# Studies toward a prebiotic protometabolism

Adam J. Coggins

UCL

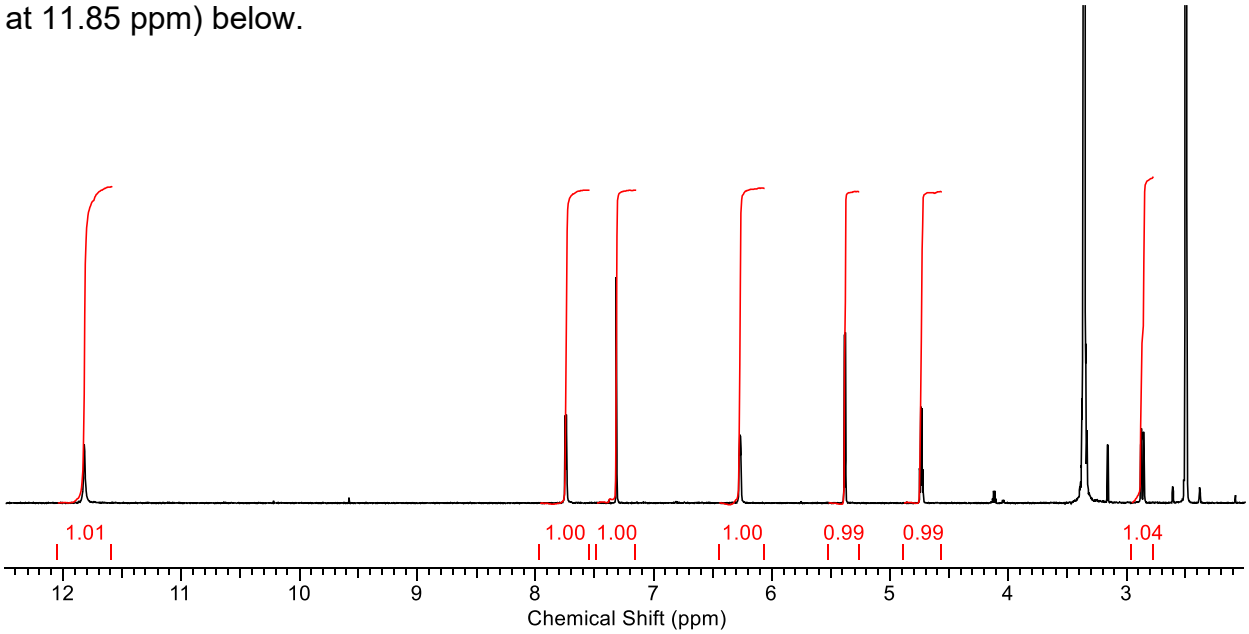
*Thesis Submitted to University College London*
*for the Degree of Doctor of Philosophy*

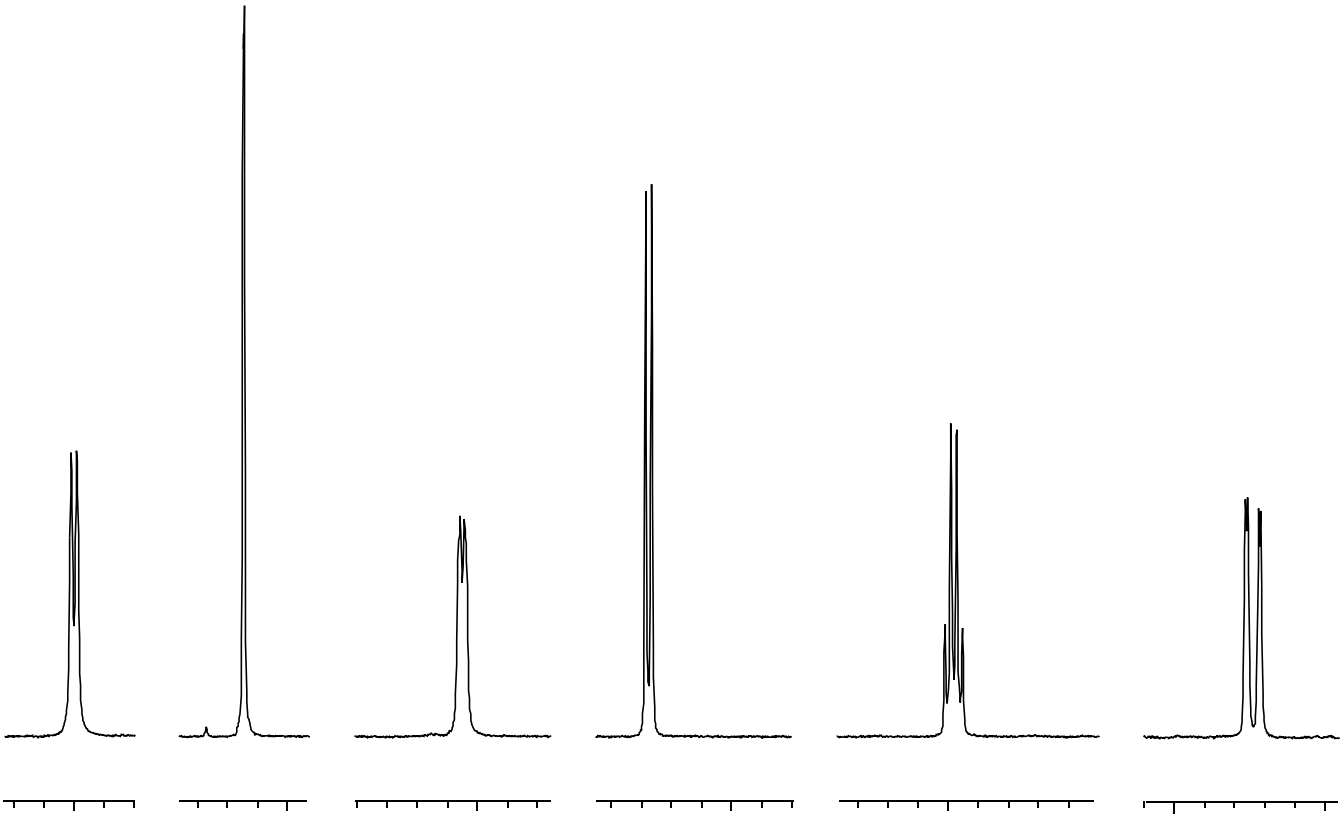
# Appendix – part 4

## ***Purine-Precursor Tethering by Imine Tautomerisation***

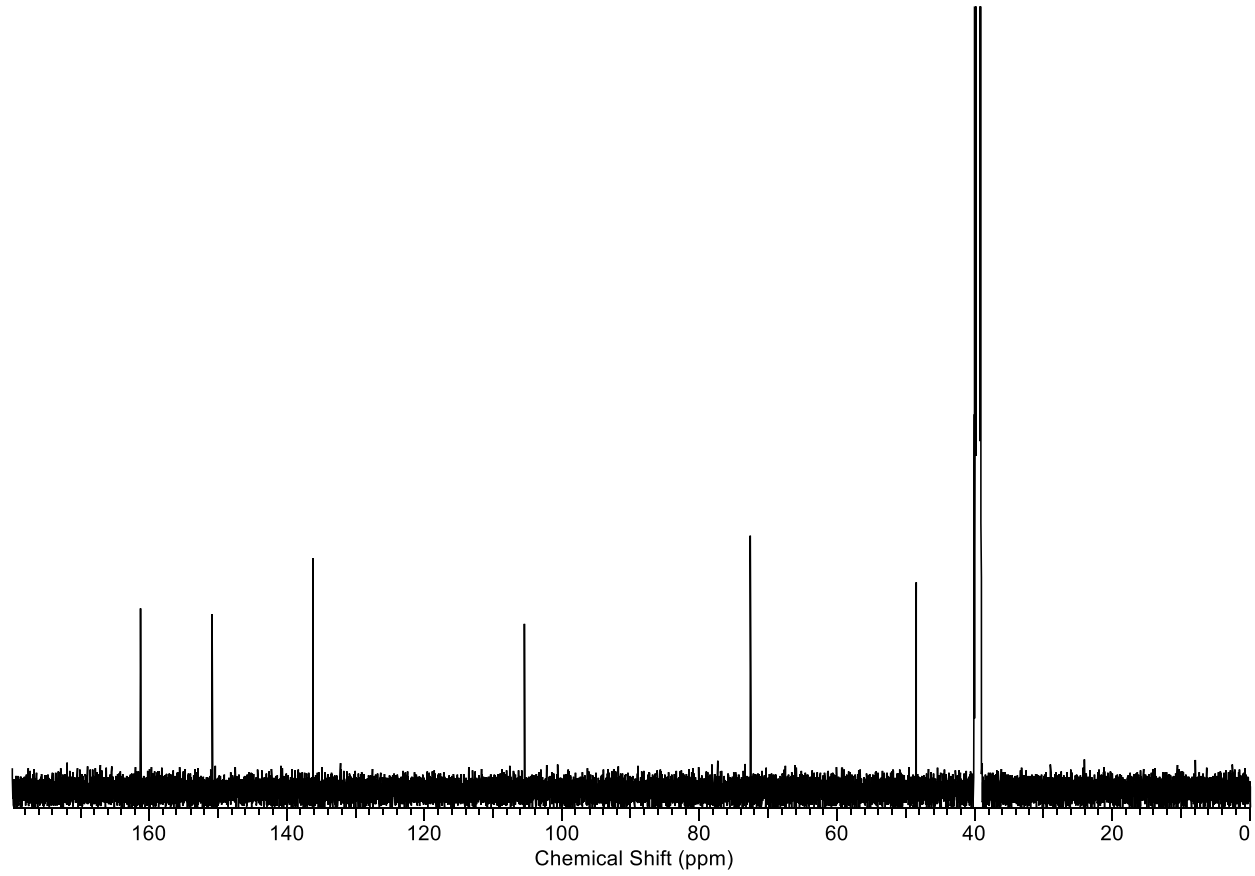
Azepinomycin (**212**)

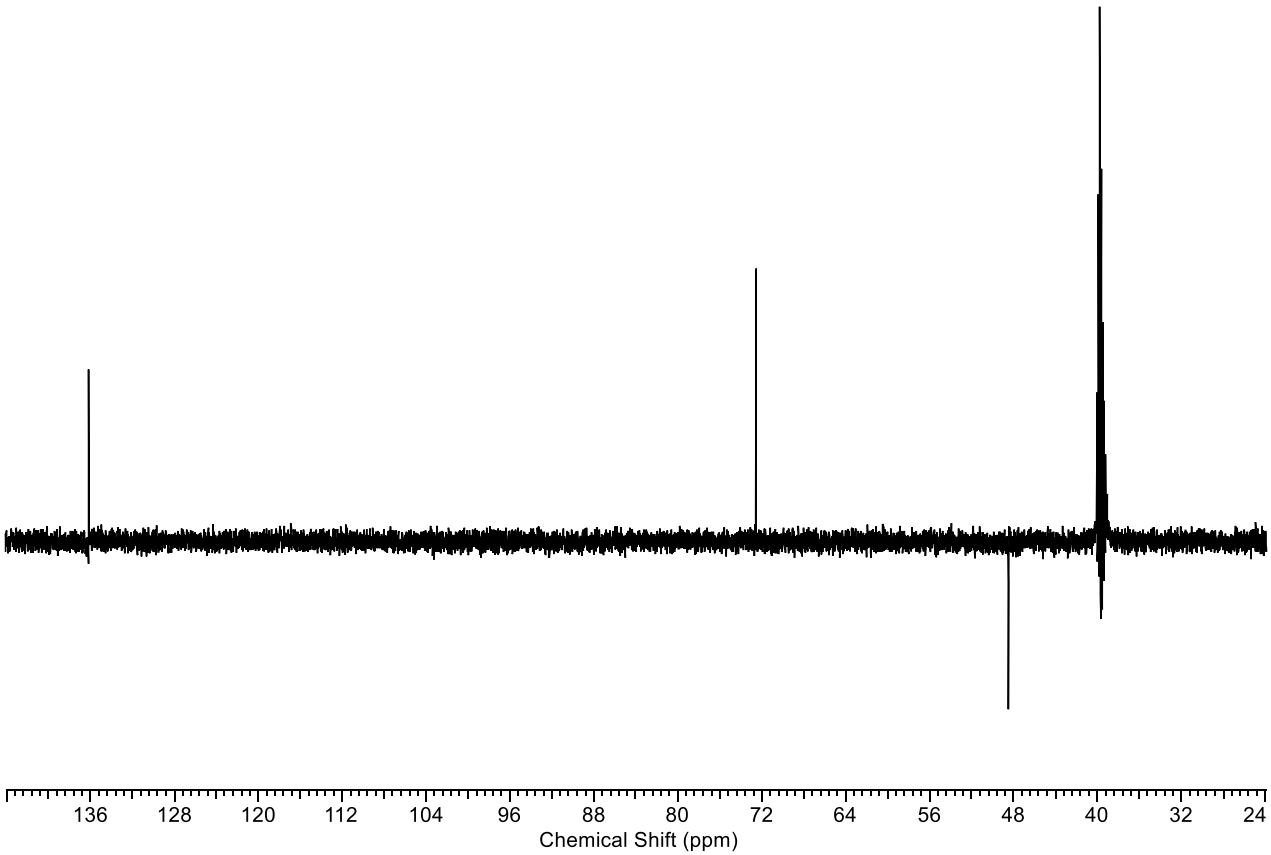
**A333**: $^{1}$H NMR spectrum (600 MHz, {DMSO-d6}, 2.0 – 12.5 ppm) of **212**, with expansions (excluding peak at 11.85 ppm) below.



