^{1}H\}$ NMR spectrum (101 MHz, $CDCl_3$, 297 K, bottom) of **B**.

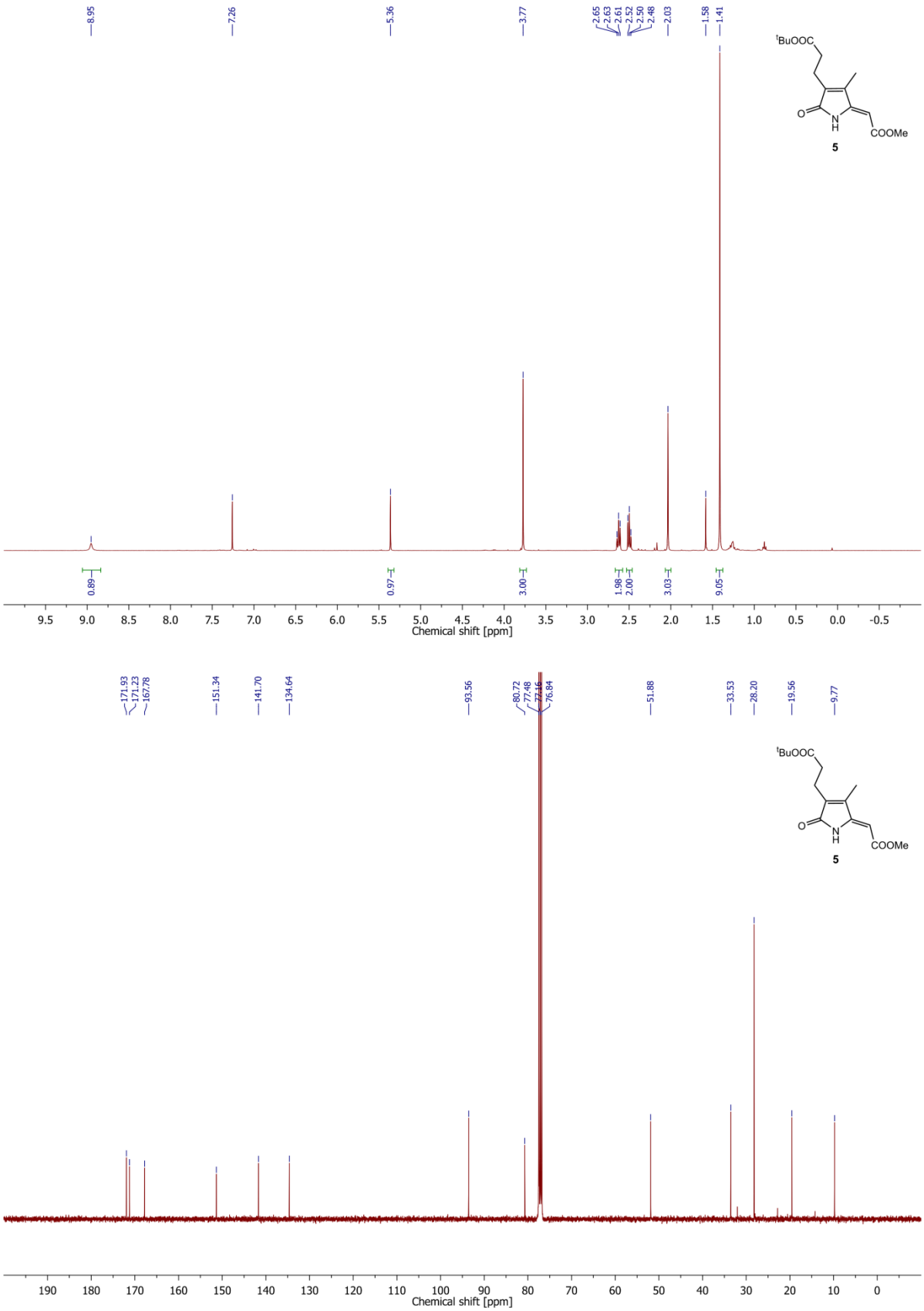

Figure S4. ^{1}H NMR spectrum (400 MHz, $CDCl_3$, 297 K, top) and $^{13}C\{^{1}H\}$ NMR spectrum (101 MHz, $CDCl_3$, 297 K, bottom) of **5**.

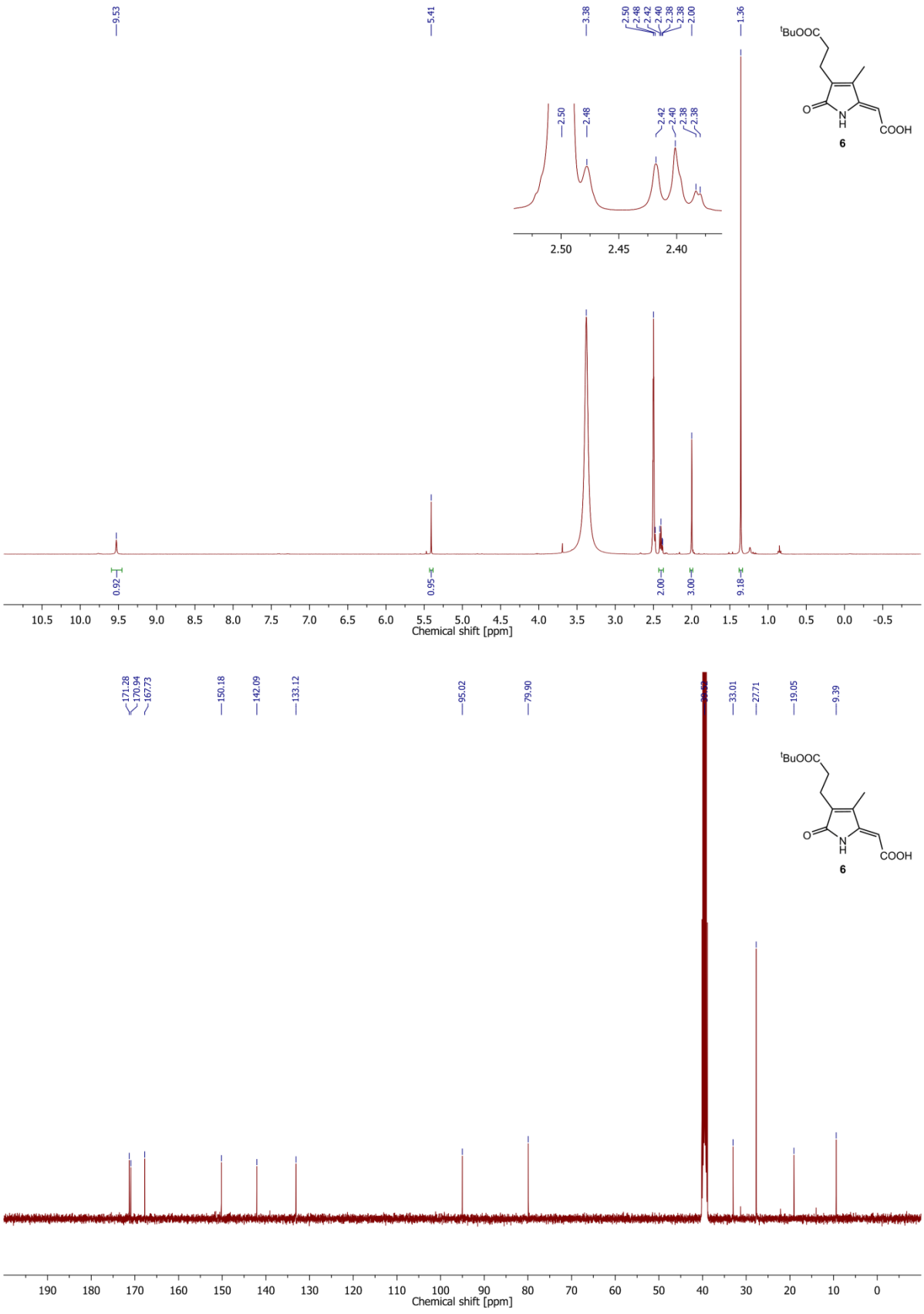

Figure S5. ^{1}H NMR spectrum (400 MHz, DMSO-d_6, 297 K, top) and ^{13}C{^{1}H} NMR spectrum (101 MHz, DMSO-d_6, 297 K, bottom) of **6**.

