Engineered Polyproline Helices for the Rational Design of Supramolecular Constructs

by

Dominic F Brightwell

A thesis submitted to the University of Kent in fulfilment of the requirements for the degree of Doctor of Philosophy

Supervised by: Aniello Palma

September 2023

School of Physical Chemistry & Forensic Science

Abstract

The purpose of this investigation is to determine if novel biocompatible supramolecular constructs with tuneable nanocavities for applications in catalysis, chemical separations, sensing and biomedical applications, can be rationally designed and synthesised utilising polyproline helices as their building blocks. The development of artificial supramolecular systems to mimic complex biological systems has been a longstanding goal of the scientific community and increasingly bioinspired building blocks such as DNA, RNA, aptamers, amino acids, peptides, and proteins have been utilised to develop complex supramolecular systems with promising applications. As such this project aimed to take advantage of the unique properties of the polyproline helix (e.g. rigidity, stability, *cis*-*trans* isomerism), and the inherent advantages of peptidic materials (e.g. chirality, periodicity, scalability), to develop peptide-based ligands that can be rationally designed to form desired complex, asymmetric, supramolecular constructs. This requires improving our understanding of the structure and assembly processes of these peptide ligands and how these can compare to the reticular design principles of traditional building-blocks. Therefore, these results will create the potential for the design of further peptide-based supramolecular constructs to produce novel, versatile, biocompatible, functional materials that can be exploited for promising applications.

A series of supramolecular peptide frameworks based on an oligoproline tetramer were synthesised. With these materials we were able to demonstrate the ability to rationally design the peptidic building block to form a variety of porous and non-porous crystalline frameworks depending on the placement of functional groups, while also providing an excellent model for the three-dimensional structure of functionalised polyproline II helices. Thus, highlighting how minimalistic peptide building-blocks can self-assemble to form desired supramolecular assemblies by tuning the supramolecular interactions. Furthermore, the reversible porosity of these peptide frameworks was shown with the ability to incorporate new guest molecules, which was not limited to solvents, with iodine vapour reinflating the peptide framework. Using a chiral guest molecule we were able to demonstrate the enantioselectivity of this absorption process with clear applications for the design and synthesis of functional materials for chemical separations and enantioselective catalysis.

Secondly, a series of functionalised non-natural prolines were synthesised with additional carboxylic acid and pyridine based functional groups. These were incorporated into a variety of peptide sequences to allow for the assembly of metal-peptide constructs *via* coordination interactions from the functional groups. By varying the functionalisation on each face of the polyproline helix both extended and discrete metal-peptide constructs were synthesised. Nanoparticles were synthesised from metal complexation of a series of polyprolines functionalised on all three helical faces, with the degree of functionalisation, placement of functional groups, and the polyproline helicity (polyproline I vs polyproline II) contributing to the topology of the assembly formed, demonstrating how these ligands can be designed to drive the formation of specific constructs. Furthermore, terminally pyridine functionalised proline tetramers were shown to form discrete palladium complexes in both organic and aqueous solutions. The dimerised peptide complexes expressed selectivity towards a single conformation, highlighting how the chirality of the helix favours a single product over a mixture of metal complex isomers. These complexes show promise for the design of discrete metal-peptide nanocavities as “molecular flasks” for enantioselective catalysis and a clear path to optimise the peptide ligands to form the desired peptide cage structures is discussed.

Declaration

I declare that the work presented herein is my own except where specific reference is made to the work of others. I certify that it has not been submitted for the purposes of a qualification at any other institution or for any other degree. This work includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text and Acknowledgements.

Dominic F Brightwell

September 2023

Acknowledgement

Firstly, I would like to express my gratitude to my supervisor Dr Aniello Palma for giving me the opportunity to work on this project. His passion and unwavering optimism were a constant source of encouragement and have helped to make a sometimes-difficult project a lot of fun. It is from his knowledge and enthusiasm to drive the project forward, as well as his guidance and support that I must thank for the skills, knowledge, and achievements I have attained over the past years.

I would also like to extend my gratitude to the University of Kent for their funding and support that made this work possible. I am especially thankful to the academics and staff of the School of Physical Chemistry & Forensic Science for their support and advice. Further thanks go to Dr Helena Sheperd and Dr Giada Truccolo for their invaluable help and training with all things crystallography, Dr Ewan Clark for all his imparted wisdom and support with anything related to NMR. Also, thanks to Dr Simon Holder for his assistance with computational analyses and Dr Chris Hawes for carrying out gas absorption studies and managing to analyse a very roughly treated crystal.

Furthermore, it would not have been the same without the people who have shared in my journey. I was fortunate to have known all the great people from the Palma group, thanks to Kushal, Athina, Elliott and everyone else who has shared in the trials and tribulations of research while being a great source of support. Beyond the group, I am grateful for all the brilliant people from the lab and School for their kindness and friendship along the way.

Also, I am thankful for my family, for their constant love and support throughout it all.

Finally, Sally, the love of my life and now my fiancé, I could not imagine doing this without you by my side. Thank you for sticking by me all these years as a poor student.

Dissemination

Peer-reviewed Publications

Brightwell, D. F.; Truccolo, G.; Samanta, K.; Fenn, E. J.; Holder, S. J.; Shepherd, H. J.; Hawes, C. S.; Palma, A. A Reversibly Porous Supramolecular Peptide Framework. *Chem. – A Eur. J.* **2022**, *28* (66), e202202368. https://doi.org/10.1002/CHEM.202202368.

Brightwell, D. F.; Truccolo, G.; Samanta, K.; Shepherd, H. J.; Palma, A. Supramolecular Self-Assembly of Engineered Polyproline Helices. *ACS Macro Lett.* **2023**, *12*, 908–914. https://doi.org/10.1021/acsmacrolett.3c00304.

Poster presentations

"Rational Design of Polyproline Supramolecular Building-Blocks" The annual RSC Macrocyclic and Supramolecular Chemistry (MASC) meeting (University of Nottingham, December 2022)

Table of Contents

List of Abbreviations

SPPS – Solid-Phase Peptide Synthesis

FCC – Flash Column Chromatography

CD – Circular Dichroism

FT-IR – Fourier-transform infrared (spectroscopy)

NMR – Nuclear-magnetic resonance (spectroscopy)

TLC – Thin-layer chromatography

HPLC – High-performance liquid chromatography

LCMS – Liquid Chromatography mass spectrometry

HRMS – High-resolution mass spectrometry

IM-MS – Ion-mobility mass spectrometry

ESI – Electro-spray ionisation

AFM – Atomic force microscopy

TEM – *trans*mission electron microscopy

SEM – Scanning-electron microscopy

RCM – Ring-closing metathesis

TFA – Trifluoroacetic acid

DIPEA- Diisopropylethylamine

Boc_2O – Di-*tert*-butyl dicarbonate

Tf_2O – Trifluoromethanesulfonic anhydride

PyBOP - benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate

DIC - Diisopropylcarbodiimide

DCM – Dichloromethane

MeOH – Methanol

EtOAc – Ethyl Acetate

Et_2O – Diethyl ether

EtOH – Ethanol

THF – Tetrahydrofuran

PrOH – Propan-1-ol

IPA – Propan-2-ol

PhEtOH – 1-Phenylethan-1-ol

DMSO – Dimethyl-sulfoxide

DMF – Dimethylformamide

MOF – Metal-organic framework

MOC – Metal-organic cage

SOF – Supramolecular-organic framework

COF – Covalent-organic framework

HOF – Hydrogen bonding-organic framework

NOE – Nuclear Overhauser Effect

NOESY – Nuclear Overhauser Enhancement Spectroscopy

ROESY – Rotating Frame Overhauser Enhancement Spectroscopy

Preface

In Chapter 1, the background and current state of the art of supramolecular chemistry is reviewed and how issues encountered therein have led in recent years to the development of bioinspired supramolecular building-blocks. From this the challenges and pitfalls are presented, while reviewing some of the most successful cases that have become somewhat common within literature. Specifically, after this discussion of general biomolecular building-blocks we focus on the potential benefits of utilising peptides and how these can lend themselves towards the rational design of supramolecular constructs. Herein, the polyproline helix is presented as a standout candidate and the previous uses and applications of the helix are discussed while looking at the prospects of the material, where further investigations can be focused, the potential design considerations and rewards from these investigations.

Chapter 2 presents the work that has been carried out on utilising minimalistic oligoprolines to synthesise supramolecular constructs, focusing on the formation of supramolecular peptide frameworks and how the design principles of an initial peptide model can be used to rationally design specific intermolecular interactions, affecting framework topologies, pore sizes and properties. The self-assembly of a series of engineered oligoproline tetramers is analysed, studied, and discussed. Furthermore, the reversible porosity and host-guest capabilities of the porous frameworks is analysed, demonstrating their enantioselective potential.

In Chapter 3 the work to utilise polyproline helices as structural units for the co-ordination driven assembly of nanostructures is presented. Functionalised proline monomers are initially synthesised, these monomers are then incorporated into a series of peptide sequences to achieve a variety of functionalised helices for the assembly of various nanostructures depending on the location, type, and degree of functionalisation. A variety of structures are targeted from nanoparticles to further extended and discrete assemblies. The assembled nanoparticles and other topologies are then analysed and studied *via* dynamic light scattering (DLS) and microscopy methods; *trans*mission electron microscopy (TEM), scanning electron microscopy (SEM), and atomic force microscopy (AFM). While discrete peptide complexes are analysed *via* NMR spectroscopy and mass spectrometry. The potential for rational design of these peptide units utilising the periodicity and tuneability of the polyproline helix is further discussed and how the switchable nature of polyproline can be used to form responsive materials dependent on external stimuli.

The experimental methods for all studies are included in Chapter 4, covering the synthesis of monomers and peptides. The supramolecular synthesis of assemblies is also presented and the analysis methods therein (e.g. microscopy x-ray diffraction, dynamic light scattering, NMR spectroscopy, and mass spectrometry).

Finally, Chapter 5 presents a summary of the work carried out and the conclusions from these findings. The future outlook for each of the projects is then analysed and the potential avenues of investigation are discussed.

Aims

A longstanding goal of the scientific community has been to develop artificial supramolecular systems that can mimic complex biological systems, and while significant progress has been made towards achieving this goal, hitherto, the ability to rationally design supramolecular constructs which can effectively mimic the capabilities of these systems (e.g. enzymes) remains elusive. The objective of this project is to take advantage of the stability and periodicity of polyproline helices to develop novel supramolecular constructs with improved biocompatibility and chemical complexity compared to traditional systems, while also improving our understanding of the assembly of these structures such that they can be rationally designed to achieve specific topologies and effects, with target structures ranging from metal-organic constructs to self-assembled supramolecular organic frameworks, driven by weaker intermolecular interactions. Thus, these results will open up the potential for the design of further peptidic constructs for the generation of versatile biocompatible materials to be exploited for catalysis, with substrate selectivity and stereoselectivity, chemical separations, sensing or biomedical applications.

Chapter 1.

Polyproline helices as structural units for the design of supramolecular constructs

1.1 Introduction

1.1.1 Background and Supramolecular Chemistry:

The field of supramolecular chemistry is at the forefront of scientific interest in recent years, with the breadth and scope of investigations into supramolecular constructs, and their applications, growing at an ever-increasing pace as the field garners greater attention from the wider scientific community. Since some of the earliest work, pioneered by Raymond and Stang,[1,2] investigating the synthesis of discrete and extended constructs such as metal-organic frameworks (MOFs) and metal-organic cages, instrumental in demonstrating the potential of *de novo* rational design of supramolecular constructs, the ideal of supramolecular building-blocks with tuneable chemical handles and predictable geometries in three-dimensional space, to allow the design of methods of assembly and construct topology, remains a lofty goal.

Despite the many difficulties encountered in the development of supramolecular constructs these have been frequently overcome to achieve novel supramolecular systems that are useful in a broad range of applications. These supramolecular constructs vary; from metal-organic frameworks (MOFs),[3,4] supramolecular organic frameworks (SOFs),[5,6] covalent organic frameworks (COFs),[7,8] hydro- and organo-gels,[9] and nanoparticles to more complex systems such as nanoscale molecular weaves and nanocages.[10–12] This broad range of materials has an equally broad expanse of applications from biomedical applications, such as drug delivery and medical treatments,[13–16] to chemical separations,[17,18] catalysis,[3,19–21] sensing,[22,23] and gas storage.[5,24,25] However, inspired by biological systems that demonstrate flawless control over self-assembly processes to achieve catalytic systems (e.g. enzymes), scientists have long sought after this same level of control in these artificial systems.

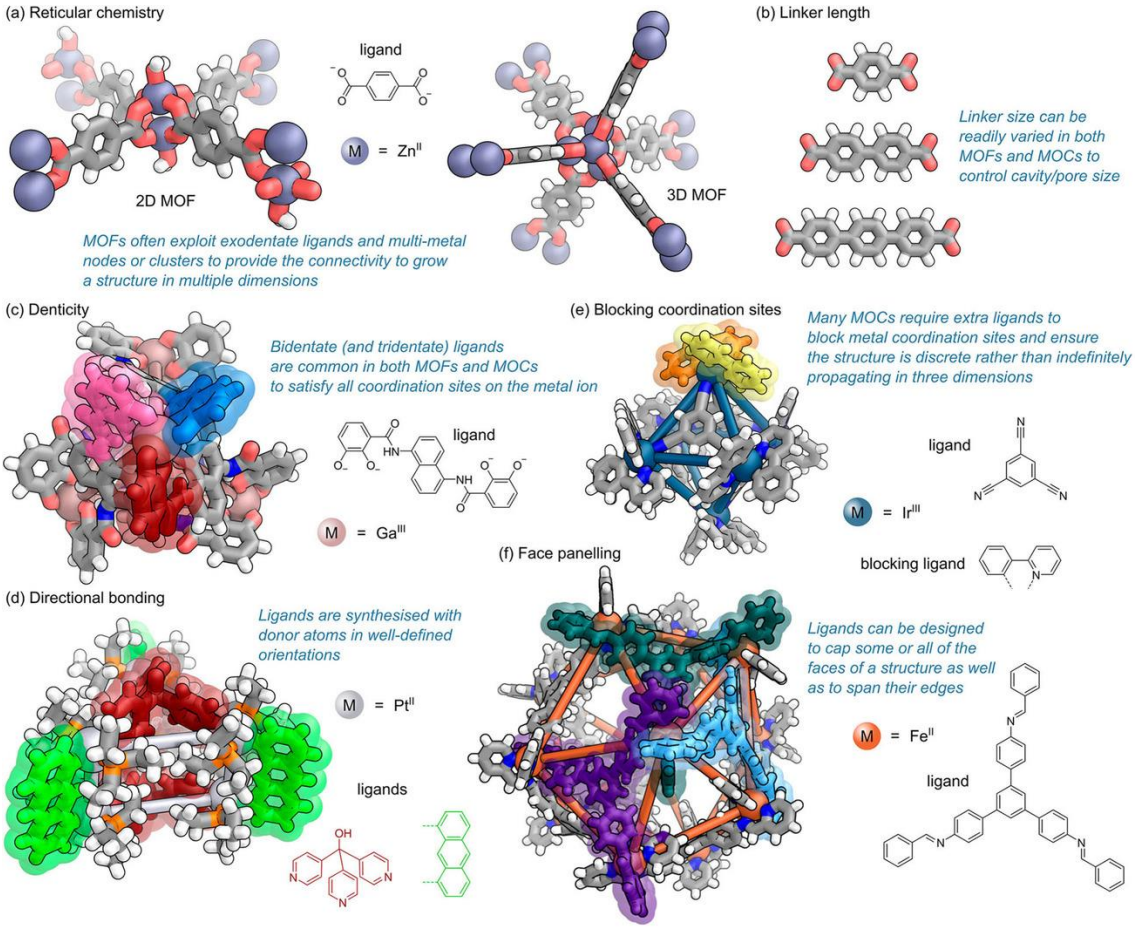


Figure 1 - Illustration from work by Dr B.Pilgrim and Prof N.Champness "Metal-Organic Frameworks and Metal-Organic Cages – A Perspective": Construction strategies: (a) The reticular chemistry of MOFs often employs both

exodentate ligands and multi-metal nodes or clusters to grow a framework in multiple dimensions; (b) Linker length can be readily varied in both MOFs and MOCs to control the size of the cavity or pore; (c) Three bidentate catecholate ligands meeting at a vertex to satisfy the octahedral coordination sphere of Ga(III) in a MOC; (d) The directional bonding approach of using well-defined ligand geometries to control the shape of MOCs;(e) Two bidentate ligands block vacant coordination sites on each vertex of an M_6L_4 octahedral MOC keeping the structure discrete; (f) Face panelling of an $M_{12}L_{12}$ icosahedral MOC, where three faces meeting at each vertex are panelled.[26]

The design and synthesis of typical MOFs or cages (MOCs)[26] is a prime example of the state of play within supramolecular chemistry (Figure 1). MOFs are crystalline frameworks assembled from metal ions with bridging organic ligands to form an extended network. While MOFs are polymeric systems, MOCs are discrete assemblies with several other naming systems coined within literature such as metal-organic polyhedral (MOPs)[27] or supramolecular coordination complexes (SCCs),[28] however there is significant overlap in their design principles and applications due to their well-defined internal cavities, useful for host-guest chemistries. Tuning the assembly of the building blocks can be guided by selection of the required metal ions and organic ligands to achieve either these extended or discrete systems.[26] MOFs have become increasingly popular as porous materials due to their high permanent porosity with large surface areas, uniform pore sizes with tuneable surfaces, and a high degree of scalability, as such they have a broad range of applications.[29] However, the discovery of new materials has previously often relied upon serendipitous discovery using an explorative approach, with a lack of defined design principles behind synthesised materials (i.e. shake and bake),[30] yet despite successful formations of important unprecedented frameworks[31–33] it has become increasingly imperative that materials designed for highly specific and cooperative functions are produced, extremely challenging without the capacity for rational design. With the properties and topologies of these materials relying on the selection of the type of metal nodes and nature of the bridging organic ligands, the rational design of MOFs has led to the emergence of reticular chemistry, linking conformationally rigid ligands with highly directional strong bonding interactions.[29] This allows the targeting of specific framework topologies and for facile tuning of framework pores to achieve desirable guest interactions. As such, these methodologies have shown significant successes in the development of desirable MOFs,[34–37] with applications for gas storage, chemical separations, sensing and in biomedicine.[38–42] While the development of similarly designed metallacages and metallacycles has shown equal successes with numerous reviews now covering the wide ranging structures discovered[26–28,43] found to have novel applications such as serving as “molecular flasks” for reactions and catalysis[44–46] or biomedical applications as selective analyte sensors and anticancer agents.[47,48]

While reticular design was first applied to the synthesis of porous MOFs these principles have since been more broadly applied to covalent organic frameworks (COFs) and cages,[49] where the reversibility of the crystallisation of covalently bonded organic molecules was initially thought to be an insurmountable challenge to the formation of porous crystalline frameworks.[50] COFs are crystalline polymers where organic molecules are linked by covalent bonds to form an ordered extended network, while cages are designed such that the ligands form a discrete assembly. Due to their strong bonding interactions, COFs are highly stable and permanently porous, however reactions to form the covalent linkages are limited to reversible formations as this allows self-healing to occur during crystallisation and polymerisation to prevent the occurrence of structural defects to form an ordered structure. Thus, to obtain ordered COFs the organic building blocks are again limited to conformationally rigid compounds with discrete directional bond formations. However, despite these limitations several successful syntheses and applications of COFs has been found with a larger focus on 2D COFs over 3D COFs, with limited examples of 3D frameworks due to crystallisation issues.[51] Although notable examples have shown applications for gas storage,[52] size selective catalysis,[53] and as optoelectronics.[54] Conversely, covalent organic 3D cages and 2D macrocycles have been more extensively studied,

yet less frequently than coordination cages. These discrete cages are notably different to extended frameworks due to the potential for cages to act as soluble porous units (SPUs) before self-assembly offering solution-processing options, unlike insoluble frameworks. Recently, several reviews have outlined the unique features of porous organic cages and their potential applications, from catalysis[55] and stabilisation of reactive compounds[56] to selective guest recognition and absorption.[57–60]

In contrast to the strong bonding interactions within MOFs and COFs, weaker non-covalent interactions such as hydrogen-bonds, π-interactions, dispersion and Van Der Waals interactions have been extensively utilised to synthesise supramolecular organic frameworks (SOFs) and assemblies (e.g. HOFs[61] and πOFs).[62] While the construction of discrete supramolecular cages from subunits *via* these weak interactions has also seen significant interest to mimic protein cages and recent promising examples have been shown.[63–66] Utilising these weaker interactions can be highly beneficial resulting in self-healing capabilities,[62] reversibility, and stimuli-responsiveness of synthesised materials.[67] HOFs, self-assembled organic molecules forming extended frameworks through reversible hydrogen bonds as well as other weak interactions, have shown a broad range of highly promising applications such as frameworks for the crystalline sponge method,[68] efficient gas storage and separation,[69–71] enantioselective separation,[72] asymmetric catalysis,[73] sensing,[74] biological applications,[75,76] and even as switchable machines.[77] These exciting applications highlight the great potential from the development of non-covalent assemblies and yet there are several challenges associated with the synthesis of these materials. Namely, H-bonding interactions, one of the stronger non-covalent intermolecular interactions used, are still relatively weak and flexible compared to covalent or coordination bonds, as such permanent porosity is difficult to achieve with guest removal often resulting in pore collapse, and stabilisation of frameworks and prediction of assembly topologies is more difficult. As such, rigid polyaromatic molecules are typically incorporated with the capacity to provide additional stabilising and directional π-interactions to the assemblies.[62,68,78]

With these examples it is no wonder that achieving rationally designed supramolecular materials is a constant topic of research within the field.[11,15,79–84] However, as mentioned previously, this has created a reliance on ligands that are rigid, polyaromatic and highly symmetrical.[29] While this methodology simplifies the principles of the assembly process allowing for prediction *via* computational methods,[85] thus allowing for rational design of these systems, it also presents several significant roadblocks. Firstly, the variety of viable building-blocks is severely limited as well as the number of functionalities that can be introduced, reducing the chemical complexity of the synthesised supramolecular assemblies. Furthermore, this class of ligands is often bio-incompatible, and these compounds lack the ability to incorporate a wide variety of functionalities, often requiring long and complex syntheses to achieve small changes in chemical structure. This is highly restricting in supramolecular chemistry where tuning and optimisation of materials *via* small changes in ligand structure to study a series of similar materials is ideal, thus the scope of any investigations are affected. However, despite substantial efforts to develop methods for the rational design of complex bioinspired supramolecular systems there yet remains significant further progress to be made. Hitherto, developed materials have yet to match the performance of nature's systems, and the great successes already achieved highlight the substantial rewards at stake, as such there remains a need for the further development of materials that can be rationally designed, are tuneable with high chemical complexity, and are inherently biocompatible.

1.1.2 Bioinspired Supramolecular Building-Blocks:

The limitations of more traditional materials have led researchers to look towards nature for inspiration to develop building-blocks that can assemble into supramolecular constructs capable of performing desirable functions such as mimicking the functionality of biological systems. Biological materials are notorious for the ability to assemble complex biopolymers to achieve reactions that are highly challenging to achieve artificially, enzymes are the prime example of this achieving highly selective and efficient chiral, catalytic *trans*formations while in an aqueous environment under ambient conditions.[86,87] Significant efforts have thus gone towards attempting to mimic the capabilities of these biological systems and while some of the functional features of an enzyme reactive site can be mimicked it is highly challenging to control the self-assembly of low symmetry systems to achieve these results. This has subsequently led to the use of bio-inspired supramolecular building blocks, ranging from proteins, peptides, and amino acids, to DNA and RNA helices.[88–93] Previously, the use of secondary structures of DNA and RNA emerged as a promising avenue for the rational design of supramolecular constructs, now well-established these methods make use of DNA's specific-base pair recognition requirements which, coupled with a toolbox of naturally occurring enzymes for manipulation and relatively cheap synthesis, makes DNA an excellent avenue for materials and nanotechnology development.[94–99] These materials have been shown to have a vast range of applications, ranging from DNA origami,[100] for the bottom-up synthesis of well-defined nanostructures, and DNAzymes, single-stranded DNA catalysts able to mimic the activity of enzymes and act as biosensors,[101,102] to DNA-based nanocarriers for intracellular biomedical applications (Figure 2).[103,104] The modular nature of DNA and RNA, combined with their biocompatibility has allowed for this rational design of programmable nanostructures leading to such an expansive range of promising functional materials.

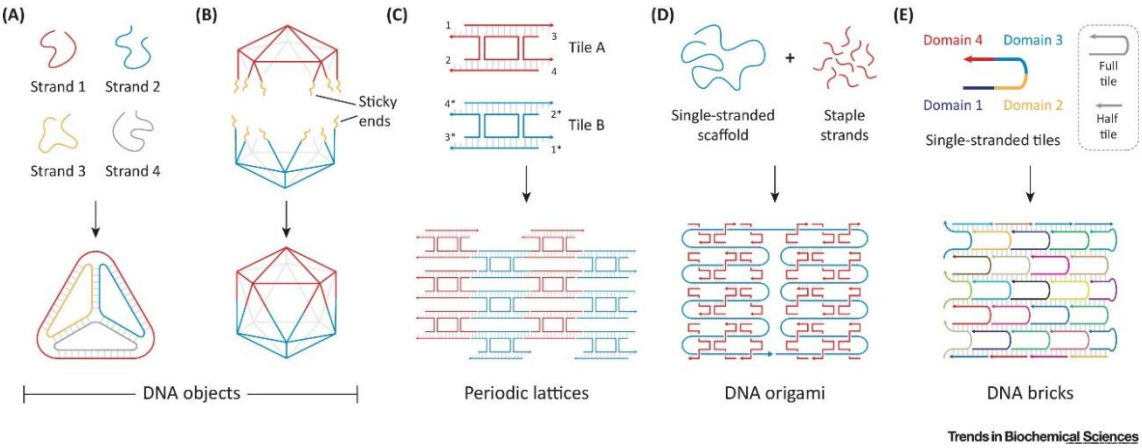


Figure 2 -Illustration from work by Dr A.Chandrasekaran "DNA Nanocarriers: Programmed to Deliver": Designing DNA Nanostructures. (A) Cooperative self-assembly of DNA strands into a DNA tetrahedron. (B) Component halves designed from five-arm branched junctions connected via sticky ends to form a DNA icosahedron. (C) Double crossover DNA motifs tailed with sticky ends alternate to form a 2D array (complementary sticky ends are denoted by N-n). (D) A long scaffold strand is folded into desired shapes by using short complementary staple strands in a method called DNA origami. (E) Single-stranded DNA tiles connect to each other via complementary domains to form a molecular canvas.*[104]

While protein-constructs have also been utilised as supramolecular building-blocks shown in the fabrication of protein-macrocycle and metal-protein frameworks,[79,105] as well as advances to create protein nanocages akin to the enzymatic reactive site, partly thanks to progress in computational methods.[106–108] Recently, further advancements in supramolecular technology have been developed using lipids, nucleic acids and peptides as building blocks to assemble novel 2D and 3D biomaterials through supramolecular interactions.[96,109,110] These molecules most commonly make use of hydrophobic interactions, π-stacking, cation-π interactions,

hydrogen bonding and metal-ligand coordination to direct self-assembly with the precise location and orientation of functional groups driving the method of assembly.[98,111] The wealth and variety of available natural materials and the ease of functionalisation and synthesis of biomaterials make them highly promising candidates for use as supramolecular building-blocks, however, complex intra- and inter-molecular interactions combined with chirality and conformational flexibility of these materials impacts their potential for rational design and has somewhat limited the successful incorporation of these materials into designed nanomaterials.[112] Efforts have been made to develop supramolecular constructs such as MOFs based on flexible ligands and these have seen some success, exhibiting unique structures and properties.[112] However, well-defined secondary and *tert*iary structures of bio-building blocks is essential to achieve *de novo* design of these bioinspired materials, ideally with minimal perturbation of their 3D structures upon functionalisation to allow tuning of these ligands to predictably change the assembly processes, topology, macromolecular properties, or reactivity. Evidence of recent successful uses of biomaterials to achieve desirable supramolecular constructs is only increasing in recent years with numerous promising examples,[15,91,94,98,111,113] emphasising the untapped potential available within this field.

1.1.3 Peptides as Supramolecular Building-Blocks:

With this desire to achieve biomaterial building-blocks; that are both easily synthesisable and tuneable, with the facile addition or removal of functional moieties, while having a well-defined structure to facilitate rational design, retaining biocompatibility, economic viability, and a vast expanse of available structures, our focus moved to peptide-based materials. Out of this selection of potential biomaterials, suitable for utilising for supramolecular assembly, peptides make highly promising candidates. Peptides are highly accessible materials with the advent of Fmoc-based solid-phase peptide synthesis (SPPS) peptides, such that they can be prepared at scale with a high purity, while incorporating natural and non-natural amino acids into the primary structure.[114] With this stepwise synthesis amino acids, and thereby functional groups, can be placed with high accuracy and can be easily rearranged for new syntheses to alter the placement and nature of these functional motifs. Also, peptides are found throughout biological systems as the constituent parts of proteins. The primary sequence of the peptide units dictate the *tert*iary structure these proteins adopt, as such peptide units present ideal candidates for the development of proteomimetic systems.[81] Additionally, peptides adopt well-known and well-defined secondary structures such as α-helices, ß-sheets, coiled-coils and polyproline helices. These well-defined secondary structures mean that the placement of amino acids within the primary structure produces highly predictable geometries of incorporated sidechains and their functional groups within three-dimensional space, which with the inherent chirality of peptides makes them ideal candidates in

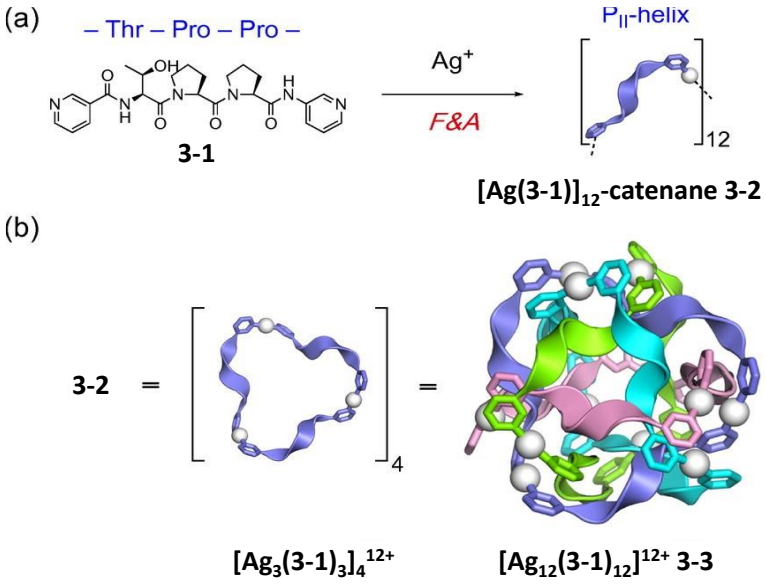


Figure 3 – Figure adapted from a report by Fujita: (a) Complexation of Ag^+ ions and tripeptide ligand 3-1 and (b) structure of $[3\text{-}2]_{12}$-catenane 3-3.[121]

stereoselective *trans*formations and separations. Thus, this information can be utilised to design peptidic ligands that can assemble through supramolecular interactions to form desired constructs with specific topologies, solubilities and reactivities. As such, interest in the development of hierarchical peptide-based supramolecular materials to mimic the capabilities of biological systems is rapidly growing.[115–117]

Recent work in this field has been pioneered by several research groups, some of the most notable examples of which are: The Fujita group, well known for the development of the crystalline sponge method and Fujita's seminal work on molecular recognition in macrocyclic complexes,[118,119] has demonstrated the formation of metal-peptide supramolecular assemblies from peptide strands (Figure 3).[21,120–124] From these peptide units Fujita is able to construct rings or strings of peptide which assemble into unprecedented orderly entangled nanostructures from a folding-and-assembly strategy (Figure 3).[121] These findings illustrate how peptides can perform as versatile and modular ligands, affording chirality, conformational adaptability and specificity to construct well-defined coordination networks. Moreover, these assemblies were shown to have potential applications in biomedicine, catalysis, and molecular recognition, as they were able to encapsulate guest molecules, catalyse reactions and bind to biomolecules.[125]

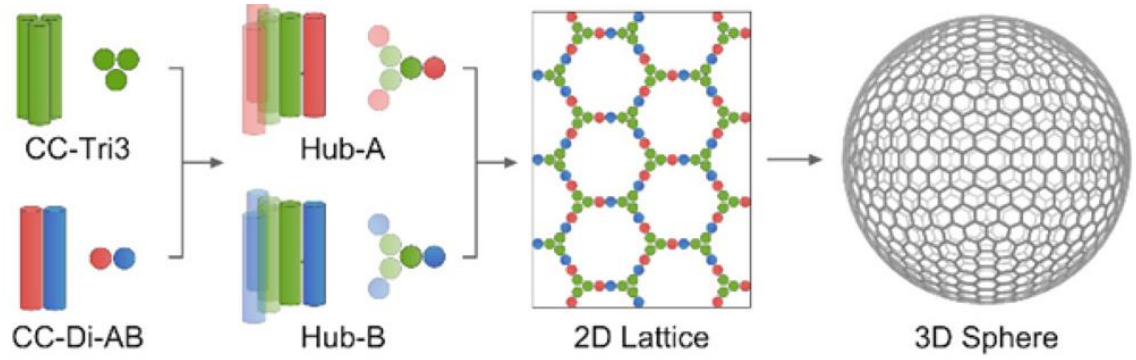


Figure 4 - Illustration adapted from a report by Prof. D. Woolfson, showing the schematics of SAGE assembly and example compositions. When mixed, these peptides assemble through heterodimer association to form a hexagonal lattice that closes to form SAGE particles.[11]

Professor Derek Woolfson's group has demonstrated the design and formation of self-assembled peptide nanocages (SAGEs) from coiled-coil peptide modules with a design strategy offering chemistry, self-assembly, and particle size control (Figure 4).[11,64] These *de novo* designed SAGEs been shown to be nontoxic nanoparticles with promising potential as scaffolds for the delivery of immunogenic components. The modularity of these systems has allowed for the development of functionalised derivatives and can be designed to be capable of internalisation into mammalian cells.[11] They can be functionalised with proteins at the C and N termini of the peptide hubs to create protein-SAGE (pSAGE) particles[65] or functionalised with antigenic peptides to drive specific cell responses,[66] and as such have clear applications as nanoreactors or cell-delivery vehicles.[126–128]

The Gazit lab has also focused on the use of peptides as supramolecular building blocks to self-assemble to form well-ordered nanostructures and has demonstrated the formation of super helical architectures from dipeptide nanostructures (i.e. minimal molecular elements still presenting stable helical structures) such as nanotubes, nanospheres, nanoplates, and hydrogels, as well as metal-peptide frameworks,[126,129–132] illustrating how short minimalistic peptides can be used for the design of complex helical self-assembling supramolecular constructs with a broad range of potential applications for the development of biomimetic functional materials.[127,128,133]

The Wennemers group has shown several examples of the application of peptides as molecular scaffolds to form supramolecular assemblies, specifically utilising the polyproline helix, notable examples range from a triaxial supramolecular weave (Figure 11) to a metal-peptide framework.[10,134–136] Another example of their work is utilising polyproline helices as supramolecular building-blocks for the formation of silver nanoparticles with defined sizes,[136] showing a clear goal of using the well-defined polyproline helix as a predictable unit to control distances within rationally designed functional materials.[137] These exciting advancements in the understanding and potential of peptidic supramolecular systems, achieving novel designed constructs with highly promising applications, highlights the wealth of untapped potential in the field and the significant promise these projects pose for the development of materials to solve as yet unresolved scientific problems with significant impact across the scientific community and wider population.[13,111,113,138,139]

Although there have been significant recent successes utilising peptides in supramolecular systems there remain significant challenges, particularly in the rational design and ordered assembly of these materials. These difficulties are in part down to the fact peptides are typically flexible, chiral, and present a multitude of chemical side chains, which results in complex intermolecular interactions and packing within the solid state. There has been significant interest in the design of flexible ligands for the creation of supramolecular constructs, however, these have been largely limited to ligands with large aromatic groups joined by single bonds capable of free rotation, thus having relatively predictable flexibility, of which some of the most notable case have been exemplified by Prof. Michael Ward with the synthesis of a family of coordination cages driven by pyridine coordination.[112,140–142] Predicting the likely method and successful assembly of these low-symmetry flexible peptide units is often difficult and more computationally challenging compared to traditional ligands. Peptide secondary structures, such as α-helices, are also often prone to perturbation upon functionalisation, relying on specific primary structures to retain helicity,[143] limiting the scope of functionalisation of these units. Also, intramolecular non-covalent interactions are present within many helices, and these groups are perfectly capable of adopting intermolecular interactions, further complicating the method of assembly with disruption of the secondary structure, and varying the potential assembly principles. These factors complicate the development and design of many peptidic building blocks and present a significant challenge to be surmounted for the field.

1.1.4 Polyproline Helices and their Potential as Supramolecular Building-Blocks:

Considering the difficulties typically encountered with peptide supramolecular building blocks, the polyproline helix, ubiquitous throughout nature, and one of the dominant helices found in proteins and peptides alongside *ß*-sheets and *α*-helices,[144,145] makes a prime candidate to reduce many of the disadvantageous properties found for other peptide building-blocks while retaining the desirable qualities. Often referred to as a "molecular ruler", due to the well-known rigidity and predictable length of the helices,[137] the polyproline helix has been used as a spacer in molecular studies and supramolecular structures to achieve ångström precision of distances between specific units.[136,137,146–148] This has also allowed the use of polyprolines in experiments such as Förster resonance energy *trans*fer (FRET) studies as a reference for the determination of the length of biomolecules.[149] These applications clearly highlight how the secondary structure of polyprolines gives a highly predictable geometry with a rigid helix that can be adjusted to produce units with precise dimensions. This alone makes the polyproline helix a highly promising candidate when compared to other peptide units.

Significantly, polyproline is known to form two different helical structures; the left-handed polyproline II helix and the right-handed polyproline I helix (Figure 5). The polyproline I conformation has a more condensed helix due to a tighter helical turn, with 3.3 residues per

turn and all its peptide bonds in the *cis* conformation covering 5.6 Å.[150] Polyproline II has a more extended structure with 3.0 residues per turn with every third residue stacked on top of each other 9.3 Å apart and amide bonds all in the *trans* conformation.[135,151] Interconversion can occur between these two types through *cis-trans* isomerism (Figure 5*d*), which can be instigated by a change in pH, temperature or solvent, with a theorised intermediate of discontinuous *trans* and *cis* amide proline bonds.[152,153] Polyproline II is the more common conformation, favoured in aqueous conditions, occurring in natural proteins and involved in many biological processes as it is generally more stable. The lack of *N*-H amide groups along the backbone means intramolecular hydrogen bonds cannot form, therefore polyproline, unlike other amino acids, is free to adopt both the *cis* and *trans* isomers of proline with N-C atoms from pyrrolidine rings forming part of the backbone. In other amino acids the *trans* conformation is significantly more favourable than the *cis* conformation, while the free energy difference in proline is significantly smaller, and a higher free energy barrier results in a slow *trans*ition rate.[154–156]

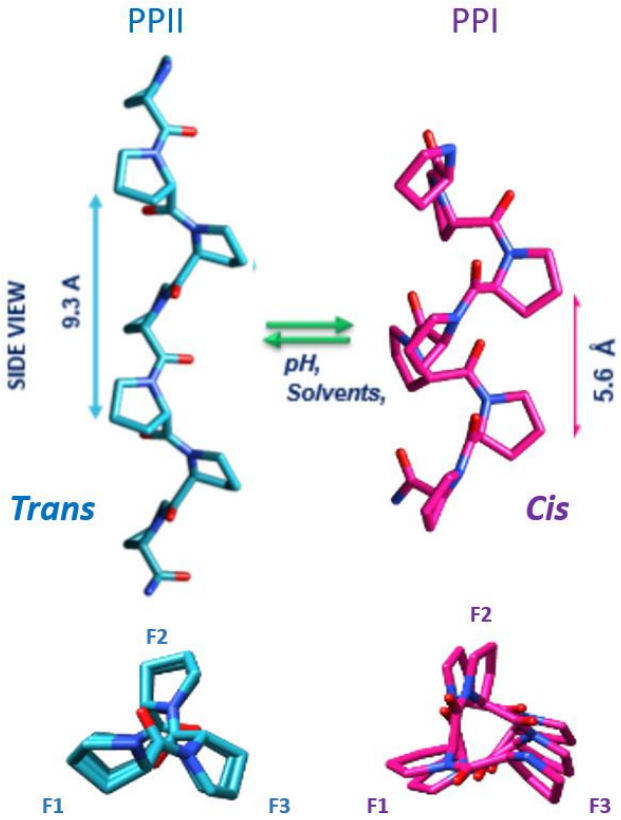


Figure 5 - Example polyproline peptide hexamer in both polyproline I and polyproline II conformations, viewed from the side and above.

The polyproline II *trans* configuration is favoured in more polar solvents such as; water, TFE, organic acids and benzyl alcohol, while the polyproline I *cis* conformation is typically favoured in less polar aliphatic alcohols such as ethanol or propan-1-ol.[157,158] In these less polar solvents the formation of the more condensed right-handed helix shields the peptide backbone, while in more polar solvents the formation of the *trans* amide bonds is stabilised by interactions of the exposed carbonyl groups with the polar solvent.[153] This makes the polyproline helix sensitive to external stimuli such that two structures, with significantly different lengths and chirality, are accessible from a single peptide unit.

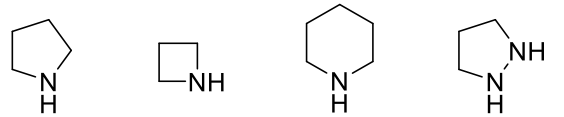


Figure 6 – Chemical structures of proline analogues

Investigations show that the interconversion between polyproline I (PPI) and polyproline II (PPII) is strongly influenced *via* two mechanisms: *via* an exothermic proton *trans*fer process (*reaction 1*) and an entropically driven endothermic conformational change resulting from *reaction 2*. These two processes, observable through kinetic investigations using IM-MS data, together, produce an extremely slow proton *trans*fer process leading to slow conversion.[159] Maximum

conversion (96%) to polyproline I occurs in propan-1-ol after six days in long chain polyprolines.[160] This is due to the former *trans*ition being caused by *cis-trans* isomerisation as propan-1-ol replaces water affecting *trans*ition states and intermediates, while the latter *trans*ition is due to hydrogen bond breakage/formation from hydration of the peptide backbone.[157] Several methods can be used to limit the interconversion of the two polyproline types and favour one over the other, such as incorporation of proline analogues (e.g. azaproline, azetidine, piperidine)[161,162] into the primary structure and substitution on the pyrrolidine ring of residues, typically at the C^{λ} position (e.g. 5-*tert*-butylproline, methanoproline and azido-, hydroxy-, fluoro-, and methoxy-proline derivatives). [151,152,163–165] Alteration of the terminal moieties of polyproline residues, such as altering the *N*-terminal capping group, has also been shown to affect the propensity to adopt either of the two helical conformations.[166,167]

(1) $PPI^{+}_{PrOH} + H_3O^{+} \longrightarrow PPII^{2+}_{PrOH/aq} + H_2O$

(2) $PPI^{+}_{PrOH} \longrightarrow PPII^{+}_{PrOH/aq}$

The stability of the polyproline helix also varies with chain length as, with greater chain lengths, the ability of the polyproline I helix to form increases, while the polyproline II helix remains stable down to short chain lengths (> $(Pro)_4$). The polyproline II helix is more favoured and stable at lower chain lengths due to the stabilising effect of *N*-π* interactions, whereby the oxygen's lone pair (n-orbital) of a neighbouring carbonyl delocalises into the antibonding orbital (π*) of the next carbonyl, following the Bürgi-Dunitz trajectory, evidenced by pyramidalization of the typically planar carbonyl group (Figure 7*d*).[135,168,169] However, the polyproline I helix is stable in propan-1-ol for a $(Pro)_{13}$ peptide and potentially as short as $(Pro)_6$,[157] with longer chain lengths the increased number of amide bonds also makes the *cis* conformation easier to detect.

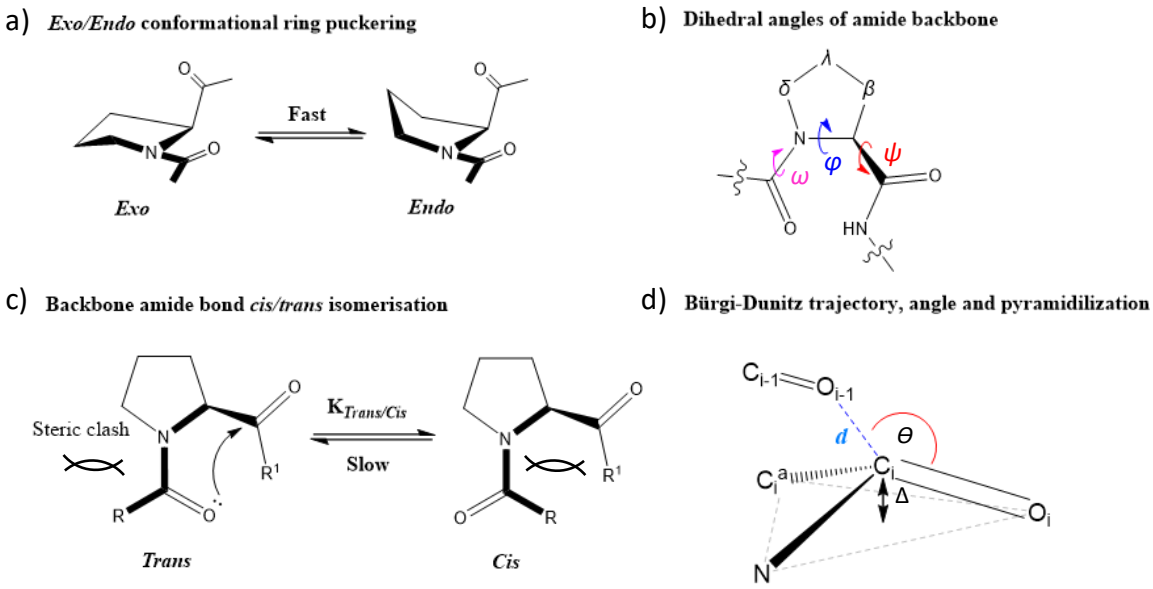


Figure 7 - a) diagram of conformational change due to endo/exo puckering of the proline ring, b) diagram showing the dihedral angles of a proline residue and the nomenclature for the pyrrolidine ring carbon atoms, c) diagram of cis/trans isomerism of proline and de/stabilising interactions, d) diagram showing the Bürgi-Dunitz trajectory and pyramidalization of the backbone amide carbonyls

This switching of polyproline can be monitored *via* several methods; circular dichroism (CD) spectrometry presents a quick and effective method for the determination of the conformation of polyproline peptides as the characteristic spectra of polyproline I and polyproline II differs significantly. The polyproline I helix presents minima at ≈200 nm, maxima at ≈214 nm and a

small negative band at ≈232 nm while the polyproline II presents reduced minima at ≈204 nm and higher intensity maxima at ≈226 nm (Figure 8a). This shift is well documented for polyproline chains 10 monomer units and larger.[135,157,160] Also, it is has been suggested that the shift between the two conformations can be observed with the intensity of the molar ellipticity at 214 nm, where an increase in signal indicates *trans*ition to polyproline I.[157]

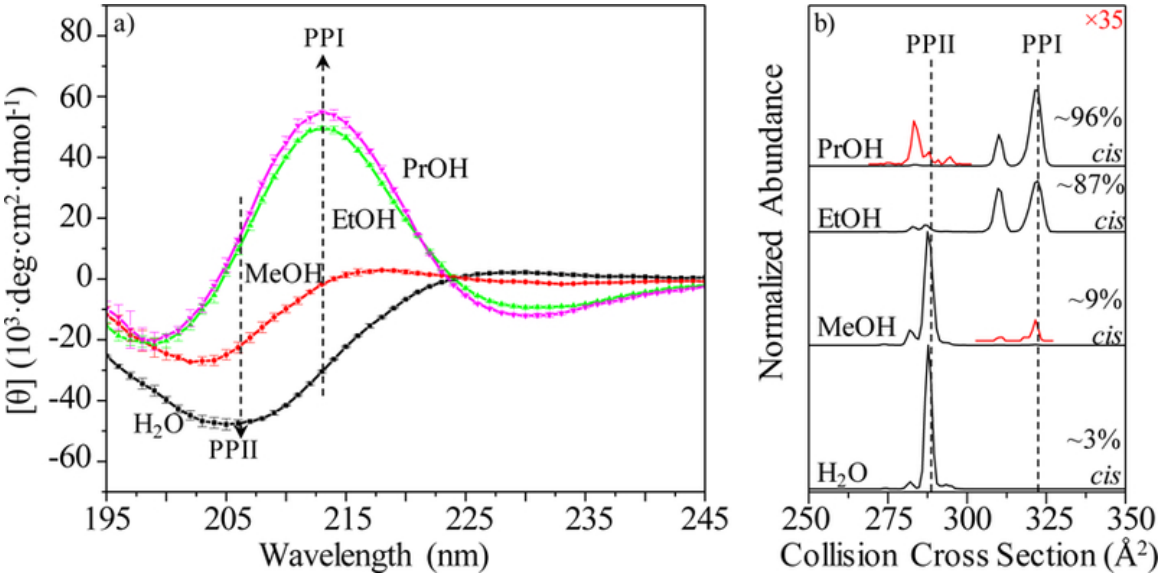


*Figure 8 - (**a**) CD and (**b**) IM-MS distributions of Pro13 incubated in water, methanol, ethanol, or 1-propanol at 298 K, from a report by Clemmer. Inset red lines in (**b**) scaled by × 35* [158]

However, the determination of the proportion of the *cis* amide bonds is more difficult to determine from CD spectroscopy for shorter oligoprolines with a weaker CD signal and other methods can be required for differentiation of the adopted helix, for example ion-mobility mass spectrometry (IM-MS) showcased by Clemmer, allows accurate size determination of peptide units from the collision cross-sectional area, as the conformation of the peptide affects the mobility through a buffer gas, and thus the proportion of *cis* vs *trans* bonds can be determined due to the peptide size difference (Figure 8). This has allowed investigation of the *trans*ition states between the two helices, mixtures of *cis* and *trans* amide bonds, and the pathways for folding and unfolding the peptide as a function of the solution environment by preserving the solution state structures.[153] FT-IR spectroscopy can also be used to monitor *cis-trans* switching, whereby, the exhibited wavenumber of the polyproline C=O amide band indicates the polyproline conformation present in the sample. Polyproline I should exhibit C=O amide stretching band at 1635 cm^{-1} and polyproline II should exhibit the band at 1623 cm^{-1}, with an intermediate band suggested to form during the *trans*ition at 1651 cm^{-1} when both conformations are present within the sample.[160] NMR data has also been previously used, showing a slight difference in obtained spectra between the conformations.[170]

These tuneable properties of the polyproline helix towards either helical confirmation may allow for the design of peptide-based building blocks that can respond to external stimuli to engender a specific supramolecular change, such as the dimensions and chirality of supramolecular structures, to achieve a desired effect. This can easily be envisioned as a useful property in chiral *trans*formations and separations as well for biomedical applications, such as the controlled release or encapsulation of molecules for drug delivery.[171]

More recently in 2014 the first crystal structure of the polyproline II helix was obtained by the Wennemer group showing the exact dimensions and conformational properties of a *p*-bromobenzoyl capped oligoproline hexamer.[135] This finding facilitates the design and use of

functionalised derivatives as both molecular rulers and scaffolds. With these specific spatial parameters from the crystal structure, and the rigidity of the polyproline II helix, supramolecular building-blocks incorporating the helix are significantly more accessible. The use of shorter polyproline helices has been demonstrated by Fujita; metal–peptide rings which assemble into highly entangled frameworks and the coordination-driven assembly of a short polyproline helix in to a small protein-like channel, showing guest uptake and chiral recognition (Figure 10)**.**[123,172] This highlights the great potential of polyproline as a ligand in metal-organic assemblies and more broadly as a supramolecular building block, with promising applications as proteomimetics. The Wennemer group has also gone on to demonstrate the formation of a metal-organic framework from an oligoproline hexamer in the polyproline II conformation, capped with an *N*-terminal benzoic acid moiety,[134] the longest peptide used to assemble a metal-peptide framework.[81] The framework formed, however, was highly unstable and degraded outside of the mother liquor and, being one of the most prominent examples of a polyproline helix utilised in an organic framework or similar structure, demonstrates the further work required to develop effective peptidic-units for the design of desirable and useful supramolecular constructs.

1.1.5 Polyproline Helix Design Considerations:

The incorporation of desirable functional groups into peptides can be readily achieved thanks to Fmoc-based solid-phase peptide synthesis (SPPS). This process can now be largely automated using a peptide synthesiser, through which the majority of synthetic peptides are now achieved. Fmoc-based amino acids are produced cheaply at an industrial scale and at ultra-high purity making them readily available. The stepwise synthesis relies on the use of a resin bound peptide allowing for sequential deprotection reactions, washes, and coupling reactions to produce peptide sequences, with microwave-assistance this gives incredibly short reaction times eliciting rapid synthesis of high purity peptides.[114] This modular synthesis also means that alternative amino acids can be easily swapped into the peptide sequence to produce different functionalities giving easy access to series of functionalised peptides for analysis.

As previously stated there exists a broad expanse of readily available Fmoc-protected amino acids for peptide synthesis, however, non-natural amino acids are less commercially available and can often require in-house synthesis. Polyproline helices are largely constituted of the amino acid proline and its derivatives (e.g. hydroxyproline), unique in that it is the only natural amino acid to have a secondary amine due to its cyclic side-chain and can much more freely adopt the *cis* isomer,[154] while also being tolerant to partial addition of some other amino acids. As such a significant focus for the synthesis of polyproline helices will be on the development of functionalised proline derivatives. Precursors to Fmoc-proline derivatives are readily available and easily functionalisable at the C_γ (C^4 position) (Figure 7) with numerous reported synthetic routes reported to achieve different functionalities.[173,174] Proline is also functionalisable at C_α and C_δ positions (Figure 7), however, these are both less reported and require more complex synthetic routes.[175] Due to the presence of the C^4 position on the faces of the polyproline helix (Figure 5 & Figure 7), facing perpendicular to the backbone, it also makes an ideal position for the incorporation of functional moieties designed for intermolecular interactions. This, combined with the pseudo-C3 rotational symmetry of the helix (Figure 5), makes the placement of intermolecular interacting groups when designing the 3-dimensional building-block a relatively simple endeavour. Some of the most typical avenues of functionalisation of the proline unit involve alkylation reactions, the widely reported Mitsunobu reaction on hydroxyproline derivatives (Figure 9), and carbon-carbon bond formation, with methods such as the Suzuki reaction.[173,176] While also utilising typical protecting group chemistry (e.g. Cbz, benzyl, Boc, *tert*-

butyl, and esters) to synthesise the Fmoc protected, free acid of proline required for SPPS Scheme.

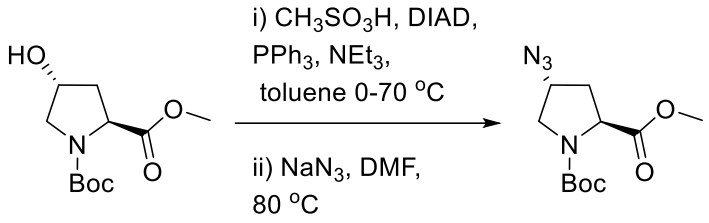


Figure 9 – Example of synthetic steps to incorporate the azide functional group onto the proline 4-position via the Mitsunobu reaction, adapted from work by Prof. Wennemers,[152]

Considerable interest has been expressed within this research field towards utilising minimalistic peptide units in the design of supramolecular constructs due to atom economy, cost, and ease of synthesis. The length of the peptide units can also have a significant impact on the stability of adopted secondary structures, with shorter peptides showing a less stable helicity yet having less conformational flexibility due to the size of the molecules, which is why short peptides (≤ 3 residues) are predominantly used in reported metal-peptide frameworks (MPFs).[17,21,123,172,177,178] However, longer peptides, which more readily adopt typical helical structures, exhibit greater conformational flexibility due to the large size of the molecules and their length also increases the cost and time of syntheses. As such, the peptide length is an important design consideration, with varied factors to consider depending on the desired use. In the case of polyproline the helix is stable to relatively short lengths when compared to other secondary structures,[123,157,172] with the accessibility of the polyproline I helix increasing with chain length. Typically, in literature 13-monomer chains and longer are reported to adopt either of the two polyproline helices depending on functionalisation and environment,[152,153,157,179] while the polyproline II helix can be adopted down to trimeric peptide chains.[123,172] Thus, to obtain switchable building-blocks slightly longer peptide units are required.

The relatively predictable chain length of the polyproline II helix can also be used to obtain fixed, predictable intermolecular distances within supramolecular constructs, this has been previously shown where the helix has been used as a spacer of variable length to obtain desired distances in molecular structures and between reactive sites on the polyproline helix.[136,148,180–182] This has clear applications for the rational design of supramolecular organic constructs with the functionalisation of specific residues on the faces of the helix, or at the termini of the peptide, achieving well defined distances between structural motifs (Figure 5). However, as is typical with supramolecular chemistry, the assembly of these peptide building blocks will vary significantly depending on the most significant factors driving either self-assembly or coordination, with the accessibility of thermodynamically and kinetically stable products. In contrast, functionalisation of the termini with interacting moieties can instead produce simpler linear ligands useful to achieve defined intermolecular distances within a supramolecular structure,[182] thanks to the specific peptide length of the polyproline helix.

1.1.6 Applications of Polyproline-based Supramolecular Constructs:

A wide variety of structures are potentially achievable with small variations in the placement of interacting moieties on the polyproline helices which, combined with the variety of functional groups, gives a massive range of potential supramolecular structures. As such, a narrow selection of desirable structures is ideal for a targeted initial investigation to provide an important proof of concept for future projects. Supramolecular organic frameworks (SOFs) are a promising initial option, through either self-assembly or co-crystallisation, guided by hydrogen bonding, π-π interactions, dispersion and Van der Waals interactions, which also allow for the use of the same or similar building blocks to be consequently taken forward for the formation of metal-organic frameworks (MOFs) or cages. The use of weaker interactions also allows for self-healing during the assembly processes, increasing the likelihood of obtaining single crystals. The self-assembly of peptide units to form useful supramolecular organic frameworks has been substantially reported, with examples of dipeptide based porous crystalline frameworks capable of selective gas absorption,[183–186] longer peptide-based frameworks capable of enantioselective separation of chiral polar drugs, self-assembly of peptides into 2D nanosheets,[187–191] and complex self-assembled structures such as molecular weaves,[10] nanocages,[182] super-helical assemblies,[126] and crystalline SOFs.[81] While examples of polyproline peptides used in the formation of supramolecular organic frameworks is more limited, however some examples exist of its use, such as the triaxial molecular weave,[10] a polyproline nanocage,[182] and a hexameric oligoproline MOF.[134] These exemplify the clear potential of polyproline based SOFs, however, further work is required to establish the ability to rationally design polyproline-based SOFs from the ideal model of a polyproline II helix using the helical faces as chemical handles for functionalisation. As such, investigation into the formation of polyproline-based SOFs is of immediate interest.

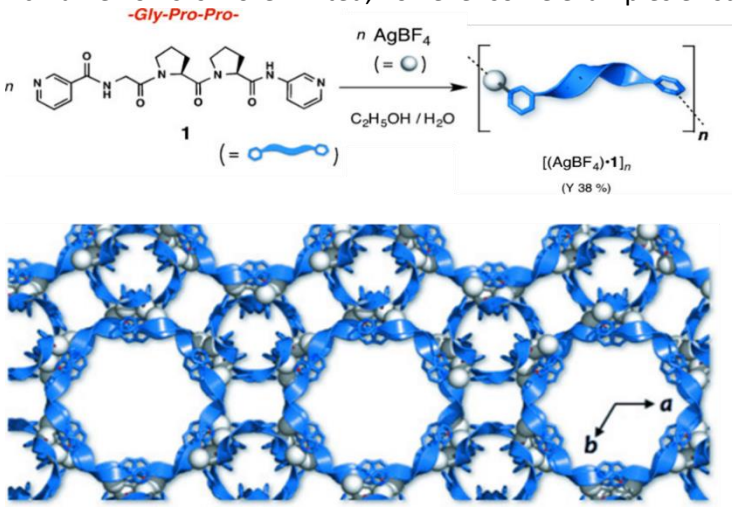


Figure 10 - Illustration adapted from a report by Prof. M Fujita; The crystallization of Gly-Pro-Pro ligand with $AgBF_4$ (top). Networked structure of $[(AgBF_4)\cdot\mathbf{1}]_n$ in the crystalline state. Py groups are represented as blue sticks, peptide helices as blue ribbons, and silver(I) ions as white spheres (bottom).[120]

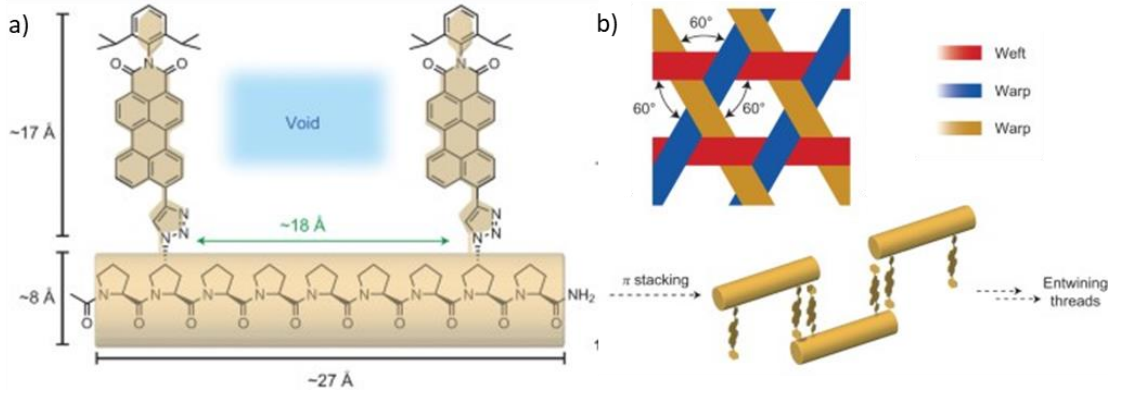


Figure 11 - Illustrations adapted from a report by Prof. H. Wennemer, showing the topologies and design principle of a triaxial weave utilising the predictable length of a polyproline helix; a) molecular structure of a peptide conjugated and molecular dimensions b) anticipated self-assembly into threads with alternating up-and-down voids for the formation of crossing points"[10].

The incorporation of peptide building-blocks into MOFs and SOFs to replace the typically synthetically challenging and bio-incompatible ligands normally utilized has immediate relevance to recent investigations, whereby this topic has seen a blooming interest within recent years with several notable successful findings.[17,111,134,139,192–194] However, as previously mentioned, this remains a significant challenge for investigators to achieve due to the flexibility and complexity of peptides compared to typical polyaromatic rigid ligands. With the hope that the unique polyproline helix may present more consistent results, a recent investigation by Wennemer's group has shown the formation of a porous crystalline metal-organic framework from an *N*-terminally functionalised hexameric polyproline helix with an unprecedented architecture of pleated nanosheets with metal ions aligned in strings,[134] demonstrating the potential success of incorporating peptidic building-blocks into MOFs to achieve unique structures. However, in this case the MOF formed was highly unstable and degrades outside of the mother liquor limiting the use for any potential applications, therefore, significant further investigation is required to develop the capability to rationally design these structures. Outside of MOFs an example of a supramolecular structure utilised a hexameric polyproline amphiphiles to achieve vesicle-like assemblies, with oppositely charge capping moieties on the peptide, that were fully reversible and has clear potential for the development of stimuli responsive drug carriers.[171] Again an additional case utilised terminal functionalisation of an oligoproline to achieve a supramolecular construct was recently shown in the synthesis of a self-assembled oligoproline nanocage, demonstrating the ability of the polyproline helix to be utilised as a rigid structural building-block.[182]

The utilisation of functionalisation of the proline residues to achieve specific geometries for supramolecular interactions has been demonstrated in SOFs[10,148,171] and to achieve specific functional group distances in other molecular structures,[136,137,147,182] however, this has yet to be attempted in the formation of any MOFs and is thus of particular interest. The potential of the ability to rationally design MOFs (or SOFs); with functionalisation of the polyproline helix with coordinating functionalities at specific angles to the peptide backbone (i.e. C3-symmetry) and predictable spacings between these groups (i → i+3 ca. 10 Å), while also containing freely functionalisable proline residues that are located between the linkers allowing for tuning of pore spaces for specific interactions, has massive implications for the design of functional materials to mimic the catalytic processes achieved by biological systems. A prime example of this ability to synthesise peptidic supramolecular frameworks with functionalisable pore spaces, due to the functionalisation of amino acid residues, was recently shown by the assembly of π-stacking helical peptides into a porous and multivariable "proteomimetic" framework (Figure 12).[81] Within this framework residues could be freely switched out with alternate amino acids containing varied sidechains while retaining the original supramolecular framework, and the framework could thus be tuned to increase affinity toward a complex organic molecule within the pores. Previous examples of porous peptidic frameworks with applications for host-guest chemistry have mostly been limited to short chain lengths,[81] while this example highlights the exciting potential of longer peptide chains to achieve more complex and tuneable porous frameworks. With this being one of the very few examples of an assembly of this kind, a similar achievement to synthesise a polyproline-based system is highly desirable and would act as a significant advancement for the field, while recent work by other groups has clearly been making progress towards polyproline based frameworks showing the clear scientific interest and as such the synthesis of these frameworks would not be a trivial achievement.

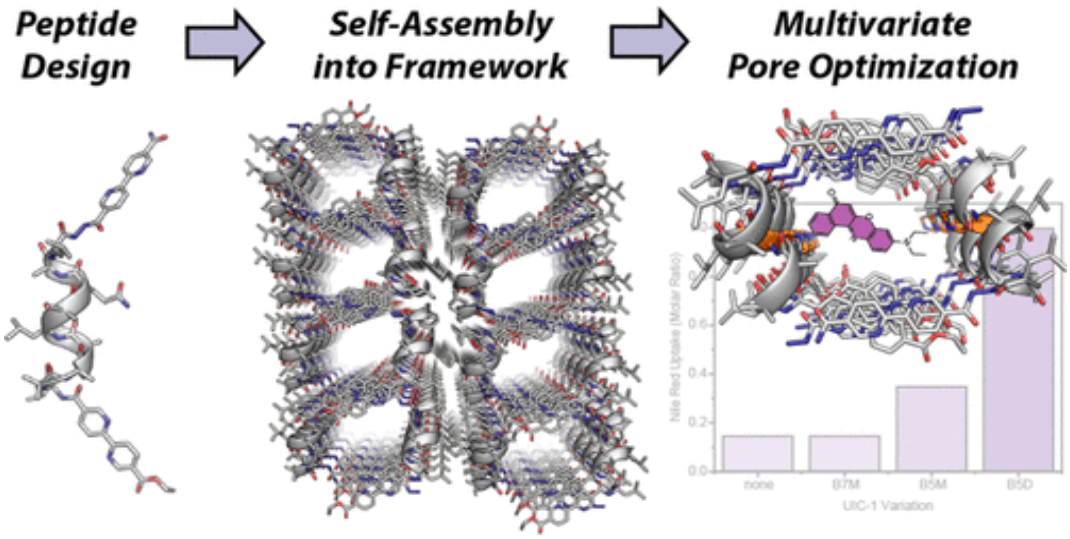


Figure 12 - Illustration from a recent report by Dr A. Nguyen, demonstrating the design and assembly of π-stacking helical peptides into porous proteomimetic frameworks, capable of optimisation to increase affinity towards complex a complex organic molecule[81]

1.2 Conclusion

Existing supramolecular materials have a major limitation associated with their poor biocompatibility and lack of chemical complexity. While bioinspired building blocks now show clear promise in supramolecular chemistry for the development of complex biocompatible materials. With an ever-growing requirement for the development of highly specific and cooperative functions, to mimic the capabilities of biological systems, the massive potential of these biomimetic supramolecular assemblies is undeniable. While peptides are being increasingly incorporated into the synthesis of novel materials with highly promising results,[81] polyproline helices are as yet minimally investigated but conversely represent an ideal candidate for the development of designed supramolecular constructs. These helices are highly tuneable, inherently chiral with well-defined periodicity, show excellent stability and are easily functionalised. Recent auspicious examples of the utilisation of the helices in supramolecular constructs,[120,134,182] demonstrate the significant untapped potential for these materials. As such, a clear goal should be to pave the way towards the discovery of novel, responsive materials that can be rationally designed from these modular and tuneable building-blocks to achieve bespoke properties for highly specific applications. Achieving this requires a better understanding of the assembly of these polyproline helices within supramolecular 3D materials and how these can be better designed and functionalised to achieve specific topologies and functions.

1.3 References

1 D. L. Caulder and K. N. Raymond, *Acc. Chem. Res.*, 1999, **32**, 975–982.

2 B. Olenyuk, J. A. Whiteford, A. Fechtenkötter and P. J. Stang, *Nature*, 1999, **398**, 796–799.

3 L. Jiao, Y. Wang, H.-L. Jiang and Q. Xu, *Adv. Mater.*, 2018, **30**, 1703663.

4 S. L. Anderson and K. C. Stylianou, *Coord. Chem. Rev.*, 2017, **349**, 102–128.

5 J. Lü, C. Perez-Krap, F. Trousselet, Y. Yan, N. H. Alsmail, B. Karadeniz, N. M. Jacques, W. Lewis, A. J. Blake, F. X. Coudert, R. Cao and M. Schröder, *Cryst. Growth Des.*, 2018, **18**, 2555–2562.

6 J. Tian, L. Chen, D. W. Zhang, Y. Liu and Z. T. Li, *Chem. Commun.*, 2016, **52**, 6351–6362.

7 S. Xiong, L. Li, L. Dong, J. Tang, G. Yu and C. Pan, *J. CO2 Util.*, 2020, **41**, 101224.

8 J. Jiang, Y. Zhao and O. M. Yaghi, *J. Am. Chem. Soc.*, 2016, **138**, 3255–3265.

9 M. Ghosh, S. Bera, S. Schiffmann, L. J. W. Shimon and L. Adler-Abramovich, *ACS Nano*, 2020, **14**, 9990–10000.

10 U. Lewandowska, W. Zajaczkowski, S. Corra, J. Tanabe, R. Borrmann, E. M. Benetti, S. Stappert, K. Watanabe, N. A. K. Ochs, R. Schaeublin, C. Li, E. Yashima, W. Pisula, K. Müllen and H. Wennemers, *Nat. Chem.*, 2017, **9**, 1068–1072.

11 J. L. Beesley, H. E. Baum, L. R. Hodgson, P. Verkade, G. S. Banting and D. N. Woolfson, *Nano Lett.*, 2018, **18**, 5933–5937.

12 J. A. Villegas, N. J. Sinha, N. Teramoto, C. D. Von Bargen, D. J. Pochan and J. G. Saven, *Molecules*, 2022, **27**, 1237.

13 X. Zhang, C. Gong, O. U. Akakuru, Z. Su, A. Wu and G. Wei, *Chem. Soc. Rev.*, 2019, **48**, 5564–5595.

14 D. G. Fatouros, D. A. Lamprou, A. J. Urquhart, S. N. Yannopoulos, I. S. Vizirianakis, S. Zhang and S. Koutsopoulos, *ACS Appl. Mater. interfacesurfaces*, 2014, **6**, 8184–8189.

15 C. Zhao, H. Chen, F. Wang and X. Zhang, *Colloids Surfaces B Biointerfaces*, 2021, **208**, 112040.

16 D. Yang, S. Gao, Y. Fang, X. Lin, X. Jin, X. Wang, L. Ke and K. Shi, *Nanomedicine*, 2018, **13**, 3159–3177.

17 J. Navarro-Sánchez, A. I. Argente-García, Y. Moliner-Martínez, D. Roca-Sanjuán, D. Antypov, P. Campíns-Falcó, M. J. Rosseinsky and C. Martí-Gastaldo, *J. Am. Chem. Soc.*, 2017, **139**, 4294–4297.

18 Y. Lu, H. Zhang, Y. Zhu, P. J. Marriott and H. Wang, *Adv. Funct. Mater.*, 2021, **31**, 2101335.

19 X. Chen, Y. Wang, H. Wang, Y. Kim and M. Lee, *Chem. Commun.*, 2017, **53**, 10958–10961.

20 A. Palma, M. Artelsmair, G. Wu, X. Lu, S. J. Barrow, N. Uddin, E. Rosta, E. Masson and O. A. Scherman, *Angew. Chemie*, 2017, **129**, 15894–15898.

21 A. Saito, T. Sawada and M. Fujita, *Angew. Chemie Int. Ed.*, 2020, **59**, 20367–20370.

22 P. D. Frischmann, V. Kunz and F. Würthner, *Angew. Chemie - Int. Ed.*, 2015, **54**, 7285–7289.

23 S. Shanmugaraju and P. S. Mukherjee, *Chem. - A Eur. J.*, 2015, **21**, 6656–6666.

24 H. Wang, B. Li, H. Wu, T. L. Hu, Z. Yao, W. Zhou, S. Xiang and B. Chen, *J. Am. Chem. Soc.*, 2015, **137**, 9963–9970.

25 W. Yang, A. Greenaway, X. Lin, R. Matsuda, A. J. Blake, C. Wilson, W. Lewis, P. Hubberstey, S. Kitagawa, N. R. Champness and M. Schröder, *J. Am. Chem. Soc.*, 2010, **132**, 14457–14469.

26 B. S. Pilgrim and N. R. Champness, *Chempluschem*, 2020, **85**, 1842–1856.

27 A. J. Gosselin, C. A. Rowland and E. D. Bloch, *Chem. Rev.*, 2020, **120**, 8987–9014.

28 T. R. Cook, Y. R. Zheng and P. J. Stang, *Chem. Rev.*, 2013, **113**, 734–777.

29 H. Furukawa, K. E. Cordova, M. O'Keeffe and O. M. Yaghi, *Science (80-.).*

30 O. M. Yaghi, M. O'Keeffe, N. W. Ockwig, H. K. Chae, M. Eddaoudi and J. Kim, *Nature*, 2003, **423**, 705–714.

31 Z. Chen, P. Li, X. Zhang, M. R. Mian, X. Wang, P. Li, Z. Liu, M. O'Keeffe, J. F. Stoddart and O. K. Farha, *Nano Res.*, 2021, **14**, 376–380.

32 P. Li, N. A. Vermeulen, C. D. Malliakas, D. A. Gómez-Gualdrón, A. J. Howarth, B. L. Mehdi, A. Dohnalkova, N. D. Browning, M. O'Keeffe and O. K. Farha, *Science (80-.).*, 2017, **356**, 624–627.

33 H. K. Chae, D. Y. Siberio-Pérez, J. Kim, Y. B. Go, M. Eddaoudi, A. J. Matzger, M. O'Keeffe and O. M. Yaghi, *Nature*, 2004, **427**, 523–527.

34 D. Feng, W. C. Chung, Z. Wei, Z. Y. Gu, H. L. Jiang, Y. P. Chen, D. J. Darensbourg and H. C. Zhou, *J. Am. Chem. Soc.*, 2013, **135**, 17105–17110.

35 R. B. Lin, F. Li, S. Y. Liu, X. L. Qi, J. P. Zhang and X. M. Chen, *Angew. Chemie Int. Ed.*, 2013, **52**, 13429–13433.

36 D. Feng, Z. Y. Gu, J. R. Li, H. L. Jiang, Z. Wei and H. C. Zhou, *Angew. Chemie Int. Ed.*, 2012, **51**, 10307–10310.

37 S. Y. Moon, Y. Liu, J. T. Hupp and O. K. Farha, *Angew. Chemie Int. Ed.*, 2015, **54**, 6795–6799.

38 M. P. Suh, H. J. Park, T. K. Prasad and D. W. Lim, *Chem. Rev.*, 2012, **112**, 782–835.

39 J. R. Li, J. Sculley and H. C. Zhou, *Chem. Rev.*, 2012, **112**, 869–932.

40 L. E. Kreno, K. Leong, O. K. Farha, M. Allendorf, R. P. Van Duyne and J. T. Hupp, *Chem. Rev.*, 2012, **112**, 1105–1125.

41 P. Horcajada, R. Gref, T. Baati, P. K. Allan, G. Maurin, P. Couvreur, G. Férey, R. E. Morris and C. Serre, *Chem. Rev.*, 2012, **112**, 1232–1268.

42 K. Sumida, D. L. Rogow, J. A. Mason, T. M. McDonald, E. D. Bloch, Z. R. Herm, T. H. Bae and J. R. Long, *Chem. Rev.*, 2012, **112**, 724–781.

43 T. R. Cook and P. J. Stang, *Chem. Rev.*, 2015, **115**, 7001–7045.

44 D. Samanta, S. Mukherjee, Y. P. Patil and P. S. Mukherjee, *Chem. – A Eur. J.*, 2012, **18**, 12322–12329.

45 T. H. Noh, W. Hong, H. Lee and O. S. Jung, *Dalt. Trans.*, 2014, **44**, 787–794.

46 H. Li, Y. F. Han, Y. J. Lin, Z. W. Guo and G. X. Jin, *J. Am. Chem. Soc.*, 2014, **136**, 2982–2985.

47 T. R. Cook, V. Vajpayee, M. H. Lee, P. J. Stang and K. W. Chi, *Acc. Chem. Res.*, 2013, **46**, 2464–2474.

48 N. Ahmad, H. A. Younus, A. H. Chughtai and F. Verpoort, *Chem. Soc. Rev.*, 2014, **44**, 9–25.

49 M. J. Bojdys, M. E. Briggs, J. T. A. Jones, D. J. Adams, S. Y. Chong, M. Schmidtmann and A. I. Cooper, *J. Am. Chem. Soc.*, 2011, **133**, 16566–16571.

50 A. P. Côté, A. I. Benin, N. W. Ockwig, M. O'Keeffe, A. J. Matzger and O. M. Yaghi, *Science (80-.).*, 2005, **310**, 1166–1170.

51 X. Guan, F. Chen, Q. Fang and S. Qiu, *Chem. Soc. Rev.*, 2020, **49**, 1357–1384.

52 H. Furukawa and O. M. Yaghi, *J. Am. Chem. Soc.*, 2009, **131**, 8875–8883.

53 Q. Fang, S. Gu, J. Zheng, Z. Zhuang, S. Qiu and Y. Yan, *Angew. Chemie Int. Ed.*, 2014, **53**, 2878–2882.

54 S. Wan, J. Guo, J. Kim, H. Ihee and D. Jiang, *Angew. Chemie Int. Ed.*, 2008, **47**, 8826–8830.

55 P. Bhandari and P. S. Mukherjee, *ACS Catal.*, 2023, 6126–6143.

56 Z. Lin, J. Sun, B. Efremovska and R. Warmuth, *Chem. – A Eur. J.*, 2012, **18**, 12864–12872.

57 M. Mastalerz, M. W. Schneider, I. M. Oppel and O. Presly, *Angew. Chemie Int. Ed.*, 2011, **50**, 1046–1051.

58 C. Ji, K. Su, W. Wang, J. Chang, E. S. M. El-Sayed, L. Zhang and D. Yuan, *CCS Chem.*, 2022, **41**, 3094–3104.

59 P. S. Huang, C. H. Kuo, C. C. Hsieh and Y. C. Horng, *Chem. Commun.*, 2012, **48**, 3227–3229.

60 C. Zhang, Q. Wang, H. Long and W. Zhang, *J. Am. Chem. Soc.*, 2011, **133**, 20995–21001.

61 P. Li, M. R. Ryder and J. F. Stoddart, *Accounts Mater. Res.*, 2020, **1**, 77–87.

62 D. Meng, J. L. Yang, C. Xiao, R. Wang, X. Xing, O. Kocak, G. Aydin, I. Yavuz, S. Nuryyeva, L. Zhang, G. Liu, Z. Li, S. Yuan, Z. K. Wang, W. Wei, Z. K. Wang, K. N. Houk and Y. Yang, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 20397–20403.

63 G. L. Li, Z. Zhuo, B. Wang, X. L. Cao, H. F. Su, W. Wang, Y. G. Huang and M. Hong, *J. Am. Chem. Soc.*, 2021, **143**, 10920–10929.

64 J. M. Fletcher, R. L. Harniman, F. R. H. Barnes, A. L. Boyle, A. Collins, J. Mantell, T. H. Sharp, M. Antognozzi, P. J. Booth, N. Linden, M. J. Miles, R. B. Sessions, P. Verkade and D. N. Woolfson, *Science (80-.).*, 2013, **340**, 595–599.

65 J. F. Ross, A. Bridges, J. M. Fletcher, D. Shoemark, D. Alibhai, H. E. V. Bray, J. L. Beesley, W. M. Dawson, L. R. Hodgson, J. Mantell, P. Verkade, C. M. Edge, R. B. Sessions, D. Tew and D. N. Woolfson, *ACS Nano*, 2017, **11**, 7901–7914.

66 C. Morris, S. J. Glennie, H. S. Lam, H. E. Baum, D. Kandage, N. A. Williams, D. J. Morgan, D. N. Woolfson, A. D. Davidson, C. Morris, H. E. Baum, D. N. Woolfson BrisSynBio, D. N. Woolfson, S. J. Glennie, H. S. Lam, D. Kandage, N. A. Williams, D. J. Morgan and A. D. Davidson, *Adv. Funct. Mater.*, 2019, **29**, 1807357.

67 I. Hisaki, Y. Suzuki, E. Gomez, Q. Ji, N. Tohnai, T. Nakamura and A. Douhal, *J. Am. Chem. Soc.*, 2019, **141**, 2111–2121.

68 C. Chen, Z. Di, H. Li, J. Liu, M. Wu and M. Hong, *CCS Chem.*, 2022, **4**, 1315–1325.

69 Y. He, S. Xiang and B. Chen, *J. Am. Chem. Soc.*, 2011, **133**, 14570–14573.

70 J. Lü, C. Perez-Krap, M. Suyetin, N. H. Alsmail, Y. Yan, S. Yang, W. Lewis, E. Bichoutskaia, C. C. Tang, A. J. Blake, R. Cao and M. Schröder, *J. Am. Chem. Soc.*, 2014, **136**, 12828–12831.

71 T. H. Chen, I. Popov, W. Kaveevivitchai, Y. C. Chuang, Y. S. Chen, O. Daugulis, A. J. Jacobson and O. Miljanić, *Nat. Commun.*, 2014, **5**, 1–8.

72 P. Li, Y. He, J. Guang, L. Weng, J. C. G. Zhao, S. Xiang and B. Chen, *J. Am. Chem. Soc.*, 2014, **136**, 547–549.

73 W. Gong, D. Chu, H. Jiang, X. Chen, Y. Cui and Y. Liu, *Nat. Commun.*, 2019, **10**, 1–9.

74 Z. Sun, Y. Li, L. Chen, X. Jing and Z. Xie, *Cryst. Growth Des.*, 2015, **15**, 542–545.

75 W. Liang, F. Carraro, M. B. Solomon, S. G. Bell, H. Amenitsch, C. J. Sumby, N. G. White, P. Falcaro and C. J. Doonan, *J. Am. Chem. Soc.*, 2019, **141**, 14298–14305.

76 Q. Yin, P. Zhao, R. J. Sa, G. C. Chen, L. Jian, T. F. Liu and R. Cao, *Angew. Chemie Int. Ed.*, 2018, **57**, 7691–7696.

77 A. Comotti, S. Bracco, A. Yamamoto, M. Beretta, T. Hirukawa, N. Tohnai, M. Miyata and P. Sozzani, *J. Am. Chem. Soc.*, 2014, **136**, 618–621.

78 J. H. Deng, J. Luo, Y. L. Mao, S. Lai, Y. N. Gong, D. C. Zhong and T. B. Lu, *Sci. Adv.*, 2020, **6**, 9976.

79 P. A. Sontz, J. B. Bailey, S. Ahn and F. A. Tezcan, *J. Am. Chem. Soc.*, 2015, **137**, 11598–11601.

80 Y. Wang, W. Qi, R. Huang, X. Yang, M. Wang, R. Su and Z. He, *J. Am. Chem. Soc.*, 2015, **137**, 7869–7880.

81 S. L. Heinz-Kunert, A. Pandya, V. T. Dang, P. N. Tran, S. Ghosh, D. McElheny, B. D. Santarsiero, Z. Ren and A. I. Nguyen, *J. Am. Chem. Soc.*, 2022, **144**, 7001–7009.

82 M. Y. Said, C. S. Kang, S. Wang, W. Sheffler, P. J. Salveson, A. K. Bera, A. Kang, H. Nguyen, R. Ballard, X. Li, H. Bai, L. Stewart, P. Levine and D. Baker, *Chem. Mater.*, 2022, **34**, 9736–9744.

83 V. V. Welborn and T. Head-Gordon, *Chem. Rev.*, 2019, **119**, 6613–6630.

84 W.-L. Li and T. Head-Gordon, *ACS Cent. Sci.*, 2021, **7**, 72–80.

85 T. K. Piskorz, V. Martí-Centelles, T. A. Young, P. J. Lusby and F. Duarte, *ACS Catal.*, 2022, **12**, 5806–5826.

86 R. Wolfenden and M. J. Snider, *Acc. Chem. Res.*, 2001, **34**, 938–945.

87 M. J. Byrne, N. R. Lees, L.-C. Han, M. W. Van Der Kamp, A. J. Mulholland, J. E. M. Stach, C. L. Willis and P. R. Race, *J. Am. Chem. Soc.*, 2016, **138**, 6095–6098.

88 C. Li, K. Deng, Z. Tang and L. Jiang, *J. Am. Chem. Soc.*, 2010, **132**, 8202–8209.

89 H. S. Wang, Y. H. Wang and Y. Ding, *Nanoscale Adv.*, 2020, **2**, 3788–3797.

90 H. Cai, Y. L. Huang and D. Li, *Coord. Chem. Rev.*, 2019, **378**, 207–221.

91 B. Sun, M. Bilal, S. Jia, Y. Jiang and J. Cui, *Korean J. Chem. Eng*, 2019, **36**, 1949–1964.

92 S. Chand, O. Alahmed, W. S. Baslyman, A. Dey, S. Qutub, R. Saha, Y. Hijikata, M. Alaamery and N. M. Khashab, *JACS Au*, 2022, **2**, 623–630.

93 V. Subramaniyam, P. V. Ravi and M. Pichumani, *J. Mol. Struct.*, 2022, **1251**, 131931.

94 S. Rojas, T. Devic and P. Horcajada, *J. Mater. Chem. B*, 2017, **5**, 2560–2573.

95 B. Sun, M. Bilal, S. Jia, Y. Jiang and J. Cui, *Korean J. Chem. Eng.*, 2019, **36**, 1949–1964.

96 J. Baillet, V. Desvergnes, A. Hamoud, L. Latxague, P. Barthélémy, S. J. Assemblies Baillet, V. Desvergnes, A. Hamoud, L. Latxague and P. Barthélémy, *Adv. Mater.*, 2018, **30**, 1705078.

97 X. Zhang, C. Gong, O. U. Akakuru, Z. Su, A. Wu and G. Wei, *Chem. Soc. Rev.*, 2019, **48**, 5564–5595.

98 Y. Huo, J. Hu, Y. Yin, P. Liu, K. Cai and W. Ji, *ChemBioChem*, 2023, **24**, e202200582.

99 F. Sheehan, D. Sementa, A. Jain, M. Kumar, M. Tayarani-Najjaran, D. Kroiss and R. V. Ulijn, *Chem. Rev.*, 2021, **121**, 13869–13914.

100 S. Dey, C. Fan, K. V. Gothelf, J. Li, C. Lin, L. Liu, N. Liu, M. A. D. Nijenhuis, B. Saccà, F. C. Simmel, H. Yan and P. Zhan, *Nat. Rev. Methods Prim. 2021 11*, 2021, **1**, 1–24.

101 D. Morrison, M. Rothenbroker, Y. Li, D. Morrison, M. Rothenbroker and Y. Li, *Small Methods*, 2018, **2**, 1700319.

102 M. P. Hendricks, K. Sato, L. C. Palmer and S. I. Stupp, *Acc. Chem. Res.*, 2017, **50**, 2440–2448.

103 S. Yoon and J. J. Rossi, *Adv. Drug Deliv. Rev.*, 2018, **134**, 22–35.

104 B. R. Madhanagopal, S. Zhang, E. Demirel, H. Wady and A. R. Chandrasekaran, *Trends Biochem. Sci.*, 2018, **43**, 997–1013.

105 K. O. Ramberg, S. Engilberge, T. Skorek and P. B. Crowley, *J. Am. Chem. Soc.*, 2021, **143**, 1896–1907.

106 K. A. Cannon, R. U. Park, S. E. Boyken, U. Nattermann, S. Yi, D. Baker, N. P. King, T. O. Yeates, | Scott, E. Boyken, U. Nattermann, S. Yi, | David Baker, N. P. King, | Todd, O. Yeates and T. O. Yeates, *Protein Sci.*, 2019, **29**, 919–929.

107 Y. Zhang, M. S. Ardejani and B. P. Orner, *Chem. – An Asian J.*, 2016, **11**, 2814–2828.

108 N. P. King, W. Sheffler, M. R. Sawaya, B. S. Vollmar, J. P. Sumida, I. André, T. Gonen, T. O. Yeates and D. Baker, *Science (80-.).*, 2012, **336**, 1171–1174.

109 C. J. J. Wilson, A. S. Bommarius, J. A. Champion, Y. O. Chernoff, D. G. Lynn, A. K. Paravastu, C. Liang, M.-C. Hsieh and J. M. Heemstra, *Chem. Rev.*, 2018, **118**, 11519–11574.

110 J. Zhuang, A. P. Young and C. K. Tsung, *Small*, 2017, **13**, 1700880.

111 J. Dong, Y. Liu and Y. Cui, *J. Am. Chem. Soc.*, 2021, **143**, 17316–17336.

112 Z. J. Lin, J. Lü, M. Hong and R. Cao, *Chem. Soc. Rev.*, 2014, **43**, 5867–5895.

113 I. Imaz, M. Rubio-Martínez, J. An, I. Solé-Font, N. L. Rosi and D. Maspoch, *Chem. Commun.*, 2011, **47**, 7287–7302.

114 R. Behrendt, P. White and J. Offer, *J. Pept. Sci.*, 2016, **22**, 4–27.

115 F. Sheehan, D. Sementa, A. Jain, M. Kumar, M. Tayarani-Najjaran, D. Kroiss and R. V. Ulijn, *Chem. Rev.*, 2021, **121**, 13869–13914.

116 R. V. Ulijn and R. Jerala, *Chem. Soc. Rev.*, 2018, **47**, 3391–3394.

117 N. J. Sinha, M. G. Langenstein, D. J. Pochan, C. J. Kloxin and J. G. Saven, *Chem. Rev.*, 2021, **121**, 13915–13935.

118 N. Zigon, V. Duplan, N. Wada and M. Fujita, *Angew. Chemie Int. Ed.*, 2021, **60**, 25204–25222.

119 M. Fujita, J. Yazaki and K. Ogura, *J. Am. Chem. Soc.*, 1990, **112**, 5645–5647.

120 T. Sawada, A. Matsumoto and M. Fujita, *Angew. Chemie*, 2014, **126**, 7356–7360.

121 T. Sawada and M. Fujita, *Bull. Chem. Soc. Jpn.*, 2021, **94**, 2342–2350.

122 Y. Inomata, T. Sawada and M. Fujita, *J. Am. Chem. Soc.*, 2021, **143**, 16734–16739.

123 T. Sawada, A. Saito, K. Tamiya, K. Shimokawa, Y. Hisada and M. Fujita, *Nat. Commun. 2019 101*, 2019, **10**, 1–7.

124 T. Sawada and M. Fujita, *Chem*, 2020, **6**, 1861–1876.

125 S. Jeong, L. Zhang, J. Kim, J. Gong, J. Choi, K. M. Ok, Y. Lee, S. Kwon and H. S. Lee, *Angew. Chemie - Int. Ed.*, , DOI:10.1002/ANIE.202108364.

126 S. Mondal, L. Adler-Abramovich, A. Lampel, Y. Bram, S. Lipstman and E. Gazit, *Nat. Commun.*, 2015, **6**, 8615.

127 L. Adler-Abramovich and E. Gazit, *Chem. Soc. Rev.*, 2014, **43**, 6881–6893.

128 P. Makam and E. Gazit, *Chem. Soc. Rev.*, 2018, **47**, 3406–3420.

129 N. Kol, L. Adler-Abramovich, D. Barlam, R. Z. Shneck, E. Gazit and I. Rousso, *Nano Lett.*, 2005, **5**, 1343–1346.

130 A. Kholkin, N. Amdursky, I. Bdikin, E. Gazit and G. Rosenman, *ACS Nano*, 2010, **4**, 610–614.

131 Y. Chen, S. Guerin, H. Yuan, J. O'Donnell, B. Xue, P. A. Cazade, E. U. Haq, L. J. W. Shimon, S. Rencus-Lazar, S. A. M. Tofail, Y. Cao, D. Thompson, R. Yang and E. Gazit, *J. Am. Chem. Soc.*, 2022, **144**, 3468–3476.

132 Y. Chen, Y. Yang, A. A. Orr, P. Makam, B. Redko, E. Haimov, Y. Wang, L. J. W. Shimon, S. Rencus-Lazar, M. Ju, P. Tamamis, H. Dong and E. Gazit, *Angew. Chemie Int. Ed.*, 2021, **60**, 17164–17170.

133 A. Levin, T. A. Hakala, L. Schnaider, G. J. L. Bernardes, E. Gazit and T. P. J. Knowles, *Nat. Rev. Chem. 2020 411*, 2020, **4**, 615–634.

134 T. Schnitzer, E. Paenurk, N. Trapp, R. Gershoni-Poranne and H. Wennemers, *J. Am. Chem. Soc.*, 2021, **143**, 644–648.

135 P. Wilhelm, B. Lewandowski, N. Trapp and H. Wennemers, *J. Am. Chem. Soc.*, 2014, **136**, 15829–15832.

136 G. Upert, F. Bouillère and H. Wennemers, *Angew. Chemie Int. Ed.*, 2012, **51**, 4231–4234.

137 S. Dobitz, M. R. Aronoff and H. Wennemers, *Acc. Chem. Res.*, 2017, **50**, 2420–2428.

138 S. Gupta, I. Singh, A. K. Sharma and P. Kumar, *Front. Bioeng. Biotechnol.*, 2020, **0**, 504.

139 A. Mantion, L. Massüger, P. Rabu, C. Palivan, L. B. McCusker and A. Taubert, *J. Am. Chem. Soc.*, 2008, **130**, 2517–2526.

140 M. D. Ward, C. A. Hunter and N. H. Williams, *Acc. Chem. Res.*, 2018, **51**, 2073–2082.

141 A. Stephenson, S. P. Argent, T. Riis-Johannessen, I. S. Tidmarsh and M. D. Ward, *J. Am. Chem. Soc.*, 2011, **133**, 858–870.

142 A. B. Sainaba, M. Venkateswarulu, P. Bhandari, K. S. A. Arachchige, J. K. Clegg and P. S. Mukherjee, *J. Am. Chem. Soc.*, 2022, **144**, 7504–7513.

143 P. Morales and M. A. Jiménez, *Arch. Biochem. Biophys.*, 2019, **661**, 149–167.

144 Z. Shi, K. Chen, Z. Liu and N. R. Kallenbach, *Chem. Rev.*, 2006, **106**, 1877–1897.

145 A. L. Rucker and T. P. Creamer, *Protein Sci.*, 2002, **11**, 980–985.

146 C. Kroll, R. Mansi, F. Braun, S. Dobitz, H. R. Maecke and H. Wennemers, *J. Am. Chem. Soc.*, 2013, **135**, 16793–16796.

147 Y. A. Nagel, S. Raschle, H. Wennemers, Y. A. Nagel, P. S. Raschle and H. W. Ennemers, *Angew. Chemie - Int. Ed.*, 2017, **56**, 22–126.

148 U. Lewandowska, W. Zajaczkowski, L. Chen, F. Bouillière, D. Wang, K. Koynov, W. Pisula, K. Müllen and H. Wennemers, *Angew. Chemie Int. Ed.*, 2014, **53**, 12537–12541.

149 E. Sobakinskaya, M. Schmidt am Busch and T. Renger, *J. Phys. Chem. B*, 2018, **122**, 54–67.

150 J. R. Stringer, J. A. Crapster, I. A. Guzei and H. E. Blackwell, *J. Am. Chem. Soc.*, 2011, **133**, 15559–15567.

151 M. Kümin, L.-S. Sonntag and H. Wennemers, *J. Am. Chem. Soc.*, 2007, **129**, 466–467.

152 E. Beausoleil and W. D. Lubell, *Biopolymers*, 2000, **53**, 249–256.

153 S. L, H. AE, G. MS, E. MA, R. DH and C. DE, *J. Am. Soc. Mass Spectrom.*, 2016, **27**, 22–30.

154 Y. K. Kang and H. Young Choi, *Biophys. Chem.*, 2004, **111**, 135–142.

155 K. P. Lu, G. Finn, T. H. Lee and L. K. Nicholson, *Nat. Chem. Biol.*, 2007, **3**, 619–629.

156 Y. K. Kang, J. S. Jhon and H. S. Park, *J. Phys. Chem. B*, 2006, **110**, 17645–17655.

157 S. Kakinoki, Y. Hirano and M. Oka, *Polym. Bull.*, 2005, **53**, 109–115.

158 T. J. El-Baba, D. R. Fuller, D. A. Hales, D. H. Russell and D. E. Clemmer, *J. Am. Soc. Mass Spectrom.*, 2019, **30**, 77.

159 S. L, H. AE, K. N, R. DH and C. DE, *J. Am. Chem. Soc.*, 2015, **137**, 8680–8683.

160 R. Dukor and T. Keiderling, *Biopolymers*, 1991, **31**, 1747–1761.

161 J. S. Jhon and Y. K. Kang, *J. Phys. Chem. B*, 2007, **111**, 3496–3507.

162 M. R. Aronoff, J. Egli, A. Schmitt and H. Wennemers, *Chem. – A Eur. J.*, 2020, **26**, 5070–5074.

163 Y.-C. Chiang, Y.-J. Lin and J.-C. Horng, *Protein Sci.*, 2009, **18**, 1967–1977.

164 G. Berger, M. Vilchis-Reyes and S. Hanessian, *Angew. Chemie Int. Ed.*, 2015, **54**, 13268–13272.

165 M. Kuemin, Y. A. Nagel, S. Schweizer, F. W. Monnard, C. Ochsenfeld and H. Wennemers, *Angew. Chemie - Int. Ed.*, 2010, **49**, 6324–6327.

166 A. K. Pandey, K. M. Thomas, C. R. Forbes and N. J. Zondlo, *Biochemistry*, 2014, **53**, 5307–5314.

167 N. A. Wenzell, H. K. Ganguly, A. K. Pandey, M. R. Bhatt, G. P. A. Yap and N. J. Zondlo, *ChemBioChem*, 2019, **20**, 963–967.

168 G. J. Bartlett, A. Choudhary, R. T. Raines and D. N. Woolfson, *Nat. Chem. Biol.*, 2010, **6**, 615–620.

169 A. Choudhary, R. W. Newberry and R. T. Raines, *Org. Lett.*, 2014, **16**, 3421–3423.

170 I. H. McColl, E. W. Blanch, L. Hecht, N. R. Kallenbach and L. D. Barron, *J. Am. Chem. Soc.*, 2004, **126**, 5076–5077.

171 C. Felip-León, F. Galindo, J. F. Miravet, V. Castelletto and I. W. Hamley, *J. Phys. Chem. B*, 2017, **121**, 7443–7446.

172 T. Sawada, A. Matsumoto and M. Fujita, *Angew. Chemie*, 2014, **126**, 7356–7360.

173 A. K. Pandey, D. Naduthambi, K. M. Thomas and N. J. Zondlo, *J. Am. Chem. Soc.*, 2013, **135**, 4333–4363.

174 K. M. Thomas, D. Naduthambi, G. Tririya and N. J. Zondlo, *Org. Lett.*, 2005, **7**, 2397–2400.

175 C. Mothes, C. Caumes, A. Guez, H. Boullet, T. Gendrineau, S. Darses, N. Delsuc, R. Moumné, B. Oswald, O. Lequin and P. Karoyan, *Molecules*, 2013, **18**, 2307–2327.

176 K. C. K. Swamy, N. N. B. Kumar, E. Balaraman and K. V. P. P. Kumar, *Chem. Rev.*, 2009, **109**, 2551–2651.

177 R. Miyake, *J. Incl. Phenom. Macrocycl. Chem.*, 2022, **102**, 711–722.

178 A. P. Katsoulidis, K. S. Park, D. Antypov, C. Martí-Gastaldo, G. J. Miller, J. E. Warren, C. M. Robertson, F. Blanc, G. R. Darling, N. G. Berry, J. A. Purton, D. J. Adams and M. J. Rosseinsky, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 193.

179 J.-C. Horng, *Protein Sci.*, 2006, **15**, 74–83.

180 R. B. Best, K. A. Merchant, I. V. Gopich, B. Schuler, A. Bax and W. A. Eaton, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 18964–18969.

181 P.-Y. Hung, Y.-H. Chen, K.-Y. Huang, C.-C. Yu and J.-C. Horng, *ACS Omega*, 2017, **2**, 5574–5581.

182 S. Matsubara, Y. Okamoto, M. Yoshikawa, S. Tsukiji and M. Higuchi, *Bioconjug. Chem.*, 2022, **33**, 1785–1788.

183 V. N. Yadav, A. Comotti, P. Sozzani, S. Bracco, T. Bonge-Hansen, M. Hennum and C. H. Görbitz, *Angew. Chemie Int. Ed.*, 2015, **54**, 15684–15688.

184 A. Comotti, S. Bracco, G. Distefano and P. Sozzani, *Chem. Commun.*, 2009, 284–286.

185 C. H. Görbitz, *Chem. – A Eur. J.*, 2007, **13**, 1022–1031.

186 A. Comotti, A. Fraccarollo, S. Bracco, M. Beretta, G. Distefano, M. Cossi, L. Marchese, C. Riccardi and P. Sozzani, *CrystEngComm*, 2013, **15**, 1503–1507.

187 E. L. Magnotti, S. A. Hughes, R. S. Dillard, S. Wang, L. Hough, A. Karumbamkandathil, T. Lian, J. S. Wall, X. Zuo, E. R. Wright and V. P. Conticello, *J. Am. Chem. Soc.*, 2016, **138**, 16274–16282.

188 T. Jiang, O. A. Vail, Z. Jiang, X. Zuo and V. P. Conticello, *J. Am. Chem. Soc.*, 2015, **137**, 7793–7802.

189 B. Dai, D. Li, W. Xi, F. Luo, X. Zhang, M. Zou, M. Cao, J. Hu, W. Wang, G. Wei, Y. Zhang and C. Liua, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 2996–3001.

190 I. W. Hamley, A. Dehsorkhi and V. Castelletto, *Chem. Commun.*, 2013, **49**, 1850–1852.

191 A. S. Parmar, J. K. James, D. R. Grisham, D. H. Pike and V. Nanda, *J. Am. Chem. Soc.*, 2016, **138**, 4362–4367.

192 R. Misra, A. Saseendran, S. Dey and H. N. Gopi, *Angew. Chemie*, 2019, **131**, 2273–2277.

193 C. Martí-Gastaldo, D. Antypov, J. E. Warren, M. E. Briggs, P. A. Chater, P. V. Wiper, G. J. Miller, Y. Z. Khimyak, G. R. Darling, N. G. Berry and M. J. Rosseinsky, *Nat. Chem.*, 2014, **6**, 343–351.

194 F. G. Cirujano, N. Martín, N. Almora-Barrios and C. Martí-Gastaldo, *Catal. Sci. Technol.*, 2021, **11**, 6053–6057.

Chapter 2.

Synthesis of Polyproline-based Supramolecular Peptide Frameworks

Investigation adapted, and expanded, from publications:

Brightwell, D. F.; Truccolo, G.; Samanta, K.; Fenn, E. J.; Holder, S. J.; Shepherd, H. J.; Hawes, C. S.; Palma, A. A Reversibly Porous Supramolecular Peptide Framework. *Chem. – A Eur. J.* **2022**, *28* (66), e202202368. https://doi.org/10.1002/CHEM.202202368

Brightwell, D. F.; Truccolo, G.; Samanta, K.; Shepherd, H. J.; Palma, A. Supramolecular Self-Assembly of Engineered Polyproline Helices. *ACS Macro Lett.* **2023**, *12*, 908–914. https://doi.org/10.1021/acsmacrolett.3c00304

2.1 A Reversibly Porous Supramolecular Peptide Framework

2.1.1 Abstract

The ability to use bio-inspired building-blocks in the assembly of novel supramolecular frameworks is at the forefront of this exciting research field. Herein, we present the first polyproline helix to self-assemble into a reversibly porous, crystalline, supramolecular peptide framework (SPF). This framework is assembled from a short oligoproline, adopting the polyproline II conformation, driven by hydrogen-bonding and dispersion interactions. Thermal activation, guest-induced dynamic porosity and enantioselective guest inclusion have been demonstrated for this novel system. The principles of the self-assembly associated with this SPF will be used as a blueprint allowing for the further development of helical peptide linkers in the rational design of SPFs and metal-peptide frameworks.

2.1.2 Introduction

Peptide based porous frameworks are an emerging class of materials, which have found applications as adaptive[1–3] and reversibly tuneable[2] porous materials, capable of capturing greenhouse gases,[4–6] and, due to their inherent chirality, have found applications in facilitating chiral events and *trans*formations,[7] and separating chiral drugs[8,9]. Peptides can be prepared at scale with high purity, have canonical and non-canonical amino acids incorporated into their primary structure with high accuracy and are biocompatible. As such, the efforts to investigate this class of compounds as chiral, tuneable ligands has seen a surge in recent years.[8–14] While efforts to post-functionalise metal-organic frameworks (MOFs) with helical peptides,[7,15] to induce chiral selectivity, and to synthesise a broad range of metal-peptide frameworks (MPFs) have seen considerable advancements,[8–10,12,16] it remains a challenge to develop peptide-based extended porous networks guided by supramolecular interactions (*e.g.* H-bonding, halogen bonding, π-interactions and host-guest interactions).[17–19] To the best of our knowledge, the use of cyclic peptides to yield porous channels in crystalline materials and the use of self-assembling hydrophobic dipeptides, are amongst the few examples of well characterised peptide-based frameworks (*i.e.* Supramolecular Peptide Frameworks; SPFs).[4,6,13,20,21] In recognition of the immense chemical space yet to be explored in this field, we set out to exploit peptides with stable secondary structures in the construction of SPFs. In particular, we focused our efforts on the use of polyproline helices as SPF building blocks due to the accessibility to rigid yet stimuli sensitive secondary structures.[22] Polyprolines can interconvert between two different secondary conformations, polyproline II with all *trans* amide bonds and polyproline I with all *cis* amide bonds. This interconversion can be controlled as a function of the environment it is exposed to (*i.e.* temperature, solvent polarity and pH) and can be monitored *via* circular dichroism (CD) spectroscopy.[23–25] Also, polyproline helices have also shown remarkable resilience to a diverse range of functionalisation, in contrast to other helices used in supramolecular chemistry,[26–28] retaining their conformation even in short sequences.[29] These properties make them promising candidates for the synthesis of novel porous SPFs. Herein, we report the first reversibly porous SPF formed *via* the self-assembly of a helical tetraproline peptide (Figure 13). This supramolecular peptide framework is formed through the self-assembly of a short helical tetraproline *via* hydrogen-bonding and Fmoc-Fmoc interactions. The remarkably resilient SPF is capable of reversibly hosting guest molecules within the channels through the guest-induced reversion from a collapsed de-solvated state. The inherently chiral nature of the peptide framework results in enantioselectivity of this adsorption process. Wennemer's group reported the first crystal structure of a polyproline hexamer in the polyproline II conformation providing insight into the stability of the polyproline helix,[30] followed by Hanessian, who reported the crystal structure of the tetrameric proline congener (*cis*-4,5-methanoproline) in the polyproline II form.[31] Previously, a crystal structure of a tetrameric oligoproline was reported by Matsuzaki,

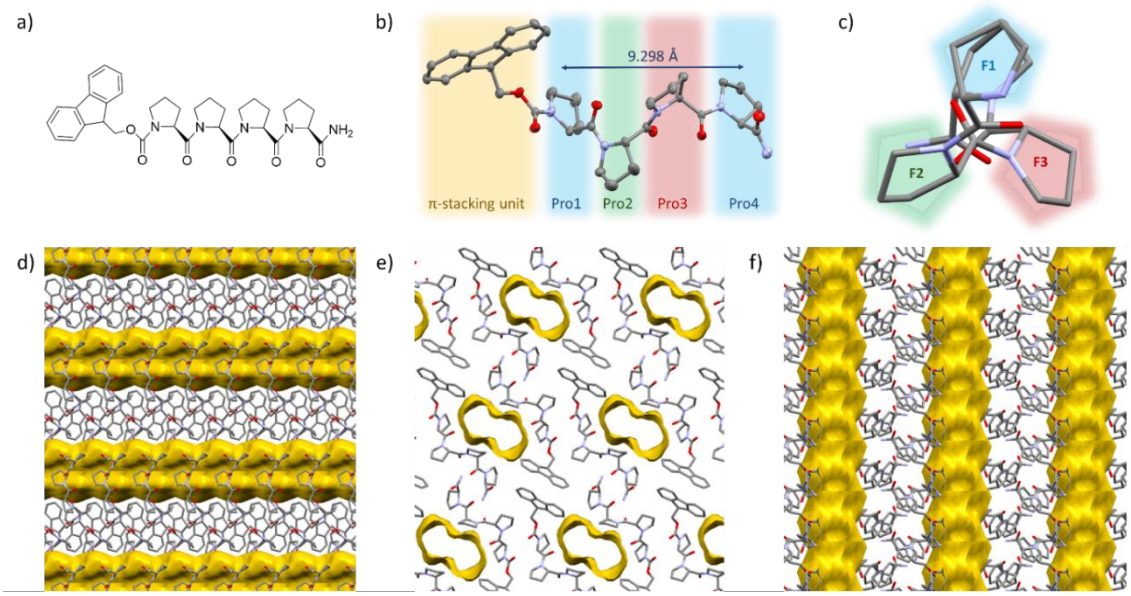


Figure 13 - a) Proline tetramer Fmoc-$(Pro)_4$-NH_2 (PP_4), b) Crystal structure of PP_4-SPF, 50% ellipsoids (Mercury), c) View along the axis displaying C3 symmetry and the 3 faces (F1-3) of the helix, Fmoc group removed for clarity, d) Crystal structure of PP4-SPF showing packing and solvent accessible voids (yellow) viewed along the a axis, ethanol molecules were removed, e) View along the b axis, f) View along the c axis.

with several of the proline units deviating from the typical polyproline II conformation.[32] However, reports have suggested that tetraproline can be found in the structured polyproline II helical conformation when in solution.[29,31] To the best of our knowledge, tetraproline is the shortest polyproline II helix reported in the literature,[29–31] therefore we decided to focus our initial efforts on this oligomer in the construction of structured SPFs.

2.1.3 Results and Discussion

2.1.3.1 Synthesis and Assembly of an Oligoproline Tetramer Framework

An oligoproline tetramer, **PP_4** was successfully synthesized using solid phase peptide methodology in a quantitative yield (SI 4.2). The *N*-terminus was capped with an Fmoc carbamate group to facilitate intramolecular assembly of peptides *via* interactions between the bulky aromatic groups, while the *C*-terminus was amidated to act as a hydrogen bond donor (**H_D**). **PP_4** was characterized in solution with far-UV circular dichroism (CD) analysis of aqueous, ethanol, and propan-1-ol solutions (0.25 mM), showing that the peptide retains the polyproline II helix in all solvents (λ_{max} at 223/229 nm and λ_{min} at 202 nm; SI 4.2.2, Figure 14).

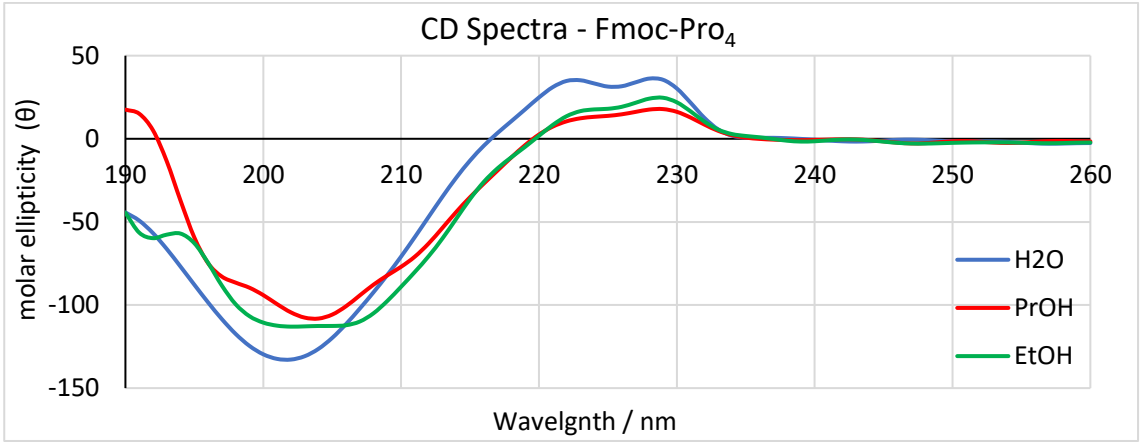


*Figure 14 – CD Spectrum overlay of **P_4** in water, ethanol, and propan-1-ol (0.25 mM) after incubating for 14 days.*

Crystals of **PP_4** suitable for single crystal X-ray diffraction analysis were reproducibly obtained *via* slow cooling of a hot super-saturated solution of **PP_4** in ethanol (SI 4.3.1). The structure (supramolecular peptide framework – SPF) was solved and refined to an atomic resolution of 0.81 Å (Figure 13) and confirmed that the peptide adopts the polyproline II helical conformation. A second structure was also obtained after flash freezing the sample (**SPF_{Flash}**). The characteristic dihedral angles, matching with the expected polyproline II helix, and an analysis of the puckering of the pyrrolidine rings, have been summarized in Table 2. We found that pyramidalization of the amide carbonyls (Δ) clearly indicated the presence of *N*-π* interactions and these interactions were more significant in prolines exhibiting the *exo* conformation with a larger degree of pyramidalization (Table 1) with only Pro3 showing minimal pyramidalization.[30,33]

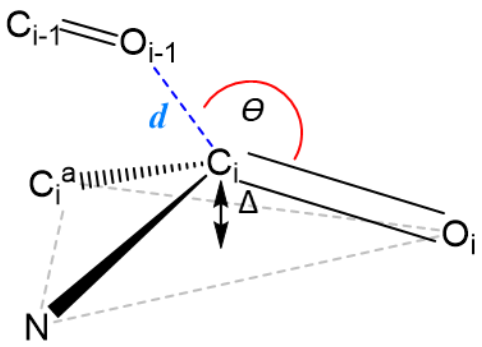


Figure 15 – Diagram depicting the Bürgi-Dunitz trajectory, angle (Θ) and pyramidalization (Δ), of the n-π interaction between adjacent carbonyl groups.*

Table 1. Trajectory angles, distances and pyramidalization of PP_4 carbonyls from PP_4-SPF and PP_4-SPF_{Flash} SC-XRD data

Residue	θ_{BD} / °		*d* / Å		Δ / Å	
	SPF	**SPF_{Flash}**	**SPF**	**SPF_{Flash}**	**SPF**	**SPF_{Flash}**
Pro1	98.7(2)	95.4(2)	3.204(3)	3.155(3)	0.019	0.018
Pro2	103.1(2)	**98.2(2)**	2.937(3)	**3.120(3)**	0.025	**0.018**
Pro3	86.9(2)	86.6(2)	3.059(3)	3.084(3)	0.004	0.002
Pro4	94.0(2)	95.3(2)	2.964(3)	2.912(3)	0.031	0.026

The polyproline II helix is stabilised by *N*-*π** interactions, between the carbonyl non-bonding orbitals to the next residue's carbonyl antibonding orbital following the Bürgi-Dunitz trajectory, this effect causes a degree of pyramidalization of the typically planar carbonyl (Figure 15). The greater the degree of pyramidalization and the closer to the typical Bürgi-Dunitz angle and distance, the greater the strength of the *N*-*π** interaction is likely to be.

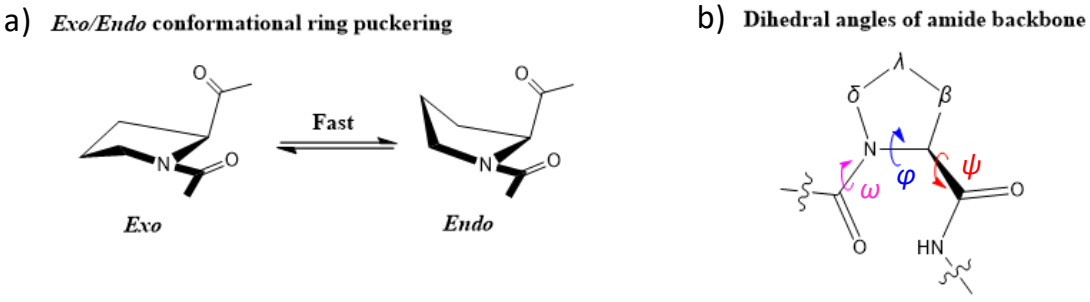


Figure 16 - a) diagram of conformational change due to endo/exo puckering of a proline ring, b) diagram showing the dihedral angles of a proline residue.

	Table 2. Helix dihedral angles and ring puckering of peptide 1 from PP_4-SPF and PP_4-SPF_{Flash} SC-XRD data				
Residue	**Crystal**	**ω/ °**	***φ/°***	***ψ / °***	**Ring Pucker**
Pro1	**SPF**	179.8(2)	-77.3(3)	154.5(2)	C^{β}-*exo*/C^{γ}-*endo* twist
	SPF_{Flash}	-175.4(2)	-74.7(3)	155.5(2)	C^{β}-*exo*/C^{γ}-*endo* twist
Pro2	**SPF**	179.5(2)	-65.4(3)	140.8(2)	**C^{γ}-*exo* envelope**
	SPF_{Flash}	-179.2(2)	-75.6(3)	154.2(2)	**C^{β}-*exo*/C^{γ}-*endo* twist**
Pro3	**SPF**	174.8(2)	-70.8(3)	163.5(2)	C^{β}-*exo*/C^{γ}-*endo* twist
	SPF_{Flash}	174.1(2)	-69.6(3)	163.2(2)	C^{β}-*exo*/C^{γ}-*endo* twist
Pro4	**SPF**	-173.3(2)	-65.1(3)	151.2(2)	C^{γ}-*exo* envelope
	SPF_{Flash}	-174.(2)	-62.1(3)	146.8(3)	C^{γ}-*exo* envelope

Analysis of the packing of **PP_4** showed an extended supramolecular peptide framework (**PP_4-SPF**) with channels extending in one dimension through the network (Void volume 226 $Å^3$, Figure 13), which were occupied with disordered solvent molecules (i.e. Ethanol; Figure 13d-f). The formation of the extended framework is driven by a combination of hydrogen-bonding and Fmoc-Fmoc interactions. The peptides are arranged in alternating anti-parallel rows, extending through the network (Figure 13d-f). Each row of peptides is offset from the next row such that the *C*-terminal amide can hydrogen bond with two other peptides. The *C*-terminal amides hydrogen bond with the carbonyl group of Pro2 on the neighbouring peptide 1 and the Pro3 carbonyl on peptide 2 (2.1478(19) Å and 2.1930(18) Å respectively, Figure 13a), therefore each peptide acts as both a **H_D** and **H_A** with another peptide extending to create 2D sheets, stacked with weaker interlayer interactions in an alternating manner creating an extended network. The crystal structure suggests that interactions between the Fmoc groups determine the second aspect of the self-assembly process. The closest distance (2.7515(19) Å) between the Fmoc moieties is that of a proton of one fluorenyl to the aromatic region of another (Figure 17), well within the possible distance to classify a potential interaction.[34–36]

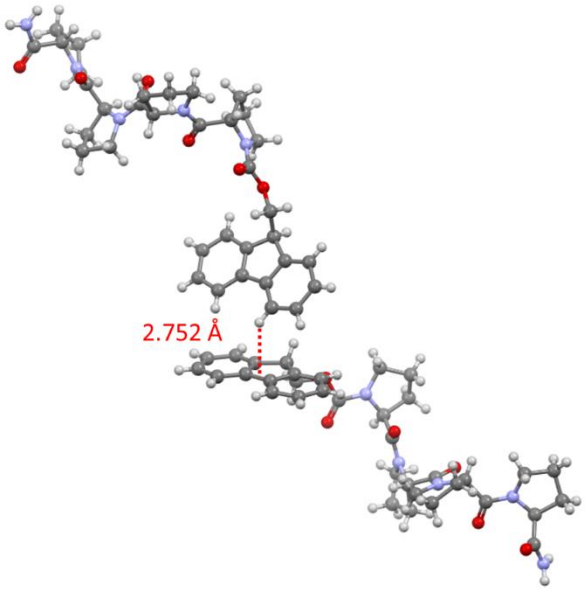


*Figure 17 - **PP$_4$-SPF** crystal structure showing the short contact (dashed red line, 2.7515(19) Å) between a Fmoc proton on one peptide to the centroid of the aromatic region of a neighbouring fluorenyl moiety.*

It is also worth noting that a reversible single crystal to single crystal *trans*ition in the unit cell of **PP$_4$-SPF** is observed (SI 4.3.1, **PP$_4$-SPF$_{Flash}$**), whereby the Pro3 puckering switches from *exo* to *endo* (Table 2). This was achieved by flash freezing crystals at 150 K and, remarkably, upon returning to room temperature the crystals readopted their initial unit cell and conformation. This was not found when slowly ramping the temperature down to 150 K, with only a slight reduction of the cell volume, but no significant conformational changes. Analysis of the *N*-π*interactions in the crystal structure data (Table 1) shows that the *trans*ition of Pro2 from *exo* to *endo*, exhibited in the flash frozen (150 K) sample (**PP$_4$-SPF$_{Flash}$**), resulted in a reduction in pyramidalization by 28% (0.025-0.018 Å), which supports the theory that the *endo* conformation disfavours *N*-π* interactions when combined with the weaker pyramidalization exhibited for all *endo* prolines,[37] compared to those in the *exo* conformation.

2.1.3.2 Computational analysis

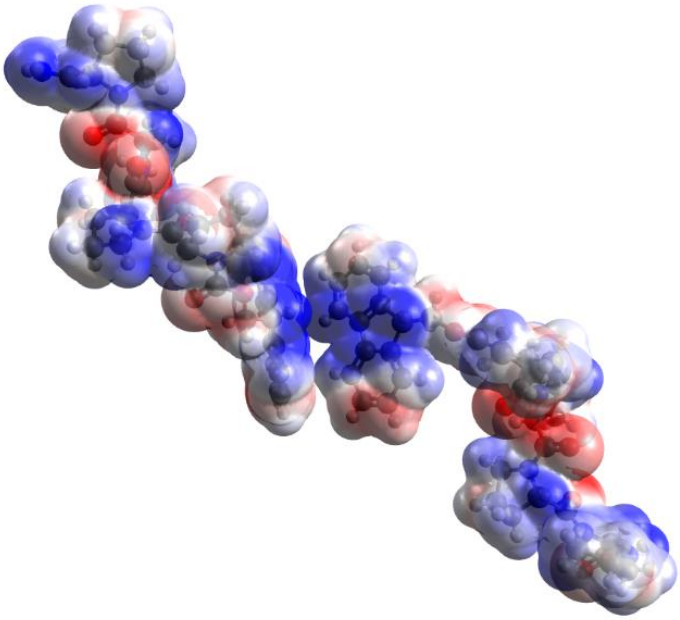


Figure 18 - Electrostatic potential map (blue – positive, red – negative) calculated at B3LYP-D3/6-311-G++(p,d) level of theory.

Computational calculations at [B3LYP-D3(BJ)/6-311++G(d,p)] on an Fmoc associated pair (geometry fixed to that of the **PP_4-SPF** crystal structure) showed orbital overlap between the groups (Figure 19). The electrostatic potential map showed relatively electron poor inner regions for the Fmoc groups, as expected, but no obvious points of significant electrostatic interaction between the groups (Figure 18). Energy decomposition analysis (EDA) of the isolated dimer system (geometry fixed to that of the **PP_4-SPF** crystal structure) was used to determine the relative breakdown of the intermolecular forces between the Fmoc groups in the crystal lattice. Calculations were run with two different functionals, BLYP-D3(BJ) and PBE-D, giving interaction energies of -44.2 and -32.4 $kJmol^{-1}$ respectively. The breakdown suggests that dispersion interactions are the predominant interaction between the Fmoc groups with smaller contributions from electrostatic and orbital interactions (and the Pauli repulsive interaction, SI 4.3.1.17, Table S5).

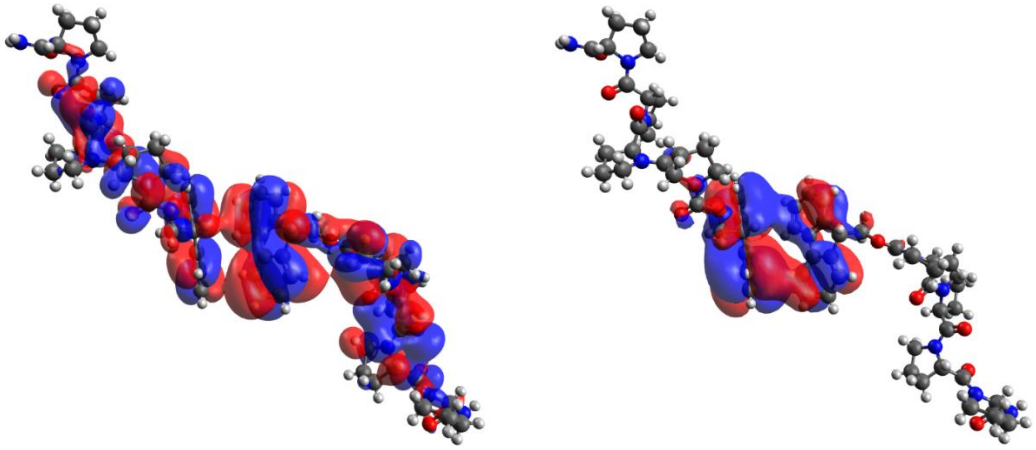


Figure 19 - HOMO-6 (left) and HOMO-19 (right) highlighting orbital interactions between discrete molecules, calculated at B3LYP-D3/6-311-G++(p,d) level of theory.

2.1.3.3 Reversible Porosity and Host-guest Interactions

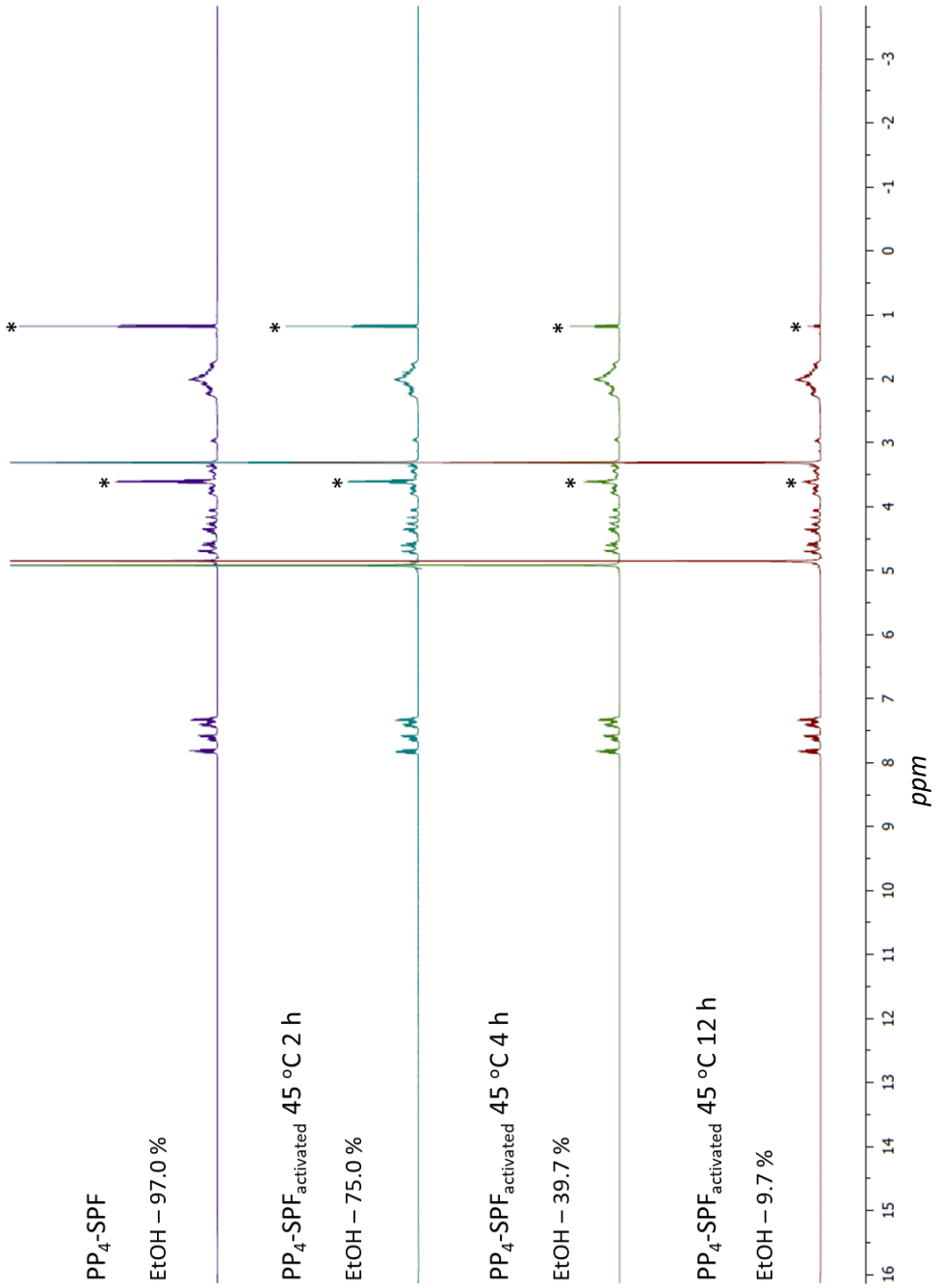


*Figure 20 – ^{1}H NMR spectra of dissolved **PP_4-SPF** with varying degrees of thermal treatment, highlighting the remaining ethanol (mol/mol%) after treatment.*

Due to the porosity of the framework, it has the potential for 'activation' by removal of the solvent from the pores, thereby allowing the introduction of other guests into the activated pores. To investigate the ability for thermal release of ethanol from the framework channels, ^{1}H-NMR analysis of **PP_4-SPF** in methanol-d_4 was carried out after activating the crystals for various timeframes under vacuum at 45 °C. These studies revealed that thermal activation could gradually reduce the ethanol content over time up to a 90% reduction after 12 h, compared to the initial content present within the framework (Figure 20). Subsequently, simultaneous thermogravimetric analysis (TGA) and differential scanning calorimetry (DSC) of **PP_4-SPF** showed a crystalline melting point with an endothermic peak (150 °C, SI 4.3.1.4, Figure S15).

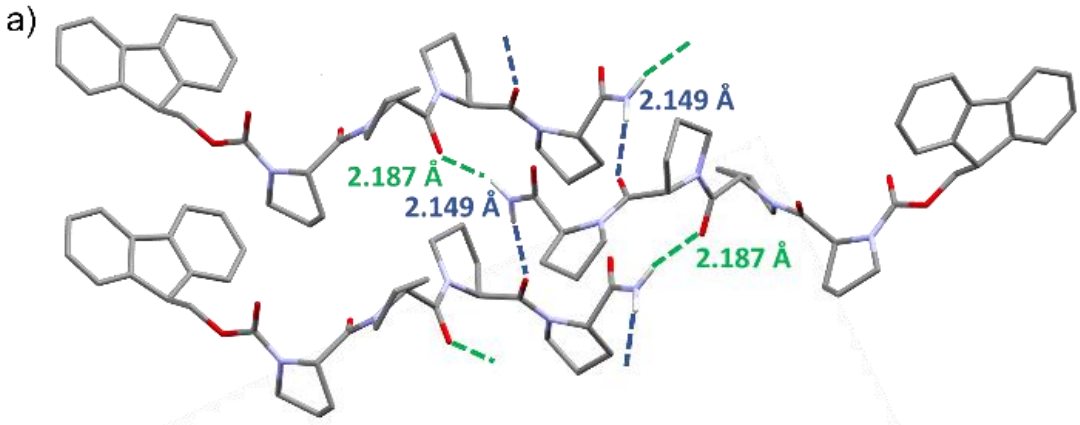


Figure 21 - Hydrogen-bonding between a single peptide unit and two opposing peptide units, through terminal amide hydrogens and Pro2 and Pro3 carbonyls (2.1930(18) Å and 2.1478(19) Å respectively), which extends through the framework (PP4-SPF)

Powder X-ray diffraction (PD-XRD) analyses revealed a change in phase upon de-solvation after heating under high vacuum, with significant differences to the original sample (Figure 22).

