

## Electronic Supplementary Information for

# The efficient and selective biocatalytic oxidation of norisoprenoid and aromatic substrates by CYP101B1 from *Novosphingobium aromaticivorans* DSM12444

Emma A. Hall and Dr. Stephen G. Bell\*

**Table S1.** Comparison of the sequence identities of CYP101B1, CYP101B2 (*N. tardaugens*), CYP101A1 (*P. putida*), CYP101D1, CYP101D2, CYP101C1 (all *N. aromaticivorans*, P450cin (*C. braakii*) CYP109D1, CYP246B1 (both *S. cellulosum*) and CYP105D1 (*S. griseus*). We note that other P450s from *S. cellulosum* had greater similarity to CYP101B1 the greatest being CYP124E1.

<b>CYP101B1</b>	<b>Identities</b>	<b>Positives</b>	<b>Gaps</b>	<b>Score</b>
<b>CYP101B2</b>	236/398(59%)	297/398(74%)	6/398(1%)	489
<b>CYP101D2</b>	184/403(46%)	254/403(63%)	9/403(2%)	348
<b>CYP101A1</b>	173/395(44%)	236/395(59%)	3/395(1%)	340
<b>CYP101D1</b>	177/401(44%)	249/401(62%)	9/401(2%)	327
<b>CYP101C1</b>	142/395(36%)	218/395(55%)	5/395(1%)	256
<b>P450cin</b>	102/373(27%)	169/373(45%)	6/373(1%)	145
<b>CYP109D1</b>	95/307(31%)	142/307(46%)	16/307(5%)	122
<b>CYP105D1</b>	88/313(28%)	144/313(46%)	13/313(4%)	107
<b>CYP264B1</b>	79/332(24%)	138/332(41%)	15/332(4%)	70.5

**Table S2** The comparison of likely residues in the active site, substrate binding channel and in close proximity of the heme of CYP101B1 (based on the CYP101C1, CYP101D1 and CYP101D2 structures) and the corresponding residues in CYP101B2, CYP101D2, CYP101A1, CYP101D1 and CYP101C1.

<b>CYP101B1</b>	<b>CYP101B2</b>	<b>CYP101D2</b>	<b>CYP101A1</b>	<b>CYP101D1</b>	<b>CYP101C1</b>
Arg72	Arg76	Glu84	Glu84	Arg85	Gln70
Ile74	Ile78	Ile86	Pro86	Ile87	Leu72
<b><u>Val75</u></b>	<b><u>Val79</u></b>	<b><u>Phe87</u></b>	<b><u>Phe87</u></b>	<b><u>Trp88</u></b>	<b><u>Ala73</u></b>
Val76	Val80	Ala92	Ala92	Ala93	Leu78
Gly77	Gly81	Gly93	Gly93	Gly94	Gly79
Glu78	Glu82	Glu94	Glu94	Glu95	Lys80
<b><u>His80</u></b>	<b><u>His84</u></b>	<b><u>Tyr96</u></b>	<b><u>Tyr96</u></b>	<b><u>Tyr97</u></b>	<b><u>Met82</u></b>
Gly81	Gly85	Gln97	Asp97	Asp98	Gln83
<b><u>Leu82</u></b>	<b><u>Leu86</u></b>	<b><u>Met98</u></b>	<b><u>Phe98</u></b>	<b><u>Met99</u></b>	<b><u>Phe84</u></b>
Thr85	Thr89	Thr101	Thr101	Thr102	Leu87
Glu144	Glu148	Glu156	Glu156	Thr157	Glu142
Arg166	Arg170	Ser178	Lys178	Gly179	Arg164
Glu170	Glu174	Arg182	Asp182	Asn183	Val168
<b>Thr173</b>	<b>Thr177</b>	<b>Thr185</b>	<b>Thr185</b>	<b>Thr186</b>	<b>Thr171</b>
<b>Arg174</b>	<b>Arg178</b>	<b>Arg186</b>	<b>Arg186</b>	<b>Arg187</b>	<b>Arg172</b>
<b>Pro175</b>	<b>Pro179</b>	<b>Pro187</b>	<b>Pro187</b>	<b>Pro188</b>	<b>Pro173</b>
Lys185	Lys189	Asn203	Lys197	Asn204	Lys183
Gln233	Gln241	Leu250	Leu244	Leu251	Asn230
<b><u>Ile236</u></b>	<b><u>Val244</u></b>	<b><u>Leu253</u></b>	<b><u>Val247</u></b>	<b><u>Leu254</u></b>	<b><u>Phe233</u></b>
Ala237	Ala245	Gly254	Gly248	Gly255	Gly234
<b>Asp240</b>	<b>Asp248</b>	<b>Asp257</b>	<b>Asp251</b>	<b>Asp258</b>	<b>Asp237</b>
<b>Thr241</b>	<b>Asp249</b>	<b>Thr258</b>	<b>Thr252</b>	<b>Thr259</b>	<b>Thr238</b>
Asn244	Asn252	Asn261	Asn255	Asn262	Ala241
<b>Val284</b>	<b>Val293</b>	<b>Val301</b>	<b>Val295</b>	<b>Val302</b>	<b>Val281</b>
Ile286	Ile294	Glu303	Asp297	Asp304	Val283
Ile384	Ile393	Ile401	Ile395	Ile402	Asn383
<b>Val385</b>	<b>Val394</b>	<b>Val402</b>	<b>Val396</b>	<b>Val403</b>	<b>Val384</b>

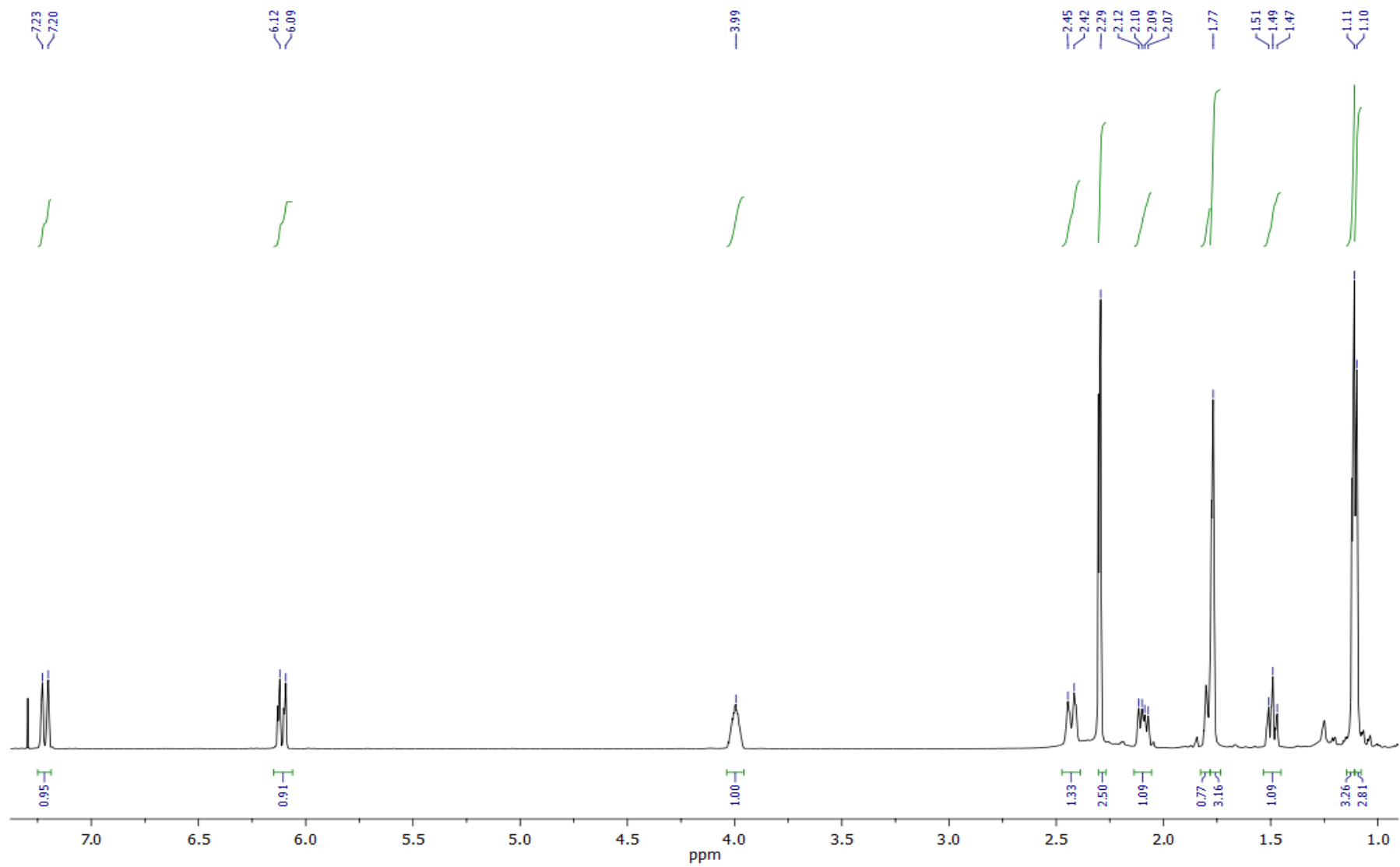
**Figure S1.** The CYP101B1 *E. coli* whole-cell oxidation system after the addition of indole.



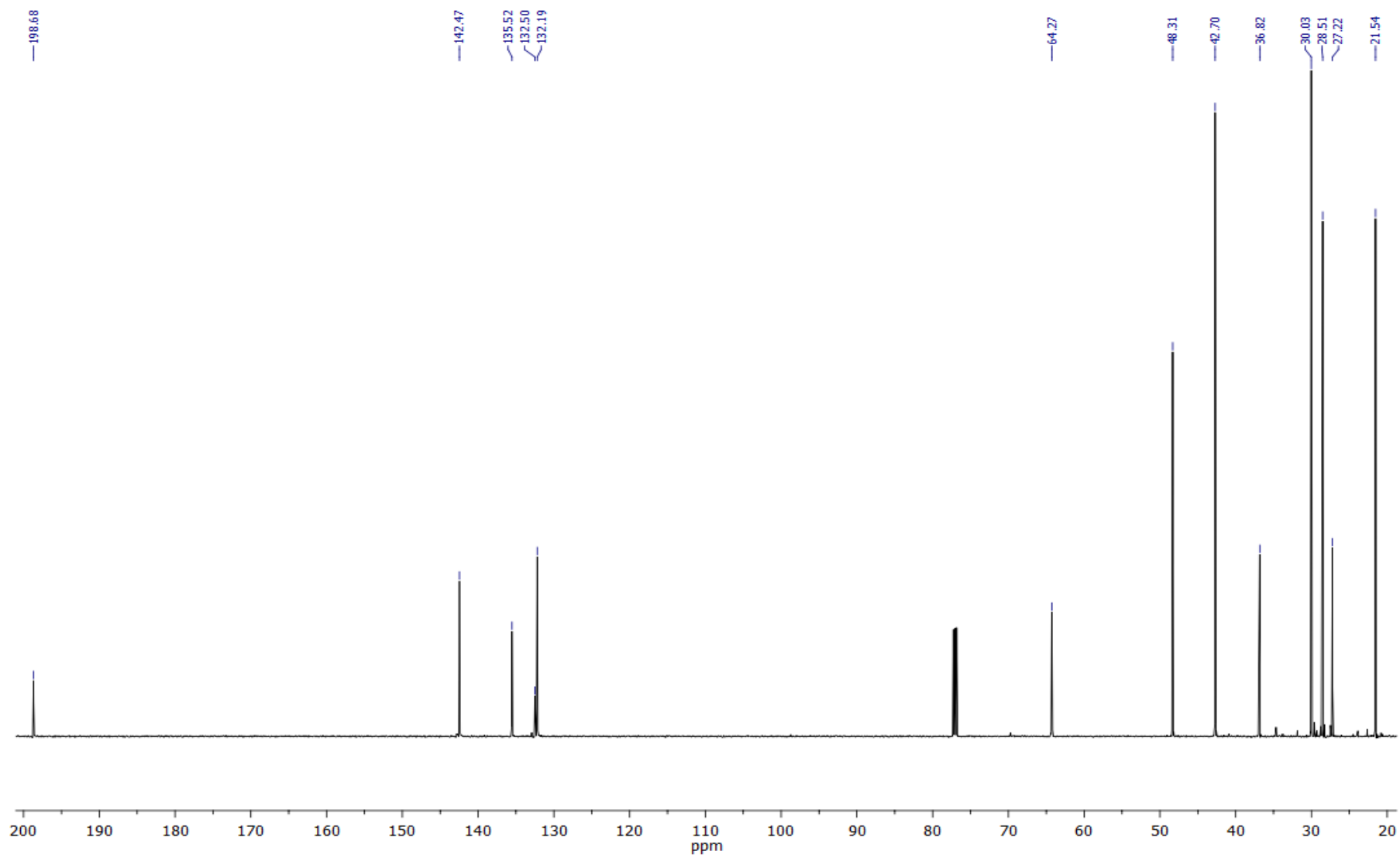
## NMR data

$^1\text{H-NMR}$ ,  $^{13}\text{C-NMR}$ , COSY and HSQC analysis of 3-hydroxy- $\beta$ -ionone, *trans*-3-hydroxy- $\alpha$ -ionone, 3-oxo-ionone, *cis*-3-hydroxy- $\alpha$ -ionone, 3-hydroxy- $\beta$ -damascone and 4-hydroxy- $\beta$ -damascone.

