IDENTIFICATION OF GLUCOSE PHOSPHATES IN THE *C. ELEGANS* METABOLOME

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Victor Baumann

December 2022

ABSTRACT

Many modular glucosides discovered in the metabolome of the model organism *C. elegans* contain a phosphate group at the 3-position of the glucose. Recent work has provided insights into how other moieties are attached to the sugar, but the point in their syntheses at which the phosphate is attached remains unknown. In this study, glucose 3-phosphate was synthesized and compared with other glucose phosphates and *C. elegans* samples via HPLC-HRMS. Glucose 3-phosphate was verified to be present in the *C. elegans* metabolome along with canonical glucose 6-phosphate and 1-phosphate, but not glucose 2-phosphate. The presence of glucose 3-phosphate suggests that the phosphate may be attached first in the biosynthesis of 3-phosphorylated glucoside derivatives.

BIOGRAPHICAL SKETCH

Victor Baumann was raised in Bowie, Maryland. He graduated from the University of Maryland, College Park, in the spring of 2021 with a Bachelor of Science degree in Chemistry and a minor in Russian. He is a member of Phi Beta Kappa and, more importantly to him, Alpha Chi Sigma. He loves Cornell University for the natural beauty of its campus and the surrounding area.

ACKNOWLEDGEMENTS

I would like to thank Brian Curtis for his aid at every step of this project, Dr. Frank Schroeder for granting me the opportunity to do this research, and Dr. David Zax for the suggestion that I consider this research group in the first place, as well as for their help editing this work. I would also like to thank Jingfang Yu, Yan Yu, Dr. Arnaud Tauffenberger, and Chester Wrobel for their contributions, and the Boyce Thompson Institute for the use of its facilities. This work was supported in part by the National Institutes of Health (grant R35GM131877 to F.C.S.) and made use of the Cornell University NMR Facility, which is supported in part by the NSF through MRI award CHE-1531632.

TABLE OF CONTENTS

BIOGRAPHICAL SKETCH iii

ACKNOWLEDGEMENTS iv

TABLE OF CONTENTS v

LIST OF ABBREVIATIONS vi

INTRODUCTION 1

METHODS 3

RESULTS AND DISCUSSION 7

CONCLUSION 9

REFERENCES 10

APPENDIX I 11

APPENDIX II 18

LIST OF ABBREVIATIONS

AcOH — acetic acid

Bn — benzyl

Cbz — carboxybenzyl

DCM — dichloromethane

DMAP — 4-dimethylaminopyridine

DMF — N,N-dimethylformamide

EtOAc — ethyl acetate

G1P — glucose 1-phosphate

G2P — glucose 2-phosphate

G3P — glucose 3-phosphate

G6P — glucose 6-phosphate

HPLC — high pressure liquid chromatography

HRMS — high resolution mass spectrometry

iPr — isopropyl

LCMS — liquid chromatography–mass spectrometry

mCPBA — *meta*-chloroperoxybenzoic acid

MeOH — methanol

NMR — nuclear magnetic resonance

Pd/C — palladium on carbon

RPM — rotations per minute

RT — room temperature

TBAF — tetrabutylammonium fluoride

THF — tetrahydrofuran

$TIPDSiCl_2$ — 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane

TsOH — *p*-toluenesulfonic acid

CHAPTER 1

INTRODUCTION

The nematode species *Caenorhabditis elegans* is a popular model organism for use in biological studies. As it is easily cultivated, fast-growing, and genetically homologous to other animals including humans, *C. elegans* has been used as a standard model organism for genetics, neurobiology, biochemistry, metabolomics, and other fields.[1-3] Focusing the efforts of thousands of scientists on this one species allows the scientific community to gain a relatively complete understanding of it, so that it is easier to understand the context of each further discovery—for example, the relationship between its genome and metabolome. This understanding of the many factors at play in an animal, even a relatively simple one, can then be applied to inform the study of more complex and important species, including humans.

Over the past two decades many ascarosides, derivatives of the 3,6-dideoxy sugar ascarylose combinatorially attached to moieties derived from various metabolic pathways, have been identified as pheromones secreted by *C. elegans*. Different arrangements of these moieties around the central sugar are synthesized very selectively and can have very different biological effects even where the structural differences between them seem minor.[2,3] More recently, *C. elegans* has been found to produce similar metabolites derived from glucose, many of which include a phosphate at the 3-position (Figure 1). These modular glucosides were initially reported as a method of neutralizing toxins, but the specificity with which they are synthesized, the decreased resilience to starvation of mutants that cannot produce some of them, and their retention in the bodies of the animals indicate that they serve other important functions inside the body, likely as

intraorganismal signaling molecules.[4-6] But while many modular glucosides were identified, our understanding of their biosynthesis remains limited.

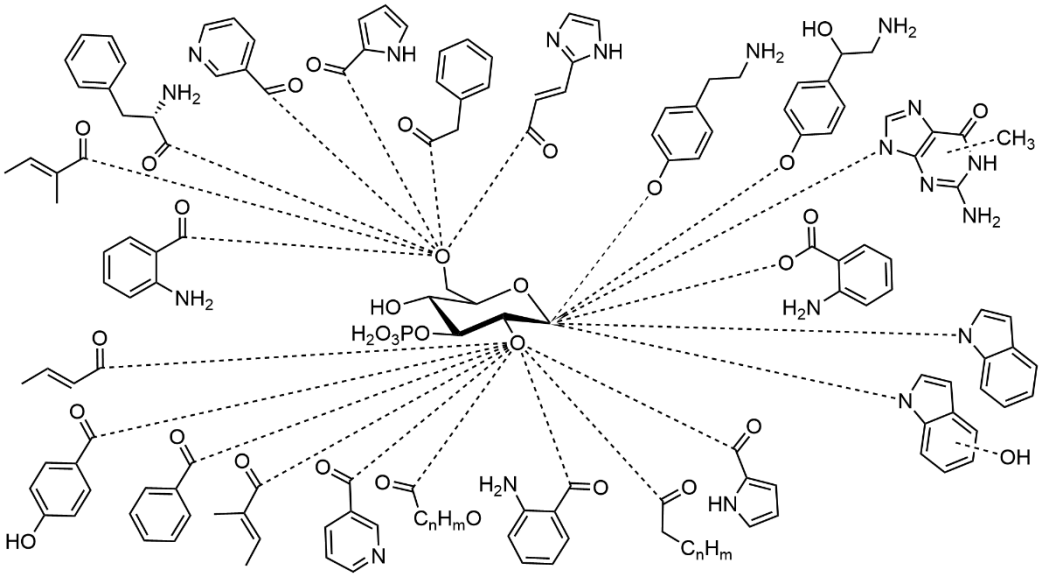


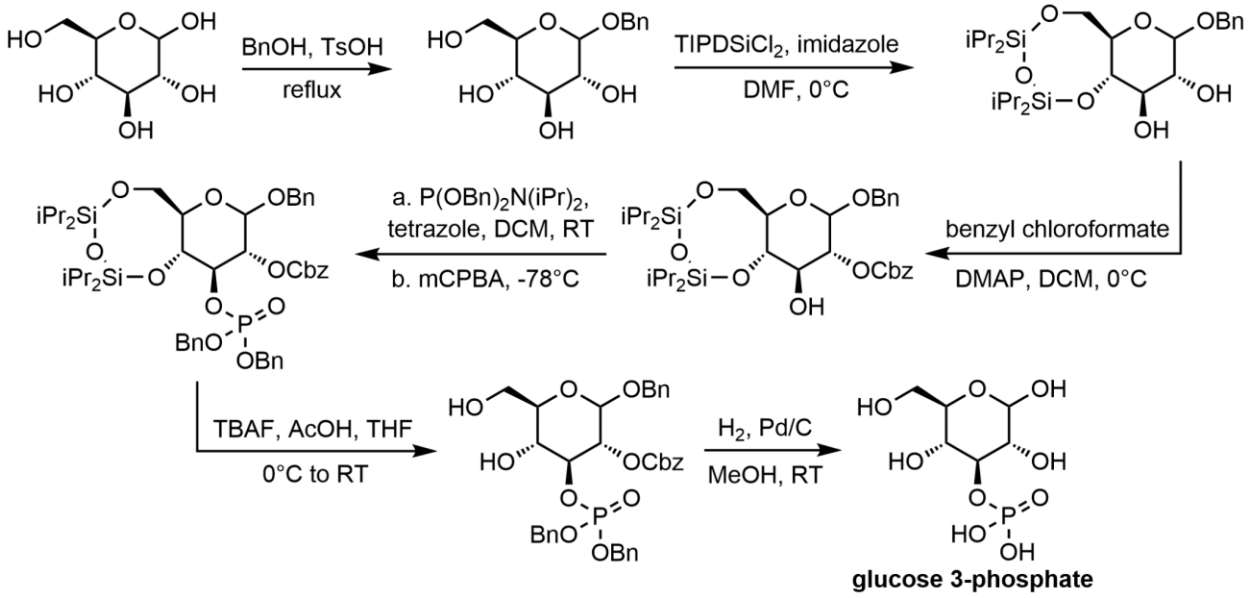
Figure 1. Examples of 3-phosphorylated modular glucosides with a variety of ester and anomeric moieties. Figure adapted from Wrobel et al.[6]

Recent studies have found that the ability of *C. elegans* to synthesize many of these modular glucosides depends on carboxylesterases, but it is not known at what stage of the process the modular glucosides are phosphorylated.[5,6] Do they attach the ester moieties to glucose and then phosphorylate it, suggesting that it has phosphotransferases that can tolerate bulky substituents? Or do they synthesize glucose 3-phosphate (G3P) first and then esterify it, suggesting that their carboxylesterases are able to tolerate a phosphate group? If the latter case is true, one would expect the metabolome of *C. elegans* to include a pool of G3P to use in such syntheses. Whether this is the case may be determined by analyzing a sample of *C. elegans* extract using high pressure liquid chromatography–high resolution mass spectrometry (HPLC-HRMS) and observing whether a peak of the correct *m/z* ratio coelutes with a sample of G3P or other glucose phosphates.

CHAPTER 2

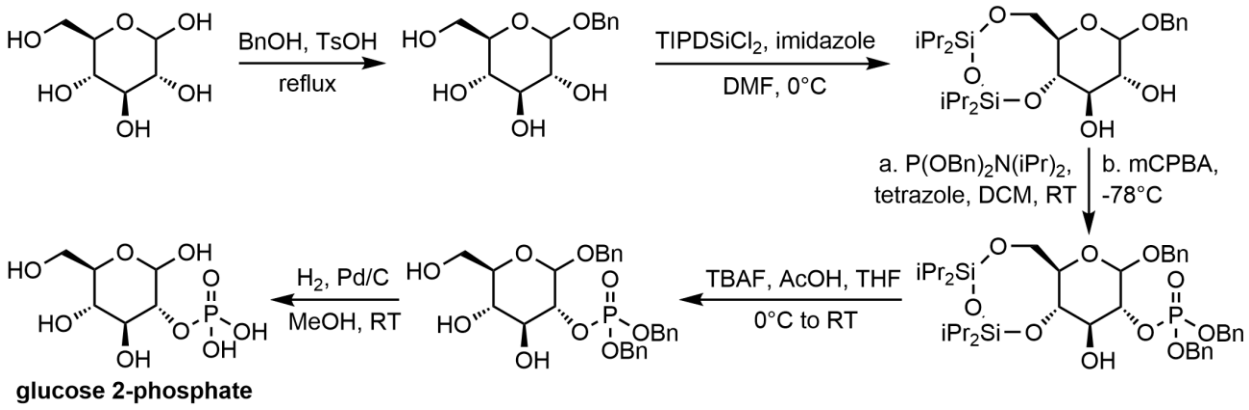
METHODS

Synthesis of Glucose 3-Phosphate (G3P) and Glucose 2-Phosphate (G2P)



Scheme 1. Synthetic route to glucose 3-phosphate from D-glucopyranose.

Unlike glucose 6-phosphate (G6P) and glucose 1-phosphate (G1P), G3P is not commercially available, so it was necessary to synthesize it for this study. This was accomplished by protecting the hydroxyl groups in the 1-, 2-, 4-, and 6-positions before phosphorylating the remaining hydroxyl group in the 3-position, followed by deprotection (Scheme 1).[6,7] The anomeric oxygen was protected with a benzyl group, the 4- and 6-positions were both protected by a single tetraisopropyldisiloxyl group, which attaches to those oxygens in order to form a ring, and then the 2-position was protected by a carboxybenzyl group. Because the oxygen in the 2-position is more reactive than the oxygen in the 3-position, both the carboxybenzyl and the phosphate will selectively attach there when it is available.



Scheme 2. Synthetic route to glucose 2-phosphate from D-glucopyranose.

Initially, we attempted to phosphorylate the glucose without protecting the oxygen in the 2-position, under the impression that treatment with tetrabutylammonium fluoride would cause the phosphate group to migrate to the 3-position. This did not go as planned, resulting in a sample mostly of benzyl-protected glucose 2-phosphate (G2P), which was subsequently reduced to afford a sample of G2P with traces of G3P (Scheme 2). The synthesis of G3P was then reattempted with the addition of the 2-O-protection step before the phosphorylation step, as shown in Scheme 1, successfully yielding G3P (Figure 2). Detailed synthetic procedures can be found in Appendix I, and NMR spectra of each intermediate and product can be found in Appendix II.

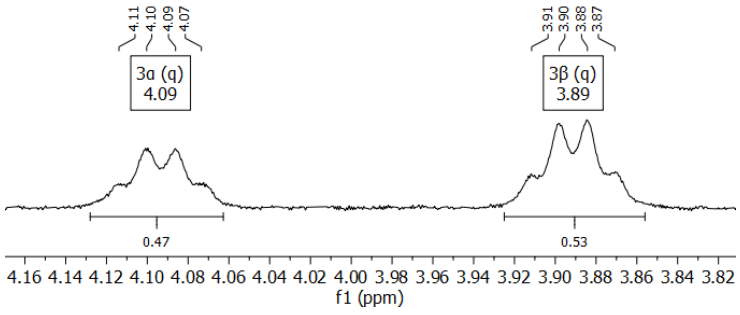


Figure 2. Portion of the ^{1}H NMR spectrum of a mixture of the α- and β-anomers of glucose 3-phosphate in deuterium oxide, showing the peaks corresponding to the proton at the 3-position of each anomer as a quartet, resulting from two ^{1}H,^{1}H-diaxial couplings and the coupling with ^{31}P.

C. elegans Extract Preparation

Wild-type animals were grown in S-complete media (100 mM NaCl, 5 μg/ml cholesterol, 50 mM potassium phosphate buffer, 10 mM potassium citrate, 10 mM trace metal solution, 3 mM $CaCl_2$, 3 mM $MgSO_4$) with OP50-1 *E. coli* (concentrated 10x). Eggs were extracted using a 20% bleach solution and washed 3 times with M9 buffer (22 mM KH_2PO_4, 42 mM Na_2HPO_4, 100 mM NaCl, 1 mM $MgSO_4$). Eggs were counted and seeded in S-complete with OP50-1. 30,000 animals were seeded in triplicates and grown for 72 hours (to day 1 of adulthood) at 20°C and 180 RPM during the entire experiment. An aliquot was taken and assessed under microscope to ensure a proper developmental stage.

For harvest, cultures were centrifuged 1 minute at 1000 rpm at room temperature. The pellets were stored frozen on dry ice, then lyophilized for 48 hours and clarified in 5 mL of methanol overnight at room temperature. The sample was then centrifuged at 1000 G for 5 minutes at room temperature. The supernatant was subsequently dried using a SpeedVac, concentrated to a final volume of 75 μl, and stored at -20°C in glass HPLC vials until analysis.

Mass Spectroscopic Analysis

G1P and G6P were purchased from Sigma-Aldrich as disodium salt hydrates. Samples of each glucose phosphate were dissolved in water and diluted with methanol. Due to delays in acquiring G1P, two sets of mass spectroscopic data were collected, one without G1P and one with it. The latter set (Figure 2, below) used a sample of *C. elegans* extract prepared as described above, while the former set (Figure 1) used another wildtype pellet sample prepared for a prior study.[8] High resolution LCMS analysis was performed on a Thermo Fisher Scientific Vanquish Horizon UHPLC System coupled with a Thermo Q Exactive HF hybrid quadrupole-orbitrap high resolution

mass spectrometer equipped with a heated electrospray ionization ion source. 1 µL of each sample of glucose phosphate or *C. elegans* extract was analyzed via a water-acetonitrile gradient on an XBridge Amide column (150 mm x 2.1 mm 3.5 µm particle size 130 Å pore size, Waters) at 40 °C.

Solvent A was 90% acetonitrile and 10% water to which was added 0.4% (v/v) of 25% ammonia in water solution and 0.1% (v/v) formic acid, and solvent B was 30% acetonitrile and 70% water to which was added 0.4% (v/v) of 25% ammonia in water solution and 0.1% (v/v) formic acid. The A/B gradient started at 1% B for 3 minutes, then rose from 1% to 60% B over 17 minutes, then rose from 60% to 100% B over 6 minutes and held for 1.5 minutes, then returned to 1% B over 0.5 minutes and equilibrated for 2 minutes.

Mass spectrometer parameters: 3.5 kV spray voltage, 380 °C capillary temperature, 300 °C probe heater temperature, 60 sheath flow rate, 20 auxiliary flow 15 rate, 1 spare gas; S-lens RF level 50.0, resolution 240,000, *m/z* range 100-1200, AGC target 3e6. The instrument was calibrated with Pierce LTQ Velos ESI positive and negative calibration solutions (Thermo-Fisher). LCMS data was analyzed using Xcalibur version 4.1.31.9 QualBrowser (Thermo Scientific).

CHAPTER 3

RESULTS AND DISCUSSION

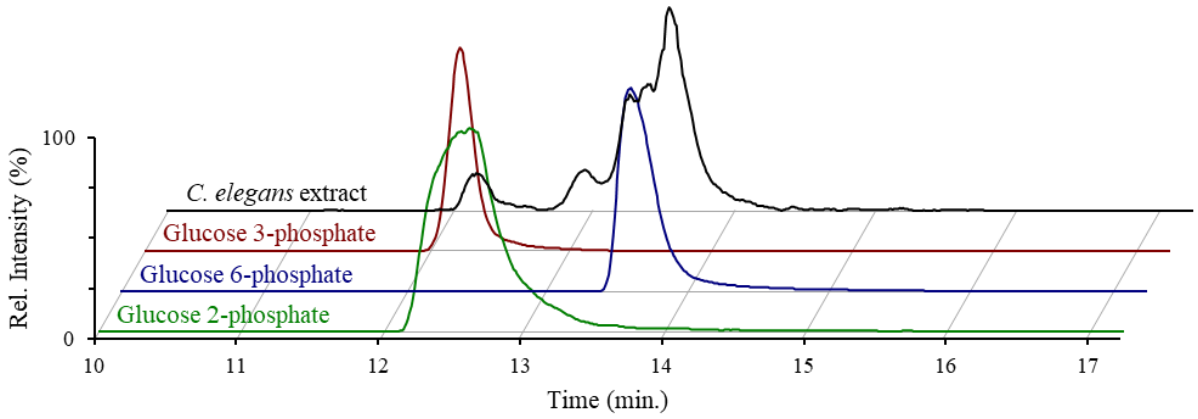


Figure 3. HPLC-HRMS ion chromatograms (*m/z* = 259.02037–259.02451) of *C. elegans* extract, synthetic samples of glucose 3-phosphate and glucose 2-phosphate, and a commercial sample of glucose 6-phosphate. Due to delays in acquiring G1P, two sets of mass spectroscopic data were collected, one without G1P and one with it.

