**An improved C-P cross-coupling route for the synthesis of novel V-shaped aryldiphosphonic acids**

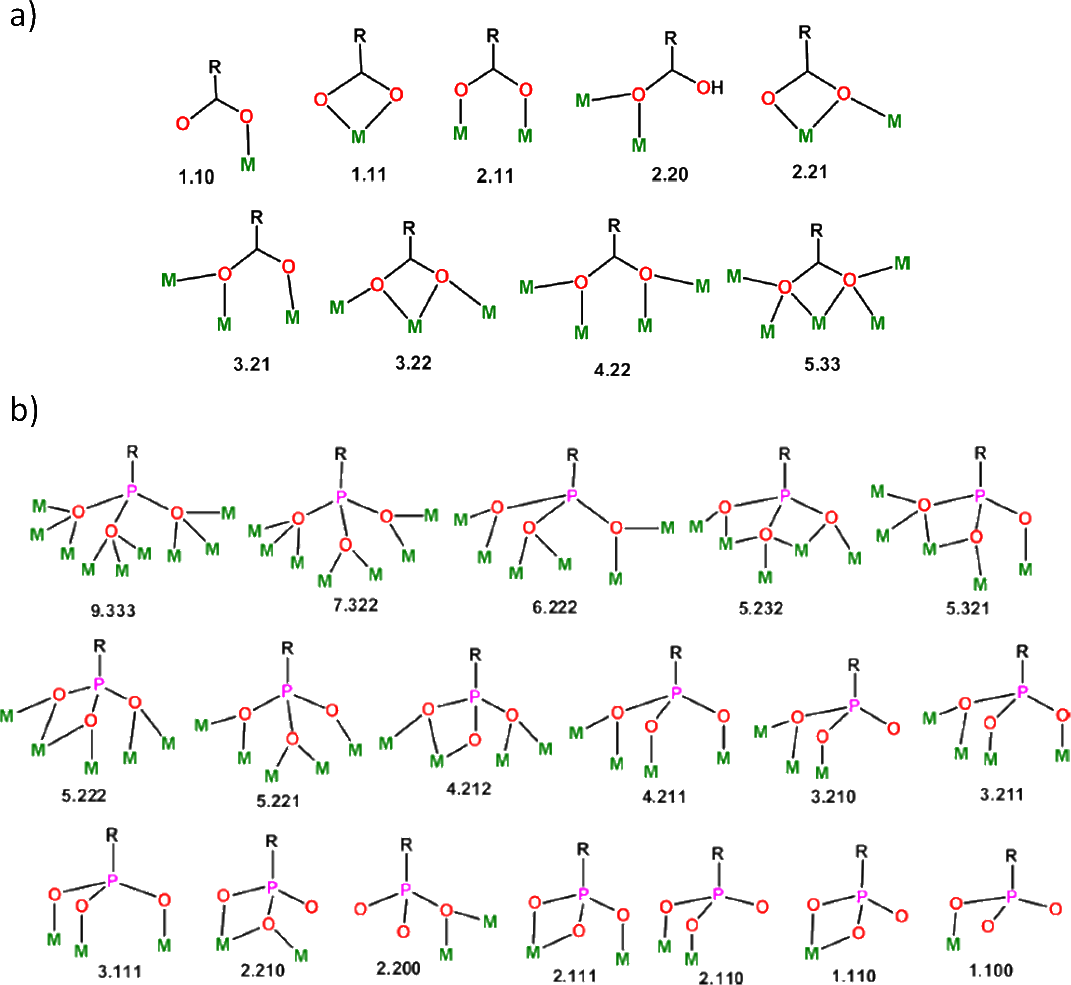
*Stephen J. I. Shearan1, Enrico Andreoli1, and Marco Taddei\*1,2*

Address: 1Energy Safety Research Institute, Swansea University, Fabian Way, Crymlyn Burrows, Skewen, Swansea SA1 8EN and 2Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Via Giuseppe Moruzzi, 13 - 56124 - Pisa (Italy)

Email: Marco Taddei – marco.taddei@unipi.it

\* Corresponding author

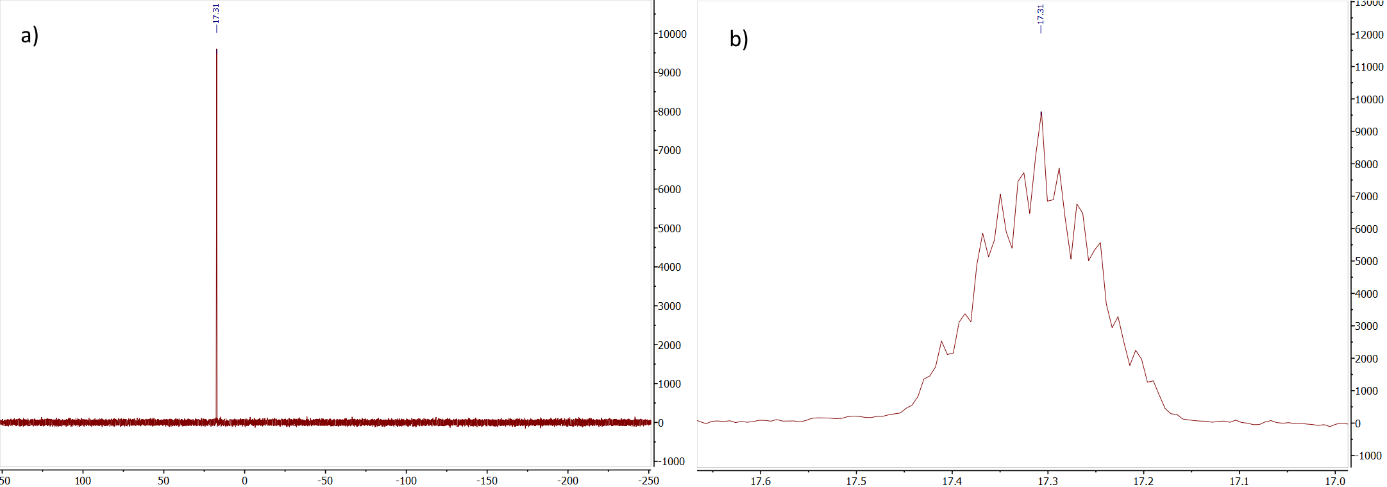
**SUPPORTING INFORMATION**

******

**Figure S1**: A comparison between the bonding modes of a) carboxylates, and b) phosphonates. Adapted with permission from Goura, J.; Chandrasekar, V. *Chem. Rev.* **2015**, *115*, 6854-6965. Copyright 2015 American Chemical Society.

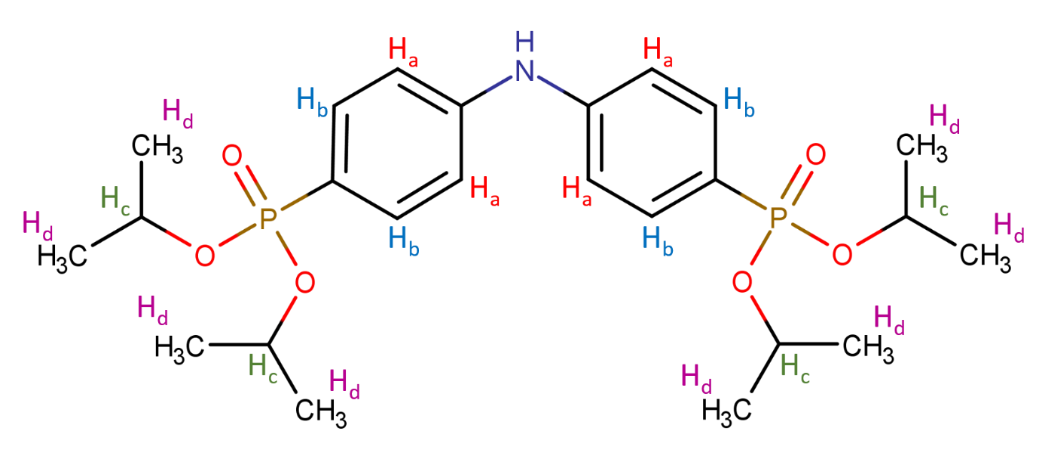
***N*,*N*-Bis(4-diisopropylphosphonophenyl)amine (iPr4BPA) (1A)**

The first synthetic target obtained was *N*,*N*-Bis(4-diisopropylphosphonophenyl) amine, known herein as iPr4BPA (1A). The procedure to obtain this compound was relatively simple and reproducible, whereby the product could be obtained consistently at yields above 80%. Before the workup of the crude reaction product, which took the form of a dark brown treacle-like mixture, thin layer chromatography (TLC) analysis often showed that some starting material was still present, alongside what is reasoned to be the mono-substituted intermediate, *N*-(4-diisopropylphosphonophenyl)-*N*-phenylamine. This is easily removed by employing flash chromatography, using a mixture of acetone and ethyl acetate in a 1:9 ratio. The first product to elute off the column is the starting material, followed by the mono-substituted intermediate. This leaves a pure iPr4BPA (1A) product in the final vials of eluate. The pure product could then be obtained quickly by rotary evaporation to remove the solvent, resulting in a white solid in yields most often between 80 to 90%. The phase purity of this was then confirmed by 31P, 1H, 13C, 2D HSQC NMR spectroscopy, and mass spectrometry. The solvent used for NMR was CDCl3.



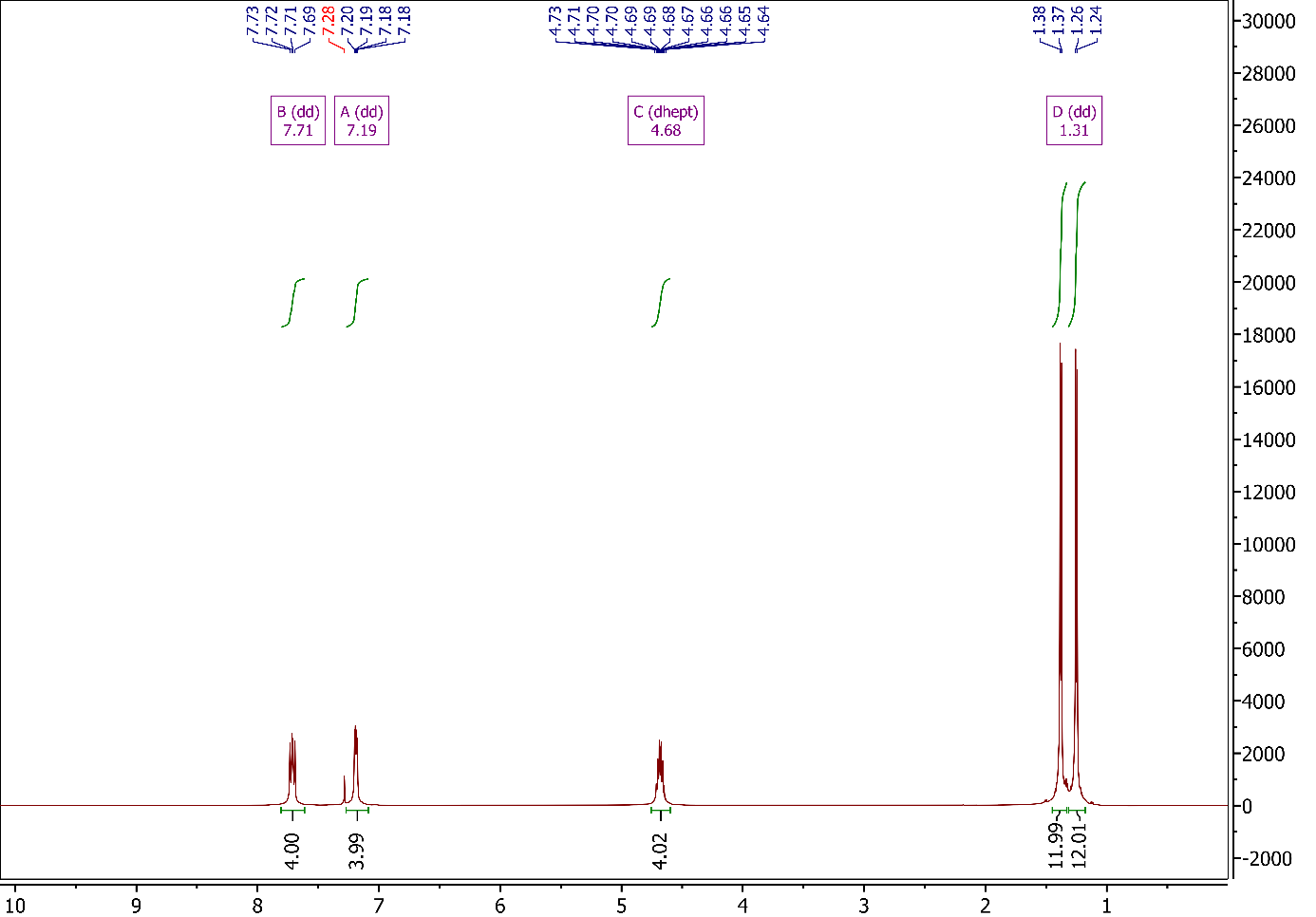
**Figure S2**: a) Full 31P NMR spectrum, and b) zoomed 31P NMR spectrum for iPr4BPA (1A).

Looking at the full 31P spectrum in Figure S2a we see a single signal at δ 17.3 ppm which corresponds to the phosphonate ester product, suggesting that no other phosphorus-containing impurities are present. A closer look at this signal, as shown in Figure S2b reveals a complex multiplet splitting pattern with coupling constants of between 4.4 and 50 Hz. Coupling constants at the lower end of this range indicate that three- and four-bond couplings are present, suggesting that phosphorus is coupling with the protons present on the methyl groups on the isopropyl moieties as well as the aromatic protons furthest away from phosphorus on the corresponding rings and those of the methine groups. At the upper end of this range, it is expected that phosphorus is coupling with the protons closest to it on the aromatic ring.



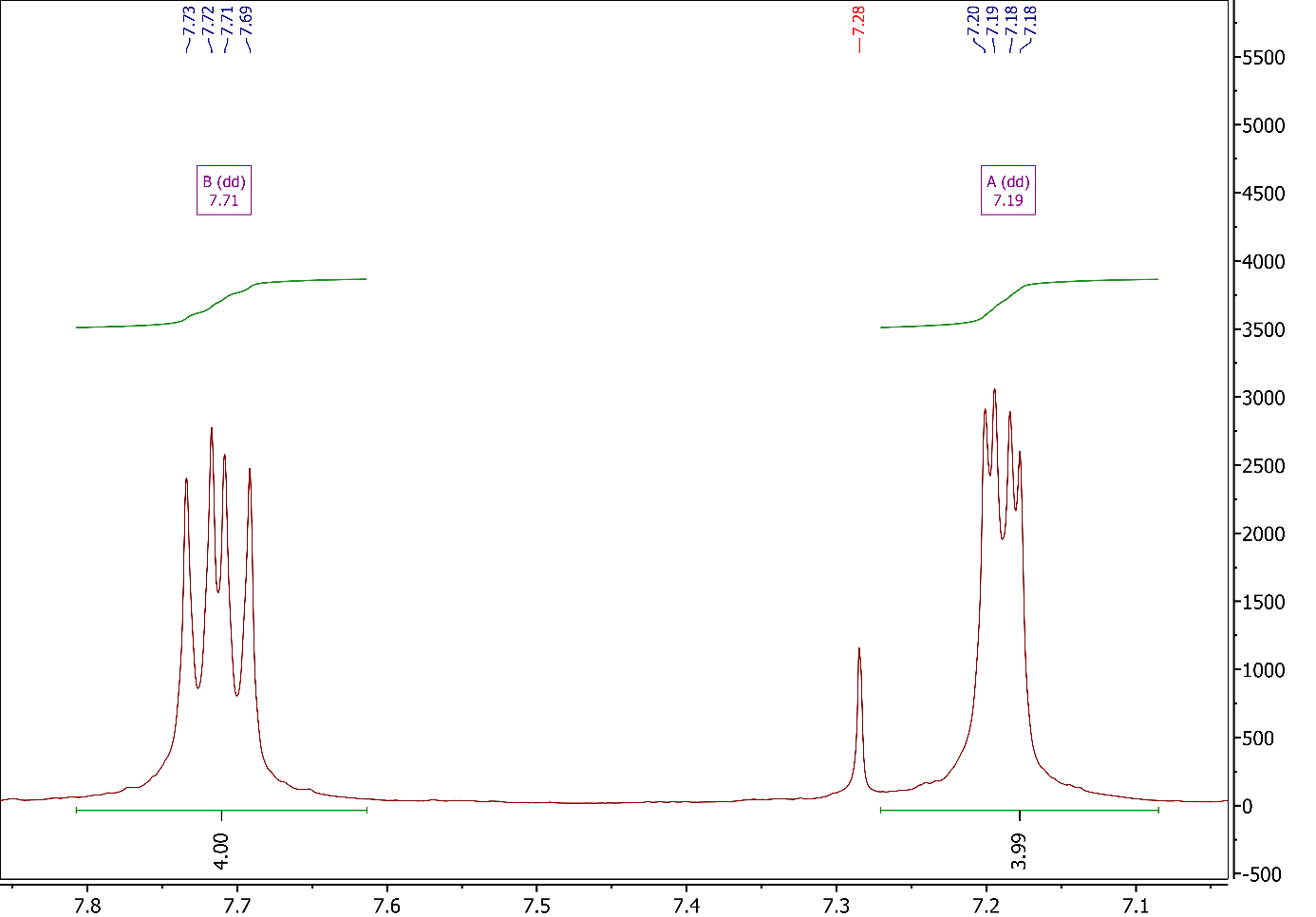
**Figure S3**: Chemical structure of iPr4BPA with proton environments labelled Ha-Hd.

With the positive result demonstrated on the 31P spectrum, 1H NMR provides a similar story. A quick investigation of the full spectrum, as shown in Figure S4, immediately reveals the correct number of proton environments for the molecule, there are no significant impurities, and we see that each of the signals integrate to give the correct number of protons in each environment.



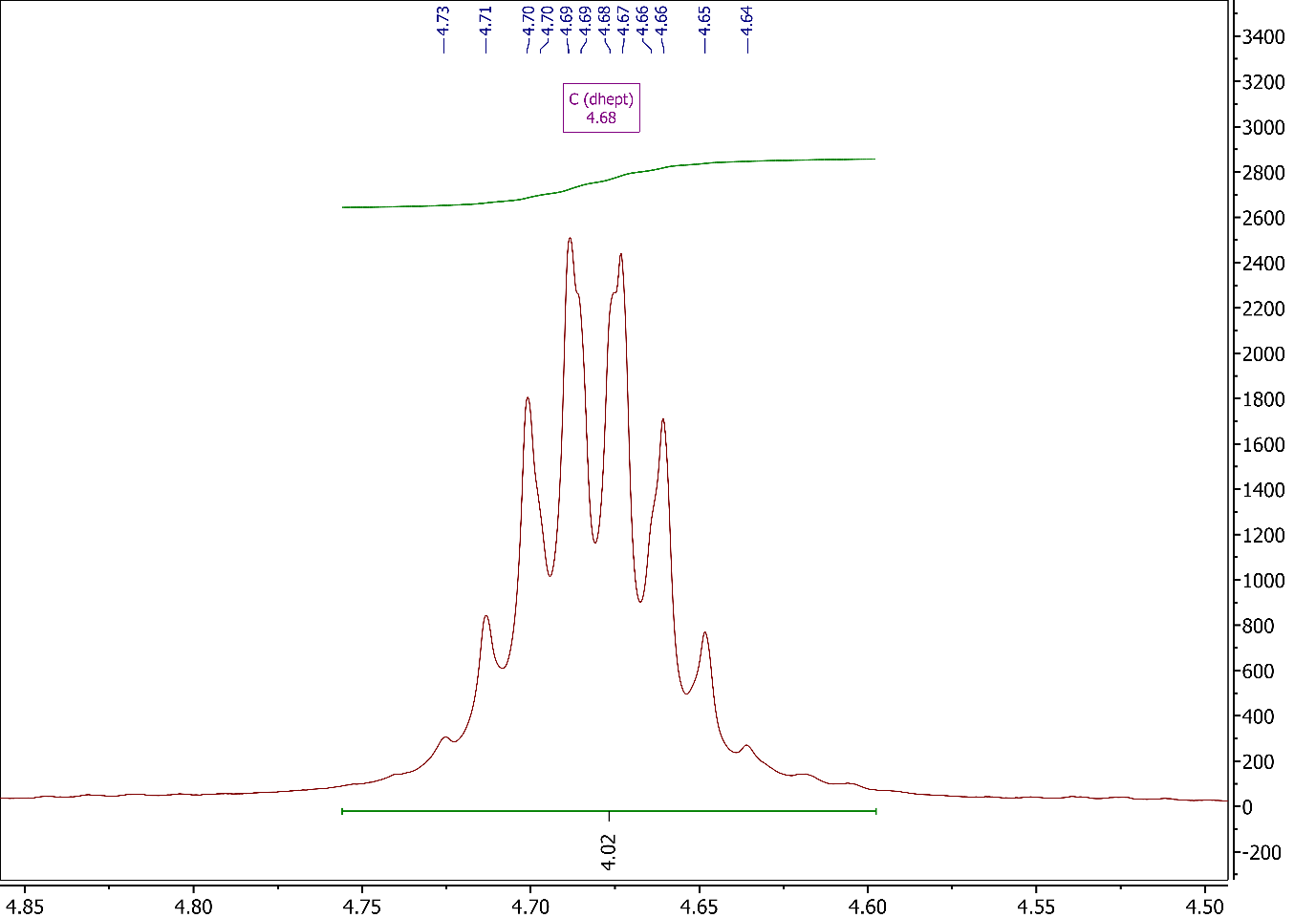
**Figure S4**: Full 1H NMR spectrum for iPr4BPA (1A).

Looking specifically at the aromatic region (Figure S5) we see three peaks, including a singlet at δ 7.28 ppm, which is due to the deuterated solvent (chloroform-*d*) and two doublets of doublets at δ 7.71 ppm and δ 7.19 ppm, which correspond to the protons on the phenyl rings. On first viewing of these signals, it was tempting to assign the more downfield of the two (δ 7.71 ppm) to the protons nearest to nitrogen, represented as Ha in Figure S3, though examination of the coupling constants reveals that the downfield signal has a larger coupling constant. This implies that the signal actually represents the protons closest to the phosphonate group (Hb), since phosphorus, like a proton, has a quantum spin number of 1/2, and will thus be involved in coupling with protons. This means that the upfield signal at δ 7.19 ppm can be assigned to the protons nearest to nitrogen (Hb).