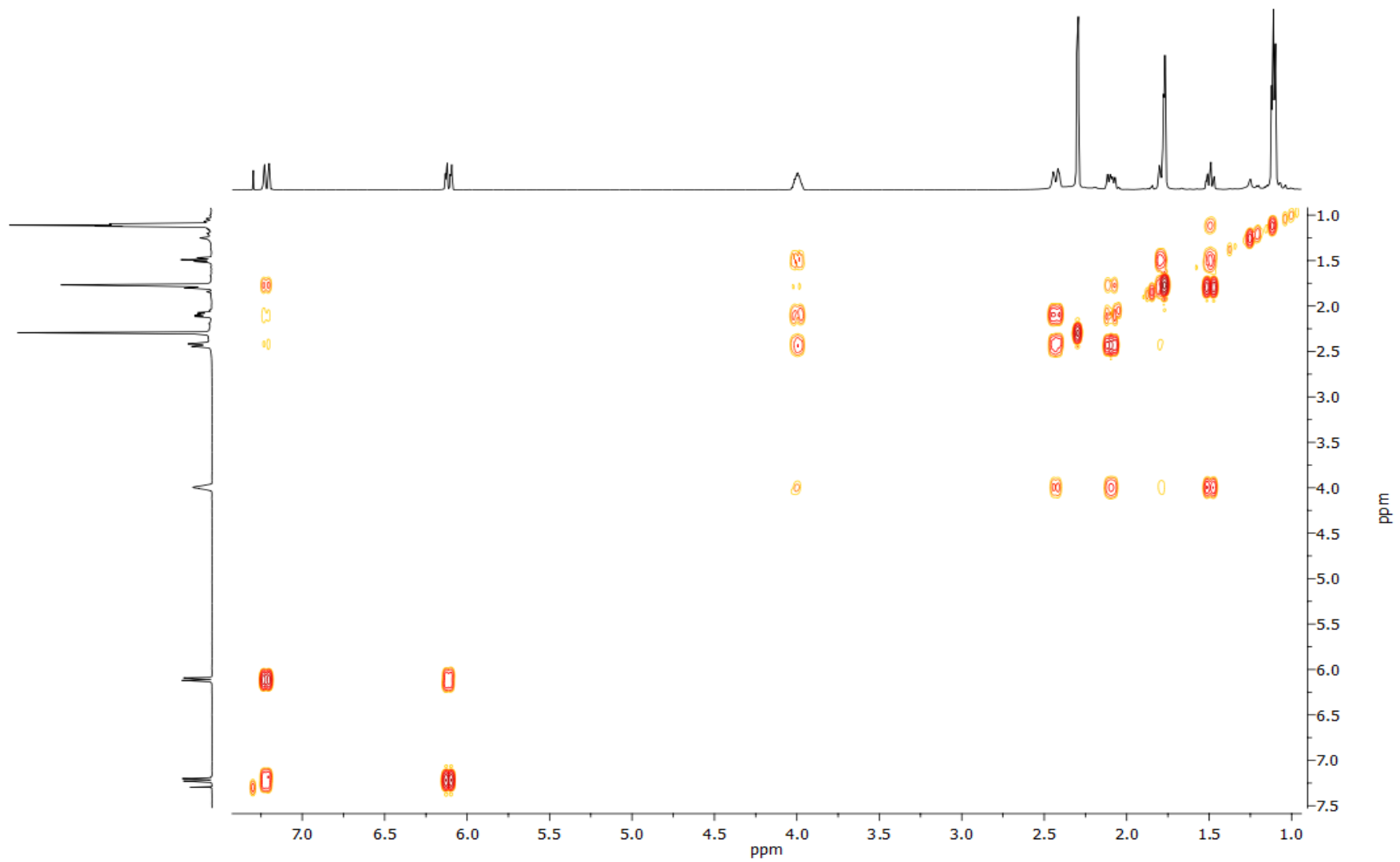
$^1\text{H-NMR}$  and COSY analysis of *trans*-4-phenylcyclohexanol



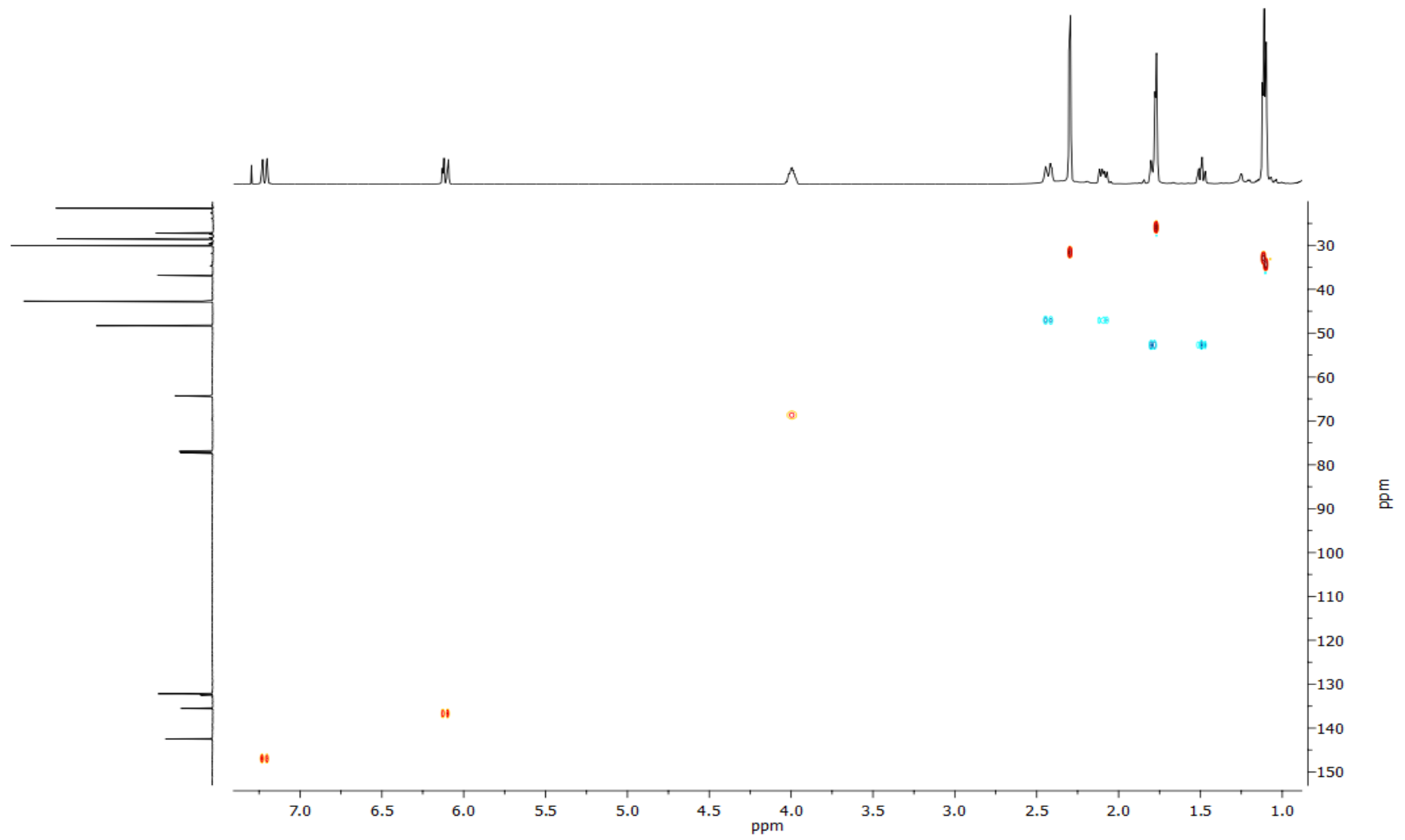
3-hydroxy- $\beta$ -ionone  $^1\text{H-NMR}$