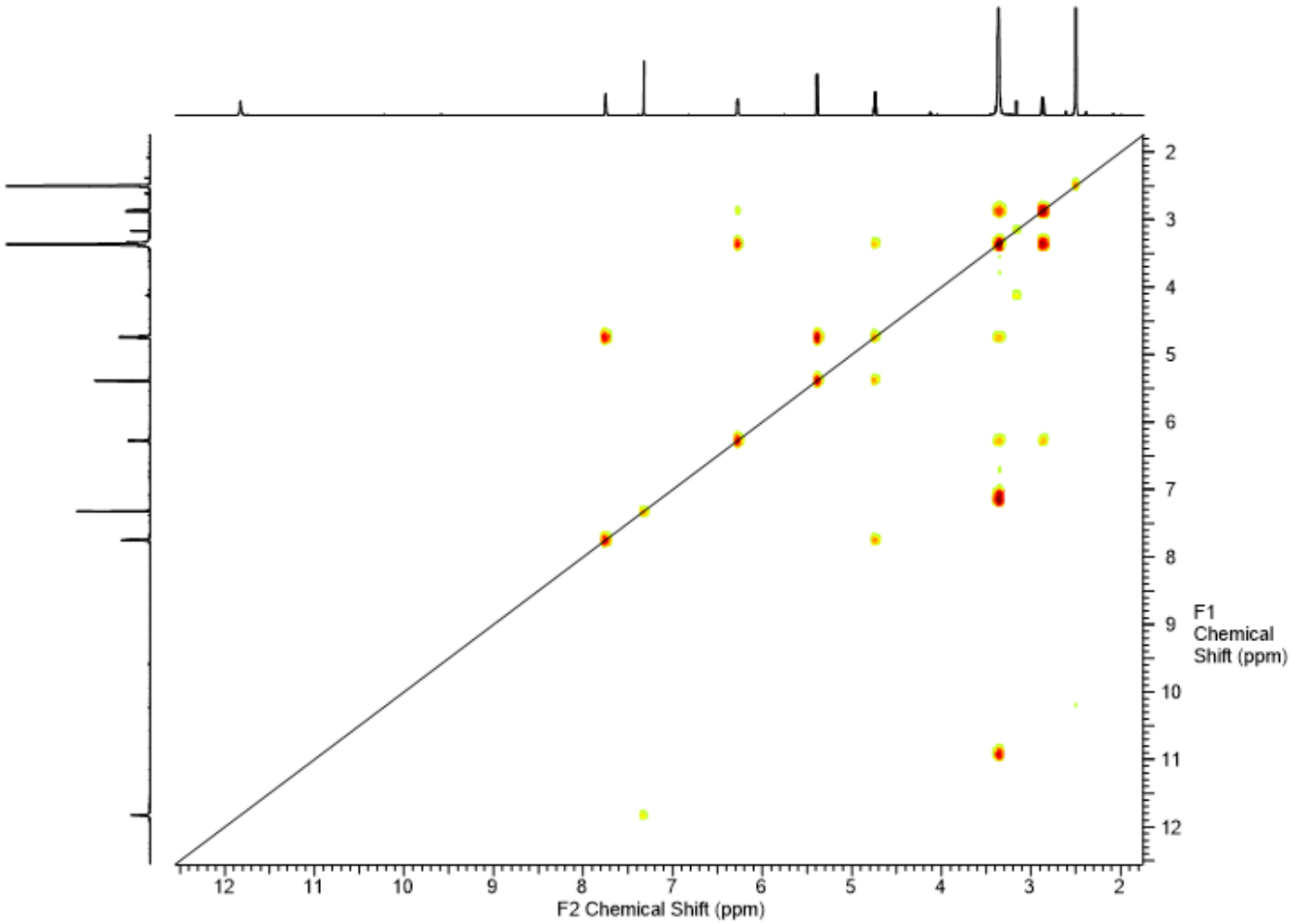



**A334**: $^{13}C$ NMR spectrum (151 MHz, {DMSO-d6}, 0 – 180 ppm) of **212** with DEPT135 spectrum (24 -144 ppm) below.

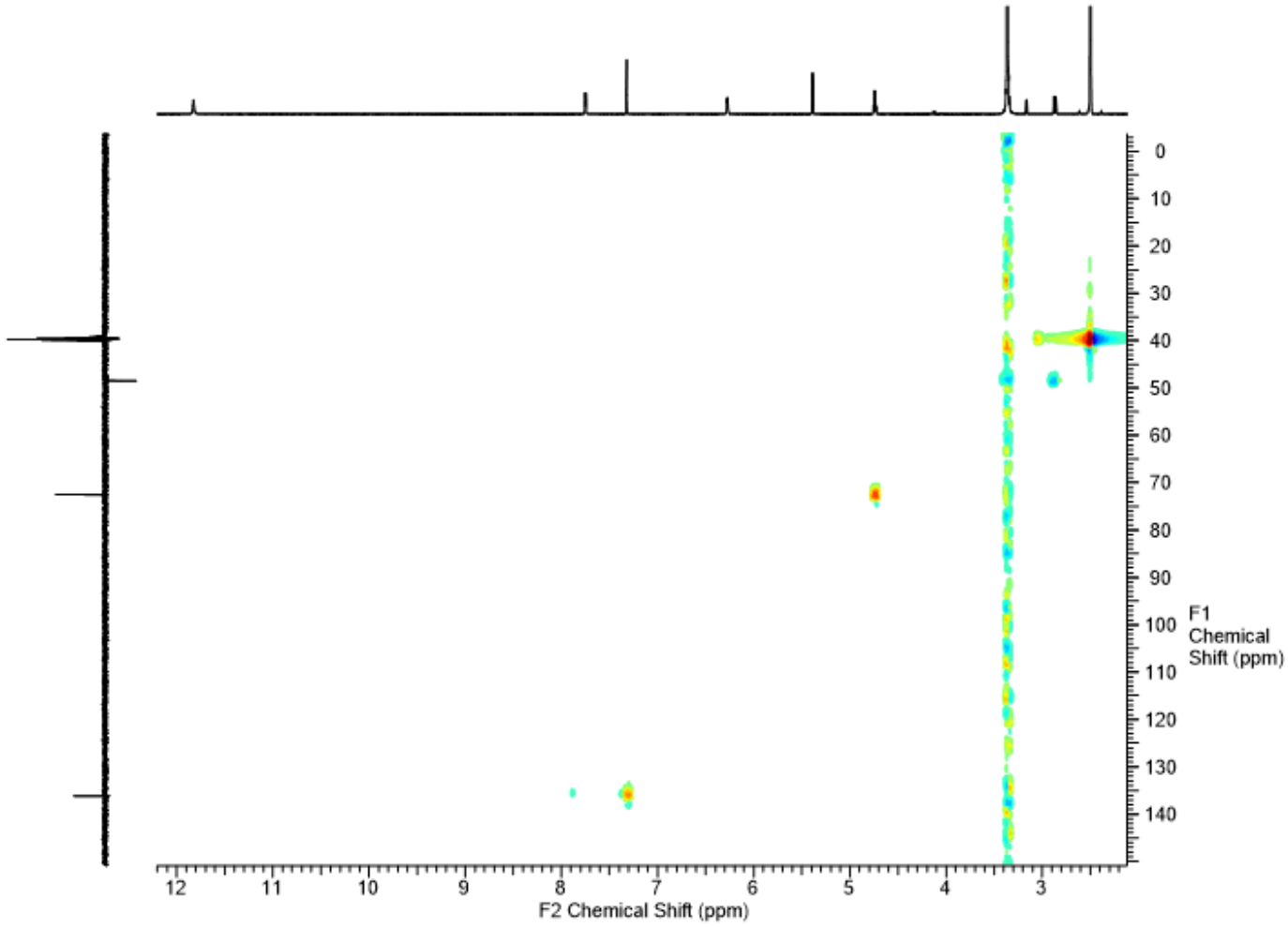


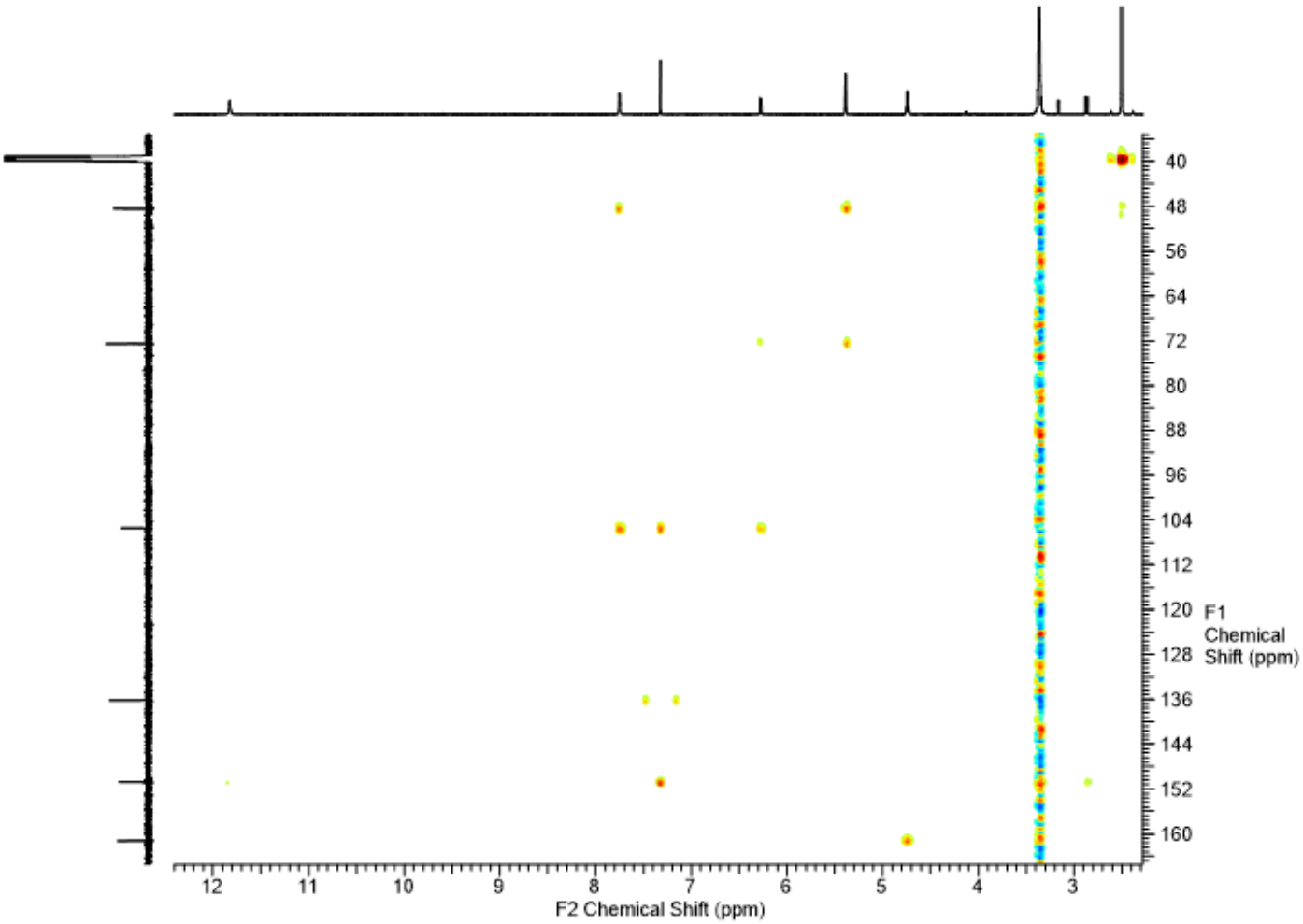
**A335**: $^{1}H$-$^{1}H$ COSY NMR spectrum (600 MHz, {DMSO-d6}) of **212.**



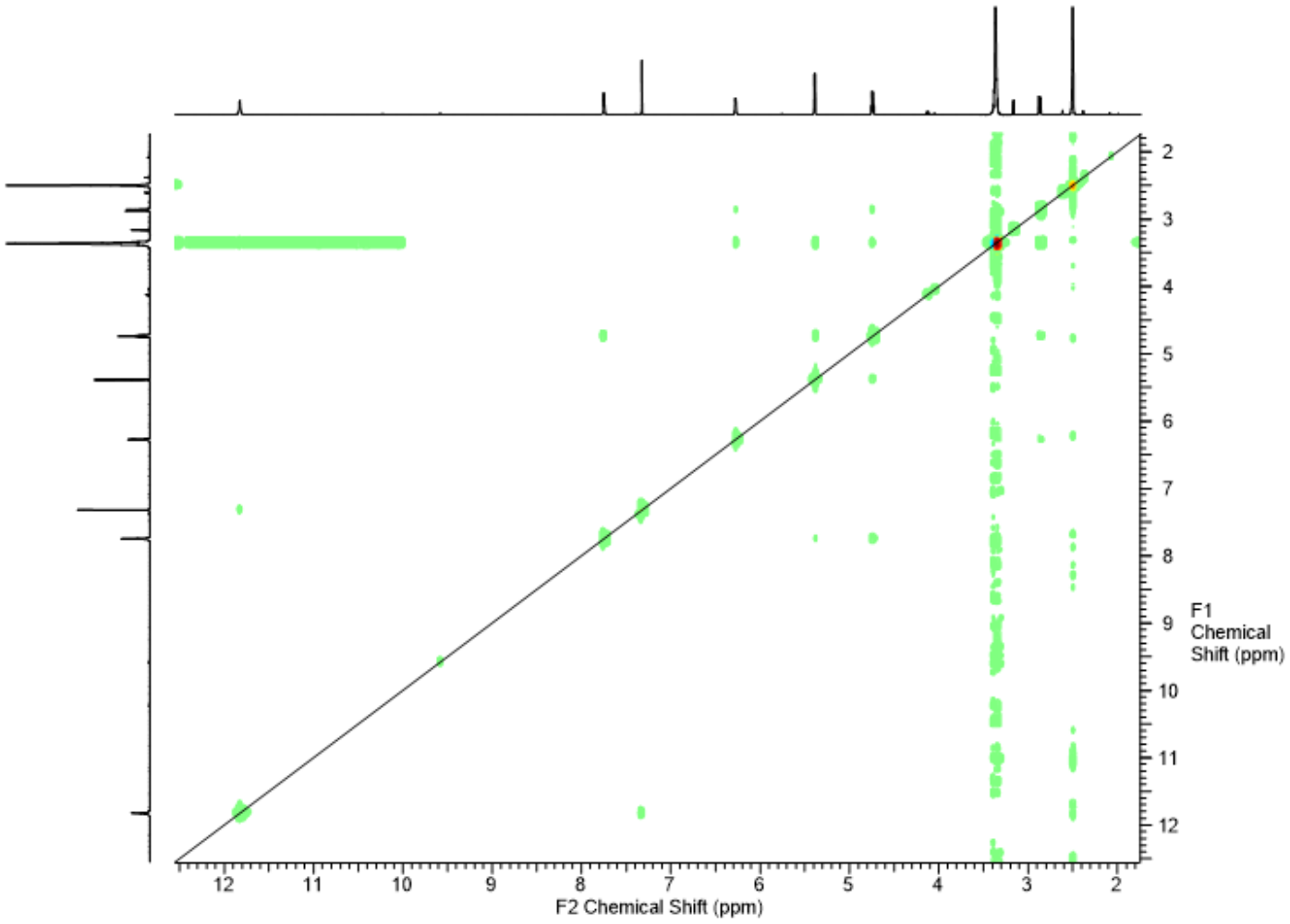
**A336**: $^{1}H$-$^{13}C$ HSQC NMR spectrum (600 MHz, {DMSO-d6}) of **212.**



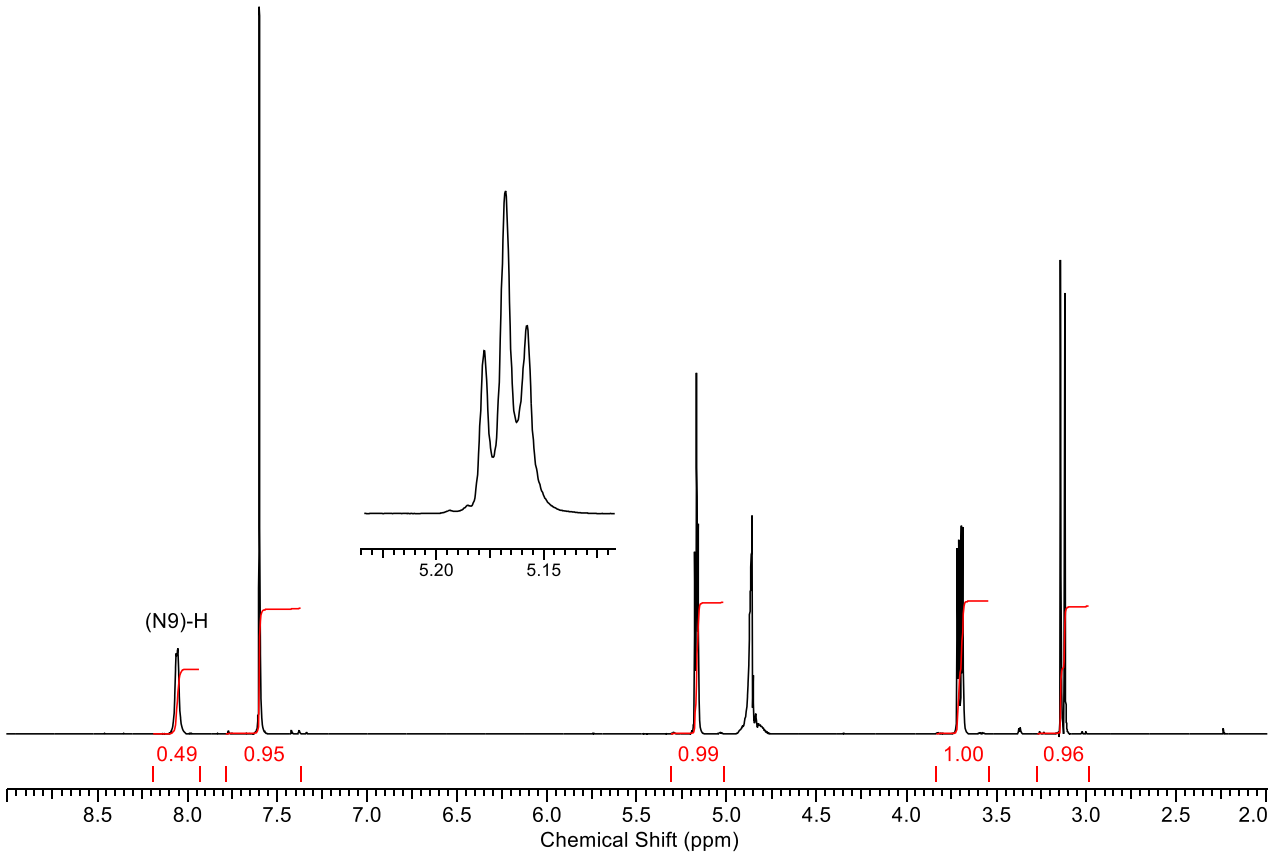
**A337**: $^{1}H$-$^{13}C$ HMBC NMR spectrum (600 MHz, {DMSO-d6}) of **212.**



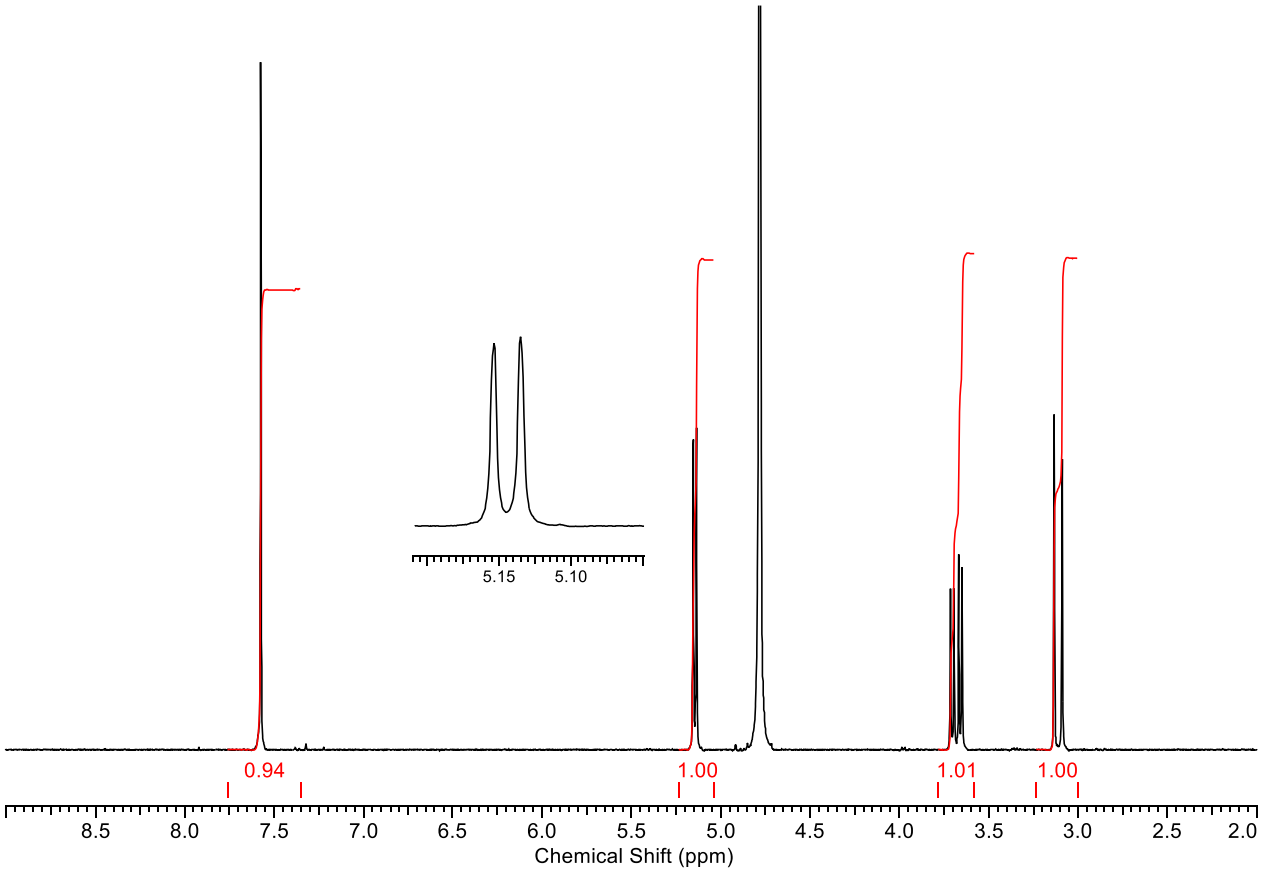
**A338**: $^{1}H$-$^{1}H$ NOESY NMR spectrum (600 MHz, {DMSO-d6}) of **212.**