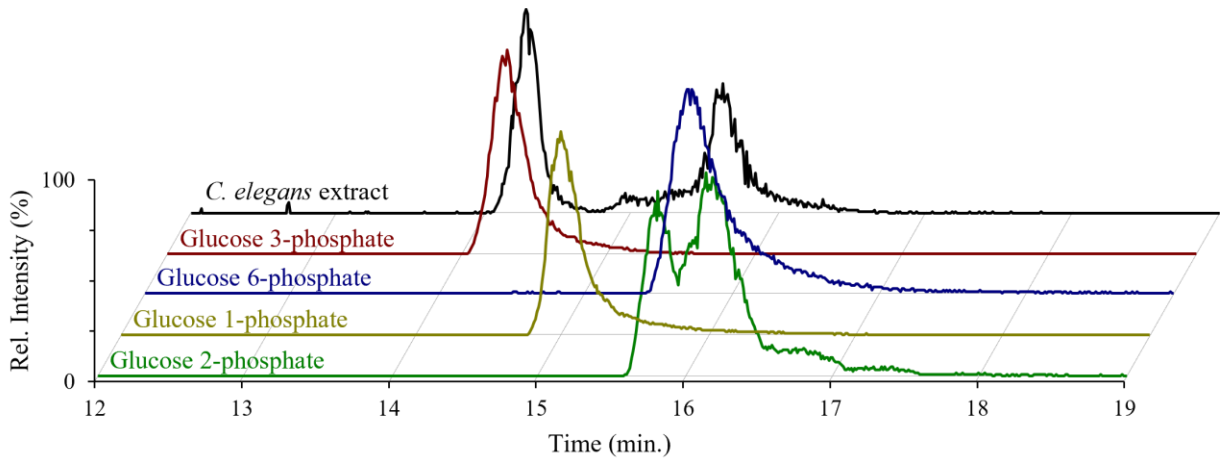


Figure 4. HPLC-HRMS ion chromatograms (*m/z* = 259.02037–259.02451) of *C. elegans* extract, synthetic samples of glucose 3-phosphate and glucose 2-phosphate, and commercial samples of glucose 6-phosphate and glucose 1-phosphate.

The synthetic sample of G3P directly correlates with a peak in the *C. elegans* sample, strongly supporting that *C. elegans* contains a pool of G3P. Thus, it is possible that *C. elegans* synthesizes modular glucosides from their existing stock of G3P or polysaccharides thereof.

The sample of G6P also coelutes with a strong peak in the extract sample, consistent with the expectation that it would be substantially present in *C. elegans* due to its role in glycolysis. The sample of G1P coelutes with a small peak in Figure 4 and likely corresponds to a larger but still small peak in the same region of Figure 3, consistent with the expectation that it would be present due to its role in glycogen metabolism.[9]

The peak of G2P varies strangely between the two data sets, but as it does not correspond to the peaks of the *C. elegans* sample in either set, it can be concluded that *C. elegans* does not produce G2P in quantities comparable with the other glucose phosphates. Considering the structural similarity of G2P to G3P and the fact that the G2P sample contains substantial traces of G3P, it is not surprising that the G2P peak overlaps with the G3P peak in Figure 3, but the reason for its change in Figure 4 is not known. It is possible that there could be an impurity in the G2P sample interfering with its retention.

HPLC-HRMS, especially using a hydrophilic column, is very sensitive to differences in solvent composition and pH, and while the solvents used for each set of data were prepared to the same specifications it is likely that slight differences between them are responsible for the observed differences in retention time and peak shape. The differences between the *C. elegans* samples are due to natural variation between animal populations, and despite their differences, the presence of peaks corresponding to G3P, G6P, and G1P in both samples supports the generalizability of these results.

CHAPTER 4

CONCLUSION

In confirming that *C. elegans* possesses a supply of G3P, and therefore may use it in the synthesis of modular glucosides, a small but important piece has been added to the puzzle of *C. elegans* metabolomics. Future studies may use HPLC-HRMS analysis of *C. elegans* samples fed with isotope-labeled G3P to confirm whether or not it truly is used as a precursor in the biosynthesis of modular glucosides. They also may conduct more analysis to verify the retention time of G2P in relation to the other glucose phosphates, and examine samples of other organisms to find which organisms the results of this study hold true for beyond *C. elegans*. Mutant strains may be examined to determine the genes and, by extension, proteins responsible for G3P production, and comparative metabolomic studies may examine factors that affect G3P supply such as age, sex, and diet. More broadly, further research may be done in the vein of Wrobel et al, searching for enzymes that facilitate the production of modular glucosides from G3P, which in combination with this work may elucidate the synthetic route by which *C. elegans* produces them with such remarkable selectivity.[6]

REFERENCES

1. Corsi, A. K.; Wightman, B.; Chalfie, M. A Transparent Window into Biology: A Primer on *Caenorhabditis elegans*. *Genetics* **2015**, *200* (2), 387-407. DOI: 10.1534/genetics.115.176099.

2. Schroeder, F. C. Modular Assembly of Primary Metabolic Building Blocks: A Chemical Language in *C. elegans*. *Chem. Biol.* **2015**, *22* (1), 7-16. DOI: 10.1016/j.chembiol.2014.10.012.

3. Salzer, L.; Witting, M. Quo Vadis *Caenorhabditis elegans* Metabolomics—A Review of Current Methods and Applications to Explore Metabolism in the Nematode. *Metabolites* **2021**, *11* (5), 284. DOI: 10.3390/metabo11050284.

4. Stupp, G. S.; von Reuss, S. H.; Izrayelit, Y.; Ajredini, R.; Schroeder, F. C.; Edison, A. S. Chemical Detoxification of Small Molecules by *Caenorhabditis elegans*. *ACS Chem. Biol.* **2013**, *8* (2), 309-313. DOI: 10.1021/cb300520u.

5. Le, H. H.; Wrobel, C. J. J.; Cohen, S. M.; Yu, J. F.; Park, H.; Helf, M. J.; Curtis, B. J.; Kruempel, J. C.; Rodrigues, P. R.; Hu, P. J.; et al. Modular metabolite assembly in *Caenorhabditis elegans* depends on carboxylesterases and formation of lysosome-related organelles. *eLife* **2020**, *9*. DOI: 10.7554/eLife.61886.

6. Wrobel, C. J. J.; Yu, J. F.; Rodrigues, P. R.; Ludewig, A. H.; Curtis, B. J.; Cohen, S. M.; Fox, B. W.; O'Donnell, M. P.; Sternberg, P. W.; Schroeder, F. C. Combinatorial Assembly of Modular Glucosides via Carboxylesterases Regulates *C. elegans* Starvation Survival. *J. Am. Chem. Soc.* **2021**, *143* (36), 14676-14683. DOI: 10.1021/jacs.1c05908.

7. Yu, J. F.; Vogt, M. C.; Fox, B. W.; Wrobel, C. J. J.; Palomino, D. F.; Curtis, B. J.; Zhang, B. S.; Le, H. H.; Tauffenberger, A.; Hobert, O.; et al. Parallel pathways for serotonin biosynthesis and metabolism in *C. elegans*. *Nat. Chem. Biol.* **2022**. DOI: 10.1038/s41589-022-01148-7.

8. Burkhardt, R. N.; Artyukhin, A. B.; Aprison, E. Z.; Curtis, B. J.; Fox, B. W.; Ludewig, A. H.; Chaturbedi, A.; Panda, O.; Wrobel, C. J. J.; Lee, S. S.; et al. Sex-specificity of the *C. elegans* metabolome. *bioRxiv* **2022**, 2022.2008.2011.503636. DOI: 10.1101/2022.08.11.503636.

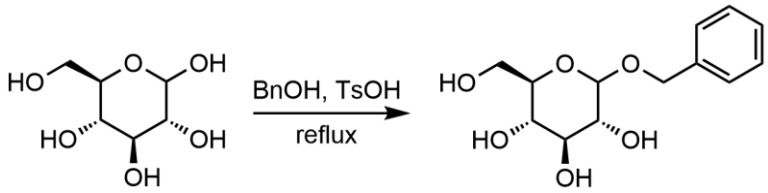
9. Huijing, F. Glycogen Metabolism and Glycogen-Storage Diseases. *Physiol. Rev.* **1975**, *55* (4), 609-658. DOI: 10.1152/physrev.1975.55.4.609.

APPENDIX I: PROCEDURES

General Synthetic Procedures

Unless stated otherwise, all chemicals and reagents used for synthetic compound preparation were purchased from Sigma-Aldrich. Benzyl chloroformate was purchased from Acros Organics, 4-dimethylaminopyridine (DMAP) from Fluka, and N,N-dimethylformamide (DMF) from BDH. Acetic acid (AcOH), acetonitrile, dichloromethane (DCM), ethyl acetate (EtOAc), formic acid, hexanes, methanol (MeOH), and toluene used for chromatography or as a reagent or solvent were purchased from Fisher Scientific. Thin-layer chromatography was performed using J. T. Baker Silica Gel IB2F plates with analysis via UV and *p*-anisaldehyde stain. Mixtures (reaction or from chromatography) were concentrated using a Buchi rotary evaporator and reaction mixtures containing dimethylformamide (DMF) were concentrated in presence of toluene for azeotropic removal. Flash chromatography was performed using Teledyne Isco CombiFlash systems and Teledyne Isco RediSep Rf silica columns. All deuterated solvents were purchased from Cambridge Isotopes. NMR spectra were recorded on a 500 MHz Bruker AVIII with BBO Prodigy cryoprobe or 600 MHz Varian INOVA 600 spectrometer at the Cornell University NMR Facility. NMR spectroscopic data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet), coupling constants (Hz), and integration. ^{1}H NMR chemical shifts are reported in ppm (δ) relative to residual solvent peaks (3.31 ppm for methanol-d_4 and 4.79 ppm for D_2O). ^{13}C NMR chemical shifts are also reported in ppm (δ) relative to residual solvent peaks (49.00 ppm for methanol-d_4 and for residual methanol signal in D_2O). All NMR data processing was done using MestReNova 14.2.3.

Benzyl D-glucopyranoside

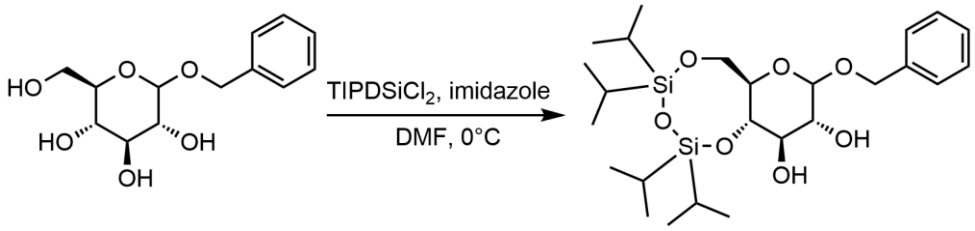


To a solution of D-glucopyranose (1.014g, 5.63 mmol, 1 equiv.) in benzyl alcohol (1.7 mL) was added *p*-toluenesulfonic acid monohydrate (322 mg, 1.70 mmol, 0.3 equiv.). The resulting solution was refluxed for 20 hours. Flash column chromatography on silica using a gradient of 0-14% MeOH in DCM was performed, affording benzyl D-glucopyranoside (465 mg, 31%) as a mixture of anomers.

^{1}H NMR (500 MHz, methanol-d_4): δ 7.42 (m, 2H), 7.33 (m, 2H), 7.27 (m, 1H), 4.93 (d, J = 11.8 Hz, 1H$_\beta$), 4.89 (d, J = 3.7 Hz, 1H$_\alpha$), 4.77 (d, J = 11.8 Hz, 1H$_\alpha$), 4.67 (d, J = 11.8 Hz, 1H$_\beta$), 4.56 (d, J = 11.8 Hz, 1H$_\alpha$), 4.35 (d, J = 7.7 Hz, 1H$_\beta$), 3.90 (dd, J = 11.9, 2.2 Hz, 1H$_\beta$), 3.79 (dd, J = 11.7, 2.2 Hz, 1H$_\alpha$), 3.71 – 3.60 (m, 2H), 3.41 (dd, J = 9.7, 3.8 Hz, 1H$_\alpha$), 3.35 – 3.23 (m, 4H).

^{13}C NMR (126 MHz, methanol-d_4): δ 129.82 – 128.67 (m), 103.28, 99.26, 78.10, 78.04, 75.14, 75.10, 73.91, 73.59, 71.84, 71.72, 71.70, 70.19, 62.82, 62.66.

Benzyl 4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside

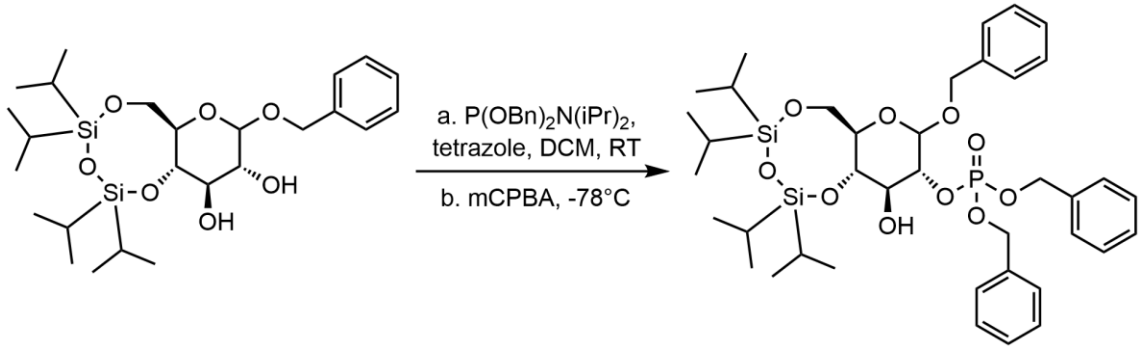


To a solution of benzyl D-glucopyranoside (101 mg, 0.373 mmol, 1 equiv.) in DMF containing imidazole (102 mg, 1.50 mmol, 4 equiv.) at 0 ˚C was added 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane (0.12 mL, 0.38 mmol, 1 equiv.). After stirring for 25 minutes cold (cooling with an ice-water bath), additional $TIPDSiCl_2$ (0.04 mL, 0.1 mmol, 0.3 equiv.) was added. After stirring for another 15 minutes. additional $TIPDSiCl_2$ (0.02 mL, 0.06 mmol, 0.2 equiv.) was added. After stirring for another 20 minutes. the reaction was quenched with DCM and water, extracted with DCM, rinsed with saturated aqueous $NaHCO_3$, and dried with Na_2SO_4. The solution was then concentrated, and then toluene was added and the solution was concentrated again, to remove DMF. Flash column chromatography on silica using a gradient of 2-6% MeOH in DCM was performed, affording benzyl 4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside (60 mg, 31%) as a mixture of anomers.

^{1}H NMR (500 MHz, methanol-d$_4$): δ 7.42 (m, 2H), 7.34 (m, 2H), 7.28 (m, 1H), 4.91 (d, *J* = 11.7 Hz, 1H$_\beta$), 4.89 (d, *J* = 4.0 Hz, 1H$_\alpha$), 4.69 (d, *J* = 12.0 Hz, 1H$_\alpha$), 4.64 (d, J = 11.8 Hz, 1H$_\beta$), 4.59 (d, J = 12.1 Hz, 1H$_\alpha$), 4.37 (d, *J* = 7.8 Hz, 1H$_\beta$), 4.16 (dd, *J* = 12.4, 2.1 Hz, 1H$_\beta$), 4.08 (dd, *J* = 12.4, 2.1 Hz, 1H$_\alpha$), 3.95 (dd, *J* = 12.4, 1.4 Hz, 1H$_\beta$), 3.77 (t, J = 9.2 Hz, 1H$_\beta$), 3.74 – 3.62 (m, 3H$_\alpha$), 3.55 – 3.47 (m, 1H$_\alpha$), 3.41 (t, *J* = 9.1 Hz, 1H$_\beta$), 3.39 – 3.36 (dd, *J* = 9.0, 4.2 Hz, 1H$_\alpha$), 3.25 – 3.19 (m, 2H$_\beta$), 1.16 – 0.95 (m, 28H)

^{13}C NMR (126 MHz, methanol-d$_4$): δ 129.39 – 128.72 (m), 103.92, 99.85, 77.84, 77.56, 75.57, 74.75, 73.93, 73.79, 71.94, 71.05, 70.74, 62.18, 62.12, 18.03 – 17.56 (m).

Benzyl 2-O-dibenzylphosphono-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside



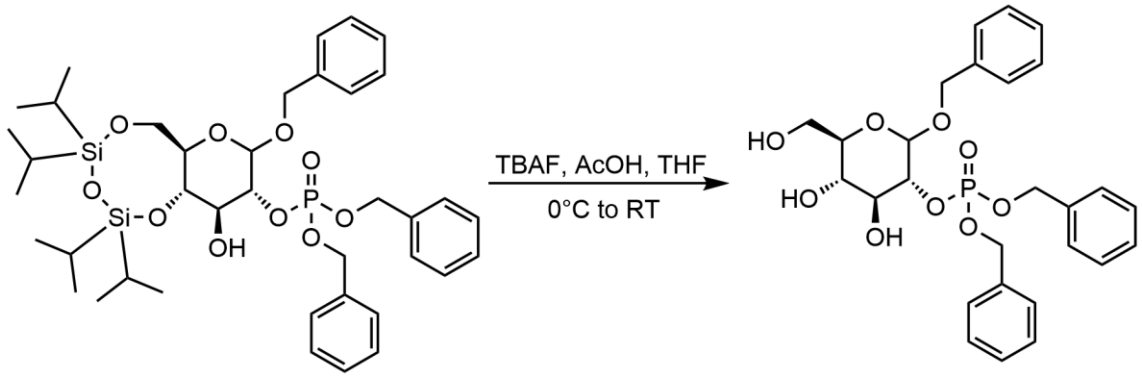
To a solution of benzyl 4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside (27 mg, 0.061 mmol, 1 equiv.) in DCM (2 mL) containing imidazolium triflate (41 mg, 0.19 mmol, 3 equiv.) was added dibenzyl N,N-diisopropylphosphoramidite (0.02 mL, 0.06 mmol, 1 equiv.).

After stirring for 20 minutes. additional dibenzyl N,N-diisopropylphosphoramidite (0.01 mL, 0.3 mmol, 0.5 equiv.) was added, and the resulting solution stirred for an additional 15 minutes. The solution was then placed in a dry ice-acetone bath, and mCPBA (40 mg, ≥77% pure, 0.18 mmol, 3 equiv.) was added. The resulting solution was stirred for 25 minutes. then removed from the bath and quenched with saturated aqueous $NaHCO_3$. The product was then extracted with DCM, dried with Na_2SO_4, and flash column chromatography on silica using a gradient of 5-33% EtOAc in hexane was performed, affording benzyl 2-O-dibenzylphosphono-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside (18 mg, 44%) as a mixture of anomers with traces of 3-O-dibenzylphosphono-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside. Due to the many overlapping NMR signals, the minor anomeric peaks are not reported.

^{1}H NMR (500 MHz, methanol-d$_4$): δ 7.38 – 7.22 (m, 15H), 5.11 (d, J = 4.0 Hz, 1H$_\alpha$), 5.10 – 4.89 (m, 4H$_\alpha$), 4.66 (d, J = 11.9 Hz, 1H$_\alpha$), 4.47 (d, J = 12.0 Hz, 1H$_\alpha$), 4.13 – 4.05 (m, 2H$_\alpha$), 3.92 (t, J = 9.3 Hz, 1H$_\alpha$), 3.79 – 3.74 (m, 2H$_\alpha$), 3.57 (dt, J = 9.6, 1.7 Hz, 1H$_\alpha$), 1.17 – 0.99 (m, 28H).

^{13}C NMR (126 MHz, methanol-d$_4$): δ 129.80 – 128.83 (m), 97.93, 79.42, 73.69, 72.79, 71.08 – 70.86 (m), 61.97, 18.18 – 17.46 (m).

Benzyl 2-O-dibenzylphosphono-D-glucopyranoside



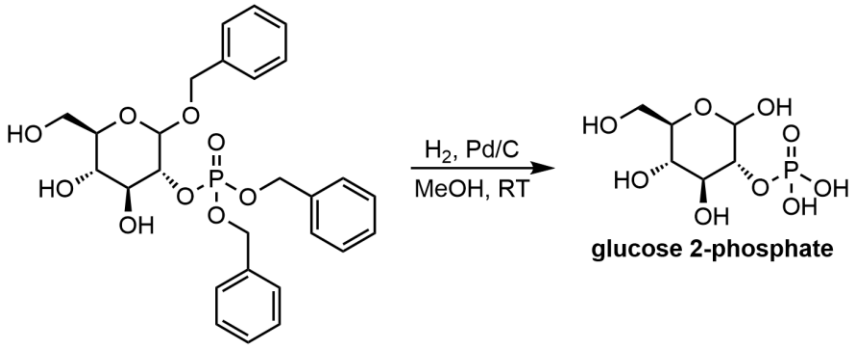
To a solution of benzyl 2-O-dibenzylphosphono-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside (14 mg, 0.018 mmol, 1 equiv.) in THF (1 mL) at 0 ˚C was added TBAF (1M in THF, 0.05 mL, 0.05 mmol, 2.5 equiv.). After stirring for 20 minutes cold (cooling with an ice-water bath), a drop of AcOH was added, and the resulting solution was then concentrated to remove THF. Flash column chromatography on silica using a gradient of 2-10% MeOH in DCM was performed, affording benzyl 2-O-dibenzylphosphono-D-glucopyranoside (5 mg, 60%).