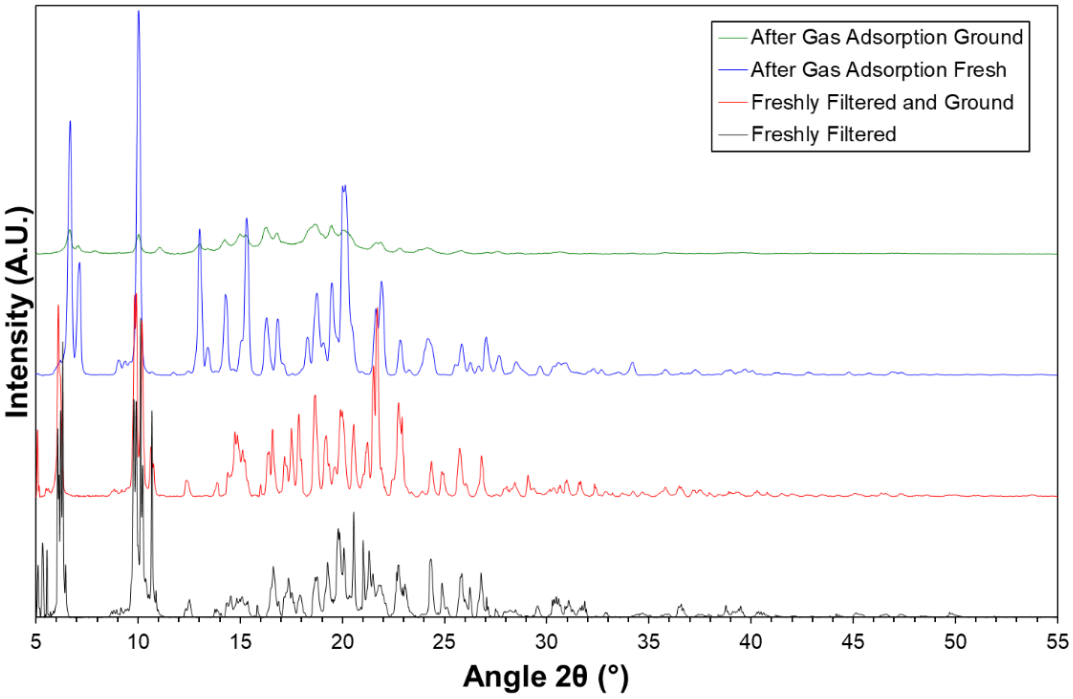


*Figure 22 – X-ray powder diffraction pattern of the freshly isolated, solvated **PP$_4$-SPF** (**black**), this sample after grinding in air (**red**), the sample immediately following de-solvation and gas adsorption experiments (**blue**) and the desolvated sample after grinding (**green**).*

Due to these changes SCXRD analysis was carried out on an activated crystal to investigate the change in porosity. Although the crystal showed very poor single crystallinity after the evacuation of the solvent, and no useful reflections were observed beyond 1 Å resolution, the diffraction data could be solved to give a connectivity model (Figure 23).

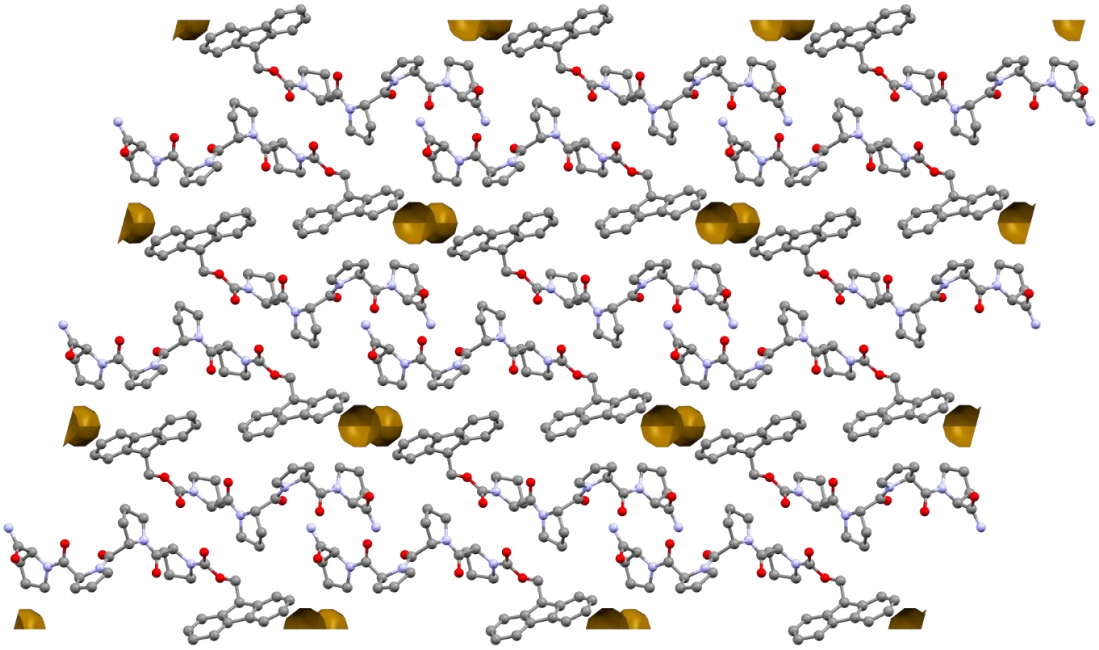


*Figure 23– Crystal structure of **PP$_4$-SPF$_{Act}$** showing the desolvated structure highlighting significantly smaller void spaces ((yellow) 32.81 Å^3, 2.0%-unit cell volume; Probe radius 1.2 Å, grid spacing 0.7 Å^2). Ball and stick model, packing 3x3x3.*

This low-resolution model is presented purely as a comparative connectivity model which suggests collapse of the porous structure with the 2D hydrogen bonded layers shifted relative to one another such that the Fmoc group fills the pore space.

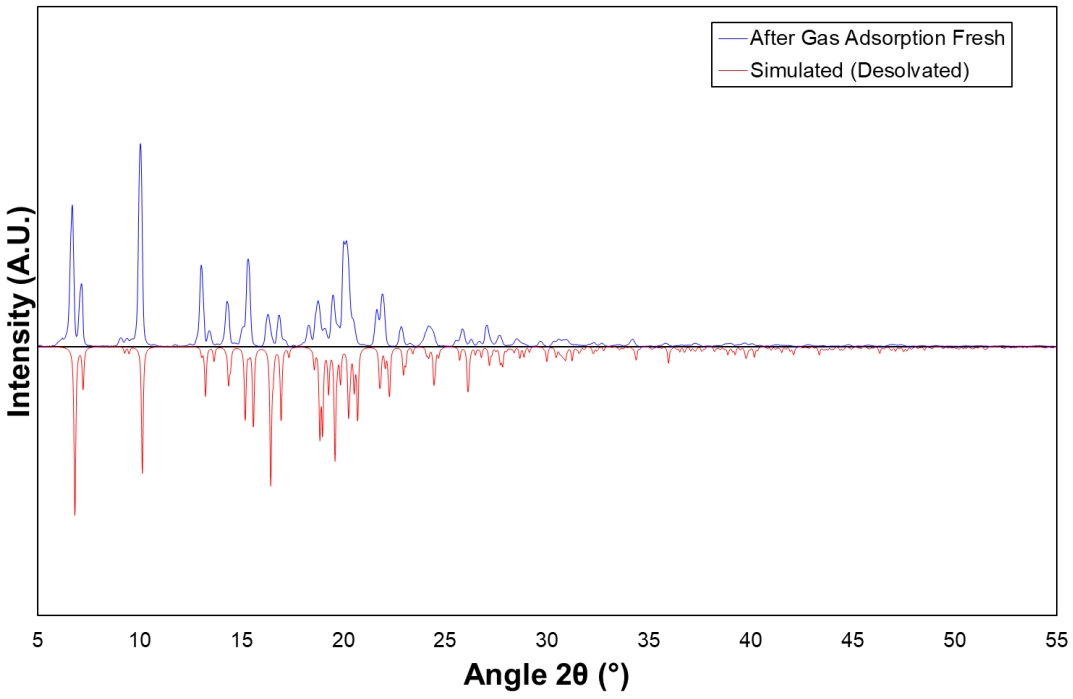


*Figure 24 – X-ray powder diffraction pattern of the desolvated **PP$_4$-SPF** following gas adsorption (**blue**), and the simulated pattern from the single crystal X-ray dataset of a crystal following de-solvation (**red**)*

However, the predicted powder pattern for the modelled de-solvated structure matches very well with the powder pattern of the bulk material (Figure 24), suggesting the correct assignment of the collapsed structure for the material. Gas absorption studies supported these findings, with negligible gas absorption due to pore collapse upon solvent removal (SI 4.3.15).

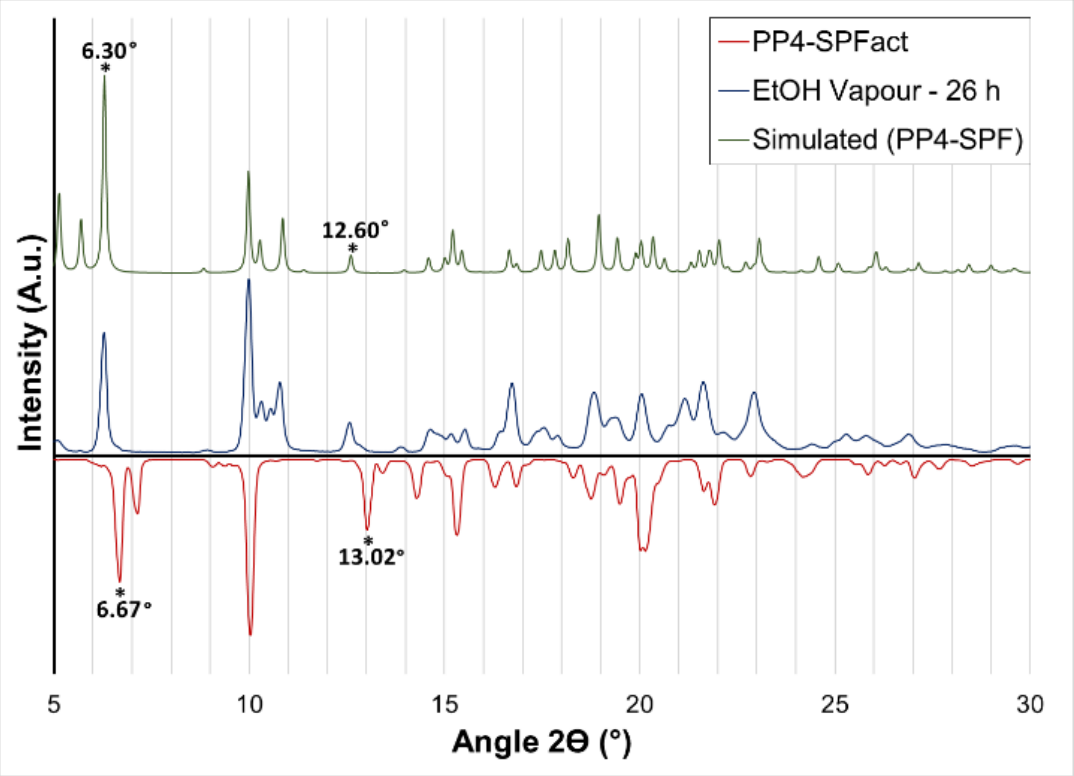


Figure 25 - PD-XRD of PP4-SPF simulated from SC-XRD data (green, top), PP4-SPF after activation at 45 °C under high vacuum (red, bottom), PP4-SPFact after incubation in a chamber saturated with ethanol vapour for 26 h (blue) showing a clear match with the solvated structure, key peaks indicative of phase change are highlighted.

Subsequently, the activated sample was then soaked in the ethanol mother liquor to investigate if this phenomenon was reversible. Remarkably within 15 minutes the crystals showed an almost complete phase change back towards the original structure with a complete return clear after 21 h (SI 4.3.1.6, Figure S17). To rule out the possibility of partial recrystallization causing the phase change, a sample of **PP_4-SPF_{act}** after PD-XRD analysis, while still on the mount, was placed in a chamber with saturated ethanol vapour for 26 h, before again recording the powder diffraction, which showed a complete change from the activated phase to the original phase (Figure 25). This type of behaviour, exhibiting a dynamic porosity whereby the framework is capable of reversible collapse, is more typically seen for MOFs or coordination polymers, whereby the conformational flexibility of a ligand or weak interactions between 2D layers allows for an adaptable *trans*ition, dependent on the guest molecules present.[1,38,39] However, there have been somewhat similar cases with supramolecular/hydrogen-bonded frameworks.[39–42] In this case the behaviour results due to the 2D hydrogen-bonded layers of peptide shifting against one another, behaving as weakly bound 2D interdigitated layers.[38] Thus negligible conformational change of the peptide units occurs and the strong hydrogen bond interactions are retained along two axes (Figure 23). This exemplifies the benefits of the exceptional rigidity of the polyproline helix compared to other biomolecules,[43] highlighting its applicability as a structural unit in constructing supramolecular frameworks.

Table 3: Determination of encapsulated guest[d]					
Entry	**Sample treatment**	**Time /h**	**guest content /mol%[c]**	**Hexane content /mol%[c]**	**Resid. EtOH /mol%[c]**
1	Washed with **hexane** (**PP_4_SPF**)	-	-	8	91
2	**Activated** under vac 45 °C (**PP_4-SPF_{act}**)	16	-	-	12
3	**EtOH 5%** /Hex	1.5	54	61	-
4	**EtOH vapour**	26	14	-	-
5	(±)**1-phenylethanol 5%** /Hex	1.5	36	20	8.
6	(±)**1-phenylethanol 5%** /Hex	16	59	36	6
7	(±)**1-phenylethanol 2 eq.** /Hex	1.5	33	17	5
8	Amorphous **PP_4** Inc. (±)**1-phenylethanol** 5%/Hex	1.5	7	16	78
9	**Acetone 5%** /Hex	16	28[a]	51	0
10	**THF 5%** /Hex	16	23[a]	48	2
11	EtOAc **5%** /Hex	16	33	50	2
12	**Hexane**	16	-	4	7
13	**Toluene 5%** /Hex	16	33	30	5
14	**1-Bromohexane 5%** /Hex	16	15	32	2
15	**I_2** saturated hexane	336	12[b]	18	2

*[a] Integral of clean spectrum of **PP_4** at same ppm range subtracted from solvent peak integral for% calculation due to overlapping peaks of peptide and guest*

[b] Iodine content determined via molecular occupancy from SC-XRD data

[c] All guest mol% calculated against moles of peptide

Due to the framework exhibiting dynamic porosity, attempts were made to see if this behaviour would apply with other guests. **PP_4- SPF_{act}** was soaked in hexane and showed no change in phase on the powder pattern, as such hexane was used as the diluent for other guest molecules. The framework was soaked in a 5% hexane solution of various guests (Table 3*, further details* SI 4.3.1.8). After soaking the SPF was washed thoroughly with fresh hexane to remove excess guest. PDXRD analysis was then carried out to determine whether a change in phase had occurred (*i.e.* reinflation) for each guest (SI 4.3.1.14, Figure S20-30). The mol% of guest retained within the pores, as well as residual EtOH was analysed *via* ^{1}H NMR, after dissolving the framework in methanol-d_4 (Table 3, SI 4.3.15).

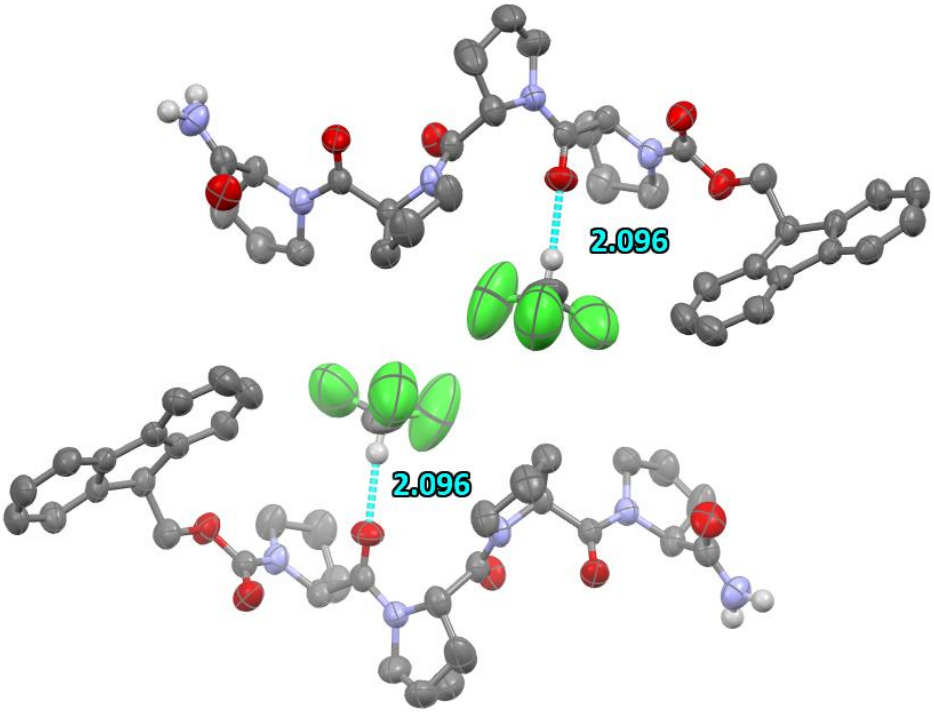


*Figure 26 - Crystal structure of **PP$_4$-SPF@CHCl$_3$** showing the CHCl$_3$ molecules contained with the channels between each row of antiparallel peptides, 50% ellipsoids.*

All the guests analysed (Acetone, THF, 1-bromohexane, ethyl acetate, toluene, (±) 1-phenylethan-1-ol, and iodine), elicited a change in phase from the de-solvated structure, after analysis of the solid SPF after soaking, suggesting they all act to reinflate the structure (SI 4.3.1.14). However, differences in the powder patterns of the samples suggests either a degree adaptation of the framework to accommodate different guests when compared to the original ethanol guest or ordering of the new guest molecules resulting in new diffraction peaks. Although, differences in peak intensities and loss of some low angle peaks are likely attributable to both loss of crystal quality through the processes of activation and guest inclusion. It should be noted that the activated material loses crystallinity when mechanically ground (SI 4.3.1.4, Figure S14), and so the measured diffraction patterns for the (unground) materials also exhibit some preferred orientation effects. Analysis of the ^{1}H NMR data showed significantly higher proportions of hexane than for the SPF soaked alone in hexane (Table 3), which suggests that once porosity is recovered hexane is adsorbed into the framework. However, the proportion of guest alters significantly (e.g. from 5% v/v), favouring adsorption of guest molecules over the hexane diluent. Interestingly, when incubating crystal of PP$_4$-SPF in a 10% chloroform solution in hexane CHCl$_3$ exchanged with the original EtOH molecules to give the new crystal structure ***PP$_4$-SPF@CHCl$_3$*** where CHCl$_3$ could be modelled within the pores, hydrogen bonding to the Pro1 carbonyl, with a 55% occupancy (Figure 26). This exemplifies how this exchange process can be utilised as a method to encapsulate desired molecules acting as a crystalline sponge akin to the coordination networks developed by Fujita,[44] and can enrich a specific molecule due to non-covalent interactions, the size and chirality of the pore space, while also avoiding the harsh vacuum drying process which impacts the crystallinity.

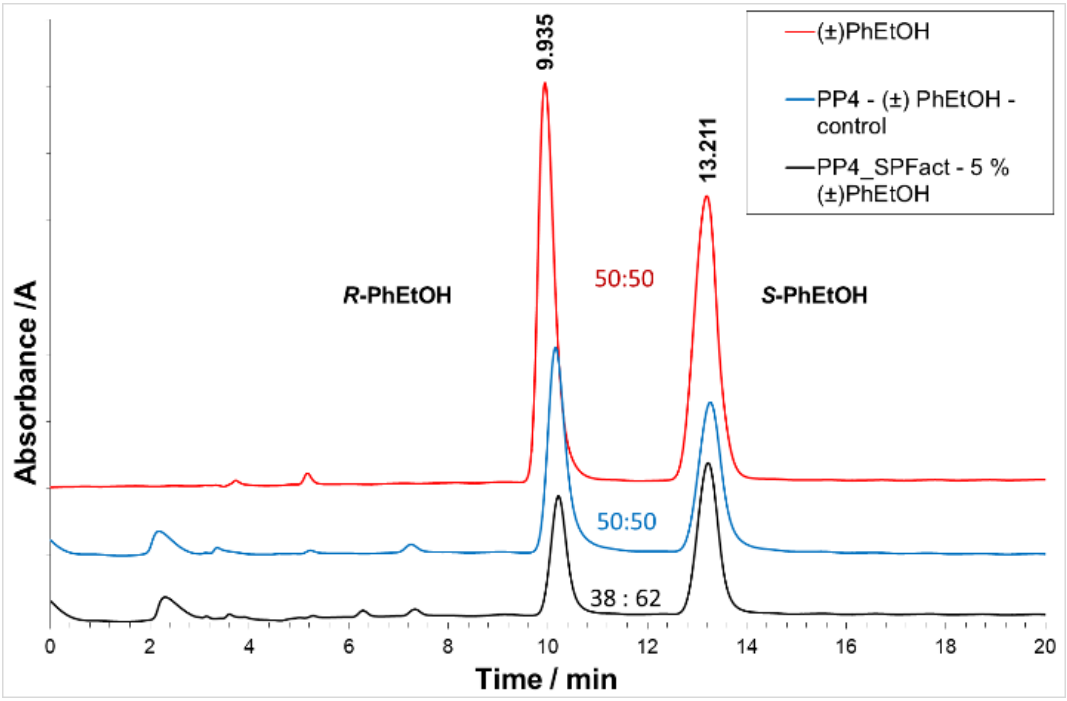


*Figure 27 - Comparative HPLC chromatogram of a solution of (±) 1-phenylethan-1-ol (**red**, top), after incubation of amorphous **PP_4** in (±) 1-phenylethan-1-ol (5% in hexane) (**blue**, middle), and after extraction from **PP_4-SPF_{act}** after 1.5 h incubation in (±) 1-phenylethan-1-ol (5% in hexane) (**black**, bottom). Highlighting the enantioselectivity towards the guest encapsulated within the de-solvated SPF. (Further details SI 4.3.1.15)*

As the SPF building blocks are inherently chiral, with the specific helicity introduced by the polyproline II conformation, we decided to investigate the framework's ability for enantioselective adsorption. With this in mind, and with the successful encapsulation of the chiral molecule (±) 1-Phenylethan-1-ol within the framework (<60 mol%, Table 3, SI 4.3.17), chiral HPLC was utilized to investigate whether any enantioselectivity was exhibited (Figure 27). These results showed a clear shift in the enantiomeric ratio from the reference racemic (±) 1-phenylethanol after exposure to **PP_4-SPF** (≈ 24% *ee*, *S*-1-Phenylethan-1-ol) (Figure 27).

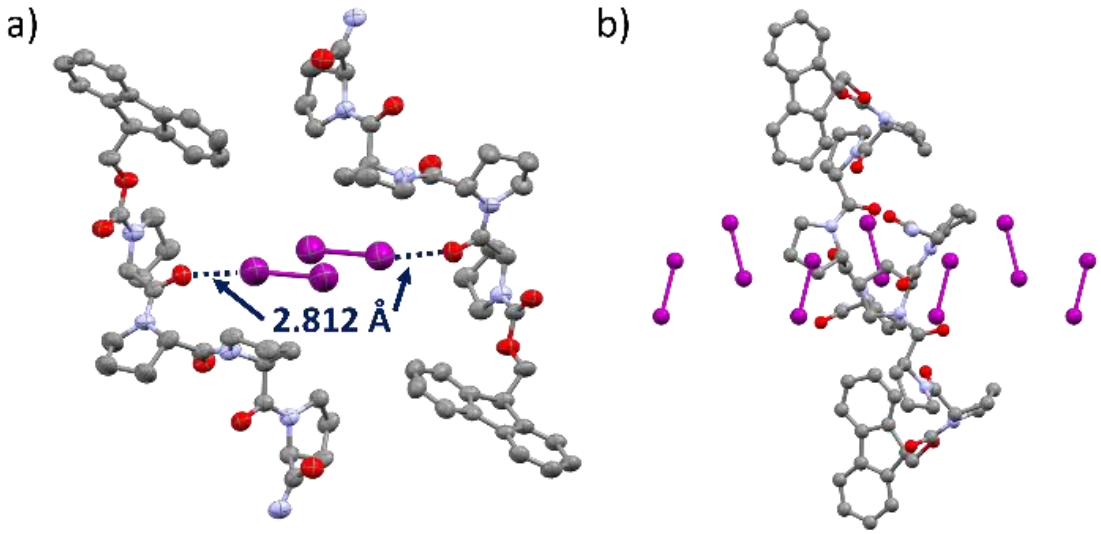


*Figure 28 - Crystal structures of **PP_4-SPF@I_2**, a) View along the b-axis, two parallel peptide units' halogen bond interactions (I---O distance of 2.813(16) Å) with two molecules of iodine (purple), 50% ellipsoids (Mercury), b) View along the a axis showing iodine filling the channels.*

Activated single crystals were then exposed to a solution of iodine in hexane to allow visualization of a guest within the reinflated pores through single crystal diffraction, even at

lower guest loadings, due to the high electron density of iodine. A colour change was evident with the crystals taking on a yellow-orange hue indicating the incorporation of iodine. SC-XRD analysis of **PP_4-SPF@I_2** was performed which showed that the SPF readopted the initial solvated structural geometry and was able to adsorb iodine with a chemical occupancy for I_2 of ≈ 12% (Figure 28, SI 4.3.1.9). The interatomic distances (2.813(16) Å) between iodine and the Pro1 carbonyl oxygens were indicative of halogen bond formation (Figure 28). To investigate the ability to perform reversible guest absorption with iodine, the same single crystal of **PP_4-SPF@I_2** was thermally treated, while mounted on the stage, using a heated flow of N_2 at 323 K for 4 hours, prior to repeating SC-XRD analysis at 150 K. The crystal was mounted without the use of any inert oil, to facilitate guest loss. Upon analysis of the thermally treated crystal, the chemical occupancy of I_2 was reduced by 50% while crystallinity of the sample was retained (SI 4.3.1.9), demonstrating the potential of the **PP_4-SPF** to perform thermally responsive guest release.

2.1.4 Conclusion

In conclusion, the formation of the first supramolecular peptide framework formed by the self-assembly of a polyproline helix is reported (**PP_4-SPF**), utilizing the resilience and rigidity of the polyproline II helix in a short oligoproline. This SPF has shown remarkable reversible porosity and the ability to reversibly host chemical guests in its channels with exhibited enantioselectivity of this adsorption process. The self-assembling principles associated with this SPF will be used as a blueprint in the synthesis of novel helical SPFs and has the exciting potential for the rational design of functionalized cavities, offering an exceptional level of positional control of easily incorporated new functionalities. These materials have clear applications in chemical separations and potential chiral catalysis.

2.1.5 Contributions

Computational calculations were carried out by S.J.Holder. Single crystal structure analyses and structure resolutions were carried out in part by G.Truccolo, D.F.Brightwell, and H.J.Shepherd. Single crystal analysis and structure resolution, gas absorption studies, and powder x-ray diffraction of activated peptide carried out by C.S.Hawes. The original manuscript was written through contributions of all authors.

CCDC-2127748, 2127749, 2127750, 2127751, 2288009 *and 2156434 contain the supplementary crystallographic data for this paper, including structure factors and refinement instructions and can be obtained free of charge from The Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK (e-mail: deposit@ccdc.cam.ac.uk), or via https://www.ccdc.cam.ac.uk/getstructures*

2.2 Supramolecular Self-assembly of Engineered Polyproline Helices

2.2.1 Abstract

The ability to rationally design biomaterials to form desired supramolecular constructs presents an ever-growing research field, with many burgeoning works within recent years providing exciting results, however, there exists a broad expanse of promising avenues of research yet to be investigated. As such we have set out to make use of the polyproline helix as a rigid, tuneable, and chiral ligand for the rational design and synthesis of supramolecular constructs. In this investigation we show how an oligoproline tetramer can be specifically designed and functionalised, allowing predictable tuning of supramolecular interactions, to engineer the formation of supramolecular peptide frameworks with varying properties. Consequently, laying the groundwork for further studies utilising the polyproline helix, with the ability to design desired supramolecular structures containing these peptide building-blocks, having tuneable structural features and functionalities.

2.2.2 Introduction

The synthesis of hierarchical supramolecular functional materials is an exciting field of research with applications in biomedicine,[45,46] separation and catalysis,[3,8,9] and sensing.[47] Pivotal to the successful design of these supramolecular constructs is the ability to synthesise building blocks with specific topologies. Pioneering work by Raymond and Stang in the field of extended and discrete metal organic frameworks were instrumental in demonstrating the importance not only of the nature of the chemical handles but also of their relative position in space.[48,49] While a high level of positional control of these handles can be achieved with relative ease on classical (poly)aromatic building blocks, the same cannot be stated if one intends to use structured peptides as supramolecular building blocks. Recently, biomolecules such as peptides, lipids and DNA/RNA, have appeared in a number of reports as interesting building blocks in the synthesis of novel 2D and 3D biomaterials which assemble using supramolecular interactions.[47,50–54] We are particularly interested in the use of structured peptides as supramolecular building blocks. Peptides can be prepared at scale with high purity, have canonical and non-canonical amino acids incorporated into their primary structure with high accuracy and are biocompatible.[55] This creates a vast array of accessible structures that are easily tuneable, thanks to the stepwise synthesis of peptides on a solid support, creating an expanse of chemical space yet to be explored, with a broad chemical diversity of potential building-blocks and secondary structures. As such, the efforts to investigate this class of compounds as chiral, tuneable ligands has seen a surge in recent years.[8–14] Peptides are typically flexible, chiral, and present a multitude of chemical side chains, which results in complex intermolecular interactions and packing within the solid state. Moreover, when structured peptides such as α-helices and β-sheets are used as supramolecular building blocks, it is known that they can suffer perturbation of their periodicity upon functionalisation (*i.e.,* deterioration of secondary structures) which leads to the inhibition of predictable self-assembly. This often means the accessibility of good quality single-crystals for solid state analysis is challenging.[47] The nature of the amino acids, their position within the sequence, and side-chain-to-side-chain interactions are all aspects that need to be carefully considered in order to minimise the risk of perturbation of the secondary structures in order to achieve a predictable periodicity.[56] With these challenges in hand, it is essential to have a thorough understanding of the peptide secondary structure, and the resulting interacting moieties, to predict the assembly of the peptide ligands within supramolecular constructs. It is within this context that we propose the use of polyproline helices as supramolecular building blocks. The polyproline II helix is both rigid and stable in short sequences, and has three repeating helical faces, creating predictable and accessible handles for functionalisation and supramolecular assembly.[30,57,58] The lack of internal hydrogen bonding further simplifies the

potential assembly principles compared to other helices, these factors combined make the polyproline helix an ideal candidate.

We have recently demonstrated that super short polyproline helices (tetrameric peptides) can assemble into a reversibly porous supramolecular peptide framework (SPF) capable of engaging in stereoselective host encapsulation.[59] Herein, we demonstrate the ability to utilise functionalised short polyproline helices as predictable ligands with an exceptional level of control, for the rational design of a series of H-bonding driven supramolecular peptide frameworks. The design principles successfully applied to these peptides, can be used to drive the design of more complex materials; the periodicity of the helix allows the expansion of the principles of assembly found in these minimalistic peptides to longer peptide chains, while the resilience of the helix to functionalisation means these principles can be applied to various functional groups.

Minimalistic peptides have the potential to play a key role in the emerging field of bionanomaterials.[42,60,61] We recently reported the first SPF formed by the self-assembly of a polyproline helical tetramer, Fmoc-$(Pro)_4$-NH_2, **P_4**. Despite the short length of this tetrapeptide, **P_4** crystallised in the polyproline II helical form, a common secondary structure found in nature.[62] This SPF showed remarkable reversible porosity, the ability to reversibly host chemical guests in its channels and exhibited enantioselectivity for this adsorption process. The formation of this porous framework was driven by hydrogen-bond donor and acceptor (H_D-H_A) interactions between the hydrogens of the primary amide at the C-terminus and the carbonyl groups on the peptide backbone, as well as, Fmoc-Fmoc association predominantly driven by dispersion interactions.[59] These results gave us an insight into the potential of short polyproline helices as supramolecular building blocks.

We aim to exploit the full potential of polyprolines as minimalistic peptides in the construction of emerging bionanomaterials. We are particularly interested in the functionalisation of position 4 (or *γ*-carbon) of the proline amino acid as this position is exposed on the exterior of the helix[30] and is capable, upon functionalisation,[63] of engaging in the formation of supramolecular interactions. Remarkably, analysing the **P_4** supramolecular framework, we were able to successfully predict the effect of a hydrogen donor (i.e. **H_D**; -OH) interactions on the supramolecular assembly for a series of polyproline peptides. Led by design principles based off the P_4 framework, a series of seven hydroxy-functionalised derivatives were synthesised using Fmoc-based solid-phase peptide (SPPS) techniques (Figure 29*, SI 4.2*). We anticipated that if the polyproline II conformation was retained for these peptides, the spatial orientation of the hydroxyl group of a hydroxyproline residue would be highly predictable, thereby enabling future endeavours in the rational design of polyproline based ligands to assemble into supramolecular bio-constructs. If successful, we would demonstrate the resilience and high level of positional control achievable using proline based minimalistic peptides. Therefore, with this information it would be feasible to rationally design polyproline-based peptides with predictable geometries of functional motifs for the incorporation of further supramolecular interactions, to be utilised in supramolecular assembly. The repeating nature of the polyproline helix also means that, from the detailed investigation of functionalisation of a minimalistic tetrameric oligoproline (four residues constituting one full turn of the polyproline II helix), it is possible to infer the structural details of longer peptide units with multiple helical turns and apply these findings as a strong foundation to design more complex polyproline-based peptides.

2.2.3 Results and Discussion

2.2.3.1 Synthesis and Analysis of Engineered Oligoproline Frameworks

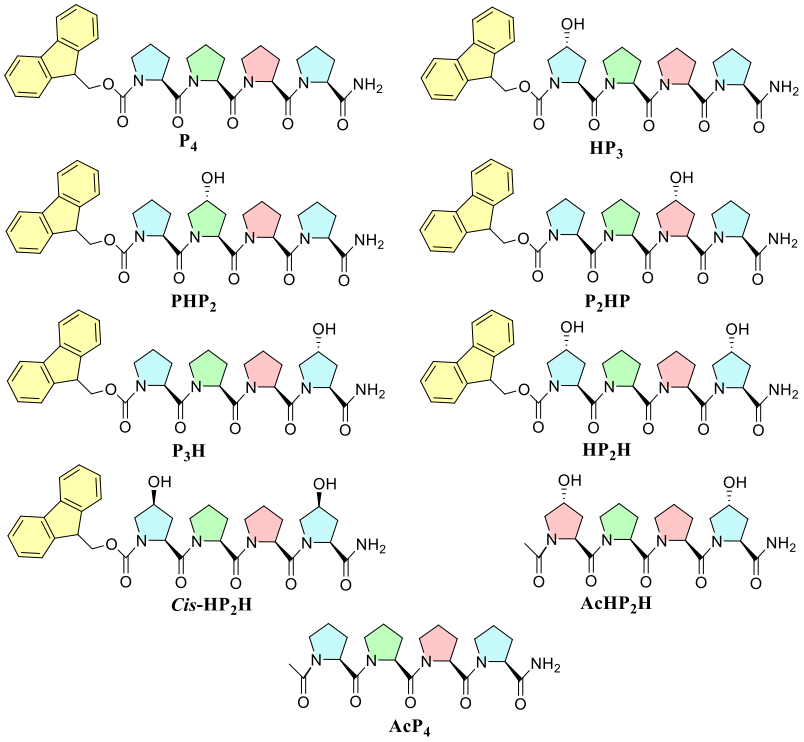


Figure 29 - Chemical structures of synthesized oligoproline peptides, helical faces are highlighted in different colours (i+3 periodicity). H in the peptide name indicates the position of the hydroxyproline starting from the N-terminus in the sequence.

Analysing the crystal structure of **P4** we were able to observe that the terminal prolines, Pro1 (N-terminus) and Pro4 (C-terminus), clearly have close contacts with the neighbouring peptide's carbonyl groups along the *b*-axis (Figure 30a).[59] We anticipated that a H-donor group such as the hydroxyl group introduced in position 4 of the first proline in the sequence, would engage in intermolecular hydrogen-bonding interactions along the *b*-axis without significantly impacting the packing topology.

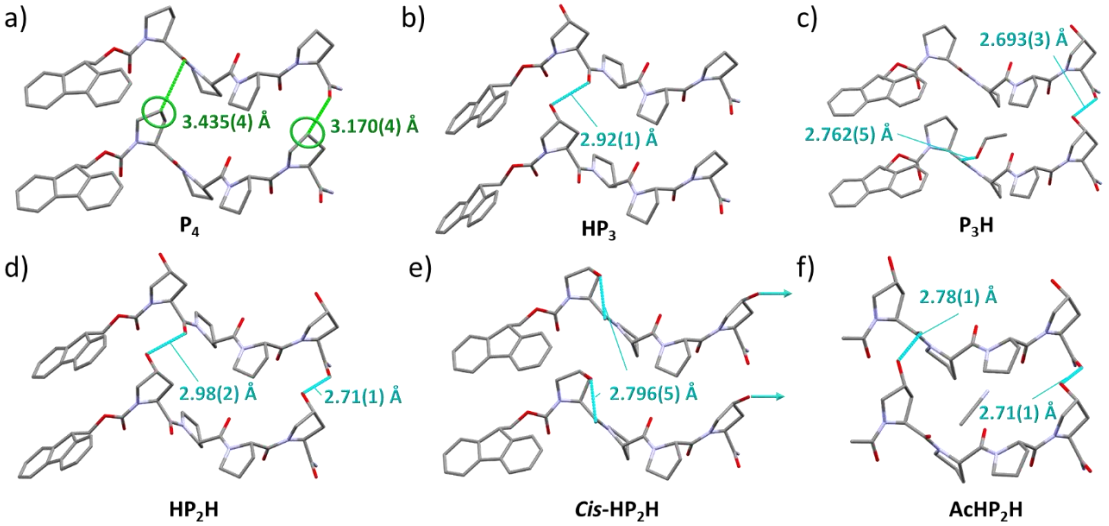


Figure 30 – Crystal structures of peptides P_4 (a), highlighting close contacts between Cγ (Pro1 and Pro4) and the adjacent peptide's closest carbonyl groups, HP_3 (b), P_3H (c), HP_2H (d), cis-HP_2H (e), and AcHP_2H (f), showing the new hydrogen bond interactions formed by the additional hydroxyl moieties. cis-HP_2H C-terminus hydroxyl has an undefined hydrogen bond to disordered solvent within the channels of the framework.

Thus, it can be inferred that the crystal structures formed should adopt, through functionalisation of the terminal residues, the same hydrogen-bonds and Fmoc-Fmoc interactions as in the **P_4** structure, with additional H-bonding (O-H---O=C) along the same plane as the amide (N-H---O=C) H-bonding interactions, to form similar 2D sheets, this packing of the peptide units is highlighted in a simplified model in Figure 32. These sheets should stack in various assemblies to form either porous, as in the **P_4** framework, or non-porous structures. With these details in mind a series of peptides were synthesised by replacing the terminal prolines with hydroxyprolines, these were peptides; **HP_3**, **P_3H**, **HP_2H**, ***cis*-HP_2H** (containing all *cis*-hydroxyprolines), **$AcHP_2H$** and **AcP_4** (Figure 29), varying not only the number of hydroxyl groups but also the *N*-terminal capping moiety. Due to the increased solubility of the functionalised peptides in polar solvents compared to peptide **P_4**, they were largely crystallised from a mixture of ethyl acetate or acetonitrile and ethanol, while peptide **$AcHP_2H$,** with the Fmoc group replaced with an acetyl group, crystallised from an acetonitrile solution.

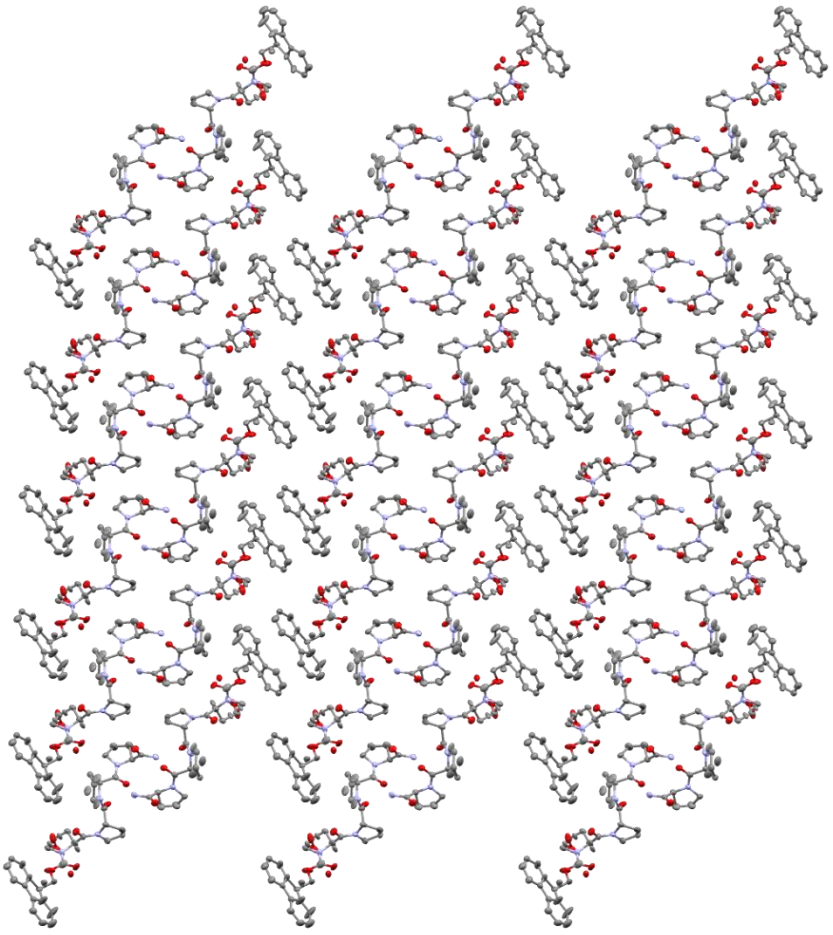


*Figure 31 - – Crystal structure of peptide **HP_3** showing the packed extended structure (3x3x3), view along the b axis, hydrogens not shown, atomic displacement parameters are shown at 50% probability.*

The first peptide tested **HP$_3$**, Fmoc-Hyp-Pro$_3$-NH$_2$ (Figure 29)**,** crystallised to form a non-porous structure (Figure 31), driven by the addition of a hydrogen-bond between the Pro1 hydroxyl to the neighbouring peptide's Pro1 carbonyl, forming 1D hydrogen bonded tapes of the peptide along the *b*-axis, similar to the tapes of peptide along the *b*-axis already present in **P$_4$**, (Figure 30*a*). The amidated *C*-terminal formed the same hydrogen-bond interactions (-NH$_2$---O=C) as found for the **P$_4$** framework, with interactions with the Pro2 and Pro3 carbonyls of adjacent antiparallel peptides. The assembly of the peptides in these hydrogen-bonded layers into 2D layers, *via* Fmoc-Fmoc interactions, is very similar to the crystal structure of **P$_4$**. However, these 2D sheets then stack with a slight displacement to form the 3D framework, significantly changing the unit cell and reducing the void volume, resulting in no solvent-accessible channels present within the framework (Figure 30a).

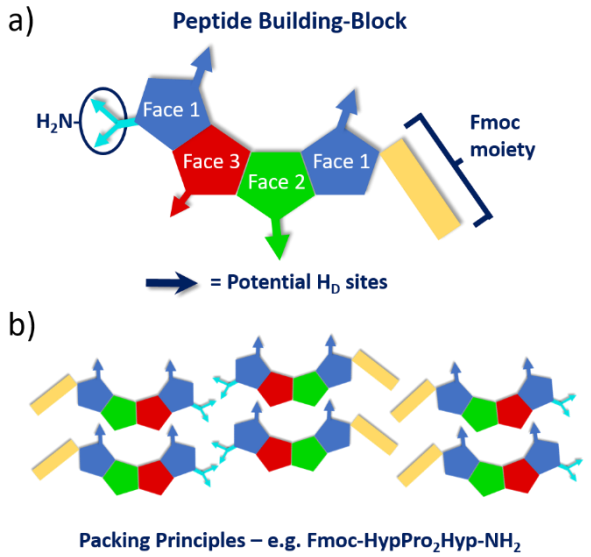


Figure 32 – a) A model of a general peptide building-block with sites of intermolecular interaction shown, b) general packing principles of hydrogen-bonded layers, example: peptide HP$_2$H shown.

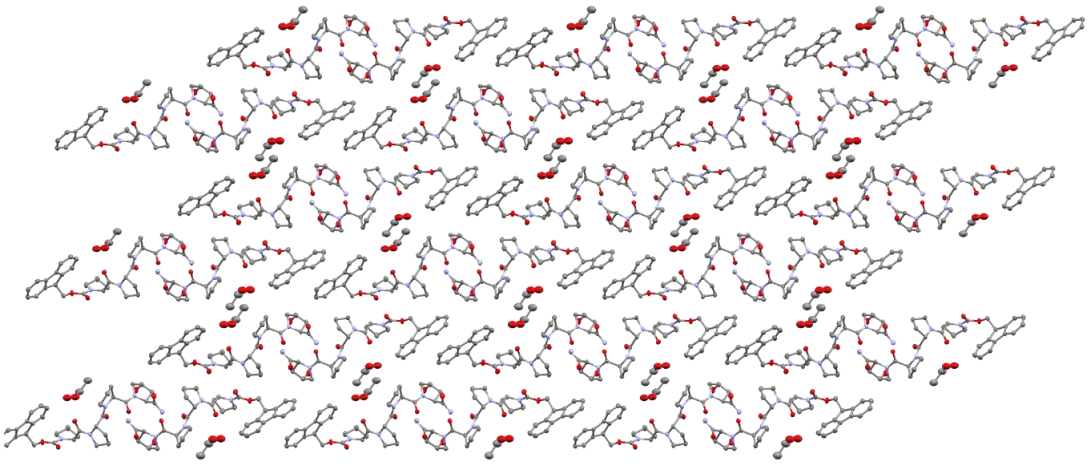


Figure 33 – Crystal structure of peptide P$_3$H showing the packed extended structure (3x3x3), view along the b axis, hydrogens not shown, atomic displacement parameters are shown at 50% probability.

The peptide **P$_3$H**, Fmoc-Pro$_3$-Hyp-NH$_2$, crystallised as a porous structure with channels (Volume 342.6 Å^3, 10% / unit cell, Probe *r* = 1.2 Å, Grid spacing 0.4 Å), isostructural to the **P$_4$** framework (Table 4) except for a doubling of the *c*-axis, to be expected from the doubling of Z' (Figure 33). These pores are a slightly larger size than those found in the original framework (Volume 245.09 Å^3, 13.8% / unit cell, Probe *r* = 1.2 Å, Grid spacing 0.4 Å; Figure 30*d*), which presents a good example of the accessibility of alternative structures *via* editing of the peptide monomers, allowing tuning of the pore environment for various functionalities. In this case the asymmetric unit was comprised of two *N*-terminally hydrogen-bonded peptides. The new hydroxyl group has a hydrogen bond interaction with the Pro4 carbonyl of the adjacent parallel peptide, occurring similarly for both peptides within the asymmetric unit (OH---O=C, 2.693(3) and 2.762(5) Å, Figure 30*c*). The precise placement of the -OH groups along the same helical face,

i.e., Pro1 and Pro4, produces H-bonding interactions towards neighbouring carbonyls of Pro1 and Pro4 units respectively, thus highlighting the exceptional control possible, producing interactions with specific predictable geometries.

Table 4: Comparative table of peptide crystal structures			
Structures	**Porous (Y/N)**	**Crystal System**	**Space Group**
P_4^*	*Y*	*Monoclinic*	$P2_1$
P_2HP^*	*Y*		
P_2HP-P_4^**(mixed)*	*Y*		
P_3H^*	*Y*		
$HP_3^‡$	*N*	*Monoclinic*	*C2*
$HP_2H^‡$	*N*		
Cis-HP_2H	*Y*	*Orthorhombic*	$P2_12_12_1$
AcHP_2H	*Y*	*Orthorhombic*	$P2_12_12_1$
AcP_4	*Y*	*Monoclinic*	$P2_1$

**Isostructural to each other, ‡Isostructural to each other,*

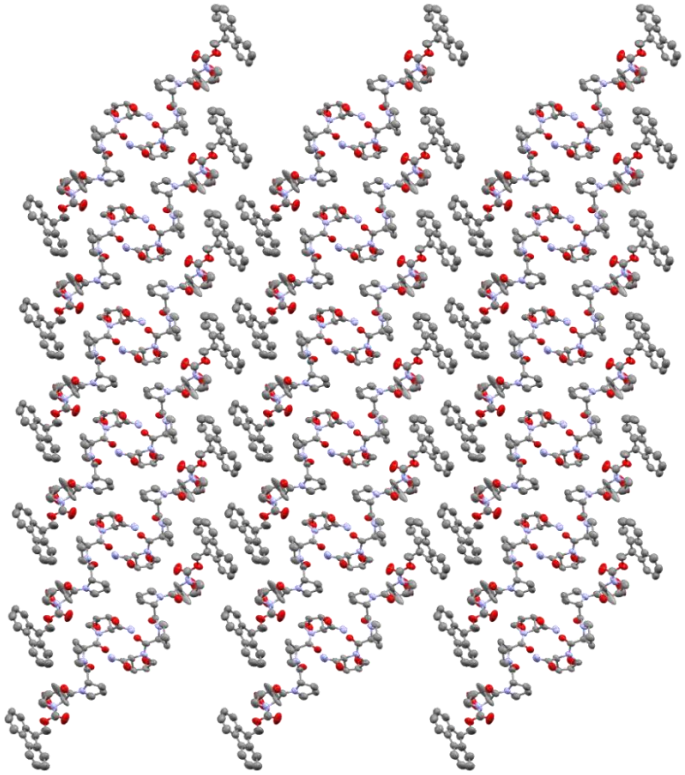


*Figure 34 – Crystal structure of peptide **HP₂H** showing the packed extended structure (3x3x3), view along the b axis, hydrogens not shown, atomic displacement parameters are shown at 50% probability.*

As both **HP₃** and **P₃H** peptides form similar hydrogen bonded layers within their frameworks (Figure 30*b*-*c*), we predicted functionalisation of both positions would allow for both hydrogen

bonds to be present simultaneously within the framework without significant disruption of the packing, as such peptide **HP$_2$H**, Fmoc-Hyp-Pro$_2$-Hyp-NH$_2$, was synthesised. Single crystals of **HP$_2$H** were successfully obtained in the same manner as **HP$_3$**/**P$_3$H**, and subsequently analysed *via* SC-XRD. The crystal structure contained both hydrogen bonds present in the previous structures as predicted, with similar hydrogen bond distances (Pro1 OH---O=C; 2.98(2) Å vs **HP$_3$**; 2.92(1) Å. Pro4; 2.71(1) Å vs **P$_3$H**; 2.693(3) Å, Figure 30*d*). However, the extended structure was isostructural with peptide **HP$_3$**, (Table 4) with only small differences in unit cell parameters (Figure 34). The successful formation of this framework and retention of the polyproline II helix, despite 50% functionalisation of the peptide clearly highlights the resilience of the polyproline helix and how the ability to predict the geometry of new functionalities can be utilised to rationally design supramolecular constructs.

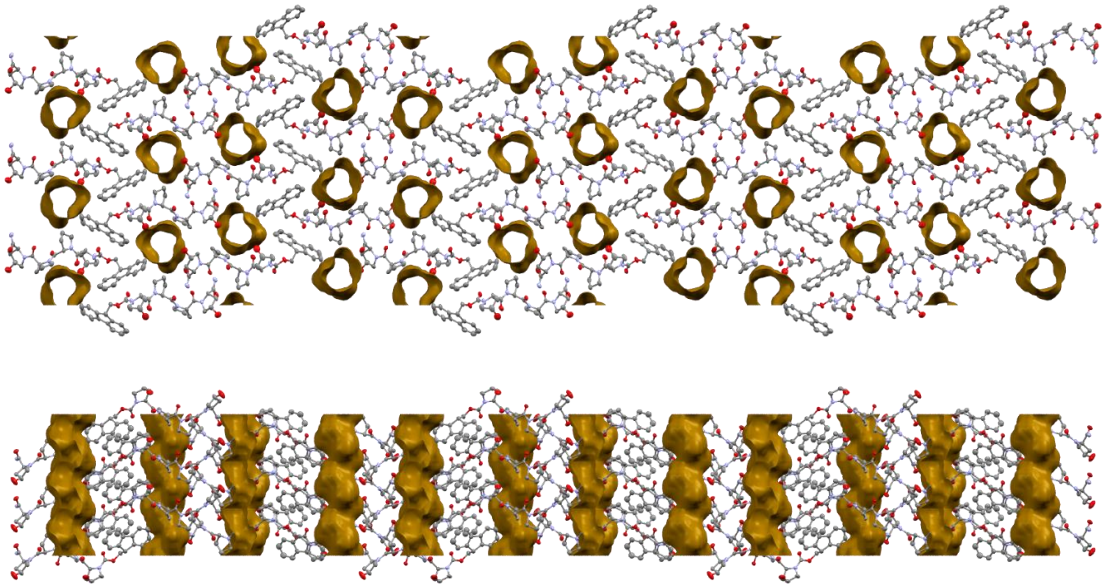


*Figure 35 – Crystal structure of peptide cis-**HP$_2$H** showing the packed extended structure (3x3x3), solvent filled channels highlighted in yellow (volume 628.9 Å^3, 17.1% / unit cell, Probe r = 1.2 Å, Grid spacing 0.4 Å). Hydrogens not shown, atomic displacement parameters are shown at 50% probability . Top: view along a-axis, bottom: view along b-axis*

To observe the impact of *cis*-hydroxy, versus the *trans*-hydroxy previously used, on the packing topology, the peptide *cis*-**HP$_2$H**, Fmoc-*cis*-Hyp-Pro$_2$-*cis*-Hyp-NH$_2$, was synthesised, with both hydroxyprolines *cis* rather than *trans*. In this case we expected the *4S*-hydroxyproline to prefer the *endo* conformation and internally hydrogen bond to the hydroxyproline's amide carbonyl, thus restricting the formation of intermolecular interactions.[64] This peptide was successfully crystallised from a mixture of acetonitrile and ethanol, forming colourless plank crystals (*SI 4.3.8*). The crystal structure obtained contains channels filled with disordered solvent (volume 628.9 Å^3, 17.1% / unit cell, Probe *r* = 1.2 Å, Grid spacing 0.4 Å, Figure 30*e, SI 4.3.8*), and was not isostructural to any of the other peptide frameworks (Table 4). Most significantly the extended structure of *cis*-**HP$_2$H** differs from the other Fmoc peptide frameworks as the Fmoc moieties of adjacent peptides no longer face one another, resulting in staggered rather than linear H-bonded layers of the peptides (Figure 35). Due to the presence of both acetonitrile and ethanol solvent molecules, and significant disorder, the solvent within the pores could not be accurately modelled as such solvent masking was used. However, the last hydroxyproline's (C-terminus) hydroxyl group is aligned into the pore space, clearly hydrogen bonding to a solvent molecule, with significant electron density adjacent to this group. While this hydroxyl was in the typical *exo* ring puckering, thus aligned mostly perpendicular to the helix, the *N*-terminal hydroxyl adopted *endo* ring puckering. The *endo* hydroxyl, now facing into the helix, was then able to

satisfy the $\mathbf{H_D}$ *via* intramolecular hydrogen bonding (OH---O=C, 2.793(4) Å, Figure 30*e*) towards the same hydroxyproline's amide carbonyl (Figure 30*e*). Therefore, the only intermolecular hydrogen-bonding between peptides is the typical *C*-terminal NH_2 amide bonding present in all the structures seen previously. This highlights how small changes in the placement of functional groups can be used to affect the assembly process, with control over even the flexible pyrrolidine ring *endo*/*exo* conformations possible by use of *4S* versus *4R* functional groups, while the polyproline II helix remains as a rigid ligand for placement of these functional groups.

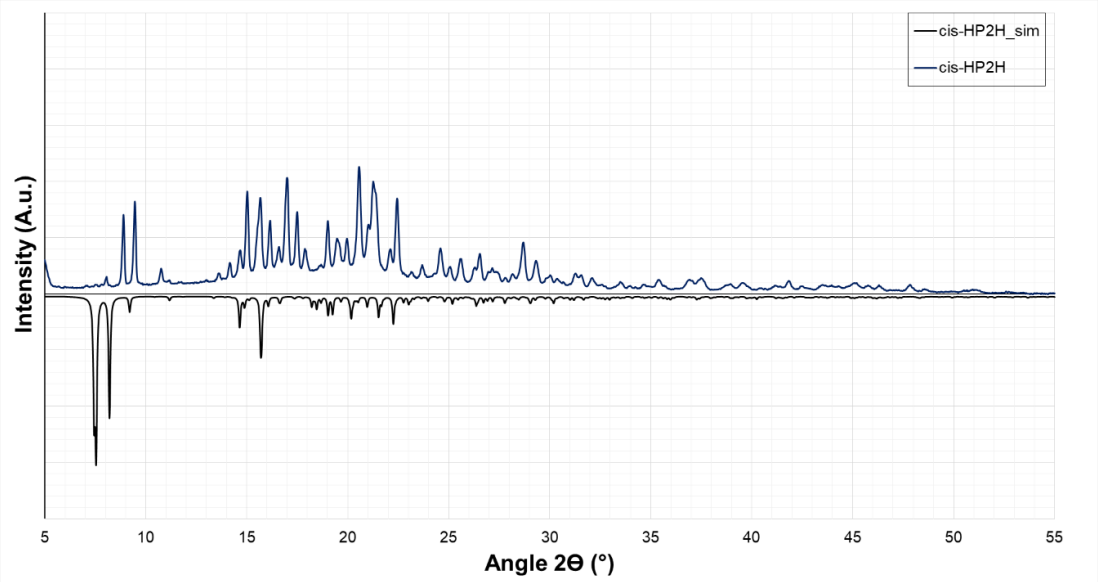


*Figure 36 – PD-XRD of **cis-HP$_2$H** simulated from SC-XRD data (**black, bottom**), **cis-HP$_2$H** experimental PD-XRD (**blue, top**)*

The significant impact of the ability to internally hydrogen bond was also observed when analysing the bulk material from crystallisation *via* powder X-ray diffraction, which differed significantly from the simulated pattern from the crystal structure (Figure 36). This suggests that the bulk material adopts a different crystal structure to that of the analysed crystals, which was not found for any other peptide crystals, and is likely a product of multiple accessible structures through the two hydroxyls adopting either *endo* (internal hydrogen bonding) or *exo* (intermolecular hydrogen bonding) ring puckering (Figure 30e) and thus changing the nature of the supramolecular interactions present in the crystal structure.

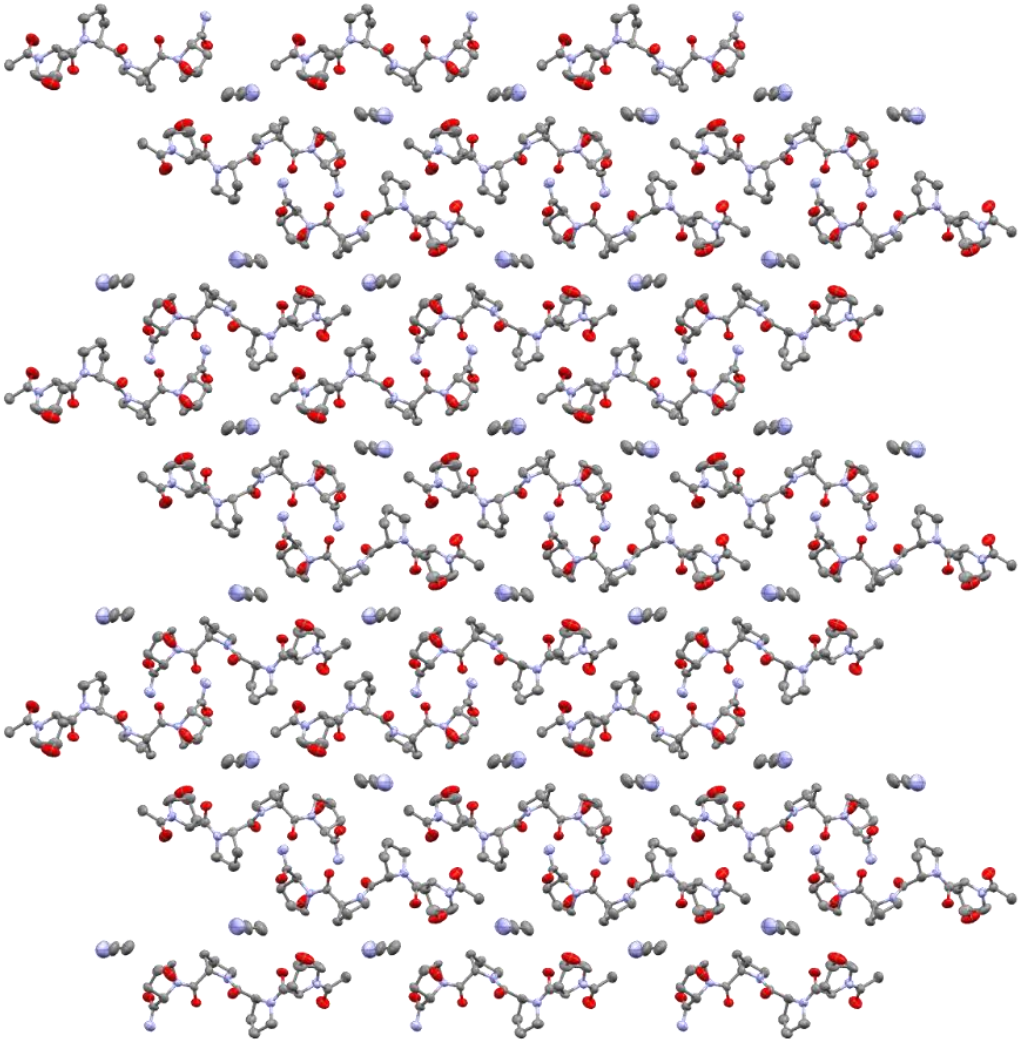


*Figure 37 – Crystal structure of peptide **AcHP$_2$H** showing the packed extended structure (3x3x3), view along the a axis, hydrogens not shown, atomic displacement parameters are shown at 50% probability.*

With the addition of two hydrogen bonding interactions in the **HP$_2$H** structure it seemed apparent that these hydrogen bonding interactions should still produce an extended 2D structure without the Fmoc-Fmoc interactions. As such the peptide **AcHP$_2$H**, Ac-Hyp-Pro$_2$-Hyp-NH$_2$, was synthesised, whereby the Fmoc protecting group on the *N*-terminus was replaced with an acetyl group (Figure 29). Initial attempts to crystallise the peptide indicated the clear formation of an extended network as the peptide formed an organogel upon sonication in a saturated solution (Figure 38). As such, a more dilute solution of the peptide in chloroform was analysed *via* atomic force microscopy (AFM) to see if initial fibre formation could be observed as the solution evaporates. These results showed the aggregation of a gel-like assembly which clearly contained trapped solvent, from spots showing areas with escaped solvent upon drying (Figure 39b), into more extended fibres (Figure 39a+c).

*Figure 38 – Sample of peptide **AcHP$_2$H** in chloroform (6 mg ml^{-1}) after sonication*

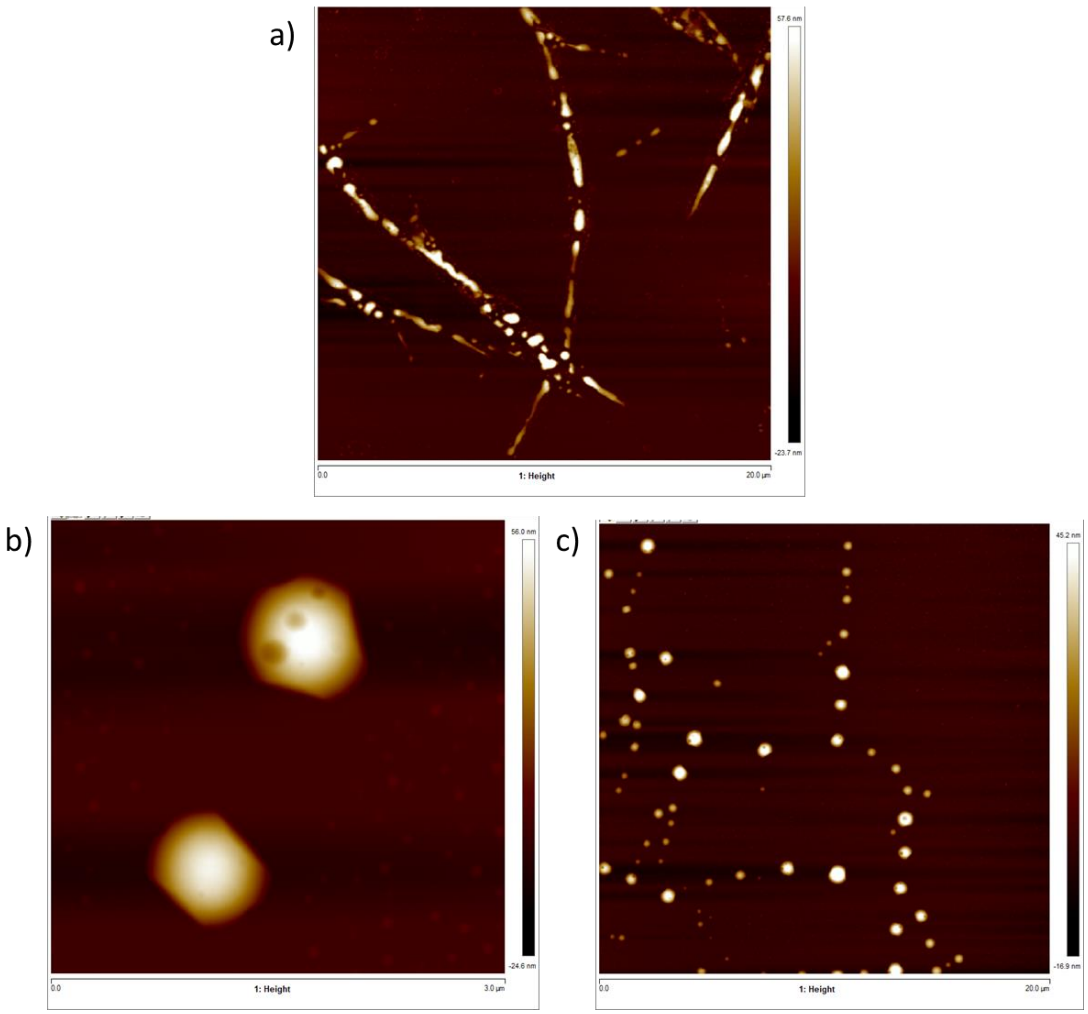


*Figure 39 - AFM image of peptide **AcHP$_2$H** in $CHCl_3$ (1.2 mg/mL), 10 µL spotted; a) max height 63 nm; b) object average height 29 nm, radius 295 nm; c) object height 22-48 nm, radius 160-370 nm.*

Heating and slowly cooling a supersaturated solution of the peptide in acetonitrile, or slow evaporation of an acetonitrile solution, produced clusters of long fibre/needle-like crystals of the peptide, unlike those produced in any of the Fmoc containing structures. These largely 1-dimensional crystals were expectedly poorly diffracting; however, a crystal structure was obtained from a suitable crystal (Figure 37), with the data giving a predicted powder pattern that matched well with PDXRD analysis of the bulk material (*SI 4.4.9*). Remarkably, the polyproline helix form II was retained for this peptide moreover, as anticipated, analysis of the single crystal data showed the retention of the same hydrogen bonding interactions present in the **HP$_2$H** structure (Figure 30*d+f*) but, with the loss of the Fmoc interactions, there are no other significant interactions extending along two of the axes (Figure 37). Channels within the framework contain well-ordered acetonitrile molecules with no apparent strong interactions between the solvent and the peptide within the structure (Volume 415.06 Å^3, 15.6% / unit cell, Probe *r* = 1.2 Å, Grid spacing 0.4 Å). Remarkably, comparing the structures of **AcHP$_2$H** and **HP$_2$H** we can clearly conclude that, while not necessary to the formation the peptide framework, nor to the stability of the polyproline helix in such short peptides, the terminal groups can also be functionalised with chemical handles (*e.g.,* Fmoc/Acetyl) to increase the level of control of the supramolecular assembly process for these short peptides.

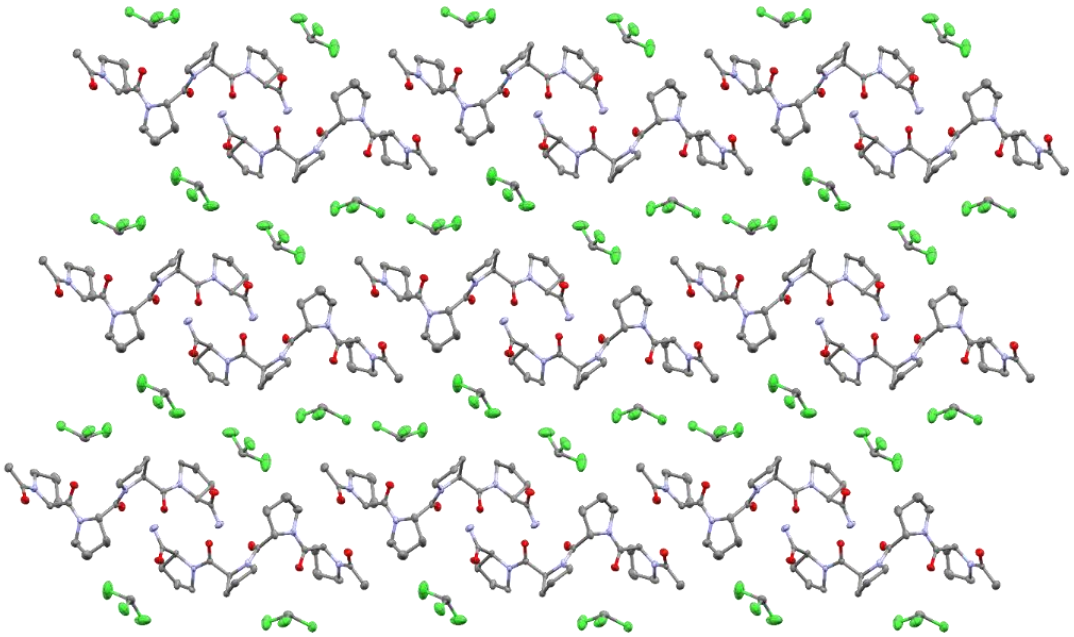


*Figure 40 – Crystal structure of peptide **AcP$_4$** showing the packed extended structure (3x3x3), view along the b axis, hydrogens not shown, atomic displacement parameters are shown at 50% probability.*

Due to the successful crystallisation of the peptide **AcHP$_2$H**, without the contribution of the Fmoc moiety, the peptide **AcP$_4$** (Figure 29), Ac-Pro$_4$-NH$_2$, was synthesised to demonstrate the contribution of the two terminal hydroxyprolines, versus the *C*-terminal amide -NH$_2$, to the self-assembly of the peptide unit. It was theorised that the structure should bear strong similarities to the **AcHP$_2$H** structure if successfully crystallised, driven by the bridging terminal -**NH$_2$** hydrogen bonding seen in all the structures, without the influence of any other moieties. However, it was uncertain whether this H-bonding would be sufficient to drive the assembly. Subsequently, the peptide did not crystallise from ACN or any of the previously attempted solvents, but readily crystallised from a $CHCl_3$ solution. Through vapour diffusion of Et_2O into a solution of the peptide in $CHCl_3$ crystalline needles were obtained suitable for SCXRD analysis. This data (SI 4.3.10) showed the formation of a similar assembly to **AcHP$_2$H** with repeated layers of peptide and solvent (Figure 40), however, adopted a significantly different unit cell.

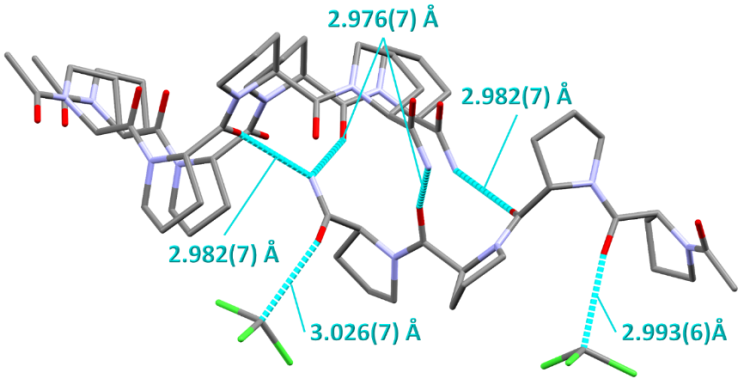


*Figure 41 Crystal structure of peptide **AcP$_4$** showing hydrogen bonding interactions of a single peptide unit, hydrogens not shown, modelled as capped sticks.*

In this case the layers of peptide are significantly further apart forming large channels (Volume 572.37 Å^3, 36.1% / unit cell, Probe *r* = 1.2 Å, Grid spacing 0.4 Å) containing two $CHCl_3$ molecules per peptide, coordinating to the Pro1 and Pro4 carbonyl groups (Cl3CH---O=C, 2.993(6) and 3.026(7) Å, SI 4.3.10 Figure 41). This clearly demonstrates the ability of the peptide to crystallise

without the presence of sterically bulky, rigid groups (e.g. Fmoc), with a minimally functionalised tetramer forming a porous framework. With this data and the previous structures we exemplify how variation of sidechain functionalities, terminal capping groups and non-coordinating/coordinating solvents can be used to achieve a variety of supramolecular frameworks, starting from an initial peptide model. Significantly, this example also highlights the stability of the polyproline II helix, with previously "inconclusive" evidence of the helix being adopted for the proline tetramer,[31,32,65] this simple peptide, and all other crystallised peptides within this work, adopted the polyproline II helix in the solid state.

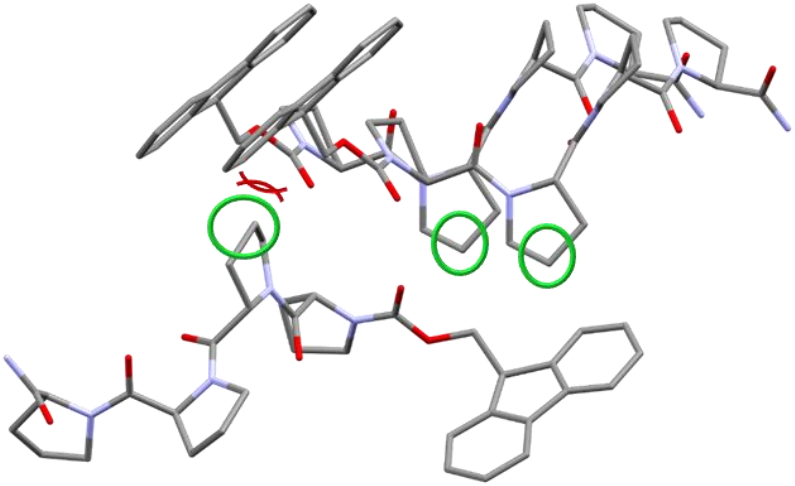


*Figure 42 – Crystal structure of **P_4**, highlighting the C^4 positions on the Pro2 residue (green), the location of the new hydroxy moiety in the **PHP_2** peptide, and the lack of close contact hydrogen bond acceptors.*

Analysing the other positions on the **P_4** backbone, the Pro2 residue on the **P_4** backbone is aligned with adjacent peptides' aromatic groups and the closest hydrogen bond acceptor (**H_A**) is 4.5 Å from the C4 position, which suggests the current crystal structure is unlikely to satisfy the hydrogen bond donation of a new hydroxyl at this position (Figure 42). Unfortunately, the peptide **PHP_2**, Fmoc-Pro-Hyp-Pro_2-NH_2, did not crystallise effectively. We hypothesise that the absence of favourably positioned **H_A** groups close to the Pro2 C4 position (closest carbonyl **H_A**, C---O 4.447(4) Å) in the **P_4** framework is impacting the assembly process. However, PDXRD of the precipitate, from the attempted crystallisation of **PHP_2** in EtOAc/EtOH, showed partial crystallinity (Figure 43).

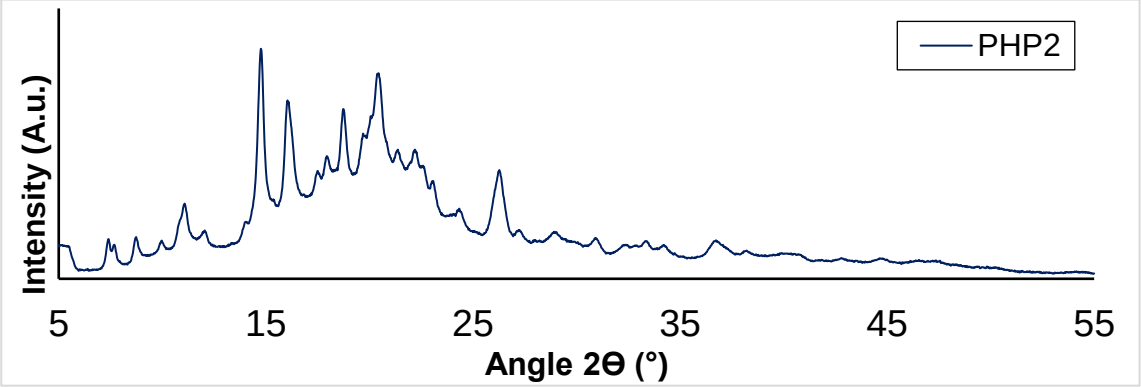


*Figure 43 - **PHP_2** experimental PD-XRD, sample was poorly crystalline and contained amorphous peptide producing a broad background peak.*

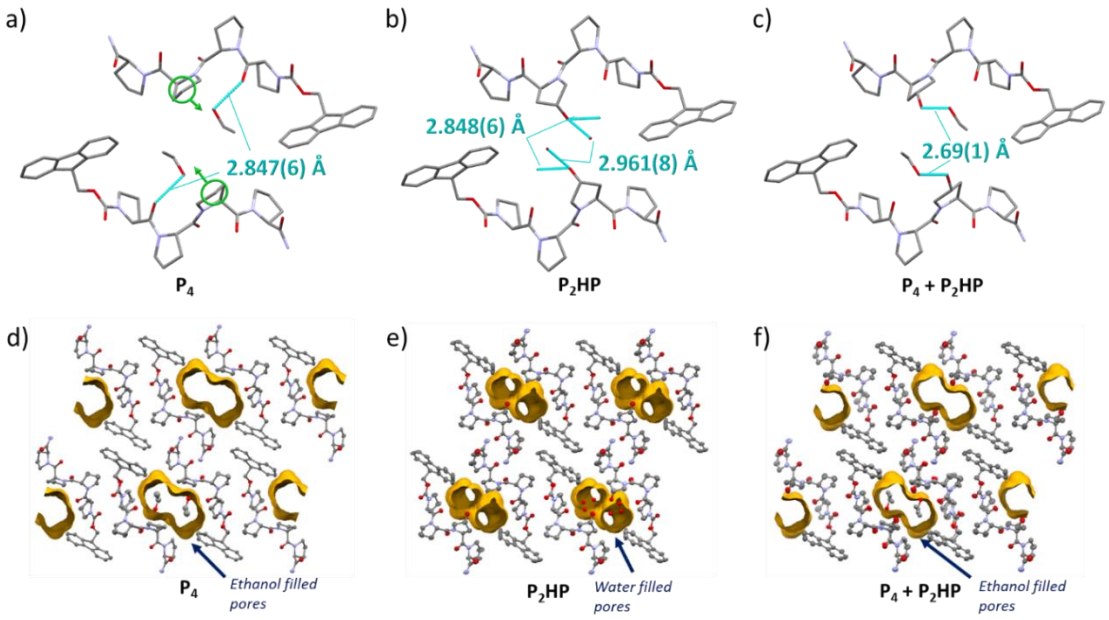


*Figure 44 – Crystal structures of **P_4** (a), **P_2HP** (b) and **P_4+P_2HP** (c) mixed crystal, depicting two peptide units adjacent to the framework channels, view along the b-axis, Mercury. Hydrogen bond interactions are shown, (c) is slightly off the b-axis to show both hydrogen bonds to different ethanol molecules. Crystal structures of **P_4** (d), **P_2HP** (e) and **P_4+P_2HP** (f) mixed crystal, showing their closely matched extended structures viewed along the b-axis, the channels are highlighted in yellow, and one solvent filled pore is shown for each, atomic displacement parameters are shown at 50% probability.*

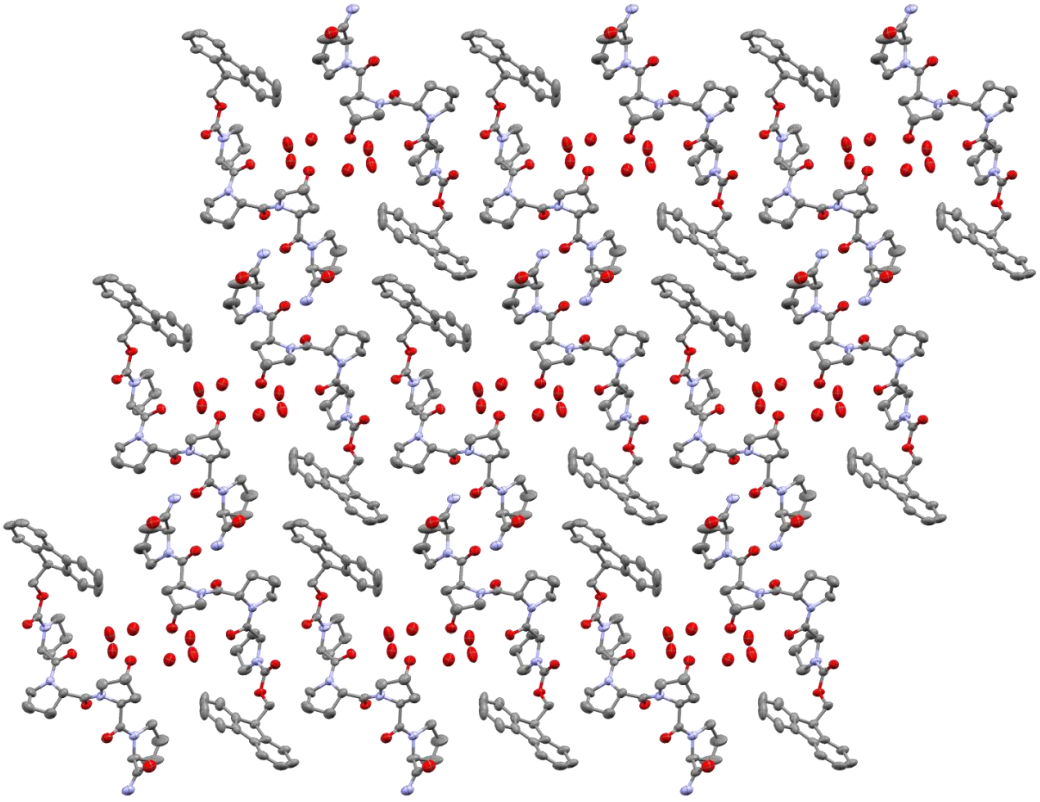


Figure 45 – Crystal structure of peptide P_2HP showing the packed extended structure (3x3x3), view along the b axis, hydrogens not shown, atomic displacement parameters are shown at 50% probability.

The Pro3 residues of the **P_4** peptide, face into the channels of the framework and as such are prime candidates for functionalisation to affect host-guest interactions (Figure 44), without disruption of any supramolecular interactions forming the extended framework. Therefore,

functionalisation at this position should have minimal impact on the adopted structure, giving rise to the same crystalline framework. As such peptide **P_2HP**, Fmoc-Pro_2-Hyp-Pro-NH_2, was synthesised and successfully crystallised with the same packing as the **P_4** framework (Figure 45), retaining polyproline II helicity. The hydroxyl groups line the previous channels of the framework, with the main structural difference being the hydroxyl forcing *exo* puckering of the attached pyrrolidine ring, typical for this functionality,[66] and the subsequent reduction in pore volume (Volume 161.8 $Å^3$, 9.2% / unit cell, Probe r = 1.2 Å, Grid spacing 0.4 Å), due to the steric bulk from the additional functional group. This reduced pore size is too small to accommodate EtOH molecules, as such the framework is selective towards H_2O molecules in the wet solvent during crystallisation, which can be clearly modelled in the crystal structure (Figure 44*b*). This change in properties of the framework was exemplified through the repetition of thermal activation studies carried out on the **P_4** framework previously, whereby encapsulated solvent could be removed resulting in the collapse of the channels forming a non-porous structure.[59] In this case heating the **P_2HP** framework under reduced pressure at 45 °C did not show the formation of a second crystalline phase (Figure 47, *Figure S57, SI 4.4.3*), showing how the solvent filled voids are now trapped in place by the additional hydroxyl moieties. This ability to functionalise the pores of the framework with no disruption of the peptide helix, thereby allowing tuning of the selectivity of the pores has clear applications towards specific host-guest interactions.

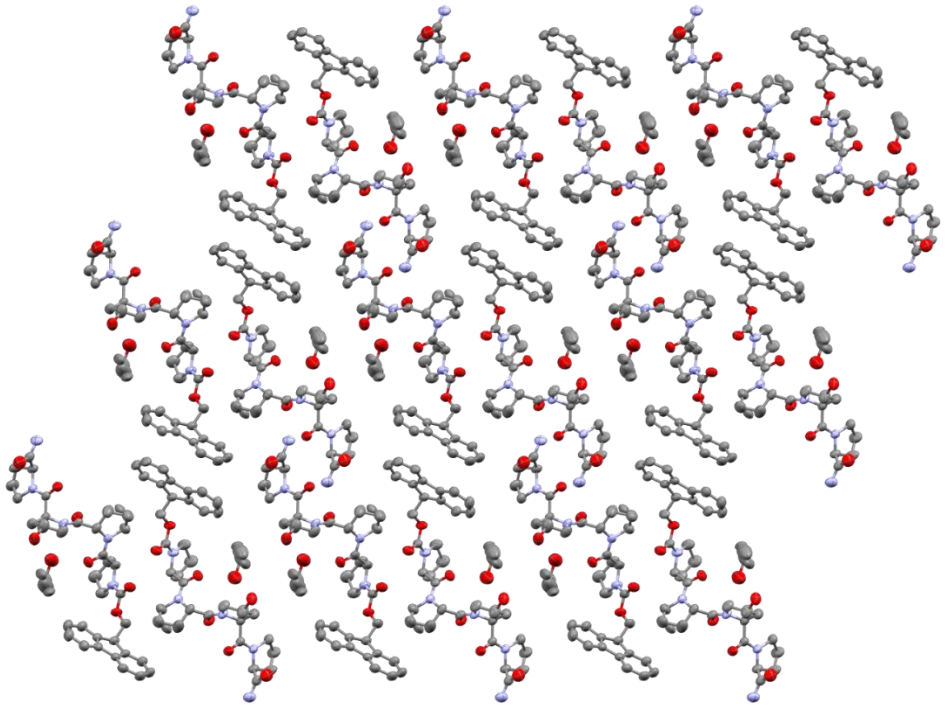


Figure 46 – Crystal structure of peptide P_4-P_2HP showing the packed extended structure (3x3x3), view along the b axis, hydrogens not shown, atomic displacement parameters are shown at 50% probability .

As **P_4** and **P_2HP** are isostructural, we theorised that both peptides would crystallise together, to yield either a co-crystalline material or a solid solution.[67] If successful, the two could be combined to create a porous framework spiked with the hydroxyproline moiety, while also reducing the effect of the limiting reduction of the channel diameter found in the **P_2HP** structure. When combined the two peptides crystallised *via* cooling a supersaturated solution of the peptides (1:1 molar ratio) in ethanol. We were pleased to see that SC-XRD analysis of the crystals formed clearly showed the presence of both peptides (Figure 46*, SI 4.3.6*). Within the pores of this SPF, the hydroxyl moiety had a reduced chemical occupancy of the oxygen atom (0.375),

indicating the lack of differentiation between the two peptides during the self-assembly process resulting in the formation of a solid solution,[67] with **P_2HP** randomly dispersed throughout the extended **P_4** framework. Interestingly this framework adopted the *endo* conformation similarly to **P_4** alone, having significantly less impact on the channel volume (Volume 233.02 Å^3, 12.8% / unit cell, Probe *r* = 1.2 Å, Grid spacing 0.4 Å), and thus contained EtOH within the pores rather than being selective towards H_2O (Figure 44*d-f*). This result is not trivial as it paves the way to the synthesis of discrete or extended supramolecular structures using different polyproline building blocks simultaneously and potentially synergistically.

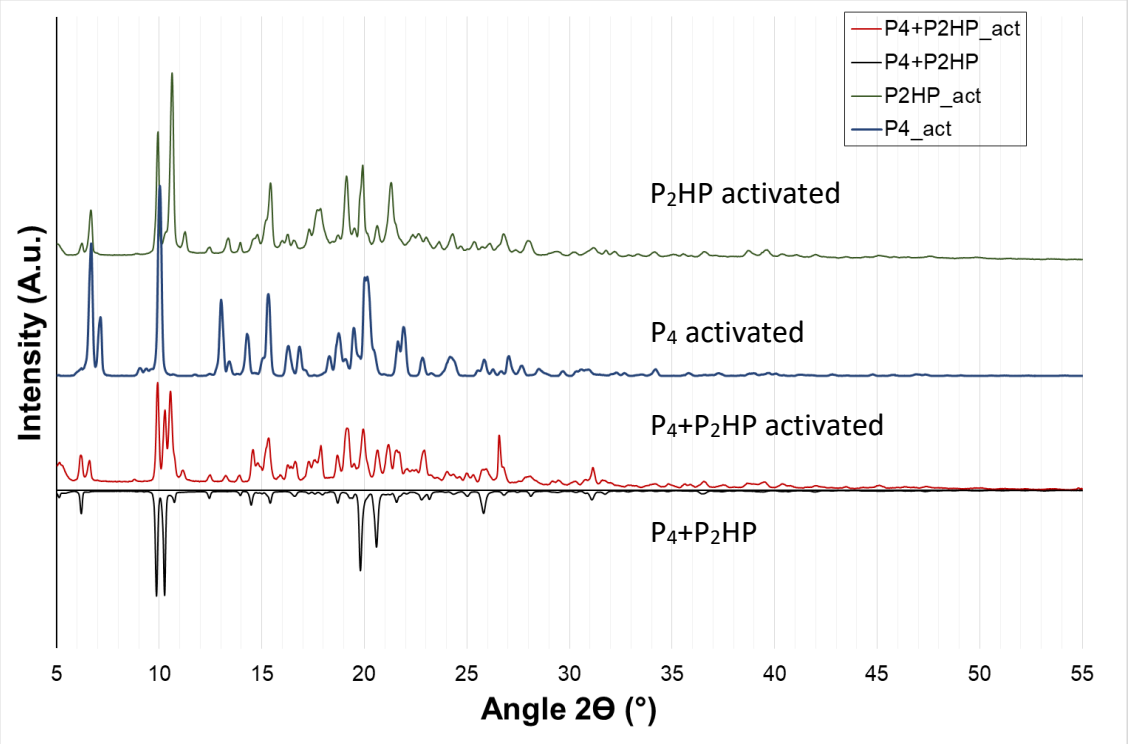


*Figure 47 – Comparative powder diffraction pattern showing; **P_4-P_2HP** experimental PD-XRD (**black, bottom**), **P_4-P_2HP** experimental PD-XRD after activation at 45 °C under vacuum (**red**), **P_4** experimental PD-XRD after activation at 45 °C under vacuum (**blue**),[59] **P_2HP** experimental PD-XRD after activation at 45 °C under vacuum (**green**). Highlights the partial activation, and subsequent change in phase, of the mixed framework compared to **P_4/ P_2HP** activation, showing similarities to both pure frameworks.*

As mentioned previously the **P_4** peptide framework could be activated by heating (45 °C) under reduced pressure,[59] with characterisation through both PDXRD and SCXRD showing encapsulated EtOH had been removed, forming a collapsed non-porous structure. As the mixed framework exhibited a very similar structure to the **P_4** framework, with EtOH encapsulated in accessible channels, while the **P_2HP** peptide framework contained water trapped in isolated pores (Figure 44*b+e*), and significantly showed no change after attempted activation (*SI 4.4.3,* Figure 47), we anticipated the mixed framework would exhibit different behaviour to the peptides alone. Interestingly, the mixed framework **P_4+P_2HP** showed hybrid behaviour compared to the two single peptide frameworks, with the powder pattern upon activation for **P_4+P_2HP** bearing similarities to both activated single peptide frameworks (i.e. **P_4** and **P_2HP**) with some formation of the new characteristic peaks seen in the **P_4** "activated" diffractogram (Figure 47*, SI 4.4.5*) while a portion of the original peaks remained. This suggested only partial "activation" and collapse of the pores, thus showing the additional hydroxyl moiety alters the pore properties, restricting the collapse of the channels.

2.2.3.2 Attempted Synthesis of Covalent Peptide Cages

With the as synthesised peptides **HP_2H** and **$AcHP_2H$** having two functional groups (-OH) on the same face, i.e. functioning as a ditopic linker, we theorised that using a multidentate linker could form a molecular cage between peptide units. As such a simple method of forming a linker between two hydroxyl groups was theorised by utilising phosphoryl chloride ($POCl_3$). This compound readily reacts with free hydroxyls to form phosphate esters($O=P(OR)_3$), in the presence of a HCl acceptor (e.g. pyridine), as such this group has the potential to act as linker between the hydroxyproline residues.[68,69] Therefore, the aim was to attempt to form a trimeric peptide cage linked by two trialkyl phosphates (Figure 48). If successful, this would be an excellent example of the potential to utilise functionalised short polyproline helices to form covalently linked supramolecular constructs. While also providing information on the degree of selectivity of the chiral helices to assemble in either anti-parallel or parallel arrangements (i.e. all peptides bonded at the same terminus to the same linker) when forming linked structures.

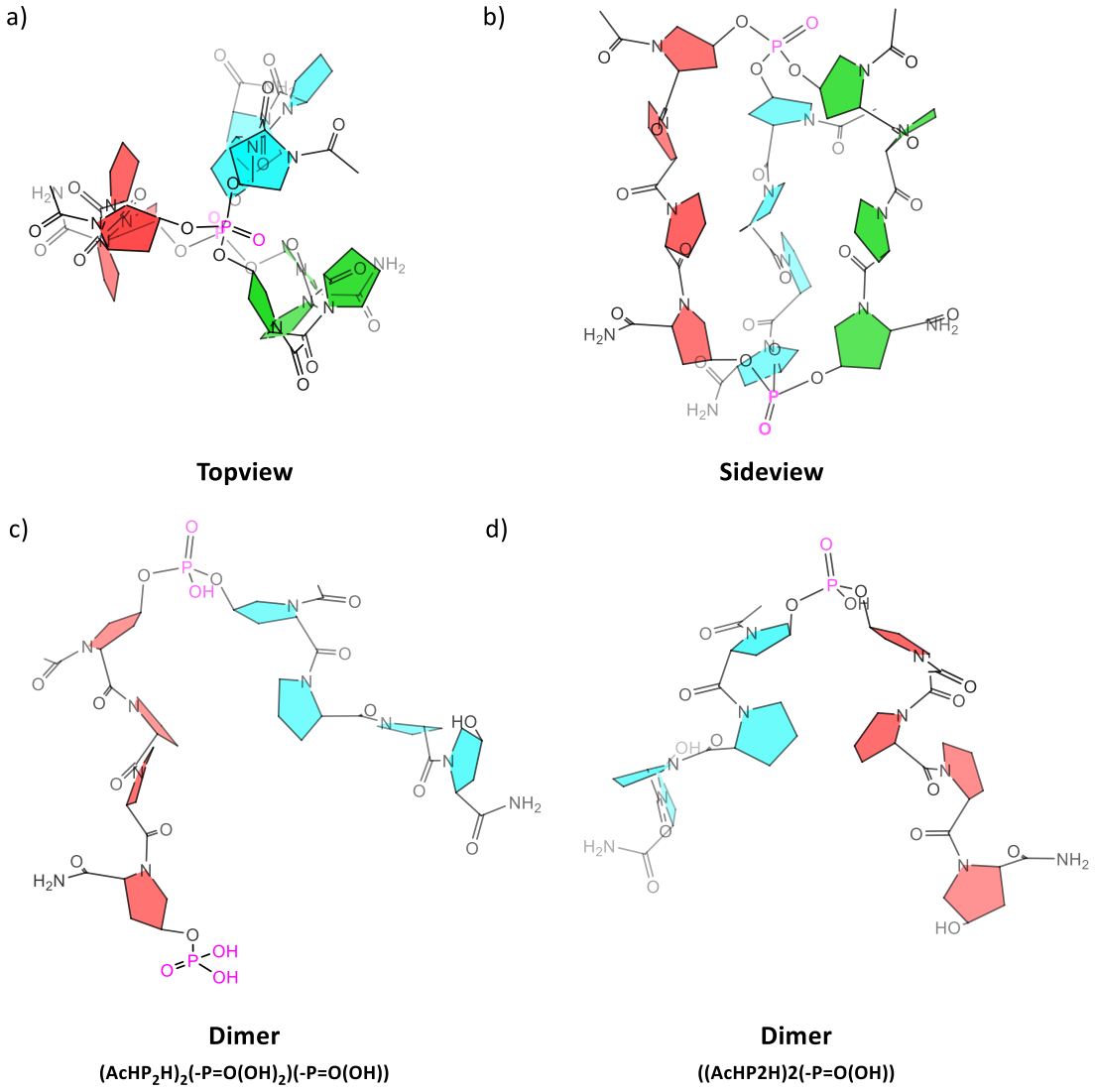


Figure 48 – (a-b) Theorised chemical structure of a trimeric covalent peptide cage of peptide $AcHP_2H$ with a phosphate linker where all peptides are parallel, (c-d) Examples of synthesised peptide phosphate ester structures determined by LCMS, further structural isomers are possible other than the examples shown.

However, repeated attempts to synthesise the cage structure were unsuccessful with only partial phosphorylation of the peptide, achieving various degrees of substitution and incomplete conversion of the starting material. The closest structure to the desired trimer achieved was isomers of a peptide dimer link with a phosphoryl group with an additional phosphorylated

hydroxyl not involved in a linkage $(AcHP_2H)_2(-P=O(OH)_2)(-P=O(OH))$ (Mass: 1101.4 $[M+H]^+$), however, the major product was a peptide dimer linked by a single phosphate ester ($(AcHP_2H)_2(-P=O(OH))$ (Mass: 1021.4 $[M+H]^+$) (Figure 48).

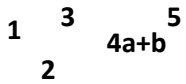


*Figure 49 - **HPLC UV-Vis** spectrum of peptide **$AcHP_2H$** after treatment with $POCl_3$ (method 1), peak 1: rt = 5.995 min **$AcHP_2H(-P=O(OH)_2)_2$**; peak 2: rt = 6.27 min, **$AcHP_2H(-P=O(OH)_2)$**;peak 3: rt = 6.755 min, **$AcHP_2H$**; peak 4a+b: rt = 7.25+7.48 min, **$(AcHP_2H)_2(-P=O(OH)_2)(-P=O(OH))$**; peak 5: rt = 8.093 min, **$(AcHP_2H)_2(-P=O(OH))$**.*

This was determined by LCMS as initial attempts to analyse *via* NMR were unsuccessful with precipitation in deuterated solvents or no visible reaction (Figure 49). This evidence suggests that subsequent esterification of the monomeric phosphate ester may be sterically hindered and becomes less energetically favourable, resulting in incomplete esterification. Alternatively, further optimisation of the reaction may be possible, with the promise from the partial completion *via* formation of a peptide dimer, to push the reaction to completion. Due to the small quantities of phosphoryl chloride involved in the reaction it is also possible trace water in the reaction is sufficient to quench the reaction forming phosphonic acid (H_3PO_3) and stalling the reaction at partial peptide substitutions, despite the use of dry solvents and attempts to scale up the reaction to larger masses. Furthermore, it is possible that the phosphate triester is readily hydrolysed in the aqueous conditions from liquid chromatography, as phosphate triesters are known to undergo more rapid hydrolysis, while diesters undergo very slow hydrolysis, which could lead to the lack of detected triester and the predominance of the diester product skewing the results.[70] Despite these setbacks this is certainly worthy of further investigation to achieve a simple route to achieve covalently linked oligoproline cages from cheap proline derivatives (hydroxyproline), with their potential application as selective nanoreactors.

2.2.3.4 Synthesis and Assembly of Further Polyproline Peptides:

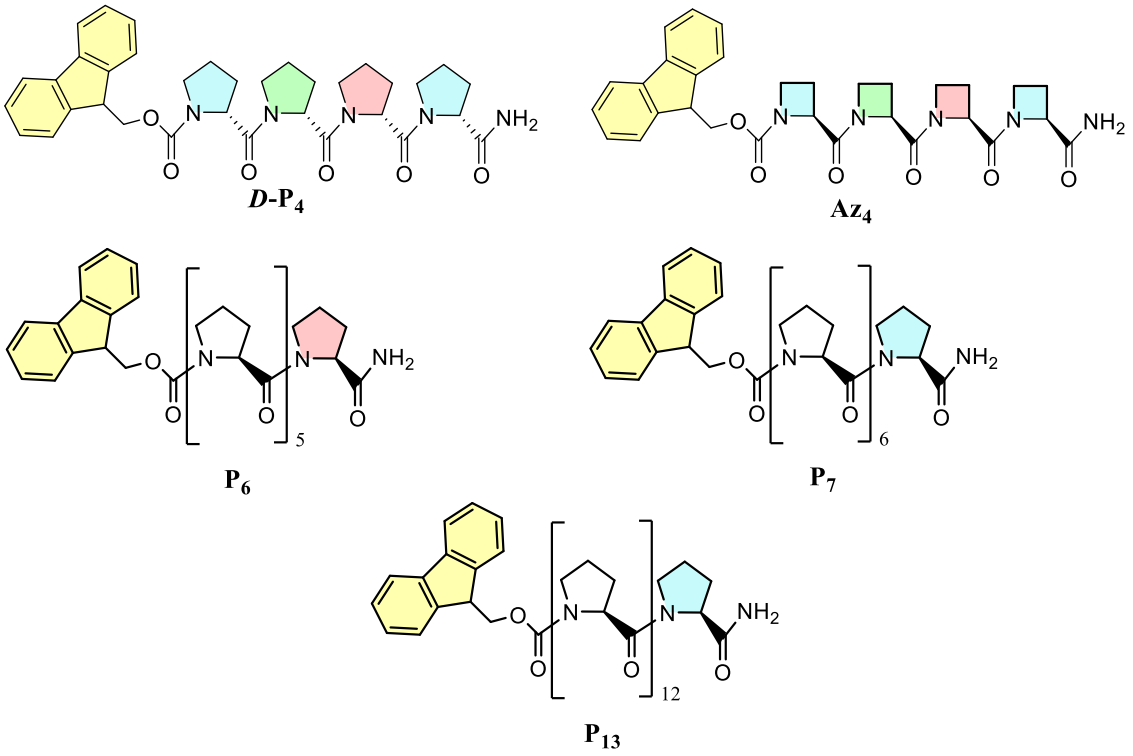


Figure 50 - Chemical structures of further synthesized oligoproline and proline derivative peptides, helical faces are highlighted in different colours (i+3 periodicity).

With the promising successes encountered with the tetrameric oligoproline and its derivatives for the construction and design of supramolecular peptide frameworks we were eager to carry this forward to alternative peptide structures. As such a series of varying lengths of Fmoc-protected polyprolines were synthesised: **P6**, **P7**, and **P13**. While two derivatives of the tetramer were also synthesised, based on the amino acids *D*-proline and azetidine: ***D*-P4** and **Az4** (Figure 50). We theorised that the peptides $P_{7, 13}$ would be able to assemble *via* the same route as the tetramer as they each contain a further helical turn so the termini of the peptides will have the same geometry of P_4 and thus minimal disruption of the packing-principles should occur as the peptide backbones assemble in rows in the crystalline framework and the bridging hydrogen bonding occurs at the terminus of the peptide. These longer assemblies, if packing using the same principles as the **P4** peptide, would then have multiple channels adjacent to each peptide with great promise for functionalisation to create differing pore environments within the same framework. To observe whether the longer peptides could switch to the polyproline I helix in solvents such as ethanol, used for the crystallisation of the oligoproline tetramers, circular dichroism (CD) spectroscopy was carried out on samples of the peptides **P6**, **P7**, and **P13**. As expected, for a polyproline with thirteen residues **P13** clearly adopted the polyproline I helix in PrOH, with the maxima (λ_{max}) shifting to 216 nm typically seen for these helices, while also showing a clear shift from the polyproline II helix in water with a λ_{max} of 228 nm to 220 nm in EtOH (Figure 51). A trough at ≈ 238 nm is also visible in PrOH and EtOH, indicative of the adoption of the polyproline I helix, alongside the shift of the minima to a lower wavelength.[29] This could potentially prevent crystallisation in a similar manner to the **P4** peptide in solvents such as EtOH as the helix has adopted the significantly different polyproline I conformation.

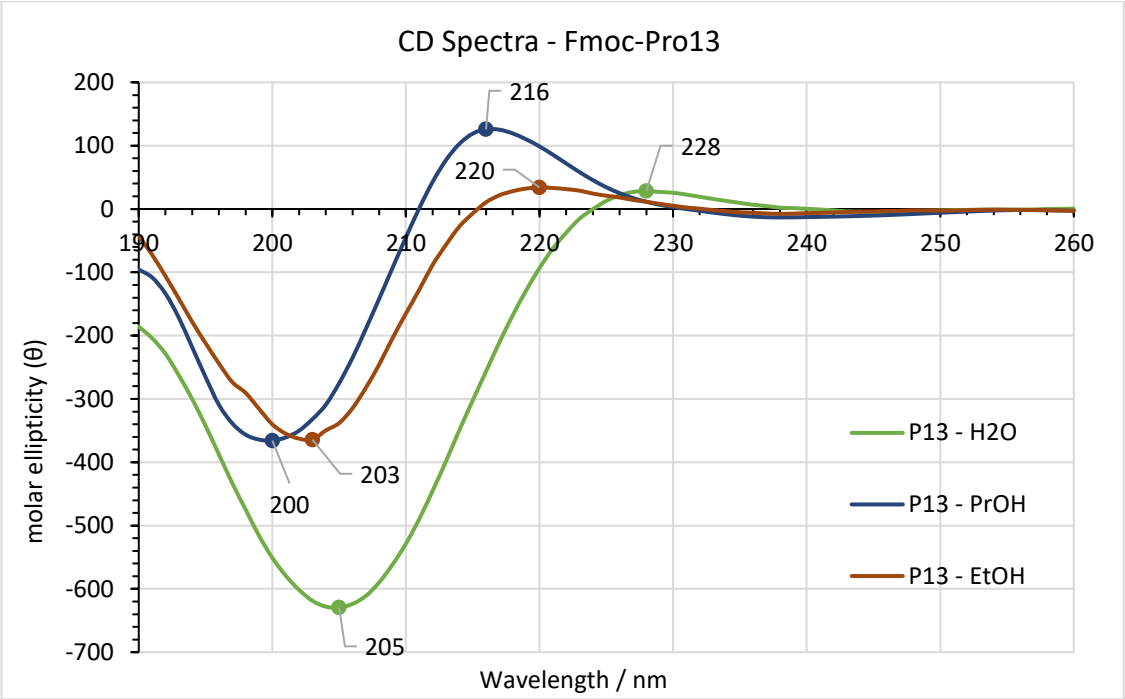


Figure 51 – CD spectra of peptide P_{13}, Fmoc-Pro_{13}, in water, EtOH and PrOH (125 µM), incubated for 15 days.

However, as with the **P_4** peptide, **P_6** and **P_7** showed no clear shift in maxima or minima when incubating in EtOH instead of water (Figure 52+52), suggesting that the polyproline II helix is retained for these peptides and their length is insufficient to stabilise the adoption of the *cis* amide bonds to form the polyproline I conformation. However, attempts to crystallise these peptides, were unsuccessful forming either glasses or oily solids. Yet it is possible these may still hold potential, and a more thorough screening of crystallisation conditions is required. The lengthening of the peptides also increases the hydrophobicity of the peptide, due to increased exposure of the hydrophobic peptide backbone, and this may well affect the crystallisation and drive self-assembly in a different manner to satisfy the increasingly significant hydrophobicity of the peptide, as hydrophobicity is well known to have a not insignificant impact on the self-assembly of peptides (e.g. amphiphiles).[71,72]

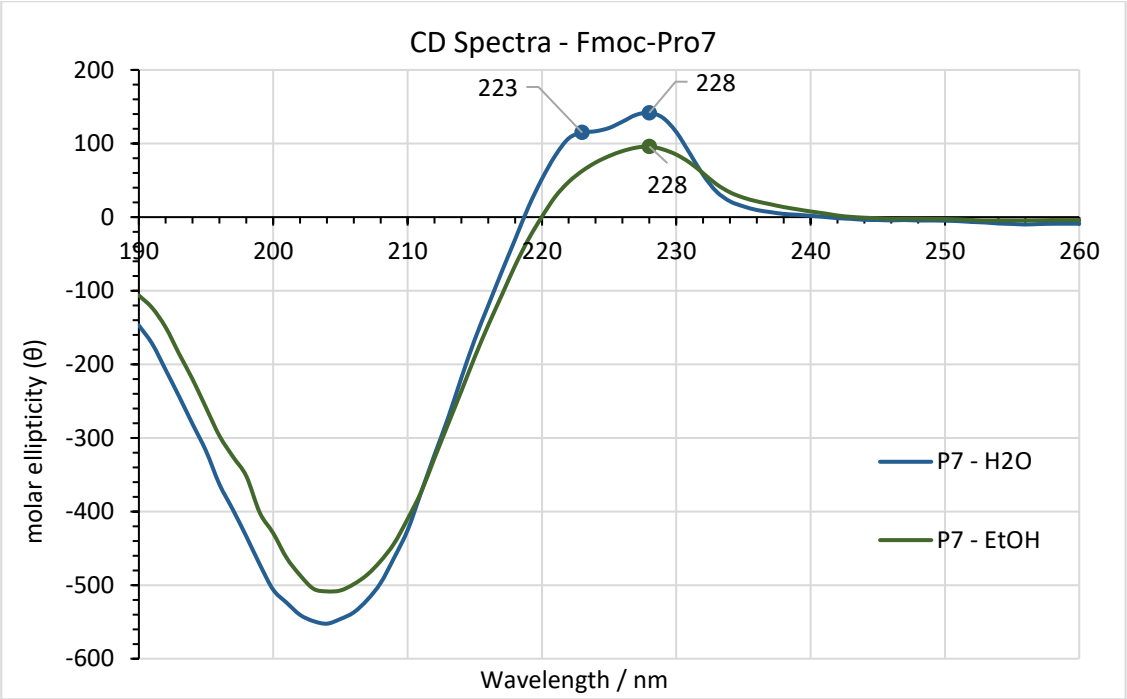


Figure 52 - CD spectra of peptide P_7, Fmoc-Pro_7, in water and EtOH (125 µM), incubated for 15 days.

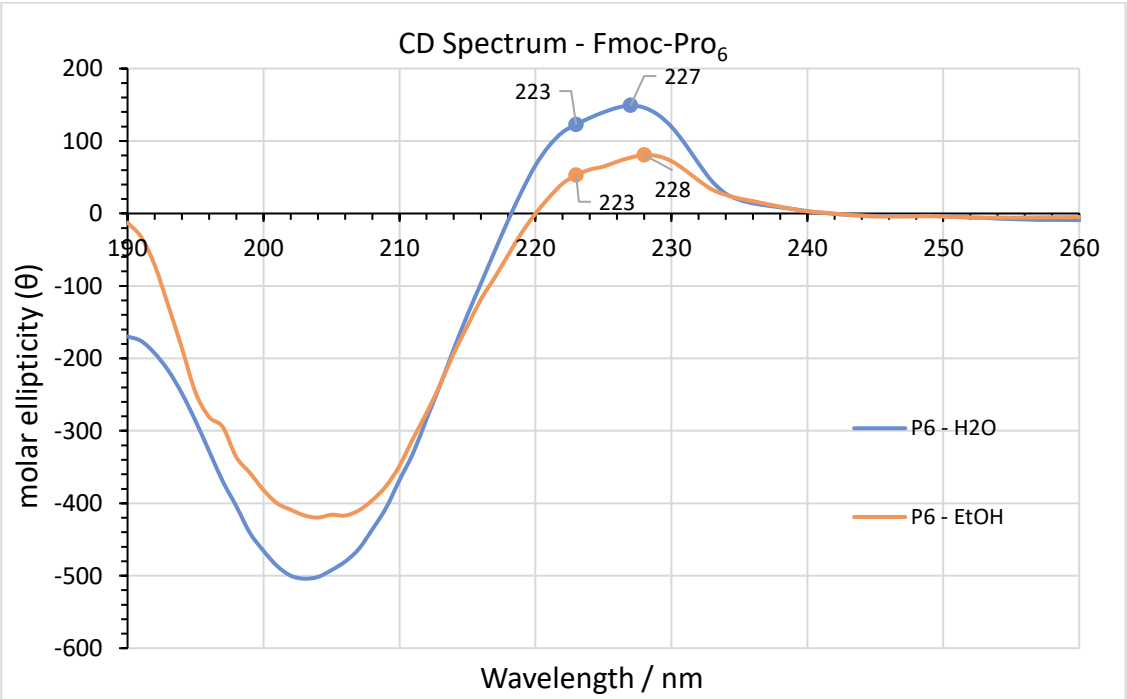


Figure 53 - CD spectra of peptide P_6, Fmoc-Pro_6, in water and EtOH (125 µM), incubated for 15 days.

We were particularly interested in the use of azetidine within the peptide framework to observe the impact of the smaller ring on the helical parameters and the adopted geometry of the monomer within the helix. It has been previously shown that azetidine residues can increase the favourability of the helix to adopt *cis* amide bonds, shifting to the polyproline I helix.[73] Thus, the incorporation of azetidine into polyproline helices could be a useful method for the tuning of the sensitivity of the peptide, increasing the propensity to switch conformations (polyproline II helix → polyproline I helix), clearly useful for optimisation of stimuli responsive supramolecular assemblies. However, more information is required on the effect of this residue on the peptide secondary structure to be able to rationally design any supramolecular building-block

successfully. With this goal in mind Fmoc-Azetidine was synthesised and used to synthesise a tetrameric peptide. Subsequently, attempts were made to crystallise the azetidine tetramer, with the additional hope that due to the lack of changes to any of the strongly interacting moieties (i.e., H-bonding) that this would form a framework isostructural to **P_4**, thus increasing the size of the channels within the framework due to the reduced sidechain ring sizes. However, all attempts to crystallise the peptide were unsuccessful forming amorphous glasses or oily solids. This suggests that changing all the residues to the azetidine derivative has a more significant impact on the self-assembly principles and potentially the adopted peptide secondary structure. As such, synthesis of partially substituted derivatives of the peptide would be more sensible, to systematically observe the impact of azetidine on the adopted structure (e.g., Az-Pro_3, Az_2-Pro_2, etc).

To investigate the effect of the helix chirality on the properties of the supramolecular assembly ***D*-P_4** was synthesised. Natural amino acids have the L configuration, except for achiral glycine, however, D-amino acids are readily available artificially. Incorporation of D-amino acids into peptide-based supramolecular frameworks can thus be used to produce the mirror-image of the framework. Where this can be applied to enantioselective processes, such as shown previously with the selectivity of the **P_4** crystalline framework towards *S*-1-phenylethanol, this can theoretically switch the selectivity towards the opposite isomers of guest molecules. This peptide was successfully crystallised as the mirror image to **P_4** and thus the selectivity of the framework towards chiral molecules can be investigated to determine that the chirality of the amino acids used, and thus the handedness of the peptide helix, is the driving force for the selectivity of the framework. This would be a highly beneficial finding for the rational design of peptide-based materials for specific host-guest interactions.

2.2.4 Conclusion

In conclusion, we have synthesised a series of hydroxyproline-based derivatives of an oligoproline tetramer, successfully forming novel H-bonding driven supramolecular peptide frameworks. Using rational design, based on the original supramolecular framework, from the placement of additional functional groups we were able to consistently predict the position of the hydroxyl groups in relation to the polyproline II helix and therefore anticipate their effect on the topology of the H-bonding supramolecular network formed. This approach is not trivial, and our results are remarkable as we use a peptide-based supramolecular building block. Moreover, engineering polyproline peptides, through sidechain functionalisation and peptide capping, we can modulate the nature of the frameworks (i.e., porous, or non-porous) as well as the size and the properties of the pores (e.g., solid solution framework), while functionalising up to 50% of the peptide's backbone. The resilience of the polyproline II helix, is crucial in our methodology as it allow us to use their predictable geometries for the rational design of discrete and extended supramolecular three-dimensional structures. This work lays the groundwork for further studies focusing on the polyproline helix, to rationally design structural units capable of forming desired supramolecular structures with tuneable structural features and functionalities.

2.2.5 Contributions

Single crystal x-ray diffraction analyses and resolutions were carried out in part by G.Truccolo and D.F.Brightwell. The original manuscript was written through contributions of all authors.

CCDC-2127750,[59] *2238152, 2238155, 2238160, 2238161, 2238180, 2238252, 2234312, 2264145, and 2178857 contain the supplementary crystallographic data for this work, including structure factors and refinement instructions, and can be obtained free of charge from The Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK (e-mail: deposit@ccdc.cam.ac.uk), or via https://www.ccdc.cam.ac.uk/getstructures.*

2.3 References

1 J. Rabone, Y.-F. Yue, S. Y. Chong, K. C. Stylianou, J. Bacsa, D. Bradshaw, G. R. Darling, N. G. Berry, Y. Z. Khimyak, A. Y. Ganin, P. Wiper, J. B. Claridge and M. J. Rosseinsky, *Science (80-.).*, 2010, **329**, 1053–1057.

2 A. P. Katsoulidis, K. S. Park, D. Antypov, C. Martí-Gastaldo, G. J. Miller, J. E. Warren, C. M. Robertson, F. Blanc, G. R. Darling, N. G. Berry, J. A. Purton, D. J. Adams and M. J. Rosseinsky, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 193.

3 A. P. Katsoulidis, D. Antypov, G. F. S. Whitehead, E. J. Carrington, D. J. Adams, N. G. Berry, G. R. Darling, M. S. Dyer and M. J. Rosseinsky, *Nature*, 2019, **565**, 213–217.

4 A. Comotti, S. Bracco, G. Distefano and P. Sozzani, *Chem. Commun.*, 2009, 284–286.

5 A. Comotti, A. Fraccarollo, S. Bracco, M. Beretta, G. Distefano, M. Cossi, L. Marchese, C. Riccardi and P. Sozzani, *CrystEngComm*, 2013, **15**, 1503–1507.

6 V. N. Yadav, A. Comotti, P. Sozzani, S. Bracco, T. Bonge-Hansen, M. Hennum and C. H. Görbitz, *Angew. Chemie Int. Ed.*, 2015, **54**, 15684–15688.

7 J. Bonnefoy, A. Legrand, E. A. Quadrelli, J. Canivet and D. Farrusseng, *J. Am. Chem. Soc.*, 2015, **137**, 9409–9416.

8 J. Navarro-Sánchez, A. I. Argente-García, Y. Moliner-Martínez, D. Roca-Sanjuán, D. Antypov, P. Campíns-Falcó, M. J. Rosseinsky and C. Martí-Gastaldo, *J. Am. Chem. Soc.*, 2017, **139**, 4294–4297.

9 A. Saito, T. Sawada and M. Fujita, *Angew. Chemie Int. Ed.*, 2020, **59**, 20367–20370.

10 T. Schnitzer, E. Paenurk, N. Trapp, R. Gershoni-Poranne and H. Wennemers, *J. Am. Chem. Soc.*, 2021, **143**, 644–648.

11 C. Martí-Gastaldo, D. Antypov, J. E. Warren, M. E. Briggs, P. A. Chater, P. V. Wiper, G. J. Miller, Y. Z. Khimyak, G. R. Darling, N. G. Berry and M. J. Rosseinsky, *Nat. Chem.*, 2014, **6**, 343–351.

12 A. Mantion, L. Massüger, P. Rabu, C. Palivan, L. B. McCusker and A. Taubert, *J. Am. Chem. Soc.*, 2008, **130**, 2517–2526.

13 S. Mehrparvar, C. Wölper, R. Gleiter and G. Haberhauer, *Angew. Chemie Int. Ed.*, 2020, **59**, 17154–17161.

14 S. Matsubara, Y. Okamoto, M. Yoshikawa, S. Tsukiji and M. Higuchi, *Bioconjug. Chem.*, 2022, **33**, 1785–1788.

15 S. Zhang, Y. Zheng, H. An, B. Aguila, C. X. Yang, Y. Dong, W. Xie, P. Cheng, Z. Zhang, Y. Chen and S. Ma, *Angew. Chemie Int. Ed.*, 2018, **57**, 16754–16759.

16 R. Misra, A. Saseendran, S. Dey and H. N. Gopi, *Angew. Chemie*, 2019, **131**, 2273–2277.

17 P. Li, M. R. Ryder and J. F. Stoddart, *Accounts Mater. Res.*, 2020, **1**, 77–87.

18 W. Yang, A. Greenaway, X. Lin, R. Matsuda, A. J. Blake, C. Wilson, W. Lewis, P. Hubberstey, S. Kitagawa, N. R. Champness and M. Schröder, *J. Am. Chem. Soc.*, 2010, **132**, 14457–14469.

19 J. Tian, L. Chen, D. W. Zhang, Y. Liu and Z. T. Li, *Chem. Commun.*, 2016, **52**, 6351–6362.

20 C. H. Görbitz, *Chem. – A Eur. J.*, 2007, **13**, 1022–1031.

21 G. Distefano, A. Comotti, S. Bracco, M. Beretta and P. Sozzani, *Angew. Chemie Int. Ed.*, 2012, **51**, 9258–9262.

22 M. R. Aronoff, J. Egli, A. Schmitt and H. Wennemers, *Chem. – A Eur. J.*, 2020, **26**, 5070–5074.

23 R. Dukor and T. Keiderling, *Biopolymers*, 1991, **31**, 1747–1761.

24 T. J. El-Baba, D. R. Fuller, D. A. Hales, D. H. Russell and D. E. Clemmer, *J. Am. Soc. Mass*

Spectrom., 2019, **30**, 77–84.

25 C. Siebler, R. S. Erdmann and H. Wennemers, *Angew. Chemie*, 2014, **126**, 10508–10512.

26 L.-S. Sonntag, S. Schweizer, C. Ochsenfeld and H. Wennemers, *J. Am. Chem. Soc.*, 2006, **128**, 14697–14703.

27 M. Kümin, L.-S. Sonntag and H. Wennemers, *J. Am. Chem. Soc.*, 2007, **129**, 466–467.

28 V. Kubyshkin and N. Budisa, *Org. Biomol. Chem.*, 2017, **15**, 619–627.

29 S. Kakinoki, Y. Hirano and M. Oka, *Polym. Bull.*, 2005, **53**, 109–115.

30 P. Wilhelm, B. Lewandowski, N. Trapp and H. Wennemers, *J. Am. Chem. Soc.*, 2014, **136**, 15829–15832.

31 G. Berger, M. Vilchis-Reyes and S. Hanessian, *Angew. Chemie Int. Ed.*, 2015, **54**, 13268–13272.

32 T. Matsuzaki, *Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem.*, 1974, **30**, 1029–1036.

33 R. W. Newberry and R. T. Raines, *Acc. Chem. Res.*, 2017, **50**, 1838–1846.

34 G. B. McGaughey, M. Gagné and A. K. Rappé, *J. Biol. Chem.*, 1998, **273**, 15458–15463.

35 M. Nishio, *Phys. Chem. Chem. Phys.*, 2011, **13**, 13873–13900.

36 S. Sakaki, K. Kato, T. Miyazaki, Y. Musashi, K. Ohkubo, H. Lhara and C. Hirayama, *J. Chem. Soc. Faraday Trans.*, 1993, **89**, 659–664.

37 P. Wilhelm, B. Lewandowski, N. Trapp and H. Wennemers, *J. Am. Chem. Soc.*, 2014, **136**, 15829–15832.

38 Susumu Kitagawa and Kazuhiro Uemura, *Chem. Soc. Rev.*, 2005, **34**, 109–119.

39 H. Wang, B. Li, H. Wu, T. L. Hu, Z. Yao, W. Zhou, S. Xiang and B. Chen, *J. Am. Chem. Soc.*, 2015, **137**, 9963–9970.

40 P. Cui, E. Svensson Grape, P. R. Spackman, Y. Wu, R. Clowes, G. M. Day, A. K. Inge, M. A. Little and A. I. Cooper, *J. Am. Chem. Soc.*, 2020, **142**, 12743–12750.

41 H. Yamagishi, H. Sato, A. Hori, Y. Sato, R. Matsuda, K. Kato and T. Aida, *Science (80-.).*, 2018, **361**, 1242–1246.

42 R. Piotrowska, T. Hesketh, H. Wang, A. R. G. Martin, D. Bowering, C. Zhang, C. T. Hu, S. A. McPhee, T. Wang, Y. Park, P. Singla, T. McGlone, A. Florence, T. Tuttle, R. V. Ulijn and X. Chen, *Nat. Mater.*, 2020, **20**, 403–409.

43 I. Imaz, M. Rubio-Martínez, J. An, I. Solé-Font, N. L. Rosi and D. Maspoch, *Chem. Commun.*, 2011, **47**, 7287–7302.

44 Y. Inokuma, T. Arai and M. Fujita, *Nat. Chem. 2010 29*, 2010, **2**, 780–783.

45 J. L. Beesley, H. E. Baum, L. R. Hodgson, P. Verkade, G. S. Banting and D. N. Woolfson, *Nano Lett.*, 2018, **18**, 5933–5937.

46 J. Y. Rho, J. C. Brendel, L. R. MacFarlane, E. D. H. Mansfield, R. Peltier, S. Rogers, M. Hartlieb and S. Perrier, *Adv. Funct. Mater.*, 2018, **28**, 1704569.

47 S. Rojas, T. Devic and P. Horcajada, *J. Mater. Chem. B*, 2017, **5**, 2560–2573.

48 D. L. Caulder and K. N. Raymond, *Acc. Chem. Res.*, 1999, **32**, 975–982.

49 B. Olenyuk, J. A. Whiteford, A. Fechtenkötter and P. J. Stang, *Nature*, 1999, **398**, 796–799.

50 B. Sun, M. Bilal, S. Jia, Y. Jiang and J. Cui, *Korean J. Chem. Eng.*, 2019, **36**, 1949–1964.

51 J. Baillet, V. Desvergnes, A. Hamoud, L. Latxague, P. Barthélémy, S. J. Assemblies Baillet, V. Desvergnes, A. Hamoud, L. Latxague and P. Barthélémy, *Adv. Mater.*, 2018, **30**, 1705078.

52 X. Zhang, C. Gong, O. U. Akakuru, Z. Su, A. Wu and G. Wei, *Chem. Soc. Rev.*, 2019, **48**, 5564–5595.

53 Y. Huo, J. Hu, Y. Yin, P. Liu, K. Cai and W. Ji, *ChemBioChem*, 2023, **24**, e202200582.

54 F. Sheehan, D. Sementa, A. Jain, M. Kumar, M. Tayarani-Najjaran, D. Kroiss and R. V. Ulijn, *Chem. Rev.*, 2021, **121**, 13869–13914.

55 R. Behrendt, P. White and J. Offer, *J. Pept. Sci.*, 2016, **22**, 4–27.

56 P. Morales and M. A. Jiménez, *Arch. Biochem. Biophys.*, 2019, **661**, 149–167.

57 S. Dobitz, M. R. Aronoff and H. Wennemers, *Acc. Chem. Res.*, 2017, **50**, 2420–2428.

58 U. Lewandowska, W. Zajaczkowski, S. Corra, J. Tanabe, R. Borrmann, E. M. Benetti, S. Stappert, K. Watanabe, N. A. K. Ochs, R. Schaeublin, C. Li, E. Yashima, W. Pisula, K. Müllen and H. Wennemers, *Nat. Chem.*, 2017, **9**, 1068–1072.

59 D. F. Brightwell, G. Truccolo, K. Samanta, E. J. Fenn, S. J. Holder, H. J. Shepherd, C. S. Hawes and A. Palma, *Chem. – A Eur. J.*, 2022, **28**, e202202368.

60 P. Makam and E. Gazit, *Chem. Soc. Rev.*, 2018, **47**, 3406–3420.

61 P.-Y. Hung, Y.-H. Chen, K.-Y. Huang, C.-C. Yu and J.-C. Horng, *ACS Omega*, 2017, **2**, 5574–5581.

62 Z. Shi, K. Chen, Z. Liu and N. R. Kallenbach, *Chem. Rev.*, 2006, **106**, 1877–1897.

63 M. Kümin, L.-S. Sonntag and H. Wennemers, *J. Am. Chem. Soc.*, 2007, **129**, 466–467.

64 A. Lesarri, E. J. Cocinero, J. C. López and J. L. Alonso, *J. Am. Chem. Soc.*, 2005, **127**, 2572–2579.

65 E. Benedetti, A. Bavoso, B. di Blasio, V. Pavone, C. Pedone, C. Toniolo and G. M. Bonora, *Biopolymers*, 1983, **22**, 305–317.

66 H. K. Ganguly and G. Basu, *Biophys. Rev.*, 2020, **12**, 25–39.

67 M. Vert, Y. Doi, K.-H. Hellwich, H. Michael, P. Hodge, R. Przemyslaw, Kubisa Marguerite and F. Schué, *Pure Appl. Chem.*, 2012, **84**, 377–410.

68 C. R. Noller and G. R. Dutton, *J. Am. Chem. Soc.*, 1933, **55**, 424–425.

69 R. A. Aitken, C. J. Collett and S. T. E. Mesher, *Synthesis (Stuttg).*, 2012, **44**, 2515–2518.

70 A. J. Kirby and F. Nome, *Acc. Chem. Res.*, 2015, **48**, 1806–1814.

71 H. Cui, M. J. Webber and S. I. Stupp, *Pept. Sci.*, 2010, **94**, 1–18.

72 M. P. Hendricks, K. Sato, L. C. Palmer and S. I. Stupp, *Acc. Chem. Res.*, 2017, **50**, 2440–2448.

73 J. S. Jhon and Y. K. Kang, *J. Phys. Chem. B*, 2007, **111**, 3496–3507.

Chapter 3.

Polyproline helices as ligands in MOFs and cages

3.1 Abstract

Synthesising supramolecular constructs capable of mimicking the catalytic activity of enzymes, to achieve enhanced reaction rates and high catalytic selectivity, has long been a pursuit of chemists. Inspired by these natural systems the concept of artificial enzyme mimetics has been developed since the earliest examples of supramolecular catalysts. While catalysis inside supramolecular cavities has been reported under mild and biomimetic conditions the chemical manipulation of these cavities is typically distinctly limited. This presents a clear limitation for the exploitation of these assemblies for a broad range of applications when utilising typical building-blocks. However, discrete nano-constructs such as metallo-organic rings and cages, as non-covalent hosts, offer a modular approach to their assembly process *via* supramolecular interactions. Seminal work by Stang, Raymond, Fujita, and Nitschke identifies several design principles allowing the prediction of the geometry of new metallo-organic constructs. Furthermore software has been developed to rationally design these types of constructs and predict their properties, however, these design methods are often limited to highly symmetric and polyaromatic ligands to simplify the assembly processes. Thus, access to highly functionalised, asymmetric constructs is restricted and presents a significant roadblock to the design and synthesis of enzyme mimetics. Significantly, some reports have been presented that move beyond these limitations, Fujita has reported the use of oligopeptides with terminal pyridines to form coordination networks and entangled metallo-peptide rings. With this research we aim to utilise peptides as a unique class of building-block for the bottom-up design and synthesis of supramolecular constructs to achieve nano-cavities with unique applications. The stepwise synthesis of peptides combined with a focus on the polyproline helix, with a uniquely rigid secondary structure capable of stimuli responsiveness, will allow unprecedented control over the morphology and cavity topology making further progress towards the synthesis of functional biomimetic systems. Herein, we demonstrate how polyproline helices can be utilised for the design and synthesis of metal-organic constructs that assemble to form a variety of nanostructures, either forming extended structures, such as nanoparticles and gels, or discrete complexes depending on the functionalisation of the peptide helix. With designs modelled from the secondary structure of the helix, functional groups can be specifically placed within the peptide sequence to predictably alter the supramolecular interactions to directly impact the topology of the achieved complexes. The stimuli responsiveness of the polyproline helix (polyproline I -> II) allowed various morphologies and topologies to be achieved from the same ligand *via* changing the solvent environment of these assemblies. With these successes we demonstrate that polyproline helices, that are both chiral and asymmetrical, can be effectively used to achieve metal-organic assemblies with the potential to form nanocavities useful as biomimetics for supramolecular catalysis and molecular *trans*port.

