Supporting Information for

# Expanding the NMR Toolkit for Biological Solids: Oxygen-17 Enriched Fmoc-Amino Acids

Brittney A. Klein[1], Dylan G. Tkachuk[1], Victor V. Terskikh[2] and Vladimir K. Michaelis[1,*]

1.Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada

2. Metrology, National Research Council Canada, Ottawa, Ontario K1A 0R6, Canada

* Corresponding Author: vladimir.michaelis@ualberta.ca

**Table S1.** Average deviations of the computed vs. experimental $^{17}O$ NMR parameters plotted in Figures 5 & S4 for the C**O** and C**O**H sites in the Fmoc-amino acids.

| $^{17}O$ Site | $\delta_{iso}$ (ppm) | Cq (MHz) | $\eta$ | $q_{zz}$ (a.u.) | $q_{yy}$ (a.u.) | $q_{xx}$ (a.u.) | $\Omega^1$ (ppm) | $\kappa$ | $\delta_{11}$ (ppm) | $\delta_{22}$ (ppm) | $\delta_{33}$ (ppm) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| C**O** | 47 | 0.4 | 0.11 | 0.07 | 0.10 | 0.07 | 77 | 0.32 | 64 | 97 | 20 |
| C**O**H | 18 | 1.2 | 0.33 | 0.20 | 0.35 | 0.16 | 55 | 0.56 | 28 | 63 | 35 |

1 The deviations were calculated using the re-evaluated experimental span values for the CO sites in Fmoc-L-valine and -leucine, which in turn were used to extract the CSA tensor components. All correlations were calculated using the equation: $[(\Sigma(\text{calc.} - \text{expt.})^2)/(n)]^{1/2}$ .

**Table S2**. Selection of basis sets for computing NMR parameters using density functional theory on crystal structure of FMOC-L-tyrosine, monohydrate.

| Basis set | $^{17}O$ site | $\sigma_{iso}$ (ppm) | $\sigma_{33}$ (ppm) | $\sigma_{22}$ (ppm) | $\sigma_{11}$ (ppm) | $C_Q$ (MHz) | $\eta$ |
|---|---|---|---|---|---|---|---|
| DZP | CO | -41.571 | 281.028 | -131.463 | -274.278 | 7.88961 | 0.12270 |
| | COH | 155.200 | 256.063 | 177.473 | 32.065 | -7.94763 | 0.37483 |
| ZORA/DZP | CO | -39.327 | 281.867 | -128.663 | -271.184 | 7.89293 | 0.11949 |
| | COH | 156.896 | 243.562 | 179.435 | 33.372 | -7.54985 | 0.37714 |
| TZ2P | CO | -88.756 | 269.898 | -188.846 | -347.315 | 8.85450 | 0.13684 |
| | COH | 132.856 | 243.562 | 164.261 | -9.255 | -9.07486 | 0.36192 |
| ZORA/TZ2P | CO | -86.815 | 270.438 | -186.307 | -344.576 | 8.85790 | 0.13550 |
| | COH | 134.330 | 245.223 | 166.099 | -8.331 | -9.07075 | 0.35039 |
| QZ4P | CO | -96.012 | 269.361 | -197.114 | -360.283 | 8.91513 | 0.13757 |
| | COH | 125.903 | 237.160 | 155.206 | -14.656 | -9.24279 | 0.34899 |
| ZORA/QZ4P | CO | -94.454 | 269.623 | -194.977 | -358.008 | 8.92528 | 0.13409 |
| | COH | 127.208 | 238.691 | 156.840 | -13.907 | -9.25906 | 0.36450 |

**Table S3**. Comparison of the computed $^{17}O$ NMR parameters[1] for Fmoc-L-tyrosine and -isoleucine obtained using their reported crystal structures with those obtained using different geometry-optimized structures.