^{1}H NMR (500 MHz, methanol-d$_4$): δ 5.12 (d, J = 3.7 Hz, 1H$_\alpha$), 5.13 – 4.92 (m, 4H), 4.97 (d, J = 11.9 Hz, 1H$_\beta$), 4.76 (d, J = 11.9 Hz, 1H$_\alpha$), 4.64 (d, J = 8.0 Hz, 1H$_\beta$), 4.61 (d, J = 11.7 Hz, 1H$_\beta$), 4.46 (d, J = 11.9 Hz, 1H$_\alpha$), 4.18 – 4.11 (m, 1H), 3.93 (dd, J = 11.9, 2.2 Hz, 1H$_\beta$), 3.89 (t, J = 9.3 Hz, 1H$_\alpha$), 3.85 – 3.79 (m, 1H$_\alpha$), 3.75 – 3.63 (m, 2H$_\alpha$, 1H$_\beta$), 3.59 (t, J = 9.0 Hz, 1H$_\beta$), 3.42 – 3.36 (m, 1H$_\alpha$, 2H$_\beta$), 3.35 (s, 1H$_\beta$).

^{13}C NMR (126 MHz, methanol-d$_4$): δ 138.71, 137.16, 129.66 – 128.79 (m), 101.24, 97.37, 81.43, (d, *J* = 7.5 Hz), 79.18, (d, *J* = 6.5 Hz), 78.15, 76.91, 73.95, 73.21, (d, *J* = 6.4 Hz), 71.74, 71.01, (d, *J* = 5.9 Hz), 70.65, (d, *J* = 5.9 Hz), 70.4, 62.59, 62.45.

^{31}P NMR (202 MHz, methanol-d$_4$): δ -2.19, -2.39.

2-O-phosphono-D-glucopyranose



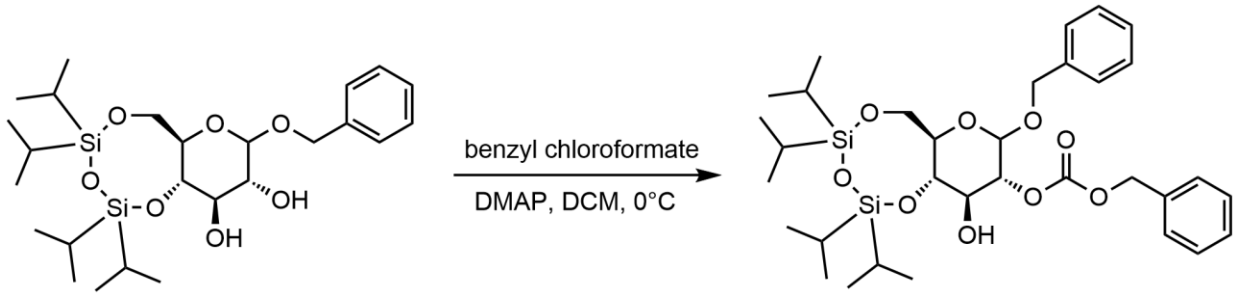
To a solution of benzyl 2-O-dibenzylphosphono-D-glucopyranoside (5 mg, 0.01 mmol) in DCM (1 mL) was added Pd/C (6 mg, 0.06 mmol). The reaction mixture was purged with argon for 10 minutes. then H_2 gas was passed over it for 90 minutes at room temperature, and the reaction vessel was again purged with argon for 10 minutes. The reaction mixture was filtered through celite, affording 2-O-phosphono-D-glucopyranose (~4 mg) with traces of 3-O-phosphono-D-glucopyranose. Due to the many overlapping NMR signals, the minor anomeric peaks are not reported.

^{1}H NMR (500 MHz, deuterium oxide): δ 5.39 (d, *J* = 3.3 Hz, 1H$_\alpha$), 5.23 (d, *J* = 3.4 Hz, 1H$_\alpha$, G3P), 4.66 (d, *J* = 11.6 Hz, 1H$_\beta$, G3P), 3.94 (td, *J* = 9.0, 3.4 Hz, 1H$_\alpha$), 3.88 – 3.72 (m, 3H$_\alpha$), 3.71 (dd, *J* = 12.4, 4.8 Hz, 1H$_\alpha$), 3.64 (t, *J* = 8.5 Hz, 1H$_\alpha$), 3.47 (t, *J* = 9.7 Hz, 1H$_\alpha$), 3.46 (s, 1H$_\alpha$).

^{13}C NMR (126 MHz, deuterium oxide): δ 95.24, 91.02, 76.03, 75.46 – 75.33 (m), 71.77 (d, *J* = 5.7 Hz), 71.42, 69.48 (d, *J* = 5.7 Hz), 60.64, 52.98.

^{31}P NMR (202 MHz, deuterium oxide): δ 2.98, 1.47, 0.24, -0.01, -0.20.

Benzyl 2-O-benzyloxycarbonyl-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside



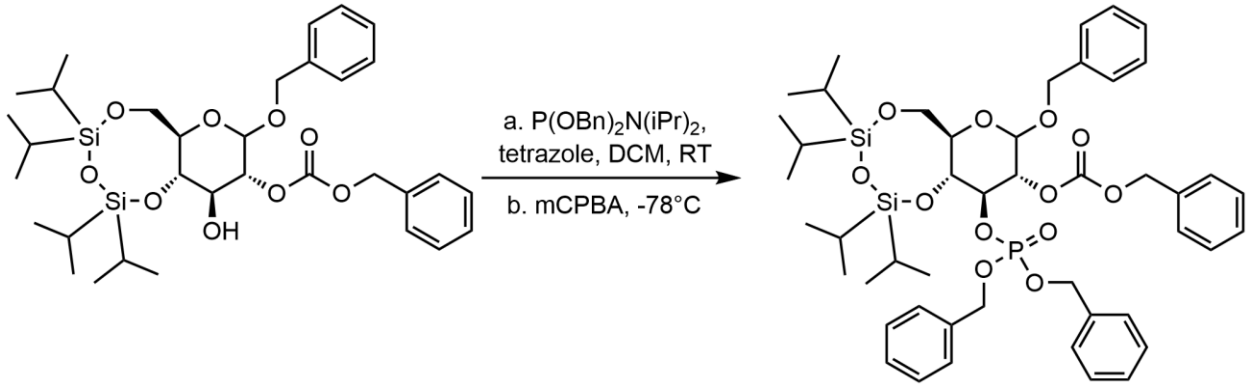
To a solution of benzyl 4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside (30 mg, 0.068 mmol, 1 equiv.) in DCM (2 mL) containing DMAP (18 mg, 0.14 mmol, 2 equiv.) at 0 ˚C

was added benzyl chloroformate (19 µL, 0.14 mmol, 2 equiv.). After stirring for 15 minutes cold (cooling with an ice-water bath), additional DMAP (9 mg, 0.07 mmol, 1 equiv.) and benzyl chloroformate (10 µL, 0.07 mmol, 1 equiv.) were added. After stirring for another 55 minutes cold up to room temp., additional DMAP (4 mg, 0.03 mmol, 0.5 equiv.) and benzyl chloroformate (5 µL, 3 mmol, 0.5 equiv.) were added. After stirring for another 15 minutes. additional DMAP (4 mg, 0.03 mmol, 0.5 equiv.) and benzyl chloroformate (5 µL, 3 mmol, 0.5 equiv.) were added. After stirring for another 15 minutes. the reaction was quenched with saturated aqueous $NaHCO_3$, extracted with DCM, and dried with Na_2SO_4. Flash column chromatography on silica using a gradient of 2-22% MeOH in DCM was performed, affording benzyl 2-O-benzyloxycarbonyl-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside (25 mg, 66%).

^{1}H NMR (500 MHz, methanol-d$_4$): δ 7.39 – 7.24 (m, 10H), 5.20 (d, *J* = 12.3 Hz, 1H$_\beta$), 5.16 – 5.16 (m, 1H$_\beta$), 5.13 (d, *J* = 1.5 Hz, 2H$_\alpha$), 5.11 (d, *J* = 3.9 Hz, 1H$_\alpha$), 4.85 (s, 1H$_\beta$), 4.67 (d, *J* = 12.3 Hz, 1H$_\alpha$), 4.60 – 4.51 (m, 3H$_\beta$), 4.51 (d, *J* = 12.3 Hz, 1H$_\alpha$), 4.36 (dd, *J* = 9.9, 3.9 Hz, 1H$_\alpha$), 4.16 (dd, *J* = 12.5, 2.1 Hz, 1H$_\beta$), 4.12 (dd, *J* = 12.7, 2.1 Hz, 1H$_\alpha$), 3.97 (dd, *J* = 12.6, 1.4 Hz, 1H$_\beta$), 3.90 (t, *J* = 9.4 Hz, 1H$_\alpha$), 3.83 (t, *J* = 9.2 Hz, 1H$_\beta$), 3.80 – 3.76 (m, 2H$_\alpha$), 3.62 (ddd, *J* = 9.2, 7.7, 1.8 Hz, 1H$_\beta$), 3.58 (dt, *J* = 9.3, 1.7 Hz, 1H$_\alpha$), 3.27 (dt, *J* = 9.5, 1.8 Hz, 1H$_\beta$), 1.19 – 0.95 (m, 28H).

^{13}C NMR (126 MHz, methanol-d$_4$): δ 129.65 – 128.67 (m), 101.65, 96.54, 79.83, 78.58, 77.89, 75.25, 73.64, 72.02, 71.81, 70.95, 70.76, 70.60, 62.00, 18.00 – 17.48 (m).

Benzyl 2-O-benzyloxycarbonyl-3-O-dibenzylphosphono-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside



To a solution of benzyl 2-O-benzyloxycarbonyl-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside (17 mg, 0.030 mmol, 1 equiv.) in DCM (2 mL) containing tetrazole (0.45 M in MeCN, 0.4 mL, 0.18 mmol, 6 equiv.) was added dibenzyl N,N-diisopropylphosphoramidite (60 µL, 0.18 mmol, 6 equiv.). After stirring for 20 minutes. additional tetrazole (0.4 mL, 0.18 mmol, 6 equiv.) and phosphoramidite (60 µL, 0.18 mmol, 6 equiv.) were added, and the resulting solution stirred for an additional 35 minutes. The solution was then placed in a dry ice-acetone bath, and mCPBA (41 mg, ≥77% pure, 0.18 mmol, 6 equiv.) was added. The resulting solution was stirred for 20 minutes. then removed from the bath and quenched with saturated aqueous $NaHCO_3$. The product was then extracted with DCM, dried with Na_2SO_4, and flash column chromatography on silica using a gradient of 2-19% EtOAc in hexane was performed, affording benzyl 2-O-

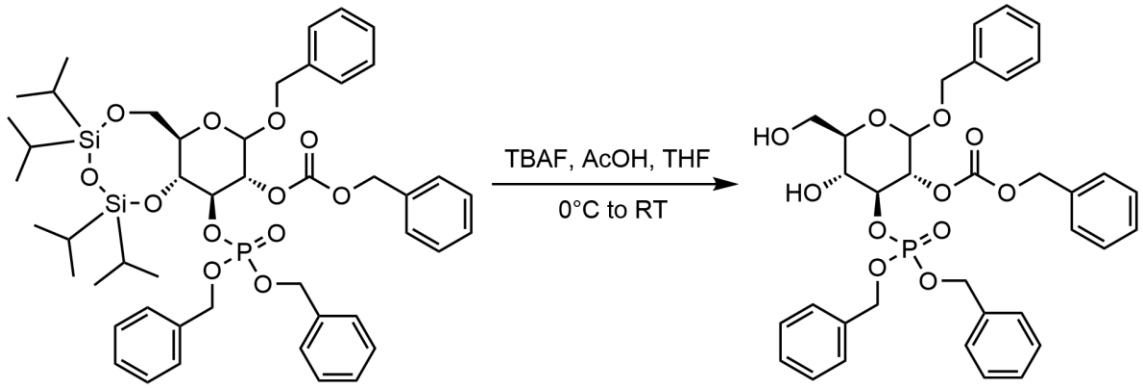
benzyloxycarbonyl-3-O-dibenzylphosphono-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside (29 mg, 120%). The high yield is attributable to residual *meta*-chlorobenzoic acid (~22% w/w). Due to the many overlapping NMR signals, the minor anomeric peaks are not reported.

^{1}H NMR (500 MHz, methanol-d$_4$): δ 7.34 – 7.23 (m, 20H), 5.16 (d, *J* = 3.8 Hz, 1H$_\alpha$), 5.13 – 4.85 (m, 6H), 4.76 (q, *J* = 9.3 Hz, 1H$_\alpha$), 4.70 (d, *J* = 12.3 Hz, 1H$_\alpha$), 4.64 (dd, *J* = 9.9, 3.8 Hz, 1H$_\alpha$), 4.54 (d, *J* = 12.3 Hz, 1H$_\alpha$), 4.12 – 4.08 (m, 1H$_\alpha$), 4.00 (t, *J* = 9.3 Hz, 1H$_\alpha$), 3.81 (dd, *J* = 12.8, 1.6 Hz, 1H$_\alpha$), 3.61 (d, *J* = 9.6 Hz, 1H$_\alpha$), 1.14 – 0.82 (m, 28H).

^{13}C NMR (126 MHz, methanol-d$_4$): δ 129.75 – 128.65 (m), 96.43, 79.95 (d, *J* = 6.5 Hz), 76.53, 73.57, 71.13 – 70.80 (m), 70.33 (d, *J* = 5.0 Hz), 61.82, 18.02 – 17.48 (m).

^{31}P NMR (202 MHz, methanol-d$_4$): -1.96, -2.03.

Benzyl 2-O-benzyloxycarbonyl-3-O-dibenzylphosphono-D-glucopyranoside

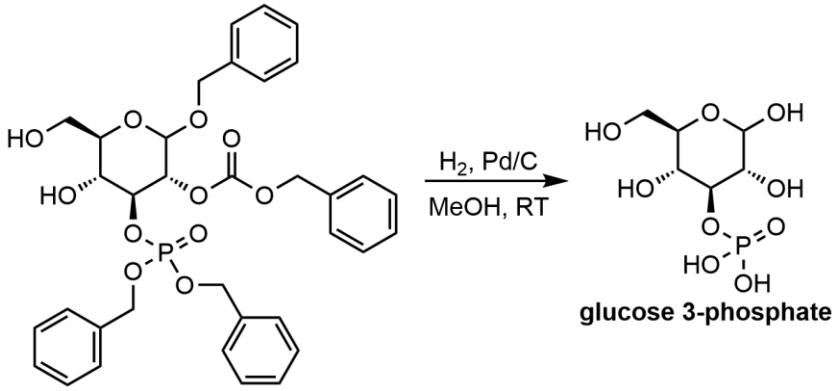


To a solution of benzyl 2-O-benzyloxycarbonyl-3-O-dibenzylphosphono-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside (29 mg, 0.025 mmol, 1 equiv.) in THF (2 mL) containing AcOH (5 µL, 0.09 mmol) at 0 ˚C was added TBAF (1M in THF, 0.09 mL, 0.09 mmol, 3.6 equiv.). After stirring for 2 hours cold (cooling with an ice-water bath), additional TBAF (0.03 mL, 0.03 mmol, 1.2 equiv) was added. After stirring for another 1.5 hours, additional AcOH (1.5 µL) was added, and the resulting solution stirred cold for an additional 17 hours up to room temp. until majority of starting material was consumed. An additional portion of AcOH was added (1.5 µL) and the solution was then concentrated to remove THF. Flash column chromatography on silica using a gradient of 2-8% MeOH in DCM was performed, affording benzyl 2-O-benzyloxycarbonyl-3-O-dibenzylphosphono-D-glucopyranoside (10 mg, 48%). Due to the many overlapping NMR signals, the minor anomeric peaks are not reported.

^{1}H NMR (500 MHz, methanol-d$_4$): δ 7.35 – 7.22 (m, 20H), 5.18 (d, *J* = 3.6 Hz, 1H$_\alpha$), 5.11 – 4.92 (m, 6H), 4.80 (dd, *J* = 17.8, 8.3 Hz, 1H$_\alpha$), 4.76 (d, *J* = 12.1 Hz, 1H$_\alpha$), 4.70 (dd, *J* = 10.0, 3.6 Hz, 1H$_\alpha$), 4.54 (d, *J* = 12.1 Hz, 1H$_\alpha$), 3.87 – 3.82 (m, 1H$_\alpha$), 3.78 – 3.72 (m, 2H$_\alpha$), 3.71 – 3.65 (m, 1H$_\alpha$).

^{31}P NMR (202 MHz, methanol-d$_4$): -2.21, -2.32.

3-O-phosphono-D-glucopyranose



glucose 3-phosphate

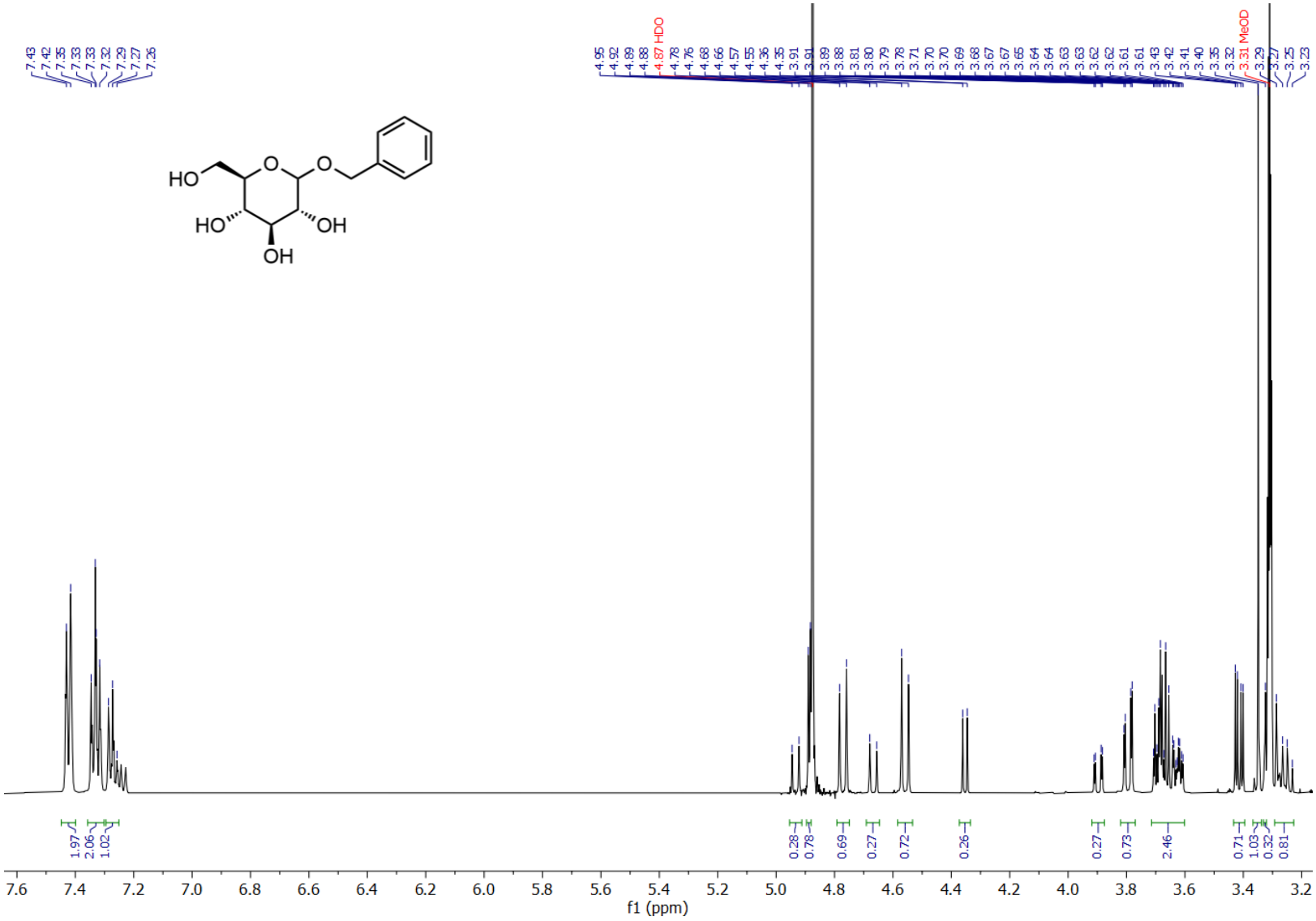
To a solution of benzyl 2-O-benzyloxycarbonyl-3-O-dibenzylphosphono-D-glucopyranoside (10 mg, 0.015 mmol) in DCM (2 mL) was added Pd/C (10 mg, 0.09 mmol). The reaction mixture was purged with argon for 10 minutes. then H_2 gas was passed over it for 90 minutes at room temperature, and the reaction vessel was again purged with argon for 10 minutes. The reaction mixture was filtered through celite and concentrated in vacuo, affording 3-O-phosphono-D-glucopyranose (~7 mg).