3.2 Introduction

Metal-organic interactions have long been used to assemble supramolecular assemblies with metal ions acting as nodes between coordinating organic ligands.[1,2] These host molecules are classified as non-covalent hosts, unlike covalent hosts such as macrocyclic structures held together by covalent interactions; cyclodextrins, cucurbiturils, and calixarenes,[3] which have been shown to bind specific guests within their cavities and even achieve catalysis under mild, biomimetic conditions.[4] These non-covalent hosts, in contrast, offer a modular assembly process *via* reversible supramolecular interactions, enabling the formation of ordered assemblies through combining components with complementary interactions while their flexibility and dynamic behaviour enables guest accommodation. Also, metal coordinate bonds can offer varying geometries and bond strengths unlike weaker interactions (Hydrogen bonds, dispersion interactions, etc) for the design of these supramolecular constructs.[5]

These metal-organic assemblies can be designed to adopt a variety of geometries. These can be extended structures such as metal-organic frameworks, nanotubes, and coordination polymers or discrete assemblies such as metal-organic cages, macrocycles, polyhedra and nanoparticles. Work by pioneering researchers such as Stang, Raymond, Fujita and Nitschke has been instrumental in identifying several key design principles to predict the topology of these metal-organic assemblies, varying the metal to ligand ratio and the geometry of the ligands about the metal nodes.[6–10] These findings combined with ever improving computational methods[11] to assist in the prediction of the assembly and properties of these constructs has allowed the rational design of metal-organic assemblies to form predictable nanocavities with the aim of tuning these spaces to mimic the catalytic effect of enzymes.[3,12]

This creation of a catalytic pocket, at least partially enclosed by the remaining supramolecular structure, functions similarly to enzyme catalysts with this nano-environment creating several important effects. The function of this restricted space is based off the "induced fit" model whereby a substrate with a specific shape and functionality induces conformational changes in the both the enzyme and substrate to achieve enzyme binding and react.[13] Thus a specific size, shape and placement of functional groups is required to bind to the flexible enzyme binding site.[14] Further important factors can be the desolvation of the reactants, destabilisation of the ground state substrate or stabilisation of *trans*ition states from interactions within the cavity, and inducing favourable conformations of the product. While the hydrophobicity of the nanopore while in an aqueous environment (or vice versa) can encapsulate the substrates increasing the concentration about the reactive site to increase the rate of reaction, alongside non-covalent interactions trapping guests within the host molecule and impacting the stability of the *trans*ition states. In enzyme catalysis these factors can generate rate enhancements of up to 10^{10} mol L^{-1} s^{-1},[3] therefore the great potential for artificial enzymes can be clearly seen and these properties can be designed into artificial confined environments.[14]

However, from the literature it is immediately evident the prediction and design of these assemblies has been largely limited to polyaromatic and symmetric organic ligands due to the simplification of the assembly processes. This common approach, despite offering a diverse platform yielding significant successes and allowing for simplification of computation modelling for prediction of structures, also raises significant restrictions and limitations. These ligands typically do not allow access to the synthesis of asymmetric constructs or highly functionalised cavities contrasting significantly with the high level of structural and functional complexity encountered in enzymatic hosts. They are also often difficult to manipulate requiring multi-step syntheses to achieve new functionalisations and limiting the placement and geometry of binding motifs, and areas for non-covalent and secondary interactions within the cavities, such as for host-guest interactions. This undoubtedly impacts the usefulness of these materials for the design of biomimetic catalytic cavities. This clear limitation of these materials has led to some encouraging reports within literature moving towards the use of asymmetric constructs with promising applications as asymmetric catalysts.[15–17] However, significant challenges still remain when attempting to synthesise asymmetric constructs that exhibit high levels of chemical complexity within their cavities, which are essential to tune the host-guest interactions to direct the catalytic effects. Thus, moving away from polyaromatic structures researchers have looked towards nature for inspiration, leading to the use of biomaterials, such as DNA, RNA, proteins, peptides, and amino acids as supramolecular building-blocks that are both inherently chiral and biocompatible.[18–23] Significantly, with the goal of achieving nanoreactors more akin to those found in nature, Fujita has extensively reported the use of simple oligopeptides terminally functionalised with pyridines which coordinate to metal centres to form either coordination networks[24,25] or metallo-peptide rings, polyhedra and entangled knots and coils[26–28] which

contain nanocavities. While promising work by Carlos Marti-Gastaldo has shown the use of peptide-metal organic frameworks for the enantioselective separation of the chiral drugs ephedrine and methamphetamine, the first example of a MOF capable of separating chiral polar drugs,[29] and that exhibit sponge-like behaviour and high catalytic activity in the Henry reaction.[30] These reports exemplify how peptides can be utilised as chiral, biocompatible ligands for the synthesis of asymmetric, highly functionalised cavities with highly promising applications and have been instrumental in redirecting this field towards these types of supramolecular building-blocks. However, the functionalisation and optimisation of these peptidic nanomaterials remains highly challenging as the design principles applied to rigid, symmetrical, polyaromatic ligands cannot be easily applied to peptide-based ligands and therefore further development of these design principles and understanding of assembly processes for this unique class of ligands is essential. Consequently, the aim of this research is to *trans*late the design principles utilised for classical metallo-organic system to peptide-based ligands in order to rationally design and synthesise chiral, highly functionalised, tuneable cavities that can act as enzyme mimetics and have further applications, such as for drug delivery vehicles.

Analysing previous studies on metal-organic constructs the high flexibility of peptides compared to traditional polyaromatic ligands, which are structurally rigid and have positional control of coordinating motifs, may evidently be a significant hindrance to forming an ordered assembly with directional supramolecular interactions. The inherent flexibility of peptide bonds makes the prediction of the geometry of supramolecular interactions and the secondary structure more difficult. Peptides are also often prone to perturbation of their secondary structure (e.g. α-helices) upon functionalisation relying on specific primary structures to retain helicity,[31] making the tuning and optimisation of these materials more complex. However, this naturally led us to consider the use of the polyproline helix as an ideal supramolecular building-block. Polyproline is ubiquitous throughout nature, found in proteins and peptides,[32,33] and is often referred to as a "molecular ruler" due to its high rigidity and predictable helix geometries.[34] As such, the polyproline II helix is treated as a rigid rod-like structure ideal for achieving angstrom precision as a spacer in molecular studies and supramolecular structures. This is clearly evidenced in reports where polyprolines are used as references for the determination of distances in biomacromolecules through Förster resonance energy *trans*fer (FRET) studies.[35] Furthermore, the polyproline II helix is extraordinarily tolerant to functionalisation, retaining their secondary structures with significant functionalisation and down to short chain lengths.[36,37] This unique stability of the polyproline helix highlights this class of peptides as a prime candidate as an organic ligand for the assembly of metal-organic constructs.

In 2014 the first crystal structure of the polyproline II helix was obtained by the Prof. Helma Wennemers, showing the exact dimensions and conformational properties of a *p*-bromobenzoyl capped oligoproline hexamer.[36] Furthermore, our other work detailed in the previous chapter, which has been recently published,[38,39] provides further detailed information on the geometry of the polyproline helix from crystallographic studies of a series of functionalised oligoproline tetramers. With these specific spatial parameters from the crystal structure, and the rigidity of the polyproline II helix, the geometry of functional groups is highly predictable facilitating the rational design and use of functionalised polyproline derivatives as both molecular rulers and scaffolds. The use of shorter polyproline helices in the design an assembly of metallo-peptide constructs has been demonstrated by Fujita; metal–peptide rings which assemble into highly entangled frameworks and the coordination-driven assembly of a short polyproline helix in to a small protein-like channel, showing guest uptake and chiral recognition.[6,28] These reports highlight the great potential of polyproline as a ligand in metal-organic assemblies with highly promising applications already evident from these few studies.

The Wennemers group has also gone on to demonstrate the formation of a metal-organic framework from an oligoproline hexamer in the polyproline II conformation, capped with an *N*-terminal benzoic acid moiety,[40] the longest peptide used to assemble a metal-peptide framework.[41] The framework formed, however, was highly unstable and degraded outside of the mother liquor lacking any useful applications which, being one of the most prominent examples of a polyproline helix utilised in an metal-organic framework, demonstrates the further work required to develop effective peptidic building-blocks for the design metallo-peptide constructs.

The polyproline helices have three functionalisable faces with a pseudo C_3 rotational symmetry axis which can be accurately functionalised by replacing specific residues within the peptide sequence with non-natural prolines. Thus, these can be used as chemical handles to induce the desired assembly of these peptide units (Figure 54). These peptide units can be envisioned as rod-like structures where functional groups have three potential directionalities, while the terminal residues can be functionalised *via* peptide capping at the *N*-terminus on the resin-bound peptide and *via* coupling at the *C*-terminus with the free peptide exemplified in Wennemers example of a polyproline based MOF with terminal carboxylic acids binding to the metal centres,[40] adding further directional handles for supramolecular assembly. Thus, polyprolines could be envisioned as multi-topic ligands with the level of functionalisation and placement of groups dictating the topology of the metal-organic assembly (Figure 54). This potential has been partially shown in the previous chapter whereby an oligoproline tetramer was functionalised with hydroxyprolines at each terminus, and thus decorated with two functional groups on the same helical face aligned in the same direction, essentially forming a C-shaped ditopic ligand with initial attempts to form a covalent-organic cage. There is clear potential for this material to form discrete cage-like structures from this design of peptide and as such this research aims to progress this work with alternate functional groups and varying chain lengths to demonstrate the possibility of utilising these peptides as scaffolds for metallo-organic assemblies.

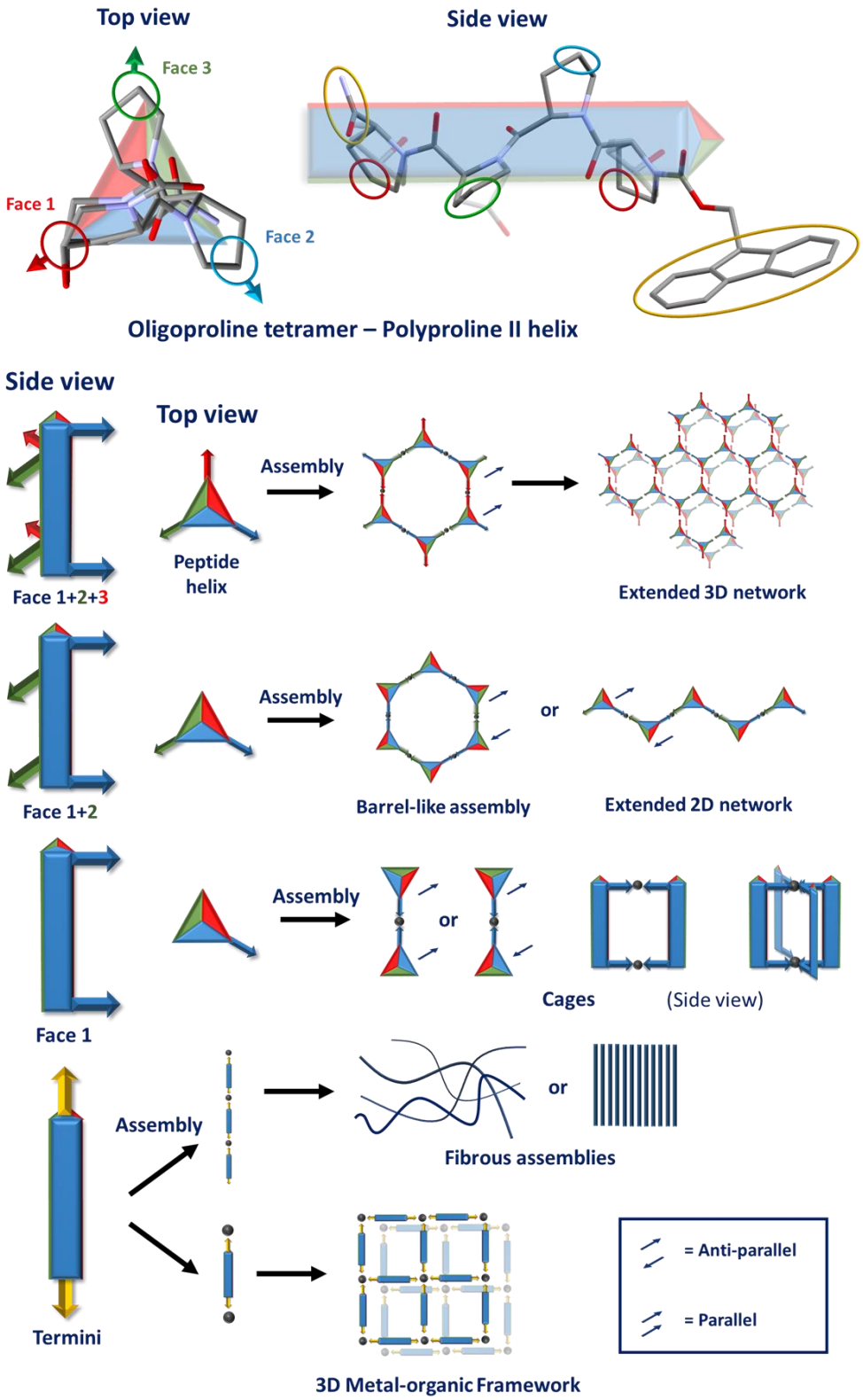


Figure 54 - Crystal structure of an oligoproline tetramer viewed along the helix after Fmoc removal with the face highlighted and a side view of the helix (top), illustration of potential supramolecular structures achievable by varying the functionalisation of a model polyproline helix (bottom)[39]

This functionalisation of the side chains of proline residues theoretically allows for the rational design of either 1-/2-/3-dimensional constructs depending on the number of faces functionalised; with the functionalisation of all three faces likely to obtain a 3-dimensional framework, and two faces capable of producing varied structures from extended frameworks to

cages and barrels, while as stated previously functionalisation of a single face is likely to form a discrete assembly such as molecular cages or nanobelts (Figure 54). The level of functionalisation of the faces is also an important consideration, especially in the multi-directional peptide ligands, with a single binding group between interacting units likely to produce a weaker, less directional interaction, while increasing the number of linkers should increase the strength of the interaction and broadly increase the likelihood of obtaining a more ordered assembly. The terminally functionalised polyproline helices can extend in 1-dimension to construct fibres forming assemblies such as hydro- or organogels, or may be used as simple linkers between metal nodes in metal-organic frameworks.[40]

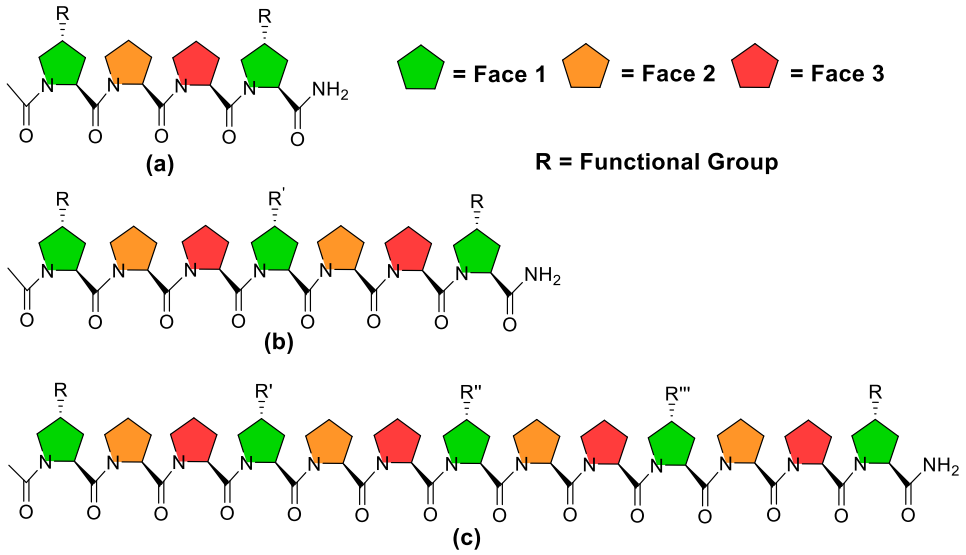


Figure 55 - Chemical structure of example polyproline helices showing the maximum number of functionalisable positions on a helical face for peptide lengths Pro4 (a), Pro7 (b), and Pro13 (c). Each proline is coloured to represent which helical face they reside on in the polyproline II helix.

Minimalistic peptide units are desirable as these smaller units are cheaper and easier to synthesise, can be simply functionalised, and have an improved atom economy, as such the use of the shortest optimal peptide units as scaffolds for metal-organic constructs is ideal. With these factors in mind there are several desirable peptide lengths to be initially investigated; very short peptide units (i.e. 4-7 monomers, Figure 55a-b) that are between 1 and 2 helical turns are liable to adopt only the polyproline II helix and be significantly simpler to rationally design and model potential assemblies, and longer peptide units (13+ monomers, Figure 55c) with more sites available for functionalisation, which are useful for tuning and optimisation of non-covalent and secondary interactions, and have the potential to easily switch between the polyproline I and II helices depending on the environment. The smaller peptides (< 6 residues) come with some limitations such as when used as a ditopic ligand, with two linker groups on the same face, these would only have two other residues for further functionalisation to produce interactions with any guest molecules and these are both on opposite faces, therefore chemical complexity and handles for manipulating host-guest interactions are somewhat limited. As such a ligand with at least two helical turns (Pro7, Figure 55b) may be desirable for further studies as this contains a 3rd residue on the same face as the terminal prolines such that functional groups can be placed directly into any nanocavity adjacent to the peptide helix, while still utilising a minimalistic peptide unit. Thus, there are varied design considerations when looking at peptide length and the placement of functionalised amino acids within the peptide sequence, depending on the desired properties of the peptide and the subsequent supramolecular constructs. Yet this especially highlights how tuneable these peptide ligands can be and functional groups can be easily incorporated through the stepwise solid-phase peptide synthesis (SPPS).

3.3 Results and Discussion

3.3.1 Synthesis of Functionalised Proline Derivatives

3.3.1.1 Synthesis of a -COOH functionalised proline

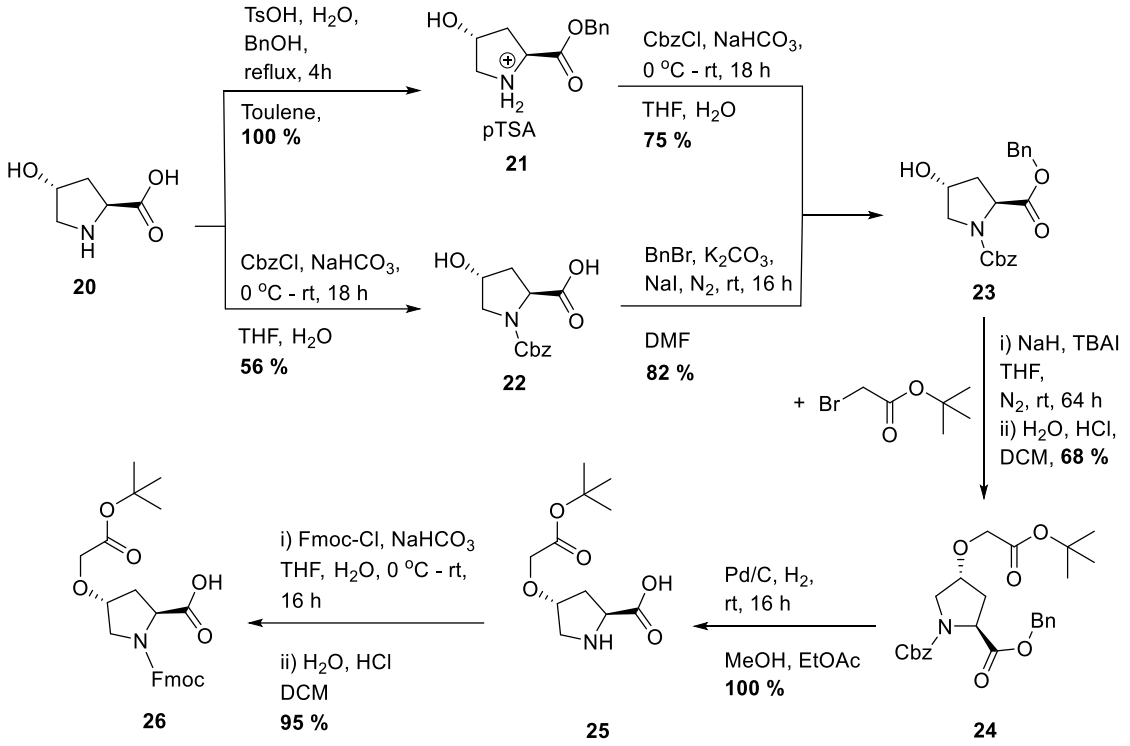


Figure 56 – Synthetic route carried out to synthesise compound 26, an Fmoc-proline with a protected carboxylic acid sidechain

The synthesis of a number of functionalised proline derivatives was attempted. The amino acid initially synthesised was functionalised with a carboxylic acid sidechain (-OCH_2COOH) at the C^4 position to yield an alkoxyproline (Figure 56). This was achieved *via* an O-alkylation reaction of a protected *trans*-L-hydroxyproline substrate with *tert*-butyl bromoacetate. Initially, the free hydroxyproline **20** was protected *via* benzylation **21** with toluenesulfonic acid and benzyl alcohol, and subsequent Cbz protection **23** to access multiple different proline derivatives from the partially protected pTSA proline salt **21**. However, future reactions and scale up reactions were synthesised *via* the Cbz protection **22** and benzylation route, both reaction routes caused minimal issues during synthesis. The functionalisation of hydroxyproline was carried out *via* the base-catalysed O-alkylation of the Cbz/Bn protected derivative **23** under dry conditions giving a good yield after purification (≈ 70%). The benzyl ester **23** was utilised to prevent alkylation at the carboxylic acid, forming an alkoxyproline ester, removing the requirement for a subsequent hydrolysis step, which together typically achieve low yields.[42] Subsequently, both Cbz and Bn protecting groups were removed simultaneously *via* hydrogenation **25** achieving quantitative yields and the product protected *via* the relatively facile Fmoc protection. Some optimisation of the Fmoc protection step **26** was required to achieve consistently high yields (50-70% → >90%). Control of the ratio of water and THF was essential, with 1:1 giving the highest yields and complete dissolution of Fmoc-Cl before addition to the amino acid solution. Also, control of the reaction pH *via* the addition of excess $NaHCO_3$ was necessary due to the production of HCl from the reaction of Fmoc-Cl. The t-butyl protecting group on the additional sidechain was retained at this point for protection of the carboxylic acid during peptide synthesis, which would

otherwise react with any free amines during the coupling reaction step. This Fmoc-protected amino acid was subsequently taken forward for peptide synthesis.

3.3.1.2 Synthesis of a rigid -COOH functionalised proline derivative

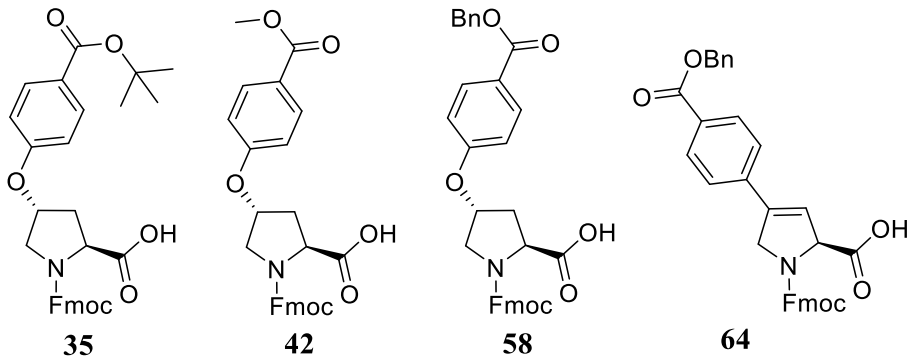


Figure 57 – Target 4-position functionalised Fmoc-proline derivatives

Subsequent, to the successful synthesis of the carboxylic acid functionalised alkoxyproline **26** a more rigid version of the carboxylic acid linker was targeted to achieve a more directional interacting moiety, in the hope that upon incorporation into a peptide sequence any subsequent assemblies are likely to be more ordered. The first aliphatic functional group is both relatively long and flexible making the prediction of the COOH geometry more difficult and likely to inhibit ordered assembly due to the flexibility, while the length also creates the potential for interactions to be less specific to a counter moiety during assembly as the group could form intramolecular bridging interactions to closely placed functionalised residues. A series of routes were attempted to achieve a more rigid functionalised derivative with varying degrees of success. The targeted functionalised prolines are shown in Figure 57.

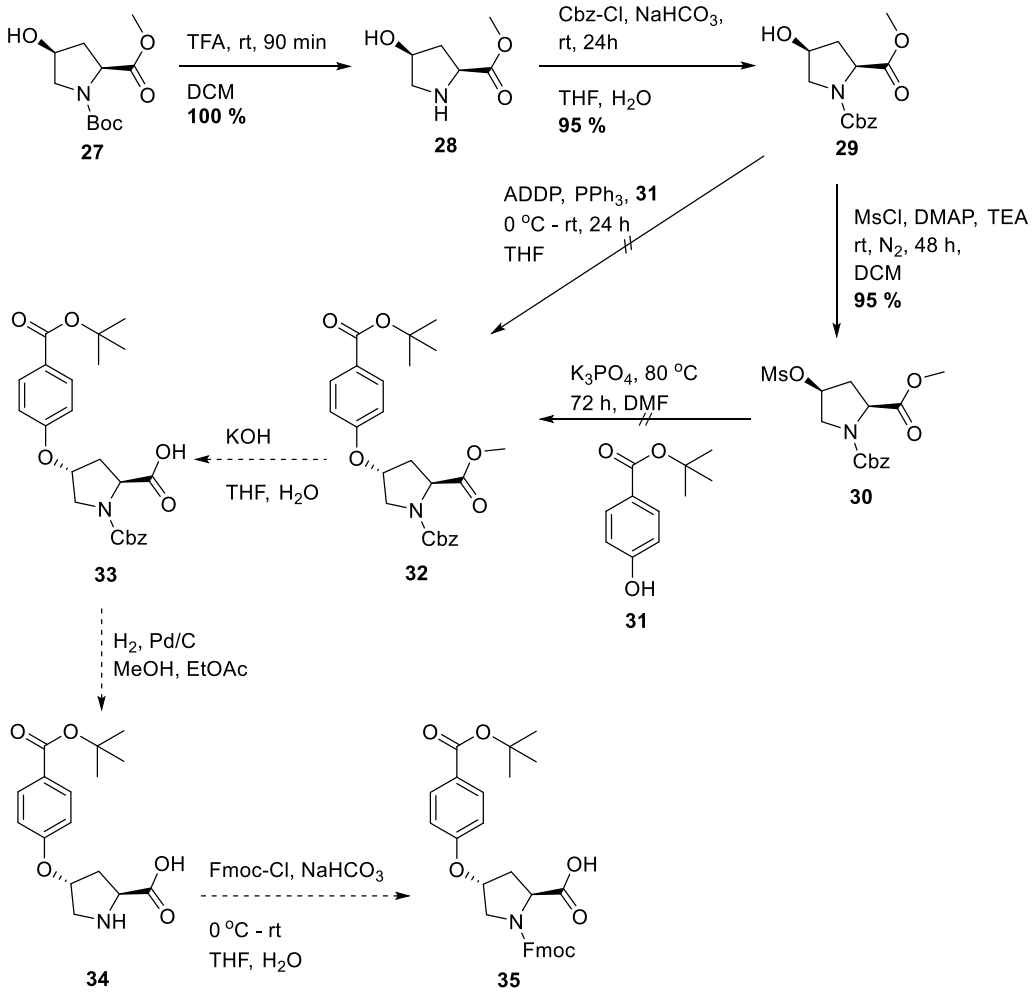


Figure 58 – Synthetic route to achieve compound 35, with an aromatic protected carboxylic acid sidechain, synthesis unsuccessful from compound 32.

The focus of the new syntheses was to achieve an aromatic carboxylate functionality on the proline with the aromatic group contributing to the rigidity of the COOH group and also increasing the directionality of any interactions due to the steric bulk and potential for π-interactions. The Mitsunobu reaction was initially used to achieve substitution at the C^4 hydroxyl group for compounds **32**, **40**, and **56** with phenolic nucleophiles to form phenol ethers. However, numerous issues were encountered during these syntheses. The Mitsunobu reaction progresses *via* a S_N2 reaction at C^4 inverting the stereochemistry at the position, this required the utilisation of *cis*-hydroxyproline derivatives prior to substitution to give the desired *trans*-functionalised derivatives to match the geometry of the previously synthesised proline, although given time the *cis*-derivatives are also a viable avenue of investigation. As such, a different starting substrate was utilised; Boc-*cis*-hydroxyproline methyl ester **27** due to access and cost compared to the fully de-protected derivative. Thus, deprotection of this substrate was required for the incorporation of new protecting groups, necessary to retain the t-butyl protecting group on the new desired aromatic carboxylic acid functionality, carried out simply *via* Boc cleavage with TFA, **28**. The amine was subsequently protected using Cbz-Cl to synthesise compound **29** which was suitable for substitution. This compound also crystallised from EtOAc and as such SC-XRD analysis was carried out to obtain the crystal structure (SI 4.8.2, CCDC-2290760). Initially, the Mitsunobu reaction to synthesise compound **32** was attempted using diisopropyl azodicarboxylate (DIAD) and triphenylphosphine (PPh_3) which, while resulting in complete consumption of the starting material, produced a very complex TLC with by-products from the DIAD reagent overlapping with the product, making purification difficult. As such, DIAD was

replaced with 1,1'-(azodicarbonyl) dipiperidine (ADDP), which is a solid at room temperature, as the hydrazine by-products are significantly more polar than for DIAD and remain at the baseline in the conditions used for TLC analysis and the azo derivative is more easily protonated by less acidic substrates due to a higher basicity, and as such is more successful with substrates with a lower pKa that are near the limit of DEAD/DIAD (≈ 11).[43] However, the attempted reaction produced multiple major products from TLC analysis suggesting significant side product formation. From this it is likely the reaction has proceeded *via* an elimination pathway instead of *via* S_N2, to give a dehydrated derivative of the substrate, which can be a common side reaction in Mitsunobu reactions.[44,45] Due to the lack of significant product formation several alternate reaction conditions were attempted. Initially the alcohol, nucleophile and PPh_3 were mixed before slow addition of the azodicarboxylate, the typical procedure for this reaction. Alternatively, the phosphine and azodicarboxylate were mixed for 30 min to pre-form the betaine intermediate before the addition of both the nucleophile and alcohol simultaneously, as well as attempting the reaction at higher temperatures (50 °C) and with sonication at high concentrations.[46] However, these attempts all produced similarly negligible product formation.

As such, the synthesis of the functionalised intermediate was attempted *via* activation of the hydroxyl with MsCl and subsequent substitution with the t-butyl 4-hydroxybenzoate nucleophile with K_3PO_4 at high temperature (80 °C). This small-scale test reaction showed the formation of the desired product for the mesylation reaction step, however the substitution reaction with the phenol showed several highly coloured product spots not expected from the reaction.

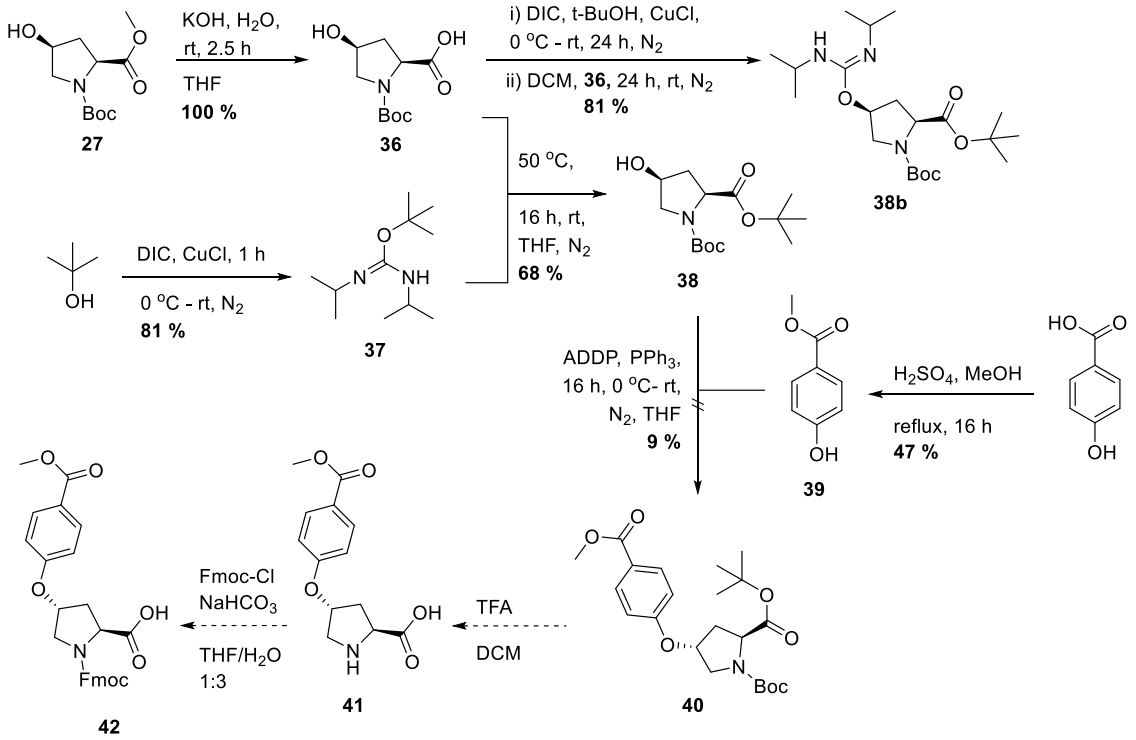


Figure 59 – Synthetic route to again incorporate an aromatic carboxylic acid sidechain, compound 42, unsuccessful Mitsunobu reaction to synthesise compound 40

Due to issues with the previously attempted monomer during the Mitsunobu reaction an alternate reaction pathway to synthesise a similar monomer, with the same functionality once inserted into a peptide sequence, **40** was targeted. This was carried out with the aim of avoiding the use of a Cbz protecting group on the proline monomer during the Mitsunobu reaction due to frequent issues encountered during this step for Cbz protected prolines. However, the use of Boc on the proline required different protecting groups on the phenolic acid. The first method

attempted used the methyl ester however it was uncertain whether this protecting group would survive the harsh conditions during peptide synthesis and furthermore, would require a hydrolysis reaction on any completed peptide.

With this new reaction pathway, a Boc protecting group was utilised on the amine, while *tert*-butyl was used to protect the carboxylic acid (**38**), allowing for full deprotection in one step *via* TFA treatment. This required deprotection of the initial methyl ester on the proline substrate and protection with the t-butyl group *via* treatment with synthesised *tert*-butyl *N,N'*-diisopropylcarbamimidate **37**. The initial one-pot method used to synthesise the *tert*-butyl protected proline did not use purification of excess reagents during the synthesis of *tert*-butyl *N,N'*-diisopropylcarbamimidate, as such this resulted in coupling of both the hydroxyl of t-BuOH and the proline hydroxyl to DIC, achieving the desired *tert*-butyl protection the carboxylic acid yet also coupling DIC to the hydroxy group, confirmed *via* ^{1}H NMR spectroscopy and LCMS analysis (**38b**). An attempt to react this product (*O*-Acylisourea) with the phenolic nucleophile to achieve the desired product for the next step was unsuccessful. As such, purification of the reactant **37** was required before reaction with the proline substrate **36**, this was achieved *via* carrying out the reaction with an excess of *tert*-BuOH versus DIC to remove the risk of residual DIC and filtering through a short alumina pad to remove excess CuCl, while also filtering off the crystallised urea by-product. The treatment of the free acid of proline **36** with this product subsequently yielded the desired t-butyl protected monomer, purified *via* FCC **40**.

The methyl ester of 4-hydroxybenzoic acid was synthesised using typical procedures and further dried in a desiccator in preparation for the Mitsunobu reaction. As such, with the Boc/t-butyl protected proline in hand the Mitsunobu reaction using ADDP was attempted with the typical procedure. However, despite complete consumption of the starting material, isolation of the predicted product produced negligible yields of the correct compound **40** (9%). Due to the potential for further problems with the methyl ester group during peptide synthesis and the less-than-ideal requirement for hydrolysis on the free peptide, combined with the difficulties of synthesis of the *tert*-butyl carboxylate and unsuccessful Mitsunobu this synthetic route was abandoned in favour of an alternate route, to synthesise compound **56**.

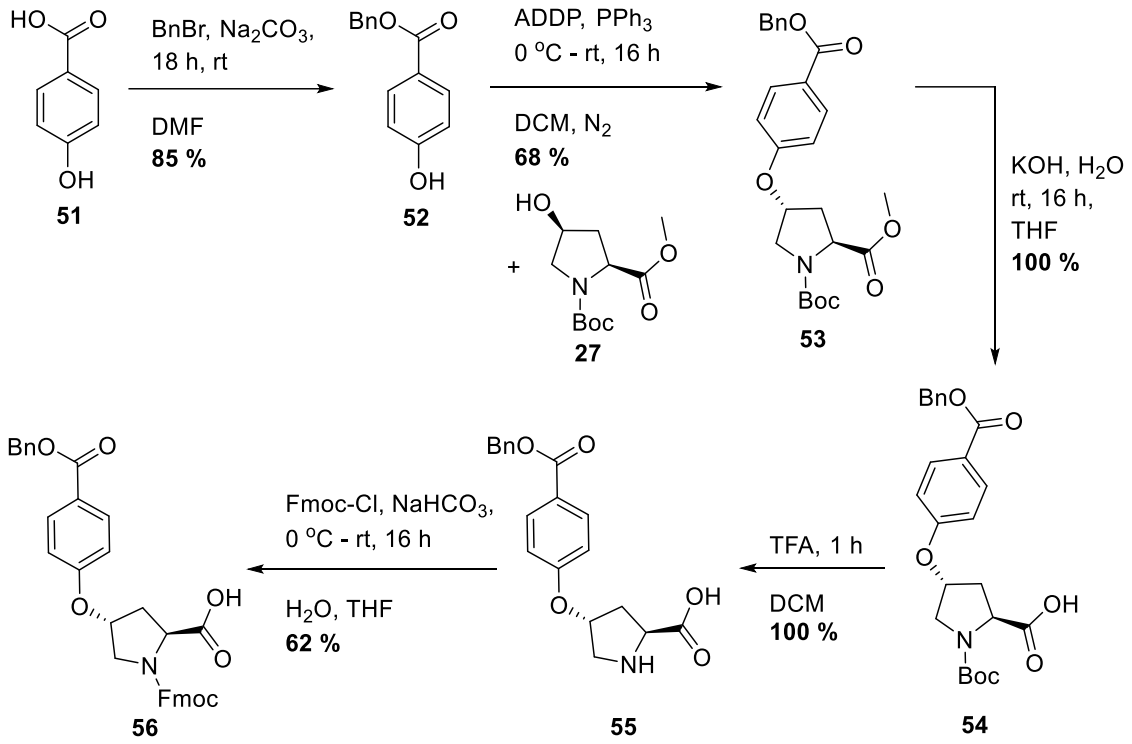


Figure 60 - Synthetic route carried out to achieve the proline derivative, 56, containing a trans -O-C_6H_4-COOH functionality at 4-position.

Compound **56** was targeted to make best use of the as-bought starting proline substrate *cis*-L-Boc-Hyp-OMe **27** such that no reactions on the proline are required prior to the Mitsunobu reaction, and significantly with the substrate being a more common example in literature with clear previous successes. In this case the phenolic acid was protected with BnBr using typical conditions and giving a high yield (85%). Subsequently, the Mitsunobu reaction was carried out using typical conditions with ADDP but in dry DCM rather than dry THF as attempted previously. Upon purification this yielded the desired product **53** in a good yield (68%). After this deprotection of the methyl ester *via* base catalysed hydrolysis **54** and Boc-deprotection with TFA **55** were carried out successfully in quantitative yields. Finally, Fmoc-protection using typical conditions was carried out to yield the desired final proline monomer **56**. However, partial deprotection of the benzyl ester occurred, requiring FCC to purify the final product. This product was taken forward for use in peptide synthesis, however, due to the incorporation of a benzyl ester protecting group any peptides synthesised will require a subsequent hydrogenation reaction to remove the benzyl protecting group.

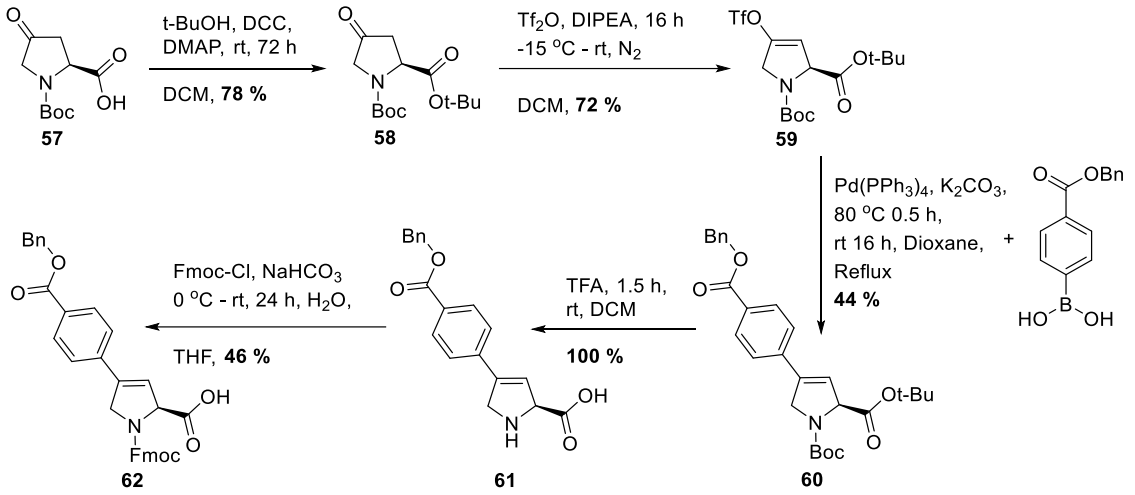


Figure 61 - Synthetic route to achieve the proline derivative, 62, with direct C-C coupling of a motif, incorporating an aromatic carboxylate sidechain

Concurrent to the synthesis of the previous proline derivative an alternate synthesis of a -RCOO-Bn functionalised proline, **62,** was attempted *via* a completely different route, using C-C bond formation from the Suzuki reaction to produce a direct Bn-Proline C-C linkage,[47] further shortening the distance from the proline ring to the -COOBn functional group, creating a highly rigid group due to the conjugated system from the alkene formed. As such, this monomer has clear potential for forming highly directional and ordered interactions compared to the significantly more conformationally flexible –OCH_2COOH group previously utilised. The starting substrate in this synthetic route was the readily available *N*-Boc-4-L-oxo-proline **57**. As this substrate has a ketone instead of the hydroxyl group, the carboxylic acid can be easily esterified with t-BuOH and DCC, catalysed by DMAP, affording the desired product in a high yield after purification (78%). To allow for the Suzuki cross-coupling reaction the ketone first had to be converted to a vinyl triflate to form a good leaving group. This was carried out *via* reaction of the carbonyl with trifluoromethanesulfonic anhydride (Tf_2O) to yield the triflate, an -OTf moiety, with a double bond between the C^{3-4} positions on the pyrrolidine ring. This was achieved in good yields (72%) after purification.

After activation of the carbonyl, the Suzuki reaction was carried out with the *para*-boronic acid of benzylbenzoate in the presence of a palladium catalyst to yield the desired product **60** in a reasonable yield (44%). This product formed a crystalline solid on which single-crystal X-ray diffraction (SCXRD) analysis was carried out. This data showed the adoption of a planar pyrrolidine ring, unlike the typical puckered pyrrolidine ring, now in-plane with the aromatic group due to the double bond forming part of the conjugated system. This gives a clear model for the geometry of the sidechain functionality compared to the proline ring, useful for the rational design of peptides for supramolecular assembly, however, it is yet unclear what impact the alternate conformation of the pyrrolidine ring and new double bond will have on the dihedral angles and helical structure of a

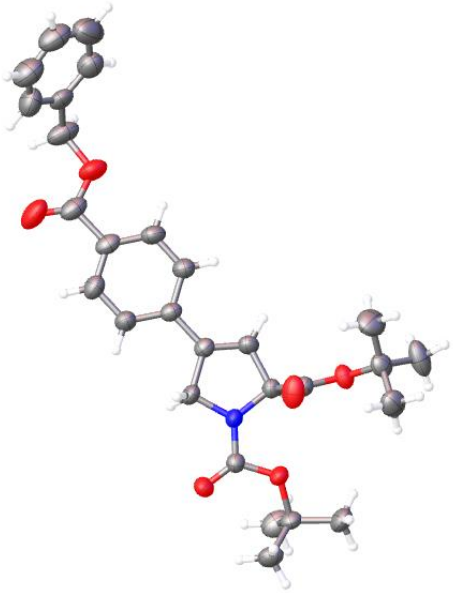


Figure 62 – Single crystal structure of compound 60, asymmetric unit, 50% ellipsoids

peptide containing this monomer. To achieve the desired Fmoc protected monomer, simple cleavage with TFA was carried out to deprotect the substrate before using the typical Fmoc protection procedure. As such this monomer was taken forward for peptide synthesis, however, upon incorporation into a peptide sequence this monomer will require hydrolysis of the benzyl ester to form the free acid sidechain. In this case the olefin can undergo catalytic dehydrogenation which should exclusively form the *cis*-enantiomer according to literature,[48] while simultaneously deprotecting the benzyl ester, a viable route if issues are encountered with saponification of the benzyl group.

3.3.1.3 Synthesis of a pyridine functionalised proline derivative

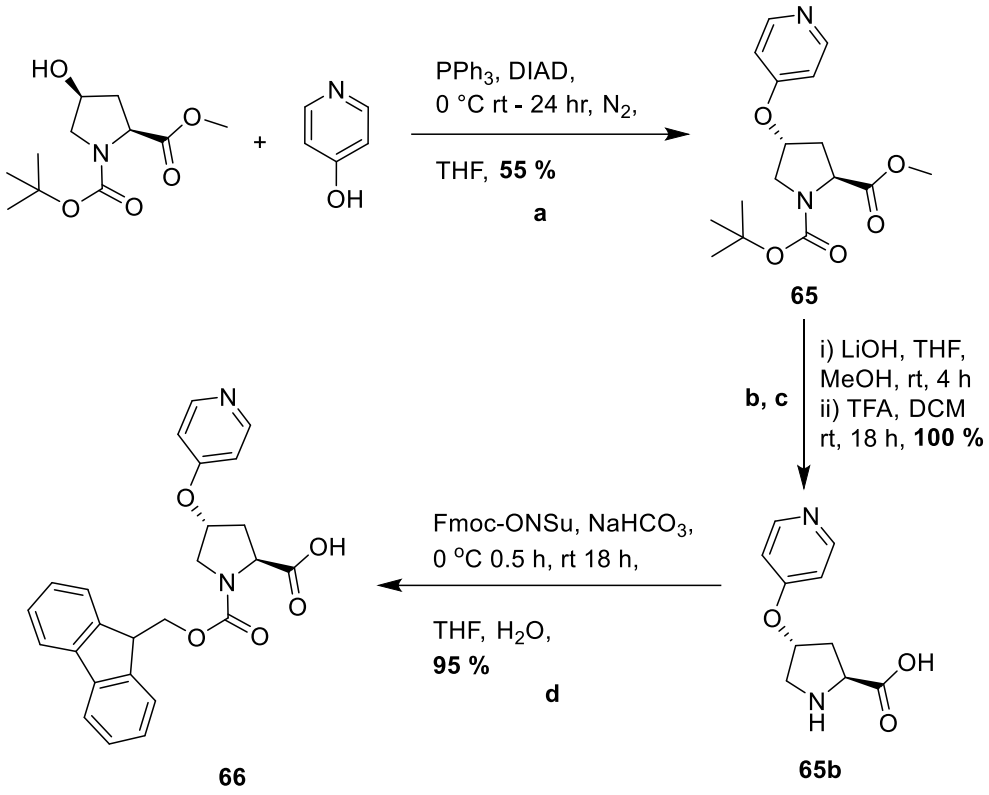


Figure 63 - Reagents and conditions: (a) PPh_3, DIAD, dry THF, addition at 0 °C, then 24 h. at RT, N_2, (b) LiOH (aq.), THF: MeOH (1:1), at RT, 2.5 h., (c) TFA, DCM, then 24 h. at RT, (d) $NaHCO_3$, THF: H_2O (1:1), addition at 0 °C, then RT, o/n. Reactions a-c carried out by Dr K.Samanta.

Moving away from the carboxylic acid functional group the pyridine functionality was instead targeted as it is well known to act as a ligand forming metal complex,[49] while it can also act as complementary group for assembly with carboxylic acids.[50]As such a pyridine functional group incorporated into a proline monomer has promising applications for supramolecular assembly when placed into a peptide sequence. To synthesise a pyridine functionalised proline monomer the Mitsunobu reaction was carried out with pyridine-4-ol and a protected *cis*-hydroxyproline to form the desired *trans*-O-functionalised proline derivative (**65**) achieved in reasonable yields (55%) after purification *via* FCC, using triethylamine as an additive to prevent binding of the basic monomer to the acidic silica. Subsequent, deprotections and protection were carried out to achieve the final Fmoc substrate required for SPPS.

3.3.2 3-Dimensional Metal-Peptide Frameworks

There exists a keen interest in the synthesis of nanoparticles as vehicles for drug delivery with the goal to incorporate bioinspired building-blocks such as peptides in the construction of these supramolecular materials. These materials have promising applications for the synthesis of biocompatible tuneable nanoparticles capable of the encapsulation of compounds for the targeted delivery and controlled release of drug molecules. As such we aimed to incorporate the polyproline helix as a structural unit in the synthesis of metal-organic nanoparticles. Initial work carried out previously within the research group evidenced highly promising results with the functionalisation of a hexameric oligoproline with three carboxylic acid functional groups allowing the synthesis of a zinc-peptide complex that assembled in water to form nanoparticles, evidenced by TEM microscopy (Figure 64).[51]

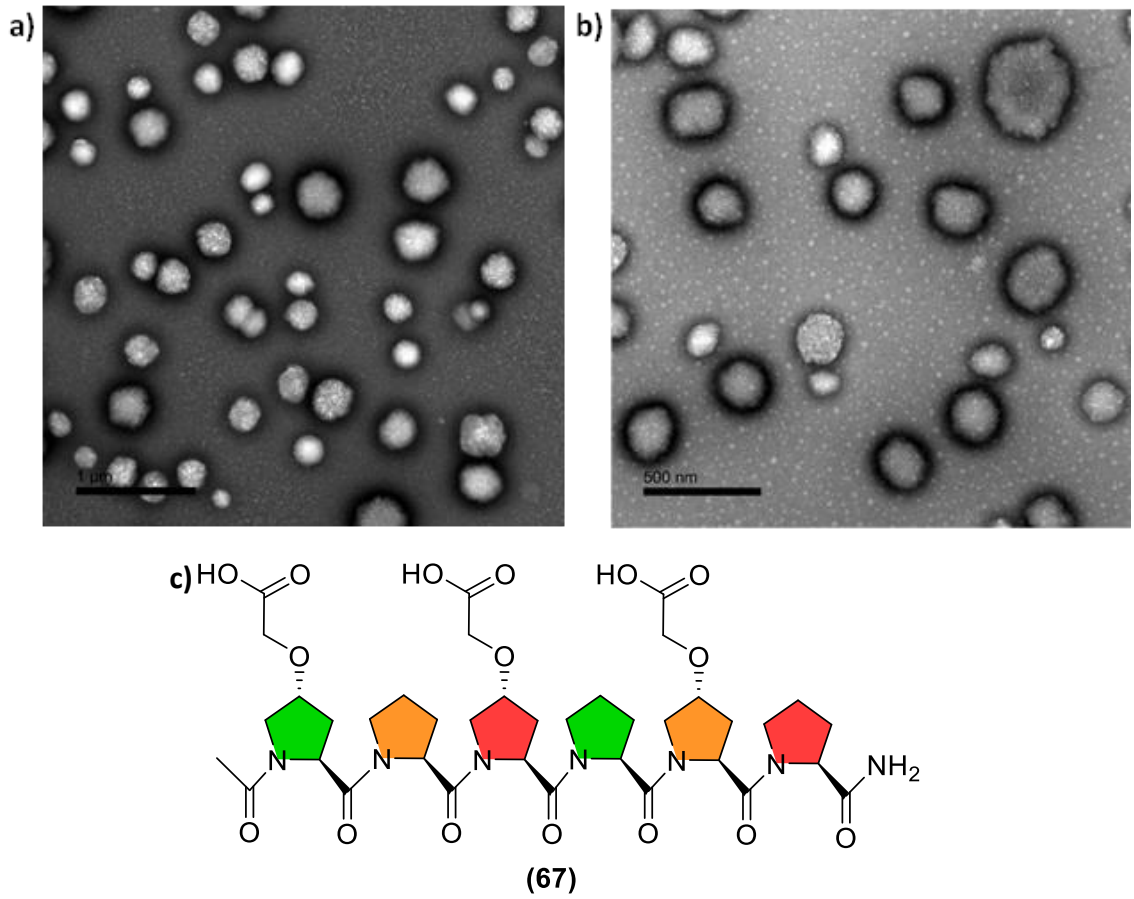


*Figure 64 – TEM images (a-b) of previously synthesised nanoparticles from peptide **67** (500 mM) and zinc nitrate (1.5 eq.) in water; (synthesised and imaged by Dr A. Palma). c) Peptide **67** used for supramolecular assembly (synthesised by J. Marsh)*[51]

3.3.2.1 Peptide Synthesis

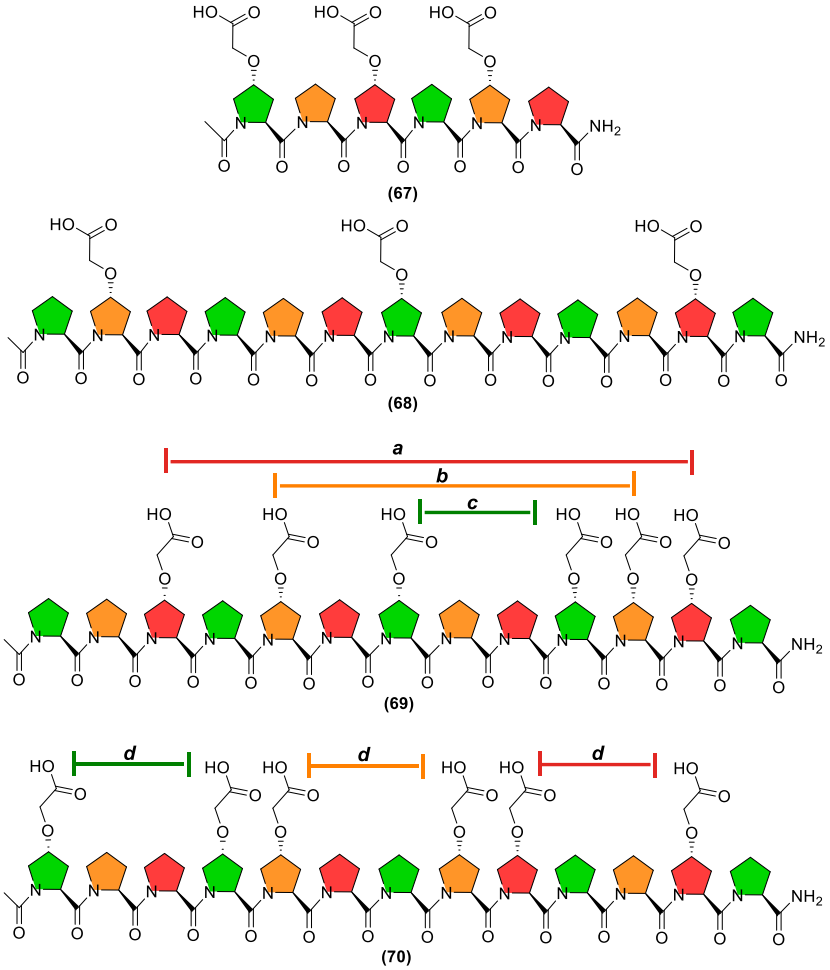


Figure 65- Synthesised polyproline peptides functionalised with -OCH_2COOH motifs on all three helical faces. Residues on the same helical face are highlighted in the same colour. distances; a = +9; b = +6; c, d = +3

Due to the previous successes with preliminary evidence of nanoparticle formation from peptide **67**-zinc complexes this peptide was again synthesised using SPPS techniques *via* a CEM liberty light peptide synthesiser. However, this short peptide hexamer lacks the capability for easy analysis of the conformational switching between the all-*cis* polyproline I and all-*trans* polyproline II helices *via* CD spectroscopy as the typical shift in maxima and minima does not occur for such a short polyproline helix.[52] This was confirmed from CD spectroscopy of the peptide in EtOH and water solutions with no diagnostic shift in the maxima or minima of the EtOH sample (Figure 66), suggesting that the polyproline II helix is retained. With this in mind an extended 13-chain version of the peptide was synthesised with the same single carboxylic acid functional group on each helical face of the peptide, with the aim of having a switchable polyproline building-block. CD spectroscopy of this peptide in EtOH and water confirmed the ability to switch between the two conformations, with a clearer shift in maxima from 227 nm in water to 220 nm in EtOH (Figure 67). Furthermore, base (ammonium hydroxide) to be used in complexation of the peptide was also added to the CD samples to observe the effect, if any, on the conformation of the polyproline helix. Notably, this increased the propensity of the peptide to switch to the polyproline I helix with a shift to the characteristic maxima of 216 nm and a trough at 235 nm, typical for the polyproline I helix. These peptides were then both taken forward for supramolecular assembly *via* complexation with metal centres.

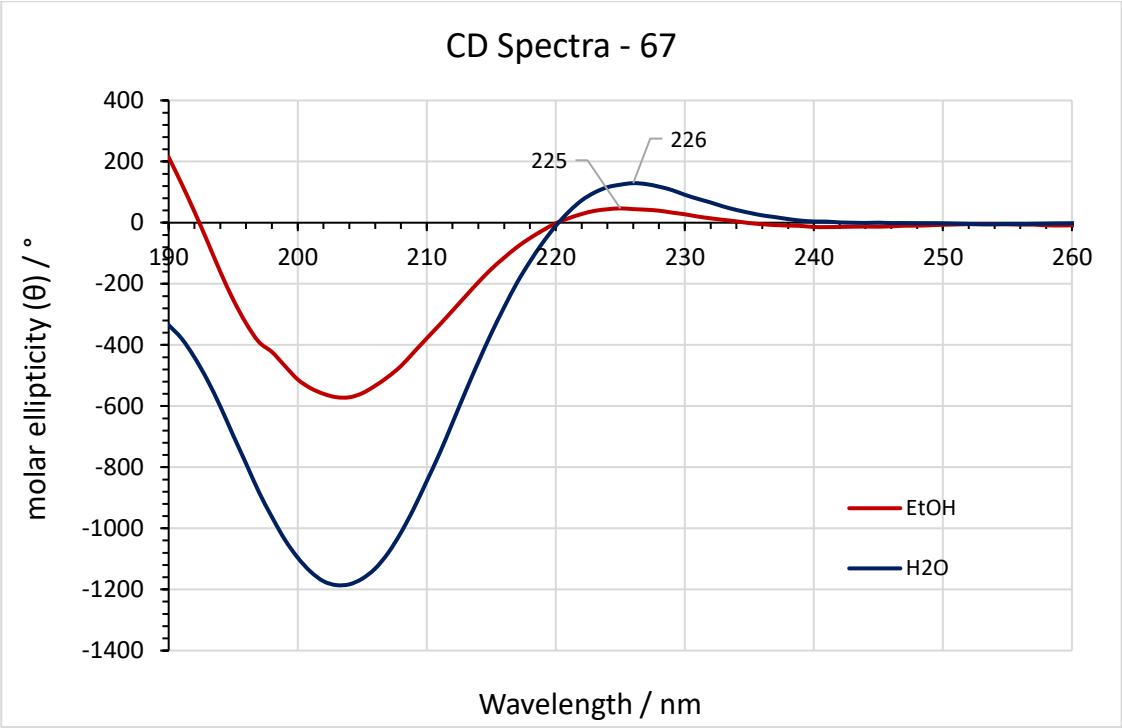


*Figure 66 – CD spectra for peptide **67** incubated in EtOH and water (250 µM), incubated for 14 days*

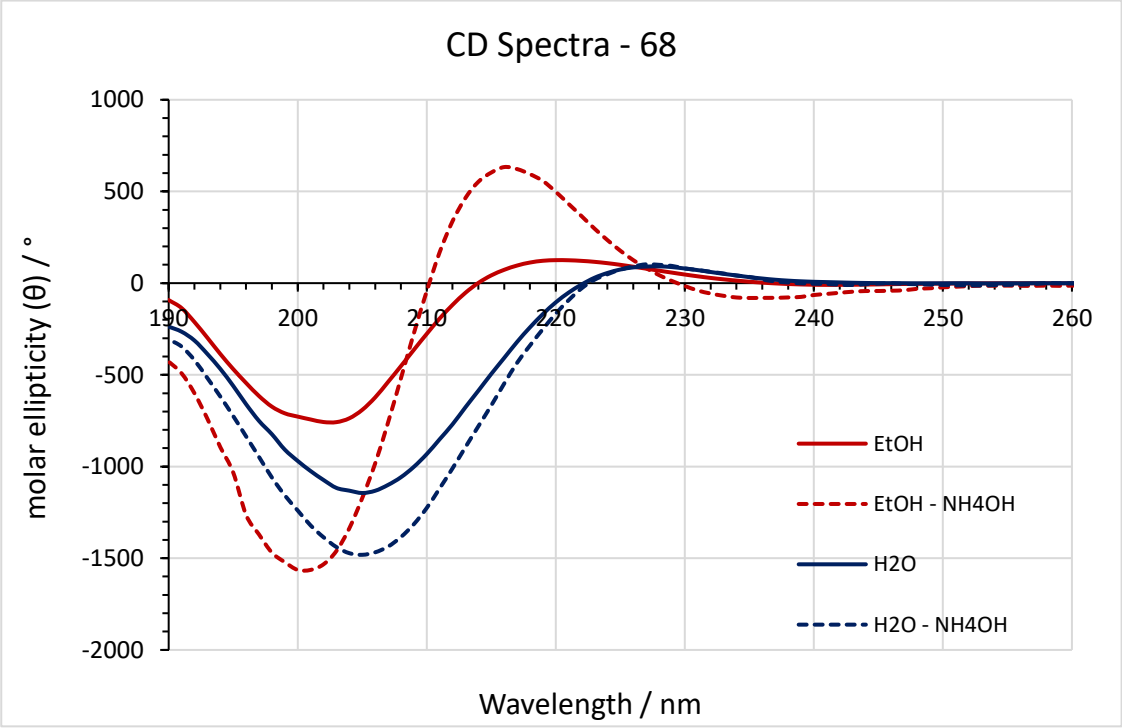


*Figure 67 – CD spectra for peptide **68** incubated in EtOH and water (250 µM), and with ammonium hydroxide (30%, 1 µL) , incubated for 14 days*

With a single binding motif on a long, relatively flexible, ligand it was theorised that the formation of an ordered assembly would be unlikely with a lack of selectivity between the faces of the helix and highly flexible interactions. As such a new peptide, **69**, was designed with the goal of achieving a stronger more directional binding between each of the helical faces. This was achieved *via* placing two functional groups on each face to increase the strength of the binding and reduce the flexibility of the complexed peptide. To introduce a degree of selectivity of this binding to prevent each face binding to any other face, as is possible for peptides **67** and **68**, the

functional groups were placed at different distances on each face (Figure 65) such that binding between non-complementary faces will not satisfy both linkers simultaneously due to the distance between placed functional motifs. Therefore, the most stable assembly should be *via* complementary binding between the same faces to achieve an extended framework, theoretically forming a more ordered assembly. CD spectroscopy analysis of peptide **69** confirmed that the peptide switches to polyproline I helix, with a clear shift in maxima to 216 nm in PrOH (Figure 68).

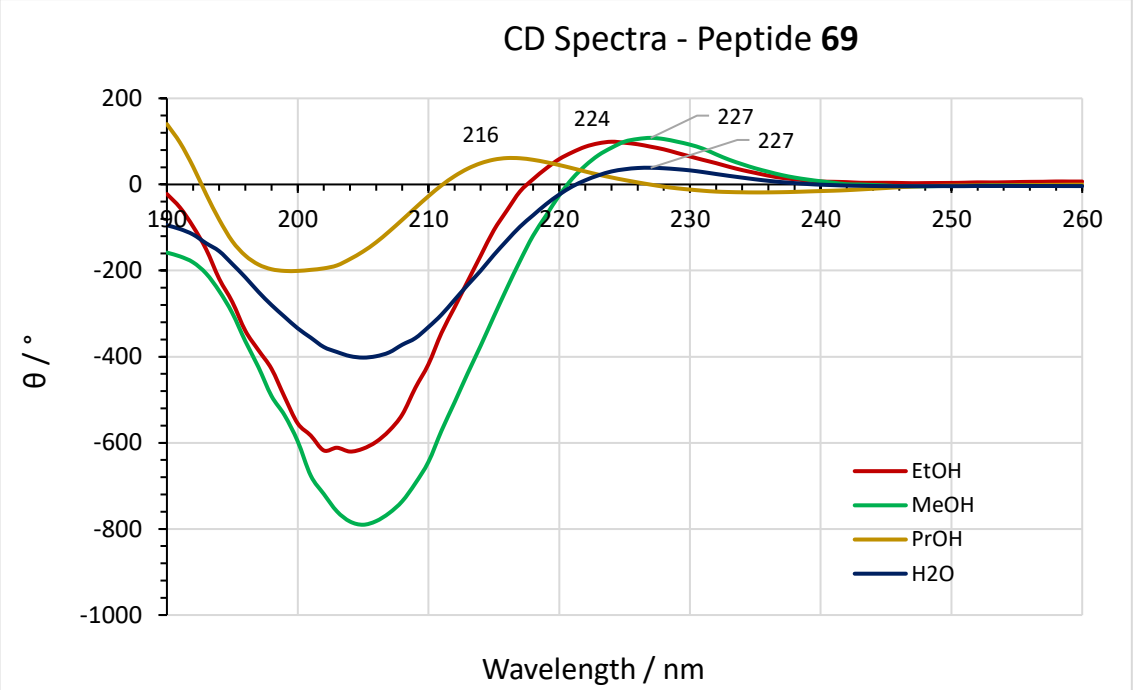


*Figure 68 – CD spectra for peptide **69** incubated in water, MeOH, EtOH, and PrOH (125 µM), incubated for 14 days*

To support the idea that the placement of the functional groups within the peptide sequence can be used to introduce selectivity of the peptide assembly a further peptide was synthesised, **70** (Figure 65), retaining the two functional groups on each face yet placing each of these the same distance apart so the peptides should be able to bind with any other face and thus be non-selective. This should therefore form a less ordered assembly and potentially disrupt the supramolecular assembly processes. However, CD spectroscopy analysis of peptide **70** showed that the peptide does not show a clear switch to the polyproline I helix, with only a small shift in maxima in PrOH and EtOH (223 nm, Figure 69). This suggests that the placement of the functional groups on the peptide backbone affects the ability for the amide bonds to switch between *cis* and *trans*, this may be due to the close proximity of the carboxylic acid in this peptide with each group one helical turn apart on the same face (Pro1-4, 5-8, and 9-12) and having the groups on adjacent prolines but on different helical faces (i.e. Pro4-5 & 8-9)(Figure 65).

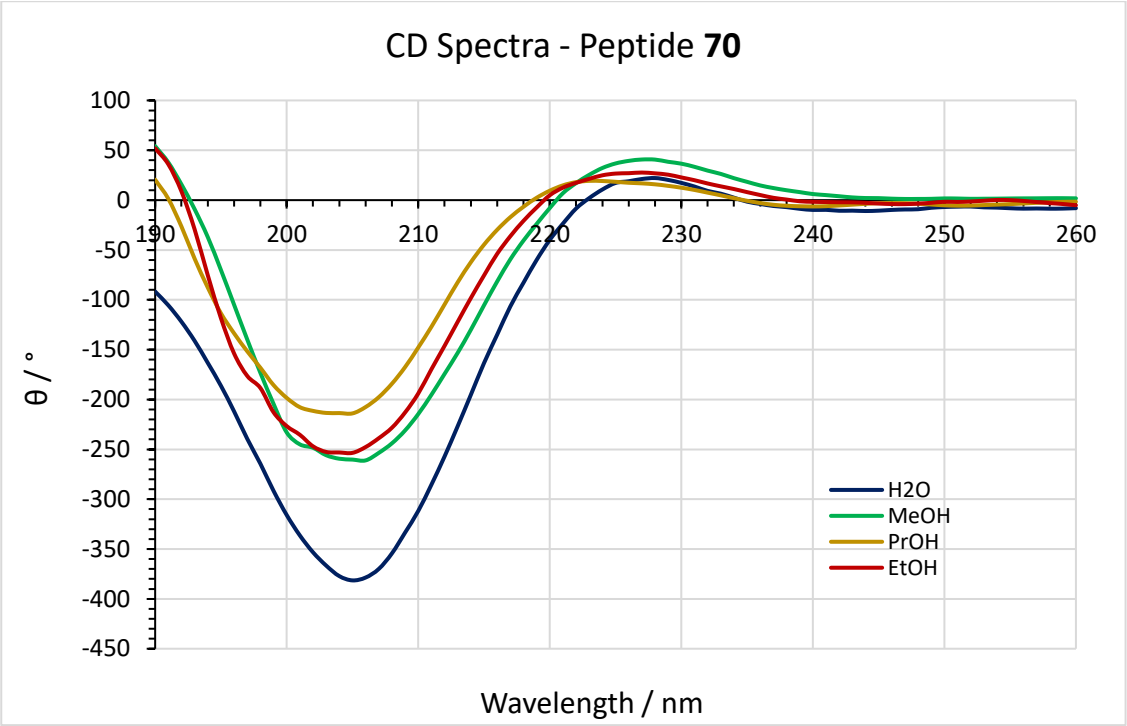


*Figure 69 – CD spectra for peptide **70** incubated in water, MeOH, EtOH, and PrOH (125 µM), incubated for 14 days*

3.3.2.2 Peptide-metal complexation – -$(OCH_2COOH)_3$

Initially, complexation of peptide **67** was attempted to reproduce the previously synthesised peptide **67**-zinc nanoparticles. These were synthesised *via* treatment of the peptide with ammonium hydroxide in water and drying to remove excess base, leaving the ammonium carboxylate of the peptide, which was then redissolved in water and combined with zinc nitrate producing a clear solution. This solution was then diluted with EtOH to produce a slightly turbid solution. Dynamic light scattering (DLS) analysis is a common method for the analysis of spherical particles in the solution state,[53,54] allowing the hydrodynamic diameter and the polydispersity of the assemblies to be analysed. As such, DLS was utilised to analyse the peptide-zinc complex in solution. These results clearly showed the formation of low polydispersity nanoparticles approximately 360 nm in diameter (Figure 70).

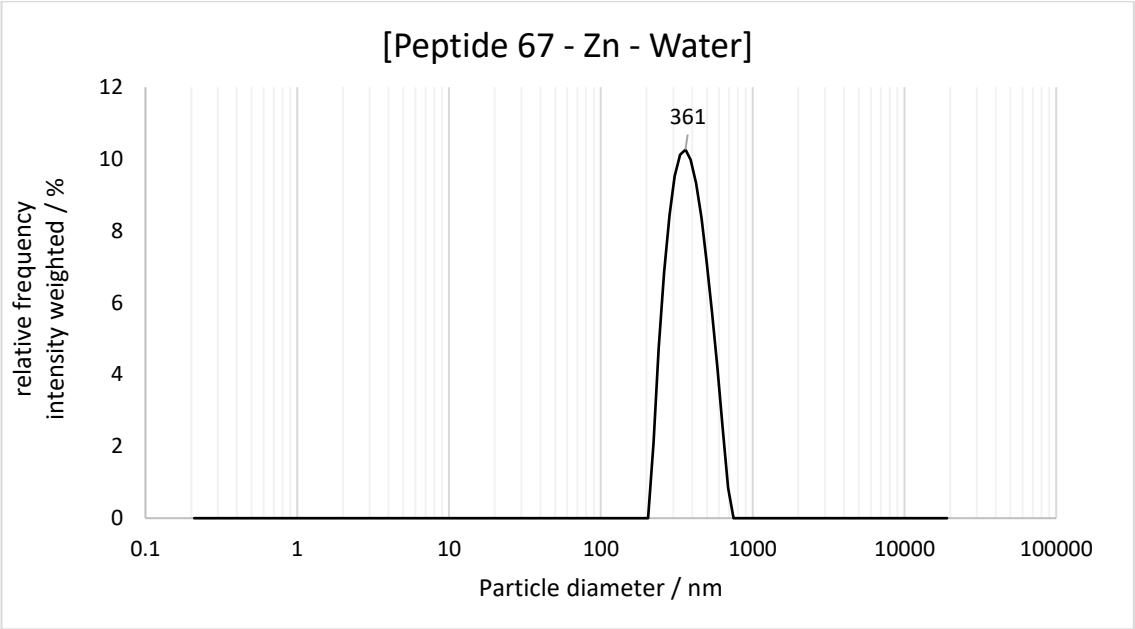


*Figure 70 – DLS intensity distribution for peptide **67** (0.5 mg/mL) dried from ammonium hydroxide and redissolved in water with zinc nitrate (1.5 eq.).*

Subsequently, *trans*mission electron microscopy (TEM) was utilised to allow for imaging of the deposited nanoparticles from solution. Samples of the complex were diluted with water to 0.5 mg/mL and drop-casted onto the TEM grids. This was then negatively stained with uranyl acetate solution to allow for imaging of the organic material, which is electron-dense to contrast the specimen on the grid. Analysis of the TEM micrographs evidenced the presence of nanoparticles however these were incorporated into a mixture of structures. Some of these nanoparticles had clearly aggregated into chain-like aggregates rather than discrete particles (Figure 71a-b), however particles are also visible dispersed through a film likely constituted of free peptide (Figure 71c-d), which prevents the negative stain from affecting some of the particles forming a dark film with light particles within and dark particles on top. Analysis of these spherical particles appears to suggest the particle may be hollow with the appearance of a collapsed sphere (Figure 71c), which is highly promising for the use of these materials allowing encapsulation of other molecules for methods such as nanotransport. This data suggests that while the peptide is forming nanoparticles these freely aggregate and are thus more polydisperse than suggested by DLS analysis. However, solution state versus solid state analysis can vary significantly as the deposited sample has been exposed to different conditions as it is the dried material. Without the use of base in this reaction the peptide instead forms 2D thin films with zinc nitrate.

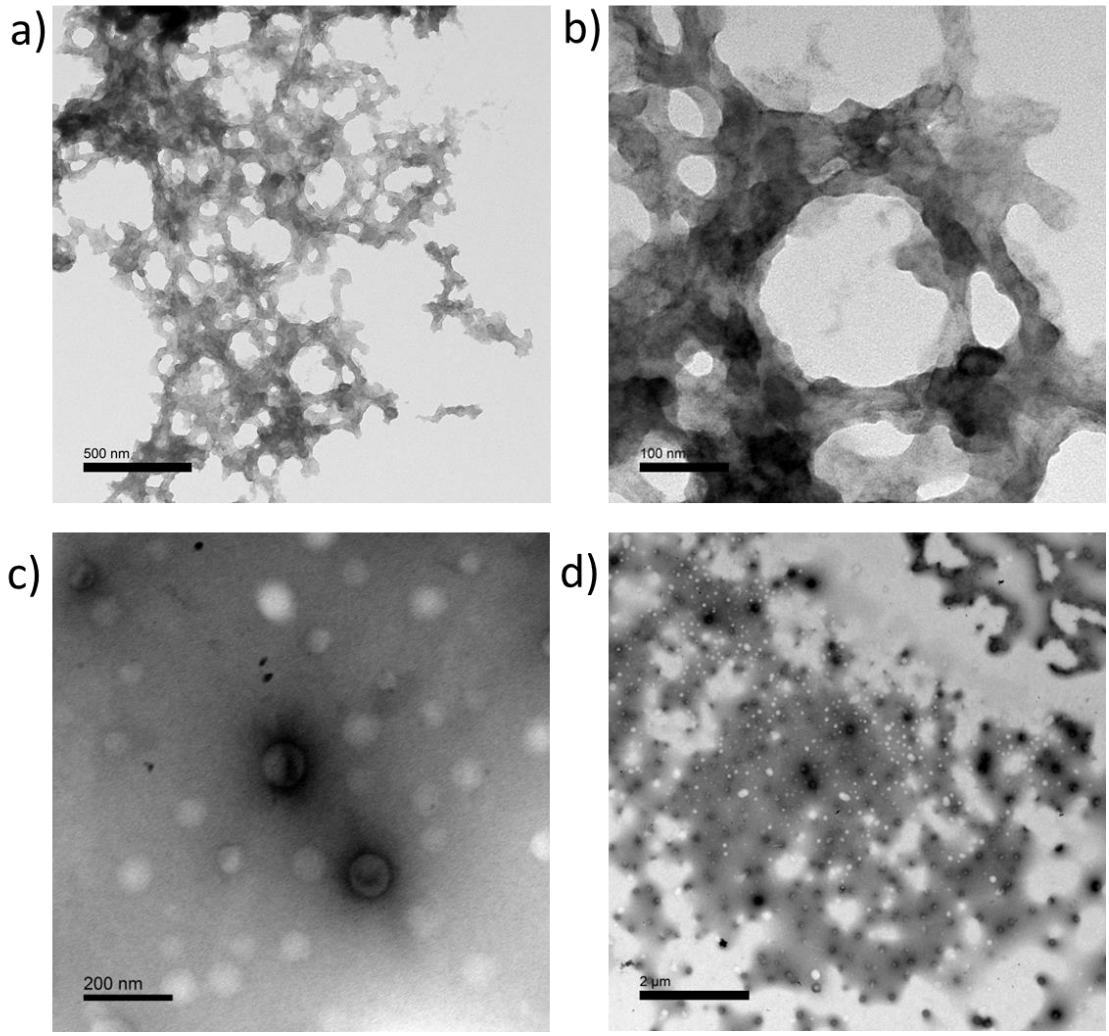


*Figure 71– TEM images of peptide **67** (1 mg) in water (100 µL) reacted with zinc nitrate (2 eq.) and diluted with EtOH (100 µL). Ammonium hydroxide (30%, 1 µL) added. Aliquot (10 µL) taken and diluted to 100 µL (H_2O) for drop-casting (0.5 mg/mL). Stained with uranyl acetate.*

Alternatively, a different method of centrifuging the peptide-zinc suspension from a mixture of EtOH/Water (1:1), decanting, and redissolving fully in water also produced nanoparticles from analysis *via* TEM. These particles were again mostly contained within a film and some aggregation of particles is clearly visible (Figure 72). However, these results are clearly indicative of the successful formation of nanoparticles from the complexation of these functionalised peptides.

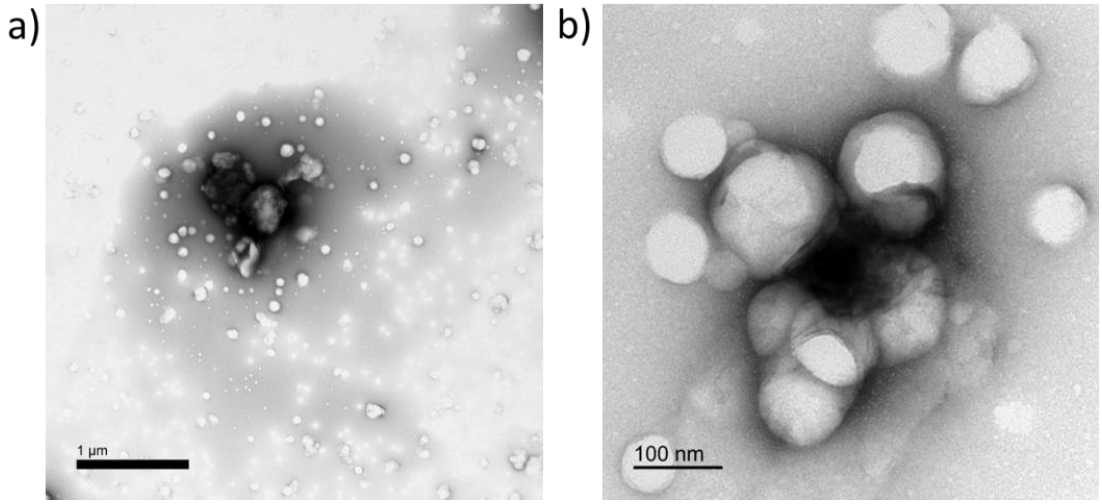


*Figure 72 – TEM images of peptide **67** (1 mg) in water (100 µL) reacted with zinc nitrate (2 eq.) and diluted with EtOH (100 µL). Sample then centrifuged and dissolved in water (200 µL). Aliquot (10 µL) taken and diluted to 100 µL (H_2O) for drop-casting (0.5 mg/mL). Stained with uranyl acetate.*

To analyse the effect of changing the solvent during the synthesis, as the polyproline helix is known to switch conformation in different solvents such as long chain alcohols,[52,55] the synthesis of the nanoparticles was carried out in an EtOH solution after incubating the peptide for longer than 10 days to ensure any conformational changes are complete due to the slow *trans->cis* switching. This resulted in a significantly larger particle size from DLS analysis (≈ 880 nm) with a worse PDI (Figure 73), suggesting solvent has a significant effect on the assembly of the particles. Subsequent, TEM analysis of the complexed peptide showed the presence of well-defined nanoparticles (Figure 73a-d) similar to those seen in previous studies in water (Figure 64) and in the water/EtOH mixtures (Figure 71). However, these particles were significantly smaller (< 200 nm) than expected from DLS analysis, showing a clear mismatch between the two methods of analysis.

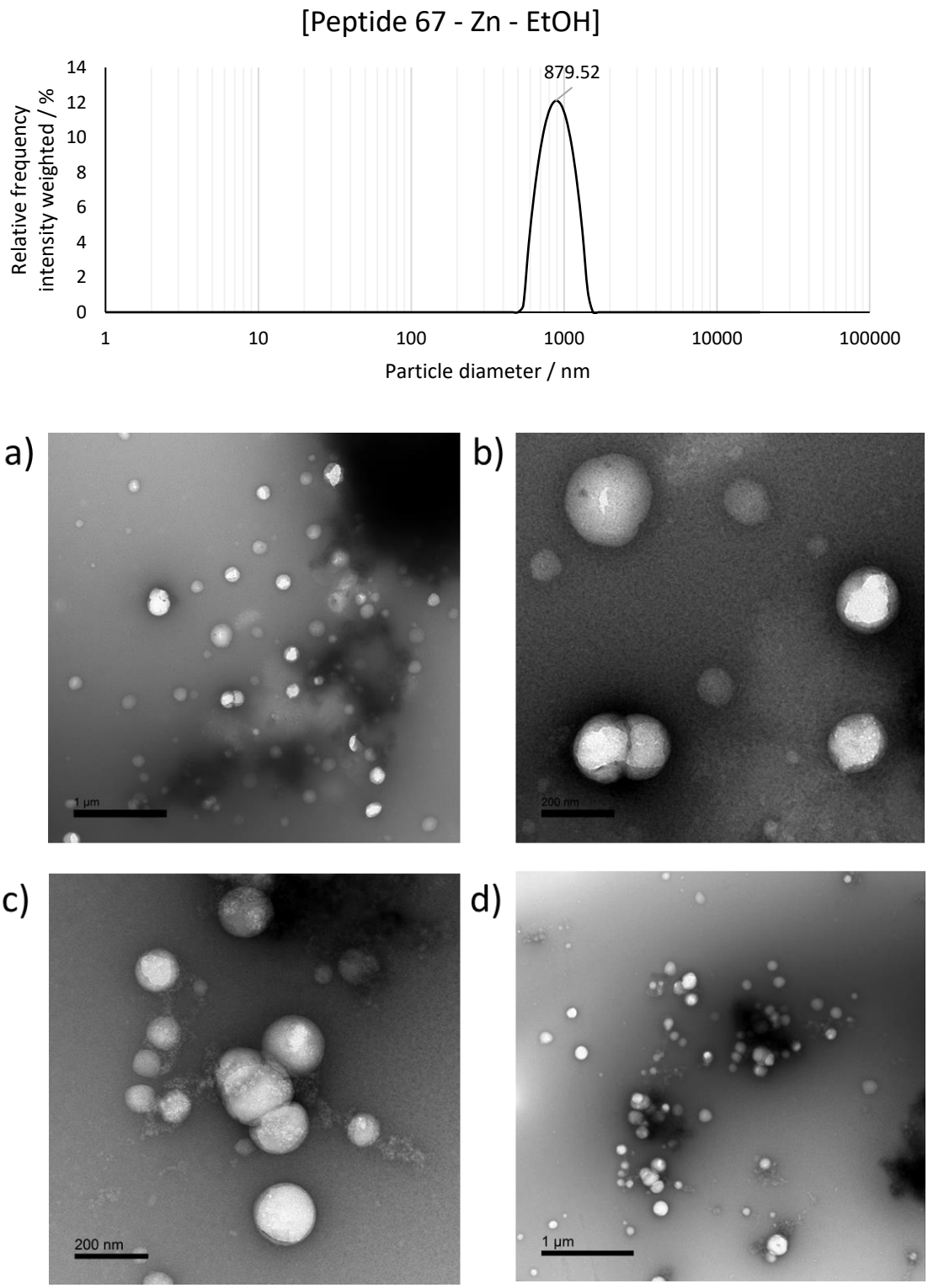


*Figure 73 - DLS intensity distribution for peptide **67** (1 mg/mL) dried from ammonium hydroxide and redissolved in EtOH with zinc nitrate (1.5 eq.). a-d) TEM images of peptide **67** (1 mg/ml) incubated for 10 days in EtOH (100 µL) and reacted with zinc nitrate (2 eq.) and diluted with EtOH (100 µL). Ammonium hydroxide (30%, 1 µL) added. Aliquot (10 µL) taken and diluted to 100 µL (EtOH) for drop-casting (0.5 mg/mL). Stained with uranyl acetate.*

Combination of the peptide with zinc without the presence of base instead produced chain-like aggregates that may have likely formed from the aggregation of spherical nanoparticles upon drying (Figure 74), however, the presence of ammonium hydroxide appears to be required to form well-defined discrete nanoparticles, with the presence of base and protonation of the peptide impacting the propensity to aggregate.

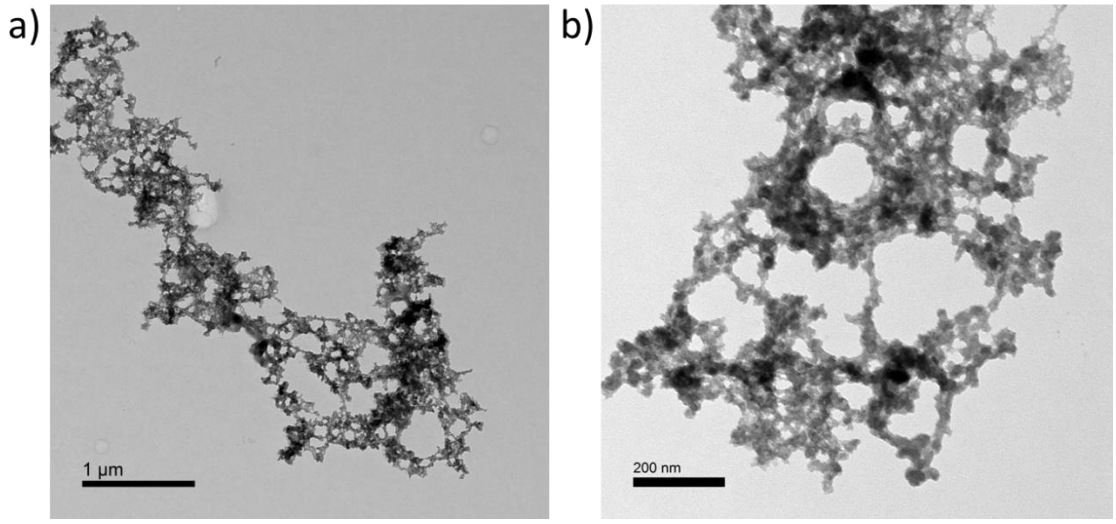


*Figure 74 –TEM images of peptide **67** (1 mg) incubated for 10 d in EtOH (100 µL) and reacted with zinc nitrate (2 eq.) and diluted with EtOH (100 µL). Aliquot (10 µL) taken and diluted to 100 µL (EtOH) for drop-casting (0.5 mg/mL). Stained with uranyl acetate.*

This data was also supported by scanning electron microscopy (SEM) showing similar clusters of aggregated nanoparticles (Figure 75). These were clearly visible without requiring a metal sputter coating. These bear a remarkable similarity to SEM images of self-assembled cages from coiled-coil peptides from work by Woolfson, where they claim to synthesise nanoparticles from peptide hubs, and describe these images as aggregated nanoparticles formed upon deposition on the stub and drying.[56]

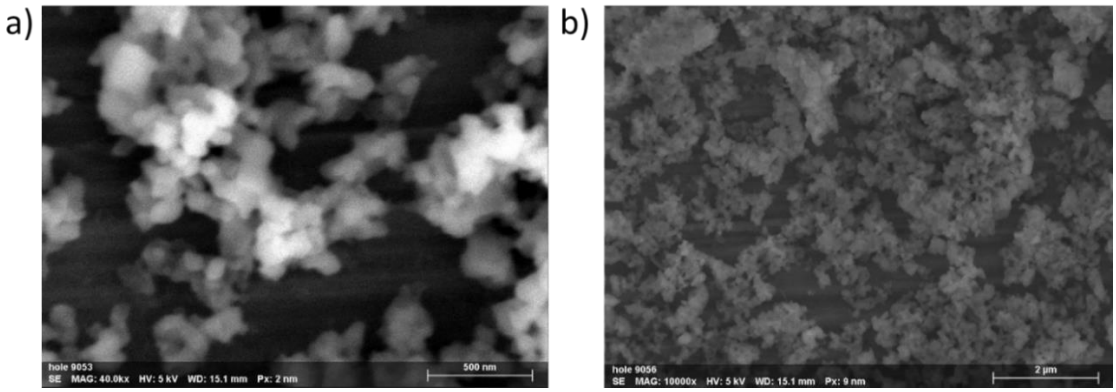


*Figure 75 - - SEM images of peptide **67** (1 mg) incubated for 10 days in EtOH (100 µL) and reacted with zinc nitrate (2 eq.) and diluted with EtOH (100 µL). Aliquot (10 µL) taken and diluted to 100 µL (EtOH) for drop-casting (0.5 mg/mL). Uncoated sample.*

With this promising data in hand, whereby nanoparticles could clearly be synthesised from a functionalised oligoproline hexamer as previously evidenced from preliminary studies, it was theorised this could be replicated in the longer 13-monomer chain, peptide **68**, giving access to different structures dependent on the solvent used due to isomerism of the helix. With the hexamer complex similar nanoparticles are achieved in both ethanol and water. Thus, a switchable peptide may give arise to a stimuli responsive supramolecular structure where the switching of the peptide results in decomposition of the assembly, which could be applied to release of encapsulated molecules or simply to access multiple structures from a single starting material with varying applications.

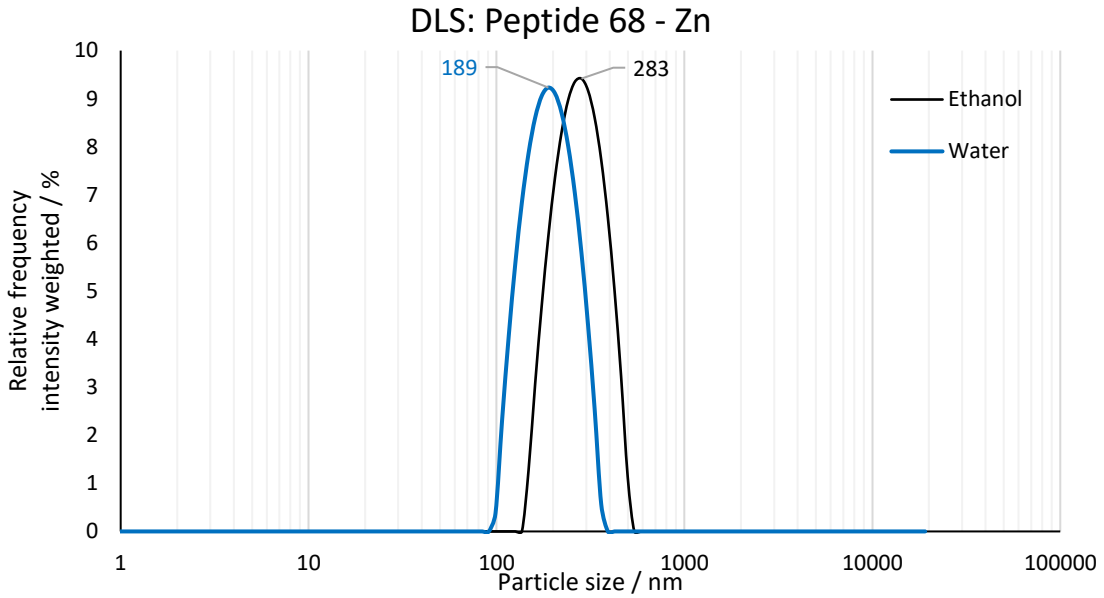


*Figure 76 - DLS intensity distribution for sample of peptide **68** (1 mg/mL) dried from ammonium hydroxide and dissolved in H_2O (1 mL) with zinc nitrate (1.5 eq.) and a sample of peptide **68** (1 mg/mL) dried from ammonium hydroxide and dissolved in EtOH (0.25 mL) with zinc nitrate (1.5 eq.).*

As such the synthesis of a peptide **68**-zinc complex was also attempted using similar methods to the previous example. Initial DLS studies indicated the formation of nanoparticles with a good PDI in both water and ethanol, however with slightly different size, 190 and 280 nm respectively (Figure 76). This supports that the longer peptide assembles similarly to the hexamer, forming nanoparticles under similar conditions.

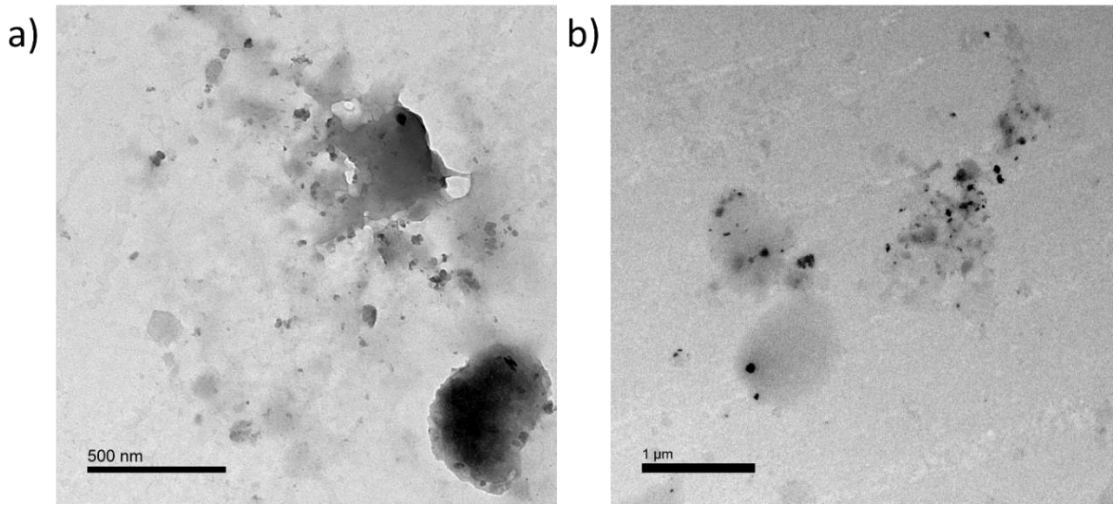


*Figure 77 – TEM images of a sample of peptide **68** (1 mg/mL) dried from ammonium hydroxide and dissolved in H_2O (1 mL) with zinc nitrate (1.5 eq.), stained with uranyl acetate (a) and unstained (b)*

TEM analysis of the peptide complex sample in water did not show the formation of any ordered assemblies (Figure 77) contradicting the results from DLS analysis. However, this shows a significant difference to the smaller peptide **67**, which clearly formed nanoparticles from TEM analysis (Figure 71). This suggests increasing the size of the peptide, or spacings of the linker groups, can significantly impact the topology of the metal-peptide assembly.

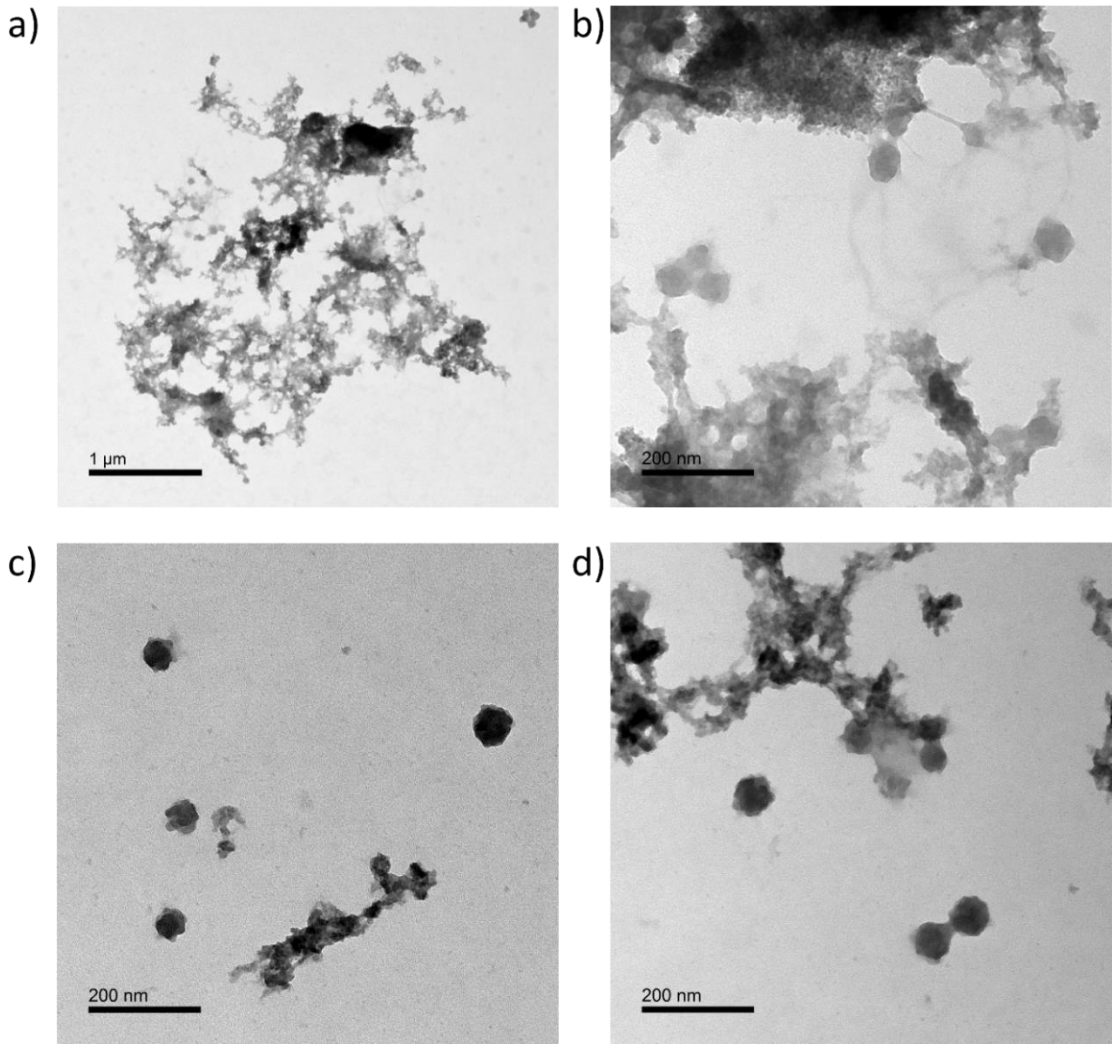


*Figure 78 – TEM images of a sample of peptide **68** (1 mg/mL) dried from ammonium hydroxide and dissolved in EtOH (1 mL) with zinc nitrate (1.5 eq.) stained with uranyl acetate (a-b) and unstained (c-d).*

Further TEM analysis of **68**-Zinc complex in EtOH clearly evidenced the formation of an ordered assembly, contrary to that seen in water, with a combination of aggregated particles forming chain-like structures (Figure 78a-b) and cluster-like nanoparticles (Figure 78c-d, 50-100 nm). The size of these nanoparticles is smaller than suggested by DLS (≈ 190 vs 50-100 nm), which can be accounted for by differences in the analysis methods and that DLS measures the hydrodynamic diameter which is expected to be larger than the actual particle size. The particles seen for this peptide are also smaller than those seen for peptide **67** and rather than containing particles dispersed in a gel-like assembly, these are freely dispersed particles and clearly aggregate upon drying to a degree (Figure 78). Also, by analysing the sample without a uranyl acetate negative stain we can see from the dark colouring that the particles are electron dense and thus contain a significant amount of zinc and are not solely peptidic (i.e. low electron density organic material) (Figure 78c-d), the less defined assemblies and increased amount of aggregation also suggests that the uranyl stain negatively impacts sample quality. These promising results show that the longer 13-chain peptide **68** can also form nanoparticles akin to the shorter peptide **67**, while the differences between the EtOH and water samples are indicative that the conformation of the helix can significantly alter the assembly of the complex as no ordered assemblies are seen in water for this peptide with polyproline II helicity. While the shorter peptide which does not switch conformation adopts similar topologies in both solvents.

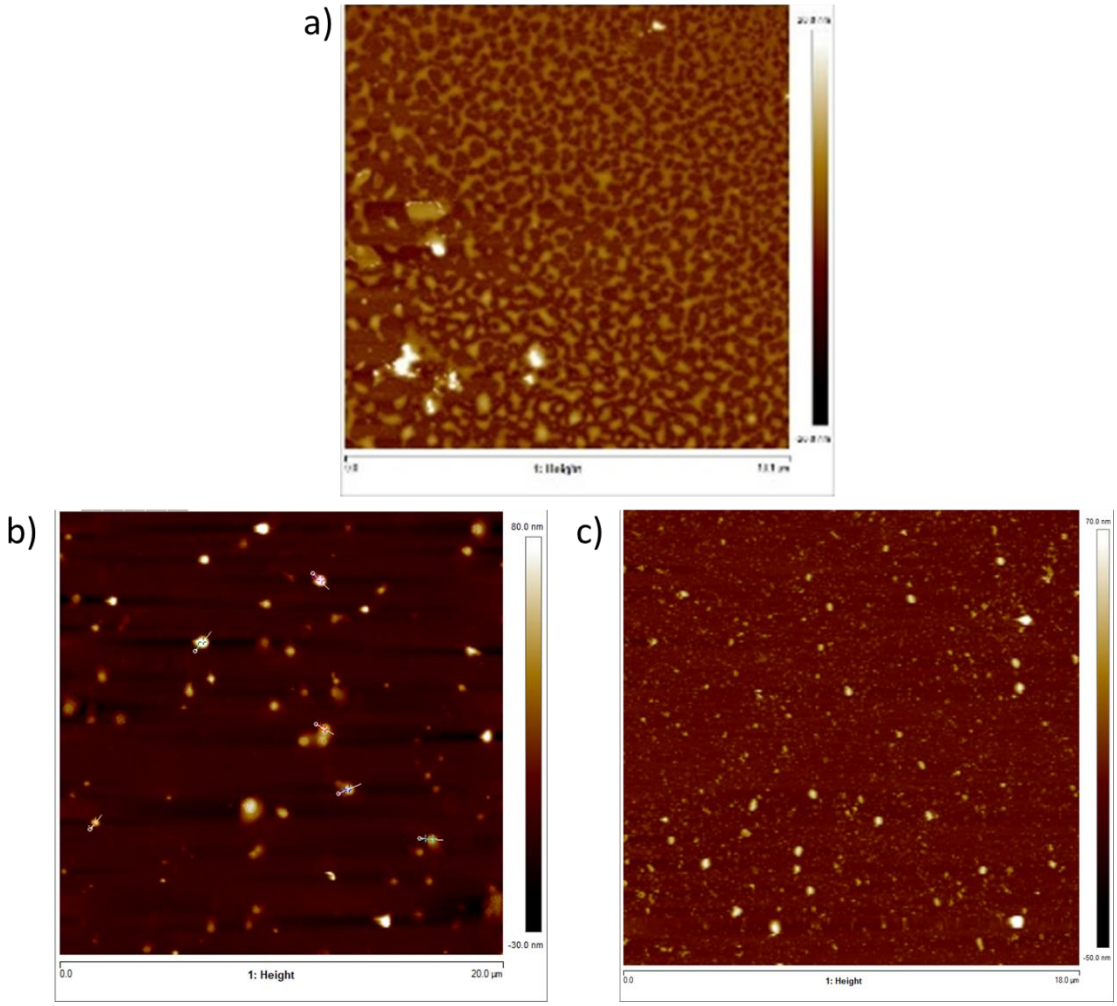


Figure 79 – a) AFM image of peptide 68 in water (1 mg/mL) with KOH (3 eq.) and zinc nitrate (1.5 eq.); b) AFM image of peptide 68 in EtOH (0.4 mg/mL) with KOH (3 eq.) 20 x 20 µm, c) AFM images of peptide 68 in EtOH (0.4 mg/mL) with KOH (3 eq.) and zinc nitrate (1.5 eq.); 18 x 18 µm,

The formation of the **68**-zinc complex was also repeated with an alternate base, utilising potassium hydroxide in place of ammonium hydroxide, so that the amount of base could be more carefully controlled. With this method equimolar amounts of base to carboxylic acid were used. AFM analysis was thus used to analyse these samples to give a clear indication of whether nanoparticles were present or not. For all samples the mica was first coated with nickel as samples were otherwise not retained, likely repelled by the negatively charged mica.

AFM analysis clearly showed that in water no nanoparticles were formed with base alone or base and zinc, instead giving 2D thin films in both cases (Figure 79a). However, samples synthesised from ethanol suggested the formation of nanoparticles with KOH alone (Figure 79b) and with both KOH and zinc (Figure 79c), with particle sizes averaging at 700 nm (height 50-100 nm) and 400 nm (height 20-60 nm) respectively. This indicates a significant difference between the assembly of the peptide in EtOH versus water, with the adoption of the polyproline I helix over the polyproline II helix a potential driving force for the formation of an alternate assembly as well as the solvent effects such as lipophilicity, volatility, and surface tension.

3.3.2.3 Peptide-metal complexation – -(OCH_2COOH)$_6$

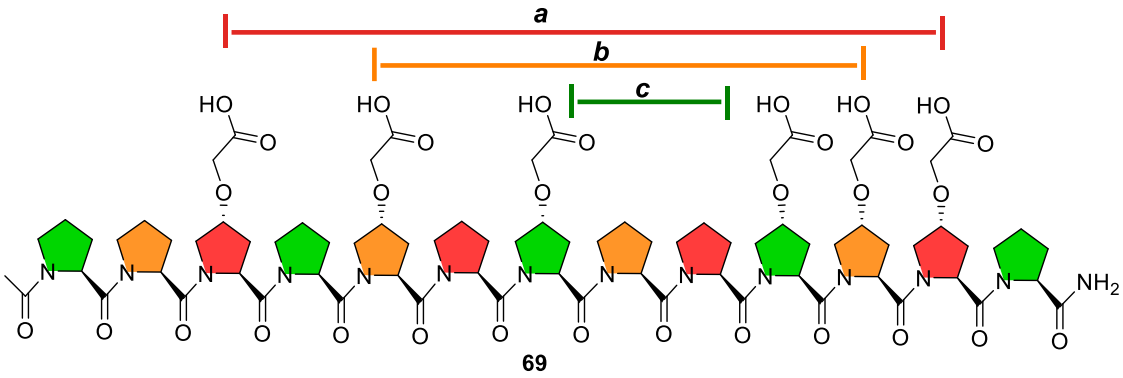


*Figure 80 – Chemical structure of peptide **69**, distances between the functional groups on each face are highlighted (a-c)*

Comparing the assemblies formed with either peptide **67** or **68** suggests the ability for the polyproline helix to switch conformations clearly impacts the assembly processes. However, the lack of formation of nanoparticles in water for peptide **68**, which is in the polyproline II conformation, the same conformation as would be adopted for the hexamer (**67**) in water and ethanol, suggests the effect of lengthening the helix also disrupts the assembly previously seen for the hexamer and nanoparticles only again form in ethanol when the more condensed polyproline I helix is adopted. These results are a positive indicator for the potential of polyproline helices a supramolecular building blocks to design stimuli responsive supramolecular constructs, yet further details of the methods of assembly and the effect of helix conformation are required to draw further conclusions from these results. As such, it was decided that the focus should be on using longer chain helices that can clearly undergo conformational switching, while refining the peptide building-block to achieve a more predictable and controlled assembly. Thus, the synthesised peptide **69** (Figure 80) was taken for complexation studies, with this peptide theoretically forming a more ordered and directional assembly due to the placement of two functional groups on each face at varying distances (*i*+3/6/9, Figure 80) as previously described.

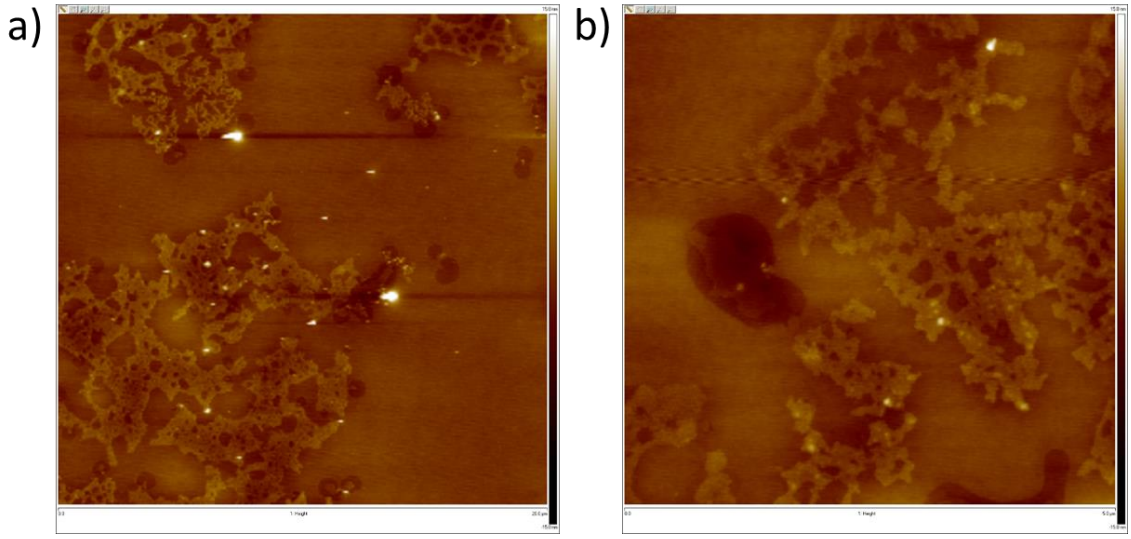


*Figure 81 – AFM images of peptide **69** (1 mg/mL) dried from ammonium hydroxide and redissolved in water with zinc nitrate (3 eq.); a) 20 x 20 µm, b) 5 x 5 µm*

A combination of AFM and DLS studies were thus used to determine the adopted topology of samples of peptide **69**-zinc complexes. Initial analysis of the free peptide and peptide with base

(NaOH/KOH/TEA) in water showed no formation of any ordered assemblies with either gel-like films or random aggregates visible as expected. Likewise, the peptide in MeOH and PrOH without zinc showed no ordered assembly with just random aggregates forming. While synthesising the complex using similarly treated peptide to previous methods, drying from ammonium hydroxide, instead formed chain-like aggregates (Figure 81). This is similar to the structures seen in some areas for the hexamer in water, and matches closely with the structures seen in ethanol, suggesting the longer peptide is assembling in a similar manner. However, individual nanoparticles are not visible only showing aggregates, suggesting some deviation from the previous samples which could be attributed to many factors, such as the change from TEM grid to the silica mica disc, dilution effects or differences between the peptides.

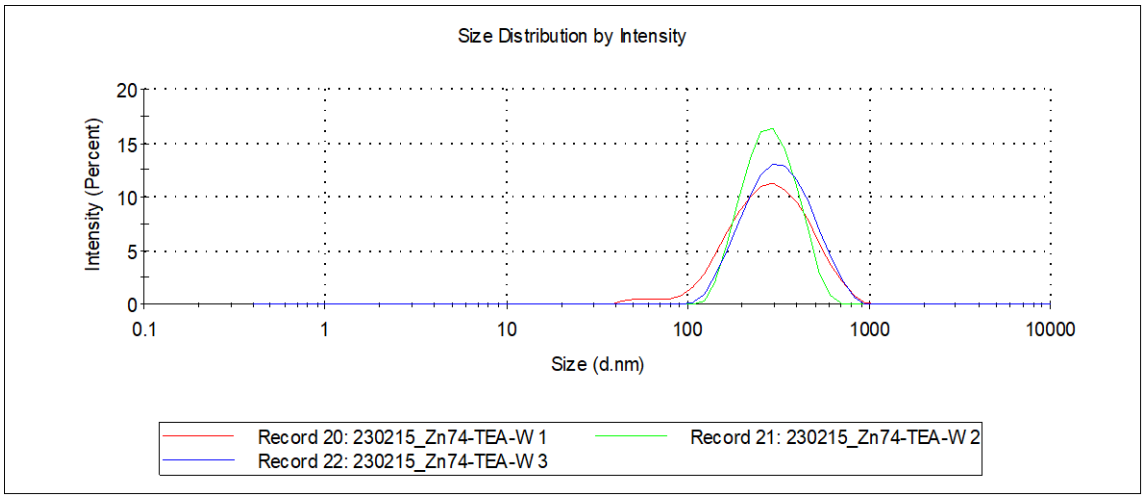


*Figure 82 – DLS size distribution for peptide **69** (1 mg/mL) in water with TEA (6 eq.) and zinc nitrate (3 eq.). Z-avg: 247.1 nm, PDI: 0.226*

Due to the lack of visibility of single nanoparticles the synthesis of the peptide complex was attempted with an alternate base, using triethylamine (TEA) in place of any of the initially tested ammonium hydroxide. DLS analysis of the complex in water indicated the presence of nanoparticles with an average size of approximately 250 nm with a reasonable PDI (0.226) (Figure 82). Subsequent AFM analysis of this complex did not show the presence of the chain-like aggregates, seen with ammonium hydroxide, matching with the findings from DLS analysis instead showing the presence of discrete particles less than 100 nm in size. These objects had relatively low height (< 8 nm) suggesting that if these are spherical nanoparticles in solution that they have collapsed upon drying on the mica surface and are therefore hollow, solvent-filled spheres (Figure 83).

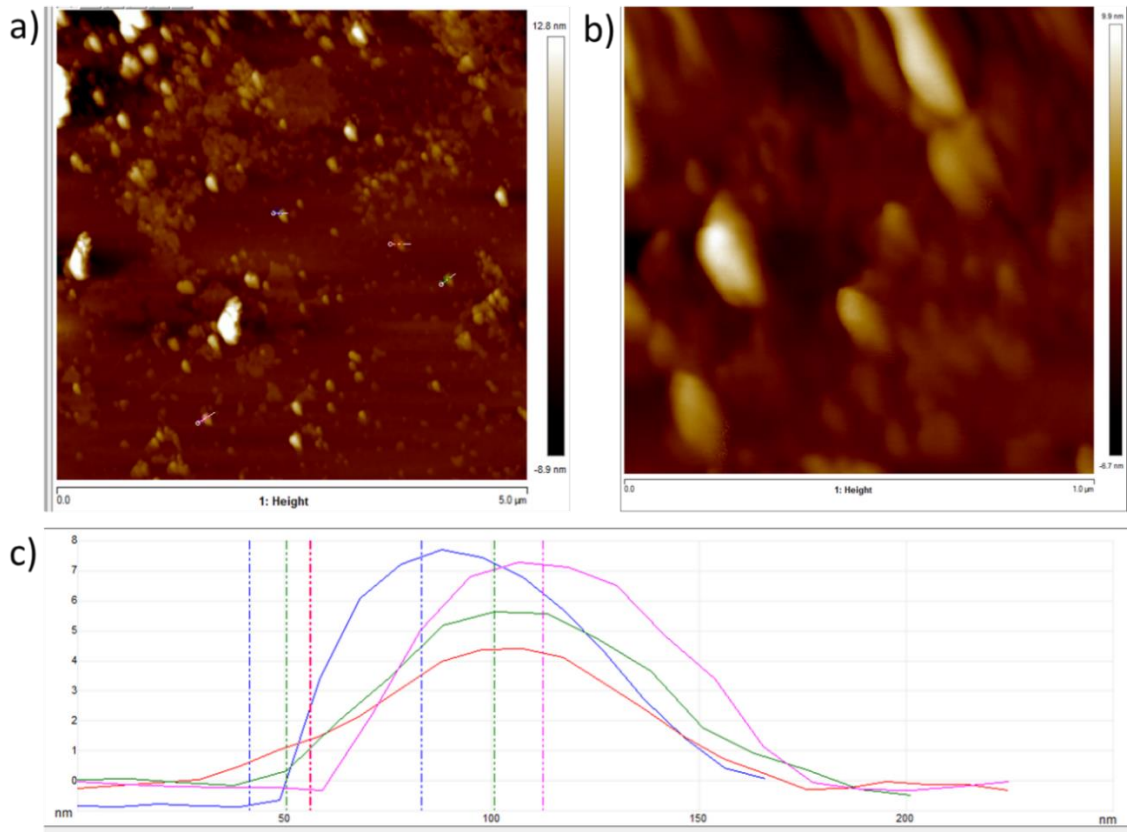
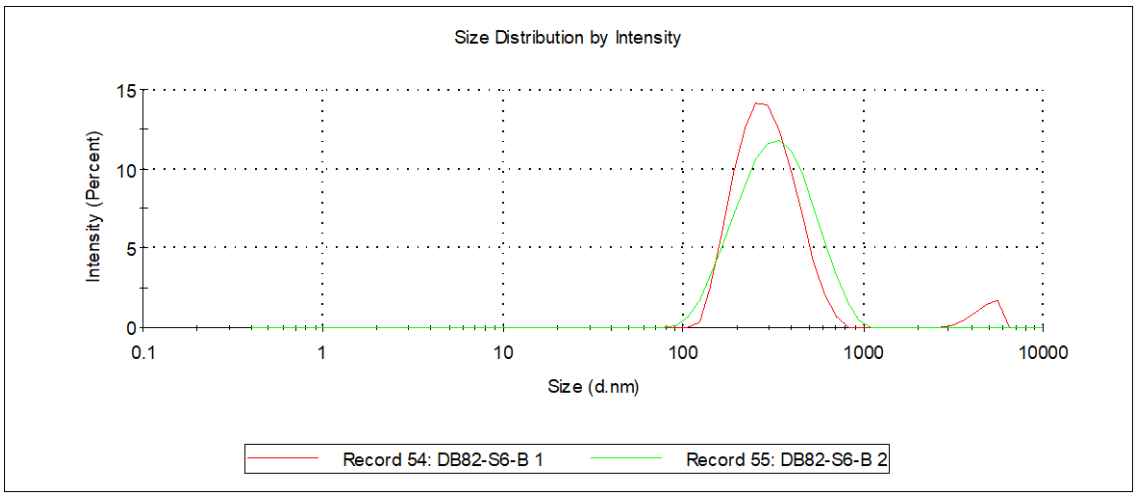


*Figure 83 – AFM images of peptide **69** (1 mg/mL) in water with TEA (6 eq.) and zinc nitrate (3 eq.); a) 5 x 5 µm, b) 1 x 1 µm, c) table showing diameter and height of spherical objects from image (a); height – 2.6-8.6 nm, diameter – approx. 100 nm*

Due to the successful formation of nanoparticles from peptide **69** with TEA and zinc in water this method was carried forward for analysis with other solvents. DLS analysis of the complex formed in methanol solution at a lower concentration (0.4 mg/mL), due to reduced solubility in organic solvent versus water, indicated the formation of similarly sized nanoparticles to water at 295 nm (Δ 48 nm) with a better PDI (0.186) (Figure 84). In methanol the peptide predominantly adopts the polyproline II helix, as shown by CD analysis, as such methanol acts as a good model of the behaviour of the complex assembled from the polyproline II helix while in organic solvent, such that the differences in topology cannot be directly attributed to differences between aqueous and organic solvents. The good match between the DLS of the complex formed in water and MeOH supports the hypothesis that the helicity is driving the adoption of this topology, not solvent effects.



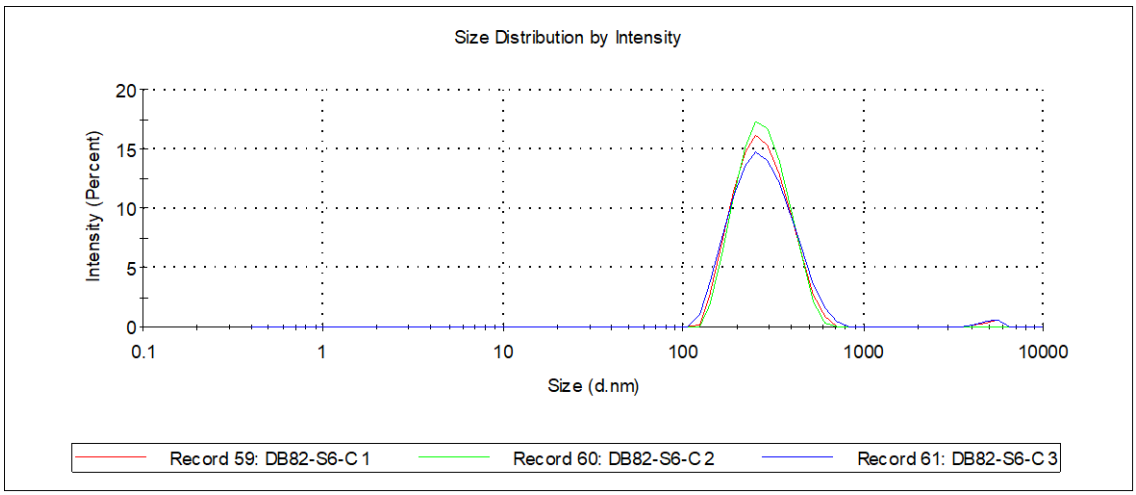


*Figure 84 – (top) DLS size distribution for peptide **69** (0.4 mg/mL) in MeOH with TEA (6 eq.) and zinc nitrate (3 eq.). Z-avg: 295.6 nm, PDI: 0.186 (bottom) DLS size distribution for peptide **69** (0.4 mg/mL) in MeOH with TEA (6 eq.) and zinc nitrate (3 eq.). Z-avg: 263.0 nm, PDI: 0.153*

In an attempt to further stabilise the nanoparticles formed to improve the PDI and prevent aggregation the particles were "capped" *via* the addition of benzoic acid subsequent to complexation. This method showed a marked improvement in the degree of aggregation over time, improved the PDI of the particles from DLS analysis (0.153 vs 0.186), and slightly reduced the determined size from DLS (263 nm) (Figure 84).

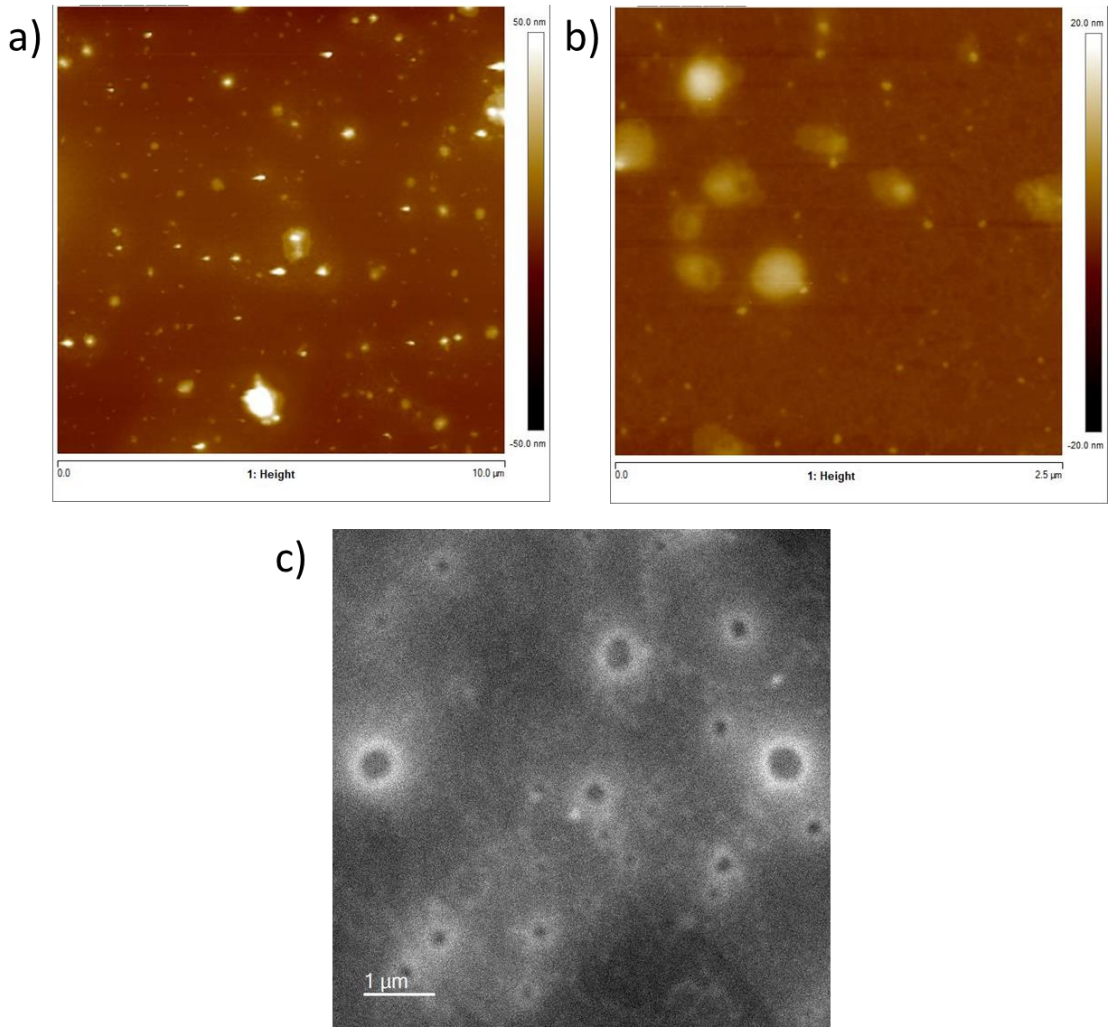


*Figure 85 – AFM images of peptide **69** in MeOH (0.4 mg/mL) with TEA (6 eq.) and zinc nitrate (3 eq.) and subsequently capped with benzoic acid (1 eq.); a) 10 x 10 µm, b) 2.5 x 2.5 µm, horizontal distance – 300-440 nm, vertical distance- 5-15 nm (larger circular objects); c) TEM image, unstained sample.*

However, AFM analysis of the capped complex formed in methanol did not show the presence of spherical nanoparticles with flat 2D circular structures forming, these were relatively large 300-440 nm and despite being circular had a vertical height that was very low for their diameter (5-15 nm) so it is unclear if these existed as nanoparticles in solution and have simply collapsed upon drying (Figure 85). TEM analysis of the complex also showed similar structures with relatively disperse circular objects clearly visible (Figure 85). Although taken with the DLS results this suggests that nanoparticles may be forming in methanol but that they do not survive the drying or washing processes of sample preparation.

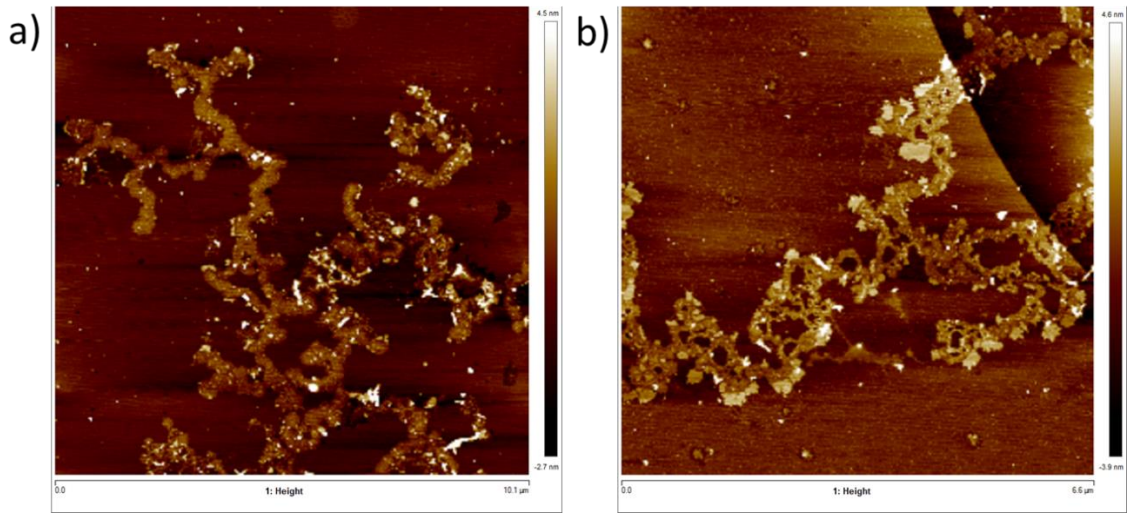


*Figure 86 – AFM images of peptide **69** dried from ammonium hydroxide and redissolved in EtOH (0.4 mg/mL); a) 10 x 10 µm, b) 6.6 x 6.6 µm*

As for previous peptides complexation of peptide **69** was repeated in EtOH, with the peptide incubated to ensure complete conversion to the polyproline I helix prior to assembly. Initial AFM studies of the peptide with ammonium hydroxide showed the formation of chain-like aggregates (Figure 86) previously seen for ammonium hydroxide-zinc complexes of the peptide in water and for the less functionalised peptides (**67-68**). With the addition of zinc nitrate however a significantly different assembly was formed.

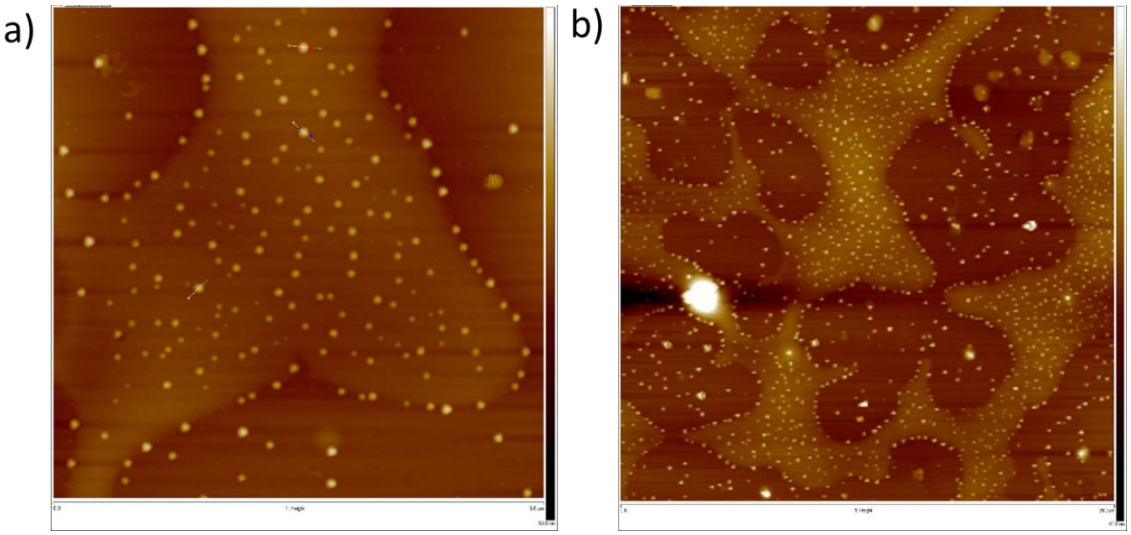


Figure 87 - AFM images of peptide 69 dried from ammonium hydroxide and redissolved in EtOH (0.4 mg/mL) with zinc nitrate (6 eq.); a) 6.6 x 6.6 µm, b) 10 x 10 µm

AFM analysis of this complex in EtOH showed the formation of two architectures, with relatively monodisperse nanoparticles and a thin film with these same nanoparticles also dispersed throughout (Figure 87). This clearly shows that while the peptide can aggregate to form supramolecular structures with some degree of order such as the chain-like aggregates, complexation of the peptide with zinc can drive the assembly towards a different topology such as the nanoparticles seen here, indicating the successful complexation of the peptide and formation of an extended metal-peptide network.

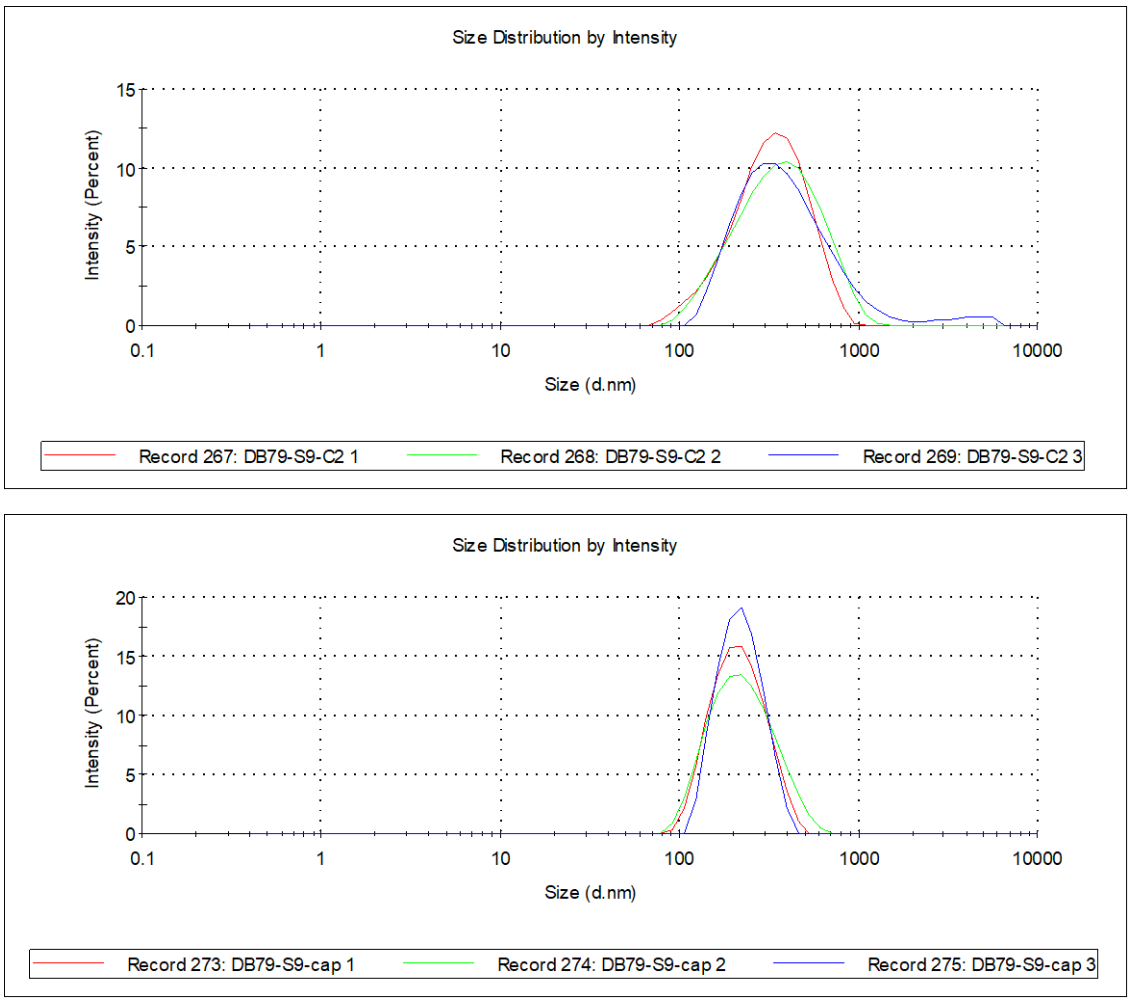


*Figure 88 – (top) DLS size distribution for peptide **69** (0.4 mg/mL) in EtOH with TEA (6 eq.) and zinc nitrate (3 eq.). Z-avg: 284.3 nm, PDI: 0.176; (bottom) DLS size distribution for peptide **69** (0.4 mg/mL) in EtOH with TEA (6 eq.) and zinc nitrate (3 eq.), capped with benzoic acid (1 eq.). Z-avg: 207.3 nm, PDI: 0.147*

Due to the preliminary evidence from DLS studies of peptide **69** complexes forming smaller nanoparticles with a better PDI and reduced aggregation potential in water after capping with benzoic acid, this study was repeated using the same procedures in EtOH. This gave similar results with a reduction in the determined size (Z-avg: 285 nm -> 210 nm) and an improved PDI (0.176 -> 0.147) (Figure 88). Thus, these results suggest capping with benzoic acid may be an effective method to prevent further aggregation of the nanoparticles to yield discrete, stable nanoparticles.