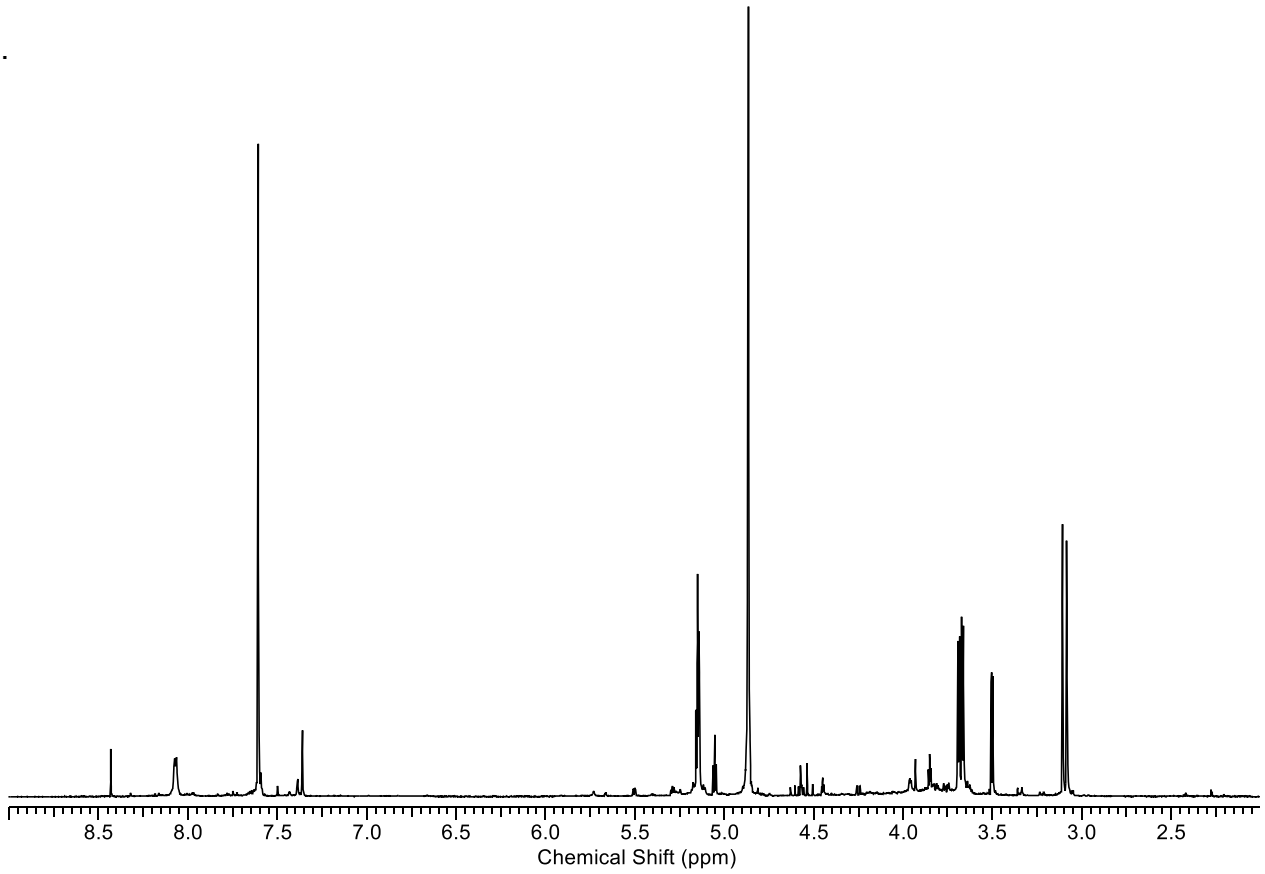
**A339**: $^{1}$H NMR spectrum (600 MHz, {$H_2O/D_2O$. 9:1}, 2.0 – 9.0 ppm) of **212** with expansion overlaid.



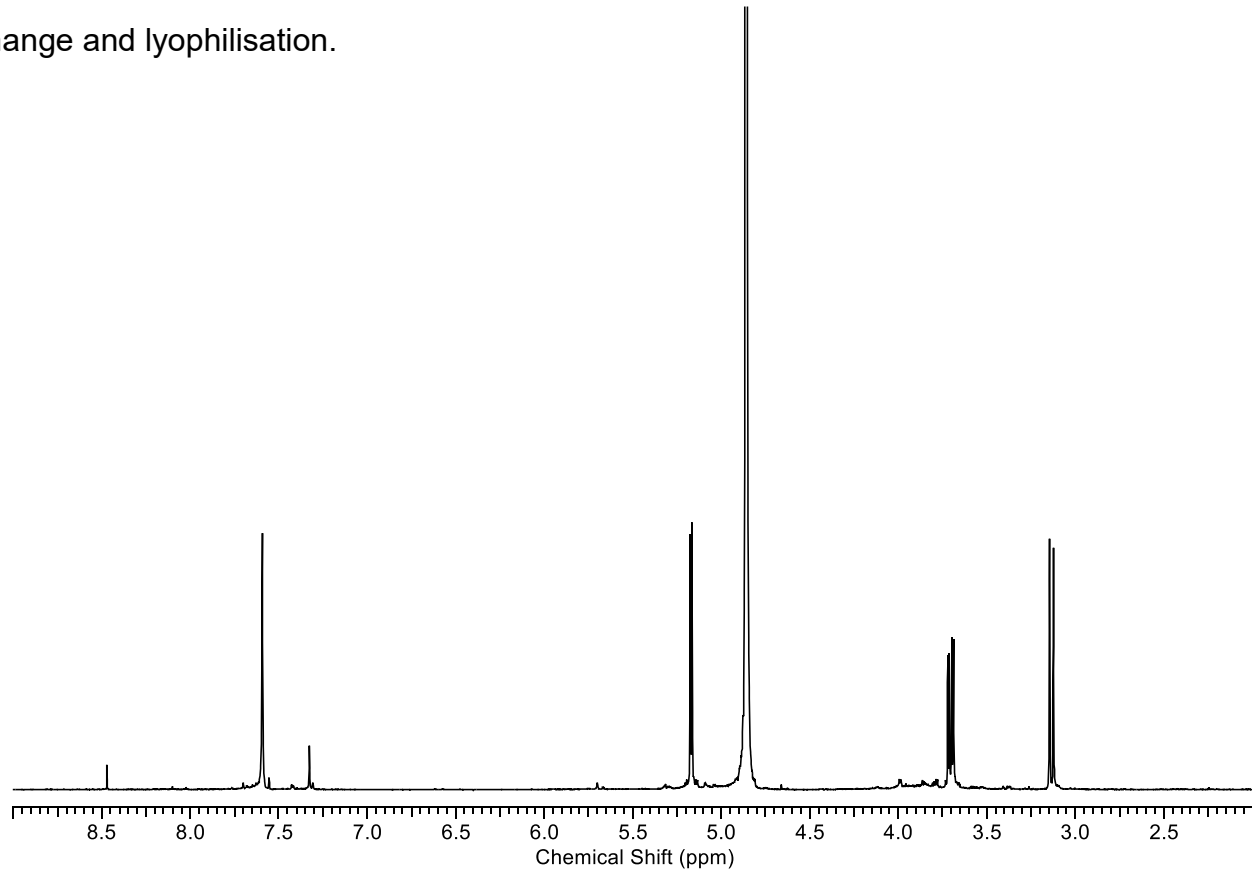

**A340**: $^{1}$H NMR spectrum (600 MHz, {$D_2O$}, 2.0 – 9.0 ppm) of **212** with expansion overlaid.



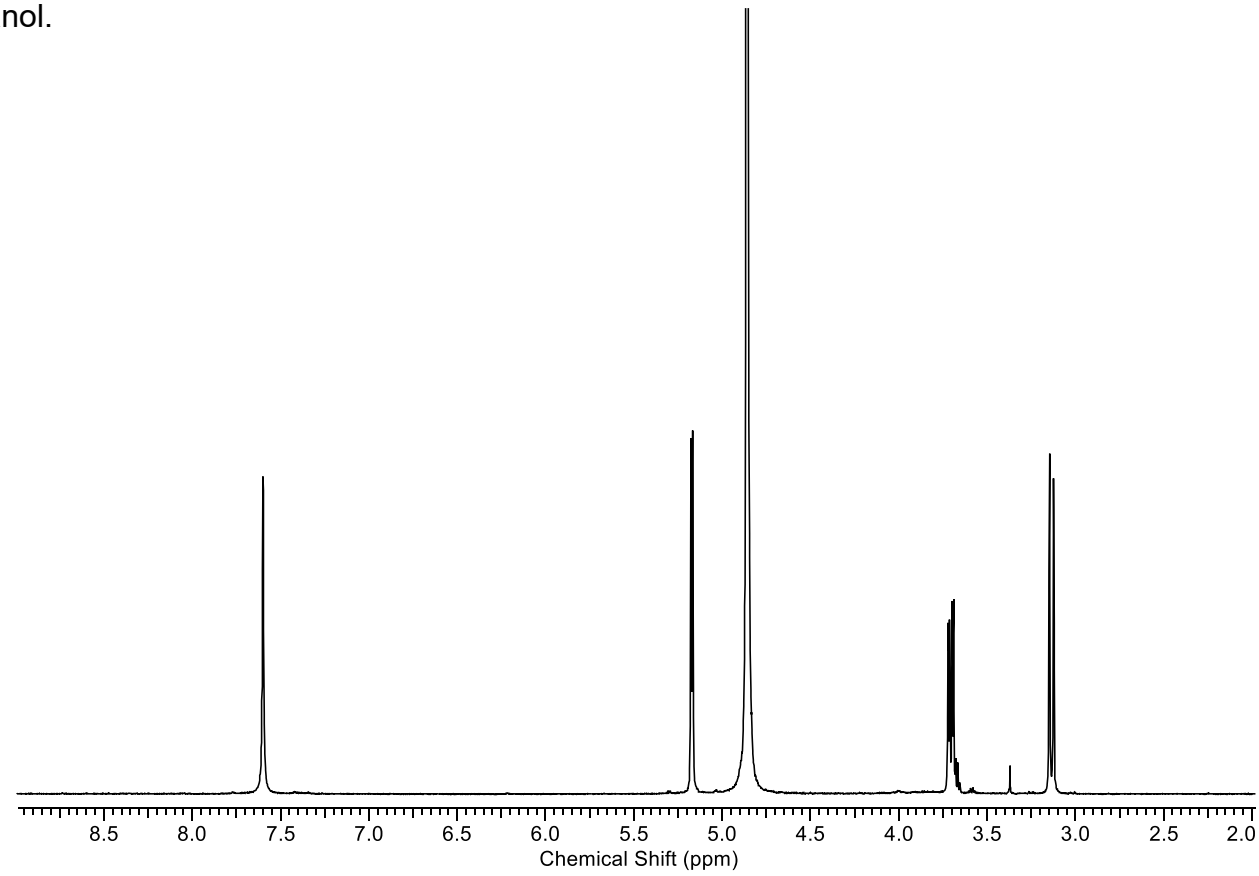
**A341**: $^{1}$H NMR spectrum (400 MHz, {$H_2O/D_2O$. 9:1 }, 2.0 – 9.0 ppm) of **212** reaction mixture after 28 h.



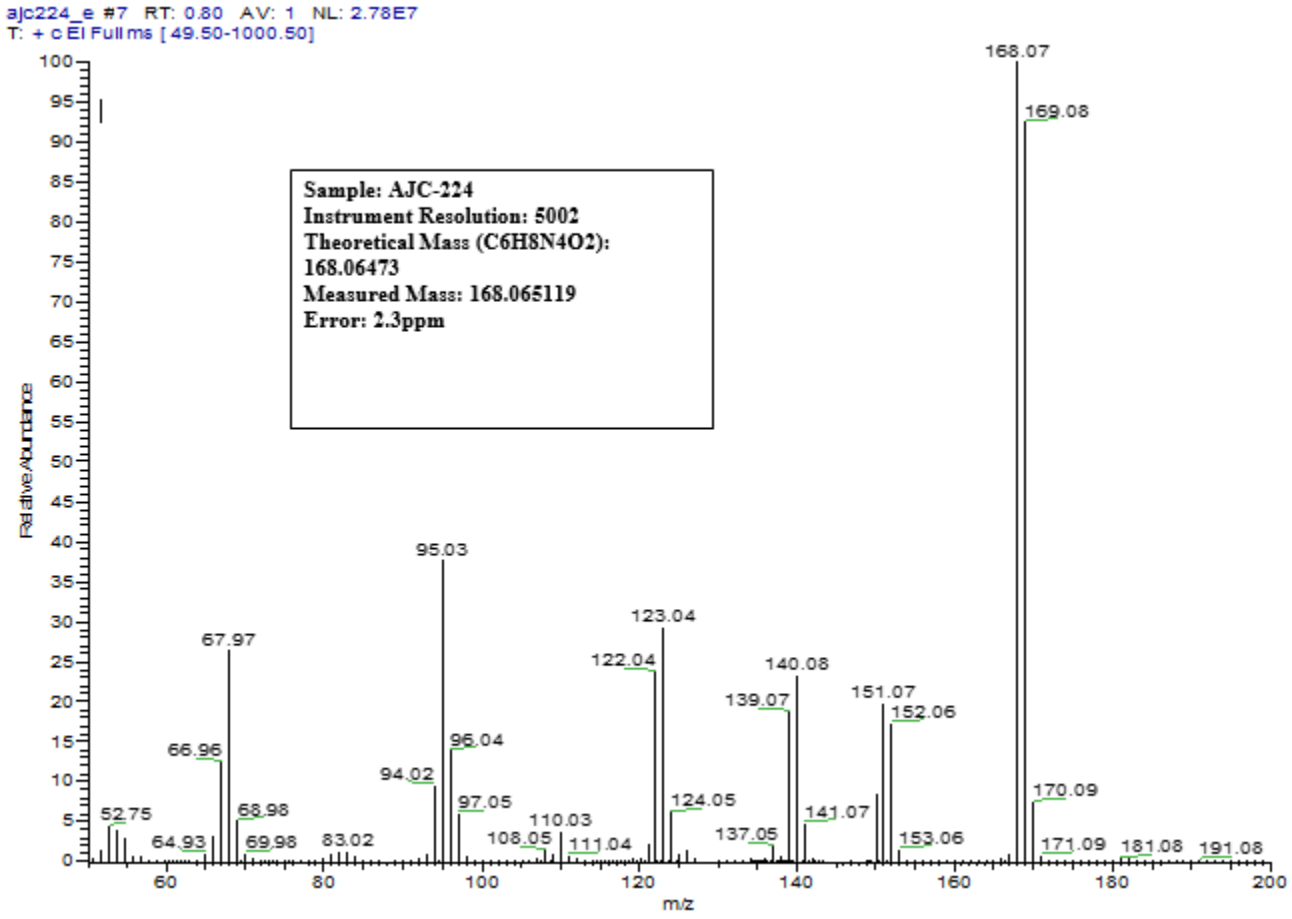

**A342**: $^{1}$H NMR spectrum (400 MHz, {$D_2O$}, 2.0 – 9.0 ppm) of **212** crude product after Dowex™ resin exchange and lyophilisation.



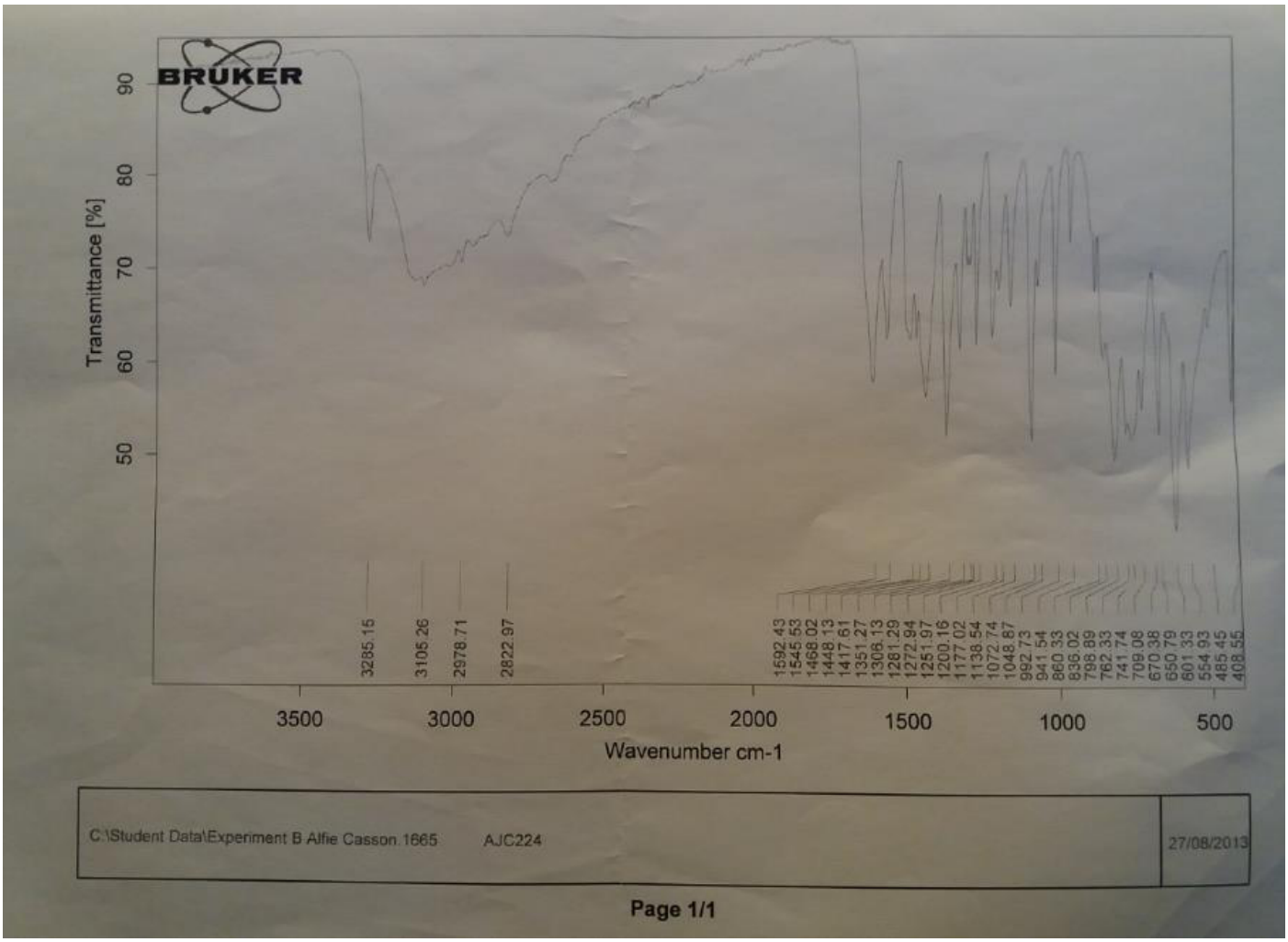
**A343**: $^{1}H$ NMR spectrum (400 MHz, {$D_2O$}, 2.0 – 9.0 ppm) of **212** after precipitation from aqueous ethanol.