| Sample | Structure Coordinates | $^{17}O$ site | $\sigma_{iso}$ (ppm) | $\sigma_{33}$ (ppm) | $\sigma_{22}$ (ppm) | $\sigma_{11}$ (ppm) | $C_Q$ (MHz) | $\eta$ |
|---|---|---|---|---|---|---|---|---|
| Fmoc-L-tyrosine | X-ray | CO | -86.815 | 270.438 | -186.307 | -344.576 | 8.85790 | 0.13550 |
| | | COH | 134.330 | 245.223 | 166.099 | -8.331 | -9.07075 | 0.35039 |
| | ZORA/DZP | CO | -93.091 | 269.163 | -206.649 | -341.787 | 8.68533 | 0.13762 |
| | | COH | 108.232 | 241.788 | 136.433 | -53.524 | -8.64595 | 0.57711 |
| | ZORA/TZ2P | CO | -93.091 | 260.277 | -220.836 | -318.558 | 8.85298 | 0.19300 |
| | | COH | 112.222 | 243.997 | 142.508 | -49.840 | -8.70477 | 0.55091 |
| | ZORA/QZ4P | CO | -84.404 | 264.308 | -219.903 | -297.617 | 8.66492 | 0.15265 |
| | | COH | 108.858 | 233.659 | 148.15 | -55.234 | -8.69706 | 0.50636 |
| Fmoc-L-isoleucine | X-ray | CO | -85.878 | 281.835 | -190.302 | -349.168 | 8.93670 | 0.1735 |
| | | COH | 151.609 | 263.304 | 207.206 | -15.684 | -8.38861 | 0.3534 |
| | ZORA/DZP | CO | -77.928 | 266.204 | -212.547 | -287.441 | 8.55007 | 0.12112 |
| | | COH | 93.889 | 227.876 | 121.964 | -68.172 | -8.54004 | 0.53998 |
| | ZORA/TZ2P | CO | -78.05 | 268.236 | -215.276 | -287.106 | 8.59598 | 0.13323 |
| | | COH | 96.387 | 229.723 | 125.097 | -65.660 | -8.59363 | 0.50872 |
| | ZORA/QZ4P | CO | -75.628 | 266.393 | -211.022 | -282.000 | 8.55424 | 0.12526 |
| | | COH | 99.299 | 230.27 | 127.273 | -59.645 | -8.62249 | 0.49025 |

[1] All NMR parameters were computed using the ZORA/TZ2P basis set. The NMR parameters for the X-ray structure were obtained from a single molecule crystal structure and did not undergo geometry optimization.

**Table S4**. Density functional theory computed NMR parameters[1] on geometry-optimized structures.

| Sample | $^{17}$O site | $\sigma_{iso}$ (ppm) | $\sigma_{33}$ (ppm) | $\sigma_{22}$ (ppm) | $\sigma_{11}$ (ppm) | $C_Q$ (MHz) | $\eta$ |
|---|---|---|---|---|---|---|---|
| Fmoc-L-tyrosine | CO | -93.091 | 269.163 | -206.649 | -341.787 | 8.68533 | 0.13762 |
| | COH | 108.232 | 241.788 | 136.433 | -53.524 | -8.64595 | 0.57711 |
| Fmoc-L-proline | CO | -88.920 | 262.934 | -214.866 | -314.829 | 8.77591 | 0.14672 |
| | COH | 104.576 | 240.112 | 130.684 | -57.068 | -8.57213 | 0.56200 |
| Fmoc-L-threonine | CO | -82.932 | 253.228 | -228.309 | -273.716 | 8.44669 | 0.10659 |
| | COH | 91.851 | 213.172 | 136.461 | -74.080 | -8.50001 | 0.55040 |
| Fmoc-L-tryptophan | CO | -93.386 | 253.942 | -213.848 | -320.252 | 8.79774 | 0.14495 |
| | COH | 101.394 | 238.317 | 136.519 | -70.654 | -8.66472 | 0.56482 |
| Fmoc-L-isoleucine | CO | -77.928 | 266.204 | -212.547 | -287.441 | 8.55007 | 0.12112 |
| | COH | 93.889 | 227.876 | 121.964 | -68.172 | -8.54004 | 0.53998 |

[1] NMR parameters were calculated using zero-order regular approximation; triple-zeta polarized basis set (ZORA/TZ2P) after geometry optimization on each structure was implemented using the ZORA, double-zeta polarized basis set (ZORA/DZP).

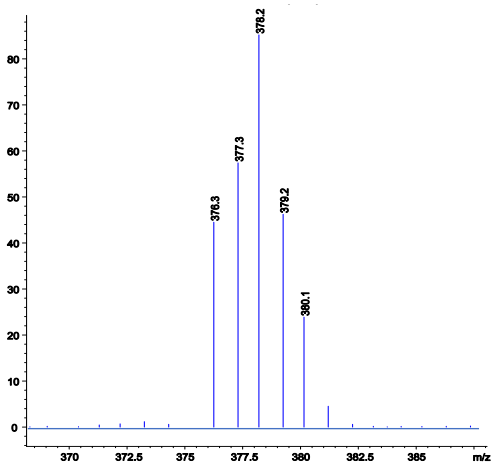


**Figure S1.** Mass spectrum of Fmoc-L-isoleucine after precipitation in acetonitrile and water. The expected MW for Fmoc-L-isoleucine is 353.41 Da and the $[M+H]^+$ series shown in the spectrum have an additional +23 Da due to the presence of $Na^+$. The α-COOH can have a mixture of contributions from $^{16}O/^{16}O$ at 376.3 m/z, $^{16}O/^{17}O$ or $^{17}O/^{16}O$ at 377.3 m/z, $^{17}O/^{17}O$ or $^{16}O/^{18}O$ or $^{18}O/^{16}O$ at 378.2 m/z, $^{18}O/^{17}O$ or $^{17}O/^{18}O$ at 379.2 and $^{18}O/^{18}O$ at 380.1 m/z