3-hydroxy- $\beta$ -ionone  $^{13}\text{C}$ -NMR

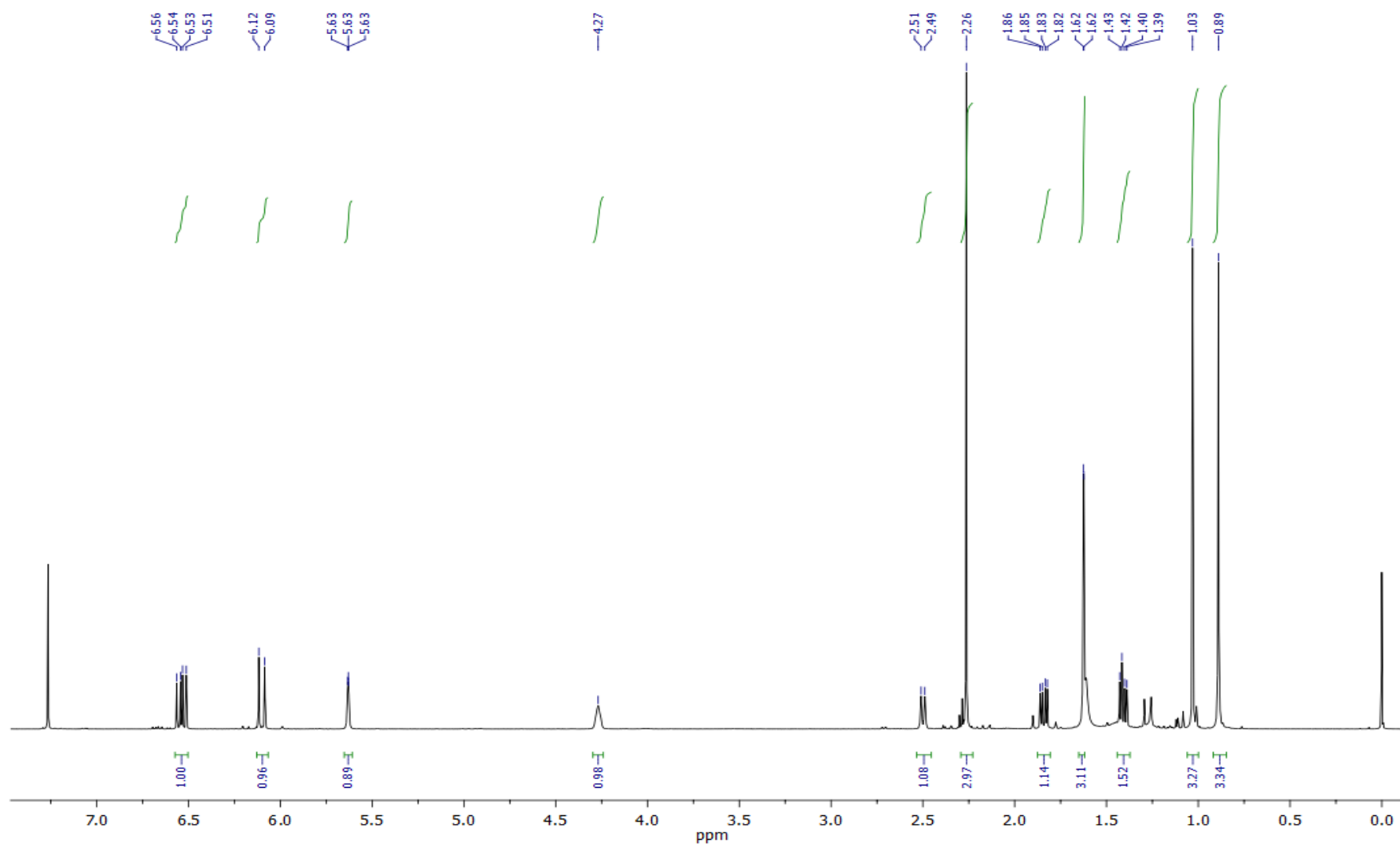


3-hydroxy- $\beta$ -ionone COSY

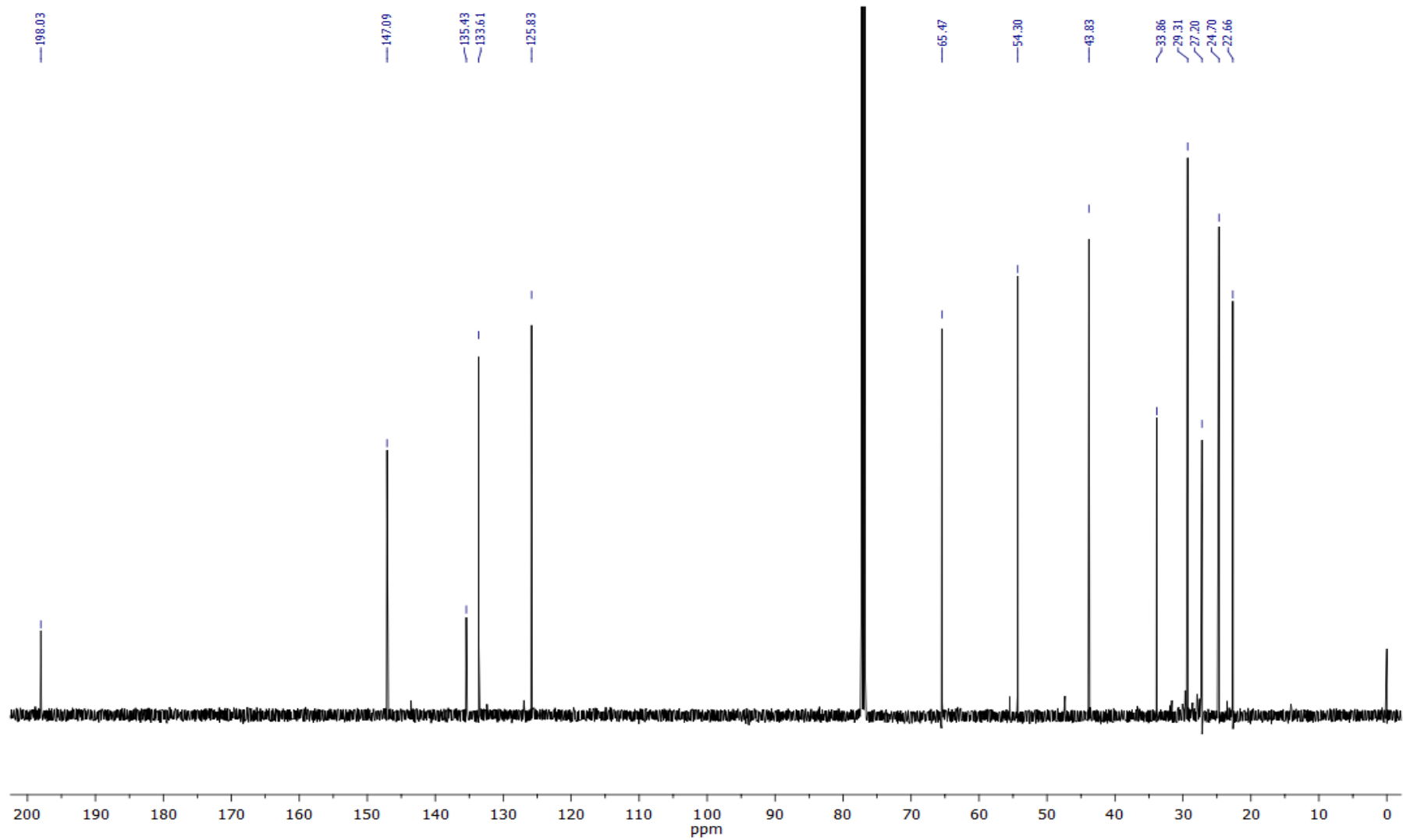


3-hydroxy- $\beta$ -ionone HSQC

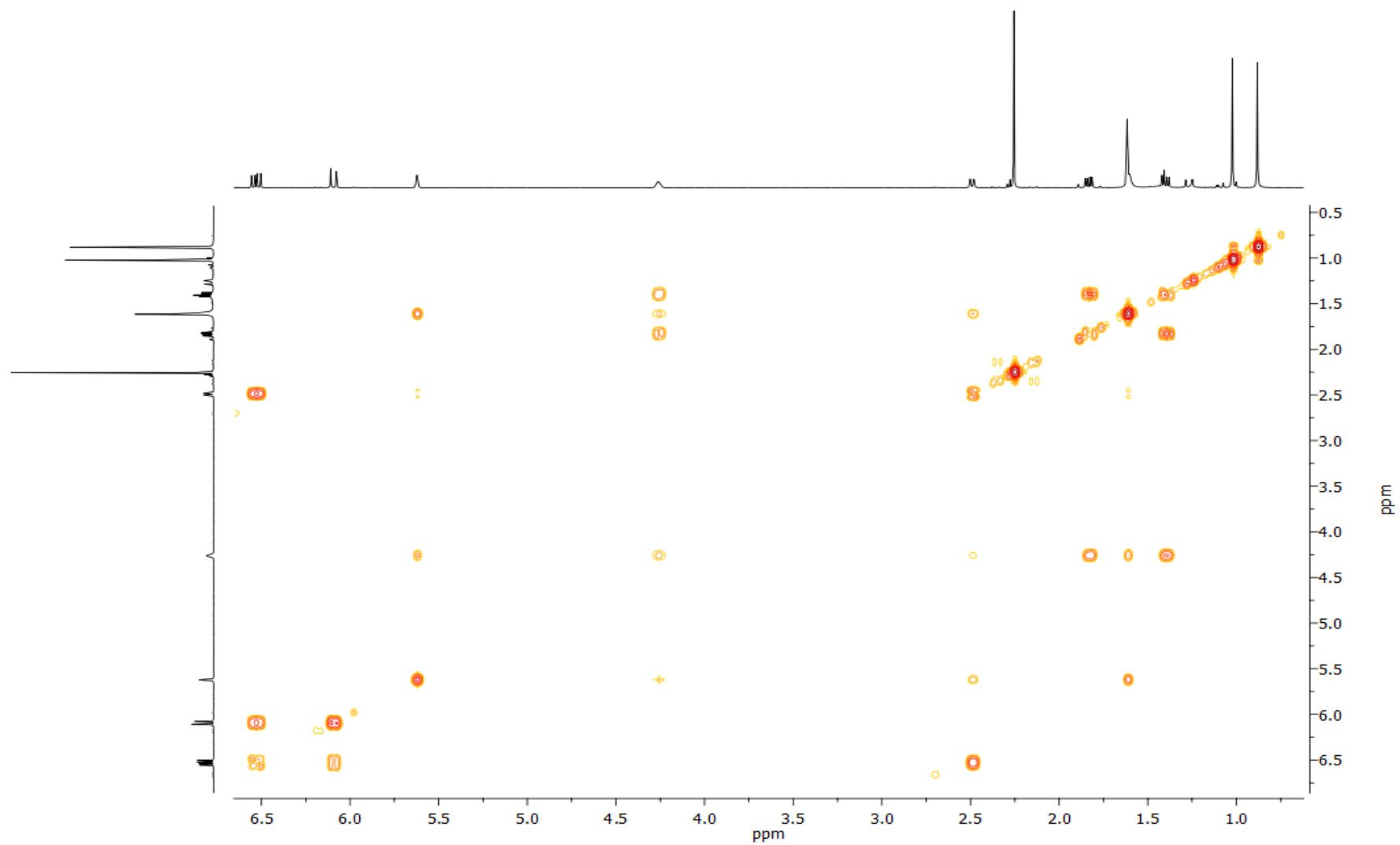




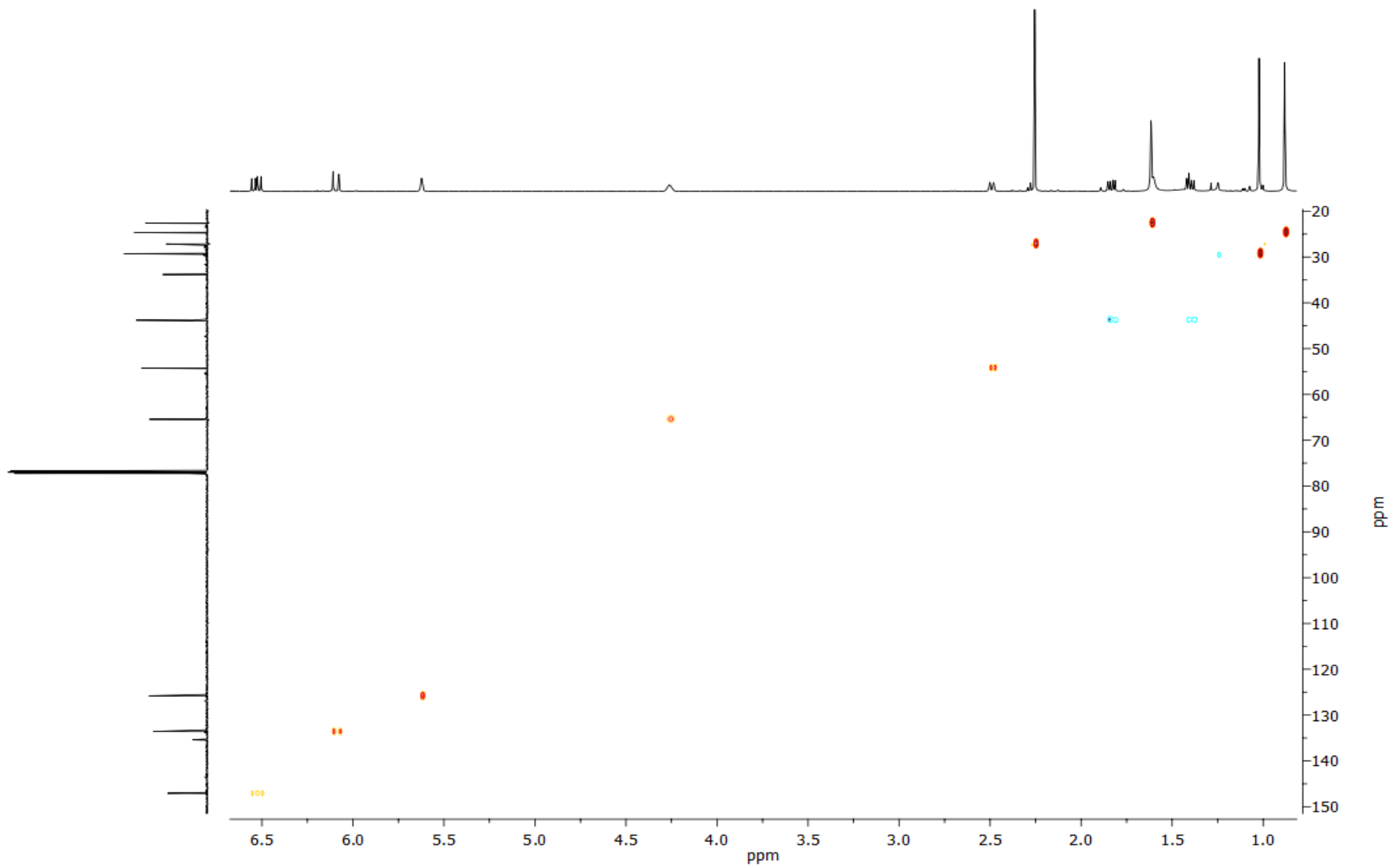
*trans*-3-hydroxy- $\alpha$ -ionone  $^1\text{H-NMR}$



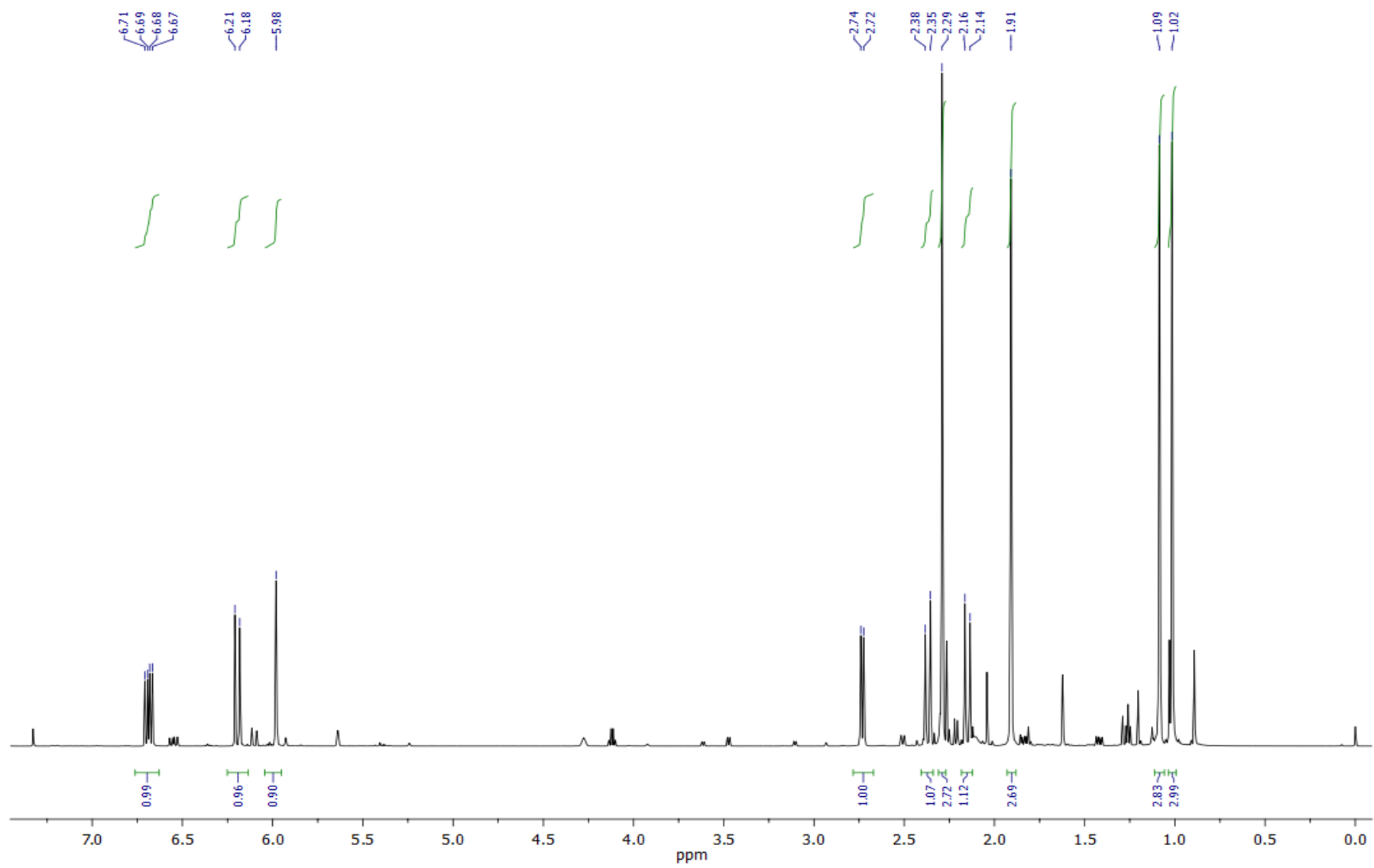
*trans*-3-hydroxy- $\alpha$ -ionone <sup>13</sup>C-NMR



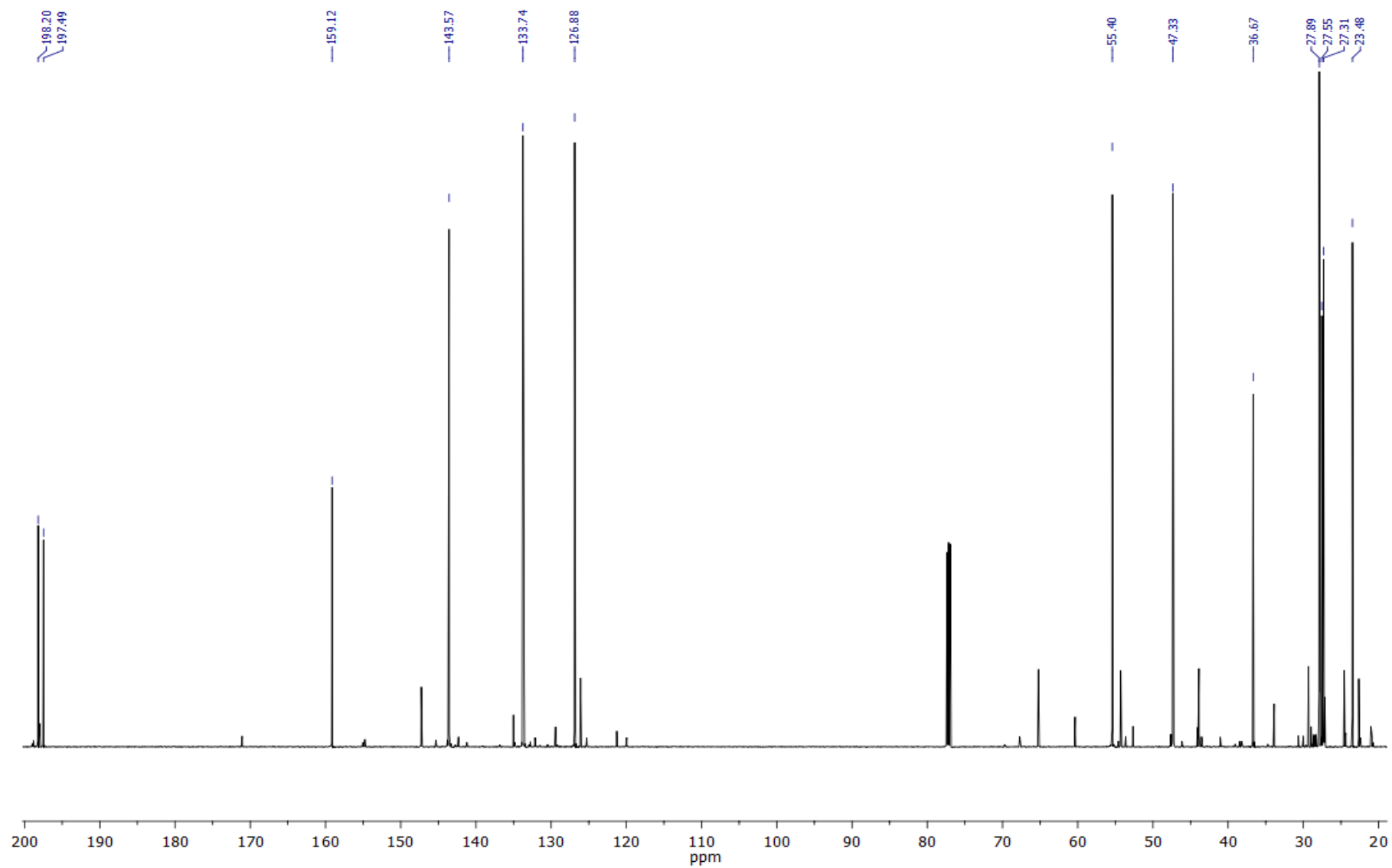
*trans*-3-hydroxy- $\alpha$ -ionone COSY