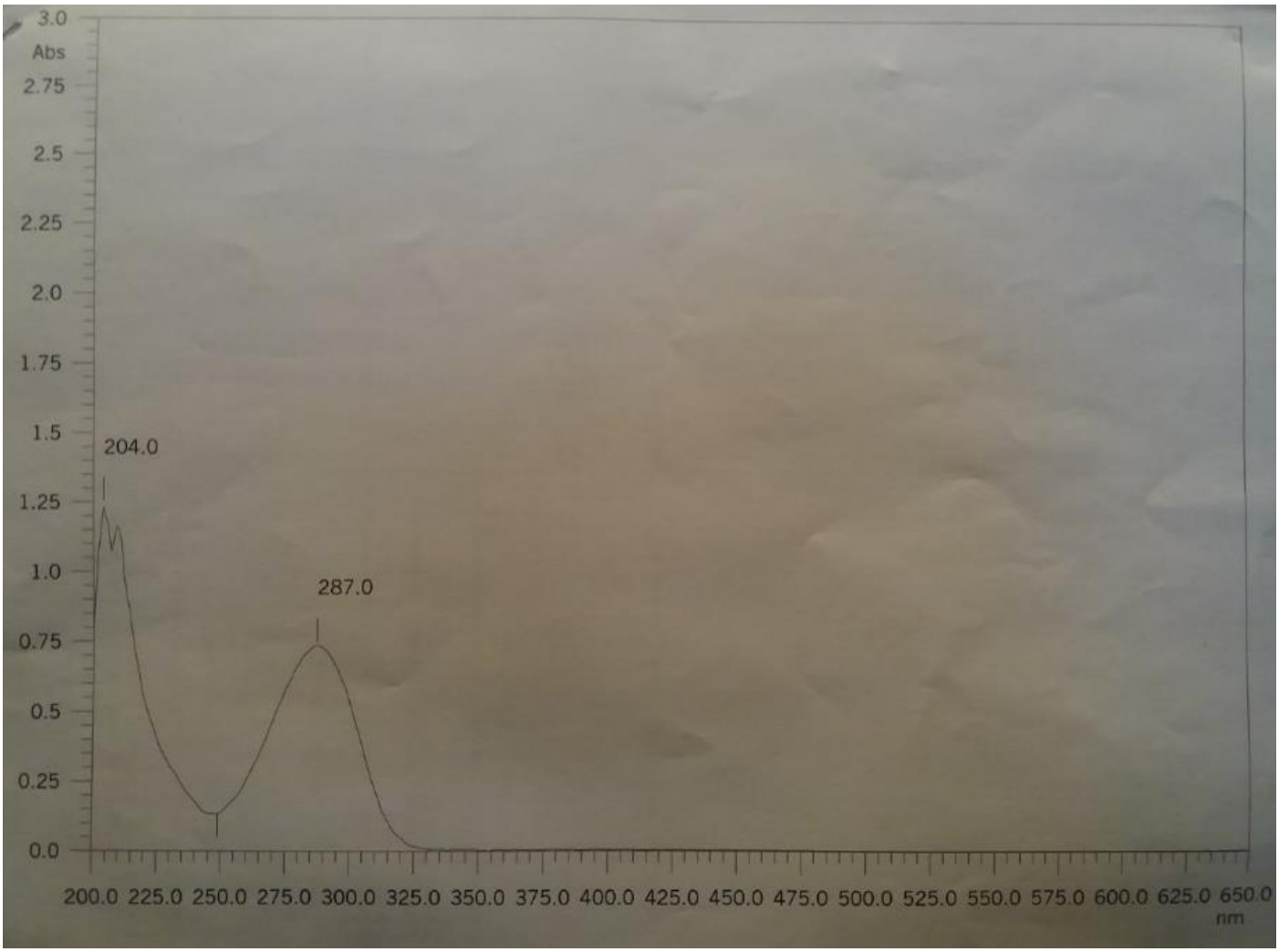
**A344**: EI+ mass spectrum of **212**.



**A345**: IR spectrum of **212**.

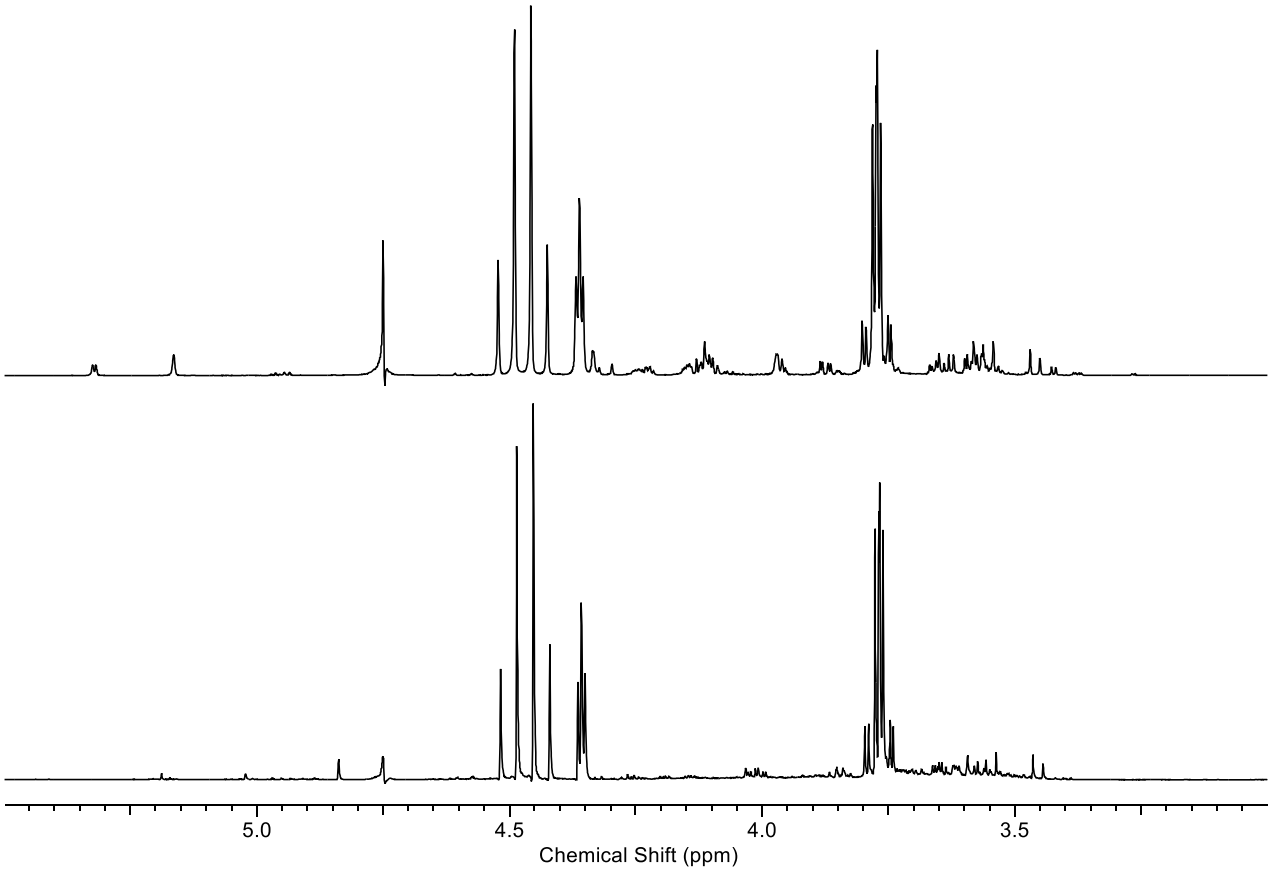


**A346**: UV-vis spectrum of **212**.



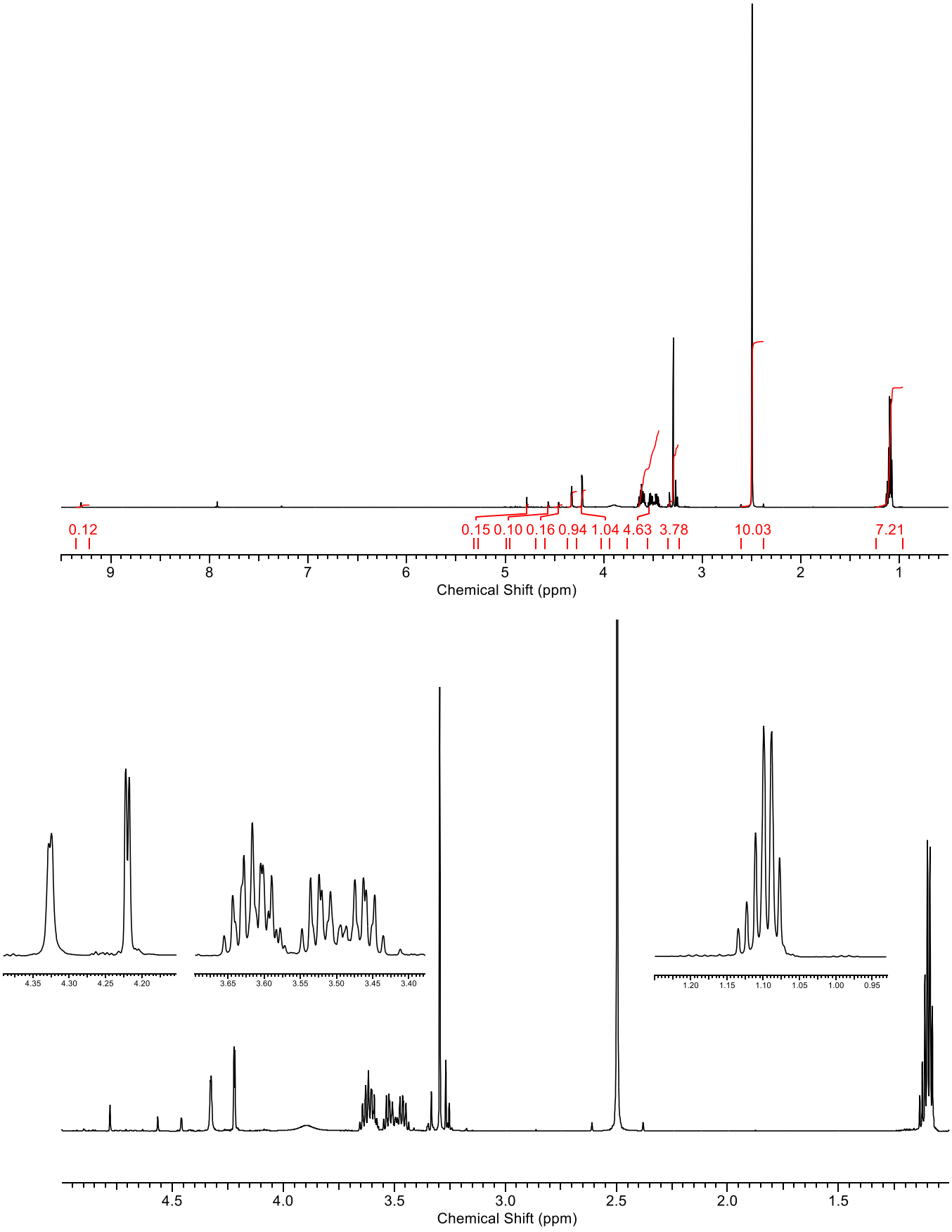
Erythrulose (**215**)

**A347**: $^{1}H$ NMR spectra (600 MHz, {$H_2O/D_2O$, 9:1}, 3.0 – 5.5 ppm) of **Top**, **215** isolated from the reaction of glycolaldehyde (500mM) in pH 7 phosphate buffer (1M) at 60 °C for 16 h and **Bottom**, commercially obtained **215**.

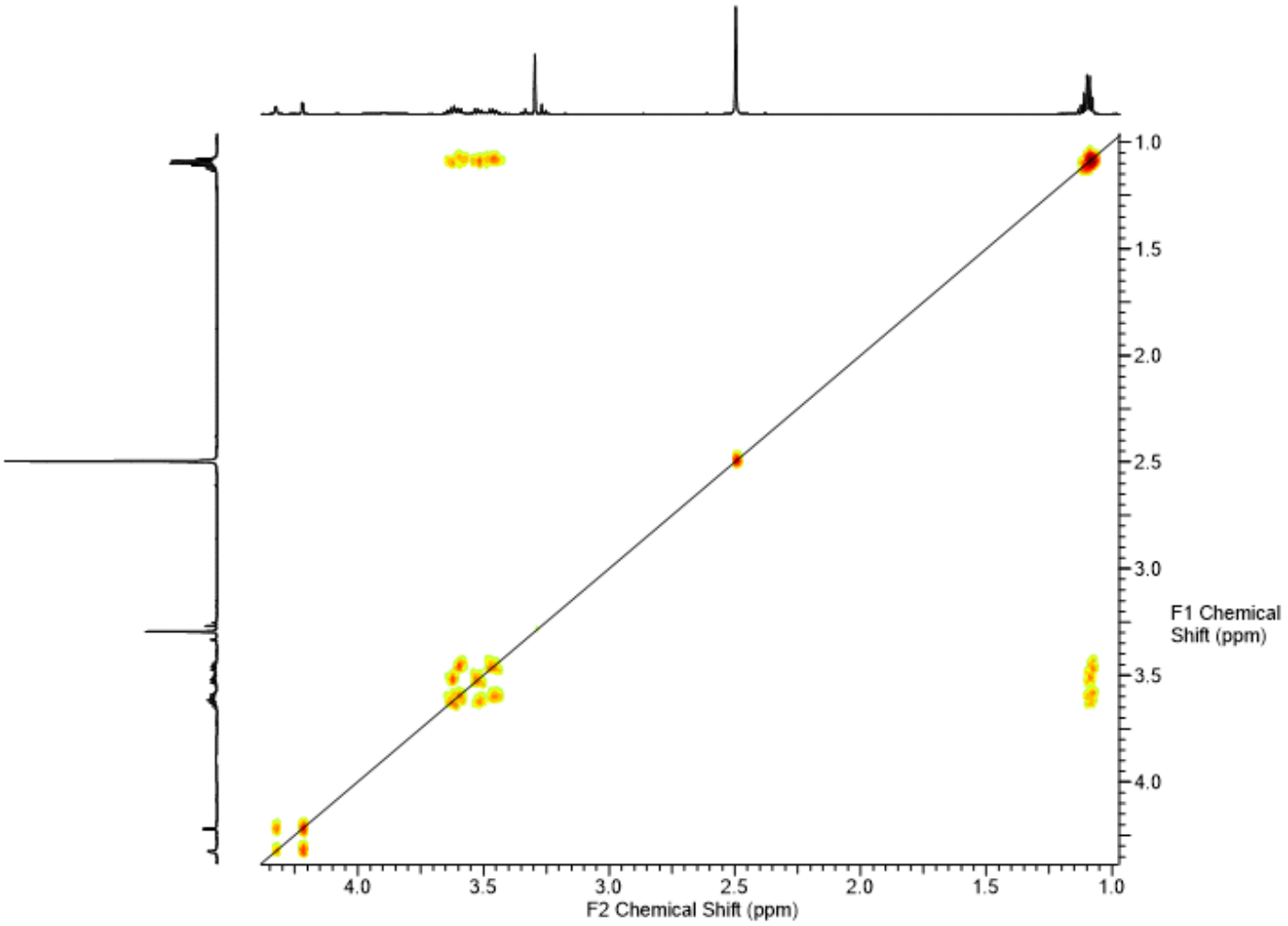


2,2-diethoxyacetaldehyde (**217**)

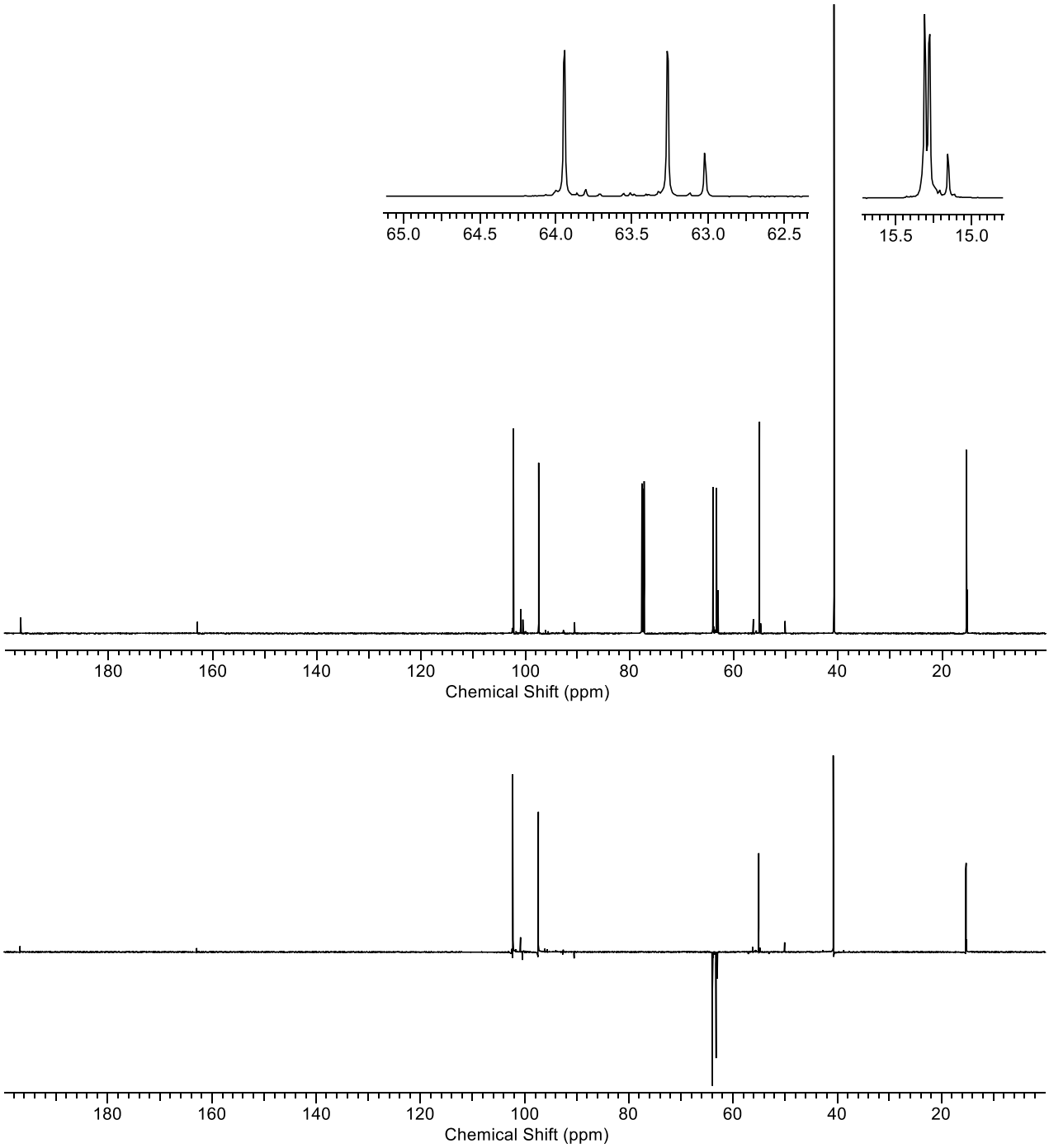
**A348**: $^{1}$H NMR spectrum (600 MHz, {$CDCl_3$}, 0.5 – 9.5 ppm) of **217** (isolated as its 1-methyl-hemi-acetal in DMSO) with expansions below.