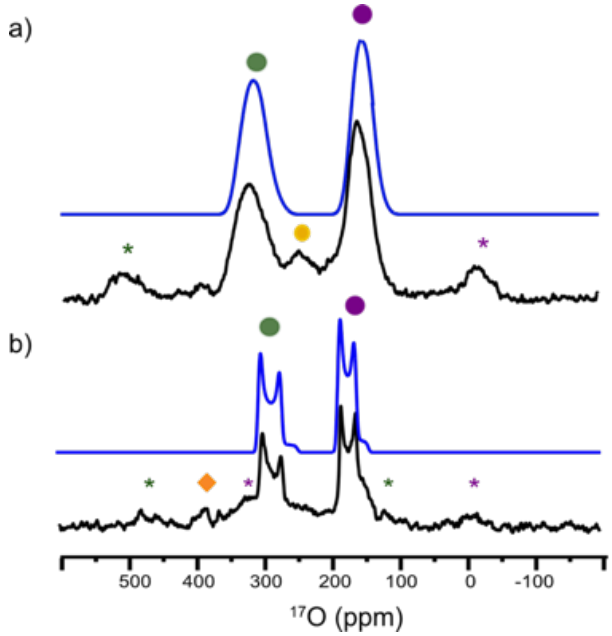


**Figure S2.** Experimental (black) and computed (blue) MAS NMR spectra of lyophilized (a) and recrystallized Fmoc-L-threonine (b) acquired at 21.1 T. The CO (green circle) and COH (purple circle) $^{17}O$ NMR parameters are as follows: $\delta_{iso}$ = 343 and 174 ppm, $C_Q$ = 8.5 and 7.4 MHz and $\eta$ = 0.50 and 0.40, respectively, for the lyophilized sample (a) and $\delta_{iso}$ = 311 and 192 ppm, $C_Q$ = 8.1 and 7.2 MHz and $\eta$ = 0.05 and 0.08, respectively, for the recrystallized sample (b). The orange diamond in (b) indicates $^{17}O$ signal from the $ZrO_2$ from the NMR rotor and the coloured asterisks indicate spinning sidebands.

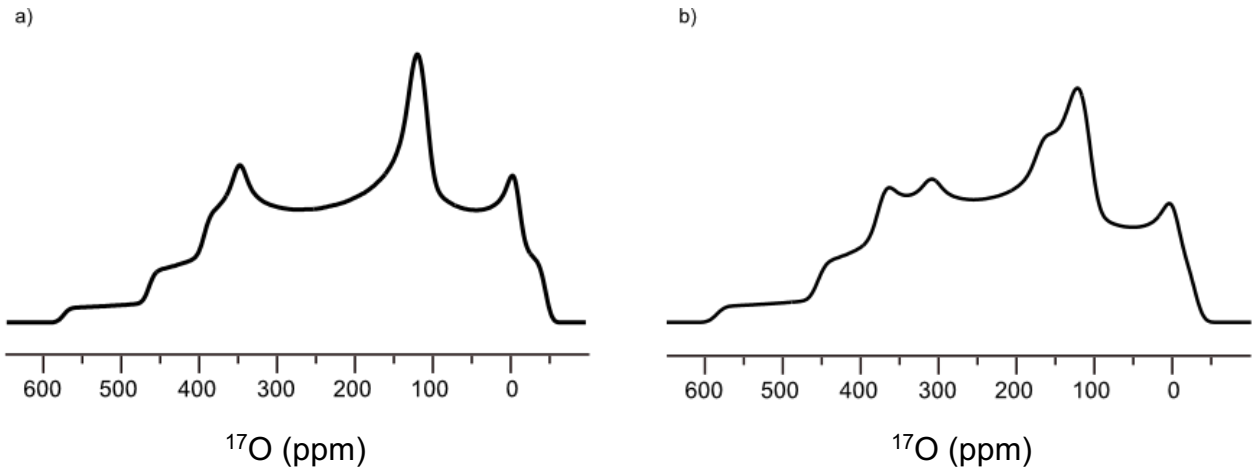


**Figure S3.** Simulation of non-spinning NMR spectra at 17.6 T for Fmoc-L-leucine (a) and Fmoc-L-valine (b) from Keeler *et al.* (2017) using spans (Ω) for the CO site of 530 and 540 ppm, respectively. All other NMR parameters are identical to what are reported in Table 1.