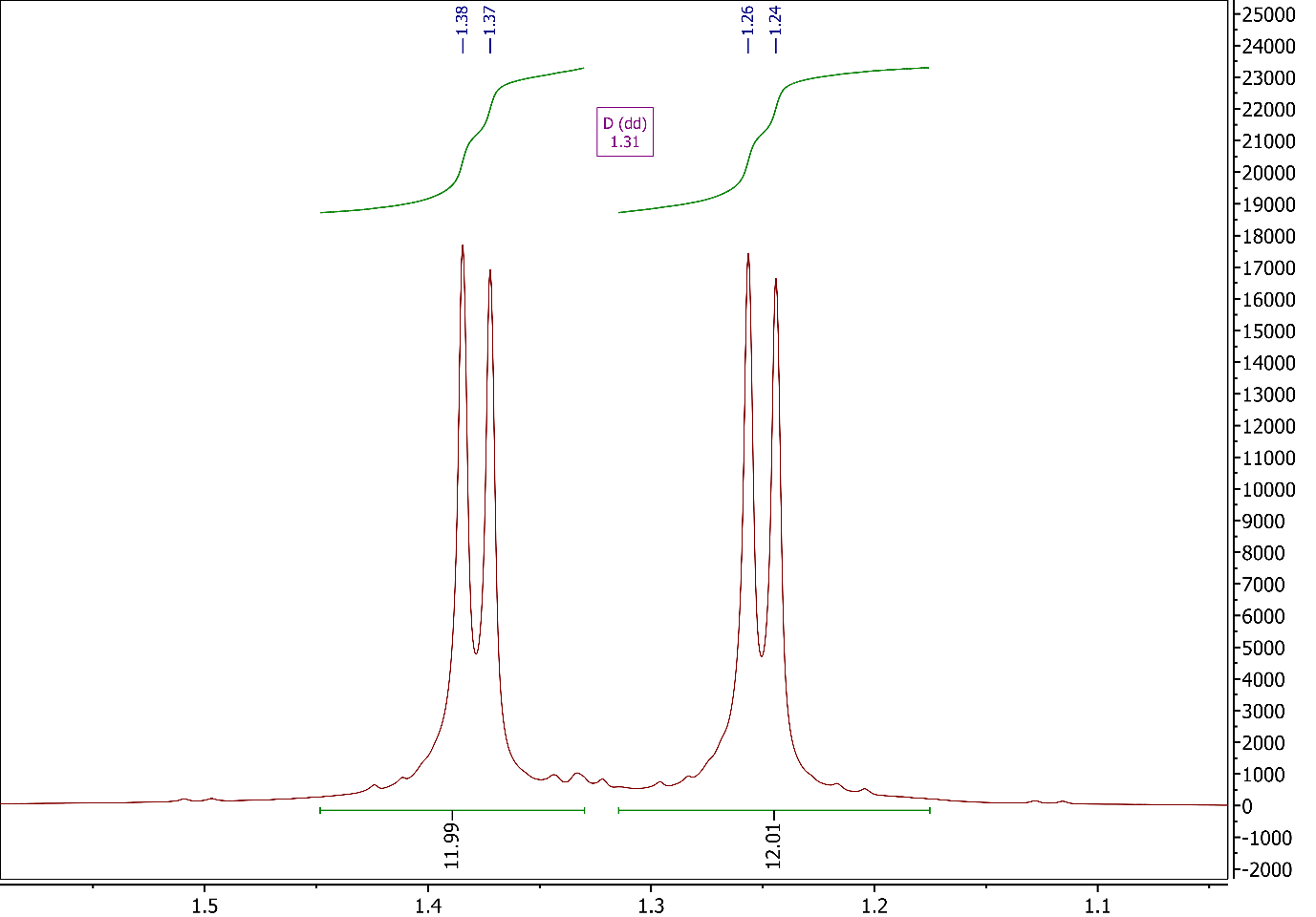
**Figure S5**: Aromatic region of the 1H NMR spectrum for iPr4BPA.

Moving further upfield (Figure S6), there is a signal at δ 4.68 ppm which is easily attributed to the methine group (Hc) of the isopropyl moiety. One might expect to see a septuplet here arising from the six aliphatic protons on each isopropyl moiety, though it appears to be a sextet. Closer analysis would appear to suggest that, in fact, the signal is a pair of overlapping sextets, since there are clear shoulders on some of the peaks. In this sense, however, multiplet analysis appears to miss two shoulders that would in fact make this a doublet of "leaning" septets. This analysis appears to be confirmed in the 1H NMR spectrum for iPr4DPC (2A), as seen in Figure S26, whereby a multiplet analysis appears to identify the previously hidden shoulders. The complicated nature of this signal is likely brought about by the coupling of 31P and 1H nuclei, as was mentioned in the discussion for the aromatic signals. Most importantly, the signal integrates to give the correct number of protons for the methine groups on the molecule.



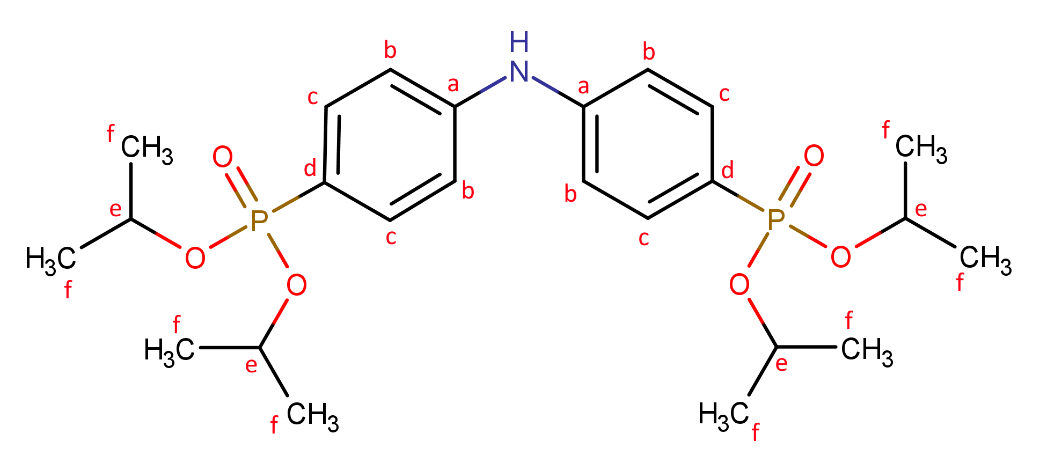
**Figure S6**: Methine region of the 1H NMR spectrum for iPr4BPA.

The finals signals on this spectrum, as seen in Figure S7, were originally thought to be a well-resolved doublet of doublets that corresponds to the terminal methyl groups (Hd) on the phosphonate ester. However, due to the relatively large coupling constant and the need to integrate the peaks separately, it was determined that the signals are more likely two doublets, indicating that there are two types of chemically inequivalent isopropyl groups present on the molecule, though how these arise has not been investigated. Working on this basis, the signal integrates to a total of 24 protons, as expected. Each of the doublets arises from the coupling with the methine proton.

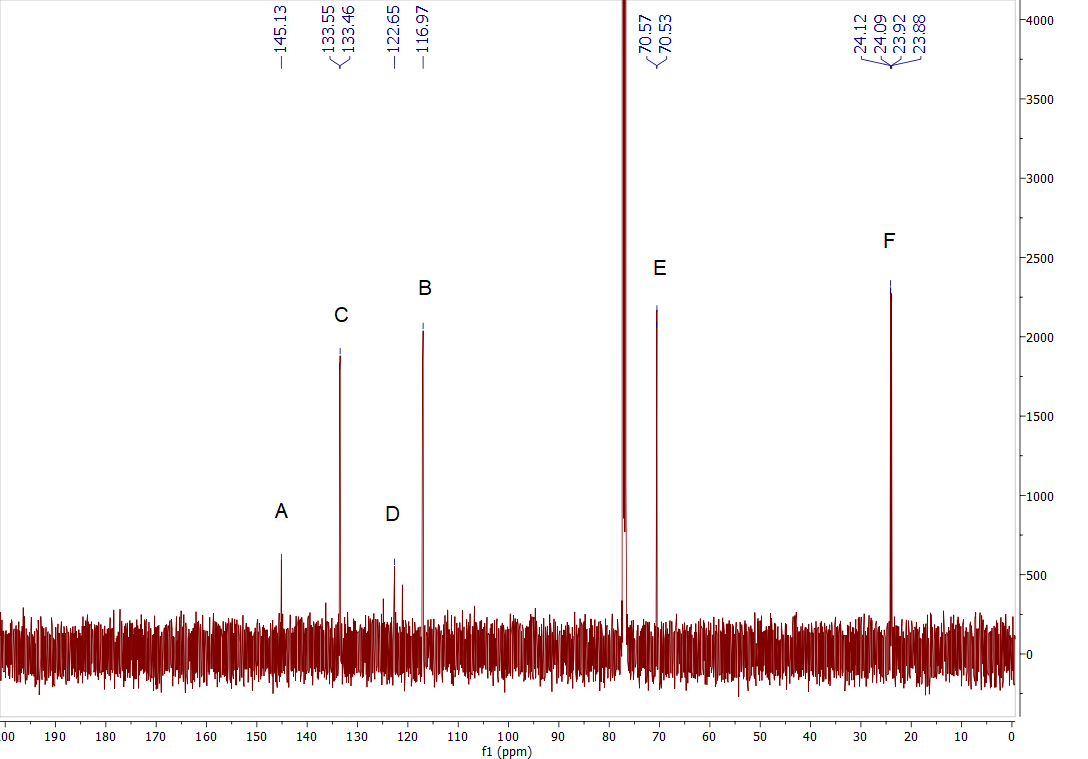


**Figure S7**: Methyl region of the 1H NMR spectrum for iPr4BPA.

Moving to the 13C NMR spectrum, as shown in Figure S9, we can immediately see the number of signals expected. Starting downfield, the signal at δ 145.13 ppm is the result of the carbons bonded to nitrogen, labelled A. Moving upfield, we see the peak at δ 133.50 ppm corresponding to the four carbons labelled C. Next is the peak at δ 122.65 ppm, which corresponds to the two carbons bonded to phosphorus, labelled D. The final peak for the aromatic region is the one labelled B at δ 116.97 ppm, which corresponds to four carbons in the β-positions relative to nitrogen. We then see a large peak at δ 76.0 ppm which goes off the scale and belongs to the solvent, CDCl3. Next to this, at δ 70.55 ppm, is the peak attributed to methine carbon on the isopropyl groups. The final peak, at δ 24.00 ppm, corresponds to the terminal methyl carbons on the isopropyl groups.

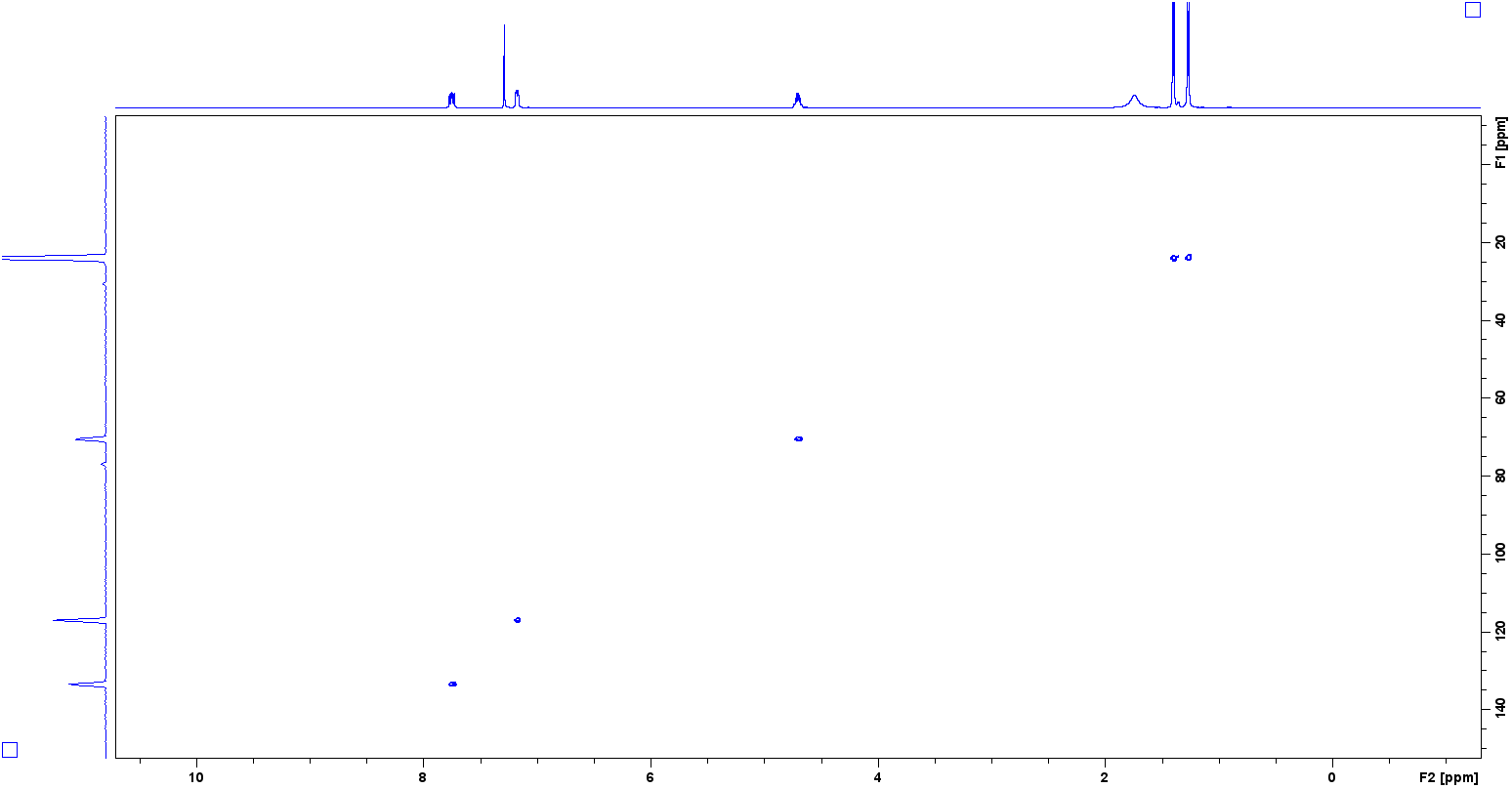


**Figure S8**: Chemical structure of iPr4BPA with carbon environments labelled Ha-Hf.



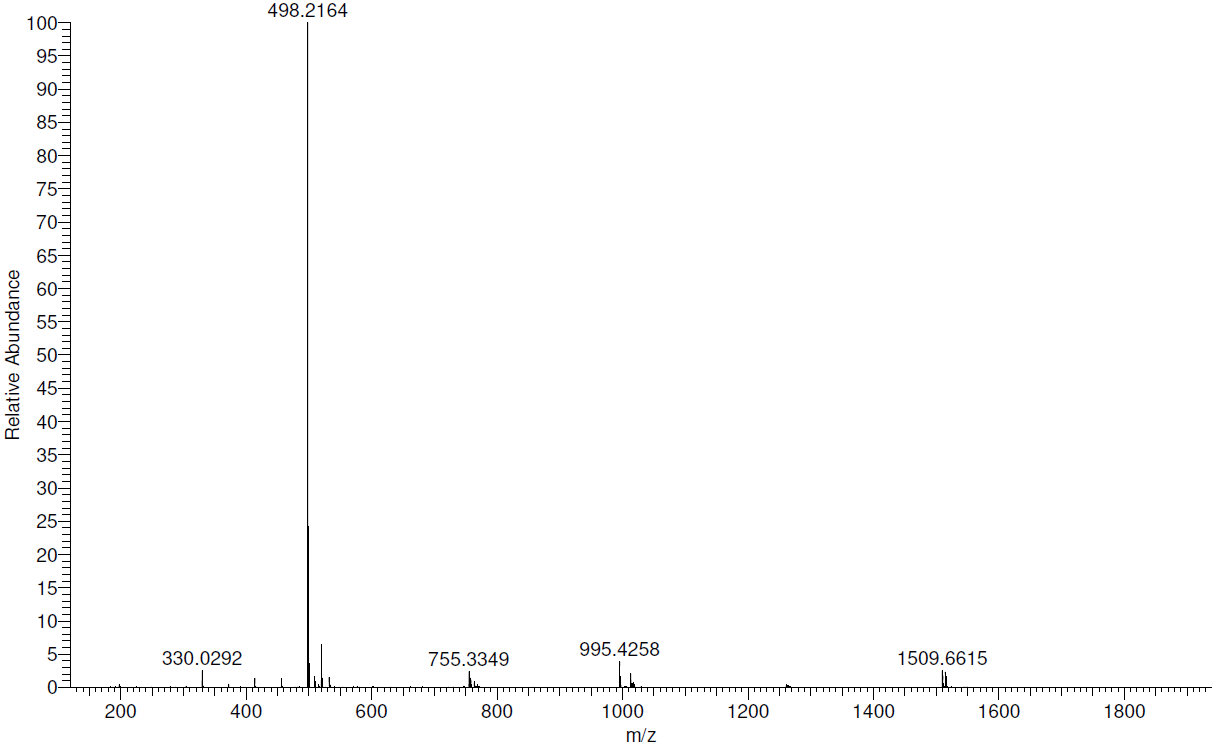
**Figure S9**: Full 13C NMR spectrum for iPr4BPA.

In order to obtain further confirmation of the NMR results, 2D H-C coupled NMR (HSQC) was carried out, which at least allows the assignment of carbon signals based on those for the proton signals. Looking at Figure S10, we can immediately see that the furthest downfield signals on both spectra, at δ 133.50 ppm, correlate to B and C on the proton and carbon NMR spectra, respectively. The other correlation in the aromatic region, at δ 116.97 ppm, is shown between B on the proton NMR spectrum and C on the carbon NMR spectrum. The next two signals, at δ 70.55 ppm and δ 24.00 ppm correctly show the H-C correlations for the methine and methyl groups.

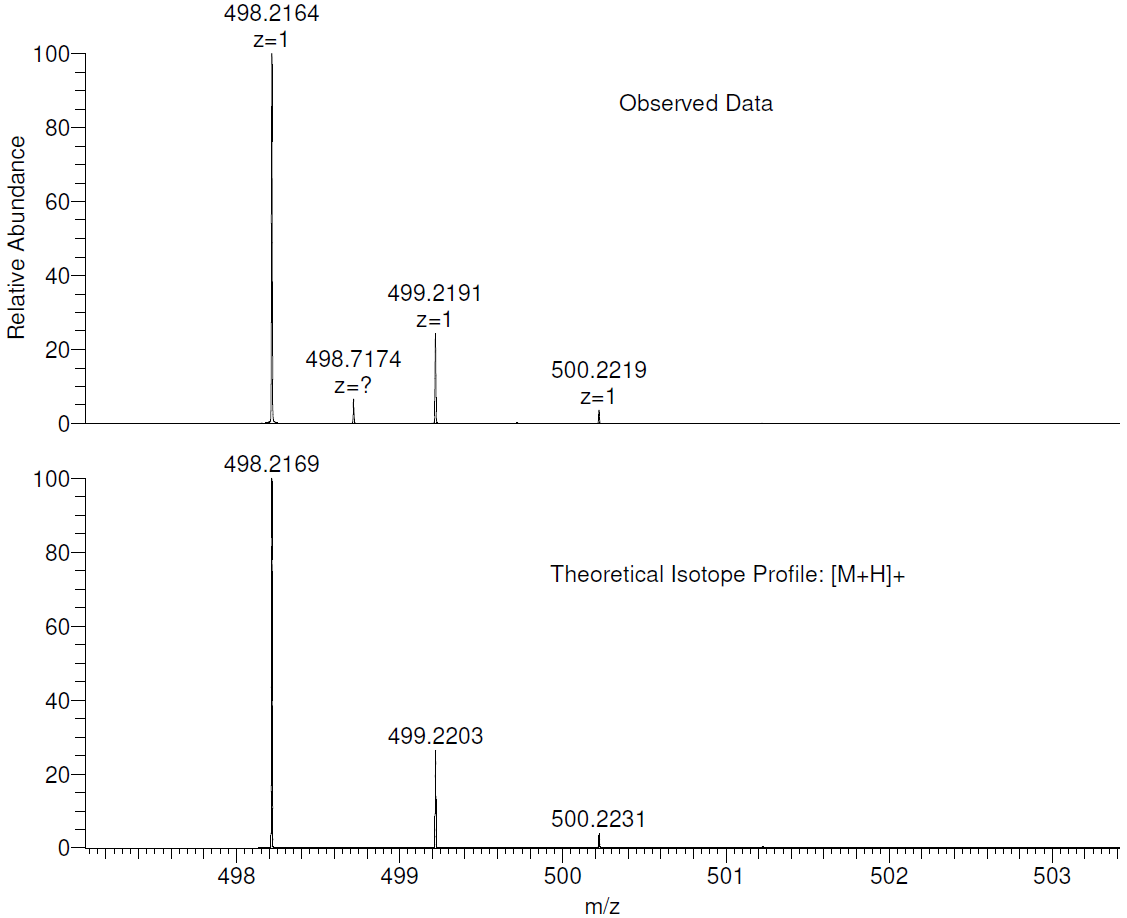


**Figure S10**: Full HSQC NMR spectrum for iPr4BPA.

With NMR data clearly showing successful synthesis of iPr4BPA, mass spectral analysis was carried out to provide further support. Figure S11 shows the full mass spectrum, where the base peak can be seen with an m/z of 498.22 and is shown in more detail in Figure S12. Here we see the theoretical isotope profile for the [M+H]+ species and the observed experimental data, with an almost exact match between the two. Moving back to the full spectrum, the peak at m/z 995.43 corresponds a dimer species, [2M+H]+, of the target molecule. The peak at m/z 330 has a clear relation to those at m/z 755.33 and 1509.66, whereby the former is likely a degradation product of the target molecule, likely corresponding to the acid form, [N(C6H4)2+H]+, and the latter two corresponding to the dimer and trimer species, respectively.



**Figure S11**: Full mass spectrum for iPr4BPA.

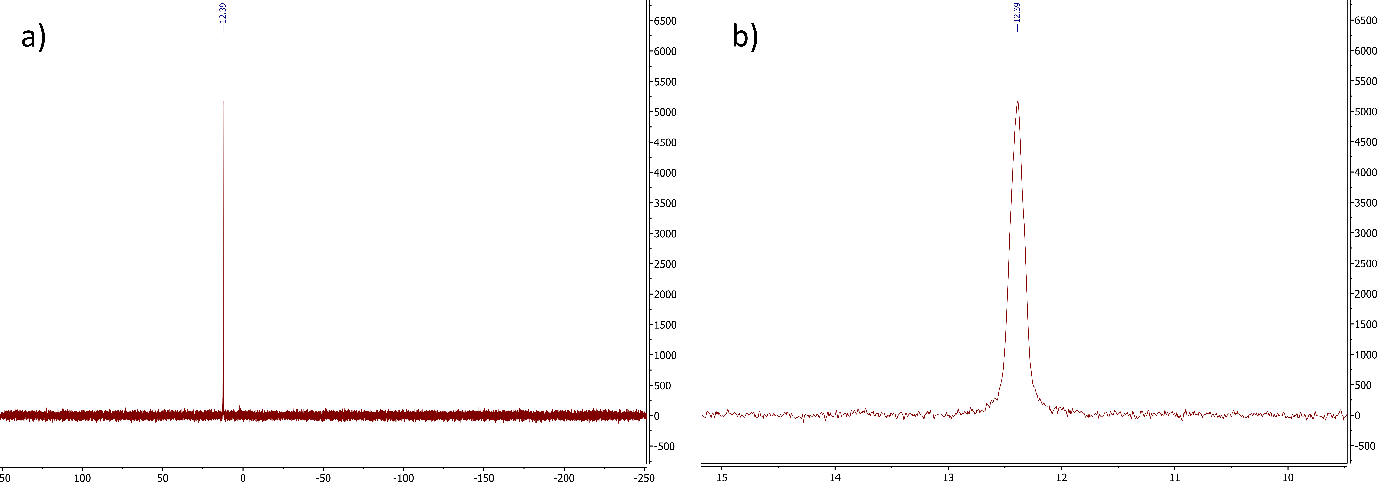


**Figure S12**: Mass spectrum for iPr4BPA showing the observed data and the theoretical isotope profile for the [M+H]+ species.