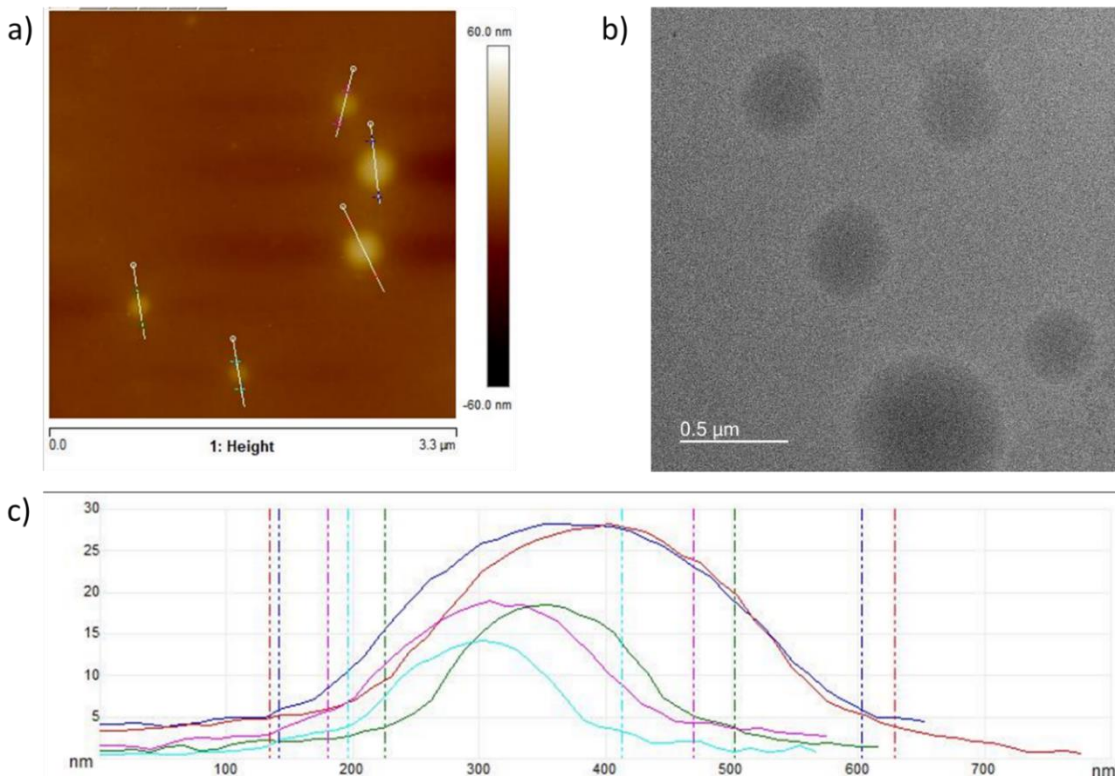


*Figure 89 – a) AFM image of peptide **69** (0.4 mg/mL) in EtOH with TEA (6 eq.) and zinc nitrate (3 eq.), capped with benzoic acid (1 eq.); b) TEM image, unstained sample; c) table showing diameter and height of spherical objects from image (a)*

Subsequent, AFM analysis of this sample did present a better match with the DLS results than seen in for the complex in water, with the apparent presence of nanoparticles (Figure 89). These were relatively polydisperse ranging from 200-400 nm, and appeared to be spherical from AFM measurements as they were perfectly circular and had a height ranging from 15-28 nm which was relative to the diameter of the particles again indicating some degree of flattening or collapsing upon deposition and drying. Again, TEM analysis of the same complex gave similar results with large circular objects clearly seen on the grid, suggesting this is the predominant structure formed from the complex. This was carried out without negative staining which suggests the presence of zinc within the structures due their dark colouring, not seen for low density organic compounds, such as the peptide alone, without staining.

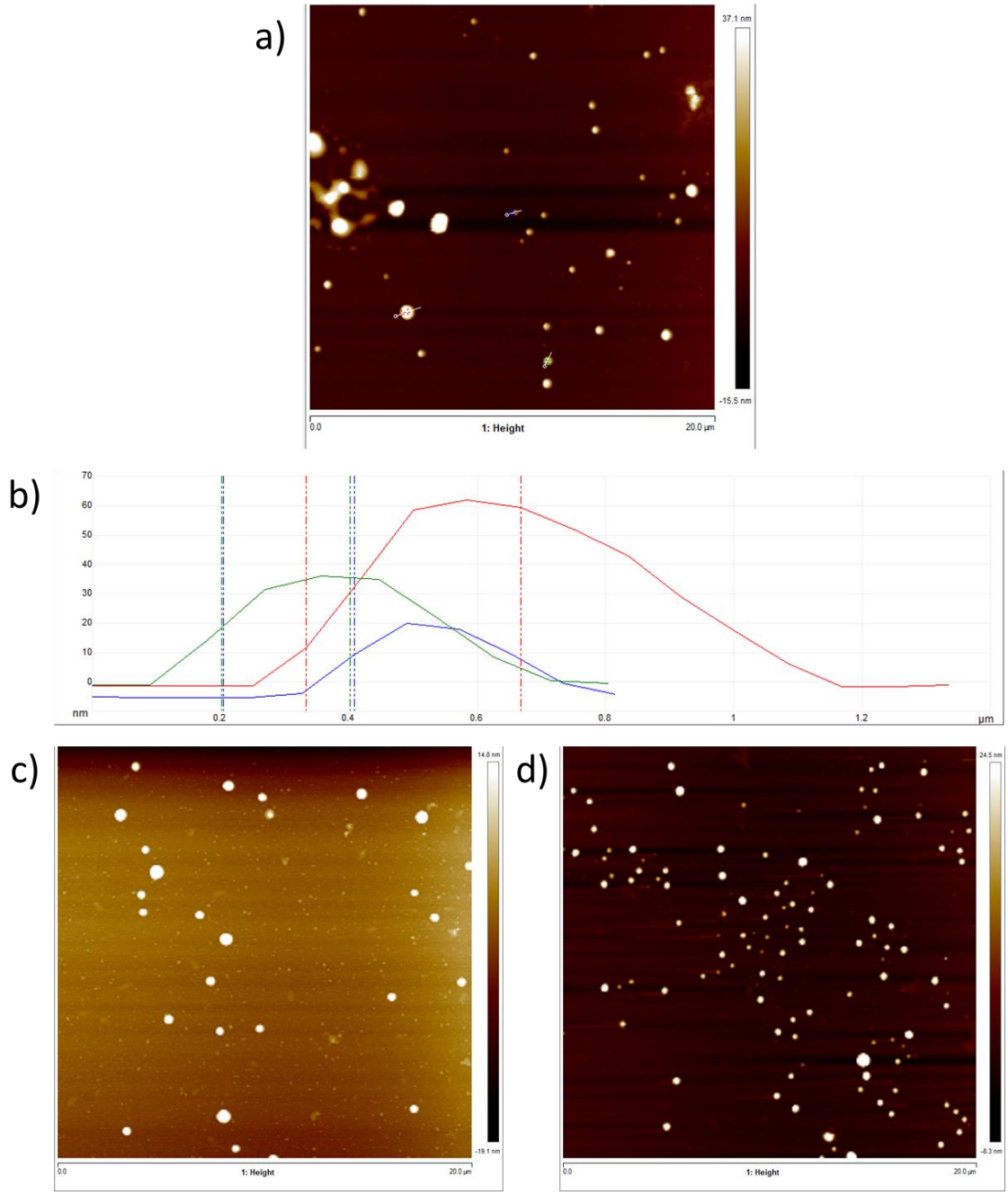


*Figure 90 - AFM images (a,c,d) of peptide **69** (1 mg/mL) in EtOH with KOH (6 eq.) and zinc nitrate (3 eq.), 20 x 20 µm; b) table showing diameter and height of spherical objects from image (a)*

With the successful formation of nanoparticles from peptide **69** in EtOH with both TEA and ammonium hydroxide the complexation was further attempted with KOH as the base with equimolar quantities to the free carboxylic acids. AFM analysis of this complex again showed the formation of nanoparticles that were relatively disperse and larger than previously seen (diameter 200-1000 nm, height 20-60 nm, Figure 90). This clearly shows that the designed peptide **69** predominantly forms nanoparticles upon complexation in EtOH, while the base used also affects the size and dispersity of the structures formed.

To provide supporting evidence for the topologies achieved using the polyproline I helix in the EtOH samples of the peptide **69**-zinc complex these experiments were repeated in propan-1-ol (PrOH). Both EtOH and PrOH have been reported to switch polyprolines to the all-cis polyproline I helix conformation over time, with PrOH achieving slightly higher percentage conversions to the *cis* amide bonds and as such is typically used in experiments to form this helix.[55] Therefore,

forming the peptide-zinc complex in PrOH should produce similar supramolecular assemblies to that seen in EtOH where the helix conformation is the driving factor for the differences in the topology formed.

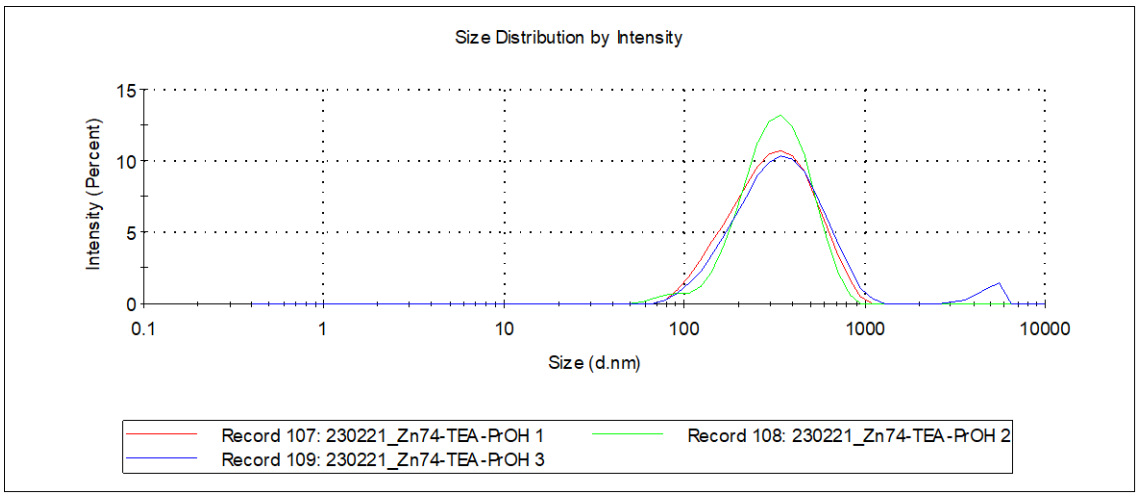


*Figure 91 - DLS size distribution for peptide **69** (0.4 mg/mL) in PrOH with TEA (6 eq.) and zinc nitrate (3 eq.) Z-avg: 311.4 nm, PDI: 0.293*

Initial DLS studies of the peptide complex in PrOH formed with TEA as the base and with zinc nitrate, again indicated the formation of nanoparticles not dissimilar to that seen for MeOH, EtOH, and water with a Z-average size of 310 nm but with a slightly worse PDI than previously seen (0.293) (Figure 91). Looking at the DLS data alone this seems to suggest that nanoparticles of a similar size range are formed in all four solvents tested when using TEA as the base, indicating that the formation of either the polyproline I or polyproline II helix does not significantly affect supramolecular assembly of the complex contrasting with the previous findings.

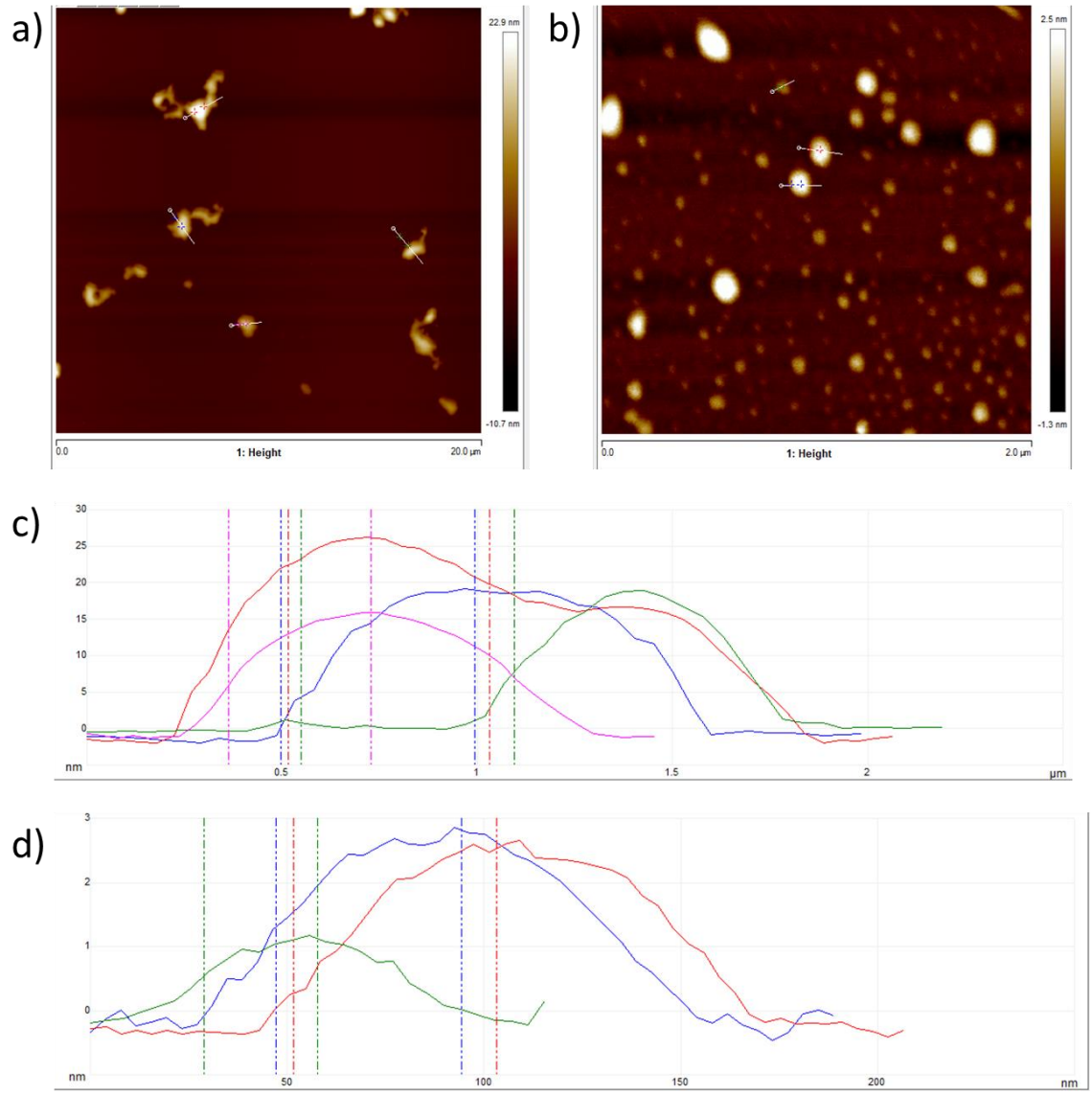
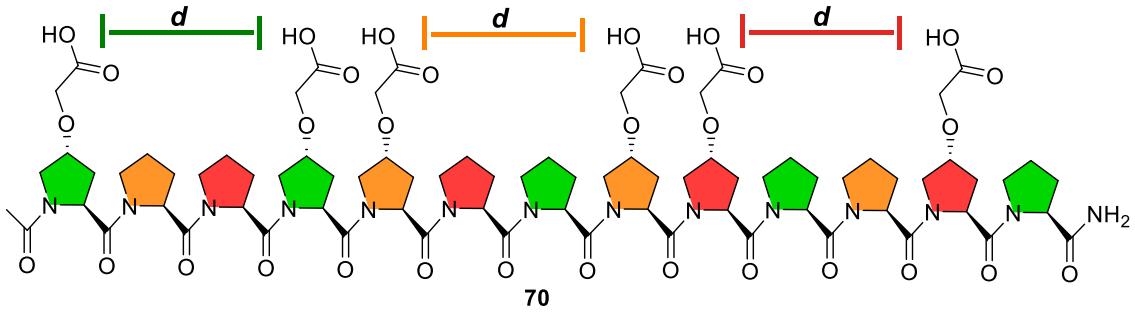


*Figure 92 – AFM images of peptide **69** (0.4 mg/mL) in PrOH with TEA (6 eq.) and zinc nitrate (3 eq.); a) 20 x 20 µm, b) 2 x 2 µm; Tables showing diameter and height of objects from AFM images; c) measurements from image (a), height 15-25 nm, diameter 0.5-2 µm; d) measurements from image (b), height 1-3 nm, diameter 50-150 nm*

However, AFM analysis of this complex instead showed the formation of amorphous gel-like assemblies and despite appearing spherical in Figure 92b these are very thin (< 3 nm) and, with a width of 50-100 nm, are likely thin films, especially when viewed in conjunction with other images (Figure 92a) which clearly show the presence of amorphous gels from the deposited sample.

Thus, from this data it is evident that solution state analysis *via* DLS measurements and solid-state analysis from microscopy methods (AFM/TEM) can provide opposing results with DLS, often indicating the presence of nanoparticles that cannot be corroborated by the microscopy results. From this it cannot be clear whether positive results for the presence of nanoparticles from DLS are either false or simply that the nanoparticles do not survive subsequent treatments for sample preparation such as deposition on the mica, drying, and washing. As such, DLS data alone cannot be used as a definite indicator for nanoparticle formation despite their promising results and could only be used where supported by evidence from microscopy studies that can corroborate these findings.

However, we have been able to show that the designed peptide **69** with six carboxylic acid functional groups is able to reliably form nanoparticles in water, EtOH and PrOH, and with circular objects in MeOH that may be collapsed nanoparticles or thin films/gels. With these samples the base used had a clear impact on the successful assembly of the complex, however successful examples were shown in specific cases for ammonium hydroxide treated peptide, and with equimolar amounts of TEA or KOH, with the most consistent success seen for samples with TEA suggesting it is an effective base to use in the construction of these nanoparticles. We were also able to demonstrate that using a capping agent, in this case the monotopic benzoic acid ligand, can be an effective method to hinder the aggregation of the preformed nanoparticles, which theoretically acts by truncating exposed zinc ions increasing the negative charge of the particles and preventing further aggregation in solution. The requirement of nickel coating the mica to allow deposition of the samples, with no visible assemblies without this treatment for zinc containing samples, clearly shows that the negatively charged mica repels the metal-peptide assemblies and thus these particles must have a negatively charged surfaced.



70

*Figure 93 – Chemical structure of peptide **70**, matching spacings of the functional groups on each face is highlighted (d)*

The data from peptide **69** alone does not conclusively show that placing the functional groups at varying distances on each face drives facial selectivity, to achieve a more ordered assembly. As such, the synthesised peptide **70**, with functionalised prolines placed two prolines apart on each face (*i*+3, Figure 93), was utilised to repeat the complexation experiments carried out on peptide **69**. This peptide should therefore have minimal facial selectivity and was therefore hypothesised to form less ordered assemblies, and thus the effect of this on the topology of the supramolecular assemblies can be determined to improve our understanding of the assembly principles of the peptide complex. However, significantly, CD spectroscopy had shown that this peptide did not adopt the polyproline I helix in EtOH or PrOH, exhibiting the characteristic polyproline II helix spectrum. This was unexpected for such a long peptide suggesting that the placement of the functional groups at one helical turn apart along the peptide sequence affects the propensity for the formation of *cis* amide bonds, preventing switching of the helix. However, this also means the solvent effects on the assembly topology can be monitored as, with the same conformation present in all solvents, any changes can be attributed to the different solvent properties.

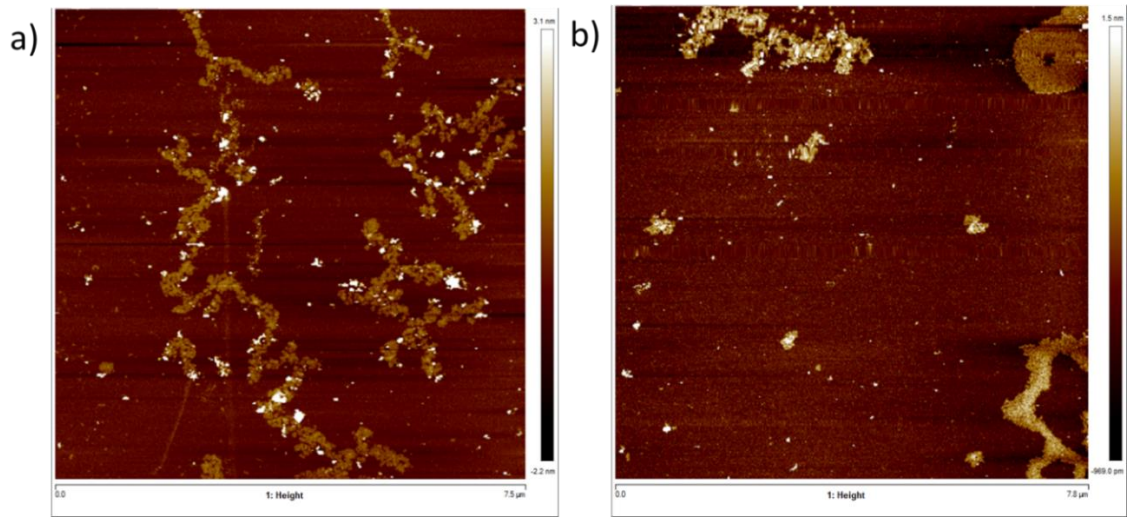


*Figure 94 - AFM images of peptide **70** in water (0.4 mg/mL) with TEA (6 eq.), 7.5 x 7.5 µm*

Analysis of the self-assembly of the peptide-TEA solution in water *via* AFM evidenced similar supramolecular assemblies to that seen for the previous peptide, **69** with ammonium hydroxide (Figure 86), with the formation of chain-like aggregates and similar morphologies (Figure 94).

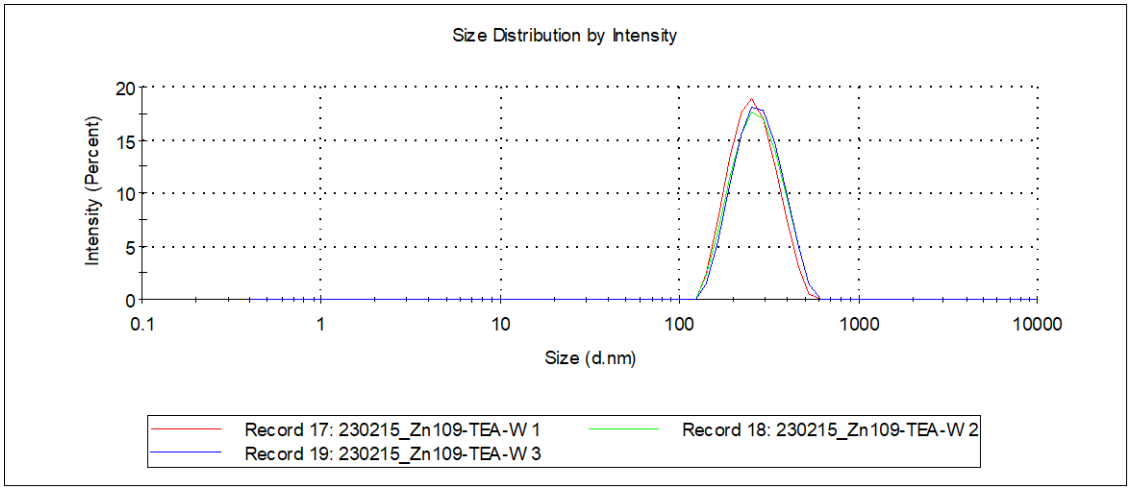


*Figure 95 – DLS size distribution of peptide **70** in water (1 mg/mL) with TEA (6 eq.) and zinc nitrate (3 eq.); Z-avg: 247.2 nm, PDI: 0.106*

The subsequent addition of zinc nitrate to the peptide-TEA solution again formed nanoparticles from DLS analysis with similar sized particles (250 nm) to that seen for the previous peptide under the same conditions with a very good PDI (0.106) (Figure 95). To corroborate these findings a sample of the complex solution was drop-cast onto nickel coated mica for AFM analysis.

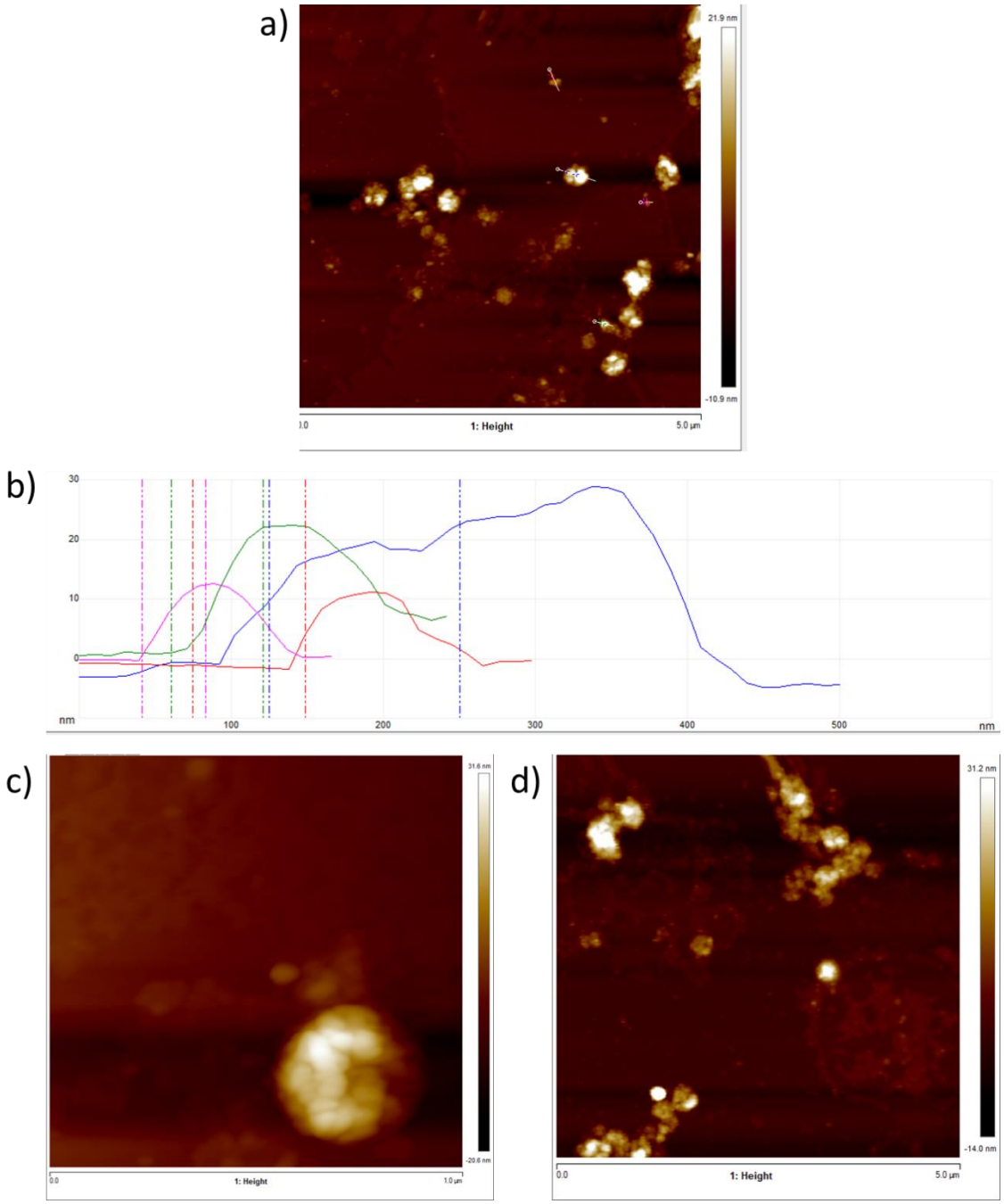


*Figure 96 – AFM images (a,c,d) of peptide **70** in water (1 mg/mL) with TEA (6 eq.) and zinc nitrate (3 eq.); b) table showing diameter and height of objects from image (a)*

The AFM images of the peptide **70**-Zn complex in water show what appear to be aggregated clusters of small particles (Figure 96). This data suggests that while nanoparticles do form in solution, similar to the previous *i*+3/6/9 peptide, these then undergo further aggregation to form cluster-like morphologies with a stark difference to the more discrete particles seen previously (Figure 90). Therefore, this supports the hypothesis that the rational design behind the placement of the functional groups can drive the supramolecular assembly to form more ordered interactions. A possible hypothesis from this evidence is that the non-selective coordination of the helical faces results in a more disordered assembly with more unreacted free carboxylates on the nanoparticle surface which promotes this cluster behaviour between the particles as these sites could bind to exposed metal centres on other particles leading to the cluster-like formations seen here.

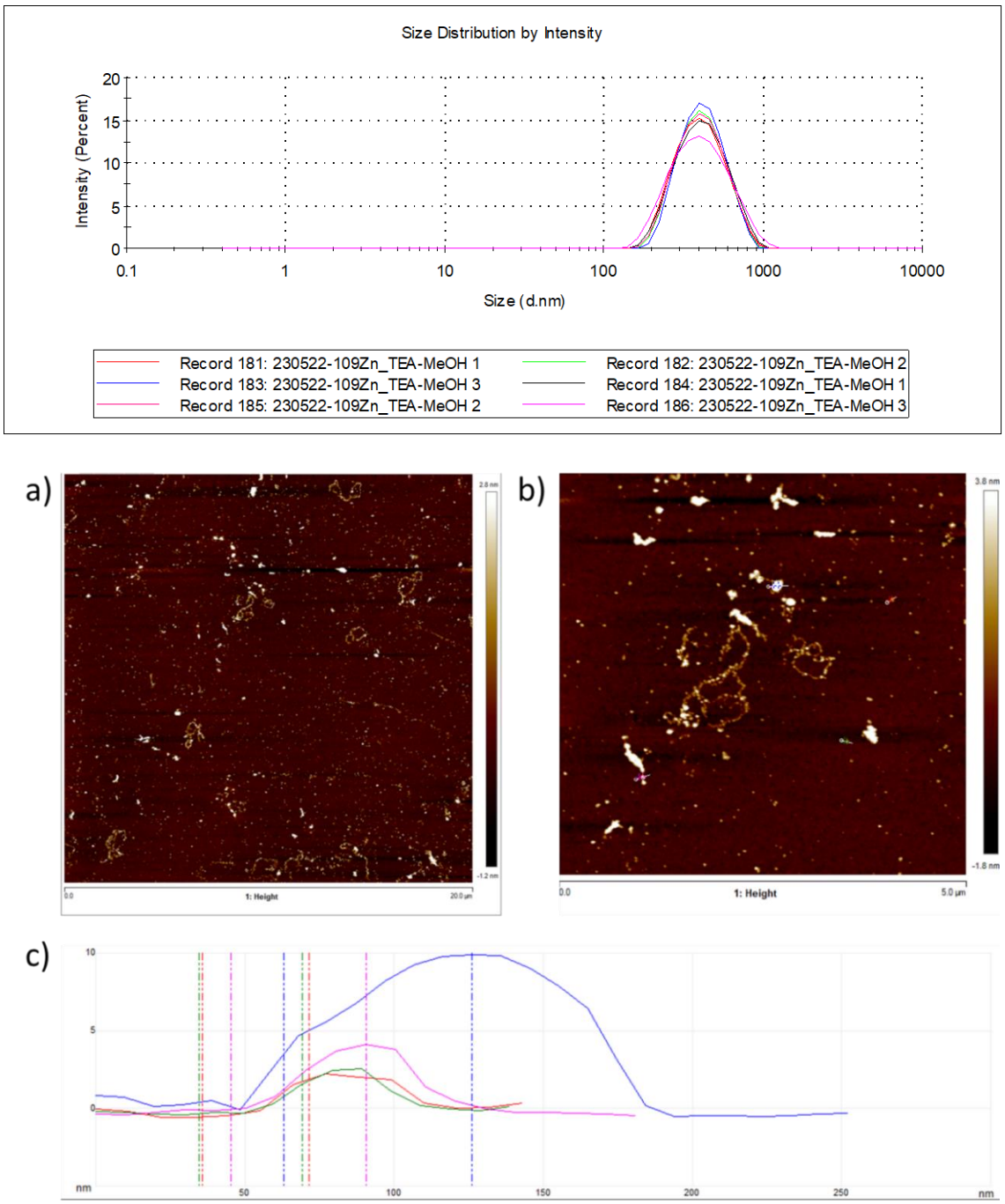


*Figure 97 – DLS size distribution of peptide **70** in MeOH (0.4 mg/mL) with TEA (6 eq.) and zinc nitrate (3 eq.); Z-avg: 247.2 nm, PDI: 0.106 (Top). AFM images (a, b) of peptide **70** in MeOH (0.4 mg/mL) with TEA (6 eq.) and zinc nitrate (3 eq.); c) table showing diameter and height of objects from image (b), height 2.5-10 nm, diameter 50-150 nm*

Subsequently, complexation was carried out in MeOH as done for the previous peptide. DLS analysis of this solution again evidenced the presence of nanoparticles in solution with a good PDI (0.146) and a slightly larger size than previously seen (370 nm). However, AFM analysis of this sample showed a mixture of morphologies, different to the low polydispersity nanoparticles seen in water (Figure 97). The AFM clearly showed areas containing fibre-like assemblies with smaller circular objects also visible (50-150 nm), however these have a low height (3-10 nm), varying in correlation with their diameter (Figure 97c), which may be simply thin-films of the complex or peptide as seen in some previous samples.

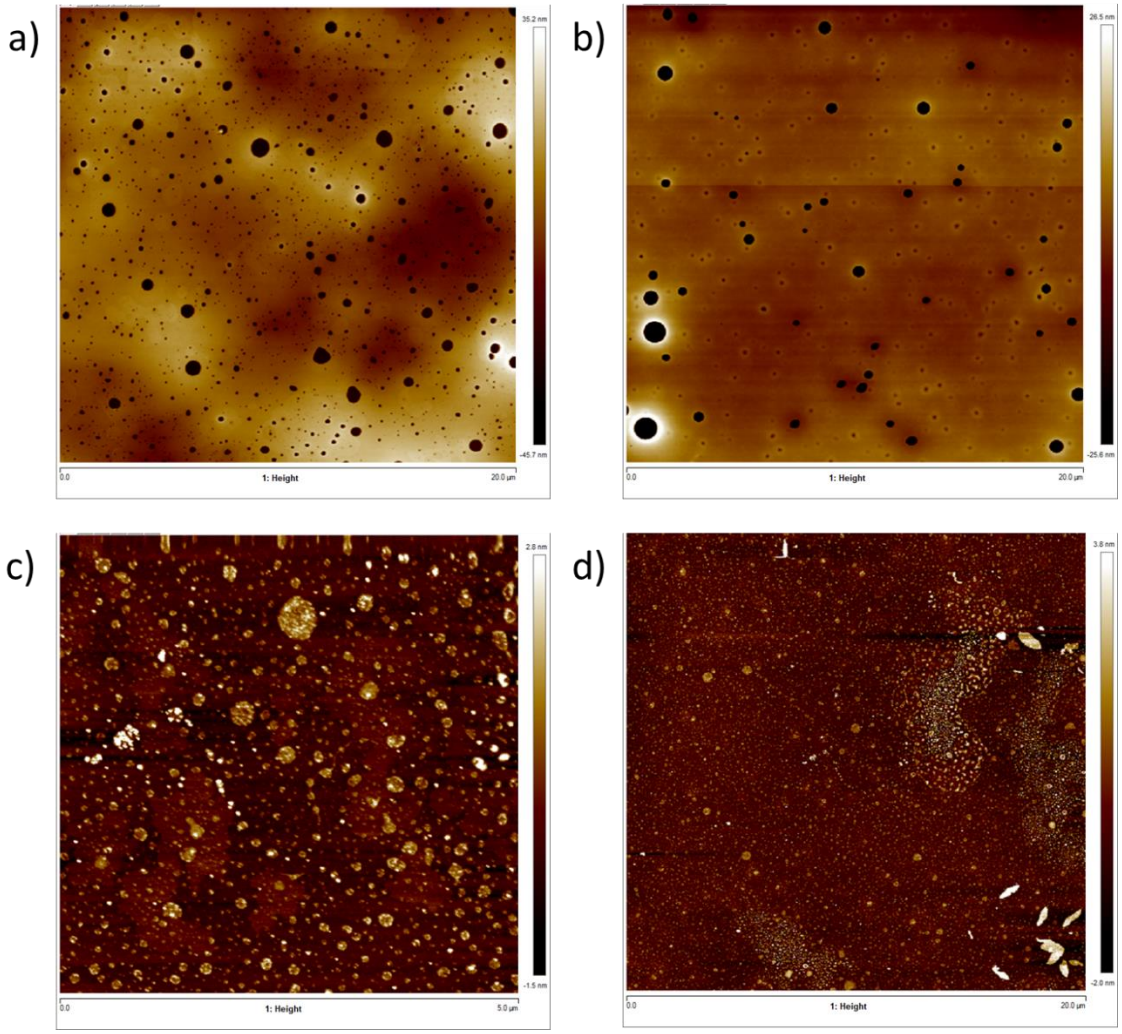


*Figure 98 - AFM images of peptide **70** in EtOH (0.4 mg/mL) with TEA (6 eq.) and zinc nitrate (3 eq.); a-b) sample without washing drop-cast sample with solvent, depth 100-150 n; c-d) drop-cast sample washed with EtOH (10 µL, x 3), depth 2-4 nm*

Analysis of peptide **70** in EtOH, where this peptide still retains the polyproline II helix, after complexation with zinc nitrate in the presence of TEA did not produce any reasonable DLS size distribution. This indicated no spherical morphologies were forming which was supported by subsequent AFM analysis of the sample. Analysis of the sample after drop-casting onto mica without subsequent washes clearly showed the formation of a gel-like film with “holes” from escaping solvent upon drying. While washing the sample clearly diluted this film to form thinner layers in smaller areas of the same film. This supports the information from DLS that no nanoparticle formation occurs for the peptide in EtOH. This differs significantly from the morphologies seen for peptide **69** in EtOH with both ammonium hydroxide and TEA where nanoparticles are clearly visible in both cases (Figure 87).

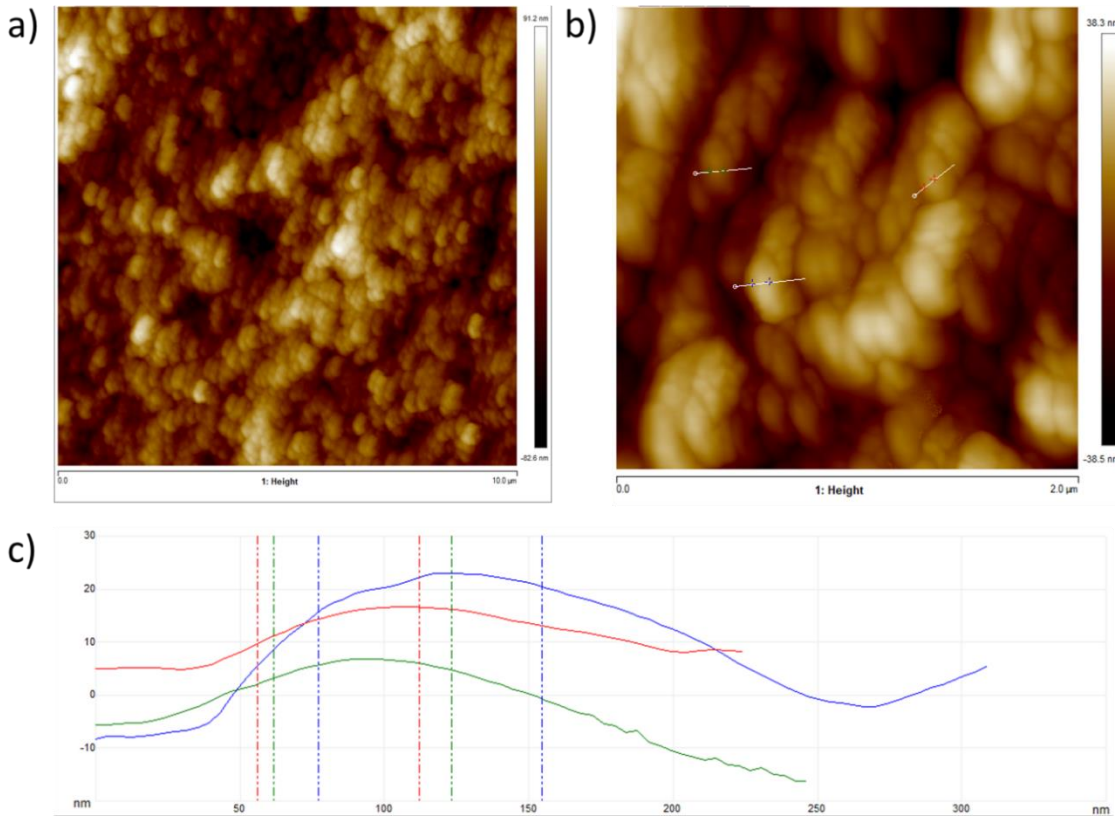
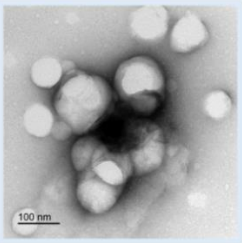
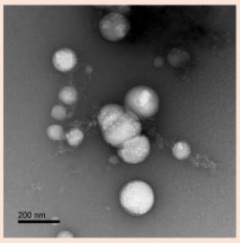
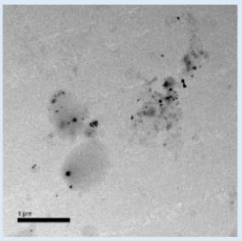
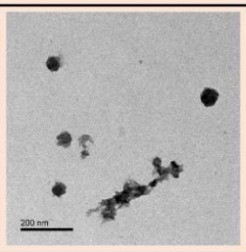
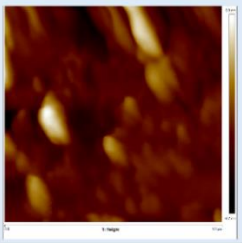
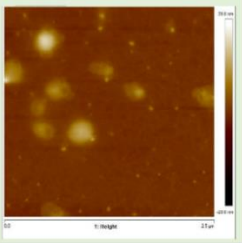
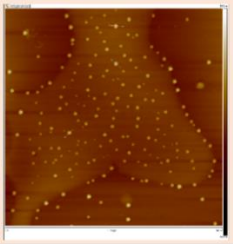
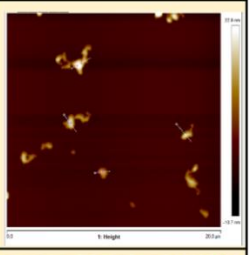
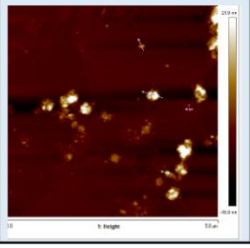
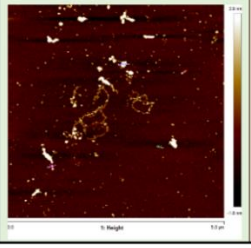
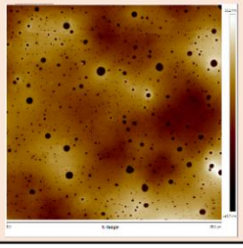
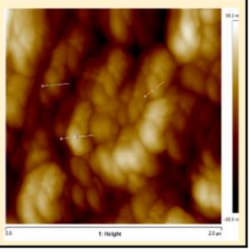


*Figure 99– AFM images (a, b) of peptide **70** in PrOH (0.4 mg/mL) with TEA (6 eq.) and zinc nitrate (3 eq.); c) table showing diameter and height of objects from image (b), height 10-35 nm, diameter 100-200 nm*

In conjunction with the complexation of the peptide in EtOH this was also carried out with the same conditions in propan-1-ol (PrOH) which also still retained the polyproline II helicity. Again, the DLS analysis of this sample did not show any reasonable size distribution indicating no detectable nanoparticles in solution. However, subsequent AFM analysis showed the formation of cluster-like aggregates (Figure 99a-b) similar to the morphology seen in water (Figure 96) yet with more random clusters than the circular clusters seen in the water sample. These dense clusters are formed of globular particles that may be nanoparticles that have quickly aggregated into these clusters. This again suggests that the placement of the carboxylic acid motifs on the peptide sequence to not drive facial selectivity, as is the case in peptide **69**, increases the likelihood of aggregation of the nanoparticles forming clusters and other linked structures.

The stark differences between the AFM results in water, MeOH, EtOH and PrOH for peptide **70** also demonstrates the significant solvent effect on the adopted topology of the assembly. As the peptide adopts the polyproline II helix in all the solvents tested the differences in structures can be attributed to solvent effect. Thus we see variation from spherical clusters, fibres, gels, to densely aggregated clusters. This is likely due to a combination of the lipophilicity and polarity effects on the assembly, and differences in the drying process as the volatility of each solvent varies, impacting the rate of deposition and affecting the morphology of the structures formed. Water is also known to solubilise the peptide backbone carbonyls, increasing the strength of n→π* inter-carbonyl interactions favouring the polyproline II helix,[57] while the longer alcohols have a less stabilising effect, thus a portion of the amide bonds may switch from *trans* to *cis* in these solvents,[58] affecting the assembly process despite the peptide adopting mostly *trans* amide bonds, as seen by CD analysis, which confirmed the polyproline II helicity of the peptide in these solvents.

Table 1: Comparison of microscopy images (TEM & AFM) of each peptide complex					
		Solvent			
		a. H_2O	b. MeOH	c. EtOH	d. PrOH
Peptide Zn-Complex	1. **67**				
	2. **68**				
	3. **69**				
	4. **70**				

In summary, these results show that the peptide length, conformation, linker spacing, degree of functionalisation, the base utilised, and solvent effects, can all be used to alter the accessible topologies from a polyproline-zinc complex. Significantly, we demonstrate how nanoparticles can be synthesised from the complexation of polyproline helices, functionalised with carboxylic acids on the proline sidechains, with zinc nitrate. This occurs for the hexameric peptide **67** in both water and EtOH, while a Pro_{13} peptide with the same number of functional groups (-$(OCH_2COOH)_3$), **68**, does not form nanoparticles in water (polyproline II) and does in EtOH (polyproline I), suggesting the ability to switch conformation directly affects the topology of the metal-peptide framework. The nanoparticles formed from the longer peptide also have a different morphology with a cluster-like structure that aggregates into chains while the shorter peptide forms distinct spheres that meld together when aggregating, which indicates a different mode of assembly for the two peptides.

Further functionalisation of the peptide to induce facial selectivity, increasing the directionality of the supramolecular interactions, with the synthesis of peptide **69**, then drove the formation of nanoparticles in water unlike the less functionalised peptide. Showing that optimisation of the peptide allows the formation of previously unobtainable assembly topologies. This peptide also formed nanoparticles in EtOH and showed some positive results for nanoparticle formation in MeOH. While the PrOH sample simply formed gel-like structures. This indicates that solvent effects are also important factors for the assembly process as the peptide adopts the polyproline

I helix in both EtOH and PrOH, yet the adopted structures are different (nanoparticles/gels), which instead may be attributable to differences in the solvent properties such as lipophilicity, viscosity, volatility, and surface tension.

Subsequent adjustment of the functional group placement was carried out to synthesise peptide **70** to test the importance of the facial selectivity. Thus, the selectivity of the peptide should be reduced such that all the peptide faces should be able to form intermolecular interactions with each other, this led to significant differences to peptide **69** in all solvents. The water sample showed the most similarities still forming nanoparticles, yet these seemed to form spherical clusters of nanoparticles showing the surface properties of the particles were different affecting the aggregation of the particles. The MeOH sample was distinctly different forming fibrous assemblies not seen in any other samples. The peptide assembled into an organogel in EtOH and, while some gel-like assemblies were visible for the previous peptide **69,** these were evenly dispersed with nanoparticles not seen in peptide **70**. This data shows that placement of functional groups significantly affects the assembly's topology and none of the adopted structures match despite retaining the same polyproline II helix, shown from CD analysis (Figure 66 to Figure 69), showing that solvent effects significantly affect the topology.

3.3.3 Metal-Peptide Cages

3.3.3.1 Synthesis of discrete carboxylate peptide-metal cages

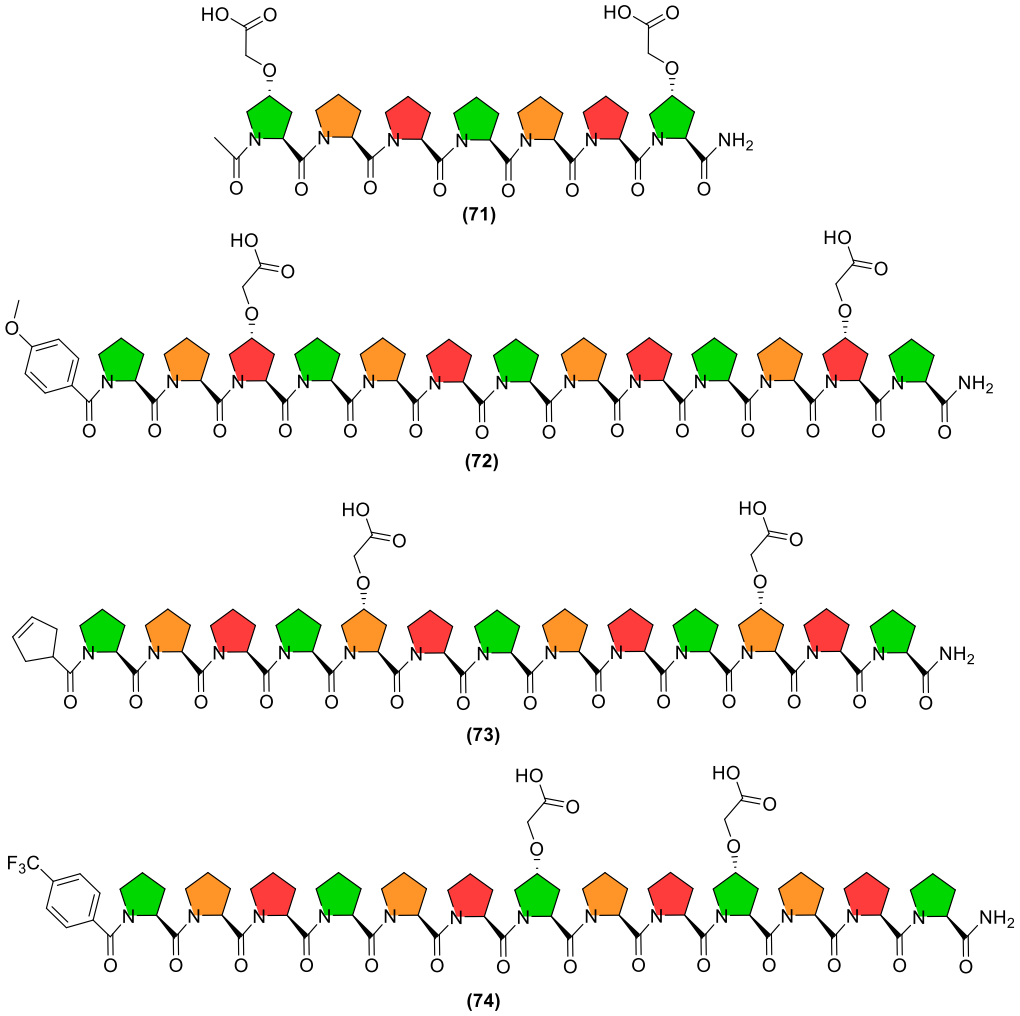


Figure 100 – Chemical structures of synthesised polyproline peptides functionalised with -OCH$_2$COOH motifs on only one of the peptide helices' faces. Residues on the same helical face are highlighted in the same colour.

With the formation of complex supramolecular assemblies (nanoparticles/gels) from multi-topic polyproline ligands (**67-70**), where all three faces of the polyproline helix were functionalised (Figure 100), the focus moved towards simpler systems to be able to form discrete supramolecular constructs with nanoscale cavities, with potential applications for host-guest chemistry and catalysis. It was also evident further elucidation of the methods of assembly and the effect of face functionalisation on the adopted interactions (i.e. selectivity of face interactions) was required. As such, a series of peptides functionalised only on one helical face, with varying distances between the two functional groups on each face, were synthesised. With the capping groups chosen due to their ease of identification *via* ^{1}H NMR spectroscopy, containing aromatic and alkene protons, with a chemical shift well outside the range of the peptide backbone protons, as well as, differing the masses of the peptides for LCMS analysis. These were peptides **71**-**74** (Figure 100) and were designed to mirror the helical faces of peptide **69**, with each peptide having the carboxylic acid sidechains placed on different helical faces and with different spacings. With these peptide structures the goal was therefore to form a complexed peptide dimer with metal nodes acting as linkers between the carboxylic acid groups, creating a nano-cavity with the peptide helices as the walls and the central proline ring of both peptides facing into the cavity. However, HPLC analysis of the crude peptide **73** showed the presence of significant impurities suggesting an issue during peptide synthesis, or the capping steps, and with the successful synthesis of peptides **72** and **74** initial studies were instead focused on these two peptides.

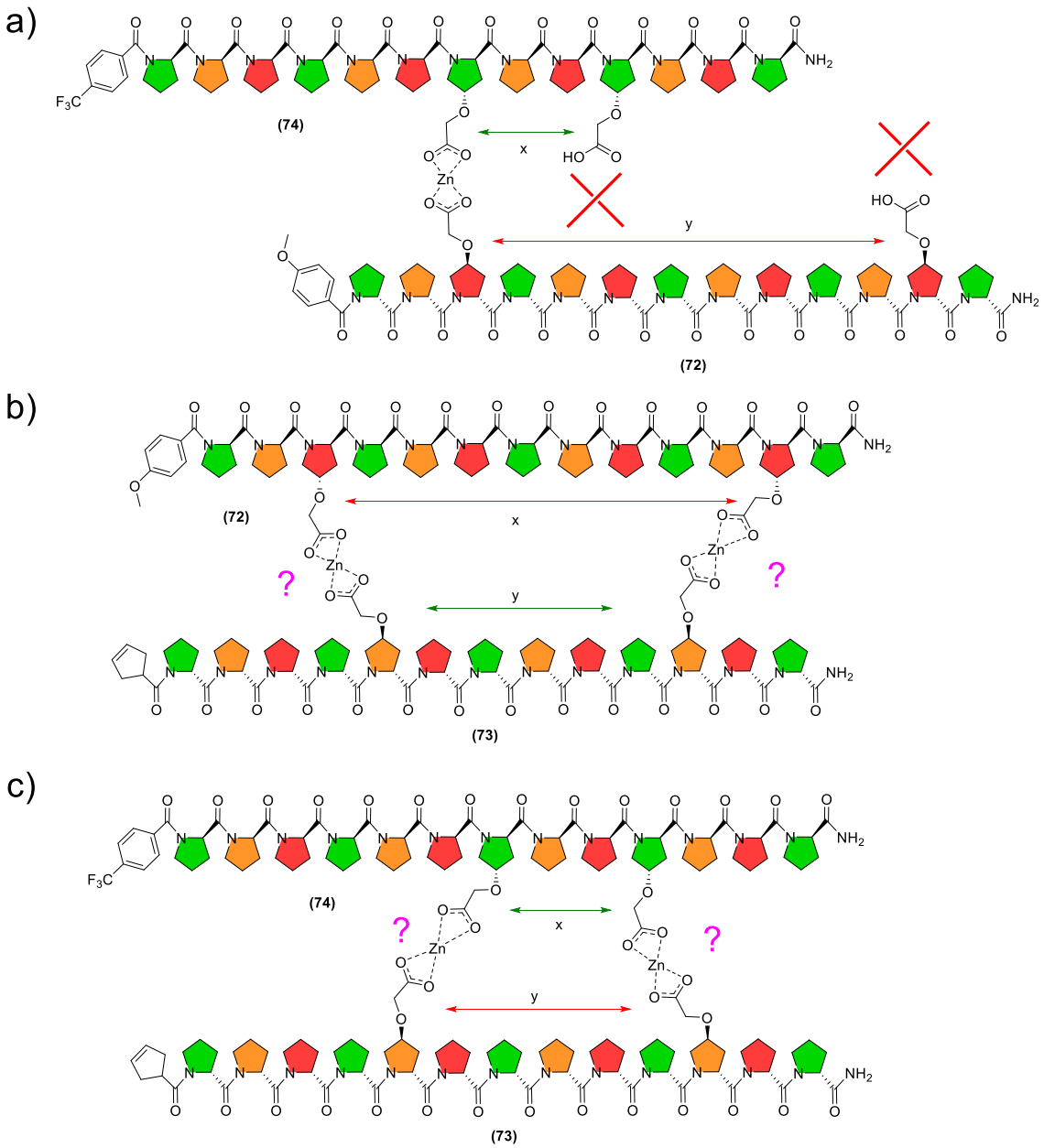


*Figure 101 – Example structures of heterogenous Zn-peptide complexation of peptides **72-74** with each other; a) Peptide **72** and **74** dimeric complex showing the impossibility to satisfy both carboxylic acid linkers simultaneously due to large differences in x and y distances, b) Peptide **72** and **73** dimeric complex showing the uncertainty that the difference in x and y is sufficient to prevent simultaneously linking of all carboxylic acid sites; b) Peptide **74** and **73** dimeric complex showing the uncertainty that the difference in x and y is sufficient to prevent simultaneously linking of all carboxylic acid sites*

With these longer peptides (**72-74**) the goal was to determine whether the previously utilised placement of two binding sites on each face could result in selectivity of the peptide complexation. For example peptides **72** and **74**, **72** having the greatest functionalised residue distance (*i*+9) and **74** having the shortest (*i*+3), should not be able to form two carboxylate-metal linkers with each other simultaneously as the formation of one linkage should result in the distance between the other free acids being too great (*i*+6 difference) to form a peptide-peptide link (Figure 101a). Thus, if the two peptides were mixed and complexation carried out with zinc then two separate complexes should form, rather than mixed complexes containing both peptides, which may be detectable by NMR analysis and mass spectrometry. However, analysing the sidechain spacings it should be considered that peptide **73** has an intermediate spacing between the two other peptides' functionalised residues (i+6), as such, it is unclear whether the

difference in spacings (i+3 difference) is sufficient to prevent linker formation at both binding sites on each molecule simultaneously (i.e. **74=73** and **73=72**, Figure 101b-c) due to the flexibility of the -OCH_2COOH group. If peptide **73** can bind in this manner this would result in peptide **69** forming a less ordered assembly than desired as all the faces can bind to their complementary face yet face 2 would also be able to bind to faces 1 and 3, reducing the selectivity of the assembly process. Thus, determining the degree of selectivity of the assembly between peptides of varying linker spacings on different faces is essential for the rational design of any supramolecular constructs.

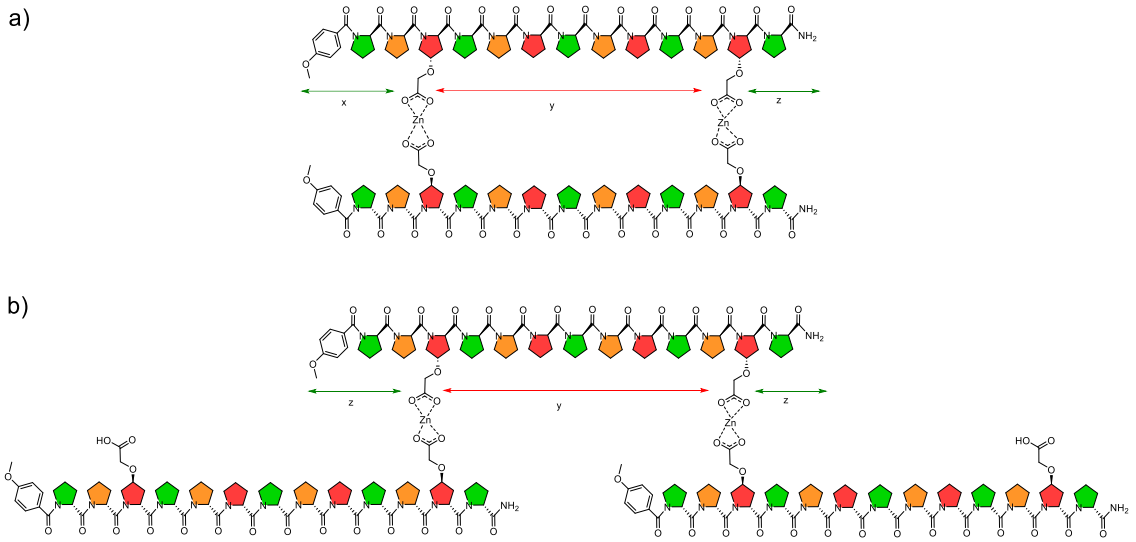


*Figure 102 - Possible structures of peptide **72** - zinc complexes; a) formation of a peptide dimer, b) formation of a peptide 1D polymer chain,*

Secondly, an important factor considered was the propensity for the peptides to form a linker at both binding sites to the same peptide or whether a third molecule can form a linker with the as formed dimer and thus lead to polymerisation into a 1D polymer chain (Figure 102b). Theoretically as the distance between the linkers (y) in peptide **72** is significantly larger than the distance of peptide remaining at either terminus (*x* & *z*) there is minimal hindrance to form an extended polymer, however, the reaction to form a single dimer should still be significantly favoured as the reaction sites are bound close together after the initial dimerization *via* one carboxylate linkage.

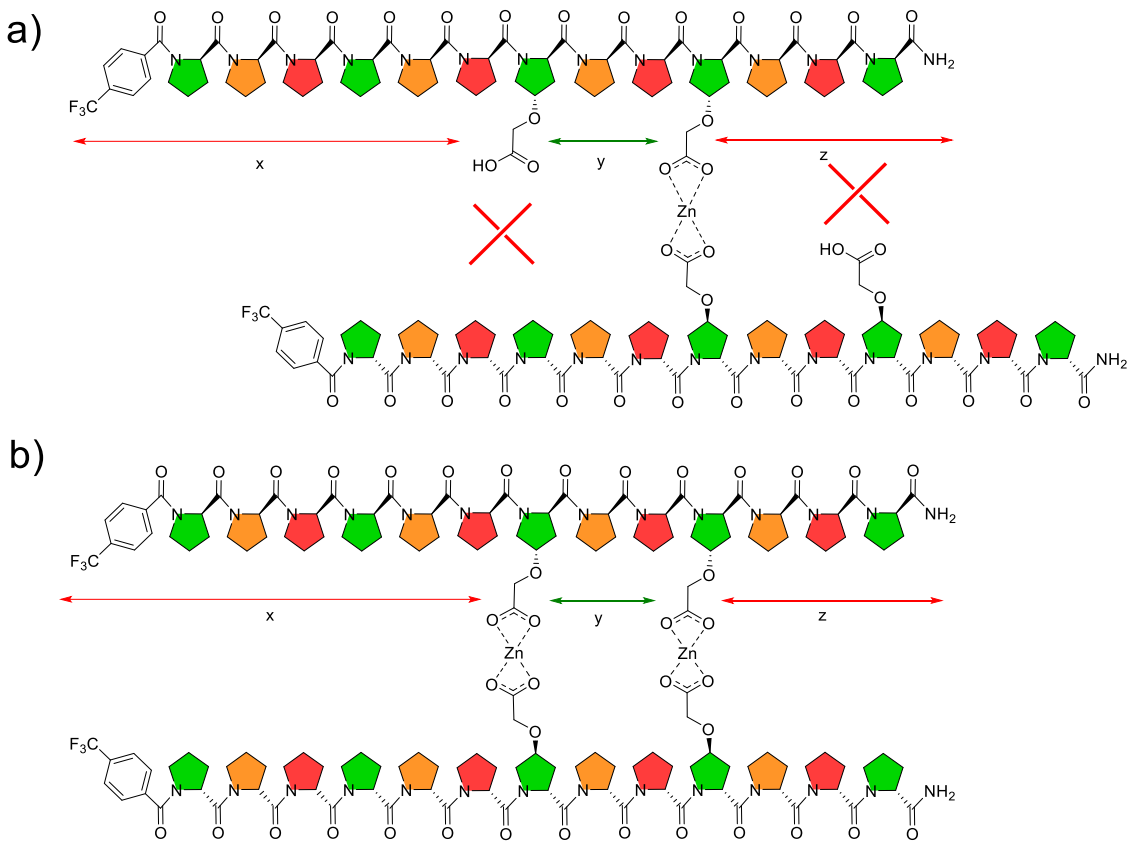


*Figure 103 - Possible structures of peptide **74** - zinc complexes; a) shows the formation of a single carboxylate link and how the significantly smaller distance of y versus x and z reduces the likelihood of polymerisation due to the steric hindrance of the peptide chain adjacent to the linker motifs hindering the angle of approach of further peptide molecules, b) Complexation of a peptide-**74**-zinc dimer, theoretically favoured over polymerisation*

In contrast, the peptide **74**, and to an extent peptide **73**, has a significantly smaller *y* distance between the linkers versus the peptide adjacent to these motifs at each terminus (*x* & *z*), this should hinder the angle of approach of further peptide molecules due to the steric bulk of the peptide backbone around the binding sites and thus prevent polymerisation of the peptide (Figure 103). If this can be determined this will clearly show that the peptide length and spacings of functional groups can be tuned to prevent the formation of specific assemblies and achieve a desired topology.

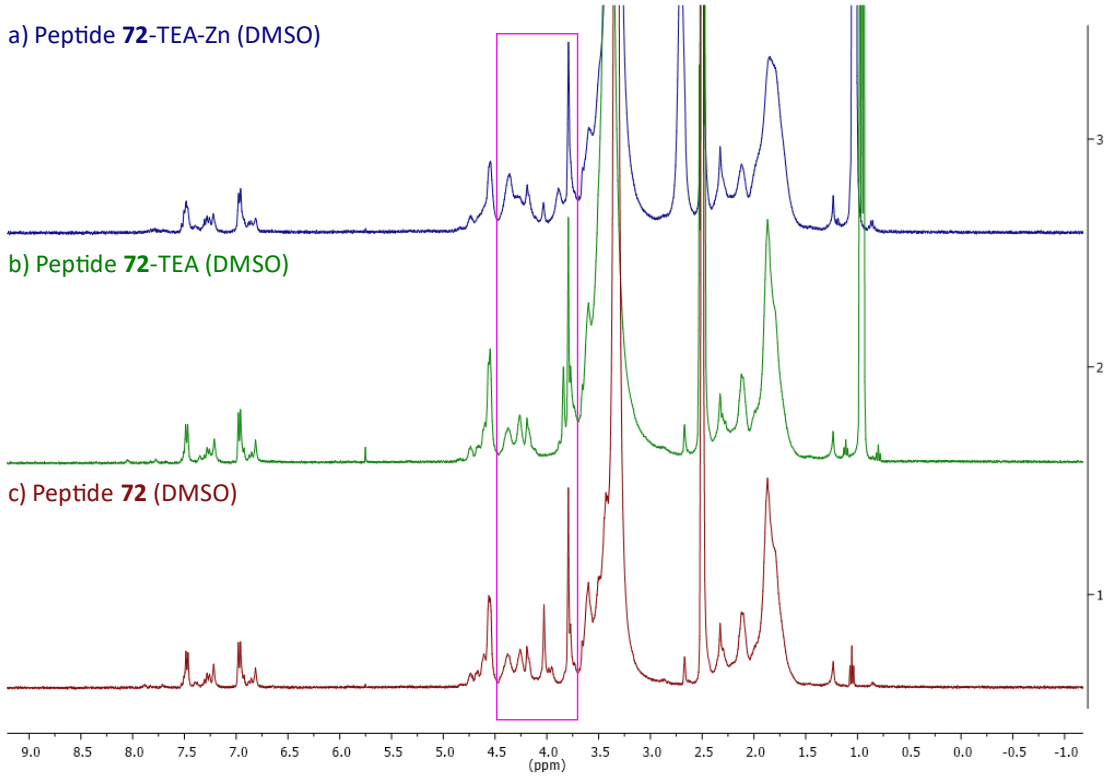


*Figure 104 – ^{1}H NMR Spectra of peptide **72** in DMSO-d_6 a) peptide with TEA (6 eq.) and zinc nitrate (1 eq.), b) peptide with TEA (6 eq.), c) peptide alone. Area with most significant differences highlighted in pink.*

Consequently complexation of these peptides was carried out using a similar method to that adopted for the nanoparticle synthesis. In this case the peptide was heated at 65 °C with equimolar amounts of zinc nitrate hexahydrate in either MeOD, DMSO-d_6 or D_2O, thus allowing for NMR analysis of the sample to assist in determination of the successful complex formation. In this case no precipitate or turbidity was observed for any of the samples, except peptide **72** in MeOD, indicating that unlike with the multi-face functionalised peptides these peptides do not form insoluble extended structures, instead forming discrete assemblies that remain soluble even in poorer solvents such as methanol. The presence of some precipitate from the peptide **72-Zn** complex in methanol suggests that an extended structure forms supporting the hypothesis that a degree of polymerisation can occur through the assembly shown in Figure 102b. However, most of the peptide remains in solution as indicated by NMR analysis. This does not occur for the closely placed carboxylic acid motifs in peptide **74** showing that these chemical handles can be used to control the topology of the adopted assembly.

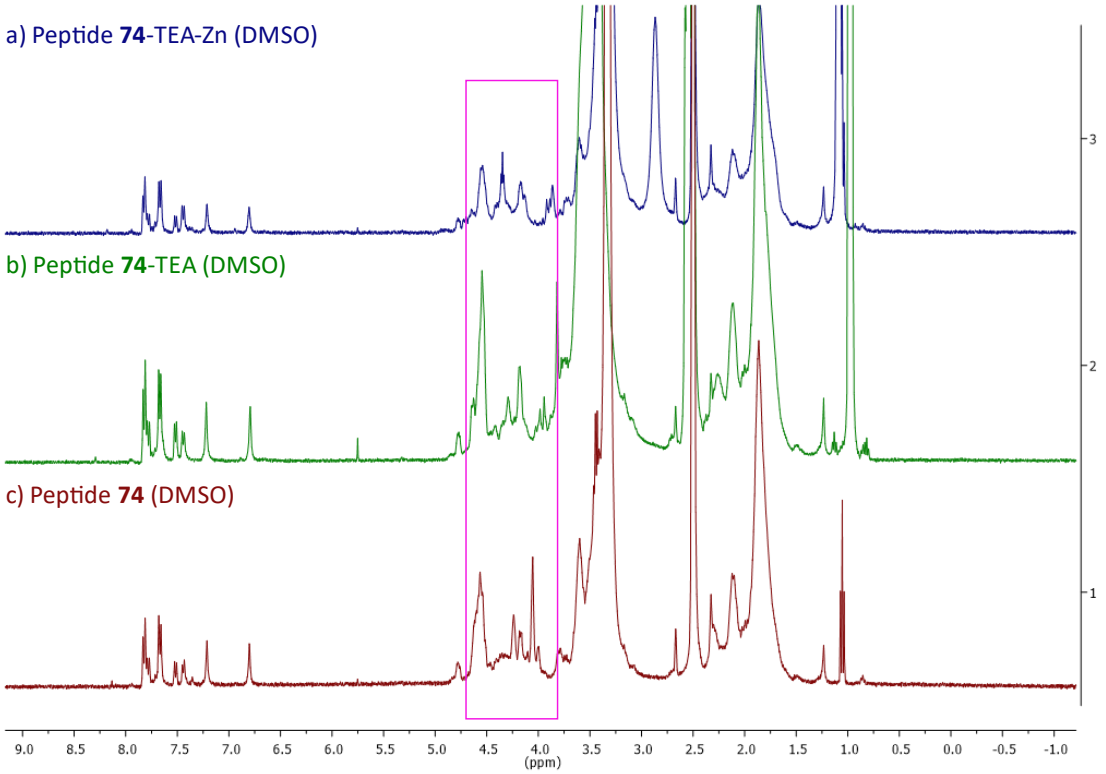


*Figure 105 – ^{1}H NMR Spectra of peptide **74** in DMSO-d_6 a) peptide with TEA (6 eq.) and zinc nitrate (3 eq.), b) peptide with TEA (6 eq.), c) peptide alone. Area with most significant differences highlighted in pink.*

Aromatic capping groups were used in the peptide synthesis of peptides **72** and **74** as the chemical shift of these protons will be well outside the range of the peptide backbone protons and, due to the size of these molecules, the ^{1}H NMR spectra are highly complex between 2-5 ppm making assignment of the protons and observing any changes significantly more difficult. However, the *N*-terminal capping group is situated at a significant distance from the carboxylic acid groups, as such, as can be seen in the ^{1}H NMR spectra (Figure 104-108), significant changes are not observed. Albeit the aromatic region appears to have an increased complexity, most evidently an increase in splitting at 7.48 ppm. Hence, the CH/CH_2 region was analysed as any diagnostic shift will be more evident due to the proximity to the Zn centre. Changes in ^{1}H spectra for the complexed peptide were observed in the 4.75-4.0 ppm region for the protons on the pyrrolidine ring adjacent to the functional groups. However, direct interpretation of any of these changes was not feasible due to the complexity of the signals in this region as such the extent of facial selectivity could not be determined upon mixing the two peptide complexes (Figure 106). As such from these results, although the structure could not be determined from NMR analysis, the evidence supports the formation of a peptide-Zn complex due to the characteristic changes seen in the ^{1}H NMR spectra of both peptides upon Zn addition.

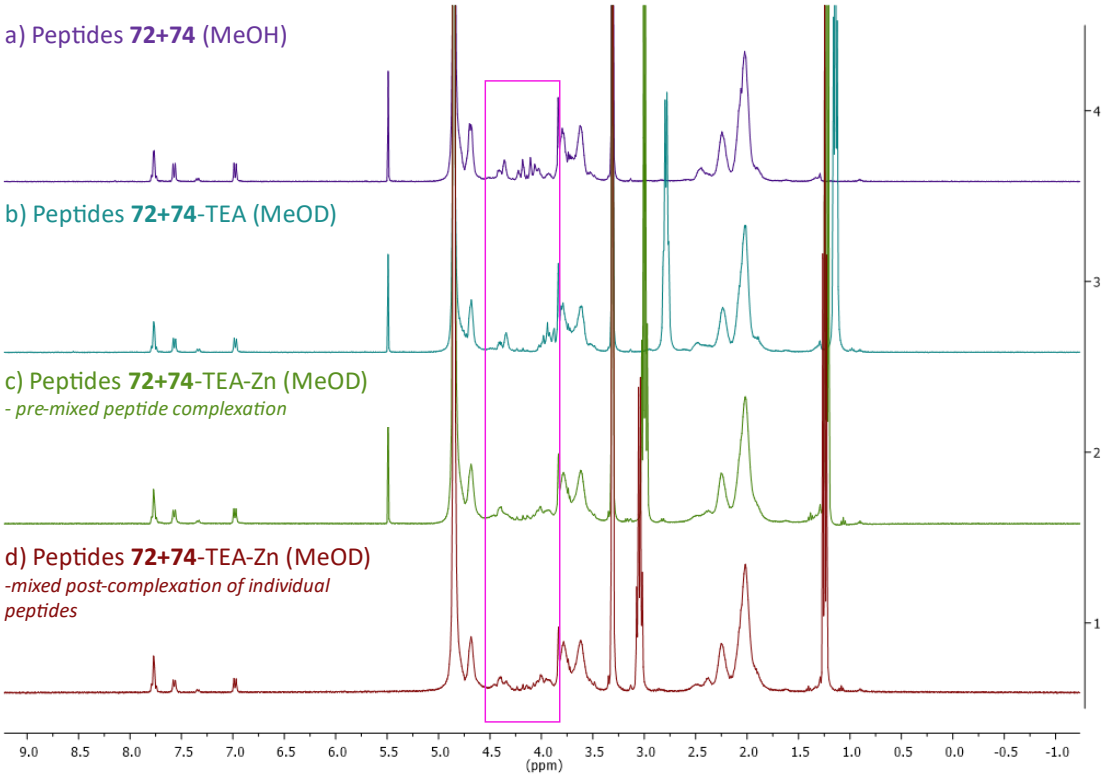


Figure 106 - ^{1}H NMR Spectra of samples of mixed peptides 72 and 74; a) peptide 72 and 74 in MeOD, b) peptide 72 (1 eq.) and 74 (1 eq.) with TEA (12 eq.) in MeOD, c) peptide 72 (1 eq.) and 74 (1 eq.) with TEA (12 eq.) and zinc nitrate (2 eq.) in MeOD, d) sample of peptides 72 and 74 complexed separately then mixed. Area with major differences highlighted in pink.

3.3.3.2 Synthesis of a Copper-peptide tetrameric cage

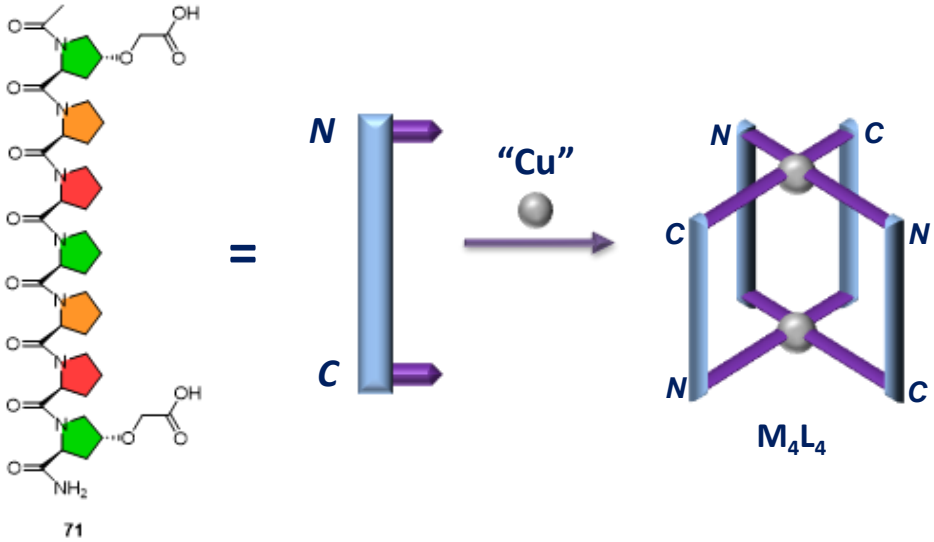


*Figure 107 – Theorised chemical structure of peptide **71** complexed with copper nitrate forming a tetramer with four linking copper centres, in this example the peptides are arranged in alternating antiparallel directions*

Subsequently, the peptide **71** was synthesised, in this case a 7-chain peptide was targeted as this is the shortest peptide with three residues on the same face of the polyproline II helix. Therefore, the terminal positions, both on the same face, could be functionalised such that the peptide acts a C-shaped ditopic ligand while the central residue (Pro4) is free to be functionalised allowing for the placement of a functional group into any cavity formed from complexation of the peptide. Thus the free residue can be utilised to tune host-guest interactions between the complexed peptide and any encapsulated molecule.

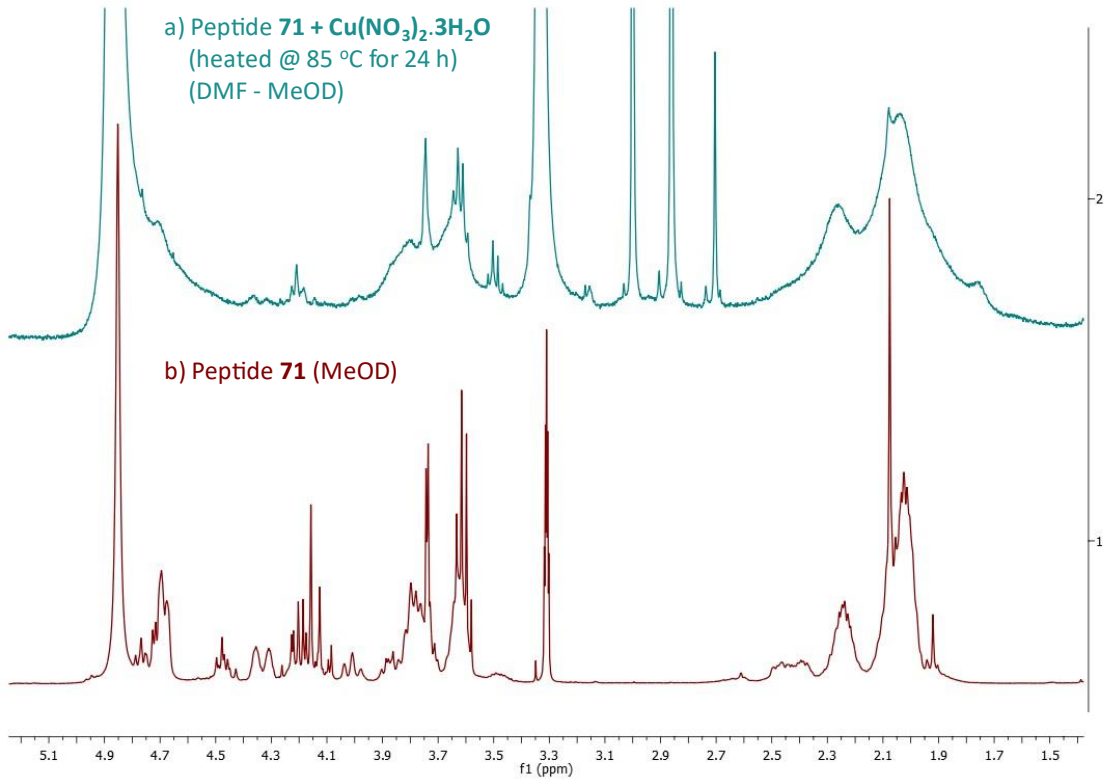


*Figure 108 – ^{1}H NMR spectra of peptide **71** in MeOD (b) and peptide **71** reacted with copper nitrate at 85 °C for 1 day in DMF, with a MeOD capillary (a)*

Initial synthesis of a peptide-copper coordination cage was attempted with peptide **71**. Copper nitrate is distorted octahedral so the complexation of this peptide with copper should form a tetrameric cage with four metal centres linking the carboxylic acid groups (Figure 107). The synthesis was carried out by heating the peptide and copper nitrate solution in DMF at 85 °C. The solution changed from the original light blue colour to dark green after 5 minutes indicating the successful reaction of the copper. No precipitate formed from this reaction unlike when using the peptides functionalised on all three faces, thus indicating the complex formed is a discrete assembly and is therefore a soluble complex. Repeated attempts were made to crystallise the formed discrete complex *via* vapour diffusion with poor solvents to determine the adopted structure *via* SC-XRD analysis, however, these were all unsuccessful only producing amorphous precipitate. Also, once precipitated the amorphous solid produced was then insoluble in DMF or EtOH, indicating that drying affects the solubility either by further complexation to form an extended network or that after exclusion of solvent from nanocavities these cannot be re-solubilised at room temperature. The solid was soluble in water, likely from disassociation of the metal complex. Subsequently, ^{1}H NMR analysis of the complex was carried out (Figure 108), as expected significant peak broadening was observed due to the presence of paramagnetic Cu (II) making determination of any diagnostic shifts difficult.

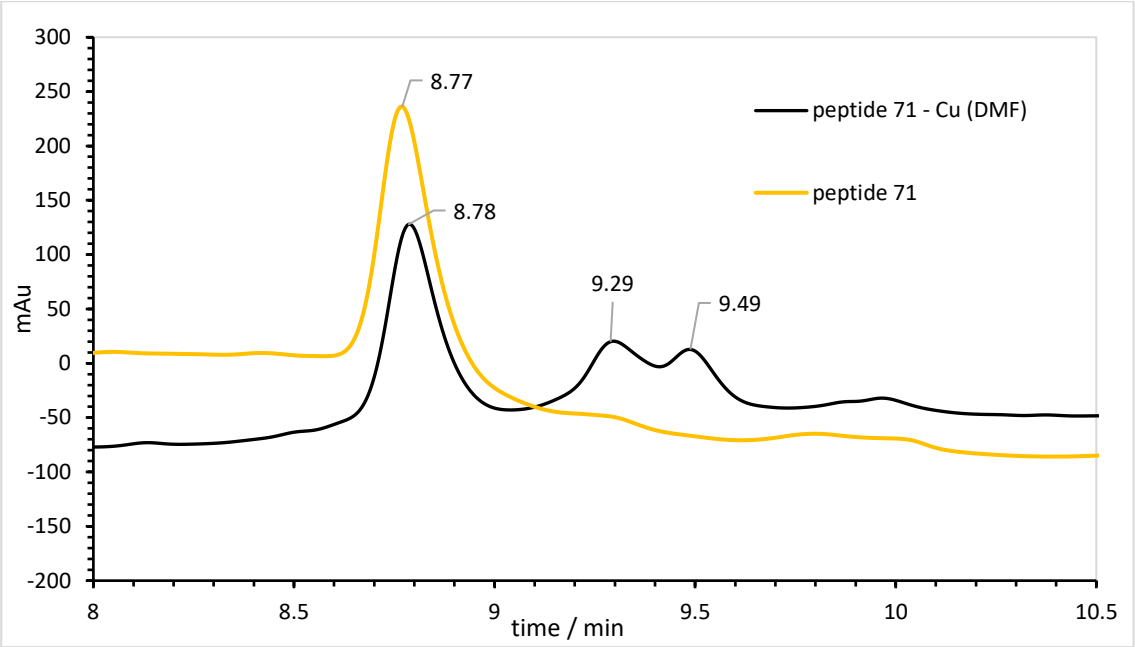


*Figure 109 - HPLC UV-Vis chromatogram of peptide **71** and peptide **71** reacted with copper nitrate, enlarged region between 8-10.5 min, (Full spectra - SI 4.10.2)*

To further determine the successful complexation of the peptide HPLC analysis was carried out on the complex solution, this showed residual unreacted peptide (rt = 8.78 min) while two small new peaks were visible (rt = 9.29-9.49 min) indicating the formation of a two new species from complexation. The large amount of residual starting material is to be expected from disassociation of complex in the aqueous media or from the decomposition of the copper complex due to the formic acid additive. For the determination of the new species HRMS analysis was carried however, the expected mass ($[M+4H]^{+4}$; 949.0789) was not found with the only identifiable mass matching the peptide ($[\mathbf{71}+Na]^{+}$; 909.404). Other larger mass peaks (e.g. 1094.467, *SI 4.10.2*), that were not present for the peptide alone, could not be identified as a reasonable fragment of the complex or peptide which indicated that an unexpected complex structure may have formed. This also suggests that the complex does not survive either contact with the aqueous media (0.1% formic acid) used in the analysis or the ionisation process, despite using the relatively soft electrospray ionisation (ESI). This is supported by the solubility of the dried complex in water and not DMF, with water causing disassociation of the copper-carboxylate bond. As such, without conclusive evidence of successful complexation from the analytical methods used, further investigation of this peptide complex was paused until further elucidation of the assembly principles of these materials could be carried out.

3.3.3.3 Synthesis of Discrete Peptide-palladium Coordination Cages

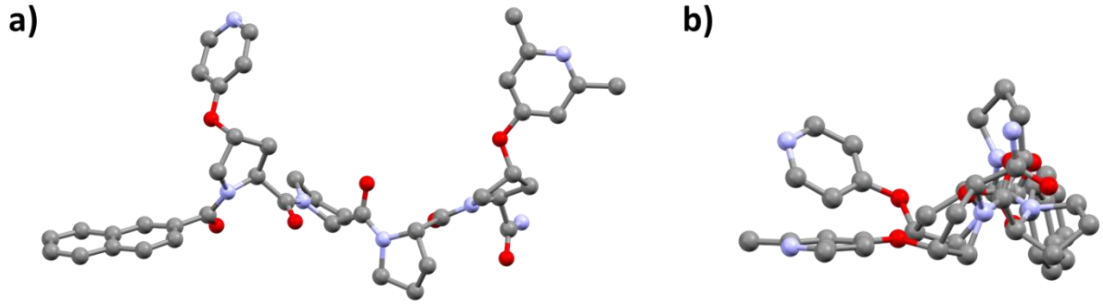
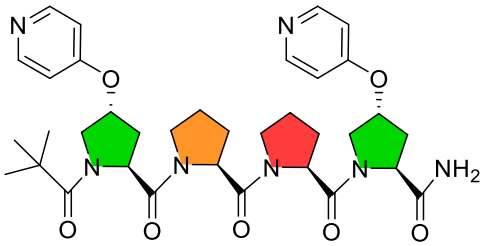


Figure 110 – Crystal structure of a proline tetramer capped with the 2-naphthoyl group at the N-terminus and functionalised with trans -O-Pyridine at the 4-position on Pro1 and trans -O-Lutidine at the 4-position on Pro4.[59]

To target a minimalistic ditopic peptide ligand for the synthesis of a peptide-metal nanocage the proline tetramer was again chosen as the ideal candidate. This peptide forms one full helical turn whereby the two terminal prolines can be functionalised for metal co-ordination. Notably within our group in as yet unpublished work a peptide tetramer, functionalised with lutidine and pyridine at the *N* and *C*-terminus respectively (*synthesised by Dr K. Samanta, University of Kent*), self-assembled to form a crystalline framework allowing for analysis *via* single crystal x-ray diffraction (*SCXRD data from Dr Lina Mardiana/Dr Michael J Hall; University of Newcastle*)[59] and thus allowing determination of the structure of the peptide helix and consequently the geometry of the functional groups (Figure 110). This data clearly showed that the polyproline II helix was retained for this structure despite significant functionalisation (50%) of the peptide backbone as in previous studies of the polyproline tetramer with hydroxyl functionalisation, again highlighting the resilience of the helix to functionalisation even at short peptide lengths. From this structure the two functional groups are located on the same helical face and with this data in hand the synthesis of a peptide functionalised with two coordinating functional groups to act as a ditopic C-shaped ligand for the synthesis of metal-peptide coordination cages could be easily envisioned.



(75)

Figure 111 – Chemical structure of peptide 75, proline rings are highlighted to show each of their locations on the three helical faces (i.e. functionalised prolines 1 and 4 are on the same face)

As such the pyridine functionality was incorporated into a peptide sequence with the 4-position functionalised monomer **66** at both termini forming a C-shaped ditopic ligand with two pyridine motifs, which are commonly used to coordinate with metal cations to form coordination complexes.[49] Peptide **75** (Figure 111) was synthesised using standard SPPS techniques in a quantitative yield. The resin bound peptide had to be treated with methanol to remove the pivalic capping group from the pyridines and the as synthesised peptide formed the TFA-pyridinium salt upon resin cleavage. Subsequent to purification of the peptide *via* RP-prepHPLC the pyridines exist as the pyridinium formate salt of the product. As such, the peptide was

treated with either a basic ion-exchange (*Purolite® A300*) resin or Si-CO_3 (0.5 mmol/g) to remove the acid and ensure the pyridines exist as the free amines, which was confirmed by either ^{1}H or ^{19}F NMR. The complete removal of formic acid was essential as formic acid is a strong reducing agent and will reduce metals such as palladium salts altering the molar ratio of the reactants.[60]

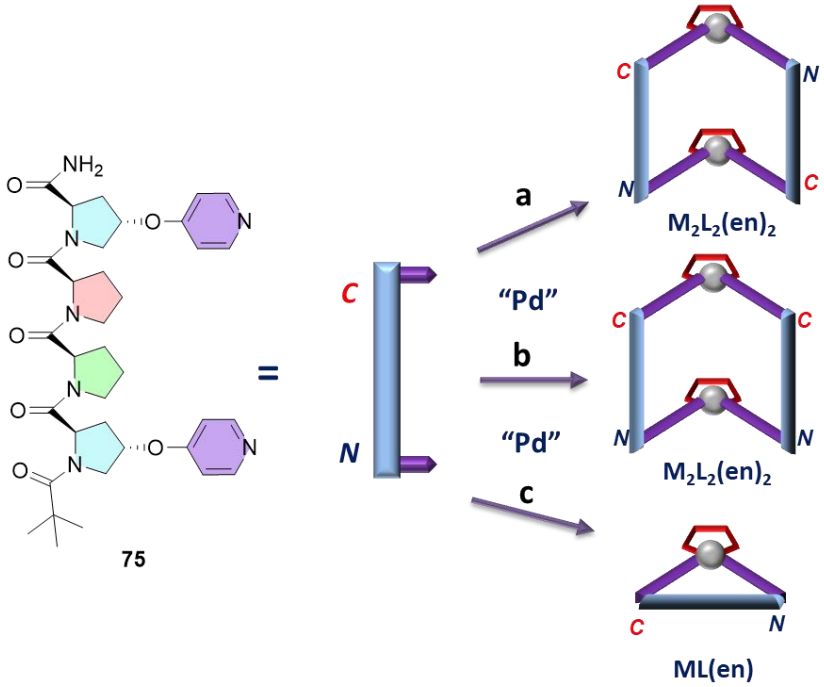


*Figure 112 – Expected chemical structures of peptide **75**-Pd(en) complexes; a) Peptides in anti-parallel conformation, b) peptides in parallel conformation, (alignment indicated by arrows); and the actual experimental structure; c) Peptide **75**-Pd monochelate complex where the pyridines coordinate from one peptide to a single Pd ion.*

With this peptide in hand it could then be used for the synthesis of discrete co-ordination complexes. For these initial trials palladium was targeted as an excellent candidate for the formation of pyridinium complexes, with the potential to form both dimeric and tetrameric cage-like structures (Figure 112) with the peptide pyridines coordinating to palladium ions in a square planar geometry. This synthesis was achieved using the palladium salts $Pd(en)(NO_3)_2$, for the attempted peptide dimer synthesis, and $Pd(CH_3CN)_4(BF_4)_2$ for the attempted synthesis of a peptide tetramer. The Pd(en) salt retains the ethylenediamine ligand which leaves only two coordination sites free on the palladium centre for coordination of the peptide pyridines (Figure 112a). The complexation was carried out *via* heating the peptide with each palladium salt to 65 ^{o}C for at least 2 h in either deuterated water or DMSO. As such, the as formed complex could be analysed *via* NMR to assist in determination of the product formation and the adopted conformation.

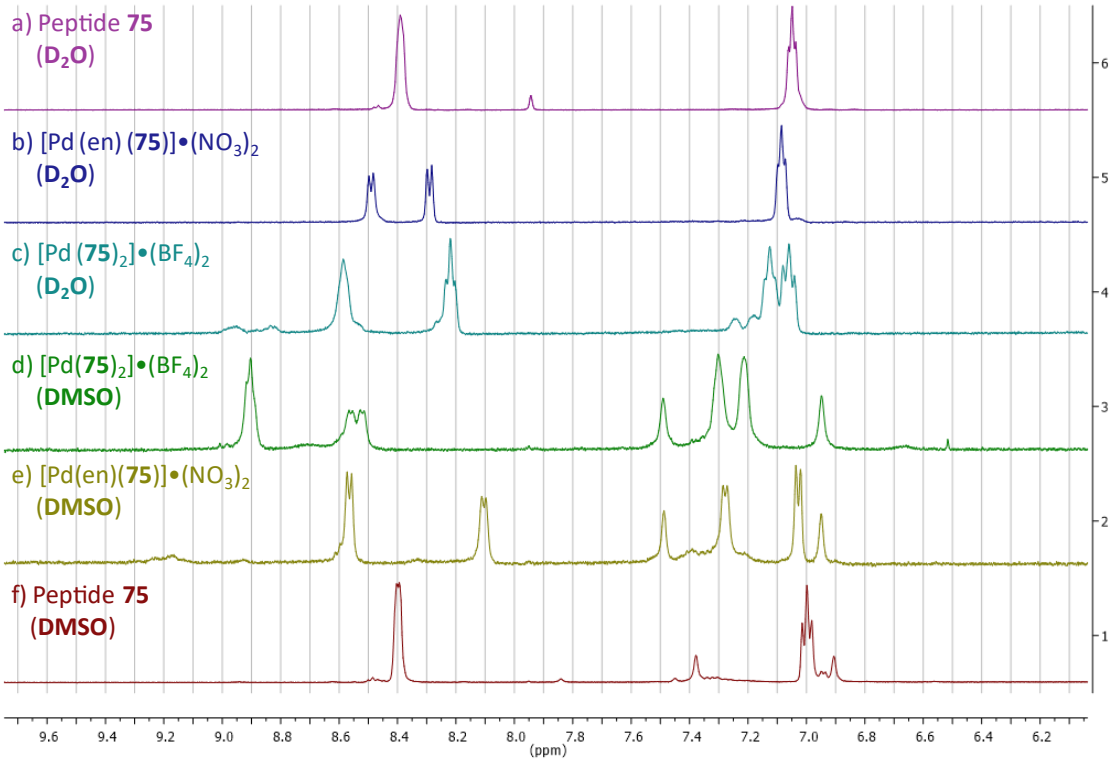


Figure 113 – ^{1}H NMR spectra in DMSO-d_6 (d-f) and D_2O (a-c) of peptide 75, enlarged aromatic proton region; a) peptide alone, b) complex with $Pd(en)(NO_3)_2$, c-d) complex with $Pd(CH_3CN)_4(BF_4)_2$, e) complex with $Pd(en)(NO_3)_2$, f) peptide alone.

Analysis of the complex formed with $Pd(en)(NO_3)_2$ in both DMSO and water *via* ^{1}H NMR showed the quantitative formation of a peptide complex, most evidently the splitting of the pyridine aromatic protons increases as the pyridines are no longer equivalent, and a downfield shift occurs due to de-shielding from the adjacent palladium (Figure 113). However, while NMR analysis confirmed that complete complexation had occurred, with no residual starting material, it was unclear which of the predicted structures had been adopted. With the capability of the peptides to form either a parallel complex (Figure 112b), antiparallel complex (Figure 112a), a mixture of both conformations, or act as a monochelating ligand (Figure 112c) it was unclear what the selectivity of this assembly process would be. However, the splitting of these aromatic protons was clearly indicative of the formation of a single major product due to the lack of complexity of the splitting. The formation a mixture of different complexes would form significantly more proton environments than the four seen in the ^{1}H NMR spectrum. To further determine the geometry of the peptides within the complex 2D NMR was utilised. *via* both 2D-ROESY and 2D-NOESY NMR the interactions through-space between protons, rather than through bond interactions, can be observed to determine the conformation.[61]

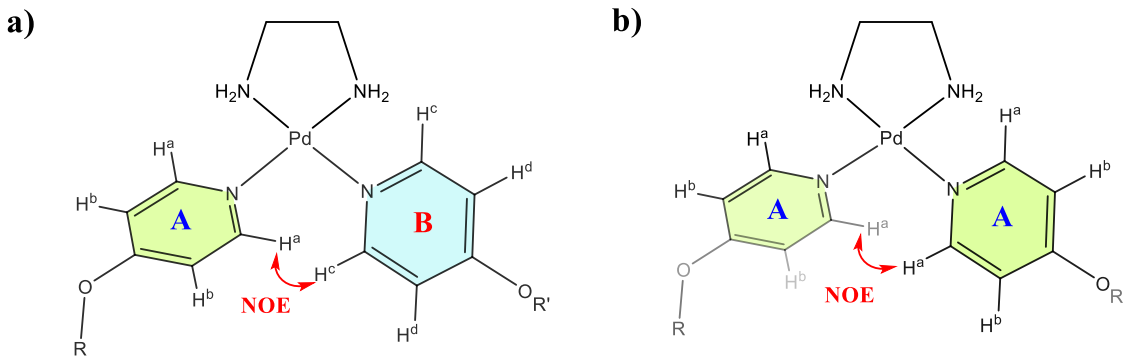


Figure 114 – a) Chemical structure of non-equivalent N-terminus and C-terminus pyridines coordinating to a single palladium ion (b) Chemical structure of equivalent pyridines coordinating to a single palladium, examples of the NOE interactions are shown between H^a or H^a-H^c protons.

From 2D NOESY and ROESY NMR of both complexes NOE interactions should be apparent as cross peaks in the 2D spectra between the non-equivalent *ortho* ($H^{a/c}$)protons of each pyridine (pyridine A and B, Figure 114) if the adopted structure is in either the antiparallel dimer structure or in the monochelate structure, both having C-terminal and *N*-terminal pyridines coordinating to the same palladium ion. While if in the parallel dimer conformation no NOE cross peaks should be seen, as the NOE interaction would be with the proton in the same chemical environment on the symmetrical pyridine motif (Figure 114b).

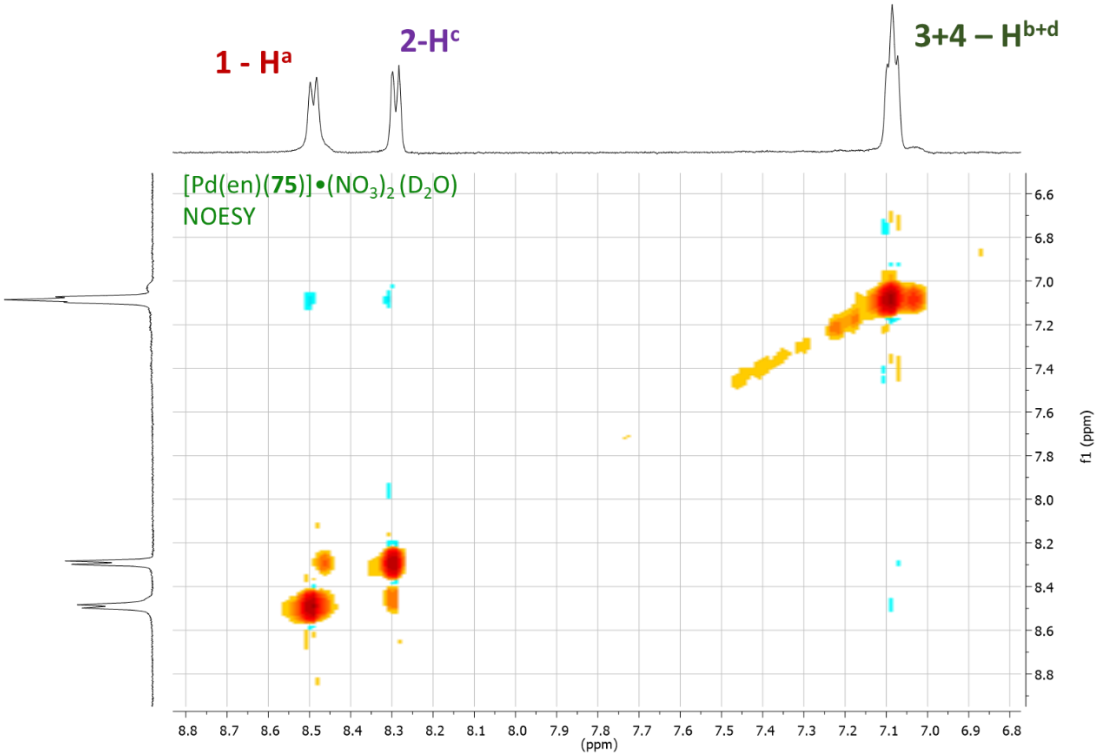


*Figure 115 – ^{1}H-^{1}H 2D NOESY NMR of peptide complex [Pd(en)(**75**)]•$(NO_3)_2$ in D_2O*

Analysing the aromatic region of the 2D NOESY and ROESY NMR of the complex in DMSO and D_2O it is evident that there are cross peaks between a shoulder on the proton region (1 – H^a) on the *ortho* protons of the pyridine at the C-terminus (A) and a region under the *ortho* protons (2-H^c, Figure 115) on the pyridine at the C-terminus (B) (Figure 114b) indicating they are close in space. However, these cross peaks do not properly align with the major peaks and the phasing is different to the H^{a-b} and H^{c-d} cross peaks and are likely from a minor product, possibly the anti-

parallel dimer complex. Further to this, to confirm whether the dimer or monomeric complex had been formed, HRMS analysis of the sample was carried out. These results indicated that the desired cage-like structure was not adopted. While the expected mass of 420.6571 was achieved, the charge of this ion (M^{2+}), as well as the isotopic pattern, indicated that the mass of the complex was half the desired mass, for the dimeric cage complex ($[Pd_2(en)_2L_2]^{4+}$), therefore instead of chelating to two palladium ions the peptide is instead acting as a monochelate ligand ($[Pd(en)L]^{2+}$, Figure 112c). This can only occur due to the added length from the pyridine functional groups with the ether linkage allowing free rotation such that both pyridines can adopt a geometry sufficient to coordinate to a single palladium ion without significant disruption of the polyproline II helix. This result matches with the sharp peaks seen *via* NMR analysis, where a degree of peak broadening would be expected from a more flexible cage-like structure. Due to this finding some of the conditions were altered to see if the formation of an alternate complex could be observed. Initially the reaction was carried out at 0 ^{O}C in D_2O to see if the high temperatures (65 ^{o}C) used previously were disrupting the polyproline helix significantly to allow this type of chelation, however, this experiment gave the same product. The concentration of the reaction was also increased (20 mg mL^{-1}) to see if intermolecular linkages could be promoted over the intramolecular bonding, however, this also had no effect with intramolecular binding to an adjacent pyridine on the same peptide clearly significantly more favourable.

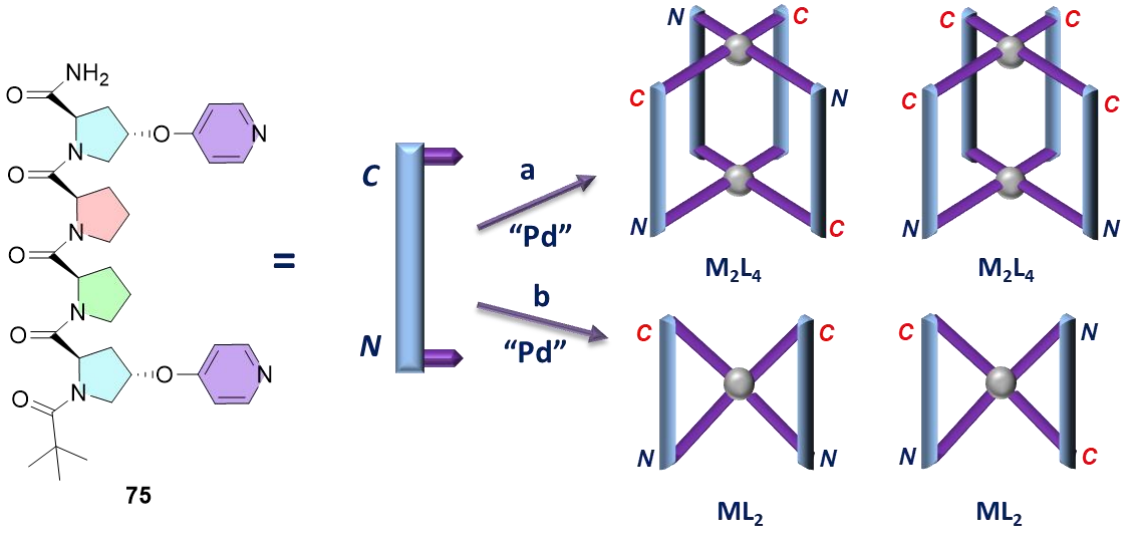


*Figure 116 - Chemical structures of potential peptide **75** $Pd(ACN)_2(BF_4)_2$ complexes; a) Peptide 75-Pd tetramer complexes with two metal centres linking four peptides via pyridine coordination with peptides in alternating anti-parallel arrangement and in the parallel conformation; b) Peptide 75-Pd complex where a single Pd ion acts as a metal centre with four coordinating pyridines from two peptides, with both peptides aligned in parallel and antiparallel. 3D-structure is not representative of the actual conformation of the peptide complexes.*

In contrast to the Pd(en) complex, complexation with $Pd(ACN)_2(BF_4)_2$, with four coordinating pyridines, can have several accessible structures from the alignment of the peptides in the complex. With the formation of the tetrameric complex if the peptides are not selective numerous combinations of peptide geometries are possible, while even if the peptide acts as monochelating ligand as seen for the Pd(en) peptide complex this can either have the peptides parallel or anti-parallel. Analysis of the ^{1}H NMR of the complex in both solvents show the formation of one major product (Figure 117) with the same splitting and downfield shift as seen previously for the Pd(en) complex, where the difference in chemical shift between the non-equivalent protons on the pyridines increases. Small impurities can be seen with a similar splitting pattern and a slight downfield shift, more evident in D_2O, which likely correspond to the alternate isomer of the complex (Figure 117). The 2D NMR analysis *via* NOESY/ROESY NMR

again shows no NOE interactions between the *ortho* protons of the *N*- and C-terminus pyridine moieties suggesting adoption of the parallel confirmation. However, cross peaks are evident for the minor product peaks in D_2O which may indicate the *ortho* protons are adjacent to the opposite terminus pyridine, therefore this minor product complex is likely the anti-parallel conformation or the cage-tetrameric product.

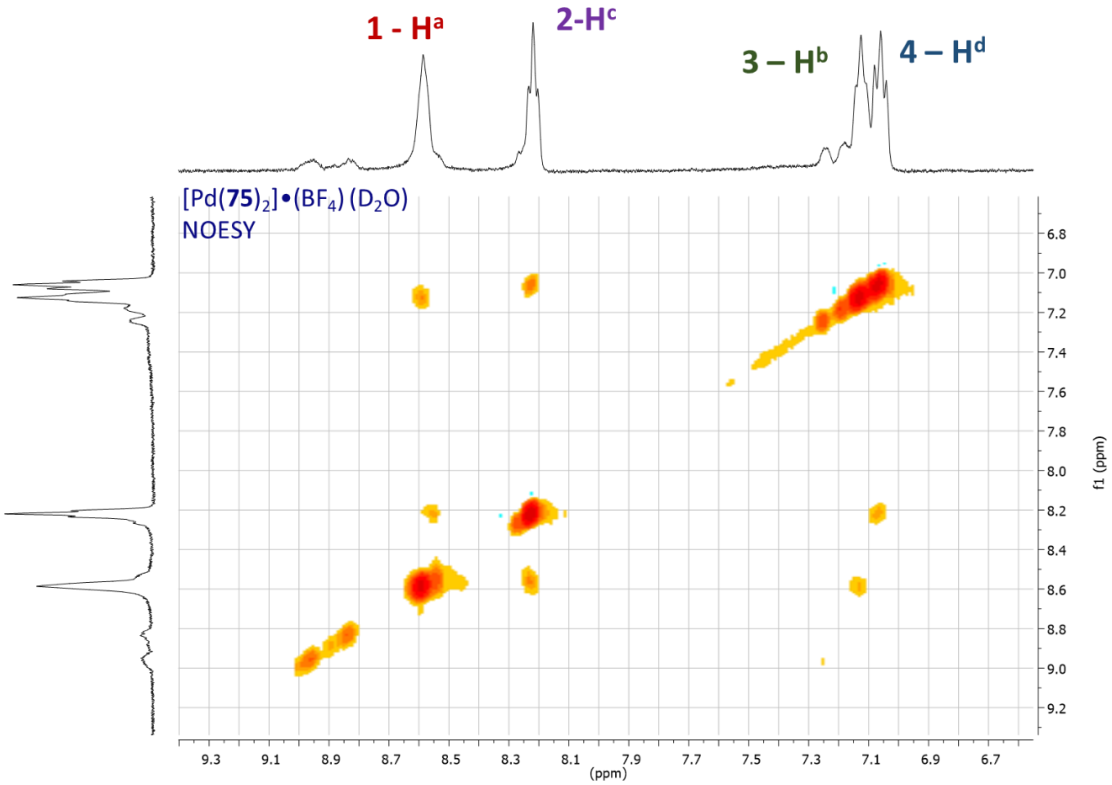


*Figure 117– ^{1}H-^{1}H 2D NOESY NMR of peptide complex [Pd (**75**)$_2$]•(BF$_4$)$_2$ in D$_2$O*

Subsequent, HRMS analysis was again carried out to indicate whether either the cage like structure or the "bowtie" like structure, with monochelating peptides, had been adopted. HRMS analysis again shows the formation of the monochelating complex (dimer) with a mass of 728.7911 (M^{+2}). The 2+ charge from HRMS is diagnostic for this complex as the predicted tetrameric structure will give the same mass but with a +4 charge, determined from the isotopic pattern of the ion peak. Therefore, the desired cage structure was not formed in this case, instead again forming a "bowtie" like complex with two peptides chelating to a single palladium (Figure 116c-d).

This shows that the incorporated pyridine functional group on the same face of the polyproline II helix can be effectively used to synthesise a metal-peptide complex with the peptides linked by metal coordination bonds, as seen for the [Pd (**75**)$_2$]•(BF$_4$)$_2$ complex. However, the functional groups used were either too long and flexible or placed too close together to form the desired intermolecular linkages as these factors allowed the peptide to act as a monochelating ligand, which is significantly more favourable over forming an intermolecular interaction to a new peptide especially at relatively low concentrations. The impact from the longer sidechain is exemplified in the previous chapter where the synthesis of organophosphate linked peptides from hydroxyproline did not form any such intramolecular bonds, due to the significantly shorter single oxygen-phosphorus bond, instead forming intermolecular bonds between two peptides with a phosphate linker (2.2.3.3 - *Attempted Synthesis of Covalent Peptide Cages*). As such, the optimisation of this peptide to achieve the desired metal-peptide cage with a nanocavity is clearly feasible. This also, has implications for the previous section using carboxylic acid side

chains which are almost as long as the pyridine group and more flexible, therefore the closely placed functional groups also have the potential to chelate with a neighbouring group to a single metal ion. Further investigation of these assemblies is therefore required and optimisation of the peptide building-blocks to allow for rational design of the supramolecular interactions. Also, recent evidence from an non-peer-reviewed preprint from the McTernan group shows the promising potential of these types of ligands as supramolecular building-blocks with the indication of the formation of cage-like structures using very similar peptide ligands only differing by the means used to attach the pyridine group to the amino acids, with ester rather than ether linkages.[62] Furthermore, several routes to optimise the peptides to prevent the chelation seen in this work are apparent, with the most evident being using a shorter, less flexible, functional motif with direct C-C coupling of a pyridine to the proline ring during amino acid synthesis, which has already been carried out using a phenolic acid sidechain previously. Also, using a sterically hindered coordinating group, such as lutidine to hinder this type of chelation, or using a longer peptide with an increased distance between the binding groups to prevent the chelation of both pyridines to the same metal ion are viable routes.

This data clearly shows how short oligoproline helices can be utilised as ligands for the synthesis of discrete chiral metal complexes with the capability for rational design from the predictable secondary structure of the polyproline II helix. We clearly evidence how previous studies on the proline tetramer provide a model from which a peptide can be designed and specifically functionalised with motifs for coordination interactions while retaining the polyproline II helix to achieve a predicted metal-peptide complex. Despite the further investigation required into these materials this undoubtedly has great potential for the rational design of nanoreactors to achieve specific host-guest interactions for enantioselective catalysis.

3.4 Conclusion

In conclusion, we have utilised polyproline helices to synthesise three different types of supramolecular constructs. By functionalising a proline monomer with a carboxylic acid sidechain we were able to incorporate this moiety into a series of different peptide sequences. Working from the polyproline secondary structure with a chemical model from the crystal structure of the previously synthesised proline tetramers we were able to place these groups at specific locations within three-dimensional space on the peptide backbone, functionalising different faces of the helix, and with varying degrees of functionalisation, to directly influence the supramolecular assembly of these peptide building blocks. With the added carboxylic acid sidechains as handles for supramolecular assembly on all three faces of the helix we were able to synthesise extended metal-peptide assemblies. These constructs then assembled into a variety of topologies depending on the solvent environment, whereby, the polyproline helix switches between the extended polyproline II helix, in either water or methanol, and the shorter polyproline I helix, in propan-1-ol or ethanol, as shown by circular dichroism spectroscopy, and the adopted helix altered the adopted topology of the metal-organic assemblies. By designing the length of the peptides and changing the placement of the functional groups we then also demonstrated that these aspects could be used to alter the topology of the assembly. From the peptides we were able to synthesise a range of assemblies; gels, fibres, chain-like aggregates, clusters, and notably nanoparticles, with nanocavities that have potential applications in enantioselective catalysis and molecular *trans*port.

To demonstrate the rationale behind the design of the peptide sequences and further investigate the assembly principles of the functionalised peptides a series of peptides functionalised on a single face were synthesized to match with each face of the peptide that was

designed to be “facially selective” during the synthesis of the extended metal-peptide frameworks. With these polyprolines discrete metal-peptide complexes were synthesised which remained in solution. However, due to the nature of the zinc carboxylate coordinate bond we were not able to demonstrate whether the spacing of the linker group could be used to control facial selectivity by mixing two peptide complexes with different carboxylic acid sidechain spacings, with the complexes not surviving the required analytical methods and NMR could not conclusively show the difference due to the complexity of the relevant region of the spectra despite indicating successful complexation.

In further attempts to synthesise nanocavities from discrete polyproline-metal cages a pyridine functionalised proline tetramer was synthesised with the functional groups on the same face at either terminus to act as a C-shaped ditopic ligand. We successfully formed metal complexes using palladium salts and by controlling the free coordination sites of the metal centre the adopted structure could be predicted. However, rather than forming the desired cage-like complexes the peptide monochelated to a single palladium ion due to the length and flexibility of the functional group. Thus, a peptide-palladium monomer and a dimer complex were both synthesised, with the dimer expressing good selectivity toward forming a single major product driven by the chirality of the peptide helix. Despite not synthesising the desired cage structure this clearly showed the potential of this functionalised peptide to form predictable metal complexes and the optimisation of the peptide building block to achieve the desired structure can be easily envisioned. This will be achieved by either shortening the functional group by direct C-C coupling of the pyridine to the proline ring, reducing the flexibility and length, or lengthening the peptide and increasing the spacing between the functional groups to prevent the type of chelation seen here. Therefore, a highly tuneable, chiral, nanocavity can be synthesised *via* peptide-metal complexation with promising applications as a nanoreactor for enantioselective catalysis or molecular *trans*port.

3.5 References

1 M. Fujita, J. Yazaki and K. Ogura, *J. Am. Chem. Soc.*, 1990, **112**, 5645–5647.

2 S. Leininger, B. Olenyuk and P. J. Stang, *Chem. Rev.*, 2000, **100**, 853–907.

3 M. Raynal, P. Ballester, A. Vidal-Ferran and P. W. N. M. Van Leeuwen, *Chem. Soc. Rev*, 2014, **43**, 1734.

4 A. Palma, M. Artelsmair, G. Wu, X. Lu, S. J. Barrow, N. Uddin, E. Rosta, E. Masson and O. A. Scherman, *Angew. Chemie*, 2017, **129**, 15894–15898.

5 M. Yoshizawa, J. K. Klosterman and M. Fujita, *Angew. Chemie Int. Ed.*, 2009, **48**, 3418–3438.

6 T. Sawada, A. Matsumoto and M. Fujita, *Angew. Chemie*, 2014, **126**, 7356–7360.

7 T. R. Cook, Y. R. Zheng and P. J. Stang, *Chem. Rev.*, 2013, **113**, 734–777.

8 C. T. McTernan, J. A. Davies and J. R. Nitschke, *Chem. Rev.*, 2022, **122**, 10393–10437.

9 A. M. Castilla, W. J. Ramsay and J. R. Nitschke, *Acc. Chem. Res.*, 2014, **47**, 2063–2073.

10 D. L. Caulder and K. N. Raymond, *J. Chem. Soc*, 1999, 1185–1200.

11 T. K. Piskorz, V. Martí-Centelles, T. A. Young, P. J. Lusby and F. Duarte, *ACS Catal.*, 2022, **12**, 5806–5826.

12 R. Breslow, *Acc. Chem. Res.*, 1995, **28**, 146–153.

13 D. E. Koshland, *Angew. Chemie Int. Ed. English*, 1995, **33**, 2375–2378.

14 T. Sawada, H. Hisada and M. Fujita, *J. Am. Chem. Soc.*, 2014, **136**, 4449–4451.

15 W. M. Bloch and G. H. Clever, *Chem. Commun.*, 2017, **53**, 8506–8516.

16 C. Tan, J. Jiao, Z. Li, Y. Liu, X. Han and Y. Cui, *Angew. Chemie Int. Ed.*, 2018, **57**, 2085–2090.

17 J. Jiao, C. Tan, Z. Li, Y. Liu, X. Han and Y. Cui, *J. Am. Chem. Soc.*, 2018, **140**, 2251–2259.

18 C. Li, K. Deng, Z. Tang and L. Jiang, *J. Am. Chem. Soc.*, 2010, **132**, 8202–8209.

19 H. S. Wang, Y. H. Wang and Y. Ding, *Nanoscale Adv.*, 2020, **2**, 3788–3797.

20 H. Cai, Y. L. Huang and D. Li, *Coord. Chem. Rev.*, 2019, **378**, 207–221.

21 B. Sun, M. Bilal, S. Jia, Y. Jiang and J. Cui, *Korean J. Chem. Eng*, 2019, **36**, 1949–1964.

22 S. Chand, O. Alahmed, W. S. Baslyman, A. Dey, S. Qutub, R. Saha, Y. Hijikata, M. Alaamery and N. M. Khashab, *JACS Au*, 2022, **2**, 623–630.

23 V. Subramaniyam, P. V. Ravi and M. Pichumani, *J. Mol. Struct.*, 2022, **1251**, 131931.

24 T. Sawada, A. Matsumoto and M. Fujita, *Angew. Chemie*, 2014, **126**, 7356–7360.

25 A. Saito, T. Sawada and M. Fujita, *Angew. Chemie Int. Ed.*, 2020, **59**, 20367–20370.

26 Y. Inomata, T. Sawada and M. Fujita, *J. Am. Chem. Soc.*, 2021, **143**, 16734–16739.

27 T. Sawada and M. Fujita, *Bull. Chem. Soc. Jpn.*, 2021, **94**, 2342–2350.

28 T. Sawada, A. Saito, K. Tamiya, K. Shimokawa, Y. Hisada and M. Fujita, *Nat. Commun. 2019 101*, 2019, **10**, 1–7.

29 J. Navarro-Sánchez, A. I. Argente-García, Y. Moliner-Martínez, D. Roca-Sanjuán, D. Antypov, P. Campíns-Falcó, M. J. Rosseinsky and C. Martí-Gastaldo, *J. Am. Chem. Soc.*, 2017, **139**, 4294–4297.

30 F. G. Cirujano, N. Martín, N. Almora-Barrios and C. Martí-Gastaldo, *Catal. Sci. Technol.*, 2021, **11**, 6053–6057.

31 P. Morales and M. A. Jiménez, *Arch. Biochem. Biophys.*, 2019, **661**, 149–167.

32 Z. Shi, K. Chen, Z. Liu and N. R. Kallenbach, *Chem. Rev.*, 2006, **106**, 1877–1897.

33 A. L. Rucker and T. P. Creamer, *Protein Sci.*, 2002, **11**, 980–985.

34 S. Dobitz, M. R. Aronoff and H. Wennemers, *Acc. Chem. Res.*, 2017, **50**, 2420–2428.

35 E. Sobakinskaya, M. Schmidt am Busch and T. Renger, *J. Phys. Chem. B*, 2018, **122**, 54–67.

36 P. Wilhelm, B. Lewandowski, N. Trapp and H. Wennemers, *J. Am. Chem. Soc.*, 2014, **136**, 15829–15832.

37 G. Berger, M. Vilchis-Reyes and S. Hanessian, *Angew. Chemie Int. Ed.*, 2015, **54**, 13268–13272.

38 D. F. Brightwell, G. Truccolo, K. Samanta, E. J. Fenn, S. J. Holder, H. J. Shepherd, C. S. Hawes and A. Palma, *Chem. – A Eur. J.*, 2022, **28**, e202202368.

39 D. F. Brightwell, G. Truccolo, K. Samanta, H. J. Shepherd and A. Palma, *ACS Macro Lett.*, 2023, **12**, 908–914.

40 T. Schnitzer, E. Paenurk, N. Trapp, R. Gershoni-Poranne and H. Wennemers, *J. Am. Chem. Soc.*, 2021, **143**, 644–648.

41 S. L. Heinz-Kunert, A. Pandya, V. T. Dang, P. N. Tran, S. Ghosh, D. McElheny, B. D. Santarsiero, Z. Ren and A. I. Nguyen, *J. Am. Chem. Soc.*, 2022, **144**, 7001–7009.

42 V. Mihali, F. Foschi, M. Penso and G. Pozzi, *European J. Org. Chem.*, 2014, **2014**, 5351–5355.

43 T. Tsunoda, Y. Yamamiya and S. Itô, *Tetrahedron Lett.*, 1993, **34**, 1639–1642.

44 D. L. Hughes, *Org. React.*, 1992, 335–656.

45 D. L. Hughes, *Org. Prep. Proced. Inc.*, 2009, **28**, 127–164.

46 S. D. Lepore and Y. He, *J. Org. Chem.*, 2003, **68**, 8261–8263.

47 A. K. Pandey, D. Naduthambi, K. M. Thomas and N. J. Zondlo, *J. Am. Chem. Soc.*, 2013, **135**, 4333–4363.

48 J. Krapcho, C. Turk, D. W. Cushman, J. R. Powell, J. M. DeForrest, E. R. Spitzmiller, D. S. Karanewsky, M. Duggan, G. Rovnvak, J. Schwartz, S. Natarajan, J. D. Godfrey, D. E. Ryono, R. Neubeck, K. S. Atwal and E. W. Petrillo, *J. Med. Chem.*, 1988, **31**, 1148–1160.

49 A. Schoedel, M. Li, D. Li, M. O'Keeffe and O. M. Yaghi, *Chem. Rev.*, 2016, **116**, 12466–12535.

50 A. Lemmerer, S. Govindraju, M. Johnston, X. Motloung and K. L. Savig, *CrystEngComm*, 2015, **17**, 3591–3595.

51 J. Marsh and A. Palma, *Unpublished Work - Synthesis of Polyproline-Zinc Nanoparticles*, 2018.

52 S. Kakinoki, Y. Hirano and M. Oka, *Polym. Bull.*, 2005, **53**, 109–115.

53 X. Chen, Y. Wang, H. Wang, Y. Kim and M. Lee, *Chem. Commun.*, 2017, **53**, 10958–10961.

54 S. Bhattacharjee, *J. Control. Release*, 2016, **235**, 337–351.

55 S. L, H. AE, G. MS, E. MA, R. DH and C. DE, *J. Am. Soc. Mass Spectrom.*, 2016, **27**, 22–30.

56 J. M. Fletcher, R. L. Harniman, F. R. H. Barnes, A. L. Boyle, A. Collins, J. Mantell, T. H. Sharp, M. Antognozzi, P. J. Booth, N. Linden, M. J. Miles, R. B. Sessions, P. Verkade and D. N. Woolfson, *Science (80-.).*, 2013, **340**, 595–599.

57 N. J. Zondlo, *Phys. Chem. Chem. Phys.*, 2022, **24**, 13571–13586.

58 T. J. El-Baba, D. R. Fuller, D. A. Hales, D. H. Russell and D. E. Clemmer, *J. Am. Soc. Mass Spectrom.*, 2019, **30**, 77.

59 K. Samanta, M. J. Hall, L. Mardiana and A. Palma, *Unpublished Work*, 2023.

60 Q. Wang, Y. Wang, P. Guo, Q. Li, R. Ding, B. Wang, H. Li, J. Liu and X. S. Zhao, *Langmuir*, 2014, **30**, 440–446.

61 D. Neuhaus, *Encycl. Magn. Reson.*, 2011, DOI:10.1002/9780470034590.EMRSTM0350.PUB2.

62 B. E. Barber, E. M. G. Jamieson, L. E. M. White and C. T. Mcternan, *Metal-Peptidic Cages-Helical Oligoprolines Generate Highly Anisotropic Nanospaces with Emergent Isomer Control*, 2023.

.

Chapter 4.

Experimental data (SI)

Chapter 4 - *Experimental (CH.2) – SI 4.1-4.7*

Materials *– All solvents were used as procured from Fischer Scientific except dry THF, which was dried by distillation using sodium wire in inert atmosphere and dry DCM, which was dried over activated molecular sieves under an inert atmosphere. For all LCMS and HPLC purposes, HPLC grade solvents were procured from Fischer Scientific and used as such. Thin layer chromatography was conducted using plates (silica gel, 250 µm, 60 Å, F254). Flash chromatography was performed using 230-400 mesh (40-63 µm, 60 Å) silica gel from Merck. Deuterated methanol (CD_3OD), chloroform ($CDCl_3$), water (D_2O) and dimethyl sulfoxide ($(CD_3)_2SO$) were purchased from Cambridge Isotope Laboratories. Rink Amide MBHA resin (100-200 mesh, 0.3 mmol/g) 1% DVB, Fmoc-L-amino acids, N,N-Diisopropylethylamine (DIPEA), acetic anhydride (Acac) and benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) were obtained from Fluorochem Ltd (Derbyshire, UK).*

LCMS - Compounds were separated *via* RP-HPLC using a HiChrom KR100 5C18 5263 column at 40 °C on a Dionex UltiMate 3000. Gradient: 5% B for 5 minutes then from 5% B to 100% B over 20 minutes, and held at 100% B for 5 minutes. Where A is Water (0.1% formic acid) and B is methanol (0.1% formic acid). Flow rate is 1.0 mL/min. Wavelength: 225 nm. The flow was directed into directed into the electrospray source of a Thermo Scientific MSQ Plus Mass Detector, operating in positive ion mode, at 75 kV and mass spectra recorded from 100-2000 m/z. Sample specific details are highlighted in the relevant text.

NMR spectroscopy - NMR spectra were collected on Bruker Avance II 400 MHz spectrometer, using Topspin software. 2D NMR experiments were measured on the Bruker Avance III 400 MHz at 298 K, using Topspin software. Chemical shifts are with reference to the residual solvent peak, with J values in Hz. For multiplicity of the peaks, the abbreviations used are (s) singlet, (d) doublet, (t) triplet, and (q) quartet.

FT-IR spectroscopy were carried out on a Shimadzu IR Affinity 1S spectrophotometer.

Single crystal XRD data were collected either on a Rigaku Oxford Diffraction SuperNova A S2 single crystal diffractometer using either Mo or Cu radiation or were collected with a Bruker D8 Quest ECO diffractometer using Mo-Kα radiation (λ = 0.71073 Å). Crystals were mounted on a Mitegen micromount in either NVH or Paratone immersion oil and temperature controlled using an Oxford Cryostream. Sample specific details can be found in the CIF files.

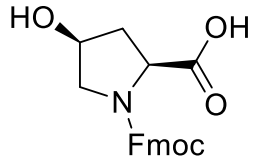
X-ray powder diffraction patterns were collected on a Rigaku Miniflex 600 using a Cu radiation source and measurements were performed at room temperature (Scan range 2Θ; 5-55°, Step; 0.02 °, scan speed; 0.8 °min^{-1}).

Further experimental details and instrumentation (e.g., gas absorption, thermogravimetric analysis, HRMS, prep-HPLC, AFM, SEM, TEM, etc.) used are described in the relevant sections.

SI 4.1 Monomer Synthesis

All other monomers, unless otherwise stated, were separated on a 2.1 x 150 mm, 3.6 um, XB-C18 Aeris Widepore column, from Phenomenex, at 30 °C on an Agilent 1100 HPLC. Gradient: 5% B for 5 minutes then from 5% B to 100% B over 25 minutes and held at 100% B for 5 minutes. Where A is Water (0.1% formic acid) and B is acetonitrile (0.1% formic acid). Flow rate is 0.2 mL/min. The flow was directed into directed into the electrospray source of a Bruker micrOTOF-QII mass spectrometer, operating in positive ion mode, at 5 kV and mass spectra recorded from 150-3000 m/z. Data was analysed with Bruker's Compass Data Analysis software.

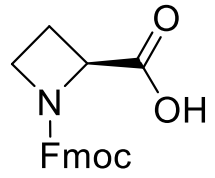
SI 4.1.1 Synthesis of (*2S,4S*)-1-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-4-hydroxypyrrolidine-2-carboxylic acid, 1:



1

(2S,4S)-4-hydroxypyrrolidine-2-carboxylic acid (0.698 g, 5.3 mmol, 1 eq.) was dissolved in THF (5 mL) and saturated aqueous $NaHCO_3$ (5 mL) in an ice bath. Fmoc-ONSu (2.7 g, 8.0 mmol, 1.5 eq.) was then added to the stirred solution and the solution stirred at 0 °C. The pH was checked after 1 h and adjusted with excess $NaHCO_3$ to pH 8. The solution was then stirred overnight at rt. The reaction was monitored by TLC (8:2, Hex/EtOAc). The reaction was quenched with deionised water and washed with ice cold diethyl ether (x 4). The aqueous solution was then acidified with 3M HCl to pH 2 and extracted with EtOAc (x 3). The combined EtOAc layers where then washed with brine (x 2), before drying over anhydrous $MgSO_4$ The solvent was removed under vacuo to yield an oil (2.25 g), this was redissolved in Et_2O and reevaporated to yield a white foamy solid of Fmoc-cis-L-Hyp-OH, 1, (1.80 g, 5.14 mmol, 97%).^{1}H NMR (400 MHz, $CDCl_3$) δ 7.73 (m, 2H), 7.61 – 7.47 (m, 2H), 7.44 – 7.27 (m, 4H), 5.82 (s, 2H), 4.55 – 4.29 (m, 4H), 4.28 – 4.08 (m, 1H), 3.72 – 3.48 (m, 2H), 2.42 – 2.22 (m, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 175.9, 172.2, 141.5, 128.0, 127.3, 125.1, 120.2, 71.0, 70.0, 68.5, 66.1, 58.4, 57.7, 55.6, 47.2, 37.3, 25.5. ESI-MS (m/z):$[M+H]^+$calcd.for $C_{20}H_{19}NO_5$, 353.1263 ;found, 353.1; $[M+Na]^+$, 376.3; $[2M+Na]^+$, 729.6.