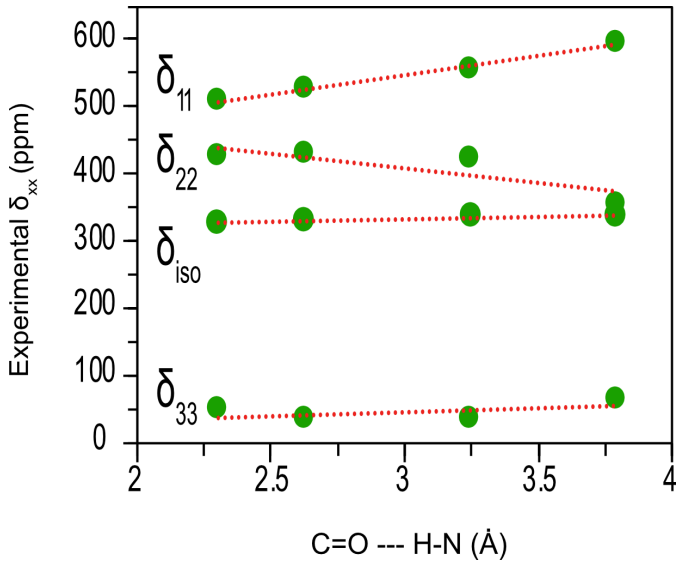


**Figure S4.** Correlation of the intramolecular hydrogen bond distance between the NH and C=O group from X-ray structures with the experimental CSA tensor components for the carbonyl oxygen site. X-ray structure of Fmoc-L serine (CIF-4514078), -tyrosine (CIF-7220755), -isoleucine (CIF-2219071) and -leucine (CIF-2218405) were used to measure the distance between the C=O and H-N groups in each molecule. The linear correlations are as follows : $\delta_{11}$ (y = 56.5x + 376.6, $R^2$ = 0.99), $\delta_{22}$ (y = -45.0x + 543.2, $R^2$ = 0.71), $\delta_{iso}$ (y = 7.3x + 311.4, $R^2$ = 0.93) and $\delta_{33}$ (y = 10.4x + 14.5, $R^2$ = 0.24).

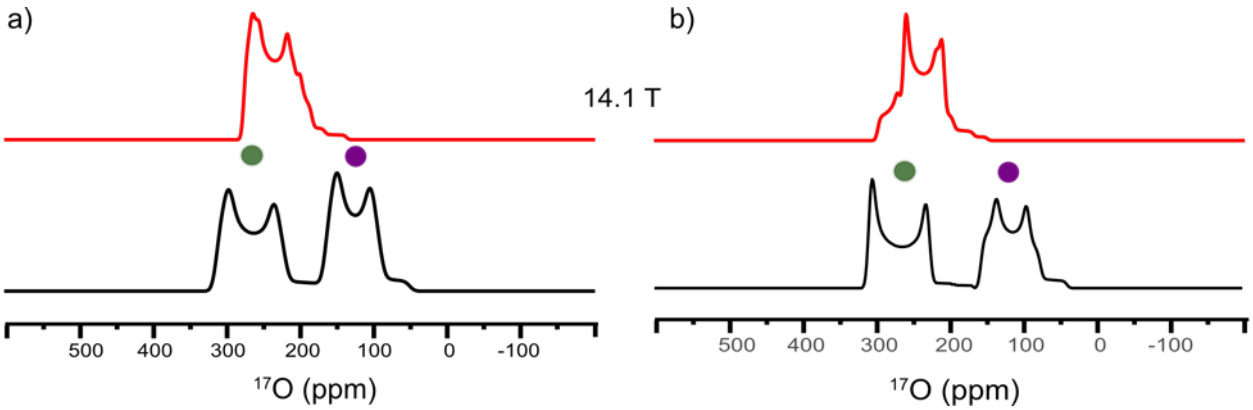


**Figure S5.** Simulated MAS NMR spectra of deprotonated (red) and protonated (black) Fmoc-L-tyrosine (a) and Fmoc-L-isoleucine (b) at 14.1 T. Deprotonated spectra were simulated using new computed NMR parameters for $\delta_{iso}$ and $C_Q$ after a proton was removed from the original geometry-optimized structures. The asymmetry values ($\eta$) were held constant between both species and the $\delta_{iso}$ and $C_Q$ for the CO (green circle) and COH (purple circle) sites in the protonated FMOC-amino acids are reported in Table 1.