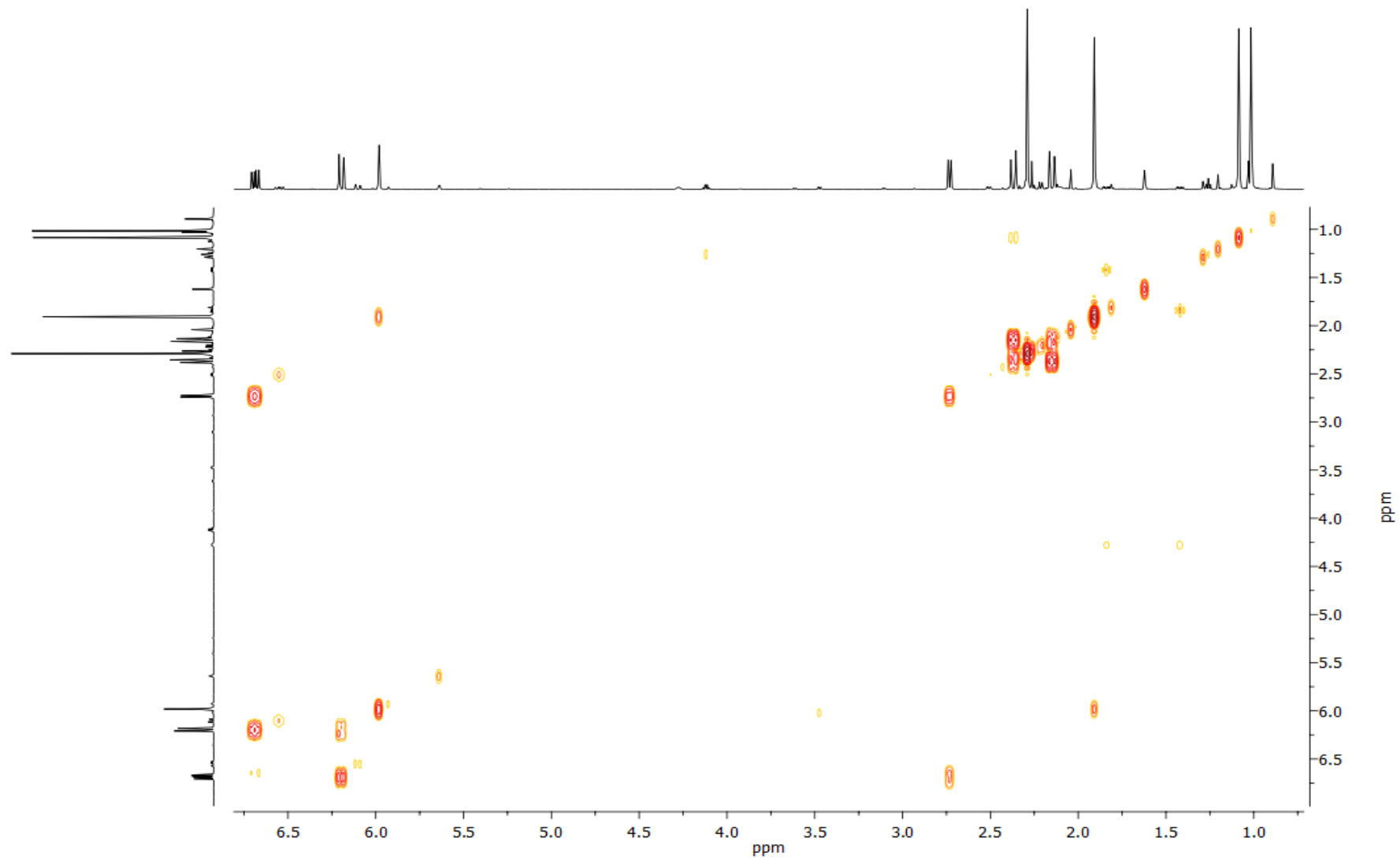
*trans*-3-hydroxy- $\alpha$ -ionone HSQC



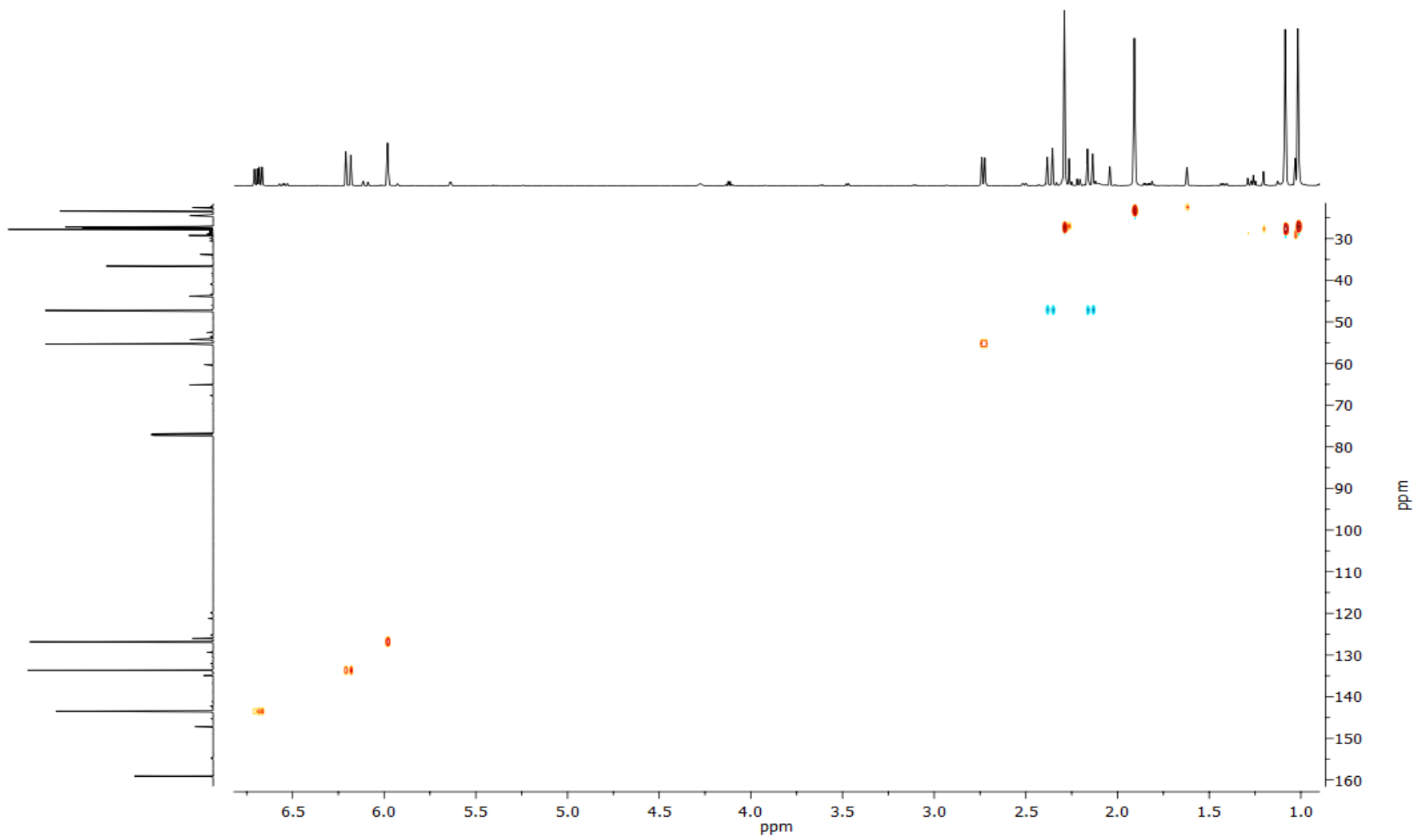
3-oxo- $\alpha$ -ionone  $^1\text{H-NMR}$