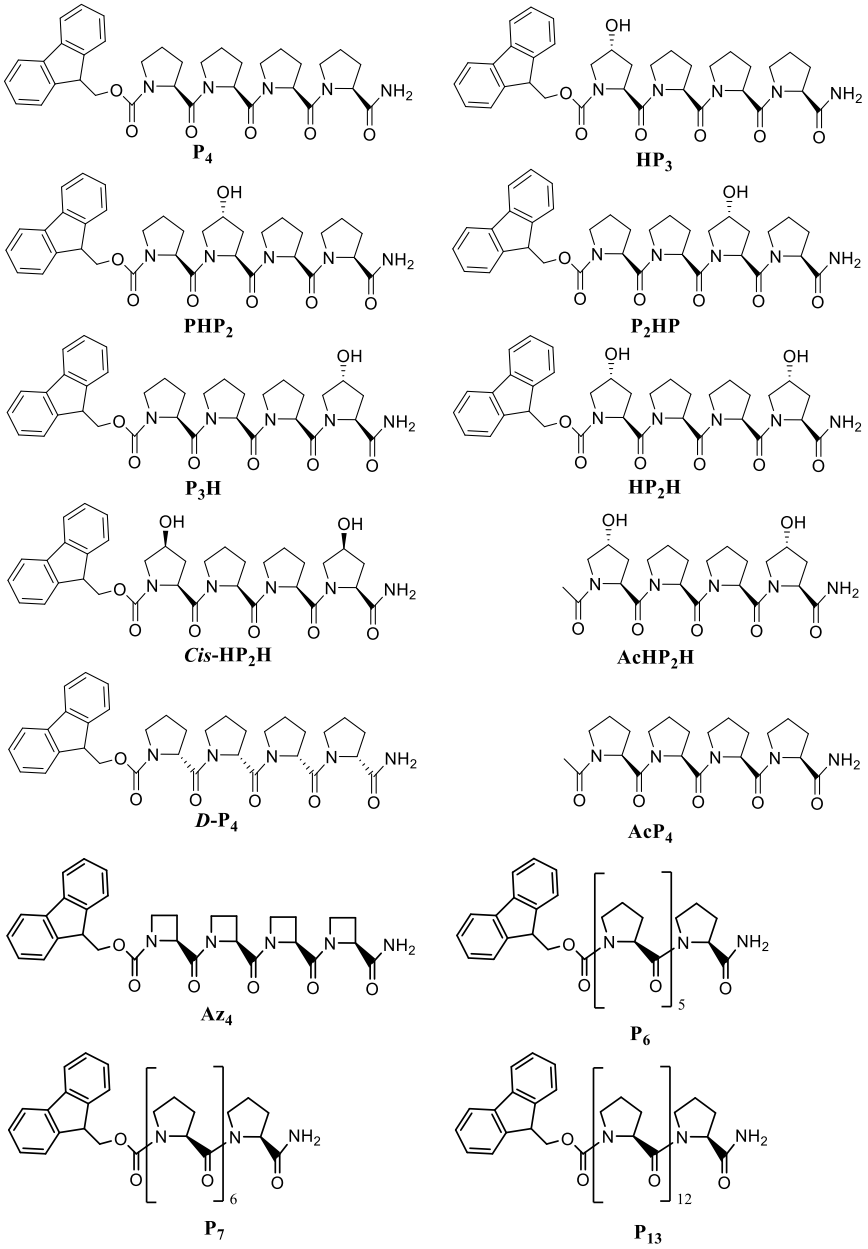
SI 4.1.2 Synthesis of (*S*)-1-(((9*H*-fluoren-9-yl)methoxy)carbonyl)azetidine-2-carboxylic acid, 2:



2

Azetidine-2-carboxylic acid (1.00 g, 9.9 mmol, 1 eq.) and K_2CO_3 (684 mg, 9.0 mmol, 0.5 eq.) were dissolved in THF/water (1:1, 5 mL). Fmoc-Cl (2.1494 g, 8.3 mmol, 0.85 eq.) was added to the stirring solution causing bubbling to occur. The reaction was stirred overnight at rt. The solution was then diluted with saturated aqueous $NaHCO_3$ and washed with Et2O (x 3). The aqueous layer was then acidified to pH 2 with 3 M HCl and extracted with DCM (x 3).The combined DCM extracts were then washed with brine (x 2) before drying over anhydrous $MgSO_4$ and filtering. The solvent as removed under vacuo to yield a white foamy solid, **2** (2.3 g, 7.1 mmol, 72%).^{1}H NMR (400 MHz, $CDCl_3$) δ 7.76 (d, *J* = 7.5 Hz, 2H), 7.56 (dd, *J* = 7.5, 1.0 Hz, 2H), 7.41 (t, *J* = 7.4 Hz, 2H), 7.32 (t, *J* = 7.4 Hz, 2H), 4.81 (s, 1H), 4.52 – 4.36 (m, *J* = 8.4 Hz, 2H), 4.23 (t, *J* = 6.2 Hz, 1H), 4.02 (t, *J* = 7.5 Hz, 2H), 2.55 (s, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 143.6, 141.5, 128.0, 127.3, 125.2, 125.1, 120.2, 68.1, 62.2, 60.8, 47.5, 47.2. ESI-MS (m/z):$[M+H]^+$calcd.for $C_{19}H_{18}NO_4^+$, 321.230 ;found; $[M+H]^+$, 320.0 ; $[M+Na]^+$, 346.0; $[M+Na_2CO_3]^+$, 429.0; $[2M+Na]^+$, 668.9

SI 4.2 Peptide synthesis:



Peptide Synthesiser Method – All peptides were synthesised on a 0.1 mmol scale on Rink amide MBHA resin using a Liberty™ microwave peptide synthesiser (CEM) utilising Fmoc solid-phase peptide synthesis techniques and repeated steps of single deprotections, and couplings interspaced with washings (4 x 4 mL DMF). The synthesis was paused after the final coupling step and the resin removed from the reaction vessel before stopping the synthesis (to prevent gradual loss of the Fmoc group). Deprotection: 20% piperidine in DMF (5 mL) for 5 min with 30 W microwave irradiation at 90 °C. Coupling: Fmoc-amino acid (1.5 mL, 0.2 M, 3 eq.), DIC (1.2 mL, 0.5 M, 6 eq.), Oxyma Pure (0.6 mL, 0.5 M, 3 eq.) in DMF, and DMF (3 mL) for 5 min at 90 °C with 30 W microwave irradiation. **Capping procedure:** Resin washed with DMF (x 5) and suspended in DMF with Acac (50 eq.) and DIPEA (50 eq.). Resin agitated for 30 min before filtering and washing the resin multiple times with DMF. Capped peptides containing hydroxyl groups were then spun with MeOH for 30 min before washing with MeOH (x 2) and DCM (x 4).

Peptide Cleavage - The resin was then washed with DCM (x 5), before the Fmoc-protected peptide was cleaved from the resin with TFA (95% in DCM) for 1.5 h. The resin was then washed with the cleavage cocktail (x 2) and the filtrate was concentrated by evaporation before

precipitation in cold Et_2O and centrifugation. The solution was then decanted and the solid repeatedly washed with cold Et_2O to isolate the peptide, as a white solid in a quantitative yield after drying under vacuum. Peptides were then used without any further purification, > 99% purity by analytical reverse-phase HPLC.

High-resolution Mass-Spectrometry (ESI+) – **P_4** and **P_2HP** were separated on a Phenomenex 2.1 x 150 mm, 3.6 um, XB-C18 Aeris Widepore column at 50 °C on a Waters H-Class Acquity UPLC. R_t = 10.49 min. Using a 0.1% formic acid/acetonitrile gradient: 5% B for 2.55 minutes then from 5% B to 95% B over 15 minutes, and held at 95% B for 2 minutes. A is 0.1% formic/Water, B is 0.1% formic/Acetonitrile. Flow rate is 0.25 mL/min. The flow is directed into the electrospray source of a Waters G2-Si mass spectrometer, operating in positive ion mode, at 2.5 kV and mass spectra recorded from 100-3000 m/z. Data was analysed with Waters Mass Lynx software.

All other peptides, unless otherwise stated, were separated on a 2.1 x 150 mm, 3.6 um, XB-C18 Aeris Widepore column, from Phenomenex, at 30 °C on an Agilent 1100 HPLC. Gradient: 5% B for 5 minutes then from 5% B to 100% B over 25 minutes, and held at 100% B for 5 minutes. Where A is Water (0.1% formic acid) and B is acetonitrile (0.1% formic acid). Flow rate is 0.2 mL/min. The flow was directed into directed into the electrospray source of a Bruker micrOTOF-QII mass spectrometer, operating in positive ion mode, at 5 kV and mass spectra recorded from 150-3000 m/z. Data was analysed with Bruker's Compass Data Analysis software.

LCMS - Peptides **P_3H**, **$AcHP_2H$**, ***cis*-HP_2H** and **AcP_4** were separated *via* RP-HPLC using a HiChrom KR100 5C18 5263 column at 40 °C on a Dionex UltiMate 3000. Gradient: 5% B for 5 minutes then from 5% B to 100% B over 20 minutes, and held at 100% B for 5 minutes. Where A is Water (0.1% formic acid) and B is methanol (0.1% formic acid). Flow rate is 1.0 mL/min. Wavelength: 225 nm. The flow was directed into directed into the electrospray source of a Thermo Scientific MSQ Plus Mass Detector, operating in positive ion mode, at 75 kV and mass spectra recorded from 100-2000 m/z.

P_4,(PP_4) **^{1}H NMR** (400 MHz, MeOD) δ 7.85 (dd, *J* = 14., 7.3 Hz, 2H), 7.70 – 7.58 (m, 2H), 7.48 – 7.31 (m, 4H), 4.76 – 4.68 (m, 1H), 4.66 – 4.61 (m, 0.5H), 4.58 (dd, *J* = 9.0, 4.1 Hz, 0.5H), 4.6 – 4.3.16 (m, 2H), 4.3.14 – 4.25 (m, 1H), 4.19 (t, *J* = 4.6 Hz, 0.5H), 4.08 (dd, *J* = 8.7, 3.2 Hz, 0.5H), 3.87 – 3.78 (m, 1H), 3.78 – 3.70 (m, 1H), 3.63 (q, *J* = 7.0 Hz, 4H), 3.59 – 3.52 (m, 1H), 3.51 – 3.42 (m, 1H), 3.42 – 3.35 (m, 1H), 2.32 – 1.74 (m, 14H). **^{13}C NMR** (101 MHz, MeOD) δ = 177.0, 172.8, 172.6, 172.3, 172.1, 156.6, 156.2, 145.4, 142.6, 128.8, 128.2, 126.2, 125.7, 120.9, 68.7, 67.0, 61.2, 59.6, 59.3, 59.1, 58.3, 48.3, 48.0, 47.8, 30.70, 30.0, 29.1, 28.8, 25.8, 25.2, 24.0, 18.4. FT-IR = ν_{max}/cm^{-1} 2956.88 (C-H), 2881.65 (C-H), 1683.86 (C=Os), 1624.06 (C=Os). m/z calcd. for $[M+H]^+$ $C_{35}H_{42}N_5O_6^+$: 628.3130; found: $[M+H]^+$ 628.3137, $[M+Na]^+$ 650.2938, $[2M+Na]^+$ 1277.6010

HP_3, ^{1}H NMR (400 MHz, MeOD) δ 7.83 (dd, *J* = 15.9, 7.7 Hz, 2H), 7.67 – 7.57 (m, 2H), 7.47 – 7.37 (m, 2H), 7.37 – 7.28 (m, 2H), 4.76 – 4.56 (m, 3H), 4.7 – 4.15 (m, 5H), 3.90 – 3.43 (m, 7H), 3.09 – 3.00 (m, 1H), 2.35 – 2.12 (m, 4H), 2.12 – 1.84 (m, 10H). ^{13}C NMR (101 MHz, MeOD) δ 177.0, 172.6, 172.6, 172.3, 172.0, 156.5, 145.2, 142.6, 128.9, 128.8, 128.2, 126.2, 125.7, 121.0, 120.9, 70.9, 69.8, 68.9, 67.3, 61.2, 59.8, 59.7, 59.6, 59.4, 58.0, 56.3, 56.1, 39.0, 30.7, 29.2, 29.1, 28.8, 25.9, 25.7. m/z calcd. for $[M+H]^+$ $C_{35}H_{42}N_5O_7^+$: 644.3079; found: $[M+H]^+$ 644.2, $[M+Na]^+$ 666.1 , $[2M+Na]^+$

PHP_2, ^{1}H NMR (400 MHz, MeOD) δ 7.82 (t, *J* = 7.5 Hz, 2H), 7.67 – 7.55 (m, 2H), 7.47 – 7.25 (m, 4H), 4.81 (t, *J* = 7.8 Hz, 0.5H), 4.73 – 4.61 (m, 1.5H), 4.57 (dd, *J* = 8.5, 3.2 Hz, 0.5H), 4.53 – 4.3.10 (m, 4.5H), 4.25 (t, *J* = 6.3 Hz, 1H), 3.89 – 3.69 (m, 3H), 3.64 (dd, *J* = 10.2, 4.8 Hz, 3H), 3.58 – 3.35 (m, 3H), 2.69 (s, 1H), 2.35 – 2.15 (m, 4H), 2.14 – 1.80 (m, 11H). ^{13}C NMR (101 MHz, MeOD) δ 177.0, 172.9, 172.6, 172.1, 145.4, 142.5, 129.9, 128.8, 128.2, 128.2, 126.1, 126.0,

122.0, 120.9, 120.7, 108.2, 71.3, 70.8, 68.8, 68.5, 61.3, 59.7, 58.6, 56.2, 55.7, 37.3, 30.9, 30.7, 29.3, 25.9. m/z calcd. for $[M+H]^+$ $C_{35}H_{42}N_5O_7^+$: 644.3079; found: $[M+H]^+$ 644.3, $[M+Na]^+$ 666.3, $[2M+H]^+$ 1287.8, $[2M+H+Na]^{2+}$ 1310.6

P_2HP, ^{1}H NMR (400 MHz, MeOD) δ 7.83 (dd, *J* = 15.5, 7.7 Hz, 2H), 7.67 – 7.55 (m, 2H), 7.47 – 7.28 (m, 4H), 4.81 – 4.67 (m, 2H), 4.63 – 4.7 (m, 2H), 4.6 – 4.3.12 (m, 2H), 4.3.12 – 4.23 (m, 1H), 4.17 (t, *J* = 4.6 Hz, 0.5H), 4.07 (dd, *J* = 8.8, 3.3 Hz, 0.5H), 3.89 – 3.69 (m, 3H), 3.69 – 3.64 (m, 1H), 3.56 – 3.38 (m, 2H), 2.97 (dd, *J* = 11.8, 5.0 Hz, 1H), 2.33 – 1.68 (m, 14H). ^{13}C NMR (101 MHz, MeOD) δ 177.0, 172.5, 145.7, 145.1, 142.7, 128.8, 128.2, 126.1, 125.6, 120.9, 71.2, 68.8, 67.0, 61.3, 59.4, 59.1, 37.5, 30.8, 30.1, 28.8, 25.9, 25.6, 24.1. m/z calcd. for $[M+H]^+$ $C_{35}H_{42}N_5O_7^+$: 644.3079; found: $[M+H]^+$ 644.3, $[M+Na]^+$ 666.2, $[2M+2H]^{2+}$ 1288.6, $[2M+H+Na]^{2+}$ 1310.2

P_3H, ^{1}H NMR (400 MHz, MeOD) δ 7.87 (dd, *J* = 16.1, 8.2 Hz, 2H), 7.71 – 7.59 (m, 2H), 7.53 – 7.44 (m, 2H), 7.43 – 7.34 (m, 2H), 4.75 – 4.61 (m, 2H), 4.60 – 4.53 (m, 1H), 4.53 – 4.3.19 (m, 2H), 4.3.11 – 4.05 (m, 2H), 3.86 – 3.74 (m, 3H), 3.71 – 3.57 (m, 2H), 3.22 (s, 1H), 2.69 (dd, *J* = 16.0, 6.3 Hz, 1H), 2.35 – 2.13 (m, 4H), 2.12 – 2.01 (m, 4H), 2.00 – 1.68 (m, 7H). ^{13}C NMR (101 MHz, MeOD) δ 172.8, 157.7, 146.5, 145.2, 142.7, 128.8, 128.2, 126.2, 125.6, 121.0, 120.9, 71.6, 71.1, 68.8, 67.8, 67.0, 60.1, 59.7, 59.4, 59.1, 56.1, 39.0, 29.1, 25.7, 24.1. m/z calcd. for $[M+H]^+$ $C_{35}H_{42}N_5O_7^+$: 644.3079; found: $[M+H]^+$ 644.2, $[2M+Na]^+$ 1309.7

HP_2H, ^{1}H NMR (400 MHz, MeOD) δ 7.83 (dd, *J* = 15.9, 7.8 Hz, 2H), 7.68 – 7.57 (m, 2H), 7.42 (dt, *J* = 12.1, 7.4 Hz, 2H), 7.37 – 7.28 (m, 2H), 4.76 – 4.56 (m, 3H), 4.8 (dt, *J* = 10.7, 7.8 Hz, 2H), 4.3.17 – 4.22 (m, 2.5H), 4.18 (t, *J* = 4.8 Hz, 0.5H), 3.89 – 3.61 (m, 5H), 3.50 (dt, *J* = 18.0, 11.3 Hz, 2H), 3.08 – 2.98 (m, 1H), 2.40 – 2.13 (m, 4H), 2.13 – 1.81 (m, 8H). ^{13}C NMR (101 MHz, MeOD) δ 176.8, 172.7, 172.6, 172.3, 172.1, 156.5, 145.5, 145.2, 142.59, 128.8, 128.2, 121.0, 71.1, 69.8, 67.2, 60.1, 59.7, 59.4, 58.0, 56.3, 56.1, 39.0, 28.9, 27.3, 25.9, 25.8, 25.7, 25.7. m/z calcd. for $[M+H]^+$ $C_{35}H_{42}N_5O_8^+$: 660.3028; found: $[M+H]^+$ 660.2, $[M+Na]^+$ 682.2, $[2M+H]^+$ 1319.6, $[2M+H+Na]^{2+}$ 1342.3

$AcHP_2H$, ^{1}H NMR (400 MHz, MeOD) δ 4.79 – 4.65 (m, 3H), 4.54 – 4.3.1 (m, 3H), 3.91 – 3.49 (m, 8H), 2.53 – 2.16 (m, 5H), 2.13 – 1.97 (m, 10H). ^{13}C NMR (101 MHz, MeOD) δ 205.0, 200.8, 200.6, 200.4, 200.3, 99.3, 99.1, 88.3, 87.9, 86.2, 85.5, 84.3, 67.2, 66.3, 57.4, 57.3, 54.0, 54.0, 50.3. m/z calcd. for $[M+H]^+$ $C_{22}H_{34}N_5O_7^+$: 480.2453 ; found: $[M+H]^+$ 480.2, $[M+Na]^+$ 502.2

***Cis*-HP_2H**, ^{1}H NMR (400 MHz, MeOD) δ 7.85 (dd, *J* = 15.0, 7.5 Hz, 2H), 7.68 – 7.58 (m, 2H), 7.43 (dt, *J* = 13.0, 6.2 Hz, 2H), 7.35 (dd, *J* = 13.4, 6.0 Hz, 2H), 4.79 – 4.71 (m, 1H), 4.64 (ddd, *J* = 15.1, 10.2, 4.3.1 Hz, 2H), 4.9 – 4.3.17 (m, 3H), 4.3.17 – 4.14 (m, 3H), 3.94 (ddd, *J* = 15.0, 10.6, 5.0 Hz, 1H), 3.88 – 3.71 (m, 2H), 3.64 (dt, *J* = 13.4, 6.4 Hz, 4H), 3.58 – 3.38 (m, 2H), 2.54 – 2.23 (m, 4H), 2.20 – 1.77 (m, 10H). ^{13}C NMR (101 MHz, MeOD) δ 177.5, 173.1, 173.0, 156.6, 145.6, 145.3, 145.1, 145.0, 142.7, 128.9, 128.8, 128.3, 128.2, 126.8, 126.1, 125.8, 125.7, 125.5, 121.0, 71.5, 70.3, 68.8, 67.2, 60.4, 59.9, 59.7, 59.5, 58.4, 57.9, 56.5, 56.0, 38.3, 38.1, 37.6, 29.3, 28.9, 25.9, 25.7. m/z calcd. for $[M+H]^+$ $C_{35}H_{42}N_5O_8^+$: 660.3028 ; found: $[M+H]^+$ 660.3, $[M+Na]^+$ 680.3

AcP_4, ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.29 (s, 0.5H), 6.78 (s, 0.5H), 5.68 (s, 0.5H), 5.41 (s, 0.5H), 4.79 – 4.61 (m, 2H), 4.58 (d, *J* = 6.3 Hz, 0.5H), 4.3.1 (t, *J* = 6.8 Hz, 0.5H), 4.28 (d, *J* = 7.9 Hz, 0.5H), 3.93 – 3.43 (m, 7H), 2.69 (s, 3H), 2.58 (dd, *J* = 12.5, 6.2 Hz, 0.5H), 2.35 – 2.26 (m, 0.5H), 2.26 – 2.10 (m, 5H), 2.09 (s, 3H), 2.07 – 1.72 (m, 8H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.3, 173.9, 172.2, 171.3, 170.7, 170.6, 170.0, 60.8, 59.5, 59.1, 58.2, 58.0, 48.5, 48.4, 47.5, 47.4, 47.3, 47.2, 46.8, 31.5, 28.8, 28.5, 28.2, 27.9, 27.0, 25.4, 25.3, 25.2, 25.0, 24.9, 24.7, 22.2. m/z calcd. for $[M+H]^+$ $C_{22}H_{34}N_5O_5^+$: 448.2554; found: $[M+H]^+$ 448.5, $[M+Na]^+$ 470.5

***D*-P_4**, ^{1}H NMR (400 MHz, MeOD) δ 7.83 (dd, J = 15.6, 7.4 Hz, 2H), 7.69 – 7.56 (m, 2H), 7.48 – 7.28 (m, 4H), 4.77 – 4.53 (m, 4H), 4.4 – 4.3.13 (m, 2H), 4.3.11 – 4.23 (m, 1H), 4.17 (t, J = 4.5 Hz, 0.5H),

4.06 (dd, J = 8.8, 3.3 Hz, 0.5H), 3.86 – 3.40 (m, 7H), 3.40 – 3.34 (m, 1H), 3.00 – 2.90 (m, 1H), 2.32 – 1.72 (m, 14H). ^{13}C NMR (101 MHz, MeOD) δ 177.0, 172.6, 172.6, 172.4, 172.1, 167.7, 145.7, 142.7, 128.7, 128.2, 125.7, 120.9, 67.0, 61.2, 59.5, 59.4, 59.1, 30.7, 30.7, 29.2, 28.9, 25.9, 25.8, 25.7, 24.1. m/z calcd. for $[M+H]^+$ $C_{35}H_{42}N_5O_7^+$: 628.3137, found: $[M+H]^+$ 628.3, $[M+Na]^+$ 650.3, $[2M+Na]^+$ 1277.6

SI 4.2.1 FT-IR Analysis of P_4:

FT-IR spectroscopy was carried out on a Shimadzu IR Affinity 1S spectrophotometer.

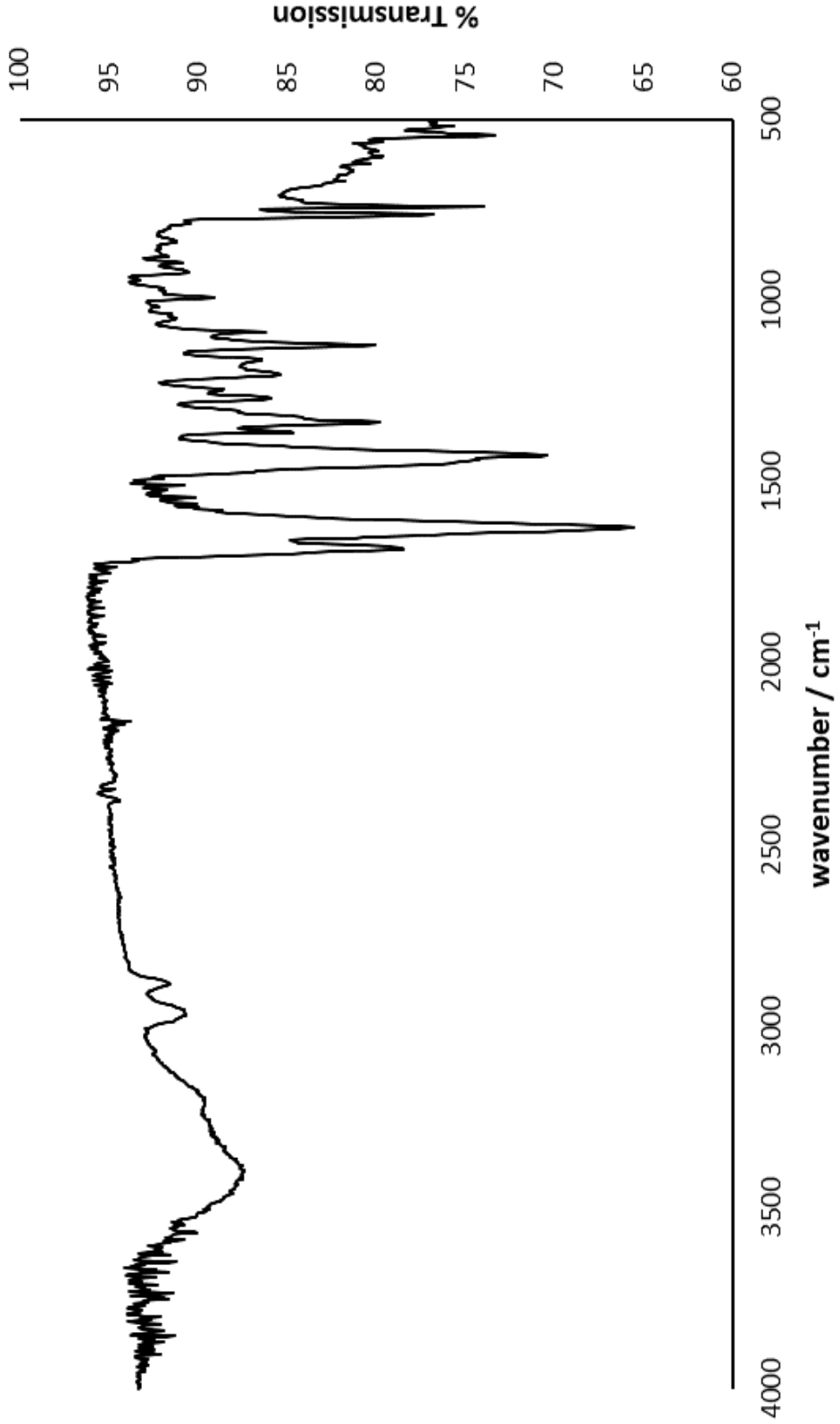
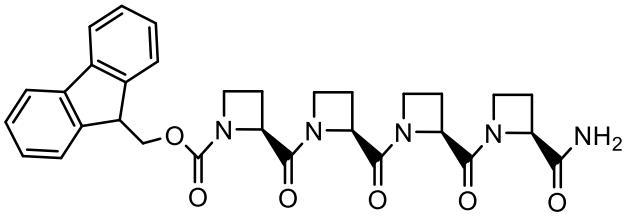


Figure S1 – FT-IR Spectrum of **P_4** as a dry solid. The C=O stretching band frequency was typical for that of a PPII helix at 1624 cm^{-1}.[1]

SI 4.2.2 Circular Dichroism Spectroscopy:

CD experiments were carried out on a Jasco J-715 spectropolarimeter. Spectra were recorded using a spectral bandwidth of 190-260 nm, at 20 °C, with a scan rate of 100 nm/min. CD data are given in ellipticity (mdeg). The spectra are formed of 4 accumulations and a spectrum of the solvent blank was subtracted from the raw CD data. A Quartz cell was used with a 1 mm path length using either 125 or 250 µM peptide solutions. All samples were kept in solution for at least 14 days prior to recording CD spectra to ensure the final stable conformation had been achieved due to slow conversion between polyproline helices (i.e. Polyproline II→ Polyproline I). The observed ellipticity has been converted to molar ellipticity (θ) for all spectra, expressed in the units $deg.cm^2.dmol^{-1}$.

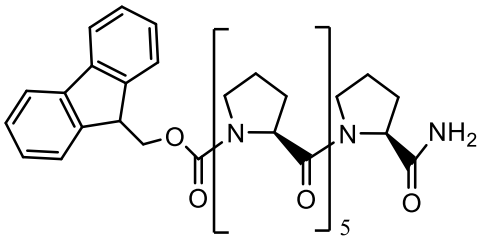
SI 4.2.3 Synthesis of Fmoc-$(Az)_4$-NH_2 (14):



(14) Az_4

Peptide **14**, **Az_4**, (Az – Azetidine) was synthesised using standard SPPS techniques on 0.1 mmol scale using the Oxyma-DIC coupling method, quantitative yield. m/z calcd. for $[M+H]^+$ $C_{31}H_{34}N_5O_6^+$: 572.2504; found: $[M+H]^+$ 572.0, $[M+Na]^+$ 594.0

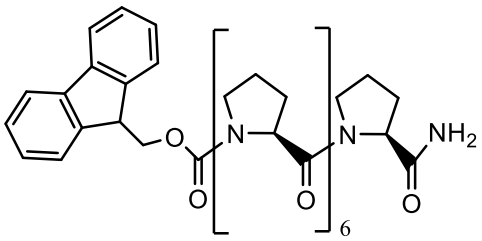
SI 4.2.4 Synthesis of Fmoc-$(Pro)_6$-NH_2 (15):



(15) P_6

Peptide **15**, **P_6**, was synthesised using standard SPPS techniques on 0.1 mmol scale using the Oxyma-DIC coupling method, quantitative yield. m/z calcd. for $[M+H]^+$ $C_{45}H_{56}N_7O_8^+$: 822.4185; found: $[M+H]^+$ 822.8, $[M+Na]^+$ 843.9

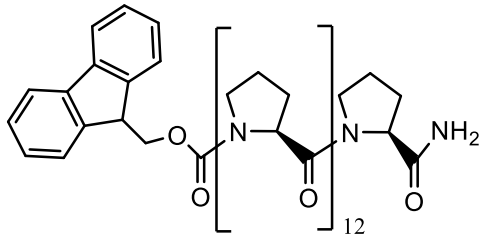
SI 4.2.5 Synthesis of Fmoc-$(Pro)_7$-NH_2 (16):



(16) P_7

Peptide **16**, **P_7**, was synthesised using standard SPPS techniques on 0.1 mmol scale using the Oxyma-DIC coupling method, quantitative yield. m/z calcd. for $[M+H]^+$ $C_{50}H_{63}N_8O_9^+$: 919.4713; found: $[M+H]^+$ 919.1, $[M+Na]^+$ 941.5

SI 4.2.6 Synthesis of Fmoc-$(Pro)_{13}$-NH_2 (18):



(18) P_{13}

Peptide **18**, **P_{13}**, was synthesised using standard SPPS techniques on 0.1 mmol scale using the Oxyma-DIC coupling method, quantitative yield. **P_{13},** m/z calc. for $C_{80}H_{105}N_{14}O_{15}^+$: 1501.7745 ,found: $[2M+H]^{2+}$ 751.2, $[2M+K]^{2+}$ 768.3, $[M+Na]^+$ 1524.6

SI 4.3 X-ray Diffraction data:

Crystallisation conditions:

Peptide **P_4** was fully dissolved (≈25-30 mg ml^{-1}) in hot EtOH before slowly cooling the solution from 60+ °C overnight, forming crystals (colourless planks) in solution.[2]

Peptide **P_2HP** was crystallised by slowly cooling from a hot EtOH/EtOAc solution, forming crystals (colourless planks) in solution.

Peptides **HP_3**, **PHP_2**, **P_3H** and **HP_2H** were crystallised by slow evaporation from an EtOH/EtOAc solution, forming crystals (colourless planks). However, no crystals of **PHP_2** suitable for SCXRD analysis were found.

Peptide **$AcHP_2H$** was dissolved in hot acetonitrile before slowly evaporating to form colourless needles. Sonication of a supersaturated solution instead forms an organogel, this occurred in acetonitrile, chloroform, and dichloromethane, with the strongest gelation in chloroform (*Figure 37*).

Crystallisation of peptide **AcP_4** was attempted in MeOH, EtOH, EtOH/EtOAc, ACN, and EtOH/ACN, producing a non-crystalline glassy solid or a viscous oil. The peptide crystallised readily from slow evaporation of a $CHCl_3$ solution to form colourless needles, however, the crystal quality was poor and unsuitable for single crystal analysis. The peptide was crystallised successfully from vapour diffusion of Et_2O into a solution of the peptide in $CHCl_3$ to produce colourless needles which were analysed *via* SCXRD analysis.

All crystalline samples were stable outside of their mother liquor. Crystals of peptide **$AcHP_2H$** melted under a stream of room temperature N_2 (290 K), likely due to the loss of ACN from the framework. Crystals of peptide *cis*-**HP_2H**, degraded under a room temperature N_2 flow (300 K) also likely due to loss of encapsulated solvent within the framework pores.

Single crystal XRD data for all peptides were collected on a Rigaku Oxford Diffraction SuperNova AS2 single crystal diffractometer using Cu Kα (λ = 1.54184) radiation, **$AcHP_2H$** and **AcP_4** used Mo Kα (λ = 0.71073). The crystals were mounted on a Mitegen micromount in Paratone immersion oil and temperature controlled using an Oxford Cryosystems 800-series Cryostream.

Using the software Olex2,[3] the structures were solved with the ShelXT structure solution program using intrinsic phasing and refined with the ShelXL refinement package using least squares minimization.[4–6]

CCDC-2127751,[2] 2234312, 2238152, 2238155, 2238160-1, 2238180, 2238252 and 2264145 contain the supplementary crystallographic data for this paper, including structure factors and refinement instructions, and can be obtained free of charge from The Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK (e-mail: deposit@ccdc.cam.ac.uk), or *via* https://www.ccdc.cam.ac.uk/getstructures.

Table S1 - Comparative table of crystallographic data for all crystal structures:

**Peptides that form isostructural structures are highlighted in the same colour*

Identification code	**PP_4-SPF**	**P2HP**	**P2HP-P4**	**P3H**	**HP3**	**HP2H**	**Cis-HP2H**	**AcHP2H**	**AcP4**
Empirical formula	$C_{37}H_{47}N_5O_7$	$C_{35}H_{45}N_5O_9$	$C_{36.5}H_{42.36}N_5O_{7.13}$	$C_{36}H_{41}N_5O_8$	$C_{35}H_{41}N_5O_7$	$C_{35}H_{40}N_5O_8$	$C_{35}H_{41}N_5O_8$	$C_{24}H_{36}N_6O_7$	$C_{24}H_{35}Cl_6N_5O_5$
Formula weight	673.79	679.76	665.19	671.74	643.73	658.72	659.73	520.59	686.27
Temperature/K	150(2)	150.15	290.0(2)	150.00(10)	150.15	295	150.15	290(5)	150.0(7)
Crystal system	monoclinic	monoclinic	monoclinic	monoclinic	monoclinic	monoclinic	orthorhombic	orthorhombic	monoclinic
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	C2	C2	$P2_12_12_1$	$P2_12_12_1$	$P2_1$
a/Å	16.4182(3)	16.6440(5)	16.6077(8)	16.8723(2)	22.1654(7)	22.476(2)	6.4160(2)	6.3773(10)	13.0167(8)
b/Å	6.27740(10)	6.3071(2)	6.3463(3)	6.45220(10)	6.3098(2)	6.2932(7)	12.0836(6)	18.232(4)	6.2940(6)
c/Å	18.2112(3)	17.8450(5)	18.3211(9)	32.1004(4)	23.0675(7)	23.245(3)	47.4708(17)	22.857(5)	19.6114(12)
α/°	90	90	90	90	90	90	90	90	90
β/°	109.152(2)	110.012(3)	109.933(5)	101.2140(10)	92.401(3)	92.036(10)	90	90	99.377(6)
γ/°	90	90	90	90	90	90	90	90	90
Volume/Å^3	1773.03(6)	1760.18(10)	1815.31(16)	3427.84(8)	3223.37(17)	3285.8(6)	3680.3(3)	2657.6(9)	1585.2(2)
Z	2	2	2	4	4	4	4	4	2
ρ_{calc}g/cm^3	1.262	1.283	1.217	1.302	1.326	1.332	1.191	1.301	1.438
μ/mm^{-1}	0.716	0.770	0.699	0.766	0.765	0.788	0.085	0.097	0.584
F(000)	720.0	724.0	707.0	1424.0	1368.0	1396.0	1400.0	1112.0	712.0
CCDC No.	2127750	2238161	2238180	2238152	2238155	2238160	2238252	2234312	2264145

SI 4.3.1 Analysis of Crystalline P_4:

Crystallisation conditions: Peptide **P_4** was fully dissolved (≈25-30 mg ml^{-1}) in hot EtOH before slowly cooling the solution from 60+ °C overnight, forming crystals (colourless planks) in solution. Slow evaporation of the EtOH solution elicits crystal formation on the walls of the vial.

All samples were stable outside of the mother liquor and no significant degradation occurred after air drying (stable at rt out of solution for ≈2-3 months before crystal deterioration).

Iodine adsorption: Single crystals of **P_4** were heated at 45 °C for 16 hours before soaking in a solution of iodine in hexane for 1 week. A colour change of colourless to orange/yellow crystals was observed over time (*further details* SI13.1).

Single crystal XRD data for the solvated structure and iodine treated sample were collected on a Rigaku Oxford Diffraction SuperNova A S2 single crystal diffractometer using Cu radiation. Sample specific details can be found in the CIF files.

Using the software Olex2,[3] the structures were solved with the ShelXT structure solution program using intrinsic phasing and refined with the ShelXL refinement package using least squares minimization.[4–6]

CCDC-2127748-2127751, and CCDC-2156434 contain the supplementary crystallographic data for this paper, including structure factors and refinement instructions, and can be obtained free of charge from The Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK (e-mail: depos-it@ccdc.cam.ac.uk), or *via* https://www.ccdc.cam.ac.uk/getstructures.

The crystal structure data was obtained from colorless crystals, crystallised by dissolving **P_4** in hot EtOH (60 °C, ≈25 mg ml^{-1}) and allowing to slowly cool to room temperature forming crystals rapidly overnight within the solution. The crystals were stable outside of solution at room temperature showing no signs of deterioration over the timeframe of the experiment. For the **PP_4-SPF** crystal, to model the disordered solvent, the ethanol molecule was split into two components with a total occupancy of 1. The distances between the oxygen-carbon and carbon-carbon atoms were set to the expected values of 1.43 Å and 1.51 Å respectively. Upon refinement, these distances were then fixed to the same value between the two components and the anisotropic displacement parameters were set to be equivalent within the molecule. For the **PP_4-SPF$_{Flash}$** the same restraints and constraints were applied as above, with the two components of the ethanol molecule set to 0.75 and 0.25. The oxygen atom of the most abundant component was further split into two, and the total occupancy for its two components was set to the sum of 0.75. A further restraint between the CH_3- carbon and the oxygen atom was necessary to model this component, whose distance was set to be 2.4 Å.

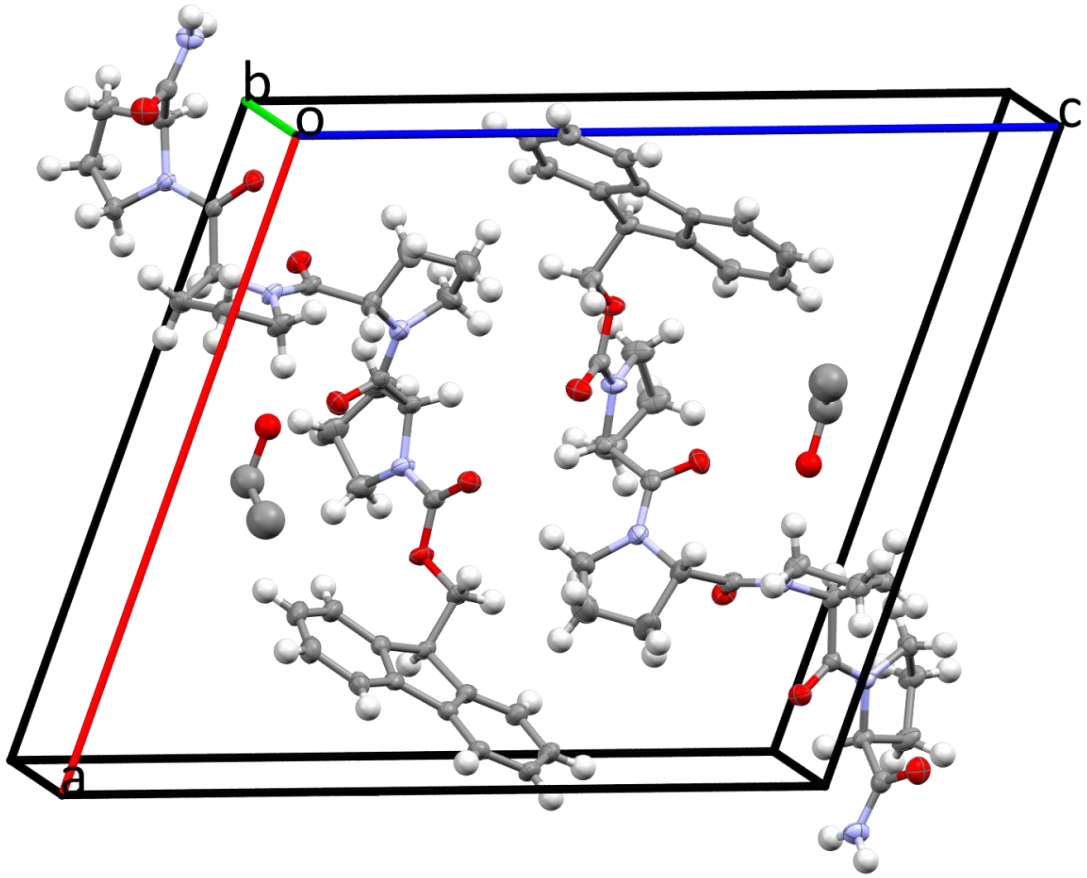


Figure S4 – Crystal structure of peptide P_4 showing the unit cell, all molecules whose centroids fit are shown, atomic displacement parameters are shown at 50% probability.

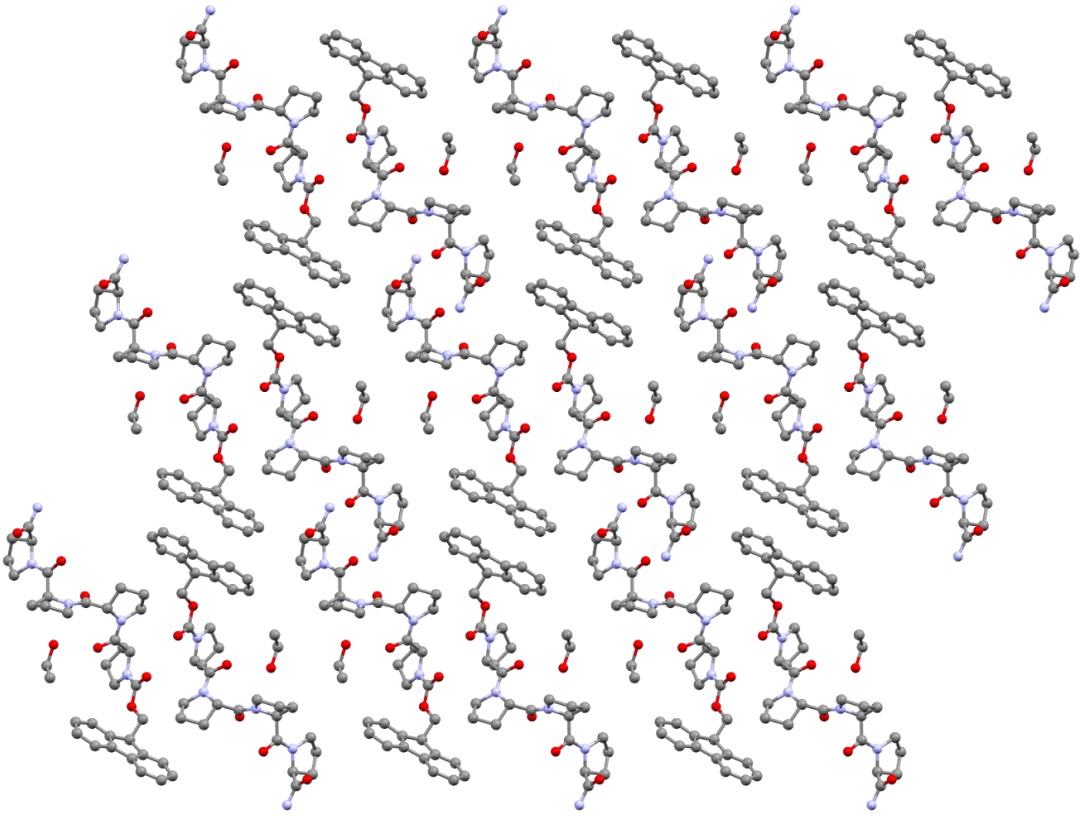


Figure S5 – Crystal structure of peptide P_4 showing the packed extended structure (3x3x3), view along the *b* axis, hydrogens not shown, atomic displacement parameters are shown at 50% probability.

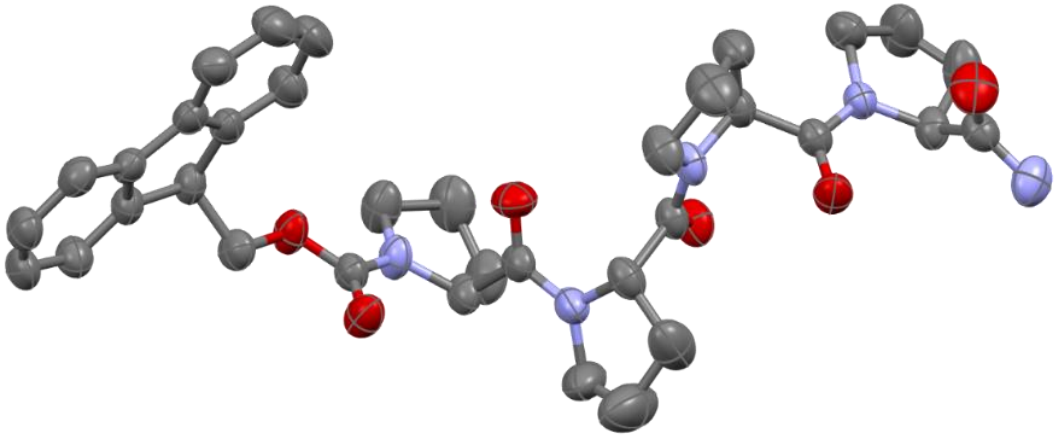


Figure S6 – Crystal structure of **PP_4-SPF** showing the asymmetric unit. Atomic displacement parameters are drawn at 50% probability. Hydrogen atoms have been removed for clarity.

PP_4-SPF was slowly cooled from 290 K to 150 K with a ramp rate of 280 K/h, while **PP_4-SPF_{Flash}** placed on the mount at 150 K (flash frozen). Upon returning to room temperature **PP_4-SPF_{Flash}** crystals readopt their initial unit cell and conformation. This was not found when slowly ramping the temperature down to 150 K, with only a slight reduction of the cell volume, but no significant conformational changes.

Identification code	**PP_4-SPF**	**PP_4-SPF$_{_Flash}$**
Empirical formula	$C_{37}H_{47}N_5O_7$	$C_{37}H_{41}N_5O_7$
Formula weight	673.79	667.75
Temperature/K	150(2)	150(2)
Crystal system	monoclinic	monoclinic
Space group	$P2_1$	$P2_1$
a/Å	16.4182(3)	15.5793(2)
b/Å	6.27740(10)	6.34080(10)
c/Å	18.2112(3)	18.5772(3)
α/°	90	90
β/°	109.152(2)	107.249(2)
γ/°	90	90
Volume/Å^3	1773.03(6)	1752.62(5)
Z	2	2
ρ_{calc}g/cm^3	1.262	1.265
μ/mm^{-1}	0.716	0.724
F(000)	720.0	708.0
Crystal size/mm^3	0.217 × 0.063 × 0.054	0.298 × 0.056 × 0.042
Radiation	Cu Kα (λ = 1.54184)	Cu Kα (λ = 1.54184)
2Θ range for data collection/°	8.84 to 146.496	9.972 to 1402
Index ranges	-16 ≤ *h* ≤ 20, -7 ≤ *k* ≤ 7, -22 ≤ *l* ≤ 19	-19 ≤ *h* ≤ 18, -7 ≤ *k* ≤ 7, -22 ≤ *l* ≤ 22
Reflections collected	20137	19739
Independent reflections	6927 [R_{int} = 0.0225, R_{sigma} = 0.0217]	6762 [R_{int} = 0.0233, R_{sigma} = 0.0225]
Data/restraints/parameters	6927/8/444	6762/9/452
Goodness-of-fit on F^2	1.040	1.055
Final R indexes [I>=2σ (I)]	R_1 = 0.0383, wR_2 = 0.1041	R_1 = 0.0449, wR_2 = 0.1271
Final R indexes [all data]	R_1 = 0.0397, wR_2 = 0.1055	R_1 = 0.0464, wR_2 = 0.1295
Largest diff. Peak/hole / e Å^{-3}	0.56/-0.55	0.67/-0.36
Flack parameter	0.01(5)	0.04(5)
CCDC No.	2127750	2127751

SI 4.3.1.1. Modelling of PP_4-SPF pores

The void space for **PP_4-SPF** was calculated with Mercury (Probe radius, 1.2 Å, grid spacing, 0.7 Å; Volume 226.19 Å^3, 12.8% of unit cell). The channel is viewed along the *b* axis and is depicted in yellow. The channel is anisotropic, so both the shortest and longest edge to edge distances were measured (shortest, 8 x 7.7 Å; largest, 5.0 x 11.7 Å).

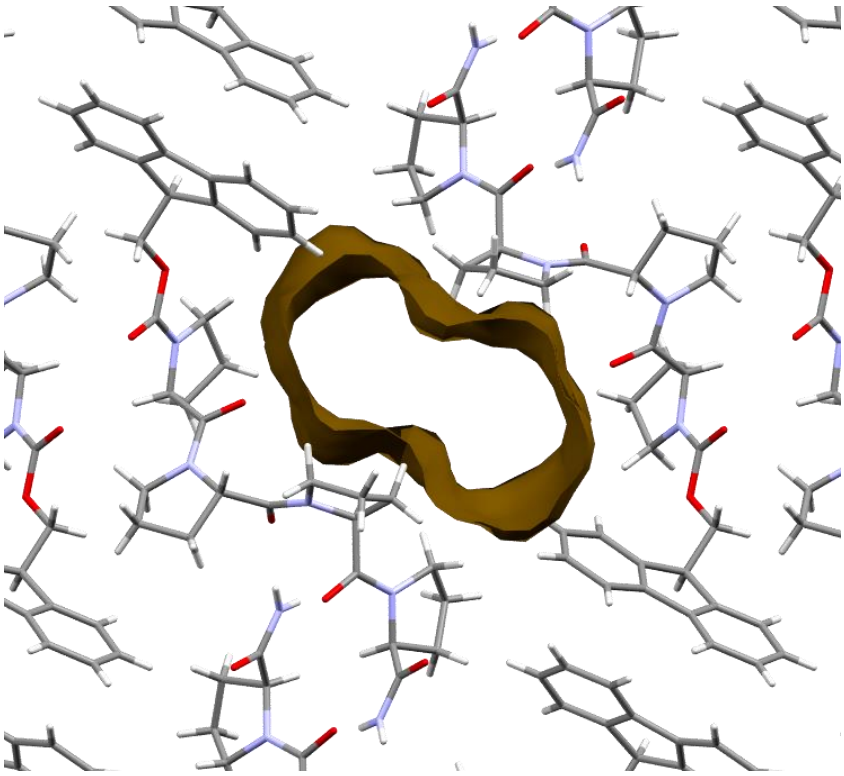


Figure S10 – Model of the channel (yellow) within the crystal structure of **PP_4-SPF** (Mercury)

SI 4.3.1.2 NMR Studies to determine thermal activation potential

Through NMR studies we were able to demonstrate that drying under high vacuum at 45 °C was sufficient for almost complete desorption of ethanol from PP_4-**SPF** (8% mol/mol, i.e. 1 g **PP_4-SPF**: 6.4 mg EtOH). Drying under vacuum at 30 °C was insufficient with significant residual ethanol (85% mol/mol, i.e. 1 g **PP_4-SPF_{act}**: 62.4 mg EtOH).

NMR analysis was carried out in MeOD from samples of **PP_4-SPF** crystals after thermal treatment. Crystals of **PP_4-SPF** was dried under vacuum at 30 °C to remove excess ethanol giving an initial ethanol content of 97% mol/mol (0.97 eq. EtOH: 1 eq. **PP_4**). Subsequent heating under high vacuum resulted in a reduction of ethanol down to 9.7% mol/mol after heating under vacuum at 45 °C for 12 hours. The recorded spectra are shown with the ethanol%, heating temperature and time. These results were reproducible with different crystal batches with only slight differences in ethanol percentage (± < 10%).

SI 4.3.1.3 Gas Adsorption Studies

The freshly isolated **PP_4-SPF** was activated by degassing at 45 °C for a total of 29 hours under dynamic vacuum, and the volumetric uptake of N_2 (77 K) and CO_2 (280 K) was measured. In both cases, no significant adsorption was observed across the loading range $P/P_0 = 10^{-7} – 1.0$ (N_2) or 0 – 1 bar (CO_2) consistent with adsorption only on the surfaces of the particles without penetration into the material, and a surface area <10m^2/g.

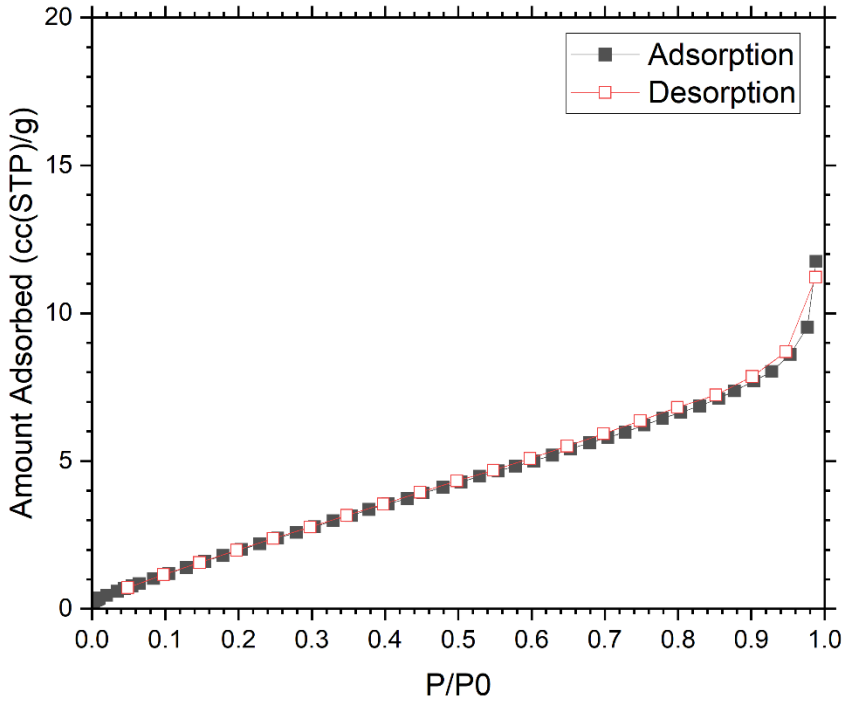


Figure S12 – N_2 adsorption isotherm for the desolvated **PP_4-SPF** (77 K)

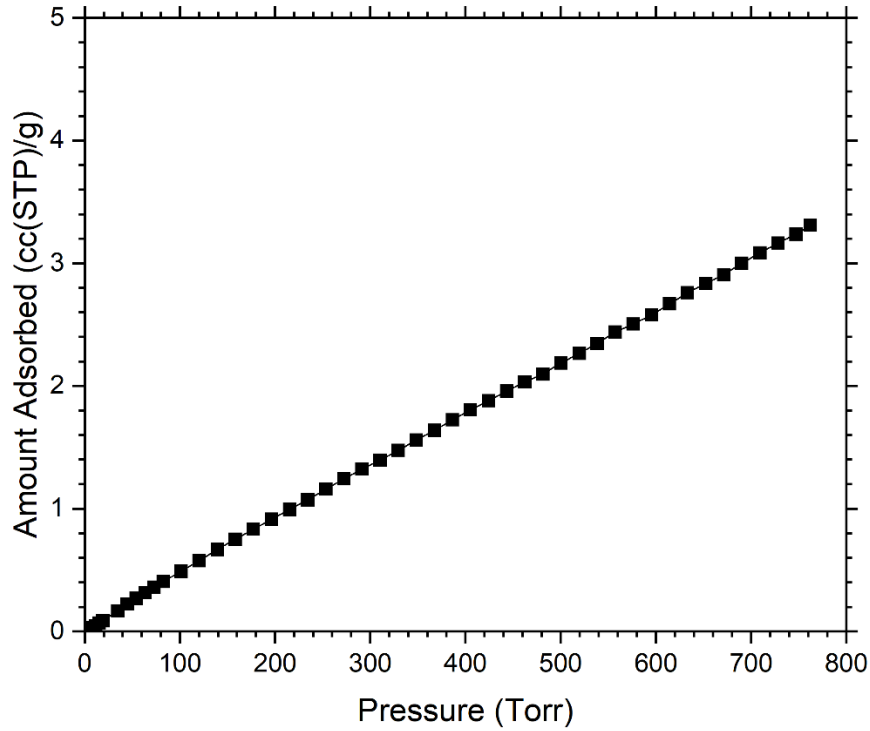


Figure S13 – CO_2 adsorption isotherm for desolvated **PP_4-SPF** (280 K).

Gas adsorption measurements were performed using a Quantachrome Autosorb iQ using N_2 and CO_2 at N5 grade or better. Temperature control was provided by a liquid nitrogen dewar for N_2 measurements and a Julabo recirculating chiller for CO_2 measurements. Samples were activated prior to the adsorption experiment at 45 °C for 6 hours dynamic vacuum provided by a rotary oil pump followed by evacuation under high vacuum provided by a turbomolecular pump at 45 °C for 15 and then 30 °C for a further 8 hours. The final desolvated sample mass was 165.2 mg.

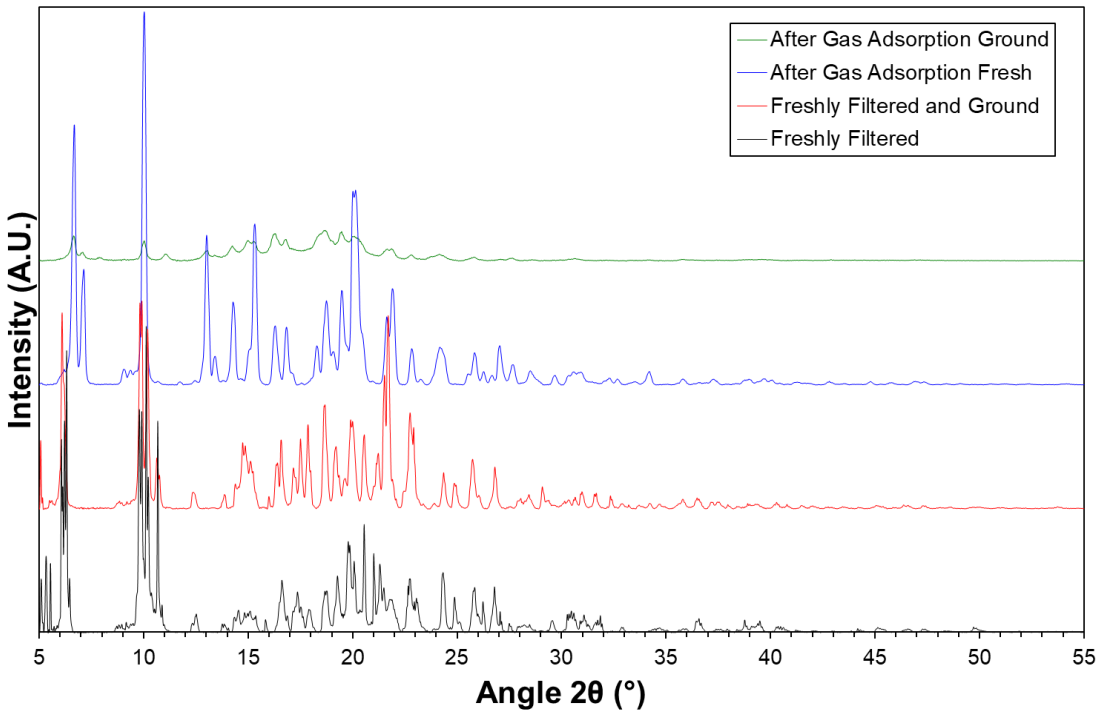


Figure S14 – X-ray powder diffraction pattern of the freshly isolated, solvated **PP_4-SPF** (**black**), this sample after grinding in air (**red**), the sample immediately following de-solvation and gas adsorption experiments (**blue**) and the desolvated sample after grinding (**green**).

SI 4.3.1.4. Simultaneous thermogravimetric analysis of P4 crystalline framework:

Thermogravimetric analysis was carried out on a STA 409 PC Luxx Simultaneous thermal analyser with N_2 as the purge gas and a temperature range of 30-600 °C and a ramp rate of 10 K/min. A powdered sample of the crystalline **P_4** was packed into an aluminium crucible after drying excess ethanol under vacuum at 20 °C for 45 min and then grinding.

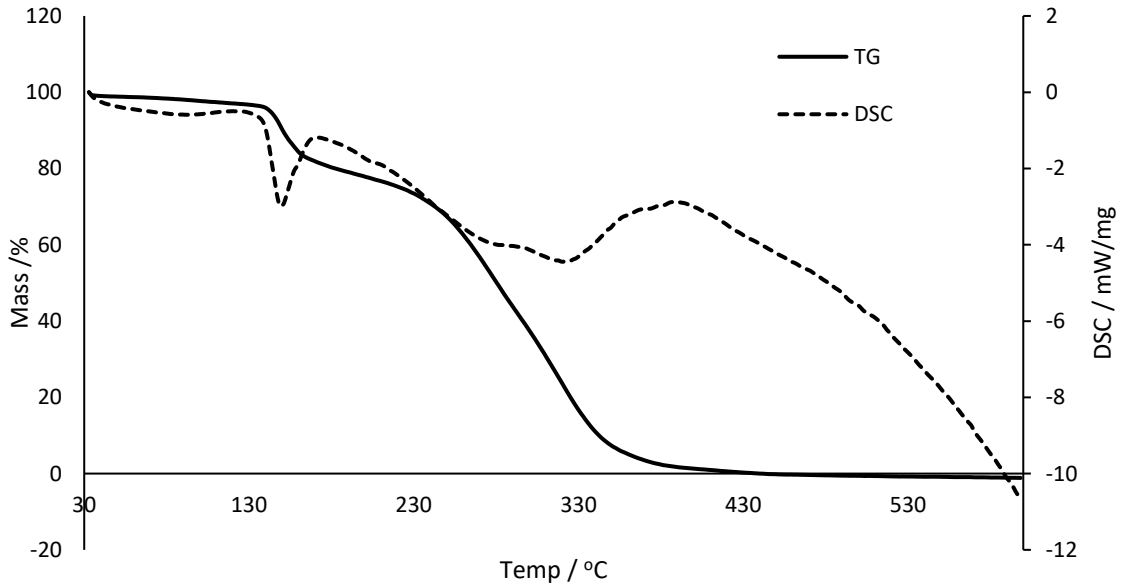


Figure S15 – TGA and DSC trace of **P_4** showing a crystalline melting point on the DSC trace at 150 °C, mass lost during melting, 12.4%. Mass loss is likely due to the release of encapsulated solvent upon melting.

SI 4.3.1.5 Single Crystal X-ray Diffraction of the Desolvated Phase (PP_4-SPF_{act})

A single crystal from the desolvated phase following gas adsorption was subjected to single crystal X-ray diffraction. The crystal quality was poor following the phase *trans*ition on de-solvation, with cracking of the crystallites and a significant decrease in both resolution and intensity. Nonetheless, a new unit cell could be indexed (**SI 12.3**) which was consistent with a *ca* 15% contraction of the *a* unit cell edge compared to the pristine sample, with smaller increases on *b* and *c* while maintaining monoclinic $P2_1$ symmetry. Despite the poor-quality diffraction an approximate structure model could be obtained suggesting a rearrangement and contraction of the individual PP4 molecules and closing of the solvent channels following de-solvation. The X-ray powder diffraction pattern simulated from this model closely matches that obtained from the bulk de-solvated material (Figure **S16**), which supports the assignment of the non-porous structure model.

SI 4.3.1.6 Re-solvation Experiments

A 30 mg sample of the desolvated material (following the gas adsorption experiments) was immersed in 5 mL of ethanol at room temperature in a sealed vial. Two samples were isolated from this suspension by filtration: one after 15 minutes, and one after 21 hours. Although both phases share an intense reflection at 10.1° (2θ), the immediate disappearance of the reflections at 6.8° and 7.2° in the desolvated phase and appearance of the reflection at 6.2°, as well as clear differences above 12° for the two phases indicate immediate regeneration of the solvated phase under these conditions.

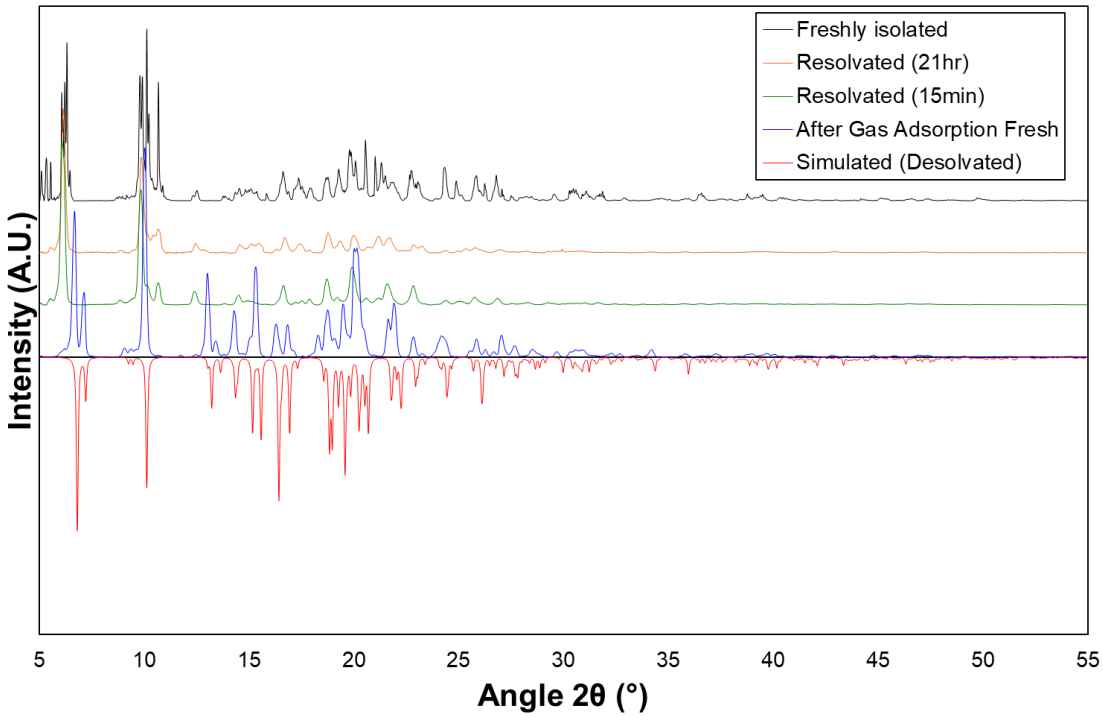


Figure S17 – Comparative X-ray powder diffraction patterns of **PP_4-SPF** desolvated (**blue**), following 15 minutes and 21 hours of soaking in ethanol (**green** and **yellow**, respectively), and the freshly isolated solvated material (**black**), compared with the simulated pattern for the desolvated sample (**red**).

SI 4.3.1.7 Crystal structure of PP_4-SPF_{Act}

The single crystal X-ray diffraction data for the desolvated material were collected with a Bruker D8 Quest ECO diffractometer using Mo-Kα radiation (λ = 0.71073 Å). A crystal was isolated from the sample following the gas adsorption experiments and mounted on a Mitegen micromount

in NVH immersion oil and cooled to 150 K using an Oxford cryostream. Data collection and reduction were controlled using the Bruker APEX-3 suite of programs.[7] Multi-scan absorption corrections were applied using SADABS.[8] All data were solved using the intrinsic phasing routine within SHELXT[6] and refined on F^2 using full-matrix least squares procedures with SHELXL[5] within the OLEX-2 package.[3] The crystal showed very poor single crystallinity after the evacuation of the solvent, and no useful reflections were observed beyond 1 Å resolution, and all observed reflections showed a high degree of mosaicity. While the diffraction data could be solved to give a connectivity model and accurate unit cell parameters, the poor data quality precludes any meaningful refinement of anisotropic displacement parameters. This model is presented purely as a comparative connectivity model for comparison with the observed X-ray powder diffraction data; no attempt has been made to restrain U_{iso} parameters to avoid over-interpretation of the model.

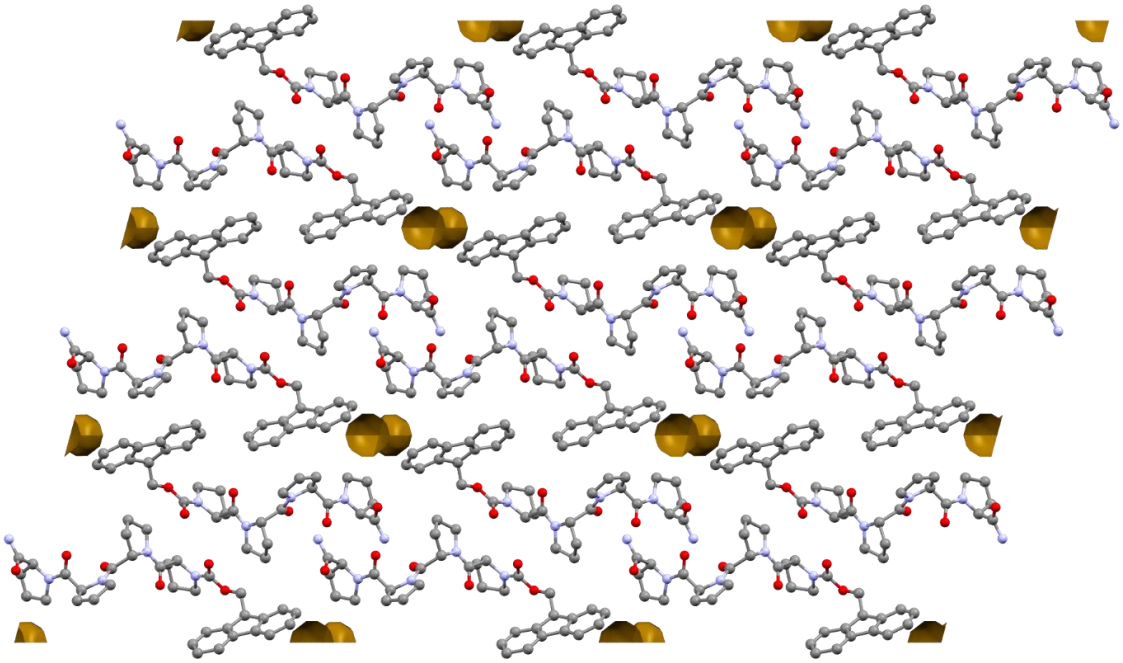


Figure S18 – Crystal structure of **PP_4-SPF$_{_Act}$** showing the desolvated structure highlighting significantly smaller void spaces ((yellow) 32.81 $Å^3$, 2.0%-unit cell volume; Probe radius 1.2 Å, grid spacing 0.7 $Å^2$). Ball and stick model, packing 3x3x3.

Crystal data and structure refinement for PP_4-SPF_{act}	
Identification code	**PP_4-SPF_{act} (**PP4dried)
Empirical formula	$C_{35}H_{41}N_5O_6$
Formula weight	627.73
Temperature/K	150.0
Crystal system	monoclinic
Space group	$P2_1$
a/Å	13.438(6)
b/Å	6.535(3)
c/Å	19.346(8)
α/°	90
β/°	105.094(9)
γ/°	90
Volume/Å^3	1640.2(12)
Z	2
ρ_{calc}g/cm^3	1.271
μ/mm^{-1}	0.088
F(000)	668.0
Crystal size/mm^3	0.34 × 0.11 × 0.03
Radiation	MoKα (λ = 0.71073)
2Θ range for data collection/°	264 to 41.794
Index ranges	-13 ≤ h ≤ 13, -6 ≤ k ≤ 6, -18 ≤ l ≤ 19
Reflections collected	14635
Independent reflections	3452 [R_{int} = 0.1698, R_{sigma} = 0.1423]
Data/restraints/parameters	3452/1/185
Goodness-of-fit on F^2	1.141
Final R indexes [I>=2σ (I)]	R_1 = 0.1758, wR_2 = 0.3106
Final R indexes [all data]	R_1 = 0.2034, wR_2 = 0.3283
Largest diff. peak/hole / e Å^{-3}	0.59/-0.45
Flack parameter	-2.6(10)
CCDC No.	2156434

SI 4.3.1.8 Framework Host-Guest chemistry:

Subsequent to de-solvation of the framework and the observation of apparent reinflation upon soaking in ethanol solution (Figure S17), reinflation on **PP_4-SPF_{act}** with various guest molecules was then attempted (Table **S3**). Both ^{1}H NMR (SI 13.3) and PD-XRD (SI 13.4) were used to confirm the degree of encapsulation of guest molecules and the change from the starting phase of the desolvated material after soaking with the guest molecule. The percentage of encapsulated guest, hexane (used as guest diluent) and residual ethanol obtained from ^{1}H NMR (SI 13.3) analyses are summarized in the table below (Table S3). Chiral HPLC was then used to determine the enantioselectivity of the SPF towards 1-Phenylethan-1-ol (SI13.5).

Table S3: Determination of encapsulated guest[d]					
Entry	**Sample treatment**	**Time /h**	**guest content /mol%[c]**	**Hexane content /mol%[c]**	**Resid. EtOH /mol%[c]**
1	Washed with **hexane** (**PP_4_SPF**)	-	-	8	91
2	**Activated** under vac 45 °C (**PP_4-SPF_{act}**)	16	-	-	12
3	**EtOH 5%** /Hex	1.5	54	61	-
4	**EtOH vapour**	26	14	-	-
5	(±)**1-phenylethanol 5%** /Hex	1.5	36	20	8.
6	(±)**1-phenylethanol 5%** /Hex	16	59	36	6
7	(±)**1-phenylethanol 2 eq.** /Hex	1.5	33	17	5
8	Amorphous **PP_4** Inc. (±)**1-phenylethanol** 5%/Hex	1.5	7	16	78
9	**Acetone 5%** /Hex	16	28[a]	51	0
10	**THF 5%** /Hex	16	23[a]	48	2
11	EtOAc **5%** /Hex	16	33	50	2
12	**Hexane**	16	-	4	7
13	**Toluene 5%** /Hex	16	33	30	5
14	**1-Bromohexane 5%** /Hex	16	15	32	2
15	**I_2** saturated hexane	336	12[b]	18	2

[a] Integral of clean spectrum of **PP_4** at same ppm range subtracted from solvent peak integral for% calculation due to overlapping peaks of peptide and guest

[b] Iodine content determined *via* molecular occupancy from SC-XRD data

[c] All guest mol% calculated against moles of peptide

[d] **General procedure for guest soaking:** A sample of **PP_4-SPF_{act}** (≈5-10 mg) was placed in a premixed solution of guest (typically 5%) in hexane (300 µL, HPLC grade) for 16 hours before decanting the solution and washing five times with fresh hexane to remove excess guest. The solid was then taken for PD-XRD analysis (**SI 13.4**). The sample was then fully dissolved in MeOD before recording the ^{1}H NMR spectra (**SI 13.3**). The peaks at 7.8-7.85 ppm for two of the Fmoc CH protons were used as the peptide reference, set as 2H, and the integrals of the guest's peaks were used to calculate the mole percentage against the peptide (i.e. 50% guest = 2 eq. peptide: 1 eq. guest).

For (±) 1-Phenylethan-1-ol studies the NMR sample was then dried under compressed air before dissolving in propan-2-ol (30 µL), and diluting with hexane (570 µL) causing the peptide to precipitate. This suspension was then filtered through a 0.2 µm syringe filter and the filtrate taken for chiral HPLC analysis (*further details **SI 13.5***).

SI 4.3.1.9 Crystal Structure of PP_4-SPF@I_2

The crystal structure was obtained from orange/yellow crystals, crystallised from a supersaturated hot ethanol solution as previously, before they were dried under vacuum at 45 °C overnight and covered with a saturated solution of iodine in hexane for a week. The crystals were stable outside of solution at room temperature showing no signs of deterioration over the timeframe of the experiment.

The experiment was carried out by mounting a crystal without using an inert oil to maximize the guest loss. The crystal was mounted at 290 K, then cooled down to 150 K with a cooling rate of 280 K/h and analyzed *via* single crystal X-ray diffraction. A second data collection was performed after heating the same crystal directly on the mount at 323 K for 4 hours and cooling it down to 150 K with the same rate used for the previous collection. The guest molecules within the pores have been modeled for each of the two sets of data in the same way: the occupancies of the

iodine atoms were allowed to freely refine while keeping the atoms isotropic and fixing the isotropic atomic displacement parameters to the same value. The analogous treatment of the two data sets allowed us to obtain the values of 12.3% and 5.8% respectively for the chemical occupancies of the iodine molecules respectively before and after the thermal treatment.

In the structures reported (CCDC-2127748 and CCDC-2127749), before and after the thermal treatment, the iodine atoms were treated anisotropically and their anisotropic displacement parameters were kept equivalent. For the **PP_4-SPF@I_2** structure the carbon atom C14 of the ring of pro2 (C_γ) was split into two components with a total occupancy of 1. The anisotropic displacement parameters were set to be equivalent for this atom and for the nearest carbon atoms. For the **PP_4-SPF@I_{2_Heated}** structure three carbon atoms (C13, C14 and C15) of the ring of pro2 ($C_{\beta,\gamma,\delta}$) were split into two components with a total occupancy of 1. The anisotropic displacement parameters were set to be equivalent for these atoms in both the components. Within the pro2 ring, both the N-C and C-C distances were fixed to be the same between the two components.

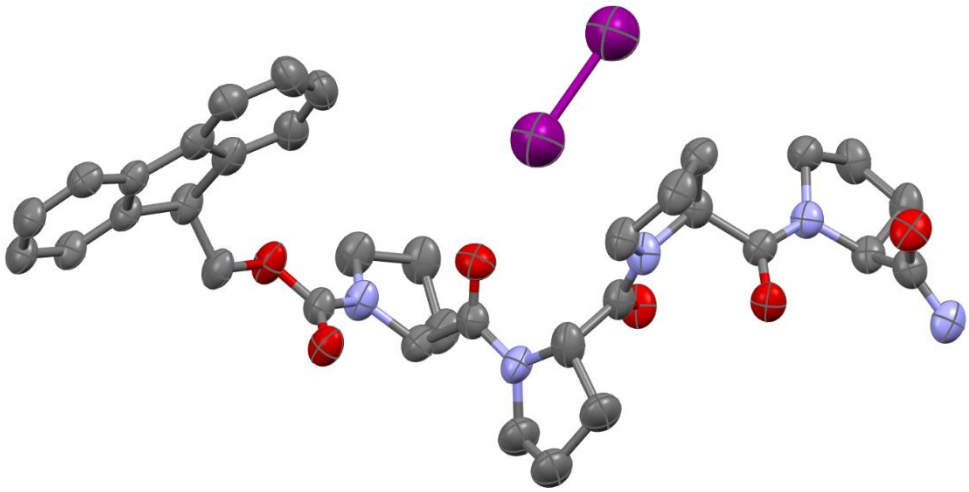


Figure S18 – Mercury ellipsoid representation (50% probability) of the asymmetric unit of **PP_4-SPF@I_2**, hydrogen atoms removed for clarity

SI 4.3.1.10 Crystal data and structure refinement for PP_4-SPF@I_2 and PP_4-SPF@I_{2_Heated}

Identification code	**PP_4-SPF_I_2**	**PP_4-SPF_I_{2_Heated}**
Empirical formula	$C_{35}H_{41}I_{0.27}N_5O_6$	$C_{35}H_{40}I_{0.14}N_5O_6$
Formula weight	661.99	645.12
Temperature/K	150(2)	150(2)
Crystal system	monoclinic	monoclinic
Space group	$P2_1$	$P2_1$
a/Å	16.523(2)	16.5376(15)
b/Å	6.3229(5)	6.2977(4)
c/Å	18.208(4)	18.2023(16)
α/°	90	90
β/°	109.803(19)	109.744(10)
γ/°	90	90
Volume/Å^3	1789.8(5)	1783.1(3)
Z	2	2
ρ_{calc}g/cm^3	1.228	1.201
μ/mm^{-1}	2.487	1.644
F(000)	697.0	681.0
Crystal size/mm^3	0.235 × 0.073 × 0.035	0.235 × 0.073 × 0.035
Radiation	Cu Kα (λ = 1.54184)	Cu Kα (λ = 1.54184)
2Θ range for data collection/°	8.882 to 143.29	8.872 to 143.6
Index ranges	-20 ≤ h ≤ 13, -7 ≤ k ≤ 7, -22 ≤ l ≤ 22	-20 ≤ h ≤ 20, -7 ≤ k ≤ 7, -22 ≤ l ≤ 13
Reflections collected	12910	13057
Independent reflections	6762 [R_{int} = 0.0564, R_{sigma} = 0.0785]	6779 [R_{int} = 0.0442, R_{sigma} = 0.0598]
Data/restraints/parameters	6762/1/421	6779/8/426
Goodness-of-fit on F^2	1.072	1.059
Final R indexes [I>=2σ (I)]	R_1 = 0.1062, wR_2 = 0.2855	R_1 = 0.0952, wR_2 = 0.2548
Final R indexes [all data]	R_1 = 0.1311, wR_2 = 0.3075	R_1 = 0.1140, wR_2 = 0.2763
Largest diff. Peak/hole / e Å^{-3}	0.85/-0.36	1.25/-0.44
Flack parameter	0.08(3)	0.31(3)
CCDC No.	2127749	2127748

SI 4.3.1.11 Crystal Structure of PP_4-SPF@$CHCl_3$

The crystal structure was obtained from colourless crystalline needles, crystallised from a supersaturated hot ethanol solution as previously, before they were then incubated in a 10% solution of chloroform in hexane. The crystals were stable outside of solution at room temperature showing no signs of deterioration over the timeframe of the experiment.

The experiment was carried out by mounting a crystal in inert oil. The crystal was mounted at 293 K analyzed *via* single crystal X-ray diffraction.

A second data collection was performed after heating the same crystal directly on the mount at 323 K for 4 hours and cooling it down to 150 K with the same rate used for the previous collection. The guest molecules within the pores have been modeled for each of the two sets of

data in the same way: the occupancies of the iodine atoms were allowed to freely refine while keeping the atoms isotropic and fixing the isotropic atomic displacement parameters to the same value. The analogous treatment of the two data sets allowed us to obtain the values of 12.3% and 5.8% respectively for the chemical occupancies of the iodine molecules respectively before and after the thermal treatment. In the structure (CCDC-)e reported the carbon atom C22 at position 4 on Pro2 was split into two components with a total occupancy of 1 and freely refined to give an occupancy of 0.34 for the carbon in with *endo* puckering (C22B) and 0.66 for the carbon in with *exo* puckering (C22A). The anisotropic displacement parameters were set to be equivalent for these atoms in both the components. The chemical occupancy of the $CHCl_3$ atoms were freely refined isotropically before fixing at an occupancy of 0.55, the atoms were then treated anisotropically and their anisotropic displacement parameters were kept equivalent.

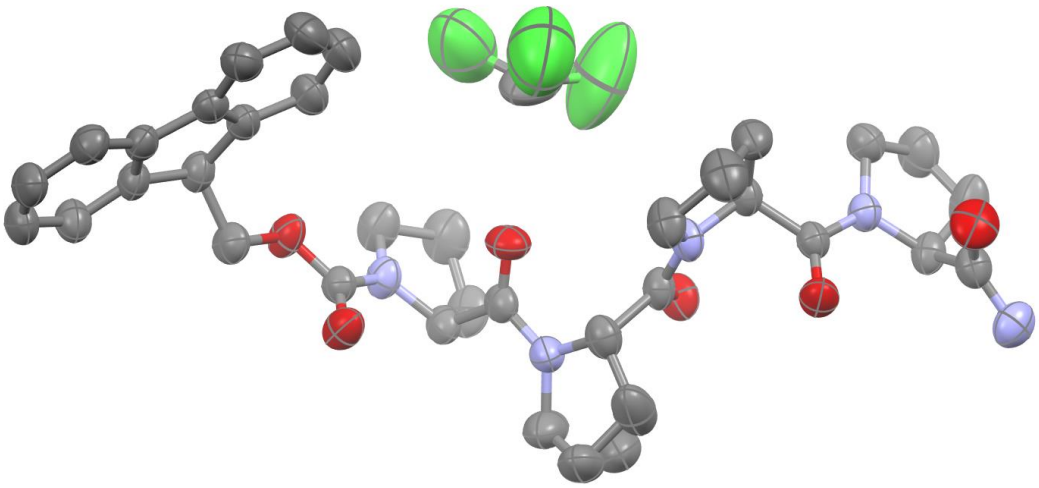


Figure S18b – Mercury ellipsoid representation (50% probability) of the asymmetric unit of **PP_4-SPF@$CHCl_3$**, hydrogen atoms removed for clarity

SI 4.3.1.12 Crystal data and structure refinement for PP_4-SPF@$CHCl_3$

Identification code	FmocP4_CHCl3
Empirical formula	$C_{35.55}H_{41.55}Cl_{1.65}N_5O_6$
Formula weight	693.38
Temperature/K	293
Crystal system	monoclinic
Space group	$P2_1$
a/Å	16.4779(4)
b/Å	6.37150(10)
c/Å	18.4425(5)
α/°	90
β/°	109.220(3)
γ/°	90
Volume/Å^3	1828.33(8)
Z	2
ρ_{calc}g/cm^3	1.259
μ/mm^{-1}	1.773
F(000)	732.0
Crystal size/mm^3	0.54 × 0.13 × 0.07
Radiation	Cu Kα (λ = 1.54184)
2Θ range for data collection/°	8.78 to 147.374
Index ranges	-20 ≤ h ≤ 16, -7 ≤ k ≤ 7, -18 ≤ l ≤ 22
Reflections collected	20453
Independent reflections	7196 [R_{int} = 0.0384, R_{sigma} = 0.0368]
Data/restraints/parameters	7196/3/461
Goodness-of-fit on F^2	1.055
Final R indexes [I>=2σ (I)]	R_1 = 0.0692, wR_2 = 0.2116
Final R indexes [all data]	R_1 = 0.0816, wR_2 = 0.2212
Largest diff. peak/hole / e Å^{-3}	0.37/-0.33
Flack parameter	0.14(5)
CCDC No.	2288009

SI 4.3.1.13 ^{1}H NMR studies of SPF after guest encapsulation

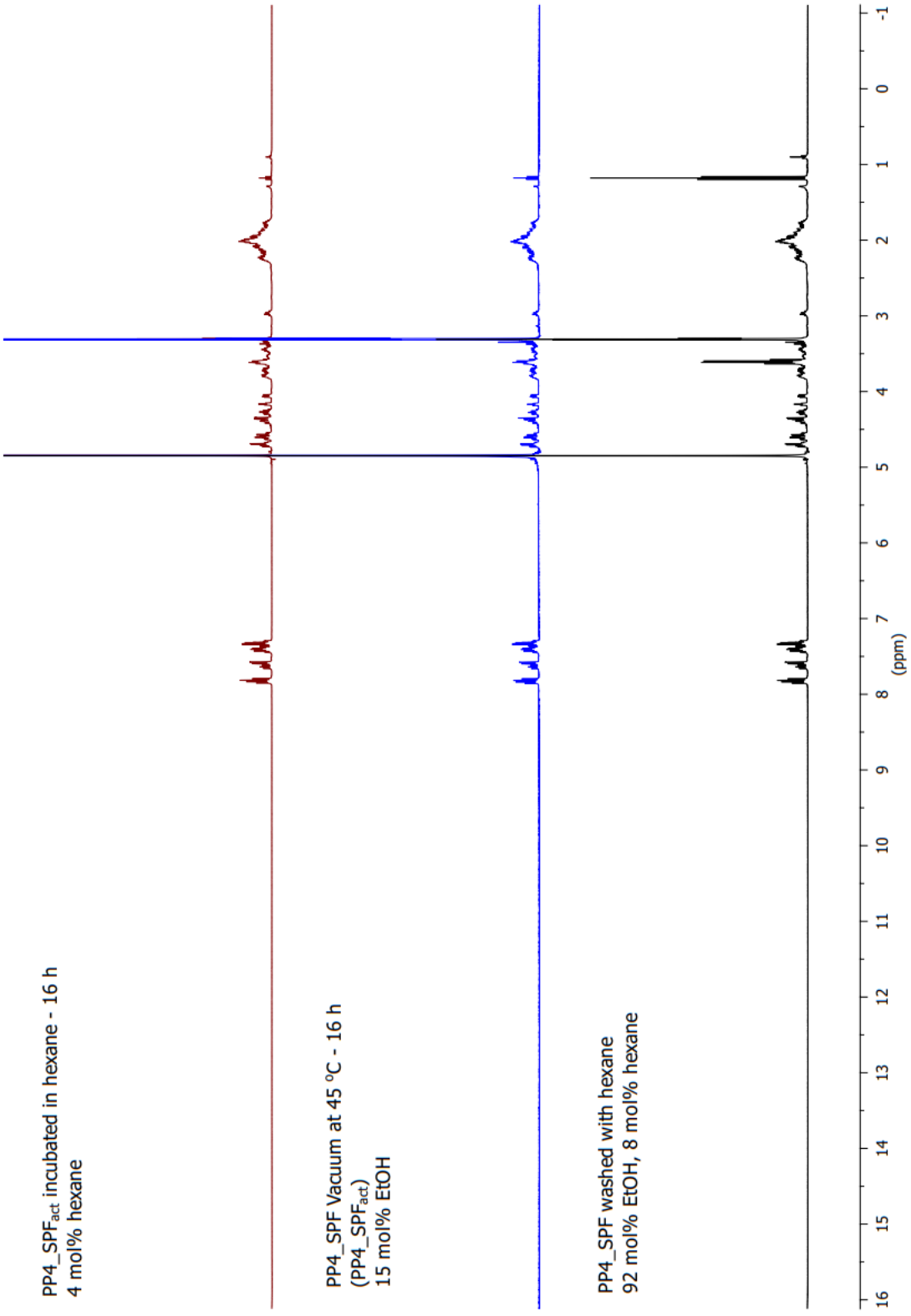


Figure S19 - ^{1}H NMR (400 MHz, MeOD) of **PP_4-SPF** washed with hexane (HPLC grade, x 5) (**black**, bottom), PP4-SPF after activation at 45 °C under vacuum for 16 h (**blue**), and **PP_4-SPF_{act}** after soaking in hexane for 16 h (**red**, top).

SI 4.3.1.14 Powder Diffraction data after SPF_{act} guest soaking

X-ray powder diffraction patterns for host-guest studies were collected on a Rigaku Miniflex 600 using a Cu radiation source and measurements were performed at room temperature (Scan range 2Θ; 5-55°, Step; 0.02 °, scan speed; 0.8 °min^{-1}).

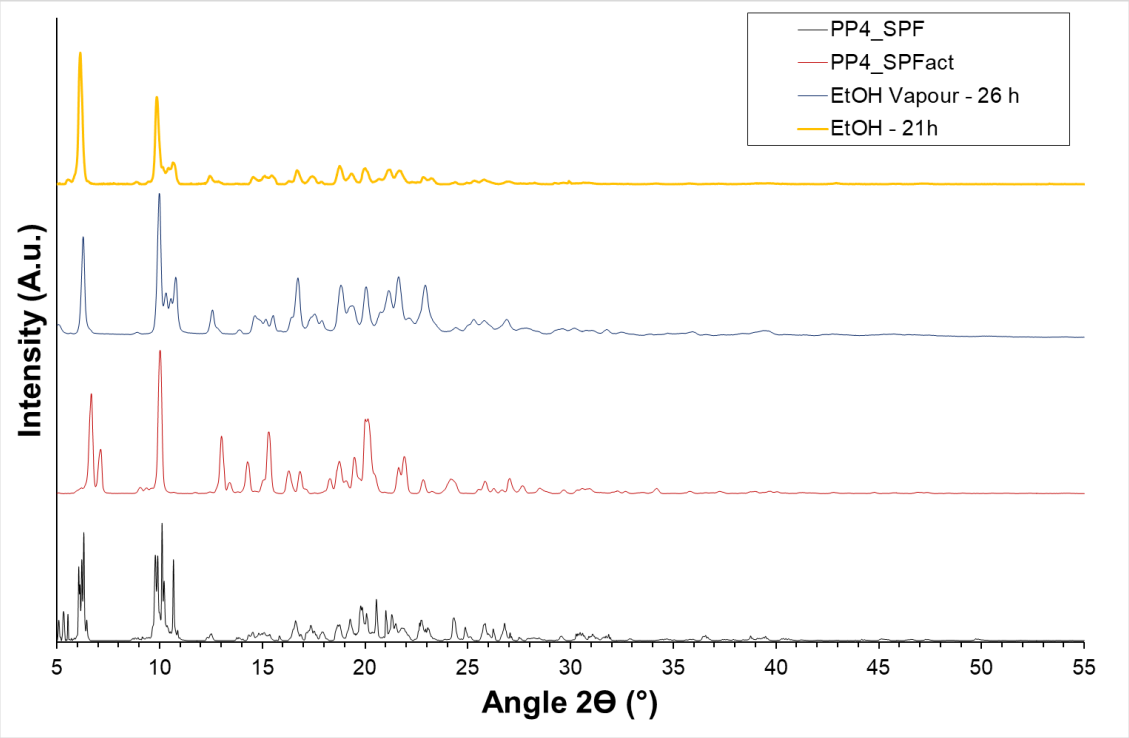


Figure S20 – PD-XRD of fresh crystals of **PP_4-SPF** unground (black, bottom), **PP_4-SPF** after activation at 45 °C under high vacuum (provided by a turbomolecular pump) (**red**), **PP_4-SPF_{act}** after soaking in a chamber saturated with ethanol vapour for 26 h (**blue**), and **PP_4-SPF_{act}** after soaking in ethanol solution for 21 h.

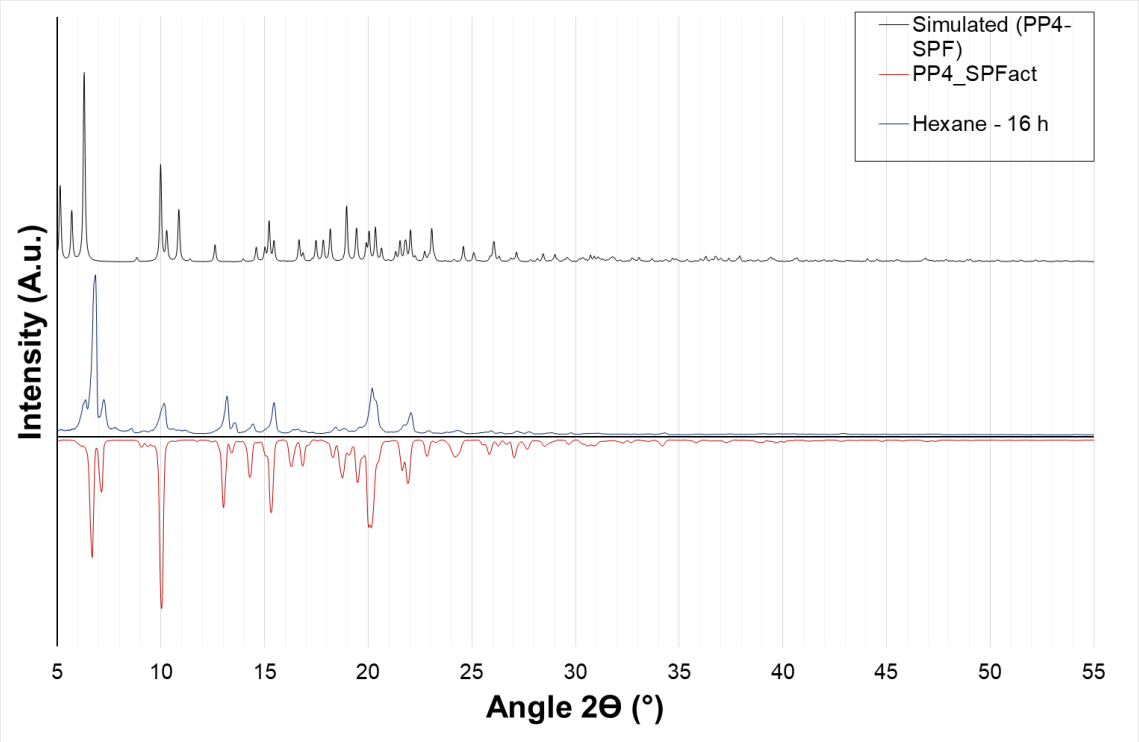


Figure S21 – PD-XRD simulated for **PP_4-SPF** (**black**, top), **PP_4-SPF** after activation at 45 °C under high vacuum (provided by a turbomolecular pump) (**red**), **PP_4-SPF_{act}** after soaking in hexane (**blue**)

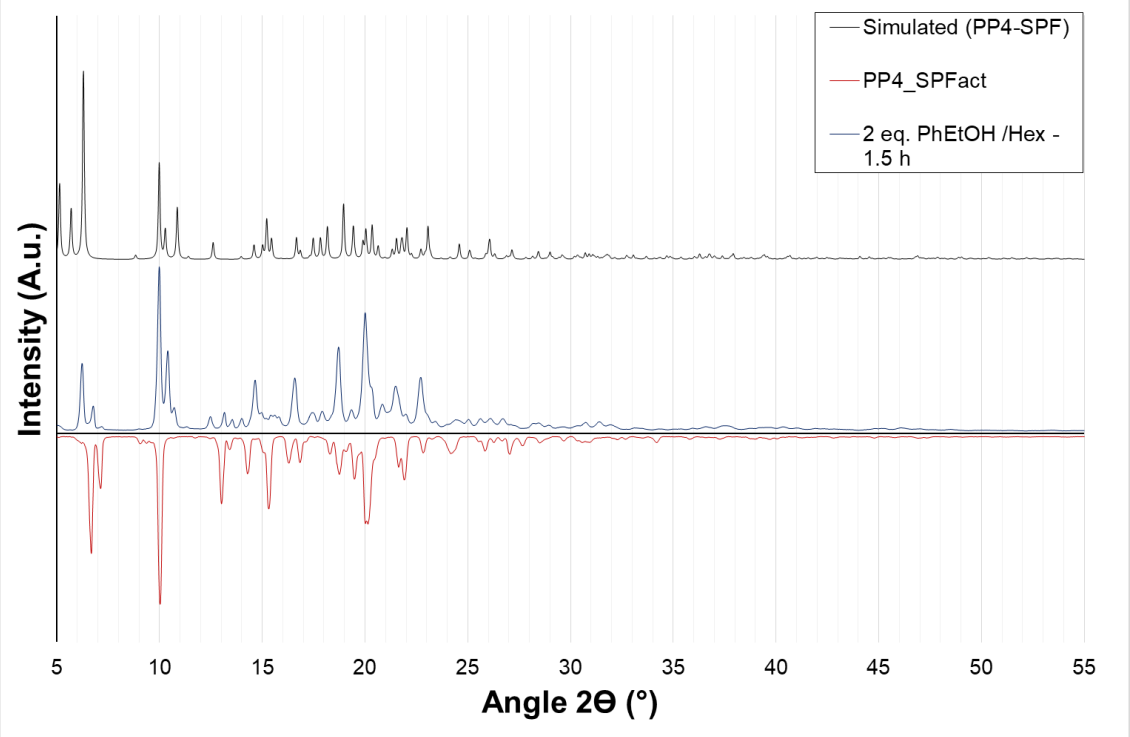


Figure S22 – PD-XRD simulated for **PP_4-SPF** (black, top), **PP_4-SPF** after activation at 45 °C under high vacuum (provided by a turbomolecular pump) (**red**), **PP_4-SPF_{act}** after soaking in a 1-phenylethanol solution (2 eq. in hexane 300 µL, HPLC grade) for 1.5 h (**blue**)

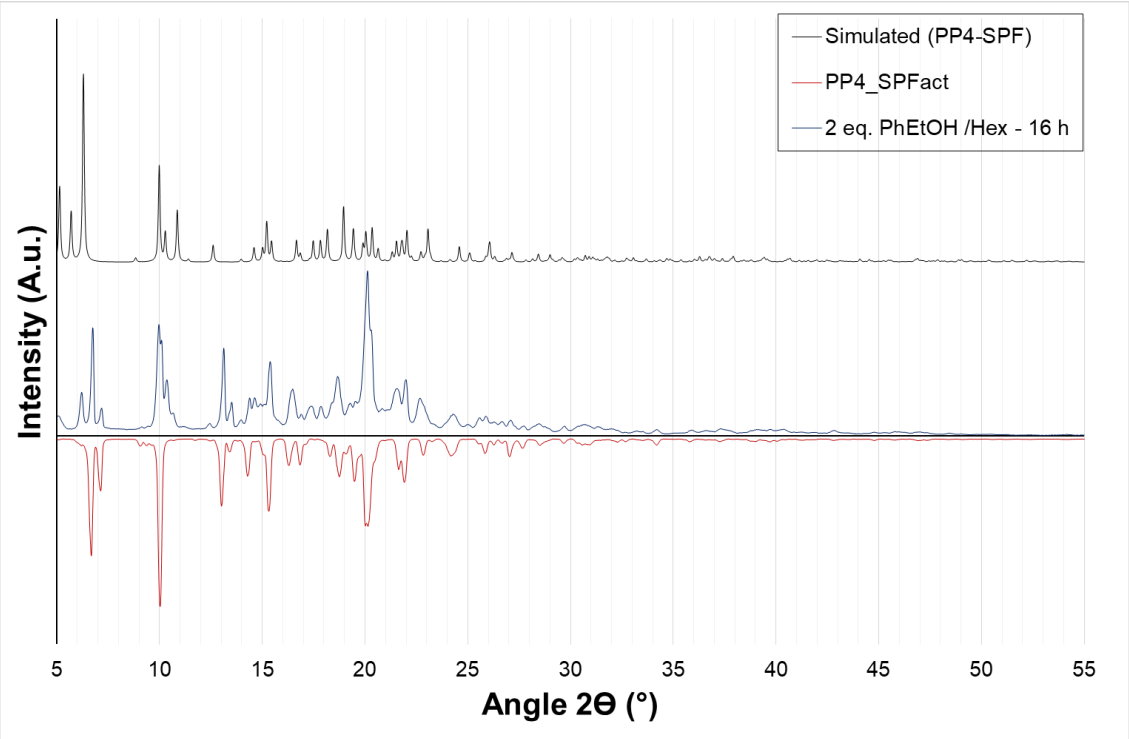


Figure S23 – PD-XRD simulated for **PP_4-SPF** (black, top), **PP_4-SPF** after activation at 45 °C under high vacuum (provided by a turbomolecular pump) (**red**), **PP_4-SPF_{act}** after soaking in a 1-phenylethanol solution (2 eq. in hexane 300 µL, HPLC grade) for 16 h (**blue**)

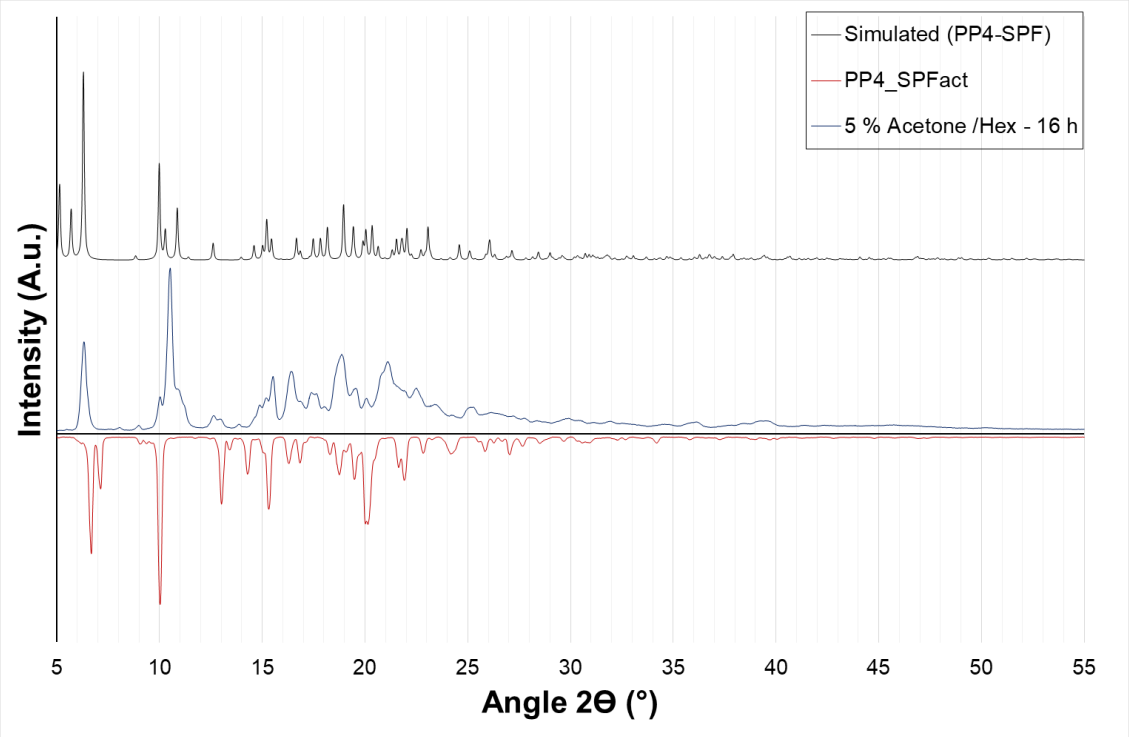


Figure S24 – PD-XRD simulated for **PP_4-SPF** (black, top), **PP_4-SPF** after activation at 45 °C under high vacuum (provided by a turbomolecular pump) (**red**), **PP_4-SPF_{act}** after soaking in a acetone solution (5% in hexane 300 µL, HPLC grade) for 16 h (**blue**)

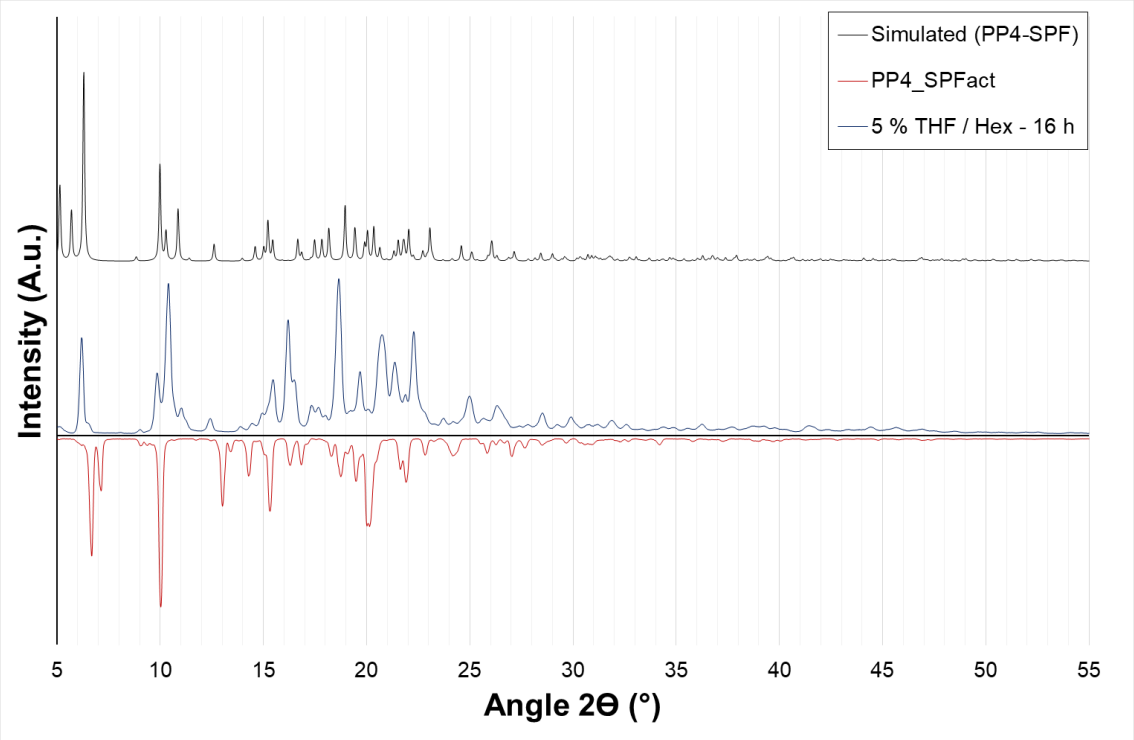


Figure S25 – PD-XRD simulated for **PP_4-SPF** (black, top), **PP_4-SPF** after activation at 45 °C under high vacuum (provided by a turbomolecular pump) (**red**), **PP_4-SPF_{act}** after soaking in a acetone solution (5% in hexane 300 µL, HPLC grade) for 16 h (**blue**)

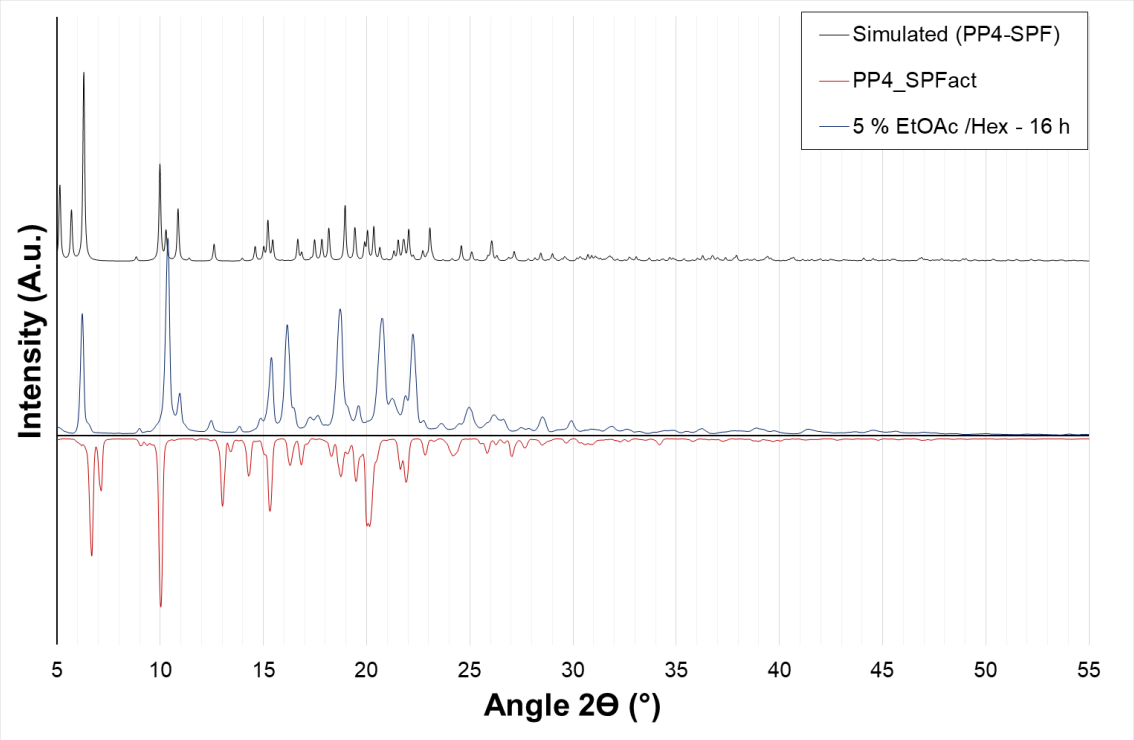


Figure S26 – PD-XRD simulated for **PP_4-SPF** (black, top), **PP_4-SPF** after activation at 45 °C under high vacuum (provided by a turbomolecular pump) (**red**), **PP_4-SPF_{act}** after soaking in a THF solution (5% in hexane 300 µL, HPLC grade) for 16 h (**blue**)

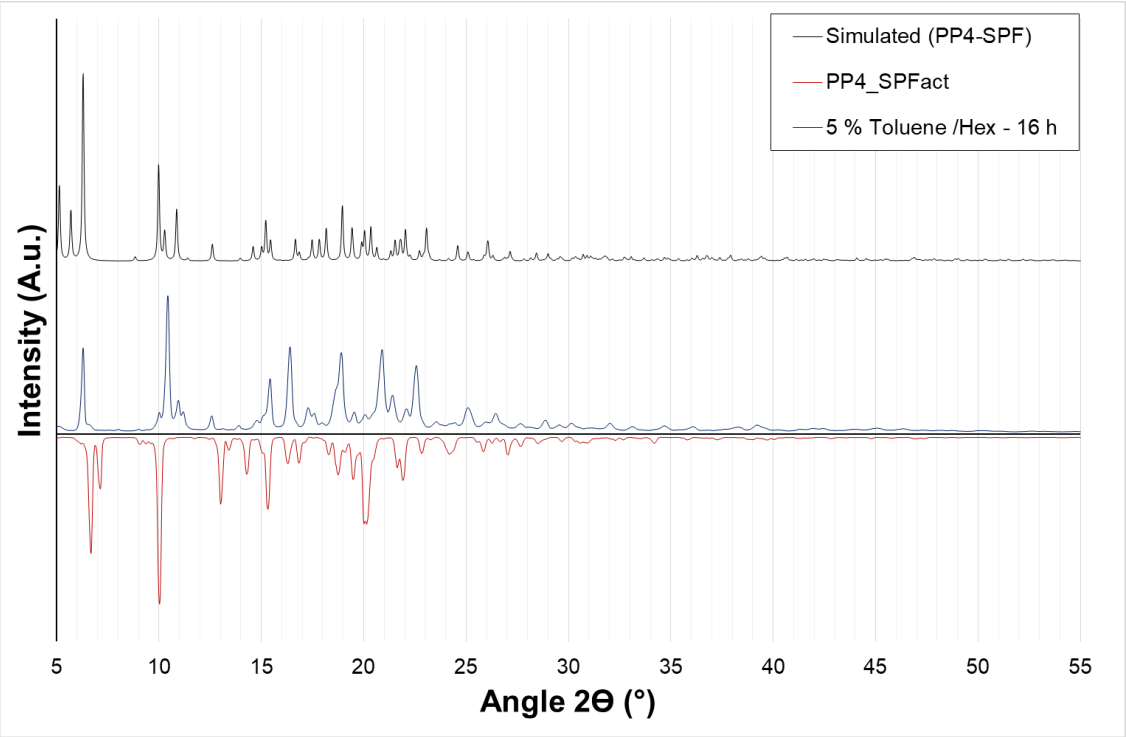


Figure S27 – PD-XRD simulated for **PP_4-SPF** (black, top), **PP_4-SPF** after activation at 45 °C under high vacuum (provided by a turbomolecular pump) (**red**), **PP_4-SPF_{act}** after soaking in a toluene solution (5% in hexane 300 µL, HPLC grade) for 16 h (**blue**)

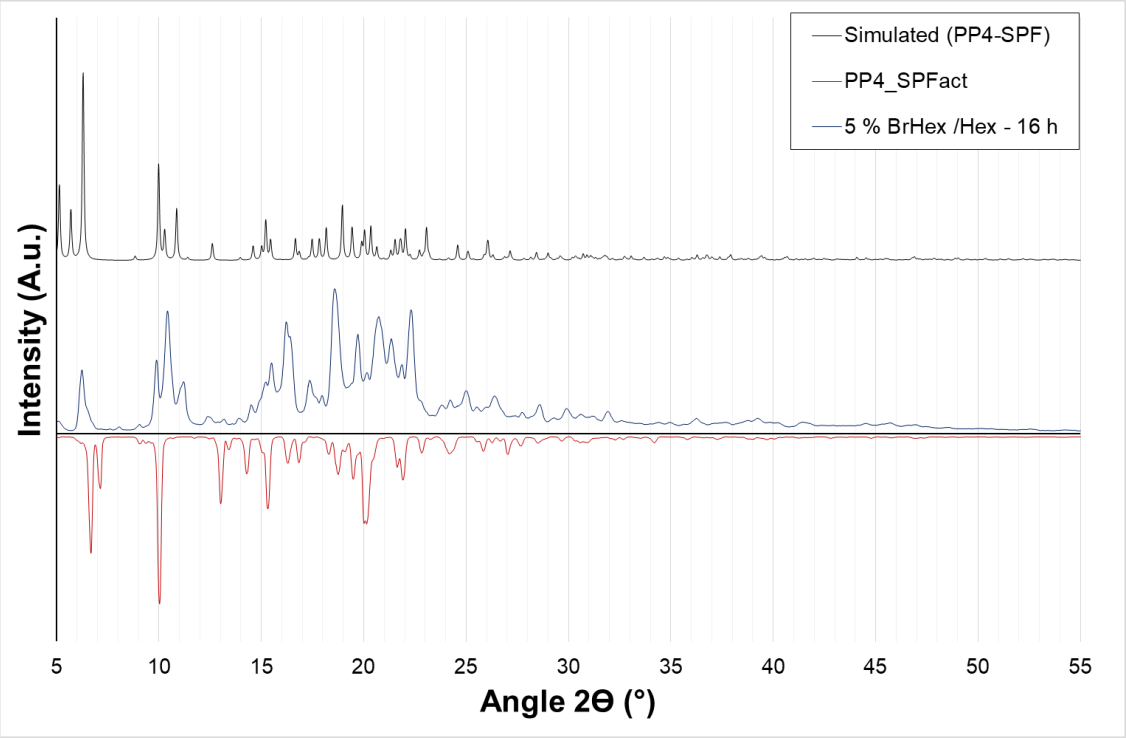


Figure S28 – PD-XRD simulated for **PP_4-SPF** (black, top), **PP_4-SPF** after activation at 45 °C under high vacuum (provided by a turbomolecular pump) (**red**), **PP_4-SPF_{act}** after soaking in a 1-bromohexane solution (5% in hexane 300 µL, HPLC grade) for 16 h (**blue**)

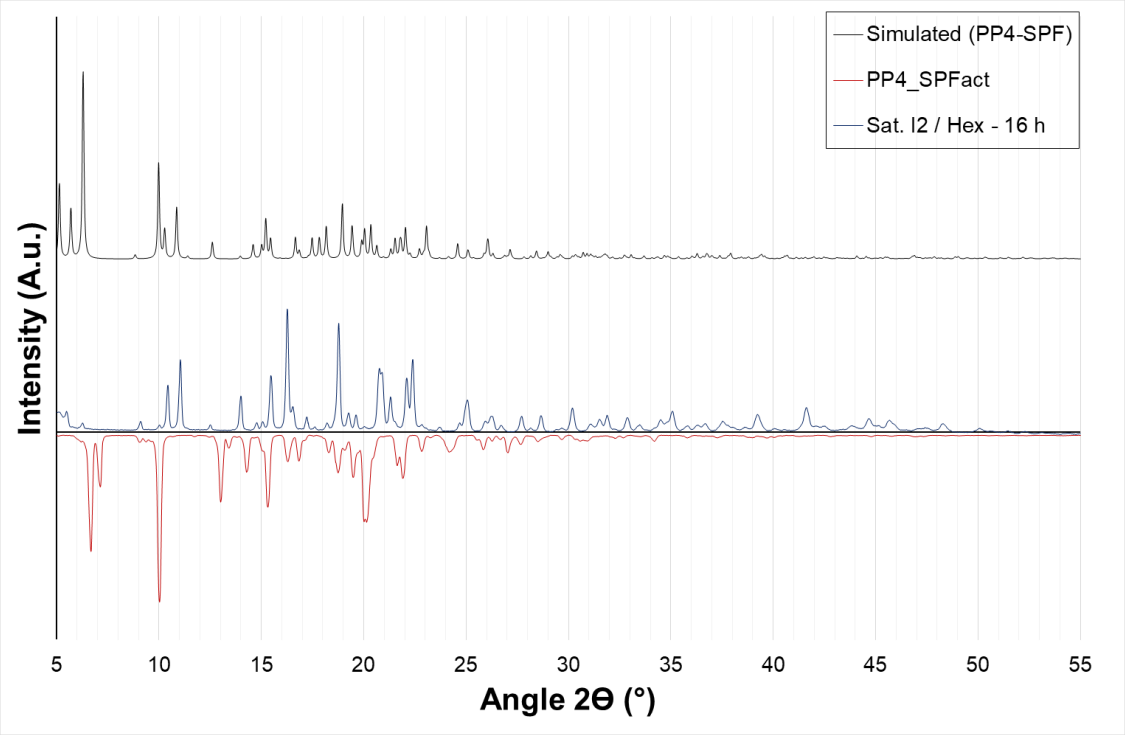


Figure S29 – PD-XRD simulated for **PP_4-SPF** (**black**, top), **PP_4-SPF** after activation at 45 °C under high vacuum (provided by a turbomolecular pump) (**red**), **PP_4-SPF_{act}** after soaking in a saturated iodine solution (hexane 300 µL, HPLC grade) for 16 h (**blue**)

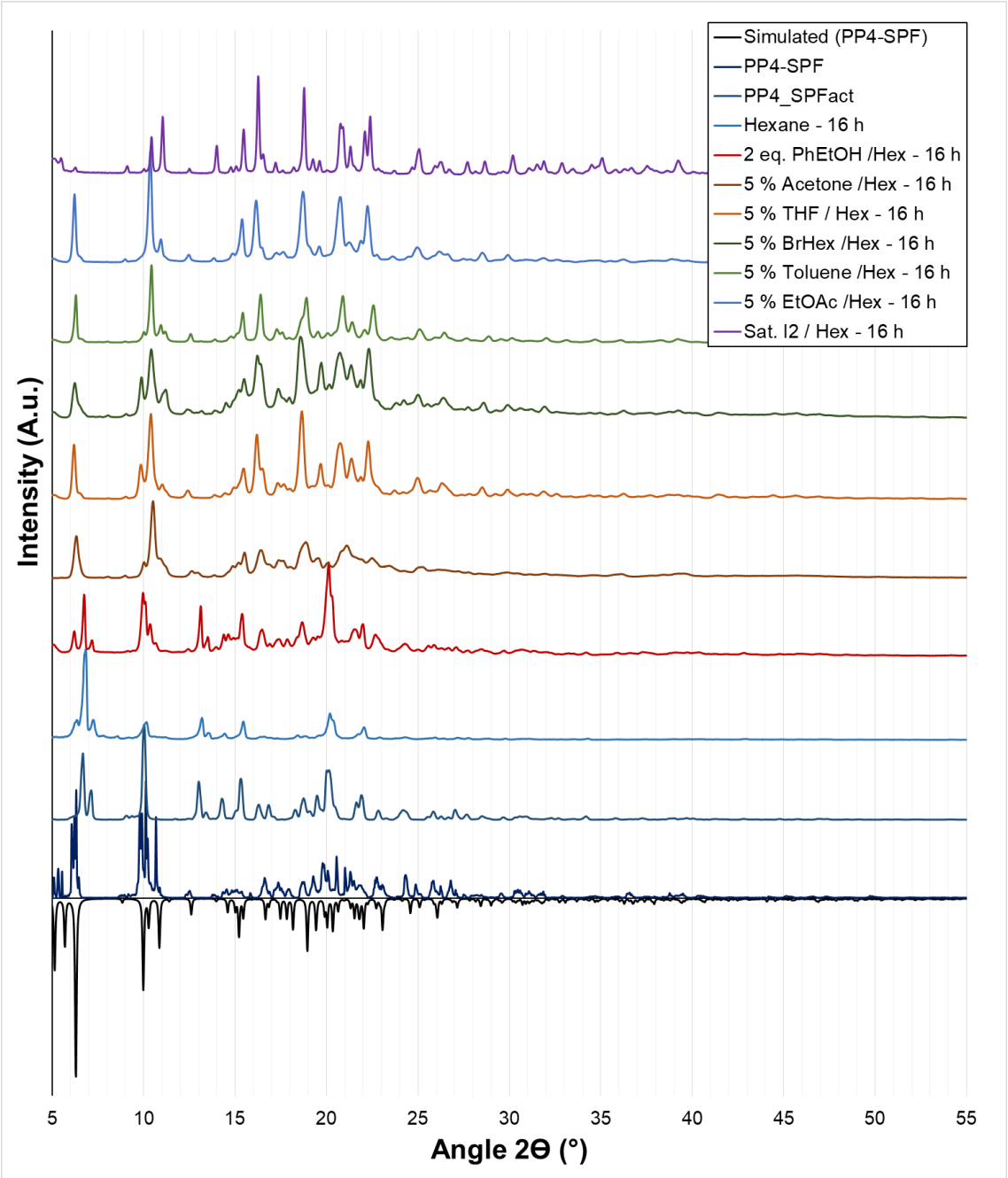


Figure S30 – Experimental PD-XRD of all studied guest molecules compared with original **PP_4-SPF** simulated, experimental, and after activation. Legend in order, with simulated (*bottom*) to iodine (*top*).

SI 4.3.1.15 Enantioselectivity studies of PP_4-SPF_{act} for (±) 1-Phenylethan-1-ol

Chiral separation – Samples of **PP_4-SPF_{act}** (5-10 mg) were soaked in a premixed solution of (**±**) 1-phenylethan-1-ol in hexane and left for a set time without agitation. The solution was then decanted and the solid washed with hexane (300 µL x 5, HPLC grade) to remove excess 1-phenylethanol. The samples were then dissolved in propan-2-ol (30 µL), to release encapsulated guest, before precipitating with hexane (570 µL). The dispersion was then filtered through a syringe filter (0.2 µm) before taking for HPLC analysis. Where the SPF was used in excess, the supernatant after soaking and first hexane wash were collected, combined, filtered and taken for HPLC analysis without dilution.

HPLC analysis - The determination of enantiomeric selectivity was investigated *via* normal phase chiral HPLC. These were carried out with an isocratic mobile phase (A: 98% hexane, B: 2% propan-2-ol, 1 mL min^{-1} flow rate, 20 min runtime, 298 K) on an Agilent 1100 Series Capillary LC system with a CHIRALCEL® OD chromatographic column (*6 mm x 250 mm*, *10 µm*). Samples used 40 µL injections in 5% propan-2-ol in hexane, the target analytes were monitored at 254 nm. *R*-1-phenylethan-1-ol rt = 10.45 min, *S*-1-phenylethan-1-ol rt = 13.07 min.

Table S4: HPLC Chromatogram data for *R* and *S*-1-phenylethanol from consecutive repeats of PP_4_SPF_{act} soaking with 1-phenylethanol (5% in hexane)

	Peak *rt* / min	**Peak area /%**	***ee* /%**
Run 1	10.111	38.783	22.434
	12.547	61.217	
Run 2	10.578	36.717	26.566
	13.220	63.283	
Run 3	10.661	38.029	23.942
	13.253	61.971	
Average	10.450	37.843	**23.114**
	13.067	62.157	

SI 4.3.1.16 Computational Methods

Molecular Orbital and Electrostatic Potential Modelling

Summary: Energy decomposition analysis (EDA) of the isolated dimer system (geometry fixed to that of the **PP_4-SPF** crystal structure) was used to determine the relative breakdown of the intermolecular forces between the Fmoc groups in the crystal lattice. Calculations were run with two different functionals, BLYP-D3(BJ) and PBE-D, giving interaction energies of -42 and -32.4 $kJmol^{-1}$ respectively. The breakdown suggests that dispersion interactions are the predominant interaction between the Fmoc groups with smaller contributions from electrostatic and orbital interactions (and the Pauli repulsive interaction, Table S5).