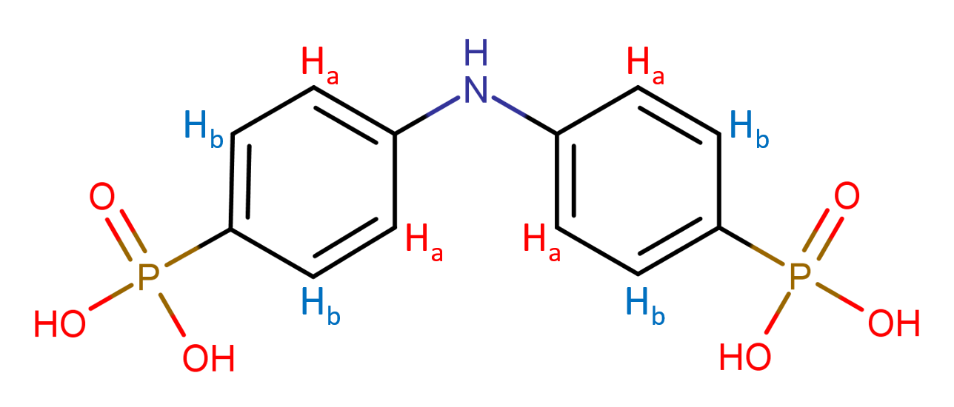
***N*,*N*-Bis(4-phosphonophenyl)amine (H4BPA) (1B)**

With a clean product obtained for iPr4BPA (1A), the attention shifted to obtaining the corresponding phosphonic acid, *N*,*N*-Bis(4-phosphonophenyl)amine H4BPA (1B). The initial method considered for this was hydrolysis using 6 M hydrochloric acid under reflux for 12+ h. The results here were quite hit and miss, as it was found that this route sometimes led to the cleavage of the C-P bond, which was confirmed by the absence of signals on the 31P NMR spectrum. As an alternative to these harsh conditions, a different route which involved silylation using trimethylsilyl bromide (TMSiBr) and subsequent hydrolysis using water was chosen. Overall, this method proved very simple and required very little work up, often just requiring separate washes with water and acetone, and leading to yields between 65 to 91%. As with the ester, purity is confirmed with both 31P, 1H, 13C, and HSQC NMR spectroscopy, and mass spectrometry. The solvent used for NMR was a 0.1 M solution of NaOH in D2O.

Looking first at the 31P spectrum (Figure S13), it is clear that the splitting pattern has completely changed this time taking the form of a singlet, which suggests that successful hydrolysis and removal of the isopropyl groups has occurred. The peak are less resolved in this case, so no splitting patterns are observed despite the peak width remaining similar to that of the ester, as shown in Figure S2. It is also clear that no phosphorus-containing impurities are present, and thus the washing protocol used is satisfactory.

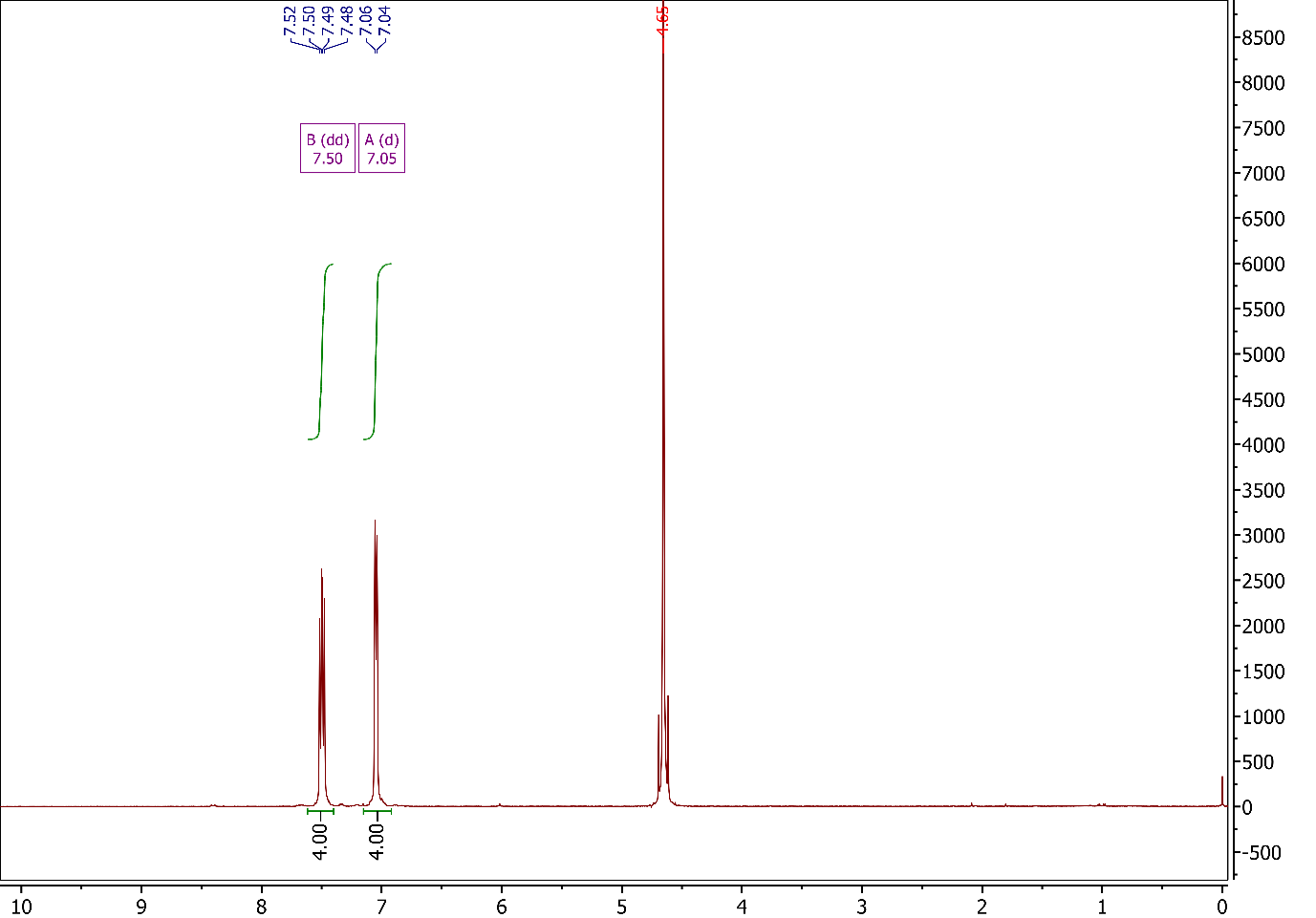


**Figure S13**: a) Full 31P NMR spectrum, and b) zoomed 31P NMR spectrum for H4BPA (1B).



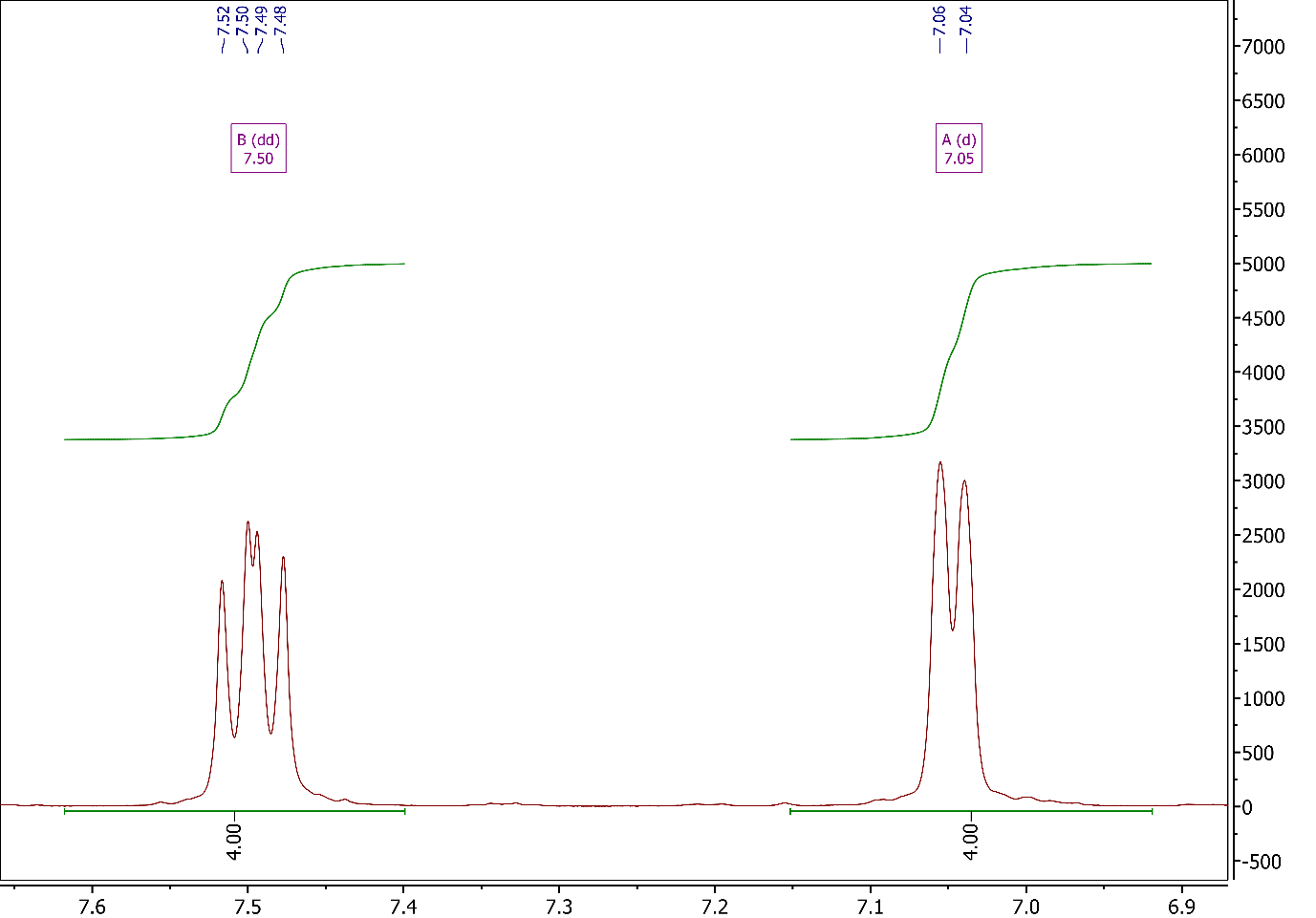
**Figure S14**: Chemical structure of H4BPA with proton environments labelled Ha-Hb.

Moving next to the 1H NMR spectrum, as shown in Figure S15, it is clear that the isopropyl groups are no longer present, since the signals for the methyl (-CH3) and methine (-CH-) groups have disappeared, giving further confirmation that hydrolysis has successfully taken place.



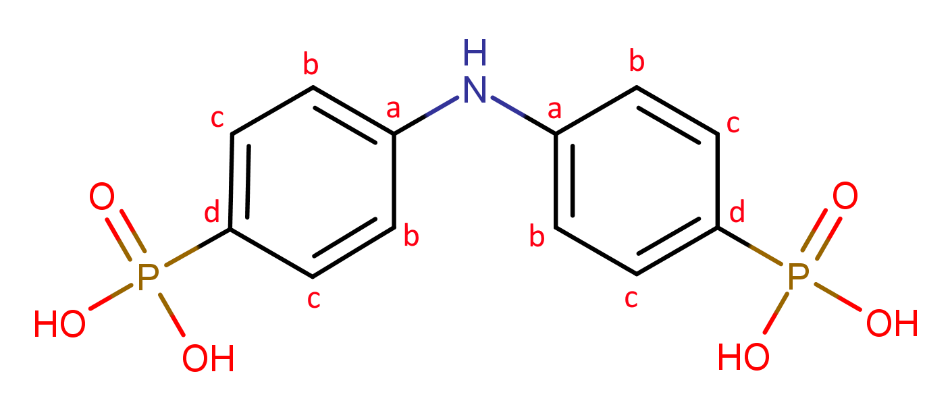
**Figure S15**: Full 1H NMR spectrum for H4BPA (1B).

Taking a closer look at the aromatic region, as shown in Figure S16, the signals can be assigned as they were for iPr4BPA (1A), whereby the downfield signal at δ 7.50 ppm (B) can be assigned to the protons nearest to phosphorus on the phenyl ring, represented by Hb, since it has a larger coupling constant. The signal at δ 7.05 ppm (A) can be assigned to the protons nearest to nitrogen on the phenyl ring, represented by Ha. This signal, while appearing as a single doublet, is more likely to be an unresolved doublet of doublets, as was seen in Figure S5 for iPr4BPA (1A).

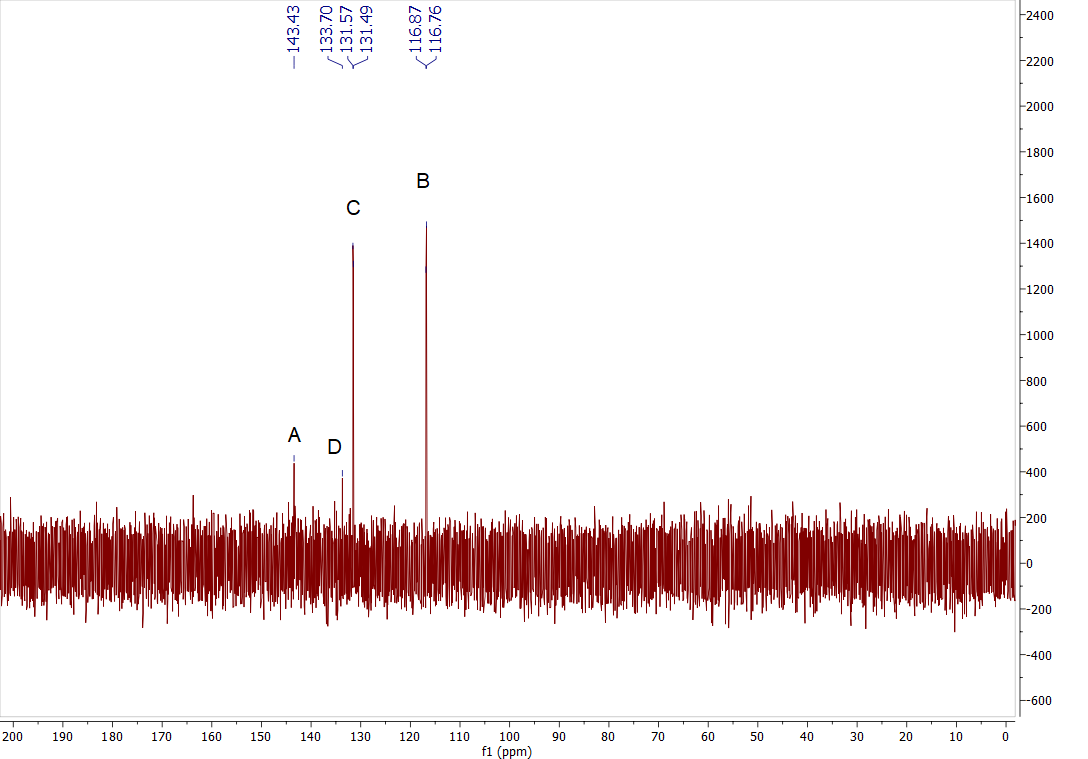


**Figure S16**: Aromatic region of the 1H NMR spectrum for H4BPA.

Moving to the 13C NMR spectrum, as shown in Figure S18, we can immediately confirm the lack of signal for the methine and methyl groups. Starting downfield, we see the peak, labelled A on the spectrum, which has been assigned to the two carbon atoms bonded to nitrogen, as was the case for iPr4BPA. The next signal, labelled D, has been assigned to the two carbons bonded to phosphorus, which has actually shifted when compared with the spectrum for iPr4BPA. The next two signals then are assigned to the carbons with bonded protons, labelled C and B respectively, with the latter of these being the more downfield. It was unfortunate here, in fact, that the acid is relatively insoluble and thus obtaining a spectrum with reasonable intensities was not possible, even though the signals were able to be assigned.

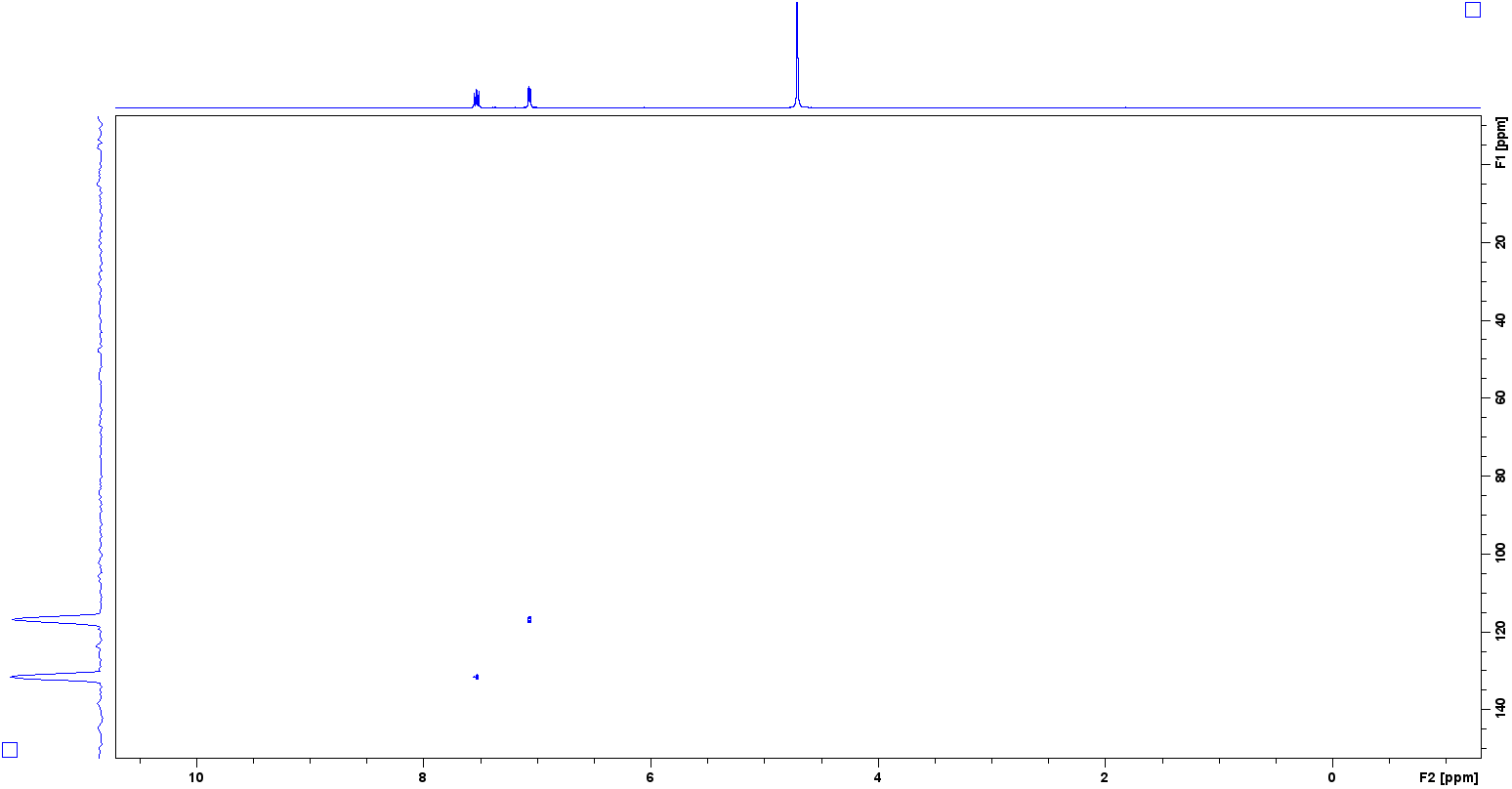


**Figure S17**: Chemical structure of H4BPA with carbon environments labelled Ha-Hf.



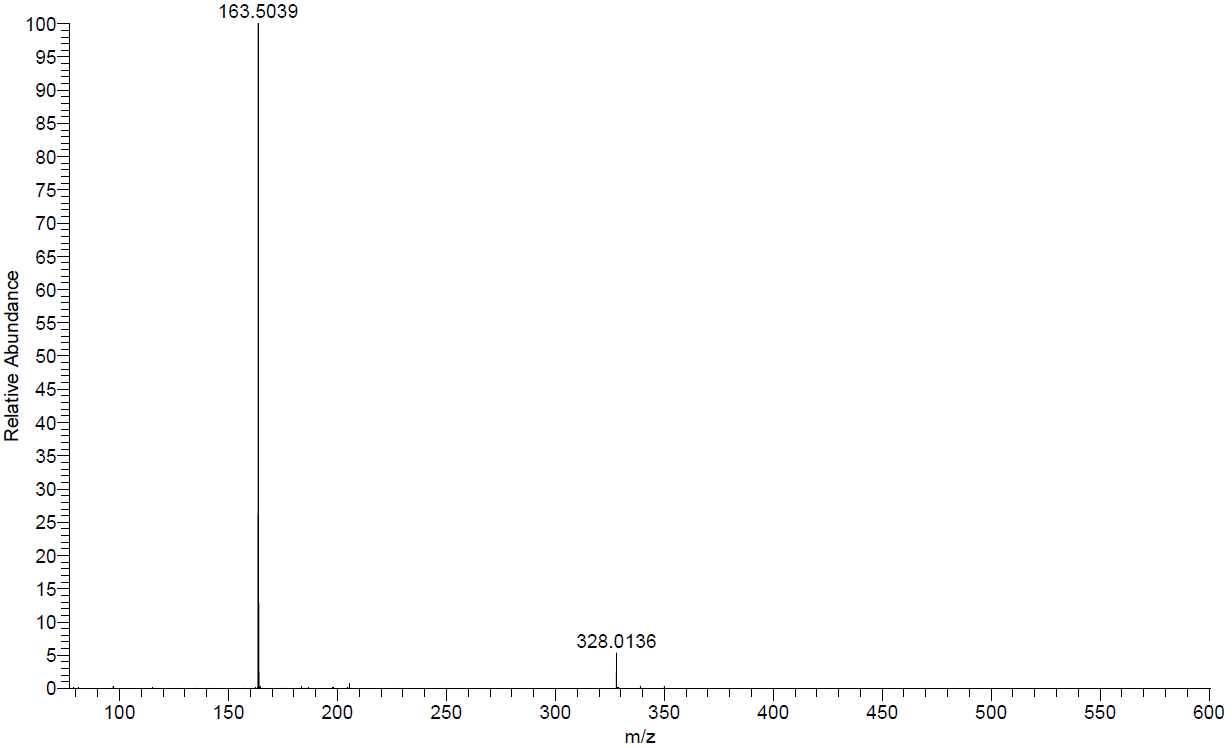
**Figure S18**: Full 13C NMR spectrum for H4BPA.