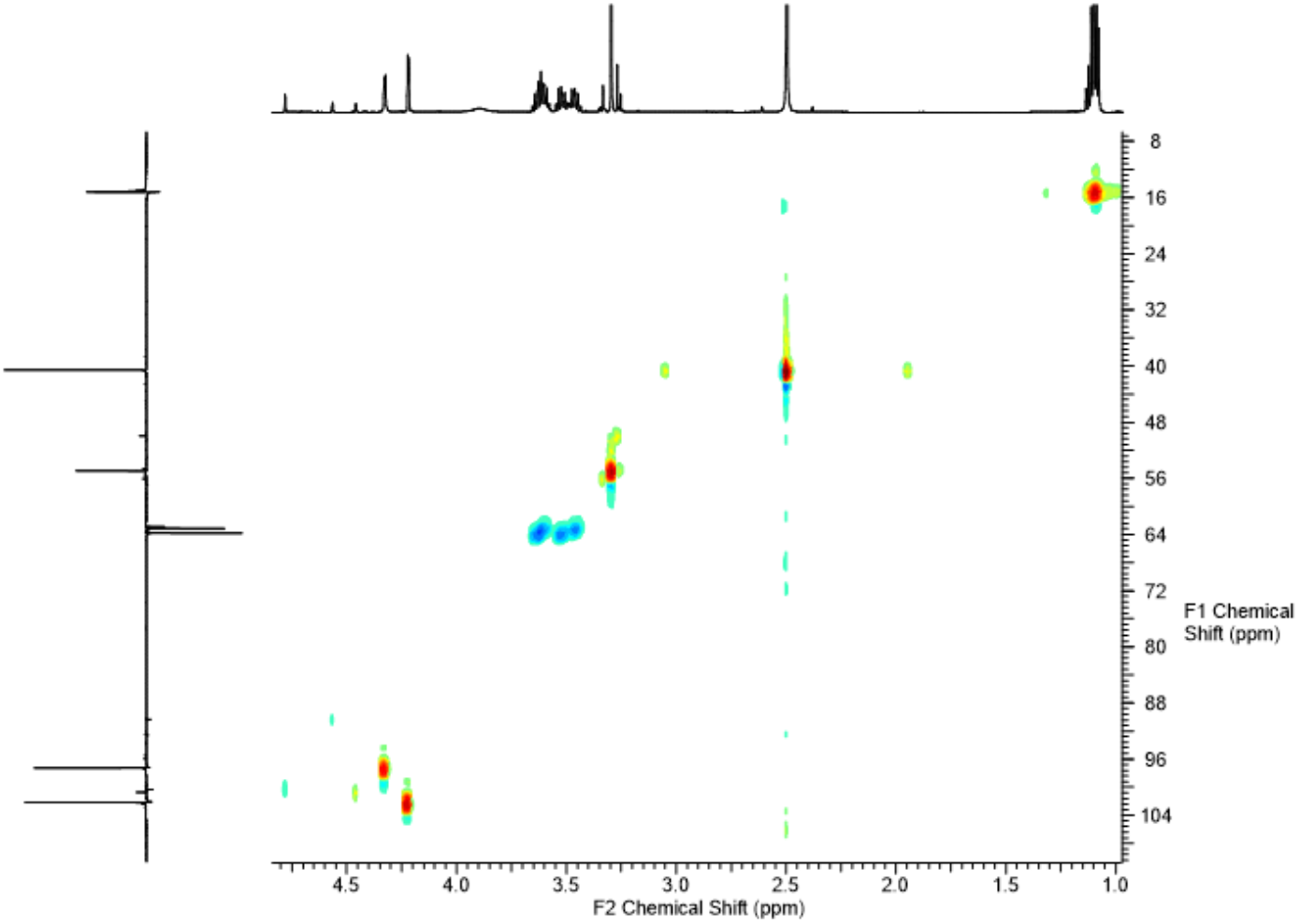
**A349**: $^{1}H$-$^{1}H$ COSY NMR spectrum (600 MHz, {$CDCl_3$}) of **217** (isolated as its 1-methyl-hemi-acetal in DMSO).



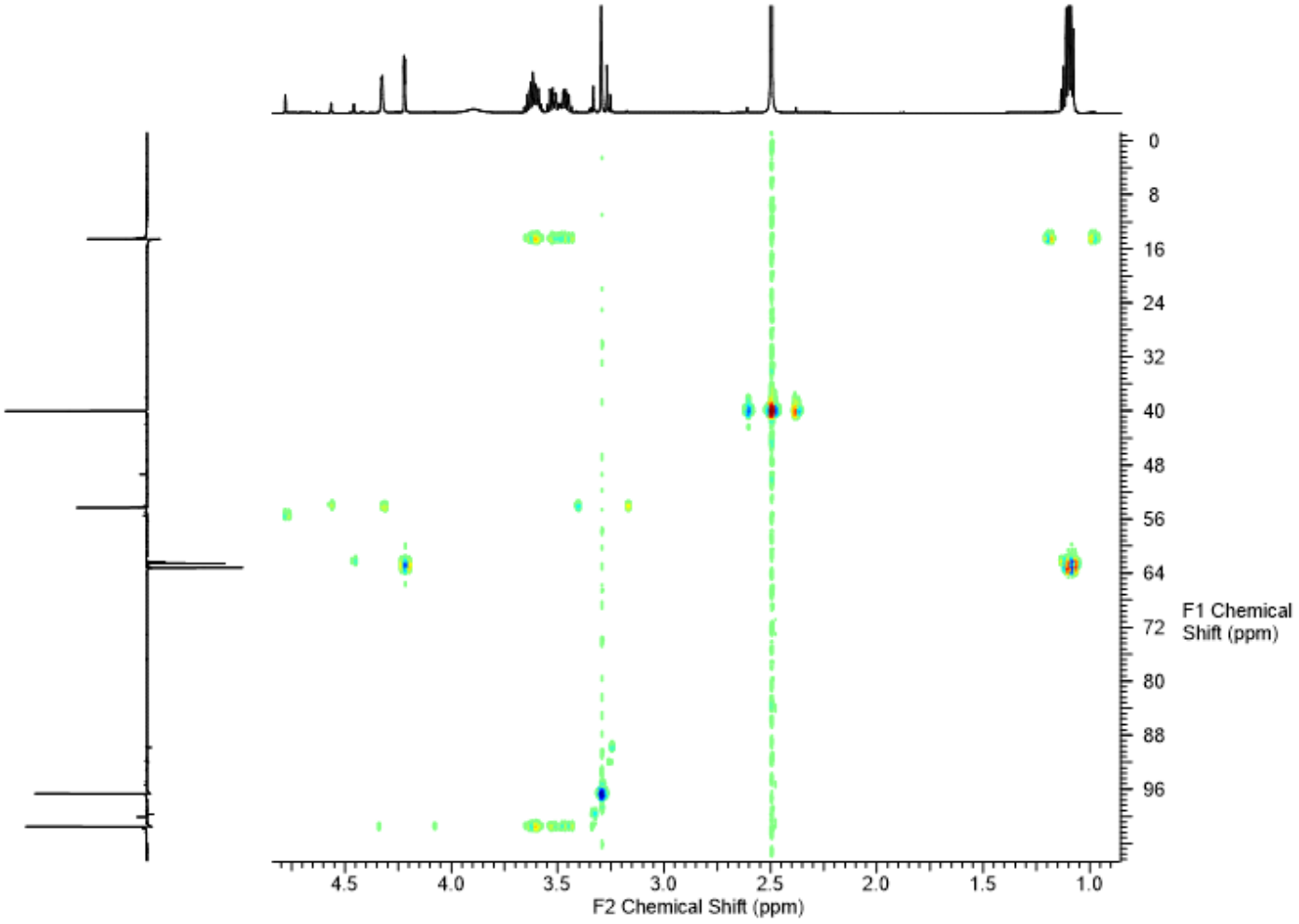
**A350**: $^{13}C$ NMR spectrum (151 MHz, {$CDCl_3$}, 0 – 200 ppm) of **217** (isolated as its 1-methyl-hemi-acetal in DMSO) with expansions overlaid and DEPT135 below.



**A351**: $^{1}H$-$^{13}C$ HSQC NMR spectrum (600 MHz, {$CDCl_3$}) of **217** (isolated as its 1-methyl-hemi-acetal in DMSO).

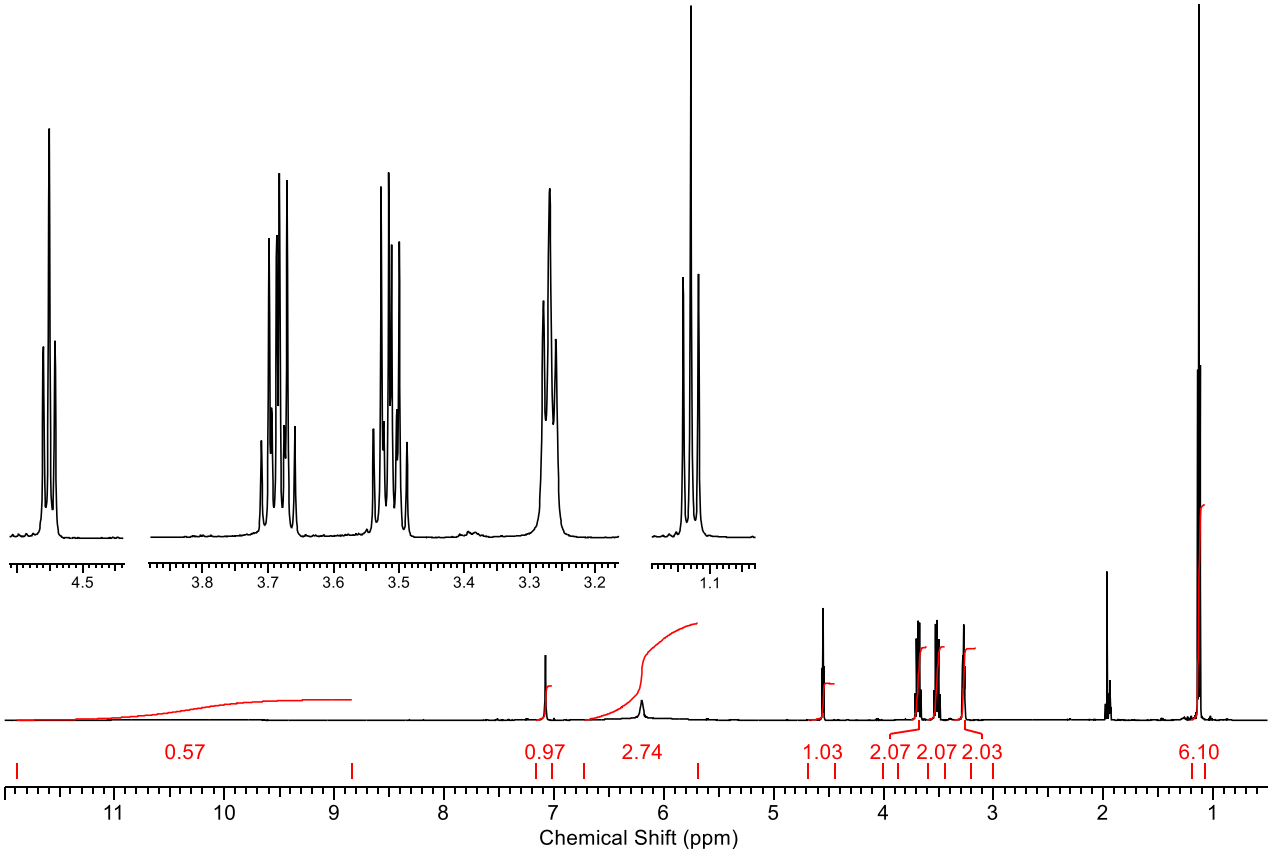


**A352**: $^{1}H$-$^{13}C$ HMBC NMR spectrum (600 MHz, {$CDCl_3$}) of **217** (isolated as its 1-methyl-hemi-acetal in DMSO).

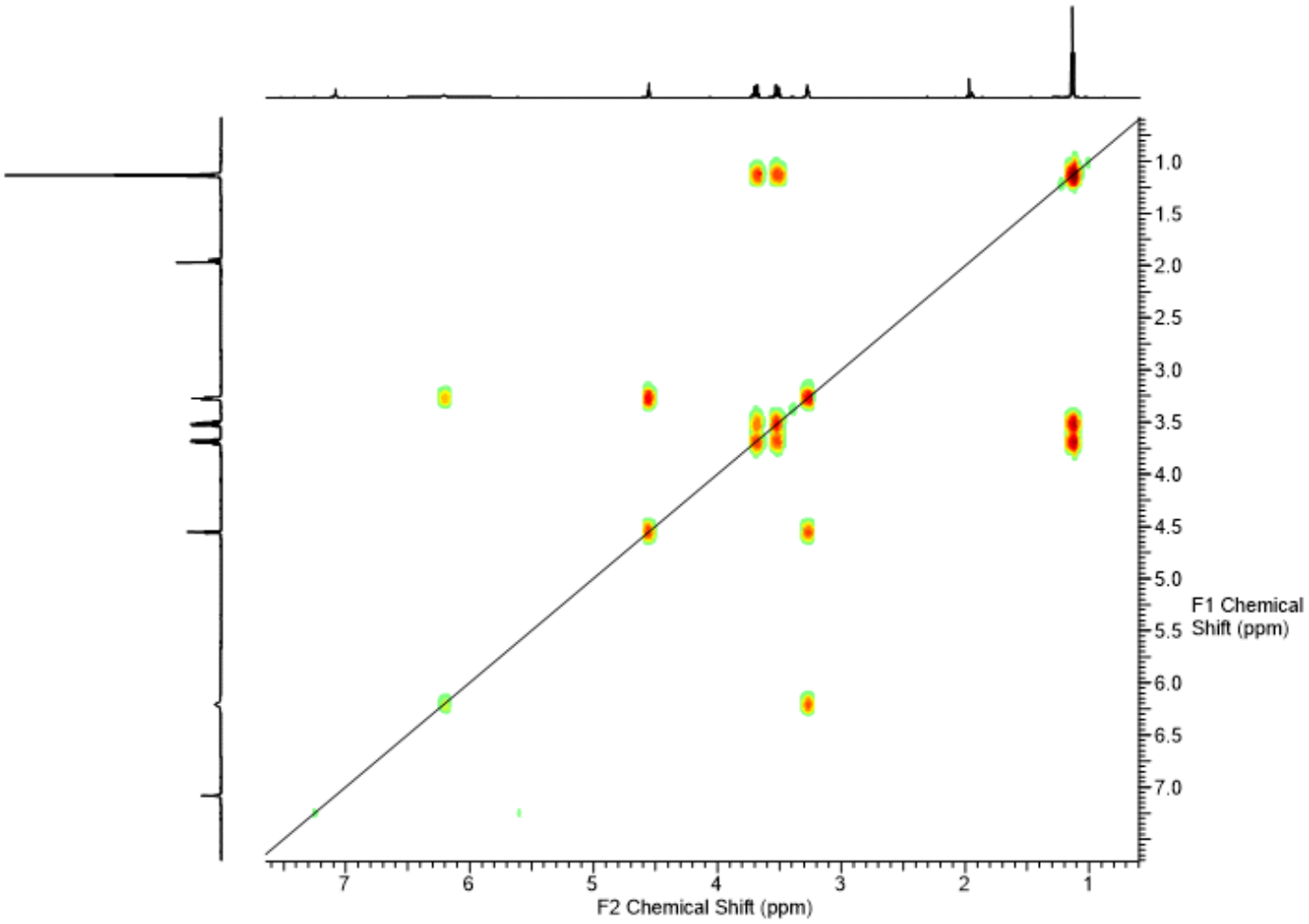


4-((2,2-diethoxyethyl)amino)-1H-imidazole-5-carboxamide (**218**)

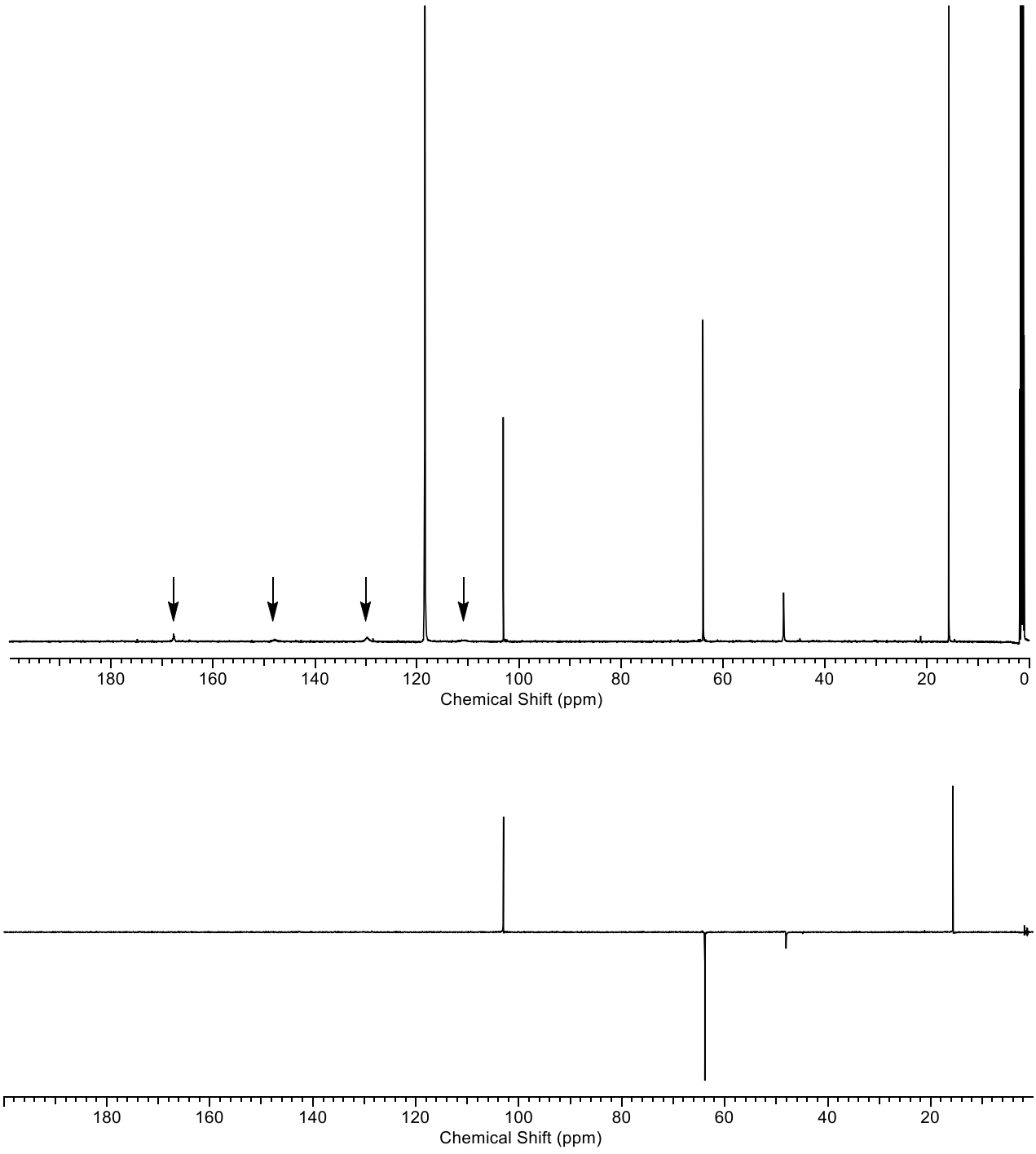
**A353**: $^{1}$H NMR spectrum (600 MHz, {MeCN-d3}, 0.5 – 12.0 ppm) of **218** with expansions overlaid.