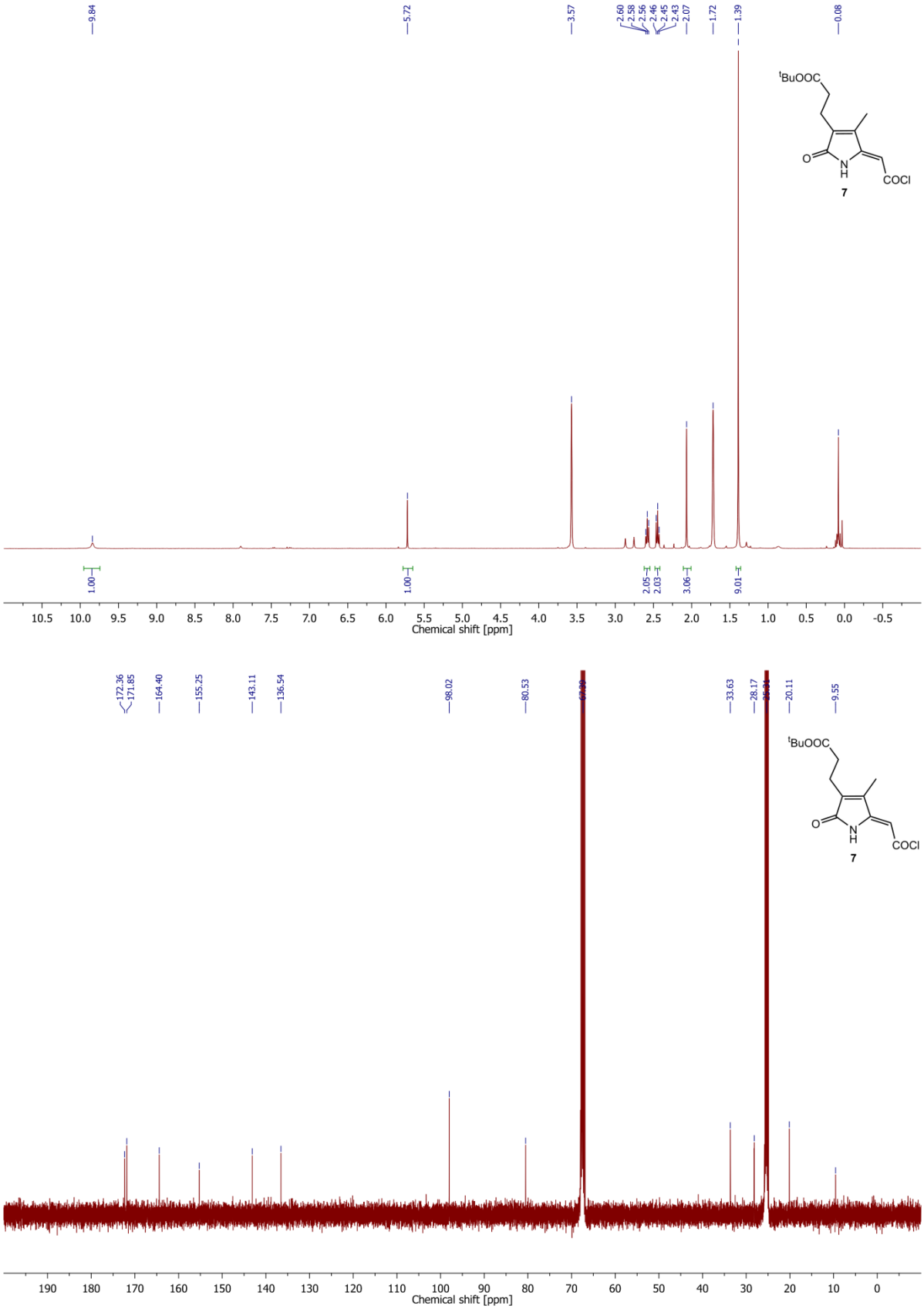

Figure S6. ^{1}H NMR spectrum (400 MHz, THF-d_8, 297 K, top) and ^{13}C{^{1}H} NMR spectrum (101 MHz, THF-d_8, 297 K, bottom) of **7**.

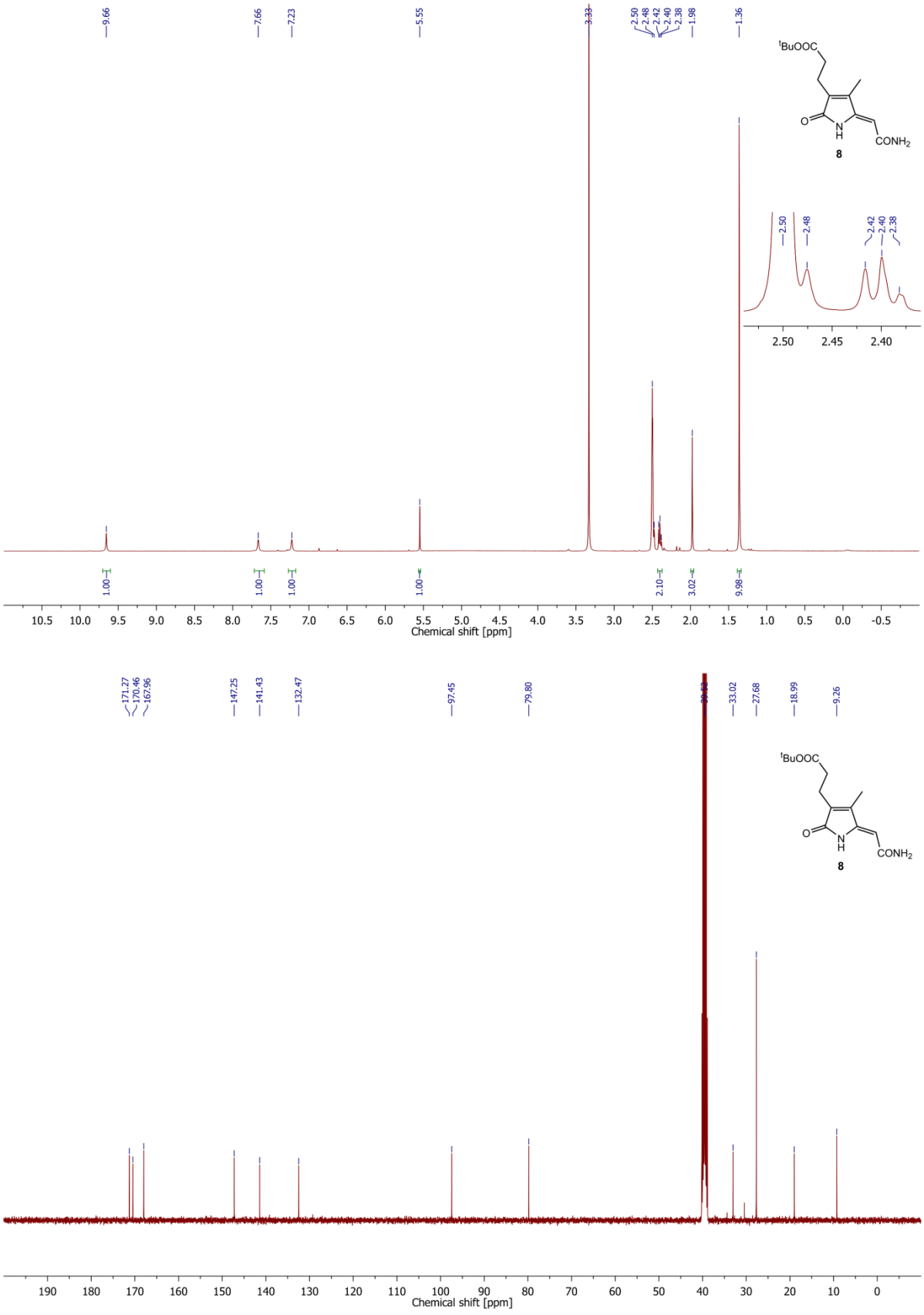

Figure S7. ^{1}H NMR spectrum (400 MHz, DMSO-d_6, 297 Ktop) and ^{13}C{^{1}H} NMR spectrum (101 MHz, DMSO-d_6, 297 K, bottom) of **8**.

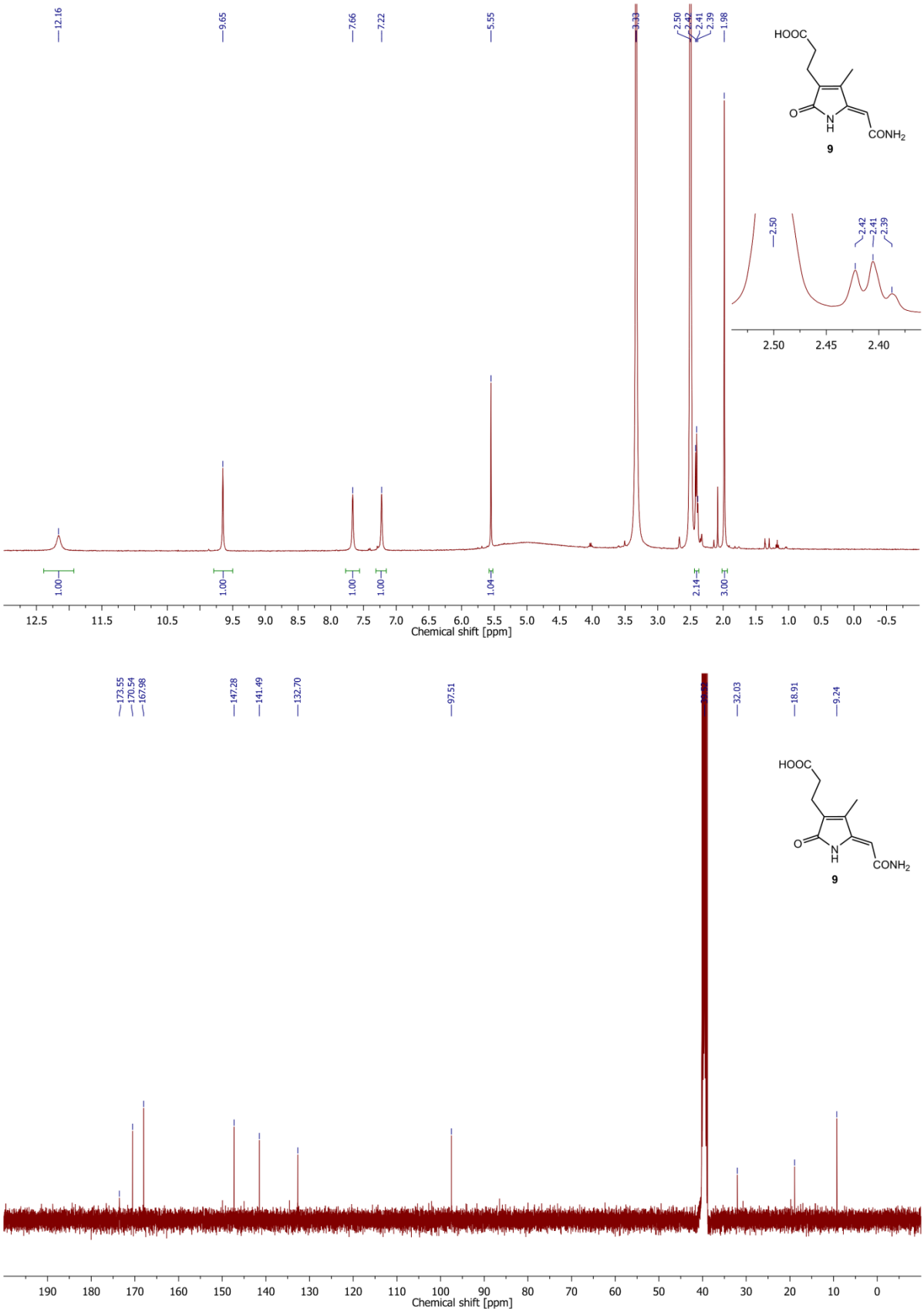

Figure S8. ^{1}H NMR spectrum (400 MHz, DMSO-d_6, 297 K, top) and ^{13}C{^{1}H} NMR spectrum (101 MHz, DMSO-d_6, 297 K, bottom) of **9**.