As was done for iPr4BPA, we again use coupled C-H NMR (HSQC) to lend support for the assignments for H4BPA NMR spectra. Looking at the spectra in Figure S19, we can again see the lack of signals associated with the isopropyl groups. We can, however, see the two expected signals for the two sets of aromatic carbons with associated protons.

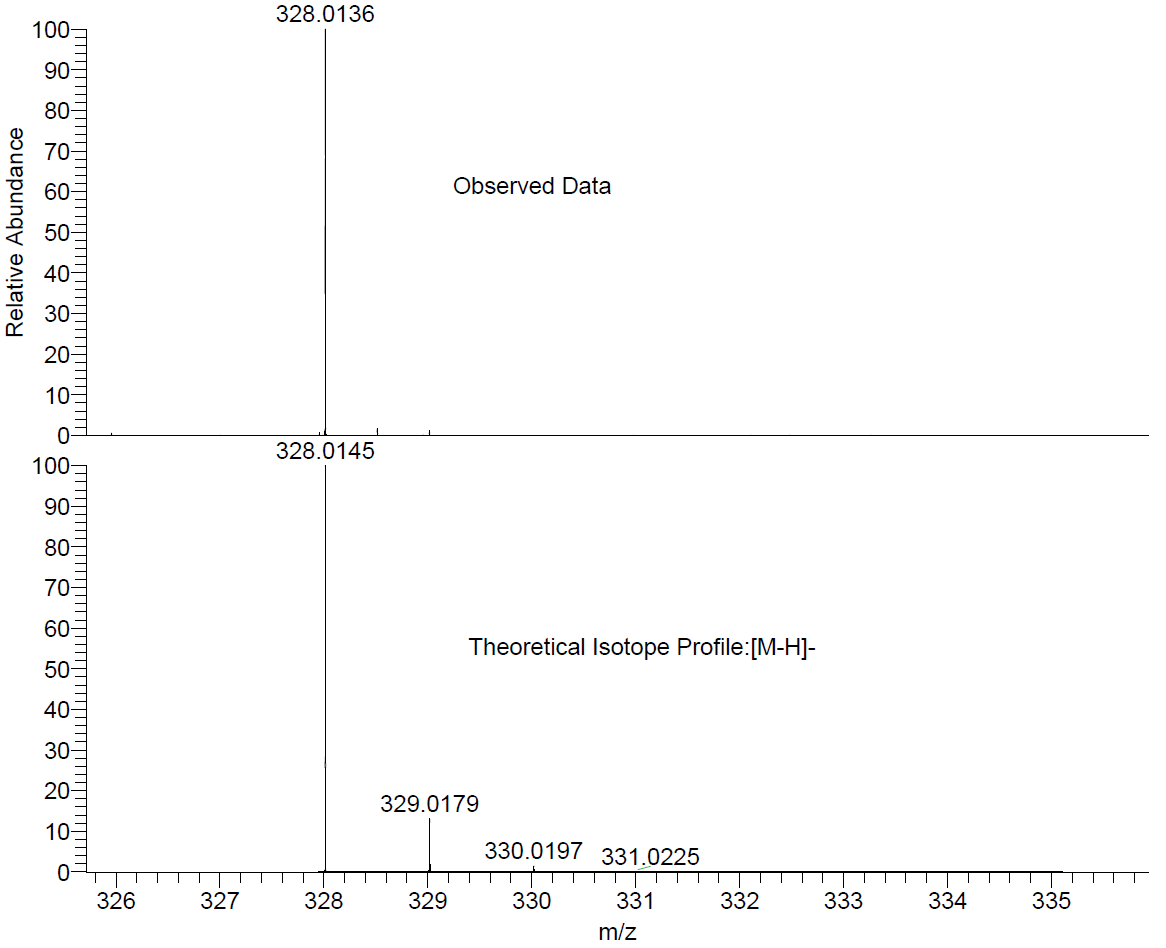


**Figure S19**: Full HSQC NMR spectrum for H4BPA.

As with the ester precursor, mass spectral analysis was carried out to confirm the results of the NMR analysis. Looking at Figure S20, the base peak is clear at a m/z of 163.50, which corresponds to the [M+2H]2+ species. Figure S21 takes a closer look at this peak, showing a close match to the theoretical isotope profile. A second peak is located with a m/z of 328.01, and corresponds to the [M+H]+ species.



**Figure S20**: Full Mass spectrum for H4BPA.

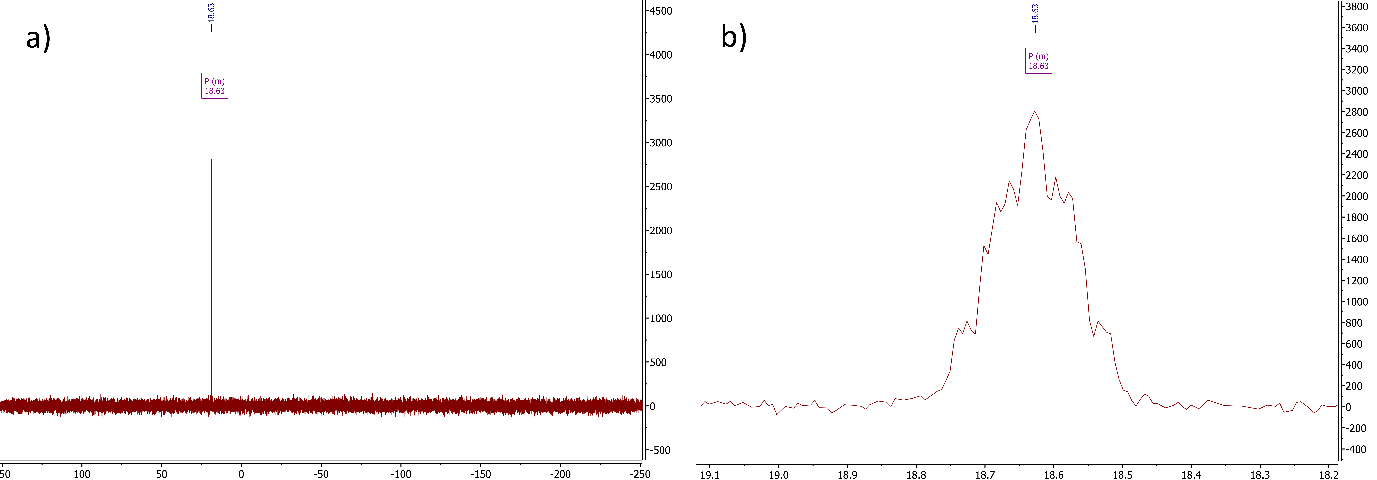
****

**Figure S21**: Mass spectrum for H4BPA showing the observed data and the theoretical isotope profile for the [M-H]- species.

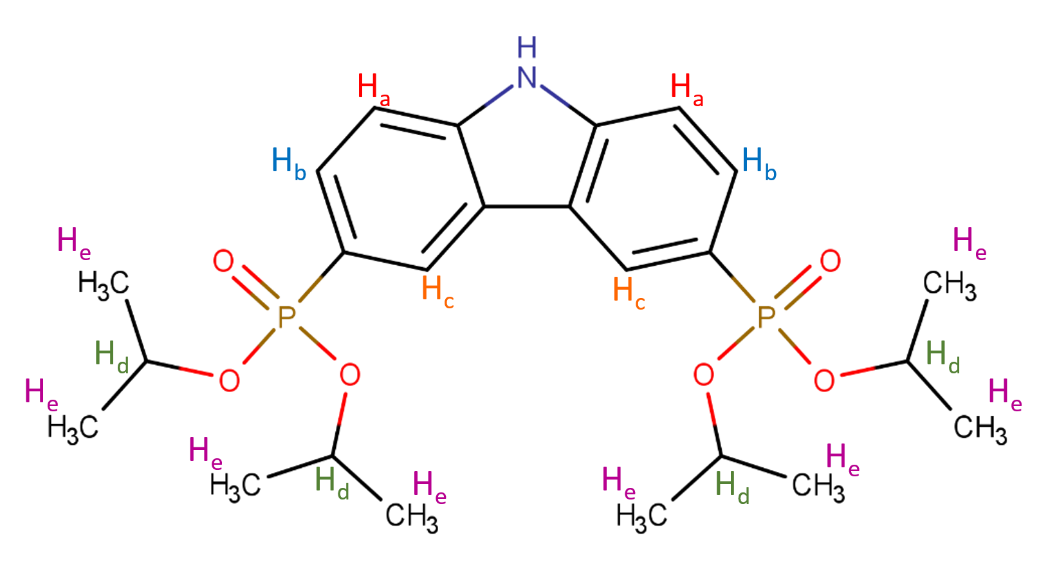
**3,6-bis(diisopropylphosphono)-9H-carbazole (iPr4DPC) (2A)**

The second synthetic target, building upon the success of the first linker, was 3,6-bis(diisopropylphosphono)-9H-carbazole (iPr4DPC) (2A). As with the previous linker, this was also a simple and reproducible procedure from which relatively high yields could be obtained. As with the amine linker, the crude reaction product was most often obtained as a dark brown treacle-like mixture, containing the target product alongside the monosubstituted side-product, some starting material, and some triisopropyl phosphite. Washing this in hexane then yielded a pinky/off-white solid, which could then be purified by flash chromatography. The yields here were often higher on average than for the previous linker, with synthetic yields obtained in the range of 82-97%. The solvent used for NMR was CDCl3.

The 31P NMR spectrum, as shown in Figure S22, again reveals a significant signal corresponding to the target product, which shows a less resolved splitting pattern than was observed for iPr4BPA (2A), though it still indicates higher order coupling with both the methine (Hd) and methyl (He) groups on the isopropyl moiety, as well as the aryl protons Ha-c.

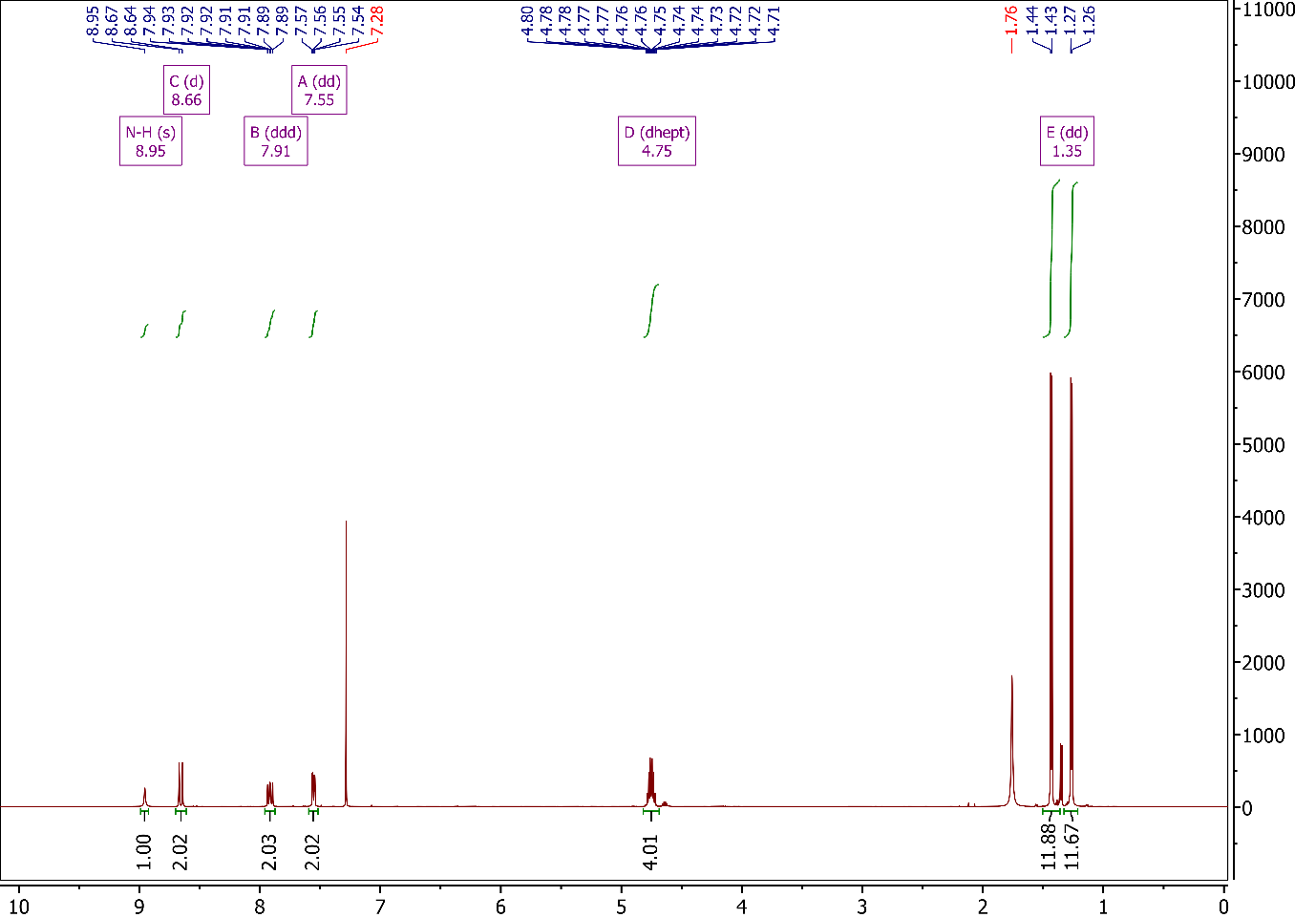


**Figure S22**: a) Full 31P NMR spectrum, and b) zoomed 31P NMR spectrum for iPr4DPC (2A).



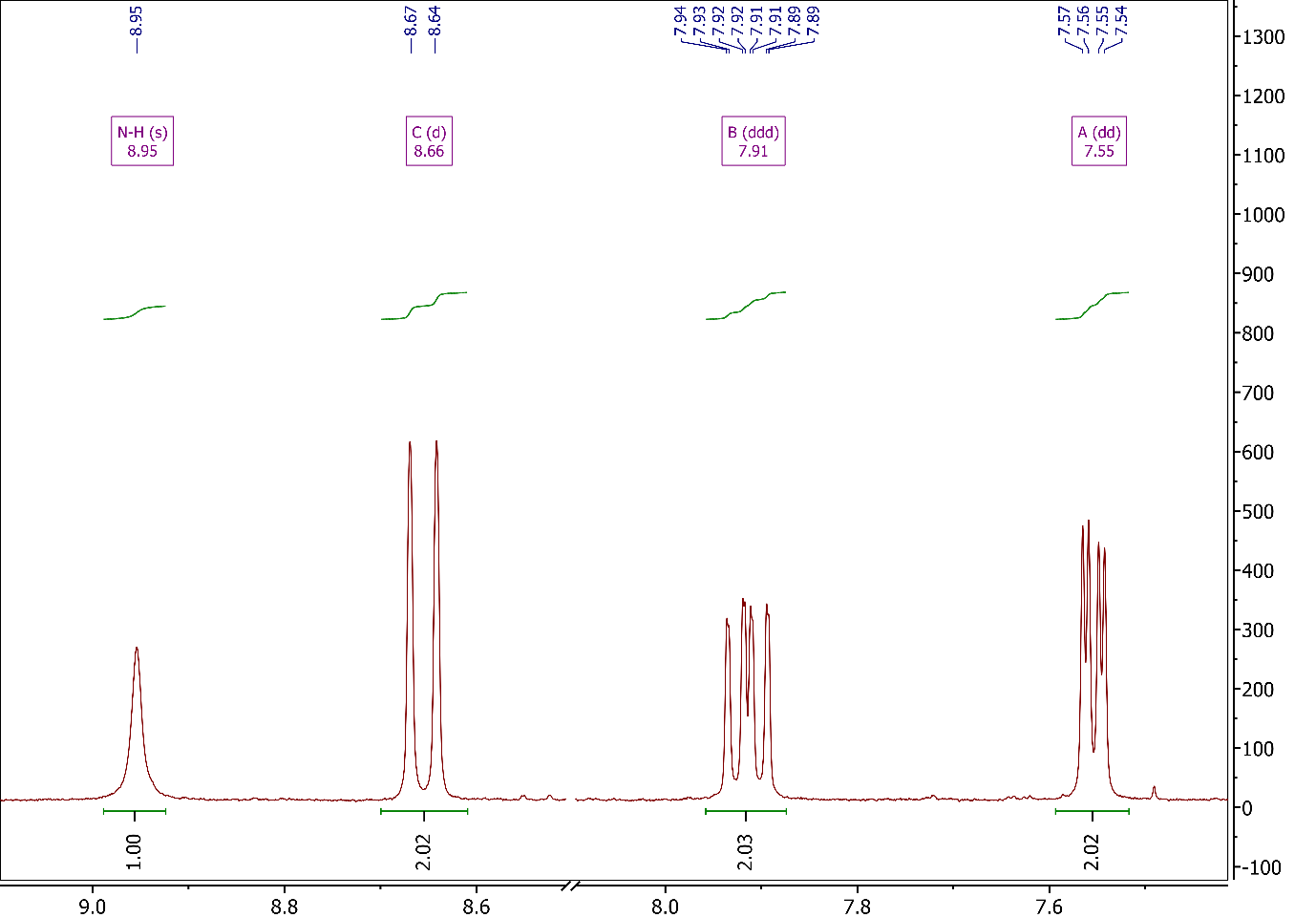
**Figure S23**: Chemical structure for iPr4DPC with proton environments labelled Ha-He.

Moving to the 1H spectrum, in Figure S24 we see that there are some unexpected signals amongst those attributed to the product. Firstly, at δ 1.76 ppm we see a slightly broadened singlet, which is likely to be water, though the chemical shift is slightly higher than you would usually expect for water, δ 1.56 ppm. It is also possible that this belongs to the methyl group of either triisopropyl phosphite [P(OiPr)3] and/or isopropyl bromide [BrCH(CH3)2]. This is also the case for the peaks at δ 4.64 ppm and δ 1.35 ppm. Despite the difficulty in identifying this component, it should be noted that this did not seem to have an effect on the synthesis of the corresponding phosphonic acid and is therefore not too much of a concern.



**Figure S24**: Full 1H NMR spectrum for iPr4DPC (2A).

Moving to Figure S25 to focus on the aromatic region, we see four signals, three of which are to be attributed to the aryl protons and one to the proton bound to nitrogen. The latter of these has been attributed to the broadened signal at δ 8.95 ppm and integrates to 1 proton. Though you might not usually expect to see a signal for these protons, the solvent used, chloroform-*d*, is both apolar and aprotic, so there is no considerable exchange with the N-H proton. What we are left with then are the three signals associated with the aryl protons. The first of these signals (C), a doublet at δ 8.66 ppm, has been attributed to the proton labelled Hc in Figure S21. The splitting here is likely due to higher order coupling with phosphorus, causing the signal to display higher coupling constant than the other signals, which also display splitting patterns that indicate neighbouring protons, which Hc does not have. The next signal (B), at δ 7.91 ppm, has been attributed to the aryl proton labelled Hb in Figure S23. Here we see a doublet of doublets of doublets (ddd), which is the result of the multiple opportunities for coupling. You would initially expect to see just a doublet caused by coupling with the aryl proton Ha, but this splits further due to coupling with both the aryl proton Hc and phosphorus, resulting in the ddd splitting pattern observed. This leaves the final signal (A) in this region, a doublet of doublets (dd) at δ 7.55 ppm, which has been attributed the proton labelled Ha. Coupling the adjacent aryl proton causes the initial splitting into a doublet, while further coupling with phosphorus causes further splitting, resulting in the observed doublet of doublets. This is again supported by the smaller coupling constants observed as you move to smaller δ shifts.



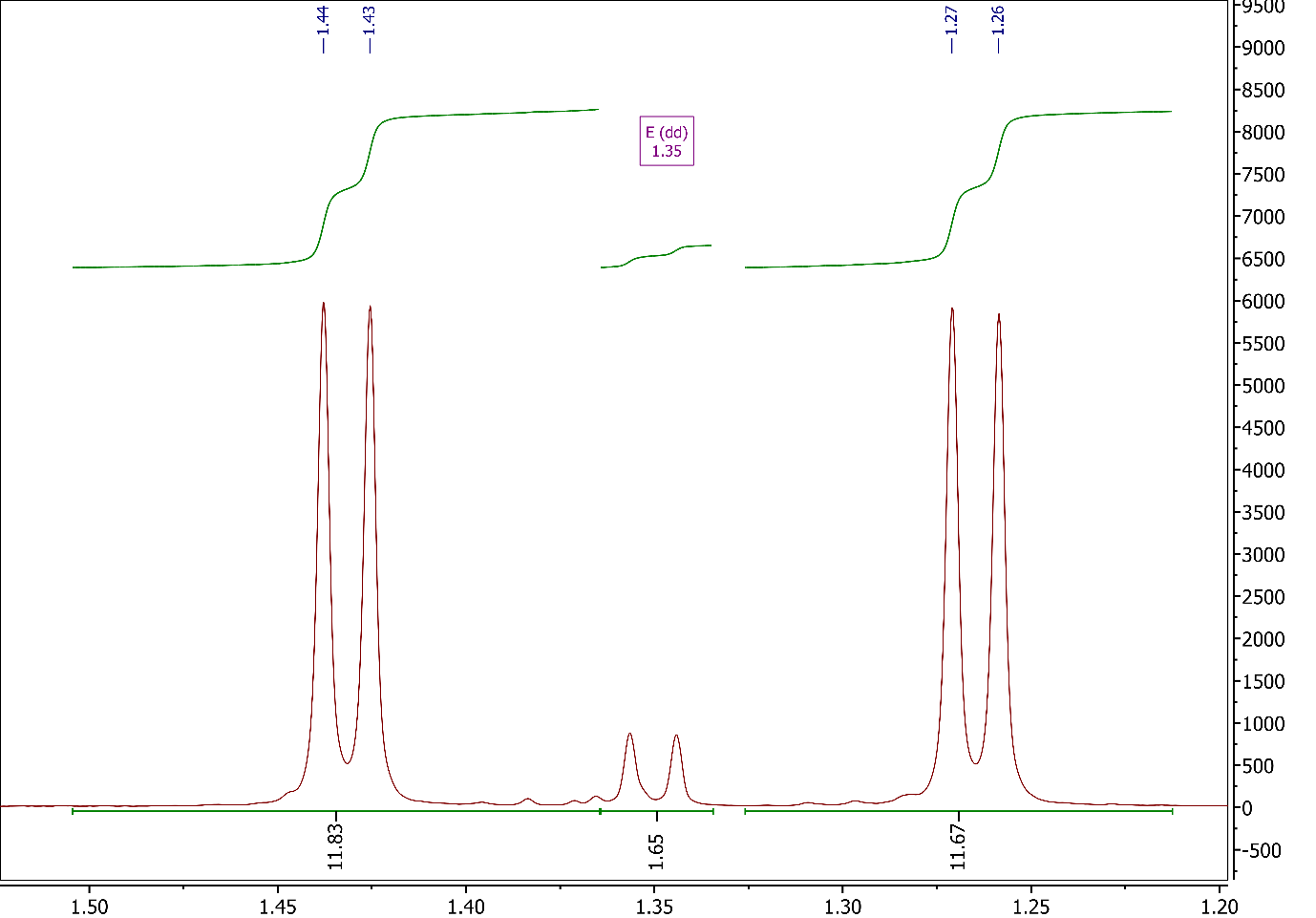
**Figure S25**: Aromatic region of 1H NMR spectrum for iPr4DPC (2A).