^{31}P NMR (202 MHz, deuterium oxide): δ 3.01, 2.93, 1.38, 0.48, -0.02.

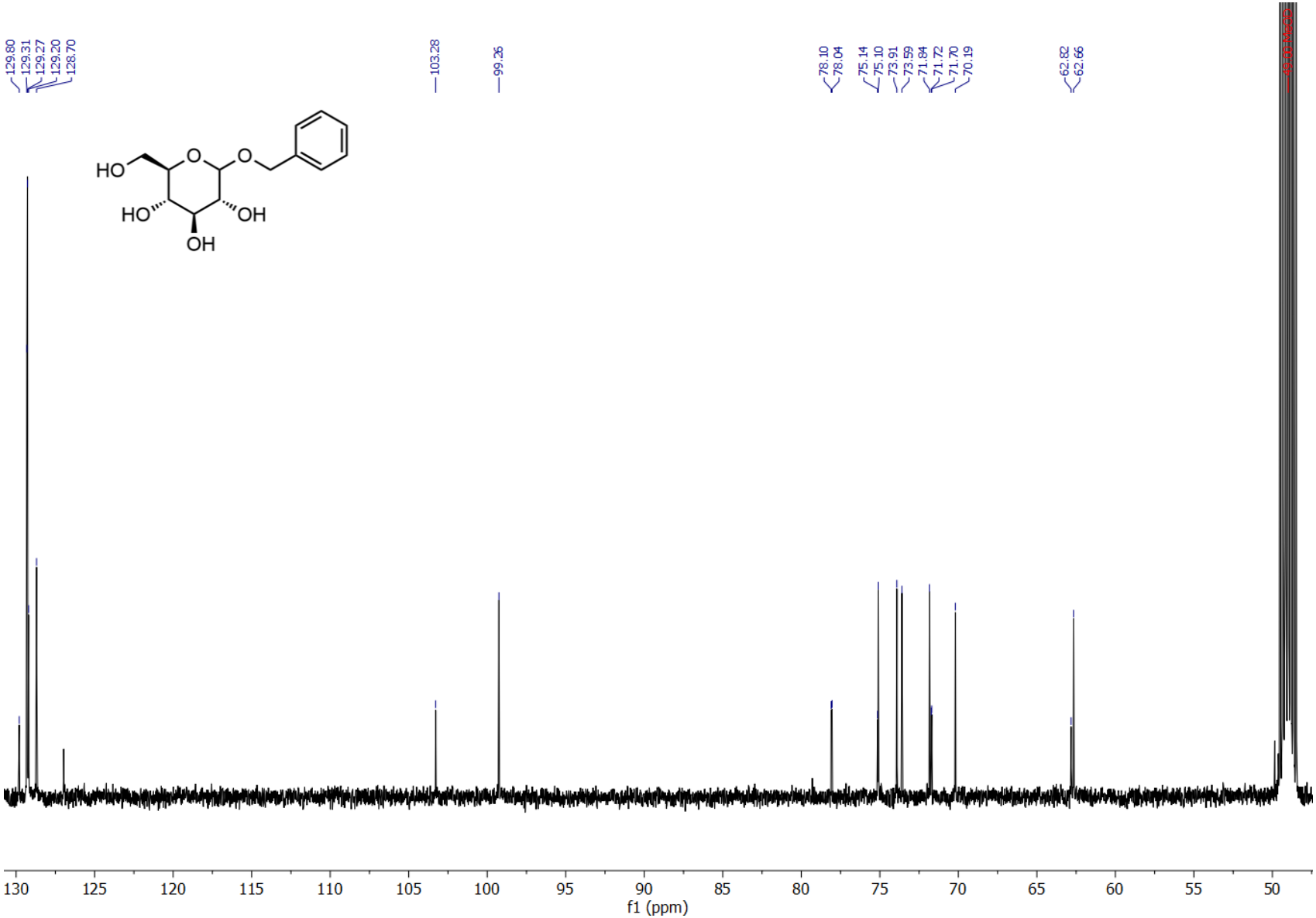
^{1}H NMR (600 MHz, deuterium oxide) and ^{13}C NMR (126 MHz, deuterium oxide):

Position	δ ^{13}C (ppm)	δ ^{1}H (ppm)	Multiplicity	J_{HH} (Hz)
1_α	92.10	5.26	d	$J_{1,2}$ = 3.4
2_α	70.90	3.68	d	$J_{2,3}$ = 8.2
3_α	79.00	4.27	q	$J_{3,2} = J_{3,4} = J_{3,P}$ = 8.6
4_α	69.10	3.59	m (overlapping)	—
5_α	71.10	3.88	m (overlapping)	—
$6a_\alpha$	60.60	3.78	dd	$J_{6a,6b}$ = 12.4, $J_{6a,5}$ = 4.3
$6b_\alpha$		3.84	d	$J_{6b,6a}$ = 12.2
1_β	95.70	4.70	d	$J_{1,2}$ = 7.8
2_β	73.60	3.40	t	$J_{2,1} = J_{2,3}$ = 8.8
3_β	81.30	4.07	q	$J_{3,2} = J_{3,4} = J_{3,P}$ = 8.5
4_β	69.10	3.58	m (overlapping)	—
5_β	75.60	3.51	m	—
$6a_\beta$	60.80	3.74	dd	$J_{6a,6b}$ = 12.2, $J_{6a,5}$ = 5.2
$6b_\beta$		3.90	d	$J_{6b,6a}$ = 12.3

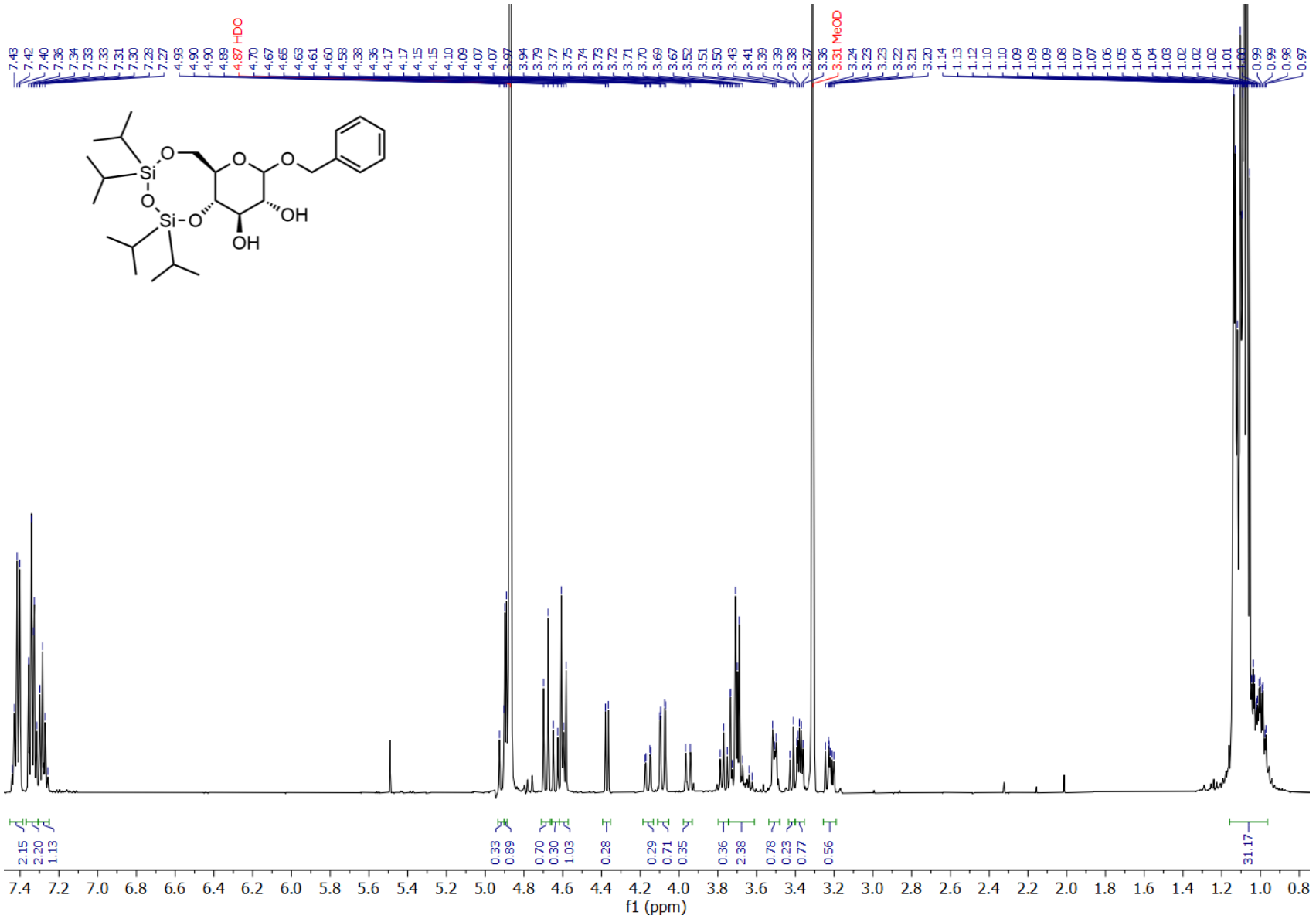
APPENDIX II: NMR SPECTRA



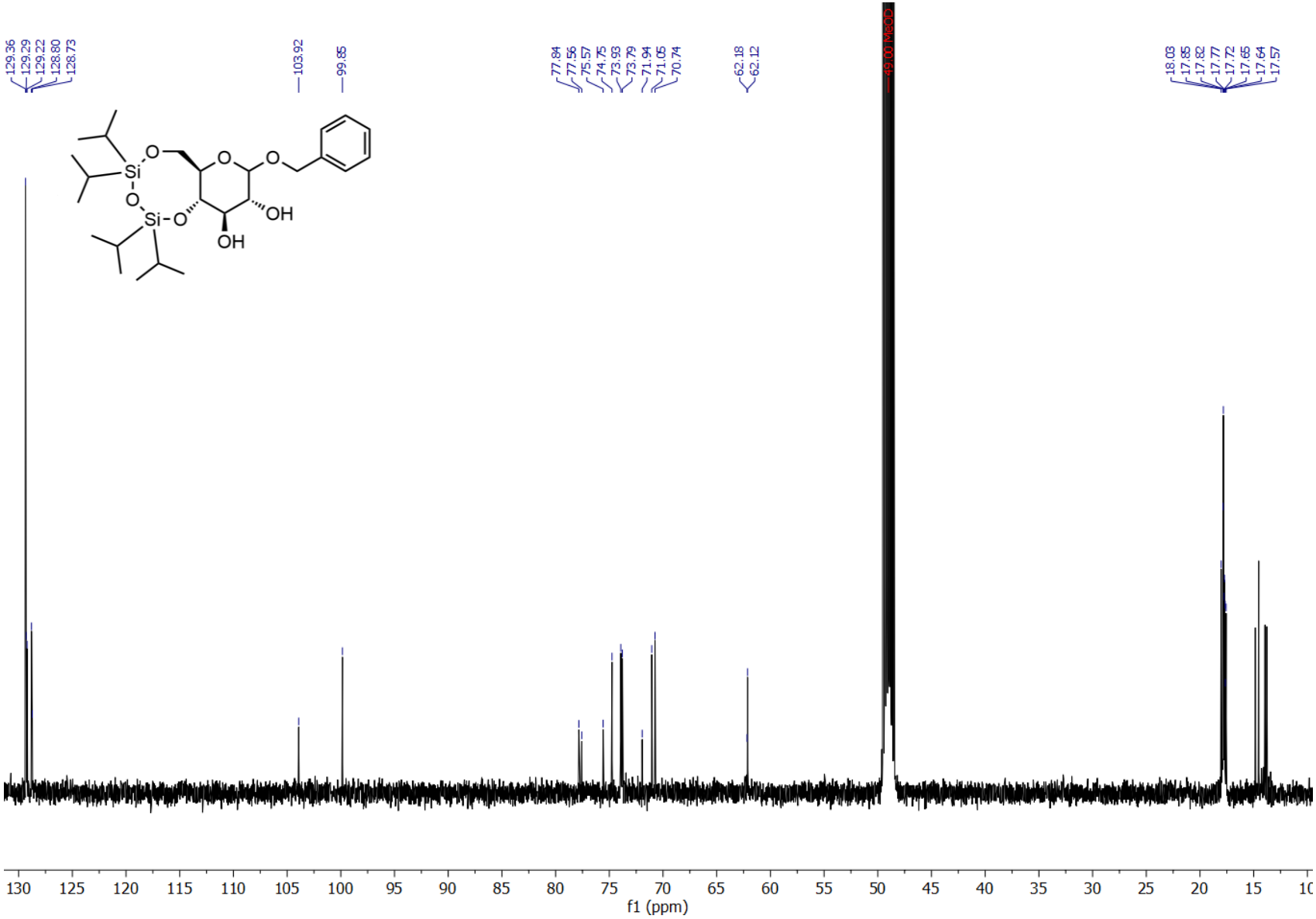
^{1}H NMR spectrum (500 MHz) of benzyl D-glucopyranoside in methanol-d$_4$.



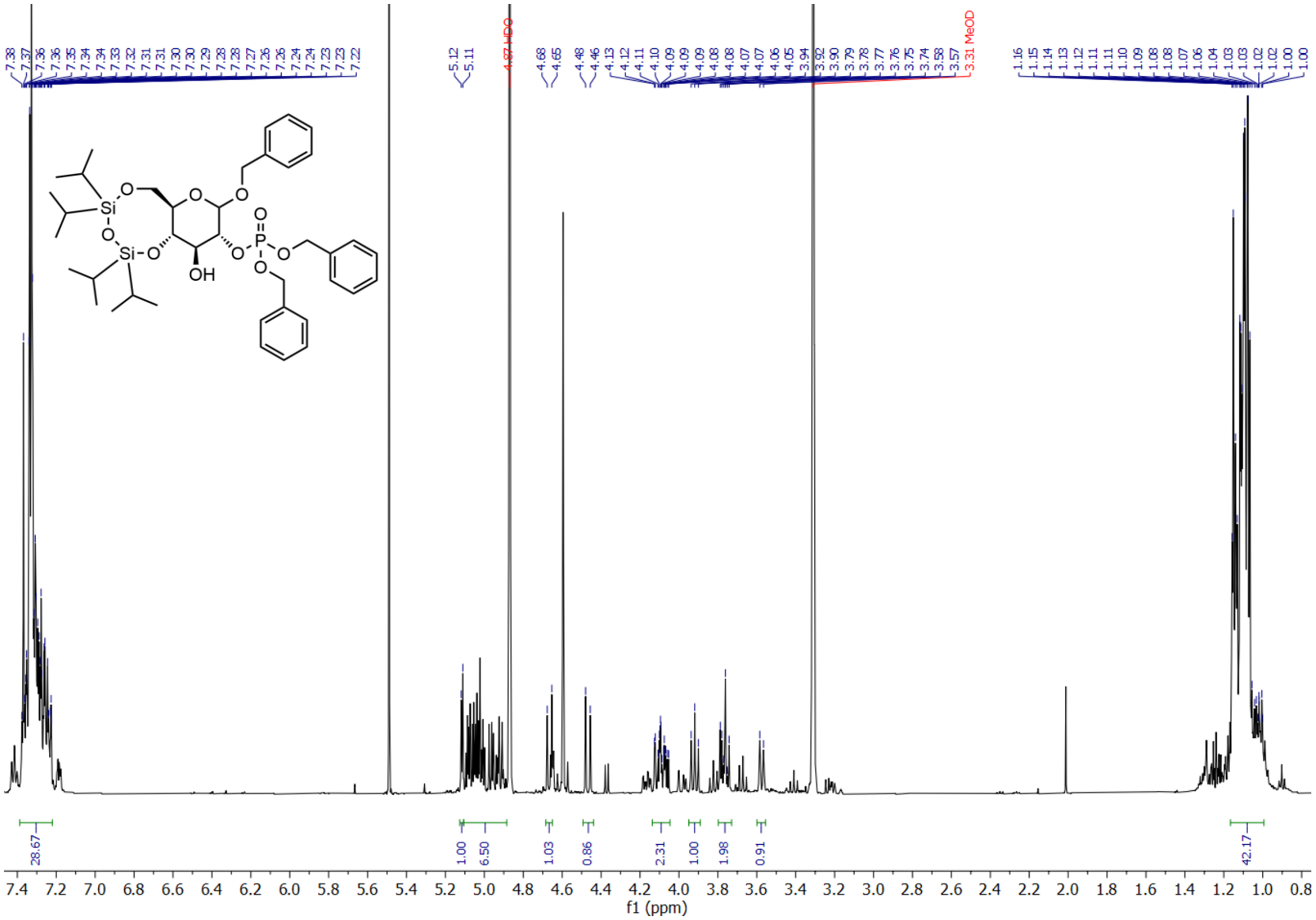
^{13}C NMR spectrum (126 MHz) of benzyl D-glucopyranoside in methanol-d_4.