The Fmoc-Fmoc dimer was modelled by density functional theory at the B3LYP-D3/6-311++G (p,d) level using GAMESS version 5 (Dec 2014) with Avogadro as the GUI(v. 1.2.7).[9,10] The input geometry was that of two **PP_4** molecules associated at the terminal Fmoc ends (Figure S31)

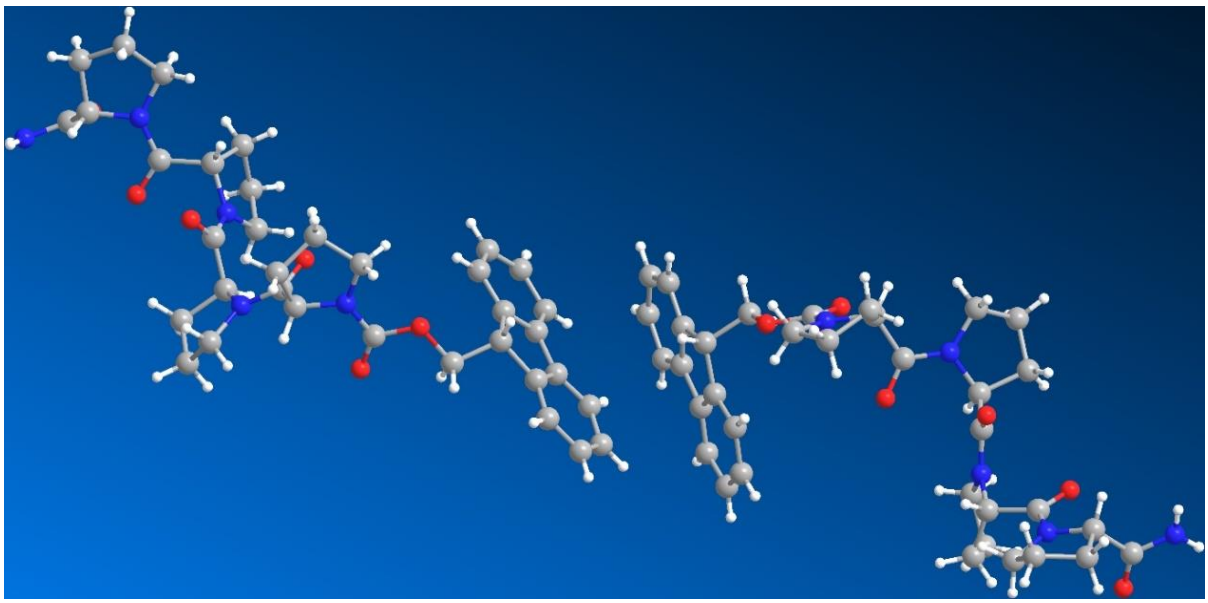


Figure S31 - Fmoc-Fmoc dimer taken from crystal structure and used as fixed geometry for subsequent molecular model calculations

The input for GAMESS is given below.

```
$BASIS GBASIS=N311 NGAUSS=6 NDFUNC=1 NPFUNC=1 DIFFSP=.TRUE. DIFFS=.TRUE. $END
 $CONTRL SCFTYP=RHF RUNTYP=ENERGY DFTTYP=B3LYP $END
 $SCF DIRSCF=.T. DAMP=.T. SOSCF=.F. $END
 $ELPOT IEPOT=1 WHERE=PDC OUTPUT=NONE $END
 $PDC PTSEL=CONNOLLY $END
 $SYSTEM MWORDS=100 $END
 $DATA
Title
C1
O   8.0    1.63200    3.86900    9.38200
O   8.0    7.19400    6400   13.19100
O   8.0    3.69000    2.92300    9.11200
O   8.0    8.76800    2.59100   16.50300
O   8.0    0400    3.62300   12.57200
O   8.0    9.52000    2.02400   19.65400
N   7.0    8.00300    3.89400   18.17000
N   7.0    3.41100    84200   10.29900
N   7.0    6.40200    2.50800   15.18600
N   7.0    6.49100    3.59300   11.76600
N   7.0   11.43500    3.04100   18.99200
H   1.0   11.93900    2.41000   19.28700
H   1.0   11.78500    3.72800   18.61000
C   6.0   10.10500    2.95600   19.12700
C   6.0    7.82000    3.14200   17.07400
C   6.0    6.83000    3.30100   120200
C   6.0    5.21300    3.99100   11.72000
C   6.0   -0.96500    1.76200    7.73800
C   6.0   -1.41200    0.56000    8.29800
C   6.0   -1.15500    0.59200    9.74000
C   6.0    2.98100    3.81000    9.56100
C   6.0    6.38900    3.01400   16.55700
H   1.0    5.94000    3.88400   16.59100
C   6.0   -1.99700   -0.41500    7.49500
H   1.0   -2.31800   -1.20600    7.86500
C   6.0   -0.53800    1.81200   10.05400
C   6.0    6.97100    52500   18.99400
H   1.0    6.64500    3.91600   19.67700
H   1.0    6.22300    82500   18.45400
C   6.0   -0.35800    2.64600    8.81600
H   1.0   -0.84800    3.49100    8.89100
C   6.0   -1.42500   -0.35300   10.73100
H   1.0   -1.84400   -1.15600   10.52100
C   6.0    9.35600    19700   18.65300
H   1.0    9.86900    66200   17.95900
C   6.0    6.92800    2.67000   12.82200
H   1.0    6.40000    1.84500   12.79000
C   6.0   -1.61800    1.00100    5.56500
H   1.0   -1.67700    1.13500    64700
C   6.0    82300    96500   10.60900
H   1.0    5.37300    82600    9.81000
C   6.0   -2.09100   -0.17800    6.12400
```

H	1.0	-2.47700	-0.82100	5.57500
C	6.0	1.10700	2.89800	8.47100
H	1.0	1.61100	2.07100	8.54100
H	1.0	1.18200	3.22400	7.56000
C	6.0	5.84600	1.15000	15.06700
H	1.0	6.42500	0.58000	153800
H	1.0	96300	1.16600	166600
C	6.0	5.56400	1.96400	17.31300
H	1.0	5.87900	1.86600	18.22600
H	1.0	62400	2.20500	17.32900
C	6.0	-1.05200	1.98600	6.37600
H	1.0	-0.73900	2.78000	6.00500
C	6.0	-0.16200	2.10400	11.36300
H	1.0	0.25100	2.91000	11.57400
C	6.0	7.71200	5.70100	19.61300
H	1.0	7.74000	6.45600	19.00500
H	1.0	7.29800	5.97800	20.44500
C	6.0	-1.05800	-0.06800	12.04400
H	1.0	-1.23500	-0.68500	12.71700
C	6.0	-0.42300	1.13900	12.35500
H	1.0	-0.16900	1.30800	13.23400
C	6.0	5.78500	0.68000	16.51900
H	1.0	5.05100	0.05900	16.65200
H	1.0	6.61400	0.24900	16.78000
C	6.0	92100	6.41400	11.10000
H	1.0	5.63900	6.51600	11.74300
H	1.0	5.07300	7.02100	10.35800
C	6.0	9.10200	5.13400	19.85300
H	1.0	9.13400	64200	20.68700
H	1.0	9.76200	5.84400	19.88100
C	6.0	3.57100	6.67400	11.74600
H	1.0	3.37400	7.62300	11.76900
H	1.0	3.54600	6.32300	12.65000
C	6.0	7.56900	3.88300	10.82600
H	1.0	8.02400	70600	11.06500
H	1.0	7.22800	3.96200	9.92100
C	6.0	2.58800	5.92600	10.83500
H	1.0	1.83500	5.58000	11.34000
H	1.0	2.25800	6.50100	10.12700
C	6.0	8.39400	2.37700	12.44000
H	1.0	9.00100	2.93300	12.95300
H	1.0	8.61200	1.44500	12.59900
C	6.0	8.48300	2.70700	10.95500
H	1.0	9.38900	2.93800	10.70000
H	1.0	8.17700	1.96400	10.41200
O	8.0	-7.60700	7.00700	7.82100
O	8.0	-13.16900	7.60300	2.81200
O	8.0	-9.66500	6.06200	8.09200
O	8.0	-174200	5.73000	0.70000
O	8.0	-10.37800	6.76100	63100
O	8.0	-15.49400	5.16300	-2.45100
N	7.0	-13.97700	7.03300	-0.96600
N	7.0	-9.38600	7.98100	6.90400
N	7.0	-12.37700	5.64700	2.01700
N	7.0	-12.46500	6.73100	5.43700
N	7.0	-17.40900	6.18000	-1.78900
H	1.0	-17.91300	5.54900	-2.08400
H	1.0	-17.76000	6.86600	-1.40600
C	6.0	-16.08000	6.09500	-1.92400
C	6.0	-13.79500	6.28100	0.12900
C	6.0	-12.80400	6.43900	3.00100
C	6.0	-11.18700	7.13000	5.48300
C	6.0	-5.01000	90100	9.46500
C	6.0	-56200	3.69900	8.90500
C	6.0	-82000	3.73100	7.46300
C	6.0	-8.95600	6.94800	7.64300
C	6.0	-12.36400	6.15200	0.64600
H	1.0	-11.91500	7.02300	0.61200
C	6.0	-3.97700	2.72400	9.70800
H	1.0	-3.65700	1.93300	9.33800
C	6.0	-5.43700	95000	7.15000
C	6.0	-12.94600	7.66300	-1.79100
H	1.0	-12.62000	7.05500	-2.47300
H	1.0	-12.19800	7.96400	-1.25100
C	6.0	-5.61700	5.78500	8.38700
H	1.0	-5.12600	6.63000	8.31300
C	6.0	-54900	2.78600	6.47200
H	1.0	-13100	1.98300	6.68300
C	6.0	-15.33000	7.33600	-1.44900
H	1.0	-15.84300	7.80100	-0.75600
C	6.0	-12.90300	5.80800	3.18200
H	1.0	-12.37400	98400	1400
C	6.0	-3.15700	14000	11.63800
H	1.0	-29700	27300	12.55700
C	6.0	-10.79800	8.10400	6.59400
H	1.0	-11.34800	7.96400	7.39300
C	6.0	-3.88400	2.96100	11.07900
H	1.0	-3.49800	2.31800	11.62800
C	6.0	-7.08100	6.03600	8.73200
H	1.0	-7.58600	5.21000	8.66200
H	1.0	-7.15700	6.36200	9.64300
C	6.0	-11.82100	28900	2.13700
H	1.0	-12.40000	3.71800	2.66600
H	1.0	-10.93800	3.10500	2.53800
C	6.0	-11.53800	5.10200	-0.11000
H	1.0	-11.85300	5.00400	-1.02300
H	1.0	-10.59900	5.34400	-0.12600
C	6.0	-92300	5.12400	10.82800

H	1.0	-5.23600	5.91800	11.19800
C	6.0	-5.81300	5.24200	5.84000
H	1.0	-6.22500	6.04900	5.62900
C	6.0	-13.68600	8.83900	-2.41000
H	1.0	-13.71500	9.59500	-1.80200
H	1.0	-13.27200	9.11700	-3.24200
C	6.0	-91600	3.07000	5.15900
H	1.0	-73900	2.45400	8600
C	6.0	-5.55200	27800	84800
H	1.0	-5.80500	4700	3.96900
C	6.0	-11.76000	3.81900	0.68500
H	1.0	-11.02600	3.19800	0.55100
H	1.0	-12.58800	3.38800	0.42400
C	6.0	-10.89500	9.55300	6.10400
H	1.0	-11.61400	9.65400	5.46100
H	1.0	-11.04700	10.16000	6.84500
C	6.0	-15.07600	8.27300	-2.64900
H	1.0	-15.10800	7.78000	-3.48400
H	1.0	-15.73700	8.98300	-2.67800
C	6.0	-9.54600	9.81200	5.45700
H	1.0	-9.34800	10.76100	5.43400
H	1.0	-9.52100	9.46100	55300
C	6.0	-13.54300	7.02200	6.37700
H	1.0	-13.99900	7.84400	6.13900
H	1.0	-13.20300	7.10000	7.28200
C	6.0	-8.56200	9.06500	6.36900
H	1.0	-7.81000	8.71800	5.86300
H	1.0	-8.23200	9.64000	7.07600
C	6.0	-13.16900	5.51600	76400
H	1.0	-197600	6.07200	25000
H	1.0	-158700	58400	60500
C	6.0	-15700	5.84600	6.24800
H	1.0	-15.36300	6.07600	6.50300
H	1.0	-115100	5.10300	6.79100
$E				

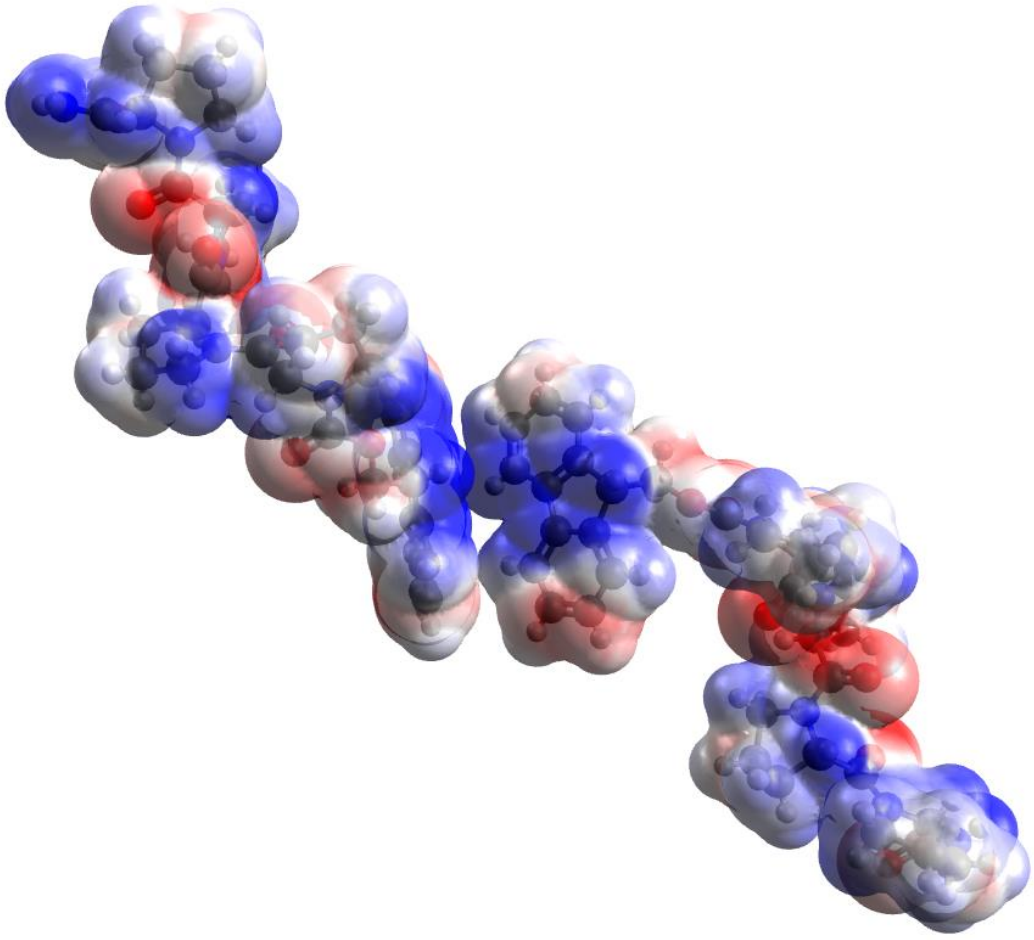


Figure S32 - Electrostatic potential map (blue – positive, red – negative) calculated at B3LYP-D3/6-311-G++(p,d) level of theory

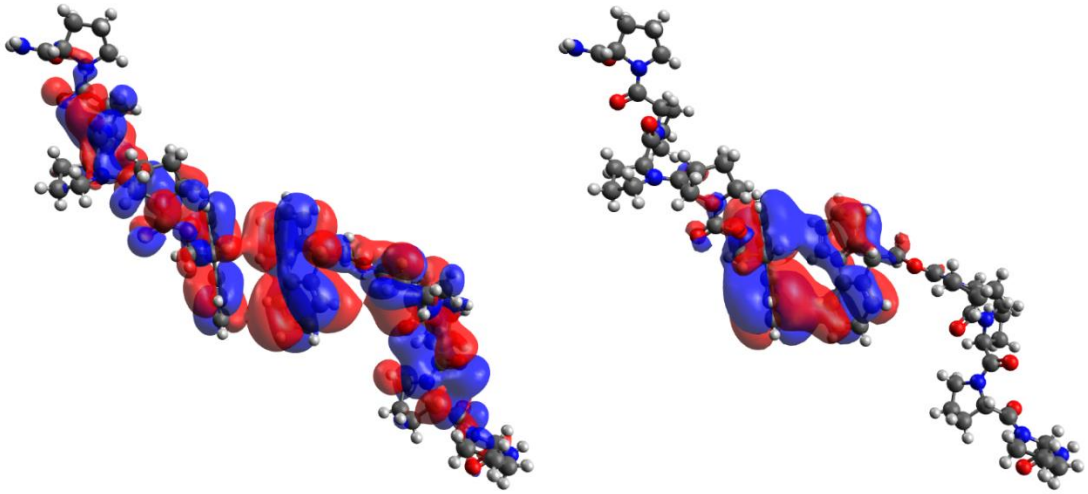


Figure S33 - HOMO-6 (left) and HOMO-19 (right) highlighting orbital interactions between discrete molecules, calculated at B3LYP-D3/6-311-G++(p,d) level of theory

SI 4.3.1.17 Energy Decomposition Analysis

EDA was conducted on the Fmoc-Fmoc dimer obtained from the crystal geometry (*Figure S43*) using the Amsterdam Density Functional (Amsterdam Modelling Suite) with the BLYP-D(BJ) and PBE-D functionals (frozen cores). [11]

Table S5. Energy components of Fmoc-Fmoc interactions from EDA (kJ mol^{-1}).

Func.	E_{elstat}	E_{Pauli}	E_{orb}	E_{disp}	E_{total}
BLYP-D3(BJ)	-15.5	32.6	-12.9	-48.3	-42
PBE-D	-13.7	17.9	-8.7	-27.9	-32.4

Model parameters were as follows.

```
DENSITY FUNCTIONAL POTENTIAL (scf)
    LDA:                                Exchange only
== Not Default ==
    Gradient Corrections:               Becke88 LYP
== Not Default ==

 SPIN  (restricted / unrestr.)
    Molecule:                           Restricted
    Fragments:                          Restricted

 OTHER ASPECTS
    Relativistic Corrections:           scalar (ZORA,MAPA)

    Nuclear Charge Density Model:       Point Charge Nuclei
    Core Treatment:                     Frozen Orbital(s)

    Hyperfine or Zeeman Interaction:    ---

 Settings for Grimme D3 dispersion correction

      damping                             BJ
      s6                               1.000
      s8                               2.700
```

```
      a1                            0.430
      a2                            236

 Other (technical) parameters

      alpha                        1000
      version                           4



DENSITY FUNCTIONAL POTENTIAL (scf)
    LDA:                                PW92
== Not Default ==
    Gradient Corrections:               PBEc PBEx
== Not Default ==

 SPIN  (restricted / unrestr.)
    Molecule:                           Restricted
    Fragments:                          Restricted

 OTHER ASPECTS
    Relativistic Corrections:           scalar (ZORA,MAPA)

    Nuclear Charge Density Model:       Point Charge Nuclei
    Core Treatment:                     Frozen Orbital(s)

    Hyperfine or Zeeman Interaction:    ---

 Settings for Grimme dispersion correction
   use heavy dispersion       T
   scaling of radii                1.100
   alpha (scaling function)       20.000
   cut off distance              100.000
   overall factor                  0.750
```

SI 4.3.2 HP_3 SC-XRD:

The crystal structure data was obtained from colourless plank crystals, the crystal quality was relatively poor, the crystal was mounted on a Mitegen micromount in Paratone immersion oil and cooled to 150 K using an Oxford Cryosystems 800-series Cryostream. The crystal structure was obtained with some regions of significant disorder. The linker to the fluorenyl group was significantly disordered as such the atom C13-14 were split and EADP applied. Split SAME was also carried out on O2. The Pro1 pyrrolidine ring was disordered, as such C16-19 were split and EADP applied, split SAME was applied to O8. The pro3 ring was also disordered as such C26-28 were split and DELU applied. Only small voids are present (40.80 $Å^3$, 1.3%-unit cell volume, Probe radius 1.2 Å, grid spacing 0.5 Å) within the structure that do not extend through the framework, forming a non-porous structure. The fmoc region is disordered and this extends to the neighbouring peptide's Pro3 sidechain, while the C-terminus is well ordered.

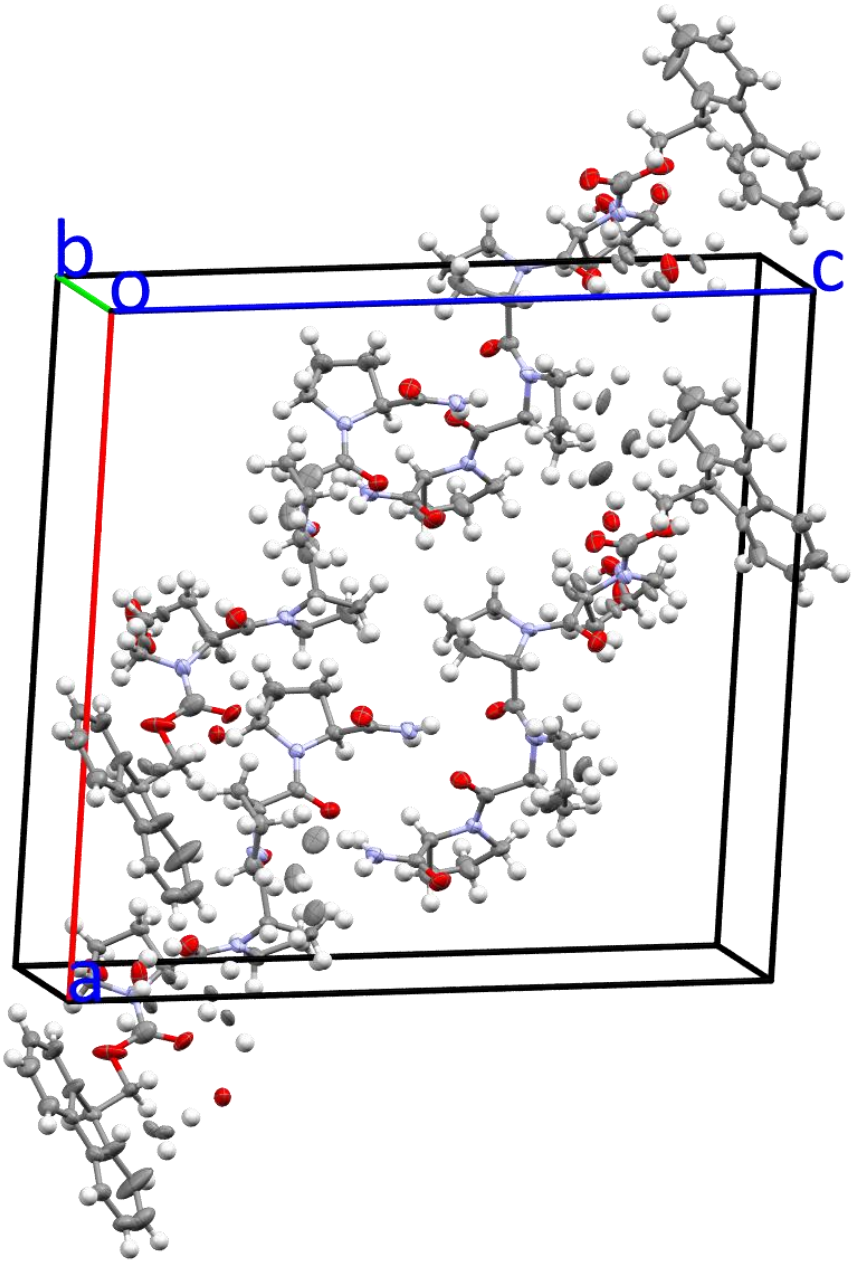


Figure S34 – Crystal structure of peptide **HP_3** showing the unit cell, Atomic displacement parameters are shown at 50% probability .

Crystal data and structure refinement for HP_3	
Identification code	HP3
Empirical formula	$C_{35}H_{41}N_5O_7$
Formula weight	643.73
Temperature/K	150.15
Crystal system	monoclinic
Space group	C2
a/Å	22.1654(7)
b/Å	6.3098(2)
c/Å	23.0675(7)
α/°	90
β/°	92.401(3)
γ/°	90
Volume/$Å^3$	3223.37(17)
Z	4
$\rho_{calc}g/cm^3$	1.326
μ/mm^{-1}	0.765
F(000)	1368.0
Crystal size/mm^3	0.21 × 0.07 × 0.06
Radiation	CuKα (λ = 1.54184)
2Θ range for data collection/°	7.672 to 136.628
Index ranges	-26 ≤ h ≤ 26, -7 ≤ k ≤ 7, -27 ≤ l ≤ 27
Reflections collected	11914
Independent reflections	5885 [R_{int} = 0.0822, R_{sigma} = 0.0962]
Data/restraints/parameters	5885/22/493
Goodness-of-fit on F^2	1.096
Final R indexes [I>=2σ (I)]	R_1 = 0.0883, wR_2 = 0.2371
Final R indexes [all data]	R_1 = 0.1103, wR_2 = 0.2675
Largest diff. peak/hole / e $Å^{-3}$	0.35/-0.38
Flack parameter	-0.5(5)
CCDC No.	2238155

SI 4.3.3 PHP_2 SC-XRD:

Crystallised from slow-evaporation of a propan-1-ol solution forming long fibrous crystals, however these were not suitable for SC-XRD analysis. Crystallised poorly from an ethanol/ethyl acetate solution and methanol/water solution. However, poor quality powder diffraction data was obtained, see 4.

Analysing the other positions on the **P_4** backbone, the Pro2 residue on the **P_4** backbone is aligned with adjacent peptides' aromatic groups and the closest hydrogen bond acceptor (**H_A**) is 5 Å from the C4 position, which suggests the current crystal structure is unlikely to satisfy the hydrogen bond donation of a new hydroxyl at this position (3.1, *Figure* S36).

SI 4.3.4 P_2HP SC-XRD:

The crystal structure data was obtained from colourless plank crystals, the crystal was mounted on a Mitegen micromount in Paratone immersion oil and cooled to 150 K using an Oxford Cryosystems 800-series Cryostream. **P_2HP** crystal structure was isostructural to **P_4**, and two water molecules were modelled within the voids of the framework. One of the water molecules was disordered as such split SAME was applied to O9. EADP was applied O2-C15. This reduced

pore size is too small to accommodate EtOH molecules, as such the framework is selective towards H_2O molecules in the wet solvent during crystallisation, which can be clearly modelled in the crystal structure (*Figure S37*).

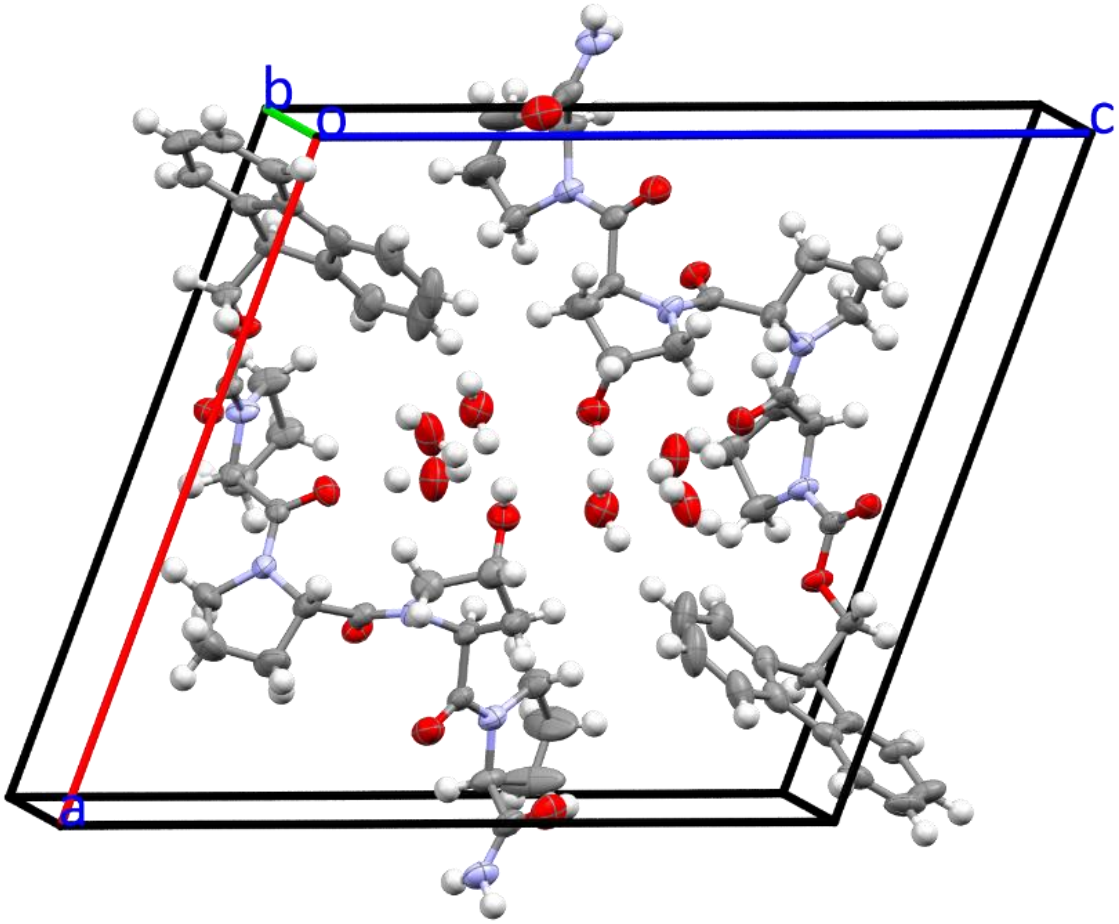


Figure S38 – Crystal structure of peptide P_2HP showing the unit cell, atomic displacement parameters are shown at 50% probability .

Crystal data and structure refinement for P_2HP	
Identification code	P2HP
Empirical formula	$C_{35}H_{45}N_5O_9$
Formula weight	679.76
Temperature/K	150.15
Crystal system	monoclinic
Space group	$P2_1$
a/Å	16.6440(5)
b/Å	6.3071(2)
c/Å	17.8450(5)
α/°	90
β/°	110.012(3)
γ/°	90
Volume/Å^3	1760.18(10)
Z	2
ρ_{calc}g/cm^3	1.283
μ/mm^{-1}	0.770
F(000)	720
Crystal size/mm^3	0.12 × 0.07 × 0.06
Radiation	CuKα (λ = 1.54184)
2Θ range for data collection/°	8.956 to 136.434
Index ranges	-20 ≤ h ≤ 20, -7 ≤ k ≤ 7, -15 ≤ l ≤ 21
Reflections collected	18801
Independent reflections	6417 [R_{int} = 0.0425, R_{sigma} = 0.0447]
Data/restraints/parameters	6417/1/470
Goodness-of-fit on F^2	1.051
Final R indexes [I>=2σ (I)]	R_1 = 0.0515, wR_2 = 0.1389
Final R indexes [all data]	R_1 = 0.0621, wR_2 = 0.1473
Largest diff. peak/hole / e Å^{-3}	0.35/-0.29
Flack parameter	-0.10(10)
CCDC No.	2238161

SI 4.3.5 P_3H SC-XRD:

The crystal structure data was obtained from colourless plank crystals, the crystal was mounted on a Mitegen micromount in Paratone immersion oil and cooled to 150 K using an Oxford Cryosystems 800-series Cryostream. The structure was resolved with two peptides comprising the asymmetric unit (Z' = 2), joined by the typical C-terminal amide NH_2 hydrogen bonds. Each of these peptides was modelled with a disordered EtOH molecules within the pores with partial occupancies (0.5). The oxygen atoms of each were disordered as such split SAME was applied to each (O8-O8X). While a model of the atomic positions of this disordered solvent is included it is likely that an atomistic model is not fully appropriate for the disordered electron density inside the pore.

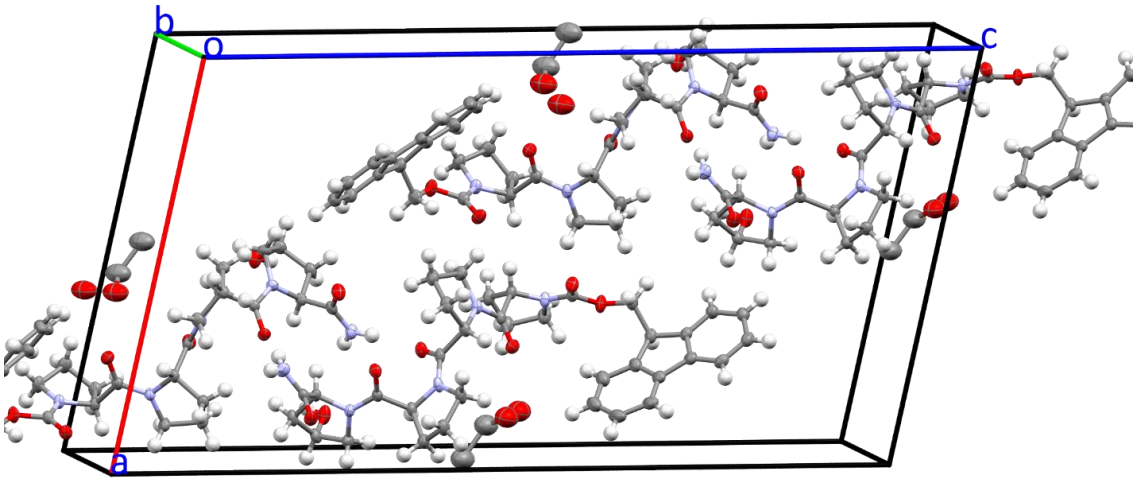


Figure S40 – Crystal structure of peptide P_3H showing the unit cell, Atomic displacement parameters are shown at 50% probability

Crystal data and structure refinement for P_3H	
Identification code	P3H
Empirical formula	$C_{36}H_{41}N_5O_8$
Formula weight	671.74
Temperature/K	150.00(10)
Crystal system	monoclinic
Space group	$P2_1$
a/Å	16.8723(2)
b/Å	6.45220(10)
c/Å	32.1004(4)
α/°	90
β/°	101.2140(10)
γ/°	90
Volume/Å^3	3427.84(8)
Z	4
ρ_{calc}g/cm^3	1.302
μ/mm^{-1}	0.766
F(000)	1420
Crystal size/mm^3	0.16 × 0.11 × 0.07
Radiation	CuKα (λ = 1.54184)
2Θ range for data collection/°	6.956 to 14164
Index ranges	-18 ≤ h ≤ 20, -7 ≤ k ≤ 7, -39 ≤ l ≤ 39
Reflections collected	38516
Independent reflections	13166 [R_{int} = 0.0224, R_{sigma} = 0.0216]
Data/restraints/parameters	13166/51/909
Goodness-of-fit on F^2	1.043
Final R indexes [I>=2σ (I)]	R_1 = 0.0416, wR_2 = 0.1251
Final R indexes [all data]	R_1 = 0.0431, wR_2 = 0.1274
Largest diff. peak/hole / e Å^{-3}	1.24/-0.86
Flack parameter	0.07(4)
CCDC No.	2238152

SI 4.3.6 P_4-P_2HP SC-XRD:

The crystal structure data was obtained from colourless plank crystals crystallised *via* slow cooling of a hot supersaturated ethanol solution of peptides **P_4** and **P_2HP** in equimolar concentrations. The crystal was mounted on a Mitegen micromount in Paratone immersion oil and kept at 290 K using an Oxford Cryosystems 800-series Cryostream. The mixed peptides crystal structure was isostructural to **P_2HP** and **P** Disordered EtOH molecules were modelled in the pores with a partial occupancy of 0.75. Split same was applied to the EtOH atoms C36-37. H atoms were not assigned to the EtOH molecules due to the high level of disorder. Interestingly this framework adopted the *endo* conformation similarly to **P_4** alone, having significantly less impact on the channel volume (Volume 233.02 Å^3, 12.8% / unit cell, Probe *r* = 1.2 Å, Grid spacing 0.4 Å), and thus contained EtOH within the pores rather than being selective towards H_2O (*Figure S42 and S43*).

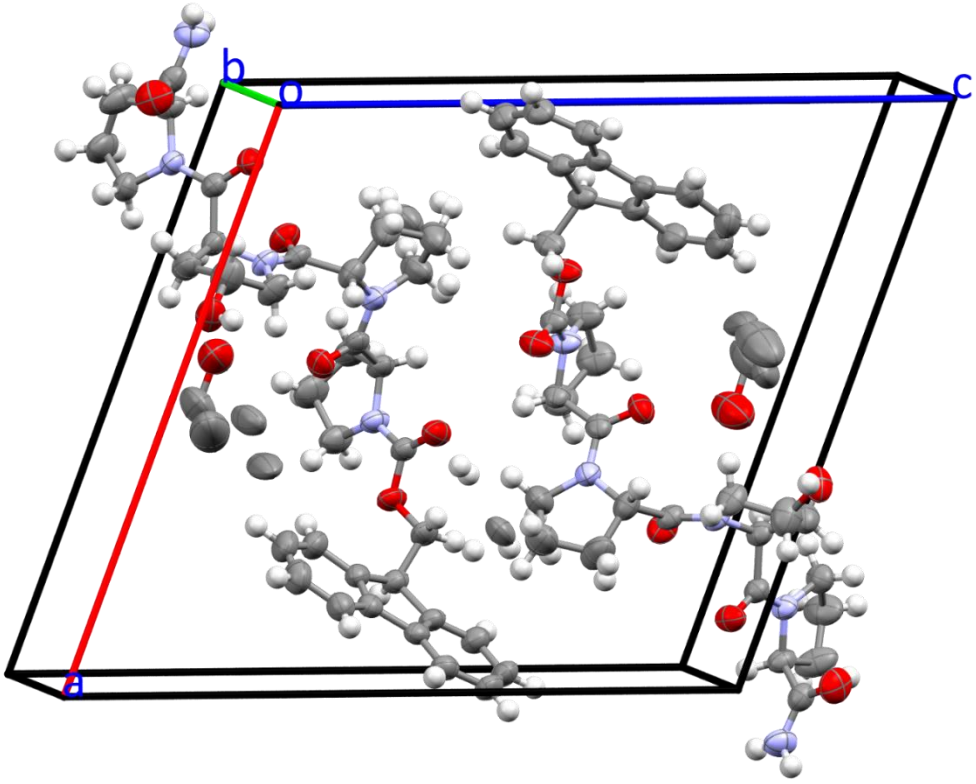


Figure S44 – Crystal structure of peptide P_4+P_2HP showing the unit cell, atomic displacement parameters are shown at 50% probability .

Crystal data and structure refinement for P_4-P_2HP	
Identification code	P2HP-P4
Empirical formula	$C_{36.5}H_{42.36}N_5O_{7.13}$
Formula weight	665.19
Temperature/K	290.0(2)
Crystal system	monoclinic
Space group	$P2_1$
a/Å	16.6077(8)
b/Å	6.3463(3)
c/Å	18.3211(9)
α/°	90
β/°	109.933(5)
γ/°	90
Volume/Å^3	1815.31(16)
Z	2
ρ_{calc}g/cm^3	1.217
μ/mm^{-1}	0.699
F(000)	707.0
Crystal size/mm^3	0.18 × 0.06 × 0.04
Radiation	Cu Kα (λ = 1.54184)
2Θ range for data collection/°	9.896 to 143.886
Index ranges	-20 ≤ h ≤ 20, -7 ≤ k ≤ 7, -22 ≤ l ≤ 19
Reflections collected	13382
Independent reflections	6926 [R_{int} = 0.0244, R_{sigma} = 0.0353]
Data/restraints/parameters	6926/29/481
Goodness-of-fit on F^2	1.039
Final R indexes [I>=2σ (I)]	R_1 = 0.0480, wR_2 = 0.1294
Final R indexes [all data]	R_1 = 0.0604, wR_2 = 0.1398
Largest diff. peak/hole / e Å^{-3}	0.40/-0.20
Flack parameter	-0.01(11)
CCDC No.	2238180

SI 4.3.7 HP_2H SC-XRD:

The crystal structure data was obtained from colourless plank crystals, crystallised *via* slow evaporation of an ethanol solution of peptide **HP_2H**. The crystal was mounted on a Mitegen micromount in Paratone immersion oil and kept at 295 K using an Oxford Cryosystems 800-series Cryostream. The first hydroxyproline residue was disordered on hydroxyl group and at the Cγ position (C18), likely to due to the presence of *endo* and *exo* ring puckering within the structure. The crystal structure is nonporous containing no significant void space. Both hydrogen bonds present in the **HP_3** and **P_3H** crystal structure are present within the crystal structure between the same atoms, extending along the *b*-axis.

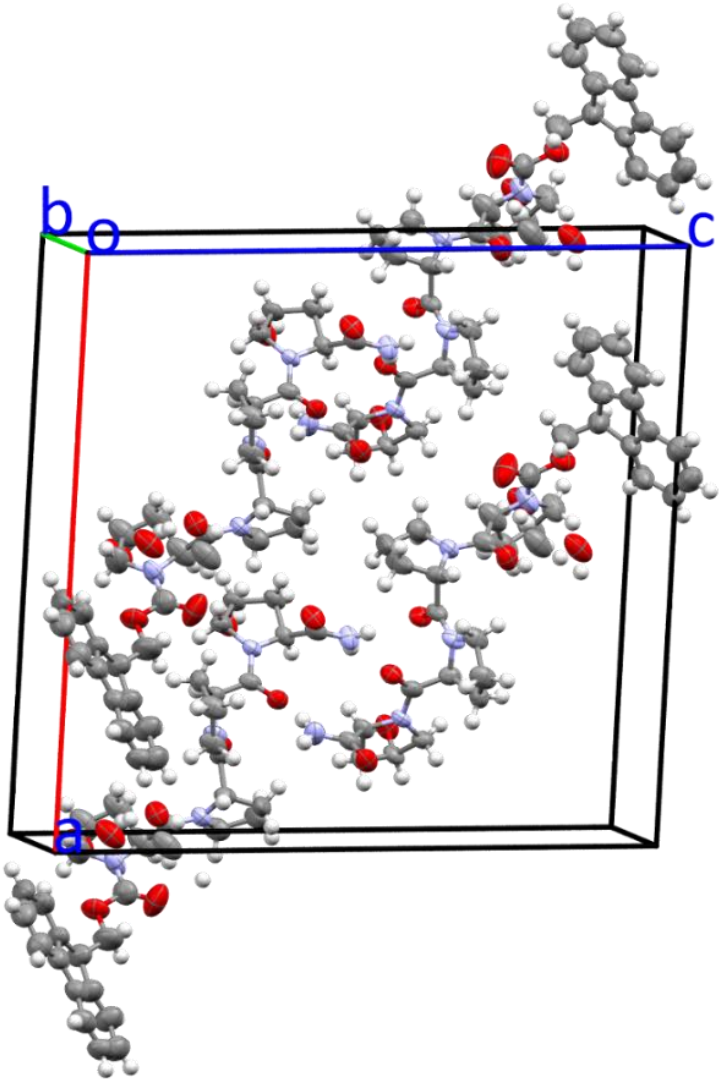


Figure S46 – Crystal structure of peptide **HP_2H** showing the unit cell, atomic displacement parameters are shown at 50% probability .

Crystal data and structure refinement for HP_2H	
Identification code	HP2H
Empirical formula	$C_{35}H_{40}N_5O_8$
Formula weight	658.72
Temperature/K	295
Crystal system	monoclinic
Space group	C2
a/Å	22.476(2)
b/Å	6.2932(7)
c/Å	23.245(3)
α/°	90
β/°	92.036(10)
γ/°	90
Volume/Å^3	3285.8(6)
Z	4
ρ_{calc}g/cm^3	1.332
μ/mm^{-1}	0.788
F(000)	1396.0
Crystal size/mm^3	0.16 × 0.03 × 0.02
Radiation	Cu Kα (λ = 1.54184)
2Θ range for data collection/°	7.612 to 136.478
Index ranges	-19 ≤ h ≤ 26, -7 ≤ k ≤ 7, -27 ≤ l ≤ 27
Reflections collected	11409
Independent reflections	6003 [R_{int} = 0.1158, R_{sigma} = 0.1952]
Data/restraints/parameters	6003/5/421
Goodness-of-fit on F^2	0.840
Final R indexes [I>=2σ (I)]	R_1 = 0.0835, wR_2 = 0.1224
Final R indexes [all data]	R_1 = 0.1873, wR_2 = 0.1585
Largest diff. peak/hole / e Å^{-3}	0.40/-0.21
Flack parameter	0.7(5)
CCDC No.	2238160

SI 4.3.8 *cis*-HP_2H SC-XRD:

The crystal structure data was obtained from colourless plank crystals, crystallised *via* slow evaporation of an ethanol/acetonitrile solution of peptide *cis*-**HP_2H**. The crystal was mounted on a Mitegen micromount in Paratone immersion oil and kept at 150 K using an Oxford Cryosystems 800-series Cryostream. The crystal structure obtained shows the presence of channels within the structure, filled with disordered solvent (volume 628.9 Å^3, 17.1% / unit cell, Figure S48) and no suitable model could be obtained for the disordered solvent (EtOH and acetonitrile) within the pores of the structure and so a solvent masking routine was used. The electron density was therefore accounted for using a solvent mask within Olex2,[3] giving a solvent accessible volume of 167 Å^3 and containing 36 electrons asymmetric unit. Screening of multiple crystallites suitable for single crystal analysis gave the same crystal structure. The extended structure of *cis*-**HP_2H** differs from the other Fmoc peptide frameworks as the Fmoc moieties of adjacent peptides no longer face one another, resulting in staggered rather than linear H-bonded layers of the peptides. The last hydroxyproline's (C-terminus) hydroxyl group is aligned into the pore space, clearly hydrogen bonding to a solvent molecule, with significant electron density adjacent to this group. Therefore, the only intermolecular hydrogen-bonding between peptides is the typical C-terminal NH_2 amide bonding present in all the structures seen previously. This highlights how small changes in the placement of functional groups can be used to affect the assembly process, with control over even the flexible pyrrolidine ring endo/exo conformations possible by use of

4S versus 4R functional groups, while the polyproline II helix remains as a rigid ligand for placement of these functional groups.

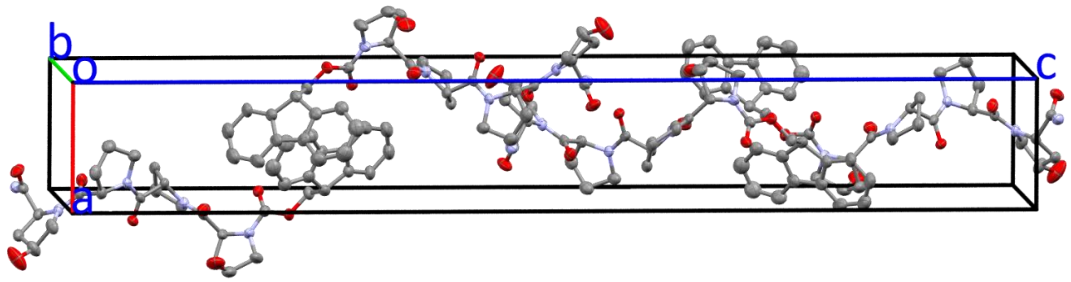


Figure S48 – Crystal structure of peptide *cis*-**HP$_2$H** showing the unit cell, Atomic displacement parameters are shown at 50% probability .

Crystal data and structure refinement for *cis*-HP$_2$H.	
Identification code	Cis-HP2H
Empirical formula	$C_{35}H_{41}N_5O_8$
Formula weight	659.73
Temperature/K	150.15
Crystal system	orthorhombic
Space group	$P2_12_12_1$
a/Å	6.4160(2)
b/Å	12.0836(6)
c/Å	47.4708(17)
α/°	90
β/°	90
γ/°	90
Volume/Å^3	3680.3(3)
Z	4
ρ_{calc}g/cm^3	1.191
μ/mm^{-1}	0.085
F(000)	1400.0
Crystal size/mm^3	0.26 × 0.06 × 0.05
Radiation	MoKα (λ = 0.71073)
2Θ range for data collection/°	6.578 to 57.246
Index ranges	-6 ≤ h ≤ 8, -16 ≤ k ≤ 15, -50 ≤ l ≤ 63
Reflections collected	18190
Independent reflections	7790 [R_{int} = 0.0457, R_{sigma} = 0.0744]
Data/restraints/parameters	7790/0/440
Goodness-of-fit on F^2	1.020
Final R indexes [I>=2σ (I)]	R_1 = 0.0600, wR_2 = 0.1215
Final R indexes [all data]	R_1 = 0.0926, wR_2 = 0.1348
Largest diff. peak/hole / e Å^{-3}	0.23/-0.25
Flack parameter	0.0(6)
CCDC No.	2238252

SI 4.3.9 AcHP$_2$H SC-XRD:

The crystal structure data was obtained from colourless needle crystals, crystallised *via* slow evaporation of an acetonitrile solution of peptide **AcHP$_2$H**. The crystals obtained formed packed fibrous assemblies that did not diffract well due to low crystal volumes requiring long exposure times. The crystal was mounted on a Mitegen micromount in Paratone immersion oil and at 29 The crystal structure obtained showed a porous structure (Volume 415.06 Å^3, 15.6% / unit cell, Probe *r* = 1.2 Å, Grid spacing 0.4 Å), in the *ortho*rhombic $P2_12_12_1$ space group, with channels extending along the *a*-axis containing ordered acetonitrile molecules with no apparent strong interactions between the solvent and the peptide within the structure. The crystal was weakly diffracting and no data above 2 I/sig was observed above 1.1 angstroms. Consequently, the data were truncated at this resolution.

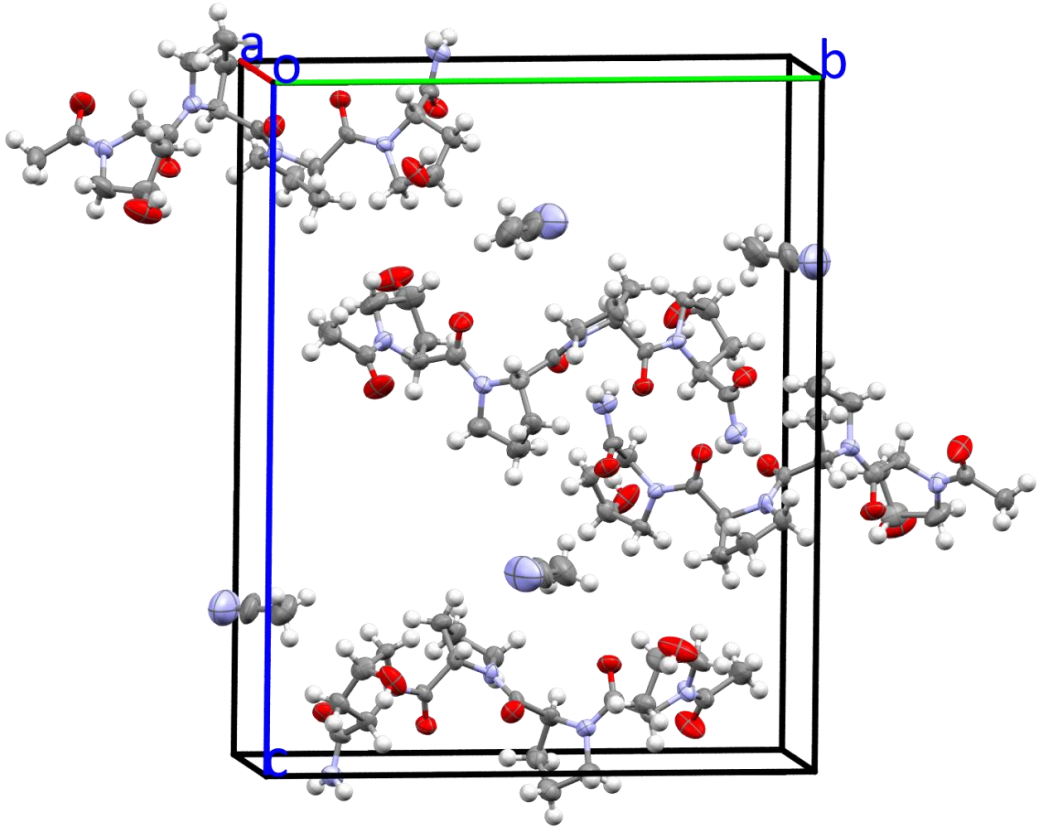


Figure S50 – Crystal structure of peptide **AcHP$_2$H** showing the unit cell, Atomic displacement parameters are shown at 50% probability .

Crystal data and structure refinement for $AcHP_2H$	
Identification code	AcHP2H
Empirical formula	$C_{24}H_{36}N_6O_7$
Formula weight	520.59
Temperature/K	290(5)
Crystal system	orthorhombic
Space group	$P2_12_12_1$
a/Å	6.3773(10)
b/Å	18.232(4)
c/Å	22.857(5)
α/°	90
β/°	90
γ/°	90
Volume/$Å^3$	2657.6(9)
Z	4
$\rho_{calc}g/cm^3$	1.301
μ/mm^{-1}	0.097
F(000)	1112.0
Crystal size/mm^3	0.19 × 0.02 × 0.02
Radiation	Mo Kα (λ = 0.71073)
2Θ range for data collection/°	6.634 to 59.564
Index ranges	-8 ≤ h ≤ 8, -25 ≤ k ≤ 23, -30 ≤ l ≤ 28
Reflections collected	26353
Independent reflections	6766 [R_{int} = 0.2746, R_{sigma} = 0.4001]
Data/restraints/parameters	6766/0/338
Goodness-of-fit on F^2	0.966
Final R indexes [I>=2σ (I)]	R_1 = 0.1033, wR_2 = 0.1305
Final R indexes [all data]	R_1 = 0.3459, wR_2 = 0.1936
Largest diff. peak/hole / e $Å^{-3}$	0.18/-0.23
Flack parameter	-1.7(10)
CCDC No.	2234312

SI 4.3.10 AcP_4 SC-XRD:

The crystal structure data was obtained from colourless crystalline needles, crystallised *via* vapour diffusion of Et_2O into a solution of **AcP_4** in $CHCl_3$. The crystal was mounted on a Mitegen micromount in Paratone immersion oil and kept at 150 K using an Oxford Cryosystems 800-series Cryostream. The crystal structure obtained showed a porous structure (Volume 572.37 $Å^3$, 36.1% / unit cell, Probe *r* = 1.2 Å, Grid spacing 0.4 Å), with in the monoclinic $P2_1$ space group, with channels extending in a 2D layer along the *b*- and *a*-axis containing ordered chloroform molecules, with two solvent molecules per asymmetric unit, hydrogen bonding to the Pro1 and Pro4 carbonyl groups.

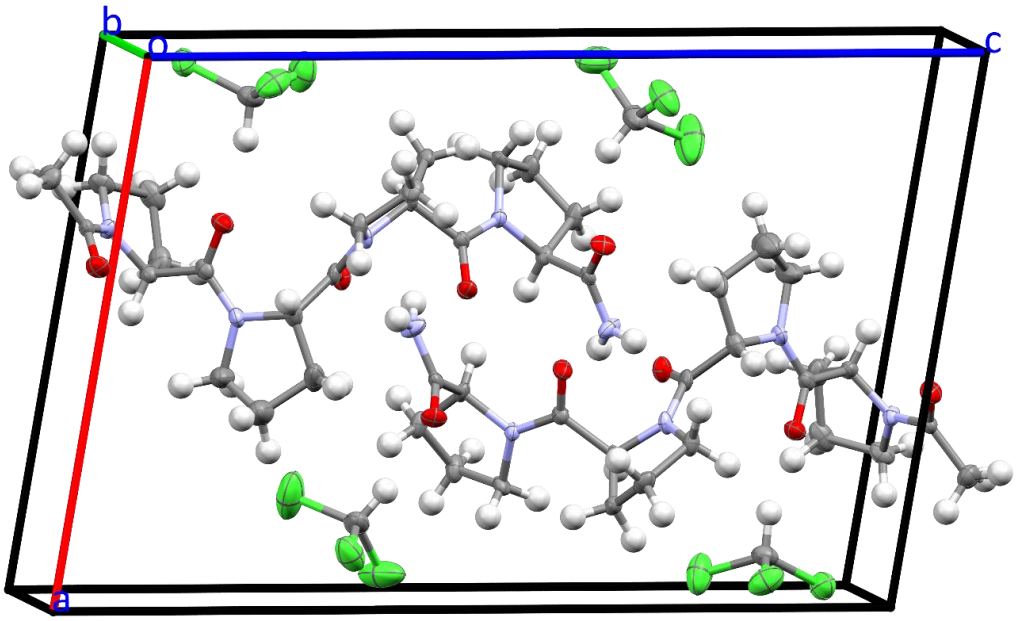


Figure S52 – Crystal structure of peptide **AcP$_4$** showing the unit cell, Atomic displacement parameters are shown at 50% probability .

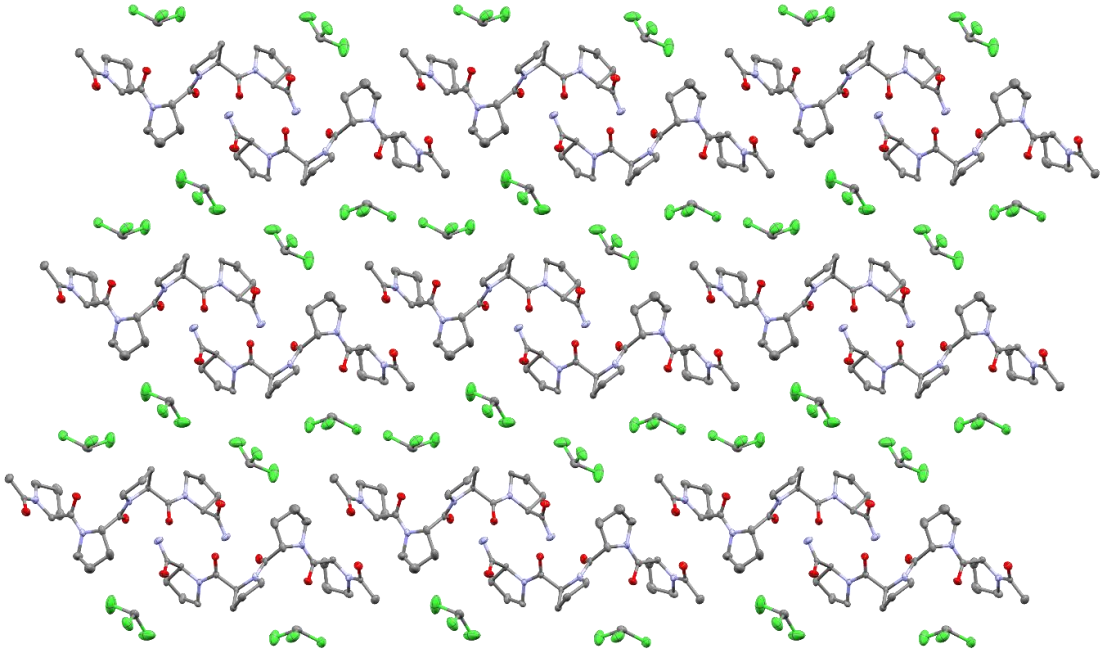


Figure S53 – Crystal structure of peptide **AcP$_4$** showing the packed extended structure (3x3x3), view along the *b* axis, hydrogens not shown, atomic displacement parameters are shown at 50% probability.

Crystal data and structure refinement for AcP_4	
Identification code	AcP4
Empirical formula	$C_{24}H_{35}Cl_6N_5O_5$
Formula weight	686.27
Temperature/K	150.0(7)
Crystal system	monoclinic
Space group	$P2_1$
a/Å	13.0167(8)
b/Å	6.2940(6)
c/Å	19.6114(12)
α/°	90
β/°	99.377(6)
γ/°	90
Volume/$Å^3$	1585.2(2)
Z	2
ρcalcg/cm^3	1.438
μ/mm^{-1}	0.584
F(000)	712.0
Crystal size/mm^3	0.51 × 0.06 × 0.04
Radiation	Mo Kα (λ = 0.71073)
2Θ range for data collection/°	6.59 to 50.038
Index ranges	-15 ≤ h ≤ 15, -7 ≤ k ≤ 7, -22 ≤ l ≤ 23
Reflections collected	14583
Independent reflections	5615 [R_{int} = 0.0679, R_{sigma} = 0.0900]
Data/restraints/parameters	5615/1/370
Goodness-of-fit on F^2	1.017
Final R indexes [I>=2σ (I)]	R_1 = 0.0512, wR_2 = 0.0969
Final R indexes [all data]	R_1 = 0.0783, wR_2 = 0.1084
Largest diff. peak/hole / e $Å^{-3}$	0.46/-0.38
Flack parameter	-0.09(5)
CCDC No.	2264145

SI 4.4 Powder XRD data:

SI 4.4.1 HP_3 PD-XRD:

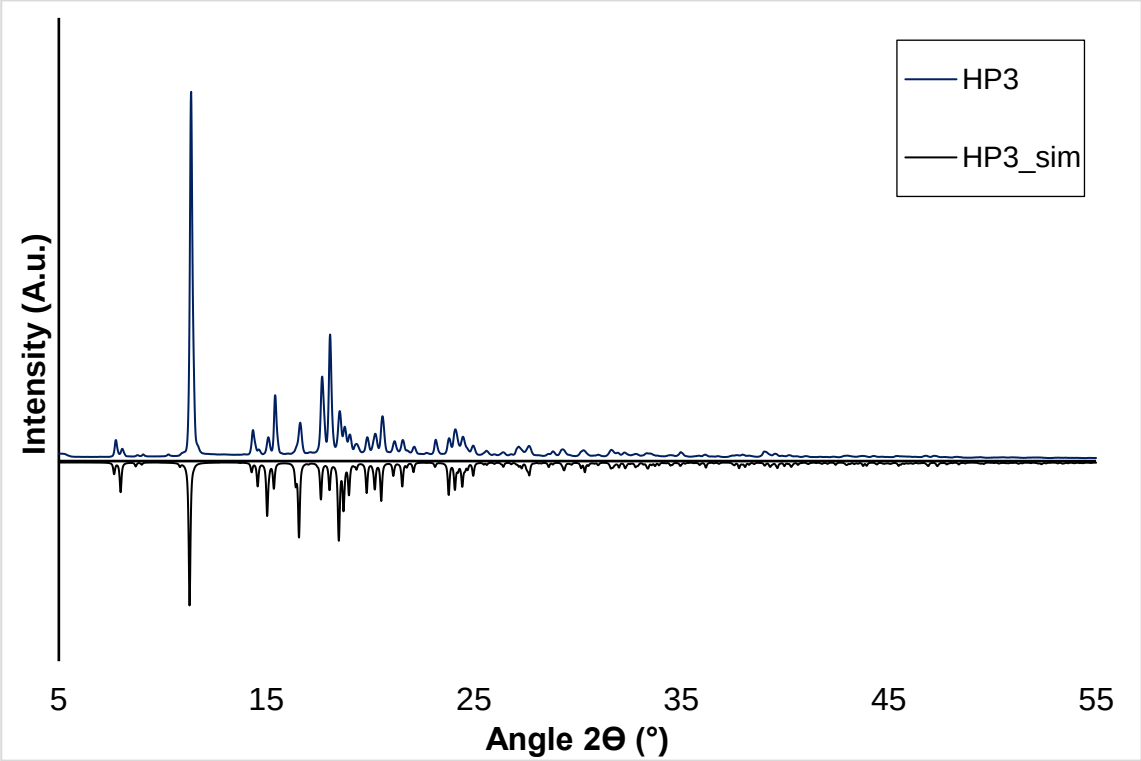


Figure S55 – PD-XRD of **HP_3** simulated from SC-XRD data (**black,** bottom), **HP_3** experimental PD-XRD (**blue,** top)

SI 4.4.2 PHP_2 PD-XRD:

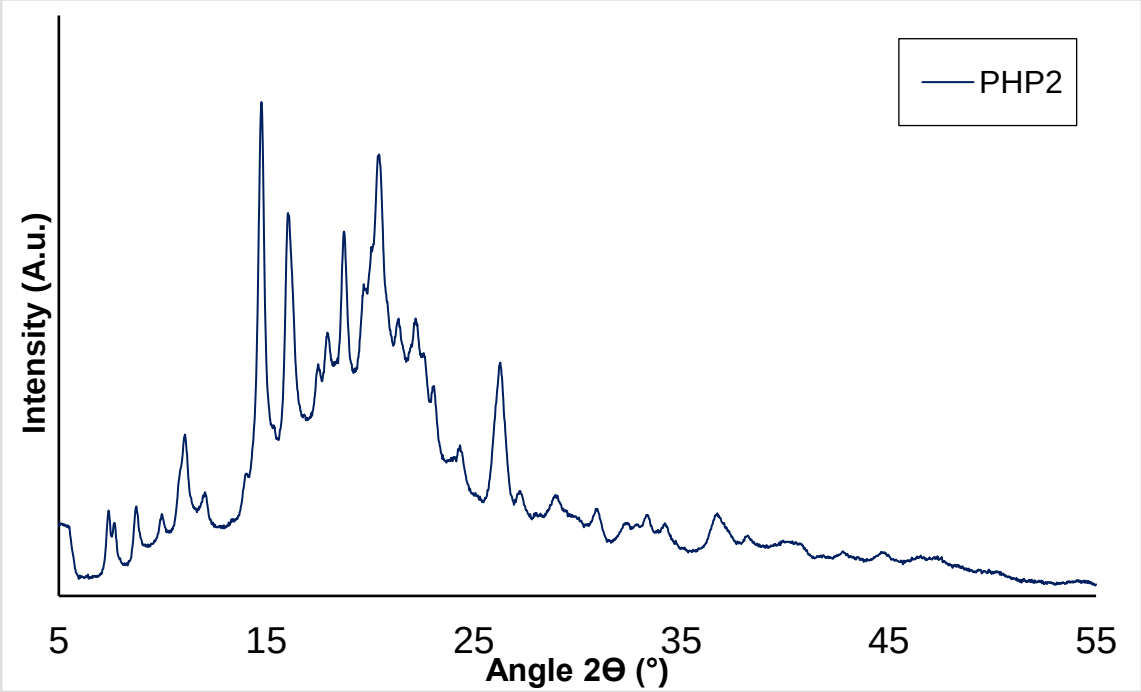


Figure S56 – PHP_2 experimental PD-XRD, sample was poorly crystalline and contained amorphous peptide producing a broad background peak.

SI 4.4.3 P_2HP PD-XRD:

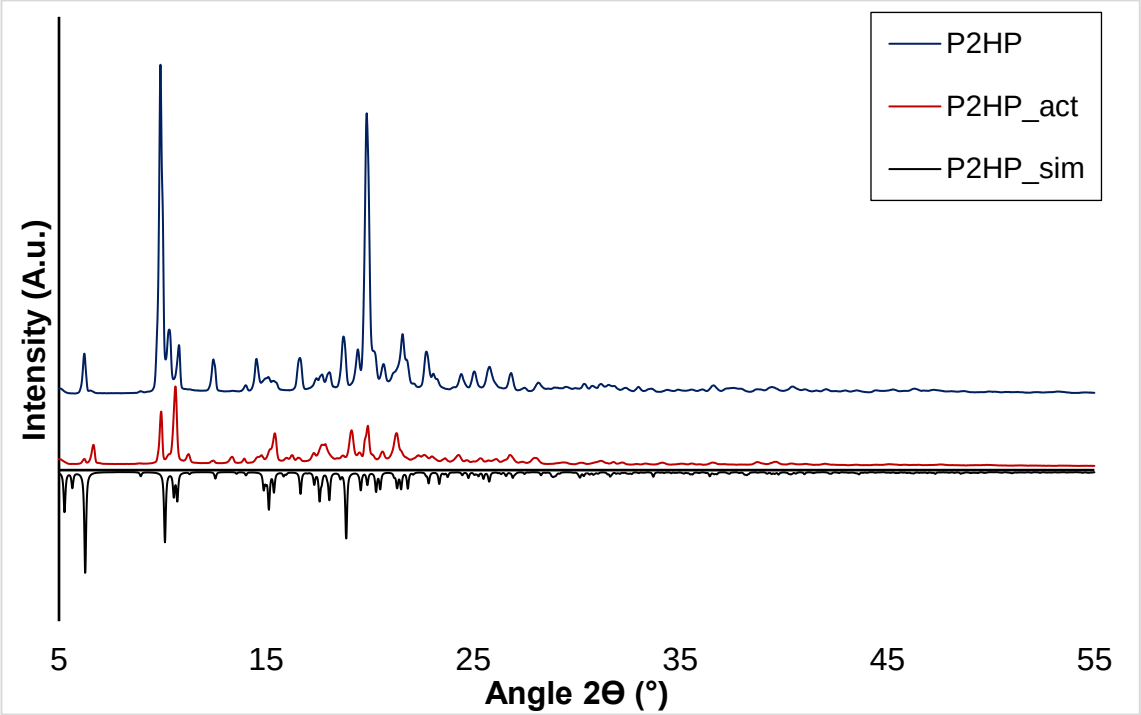


Figure S57 – PD-XRD of **P_2HP** simulated from SC-XRD data (**black, bottom**), **P_2HP** experimental PD-XRD (**blue, top**), **P_2HP** after activation at 45 °C under high vacuum (**red**)

SI 4.4.4 P_3H PD-XRD:

Peptide **P_3H** did not crystallise well as once precipitated from solution an insoluble white solid formed preventing redissolution, as such a suitable powder pattern could not be obtained, despite the crystallisation of some single crystals suitable for SCXRD, sufficient crystalline material was not obtained for PDXRD analysis.

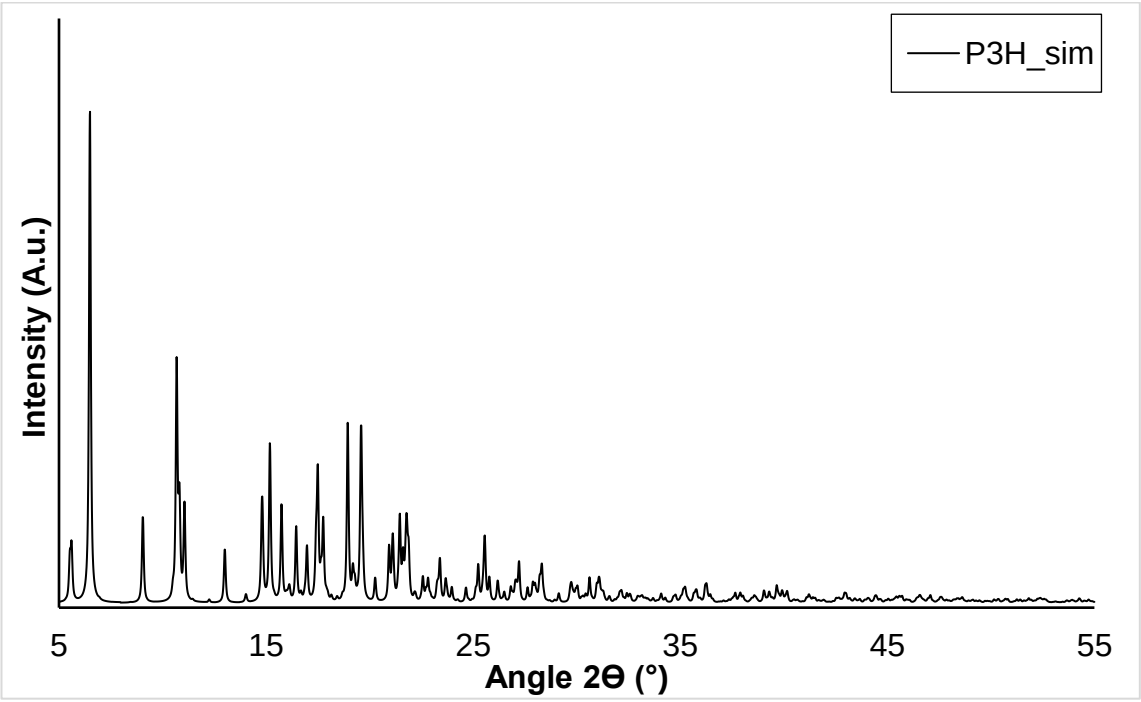


Figure S58 – PD-XRD of **P_3H** simulated from SC-XRD data

SI 4.4.5 P_4+P_2HP PD-XRD:

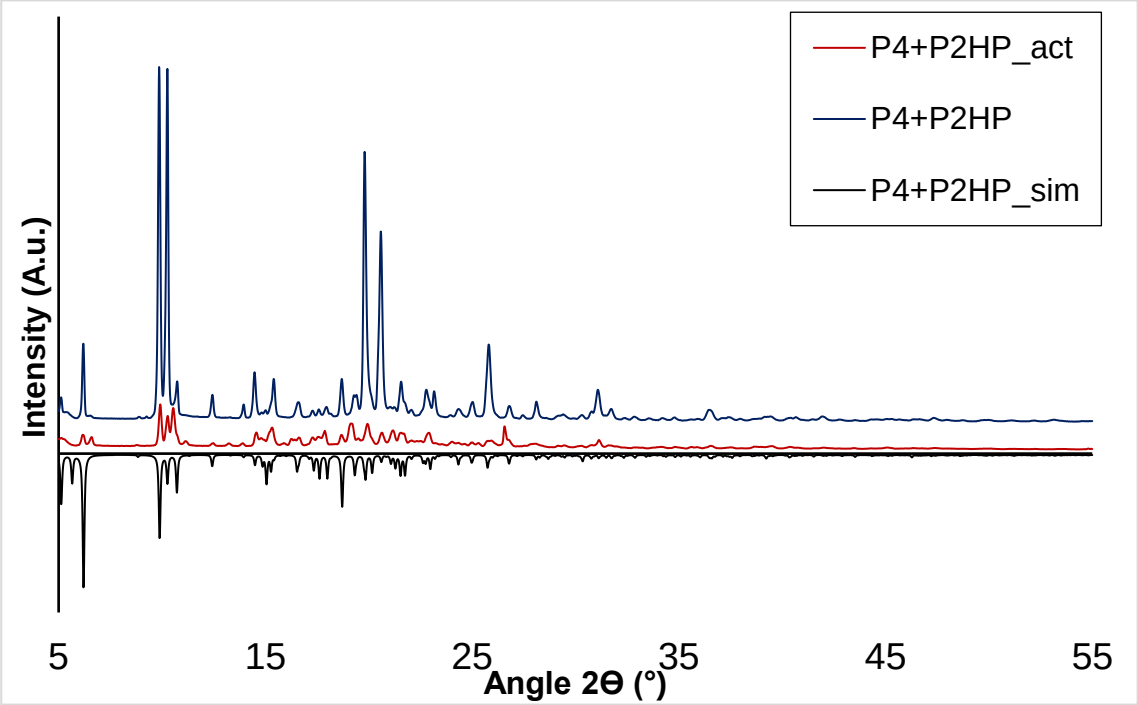


Figure S59 – PD-XRD of **P_4+ P_2HP** simulated from SC-XRD data (**black, bottom**), **P_4+ P_2HP** experimental PD-XRD (**blue, top**), **P_4+ P_2HP** after activation at 45 °C under high vacuum (**red**)

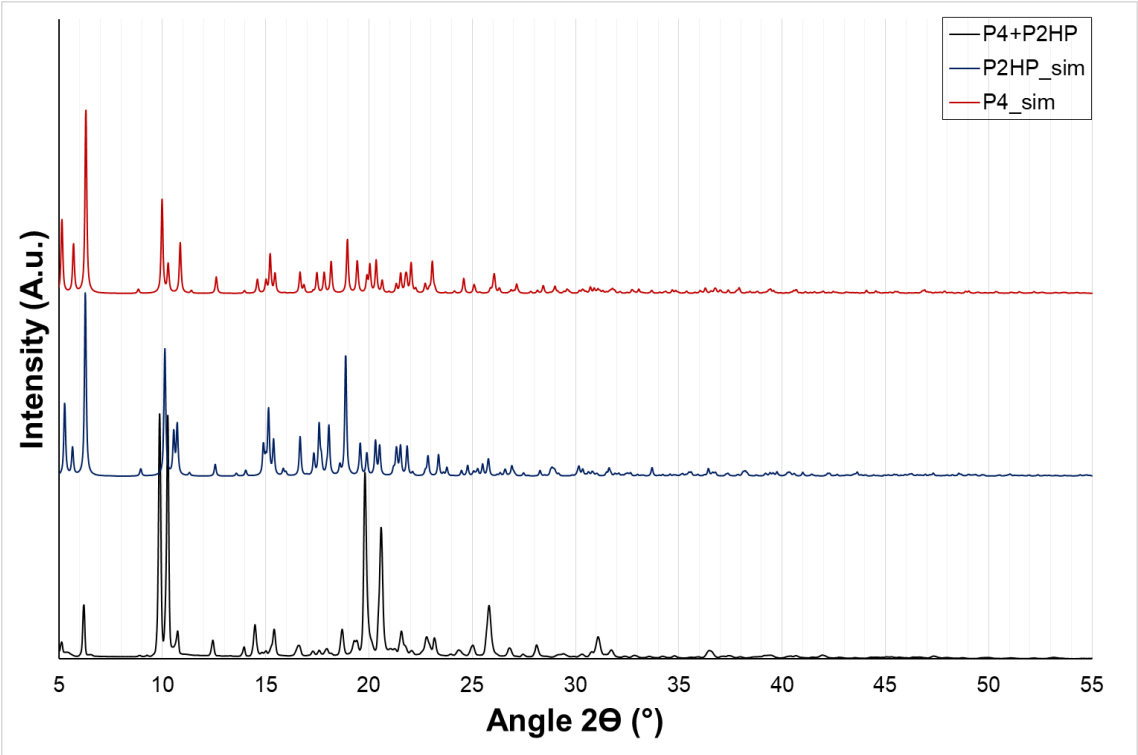


Figure S60 – Comparative powder diffraction pattern showing; **P_4+ P_2HP** experimental PD-XRD (**black, bottom**), **P_2HP** powder pattern simulated from SC-XRD data (**blue**), **P_4** powder pattern simulated from SC-XRD data (**red, top**).

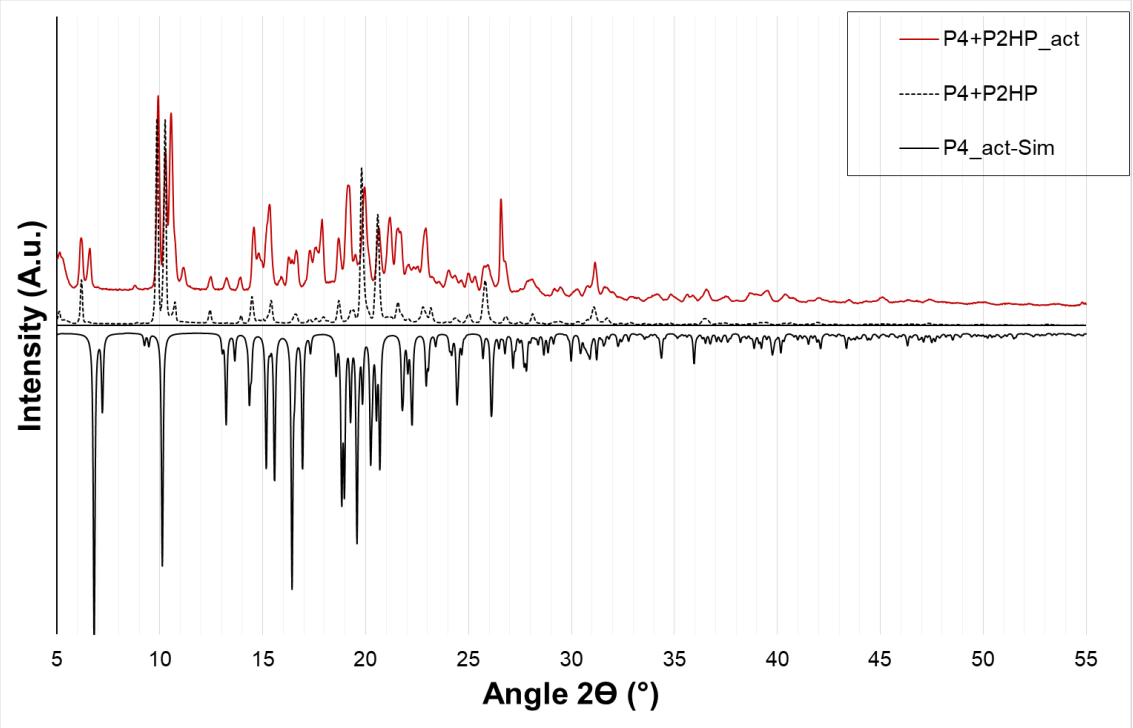


Figure S61 – Comparative powder diffraction pattern showing; **P_4+ P_2HP** experimental PD-XRD (**dotted**), **P_4+ P_2HP** experimental PD-XRD after activation at 45 °C under vacuum (**red**, top), **P_4** activated powder pattern simulated from SC-XRD data (**black, bottom**).[2] Highlights the partial activation and change in phase of the mixed peptide framework.

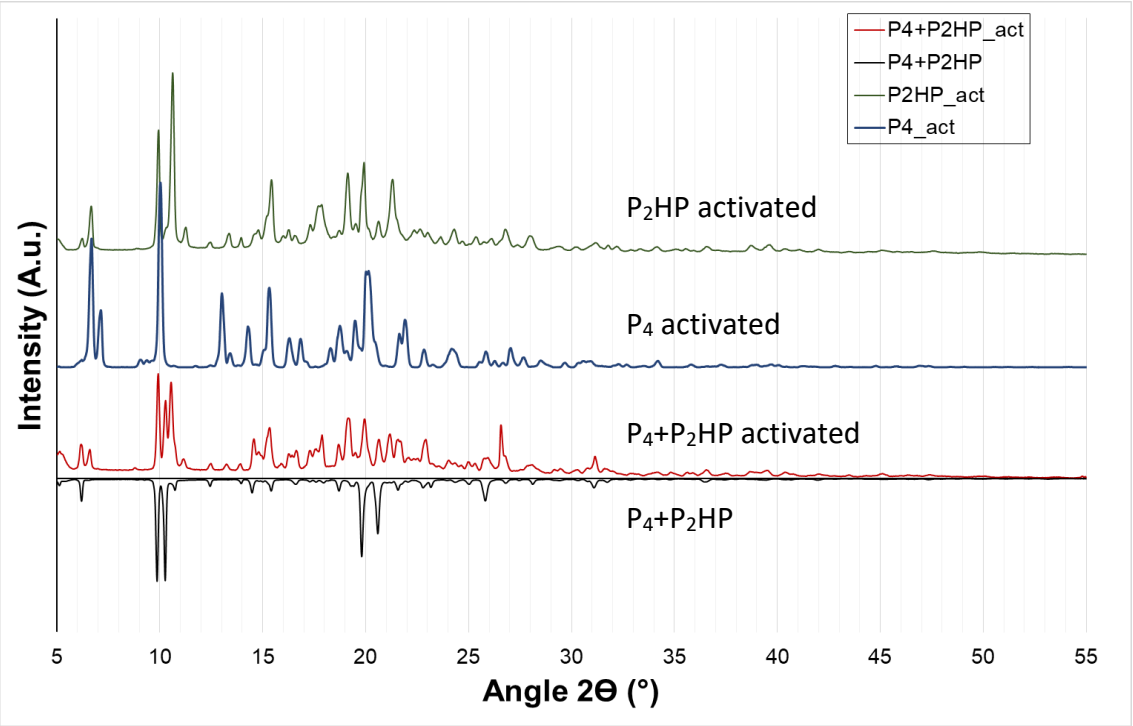


Figure S62 – Comparative powder diffraction pattern showing; **P_4+ P_2HP** experimental PD-XRD (**black, bottom**), **P_4+ P_2HP** experimental PD-XRD after activation at 45 °C under vacuum (**red**), **P_4** experimental PD-XRD after activation at 45 °C under vacuum (**blue**),[2] **P_2HP** experimental PD-XRD after activation at 45 °C under vacuum (**green**). Highlights the partial activation, and subsequent change in phase, of the mixed framework compared to **P_4/ P_2HP** activation, showing similarities to both pure frameworks.

SI 4.4.6 HP_2H PD-XRD:

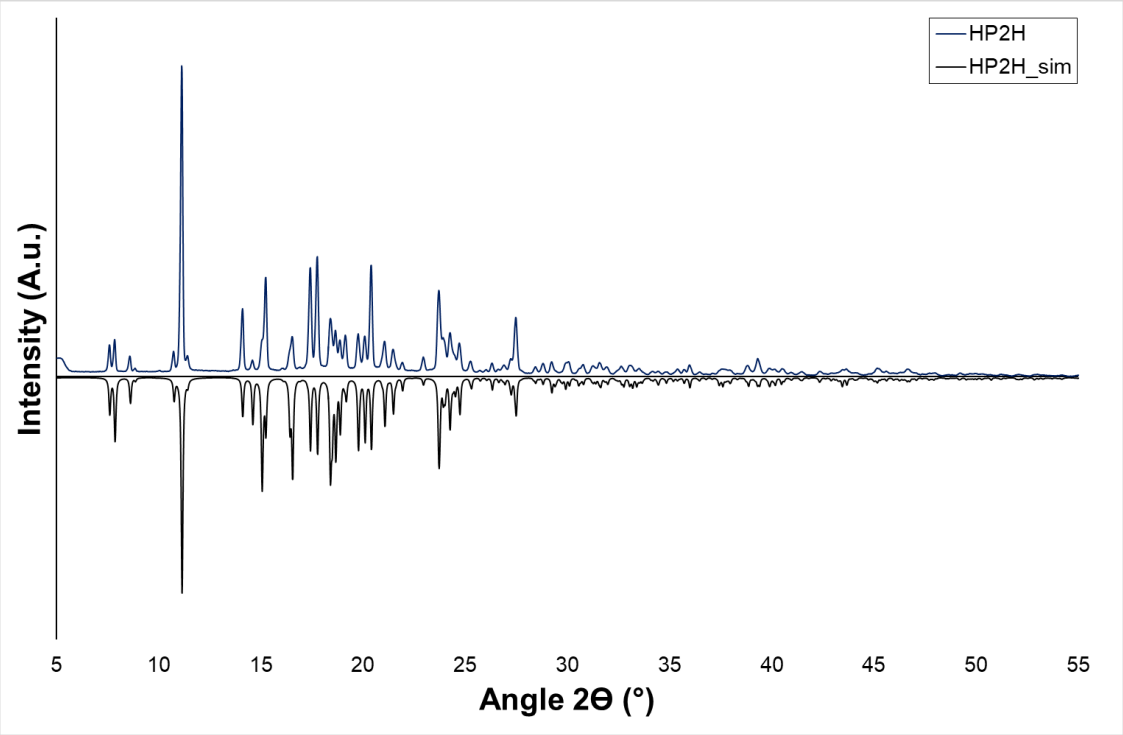


Figure S63 – PD-XRD of **HP_2H** simulated from SC-XRD data (**black, bottom**), **HP_2H** experimental PD-XRD (**blue, top**)

SI 4.4.7 *cis*-HP_2H PD-XRD:

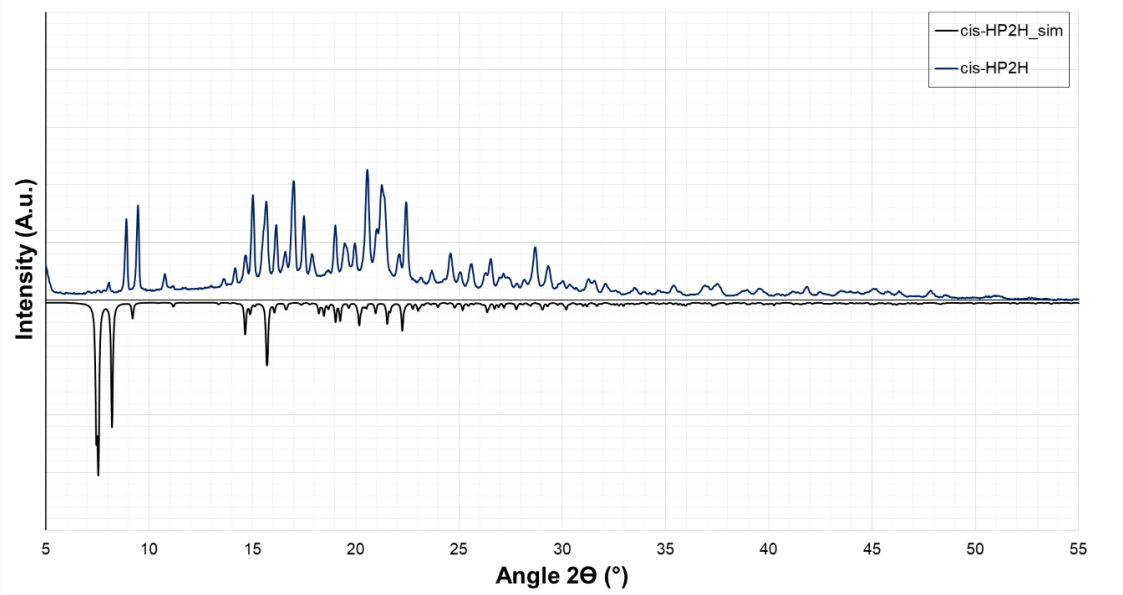


Figure S64 – PD-XRD of ***cis*-HP_2H** simulated from SC-XRD data (**black, bottom**), ***cis*-HP_2H** experimental PD-XRD (**blue, top**)

The experimental spectra for ***cis*-HP_2H** differs significantly from the simulated powder pattern (*Figure S64*), from the crystal structure obtained, suggesting the bulk material differs from the singly crystalline material. Screening of crystals for alternate crystal structures was unsuccessful, all crystals suitable for single crystal analysis exhibited the previously obtained structure.

SI 4.4.8 $AcHP_2H$ PD-XRD:

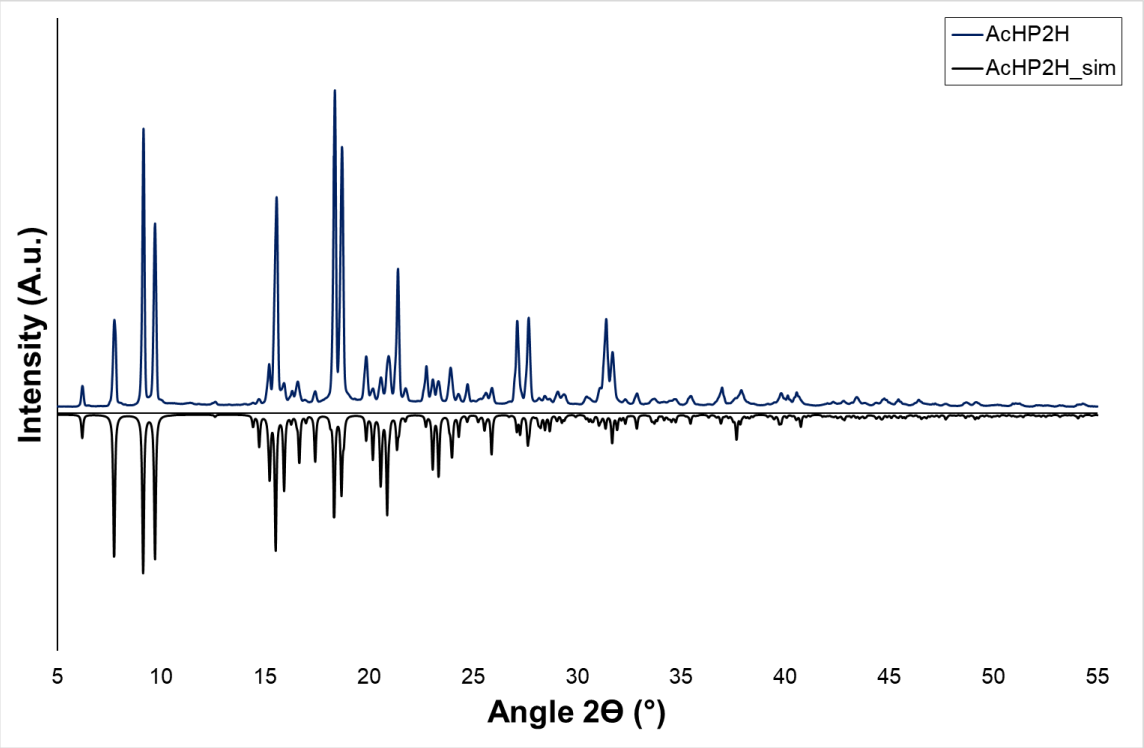


Figure S65 – PD-XRD of **$AcHP_2H$** simulated from SC-XRD data (**black, bottom**), **$AcHP_2H$** experimental PD-XRD (**blue, top**)

SI 4.4.9 AcP_4 PD-XRD:

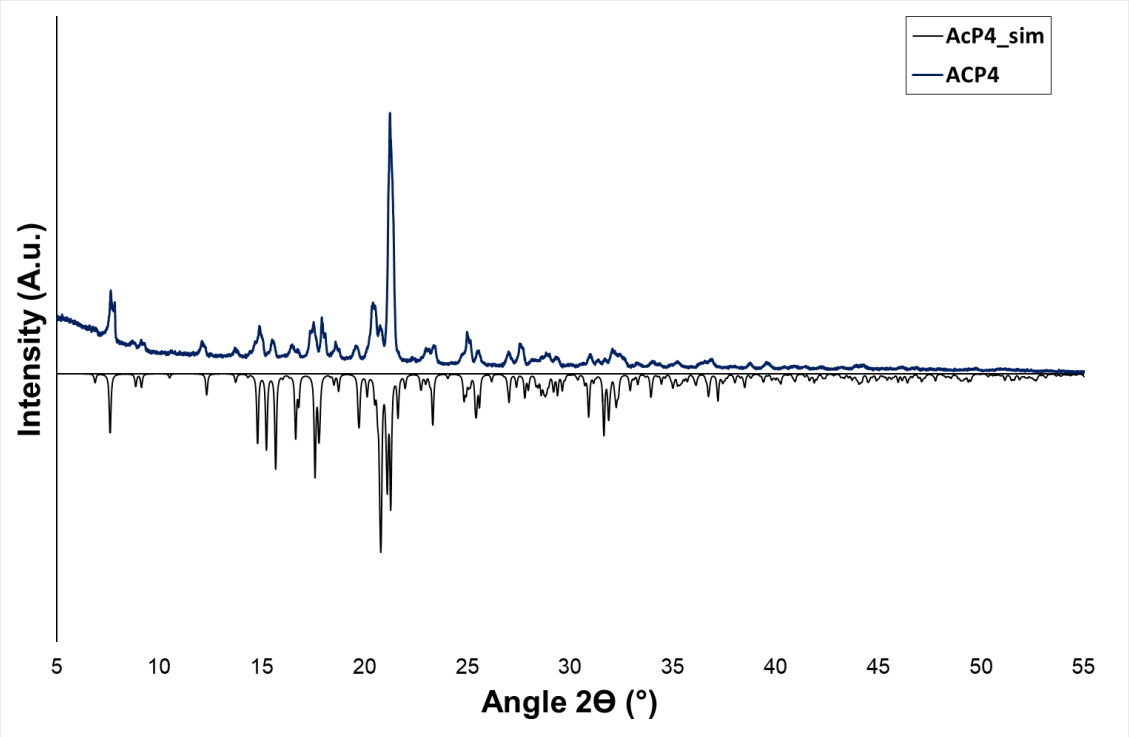


Figure S66 – PD-XRD of **AcP_4** simulated from SC-XRD data (**black, bottom**), **AcP_4** experimental PD-XRD (**blue, top**)

SI 4.5 AFM of self-assembled $AcHP_2H$

Atomic force microscopy (AFM) analysis: Samples were diluted as specified for each sample and 10 µL droplets were deposited on freshly cleaved mica discs (Agar Scientific F7013). For aqueous samples after 10-min incubation at room temperature, excess sample was removed by washing with 1 mL of 0.2-µm syringe-filtered mQ H_2O, and the specimens were then dried under a gentle stream of $N_{2\,(g)}$. For volatile solvents, after 1-min incubation at room temperature excess sample was removed and the mica discs washed with 10 µL (x 3) of 0.2-µm syringe-filtered solvent, and the specimens were then dried under a gentle stream of $N_{2\,(g)}$.Samples were imaged using a Bruker Multimode AFM with a Nanoscope V controller and a ScanAsyst probe (Silicone nitride tip with nominal tip radius = 2 nm, nominal spring constant 0.4 N/m, and nominal resonant frequency 70 kHz). Images were captured at a resolution of 88 nm per pixel scanned. All images were processed using the Nanoscope analysis software (version 1.5, Bruker).

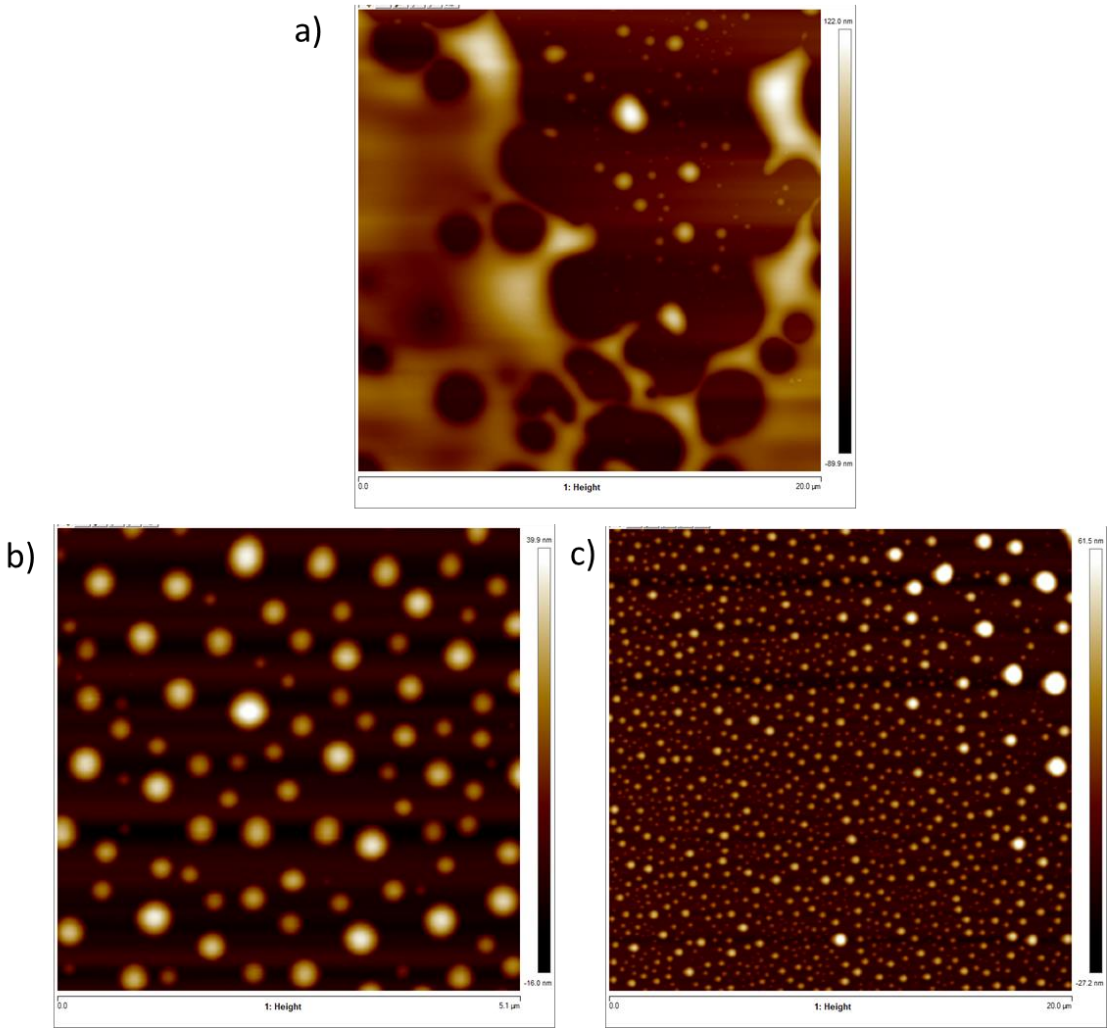


Figure S67 - AFM image of peptide **$AcHP_2H$** in $CHCl_3$ (6 mg/mL), 10 µL spotted, washed with $CHCl_3$ (x 3); a) average height 50 nm, b) object average height 19.5 nm, radius 80-170 nm; c) object height 17-50 nm, radius 160-550 nm.

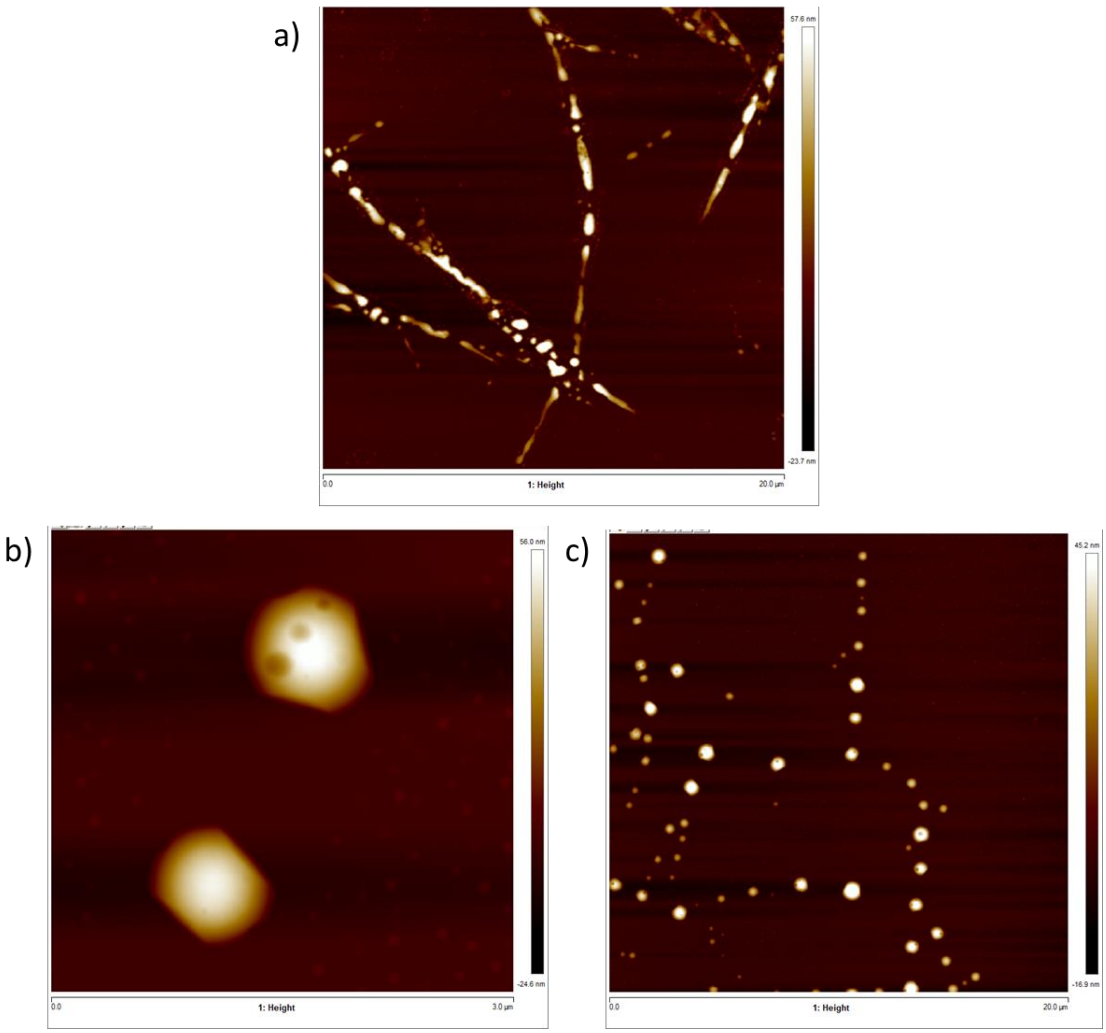


Figure S68 - AFM image of peptide **AcHP$_2$H** in $CHCl_3$ (1.2 mg/mL), 10 µL spotted; a) max height 63 nm; b) object average height 29 nm, radius 295 nm; c) object height 22-48 nm, radius 160-370 nm.

SI 4.6 Synthesis of phosphate ester peptide cages

SI 4.6.1 Reaction conditions tested:

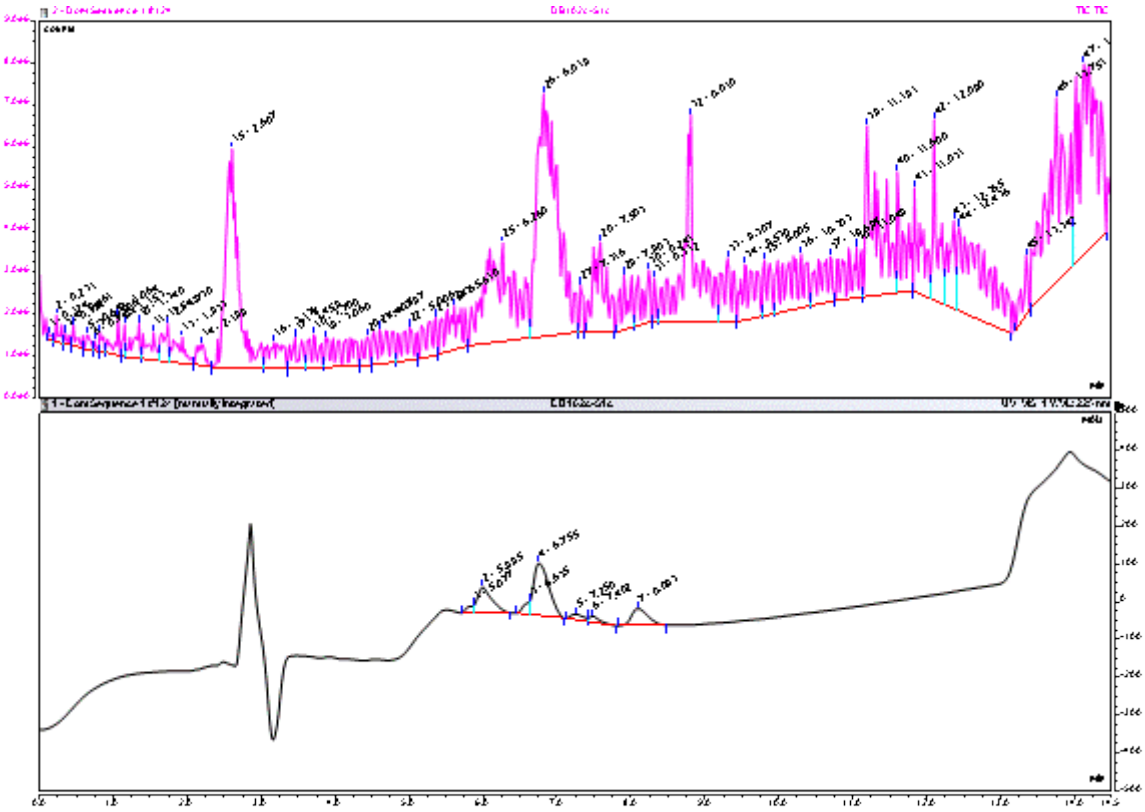
1) Peptide **AcHP$_2$H** (1 eq, 5.5 mg) was dissolved in pyridine-d_5 (600 µL) under a N_2 atmosphere. $POCl_3$ (0.6 eq, 0.65 µL) in dry DCM (50 µL) was then added to the solution causing a white precipitate to form and the solution was stirred for 1 h. The supernatant was then precipitated with cold Et_2O and the Et_2O precipitate washed with Et_2O (x 3) to remove excess pyridine. The yellow/white solid was then dissolved in 10% MeOH (aq.) and filtered through a 0.2 µm filter. This solution was then analysed *via* LCMS, showing a mixture of products. The precipitate from the reaction was analysed *via* ^{1}H NMR.
2) Peptide **AcHP$_2$H** (1 eq, 9.5 mg) was dissolved in dry pyridine (1 mL) under a N_2 atmosphere. $POCl_3$ (0.6 eq, 1.11 µL) in dry pyridine (1.5 mL) was then added dropwise over 1 h, no precipitate formed. The solution was then stirred at 60 °C for 4 h. HPLC analysis of an aliquot of the sample showed only starting material.
3) Peptide **AcHP$_2$H** (26.2 mg) was dissolved in pyridine-d_5 (650 µL) under a N_2 atmosphere. $POCl_3$ (0.66 eq, 3.4 µL) in dry DCM (50 µL) was then added to the solution causing a white precipitate to form, and the solution was stirred for 1 h. The supernatant was then precipitated with cold Et_2O and the Et_2O precipitate was washed with DCM (x 3). The yellow/white solid was then dissolved in 10% MeOH (aq.) and filtered through a 0.2 µm

filter. This solution was then separated *via* semi-prep HPLC (5-95% B, MeOH 0.1% formic acid, 9 min, 225 nm). NMR analysis was carried out on the two major products (rt = 7.009, 7.500).

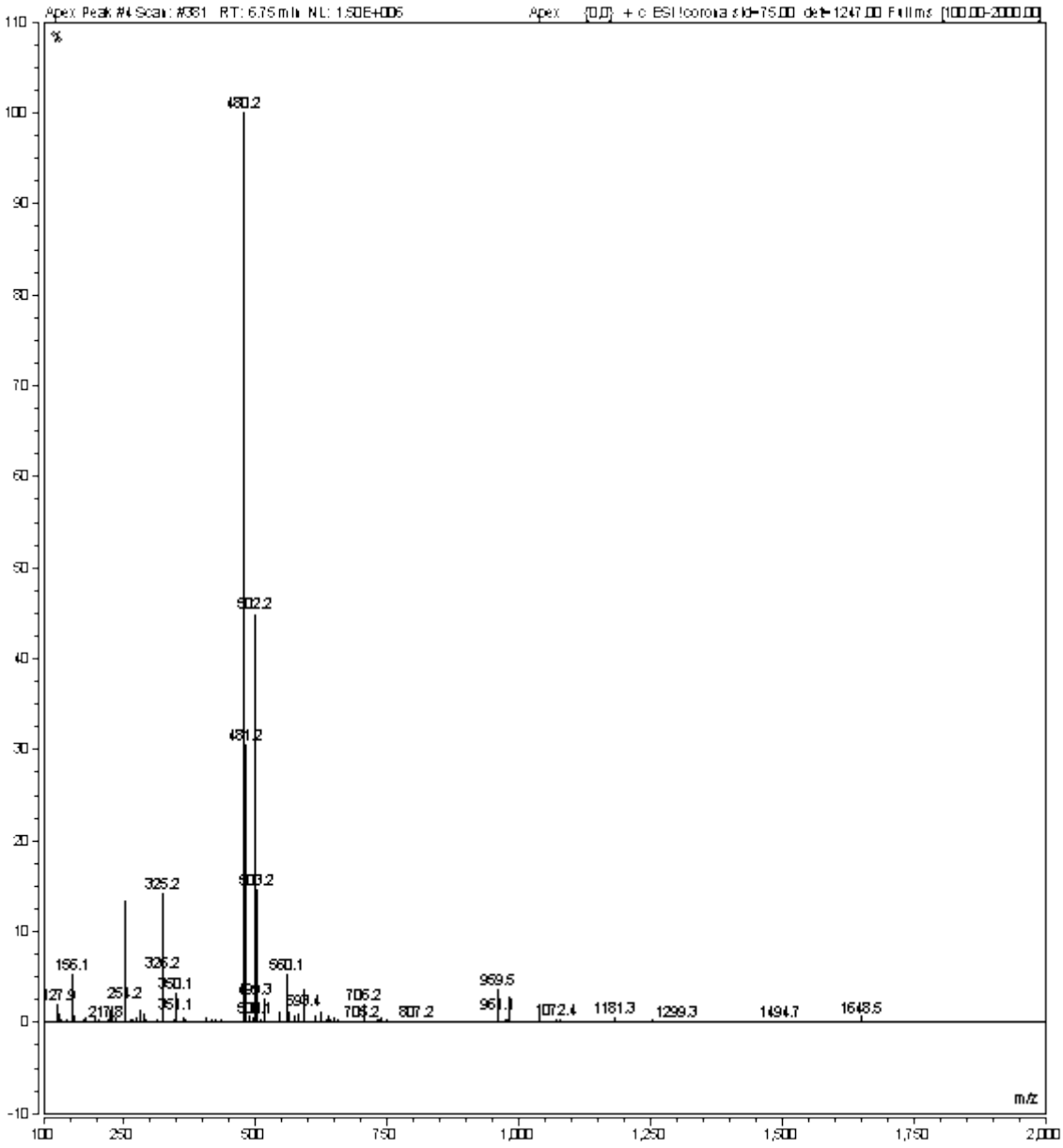
Mass and chemical formula for a trimeric AcHP$_2$H phosphate ester cage:

Chemical Formula: $C_{66}H_{93}N_{15}O_{23}P_2$ Exact Mass: 1525.6044

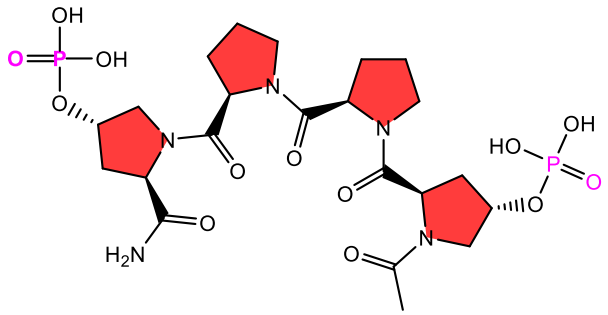
SI 4.6.2 HPLC-MS results from reaction conditions 1:



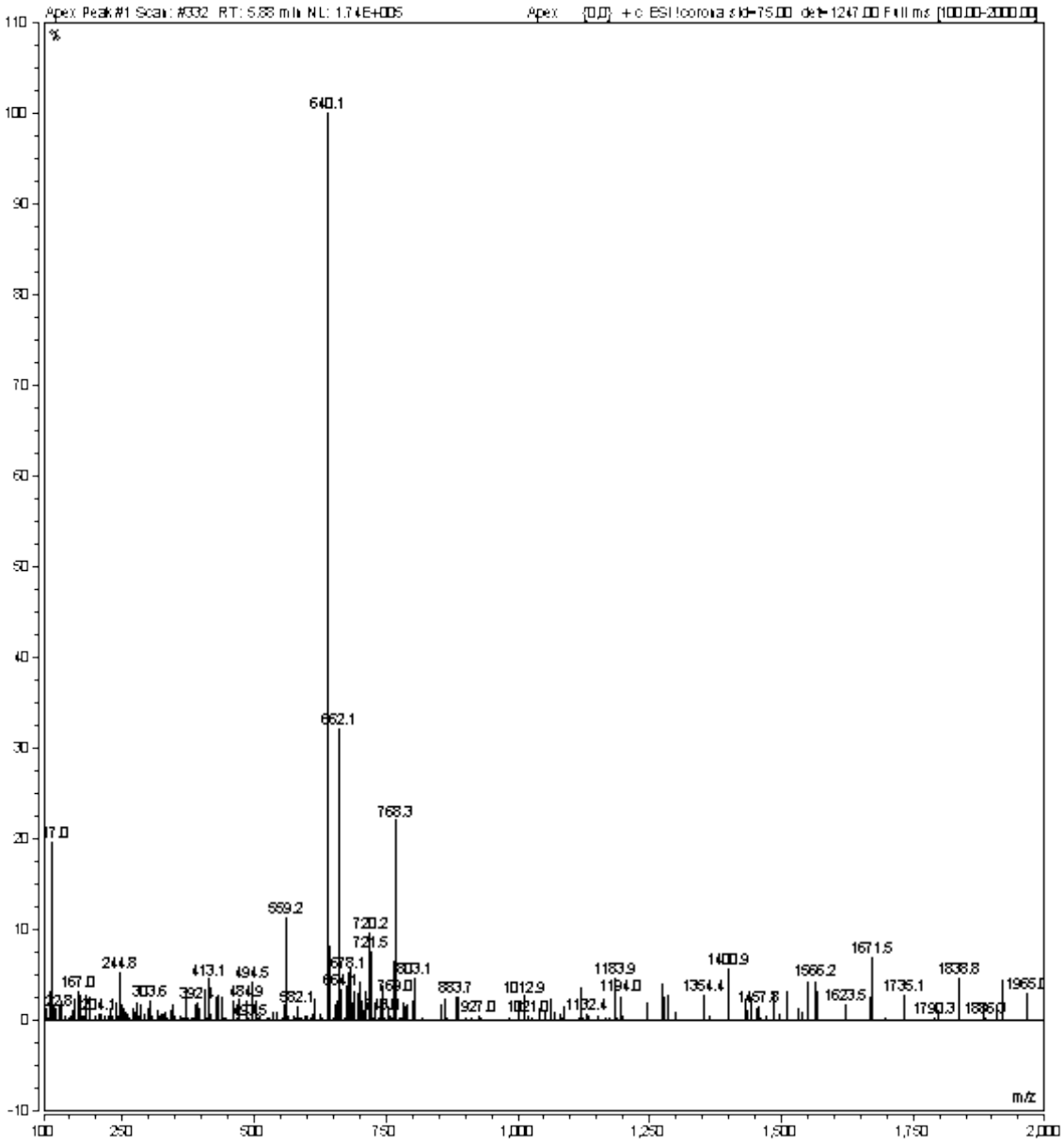
HPLC UV-Vis spectrum of peptide **AcHP$_2$H** after treatment with $POCl_3$ (method 1), rt 5.995 min, 6.755 min, 8.093 min



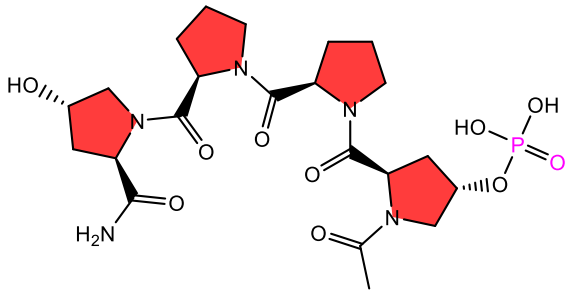
Mass spectrum (100-200 m/z), peptide **AcHP$_2$H (from POCl$_3$ reaction)**, peak at 6.755 min, m/z calcd. for $[M+H]^+$ $C_{22}H_{34}N_5O_7^+$: 480.2453 ; found: $[M+H]^+$ 480.2, $[M+Na]^+$ 502.2



Chemical Formula: $C_{22}H_{35}N_5O_{13}P_2$
Exact Mass: 639.1707

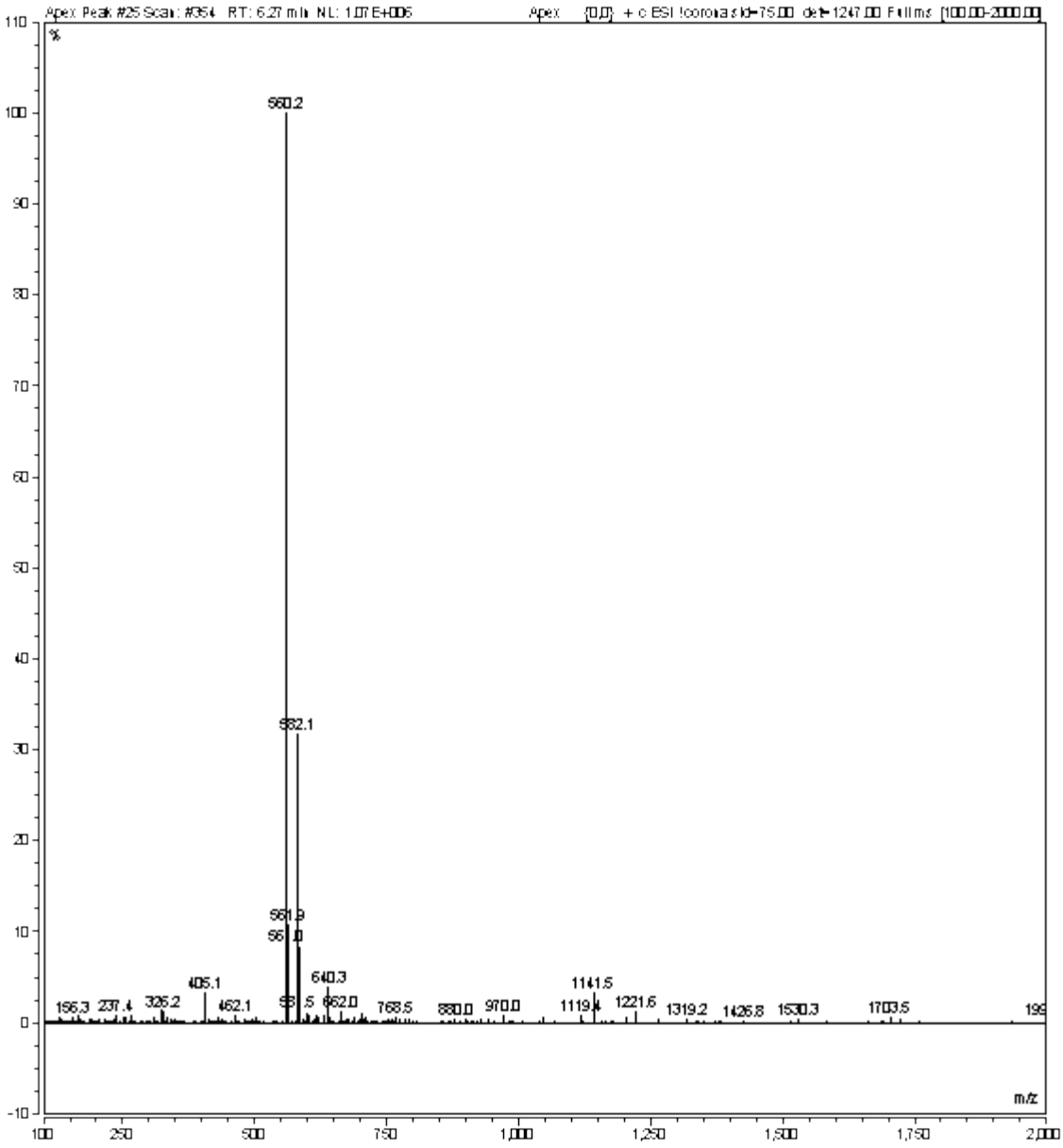


Mass spectrum (100-200 m/z), **AcHP$_2$H(-P=O(OH)$_2$)$_2$**, peak at 5.88 min, m/z calcd. for $[M+H]^+$ $C_{22}H_{36}N_5O_{13}P_2^+$: 640.1779 ; found: $[M+H]^+$ 640.1, $[M+Na]^+$ 662.1

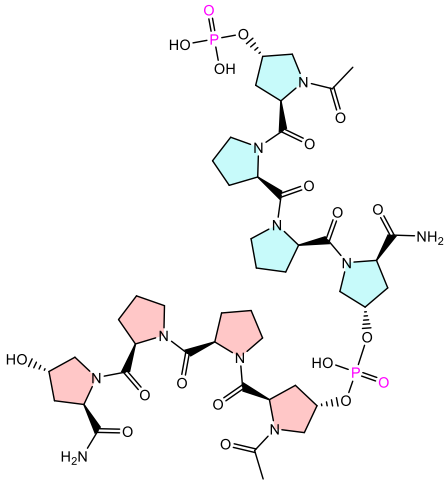


Chemical Formula: $C_{22}H_{34}N_5O_{10}P$
Exact Mass: 559.2043

(Either hydroxyl substitution possible 1 or 4)

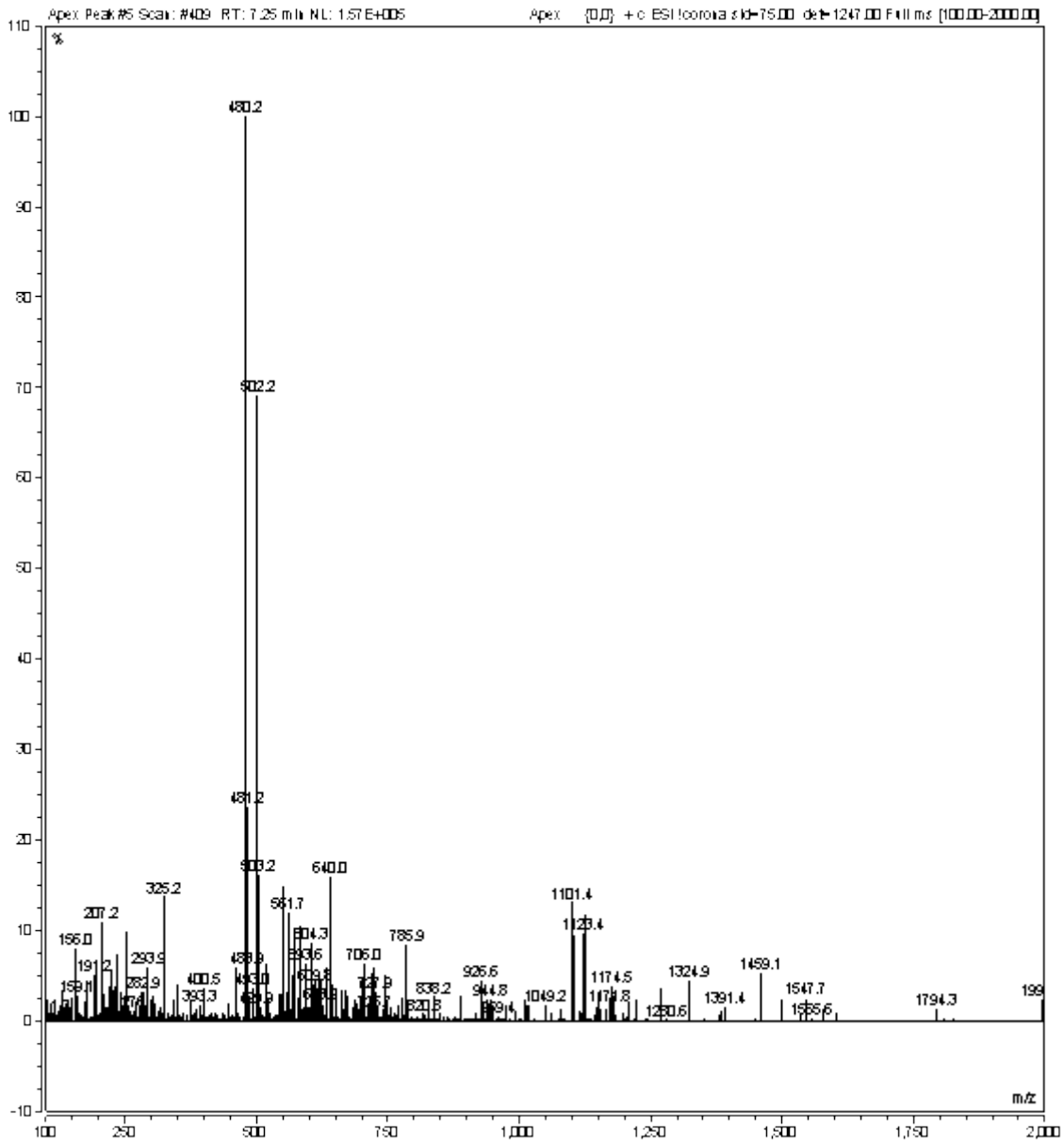


Mass spectrum (100-200 m/z), **$AcHP_2H(-P=O(OH)_2)$**, peak at 6.27 min, m/z calcd. for $[M+H]^+$ $C_{22}H_{36}N_5O_{10}P^+$: 560.2116; found: $[M+H]^+$ 560.2, $[M+Na]^+$ 582.1

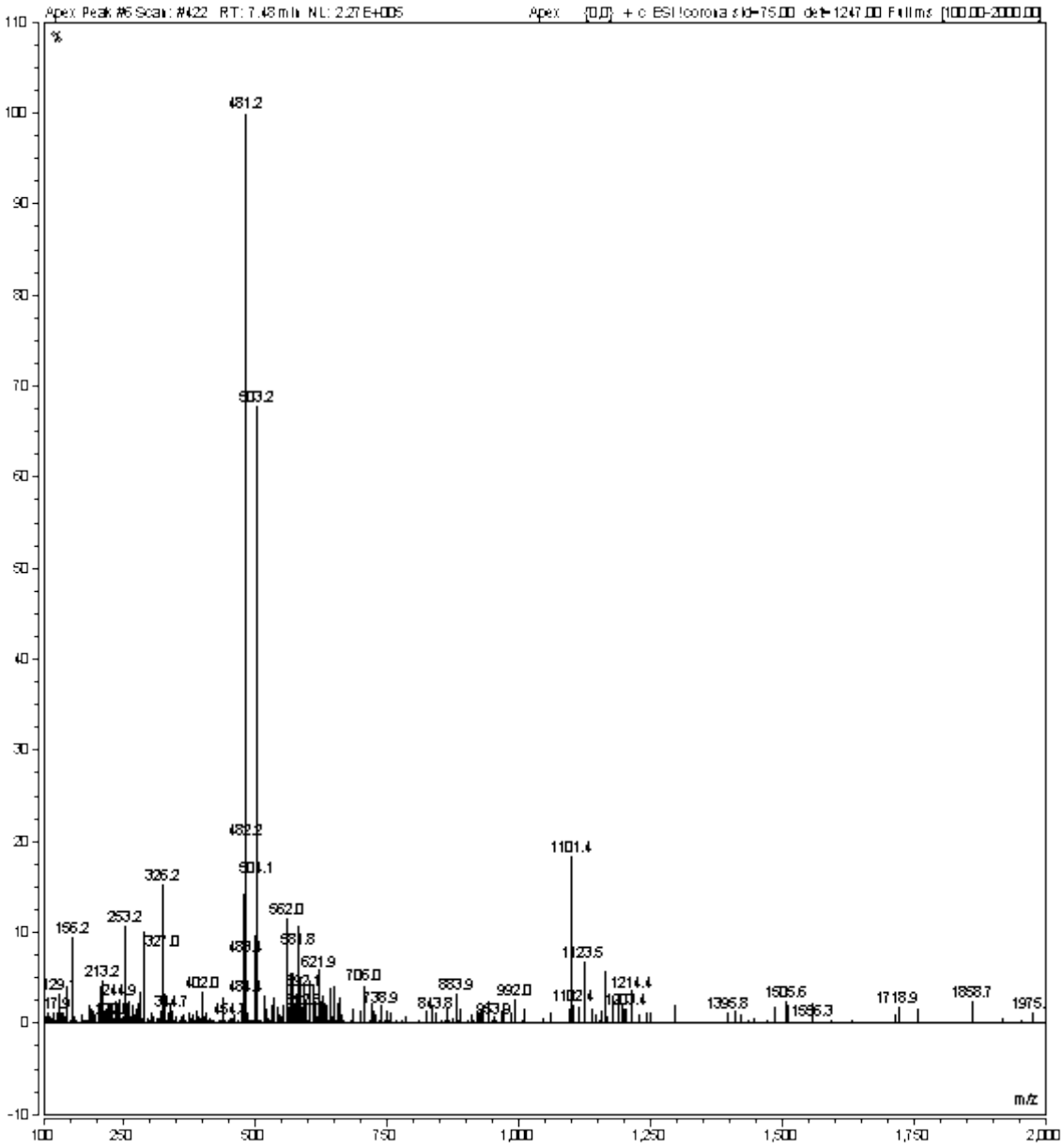


Chemical Formula: $C_{44}H_{66}N_{10}O_{19}P_2$
Exact Mass: 1100.3981

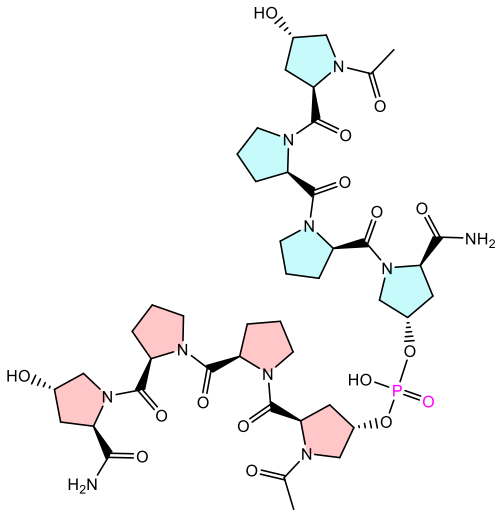
(Multiple stereoisomers of this compound – multiple peaks)



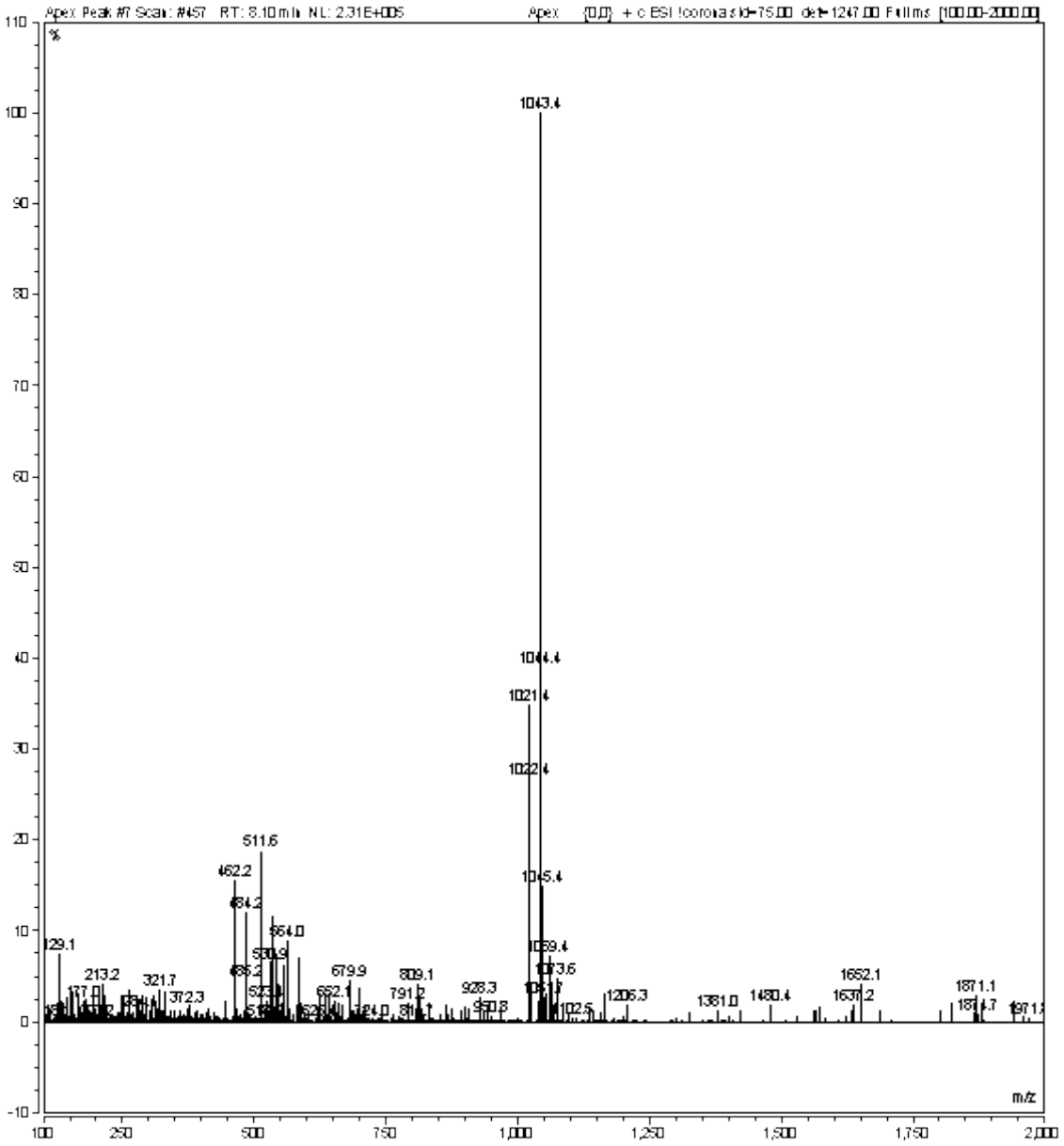
Mass spectrum (100-200 m/z), **$(AcHP_2H)_2$ $(-P=O(OH)_2)(-P=O(OH))$**, peak at 7.25 min, m/z calcd. for $[M+H]^+$ $C_{44}H_{67}N_{10}O_{19}P_2^+$: 1101.4054; found: $[M+H]^+$ 1101.4, $[M+Na]^+$ 1123.4



Mass spectrum (100-200 m/z), **$(AcHP_2H)_2(-P{=}O(OH)_2)(-P{=}O(OH))$**, peak at 7.48 min, m/z calcd. for $[M+H]^+$ $C_{44}H_{67}N_{10}O_{19}P_2^+$: 1101.4054; found: $[M+H]^+$ 1101.4, $[M+Na]^+$ 1123.5

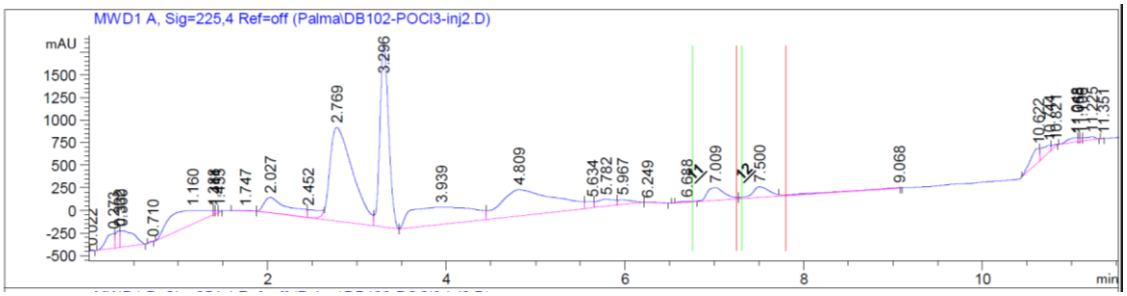


Chemical Formula: $C_{44}H_{65}N_{10}O_{16}P$
Exact Mass: 1020.4318

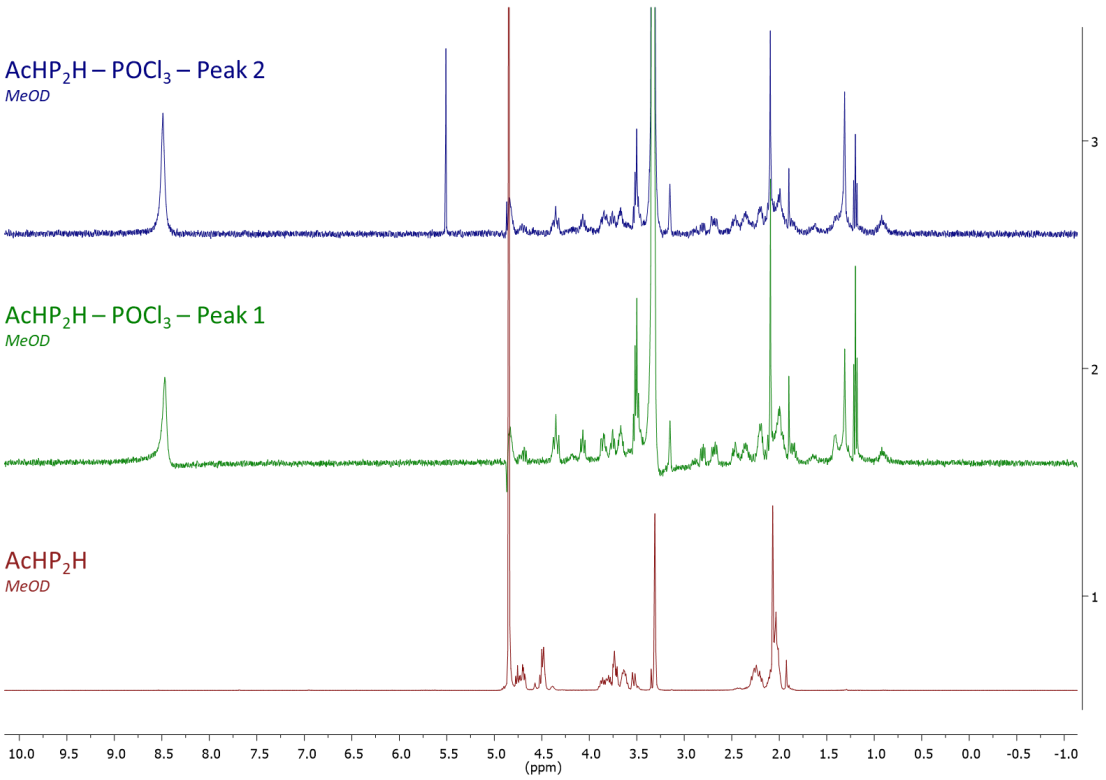


Mass spectrum (100-200 m/z), **$(AcHP_2H)_2$(-P=O(OH))**, peak at 8.10 min, m/z calcd. for $[M+H]^+$ $C_{44}H_{66}N_{10}O_{16}P^+$: 1021.4390; found: $[M+H]^+$ 1021.4, $[M+Na]^+$ 1043.4 $[M/2]^{2+}$ 511.6

SI 4.6.3 Prep-HPLC and 1H NMR – reaction conditions 3:



Prep-HPLC UV-Vis spectrum – Reaction conditions 3, **$AcHP_2H$** and $POCl_3$ in pyridine d_5, rt = 7.009 min (Fraction 11) and 7.500 min (Fraction 12), 225 nm



^{1}H NMR spectra – AcHP$_2$H, **AcHP$_2$H** reaction with $POCl_3$ (Conditions 3) peak 1 and 2 from prep-HPLC (MeOD, 400 MHz)

SI 4.7 References – 1

1 R. Dukor and T. Keiderling, *Biopolymers*, 1991, **31**, 1747–1761.

2 D. F. Brightwell, G. Truccolo, K. Samanta, E. J. Fenn, S. J. Holder, H. J. Shepherd, C. S. Hawes and A. Palma, *Chem. – A Eur. J.*, 2022, **28**, e202202368.