As was previously discussed for the methine signal for iPr4BPA, the 1H NMR spectrum for iPr4DPC, as shown in Figure S26, also shows a signal for the methine protons, though at a slightly higher shift of δ 4.75 ppm. In the case of iPr4BPA, it was mentioned that the signal is likely a pair of unresolved septets, which is something that is much clearer here, showing that the original designation as a pair of leaning septets to be quite likely. Most importantly, the signal integrates to the correct number of protons. There is also a similar, though much less intense, signal at ~ δ 4.64 ppm which can be attributed to the methine group of leftover triisopropyl phosphite, which is not unexpected considering the signal at δ 1.35 ppm, which is attributed to the methyl group of triisopropyl phosphite.



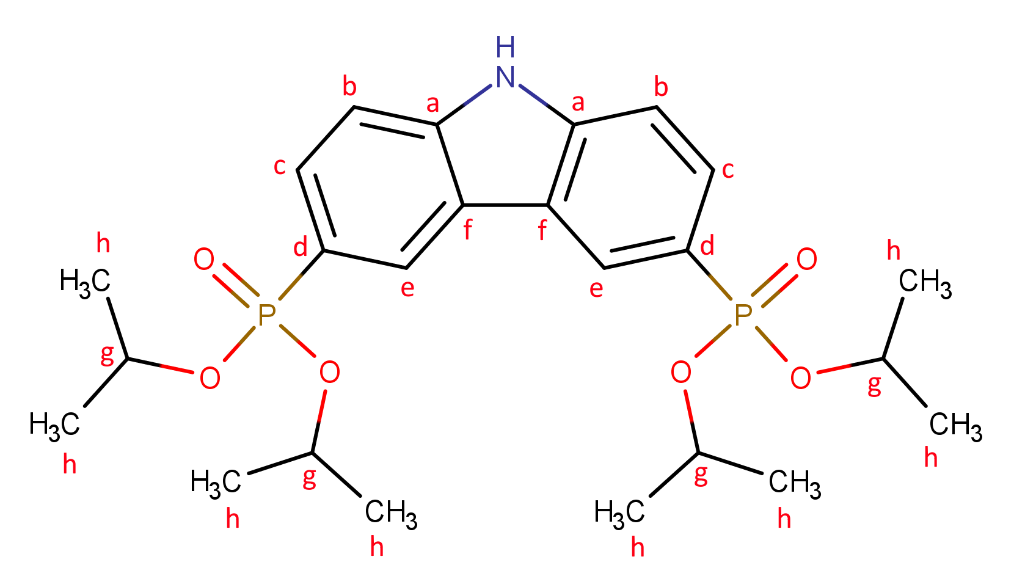
**Figure S26**: Methine region of 1H NMR spectrum for iPr4DPC (2A).

The final signals of concern in this spectrum, which again, was initially thought to be a doublet of doublets, can be seen at δ 1.435 and 1.265 ppm in Figure S27. As was established for iPr4BPA, however, this is in fact more likely to be two doublets which are the result of two chemically distinct methyl groups, and the splitting in each of the doublets arises through coupling with the methine proton. There is also a small impurity present between the methyl peaks at δ 1.35 ppm, which, as mentioned previously, is suspected either isopropyl bromide or triisopropyl phosphite. If quantified using the integration based on the assumption that phosphite is present, this would approximate to a 1:9 ratio of phosphite to iPr4DPC.

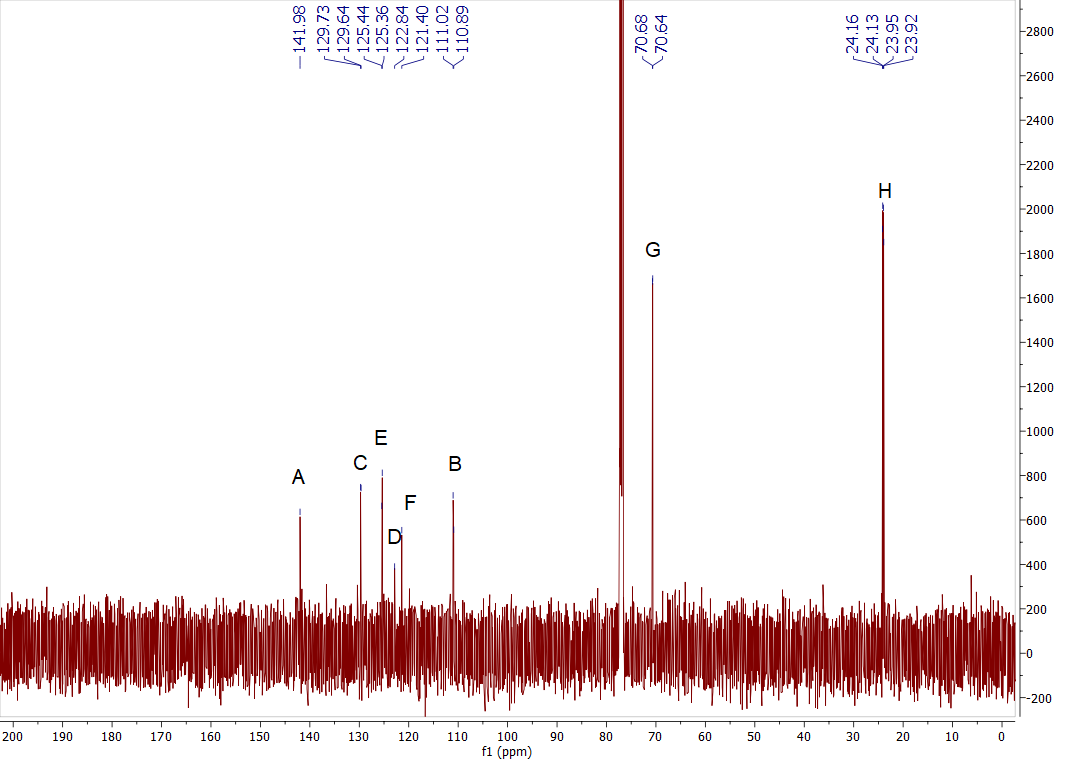


**Figure S27**: Methyl region of 1H NMR spectrum for iPr4DPC (2A).

Looking to the 13C NMR spectrum shown in Figure S29 for iPr4DPC, we can immediately see it is slightly more complicated than that of iPr4BPA. Starting downfield, the first signal we see is the one at δ 141.98 ppm, which has been assigned to the carbons bonded to nitrogen, labelled A. Next we see two signals, labelled C (δ 129.69 ppm) and E (δ 125.40 ppm), have been assigned to the aromatic carbons bearing protons, labelled accordingly, with the third of these aromatic signals (δ 110.96 ppm) labelled B. Then we have the peak labelled D at δ 122.84 ppm, assigned to the carbon bonded to phosphorus, followed by the peak at δ 121.40 ppm which is assigned to carbon in the β-position, relative to nitrogen on the pyrrole-like ring. The final two peaks of interest labelled G (δ 70.66 ppm) and H (δ 24.04 ppm) are assigned to the carbons on the isopropyl groups.

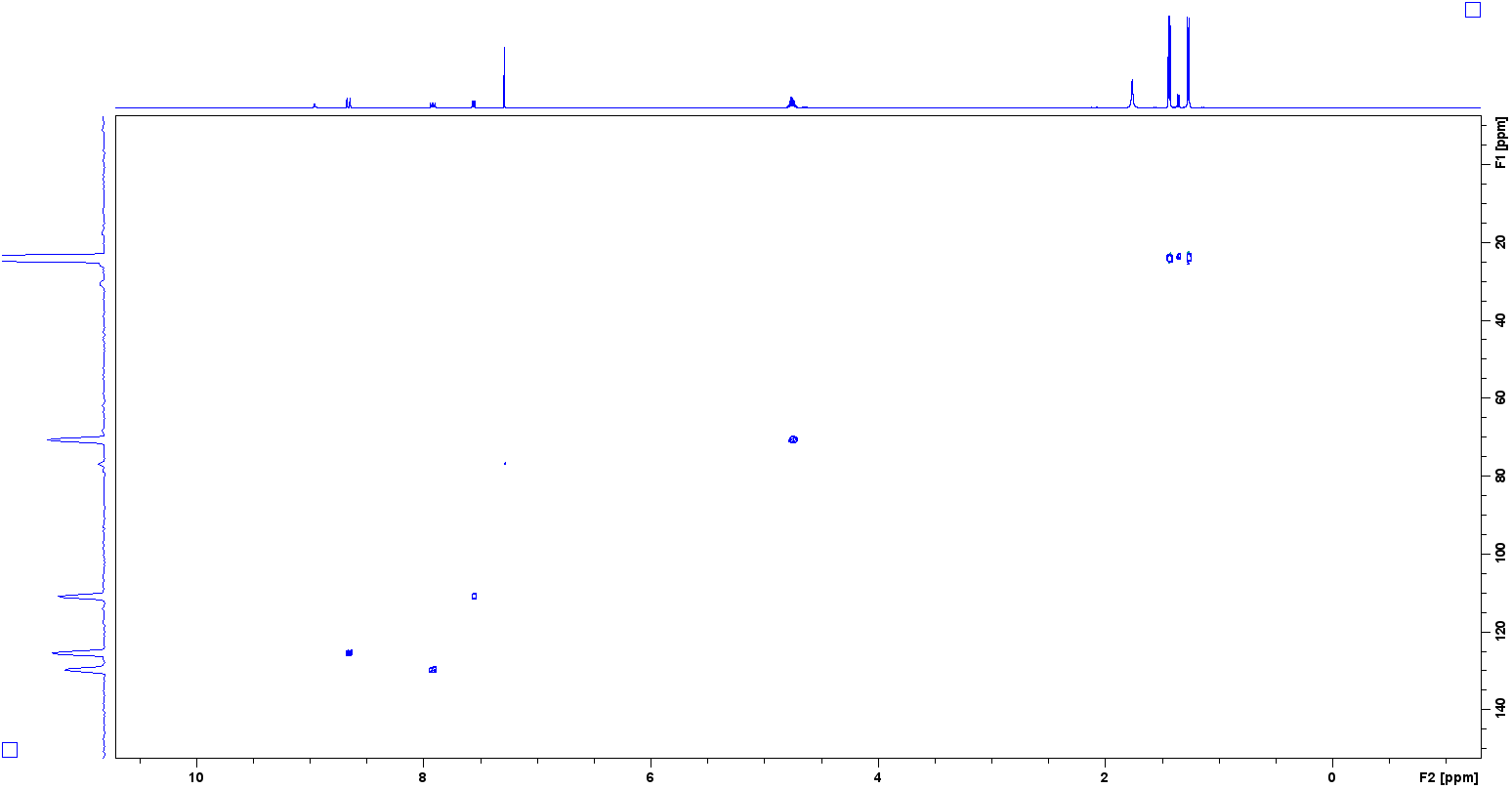


**Figure S28**: Chemical structure of H4BPA with carbon environments labelled Ha-Hf.



**Figure S29**: Full 13C NMR spectrum for iPr4DPC.

Looking to the HSQC NMR spectra for clarification, as shown in Figure 30, we can confirm that the peaks labelled C, E, and B on the 13C NMR spectrum, are indeed correlated with those labelled B, C, and A on the 1H NMR spectrum. We also see the expected correlations for the methine and methyl groups.



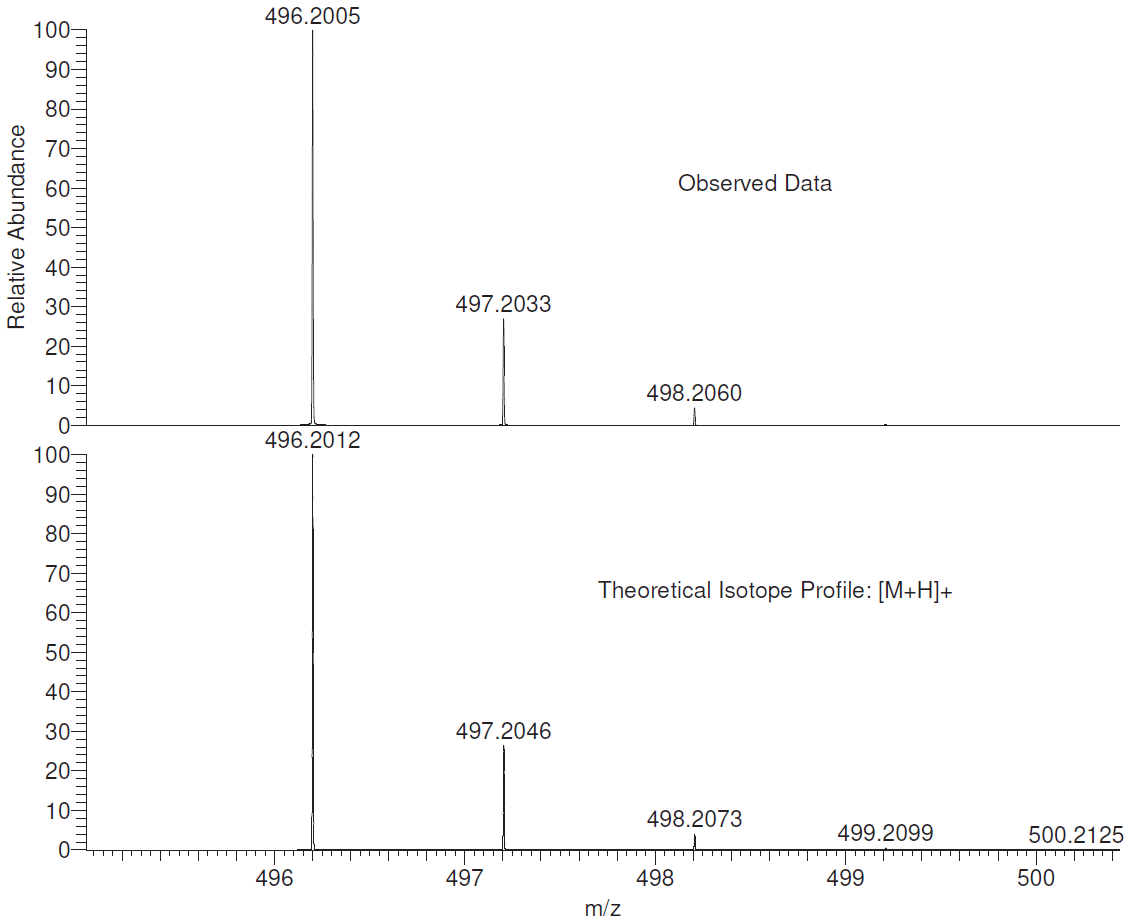
**Figure S30**: Full HSQC NMR spectrum for iPr4DPC.

Moving the mass spectroscopy data, Figure S31 clearly shows the base peak at m/z 496.20 which corresponds to the [M+H]+ species, as confirmed in Figure S32. A second major peak can be seen at 991.39, which corresponds to the [2M+H]+ dimer species.

Chart

Description automatically generated

**Figure S31**: Full Mass spectrum for iPr4DPC.

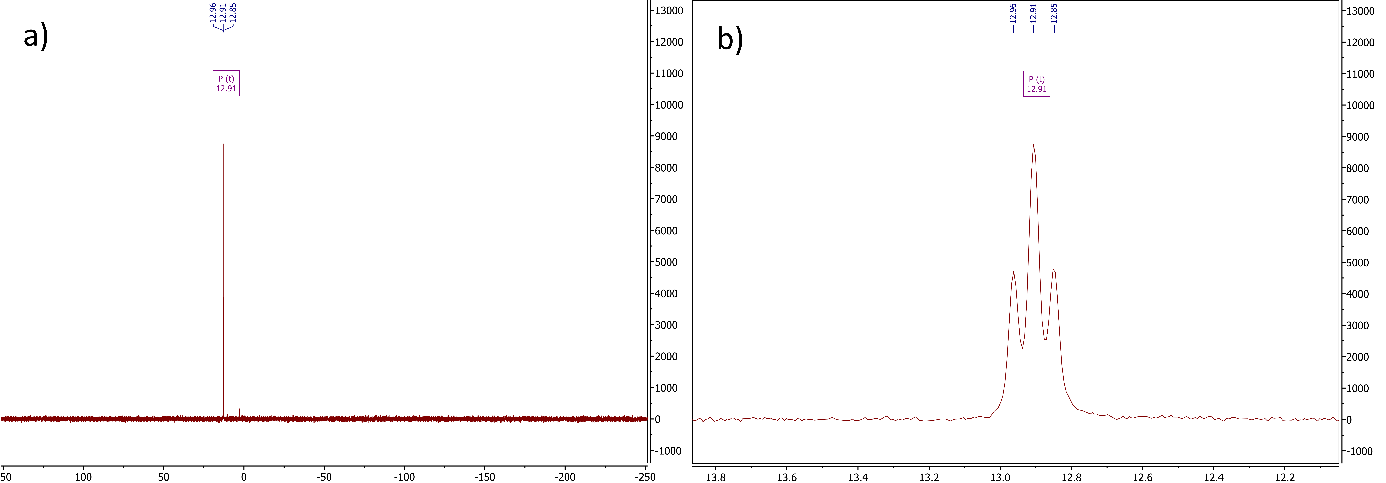


**Figure S32**: Mass spectrum for iPr4DPC showing the observed data and the theoretical isotope profile for the [M+H]+ species.

**3,6-diphosphono-9H-carbazole (H4DPC) (2B)**

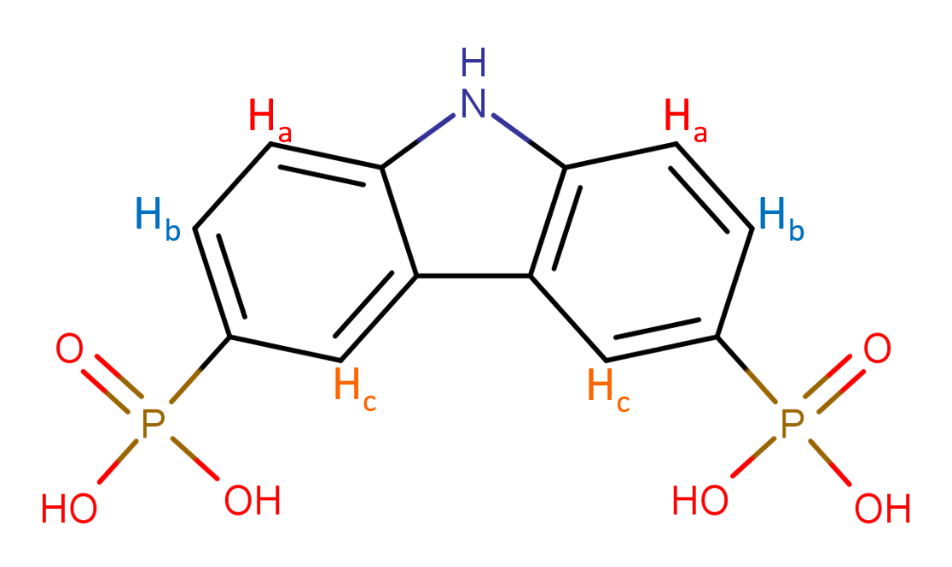
Having successfully obtained iPr4DPC, attention again shifted to obtaining the corresponding phosphonic acid, 3,6-diphosphono-9H-carbazole (H4DPC). As with the amine, this was a relatively simple procedure and the product was obtained in yields between 87-98% when using the silylation and hydrolysis using TMSiBr, since hydrolysis using HCl again seemed to cleave the P-C bond. The solvent used for NMR was a 0.1 M solution of NaOH in D2O.

Looking first at the 31P NMR spectrum, as seen in Figure S33, there is a major signal at δ 12.91 ppm, which corresponds to the target product, H4DPC. There is also a clear change in splitting pattern here, whereby a triplet is observed, though this can be attributed to coupling with the two nonequivalent aryl protons (Hb & Hc). There is also a minor signal upfield at δ 2.6 ppm which may correspond with hydrolysis of leftover triisopropyl phosphite, though this signal is relatively weak in comparison and can reasonably be ignored.

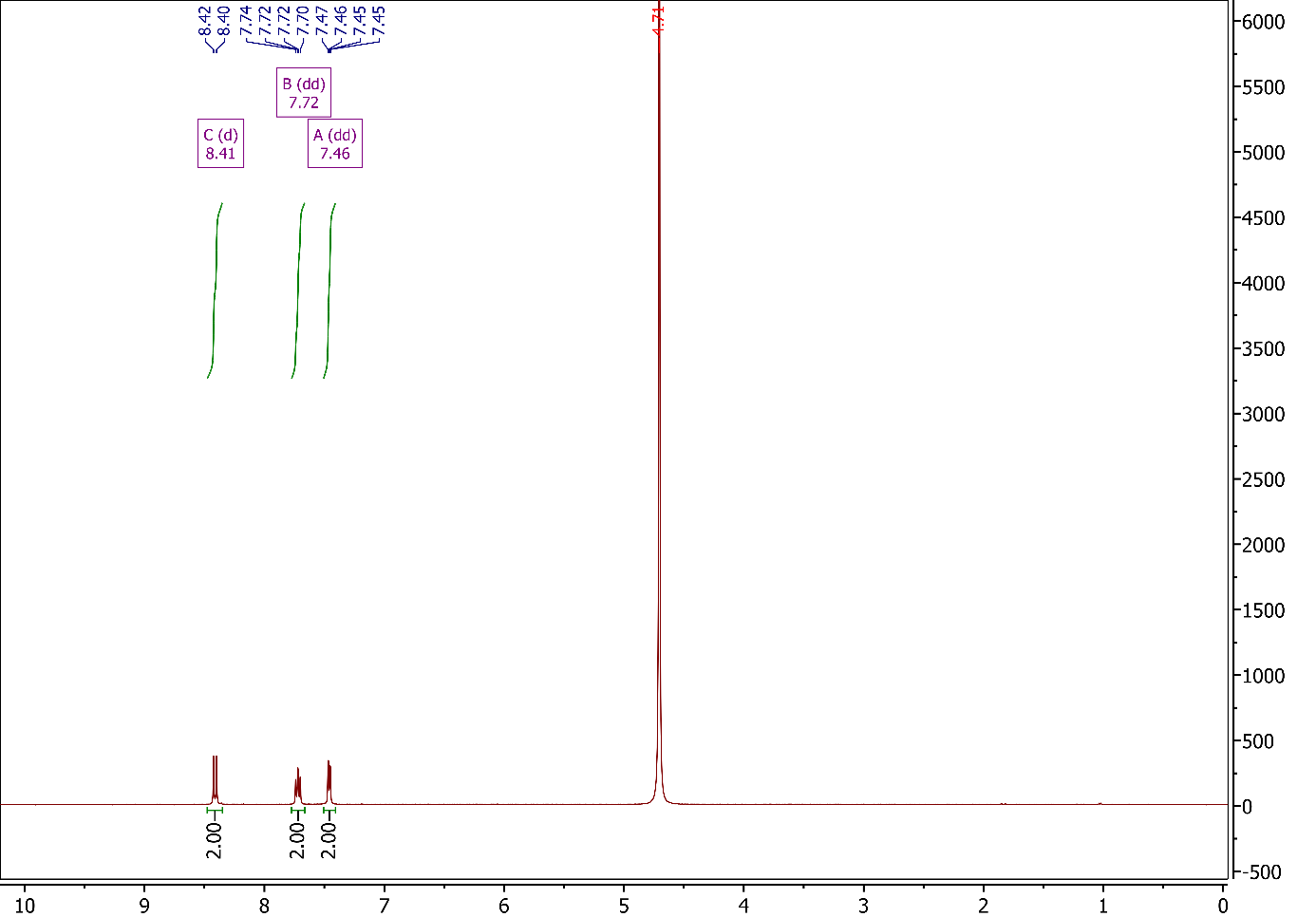


**Figure S33**: a) Full 31P NMR spectrum, and b) zoomed 31P NMR spectrum for H4DPC (2B).

Moving next to the 1H NMR spectrum as shown in Figure S35, it is clear that there are no signals corresponding to the isopropyl groups, thus the hydrolysis has been successful, and there seem to be no signals that indicate other impurities. It is also clear that the signal observed previously for N-H has now disappeared, since a protic solvent is now being used.