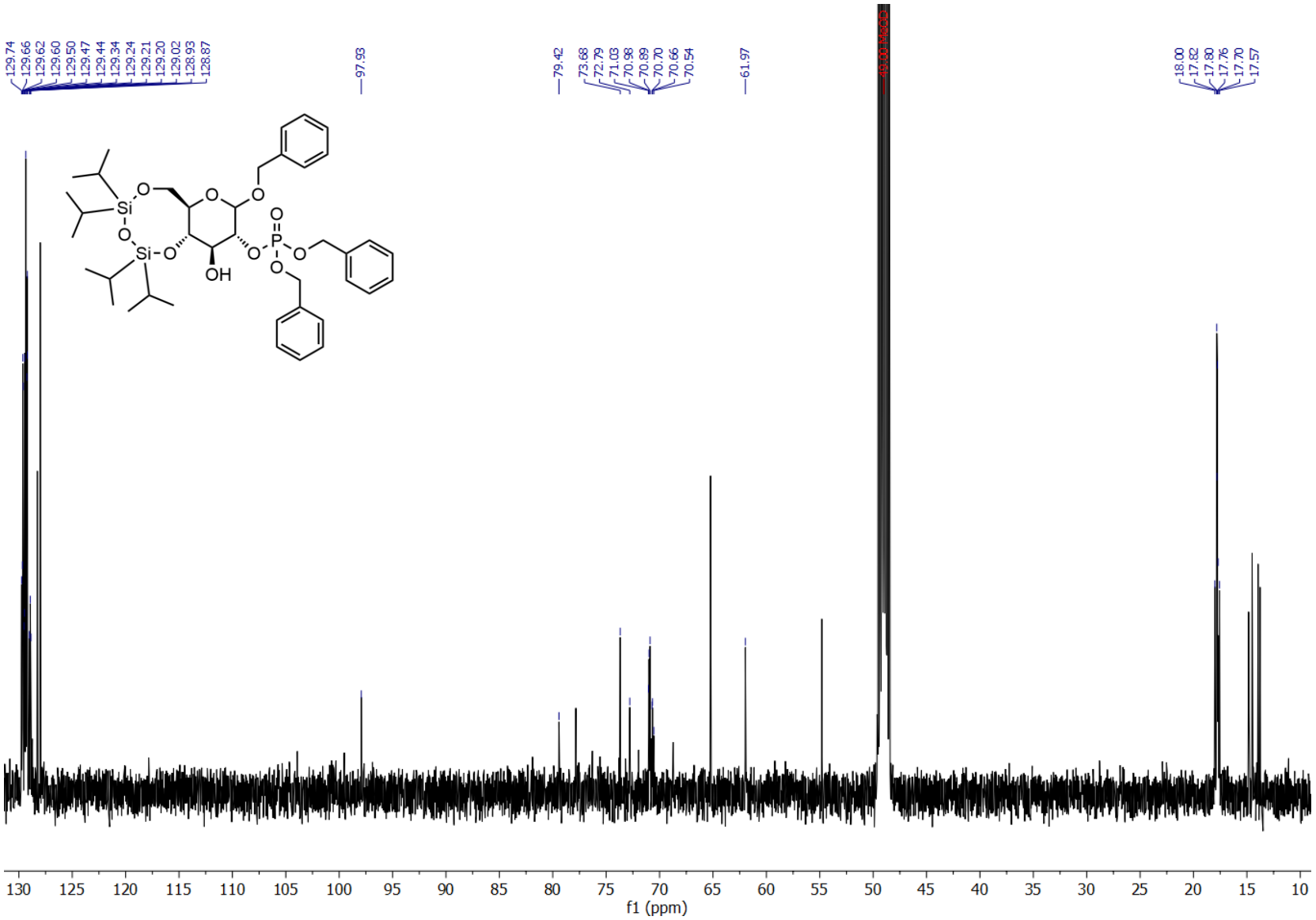
^{1}H NMR spectrum (500 MHz) of benzyl 4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside in methanol-d_4.



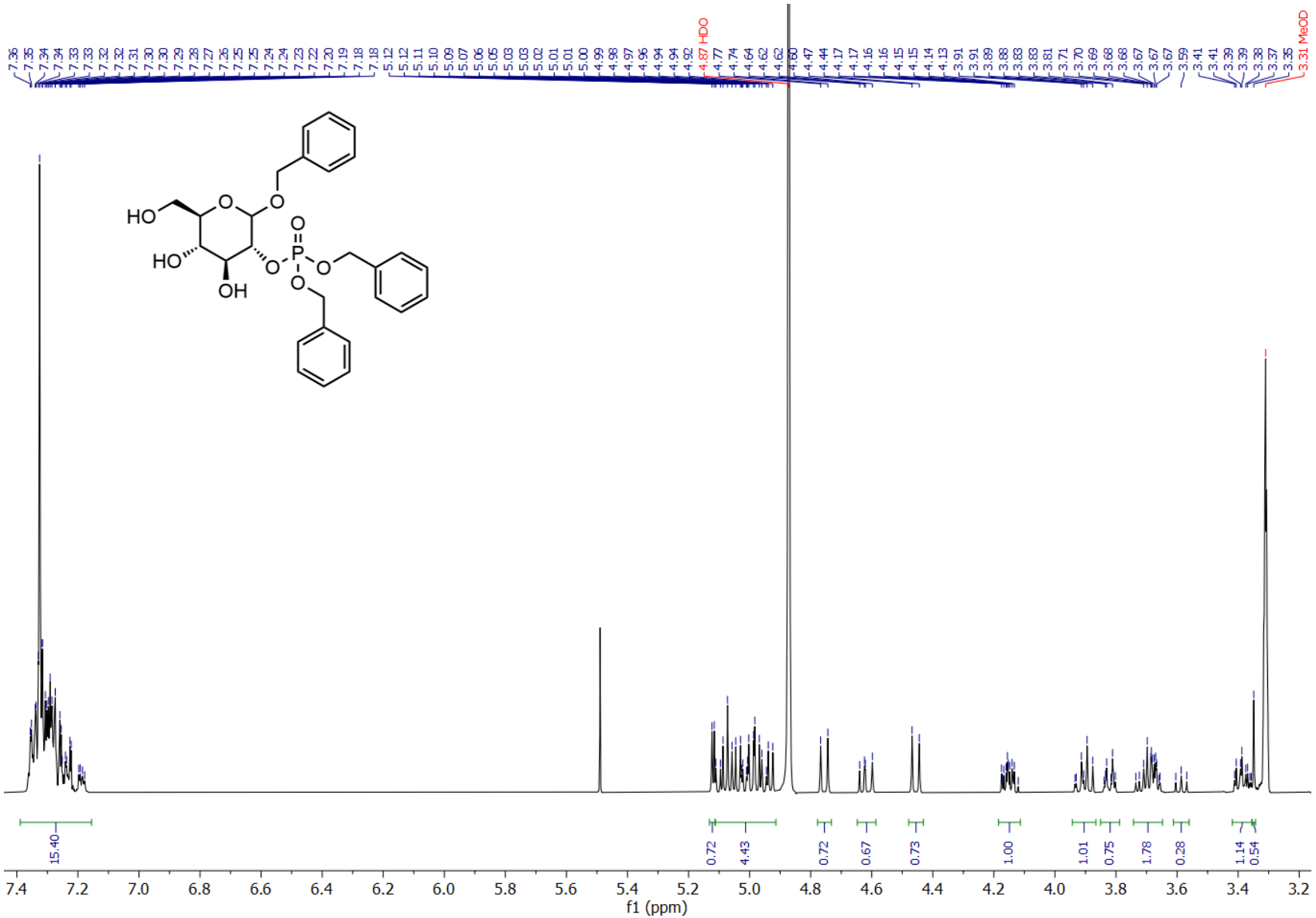
^{13}C NMR spectrum (126 MHz) of benzyl 4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside in methanol-d_4.



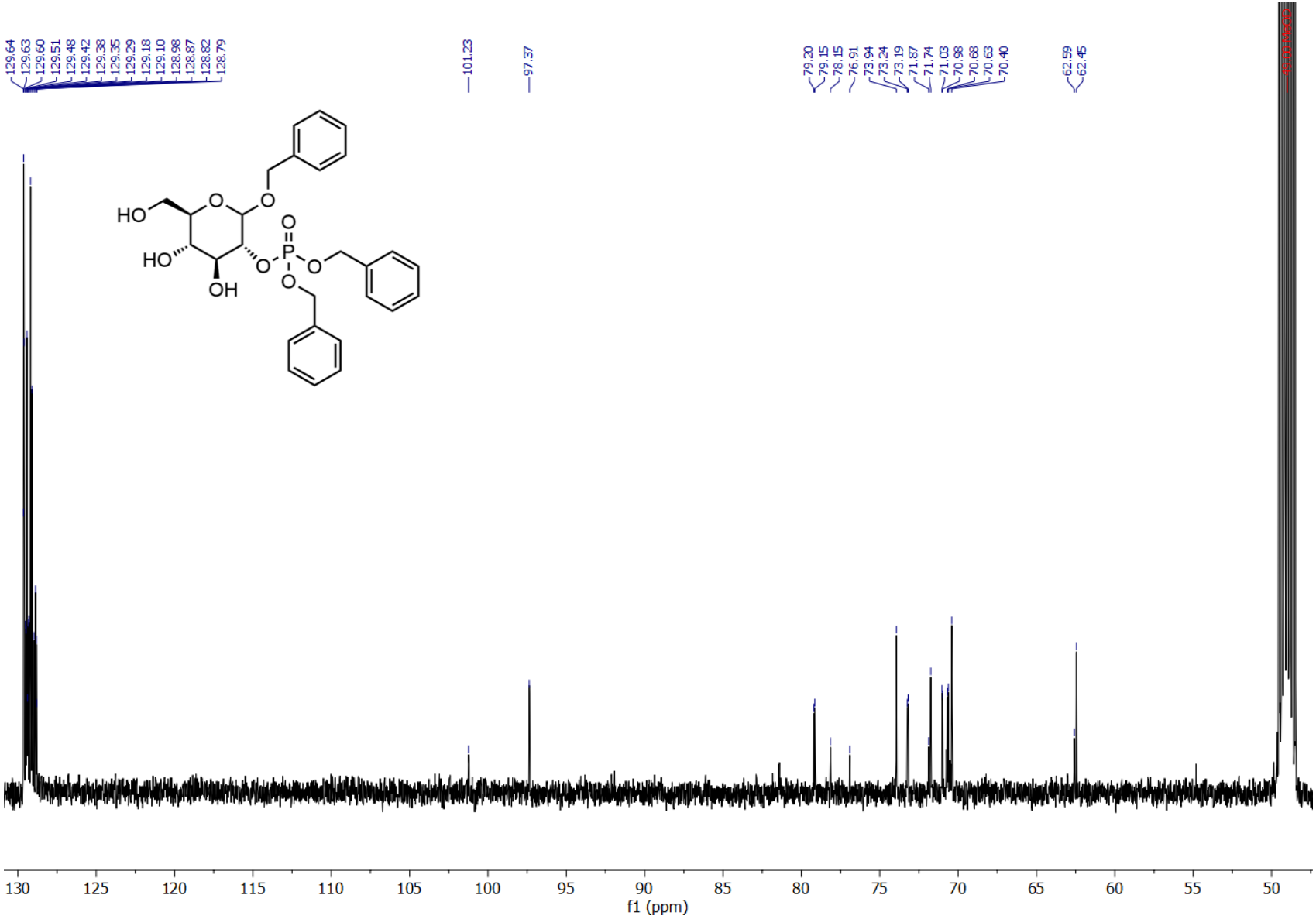
^{1}H NMR spectrum (500 MHz) of benzyl 2-O-dibenzylphosphono-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside in methanol-d_4.



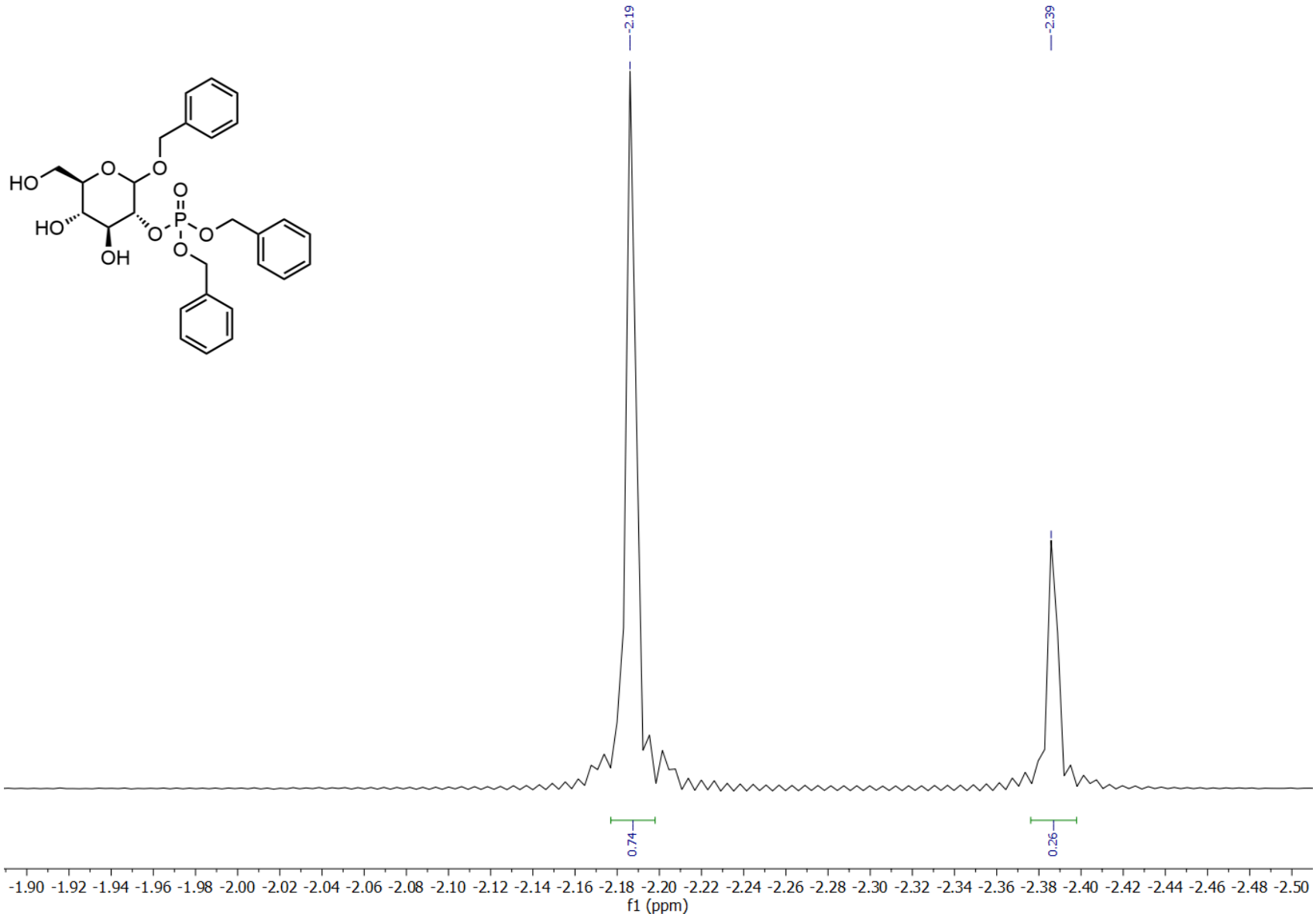
^{13}C NMR spectrum (126 MHz) of benzyl 2-O-dibenzylphosphono-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside in methanol-d_4.



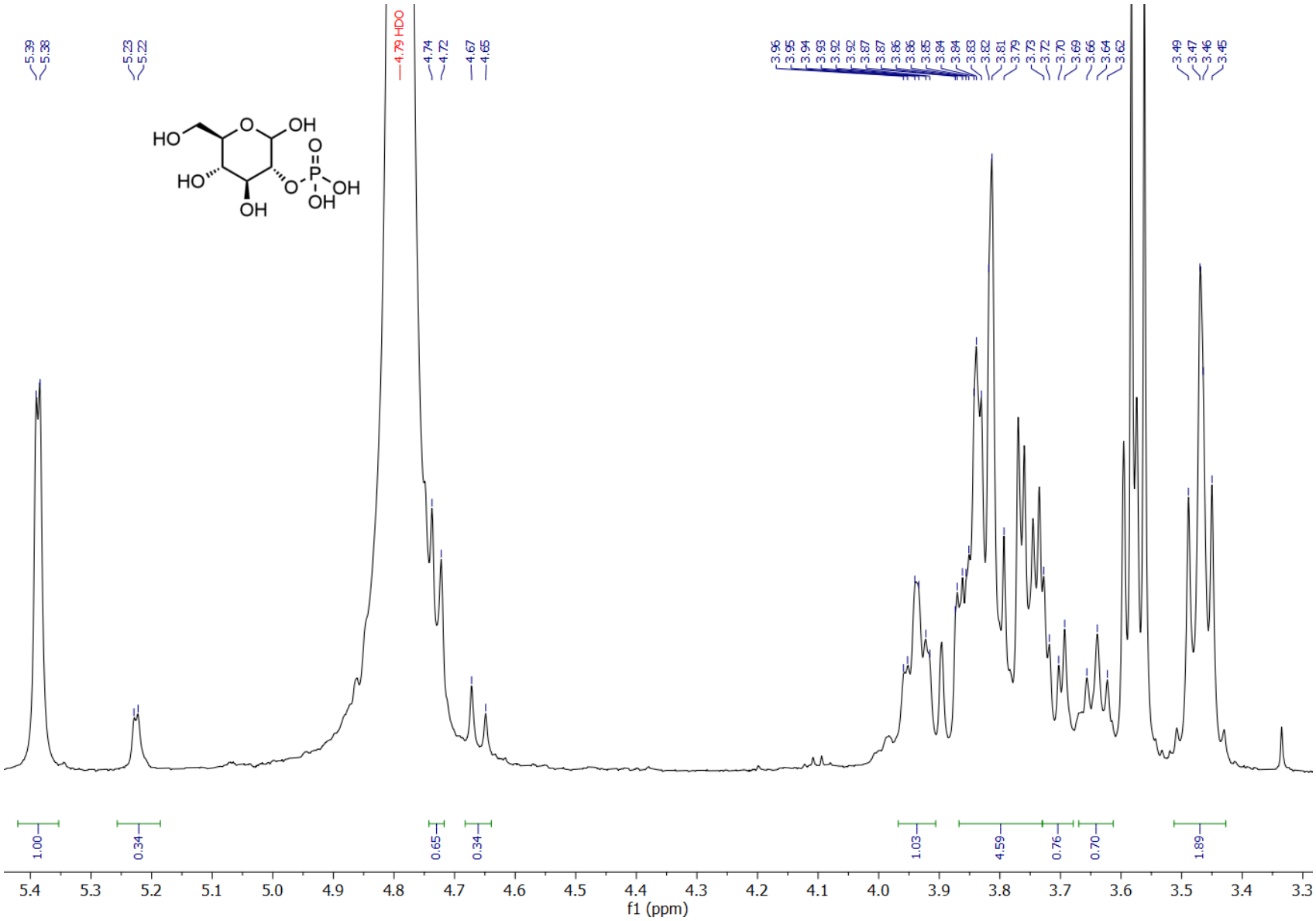
^{1}H NMR spectrum (500 MHz) of benzyl 2-O-dibenzylphosphono-D-glucopyranoside in methanol-d_4.



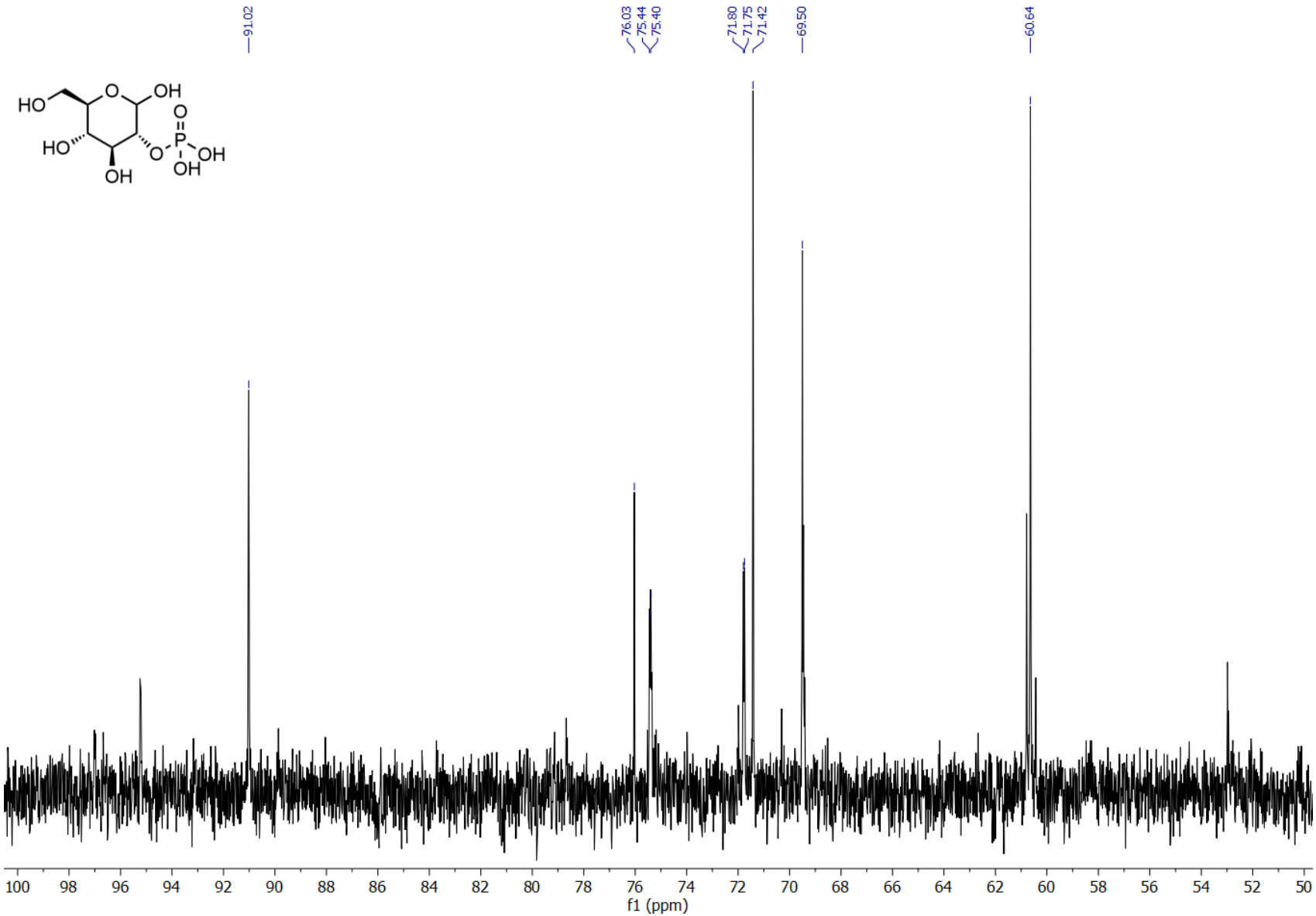
^{13}C NMR spectrum (126 MHz) of benzyl 2-O-dibenzylphosphono-D-glucopyranoside in methanol-d_4.