3 O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard and H. Puschmann, *J. Appl. Crystallogr.*, 2009, **42**, 339–341.

4 G. M. Sheldrick, *Acta Crystallogr. Sect. A Found. Crystallogr.*, 2008, 64, 112–122.

5 G. M. Sheldrick, *Acta Crystallogr. Sect. C Struct. Chem.*, 2014, **71**, 3–8.

6 G. M. Sheldrick, *Acta Crystallogr. Sect. A Found. Adv.*, 2015, **71**, 3–8.

7 Bruker APEX-3, 2016, Bruker-AXS Inc., Madison, WI.

8 SADABS 2016/2, 2016, Bruker-AXS Inc., Madison, WI.

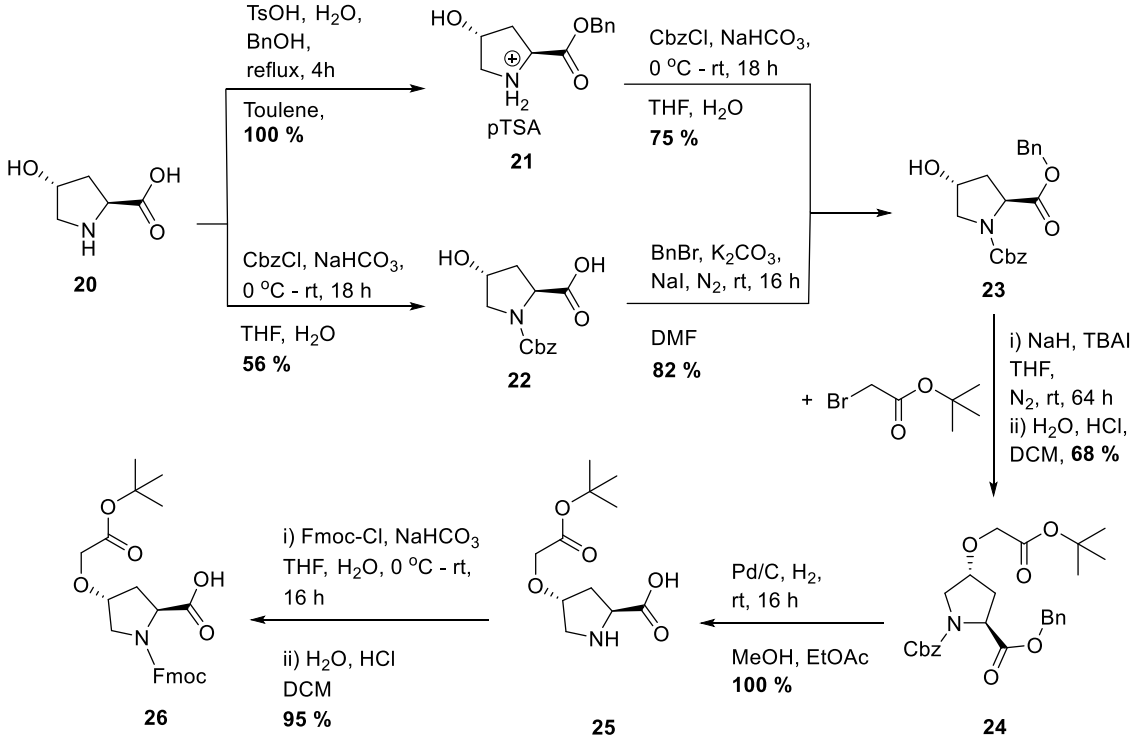
9 M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis and J. A. Montgomery, *J. Comput. Chem.*, 1993, **14**, 1347–1363.

10 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeerschd, E. Zurek and G. R. Hutchison, *J. Cheminform.*, 2012, **4**, 1–17.

11 G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders and T. Ziegler, *J. Comput. Chem.*, 2001, **22**, 931–967.

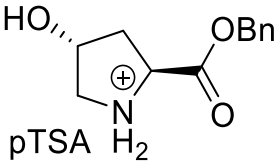
SI 4.8 Monomer Synthesis - 2

SI 4.8.1 Synthesis of 4-position carboxylic acid functionalised trans-fmoc-proline (26):



The target molecule for this synthesis, **26**, was desired for the synthesis of a polyproline peptide functionalised with carboxylic acid side chains. Work was initially carried out *via* compound **21**, due to access to other protecting groups for other syntheses. Further repeats and reaction scale ups were carried out with Cbz protection as the initial step, (**21**), due to higher overall yields.

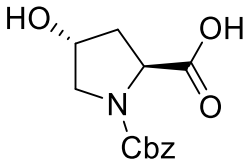
Synthesis of tosylate benzyl-(2S,4R)-4-hydroxypyrrolidine-2-carboxylate (21):



21

Trans-L-Hyp-OH, **20**, (1.00 g, 7.63 mmol, 1 eq.) was dissolved in toluene (10 mL). Benzyl alcohol (3.2 mL, 30.5 mmol, 4 eq.) was then added to this solution with p-TsOH.H_2O (1.17 g, 9.156 mmol, 1.2 eq.), the reaction was then heated at reflux for 4 h. The solution was then diluted with Et_2O (20 mL) and the resulting precipitate was isolated *via* filtration and washed with Et_2O to yield the product, **21**, as a white foam in a quantitative yield, this was used as is for the next step.

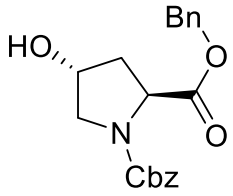
Synthesis of benzyl-1-(2S,4R)-4-hydroxypyrrolidine-1-carboxylate-2-carboxylic acid (22):



22

Trans-L-Hyp-OH, **20**, (7.0125 g, 53.48 mmol, 1 eq.) was dissolved aq. $NaHCO_3$ (30 mL). The resulting solution was stirred at 0 °C for 30 minutes before adding Cbz-Cl (9.2 mL, mmol, 1.2 eq.). The solution was stirred at 0 °C for 1h and then stirred for 18 h at rt. The reaction was monitored *via* TLC (6:4, Hex/EtOAc). Upon completion the reaction was diluted with deionised water (100 mL) and washed with cold Et_2O (x 3) while at pH 8. The solution was then acidified to pH 2 with 3 M HCl and extracted with EtOAc (x 3). The combined EtOAc extracts were then washed with brine (x 2), dried over anhydrous $MgSO_4$, filtered and the solvent was removed under vacuo to yield **22** as an oil (7.923 g, mmol, 56%). ^{1}H NMR (400 MHz, MeOD) δ 7.39 – 7.25 (m, 5H), 5.16 – 5.03 (m, 2H), 4.47 – 4.36 (m, J = 16.0, 8.0 Hz, 2H), 3.65 – 3.49 (m, 2H), 2.38 – 2.23 (m, J = 12.9, 9.7, 8.1, 2.9, 1.5 Hz, 1H), 2.16 – 2.01 (m, 1H). ^{13}C NMR (101 MHz, MeOD) δ 176.3, 176.0, 156.8, 156.5, 137.9, 137.7, 129.4, 129.1, 128.9, 128.5, 70.7, 70.0, 68.3, 59.3, 59.02, 56.0, 55.7, 40.2, 39.3. LCMS (ESI-MS) m/z: $[M+Na]^+$ calcd for $C_{13}H_{15}NO_5$ = 288.08, found = 288.00.

Synthesis of dibenzyl (*2S,4R*)-4-hydroxypyrrolidine-1,2-dicarboxylate (23):

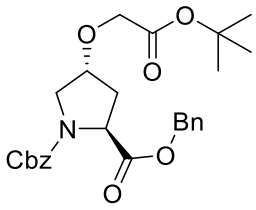


23

4-*trans*-L-Hydroxyproline benzyl ester tosylate salt, **21**, (8.444 g, 21.5 mmol, 1 eq.) and $NaHCO_3$ (4.52 g, 53.8 mmol, 2 eq.) were dissolved in water (32 mL) and after 5 min acetone (32 mL) was added followed by Cbz-Cl (3.2 mL, 22.5 mmol, 1.1 eq.). The resulting slurry was stirred overnight at room temperature. The reaction mixture was quenched adjusting to pH 1-2 with 3 M HCl and diluted with H_2O (300 mL), the solution was then extracted with Et_2O (3 x 350 mL). The combined organic extracts were washed with saturated aqueous $NaHCO_3$ (250 mL) and 5% aqueous NaCl (150 mL), dried over $MgSO_4$, filtered, and concentrated on a rotary evaporator. The residue was purified by flash chromatography (3:2 Hex/EtOAc) to give the title compound, **23** (5.69 g, 16.13 mmol, 75%), as a clear and colourless oil.

Compound **22** (7.923 g, 29.87 mmol, 1 eq.), K_2CO_3 (9.0814 g, 65.71 mmol, 2.2 eq.) and NaI (0.4477 g, 2.99 mmol, 0.1 eq.) were dissolved in DMF (22 mL) under a N_2 atmosphere. BnBr (8.82 mL, 89.6 mmol, 3 eq.) was added dropwise over 5 min to the stirring solution. The solution was stirred overnight at rt forming a yellow solution with a white precipitate. The reaction was monitored *via* TLC (3:2, Hex/EtOAc). The reaction was diluted with EtOAc and washed with deionised water (x 4) and brine (x 4), before drying over anhydrous $MgSO_4$ and filtering. The solvent was removed under vacuo to yield an orange oil, this was then washed with hexane (x 5) forming a pink extract. The washed oil was then purified *via* FCC (3:2, Hex/EtOAc) and concentrated to yield a yellowish oil, **23** (8.68 g, 24.4 mmol, 82%).^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.16 (m, 10H), 5.25 – 4.94 (m, 4H), 4.55 (dt, *J* = 15.6, 7.9 Hz, 1H), 4.45 – 4.34 (m, 1H), 3.68 – 3.49 (m, 2H), 3.08 (s, 1H), 2.35 – 2.21 (m, 1H), 2.08 – 1.97 (m, *J* = 8.1, 7.2, 4.7 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.68, 172.48, 155.18, 154.71, 136.40, 136.19, 135.57, 135.36, 128.57, 128.49, 128.44, 128.38, 128.29, 128.15, 128.09, 128.05, 128.01, 127.85, 127.78, 69.88, 69.15, 67.30, 67.27, 66.98, 66.85, 60.52, 58.14, 57.90, 55.23, 54.63, 39.11, 38.31, 21.06, 14.19.

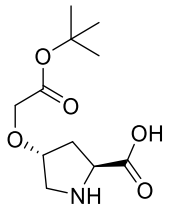
Synthesis of dibenzyl (2S,4R)-4-(2-(tert-butoxy)-2-oxoethoxy)pyrrolidine-1,2-dicarboxylate (24):



24

NaH, 60% in mineral oil, (1.2656 g, 31.62 mmol, 2.5 eq.) was placed in an oven dried flask under a N_2 atmosphere. Dry THF (30 mL) was then added slowly. *Tert*-butyl bromoacetate (7.46 mL, 50.50 mmol, 4 eq.) and TBAI (3.75 g, 10.2 mmol, 0.8 eq.) in dry THF (5 mL) was then added in one portion. The solution was stirred at rt for 20 min. Compound **23** (4.4870 g, 12.6 mmol, 1 eq.) in dry THF (18 mL) was added dropwise over 2 h. This formed an off-white cloudy solution and was left stirring at rt for 64 h hours. TLC analysis (6:4, Hex/EtOAc) was carried out to confirm reaction completion. The reaction was then diluted with deionised water (200 mL), acidified to pH 6-7 with 3 M HCl and quickly extracted with DCM (150 mL x 3). The combined organic layers were then washed with brine and dried over anhydrous $MgSO_4$. The remaining solvent was then removed under vacuo. The crude yellow oily product (16.5706 g) was then placed under high vac at 65 °C with a nitrogen trap, to remove residual *tert*-butyl bromoacetate yielding an oily partially yellow crystalline crude product (10. 9579 g). This was then purified by FCC (3:1, Hex/EtOAc), yielding compound **24** as an oil (4.0324g, 8.6 mmol, 68%).Spectral Data ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.53 – 7.37 (m, 9H), 7.37 – 7.31 (m, 1H), 5.38 (s, 0.13H), 5.35 (s, 0.3H), 5.30 (s, 0.32H), 5.28 (s, 0.17H), 5.27 (s, 1H), 5.18 (s, 1H), 5.13 (s, 1H), 4.67 (dt, 1H), 4.37 – 4.31 (m, 1H), 4.08 (dt, 2H), 3.95 – 3.72 (m, 2H), 2.66 – 2.49 (m, 1H), 2.29 – 2.18 (m, 1H), 1.60 (d, *J* = 3.4 Hz, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.49, 172.26, 169.18, 154.86, 154.34, 136.49, 136.37, 135.64, 135.42, 128.59, 128.50, 128.44, 128.37, 128.30, 128.14, 128.09, 128.05, 127.98, 127.93, 127.85, 82.00, 78.00, 77.26, 67.26, 67.20, 67.01, 66.96, 66.92, 66.82, 58.10, 57.86, 51.90, 51.79, 36.77, 35.48, 28.11.

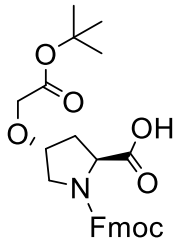
Synthesis of (*2S,4R*)-4-(2-(*tert*-butoxy)-2-oxoethoxy)pyrrolidine-2-carboxylic acid (25):



25

Pd/C 5% (0.2 g, 20% w/w) was placed in a three-neck flask with stirrer bar under a N_2 atmosphere. The catalyst was then covered with ethyl acetate (10 mL) and methanol (10 mL) was then added in a stream down the side of the flask. Compound **24** (0.7965 g, 1.7 mmol, 1 eq.) in methanol (10 mL) was then added. Stirring was then started and the flask was evacuated till the solvent was bubbling and refilled with nitrogen three times. The vessel place under a H_2 atmosphere with a H_2 balloon attached. The balloon was then left open to the flask with the needle underneath the solvent line and the reaction stirred overnight. The hydrogen balloon was refilled as needed. The reaction was monitored *via* TLC (8:2, Hex/EtOAc). After completion of the reaction the solution was filtered through a celite filter and the solvent removed under vacuo to yield compound **25** as a white crystalline solid (0.416 g, 1.7 mmol, 100%). Spectral Data: ^{1}H NMR (400 MHz, MeOD) δ 4.39 – 4.29 (m, 1H), 4.20 – 4.13 (m, 1H), 4.13 – 3.97 (m, 2H), 3.50 – 3.35 (m, 2H), 2.64 – 2.51 (m, 1H), 2.11 – 2.00 (m, 1H), 1.48 (d, *J* = 4.3 Hz, 9H). ^{13}C NMR (101 MHz, MeOD) δ 173.66, 171.49, 83.19, 80.33, 67.60, 61.36, 52.10, 36.34, 28.30.

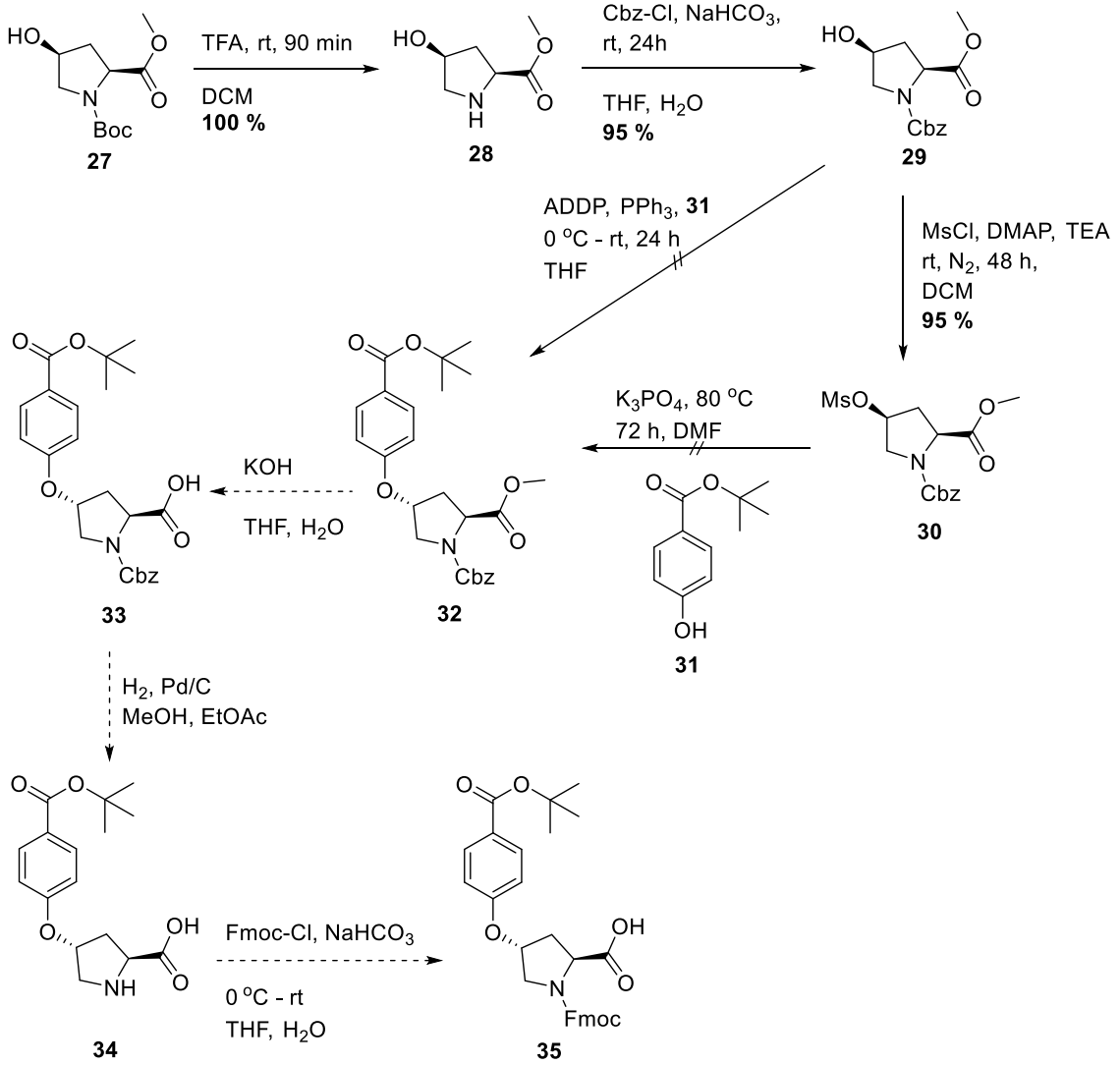
Synthesis of (*2S,4R*)-1-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-4-(2-(*tert*-butoxy)-2-oxoethoxy)pyrrolidine-2-carboxylic acid (26):



26

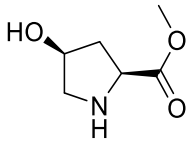
Compound **25** (1.43 g, 5.83 mmol, 1 eq.) was dissolved in THF (40 mL) and saturated aq. $NaHCO_3$ (40 mL) in an ice bath. Fmoc-Cl (1.81 g, 7.00 mmol, 1.2 eq.) was dissolved in THF (7 mL) and added to the solution at 0 °C. The pH was checked after 1h and adjusted with excess $NaHCO_3$ to pH 8. The solution was then left stirring overnight at rt. The reaction was monitored *via* TLC (6:4, Hex/EtOAc). The reaction was quenched with ice cold water (250 mL) and washed with cold Et_2O (150 mL x 3). The aqueous solution was then acidified to pH 2 with 3 M HCl and quickly extracted with DCM (500 mL x 3). The combined DCM layers were then washed with brine (x 2), dried over anhydrous $MgSO_4$, filtered, and the solvent removed under vacuo to yield a white crystalline solid, **26** (2.59 g, 5.54 mmol, 95%). Spectral Data: ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.76 (d, *J* = 7.5 Hz, 1.25H), 7.69 (d, *J* = 7.5 Hz, 0.75H), 7.60 – 7.50 (m, 2H), 7.43 – 7.23 (m, 4H), 6.57 (s, 1H), 4.56 – 4.29 (m, 3H), 4.28 – 4.09 (m, 2H), 4.02 – 3.86 (m, 2H), 3.81 – 3.57 (m, 2H), 2.57 – 2.48 (m, 0.38H), 2.47 – 2.37 (m, 0.62H), 2.30 – 2.21 (m, 0.62H), 2.21 – 2.11 (m, 0.38H), 1.48 (d, *J* = 8.8 Hz, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 177.21, 175.54, 169.34, 155.95, 154.69, 144.07, 143.82, 143.78, 141.40, 141.38, 141.32, 127.87, 127.74, 127.23, 127.18, 125.19, 125.05, 120.11, 120.00, 82.30, 77.85, 68.13, 68.04, 67.90, 67.05, 66.99, 58.14, 57.48, 51.92, 51.83, 47.20, 47.15, 36.99, 34.98, 28.21, 28.18, 25.67. Mass spectrometry: ESI-MS(m/z):[M-Na]+calcd.for$C_{26}H_{28}NNO_7Na$, 490.52;found, 490.1.

SI 4.8.2 Synthesis of (*2S,4R*)-1-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-4-(4-(*tert*-butoxycarbonyl)phenoxy)pyrrolidine-2-carboxylic acid (35):



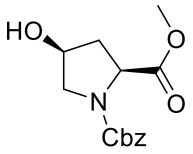
Synthetic route to achieve a benzoic acid sidechain on a Fmoc proline (**35**). This monomer is required to functionalise peptides, with the benzene ring adding rigidity and directionality to the interaction that may be lacking in the previously synthesised aliphatic carboxylic acid group. Two different routes were attempted for the substitution of the phenolic group as difficulties were encountered using the Mitsunobu reaction.

Synthesis of methyl (2S,4S)-4-hydroxypyrrolidine-2-carboxylate (28):



Compound **27** (2.0 g, 1 eq, 8.15 mmol) was dissolved in DCM (2 mL) and TFA (6.3 mL, 10 eq.), the resulting solution was then stirred for 90 min. TLC (8:2, Hex/EtOAc) confirmed no residual starting material. The solution was then concentrated under reduced pressure before drying with compressed air overnight to yield a viscous oil of **28** (2.4862 g), which was carried forward for the next step with no further purification.

Synthesis of 1-benzyl 2-methyl (2S,4S)-4-hydroxypyrrolidine-1,2-dicarboxylate (29):



29

Compound **28** (10.3 mmol, 1 eq.) was dissolved in THF (6 mL) and the pH was adjusted to pH 7-8 with aqueous sodium bicarbonate (sat.). The solution was stirred for 10 minutes before adding Cbz-Cl (1.69 mL, 11.9 mmol, 1.15 eq.) in one portion. The solution was stirred for 24 h before diluting with deionised water (90 mL) and then acidified to pH 3 with conc. HCl before extracting with EtOAc (200 mL x3). The combined organic layers were washed with aqueous sodium bicarbonate (x 2) and brine (x 2). They were then dried over anhydrous magnesium sulphate before filtering and concentrating under vacuum to yield a white crystalline solid, **29** (2.74 g, 9.8 mmol, 95%).

Single-crystal x-ray diffraction analysis was then carried out (CCDC – 2290760). The crystal structure data was obtained from colorless block crystals, crystallised by evaporation of an EtOAc solution. The crystals were stable outside of solution at room temperature showing no signs of deterioration over the timeframe of the experiment. Crystal data and structure refinement are available in the appendices.

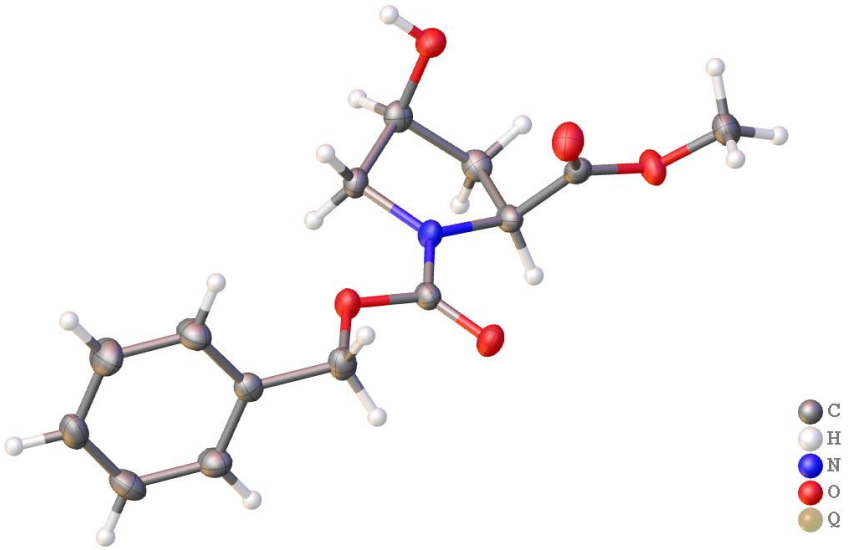
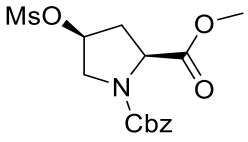


Figure 118 – Crystal structure of compound 29, asymmetric unit, 50% ellipsoids (Olex 2)

Synthesis of 1-benzyl 2-methyl (2S,4R)-4-((methylsulfonyl)oxy)pyrrolidine-1,2-dicarboxylate (30):

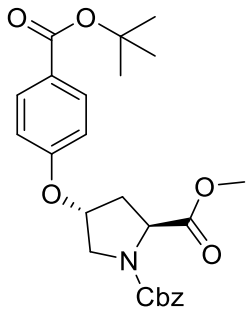


30

Compound **29** (571.8 mg) was dissolved in DCM (5 mL) under a nitrogen atmosphere. TEA (320 µL, 1.1 eq.), MsCl (170 µL, 1.05 eq.) and DMAP (27.8 mg, 0.1 eq.) were then added in sequence to the stirred solution. A white precipitate forms which dissolved over time, reaction left stirring for 48 h. Reaction monitored *via* TLC (1:1, Hex/EtOAc).The solution was concentrated under reduced pressure and partitioned between EtOAc (25 mL) and deionised water (10 mL). The

EtOAc extract was then washed with deionised water (x 4) and dried over anhydrous $MgSO_4$ before concentrating under reduced pressure to yield an oil, **30** (678.1 mg, 93%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40 – 7.27 (m, 5H), 5.27 – 5.05 (m, 3H), 4.54 (ddd, *J* = 27.9, 9.1, 2.3 Hz, 1H), 3.90 – 3.79 (m, 2H), 3.70 (m, 3H), 2.99 (s, 3H), 2.62 – 2.42 (m, 2H).

Synthesis of 1-benzyl 2-methyl (2S,4R)-4-(4-(tert-butoxycarbonyl)phenoxy)pyrrolidine-1,2-dicarboxylate (32):

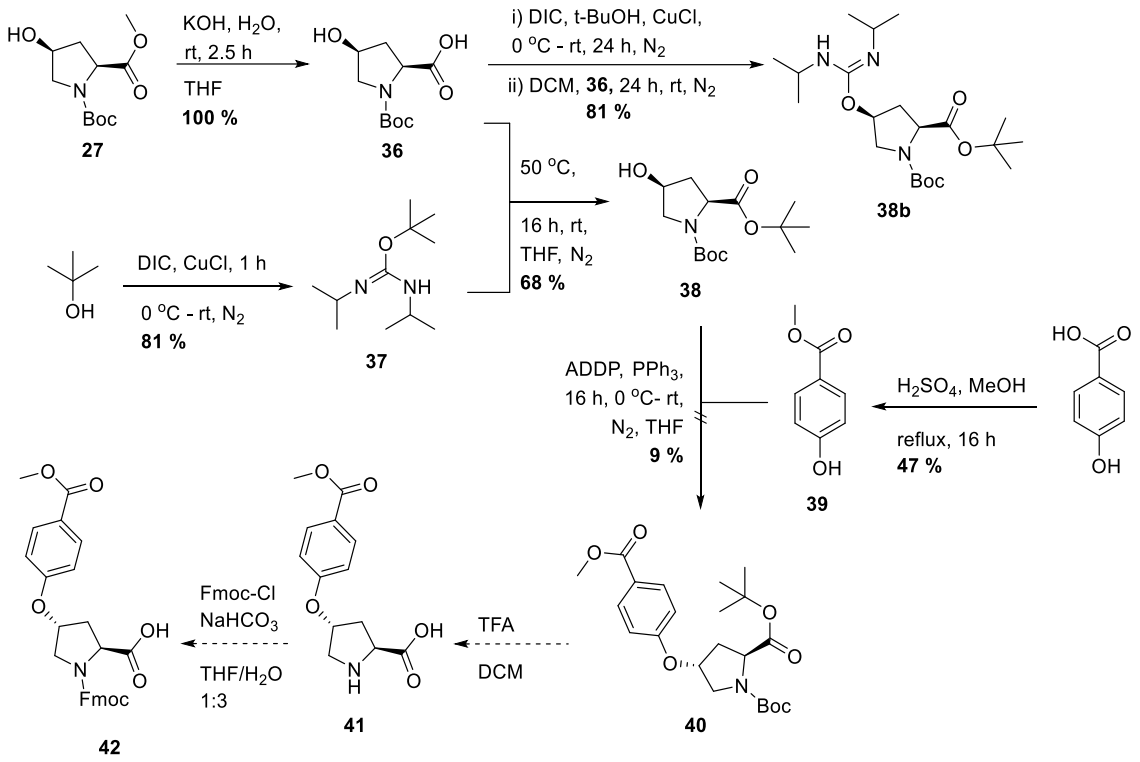


32

PPh_3 (751.3 mg, 2.86 mmol, 2 eq.), *tert*-butyl 4-hydroxybenzoate (556.4 mg, 2.86 mmol, 2 eq.) and compound **29** (400 mg, 1.43 mmol, 1 eq.) were dissolved in dry THF (5 mL) in a flask under N_2. DIAD (564 µL, 2.86 mmol, 2 eq.) dissolved in dry THF (4 mL) was then added dropwise over 40 min at 0 °C to the stirring solution of reactants. The resulting solution was then stirred at 0 °C for 30 min before stirring for 1 day at rt. TLC (1:1, Hex/EtOAc) was used to monitor the reaction. After completion the solution was concentrated under reduced pressure. Diethyl ether (100 mL) was then added causing a white crystalline solid to precipitate from the solution. The white solid was then filtered off and the resulting filtrate concentrated under reduced pressure to yield an orange oily solid. Multiple major products were present within the crude product suggesting the occurrence of significant side reactions.

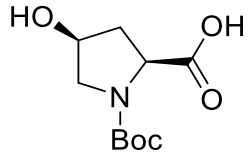
Compound **30** (628.7 mg, 1.8 mmol, 1 eq.) was dissolved in DMF (25 mL), to this stirred solution *tert*-butyl 4-hydroxybenzoate (683.4 mg, 3.5 mmol, 2 eq.) and K_3PO_4 (1.1203 g, 5.3 mmol, 3 eq.) was added. The solution was heated at 80 °C for 72 h. TLC (1:1, Hex/EtOAc) was used to monitor the reaction. The solution was diluted with deionised water (150 mL) causing a white precipitate to form. The suspension was then extracted with EtOAc (x 3) and the combined extracts were washed with water (x2) and brine (x 2). The organic extracts were then dried over anhydrous $MgSO_4$ and concentrated under reduced pressure to yield an orange oil (1.09 g). TLC analysis of the concentrated material shows several more highly coloured spots. This was confirmed by HPLC analysis with multiple major products. As such the product was not taken forward for continued synthesis.

SI 4.8.3 Synthesis of (*2S,4R*)-1-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-4-(4-(methoxycarbonyl)phenoxy)pyrrolidine-2-carboxylic acid (42):



Synthetic route to achieve a benzoic acid sidechain functionality on a proline monomer for incorporation into a peptide, adding a more rigid, directional group.

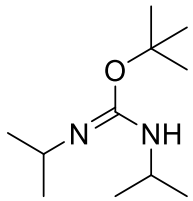
Synthesis of (2S,4S)-1-(tert-butoxycarbonyl)-4-hydroxypyrrolidine-2-carboxylic acid (36):



36

Compound **27** (5.003 g, 20.4 mmol, 1 eq.) was dissolved in THF (15 mL). KOH (2.3 g, 40.8 mmol, 2 eq.) was then added in one portion and deionised water (15 mL) added subsequently. The solution was stirred at rt for 2.5 h. The reaction was monitored *via* TLC (6:4, Hex/EtOAc). Upon completion the reaction was diluted with deionised water (100 mL), acidified to pH 4 and quickly extracted with EtOAc (x 3). The combined organic layers were then washed with brine (x 2), dried over anhydrous $MgSO_4$ and filtered. The solvent was then removed under vacuo to yield an oil that the formed a white crystalline solid, **36** (4.714 g, 20.4 mmol, 100%). ^{1}H NMR (400 MHz, MeOD) δ 5.00 (s, 2H), 4.40 – 4.24 (m, 2H), 3.66 – 3.58 (m, 1H), 3.36 (dd, *J* = 5.1, 2.3 Hz, 1H), 3.35 – 3.32 (m, 1H), 2.52 – 2.37 (m, 1H), 2.08 (ddd, *J* = 13.1, 8.9, 4.1 Hz, 1H), 1.47 (d, *J* = 16.2 Hz, 9H). ^{13}C NMR (101 MHz, MeOD) δ 176.32, 175.98, 156.27, 155.91, 81.56, 81.37, 70.80, 69.96, 59.20, 58.76, 55.57, 54.89, 39.57, 38.96, 28.70, 28.52.

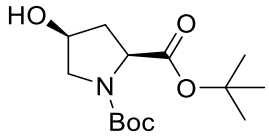
Synthesis of tert-butyl N,N'-diisopropylcarbamimidate (37):



37

$CuCl_2$ (127 mg, 1.3 mmol, 0.02 eq.) was dissolved in DIC (10 mL, 64 mmol, 1 eq.). The resulting solution was chilled to 0 °C and purged with N_2 before adding t-BuOH (6.36 mL, 67 mmol, 1.05 eq.). The solution was then stirred for 1h at 0 °C and overnight at rt, forming a green solution with precipitate. The solution was then diluted with hexane and filtered through a short alumina pad. Crystallised urea was then filtered off and the solvent removed under reduced pressure to yield a liquid, product **37** (10.34 g, 51 mmol, 81%). The product was then taken forward without any further purification. ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.63 (s, 0.1H), 3.50 (s, 0.3H), 3.43 (s, 0.5H), 3.00 (s, 0.4H), 2.91 (s, 0.55H), 1.24 (s, 4H), 1.15 (s, 2H), 1.04 (s, 4H), 0.85 (s, 9H), 0.65 (s, 1H).

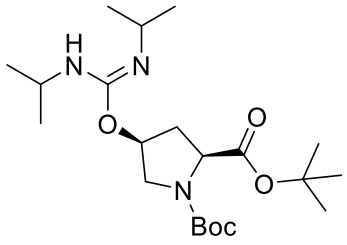
Synthesis of di-tert-butyl (2S,4S)-4-hydroxypyrrolidine-1,2-dicarboxylate (38):



38

Compound **36**, (2.32 g, 10.3 mmol, 1 eq.) was dissolved in dry THF (31 mL) under a N_2 atmosphere and compound **37** (2.4 mL, 10 mmol, 1 eq.) was added. The solution was stirred for 30 min forming a precipitate. Further **37** (1.2 mL, 5 mmol, 0.5 eq.) was then added and the solution stirred overnight. TLC (1:1, Hex/EtOAc) was used to monitor the reaction. The solution was then filtered and concentrated under vacuum to yield an oil (3.1102 g). This was purified *via* FCC (1:1, Hex/EtOAc) to yield an oily product, **38** (2.012 g, 7.0 mmol, 68%).

Synthesis of di-tert-butyl (2S,4S)-4-(((E)-N,N'-diisopropylcarbamimidoyl)oxy)pyrrolidine-1,2-dicarboxylate (38b):

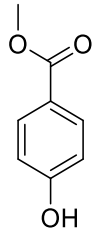


38b

t-BuOH (3.7 mL, 39 mmol, 7 eq.) was added to CuCl (34 mg, 0.3 mmol, 0.06 eq.) and DIC (5.25 mL, 6 eq, 33 mmol) under a N_2 atmosphere. The resulting solution was stirred for 24 h forming a cloudy green solution. The solution was diluted with dry DCM (19.2 mL) and added to a stirred solution of compound **27** (1.2877 g, 5.6 mmol, 1 eq.) in dry DCM (19.2 mL) under N_2. The solution was stirred for a further 24 h forming a light green solution with some precipitate. TLC (1:1, EtOAc/Hex) was used to monitor the reaction. The solution was then concentrated and filtered through a short celite pad. The product was then isolated *via* FCC to afford the byproduct, **38b** , as an oil (1.62g, 70%).^{1}H NMR (400 MHz, $CDCl_3$) δ 5.23 (d, *J* = 2.7 Hz, 1H), 4.28 (dd, *J* = 9.6, 3.2 Hz, 0.3H), 4.18 (dd, *J* = 9.8, 3.0 Hz, 0.7H), 3.78 – 3.62 (m, 2H), 3.54 (dd, *J* = 12.2, 2.2 Hz, 0.65H),

3.42 (dd, *J* = 12.0, 2.8 Hz, 0.35H), 3.31 (s, 1H), 3.11 (d, *J* = 6.0 Hz, 1H), 2.53 – 2.34 (m, 1H), 2.22 (dd, *J* = 6.4, 3.3 Hz, 0.65H), 2.19 (t, *J* = 3.0 Hz, 0.35H), 1.85 (d, *J* = 20.4 Hz, 1H), 1.45 (s, 9H), 1.44 (s, 3H), 1.42 (s, 6H), 1.04 (dd, *J* = 9.8, 5.4 Hz, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.04, 154.25, 154.02, 150.21, 80.85, 80.66, 79.70, 79.65, 72.27, 71.41, 58.48, 58.22, 52.50, 52.34, 46.12, 43.18, 36.16, 35.37, 28.45, 28.36, 28.01, 24.42, 24.12, 23.87. ESI-MS: (m/z):$[M+H]^+$calcd. for $C_{21}H_{40}N_3O_5^+$, 414.2963; found, $[M+H]^+$, 414.2; $[2M+H]^+$, 827.4

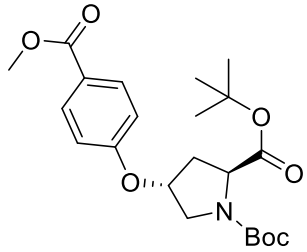
Synthesis of methyl 4-hydroxybenzoate (39):



39

4-Hydroxybenzoic acid (5.0094 g, 26.2 mmol, 1 eq.) was suspended in MeOH (33 mL). Conc. H_2SO_4 (5 mL) was slowly added to the stirred solution and the resulting solution was heated at reflux overnight. The solution was then cooled to rt, diluted with deionised water (700 mL) and neutralised with $NaHCO_3$ to pH7-8. The white precipitated formed was then filtered off and washed with deionised water. The resulting white solid was then dissolved in acetone and concentrated under reduced pressure to yield a white solid, **39** (2.5775 g, 12.3 mmol, 47%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.95 (d, *J* = 8.9 Hz, 2H), 6.89 (d, *J* = 8.9 Hz, 2H), 3.90 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 167.46, 160.26, 131.99, 122.30, 115.31, 52.15.

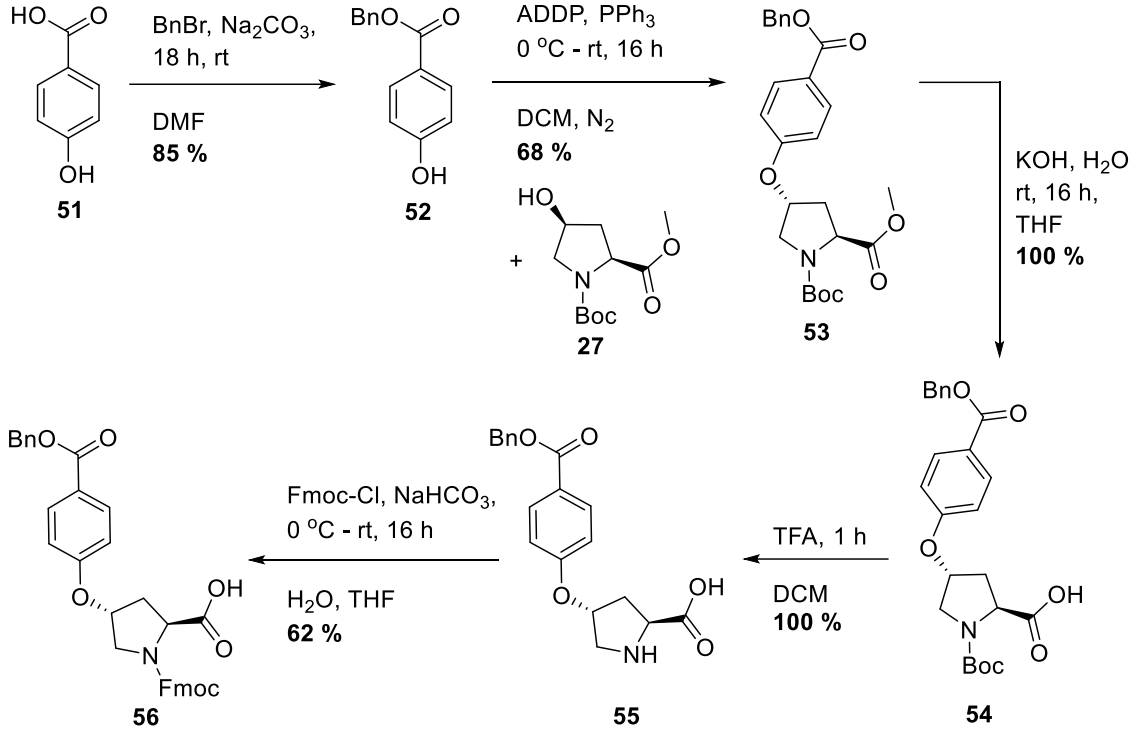
Synthesis of di-tert-butyl (2S,4R)-4-(4-(methoxycarbonyl)phenoxy)pyrrolidine-1,2-dicarboxylate (40):



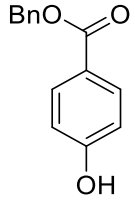
40

ADDP (772 mg, 3.1 mmol, 1.4 eq.) and PPh_3 (802 mg, 3.1 mmol, 1.4 eq.) were dissolved in dry THF (12 mL) at 0 °C under a N_2 atmosphere. The solution was stirred for 30 min before adding compound **38** (627.5 mg, 2.2 mmol, 1 eq.) and **39** (400 mg, 2.6 mmol, 1.2 eq.) dropwise over 15 min in dry THF (6 mL) at 0 °C. The reaction was stirred overnight forming a yellow precipitate. TLC (8:2, Hex/EtOAc) was used to monitor the reaction. The solution was concentrated under vacuo and precipitated in cold Et_2O. The precipitate was then filtered off and the crude oil was then purified *via* FCC returning a negligible yield of the product.

SI 4.8.4 Synthesis of (*2S,4R*)-1-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-4-(4-((benzyloxy)carbonyl)phenoxy)pyrrolidine-2-carboxylic acid (56):

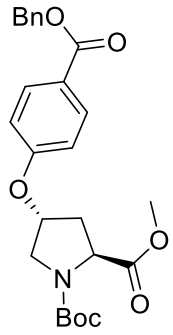


Synthesis of benzyl 4-hydroxybenzoate (52):



4-hydroxybenzoic acid, **51**, (1.986 g, 14.4 mmol, 1 eq.) was dissolved in DMF (20 mL) with Na_2CO_3 (1.524 g, 14.4 mmol, 1 eq.). Benzyl bromide (1.76 mL, 14.4 mmol, 1 eq.) was then added to the stirred solution and the solution stirred for 18 h at rt. The solution was then diluted with deionised water (150 mL) causing a white precipitate to form and extracted with Et_2O (x 4). The combined organic extracts were then washed with saturated aqueous $NaHCO_3$ (x 2), deionised water (x 5), brine (x 2) and dried for 1 h over anhydrous $MgSO_4$. The solvent was then removed under vacuo to yield a white crystalline solid, **52** (2.7894 g, 85%). ESI-MS (m/z):$[M+H]^+$calcd. for $C_{14}H_{13}O_3$, 229.0859; found, $[2M+Na]^+$, 479.5.

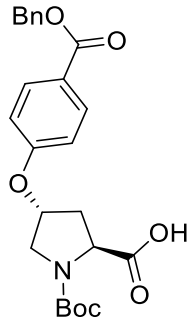
Synthesis of 1-(tert-butyl) 2-methyl (2S,4R)-4-(4-((benzyloxy)carbonyl)phenoxy)pyrrolidine-1,2-dicarboxylate (53):



53

ADDP (730.5 mg, 2.90 mmol, 1.5 eq.), PPh_3 (784.7 mg, 3.00 mmol, 1.55 eq.), compound **27** (473.4 g, 1.93 mmol, 1 eq.), and compound **52** (528.6 mg, 2.32 mmol, 1.2 eq.) were placed in an oven dried flask under a N_2 atmosphere. The reactants were then partially dissolved in dry DCM (7 mL) at 0 °C, and the resulting solution stirred for 1h and then warmed to room temperature, fully dissolving to form an orange solution. The solution was then stirred overnight forming a cloudy yellow solution. The reaction was monitored *via* TLC (8:2, Hex/EtOAc). The solution was then concentrated under reduced pressure and precipitated with ice cold Et_2O. The solution was then filtered and concentrated to yield a crude oil (1.9495 g) and subsequently purified by FCC (8:2, Hex/EtOAc) to yield a cloudy oil, **53** (530 mg, 1.3 mmol, 68%).^{1}H NMR (400 MHz, $CDCl_3$) δ 8.08 – 7.97 (m, 2H), 7.49 – 7.29 (m, 5H), 6.86 (d, *J* = 8.3 Hz, 2H), 5.34 (s, 2H), 4.97 (s, 1H), 4.46 (dt, *J* = 30.7, 7.8 Hz, 1H), 3.85 – 3.79 (m, 1.5H), 3.76 (d, *J* = 4.0 Hz, 3H), 3.67 (d, *J* = 12.0 Hz, 0.5H), 2.60 – 2.47 (m, 1H), 2.31 – 2.19 (m, 1H), 1.43 (d, *J* = 9.6 Hz, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.41, 166.12, 160.90, 136.35, 132.04, 131.97, 128.73, 128.33, 128.26, 123.40, 115.05, 80.76, 77.36, 74.92, 66.64, 58.08, 52.32, 51.99, 36.65, 28.38. ESI-MS (m/z):$[M+H]^+$calcd. for $C_{25}H_{29}NO_7$, 456.2017; found, $[M+Na]^+$, 478.3.

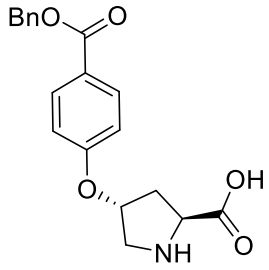
Synthesis of (2S,4R)-4-(4-((benzyloxy)carbonyl)phenoxy)-1-(tert-butoxycarbonyl)pyrrolidine-2-carboxylic acid (54):



54

Compound **53** (490 mg, 1.1 mmol, 1 eq.) was dissolved in THF (10 mL) and deionised water (5 mL). KOH (93.8 mg, 1.67 mmol, 1.5 eq.) was added and the solution stirred overnight at rt. The reaction was monitored by TLC (6:4, Hex/EtOAc). Upon completion the solution was diluted with deionised water (100 mL) and acidified to pH 4, forming a white precipitate. The solution was then extracted with EtOAc (x 3), the combined organic extracts were then washed with brine (x 2), dried over anhydrous $MgSO_4$ and filtered. The filtrate was then concentrated under reduced pressure to yield a white solid, **54** (474.5 mg, 1.1 mmol, 100%).

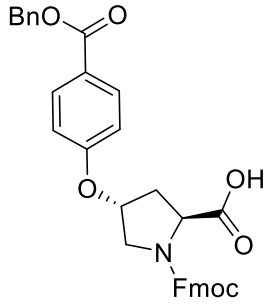
Synthesis of (2S,4R)-4-(4-((benzyloxy)carbonyl)phenoxy)pyrrolidine-2-carboxylic acid (55):



55

Compound **54** (474.5 mg, 1.1 mmol) was dissolved in TFA (5 mL) and DCM (500 µL). This was stirred at rt for 60 min before evaporating, washing with cold Et_2O and drying to yield a white foamy solid, **55**. The crude material was then taken forward for the next step with no further purification.

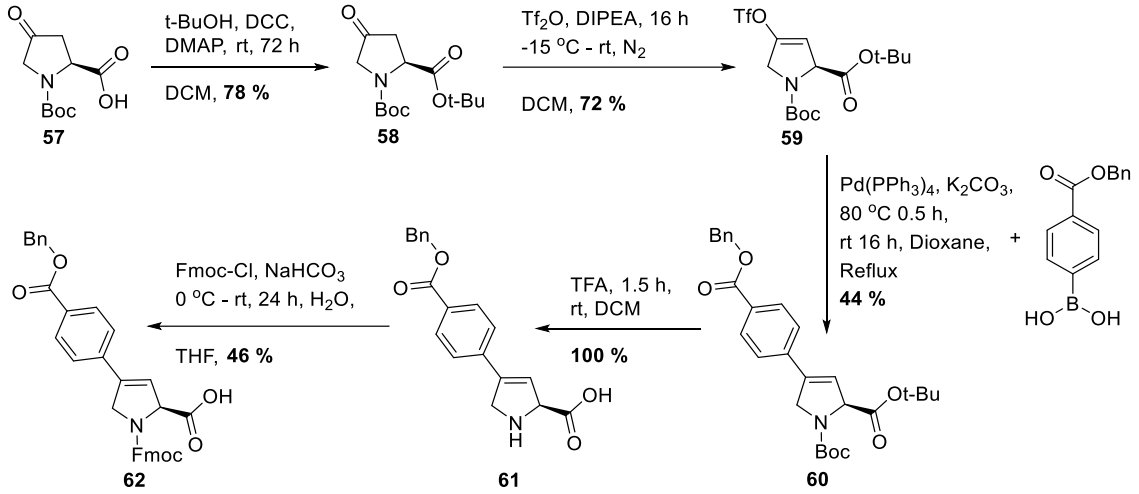
Synthesis of (2S,4R)-1-(((9H-fluoren-9-yl)methoxy)carbonyl)-4-(4-((benzyloxy)carbonyl)phenoxy) pyrrolidine-2-carboxylic acid (56):



56

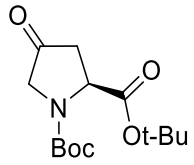
The compound **55** (474.5 mg, 1.1 mmol, 1 eq.) was dissolved in THF (10 mL) and saturated aqueous $NaHCO_3$ (10 mL). The resulting solution was stirred at 0 °C for 15 min before adding Fmoc-Cl (343 mg, 1.33 mmol, 1.2 eq.) in THF (10 mL). The reaction was monitored by TLC (6:4, Hex/EtOAc). Upon completion the solution was diluted with deionised water (100 mL) and washed with ice cold Et_2O (x 2). The solution was then acidified to pH 2 with 3 M HCl, causing a white precipitate to form, and extracted with EtOAc (x 3). The combined organic extracts were then washed with brine (x 2), dried over anhydrous $MgSO_4$ and filtered. The filtrate was then concentrated under reduced pressure to yield a white solid, containing both **56** and the hydrolysed benzyl ester (546.8 mg). The product was purified *via* FCC (8:2, Hex/EtOAc) to yield, **56**, as a white solid (383.8 mg, 0.68 mmol, 62%). ESI-MS (m/z):$[M+H]^+$calcd. for $C_{34}H_{29}NO_7$, 564.2017; found, $[M+H]^+$ 564.7 $[M+Na]^+$, 586.5. $[2M+Na]^+$, 1151.4. **56 minus Bn**, (m/z):$[M+H]^+$calcd. for $C_{27}H_{24}NO_7^+$, 474.1548; found, $[M+H]^+$, 474.5; $[M+Na]^+$, 496.5; $[2M+Na]^+$, 969.8

SI 4.8.5 Synthesis of (S)-1-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-4-(4-((benzyloxy)carbonyl)phenoxy)-2,5-dihydro-1*H*-pyrrole-2-carboxylic acid, (62):



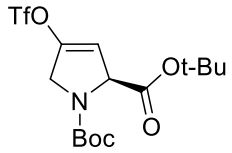
Synthetic route to achieve compounds **62**, for the introduction of a benzoic acid functionality as a peptide sidechain, whereby the benzoic acid can act as a more directing, rigid complexation handle compared to the previously synthesised alkyl carboxylic acid. This method makes use of a Suzuki coupling due to difficulties with success in Mitsunobu reaction attempts to add a benzo ester. Subsequent saponification of synthesised peptides utilising this monomer was required to remove the benzyl protecting group.

Synthesis of di-tert-butyl (S)-4-oxopyrrolidine-1,2-dicarboxylate, (58):



N-Boc-4-oxo-L-proline, 57 (1.499 g, 6.54 mmol, 1 eq.) and DMAP (79 mg, 0.65 mmol, 0.1 eq.) were added to DCM (35 mL). DCC (1.3502 g, 6.54 mmol, 1 eq.) was added to the stirred solution. Subsequently t-BuOH (1.87 mL, 19.70 mmol, 3 eq.) was added, forming a brown solution which was stirred for 72 h at rt. The solution was then filtered, and the filtrate was washed with aqueous HCl (1M, x 2), sat. aqueous $NaHCO_3$ (x 2), deionised water (x 2) and sat. aqueous NaCl (x 2) before drying over anhydrous $MgSO_4$. The organic phase was then concentrated under reduced pressure yield an oil which formed yellowish crystals (1.7221 g). The product was then purified by FCC (1:9, EtOAc/Hex) which yielded a colourless oil which crystallised overnight to form colourless crystals of **58**, (1.45 g, 5.08 mmol, 78%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.65 (d, *J* = 10.3 Hz, 0.42H), 4.57 (dd, *J* = 10.4, 1.6 Hz, 0.58H), 3.88 (d, *J* = 5.8 Hz, 1H), 3.83 (s, 1H), 2.98 – 2.81 (m, 1H), 2.53 (d, *J* = 2.2 Hz, 0.57H), 2.48 (d, *J* = 2.2 Hz, 0.43H), 1.48 – 1.42 (m, 18H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 209.10, 208.28, 170.89, 154.38, 153.70, 82.41, 81.07, 77.26 (CH), 57.06 (CH), 56.60 (CH), 52.99 (CH_2), 52.56 (CH_2), 41.44 (CH_2), 40.92 (CH_2), 28.27 (CH_3), 27.92 (CH_3).

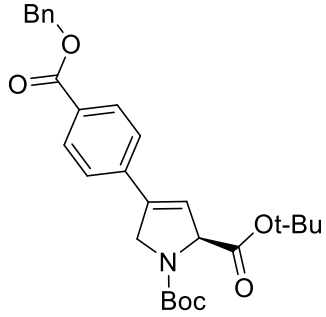
Synthesis of di-tert-butyl (S)-4-(((trifluoromethyl)sulfonyl)oxy)-2,5-dihydro-1H-pyrrole-1,2-dicarboxylate, (59):



59

Compound **58** (1.44 g, 5.05 mmol, 1 eq.) was dissolved in dry DCM (25 mL) under a N_2 atmosphere. DIPEA (4.35 mL, 24.97 mmol, 5 eq.) was added and the solution chilled to -15 °C and stirred. Tf_2O (1.3 mL, 7.74 mmol, 1.5 eq.) was then added dropwise over 10 minutes. The solution was kept at -15 °C for 30 min before stirring at room temperature for 24 h, forming a dark reddish solution. TLC (2:8 EtOAc/Hex) was used to monitor the reaction progress. The solution was then quenched by adding sat. aqueous $NaHCO_3$, the aqueous phase was separated and extracted with DCM. The combined organic phases were then washed with sat. aqueous NaCl, before drying over anhydrous $MgSO_4$ and concentrating under vacuo to yield a dark reddish oil (2.3221 g). The crude product was then isolated by FCC (95:5, Hex/EtOAc) to yield a yellow oil which formed a crystalline solid, **59** (1.52 g, 3.6 mmol, 72%). NMR (400 MHz, $CDCl_3$) δ 5.73 (dd, *J* = 4.2, 2.0 Hz, 0.38H), 5.69 (dd, *J* = 4.3, 1.9 Hz, 0.62H), 4.90 (dt, *J* = 6.1, 2.5 Hz, 0.38H), 4.85 (dt, *J* = 5.9, 2.5 Hz, 0.62H), 4.39 – 4.20 (m, 2H), 1.47 (d, *J* = 5.6 Hz, 9H), 1.44 (d, *J* = 1.0 Hz, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 168.11, 167.97, 153.22, 152.85, 146.17, 145.78, 118.59 (q, *J* = 321.0 Hz),112.27, 112.02, 82.69, 81.28, 81.22, 64.53, 64.27, 50.47, 50.19, 28.41, 28.33, 28.01, 27.98.

Synthesis of di-tert-butyl (S)-4-(4-((benzyloxy)carbonyl)phenyl)-2,5-dihydro-1H-pyrrole-1,2-dicarboxylate, (60):

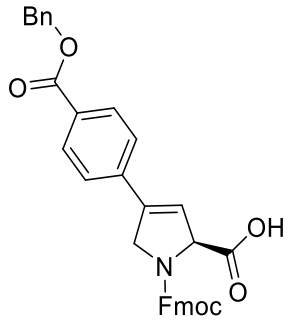


60

Compound **59** (1.45 g, 3.47 mmol, 1 eq.) and 4-benzyloxycarbonylphenylboronic acid (1.3439 g, 5.25 mmol, 1.5 eq.) were dissolved in dioxane (40 mL) with tetrakis(triphenylphosphine)palladium (0) (311 mg, 0.27 mmol, 0.08 eq.). A potassium carbonate solution (7.7 mL, 0.254 mmol, 4.4 eq.) was then added and the yellow solution heated at 80 °C and stirred for 30 min, forming a black solution. The solution was then left stirring overnight at rt, forming a black foamy precipitate. TLC (5:95, EtOAc/Hex) was used to monitor the reaction. The solution was then diluted with a brine solution (150 mL) and extracted with EtOAc (x 3). The combined organic phases were then washed with brine (x 2), dried over anhydrous magnesium sulphate and filtered to afford a yellow solution. This was concentrated under reduced pressure to yield a mixture of brown precipitate and white crystalline solids (2.0894 g). The crude product was purified *via* column chromatography to afford **60** as a white crystalline solid (735.5 mg, 1.54 mmol, 44%). ^1H NMR (400 MHz, $CDCl_3$) δ 8.09 – 8.03 (m, 2H), 7.48 – 7.42 (m, 4H), 7.42 – 7.32 (m, 3H), 6.22 (dd, J = 4.3, 1.9 Hz, 0.32H), 6.18 (dd, J = 4.3, 1.9 Hz, 0.68H), 5.37 (d, J = 2.4 Hz, 2H), 5.09 – 5.04 (m, 0.32H), 5.00 (dt, J = 5.4, 2.6 Hz, 0.68H), 4.69 –

4.47 (m, 2H), 1.59 (s, 1H), 1.52 (s, 3H), 1.50 – 1.44 (m, 15H). (m/z):$[M+H]^+$calcd. for $C_{28}H_{34}NO_6^+$, 480.2381; found, $[M+H]^+$, 479.5

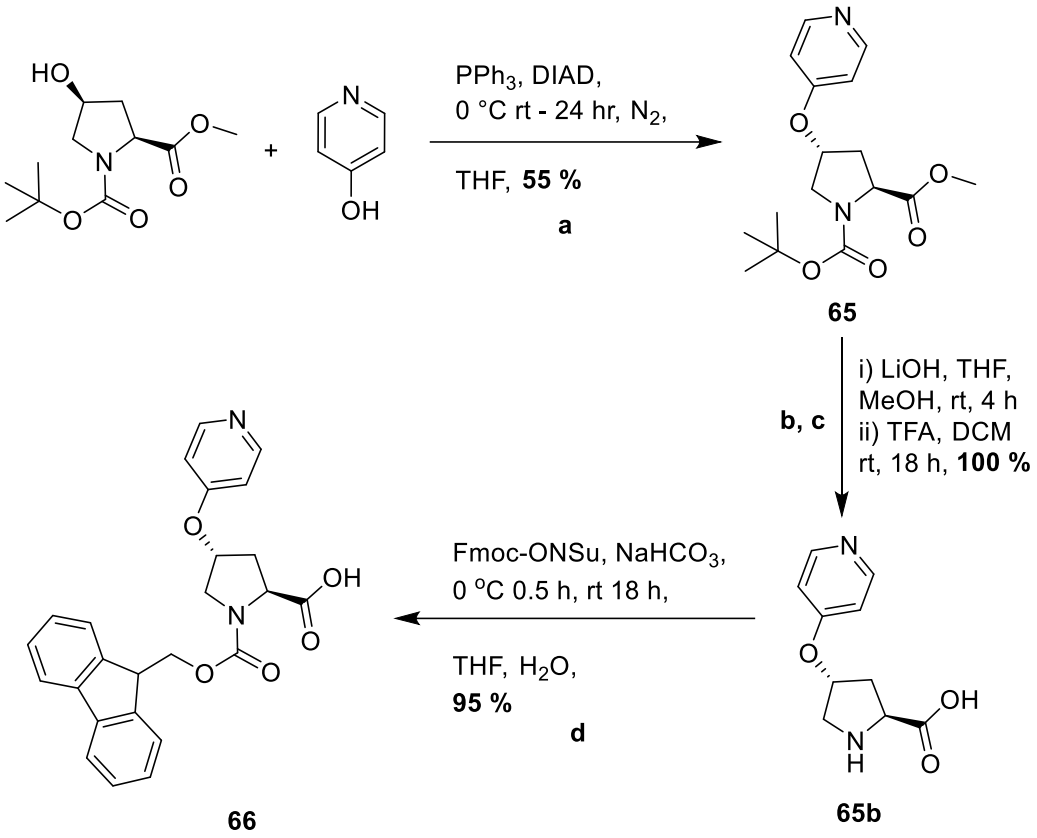
Synthesis of (S)-1-(((9H-fluoren-9-yl)methoxy)carbonyl)-4-(4-((benzyloxy)carbonyl)phenyl)-2,5-dihydro-1H-pyrrole-2-carboxylic acid, (62):



62

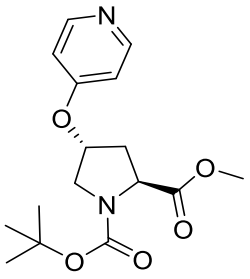
Compound **60** (420 mg, 0.88 mmol, 1 eq.) was dissolved in TFA (5 mL, 95% in DCM) and stirred for 90 min before evaporating. The subsequent oily residue, **61**, was dissolved in THF (2 mL) and sat. aq. $NaHCO_3$ (2 mL) was added at 0 °C. Excess $NaHCO_3$, and Fmoc-Cl (272 mg, 2.19 mmol, 1.2 eq.) in THF (1 mL), was then added, the resulting solution was stirred at 0 °C for 1 h and at rt for a further day. TLC (3:7, EtOAc/Hex) was used to monitor the reaction. The solution was diluted with deionised water (50 mL) and washed with cold Et_2O (x 2). The aqueous solution was then acidified to pH 2 with 3 M HCl causing a white precipitate to form. This was then extracted with EtOAc (x 3), the combined organic extracts were washed with brine (x 2), dried over anhydrous $MgSO_4$, filtered, and concentrated under reduced pressure to yield a white foam solid (530 mg). TLC showed the presence of some Fmoc sideproduct so FCC was carried out (7:3, Hex/EtOAc) to yield the pure compound as a white foam solid after removing solvent under vacuo, **62** (220 mg, 46%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.07 (dd, *J* = 16.3, 8.4 Hz, 5H), 7.75 (dd, *J* = 18.0, 7.0 Hz, 2H), 7.64 – 7.51 (m, 2H), 7.49 – 7.27 (m, 12H), 6.29 (d, *J* = 1.8 Hz, 0.5H), 6.16 (d, *J* = 2.2 Hz, 0.5H), 5.37 (d, *J* = 8.7 Hz, 2H), 5.33 – 5.28 (m, 0.5H), 5.06 (dd, *J* = 6.4, 3.7 Hz, 0.5H), 4.70 – 4.66 (m, 1H), 4.66 – 4.43 (m, 3H), 4.33 (t, *J* = 6.8 Hz, 0.5H), 4.19 (t, *J* = 6.0 Hz, 0.5H), 2.12 (s, 3H), 2.07 (s, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 177.56, 174.74, 174.05, 172.19, 166.09, 155.27, 154.46, 143.78, 143.64, 141.49, 140.47, 140.12, 136.29, 136.16, 135.86, 130.50, 130.38, 128.80, 128.79, 128.56, 128.53, 128.44, 128.40, 128.01, 127.84, 127.32, 127.23, 125.88, 125.84, 125.15, 125.08, 124.89, 124.80, 120.19, 120.13, 119.95, 119.54, 77.48, 77.36, 77.16, 76.84, 68.44, 67.97, 67.29, 67.19, 67.15, 66.68, 60.93, 54.07, 53.57, 47.21, 21.23, 20.85, 14.27.

SI 4.8.6 Synthesis of (2S,4R)-1-(((9H-fluoren-9-yl)methoxy)carbonyl)-4-(pyridin-4-yloxy)pyrrolidine-2-carboxylic acid (66):



Reagents and conditions: (a) PPh_3, DIAD, dry THF, addition at 0 °C, then 24 h. at RT, N_2, (b) LiOH (aq.), THF: MeOH (1:1), at RT, 2.5 h., (c) TFA, DCM, then 24 h. at RT, (d) $NaHCO_3$, THF: H_2O (1:1), addition at 0 °C, then RT, o/n. Reactions a-c carried out by Dr K.Samanta.

Synthesis of 1-(*tert*-butyl) 2-methyl (2R,4S)-4-(pyridin-4-yloxy)pyrrolidine-1,2-dicarboxylate (65):

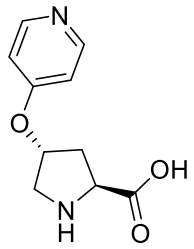


65

To a stirred solution of *N*-Boc-cis-4-hydroxy-L-proline methyl ester (2 g, 8.2 mmol, 1 eq.) in anhydrous THF (30 mL) at room temperature was added 4-hydroxy pyridine (0.9301 g, 9.78 mmol, 1.2 eq.) under an inert atmosphere. PPh_3 (12.2 mmol, 1.5 eq.) was then added followed by drop wise addition of diisopropyl azodicarboxylate (12.2 mmol, 1.5 eq.) over 30 minutes at 0 °C. The reaction was then allowed to warm up to room temperature and stirred at ambient temperature overnight. After completion, the reaction mixture was concentrated and the residue treated with EtOAc (100 mL) and then Hexane (100 mL), to remove excess 4-hydroxy pyridine/lutidine as yellow solid by filtration. The filtrate was concentrated to one fourth of its volume to initiate crystallization of triphenylphosphine oxide by-product and filtered off. The filtrate was then evaporated and purified by column chromatography (EtOAc/n-hexane, 60/40 to EtOAc/Methanol/Et_3N 88/10/2) affording the product, **65**, as a yellow oil (1.4538 g, 4.51

mmol, 55%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.43 (dt, J = 4.9, 3.1 Hz, 2H), 6.79 – 6.73 (m, 2H), 4.9 (bs, 1H), 4.45 (dt, J = 31.7, 7.8 Hz, 1H), 3.81 (d, J = 4.1 Hz, 2H), 3.75 (d, J = 2.2 Hz, 3H), 2.61 – 2.47 (m, 1H), 2.28 (ddd, J = 13.5, 8.2, 5.0 Hz, 1H), 1.43 (d, J = 11.0 Hz, 9H).

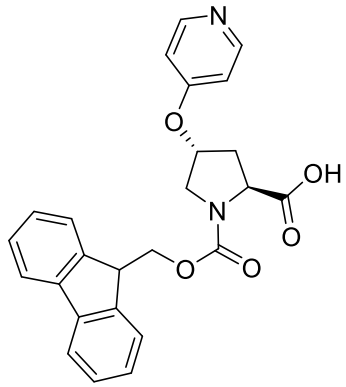
Synthesis of (*2S,4R*)-4-(pyridin-4-yloxy)pyrrolidine-2-carboxylic acid (65b):



65b

Compound **65** (0.75 g, 1 eq, 2.14 mmol) was dissolved in a 2:1 vol/vol mixture THF and MeOH (total 6 mL) with stirring followed by addition of LiOH (3.2 mmol, 1.5 eq.) in water (2 mL) and left to stir for 2-4 h. TLC was used to monitor the reaction (hexane/EtOAc, 25/75,) and once complete, organic layer was removed under reduced pressure. The crude product was dissolved in TFA (32 mmol, 15 eq.) and DCM (10 mL) before being left to stir at ambient temperature overnight. After completion of the reaction, the combined organic layer was removed under reduced pressure and co-evaporated with methanol (5 x volume of TFA) thrice. The resultant yellowish oil was dissolved in 1M HCl and stirred for an hour at 50 °C. The insoluble residue was removed by dissolving and extracting with EtOAc (3 x 25 mL). The product was used without further purification. ^{1}H NMR (400 MHz, CD_3OD) δ 8.76 – 8.68 (m, 2H), 7.69 – 7.55 (m, 2H), 5.67 (t, J = 4.4 Hz, 1H), 4.74 (dd, J = 10.3, 7.8 Hz, 1H), 3.95 – 3.67 (m, 2H), 2.83 (ddt, J = 14.9, 7.8, 1.6 Hz, 1H), 2.66 (ddd, J = 14.9, 10.3, 4.8 Hz, 1H).

Synthesis of (*2S,4R*)-1-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-4-(pyridin-4-yloxy)pyrrolidine-2-carboxylic acid (66):



66

Compound **65b** (0.78 g, 3.3 mmol) was dissolved in H_2O (10 mL) with $NaHCO_3$ (4 mmol, 1.2 eq.). The solution was stirred for 15 min at 0 °C. Fmoc-ONSu (4 mmol, 1.2 eq.) in THF (10 mL) was then added and the solution stirred at 0 °C for 30 min. The solution was then stirred overnight at rt, the THF was then evaporated, and reaction quenched with methanol (10 mL) and stirred for an hour. The methanol was evaporated and 2.5% $NaHCO_3$ (15 mL) was added to the solution. The solution was extracted with cold Et_2O (4 x 25 mL), the combined Et_2O was then extracted once with 2.5% $NaHCO_3$ and H_2O (25 mL). Upon neutralization of the combined aqueous extracts with 1 M HCl, a white precipitate formed. The aqueous solution was then extracted with EtOAc (6 x 50 mL). The combined organic extracts were then washed with brine (2 x 30 mL), dried over anhydrous $MgSO_4$ and evaporated to obtain crude product. The product, **66**, was further re-

crystallized from hot MeOH to obtain a white crystalline solid (1.0654 g, 2.475 mmol, 75%). ^{1}H NMR (400 MHz, CD_3OD) δ 6.93 – 6.85 (m, 4H), 6.24 (d, J = 7.5 Hz, 2H), 6.20 – 6.06 (m, 4H), 5.98 (m, 2H), 5.88 – 5.65 (m, 7H), 5.64 – 5.55 (m, 3H), 5.54 – 5.48 (m, 2H), 3.36 (s, 2H), 3.00 – 2.90 (m, 2H), 2.90 – 2.76 (m, 3H), 2.67 (td, J = 6.7, 3.0 Hz, 2H), 2.22 – 2.07 (m, 2H), 1.76 (p, J = 1.6 Hz, 18H). ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.42 (m, 2H), 7.92 – 7.78 (m, 2H), 7.67 (t, J = 7.9 Hz, 1H), 7.57 – 7.49 (m, 1H), 7.45 – 7.27 (m, 3H), 7.20 (dtd, J = 23.6, 7.5, 1.1 Hz, 1H), 6.99 (dt, J = 4.7, 1.5 Hz, 2H), 5.16 (dt, J = 5.1, 2.4 Hz, 1H), 4.46 (t, J = 8.0 Hz, 1H), 4.37 – 4.10 (m, 4H), 3.76 – 3.58 (m, 4H), 3.16 (s, 1H), 2.63 – 2.53 (m, 1H), 2.46 (t, J = 2.0 Hz, 1H), 2.41 – 2.29 (m, 1H), 2.24 (ddd, J = 13.7, 8.4, 4.9 Hz, 1H). In MeOH, **66** exists as a 1:1 mixture of monomer and H-bonded dimer, whereas in DMSO it only exists as a monomer due to disruption of H-bonding.

SI 4.9 Peptide Synthesis - 2

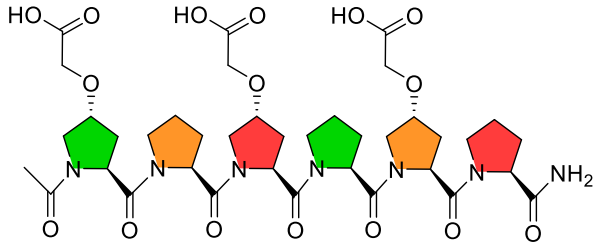
Materials *– Rink Amide MBHA resin (100-200 mesh, 0.3 mmol/g) 1% DVB, Fmoc-amino acids, N,N-Diisopropylethylamine (DIPEA), acetic anhydride (Acac), pivalic anhydride, Trifluoroacetic acid (TFA), Diisopropylcarbodiimide (DIC), Ethyl cyano(hydroxyimino)acetate (Oxyma Pure), and benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) were obtained from Fluorochem Ltd (Derbyshire, UK).*

Peptide Synthesiser Method – All peptides were synthesised on a 0.1 mmol scale on Rink amide MBHA resin using a Liberty^{TM} microwave peptide synthesiser (CEM) utilising Fmoc solid-phase peptide synthesis techniques and repeated steps of single deprotections, and couplings interspaced with washings (4 x 4 mL DMF). The synthesis was paused after the final coupling step and the resin removed from the reaction vessel before stopping the synthesis (to prevent gradual loss of the Fmoc group). Deprotection: 20% piperidine in DMF (4.5 mL) for 5 min with 30 W microwave irradiation at 90 °C. Coupling: Fmoc-amino acid (1.5 mL, 0.2 M, 3 eq.), DIC (1.2 mL, 0.5 M, 6 eq.), Oxyma Pure (0.6 mL, 0.5 M, 3 eq.) in DMF, and DMF (3 mL) for 5 min at 90 °C with 30 W microwave irradiation. **Capping procedure:** Resin washed with DMF (x 5) and suspended in DMF with Acac (50 eq.) and DIPEA (50 eq.). Resin agitated for 30 min before filtering and washing the resin multiple times with DMF.

Peptide Cleavage - The resin was then washed with DCM (x 5), before the Fmoc-protected peptide was cleaved from the resin with TFA (95% in DCM) for 1.5 h. The resin was then washed with the cleavage cocktail (x 2) and the filtrate was concentrated by evaporation before precipitation in cold Et_2O and centrifugation. The solution was then decanted and the solid repeatedly washed with cold Et_2O to isolate the peptide, as a white solid in a quantitative yield after drying under vacuum. Peptides were then purified *via* RP-HPLC, > 99% purity by analytical reverse-phase HPLC.

HPLC: Semi Preparative HPLC was performed on a 1260 Infinity II (Agilent) HPLC, equipped with a C18 (Kormasil 100-5-C18, 10 x 250 mm) column at a 4.73 mL/min flow rate monitored at 205/225/254 nm wavelengths.

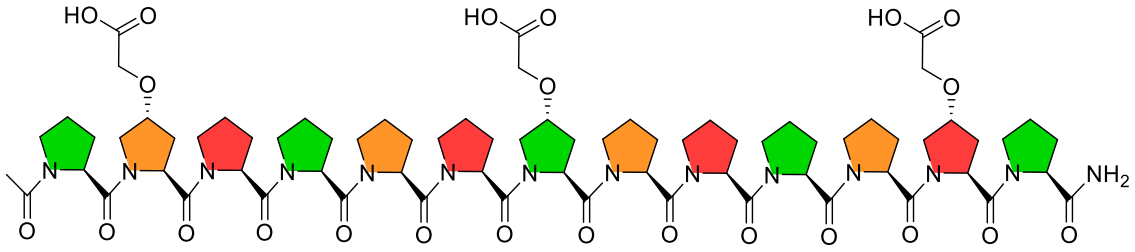
SI 4.9.1 Synthesis of Ac-Pro$_6$-NH$_2$ (-OCH$_2$COOH)$_3$(i: 1, 3, 5) (67):



(67)

Peptide **67** was synthesised using standard SPPS techniques on 0.1 mmol scale using the standard Oxyma-DIC coupling method, quantitative yield. The peptide was purified *via* semi-prep RP-HPLC. **67**,

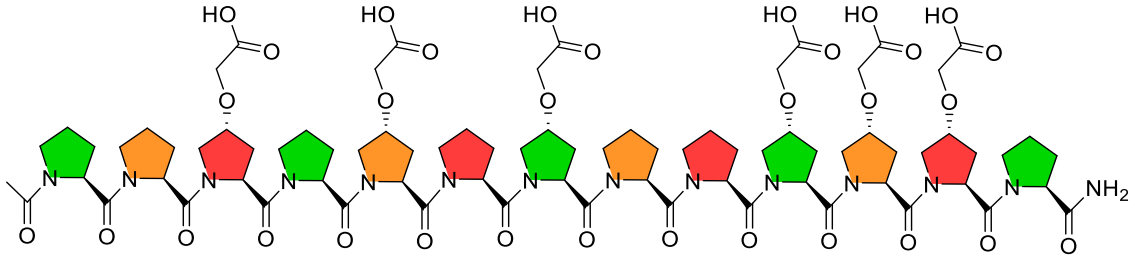
SI 4.9.2 Synthesis of Ac-Pro$_{13}$-NH$_2$ (-OCH$_2$COOH)$_3$(i: 2, 7, 12) (68):



(68)

Peptide **68** was synthesised using standard SPPS techniques on 0.1 mmol scale using the standard Oxyma-DIC coupling method, quantitative yield. **68**, ^{1}H NMR (400 MHz, MeOD) δ 4.75 – 4.63 (m, 5H), 4.62 – 4.44 (m, 4H), 4.44 – 4.33 (m, 3H), 4.32 – 4.07 (m, 8H), 4.06 – 3.86 (m, 3H), 3.86 – 3.44 (m, 26H), 2.85 – 2.55 (m, 1H), 2.55 – 2.37 (m, 5H), 2.37 – 2.14 (m, 10H), 2.14 – 1.77 (m, 31H). m/z calcd for $[M+H]^+$: $C_{73}H_{103}N_{14}O_{23}^+$, 1543.7315; found; $[M+H]^+$ 1543.7315

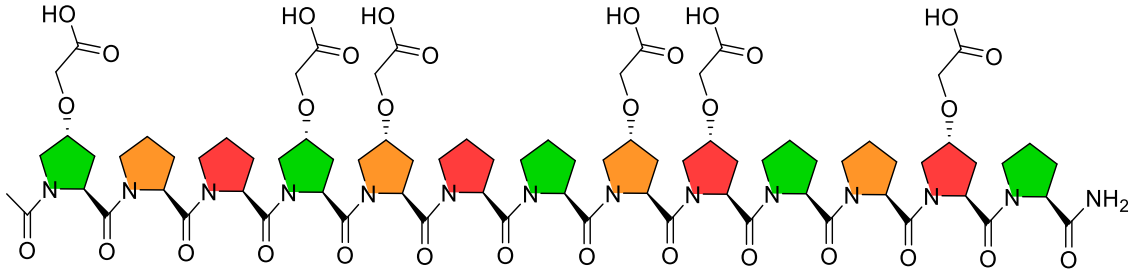
SI 4.9.3 Synthesis of Ac-Pro$_{13}$-NH$_2$ (-OCH$_2$COOH)$_6$(i: 3, 5, 7, 10, 11, 12) (69):



(69)

Peptide **69** was synthesised using standard SPPS techniques on 0.1 mmol scale using the standard Oxyma-DIC coupling method, quantitative yield. **69**, ^{1}H NMR (400 MHz, MeOD) δ 4.83 – 4.60 (m, 7H), 4.45 – 4.32 (m, 4H), 4.29 – 3.94 (m, 12H), 3.88 – 3.55 (m, 14H), 2.68 – 1.86 (m, 27H). m/z calcd for $[M+H]^+$: $C_{79}H_{109}N_{14}O_{32}^+$, 1765.7327; found; $[M+Na]^+$ 1787.3; $[M/2+H]^{2+}$ 883.5

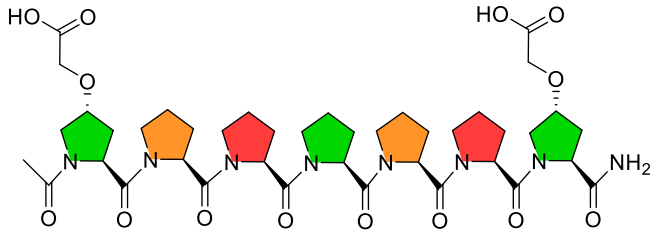
SI 4.9.4 Synthesis of Ac-Pro$_{13}$-NH$_2$ (-OCH$_2$COOH)$_6$(i: 1, 4, 5, 8, 9, 12) (70):



(70)

Peptide **70** was synthesised using standard SPPS techniques on 0.1 mmol scale using the standard Oxyma-DIC coupling method, quantitative yield. ^{1}H NMR (400 MHz, MeOD) δ 4.73 – 4.62 (m, 3H), 4.45 – 4.28 (m, 4H), 4.28 – 3.93 (m, 10H), 3.92 – 3.46 (m, 12H), 2.57 – 2.33 (m, 4H), 2.32 – 2.15 (m, 4H), 2.14 – 1.86 (m, 16H). **70**, m/z calcd for $[M+H]^+$: $C_{79}H_{109}N_{14}O_{32}^+$, 1765.7327; found; $[M+2H]^{2+}$ 883.6

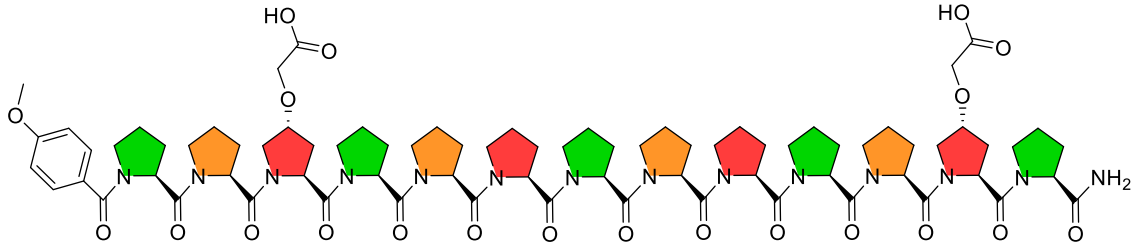
SI 4.9.5 Synthesis of Ac-Pro$_7$-NH$_2$ (-OCH$_2$COOH)$_2$(i: 1, 7) (71):



(71)

Peptide **71** was synthesised using standard SPPS techniques on 0.1 mmol scale using the standard Oxyma-DIC coupling method, quantitative yield. ^{1}H NMR (400 MHz, MeOD) δ 4.76 (dd, *J* = 10.6, 5.0 Hz, 1H), 4.70 (t, *J* = 8.1 Hz, 4H), 4.51 – 4.42 (m, 1H), 4.39 – 4.33 (m, 1H), 4.33 – 4.28 (m, 1H), 4.27 – 4.07 (m, 4H), 4.01 (t, *J* = 12.0 Hz, 1H), 3.92 – 3.68 (m, 8H), 3.68 – 3.55 (m, 6H), 2.52 – 2.32 (m, 2H), 2.32 – 2.16 (m, 4H), 2.15 – 1.89 (m, 16H). **71**, m/z calcd for [M+H]$^+$: $C_{41}H_{59}N_8O_{14}^+$, 887.4145; found; [M+H]$^+$ 887.2; [M+Na]$^+$ 909.1

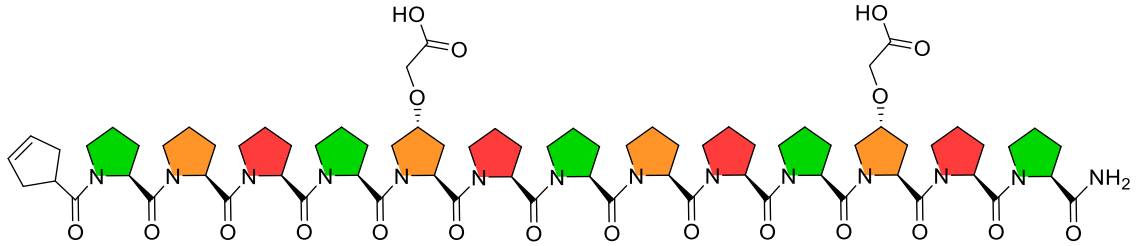
SI 4.9.6 Synthesis of CH$_3$O-C$_6$H$_4$-CO-Pro$_{13}$-NH$_2$ (-OCH$_2$COOH)$_2$(i: 3, 12) (72):



(72)

Peptide **72** was synthesised using standard SPPS techniques on 0.1 mmol scale using the standard Oxyma-DIC coupling method, coupling of 4-methoxybenzoic acid was carried out on the peptide synthesiser, no final deprotection was required, quantitative yield. ^{1}H NMR (400 MHz, MeOD) δ 7.59 (d, *J* = 8.5 Hz, 1.5H), 7.36 (d, *J* = 8.4 Hz, 0.5H), 7.00 (d, *J* = 8.6 Hz, 2H), 4.77 – 4.65 (m, 7H), 4.58 – 4.48 (m, 1H), 4.47 – 4.34 (m, 3H), 4.27 – 4.18 (m, 2H), 4.18 – 4.02 (m, 4H), 3.96 (dd, *J* = 20.9, 11.6 Hz, 2H), 3.90 – 3.77 (m, 11H), 3.77 – 3.68 (m, 5H), 3.68 – 3.47 (m, 9H), 2.55 – 2.34 (m, 4H), 2.34 – 2.17 (m, 10H), 2.17 – 1.83 (m, 33H). **72**, m/z calcd for [M+H]$^+$: $C_{77}H_{105}N_{14}O_{21}^+$, 1561.7573; found; [M+Na]$^+$ 1583.5

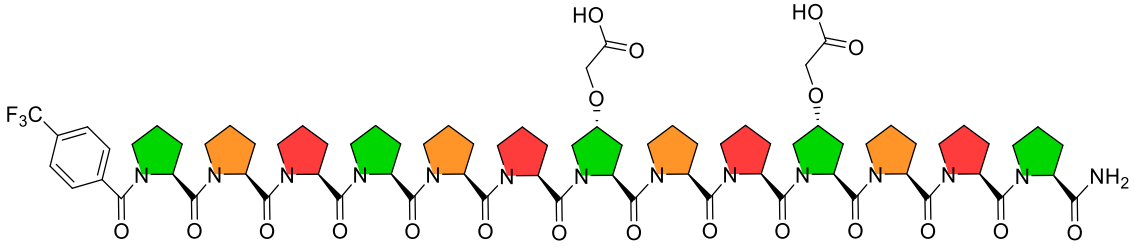
SI 4.9.7 Synthesis of Cyclopent-3-ene-Pro$_{13}$-NH$_2$ (-OCH$_2$COOH)$_2$(i: 5, 11) (73):



(73)

Peptide **73** was synthesised using standard SPPS techniques on 0.1 mmol scale using the standard Oxyma-DIC coupling method, coupling of cyclopentene carboxylic acid was carried out on the peptide synthesiser, no final deprotection was required, significant impurities were observed for the crude peptide as such resynthesis is required to achieve a good yield of the target molecule. The synthesis of this peptide was placed on hold until results from peptide 72 and 74 complexation could be obtained.

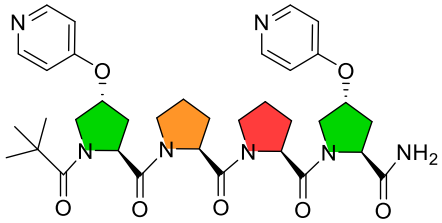
SI 4.9.8 Synthesis of F_3C-C_6H_4-CO-Pro_{13}-NH_2 (-$OCH_2COOH)_2$(i: 7, 10) (74):



(74)

Peptide **74** was synthesised using standard SPPS techniques on 0.1 mmol scale using the standard Oxyma-DIC coupling method, coupling of 4-(triflouromethyl)benzoic acid was carried out on the peptide synthesiser, no final deprotection was required, quantitative yield. ^{1}H NMR (400 MHz, MeOD) δ 7.81 – 7.73 (m, 3.5H), 7.57 (d, *J* = 8.6 Hz, 0.5H), 4.74 – 4.57 (m, 8H), 4.54 – 4.43 (m, 1H), 4.41 (dd, *J* = 8.4, 4.3 Hz, 1H), 4.35 (s, *J* = 1.6 Hz, 2H), 4.21 (d, *J* = 16.8 Hz, 2H), 4.12 (s, 1H), 4.07 (s, 1H), 4.03 (d, *J* = 10.8 Hz, 2H), 3.97 – 3.75 (m, 10H), 3.75 – 3.68 (m, 4H), 3.68 – 3.55 (m, 10H), 3.55 – 3.46 (m, 2H), 2.69 – 2.54 (m, 1H), 2.52 – 2.35 (m, 4H), 2.34 – 2.17 (m, 10H), 2.16 – 1.79 (m, 35H). **74**, m/z calcd for $[M+H]^+$: $C_{35}H_{46}N_7O_7^+$, 676.3453; found; $[M+H]^+$ 676.7; $[2M+H]^+$ 1352.7

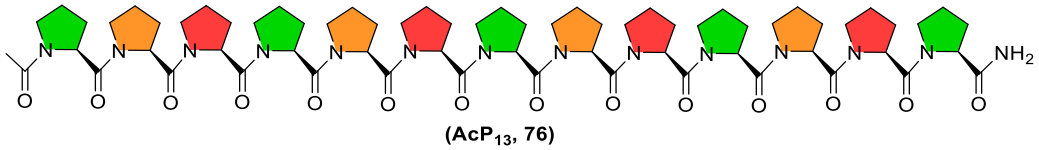
SI 4.9.9 Synthesis of Piv-Pro_4-NH_2 (-$OC_5N)_2$(i: 1, 4) (75):



(75)

Peptide **75** was synthesised using standard SPPS techniques on a 0.1 mmol scale using the standard Oxyma-DIC coupling method, at couplings 1 and 4 the pyridine functionalised proline derivative compound **65b** was used in place of proline, final deprotection was carried out. Subsequent capping was carried out on the bench using the standard capping procedure with pivalic anhydride in place of acetic anhydride, the peptide bound resin was treated with methanol (10 mL) for 1 hour at room temperature followed by washing with MeOH (2 x 5 mL) and DCM (3 x 5 mL) before cleavage. quantitative yield. **75**, ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.41 (dt, *J* = 4.8, 2.4 Hz, 4H), 7.39 (s, 1H), 7.05 – 6.97 (m, 4H), 6.91 (s, 1H), 5.22 (d, *J* = 19.6 Hz, 2H), 4.75 (t, *J* = 8.1 Hz, 1H), 4.61 (dd, *J* = 8.8, 4.2 Hz, 3H), 4.32 (t, *J* = 8.0 Hz, 1H), 4.06 (d, *J* = 11.7 Hz, 1H), 3.98 – 3.80 (m, 2H), 3.76 – 3.57 (m, 3H), 3.46 (q, *J* = 7.8, 7.3 Hz, 2H), 2.41 – 2.27 (m, 3H), 2.13 (ddd, *J* = 14.0, 8.4, 5.2 Hz, 3H), 1.94 (dq, *J* = 44.0, 5.7 Hz, 8H), 1.09 (s, 9H). ^{1}H NMR (400 MHz, D_2O) δ 8.39 (s, 4H), 7.05 (t, *J* = 4.7 Hz, 4H), 5.33 (d, *J* = 15.5 Hz, 2H), 4.90 (t, *J* = 8.5 Hz, 1H), 4.73 – 4.66 (m, 1H), 4.59 (t, *J* = 8.6 Hz, 1H), 4.24 (dd, *J* = 11.8, 6.7 Hz, 2H), 4.06 (dd, *J* = 11.9, 3.5 Hz, 1H), 3.95 (d, *J* = 9.6 Hz, 1H), 3.91 – 3.79 (m, 2H), 3.69 – 3.55 (m, 2H), 3.03 (s, 1H), 2.87 (s, 1H), 2.68 (dd, *J* = 14.2, 7.6 Hz, 1H), 2.55 (dd, *J* = 13.7, 7.9 Hz, 1H), 2.44 – 2.23 (m, 3H), 2.12 – 2.02 (m, 4H), 2.02 – 1.86 (m, 3H), 1.17 (s, 8H). m/z calcd for $[M+H]^+$: $C_{35}H_{46}N_7O_7^+$, 676.3453; found; $[M+H]^+$ 676.3477; $[M+Na]^+$ 698.3629

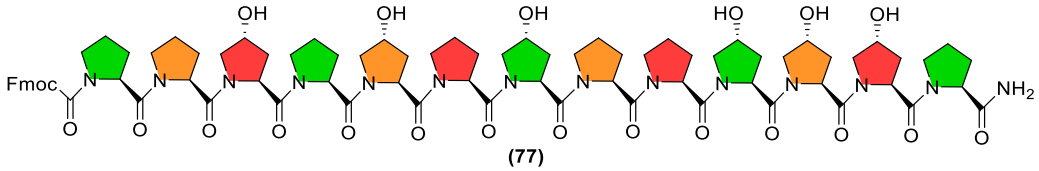
SI 4.9.10 Synthesis of Ac-Pro_{13}-NH_2 (76):



(AcP_{13}, 76)

Peptide **76** was synthesised using standard SPPS techniques on 0.1 mmol scale using the standard Oxyma-DIC coupling method, and deprotection/capping procedures. Quantitative yield. m/z calcd for $[M+H]^+$: $C_{67}H_{97}N_{14}O_{14}^+$, 1321.7303; found; $[M+H]^+$ 1322.8; $[M+Na]^+$ 1345.6; $[M+2H]^{2+}$ 662.3

SI 4.9.11 Synthesis of Fmoc-Pro$_{13}$-NH$_2$ (-OH)$_6$(i: 3, 5, 7, 10, 11, 12) (77):



(77)

Peptide **77** was synthesised using standard SPPS techniques on 0.1 mmol scale using the standard Oxyma-DIC coupling method, and deprotection/capping procedures, the resin-bound peptide was agitated in a MeOH solution before washing with MeOH and DCM before TFA cleavage. m/z calcd for $[M+H]^+$: $C_{65}H_{84}N_{11}O_{15}^+$, 1258.6143; found; $[M+Na]^+$ 1280.7; $[M+H+Na]^{2+}$ 640.9

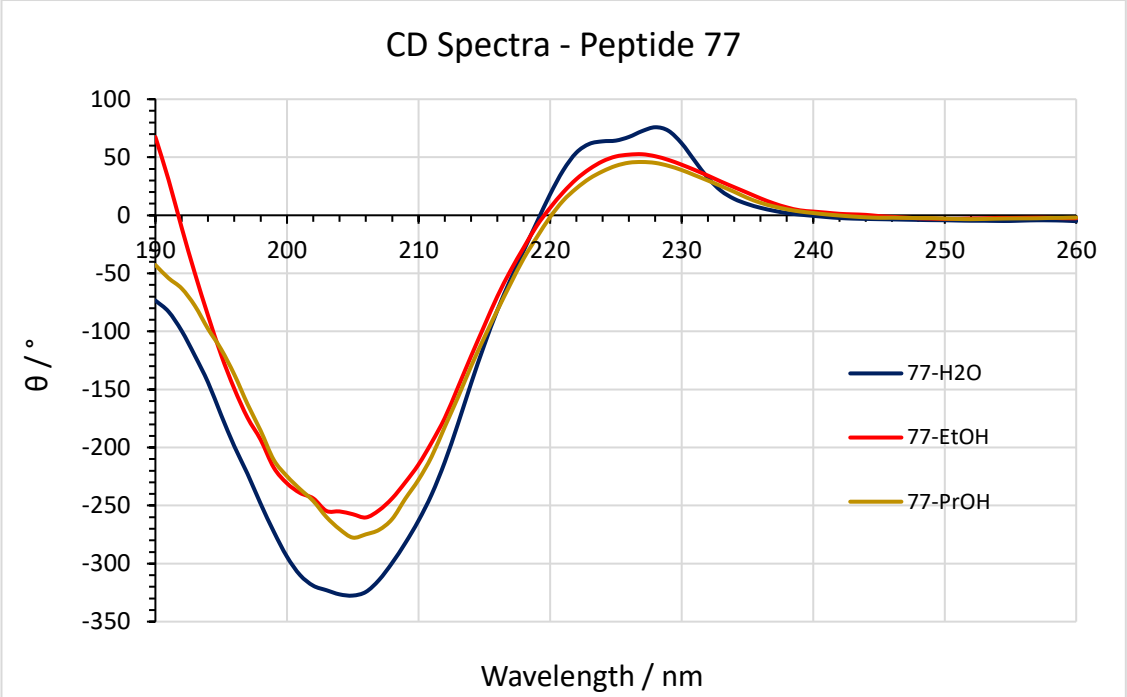
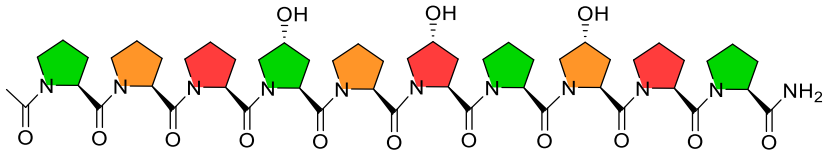


*Figure 119 – CD spectra of peptide **77** in water, EtOH and PrOH (250 µM) incubated for 14 days*

SI 4.9.12 Synthesis of Ac-Pro$_{10}$-NH$_2$ (-OH)$_3$(i: 4, 6, 8) (78):



(78)

Peptide **78** was synthesised using standard SPPS techniques on 0.1 mmol scale using the standard Oxyma-DIC coupling method, and deprotection/capping procedures, the resin-bound peptide was agitated in a MeOH solution before washing with MeOH and DCM before TFA cleavage. **78**, m/z calcd for $[M+H]^+$: $C_{52}H_{76}N_{11}O_{14}^+$,1078.5568; found; $[M+H]^+$ 1078.5, $[M+2H]^{2+}$ 539.3

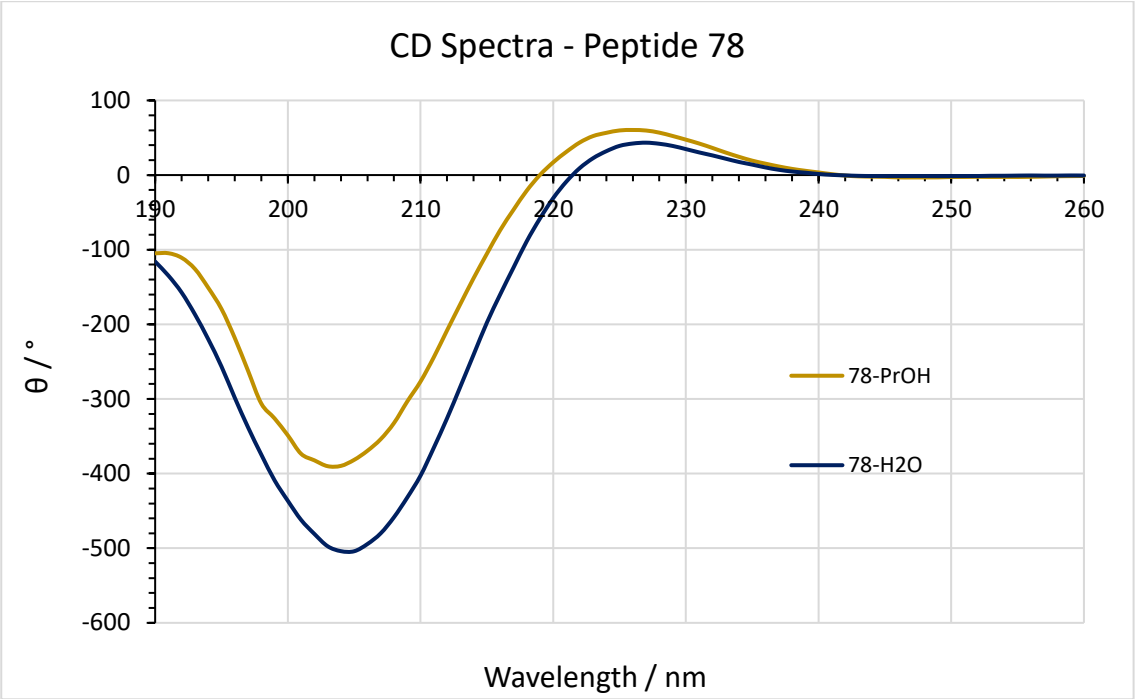


*Figure 120 – CD spectra of peptide **78** in water, EtOH and PrOH (250 µM) incubated for 14 days*

SI 4.10 Complexation reactions:

Materials: *$Cu(NO_3)_2.3H_2O$ was procured from Sigma-Aldrich. $Zn(NO_3)_2.6H_2O$ was procured from Acros Organics. Strong basic Ion-exchange resin, Purolite® A300 has been received as a generous gift from Purolite Ltd. $Pd(en)(NO_3)_2$ has been synthesized from commercially available K_2PdCl_4 by following literature procedure.*[1] Triethylamine (TEA) anhydrous, *K_2PdCl_4 and $Pd(CH_3CN)_4(BF_4)_2$ were procured from Fluorochem Ltd (Derbyshire, UK).* Ref[1]: J. Am. Chem. Soc. 2009, 131, 12, 4505–4512

Dynamic Light Scattering (DLS) – All solvents used for DLS samples were filtered through 0.2 µm syringe filters. Data for samples 67 and 68 were collected on an Anton Paar Litesizer. Data was made up from an accumulation of 3 runs with a minimum of 10 scans per run. Data for all other samples were collected on a Malvern Pananalytical Zetasizer NanoZS. Data was made up from an accumulation of 3 runs with a minimum of 10 scans per run.

AFM Analysis: Samples were diluted as specified for each sample and 10 µL droplets were deposited on freshly cleaved mica discs (Agar Scientific F7013). For samples with zinc or base the mica was first coated with Ni. A solution of $NiCl_2$ (2 mM) in water (10 µL) was dropcast onto the mica and incubated for 10 min, excess solution was removed by washing with 1 mL of 0.2-µm syringe-filtered mQ H_2O, and the mica were then dried under a gentle stream of $N_{2\,(g)}$. For aqueous samples after 10-min incubation at room temperature, excess sample was removed by washing with 1 mL of 0.2-µm syringe-filtered mQ H_2O, and the specimens were then dried under a gentle stream of $N_{2\,(g)}$. For volatile solvents, after 1-min incubation at room temperature excess sample was removed and the mica discs washed with 10 µL (x 3) of 0.2-µm syringe-filtered solvent, and the specimens were then dried under a gentle stream of $N_{2\,(g)}$.Samples were imaged using a Bruker Multimode AFM with a Nanoscope V controller and a ScanAsyst probe (Silicone nitride tip with nominal tip radius = 2 nm, nominal spring constant 0.4 N/m, and nominal resonant frequency 70 kHz). Images were captured at a resolution of 4.88 nm per pixel scanned. All images were processed using the Nanoscope analysis software (version 1.5, Bruker).

TEM Analysis: Samples were prepared by dilution to 0.5 mg/mL and 5 µL droplets were deposited onto 3 mm 400 mesh copper microscope grids covered with holey carbon film and allowed to settle on the grid for 5 min. Samples were then air dried and negative stained in 2% aqueous uranyl acetate unless otherwise stated. Samples were viewed using a Jeol 1230 *trans*mission electron microscope at 80 kV and images were recorded on a Gatan OneView 16-megapixel digital camera.

SEM Analysis: Samples were diluted to 0.5 mg/mL and 5 µL droplets were deposited on an aluminium plate mounted on a carbon tab on an aluminium stub. Samples were then dried. Imaging was achieved *via* an Hitachi S-3400N scanning electron microscope and analysed with Oxford Instruments Aztec analysis software.

SI 4.10.1 Synthesis of Peptide-Metal Complexes based on carboxylic-acid functionalised peptides:

SI 4.10.1.1 Peptide 67 and 68 Complexes (-COOH)$_3$:

General procedures for synthesizing Zn-Peptide (COOH)$_3$ complexes:

1) 1 mg (1 eq.) of peptide (**67-68**) was dissolved in EtOH/H_2O (950 µL) in an epindorph vial (samples in EtOH were incubated for 10 d) and TEA (6 eq.) was added. $Zn(NO_3)_2.6H_2O$ (1.5 eq.) in the respective solvent (50 µL) was then added to the solution. The resultant solution was sonicated for 2 min.

2) 1 mg (1 eq.) of peptide (**67-68**) was dissolved in EtOH/H_2O (950 µL) in an epindorph (samples in EtOH were incubated for 10 d) and excess ammonium hydroxide was added, the sample was then evaporated to dryness and the solid redissolved in the respective solvent. $Zn(NO_3)_2.6H_2O$ (1.5 eq.) in the respective solvent (50 µL) was then added to the solution. The resultant solution was sonicated for 2 min.

Synthesis of peptide 67 – copper complexes: Peptide **67** (5 mg, 5.6 µmol, 1 eq.) was dissolved in DMF (800 µL) and $Cu(NO_3)_2.3H_2O$ (1.5 eq, 1.7 mg, 6.6 µmol) was added forming a blue solution. The solution was heated at 85 ^oC for 5 min forming a green solution with precipitate. After 2h the solution turns green with further precipitate. Precipitate forms upon addition of EtOAc or Et_2O. The addition of EtOH does not cause a precipitate to form. Recrystallizations were attempted with DMF/EtOAc, DMF/Et_2O, EtOH/EtOAc, and EtOH/Et_2O, with the formation of only amorphous precipitates. These were attempted on the solution as is, and on precipitated solid washed with EtOAc (x 3) before redissolving in DMF or EtOH. The precipitate was fully soluble in H_2O.

Synthesis of peptide 67-Zn complex in water: Peptide **67** (1 mg) was dissolved in ammonium hydroxide (20 µL, 30%) in H_2O (200 µL).The solution was then dried and the solid redissolved in H_2O (1 mL). Zinc nitrate hexahydrate (1.5 eq.) was then added causing a white precipitate to form.

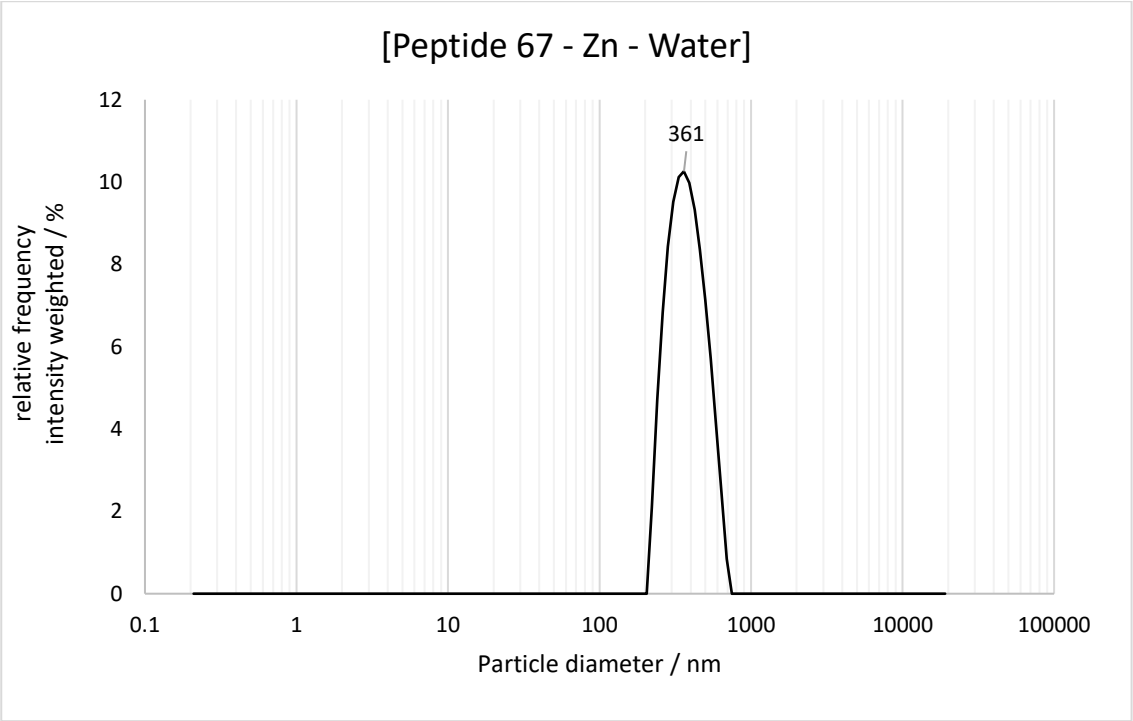


*Figure 121 – DLS intensity distribution for peptide **67** (0.5 mg/mL) dried from ammonium hydroxide and redissolved in water with zinc nitrate (1.5 eq.).*

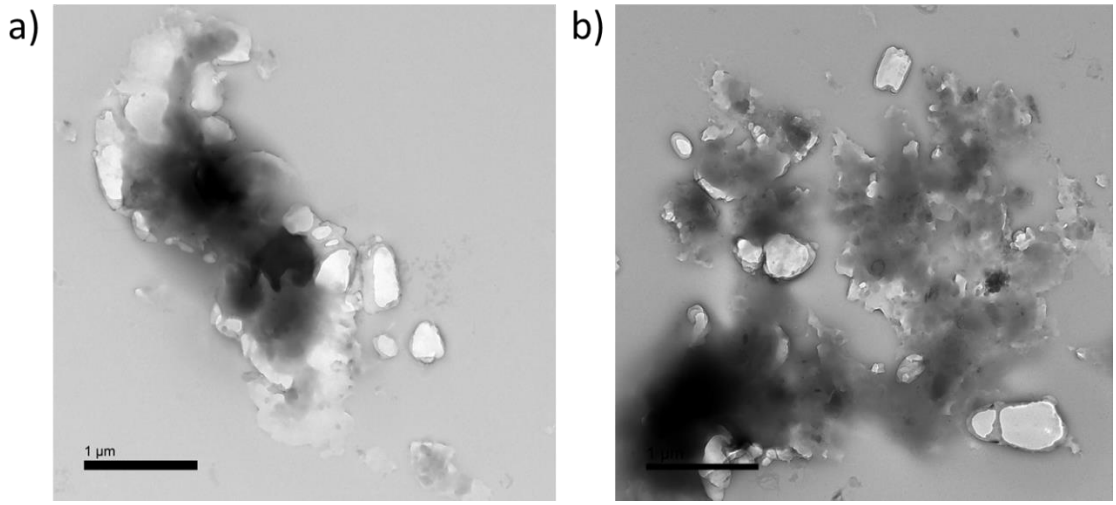


Figure 122 – TEM images of peptide 67 (1 mg) in water (100 µL) reacted with zinc nitrate (2 eq.) and diluted with EtOH (100 µL). Aliquot (10 µL) taken and diluted to 100 µL (H_2O) for drop-casting (0.5 mg/mL). Stained with uranyl acetate.

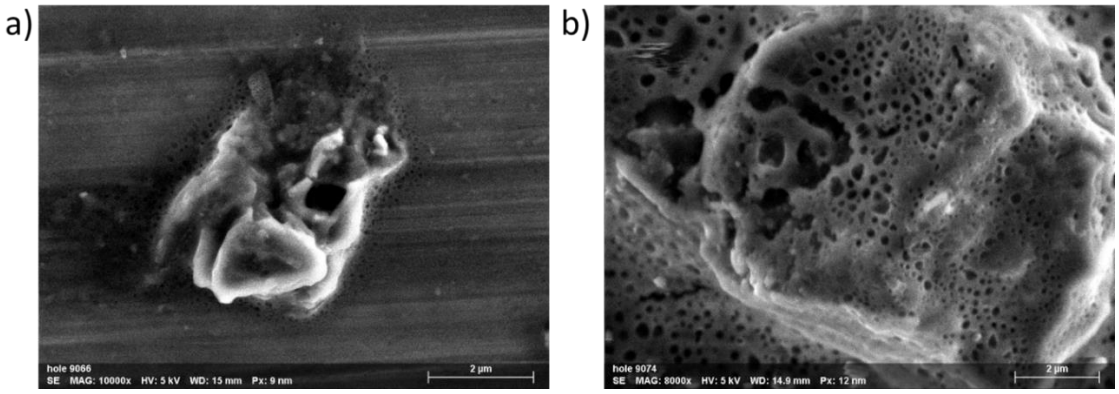


Figure 123 - SEM images of peptide 67 (1 mg) in water (100 µL) reacted with zinc nitrate (2 eq.) and diluted with EtOH (100 µL). Aliquot (10 µL) taken and diluted to 100 µL (H_2O) for drop-casting (0.5 mg/mL). Uncoated sample.

Synthesis of peptide 67-Zn complex in EtOH: Peptide **67** (2 mg) was dissolved in ammonium hydroxide (20 µL, 30%) in EtOH (200 µL).The solution was then dried and the solid redissolved in EtOH (2 mL). To an aliquot of this solution (1 mL, 1 eq.) zinc nitrate hexahydrate (1.5 eq.) was added causing the solution to become turbid. DLS analysis shows the initial presence of large particles (880 nm) that then aggregated over time.

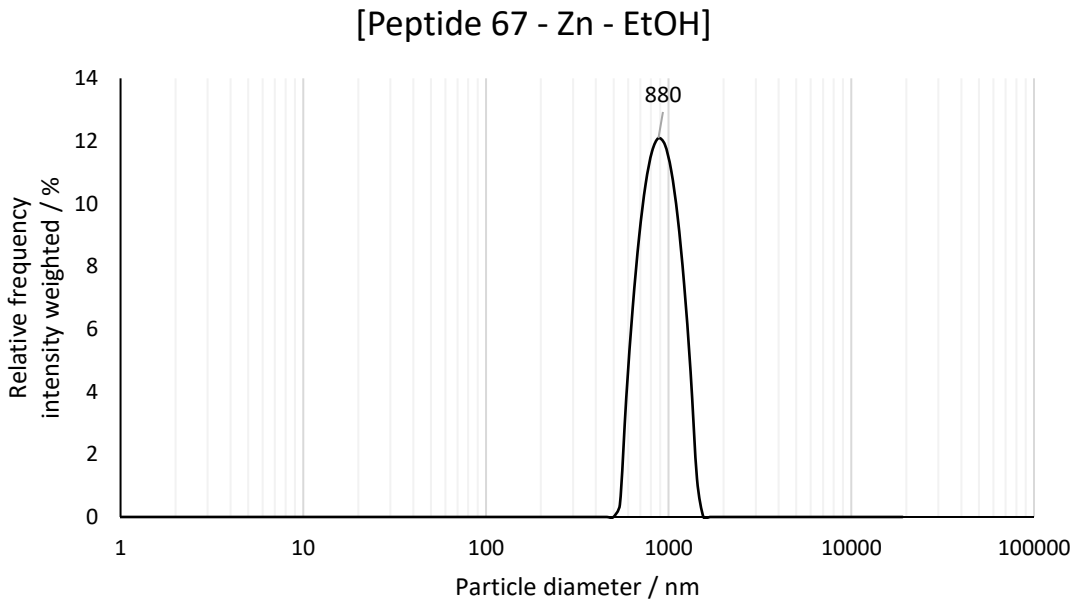


*Figure 124 – DLS intensity distribution for peptide **67** (1 mg/mL) dried from ammonium hydroxide and redissolved in EtOH with zinc nitrate (1.5 eq.).*

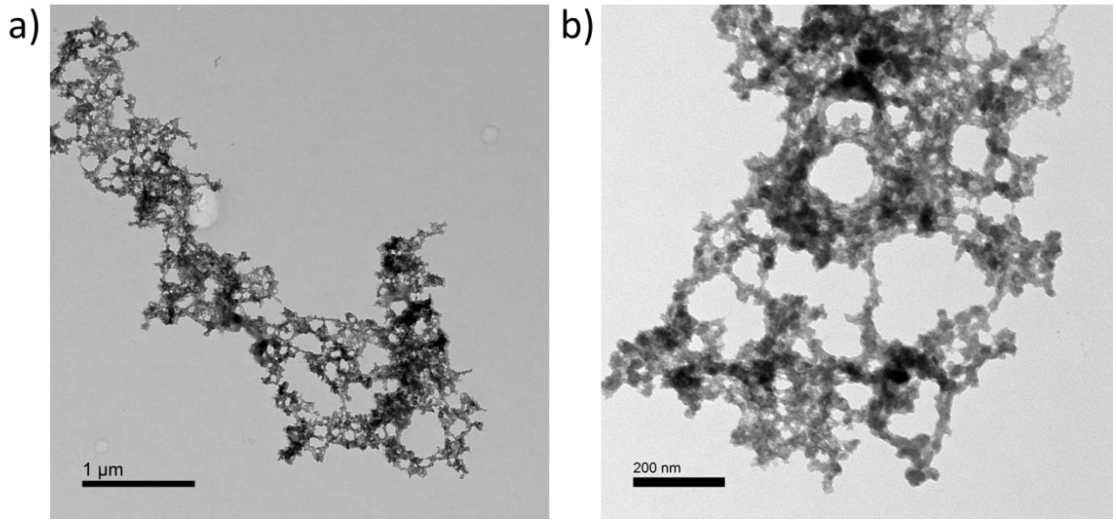


Figure 125 – TEM images of peptide 67 (1 mg) incubated for 10 d in EtOH (100 µL) and reacted with zinc nitrate (2 eq.) and diluted with EtOH (100 µL). Aliquot (10 µL) taken and diluted to 100 µL (EtOH) for drop-casting (0.5 mg/mL). Stained with uranyl acetate.

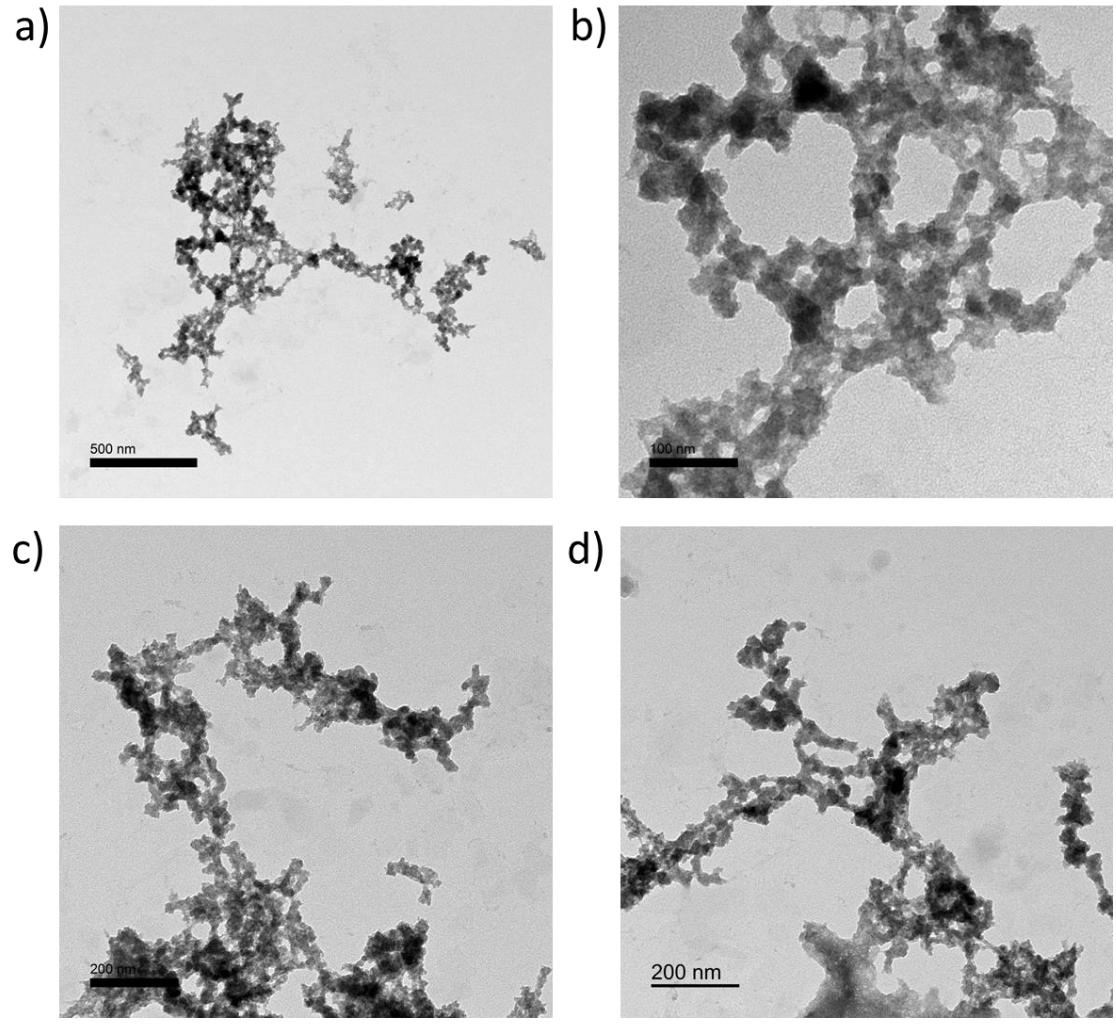


Figure 126 - TEM images of peptide 67 (1 mg) incubated for 10 d in EtOH (100 µL) and reacted with zinc nitrate (2 eq.) and diluted with EtOH (100 µL). Sample was then centrifuged and decanted before redispersing in water (200 µL). Aliquot (10 µL) taken and diluted to 100 µL (EtOH) for drop-casting (0.5 mg/mL). Stained with uranyl acetate.

Synthesis of peptide 68-Zn complex (H_2O): Peptide **68** (1 mg) was dissolved in ammonium hydroxide (20 µL, 30%) in H_2O (200 µL).The solution was then dried and the solid redissolved in

H_2O (1 mL). Zinc nitrate hexahydrate (1.5 eq.) was then added, no precipitate formed. DLS analysis shows particles with good PDI (190 nm).

Synthesis of peptide 68-Zn complex – EtOH: Peptide **68** (1 mg) was dissolved in ammonium hydroxide (20 µL, 30%) in EtOH (200 µL).The solution was then dried and the solid redissolved in EtOH (0.25 mL). Zinc nitrate hexahydrate (1.5 eq.) was then added, the solutions became slightly turbid. DLS analysis shows a good PDI with an average particle size of 280 nm. Diluting the sample up to 1 mL shows large aggregating particles.

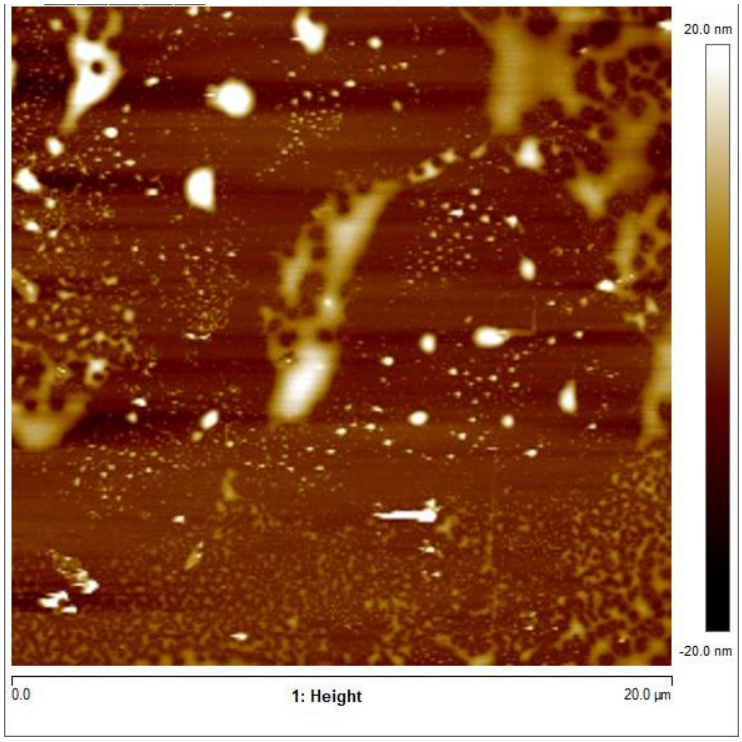


Figure 127 – AFM image of peptide 68 in water (1 mg/mL) with KOH (3 eq.)

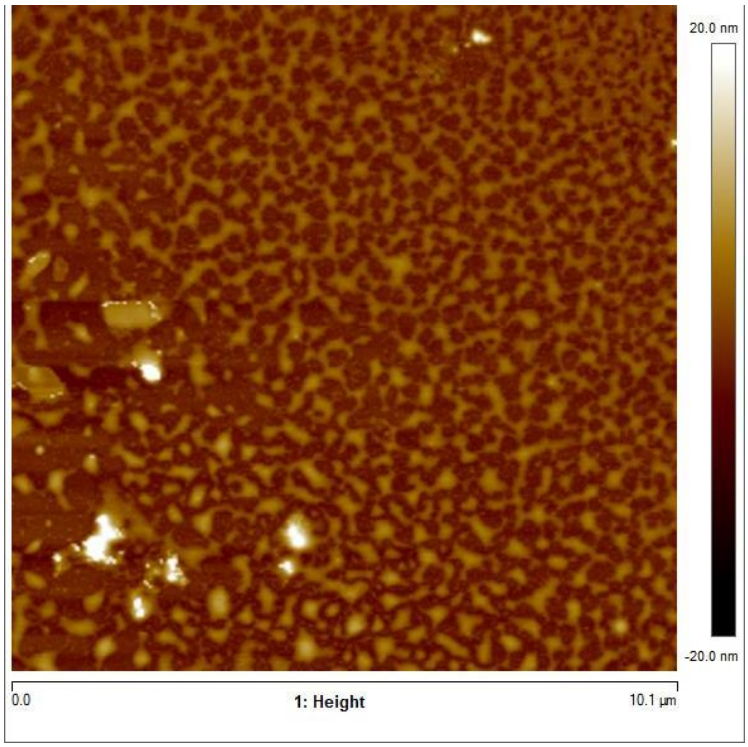
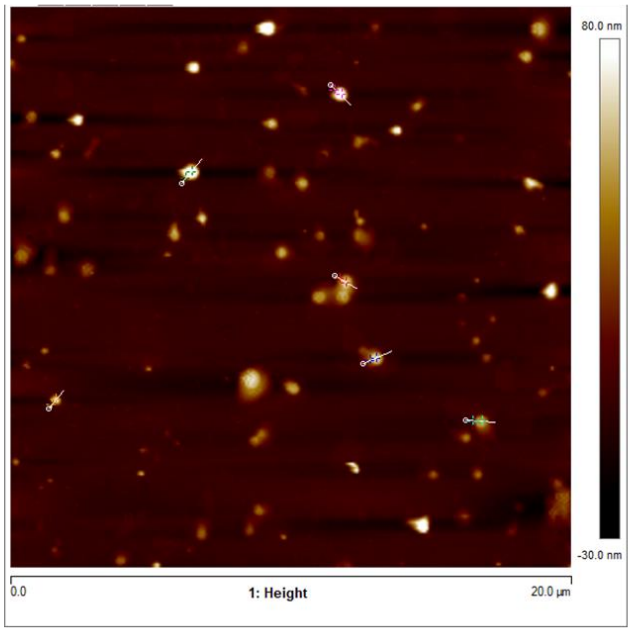


Figure 128 - AFM image of peptide 68 in water (1 mg/mL) with KOH (3 eq.) and zinc nitrate (1.5 eq.)

a)



b)

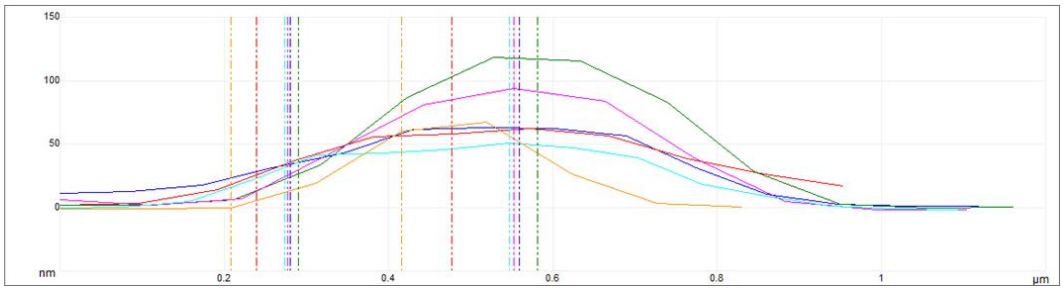


Figure 129 – a) AFM image of peptide 68 in EtOH (0.4 mg/mL) with KOH (3 eq.) 20 x 20 µm, b) table showing diameter and height of a selection of particles from image (a)

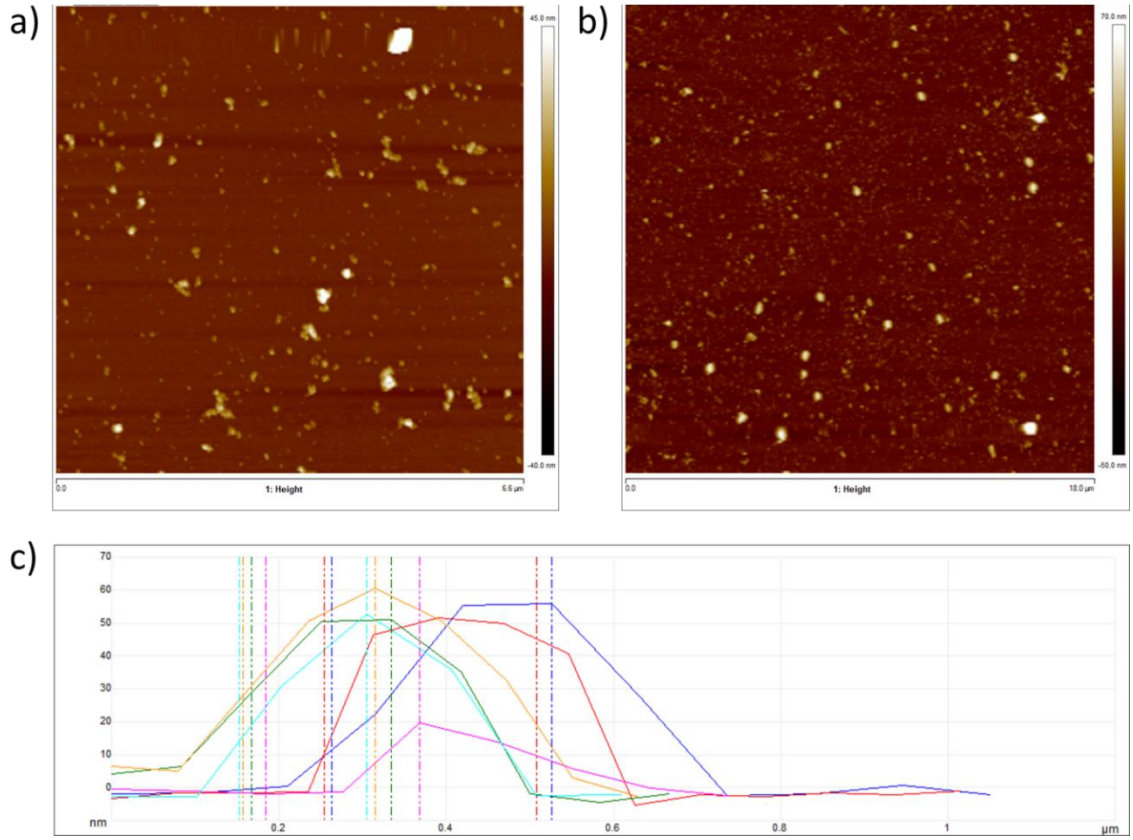


Figure 130 - AFM images (a-b) of peptide 68 in EtOH (0.4 mg/mL) with KOH (3 eq.) and zinc nitrate (1.5 eq.); a) 6.6 x 6.6 µm, b) 18 x 18 µm, c) table showing diameter and height of a selection of particles from image (b)

*SI 4.10.1.2 Peptide **69** and **70** Complexes **(-COOH)$_6$**:*

General procedure for synthesizing Zn-Peptide (COOH)$_6$ complexes: 0.4/1 mg (1 eq.) of peptide (**69-70**) was dissolved in MeOH/EtOH/PrOH/H_2O (950 µL) in an epindorph and TEA (6 eq.) was added. $Cu(NO_3)_2.3H_2O$ (3 eq.) or $Zn(NO_3)_2.6H_2O$ (3 eq.) in the respective solvent (50 µL) was then added to the solution. The resultant solution was sonicated for 2 min.