**Figure S34**: Chemical structure for H4DPC with proton environments labelled Ha-Hc.



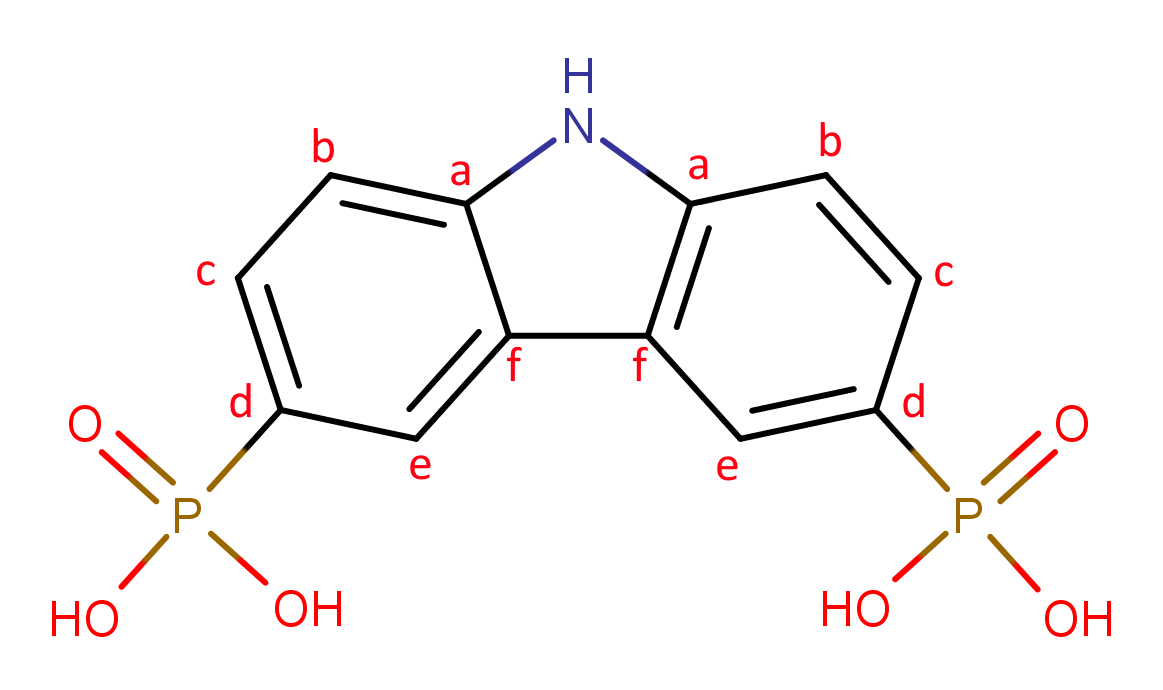
**Figure S35**: Full 1H NMR spectrum for H4DPC (2B).

Taking a closer look at the aromatic region, as seen in Figure S36, the signals present can be assigned exactly as they were for iPr4DPC, whereby protons labelled Ha-c correspond to signals A-C respectively. One clear difference, however, is the less resolved nature of the spectrum, which causes overlap of signals. This is immediately clear for the signal at δ 7.72 ppm, which was previously classified as a doublet of doublets of doublets (ddd), though now seems to be a triplet. This increased intensity of the central peak is caused by the overlap of multiple peaks which, when resolved, would represent the expected ddd splitting pattern as caused by coupling to the other aryl protons and phosphorus.

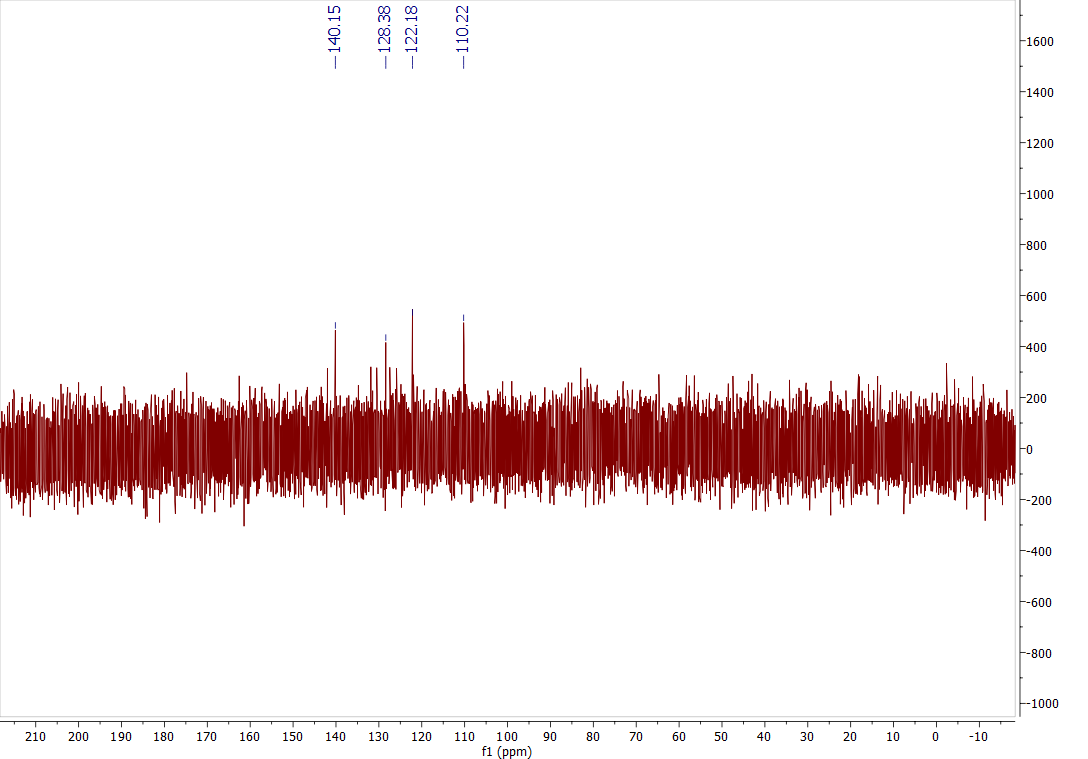


**Figure S36**: Aromatic region of 1H NMR spectrum for H4DPC (2B).

Further characterisation using 13C NMR was less successful, most likely due to the insolubility of H4DPC and the subsequent inability to obtain a spectrum with sufficient peak intensities to accurately assign peaks, at least fully. Looking at Figure S38, it is clear that some of the potential signals in the area of interest are barely above the rather noisy background.

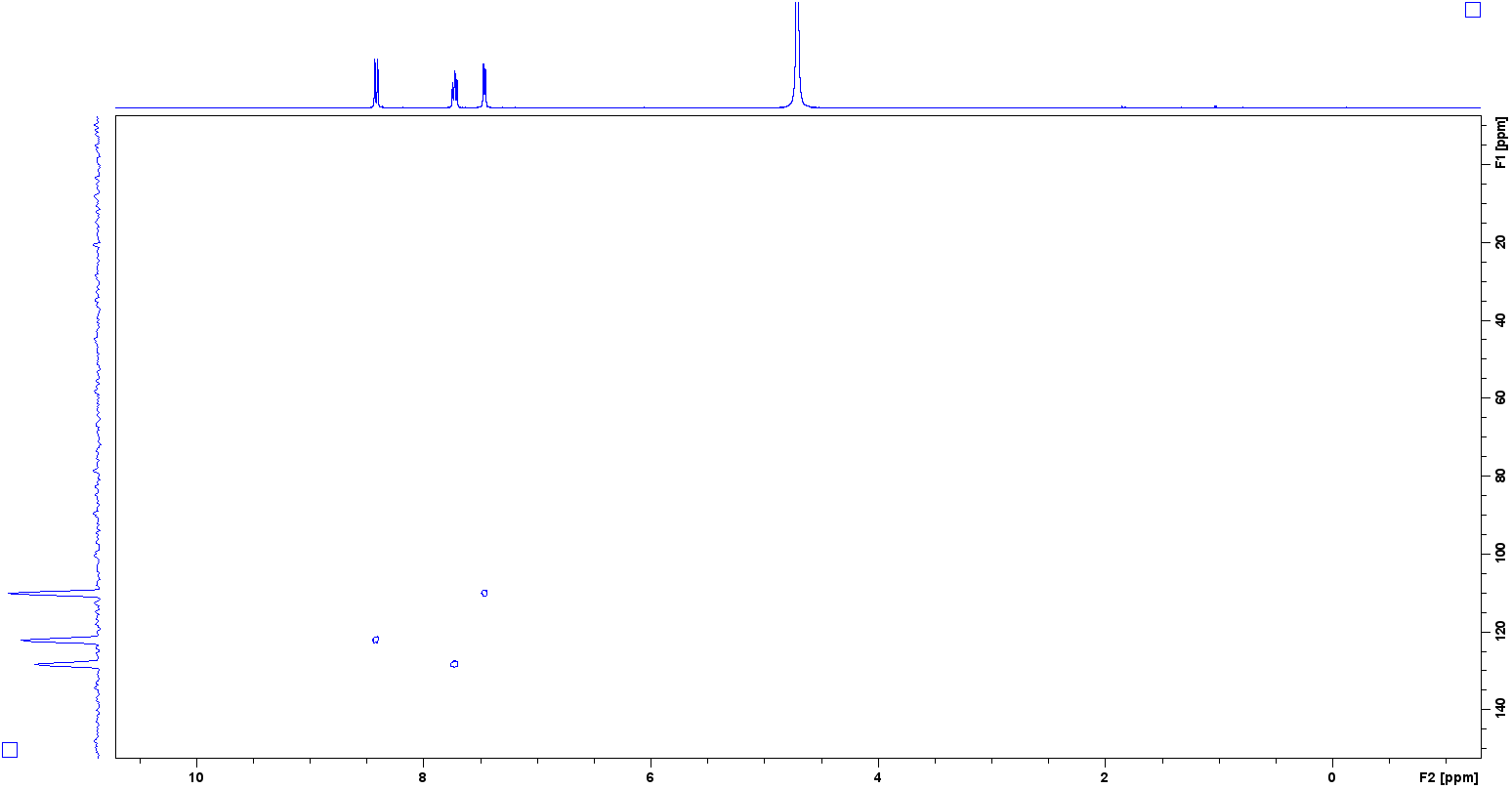


**Figure S37**: Chemical structure of H4DPC with carbon environments labelled Ha-Hf.



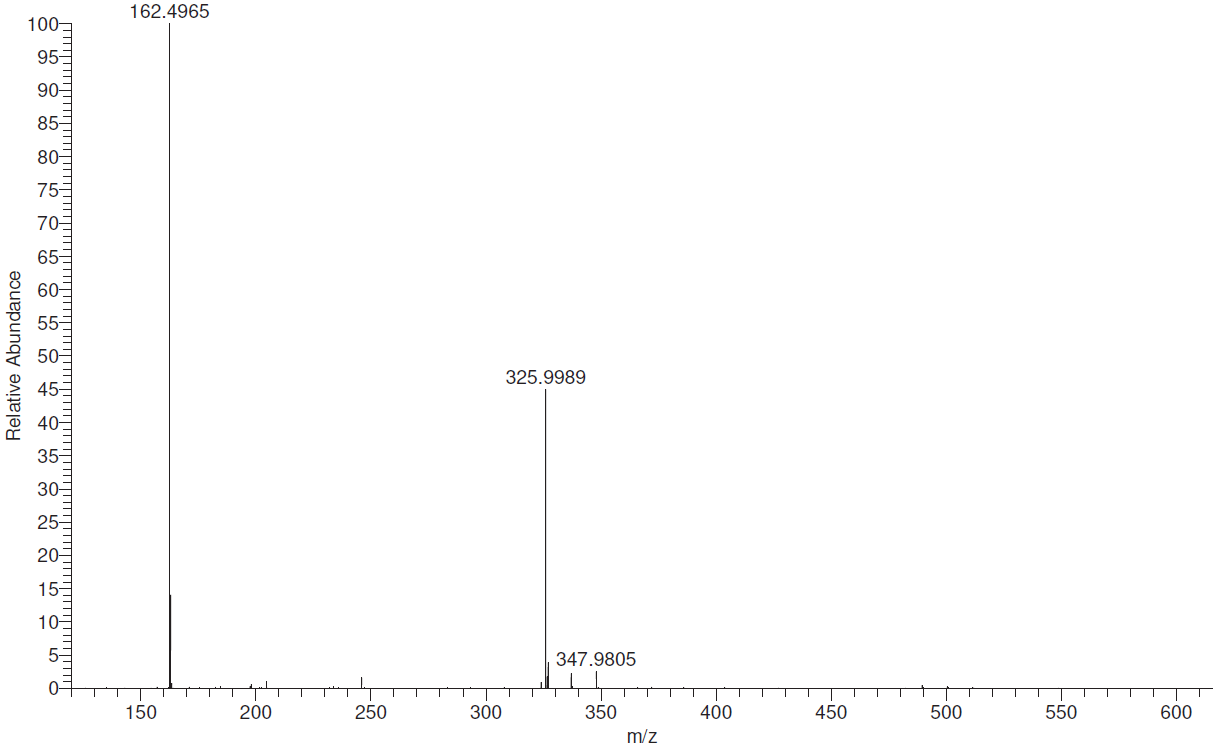
**Figure S38**: Full 13C NMR spectrum for H4DPC.

Referring to the HSQC NMR spectra, as shown in Figure 39, it is possible to at least assign the peaks for carbons with bonded protons. With this in mind, it is clear that the 13C peaks at δ 128.38, δ 122.18, and δ 110.22 ppm can be assigned to the carbons labelled E, C, and B, respectively.

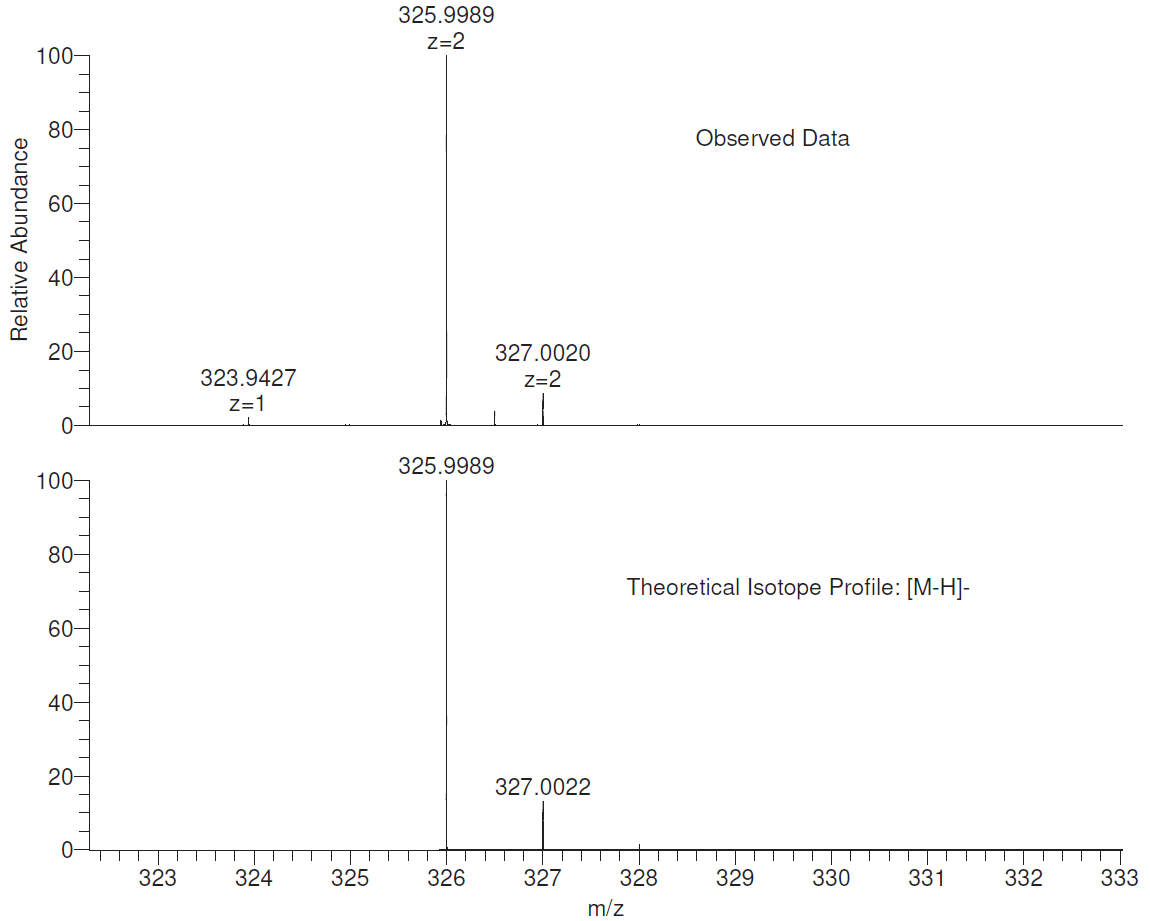


**Figure S39**: Full HSQC NMR spectrum for H4DPC.

Looking at the mass spectral data, Figure S40 clearly shows the base peak at m/z 162.50 corresponding to the [M-2H]2- species and can be seen in more detail in Figure S41. Another peak is shown at m/z 325.99, corresponding to the [M-H]- species.



**Figure S40**: Mass spectrum for iPr4DPC.

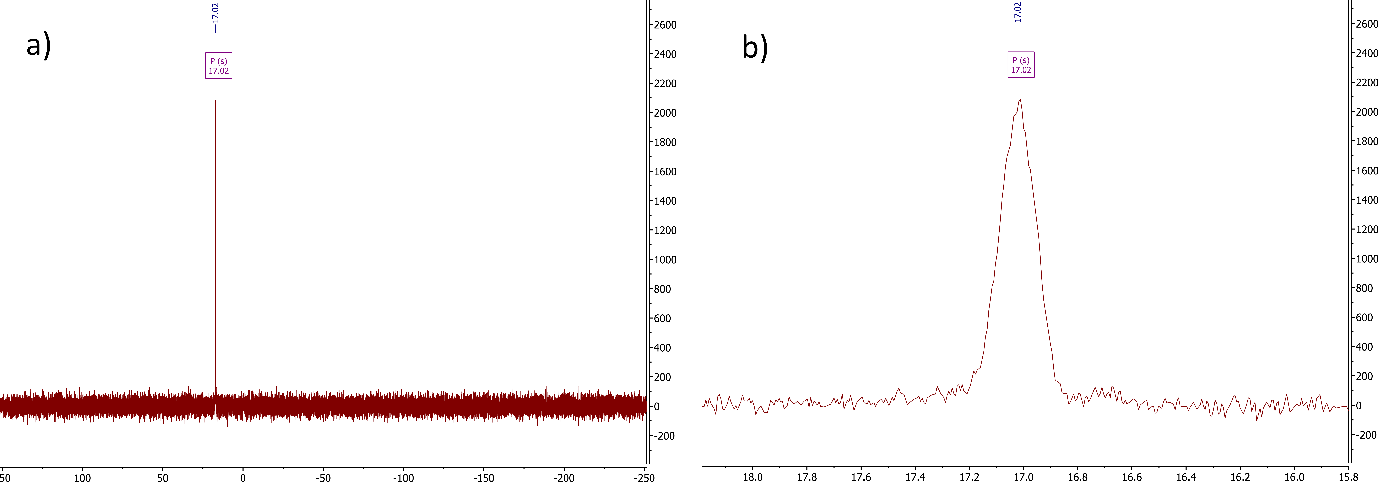


**Figure S41**: Mass spectrum for iPr4DPC showing the observed data and the theoretical isotope profile for the [M-H]- species.

**4-diisopropylphosphono-*N*-(4-diisopropylphosphonophenyl)-*N*-phenylaniline [iPr4DPPA] (3A)**

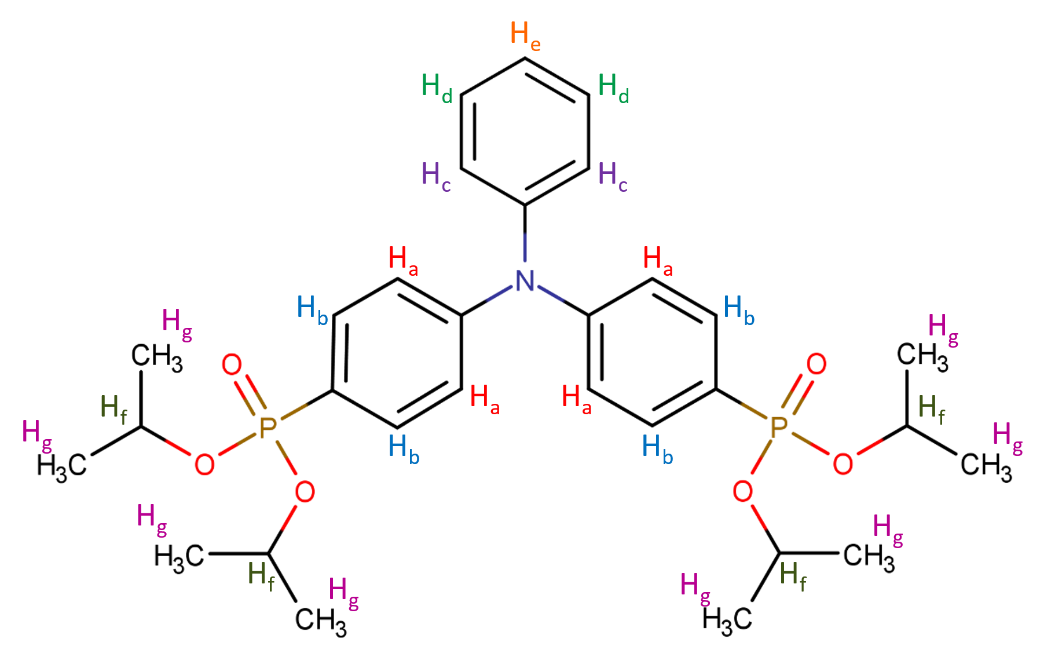
The third synthetic target obtained was 4-diisopropylphosphono-*N*-(4-diisopropyl-phosphonophenyl)-*N*-phenylaniline (iPr4DPPA) (3A). As with the two previous linkers, this synthesis was also straightforward and could be achieved in yields above 80%. The procedure itself was essentially identical, with a simple workup in hexane to obtain the linker. The solvent used for NMR was CDCl3.