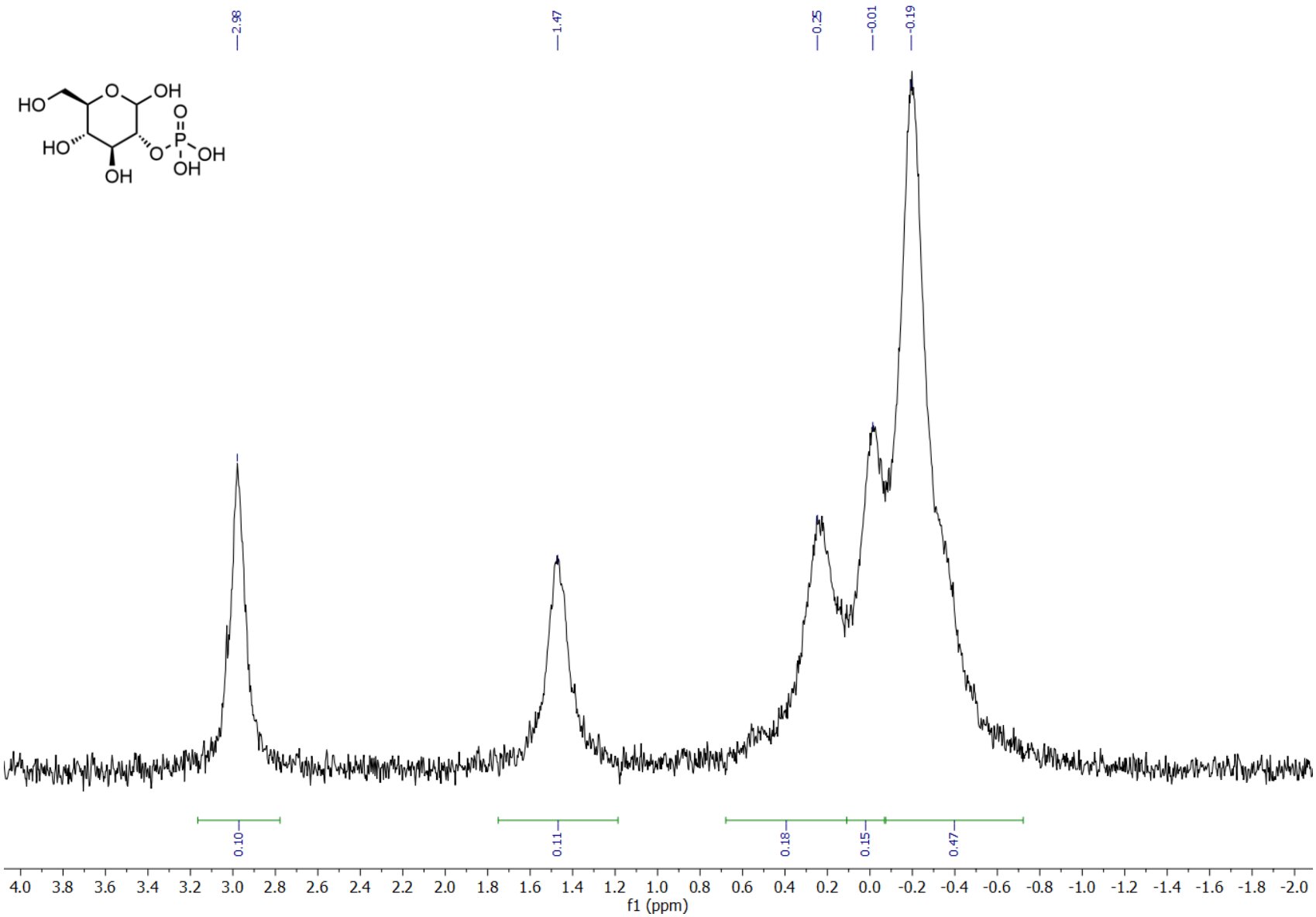
^{31}P NMR spectrum (202 MHz) of benzyl 2-O-dibenzylphosphono-D-glucopyranoside in methanol-d_4.



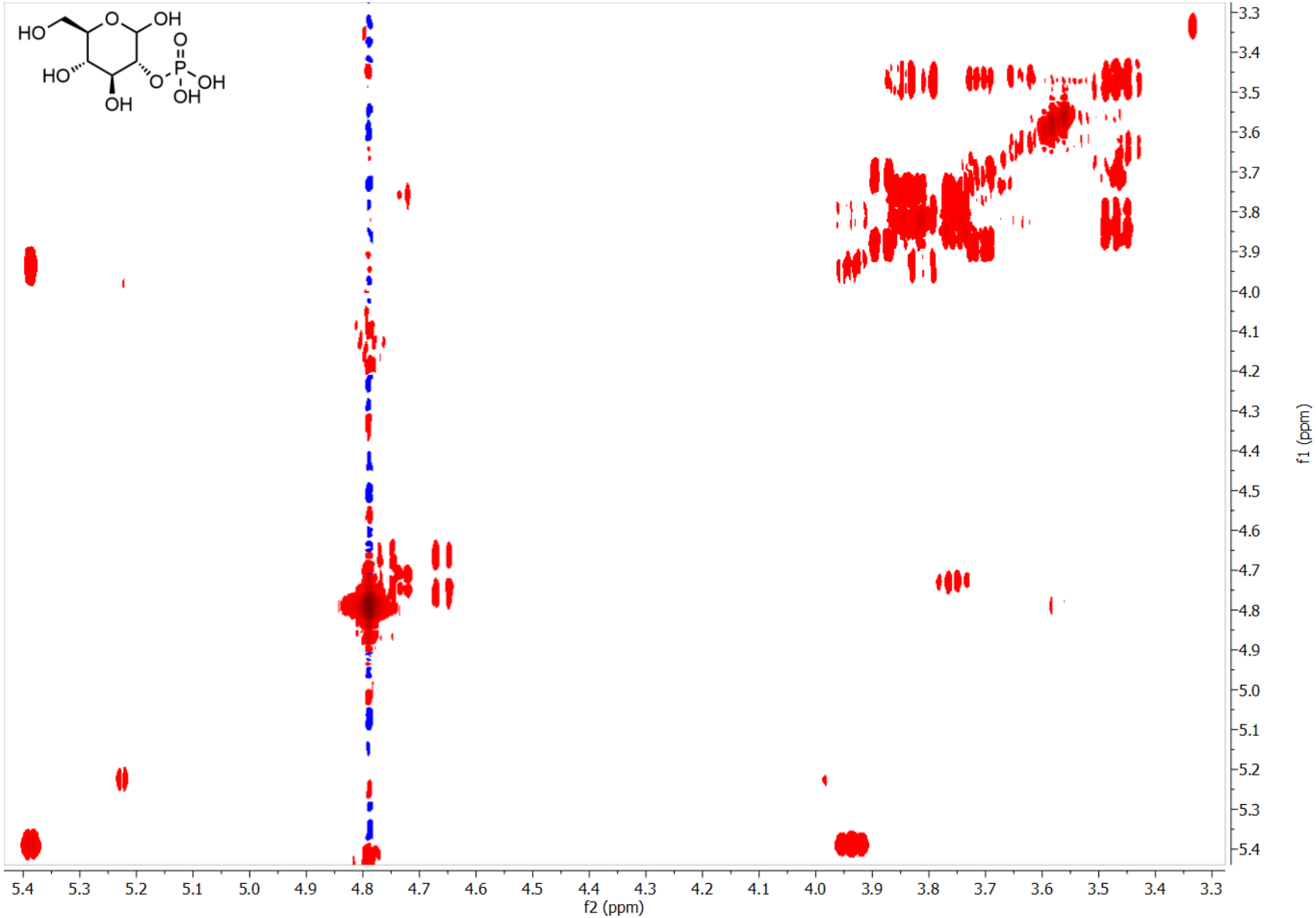
^{1}H NMR spectrum (500 MHz) of 2-O-phosphono-D-glucopyranose in deuterium oxide.



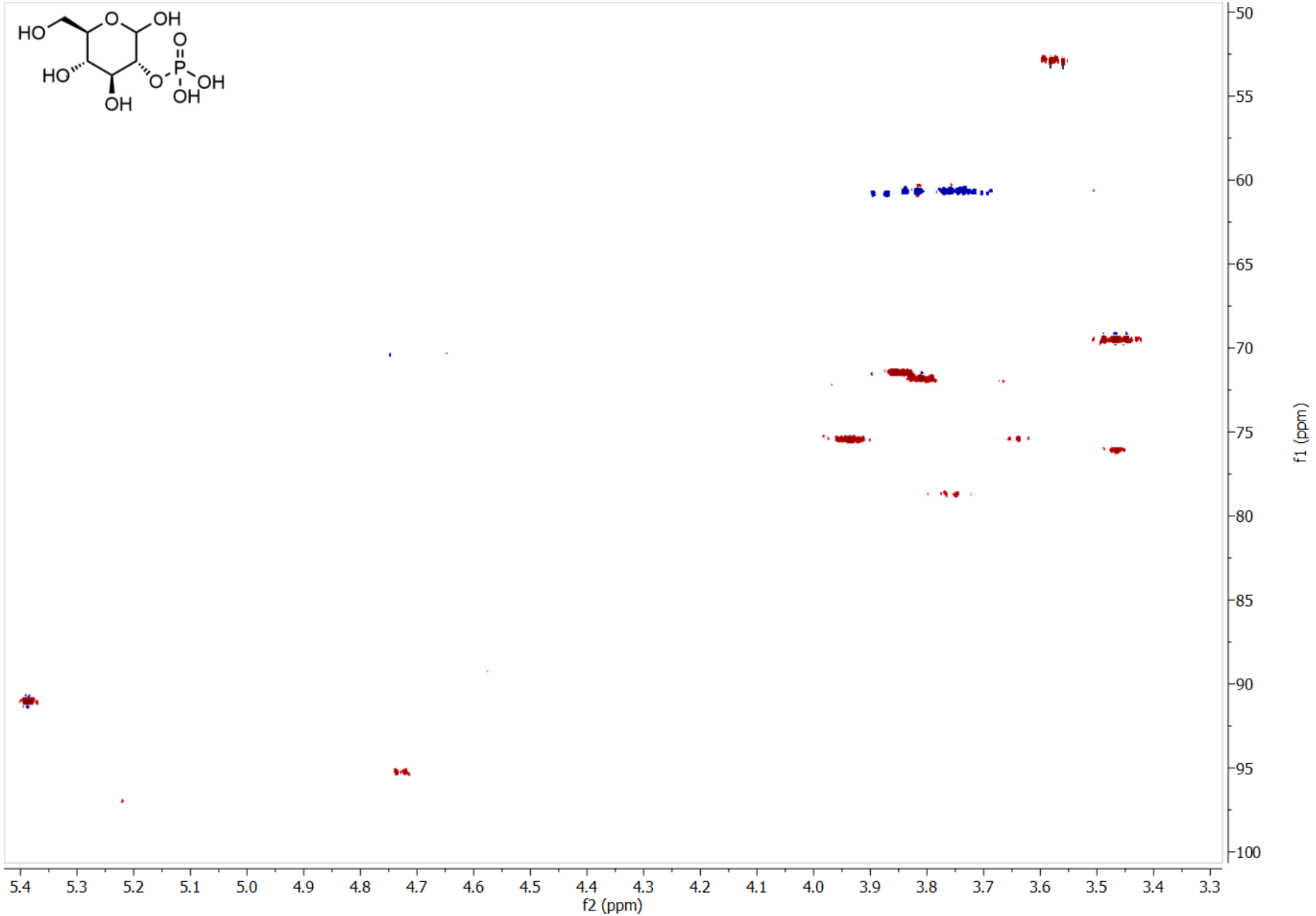
^{13}C NMR spectrum (126 MHz) of 2-O-phosphono-D-glucopyranose in deuterium oxide.



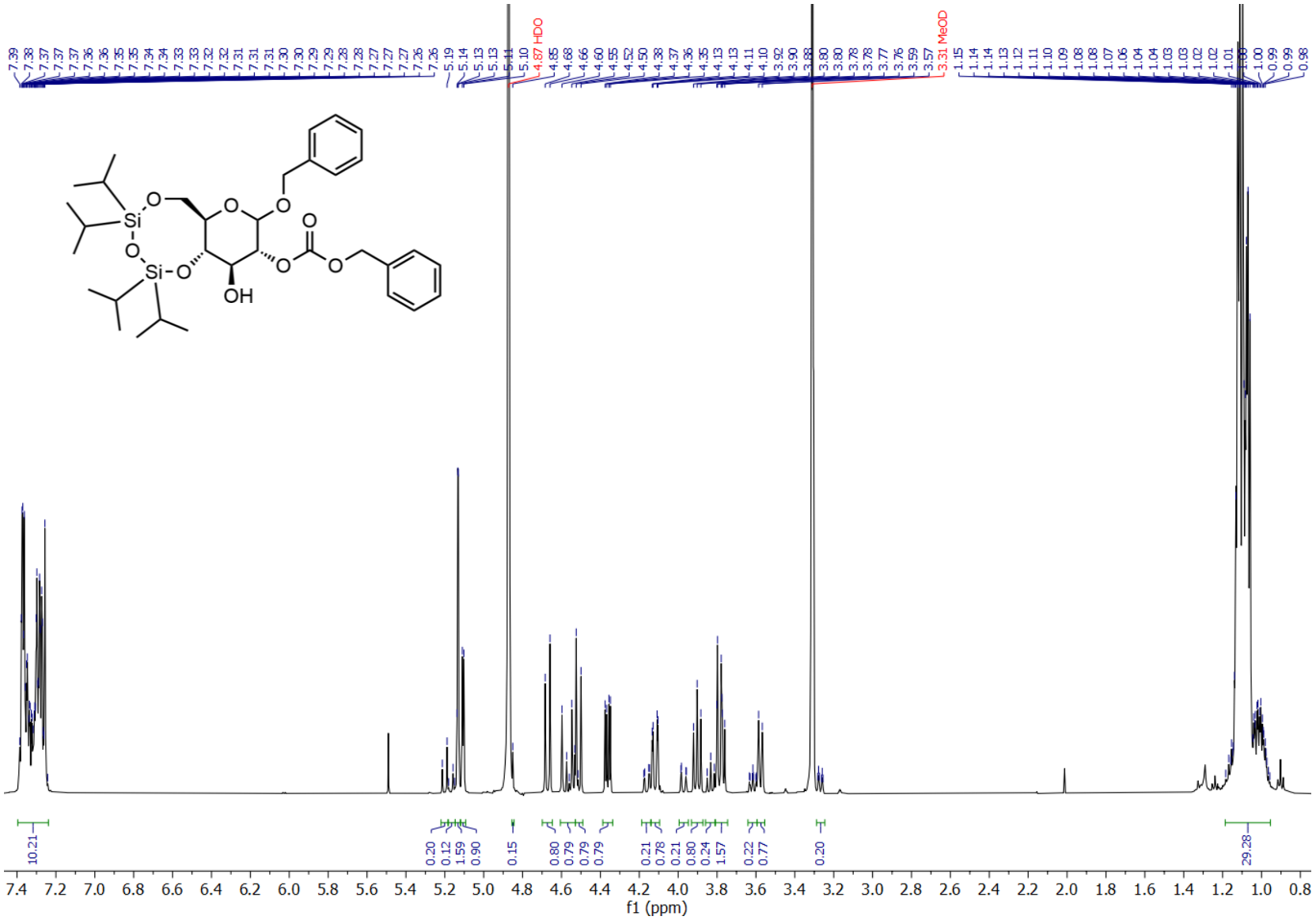
^{31}P NMR spectrum (202 MHz) of 2-O-phosphono-D-glucopyranose in deuterium oxide.



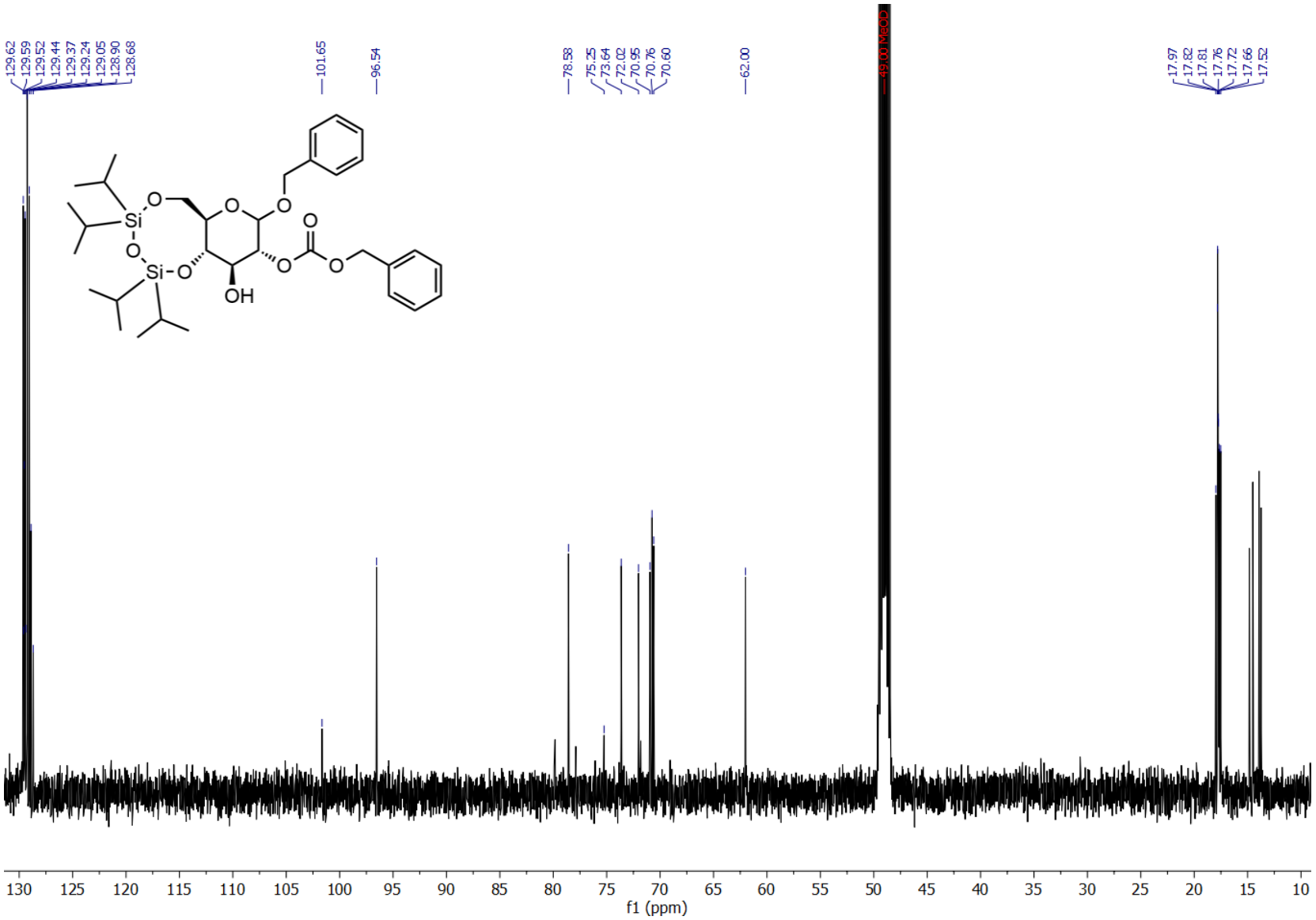
gCOSY NMR spectrum (500 MHz) of 2-O-phosphono-D-glucopyranose in deuterium oxide.