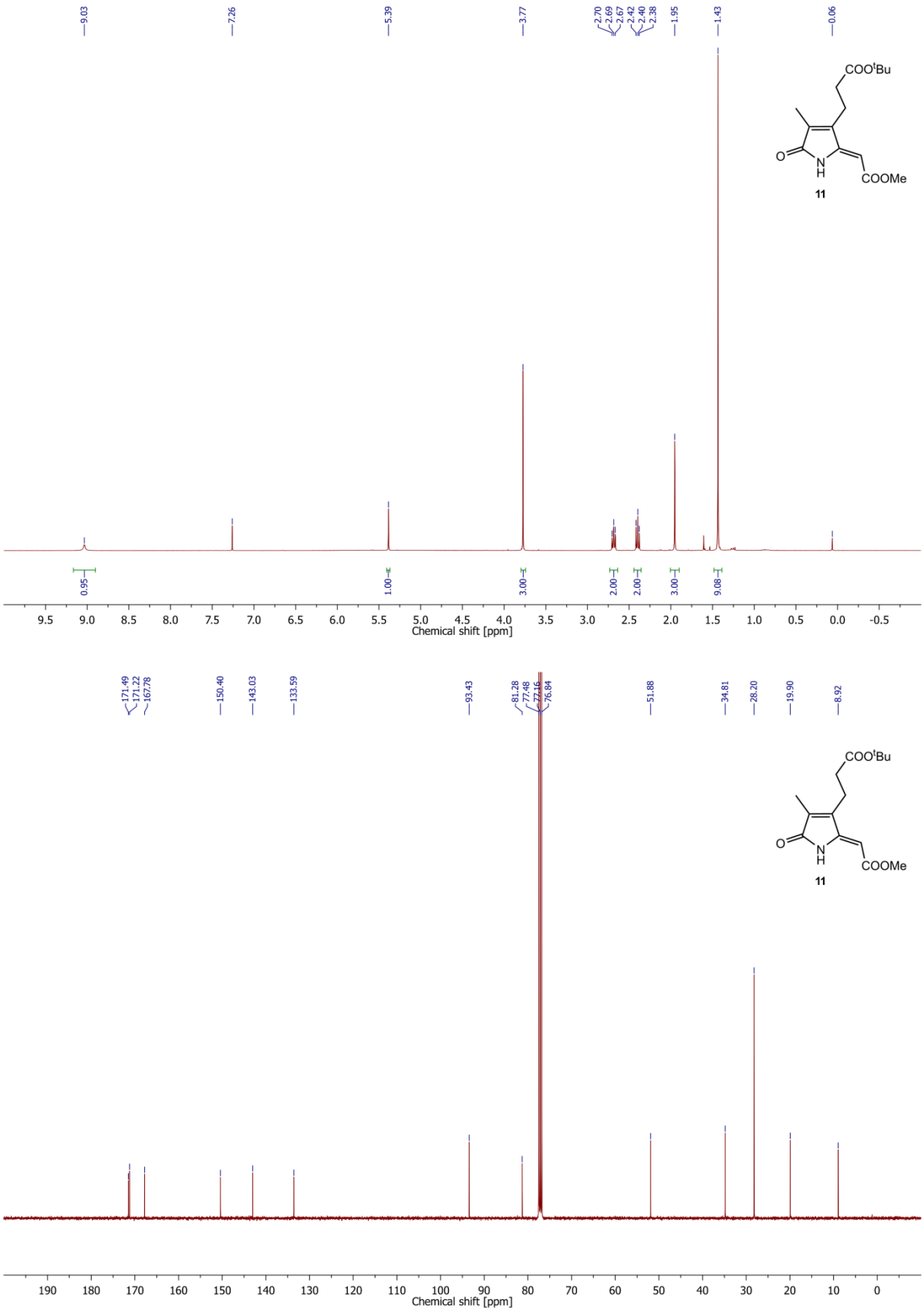

Figure S9. ^{1}H NMR spectrum (400 MHz, $CDCl_3$, 297 K, top) and $^{13}C\{^{1}H\}$ NMR spectrum (101 MHz, $CDCl_3$, 297 K, bottom) of **11**.

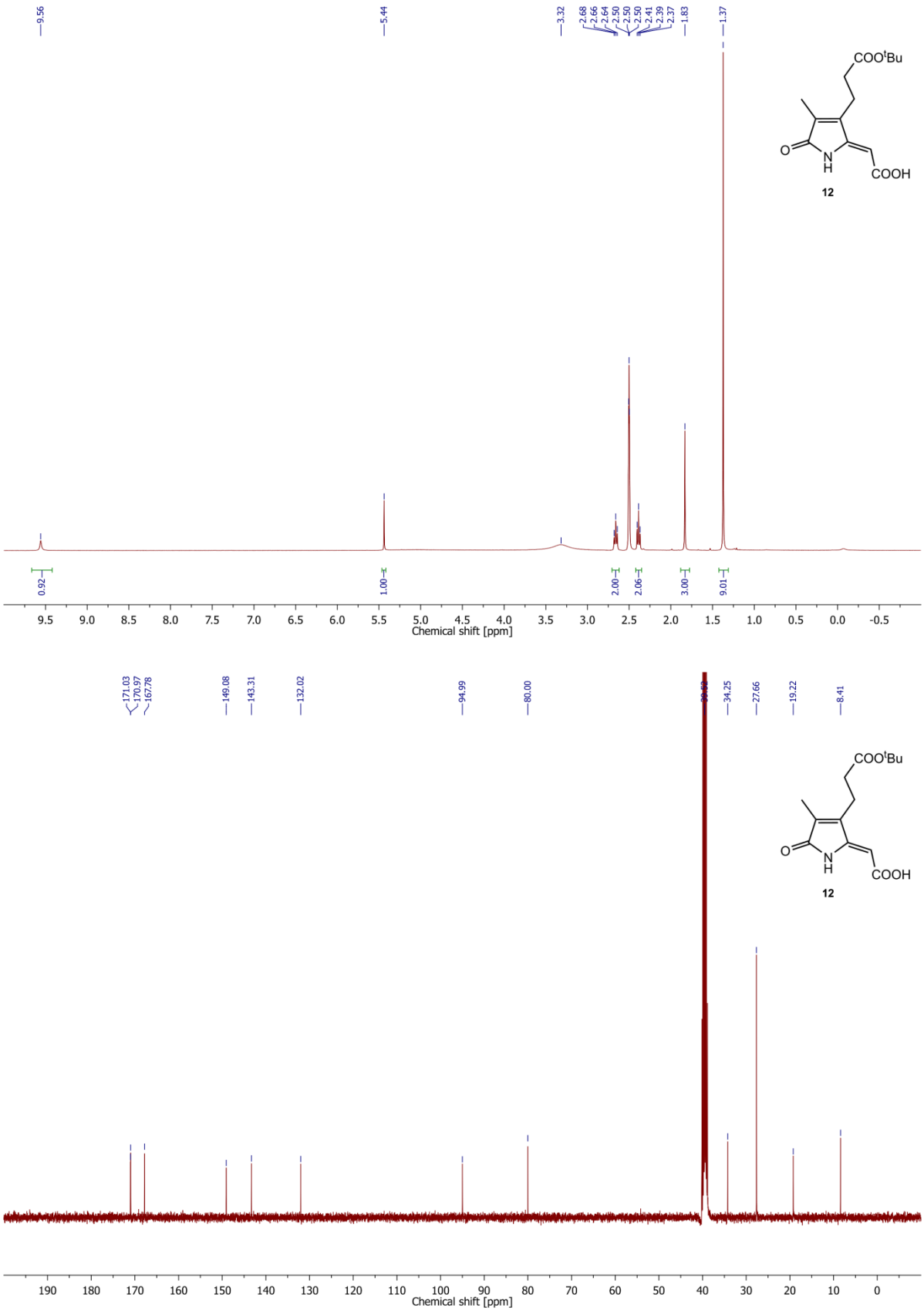

Figure S10. ^{1}H NMR spectrum (400 MHz, DMSO-d_6, 297 K, top) and ^{13}C{^{1}H} NMR spectrum (101 MHz, DMSO-d_6, 297 K, bottom) of **12**.