3-oxo- $\alpha$ -ionone  $^{13}\text{C}$ -NMR

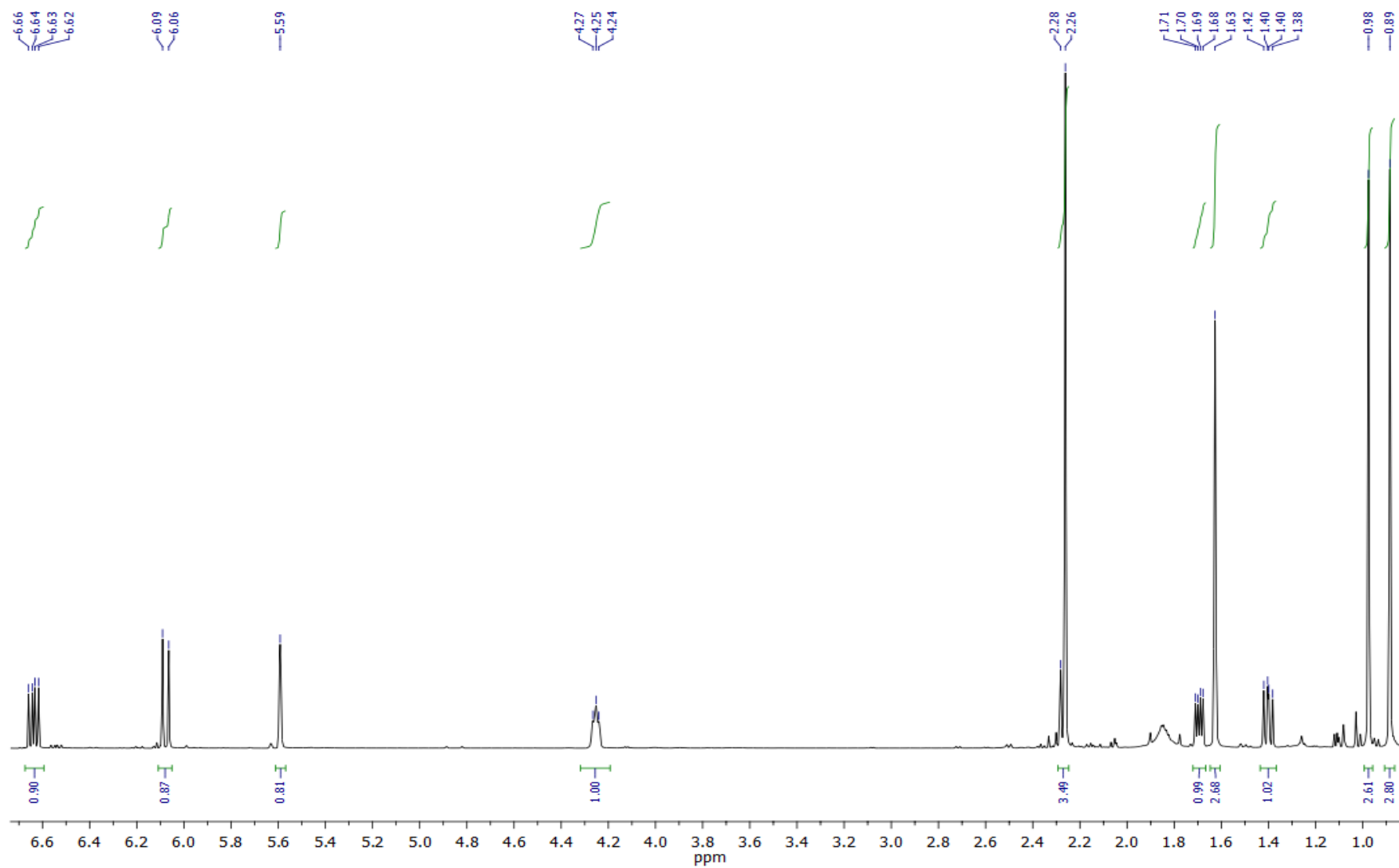


3-oxo- $\alpha$ -ionone COSY

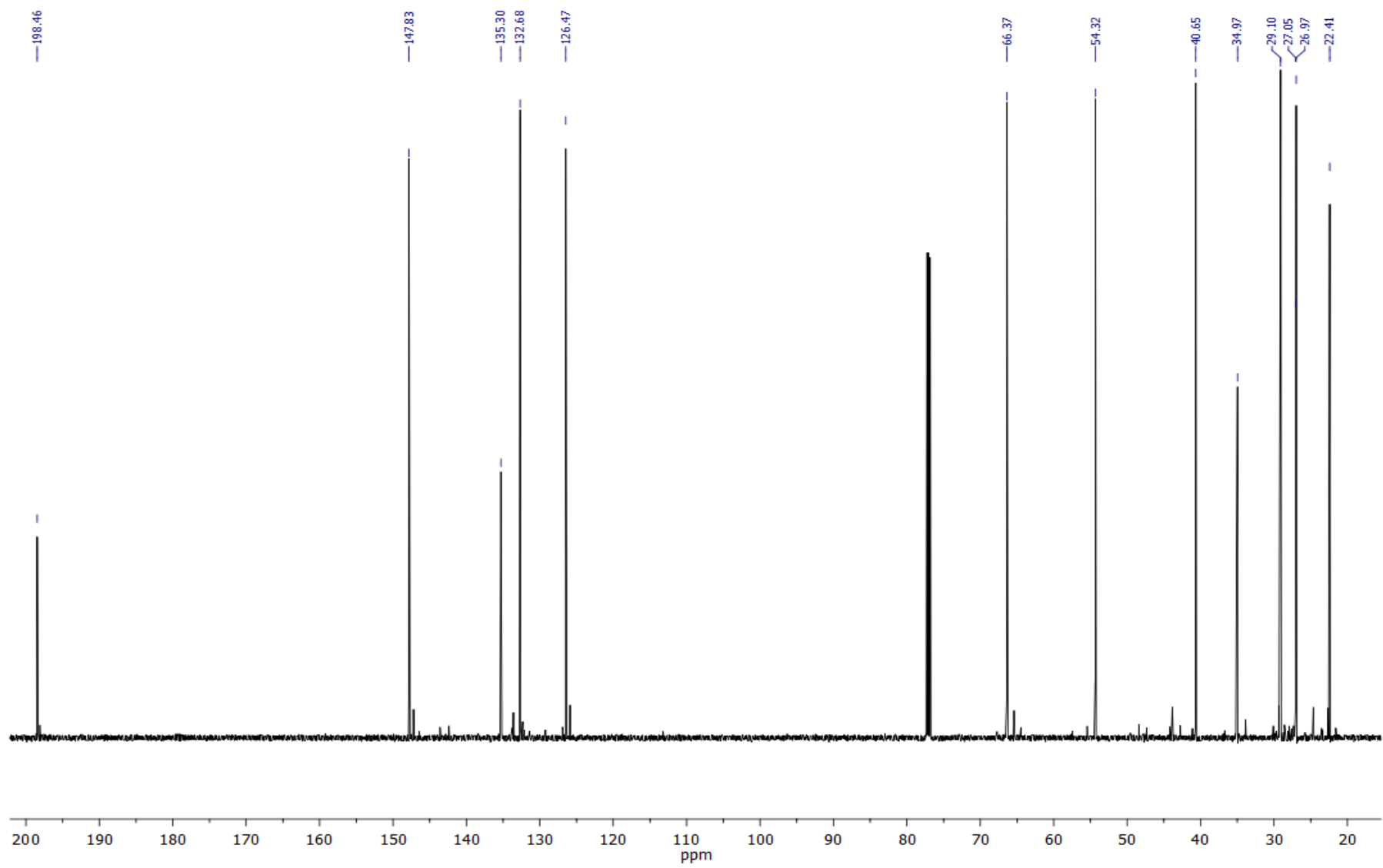


3-oxo- $\alpha$ -ionone HSQC

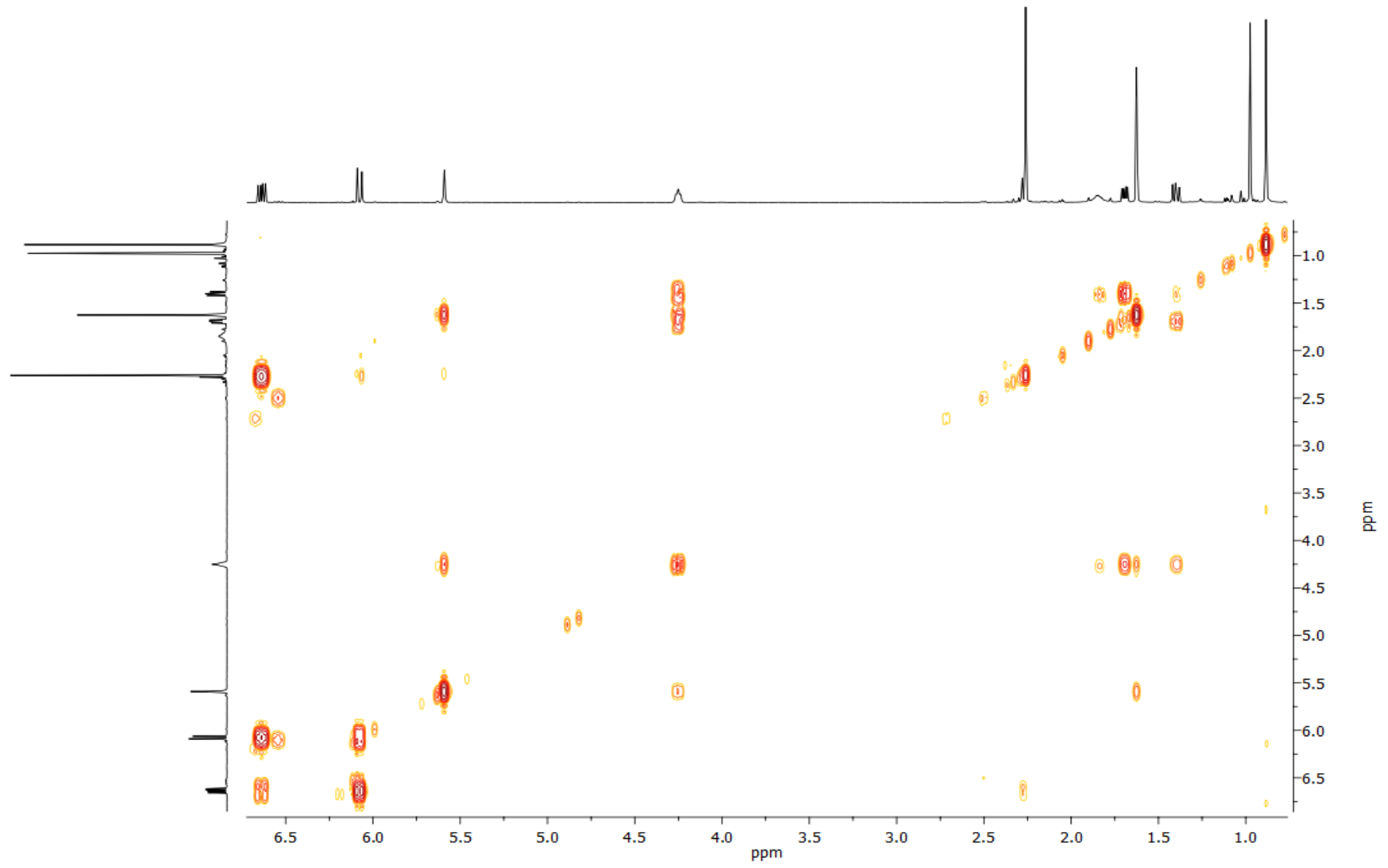




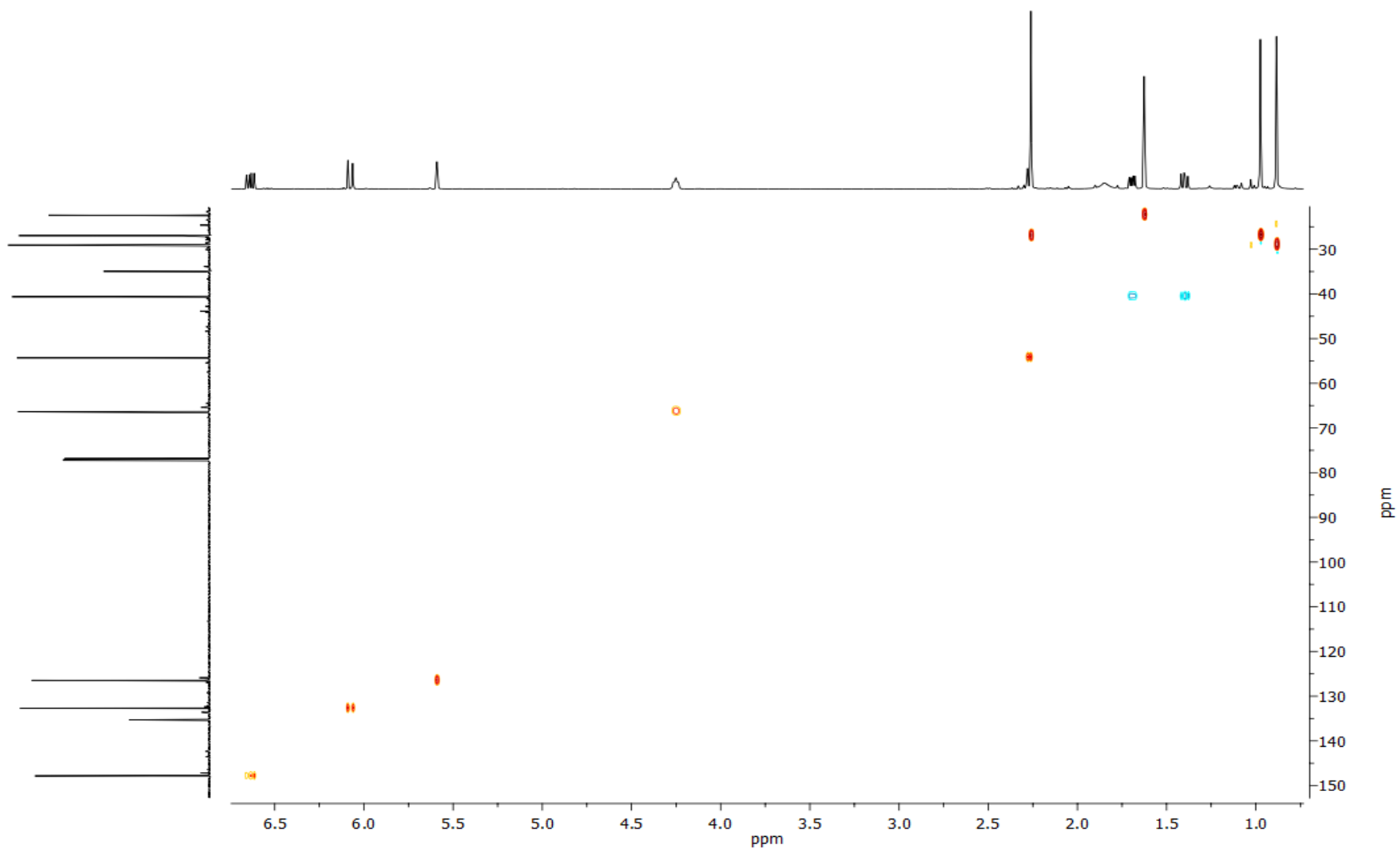
*cis*-3-hydroxy- $\alpha$ -ionone  $^1\text{H-NMR}$



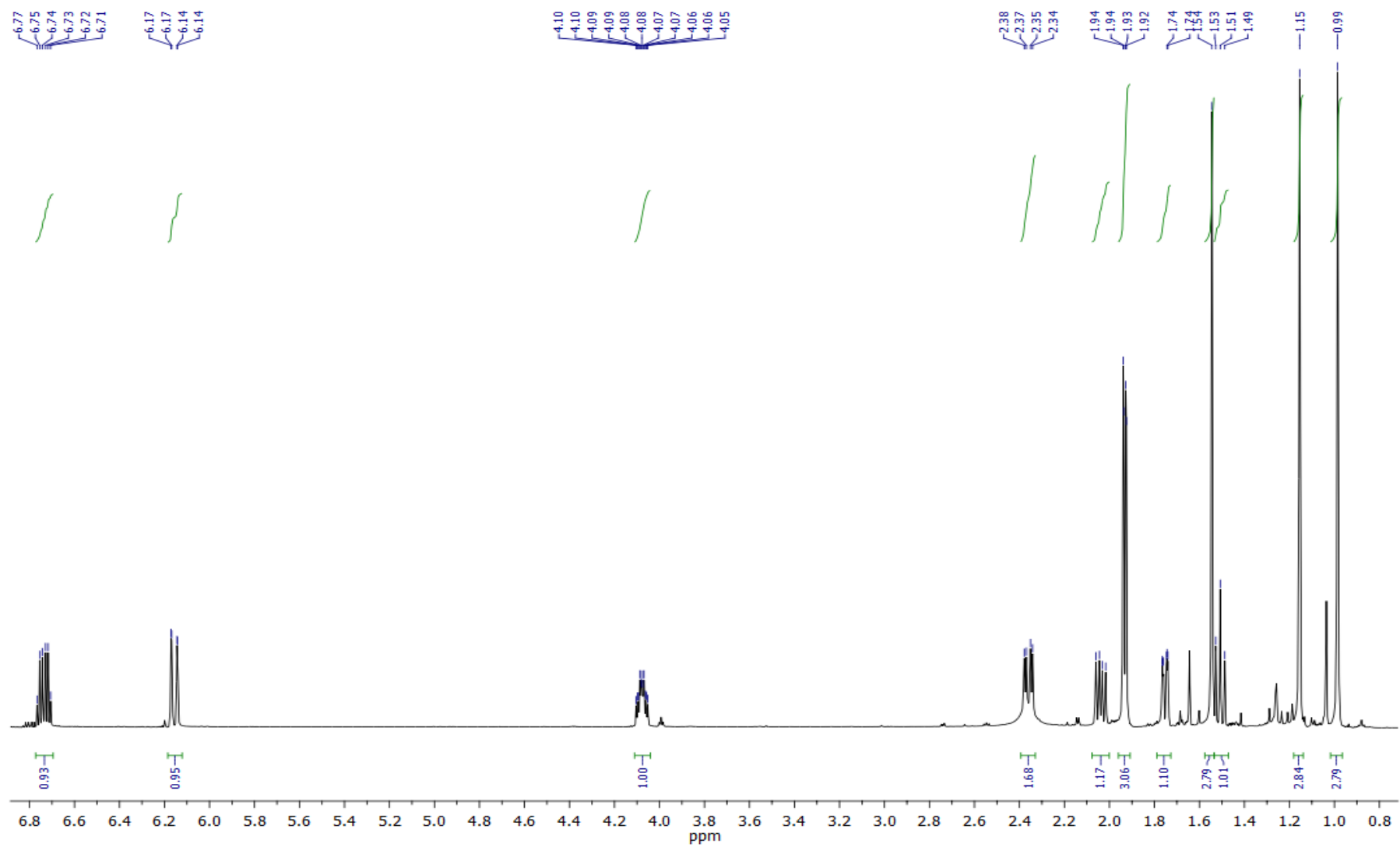
*cis*-3-hydroxy- $\alpha$ -ionone  $^{13}\text{C}$ -NMR