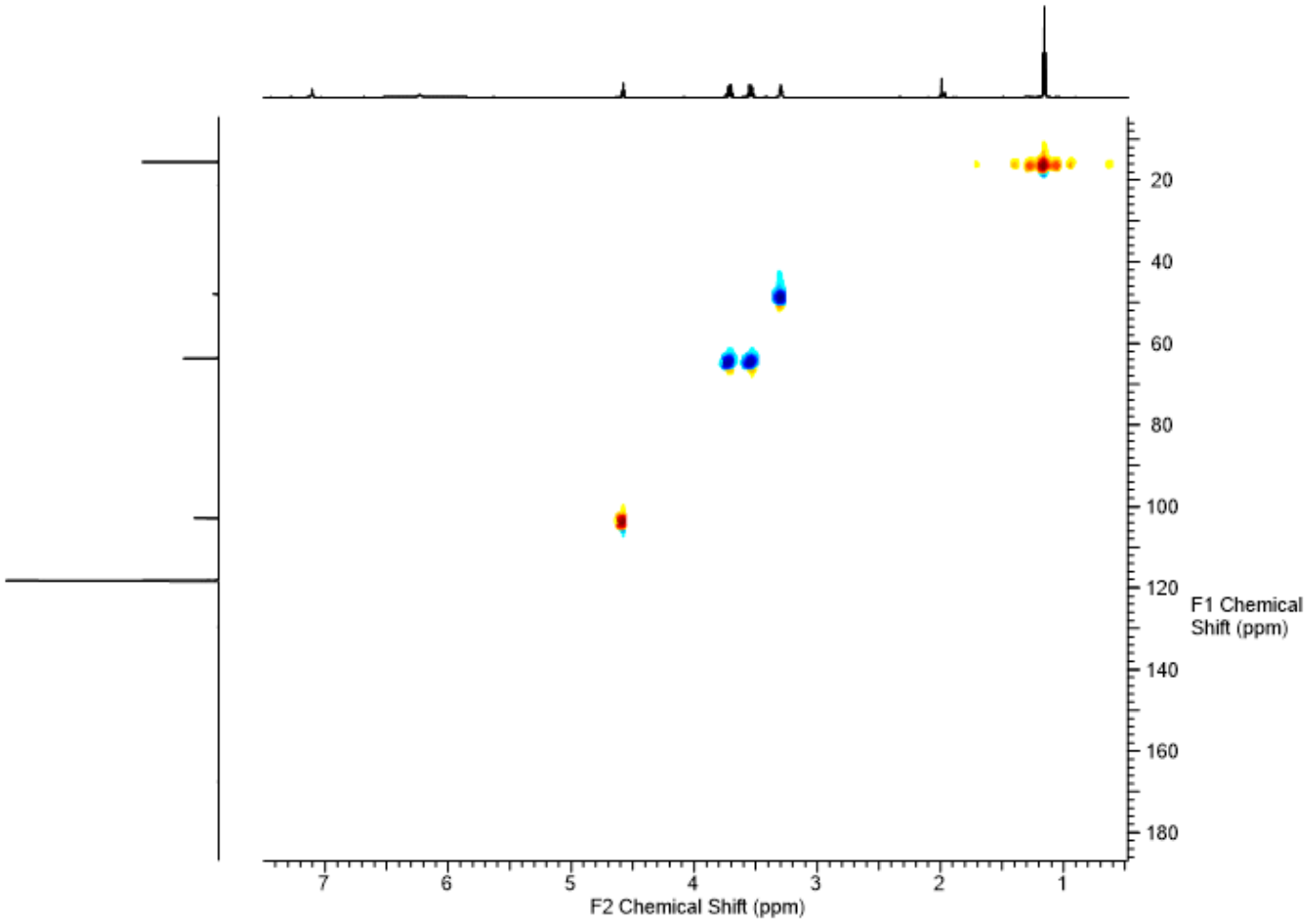
**A354**: $^{1}H$-$^{1}H$ COSY NMR spectrum (600 MHz, {MeCN-d3}) of **218**.



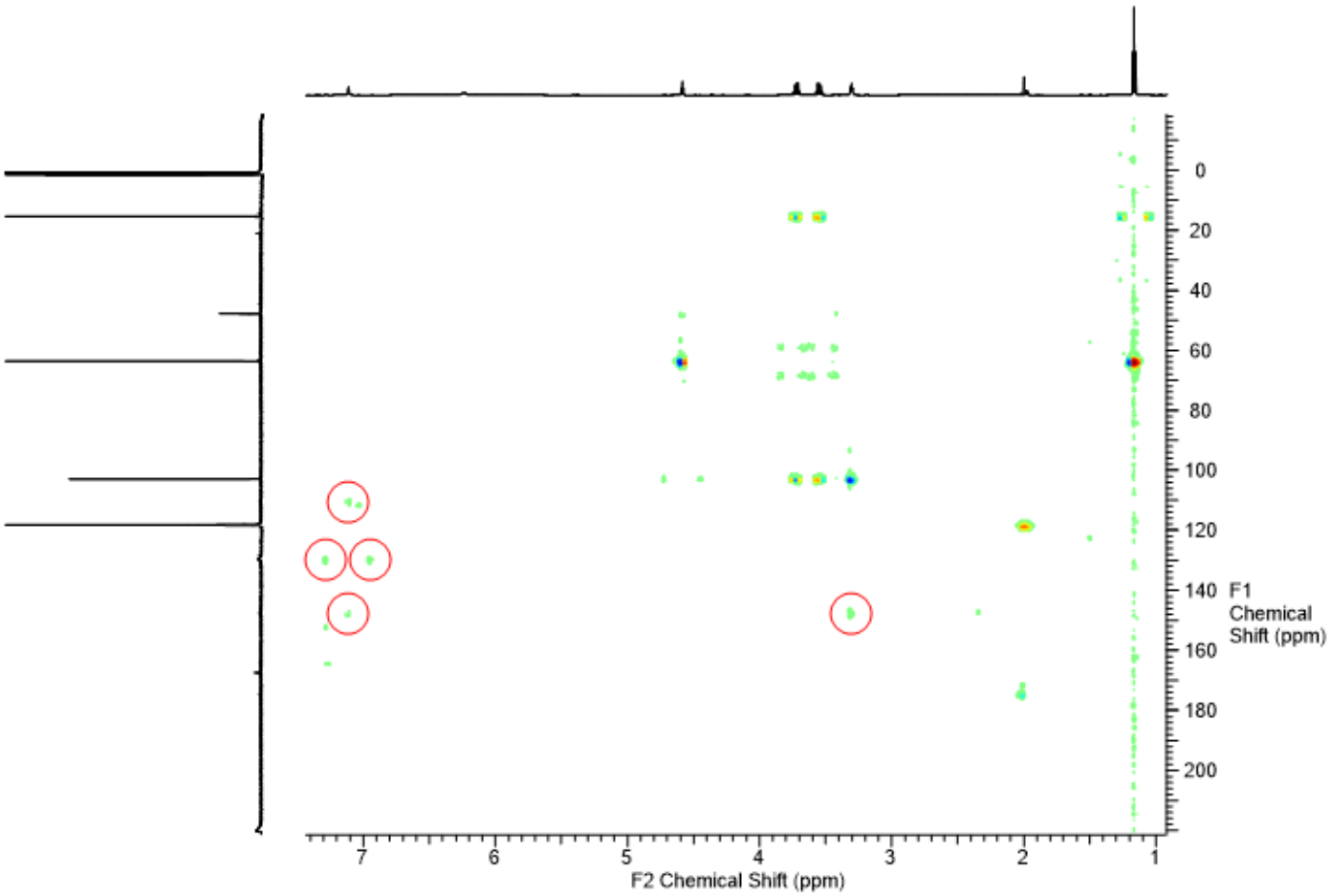
**A355**: $^{13}C$ NMR spectrum (151 MHz, {$CDCl_3$}, 0 – 200 ppm) of **218** with expansions overlaid and DEPT135 spectrum below. The highly broadened imidazolyl signals are indicated with arrows.



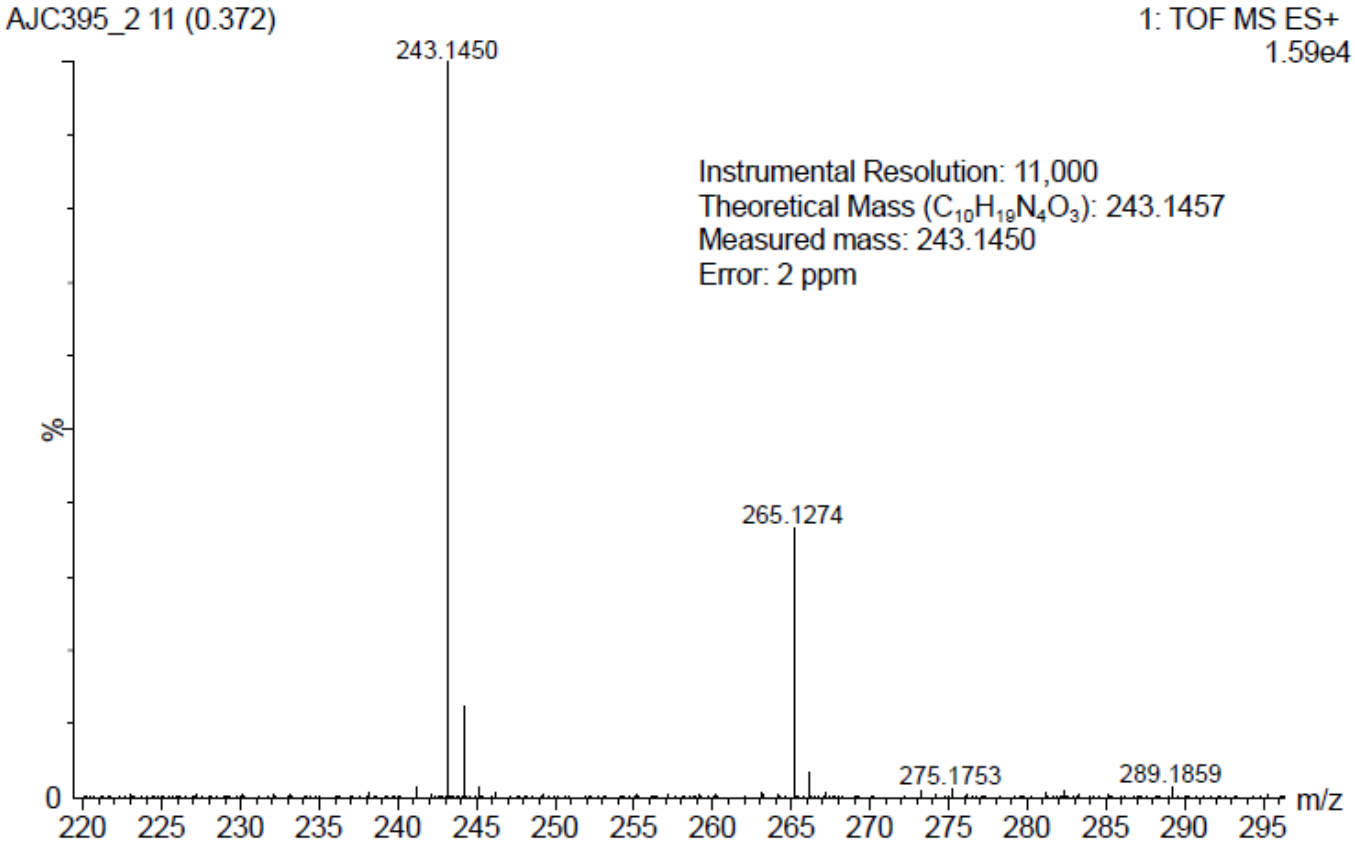
**A356**: $^{1}H$-$^{13}C$ HSQC NMR spectrum (600 MHz, {MeCN-d3}) of **218**.



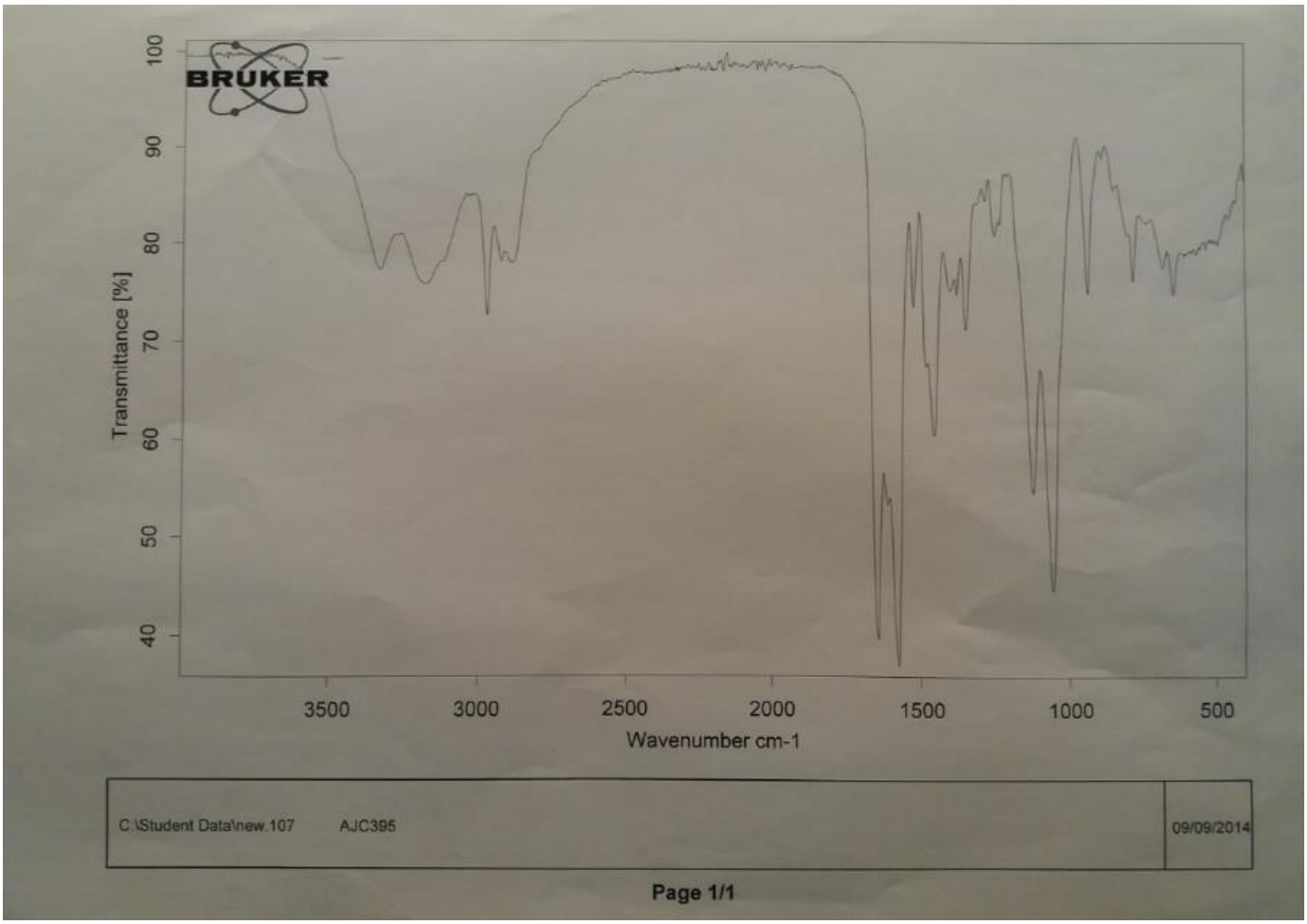
**A357**: $^{1}H$-$^{13}C$ HMBC NMR spectrum (600 MHz, {MeCN-d3}) of **218**. The correlations used to confirm that the broadened peaks in the $^{13}C$ spectrum were due to genuine signals have been circled.