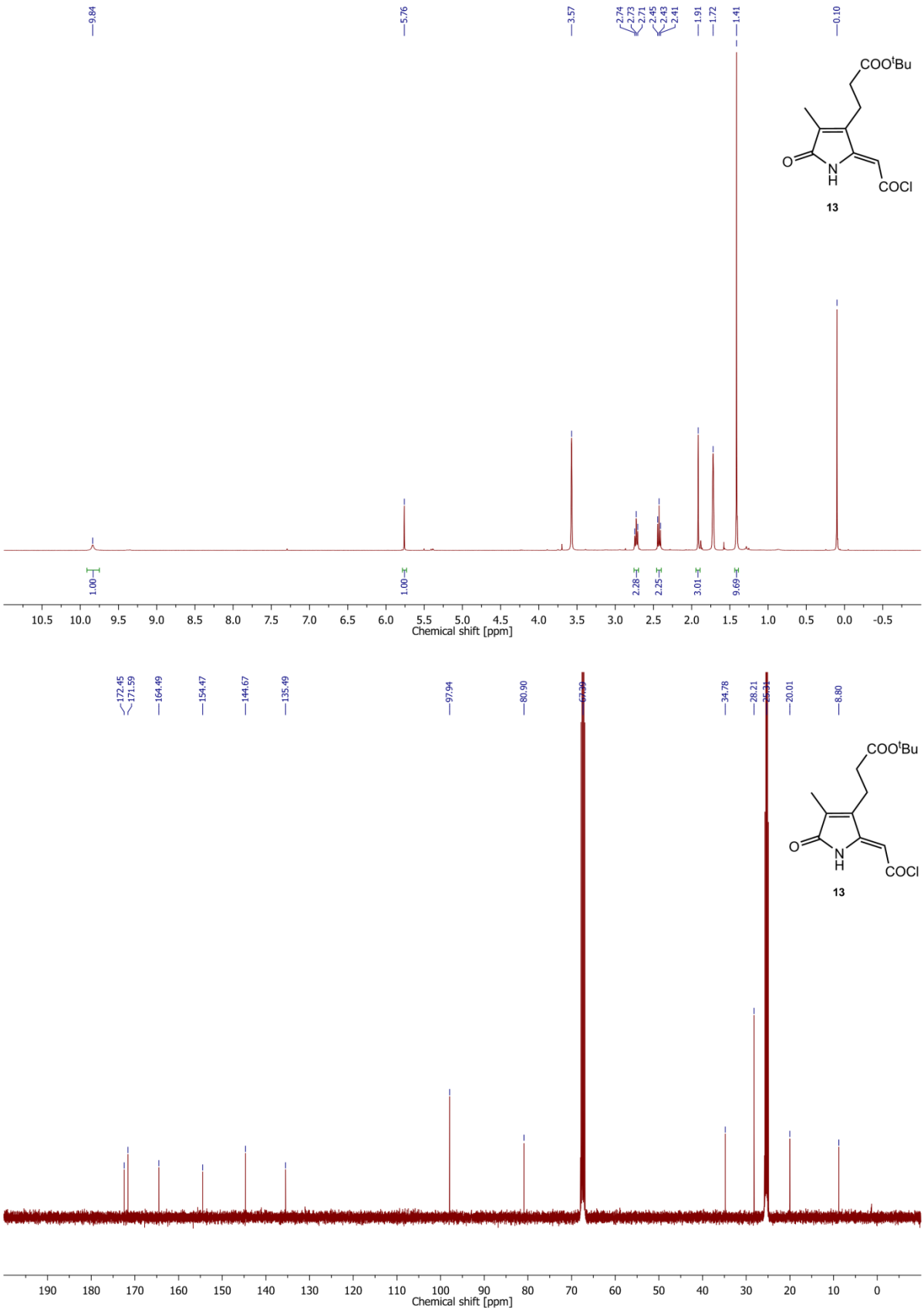

Figure S11. ^{1}H NMR spectrum (400 MHz, THF-d_8, 296 K, top) and ^{13}C{^{1}H} NMR spectrum (101 MHz, THF-d_8, 296 K, bottom) of **13**.

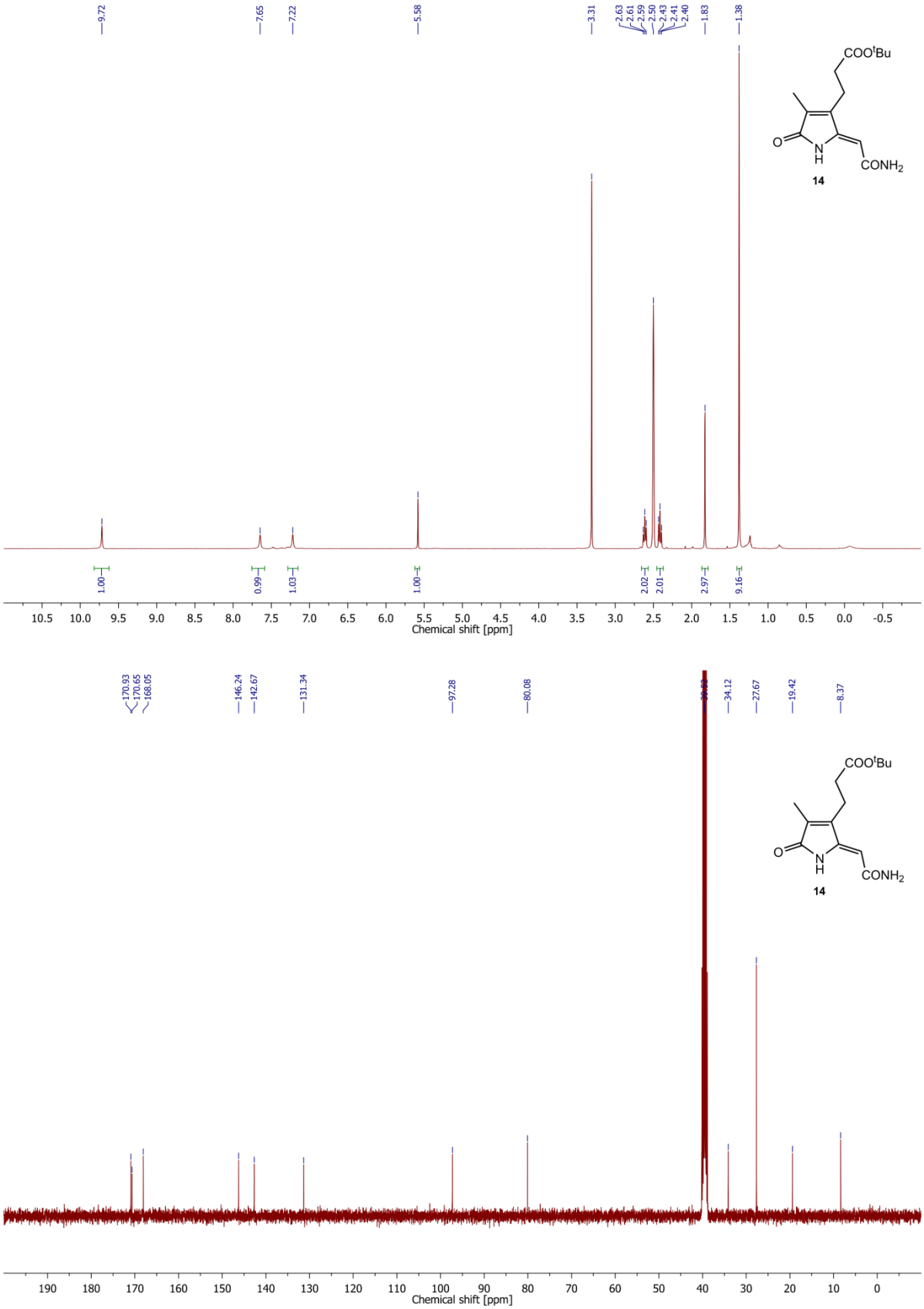

Figure S12. ^{1}H NMR spectrum (400 MHz, DMSO-d_6, 297 K, top) and ^{13}C{^{1}H} NMR spectrum (101 MHz, DMSO-d_6, 297 K, bottom) of **14**.

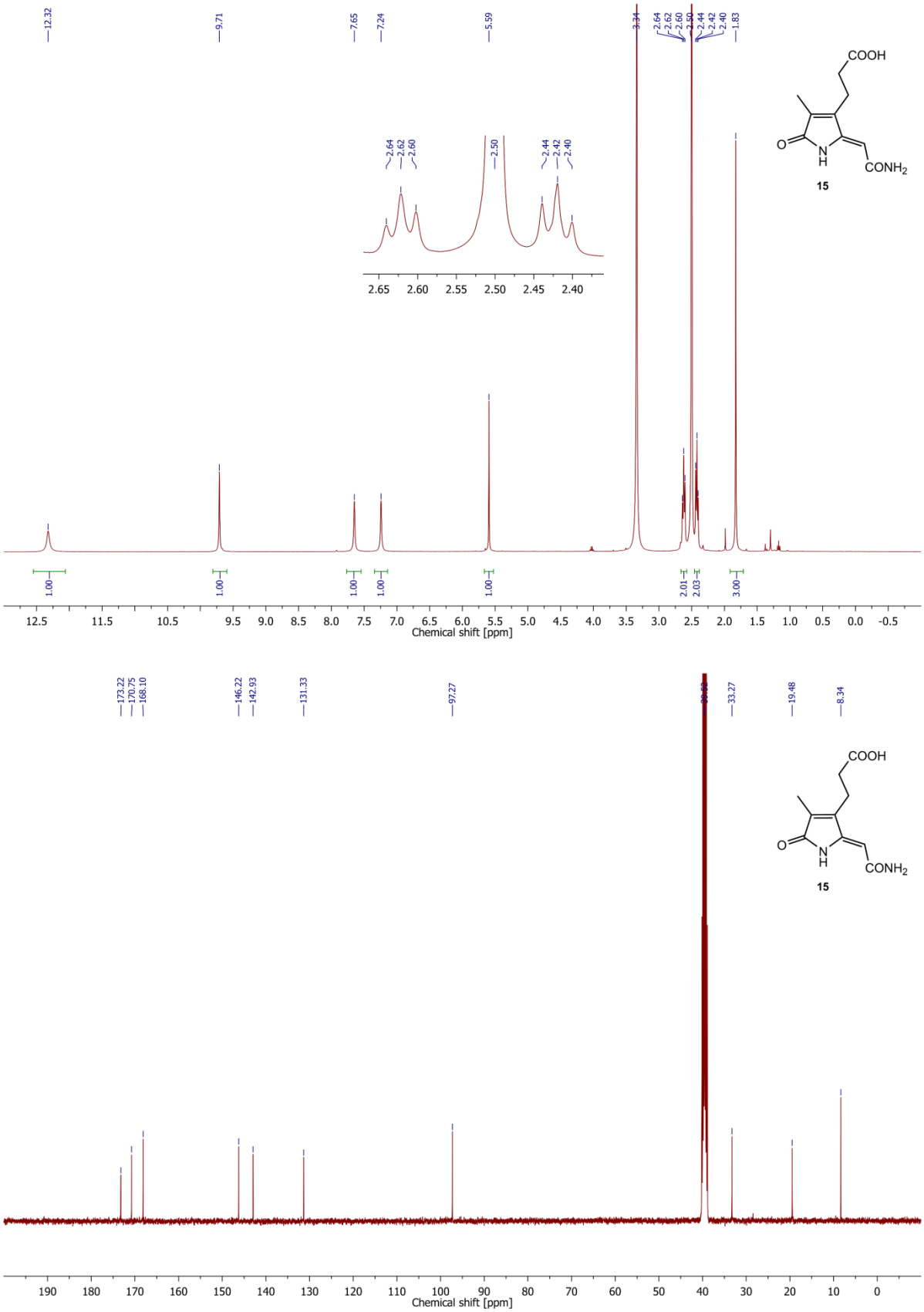

Figure S13. ^{1}H NMR spectrum (400 MHz, DMSO-d_6, 297 K, top) and ^{13}C{^{1}H} NMR spectrum (101 MHz, DMSO-d_6, 297 K, bottom) of **15**.