Looking first at the 31P NMR spectrum, see Figure S42, we see just a single signal which can be attributed to the target product, iPr4DPPA. Taking a closer look at the signal, it seems that the splitting observed for the two previous linkers is not present, though this might be due to a reduction in the signal to noise ratio.



**Figure S42**: a) Full 31P NMR spectrum, and b) zoomed 31P NMR spectrum for iPr4DPPA (3A).

Moving to the 1H NMR spectrum, as shown in Figure S44, we again see the impurities suspected to be either triisopropyl phosphite and/or isopropyl bromide, as was indicated for iPr4DPC. Beyond this, there are no other unexpected signals, and the only other features of note are the slightly unresolved signals in the aromatic region.



**Figure S43**: Chemical structure for iPr4DPPA with proton environments labelled Ha-Hg.

Diagram, schematic

Description automatically generated

**Figure S44**: Full 1H NMR spectrum for iPr4DPPA (3A).

Zooming in on the aromatic region, as seen in Figure S45, it is possible to identify five signals in total. The first of these signals has a δ-shift of 7.86 ppm, which can be assigned similarly to the iPr4BPA linker, whereby the peak corresponds to protons nearest to phosphorus on the phenyl rings, labelled Hb in Figure S43, again indicated by the large coupling constants. the integration also indicates that it at least belongs to the rings containing phosphonate groups. The next signal, at δ 7.36 ppm, has been assigned to the protons labelled Hc on the phenyl ring with no phosphonate group. The splitting on the spectrum shows a triplet, though you might expect a doublet of doublets for the associated proton, therefore it is likely that this is what we are seeing, simply unresolved. Next, we see the solvent peak at δ 7.28 ppm. Moving upfield, we see a relatively easy to assign the signal at δ 7.20 ppm, and in this case has been attributed to the proton labelled He and is supported by the 1.13 integration value. The final two signals in the aromatic region, while not fully resolved, can be assigned and the splitting can also be identified. The first one, at δ 7.15 ppm, appears to be a simple doublet, and integrates approximately to two, and should therefore be assigned to the remaining protons on the non-phosphonate-containing ring labelled Hd. By process of elimination, this leaves the final aromatic signal at δ 7.12 ppm to be assigned to the protons labelled Ha, though this assignment is supported by the doublet of doublet splitting pattern as well as the integration of approximately four.

Chart

Description automatically generated with medium confidence

**Figure S45**: Aromatic region of 1H NMR spectrum for iPr4DPPA (3A).

Moving to the signal at δ 4.73 ppm, as shown in Figure S46, we see what initially looked like a sextet, though as we have previously explored, is most likely an unresolved doublet of septets, as it would be expected for the methine moiety on the isopropyl groups, labelled Hf in Figure 39. The integration of four is also in favour of this assignment.

Histogram

Description automatically generated with medium confidence

**Figure S46**: Methine region of 1H NMR spectrum for iPr4DPPA (3A).

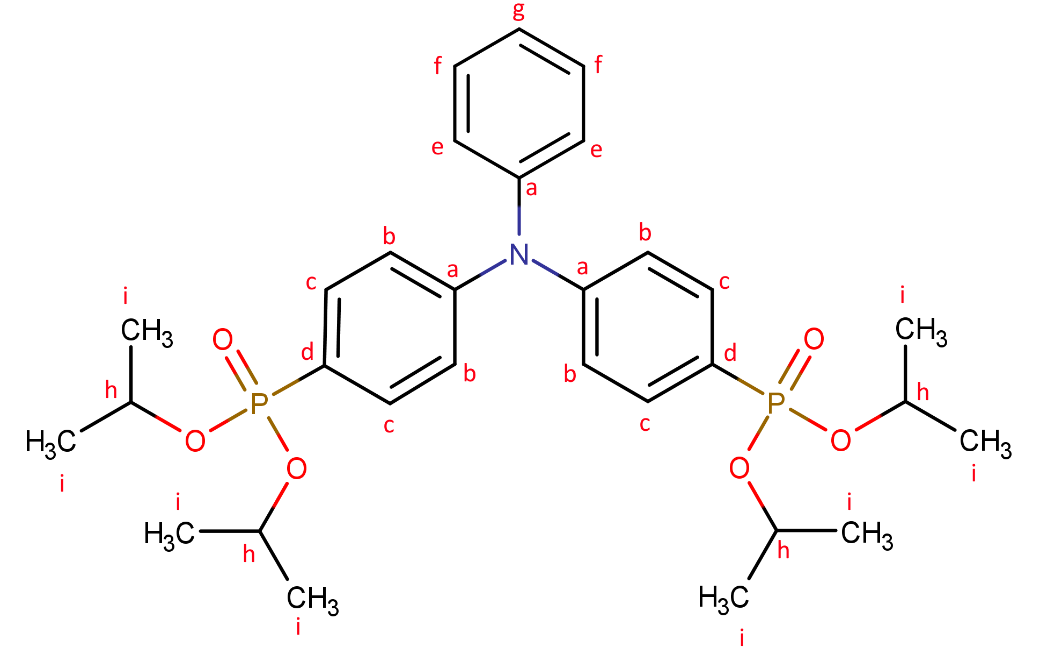
The final signals, as shown in Figure S47, appears at δ 1.34 ppm as a pair of doublets, as was established for iPr4BPA and iPr4DPC, and corresponds to the methyl groups belonging to the isopropyl moiety. The integration for each of the doublets is 12, together totalling 24, as expected.

A picture containing chart

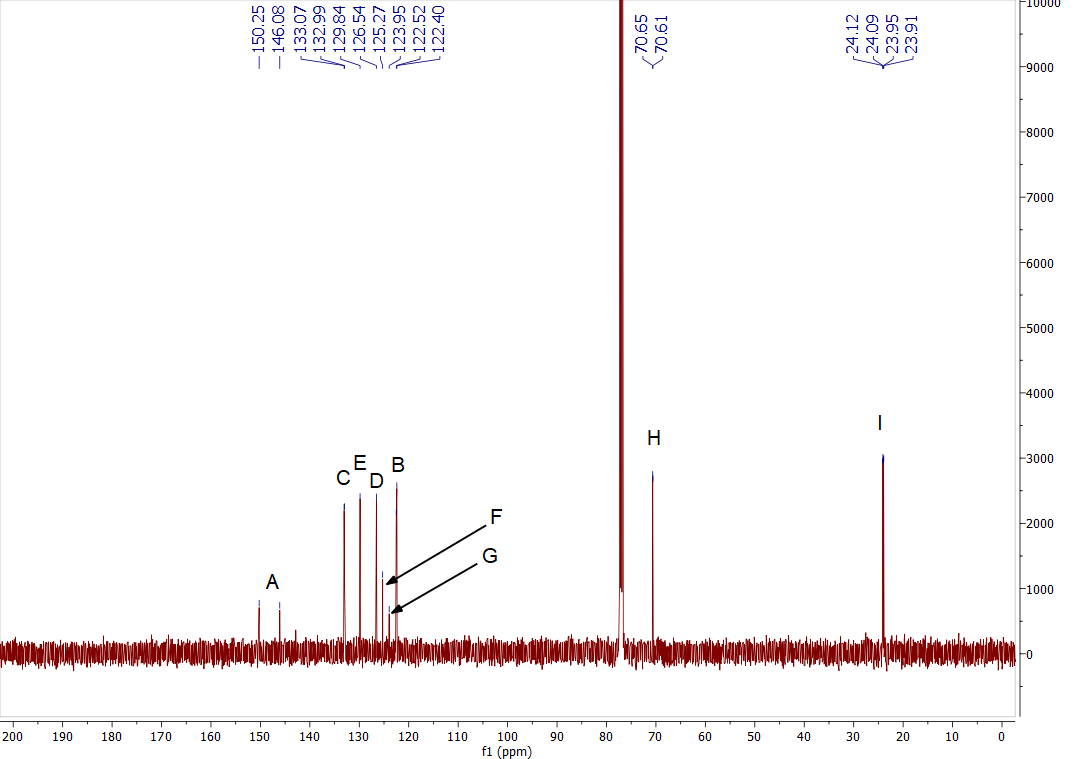
Description automatically generated

**Figure S47**: Methyl region of 1H NMR spectrum for iPr4DPPA (3A).

Moving on to the 13C NMR spectra, as shown in Figure S49, we can immediately assign the methyl (i) and methine (h) carbons to the signals at δ 24.02 ppm and δ 70.63 ppm, respectively. Moving downfield, things start to get slightly more complicated. The first two signals at δ 150.25 ppm and δ 145.08 ppm have both been assigned to the carbon atoms bonded to nitrogen, labelled (a) in Figure S48, with one representing the rings containing phosphonate groups and the other representing the ring with no phosphonate group. Following this, we have the signal at δ 133.03 ppm, which has been assigned to the carbons labelled (c). Next is the signal at δ 129.84 ppm, which has been assigned to the carbons on the non-phosphonate-containing ring, labelled (e). The next signal, at δ 126.54 ppm, has been assigned to the carbons directly bonded to phosphorus, labelled (d). The next two signals at δ 125.27 ppm and δ 123.95 ppm, which are of slightly lower intensity, have been assigned to the carbons on the non-phosphonate-containing ring, labelled (f) and (g), respectively. The final signal on the spectrum, δ 122.46 ppm, has been assigned to the carbons in the β-positions, relative to nitrogen, on the phosphonate containing rings, labelled (b).

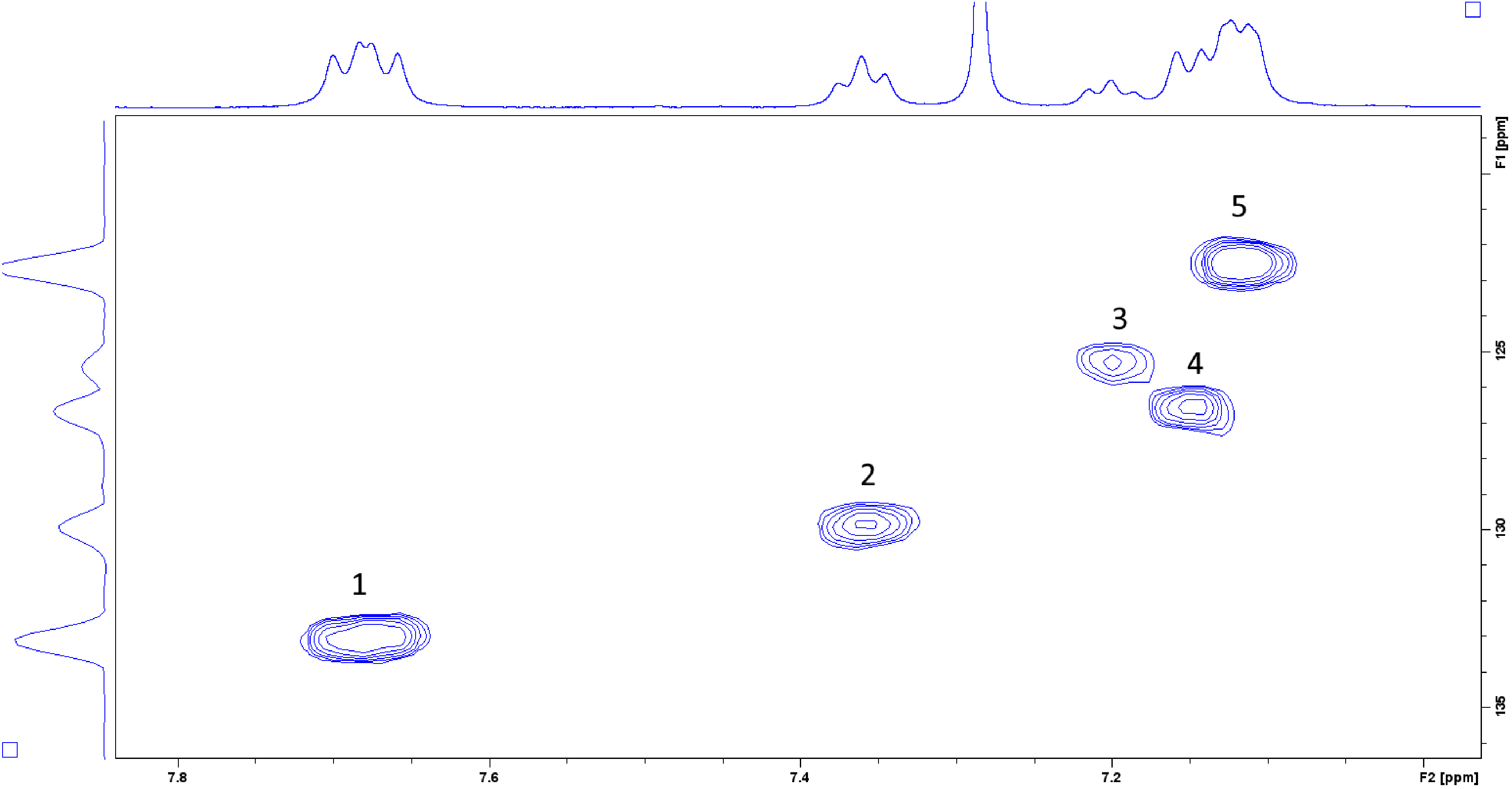


**Figure S48**: Chemical structure for iPr4DPPA with carbon environments labelled Ha-Hi.



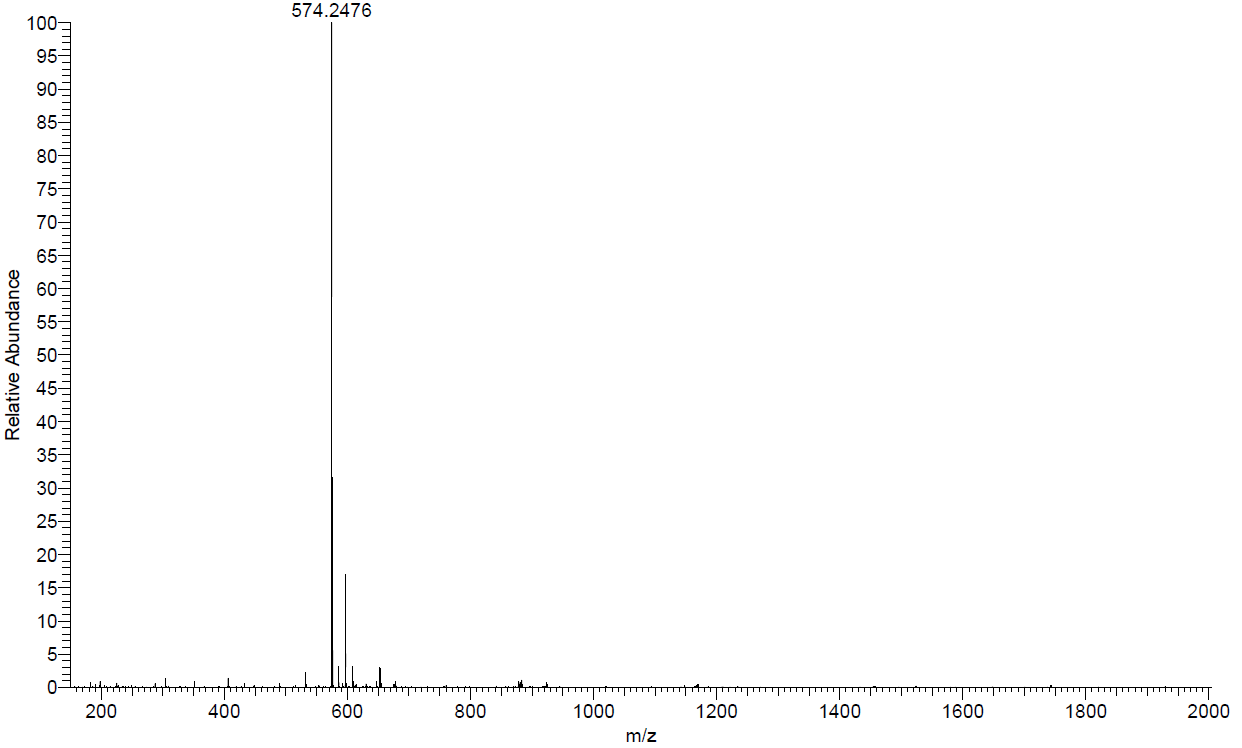
**Figure S49**: Full 13C NMR spectrum for iPr4DPPA (3A).

Lending support for these assignments as well as those for the proton NMR, is the HSQC NMR spectrum shown in Figure S50, in which the signals have been labelled 1-5 for clarity. The signals show the correlations between Cc-Hb (1), Ce-Hc (2), Cg-He (3), Cf-Hd (4), and Cb-Ha (5).

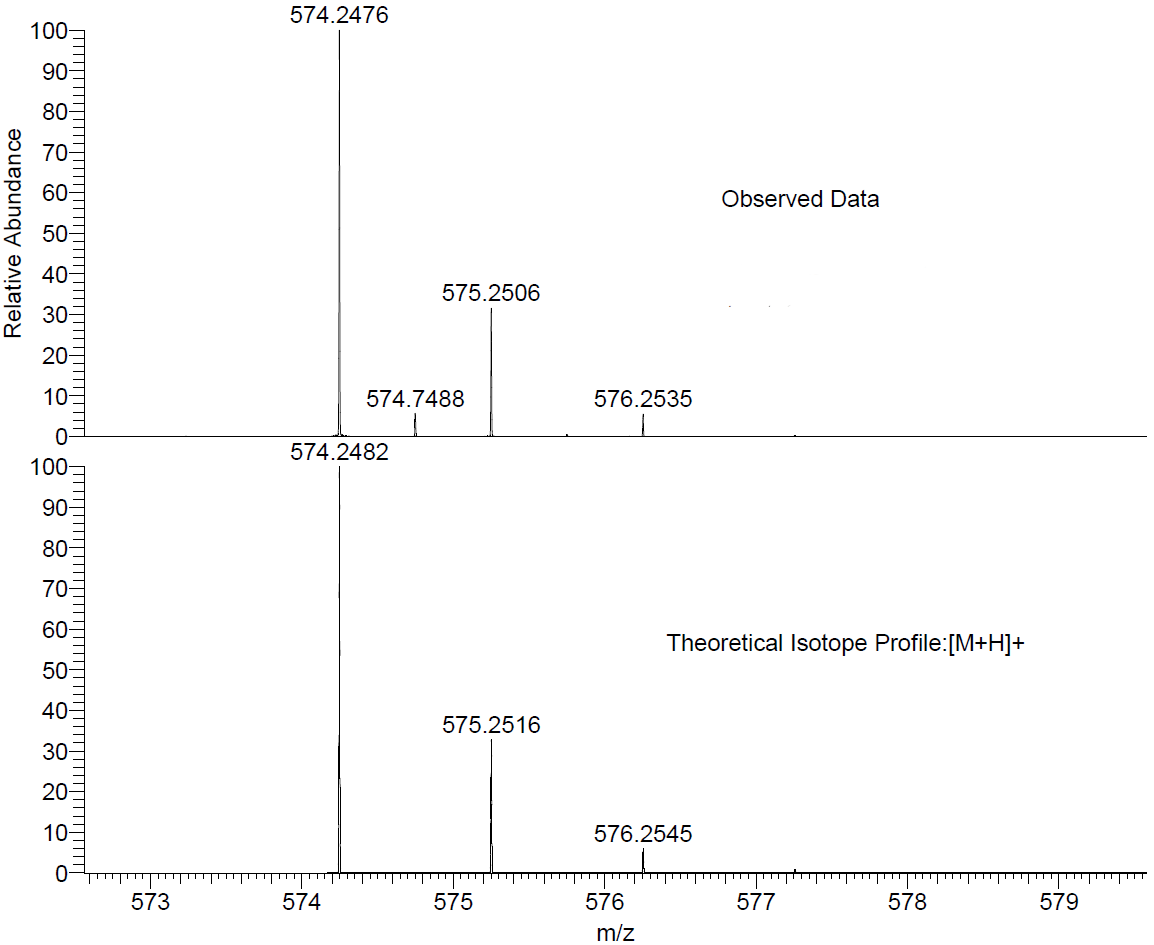


**Figure S50**: Full HSQC NMR spectrum for iPr4DPPA (3A).

Mass spectral analysis was carried out, and Figure S51 clearly shows the base peak at m/z 574/25, which corresponds to the [M+H]+ species. This can be seen in greater detail in Figure S52, which compares the theoretical isotope profile with observed data, shown a good match between the two.



**Figure S51**: Full Mass spectrum for iPr4DPPA.



**Figure S52**: Mass spectrum for iPr4DPPA showing the observed data and the theoretical isotope profile for the [M+H]+ species.

**4-phosphono-*N*-(4-phosphonophenyl)-*N*-phenylaniline [H4DPPA] (3B)**

Although a rather repetitive statement at this point, the hydrolysis of iPr4DPPA to obtain H4DPPA was, as was the case for the other two linkers, a relatively simple procedure, and pretty reasonable yields of between 70-95% could be achieved. Unlike the previous two, this linker was not assessed for its suitability for hydrolysis using HCl, instead opting to go straight for the less harsh procedure involving silylation through the use of TMSiBr. The solvent used for NMR was a 0.1 M solution of NaOH in D2O.

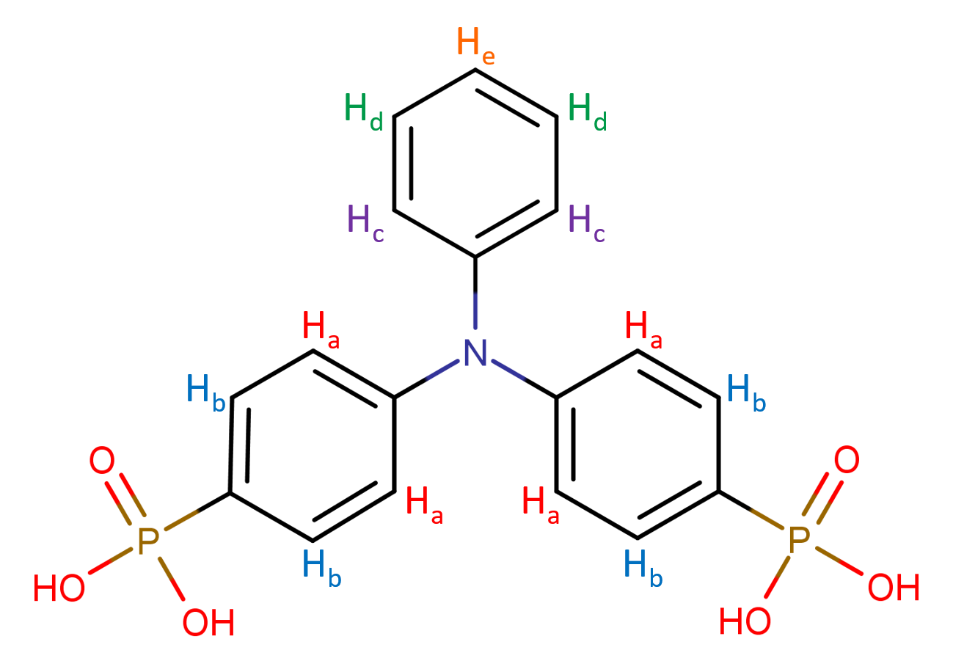
Starting with the 31P NMR spectrum, as shown in Figure S53, we can immediately see the presence of only one signal, indicating relative purity with regards to phosphorus-containing compounds. We can already observe the loss of the isopropyl groups here from the huge change in the splitting pattern. Here, we see only a triplet, likely caused by coupling with the aromatic protons, whereas the pattern for the ester presented a much more complicated and much less resolved pattern due to the amount of coupling taking place.