**A358**: ESI+ mass spectrum of **218**.



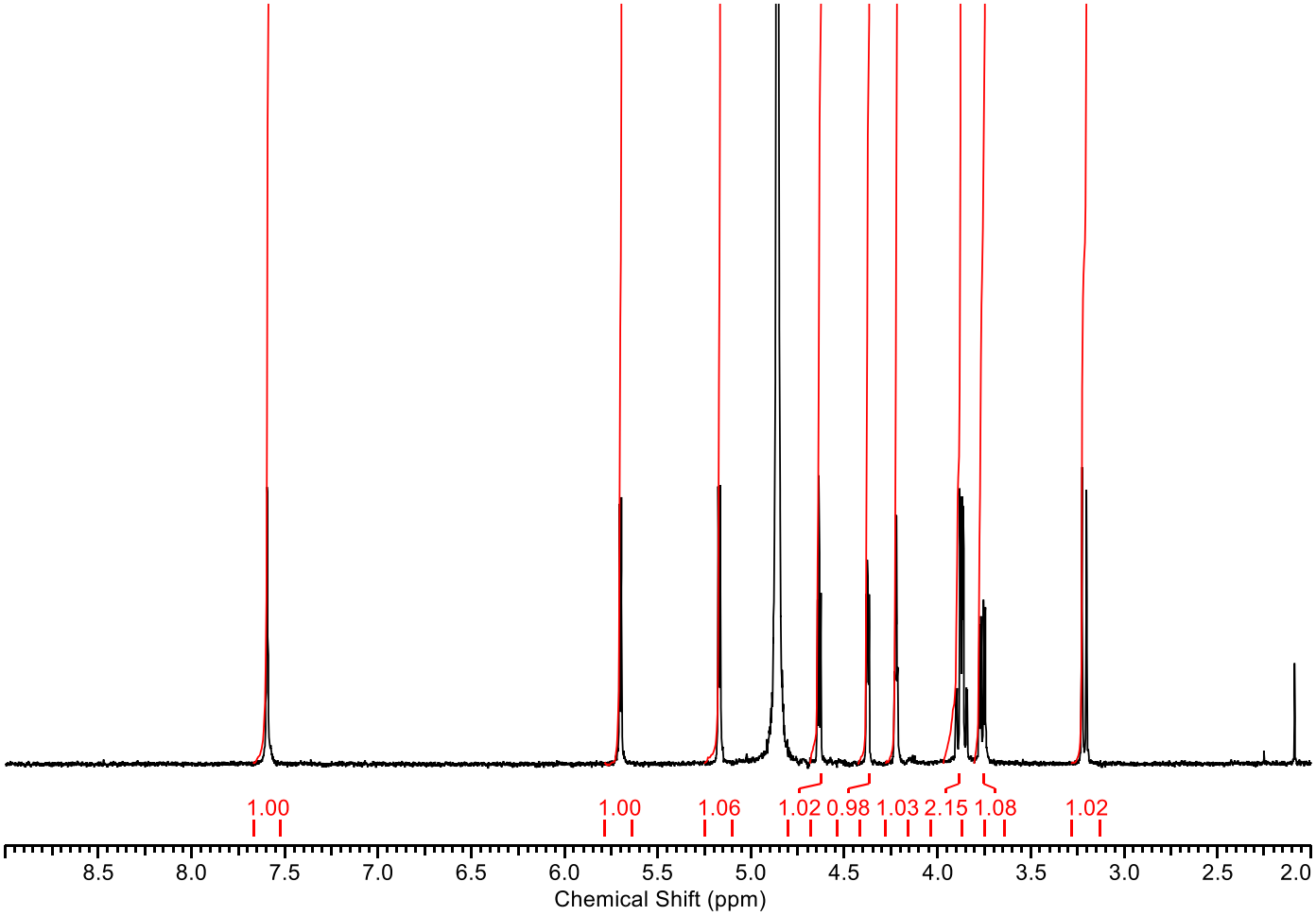
**A359**: IR spectrum of **218**.

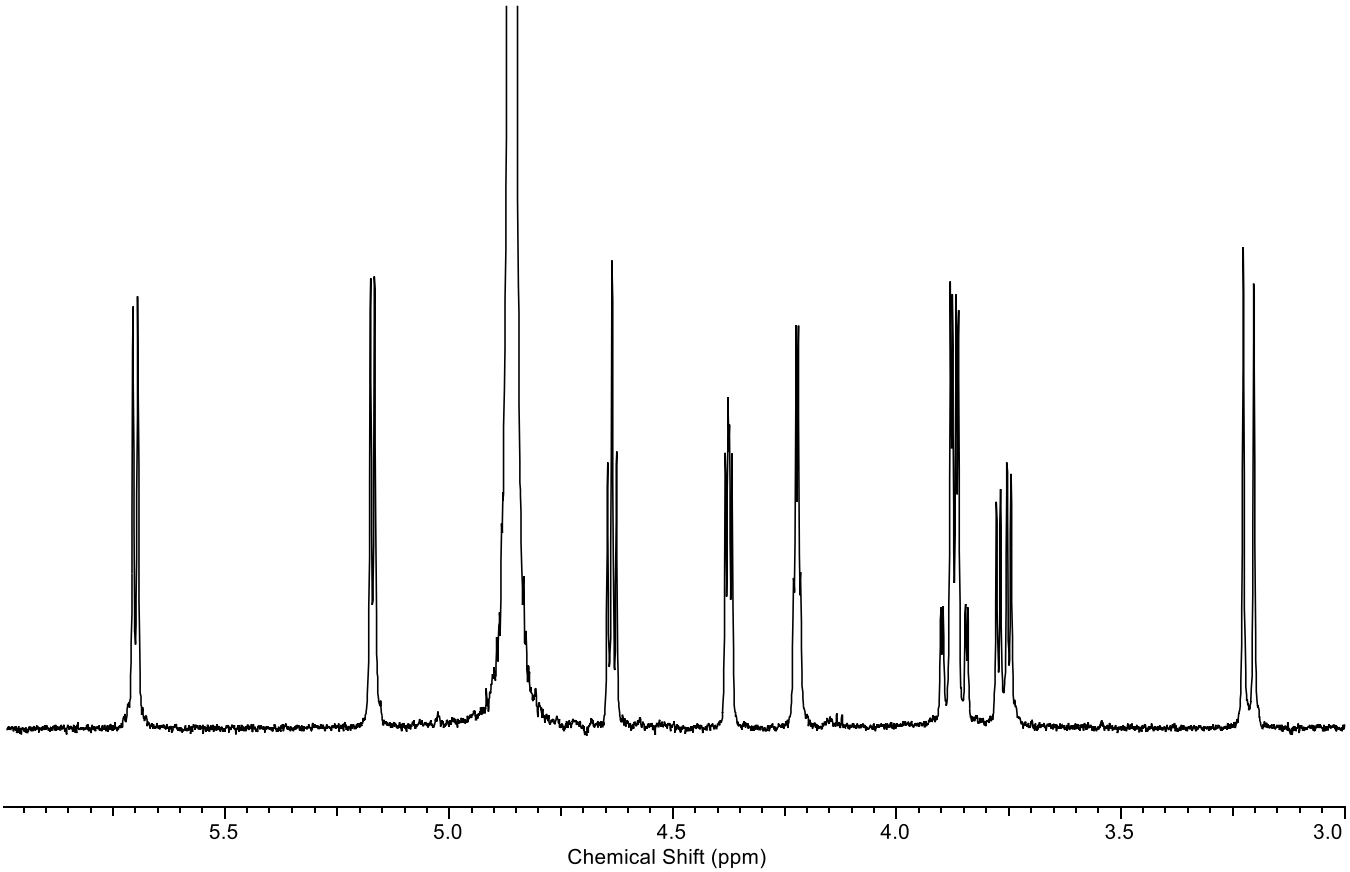


N1-(β-D-ribofuranosyl)-(8R)-azepinomycin (**214$_A$**)

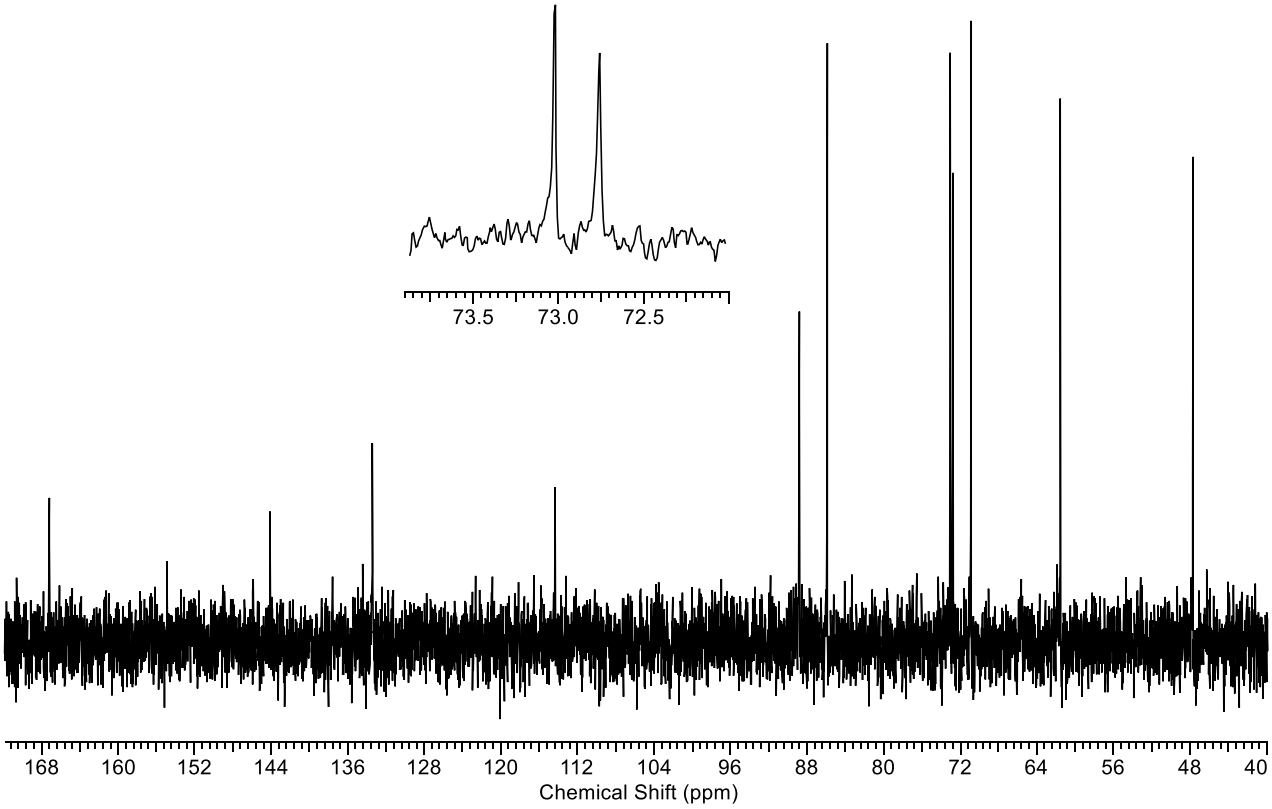
and N1-(β-D-ribofuranosyl)-(8S)-azepinomycin (**214$_B$**)

**A360**: $^1$H NMR spectrum (600 MHz, {$D_2O$}, 2.0 – 9.0 ppm) of **214$_A$** with expansion below.




5.5
5.0
4.5
4.0
3.5
3.0
Chemical Shift (ppm)

**A361**: $^{13}C$ NMR spectrum (151 MHz, {$D_2O$}, 40 – 170 ppm) of **214$_A$** with expansion overlaid.



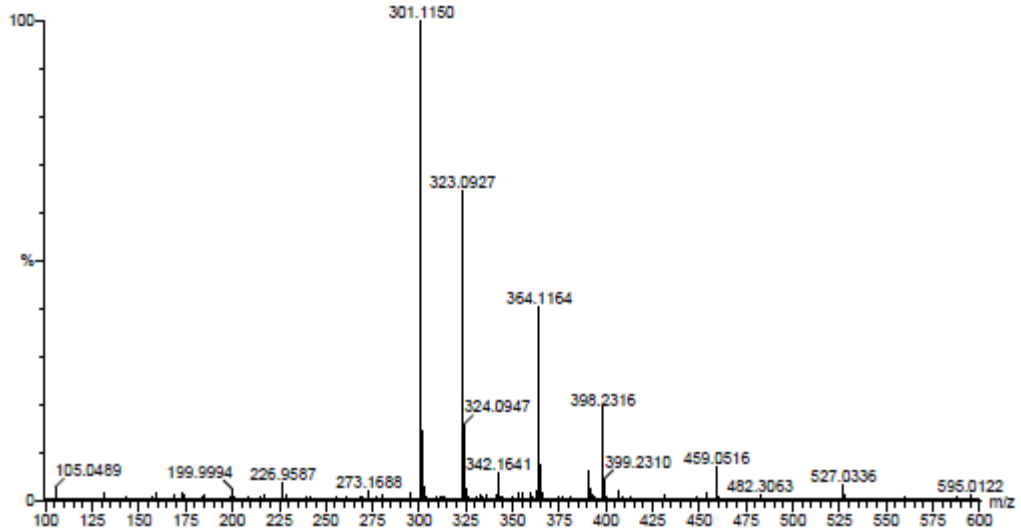
**A362**: ESI+ mass spectrum of **214A**.

**Elemental Composition Report** **Page 1**

**Single Mass Analysis**
Tolerance = 5.0 PPM / DBE: min = -1.5, max = 50.0
Element prediction: Off
Number of isotope peaks used for i-FIT = 3

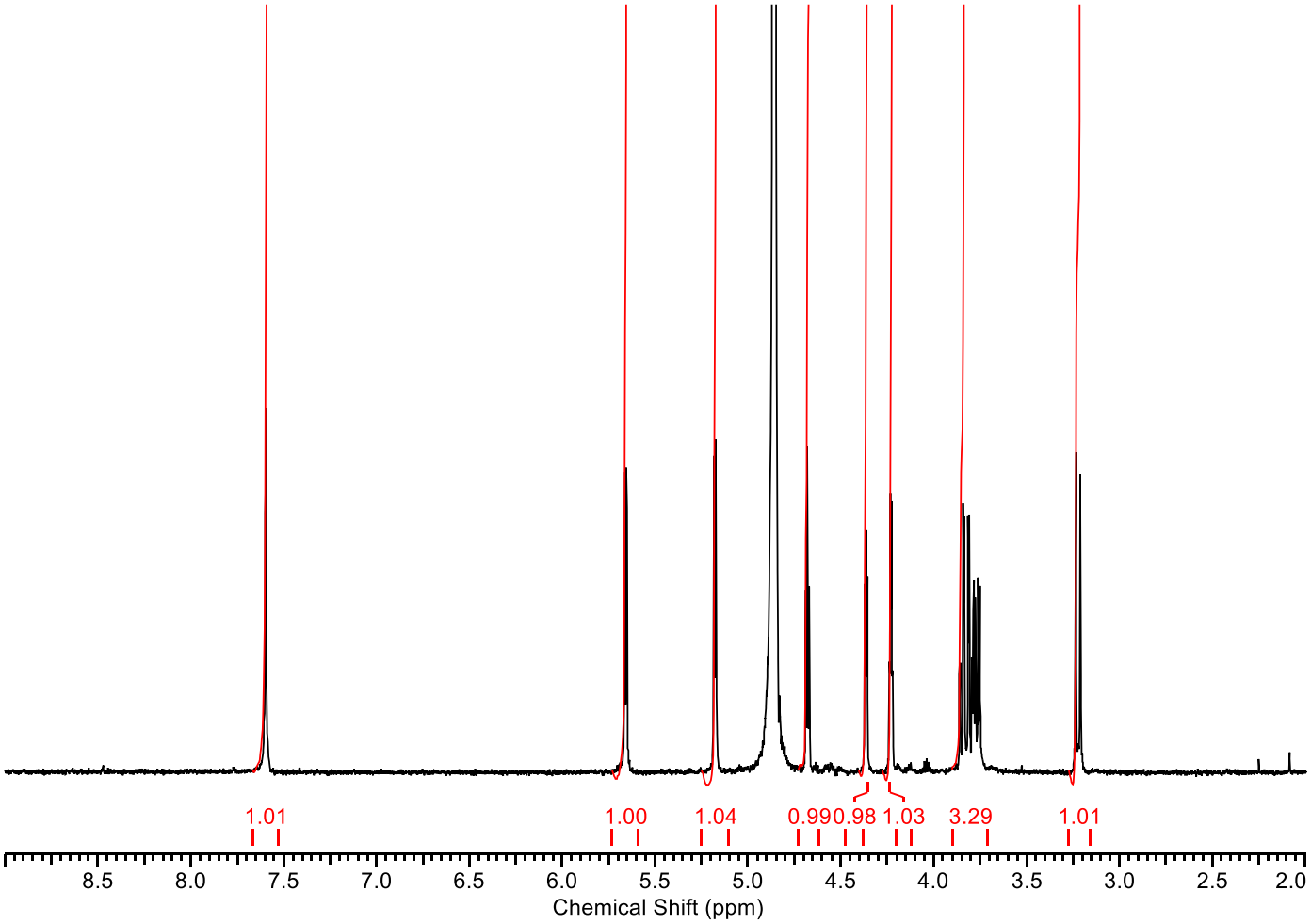
Monoisotopic Mass, Even Electron Ions
157 formula(e) evaluated with 1 results within limits (all results (up to 1000) for each mass)
Elements Used:
C: 0-20 H: 0-30 N: 0-5 O: 0-10
AJC437A
AJC437A_2 17 (0.598)