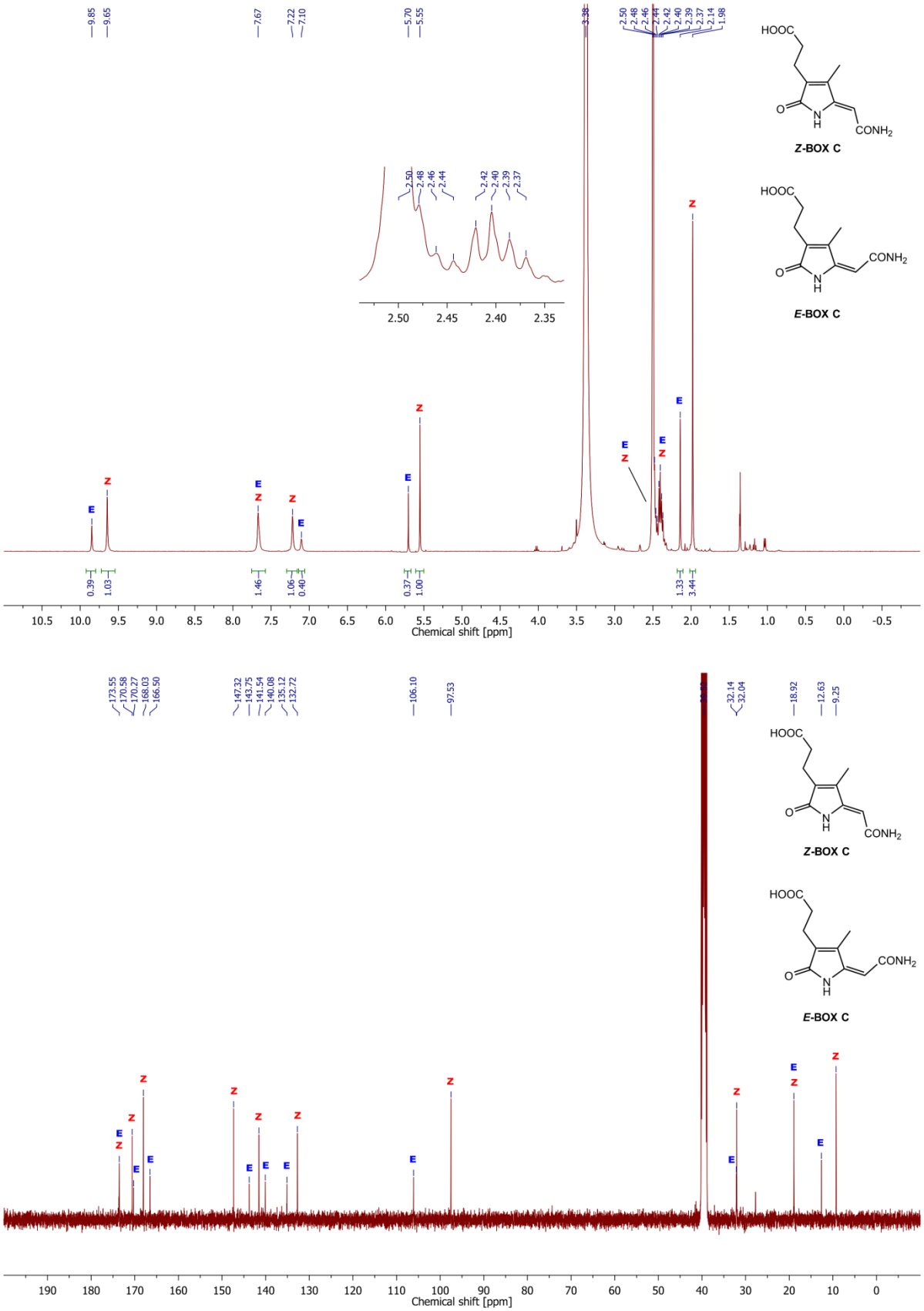

Figure S14. ^{1}H NMR spectrum (400 MHz, DMSO-d_6, 297 K, top) and ^{13}C{^{1}H} NMR spectrum (101 MHz, DMSO-d_6, 297 K, bottom) of **9** after irradiation with bright sunlight for 21 d, showing a mixture of *Z*-BOX C and *E*-BOX C.

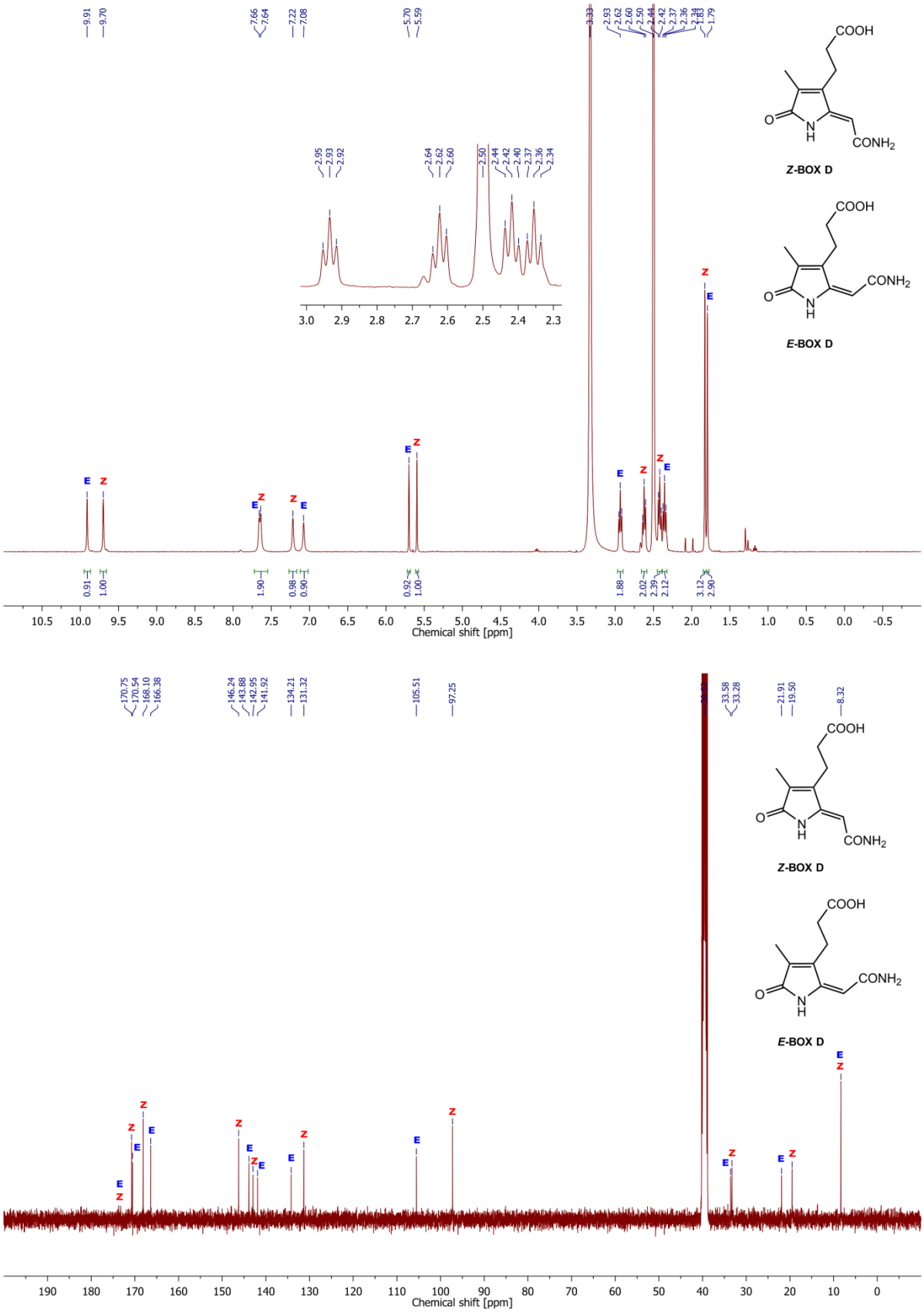

Figure S15. ^{1}H NMR spectrum (400 MHz, DMSO-d_6, 297 K, top) and ^{13}C{^{1}H} NMR spectrum (101 MHz, DMSO-d_6, 297 K, bottom) of **15** after irradiation with bright sunlight for 14 d, showing a mixture of *Z*-BOX D and *E*-BOX D.