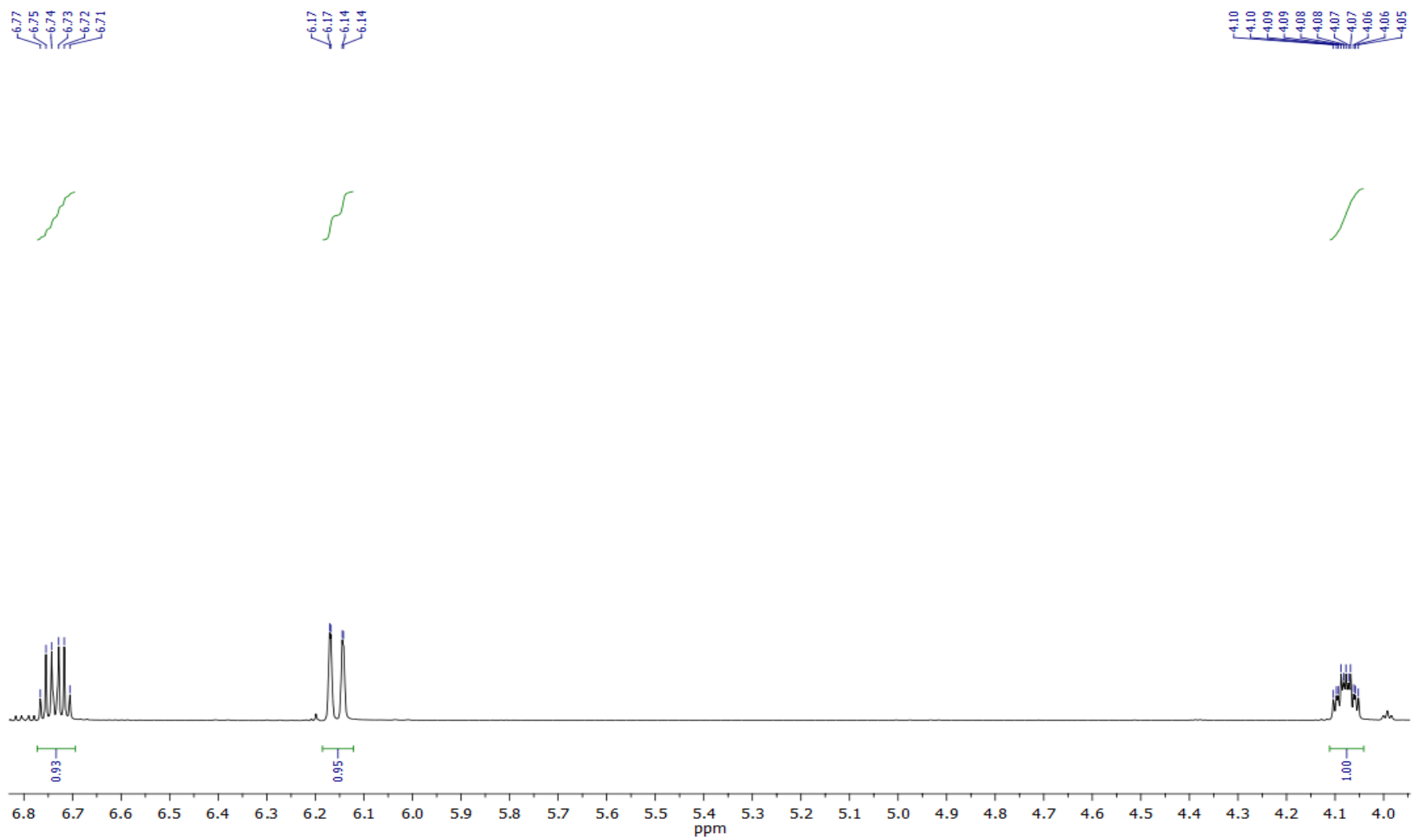
*cis*-3-hydroxy- $\alpha$ -ionone COSY



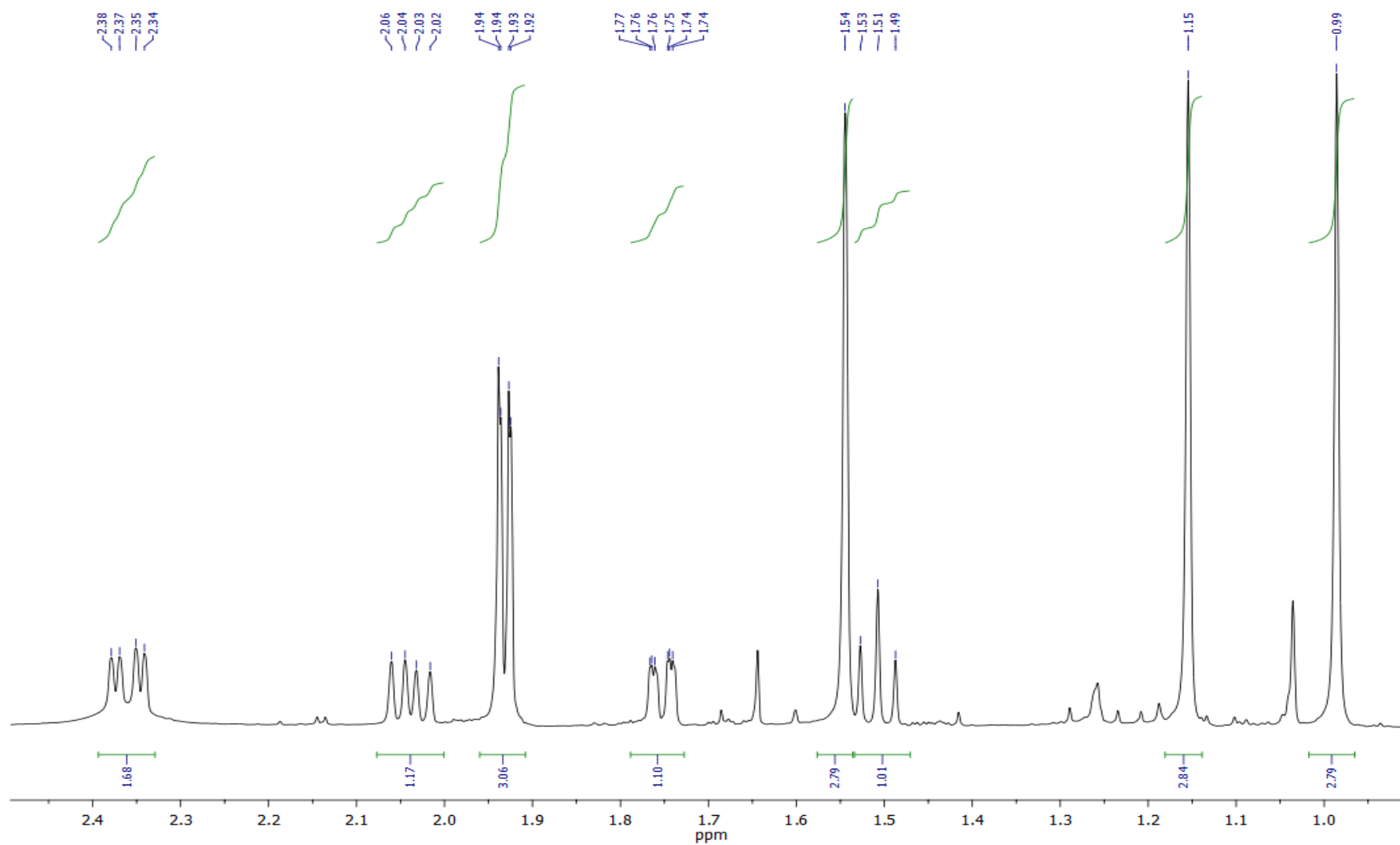
*cis*-3-hydroxy- $\alpha$ -ionone HSQC



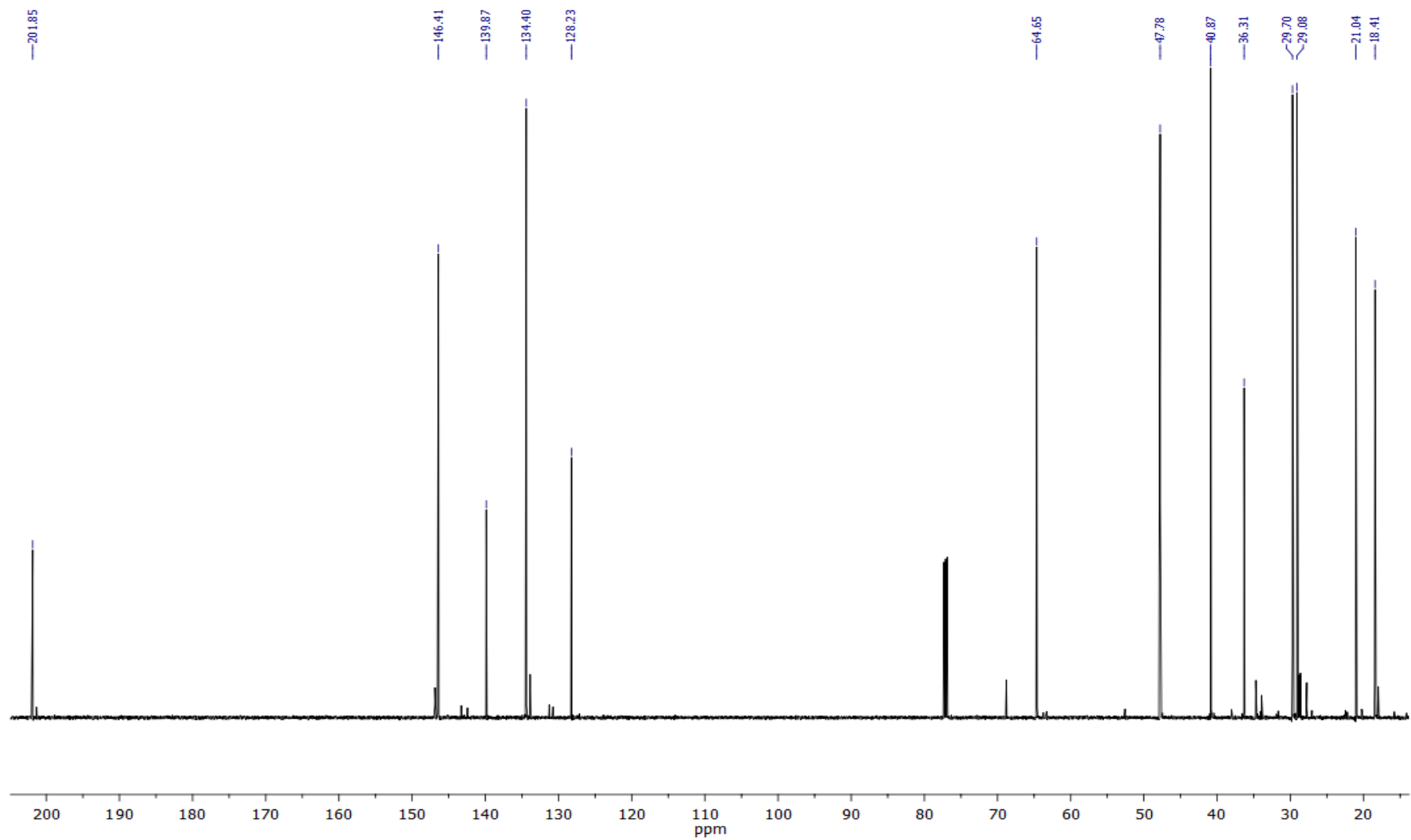
$^1\text{H-NMR}$  for 3-hydroxy- $\beta$ -damascone