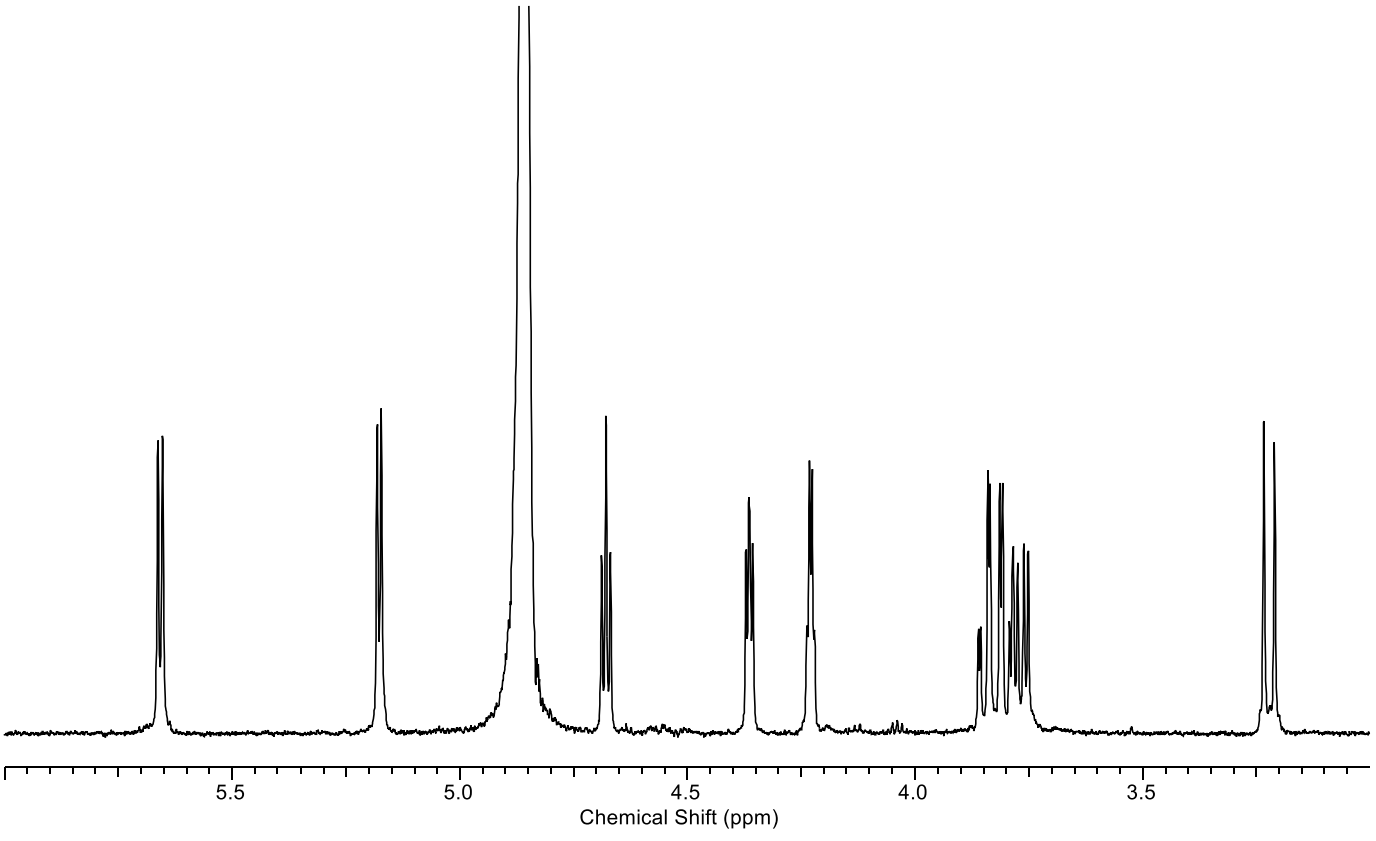
1: TOF MS ES+
4.30e+004



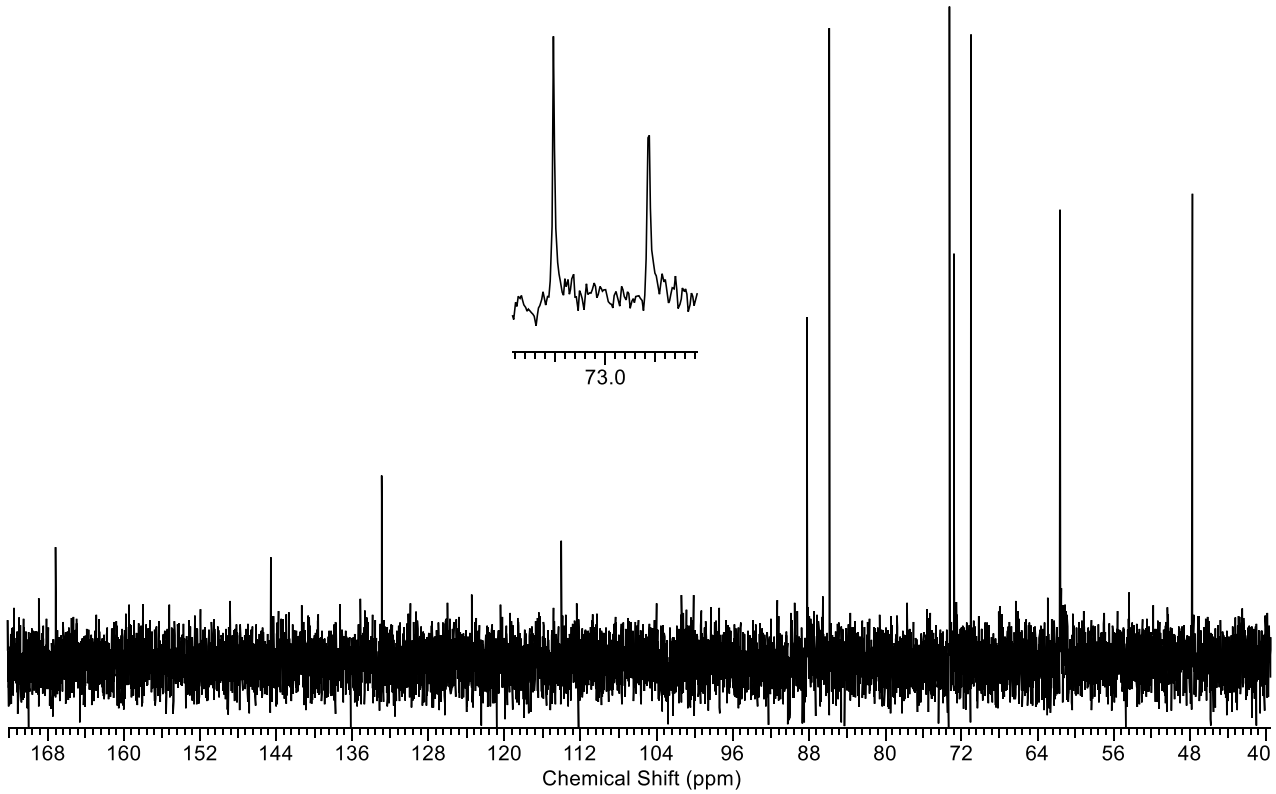

| | | | | | | |
|---|---|---|---|---|---|---|
| Minimum: | | | | -1.5 | | |
| Maximum: | | 5.0 | 5.0 | 50.0 | | |
| Mass | Calc. Mass | mDa | PPM | DBE | i-FIT | Formula |
| 301.1150 | 301.1148 | 0.2 | 0.7 | 5.5 | 28.2 | C11 H17 N4 O6 |

**A363**: $^{1}$H NMR spectrum (600 MHz, {$D_2O$}, 2.0 – 9.0 ppm) of **214$_B$** with expansion below.



**A364**: $^{13}C$ NMR spectrum (151 MHz, {$D_2O$}, 40 – 170 ppm) of **214B** with expansion overlaid.



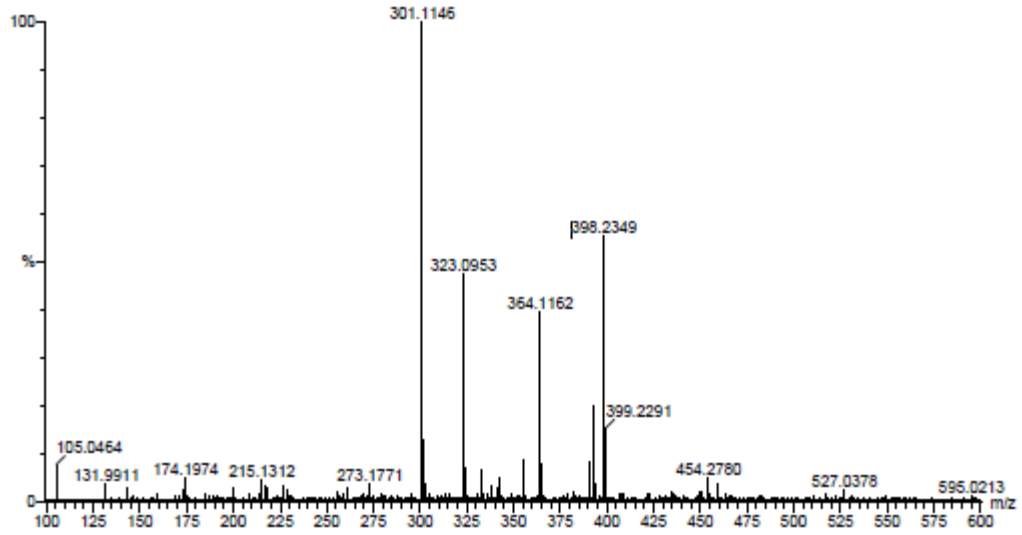
**A365**: ESI+ mass spectrum of **214B**.

**Elemental Composition Report** Page 1

**Single Mass Analysis**
Tolerance = 5.0 PPM / DBE: min = -1.5, max = 50.0
Element prediction: Off
Number of isotope peaks used for i-FIT = 3

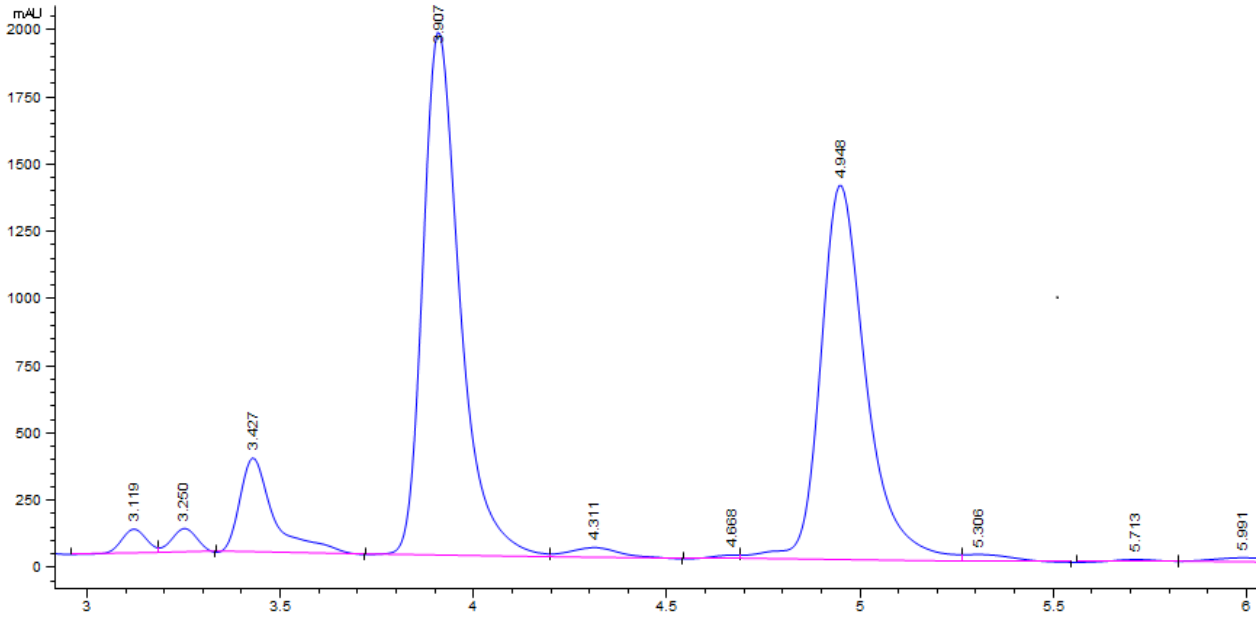
Monoisotopic Mass, Even Electron Ions
157 formula(e) evaluated with 2 results within limits (all results (up to 1000) for each mass)
Elements Used:
C: 0-20 H: 0-30 N: 0-5 O: 0-10
AJC437B
AJC437B_2 16 (0.541)

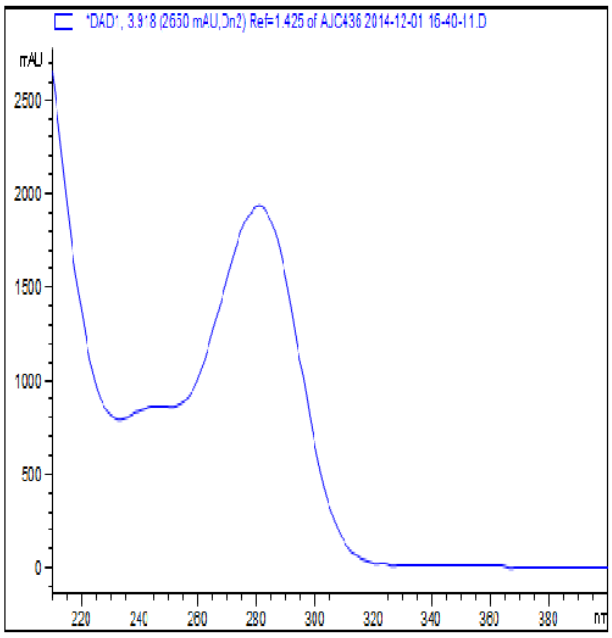
1: TOF MS ES+
1.44e+004

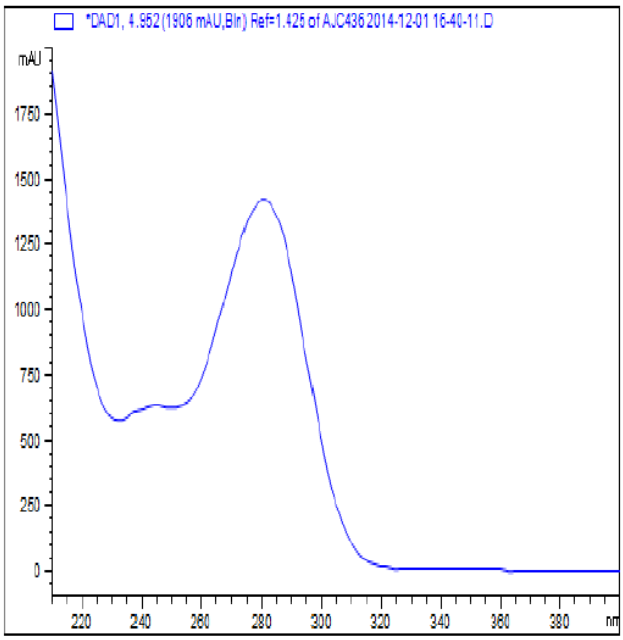



| | | | | | | |
|---|---|---|---|---|---|---|
| Minimum: | | 5.0 | 5.0 | -1.5 | | |
| Maximum: | | | | 50.0 | | |
| Mass | Calc. Mass | mDa | PPM | DBE | i-FIT | Formula |
| 301.1146 | 301.1148 | -0.2 | -0.7 | 5.5 | 51.5 | C11 H17 N4 O6 |
| | 301.1135 | 1.1 | 3.7 | 0.5 | 18.9 | C10 H21 O10 |

**A366**: HPLC trace of crude **214** (Diastereomer **A** = 3.91 min; Diastereomer **B** = 4.95 min) with corresponding UV traces of the major peaks below.



***Single crystal X-ray diffraction studies***

General information

Single cyrstal X-ray diffraction data were collected using an *Agilent SuperNova (Dual Source)* single crystal X-ray diffractometer equipped with an Atlas CCD Detector. All data sets collected at 150 K using Cu$K_\alpha$ radiation ($\lambda$ = 1.54184 Å). The data were acquired and processed using the *CrysAlisPro* program and the datasets were corrected for Lorentz and polarization effects. Structure solution and refinement was accomplished using SHELXS-97 and SHELXL-97, respectively. The crystal structures were solved using direct methods. All non-hydrogen atoms were refined anisotropically. Hydrogen atoms affiliated with oxygen and nitrogen atoms were refined isotropically in positions identified in the difference Fourier map or in geometrically constrained positions, while hydrogen atoms associated with carbon atoms were refined isotropically in geometrically constrained positions. Crystallographic and refinement parameters for all crystal structures are given in the table below.

**A367**: Crystallographic and refinement parameters for **108**, **184**, **185**, **198** and **212.**

| compound | 212 | 185 | 184 | 108 | 198 |
|---|---|---|---|---|---|
| chemical formula | $C_6H_8N_4O_2$ | $(C_{11}H_{10}N_6O_3)_2(H_2O)$ | $C_{11}H_{12}N_6O_4$ | $C_{12}H_{38}Ca_3N_{12}O_{29}P_6$ | $(C_{13}H_{15}N_6O_5)(H_2O)_{0.83}$ |
| $M_r$/gmol$^{-1}$ | 168.16 | 566.52 | 292.27 | 1120.60 | 349.27 |
| crystal system | monoclinic | monoclinic | monoclinic | monoclinic | triclinic |
| space group | $P2_1/c$ | $P2_1/n$ | $P2_1/c$ | $P2_1/c$ | *P-1* |
| *a*/Å | 6.68948(7) | 18.7664(4) | 7.6237(2) | 10.59420(10) | 11.3819(2) |
| *b*/Å | 14.04389(15) | 5.30070(10) | 7.24320(10) | 22.20680(10) | 14.1561(3) |
| *c*/Å | 7.22458(9) | 23.3702(4) | 21.8767(4) | 18.34810(10) | 16.2079(3) |
| *α*/° | 90 | 90 | 90 | 90 | 67.126(2) |
| *β*/° | 93.8750(11) | 93.478(2) | 1199.08(4) | 102.1080(10) | 76.398(2) |
| *γ*/° | 90 | 90 | 90 | 90 | 74.371(2) |
| *V*/Å$^3$ | 677.171(13) | 2320.47(8) | 1199.08(4) | 4220.61(5) | 2291.97(9) |
| *Z* | 4 | 4 | 4 | 4 | 6 |
| $\rho_{calc}$ / mg mm$^3$ | 1.649 | 1.627 | 1.619 | 1.727 | 1.518 |
| *F*(000) | 352.0 | 1176 | 608 | 2304 | 1094 |
| $\mu$(Cu$K_\alpha$)/mm$^{-1}$ | 1.090 | 1.154 | 1.084 | 6.556 | 1.044 |
| *T*/K | 150.0(1) | 150.0(1) | 150.0(1) | 150.0(1) | 150.0(1) |
| crystal size/mm | 0.28 × 0.26 × 0.15 | 0.33 × 0.27 × 0.13 | 0.14 × 0.12 × 0.09 | 0.22 × 0.18 × 0.13 | 0.32 × 0.22 × 0.11 |
| index range | -8 ≤ h ≤ 8<br>-17 ≤ k ≤ 17<br>-8 ≤ l ≤ 8 | -22 ≤ h ≤ 21<br>-6 ≤ k ≤ 4<br>-27 ≤ l ≤ 27 | -9 ≤ h ≤ 7<br>-7 ≤ k ≤ 8<br>-18 ≤ l ≤ 26 | -13 ≤ h ≤ 13<br>-27 ≤ k ≤ 27<br>-22 ≤ l ≤ 22 | -13 ≤ h ≤ 13<br>-17 ≤ k ≤ 17<br>-20 ≤ l ≤ 20 |
| collected reflections | 9206 | 8245 | 4244 | 72333 | 47318 |
| unique reflections | 1328 | 4113 | 2123 | 8486 | 9144 |
| $R_{int}$ | 0.0198 | 0.0251 | 0.0135 | 0.0325 | 0.0328 |
| reflections with $I > 2\sigma(I)$ | 1294 | 3582 | 1969 | 8083 | 8409 |
| no. parameters | 110 | 405 | 207 | 564 | 750 |
| $R(F)$, $F > 2\sigma$(F) | 0.0383 | 0.0357 | 0.0317 | 0.0471 | 0.0371 |
| $wR(F^2)$, $F > 2\sigma$(F) | 0.1004 | 0.0420 | 0.0339 | 0.0487 | 0.0399 |
| $R(F)$, all data | 0.0389 | 0.0934 | 0.0826 | 0.1270 | 0.0988 |
| $wR(F^2)$, all data | 0.1007 | 0.1002 | 0.0846 | 0.1287 | 0.1020 |
| $\Delta_r$ (min., max.) eÅ$^{-3}$ | 0.69, -0.66 | 0.23, -0.21 | 0.22, -0.26 | 2.20, -1.34 | 0.59, -0.24 |