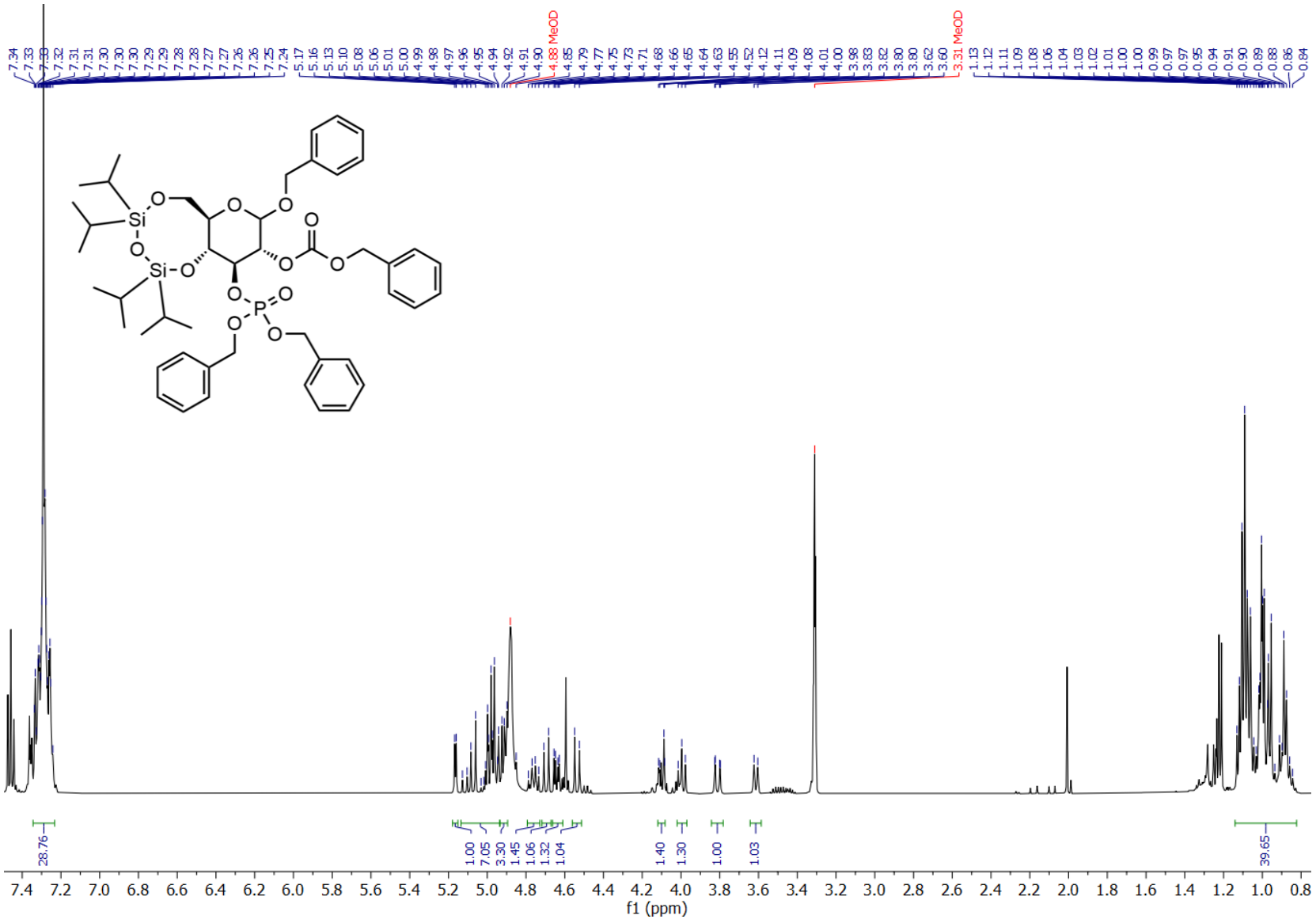
HSQC NMR spectrum (500 MHz) of 2-O-phosphono-D-glucopyranose in deuterium oxide.



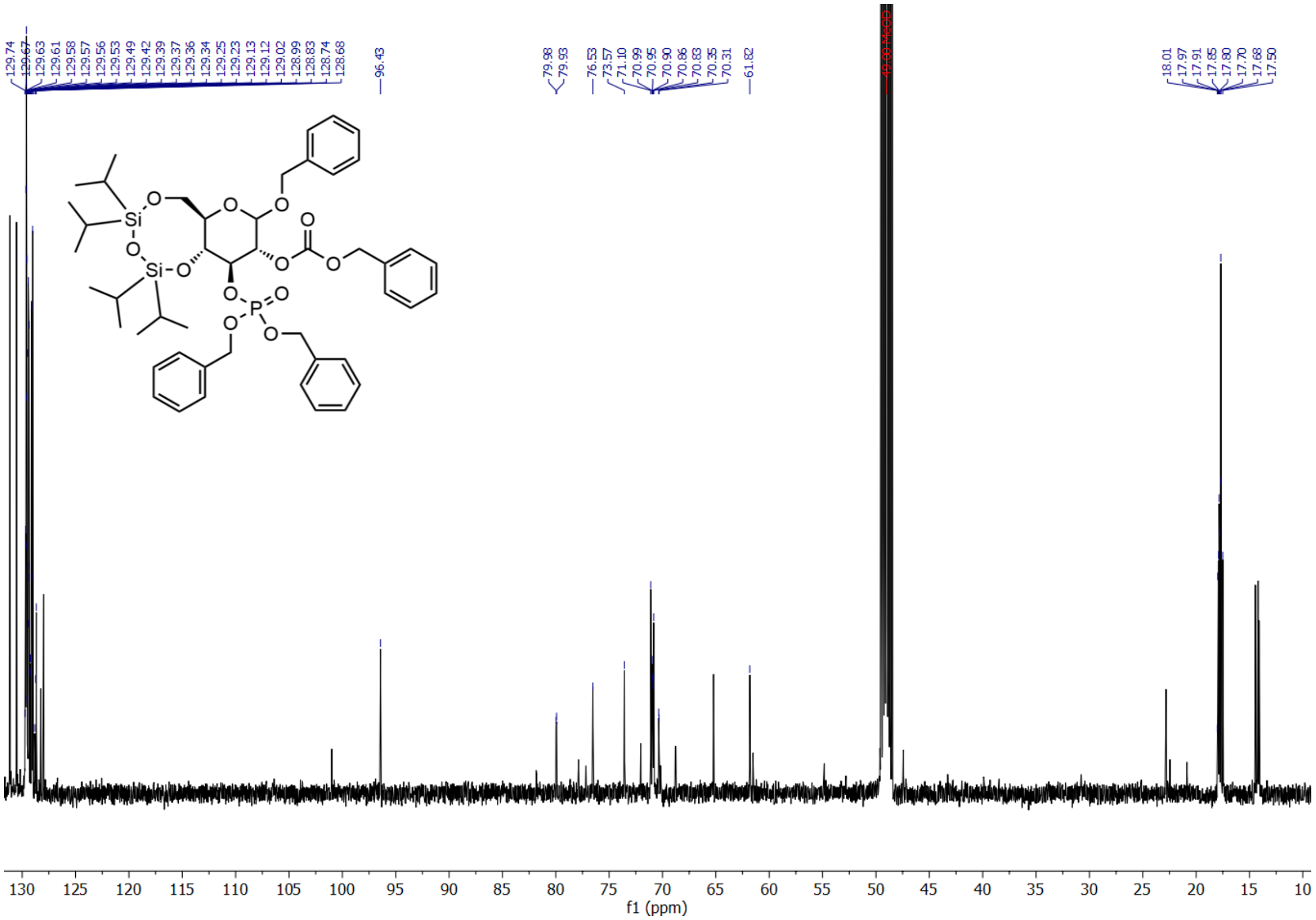
^{1}H NMR spectrum (500 MHz) of benzyl 2-O-benzyloxycarbonyl-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside in methanol-d_4.



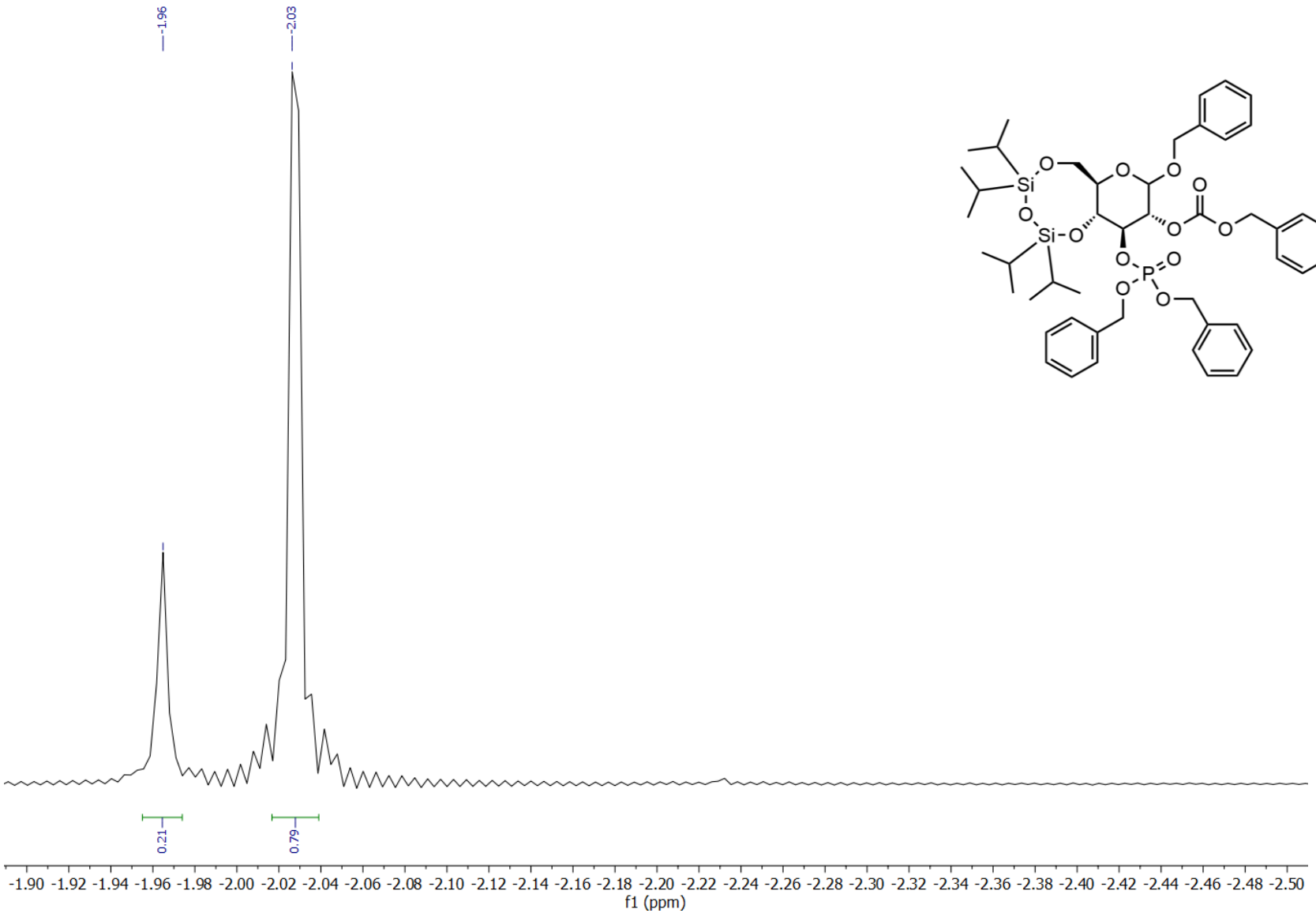
^{13}C NMR spectrum (126 MHz) of benzyl 2-O-benzyloxycarbonyl-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside in methanol-d_4.



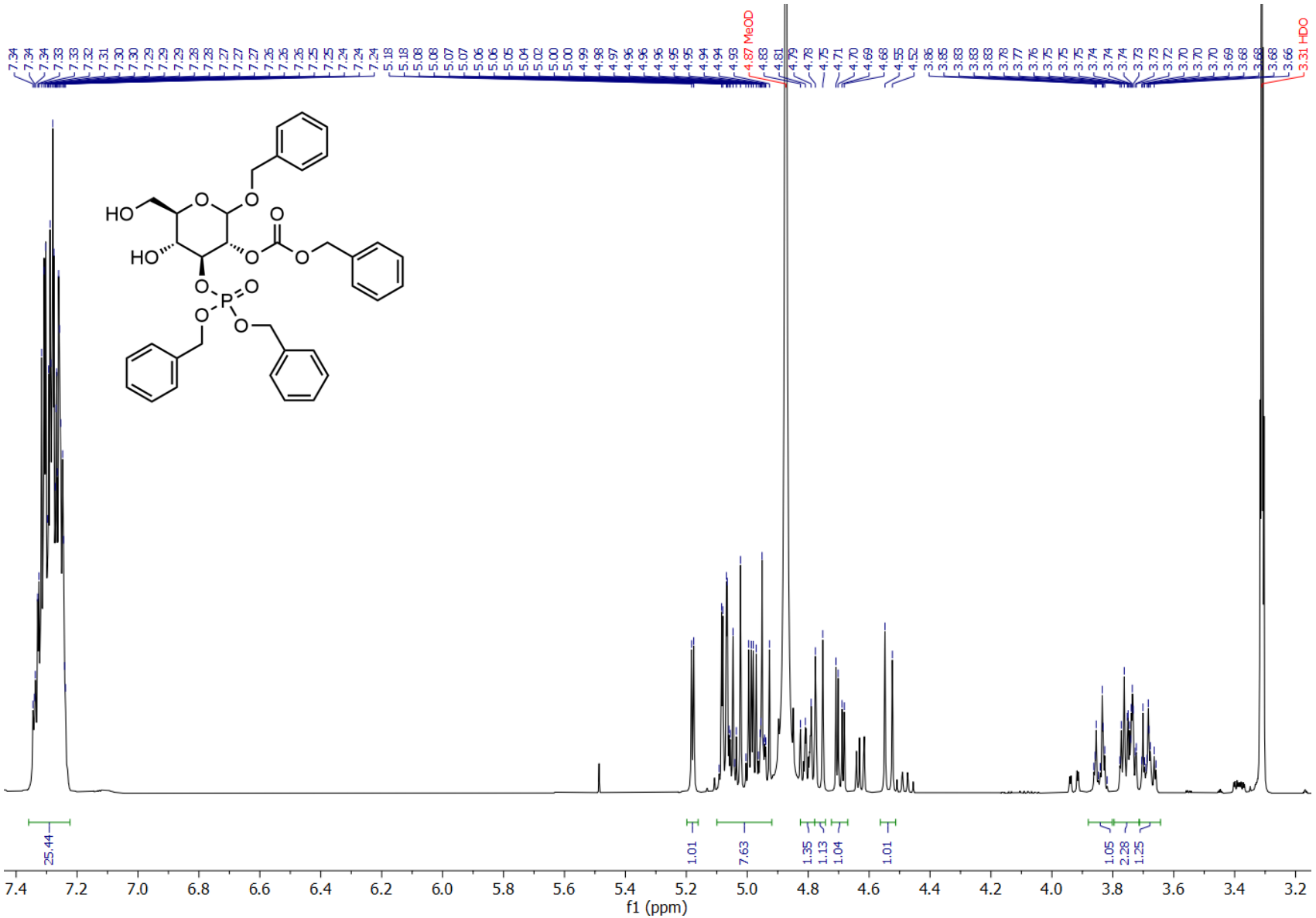
^{1}H NMR spectrum (500 MHz) of benzyl 2-O-benzyloxycarbonyl-3-O-dibenzylphosphono-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside in methanol-d_4.



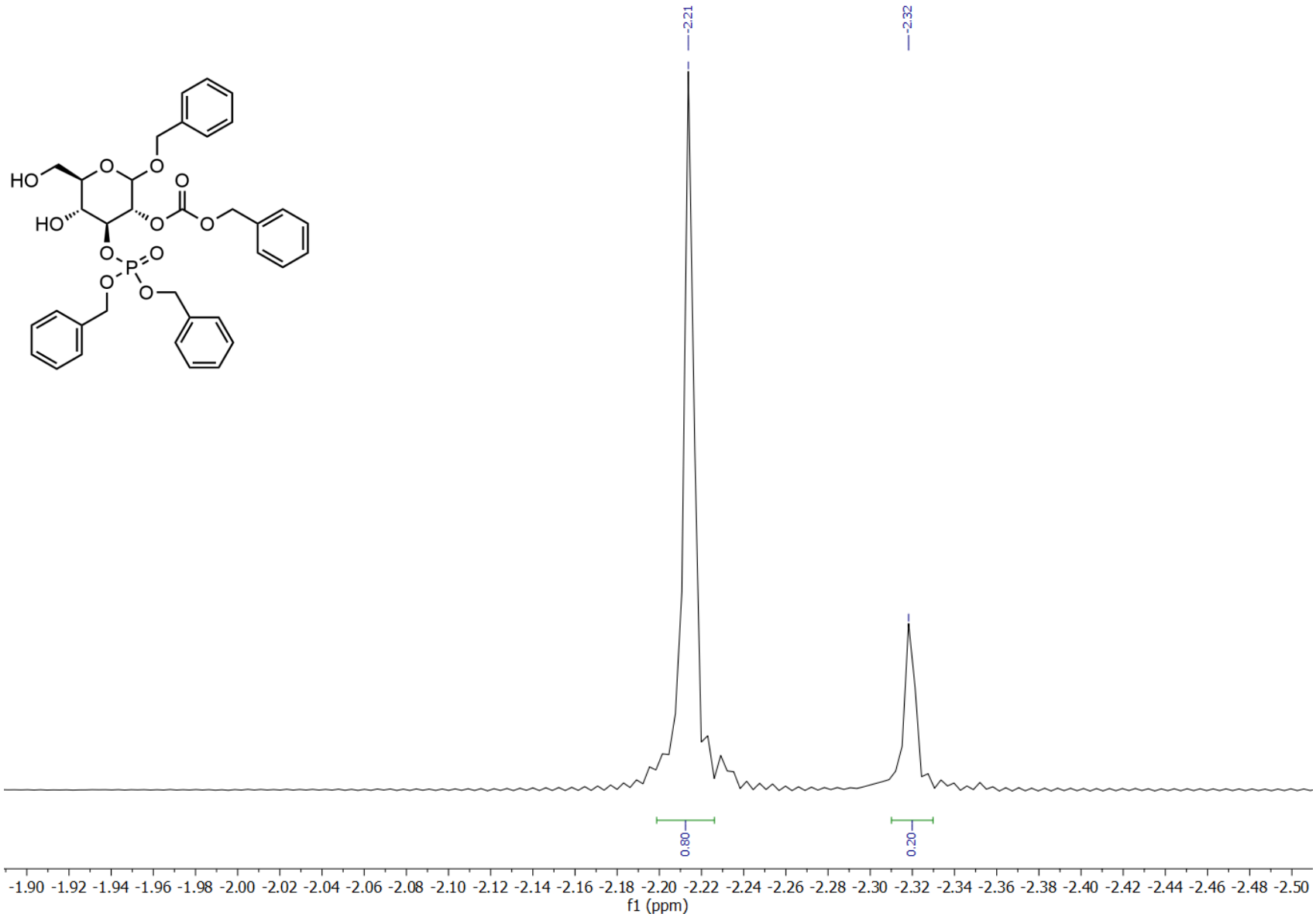
^{13}C NMR spectrum (126 MHz) of benzyl 2-O-benzyloxycarbonyl-3-O-dibenzylphosphono-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside in methanol-d_4.