$^1\text{H-NMR}$  of  $\text{sp}^2$  carbons for 3-hydroxy- $\beta$ -damascone

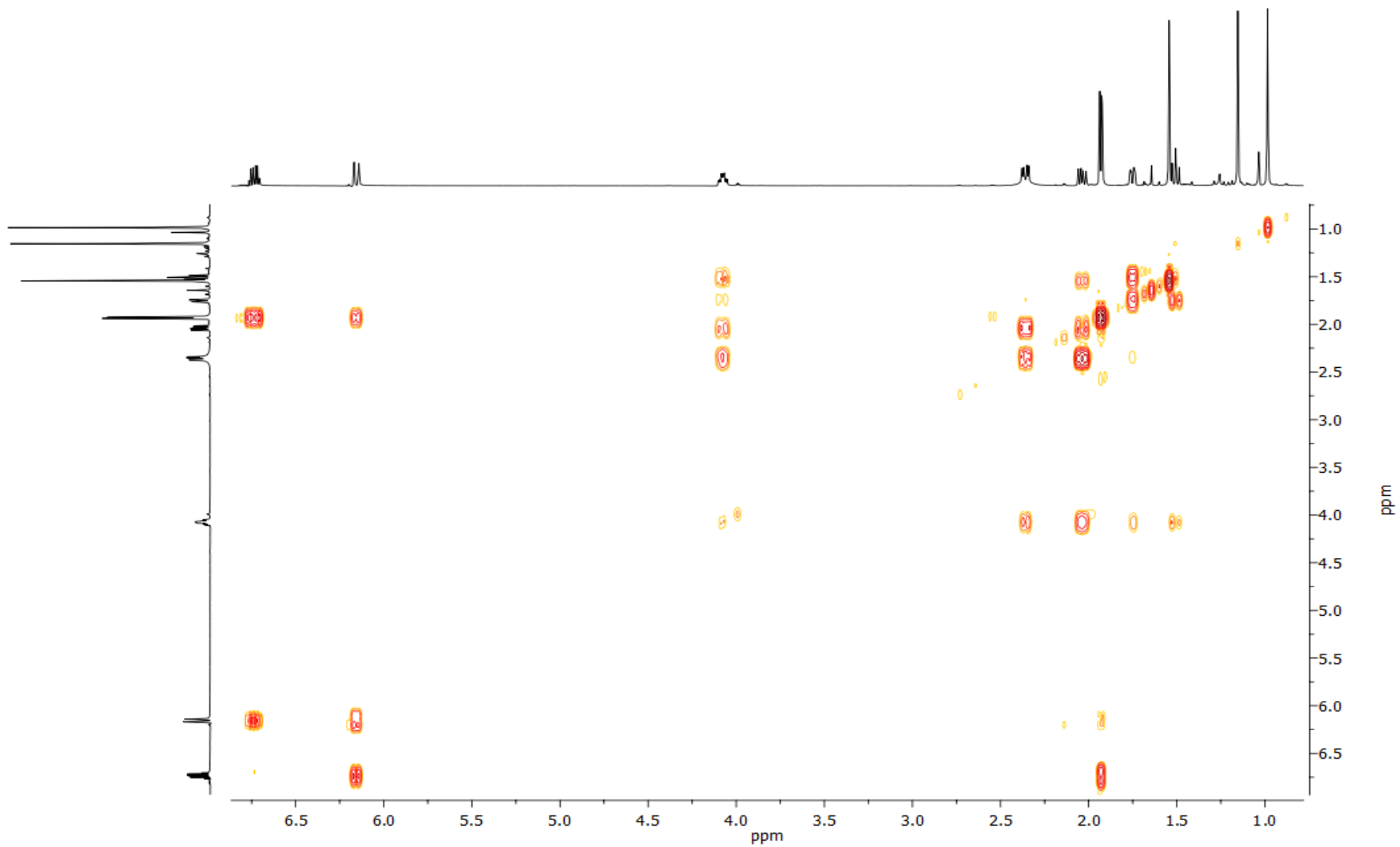


$^1\text{H-NMR}$  of  $\text{sp}^3$  carbons for 3-hydroxy- $\beta$ -damascone

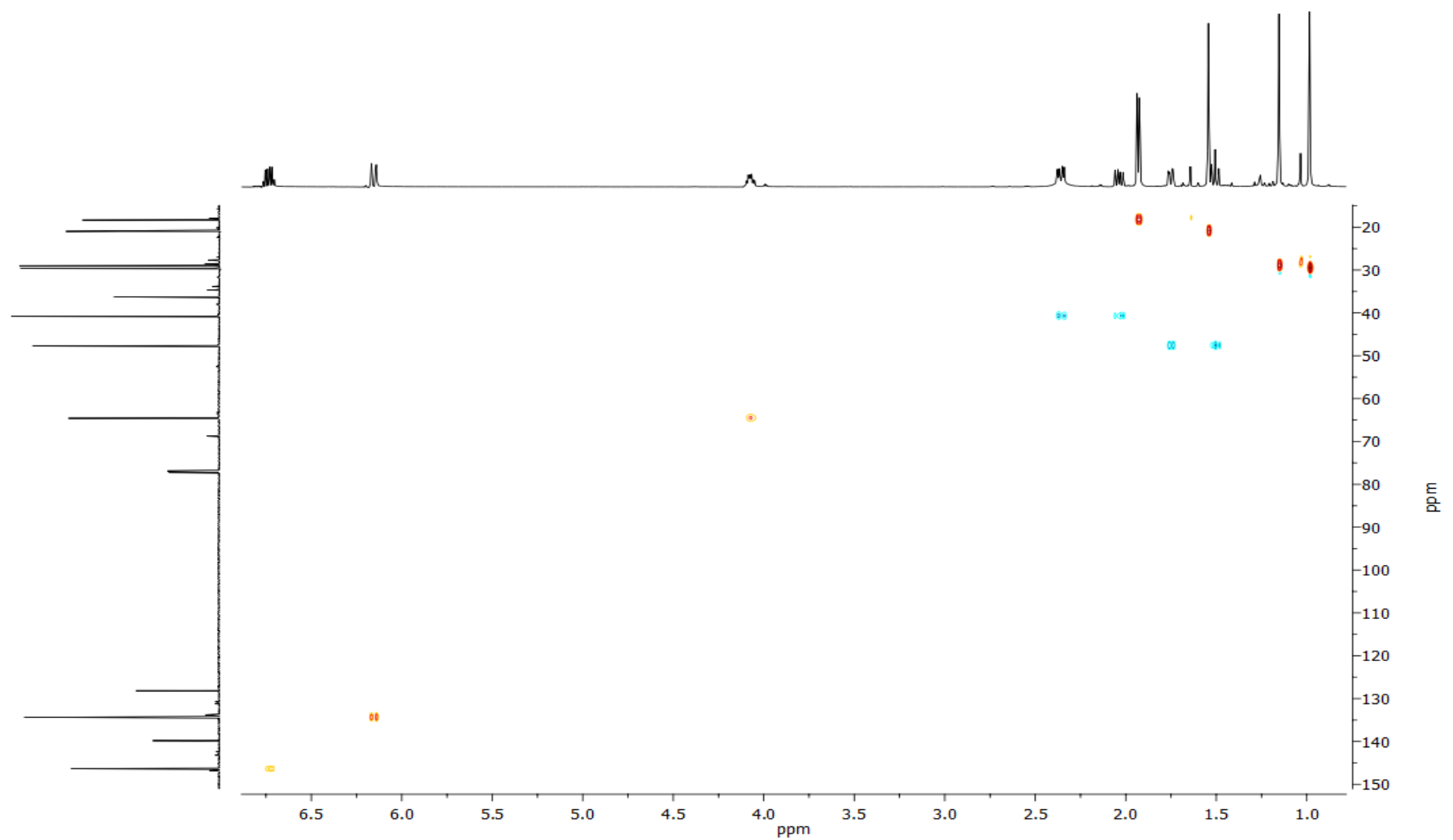


<sup>13</sup>C NMR for 3-hydroxy-β-damascone

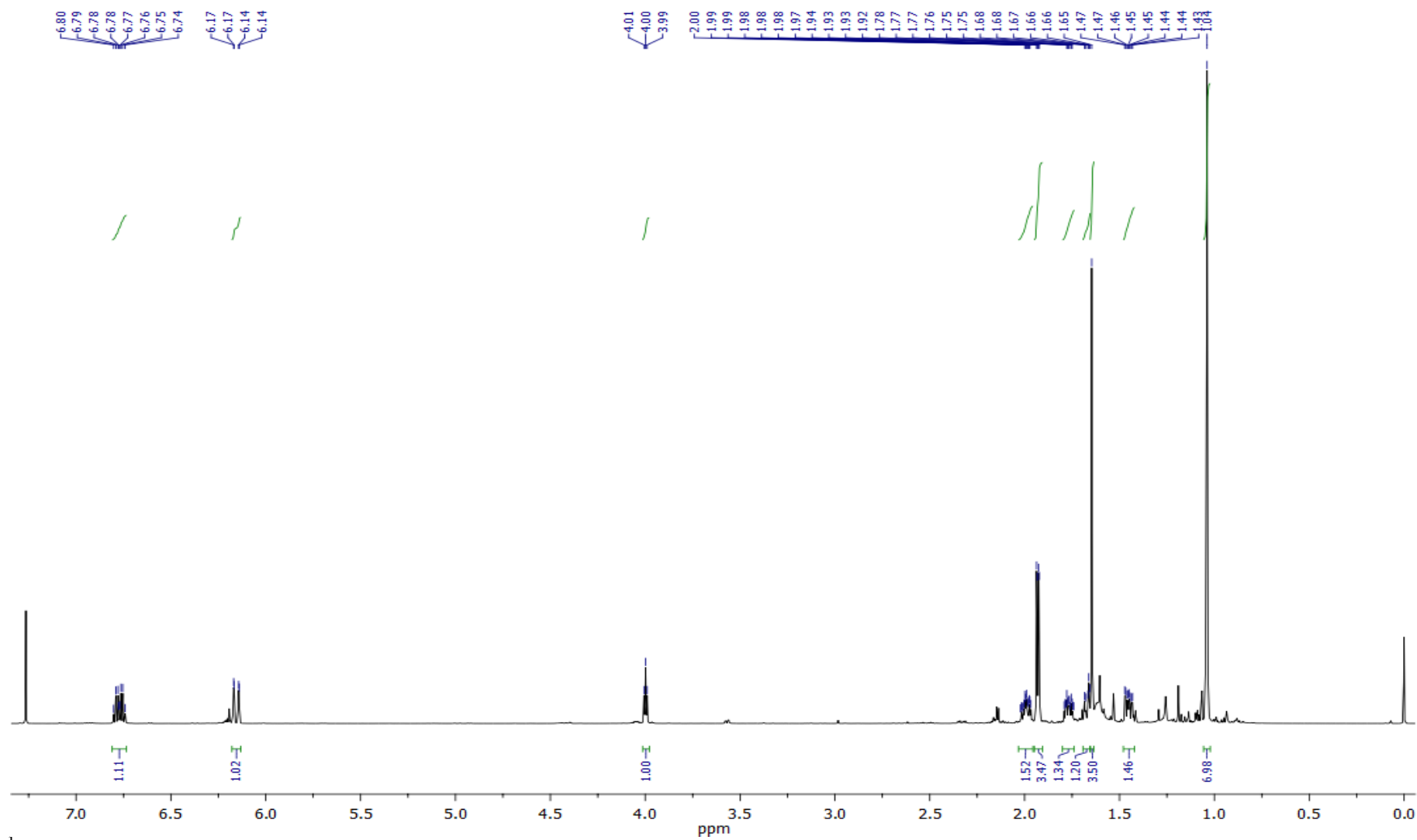




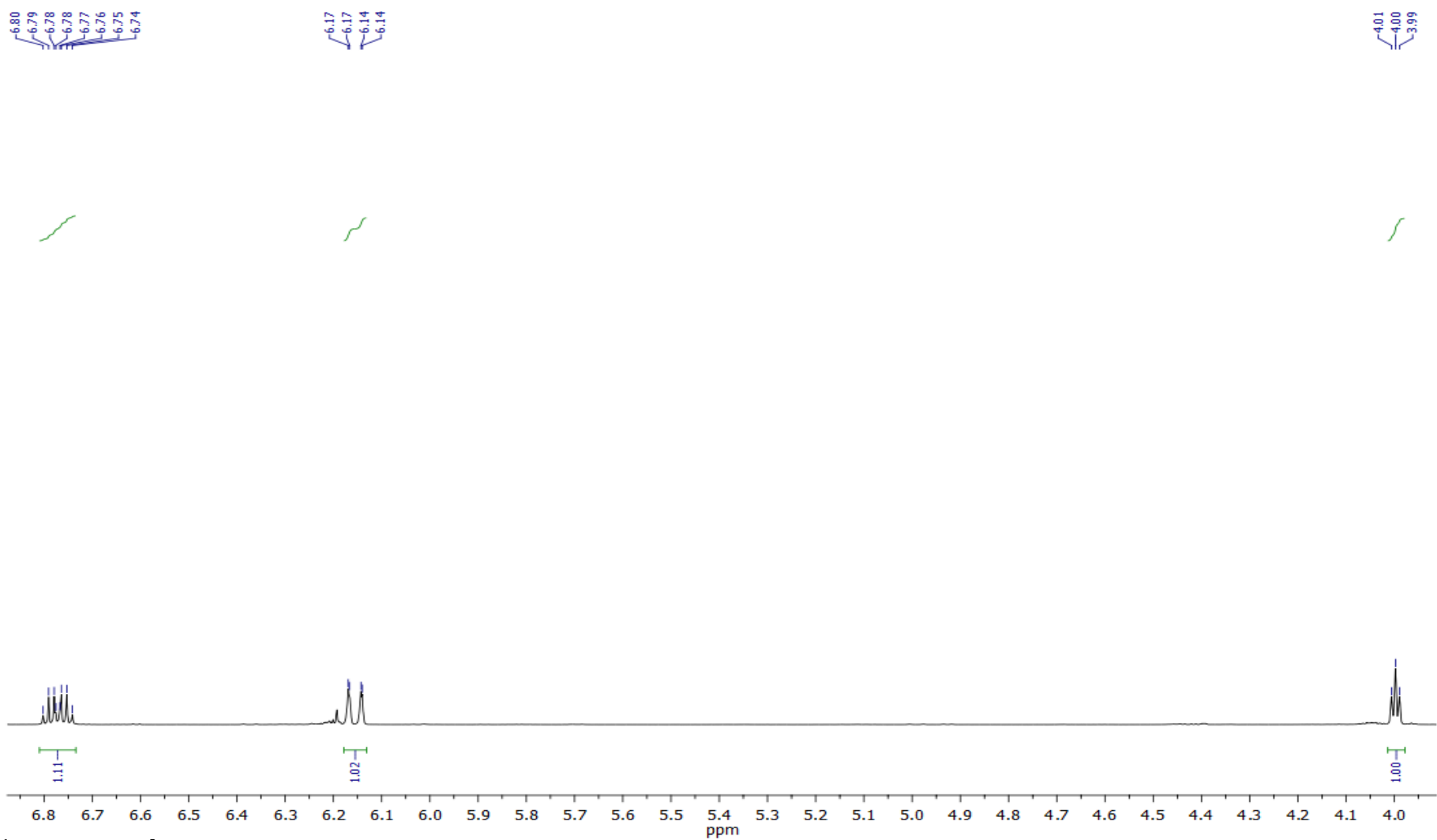
COSY-NMR for 3-hydroxy- $\beta$ -damascone



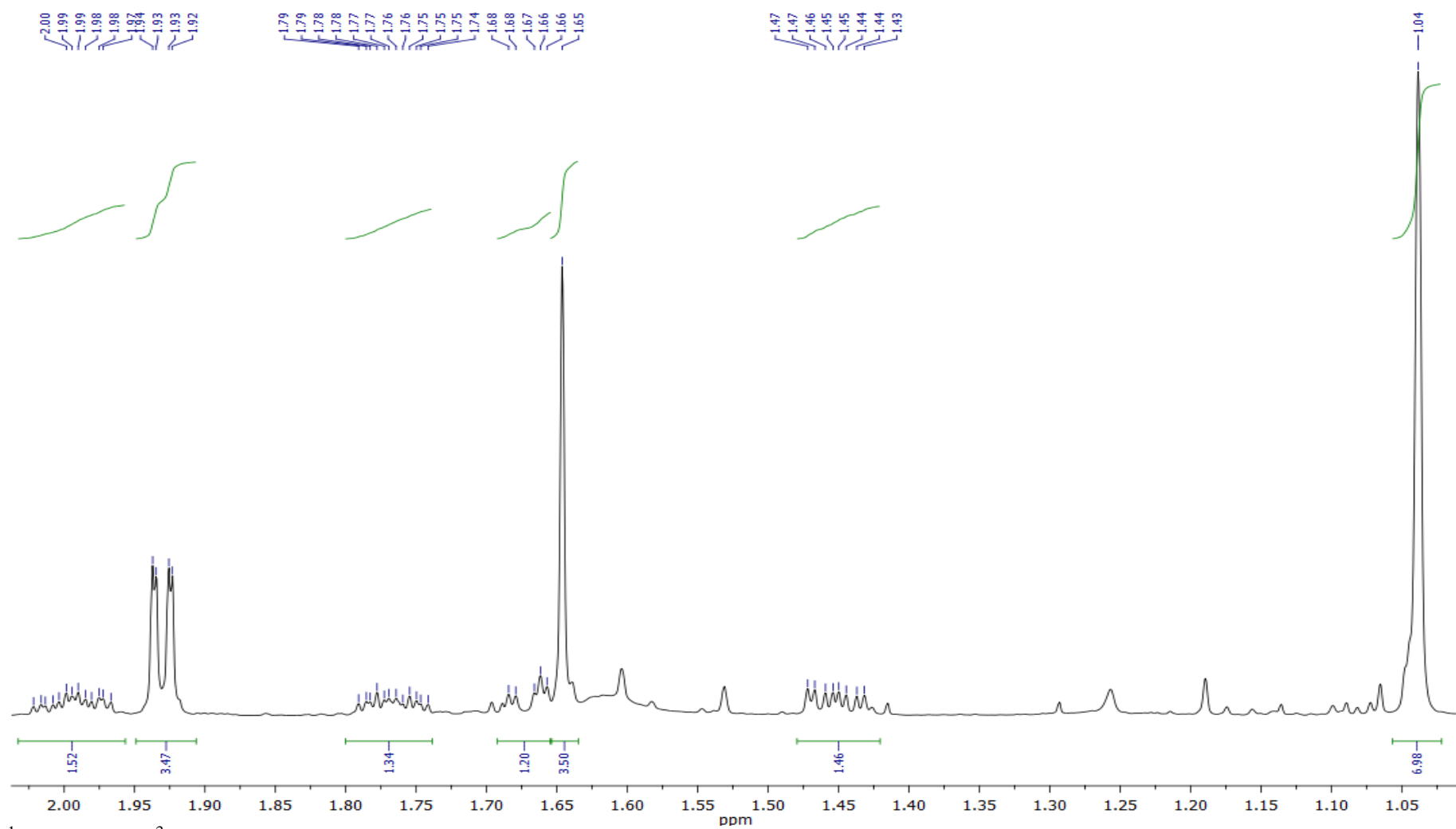
HSQC-NMR for 3-hydroxy- $\beta$ -damascone



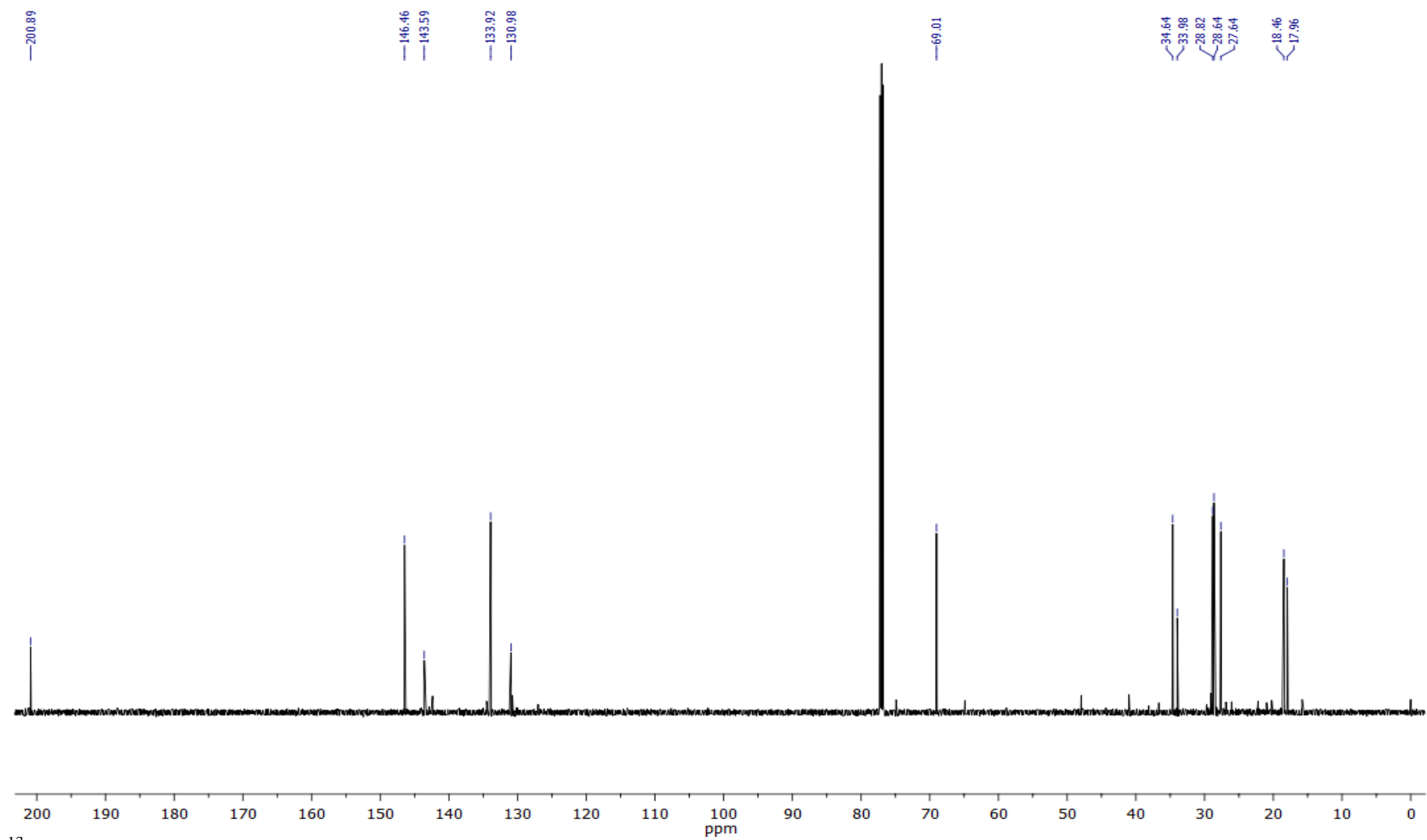
<sup>1</sup>H NMR for 4-hydroxy-β-damascone