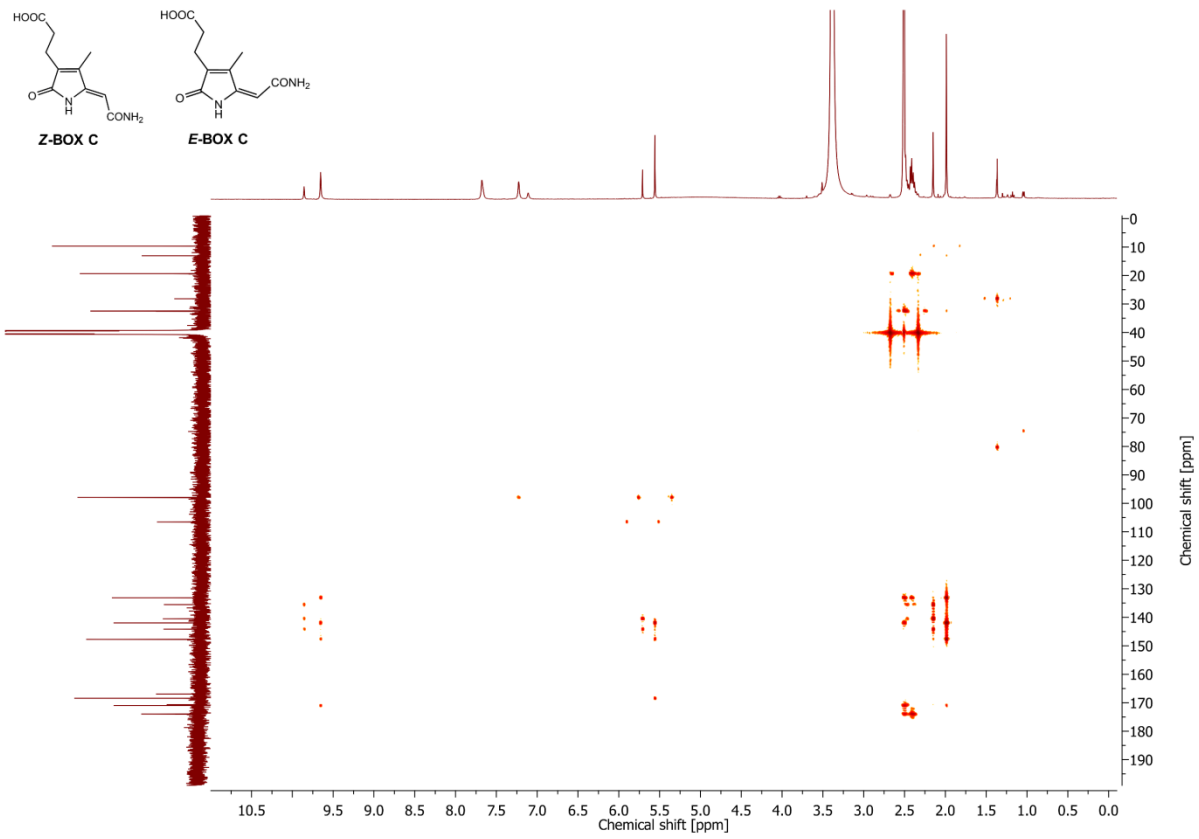

Figure S16. ^{13}C, ^{1}H HMBC spectrum (400 MHz, 101 MHz, DMSO-d_6, 297 K) of **9** after irradiation with sunlight for 21 d.

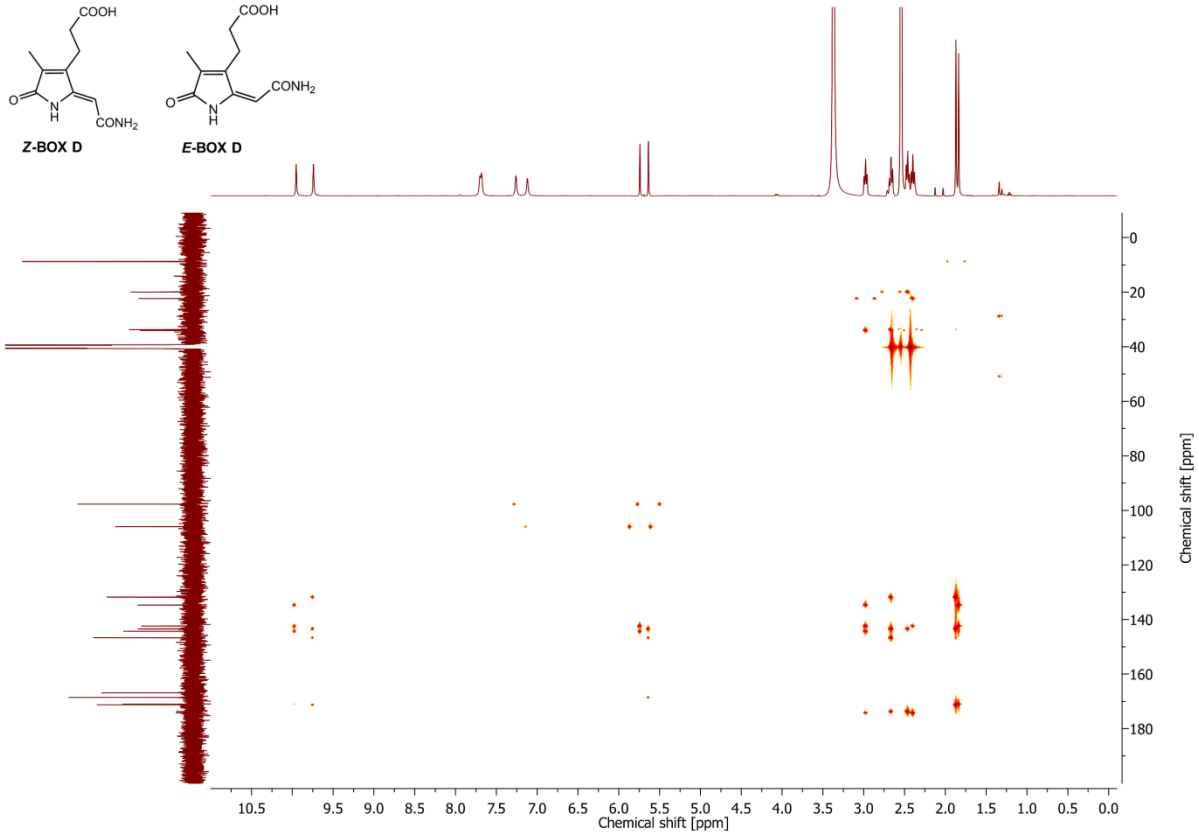

Figure S17. ^{13}C, ^{1}H HMBC spectrum (600 MHz, 151 MHz, DMSO-d_6, 297 K) of **15** after irradiation with sunlight for 14 d.

II. Mass spectrometric comparison of synthetic and bilirubin-derived BOX C

For the comparison of the synthetic and bilirubin-derived BOX C (**9**), a Dionex UltiMate 3000 UHPLC (Thermo Fisher Scientific, Leicestershire, United Kingdom), with a Acquity UHPLC BEH C18 column (1.7 µm, 100 × 2.1 mm) coupled with a Q-Exactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific, Leicestershire, United Kingdom) with electrospray ionization in positive ion mode was used. Solvents A (water, 2% acetonitrile, 0.1% HCOOH) and B (acetonitrile, 0.1% HCOOH) were used with a flow rate of 0.4 mL/min and the following solvent program: 0.0 min, 0% B; 0.5 min, 0% B; 1.0 min, 18% B; 8.0 min, 18% B; 9.0 min, 100% B; 10.9 min, 100% B; 11.0 min, 0% B; 13.0 min, 0% B. Synthetic and degraded BOX C were dissolved in water and further diluted for injection. For direct comparison, both standards were combined and diluted as the separate samples, resulting in a doubled peak area in the shown extracted ion chromatogram (+/- 5 ppm of exact mass, TOC).

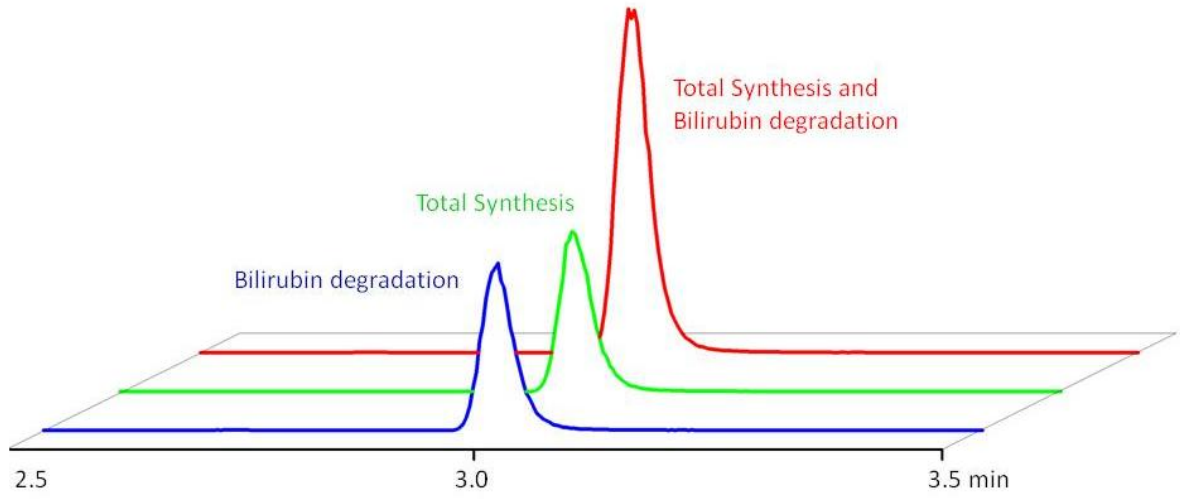

Figure S18. The identity of synthetic BOX C (green) with the derivative, obtained from an *in vitro* oxidation of bilirubin (blue), was verified by ultrahigh performance liquid chromatography (UHPLC) ESI Orbitrap HR-MS method. Identical retention times were found for synthetic BOX C, bilirubin-derived BOX C as well as for a combination of both (red).

III. X-Ray crystal structure determinations

Table S1: Crystal data and refinement details for the X-ray structure determinations.