Chart, line chart

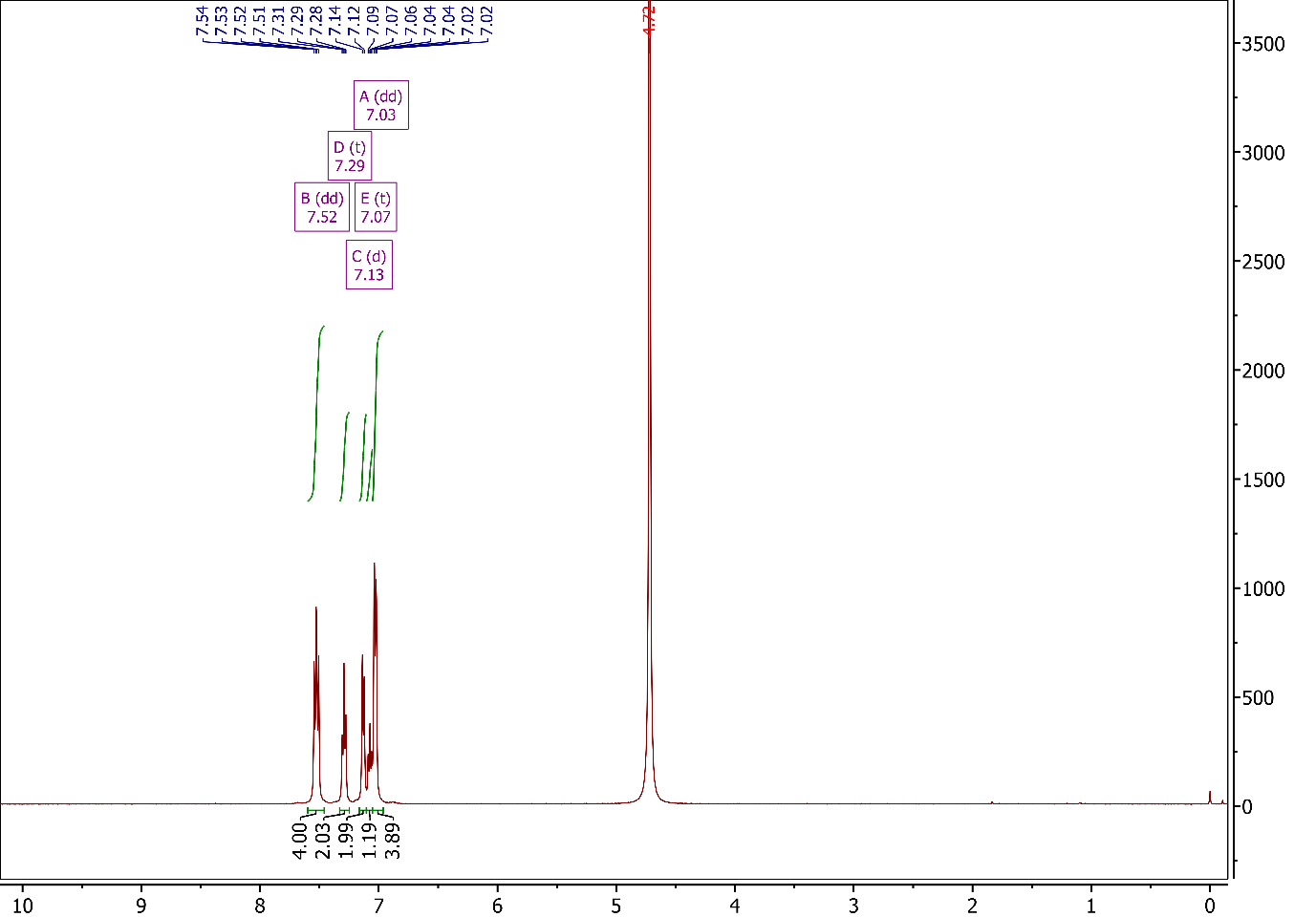
Description automatically generated

**Figure S53**: a) Full 31P NMR spectrum, and b) zoomed 31P NMR spectrum for H4DPPA (3B).

Moving to the 1H NMR spectrum as shown in Figure 55, we can immediately confirm the loss of the isopropyl groups through the lack of methine and methyl signals upfield. Beyond this, the only potential impurity shown at δ 0.0 ppm, which can be attributed to phosphorus acid (H3PO3), though the amount present must be quite small as it is not seen in 31P NMR spectrum and could easily be removed through further washing in small amounts of water or ethanol.

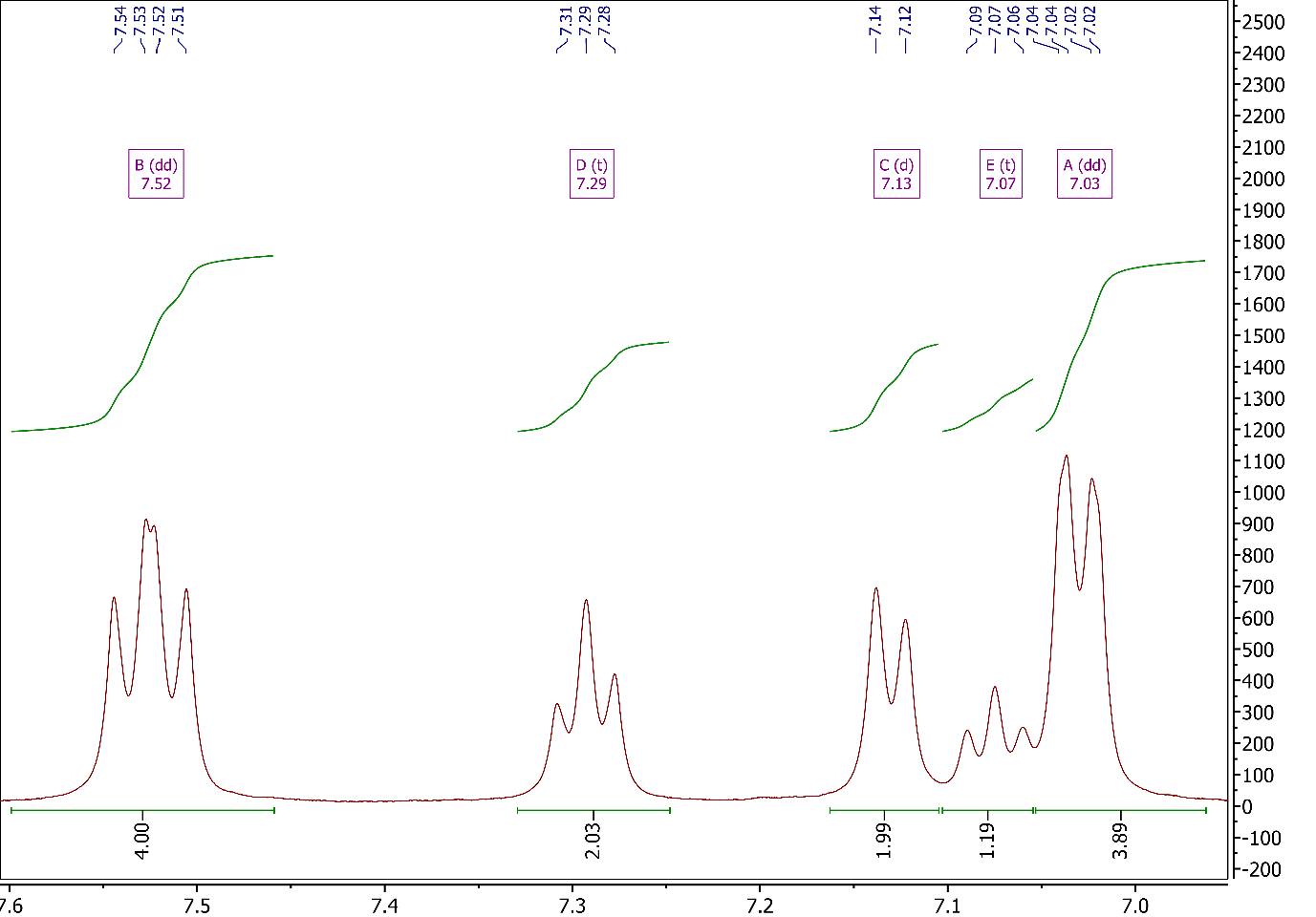


**Figure S54**: Chemical structure for H4DPPA with proton environments labelled Ha-He.



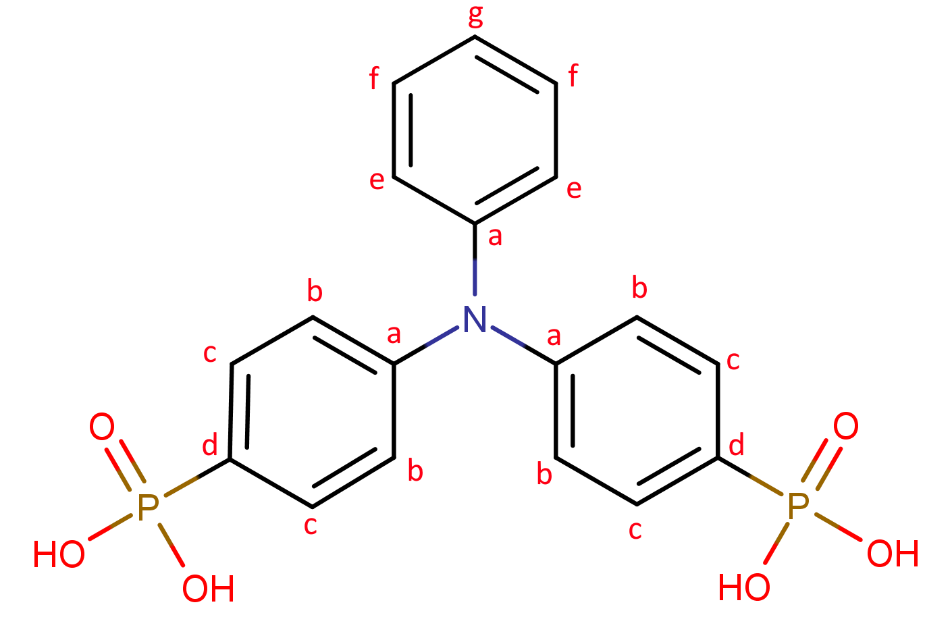
**Figure S55**: Full 1H NMR spectrum for H4DPPA (3B).

Focusing on the aromatic region, as shown in Figure S56, and comparing it to that of the ester in Figure S45, the assignment of the signals is clear, though there is one striking difference, which is that two of the signals, C and E, have switched places on the spectrum. Turning back to the assignment of the signals, as was the case for the ester, the signal at δ 7.52 ppm has been assigned to the protons nearest to phosphorus, labelled Hb. The next signal, at δ 7.29 ppm, has been assigned to the protons labelled Hd on the non-phosphonate-containing ring. The signal at δ 7.13 ppm has been assigned to the protons nearest nitrogen on the non-phosphonate-containing ring, labelled Hc. The next signal, and generally the easiest to assign due to the integration, has been assigned to the proton labelled He on the non-phosphonate-containing ring. This leaves the final signal at δ 7.13 ppm, which through process of elimination, supported by splitting pattern and integration, be attributed to the protons labelled Ha.

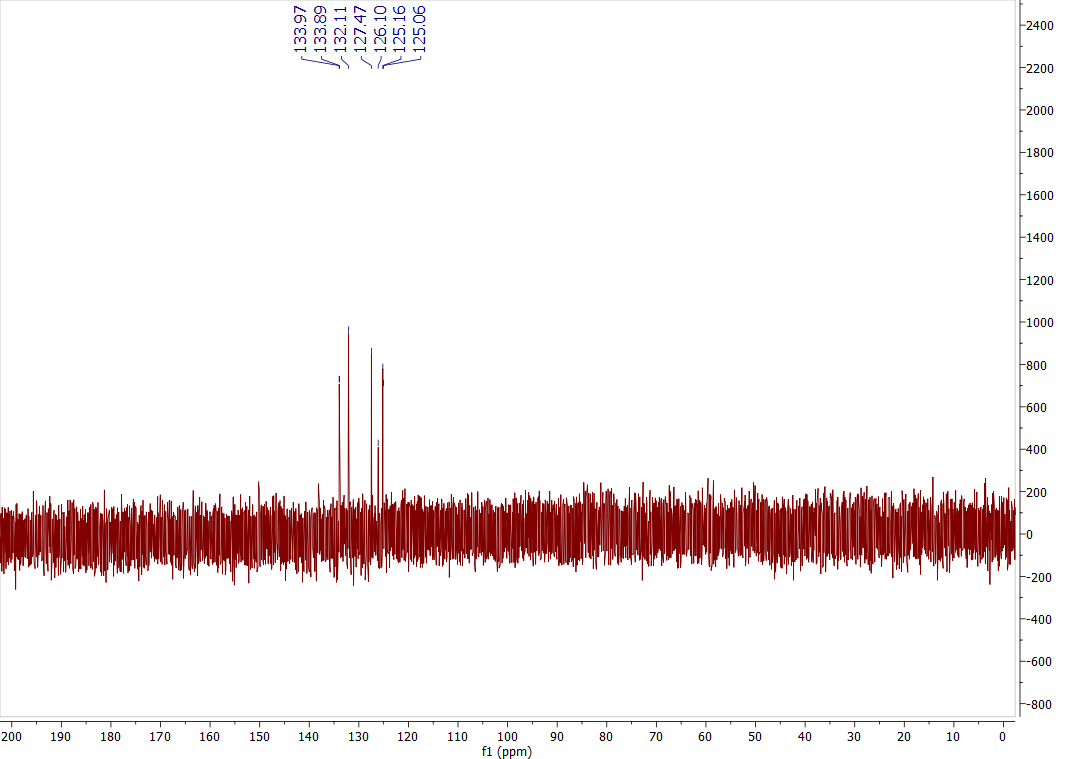


**Figure S56**: Full 1H NMR spectrum for H4DPPA (3B).

As was the case for the ester, 13C NMR, as shown in Figure S58, has proven insufficient in aiding full characterisation of the linker. This is mostly due to the insolubility of the product, which has resulted in a poor signal-to-noise ratio. With this in mind, HSQC should help with assigning some of the 13C NMR peaks. It should be noted here that the 13C NMR spectrum clearly shows the lack of isopropyl group peaks.

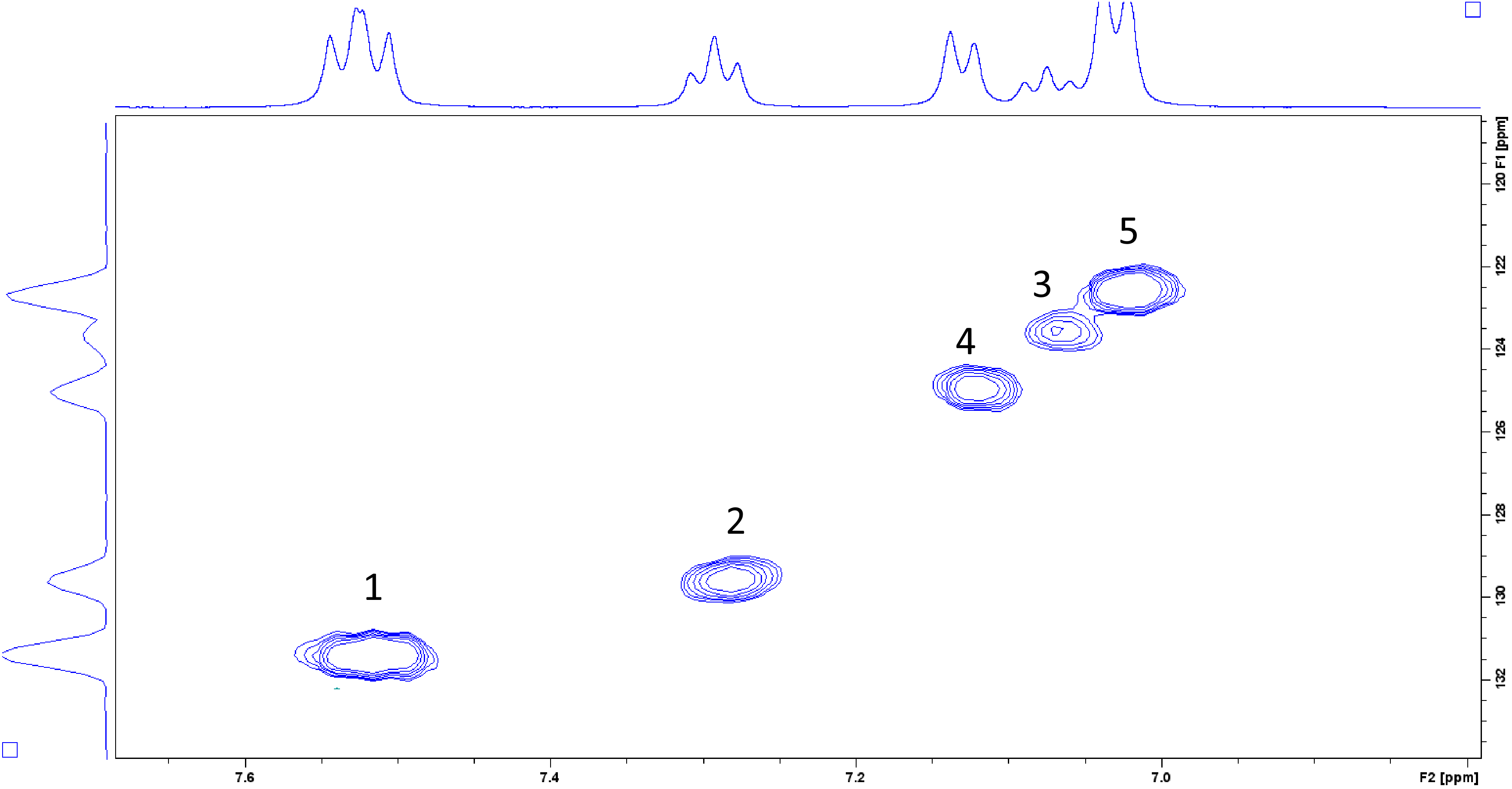


**Figure S57**: Chemical structure for H4DPPA with carbon environments labelled Ha-Hg.



**Figure S58**: Full 13C NMR spectrum for H4DPPA (3B).

Looking at the HSQC NMR spectra in Figure S59, we can indeed see the switching of signals 3 and 4 when comparing with that of the ester in Figure S50. Beyond that, the assignments are identical, with the signals showing correlations between Cc-Hb (1), Ce-Hc (2), Cg-He (3), Cf-Hd (4), and Cb-Ha (5).



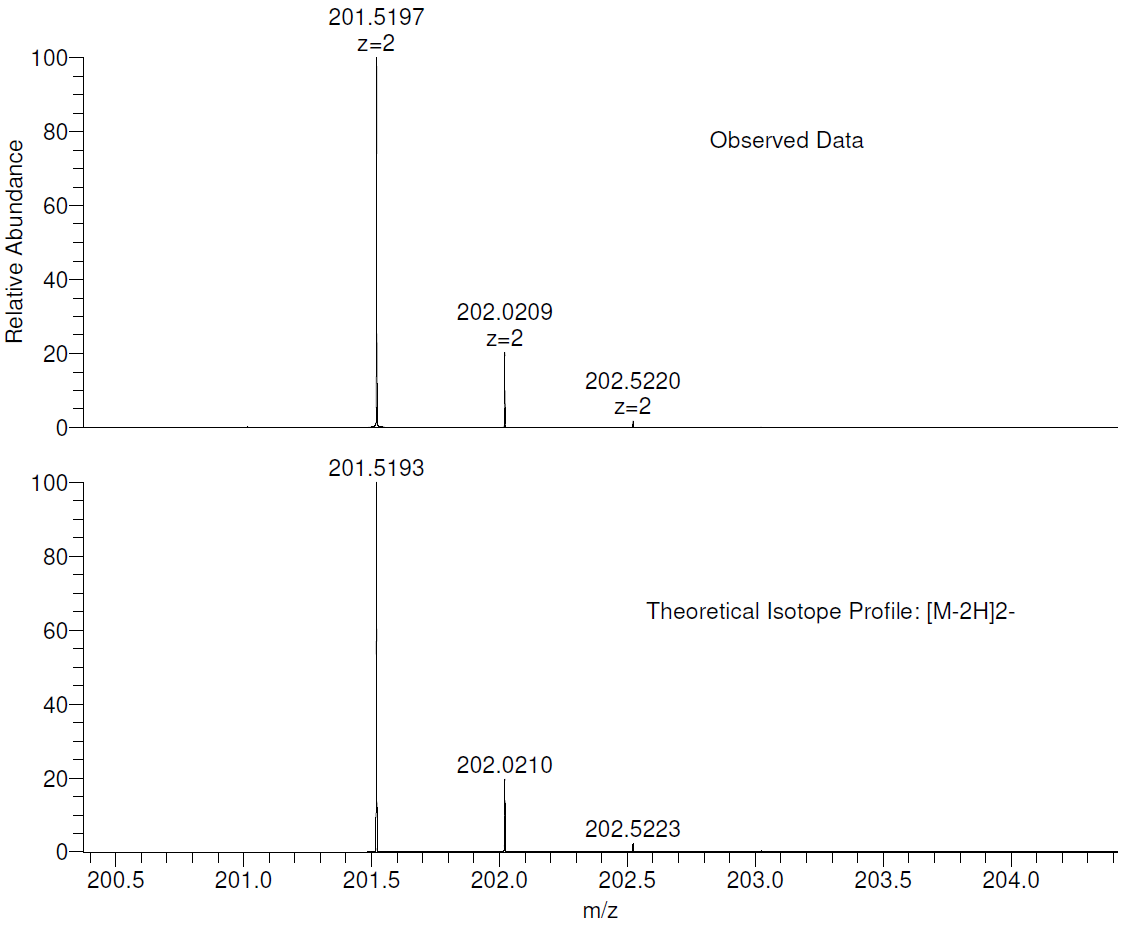
**Figure S59**: Full HSQC NMR spectrum for H4DPPA (3B).

Mass spectral analysis was carried out, and Figure S60 clearly shows the base peak at m/z 201.52, which corresponds to the [M-2H]2- species. This can be seen in greater detail in Figure S61, which compares the theoretical isotope profile with observed data, shown a good match between the two.

Chart

Description automatically generated

**Figure S60**: Full Mass spectrum for H4DPPA.



**Figure S61**: Mass spectrum for H4DPPA showing the observed data and the theoretical isotope profile for the [M-2H]2- species.