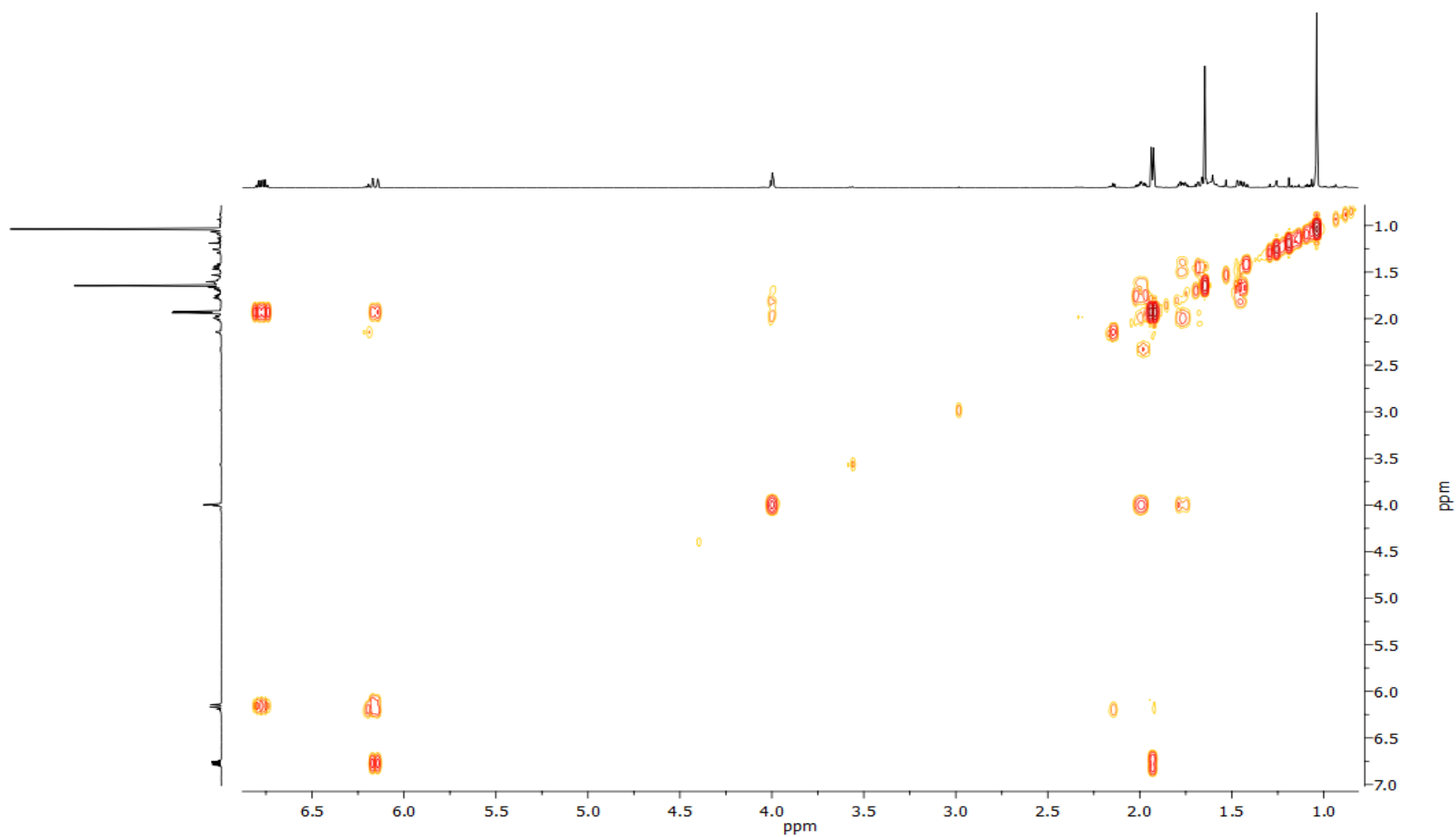


<sup>1</sup>H NMR of sp<sup>2</sup> carbons for 4-hydroxy-β-damascone

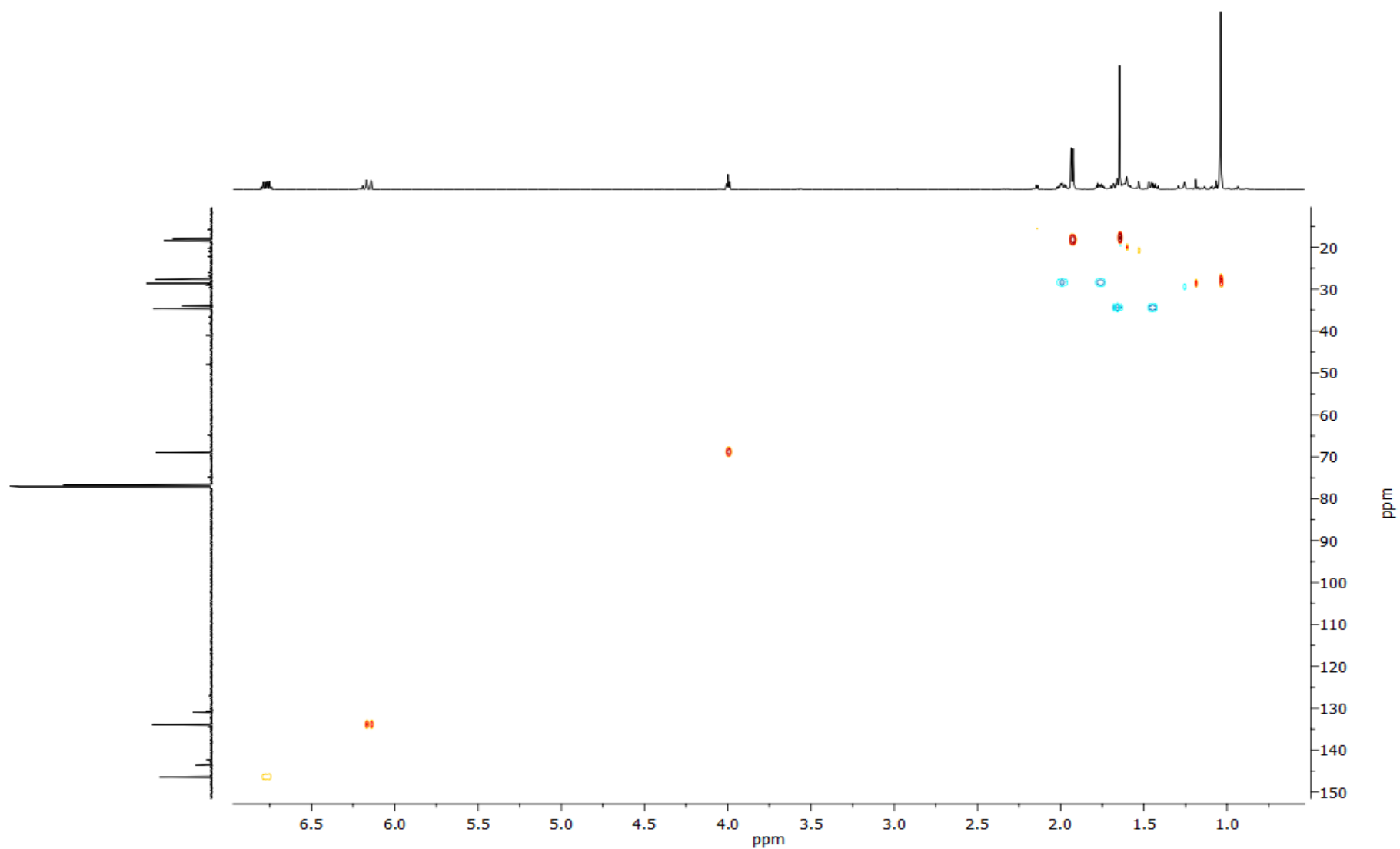


$^1\text{H}$  NMR of  $\text{sp}^3$  carbons for 4-hydroxy- $\beta$ -damascone



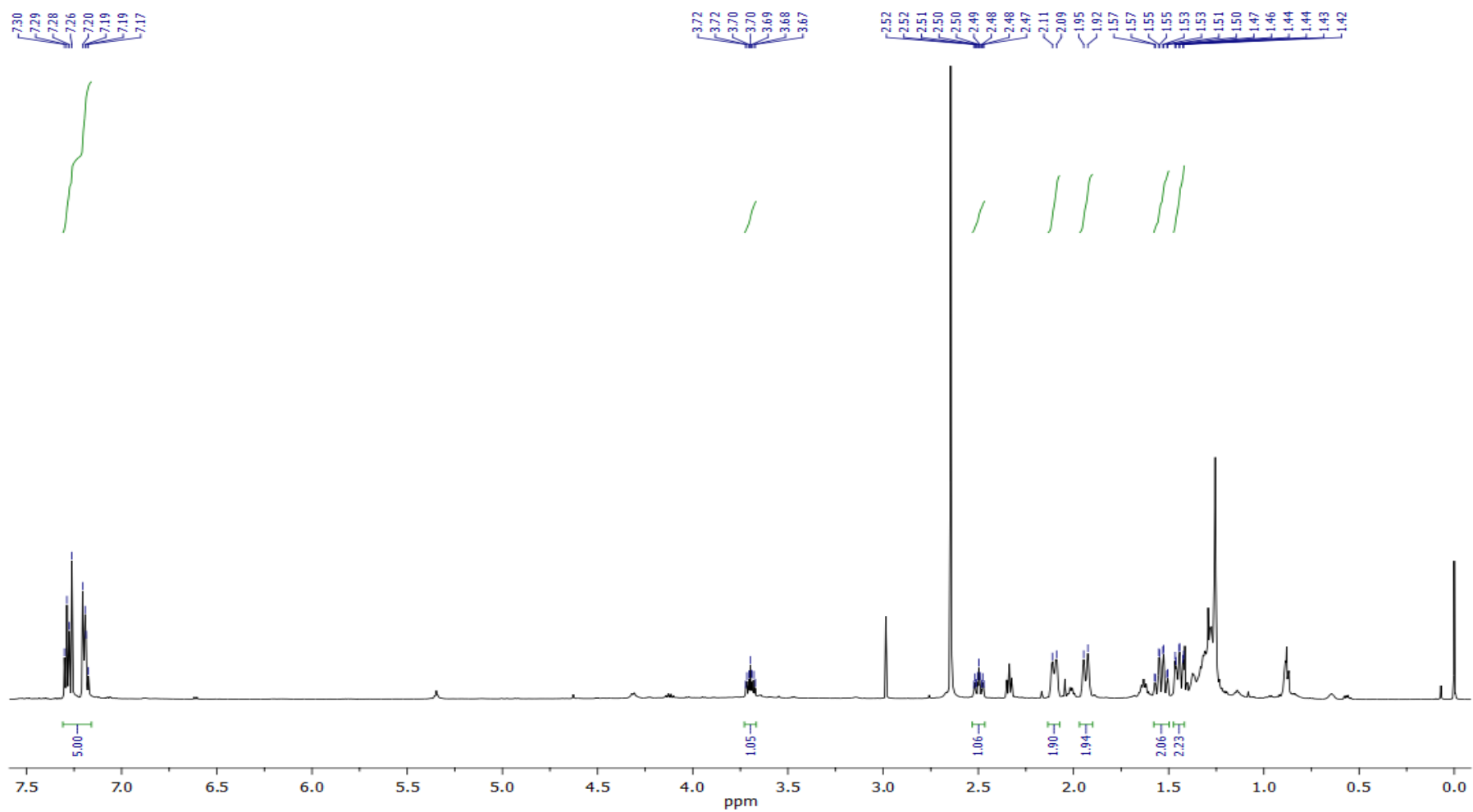


COSY for 4-hydroxy- $\beta$ -damascone

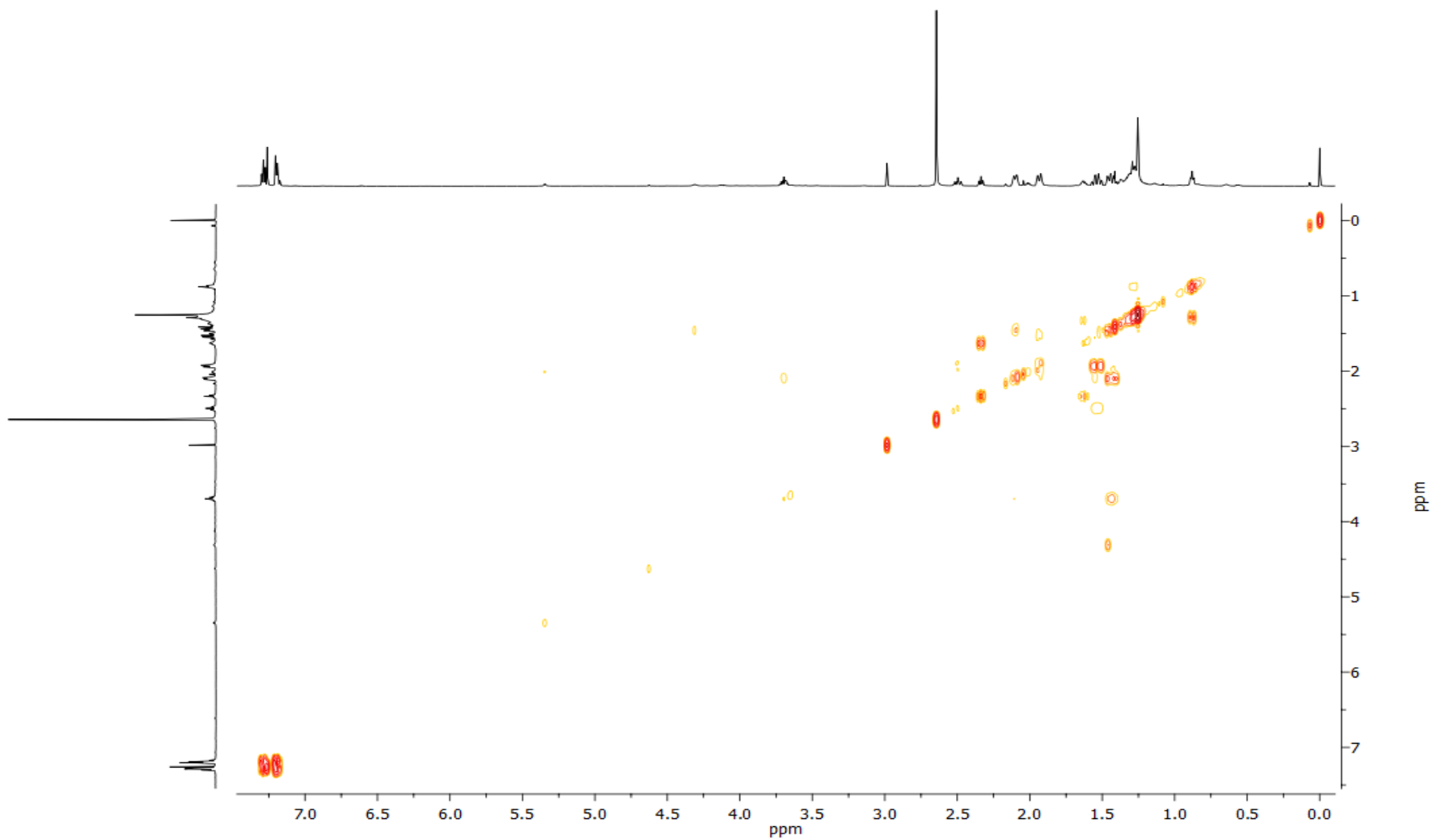


HSQC for 4-hydroxy- $\beta$ -damascone





$^1\text{H-NMR}$  of *trans*-4-phenylcyclohexanol



COSY of *trans*-4-phenylcyclohexanol