Compound	**B**	**3**	**9**	**15**
formula	$C_{15}H_{17}NO_7S$	$C_{15}H_{15}NO_6S$	$C_{10}H_{12}N_2O_4$	$C_{10}H_{12}N_2O_4$
fw ($g \cdot mol^{-1}$)	355.36	337.34	224.22	224.22
T *(°C)*	-140(2)	-140(2)	-140(2)	-140(2)
crystal system	monoclinic	triclinic	triclinic	monoclinic
space group	P 2_1/n	P $\bar{1}$	P $\bar{1}$	P 2_1/n
a (Å)	7.2163(2)	7.7698(4)	7.6784(5)	7.8544(3)
b (Å)	12.4415(2)	9.8328(6)	7.8448(5)	14.0260(6)
c (Å)	17.8306(4)	11.4487(7)	9.1114(6)	9.9668(4)
A(°)	90	68.144(3)	91.933(4)	90
B(°)	96.925(1)	73.035(3)	101.720(2)	110.854(2)
Γ(°)	90	76.069(3)	105.491(4)	90
V($Å^3$)	1589.18(6)	767.82(8)	515.61(6)	1026.07(7)
Z	4	2	2	4
ρ ($g \cdot cm^{-3}$)	1.485	1.459	1.444	1.451
μ (cm^{-1})	2.42	2.42	1.13	1.14
measured data	9494	6957	6485	13375
data with I > 2σ(I)	3350	3006	1882	2074
unique data / R_{int}	3633/0.0277	3396/0.0214	2328/0.0346	2339/0.0413
wR_2 (all data, on F^2)[a)]	0.1231	0.0985	0.1377	0.0937
R_1 ($I > 2\sigma(I)$) [a)]	0.0516	0.0407	0.0586	0.0405
s [b)]	1.204	1.109	1.128	1.115
Res.dens.$_{max/min}$ ($e \cdot Å^{-3}$)	0.434/-0.440	0.325/-0.450	0.376/-0.328	0.352/-0.191
absorpt method	multi-scan	multi-scan	multi-scan	multi-scan
absorpt corr $T_{min/max}$	0.6978/0.7456	0.6763/0.7456	0.6270/0.7456	0.6291/0.7456
CCDC No.	1904792	1904793	1904794	1904795

[a)] Definition of the *R* indices: $R_1 = (\Sigma \| |F_o| - |F_c| \|)/\Sigma |F_o|$;
$wR_2 = \{\Sigma[w(F_o^2-F_c^2)^2]/\Sigma[w(F_o^2)^2]\}^{1/2}$ with $w^{-1} = \sigma^2(F_o^2) + (aP)^2+bP$; $P = [2F_c^2 + Max(F_O^2]/3$;
[b)] $s = \{\Sigma[w(F_o^2-F_c^2)^2]/(N_o-N_p)\}^{1/2}$.

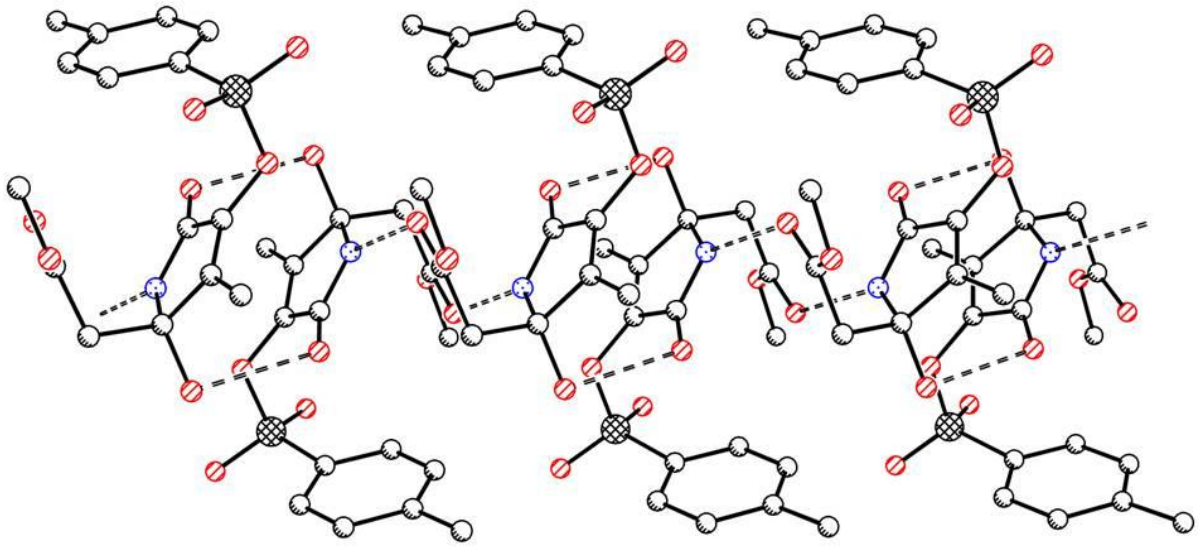

Figure S19. Hydrogen bridge network of compound **B**. The atoms are drawn with arbitrary radii, hydrogen atoms are omitted for the sake of clarity. Hydrogen bridges are depicted with dotted lines.

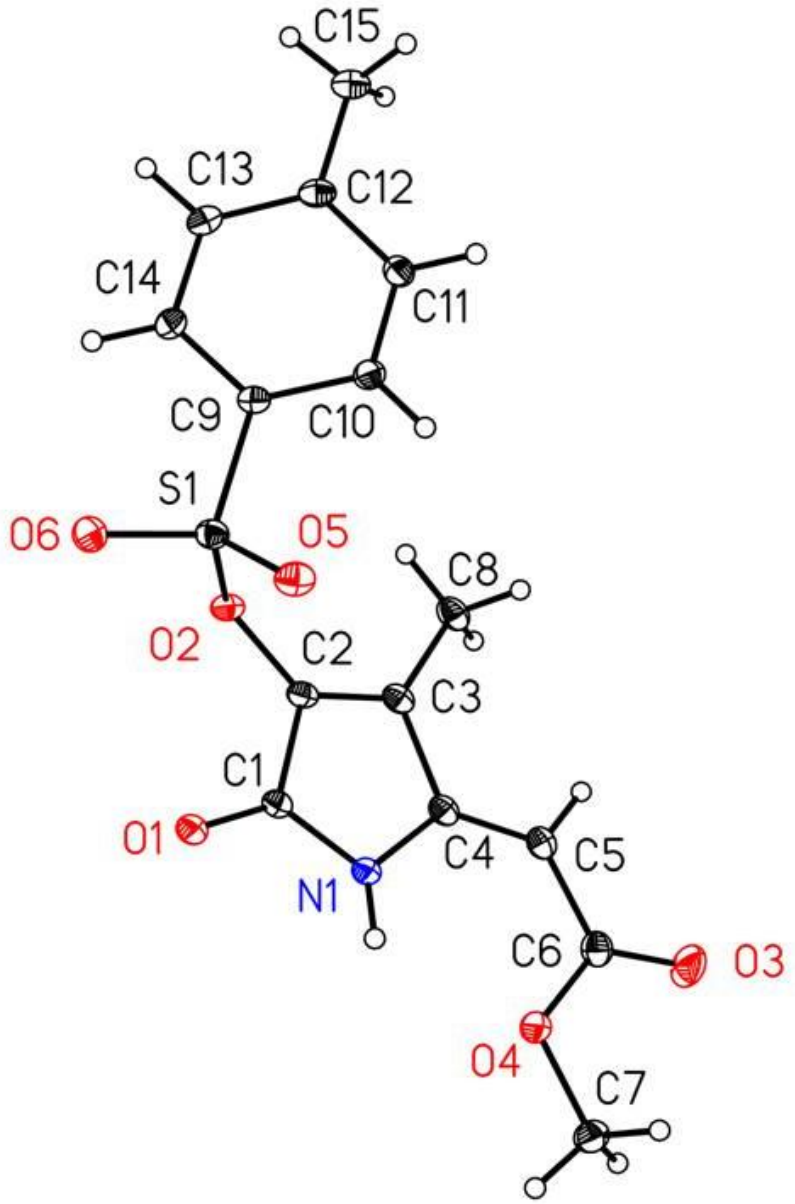

Figure S20. Representation of the molecular structure and atom labeling scheme of **3**. The ellipsoids represent a probability of 30 %, H atoms are drawn with arbitrary radii. Selected bond lengths (pm): S1-O2 162.74(13), S1-O5 142.71(13), S1-O6 142.26(14), C1-N1 136.9(2), C1-O1 122.4(2), C1-C2 147.9(2), C2-O2 138.0(2), C2-C3 133.1(3), C3-C4 148.0(2), C3-C8 149.2(2), C4-N1 139.4(2), C4-C5 134.2(3), C5-C6 146.2(3), C6-O3 120.7(2), C6-O4 133.6(2), O4-C7 145.1(2); angles (deg.): C1-N1-C4 110.57(15), N1-C1-C2 105.04(14), C1-C2-C3 111.02(15), C2-C3-C4 106.04(15), C3-C4-N1 107.22(15).

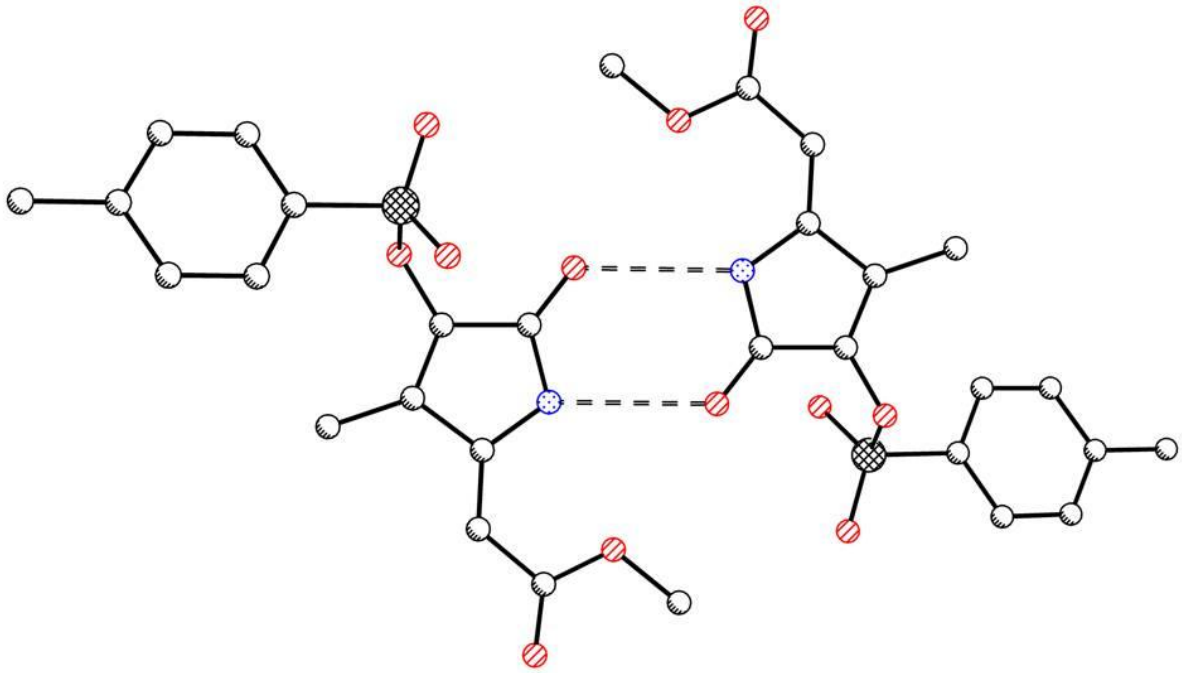

Figure S21. Formation of a dimer in the crystalline state of compound **3**. The atoms are shown with arbitrary radii, hydrogen atoms are omitted for the sake of clarity. N-H···O Hydrogen bridges are depicted with dotted lines.