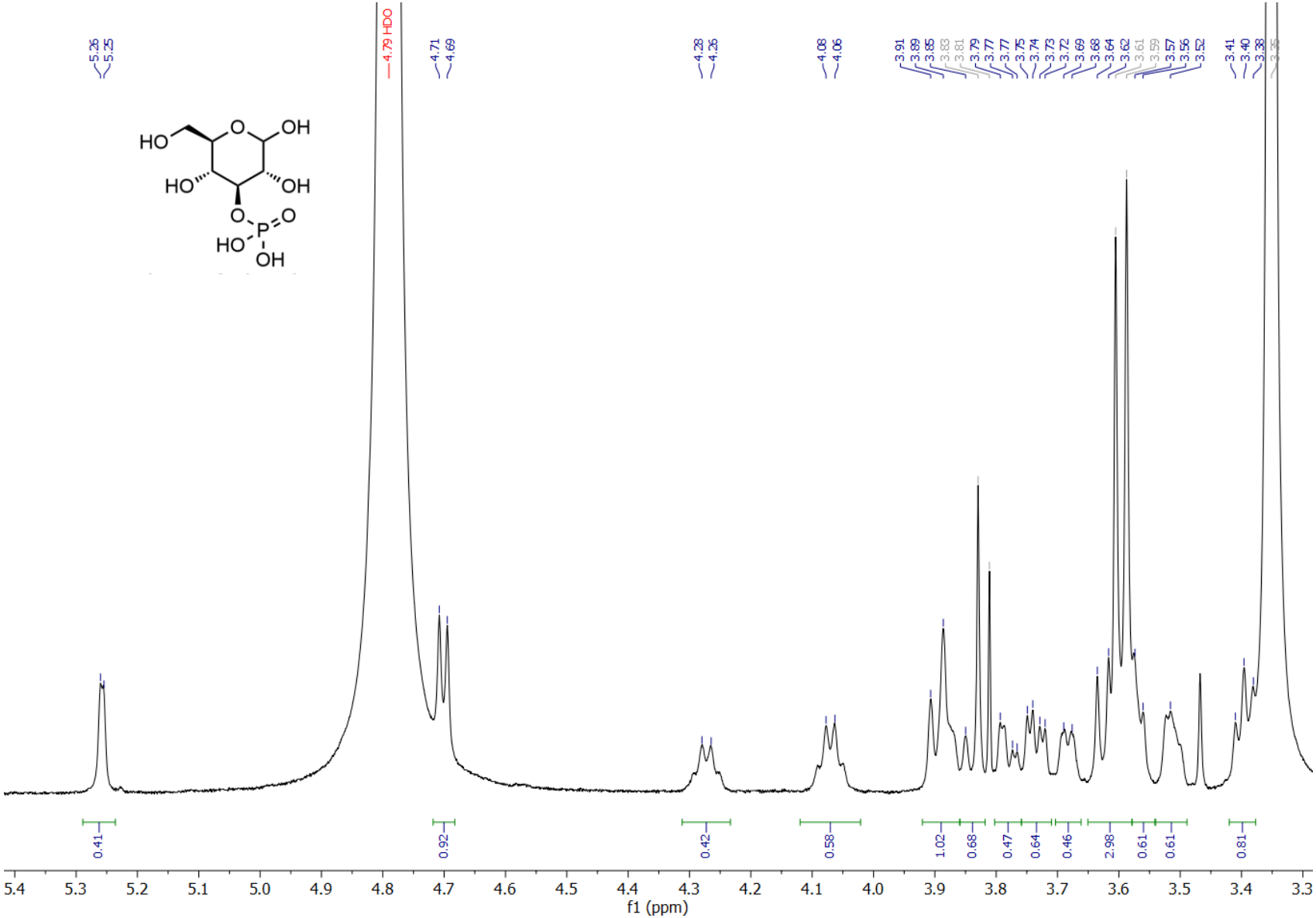
^{31}P NMR spectrum (202 MHz) of benzyl 2-O-benzyloxycarbonyl-3-O-dibenzylphosphono-4-O,6-O-(1,1,3,3-tetraisopropyldisiloxyl)-D-glucopyranoside in methanol-d_4.



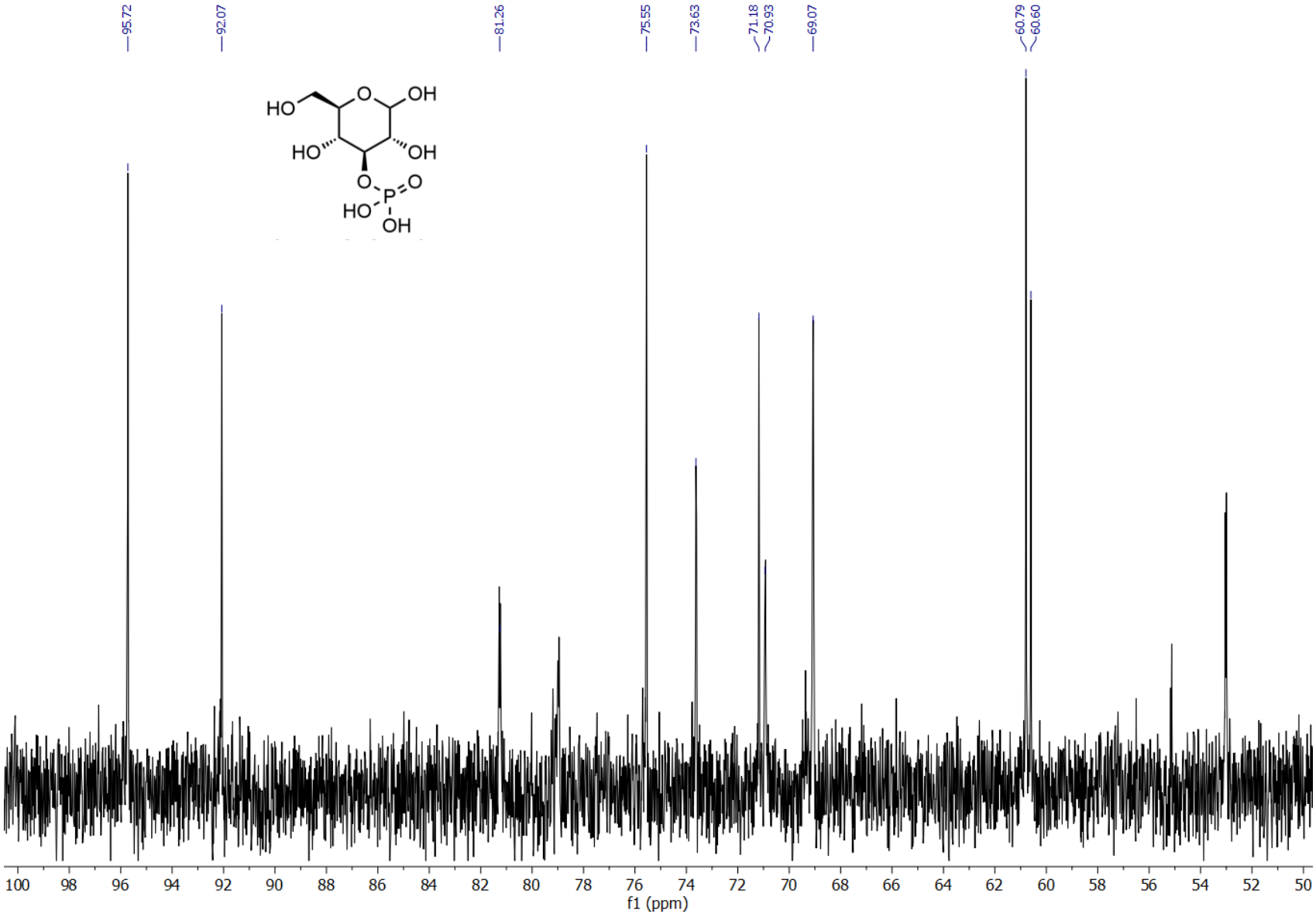
^{1}H NMR spectrum (500 MHz) of benzyl 2-O-benzyloxycarbonyl-3-O-dibenzylphosphono-D-glucopyranoside in methanol-d_4.



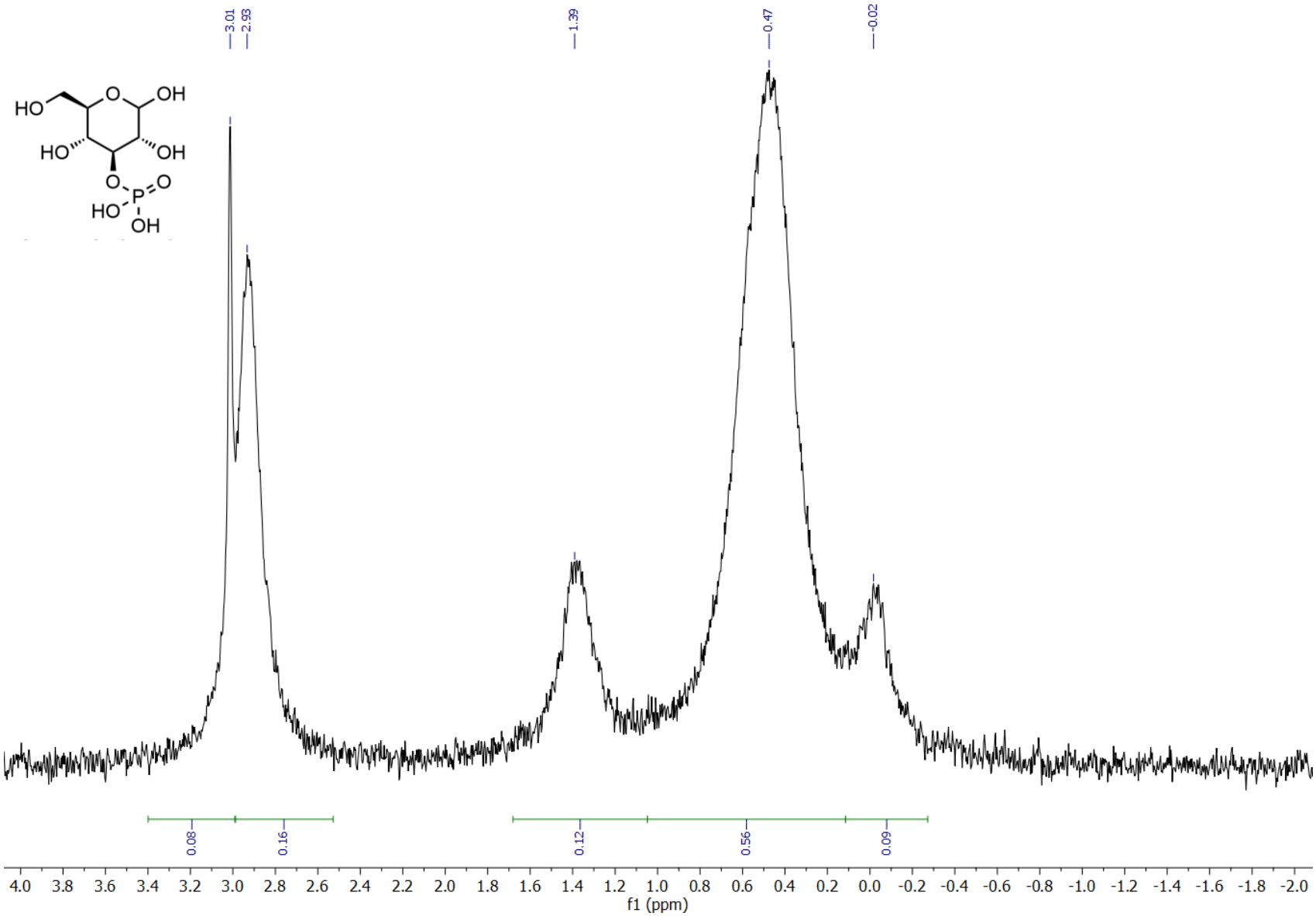
^{31}P NMR spectrum (202 MHz) of benzyl 2-O-benzyloxycarbonyl-3-O-dibenzylphosphono-D-glucopyranoside in methanol-d_4.



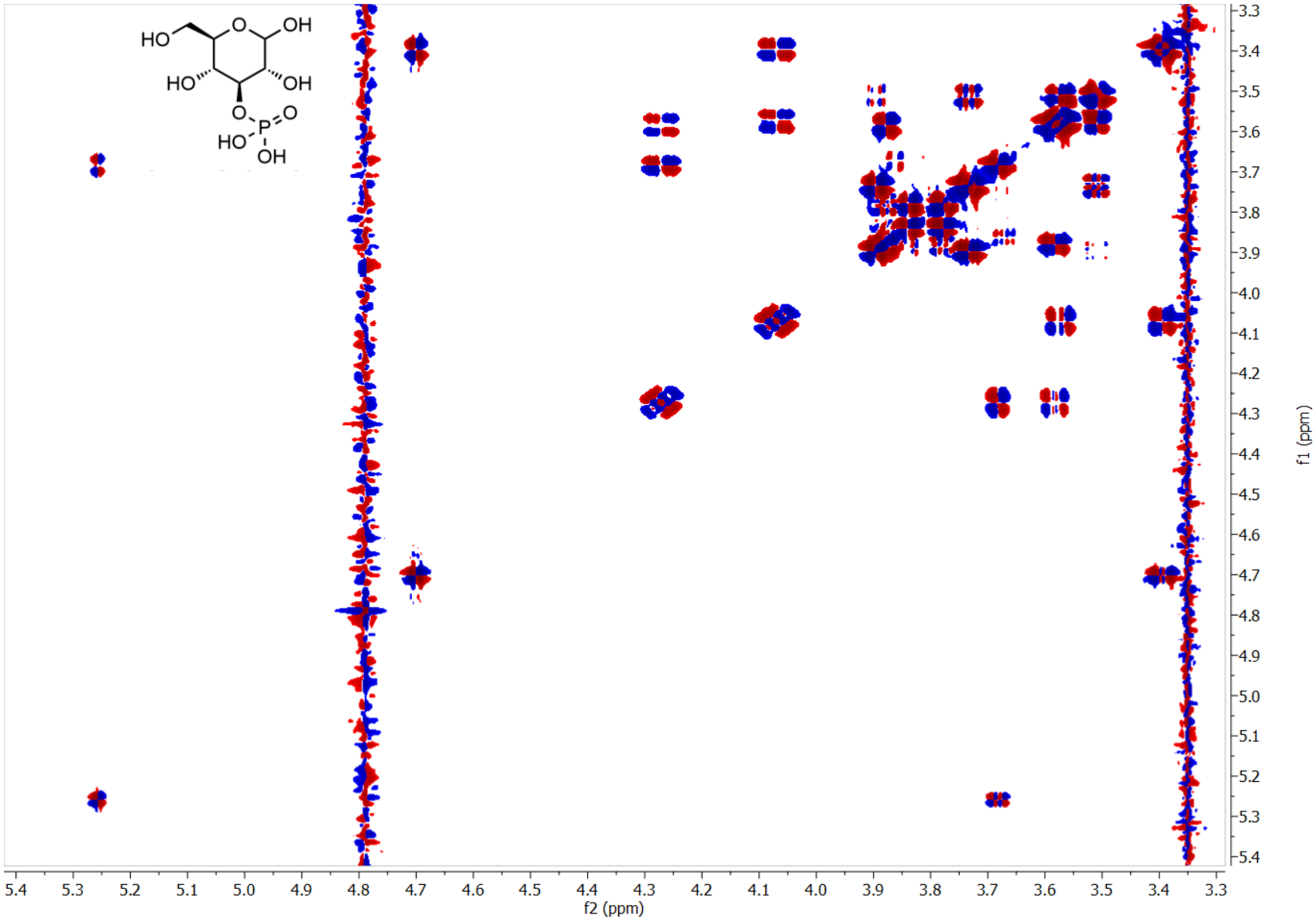
^{1}H NMR spectrum (600 MHz) of 3-O-phosphono-D-glucopyranose in deuterium oxide.



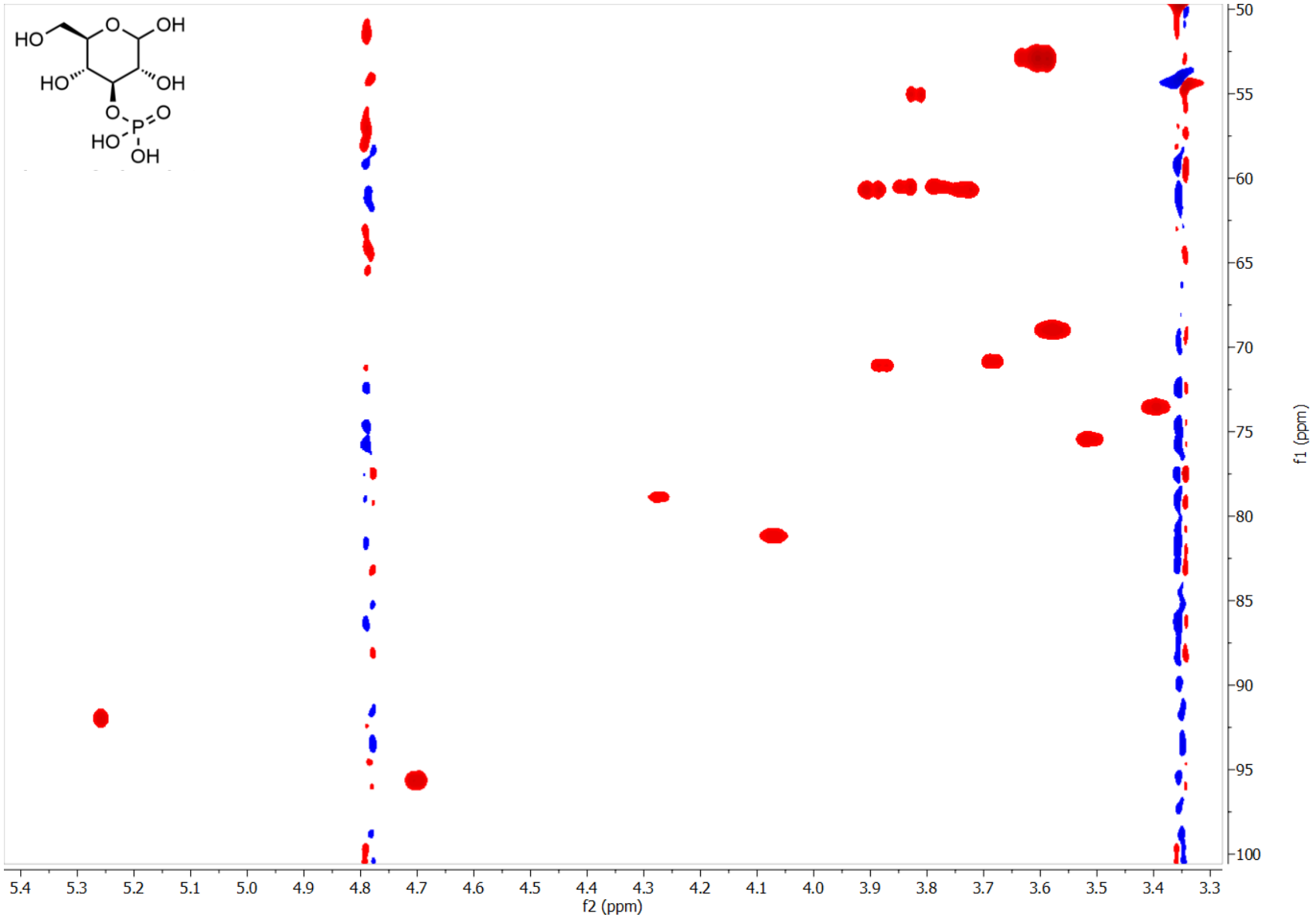
^{13}C NMR spectrum (126 MHz) of 3-O-phosphono-D-glucopyranose in deuterium oxide.



^{31}P NMR spectrum (202 MHz) of 3-O-phosphono-D-glucopyranose in deuterium oxide.



dqfCOSY NMR spectrum (600 MHz) of 3-O-phosphono-D-glucopyranose in deuterium oxide.



HMQC NMR spectrum (600 MHz) of 3-O-phosphono-D-glucopyranose in deuterium oxide.