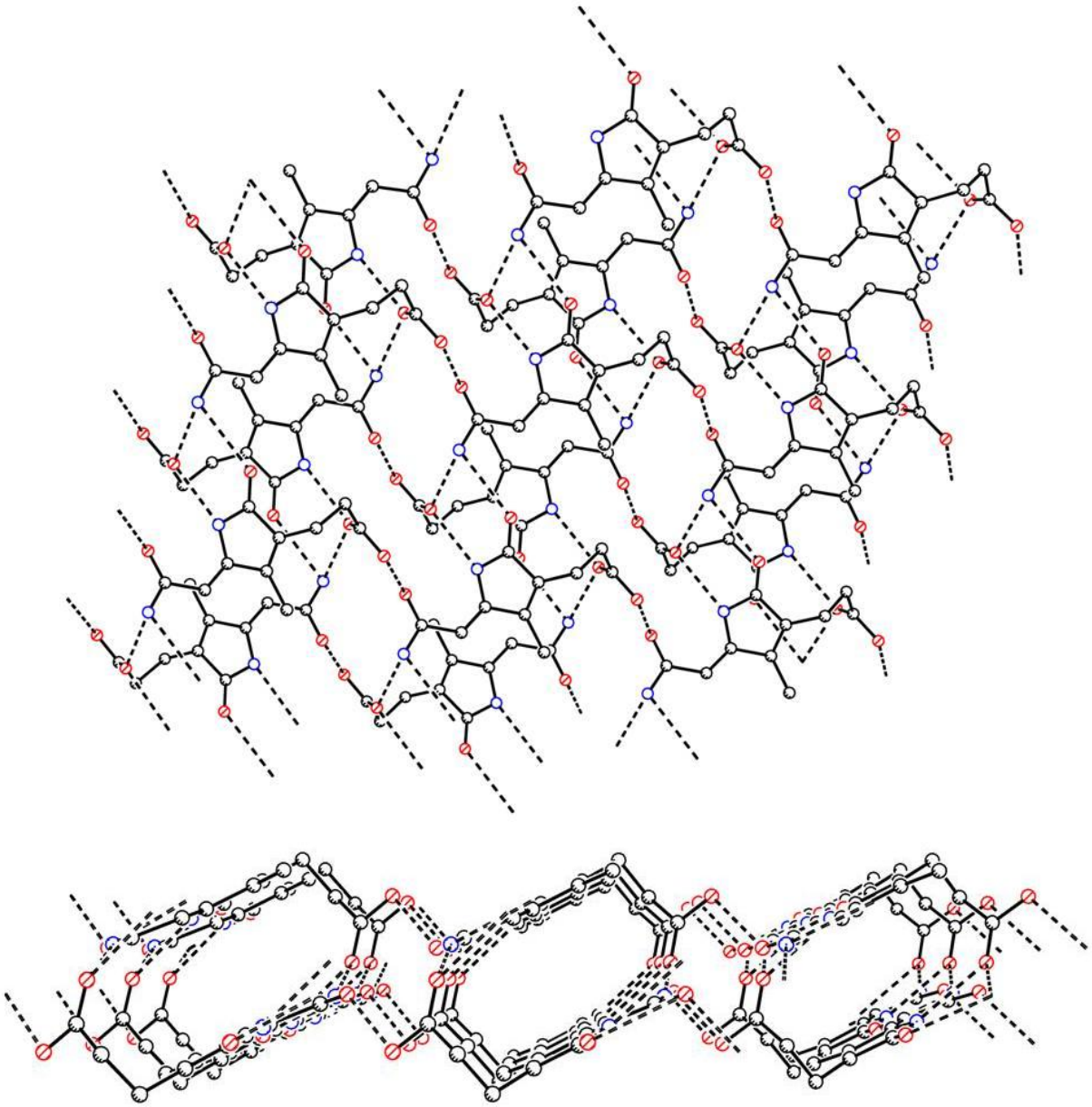

Figure S22. Two views of the hydrogen bridge network of compound *Z*-BOX C (**9**). The atoms are drawn with arbitrary radii, hydrogen atoms are omitted for clarity reasons. Hydrogen bridges are depicted with dotted lines.

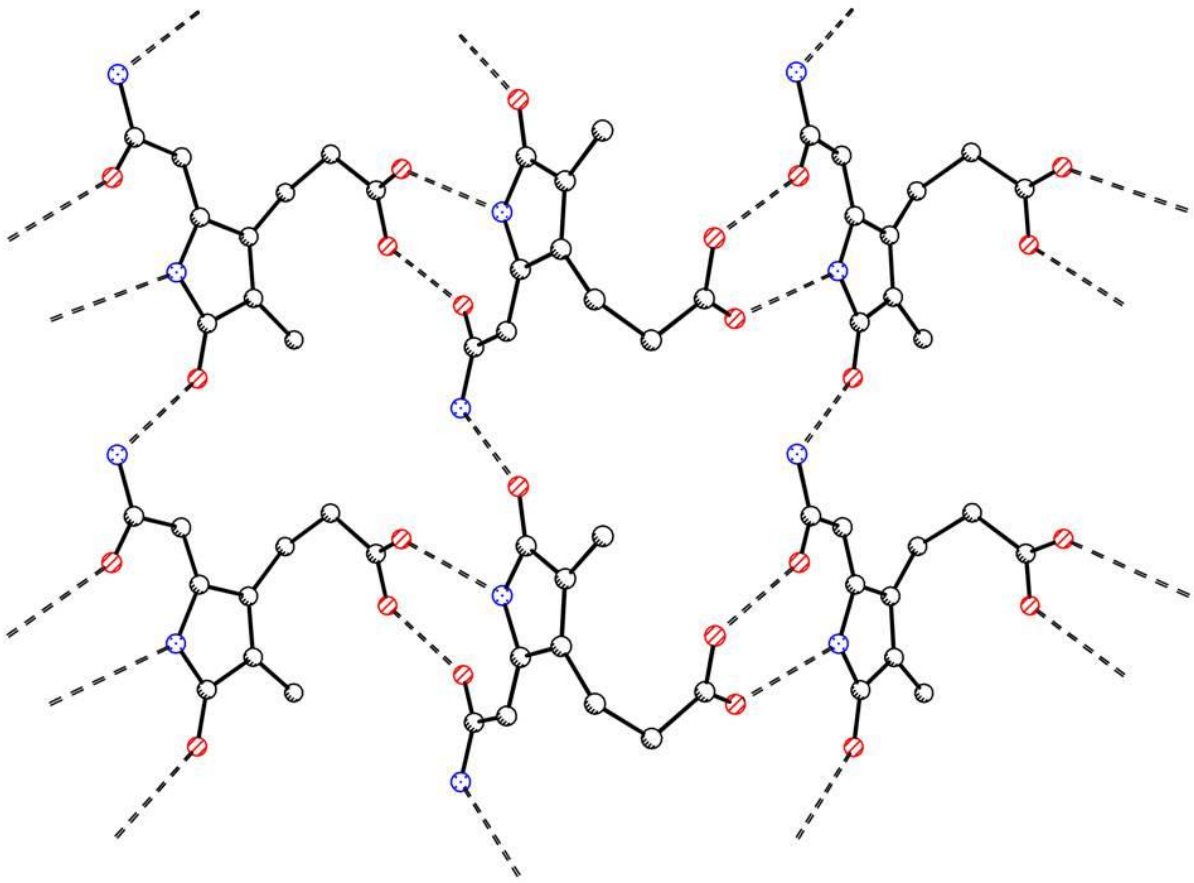

Figure S23. Hydrogen bridge network of the layer structure of compound *Z*-BOX D (**15**). The atoms are drawn with arbitrary radii, hydrogen atoms are omitted for the sake of clarity. Hydrogen bridges are depicted with dotted lines.

Table S2. Hydrogen bridges for *Z*-BOX C (**9**, [Å and deg.]).

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
O(2)-H(1O2)...O(4)#1	0.88(4)	1.80(4)	2.681(2)	175(3)
N(1)-H(1N1)...O(3)#2	0.89(3)	1.97(3)	2.813(2)	158(3)
N(1)-H(1N1)...O(4)	0.89(3)	2.42(3)	2.857(2)	111(2)
N(2)-H(2N2)...O(1)#3	0.95(3)	2.06(3)	2.961(3)	158(2)
N(2)-H(1N2)...O(4)#4	0.92(3)	2.23(3)	3.144(3)	172(3)

Symmetry transformations used to generate equivalent atoms:
#1: x, y, z+1; #2: -x+1 ,-y, -z+1; #3: x, y+1, z; #4: -x+1, -y+1, -z

Table S3. Hydrogen bridges for *Z*-BOX D (**15**, [Å and deg.]).

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
O(2)-H(1O2)...O(4)#1	0.93(3)	1.71(3)	2.6423(15)	178(3)
N(1)-H(1N1)...O(3)#2	0.86(2)	2.26(2)	3.0556(16)	154.5(17)
N(1)-H(1N1)...O(4)	0.86(2)	2.328(19)	2.8096(16)	115.8(15)
N(2)-H(1N2)...O(4)#3	0.86(2)	2.34(2)	3.0877(17)	145.9(17)
N(2)-N(2N2)...O(1)#4	0.88(2)	1.98(2)	2.8438(17)	164.8(18)
C(9)-H(9)...O(1)#4	0.956(18)	2.421(18)	3.2123(16)	139.9(14)

Symmetry transformations used to generate equivalent atoms:
#1: -x+3/2, y-1/2, -z+3/2; #2: -x+3/2, y+1/2, -z+3/2;
#3: -x+1, -y+1, -z+2; #4: x-1, y, z