Synthesis of peptide 69 – copper complexes: Peptide **69** (10 mg, 5.7 µmol, 1 eq.) was dissolved in DMF (800 µL) and $Cu(NO_3)_2.3H_2O$ (3 eq, 4.9 mg, 20.4 µmol) was added forming a blue solution. The solution was heated at 85 °C for 30 min forming a white precipitate. After 2h the solution turns green with further precipitate formed.

At maximum magnification (x 1000) very small particles that appeared partially crystalline were observable however these were unsuitable for SCXRD. The solution was then centrifuged and decanted, the solid was then washed with DMF (x 3) to yield a white solid. The solid was then dissolved in H_2O to yield a colourless solution. The slow addition of EtOH caused a white precipitate to form at the solvent interface, no crystal formation was observed. The complexation was repeated on several smaller batches and diffusion recystallisations were attempted with DMF/EtOAc, DMF/Et_2O, DMF/EtOH with no success.

Control sample: $Cu(NO_3)_2.3H_2O$ (3 eq, 4.9 mg, 20.4 µmol) was dissolved in DMF (800 µL) forming a blue solution. The solution was heated at 85 °C for 5 min forming a green solution, after a further 5 min the solution turns brown-green, gradually turning dark brown after 2 h. No precipitate was formed.

SI 4.10.1.3 AFM and DLS analysis of peptide 69 and complexes:

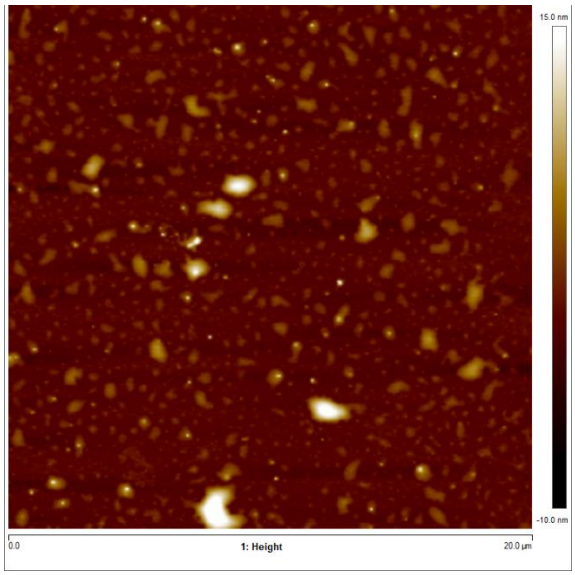


Figure 131 – AFM image of peptide 69 in water (1 mg/mL), 20 x 20 µm

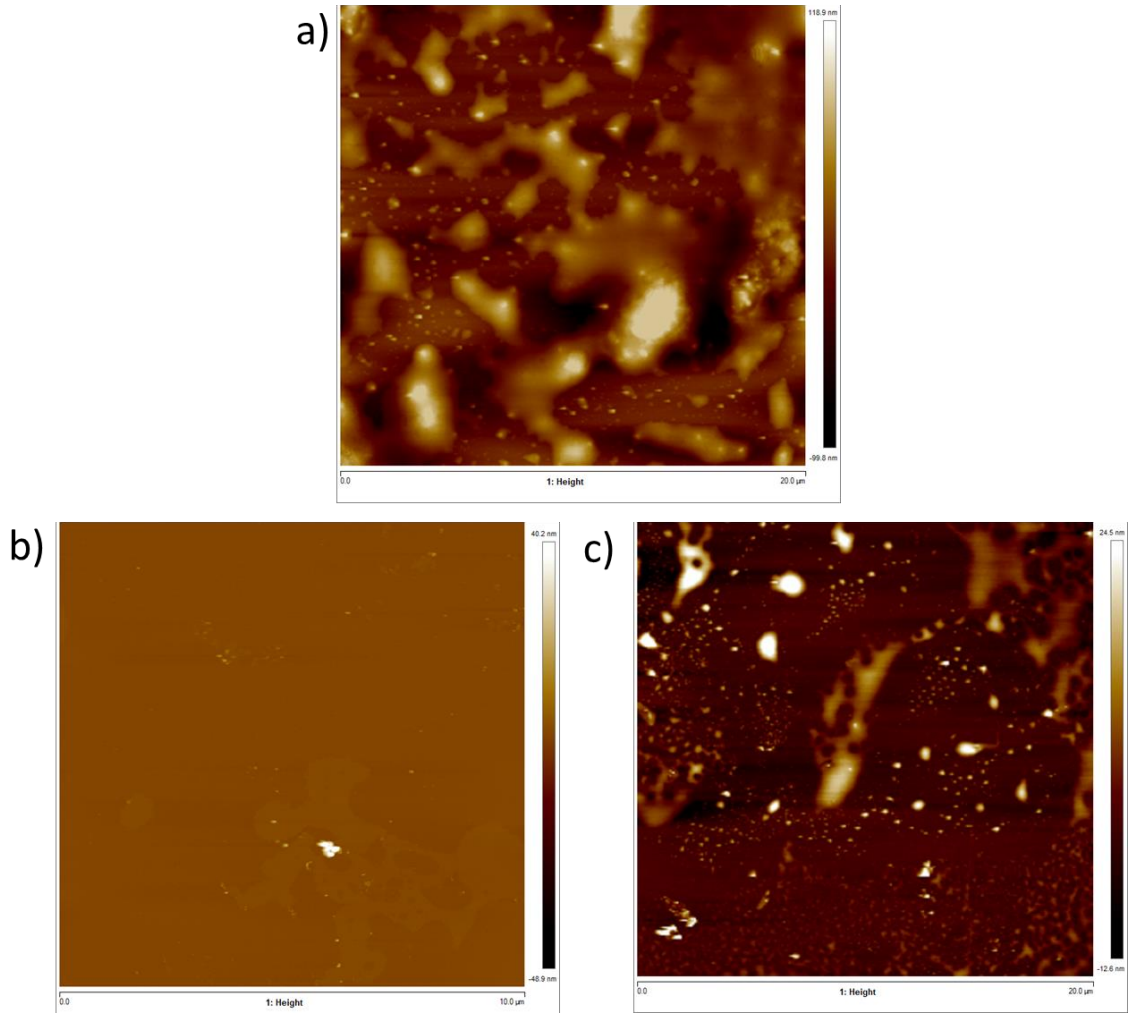


Figure 132 - AFM images of peptide 69 in water with base (1 mg/mL); a) NaOH (6 eq.), 20 x 20 µm, b) TEA (6 eq.), 10 x 10 µm, c) KOH (6 eq.), 20 x 20 µm

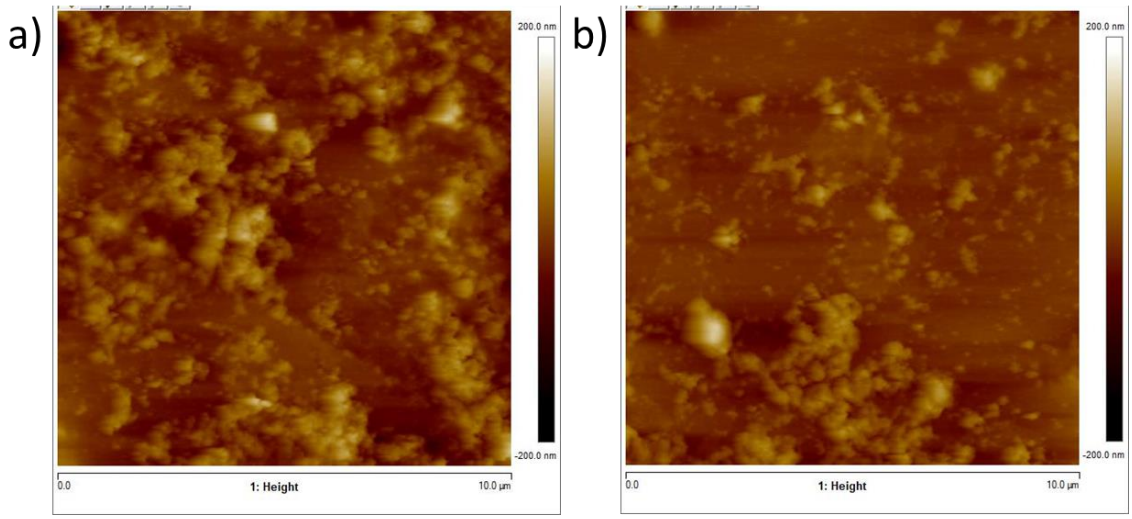


Figure 133 - AFM images of peptide 69 in MeOH (0.4 mg/mL), 10 x 10 µm

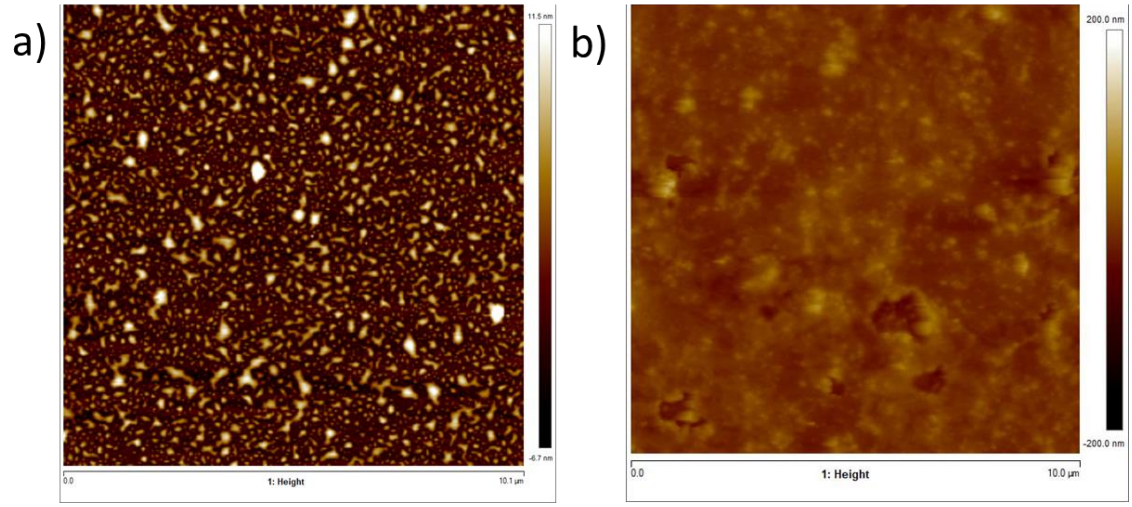


Figure 134 – AFM images of peptide 69 in EtOH (0.4 mg/mL), 10 x 10 µm; a) Ni coated mica, b) uncoated mica

SI 4.10.1.4 AFM and DLS analysis of peptide 70 and complexes

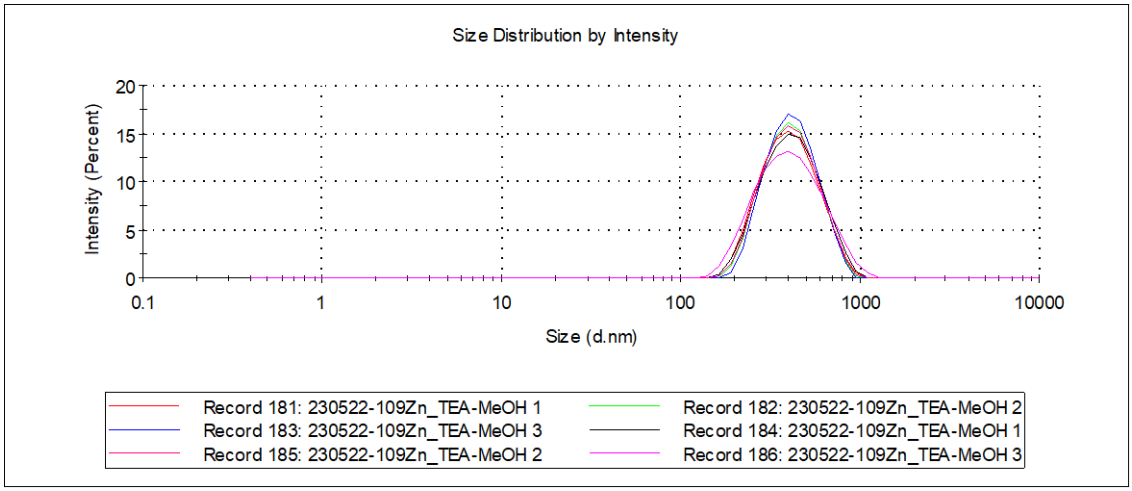


Figure 135 - DLS size distribution of peptide 70 in MeOH (0.4 mg/mL) with TEA (6 eq.); Z-avg: 368.9 nm, PDI: 0.146

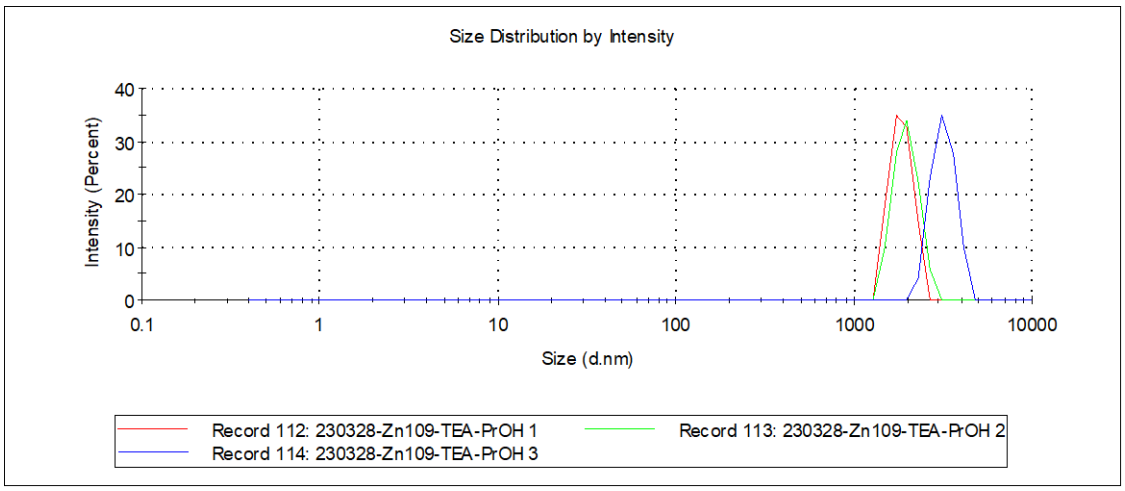


Figure 136- DLS size distribution of peptide 70 in PrOH (0.4 mg/mL) with TEA (6 eq.); Z-avg: 2512 nm, PDI: 0.291. Aggregate over time

SI 4.10.1.5 CD Spectroscopy

CD experiments were carried out on a Jasco J-715 spectropolarimeter. Spectra were recorded using a spectral bandwidth of 190-260 nm, at 20 °C, with a scan rate of 100 nm/min. CD data are given in ellipticity (mdeg). The spectra are formed of 4 accumulations and a spectrum of the solvent blank was subtracted from the raw CD data. A Quartz cell was used with a 1 mm path length using either 125 or 250 µM peptide solutions. All samples were kept in solution for at least 14 days prior to recording CD spectra to ensure the final stable conformation had been achieved due to slow conversion between polyproline helices (i.e. Polyproline II→ Polyproline I). The observed ellipticity has been converted to molar ellipticity (θ) for all spectra, expressed in the units $deg.cm^2.dmol^{-1}$.

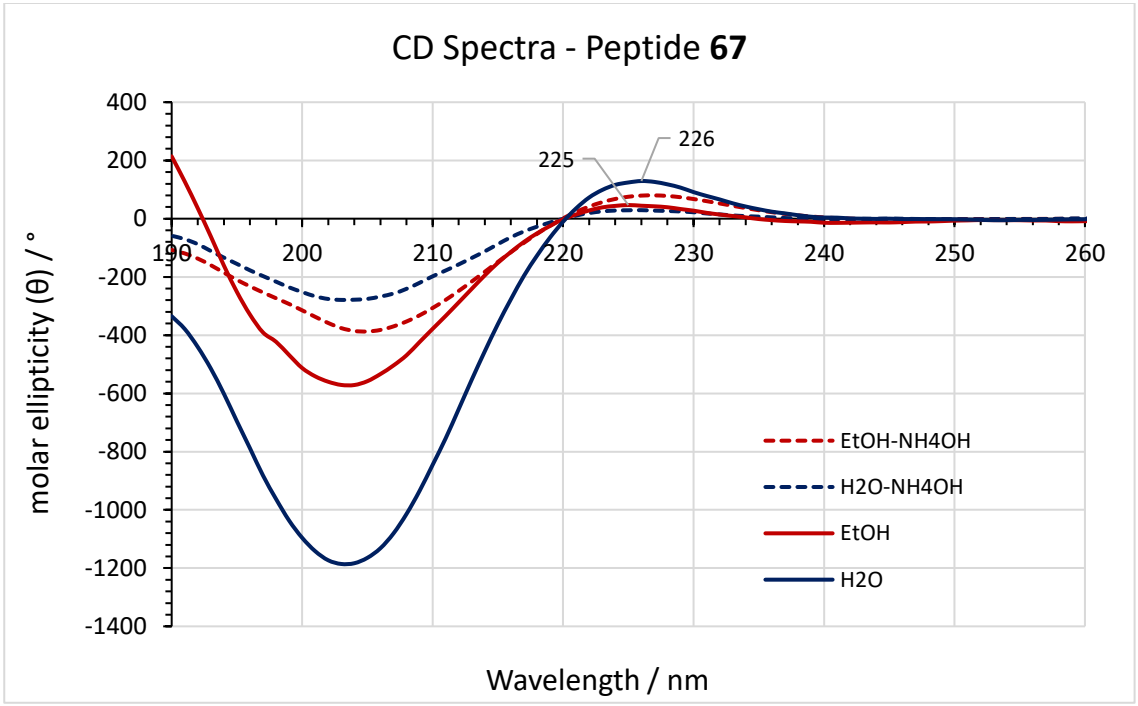


*Figure 137 – CD spectra for peptide **67** incubated in EtOH and water (250 µM), and with ammonium hydroxide (30%, 1 µL)*

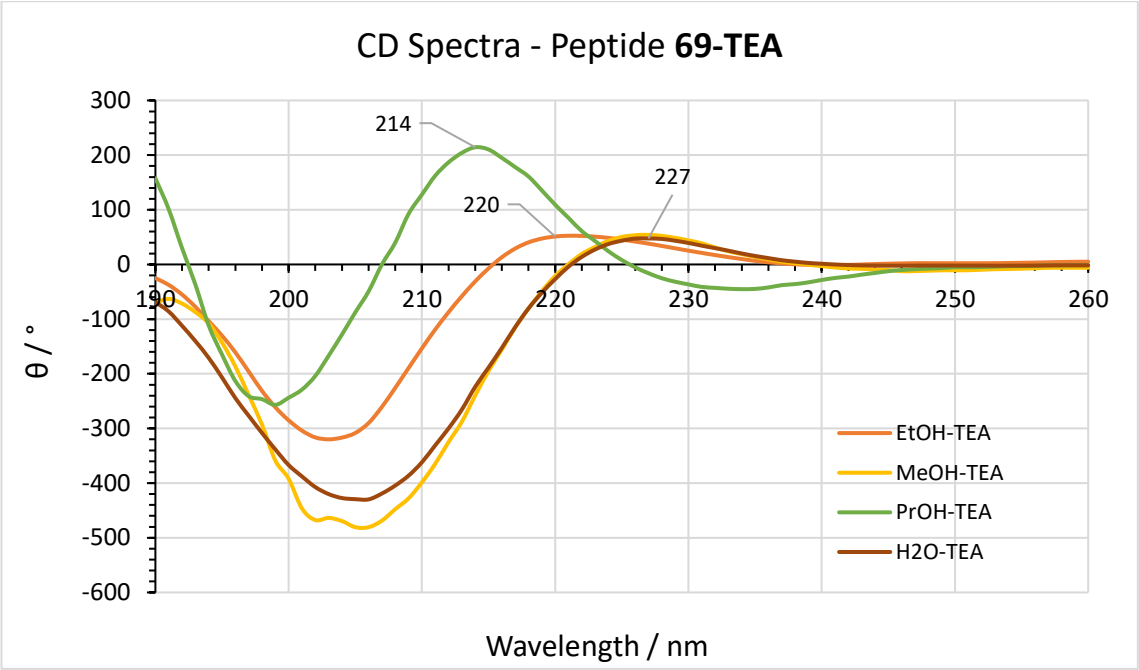


*Figure 138 – CD spectra for peptide **69** incubated in MeOH, EtOH, PrOH, and water (125 µM) with triethylamine (6 eq.), incubated for 14 days*

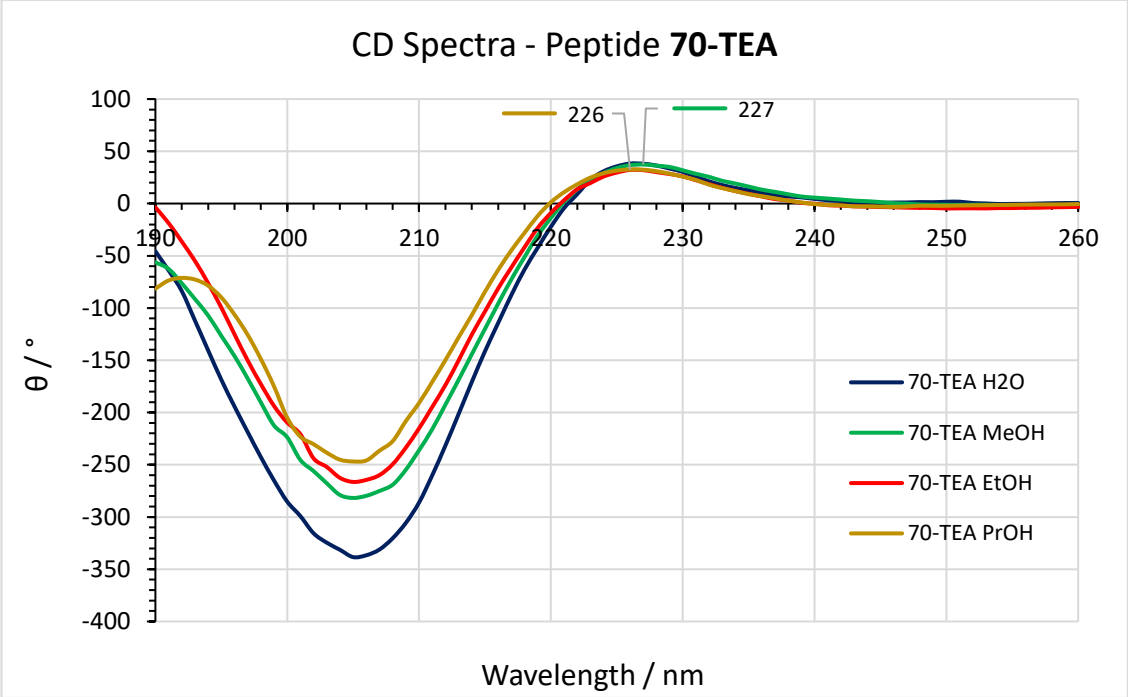


*Figure 139 – CD spectra for peptide **70** incubated in MeOH, EtOH, PrOH, and water (125 µM) with triethylamine (6 eq.), incubated for 14 days in solvent and further 14 days with base.*

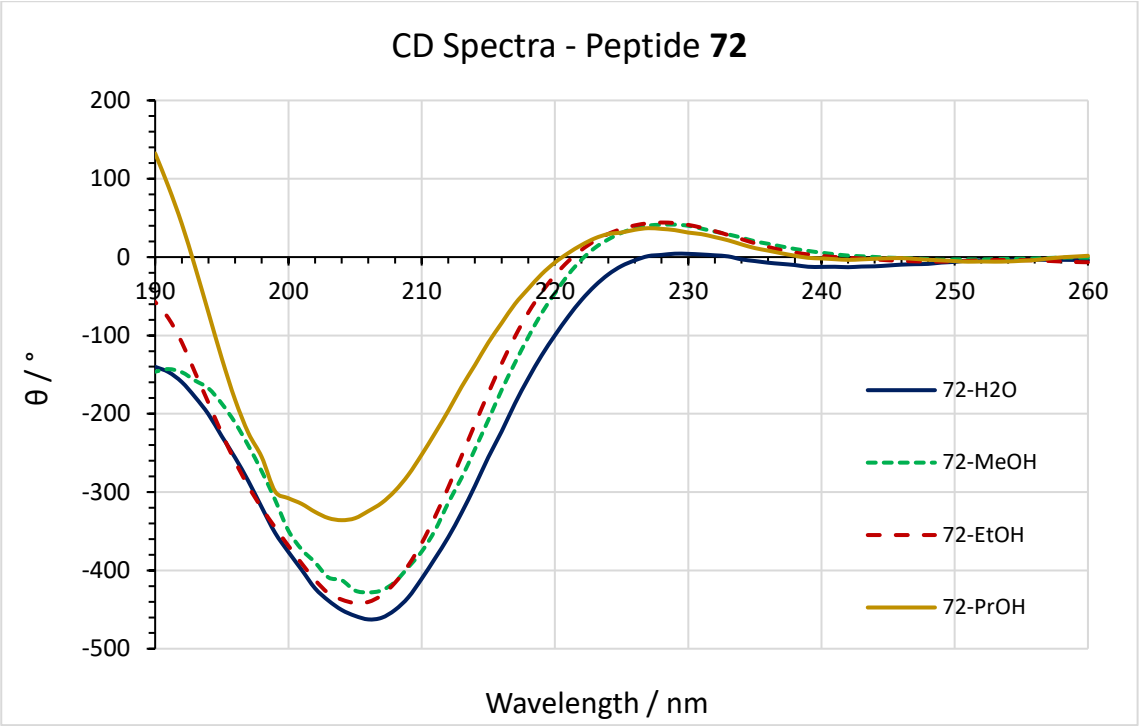


*Figure 140 – CD spectra of peptide **72** in water, MeOH, EtOH, and PrOH (125 µM), incubated for 14 days*

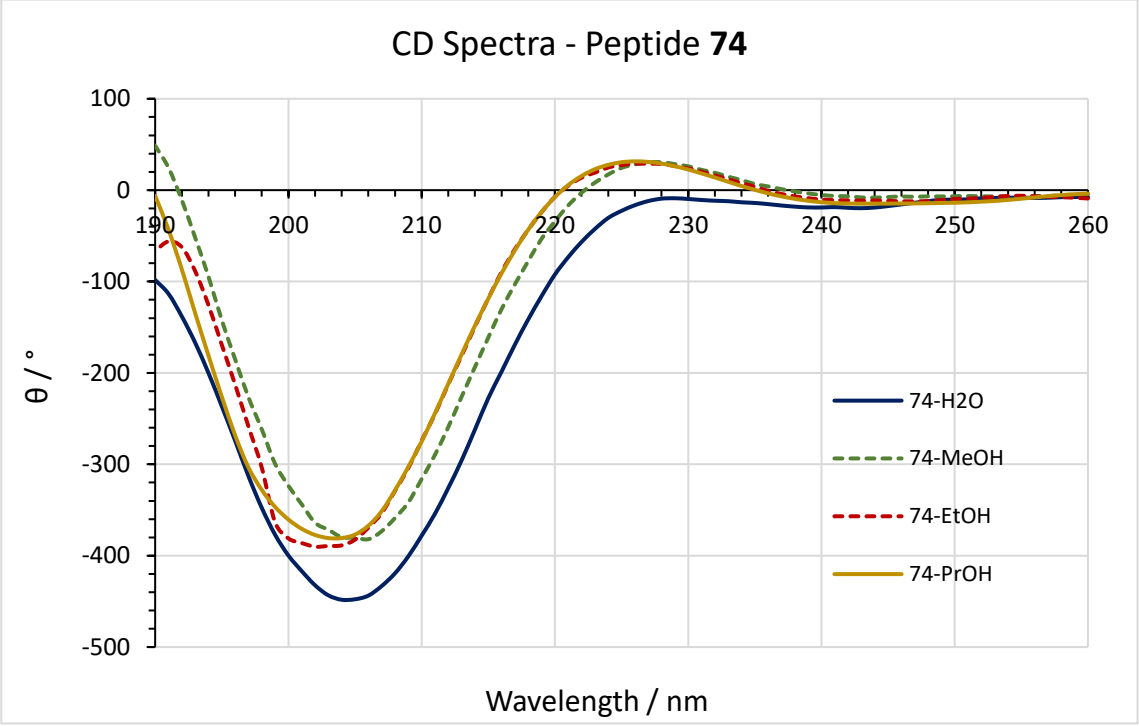


*Figure 141 – CD spectra of peptide **74** in water, MeOH, EtOH, and PrOH (125 µM), incubated for 14 days*

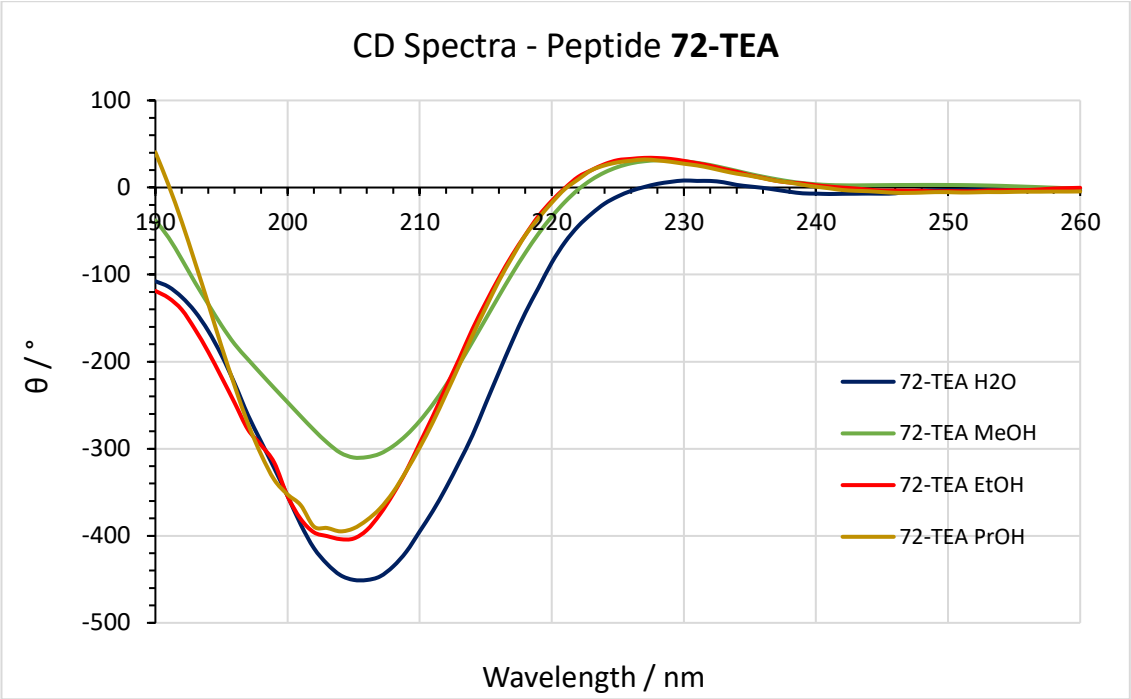


*Figure 142 – CD spectra for peptide **72** incubated in MeOH, EtOH, PrOH, and water (125 µM) with triethylamine (6 eq.), incubated for 14 days in solvent and further 14 days with base.*

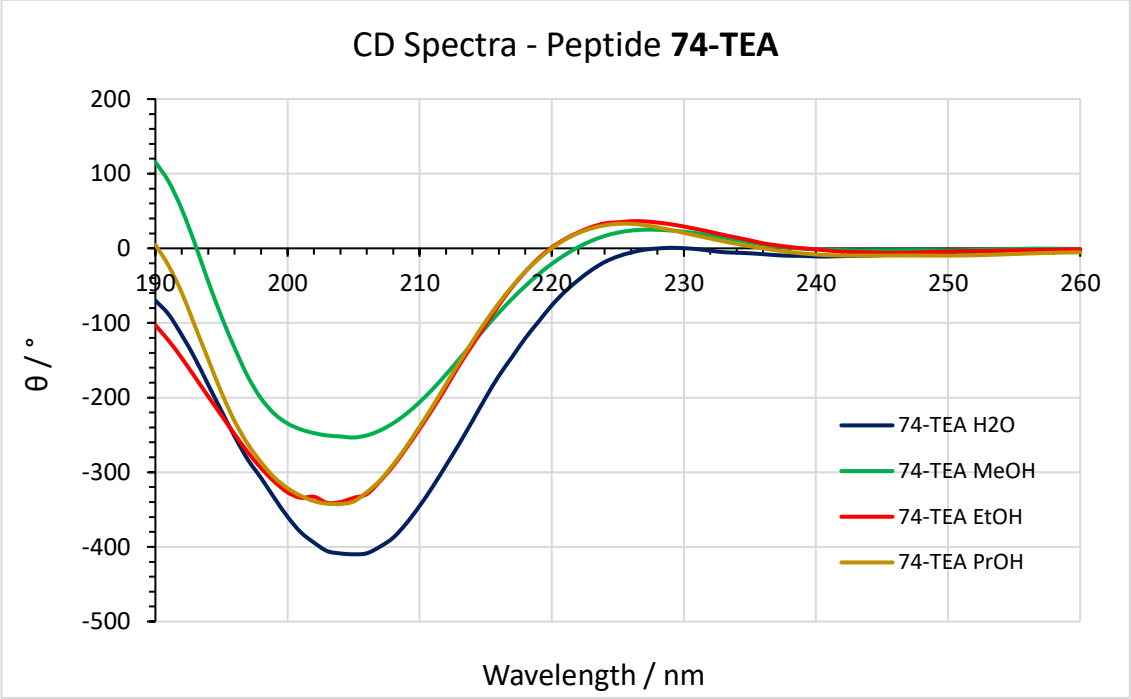


*Figure 143 – CD spectra for peptide **74** incubated in MeOH, EtOH, PrOH, and water (125 µM) with triethylamine (6 eq.), incubated for 14 days in solvent and further 14 days with base.*

SI 4.10.2 Synthesis of discrete peptide-metal cages based on carboxylic-acid functionalised peptides:

General procedure for synthesizing M_2L_2 M-Peptide complexes: 1.5-3 mg (1 eq.) of peptide (**71-74**) was dissolved in MeOD/DMSO-d_6/D_2O (540 µL) in an NMR tube and $Cu(NO_3)_2.3H_2O$ (1 eq.) or $Zn(NO_3)_2.6H_2O$ (1 eq.) was added to the solution. The resultant solution was heated at 65 °C for at least 2 h.

Synthesis of peptide 71 – copper complexes: Peptide **71** (5 mg, 5.6 µmol, 1 eq.) was dissolved in DMF (800 µL) and $Cu(NO_3)_2.3H_2O$ (1 eq, mg, µmol) was added forming a blue solution. The solution was heated at 85 °C for 5 min forming a green solution with no precipitate. This was heated for a further 1 day and remained green. (A control sample at the same concentration without peptide followed the same colour changes but continues to turn brown overtime) Precipitate formed upon addition of EtOAc or Et_2O to the peptide-Cu solution. The addition of EtOH does not cause a precipitate to form. Recrystallizations were attempted with DMF/EtOAc, DMF/Et_2O, EtOH/EtOAc, and EtOH/Et_2O, with the formation of only amorphous precipitates. These were attempted on the solution as is, and on precipitated solid washed with EtOAc (x 3) before redissolving in hot DMF or EtOH. The precipitate was fully soluble in H_2O. ^1H NMR analysis was carried out on the sample with a MeOD capillary insert. HPLC analysis was also a carried out on the sample: separated *via* RP-HPLC using a HiChrom KR100 5C18 5263 column at 40 °C on a Dionex UltiMate 3000. Gradient: 5% B for 5 minutes then from 5% B to 100% B over 20 minutes, and held at 100% B for 5 minutes. Where A is Water (0.1% formic acid) and B is methanol (0.1% formic acid). Flow rate is 1.0 mL/min. Wavelength: 225 nm. The complex was also submitted for HRMS analysis. (**71**)$_4$**Cu**$_4$, m/z calcd for $[M+H]^+$: $C_{164}H_{225}Cu_4N_{32}O_{56}^+$,3790.2921; $[M+4H]^{+4}$ 949.0789; found; [**71**+Na]$^+$ 909.404

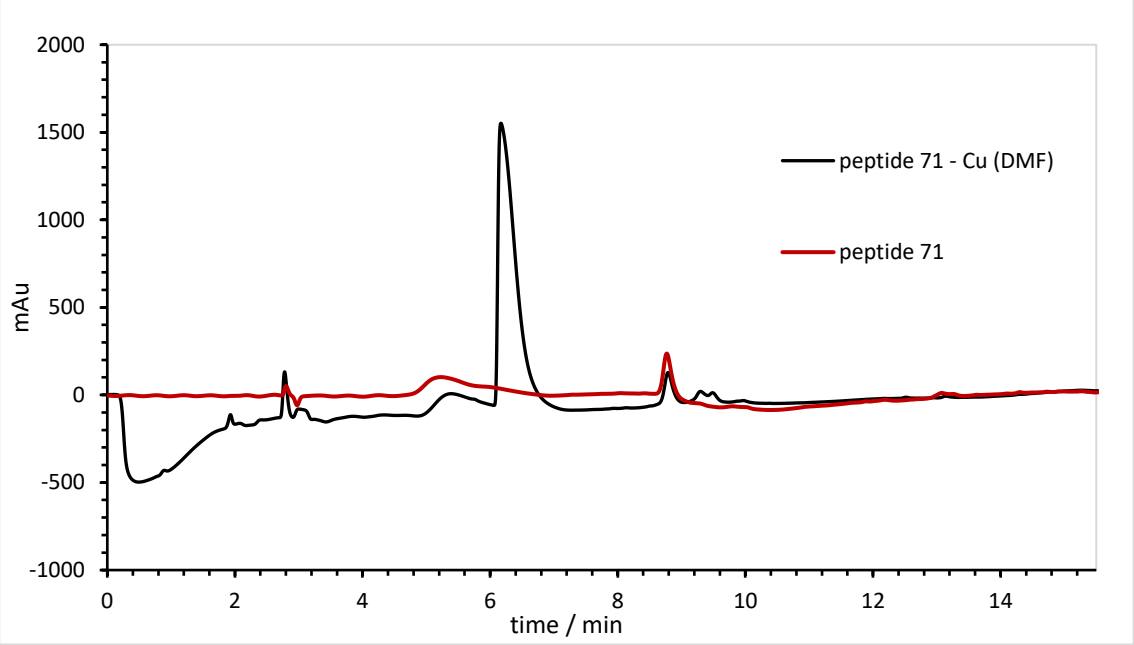


Figure 144 – HPLC UV-Vis chromatogram of peptide 71 and peptide 71 reacted with copper nitrate, large peak at rt 6.15 min is DMF

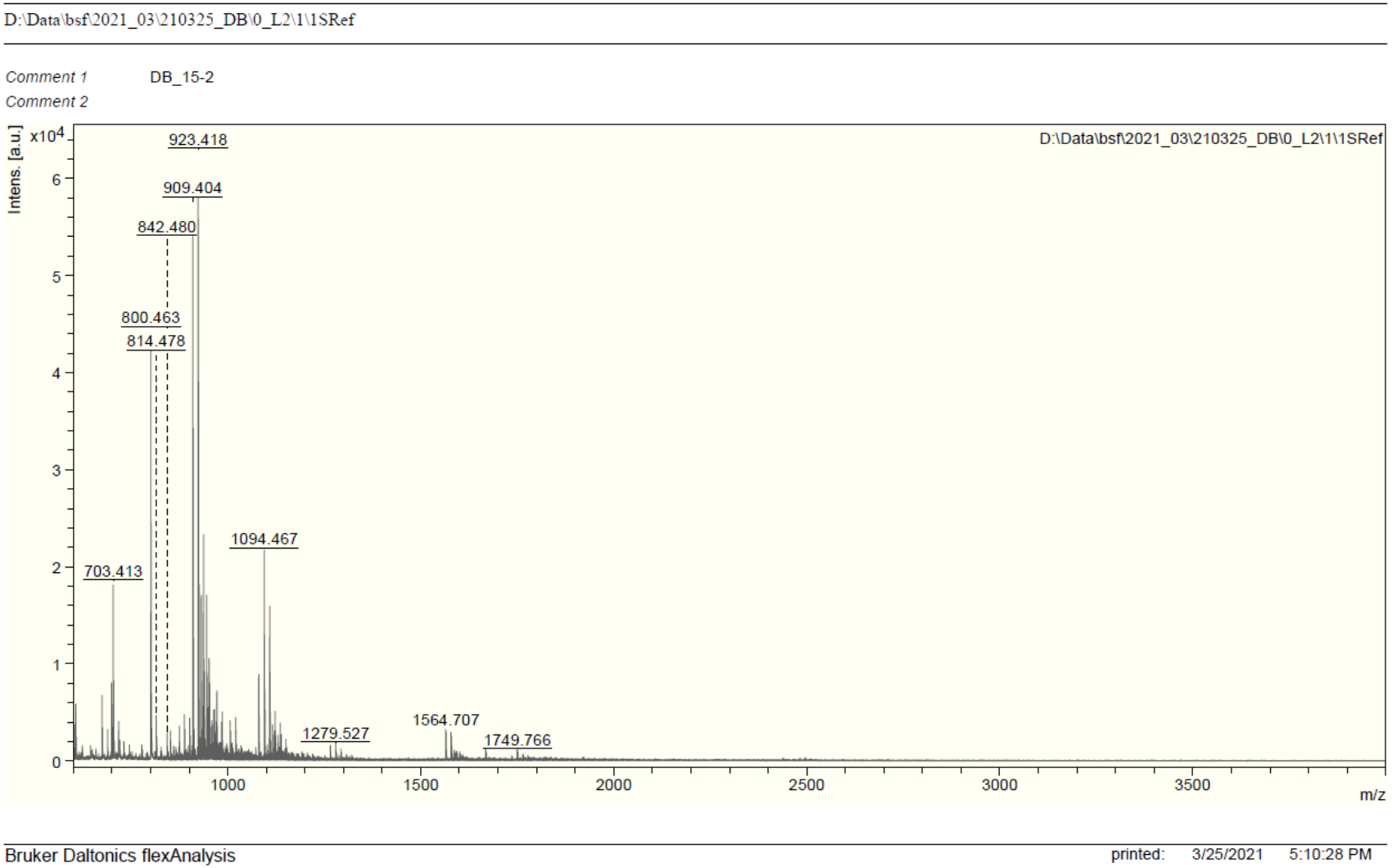


*Figure 145 – Mass spectrum (m/z 600-4000) of peptide **71** reacted with copper nitrate at 85 °C for 1 day in DMF*

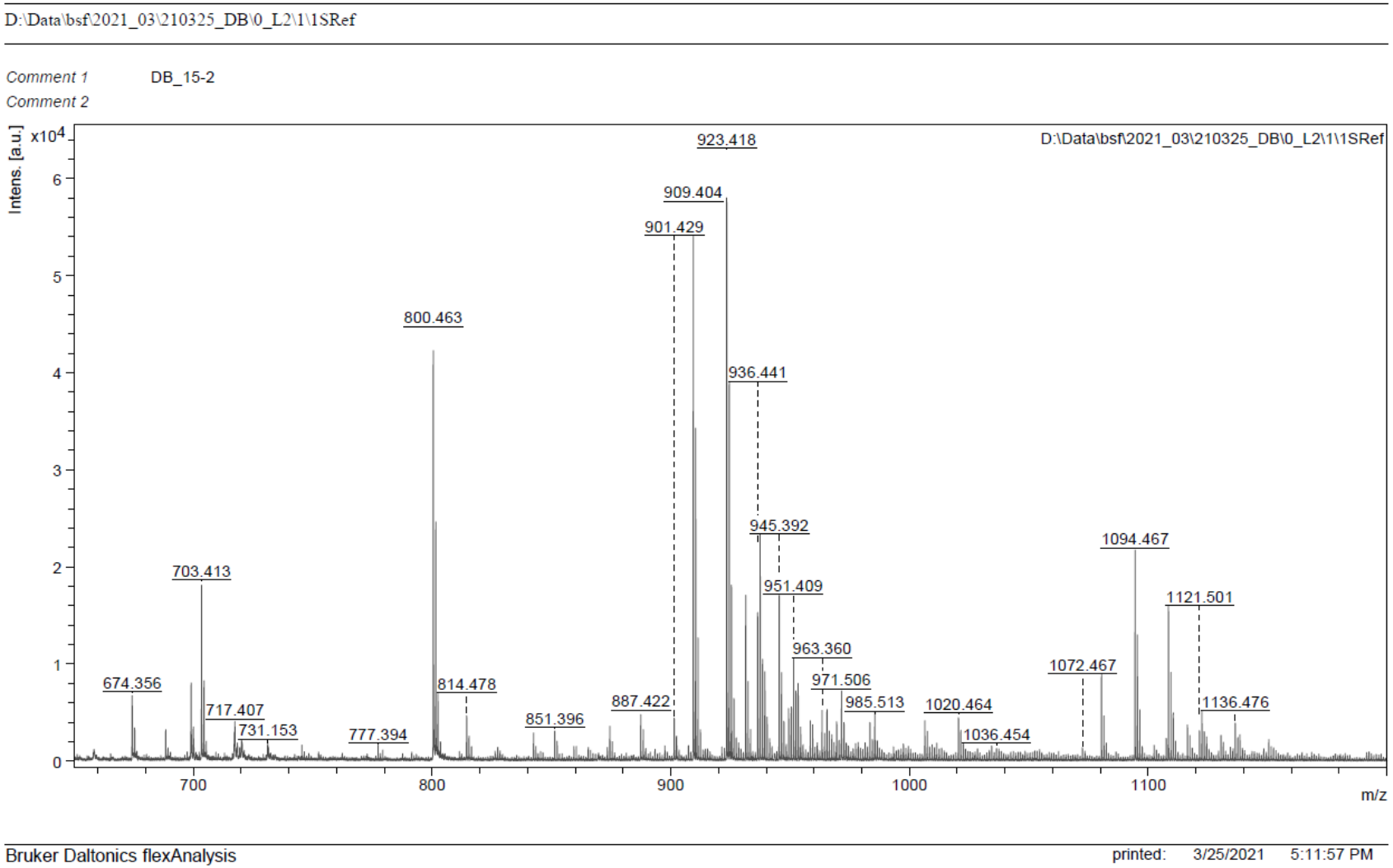


*Figure 146 - Mass spectrum (m/z 0-4000) enlarged m/z 650-1200 region of peptide **71** reacted with copper nitrate at 85 °C for 1 day in DMF*

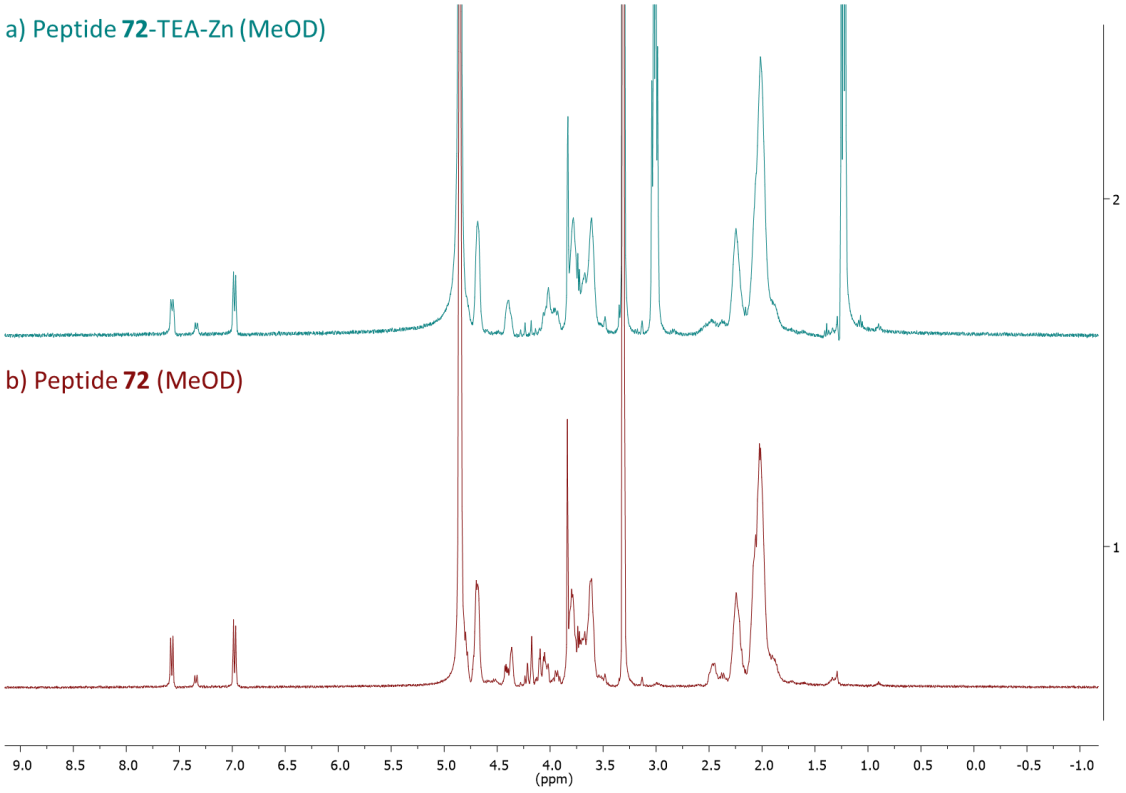


Figure 147 – ^{1}H NMR Spectra of peptide 72 in MeOD a) peptide with TEA (6 eq.) and zinc nitrate (1 eq.), b) peptide alone

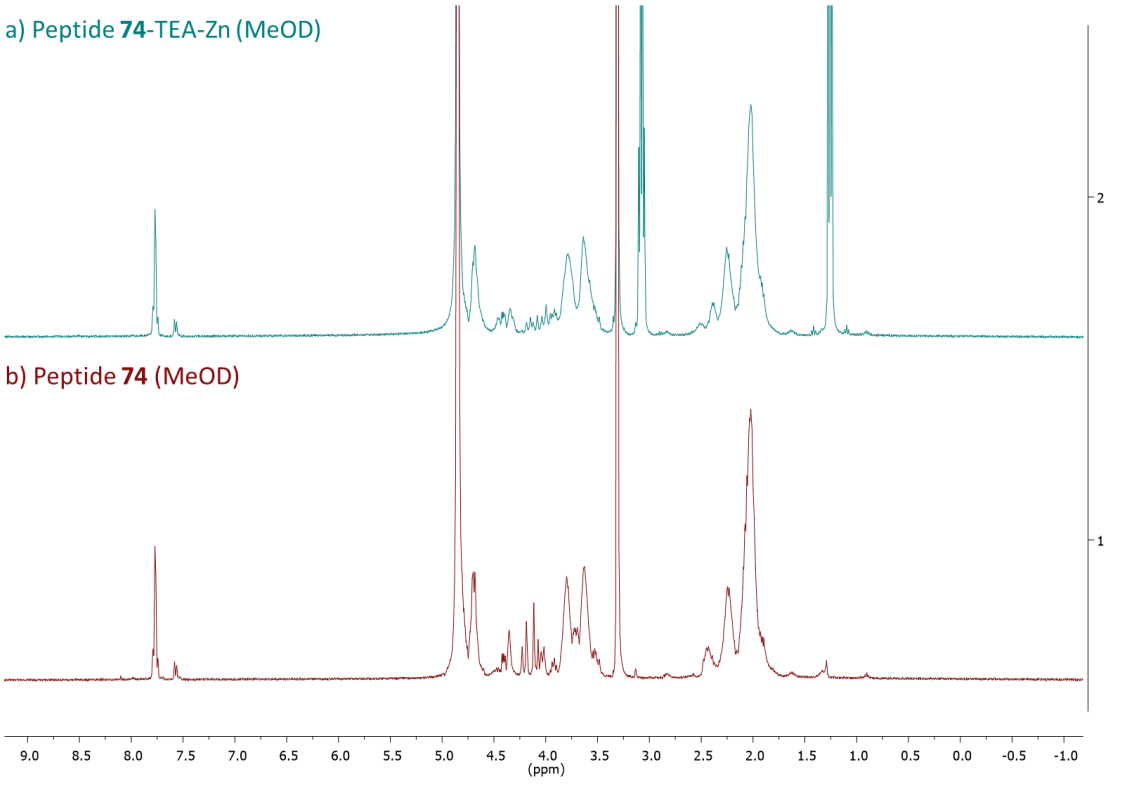


Figure 148 – ^{1}H NMR Spectra of peptide 74 in DMSO-d_6 a) peptide with TEA (6 eq.) and zinc nitrate (1 eq.), b) peptide alone

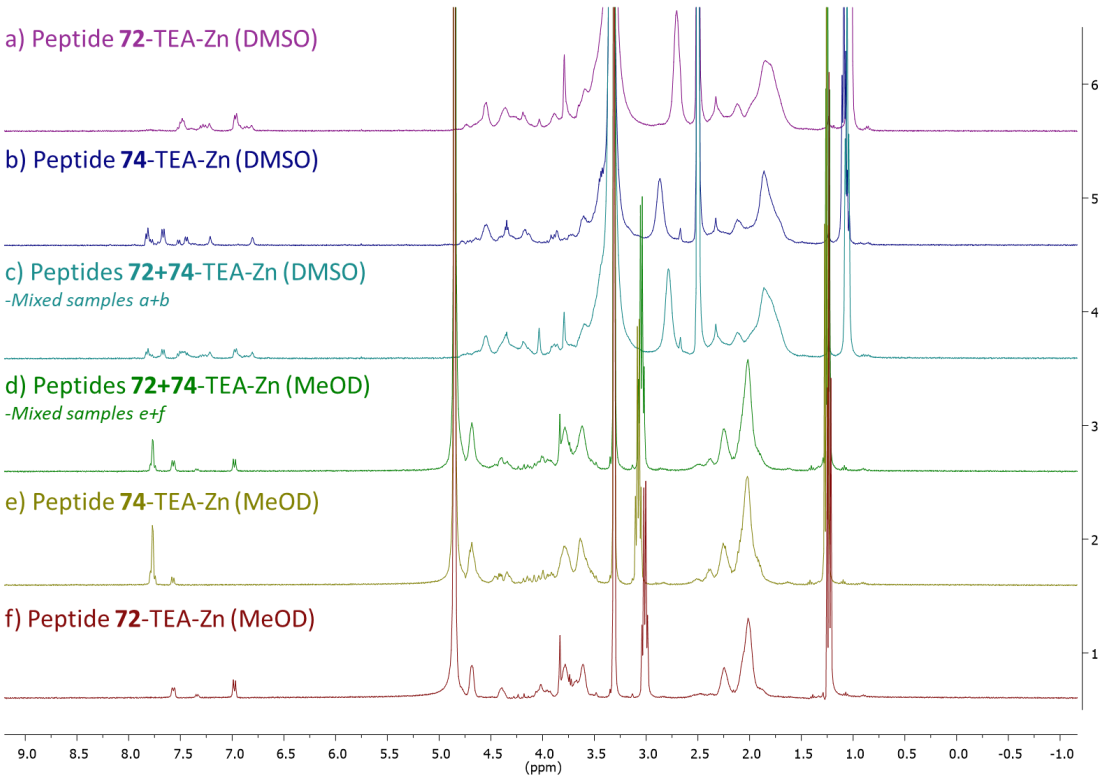


Figure 149 –^{1}H NMR Spectra of peptide 72 and 74 zinc complexes; a) peptide 72 with TEA (6 eq.) and zinc nitrate (1 eq.) in DMSO-d_6, b) peptide 74 with TEA (6 eq.) and zinc nitrate (6 eq.) in DMSO-d_6, c) complexes from samples (a) and (b) mixed, d) complexes from samples (e) and (f) mixed, e) peptide 74 with TEA (6 eq.) and zinc nitrate (1 eq.) in MeOD, f) peptide 72 with TEA (6 eq.) and zinc nitrate (1 eq.) in MeOD

SI 4.10.3 Pyridine-based Discrete Peptide-Metal Complexes:

Synthesis of Pd(en)Cl_2: K_2PdCl_4 (0.4 g, 1.22 mmol) was dissolved in 0.1 N HCl (8 mL) under stirring. Any insoluble salt was removed at this stage by filtration. To this, ethylenediamine (1.35 mmol, 1.1 eq.) in 0.05 N HCl (40 mL) was added dropwise. The resultant mixture was stirred at 60 °C for 2 h. A yellowish-orange precipitate of Pd(en)Cl_2 was formed and stirred for a further 30 min at 60 °C. Precipitate was collected by filtration and was thoroughly washed with H_2O, ethanol, and diethyl ether. A yellow powder was obtained (88%).

Synthesis of Pd(en)$(NO_3)_2$: Pd(en)Cl_2 (0.26 g, 1.09 mmol,) was dissolved in water (280 mL) and $AgNO_3$ (2.3 mmol, 2.1 eq.) was added. The resulting solution was stirred in the dark at 60 °C for 2 h and then at room temperature for a further 24 h. The AgCl precipitate was removed by gravity filtration and the filtrate concentrated by gentle heating on a hot plate to yield a residue of [Pd(en)$(ONO_2)_2$] as a yellow solid (73%). Analytical data matches the previous reports.

Treatment with ion exchange resin: 1 g (per 50 mg of peptide) of highly basic Purolite A300 resin was treated with DI water (250 mL), 4% aqueous NaOH solution (250 mL), DI water (250 mL), and MeOH (250 mL). The peptide was then added to the resin in MeOH and spun for 2 days. The solvent was then collected and the resin flushed with MeOH (250 mL) before removing the solvent under vacuo to yield the free peptide, no longer the TFA/formate salt of the peptide.

General procedure for synthesizing $mL_{1\text{-}2}$Pd-Peptide complexes: 1.5-3 mg (1 eq.) of **75** was dissolved in DMSO-d_6 or D_2O (540 µL) in an NMR tube and Pd(en)$(NO_3)_2$ (1.1 eq.) or Pd$(CH_3CN)_4(BF_4)_2$ (0.55 eq.) was added to it. The solution was then heated at 65 °C for at least 2 h. The quantitative complex formation was confirmed by ^{1}H NMR.

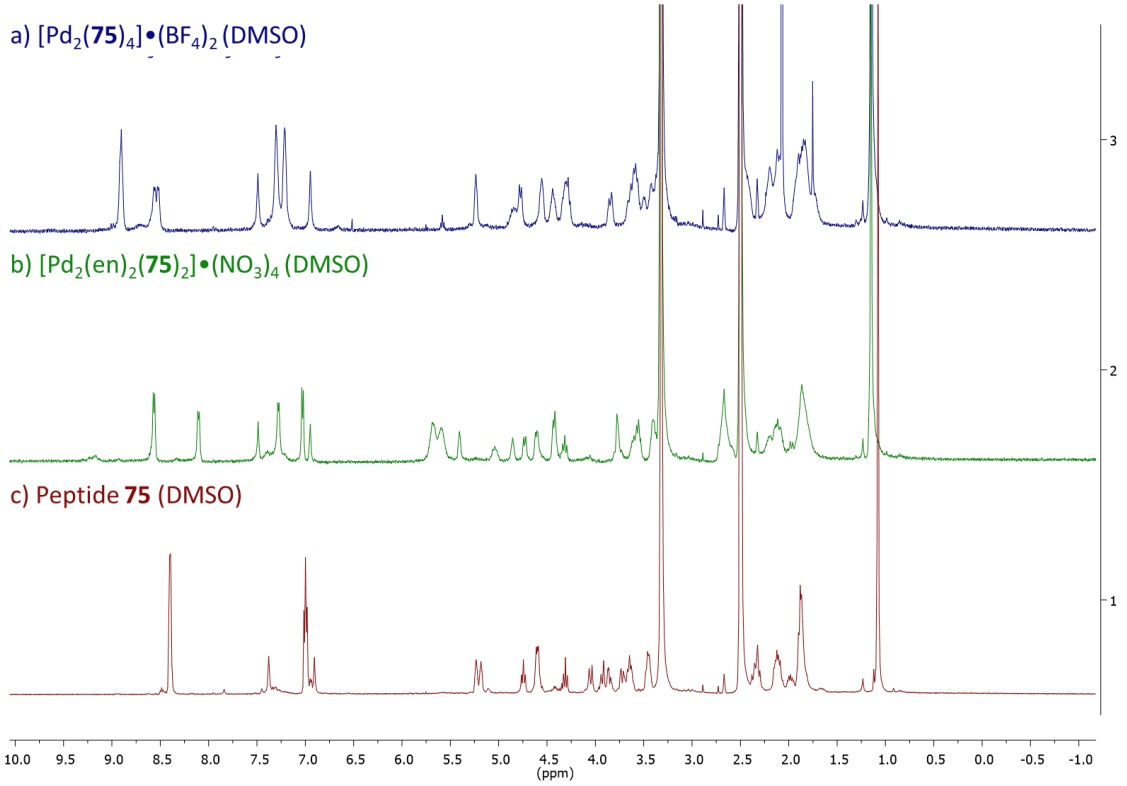


Figure 150 – ^{1}H NMR spectra in DMSO-d_6 of a) peptide 75 complex with $Pd(CH_3CN)_4(BF_4)_2$, b) peptide 75 complex with $Pd(en)(NO_3)_2$, c) Peptide 75 alone

ML complex: [Pd(en)(**75**)]•$(NO_3)_2$: ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.57 (d, J = 6.2 Hz, 2H), 8.11 (d, J = 5.4 Hz, 2H), 7.49 (s, 1H), 7.28 (d, J = 5.3 Hz, 3H), 7.03 (d, J = 6.4 Hz, 2H), 6.95 (s, 1H), 5.64 (m, 6H), 5.41 (s, 1H), 5.04 (s, 1H), 4.86 (s, 1H), 4.73 (d, J = 9.2 Hz, 1H), 4.61 (d, J = 8.5 Hz, 2H), 4.43 (d, J = 8.4 Hz, 2H), 4.32 (t, J = 8.5 Hz, 1H), 3.79 (d, J = 12.7 Hz, 2H), 3.67 – 3.50 (m, 3H), 3.40 (d, J = 6.9 Hz, 3H), 2.67 (s, 6H), 2.15 (m, 5H), 1.97 (d, J = 10.6 Hz, 1H), 1.86 (s, 7H), 1.15 (s, 9H).

^{1}H NMR (400 MHz, D_2O) δ 8.49 (d, *J* = 6.0 Hz, 2H), 8.29 (d, *J* = 6.2 Hz, 2H), 7.08 (t, *J* = 5.0 Hz, 4H), 5.41 (s, 1H), 5.15 (dt, *J* = 17.1, 8.5 Hz, 1H), 4.87 (d, *J* = 9.1 Hz, 1H), 4.74 – 4.68 (m, 1H), 4.62 (dd, *J* = 16.2, 7.8 Hz, 2H), 4.57 – 4.49 (m, 1H), 4.10 (d, *J* = 12.2 Hz, 1H), 3.99 (d, *J* = 12.8 Hz, 1H), 3.87 – 3.71 (m, 3H), 3.67 – 3.47 (m, 3H), 2.89 (s, 4H), 2.79 – 2.69 (m, 2H), 2.65 (s, 2H), 2.44 – 2.26 (m, 3H), 2.25 – 1.82 (m, 9H), 1.25 (s, 9H).

[Pd(en)(75)]•$(NO_3)_2$, m/z calcd for $[M+H]^+$: $C_{37}H_{53}N_9O_7Pd^{2+}$, 841.3092 ; found; M^{2+}, 420.6571

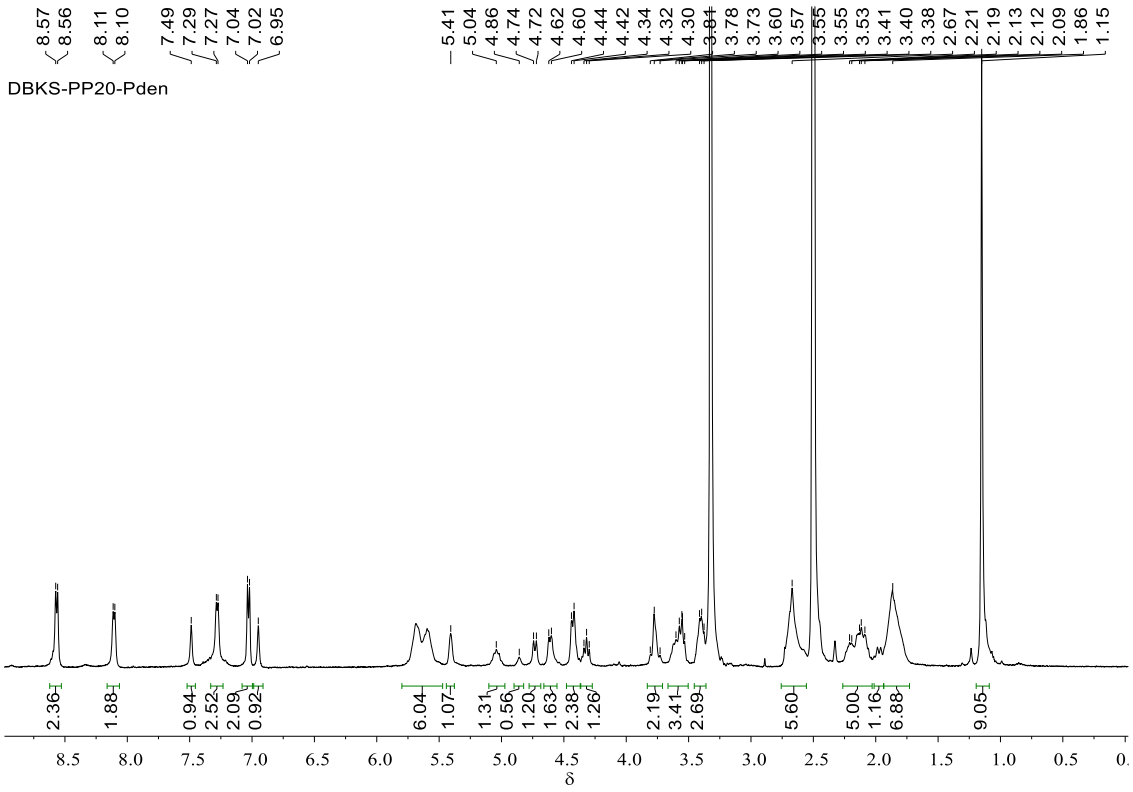


*Figure 151 - ^{1}H NMR spectrum of **[Pd(en)(75)]•$(NO_3)_2$** (400 MHz, DMSO-d_6, 298 K).*

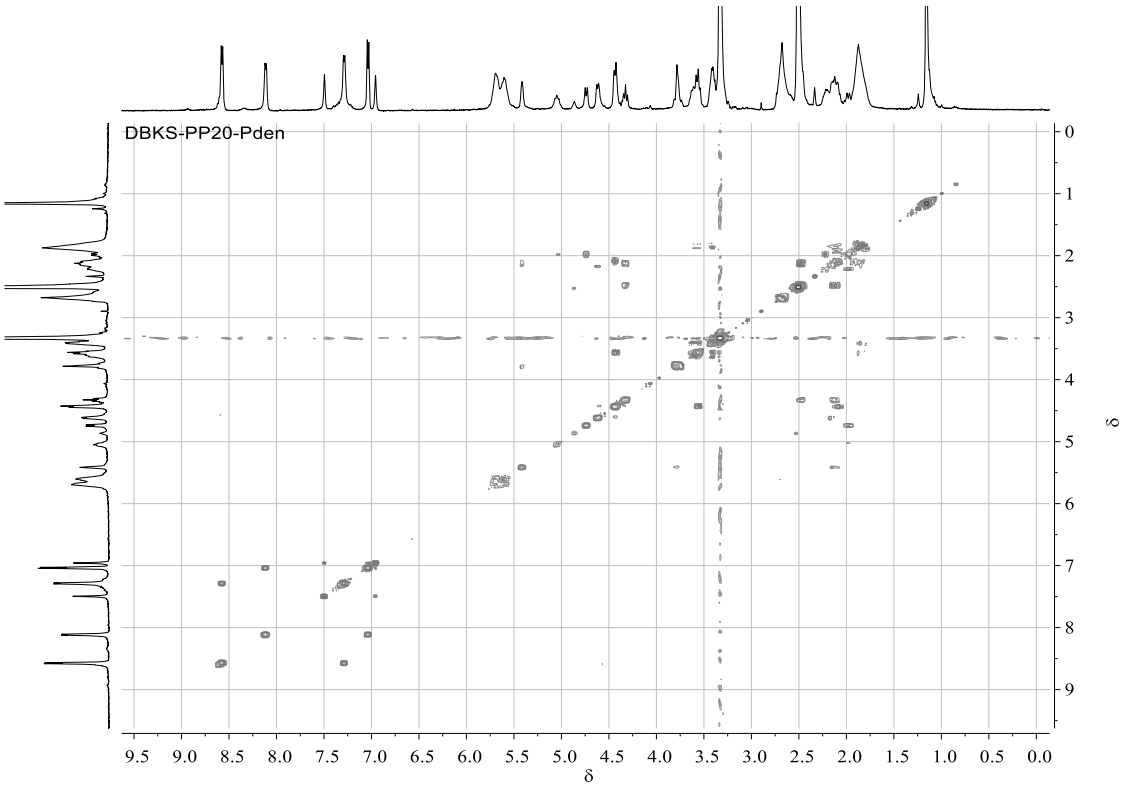


*Figure 152 - ^{1}H-^{1}H COSY spectrum of **[Pd(en)(75)]•$(NO_3)_2$** (400 MHz, DMSO-d_6, 298 K).*

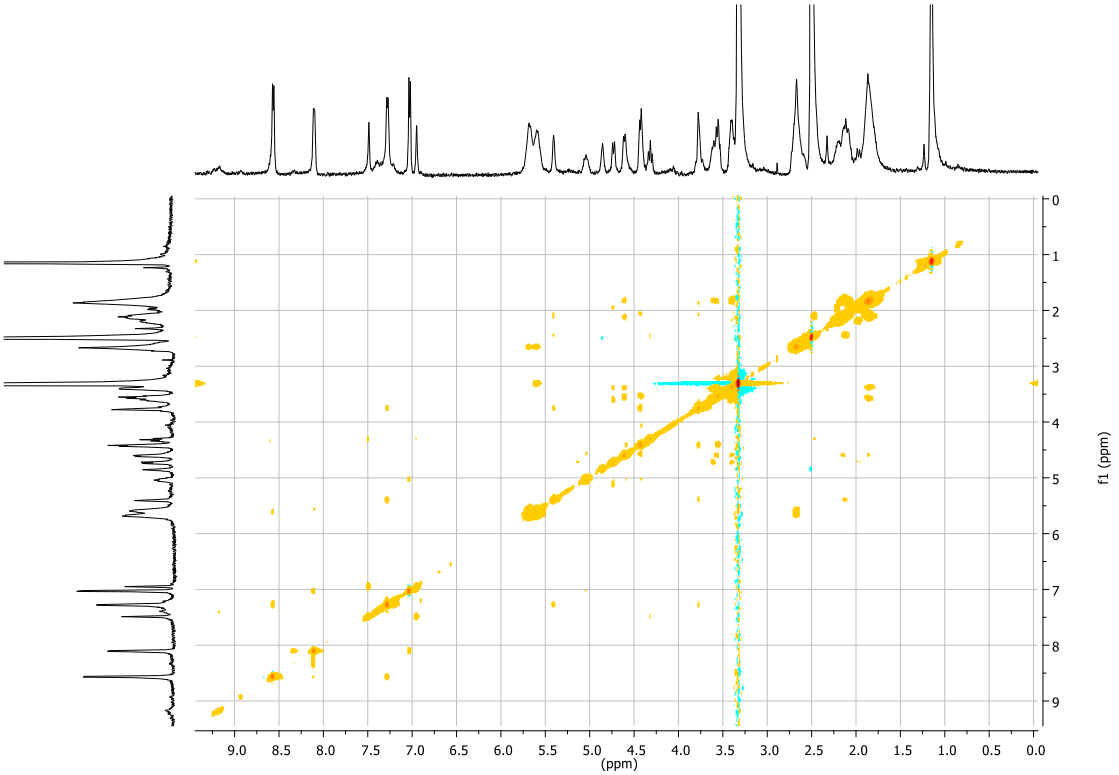


*Figure 153 - ^{1}H-^{1}H NOESY spectrum of **[Pd(en)(75)]•(NO₃)₂** (400 MHz, DMSO-d_6, 298 K).*

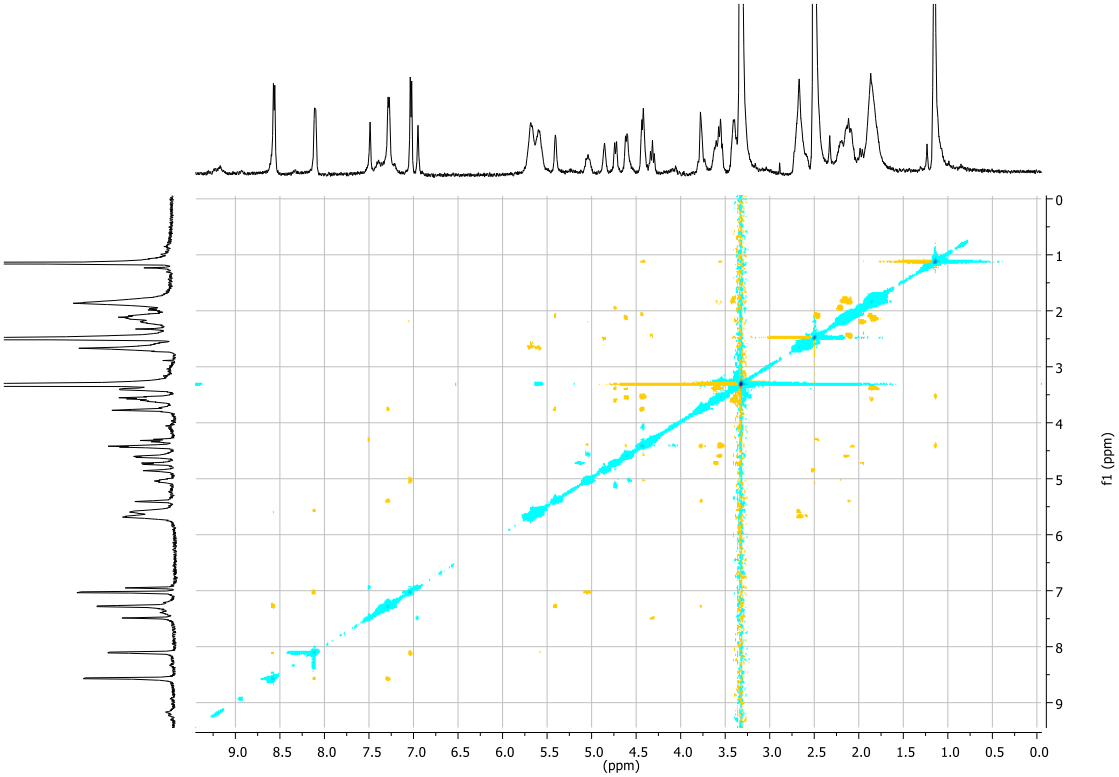


*Figure 154 ^{1}H-^{1}H ROESY spectrum of **[Pd(en)(75)]•(NO₃)₂** (400 MHz, DMSO-d_6, 298 K).*

ML_2 complex: [Pd(**75**)$_2$]•(BF_4)$_2$: ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.90 (s, 1H), 8.54 (d, J = 16.5 Hz, 1H), 7.49 (s, 1H), 7.30 (s, 1H), 7.21 (s, 1H), 6.95 (s, 1H), 5.24 (s, 1H), 4.91 – 4.71 (m, 1H), 4.56 (s, 1H), 4.44 (s, 1H), 4.30 (d, J = 13.5 Hz, 1H), 3.85 (d, J = 11.7 Hz, 1H), 3.60 (q, J = 10.0, 8.5 Hz, 2H), 3.46 (m, 2H), 2.20 (s, 1H), 2.07 (d, J = 1.4 Hz, 4H), 1.96 – 1.68 (m, 4H), 1.15 (s, 5H).

^{1}H NMR (400 MHz, D_2O) δ 9.02 – 8.91 (m, 0.5H), 8.91 – 8.78 (m, 0.5H), 8.59 (s, 2H), 8.22 (t, *J* = 6.0 Hz, 2H), 7.28 – 7.16 (m, 1H), 7.09 (dt, *J* = 15.7, 7.5 Hz, 3H), 5.48 – 5.36 (m, 1H), 5.36 – 5.19 (m, 2H), 4.69 – 4.59 (m, 2H), 4.58 – 4.48 (m, 1H), 4.45 – 4.22 (m, 1H), 4.11 (d, *J* = 12.3 Hz, 1H), 3.98 (d, *J* = 11.4 Hz, 2H), 3.91 – 3.69 (m, 3H), 3.69 – 3.47 (m, 3H), 2.79 – 2.67 (m, 1H), 2.66 – 2.55 (m, 1H), 2.44 – 2.25 (m, 3H), 2.25 – 2.14 (m, 2H), 2.13 – 1.83 (m, 11H), 1.26 (d, *J* = 8.6 Hz, 9H).

[Pd(75)$_2$]•(BF$_4$)$_2$, m/z calcd for [M]$^{+}$: $C_{70}H_{89}N_{14}O_{14}Pd^{2+}$, 1455.5785; found; [M]$^{+2}$, 728.7911;

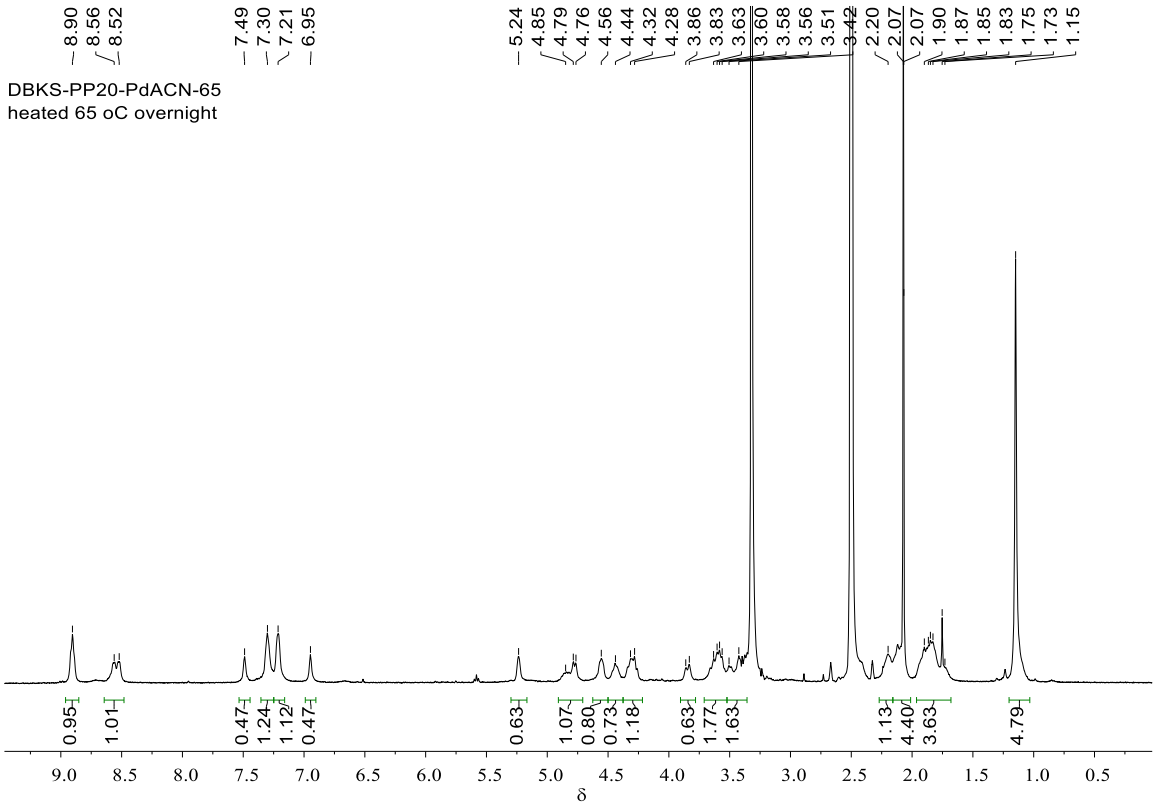


*Figure 155 - ^{1}H NMR spectrum of **[Pd(75)$_2$]•(BF$_4$)$_2$** (400 MHz, DMSO-d$_6$, 298 K).*

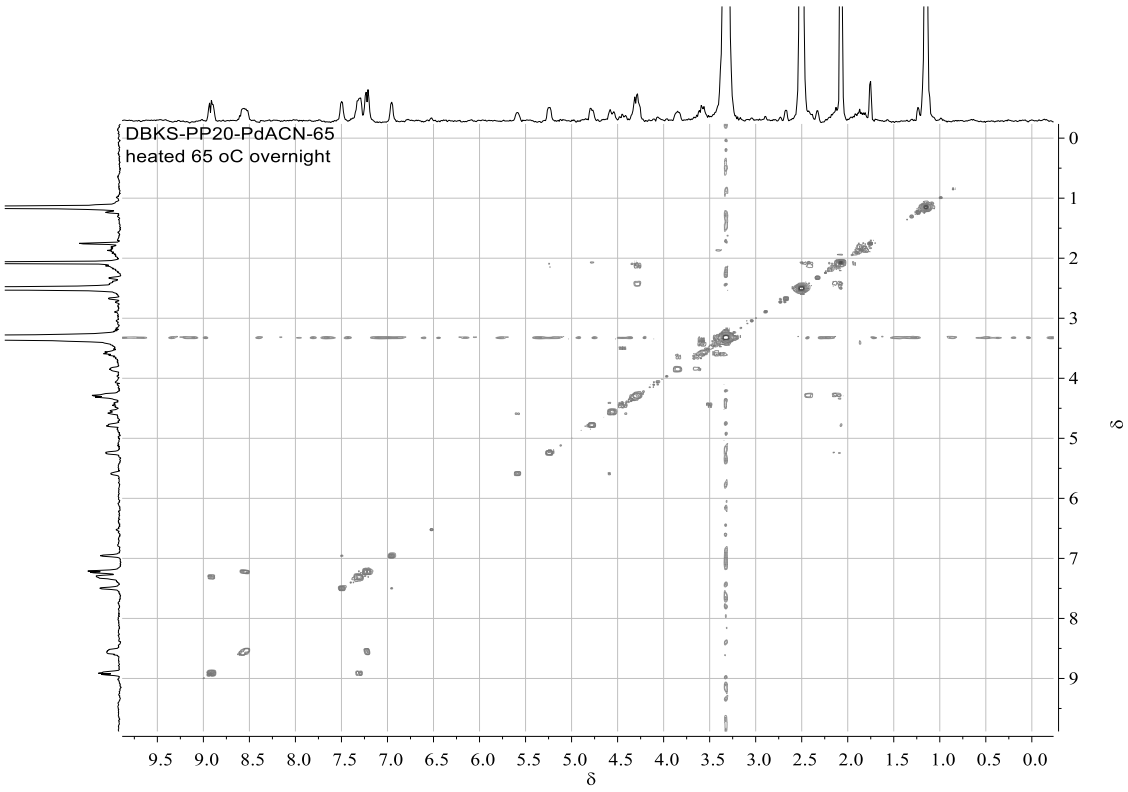


*Figure 156 - ^{1}H-^{1}H COSY spectrum of **$[Pd(75)_2]\bullet(BF_4)_2$** (400 MHz, DMSO-$d_6$, 298 K).*

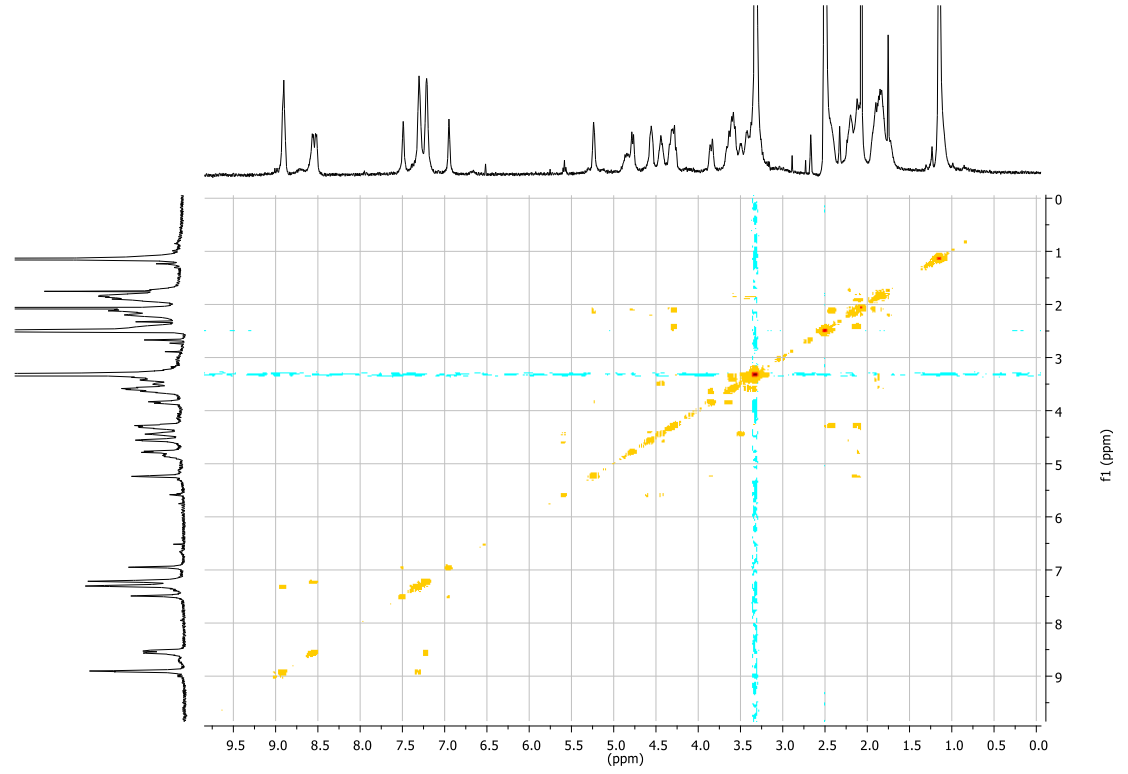


*Figure 157 - ^{1}H-^{1}H NOESY spectrum of **$[Pd(75)_2]\bullet(BF_4)_2$** (400 MHz, DMSO-$d_6$, 298 K).*

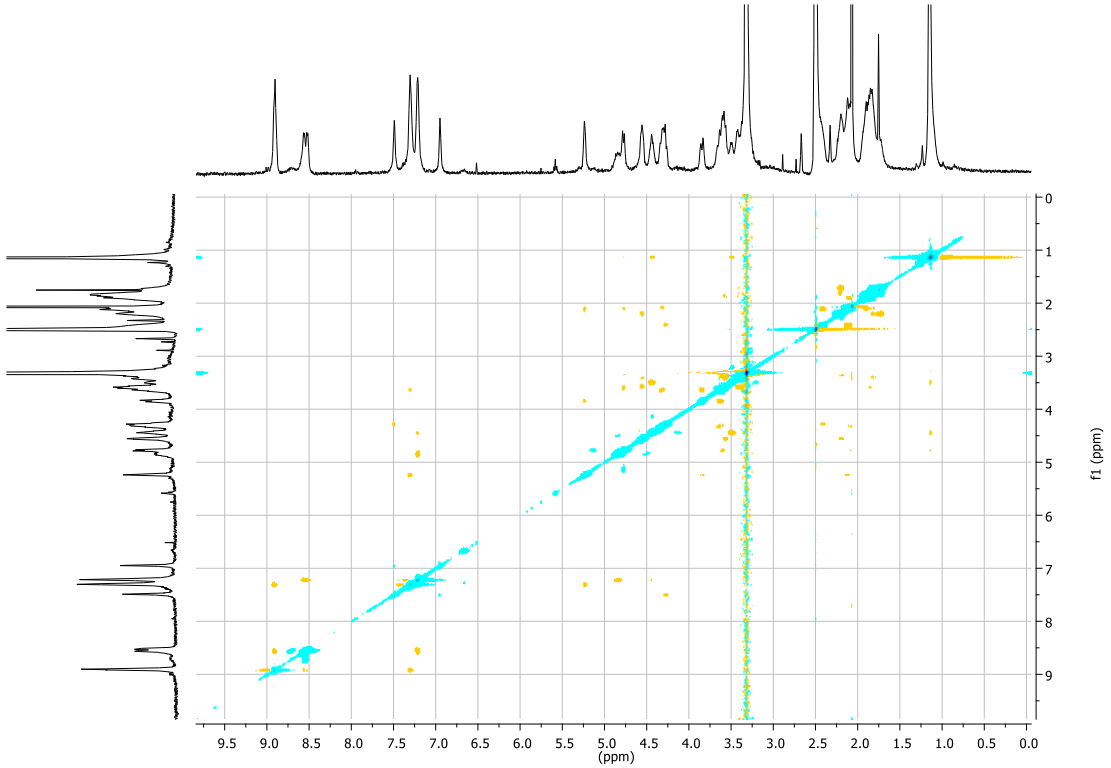


*Figure 158 - ^{1}H-^{1}H ROESY spectrum of **[Pd(75)$_2$]•(BF$_4$)** (400 MHz, DMSO-d_6, 298 K).*

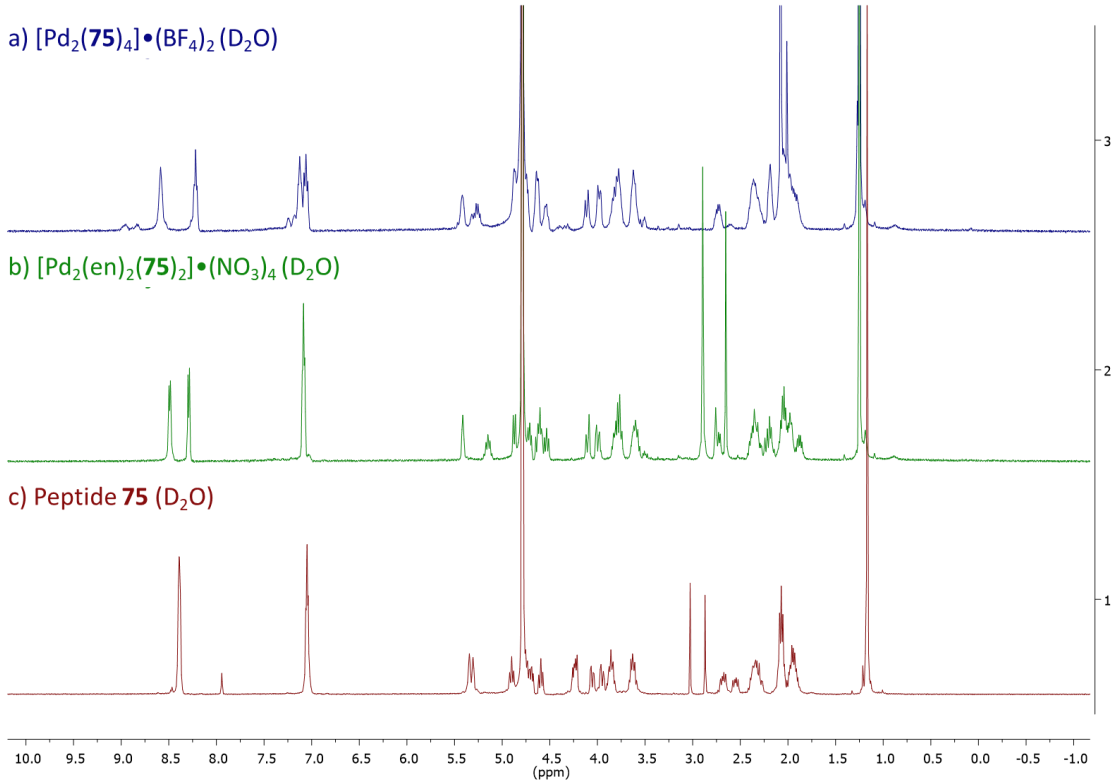


Figure 159 - ^{1}H NMR spectra in D_2O of; a) peptide 75 complex with $Pd(CH_3CN)_4(BF_4)_2$, b) peptide 75 complex with $Pd(en)(NO_3)_2$, c) Peptide 75 alone

SI 4.11 References - 2

1 M. T. Ma, H. N. Hoang, C. C. G. Scully, T. G. Appleton and D. P. Fairlie, *J. Am. Chem. Soc.*, 2009, **131**, 4505–4512.

Chapter 5.

Summary and Future Outlook

5.1 Summary of Work

There is an increasing need to develop ever more targeted systems for a wide range of applications *via* the use supramolecular constructs. This has led to a desire to develop artificial supramolecular systems to mimic complex biological systems such as enzymes such that these materials can be designed and tuned towards specific functions. However, despite progress towards developing such systems there remain significant challenges, ranging from their predictable assembly to the range of accessible building blocks and structures. Increasingly, biomaterials are seen as promising building-blocks for the synthesis of biocompatible, chemically complex, and tuneable supramolecular constructs.

Inspired by successes in the development of bioinspired supramolecular constructs whereby this new generation of supramolecular materials has demonstrated applications in catalysis with substrate selectivity and stereoselectivity, chemical separations and biomedical applications, we aimed to specifically make use of peptides as supramolecular building-blocks. Their ease of synthesis, periodicity, biocompatibility, economic viability, scalability, and vast array of accessible structures made peptide ideal candidates for this endeavour. However, the secondary structures of peptides are often prone to perturbation upon functionalisation and are highly flexible, compared to traditional rigid polyaromatic ligands, limiting the potential of these materials with such obstacles to their rational design. As such, we sought to incorporate polyproline helices as building-blocks, as these peptides are highly stable to functionalisation and are notably rigid compared to other secondary structures.

The first part of chapter 1 explored the state of play within supramolecular chemistry and how increasingly more complex supramolecular structures have been developed with promising applications as researchers have moved from traditional building-blocks for the synthesis of constructs in MOFs, COFs, SOFs and other nanomaterials towards more chemically complex bioinspired building-blocks such as DNA, RNA, aptamers, amino acids, proteins, and recently peptides. This discussion highlights how investigations into the use of peptides in supramolecular materials is still limited, especially for MOFs, SOFs, and COFs, and thus the development of novel constructs based on these building-blocks is an essential avenue to be explored within the field. Polyproline helices have also been largely unexplored as potential supramolecular building blocks and recent work within literature exploring their use in a MOF and some metal-organic cages (MOCs) or entangled polyhedra are described, highlighting the potential applications of this secondary structure.

The structure and unique features of the polyproline helix are detailed, such as its' unique ability switch from *trans* to *cis* amide bonds, due to the proline secondary amine, with a cyclic sidechain, and thus can switch between the extended left-handed polyproline II helix and the shortened right-handed polyproline I helix. The stabilising factors, *trans*ition processes, and the effects of functionalisation and external stimuli on this process are further described.

With these aspects in mind the final part of the chapter looks at the design considerations of a polyproline helix building-block for various applications and discusses the potential chemical handles available for supramolecular interactions. Furthermore, the synthesis of non-natural proline monomers to incorporate functional groups into the polyproline helix *via* several routes such as the Mitsunobu reaction or Suzuki cross-coupling is investigated. With these considerations and potential accessible structures the applications of synthesised supramolecular materials is discussed, while highlighting recent examples within literature of relevant peptide and polyproline-based supramolecular constructs that influence the avenues of investigation of this project.

Chapter 2 presents the recently published work in Chemistry: A European Journal *("A Reversibly Porous Supramolecular Peptide Framework"*) and ACS Macro Letters (*"Supramolecular Self-Assembly of Engineered Polyproline Helices*") expanding on these articles with further experimental data which is as yet unpublished. The first section of chapter 2 focuses on the initial publication, in Chemistry A European Journal, on the synthesis of a proline tetramer which self-assembles to form a porous supramolecular peptide framework. This was analysed *via* SC-XRD to give a crystal structure, this determined the peptide adopted the polyproline II helix, the shortest unfunctionalized polyproline to adopt the secondary structure, giving a useful model for the geometries of the proline residues in a full helical turn.

We were then able to demonstrate that the supramolecular peptide framework exhibited reversible porosity where thermal activation caused removal of encapsulated ethanol molecules resulting in pore collapse. Significantly, re-exposure to ethanol vapour then reformed the original structure, confirmed by single crystal and powder x-ray diffraction. As the structure could be reinflated, we sought to replace the encapsulated molecule with a new guest. We successfully demonstrated this with several different solvent molecules analysed from the material powder patterns and ^{1}H NMR to confirm solvent encapsulation. However, to visualise the new encapsulated guest we carried out vapour diffusion of iodine into the desolvated framework, with the high electron density of iodine allowing SC-XRD to confirm the presence of the new guest and the slight adaption of the framework to accommodate the new guest and forming host-guest halogen bonds. Due to the chirality of the framework, inherent from the peptide helix, we subsequently sought to incorporate a chiral molecule, 1-phenylethan-1-ol, which demonstrated that the framework was enantioselective, preferential to absorption of the S-enantiomer. Thus, highlighting the potential applications of these porous peptidic materials in chemical separations with chiral pores and host-guest interactions.

The second part of the chapter focuses on the more recently published paper, in ACS Macro Letters, on the rational design of functionalised oligoproline tetramers by using the peptide structure determined from the crystal structure of the Fmoc-protected proline tetramer to predict the geometry of new hydroxyl groups on each residue and how this placement would affect the packing principles. A series of hydroxyl functionalised derivatives were initially synthesised with which the packing could be altered to form either isostructural frameworks or new porous and non-porous peptide frameworks depending on which positions were functionalised with hydroxy groups. We were also able to demonstrate that altering the capping group (Fmoc or acetyl) could alter the adopted assembly, either adding or removing a dimension of stronger intermolecular interactions, such that predominantly one-dimensional crystals and gels could be formed from acetyl capped peptides without the presence of Fmoc-Fmoc interactions. This also demonstrated that polyproline II helix was retained with different capping groups and degrees of functionalisation, demonstrating how this peptide helix can be utilised as a highly predictable ligand for the rational design of supramolecular constructs.

In chapter 3 our investigation into the use of polyproline helices as supramolecular building-blocks moved towards more complex peptides and supramolecular constructs. Herein, the focus was on synthesising functionalised polyprolines with coordinating groups to form metal-peptide bonds, constructing metal-organic frameworks, cages, and assemblies. The initial introduction of the chapter focuses on the background of metal-organic assemblies and recent examples constructed from bioinspired building blocks such as peptides, with a few examples utilising polyproline-based ligands highlighting the growing interest in these materials and their potential. Previous work from within the research group in the synthesis of a carboxylic acid functionalised oligoproline was also highlighted with preliminary results suggesting the

formation of nanoparticles from assembly with zinc salts, suggesting a promising avenue of investigation.

The initial experimental work in this chapter focused on the synthesis of non-natural functionalised proline monomers to incorporate the desired functional groups. The first monomer synthesised was a 4-position *trans*-(-OCH_2COO-tBu) functionalised Fmoc-proline while further attempts were carried out to synthesise alternative sidechains with aromatic carboxylic acid groups. While the synthesis of the alkyl carboxylate functionalised proline was successful, several attempts to incorporate aromatic groups, to reduce the level of flexibility and increase the directionality of any supramolecular interactions, faced setbacks especially using the Mitsunobu reaction and several different reaction conditions and starting materials with different protecting groups were attempted. Successfully incorporating a benzoic acid functional group was eventually achieved *via* carbon-carbon cross-coupling using the Suzuki reaction, which had the added benefit of incorporating a double bond in the proline to create an extended conjugated system such that the functional group is both rigid and linear with the proline ring, confirmed by single-crystal x-ray diffraction of the monomer. While this also reduced the length of the sidechain by removing the previous ether linkage, thus creating a highly directional functional group on the proline monomer. Subsequently, synthesis of an alternatively functionalised proline monomer was successfully carried out *via* the Mitsunobu reaction to achieve a pyridine-functionalised monomer, frequently used as a coordinating ligand in metal-organic assembly.

Thus, with these monomers in hand the next section of the chapter focused on their incorporation into a variety of peptide sequences and the rational design of these building-blocks to achieve specific types of supramolecular assembly. Peptide synthesis was carried out using Fmoc-based solid-phase peptide synthesis on an automated peptide synthesiser (Liberty Lite CEM), allowing the rapid synthesis of high purity peptides. Initially a series of carboxylic acid functionalised peptides were synthesised incorporating the -OCH_2COOH functionalised proline monomer. These were functionalised on all three helical faces with the goal of forming nanoparticles based on an extended metal-peptide framework, driven by preliminary results showing nanoparticle formation from an oligoproline hexamer with three carboxylic acid groups. From the synthesis of a peptide hexamer and three 13-chain polyprolines with varying degrees of functionalisation and positioning of these groups we were able to demonstrate the synthesis of several zinc-peptide assembly topologies such as nanoparticles, gels, and chain-like aggregates. These were determined by several methods of analysis: dynamic light scattering (DLS), *trans*mission electron microscopy (TEM), scanning electron microscopy (SEM), and atomic force microscopy (AFM). Combined with controlling the helicity of the longer chain peptides by changing the solvents and altering the base used we could demonstrate how the adopted topology could be controlled *via* several factors and, with further knowledge of these systems, the ability to use these polyprolines as versatile ligands in metal-organic constructs can be easily envisioned.

Consequently, the next part of the chapter looked at simpler peptide ligands extracted from the design of one of the previous 13-chain peptides with six carboxylic acid sidechains with varying spacings on each helical face, hypothesised to induce facial selectivity, such that this hypothesis could be investigated and to further elucidate the self-assembly principles. Three peptides were therefore synthesised to mirror each face of this peptide with two carboxylic acid groups, varying the capping groups to aid differentiation between the peptides from NMR and mass. With these peptides complexation was carried out with zinc as previously, with goal of forming discrete cages that could be studied *via* NMR and mass spectrometry. We successfully demonstrated that complexation to form discrete structures by only functionalising a single

helical face to prevent the formation of an extended network was successful. However, elucidation of further details of these structures were hampered by the complexity of the NMR spectra and decomposition of the complex during mass spectrometry analytical methods. Attempts to crystallise a smaller 7-chain peptide-copper cage to directly determine the structure of the construct was also unsuccessful. However, these complexes hold promise of forming discrete chiral nanocavities with potential applications for enantioselective catalysis. As such further investigation is required to determine the successful application of the design principles to achieve the hypothesised selectivity of these peptides to form complexes with only themselves and disfavour complexation with peptides with different functional group spacings.

Finally, this chapter looked at the synthesis of a pyridine functionalised oligoproline tetramer, with *trans* -O-Pyr sidechains on the terminal prolines. This structure was based off a model derived from the previously crystallised oligoproline tetramer with terminal hydroxyprolines. The functionalised peptide could then be treated as a ditopic C-shaped ligand in the design of discrete metal-peptide cages with asymmetric nanocavities. This was achieved by reacting the peptide with palladium salts $Pd(en)(NO_3)_2$ and $Pd(ACN)(BF_4)_2$ with the aim of forming a peptide dimer linked by two palladium centres for the former and a peptide tetramer, linked in the same manner, for the latter. The Pd(en) salt retains the ethylenediamine ligand after complexation limiting the number of free coordination sites.

Thus, with the synthesis of these peptide-palladium complexes carried out, analysis out of the complexes was carried out to determine the adopted three-dimensional structure. Initially ^1H NMR analysis was carried out which clearly indicated complete complexation of the peptide starting material. 2D NOESY and ROESY NMR was then carried out to determine the conformation of the peptides in the structures *via* NOE interactions between protons through space. This suggested the peptides were in an anti-parallel arrangement due to NOE interactions between the pyridine hydrogens on each peptide terminus and that a single major product was formed. However, subsequent high-resolution mass spectrometry analysis of the peptide complexes indicated that the expected cage-like complexes had not formed, as half the expected mass was found. Instead these peptides were acting as monochelating ligands to a single palladium, with the length and flexibility of the pyridine sidechain allowing both groups to adopt a geometry to coordinate to the same metal centre. Thus, the complexes formed a “bowtie” like structure with the $Pd(ACN)(BF_4)_2$ complex forming a peptide dimer. Despite this setback, this clearly evidences how these peptides can be used to form discrete metal-peptide complexes, with the chirality of the peptide helix favouring a single isomer of the product, instead of forming a 50:50 mixture of anti-parallel and parallel peptides for the dimeric complex. However, further optimisation of the peptide functional groups is required to prevent this type of chelation and drive the formation of the desired complex with a nanocavity for use in catalysis.

Chapter 4 presents the experimental work from the investigations in chapter 2 and then chapter 3, with the relevant synthetic steps, analytical methods, and the resulting data. More routine analytical data with less relevance to the findings of the work, such as NMR spectra and mass spectra of synthesised compounds, has been placed in the appendices.

5.2 Future Outlook

The study investigating the use of minimalistic oligoprolines to form tuneable self-assembled supramolecular peptide frameworks demonstrated the potential of these porous materials to act as an enantioselective host towards the polar 1-phenylethan-1-ol molecule. With this promising ability, further work should look to investigate the applications for the chiral channels in chemical separations. The use of the crystalline peptide in a chiral SPE cartridge to separate

further test racemic mixtures would further demonstrate the applicability of the peptide framework, while the enantioselectivity, time efficiency, and stability of the framework could be tested. This could also be expanded to test the new porous crystalline frameworks synthesised with hydroxyprolines, different capping groups, and *D*-proline, to determine the tuneability of the pore spaces and how this affects the enantioselectivity.

Furthermore, investigation of the host-guest interactions and the degree of adaptation of the framework to include the new guests should be carried out, ideally by single crystal analysis with encapsulated molecules and computational methods. Thus, new porous frameworks could be more accurately targeted towards specific host-guest interactions, allowing tuning of the selectivity.

The subsequent investigation into the use of functionalised polyproline helices to synthesise metal-peptide assemblies demonstrated the ability to form nanoparticles of varying sizes with potential in molecular *trans*port and chemical separations. However, further characterisation of these assemblies is required to determine the porosity of the nanoparticles, their tolerance to guest molecules, and cryo-TEM of each to demonstrate their internal structure. With this in hand studies can be carried out to analyse their capability to incorporate fluorescent molecules and whether these can be released in controlled manner. This can then be applied to test the interaction of the particles with cells, whether they are internalised, their cytotoxicity, and if encapsulated molecules previously impermeable to the cell membrane can be carried into a cell, which can be visualised by confocal microscopy. This would demonstrate whether these particles can be used as tuneable cell-delivery vehicles, having biomedical applications.

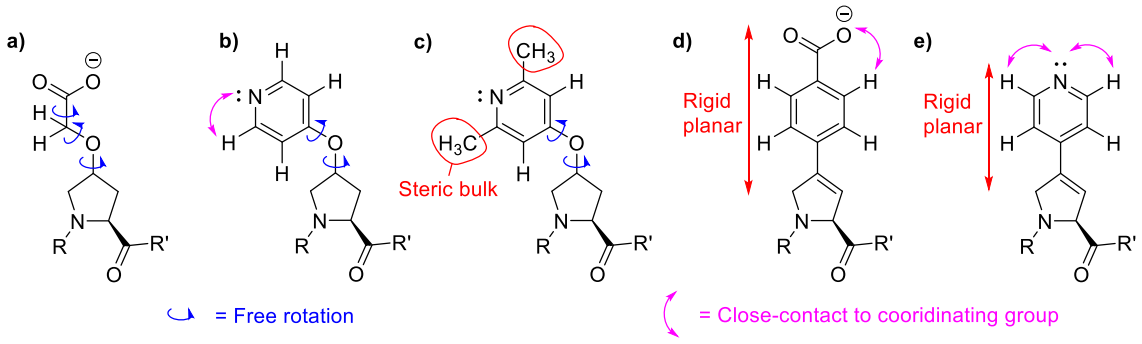


Figure 160 – Chemical structures of different functional groups on non-natural proline residues. Synthesised functionalised prolines with flexible linkers utilised in the synthesis of peptides for complexation reactions (a-b), lutidine functionalised proline with added steric bulk adjacent to the amine coordinating group (c), synthesised benzoic acid functionalised proline to be utilised in peptide synthesis, with a rigid planar functional group for highly directional coordination bonds (d), theorised structure of a pyridine functionalised proline with a direct C-C bond to the pyrrolidine ring creating a shorter, rigid and planar functional group

However, further optimisation of the peptide functionalisation is required as further studies with a pyridine functionalised monomer (Figure 160b) demonstrated that longer, flexible sidechains can chelate to the same metal node affecting the predictability of the adopted structure. As such futures studies should aim to incorporate the benzoic acid functionalised proline derivative (Figure 160d), coupled by direct C-C bond coupling to the pyrrolidine ring. This shorter, more rigid, group should allow for the synthesis of more predictable intermolecular interactions and assist with the determination of the adopted structure and assembly principles. This can also be applied to the synthesis of discrete cages from similarly functionalised peptides to determine the facial selectivity of the complexation of these peptides and conformation of the peptides within these structures. The *ortho* protons adjacent to the carboxylic acid group on the benzene ring should also allow elucidation of NOE interactions between adjacent ligands around a metal centre to determine the adopted structure, as seen for the pyridine

functionalised monomer (Figure 160b+d). Thus, with a more thorough understanding of the structure and assembly principles these constructs can be more readily designed and reproduced.

The synthesis of palladium-polyproline complexes with pyridine coordinating functional groups was also demonstrated, however the desired structures were not synthesised due to the monochelation of the two pyridine groups. Thus, further studies into these constructs to achieve polyproline-based metallo-cages should focus on the optimisation of the peptide ligand to prevent the formation of intramolecular linkages. This could be achieved by additional steric bulk around the coordination amine by replacing pyridine with lutidine (Figure 160c), thus hindering the coordination of pyridines from the same peptide. Additionally, the flexibility and length of the functional group could be reduced which should hinder the adoption of the geometry required to satisfy the palladium coordination sites. This can be achieved using a similar route carried out for the synthesis of a benzoic acid functionalised proline previously (Figure 160d), by direct C-C bond coupling of the pyridine to the pyrrolidine ring (Figure 160e). This removes the ether linkage and forms a rigid, planar group that cannot reasonably monochelate to single metal centre from the terminal proline residues. Finally, the distance between the functional groups could be increased by synthesising a longer peptide with another helical turn, such as a 7-residue peptide. Therefore, the increased distance should be too great to bind to a single metal centre, while having the added benefit of a 3rd residue on the same helical face on Pro4 that may be utilised for tuning host-guest interactions if facing into any cavity created by the formation of a peptide cage. With these steps there is a clear path forward to develop the desired polyproline-based metallo-cages, to achieve the synthesis of tuneable, chiral nanocavities.

Appendices: Chapter 2

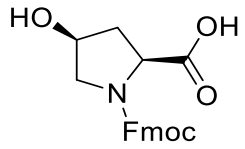
NMR experiments were recorded on a Bruker Avance II 400 MHz spectrometer in deuterated solvents.

High-resolution Mass-Spectrometry (ESI+) – **P_4** and **P_2HP** were separated on a Phenomenex 2.1 x 150 mm, 3.6 um, XB-C18 Aeris Widepore column at 50 °C on a Waters H-Class Acquity UPLC. R_t = 10.49 min. Using a 0.1% formic acid/acetonitrile gradient: 5% B for 2.55 minutes then from 5% B to 95% B over 15 minutes, and held at 95% B for 2 minutes. A is 0.1% formic/Water, B is 0.1% formic/Acetonitrile. Flow rate is 0.25 mL/min. The flow is directed into the electrospray source of a Waters G2-Si mass spectrometer, operating in positive ion mode, at 2.5 kV and mass spectra recorded from 100-3000 m/z. Data was analysed with Waters Mass Lynx software.

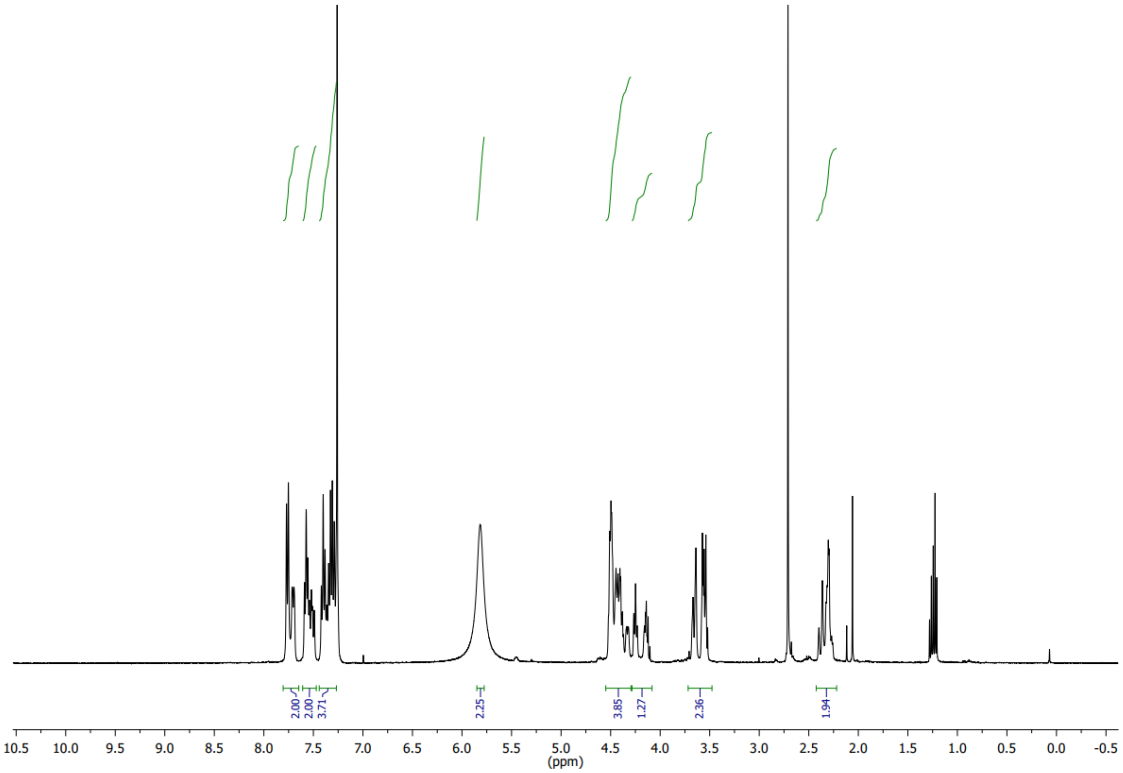
All other peptides, unless otherwise stated, were separated on a 2.1 x 150 mm, 3.6 um, XB-C18 Aeris Widepore column, from Phenomenex, at 30 °C on an Agilent 1100 HPLC. Gradient: 5% B for 5 minutes then from 5% B to 100% B over 25 minutes, and held at 100% B for 5 minutes. Where A is Water (0.1% formic acid) and B is acetonitrile (0.1% formic acid). Flow rate is 0.2 mL/min. The flow was directed into directed into the electrospray source of a Bruker micrOTOF-QII mass spectrometer, operating in positive ion mode, at 4.5 kV and mass spectra recorded from 150-3000 m/z. Data was analysed with Bruker's Compass Data Analysis software.

LCMS - Peptides **P_3H**, **$AcHP_2H$**, ***cis*-HP_2H** and **AcP_4**, as well as all monomers, were separated *via* RP-HPLC using a HiChrom KR100 5C18 5263 column at 40 °C on a Dionex UltiMate 3000. Gradient: 5% B for 5 minutes then from 5% B to 100% B over 20 minutes, and held at 100% B for 5 minutes. Where A is Water (0.1% formic acid) and B is methanol (0.1% formic acid). Flow rate is 1.0 mL/min. Wavelength: 225 nm. The flow was directed into directed into the electrospray source of a Thermo Scientific MSQ Plus Mass Detector, operating in positive ion mode, at 75 kV and mass spectra recorded from 100-2000 m/z.

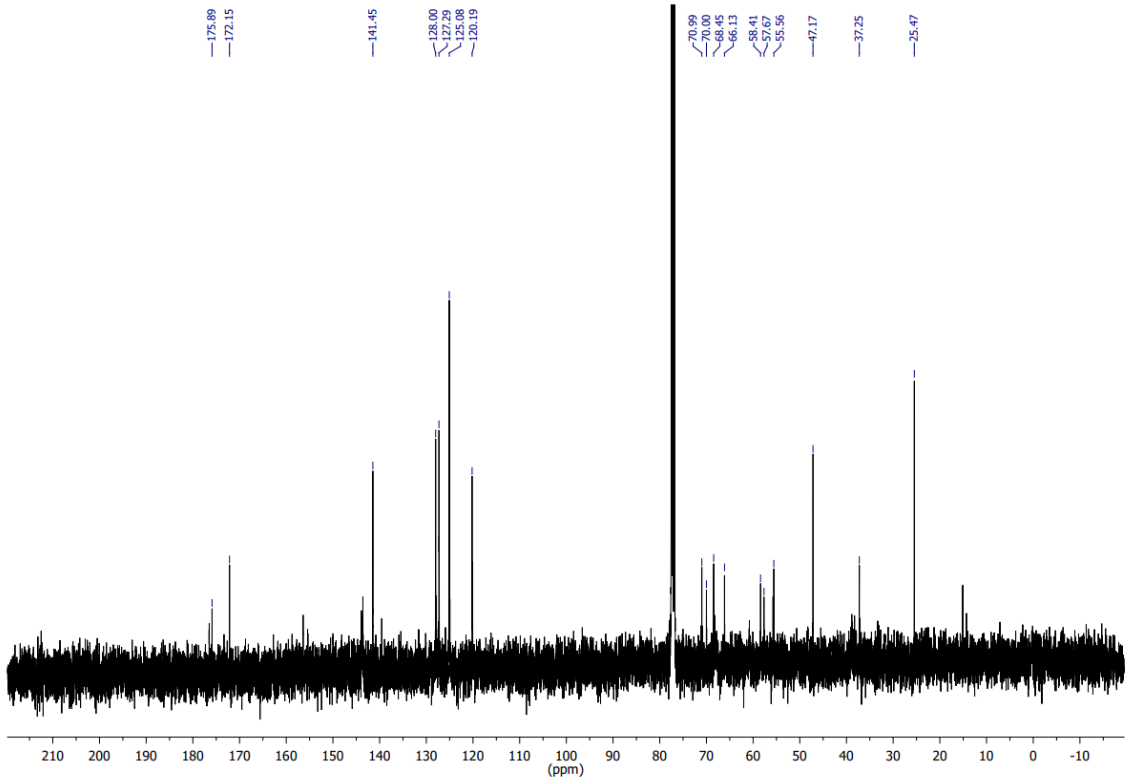
6.1 Synthesis of (2S,4S)-1-(((9H-fluoren-9-yl)methoxy)carbonyl)-4-hydroxypyrrolidine-2-carboxylic acid, 1:



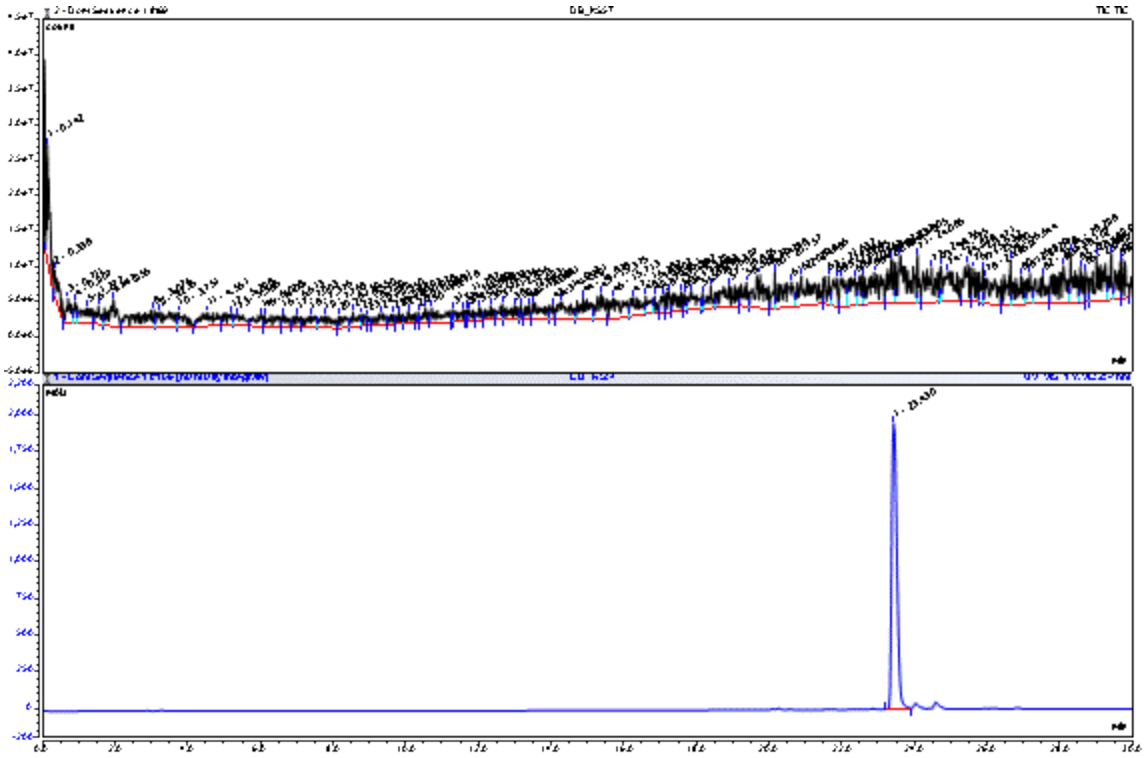
1



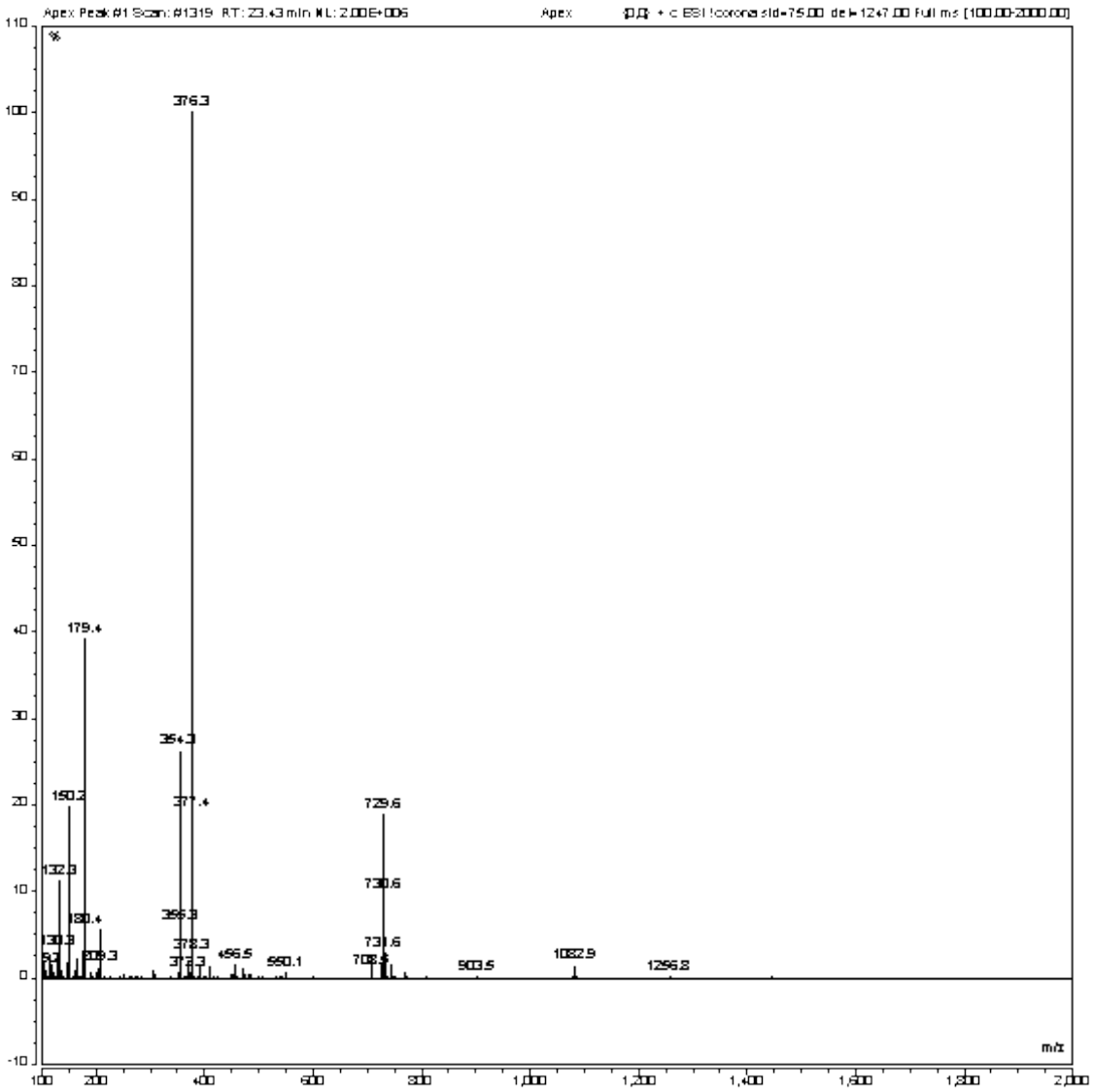
^{1}H NMR of compound 1 (400 MHz, $CDCl_3$)



^{13}C NMR of compound 1 (101 MHz, $CDCl_3$)

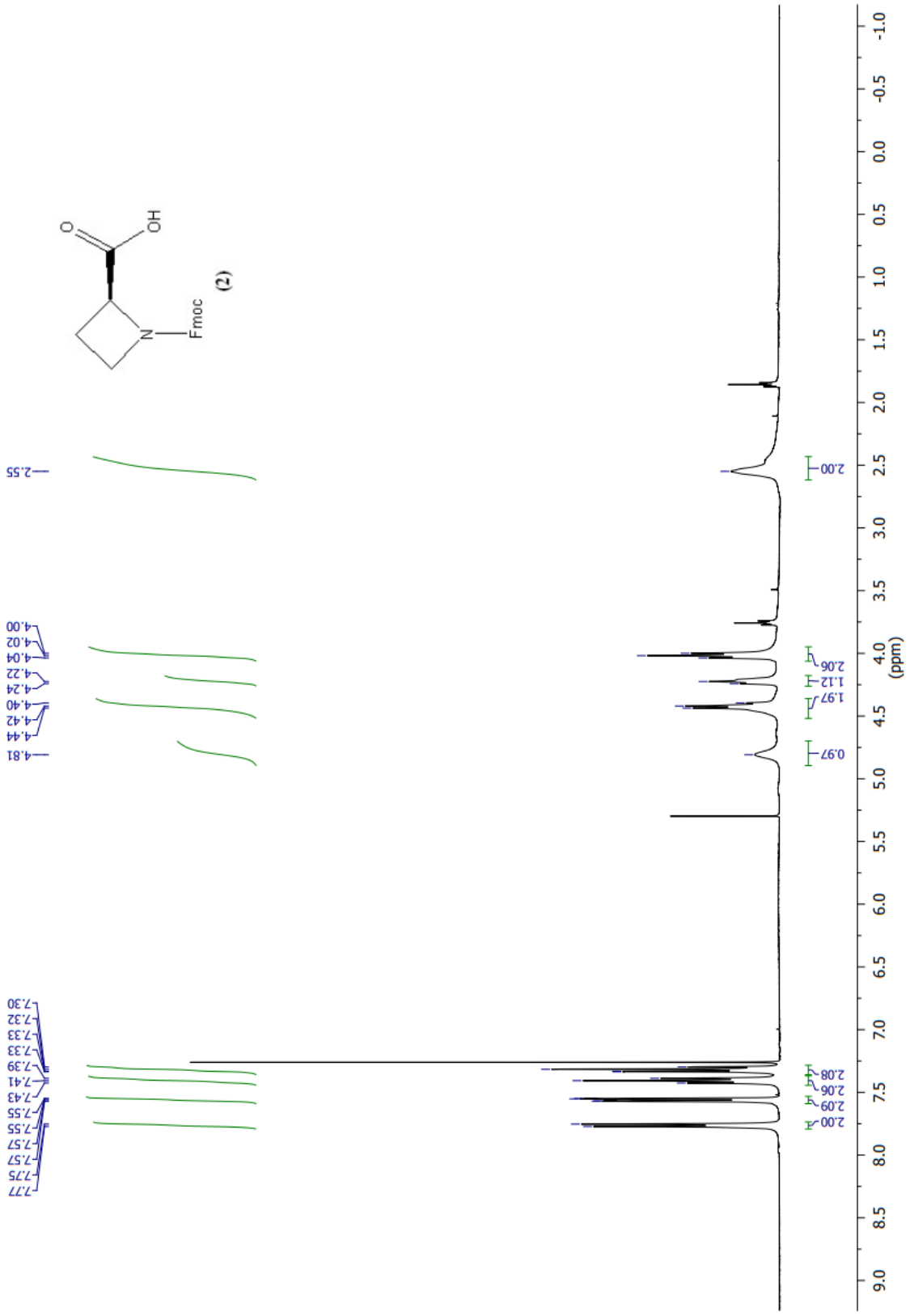


HPLC UV-Vis trace for compound 1, rt = 23.430 min, 254 nm

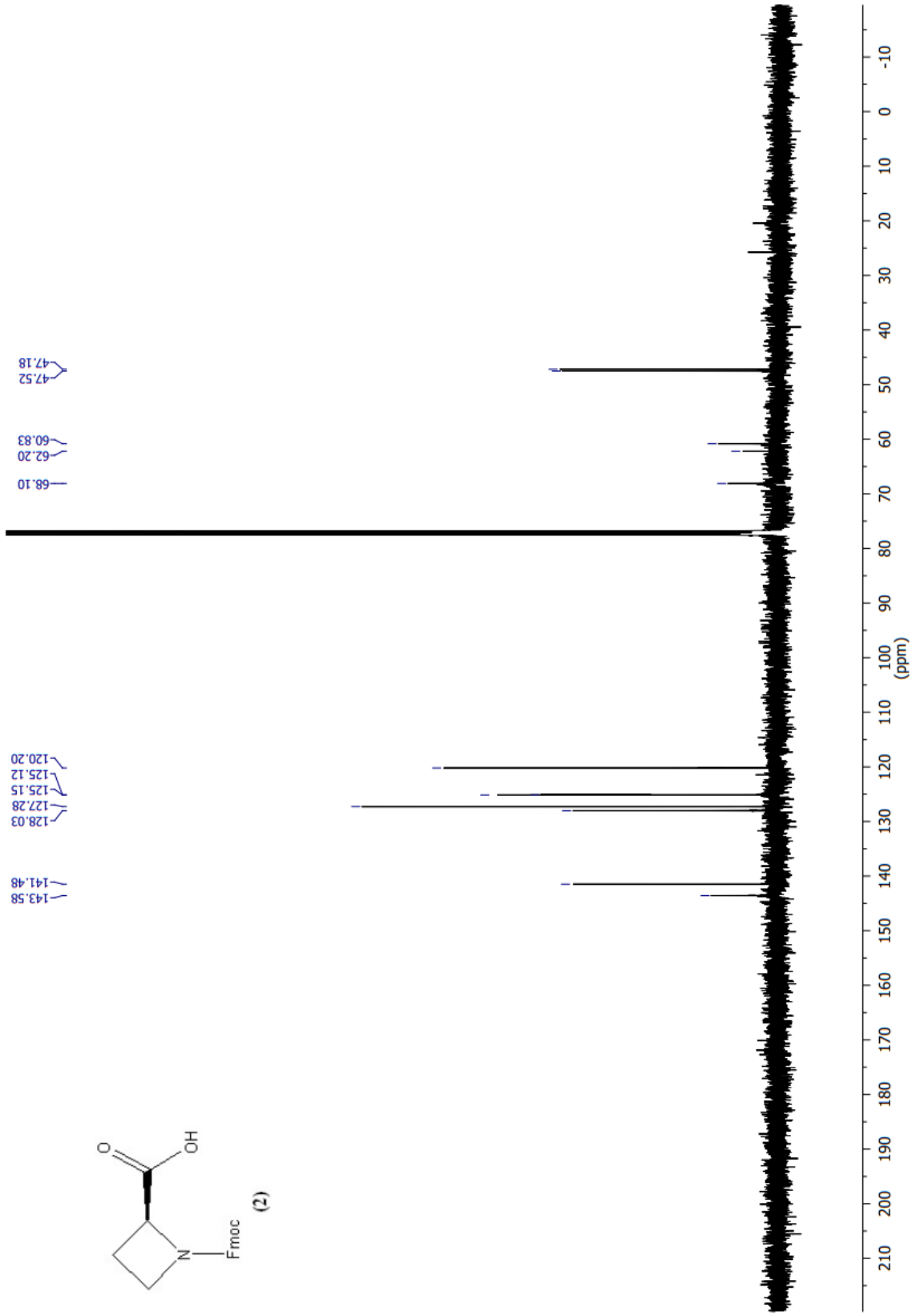


Mass spectrum (100-2000 m/z), compound 1, of peak at 23.430 min

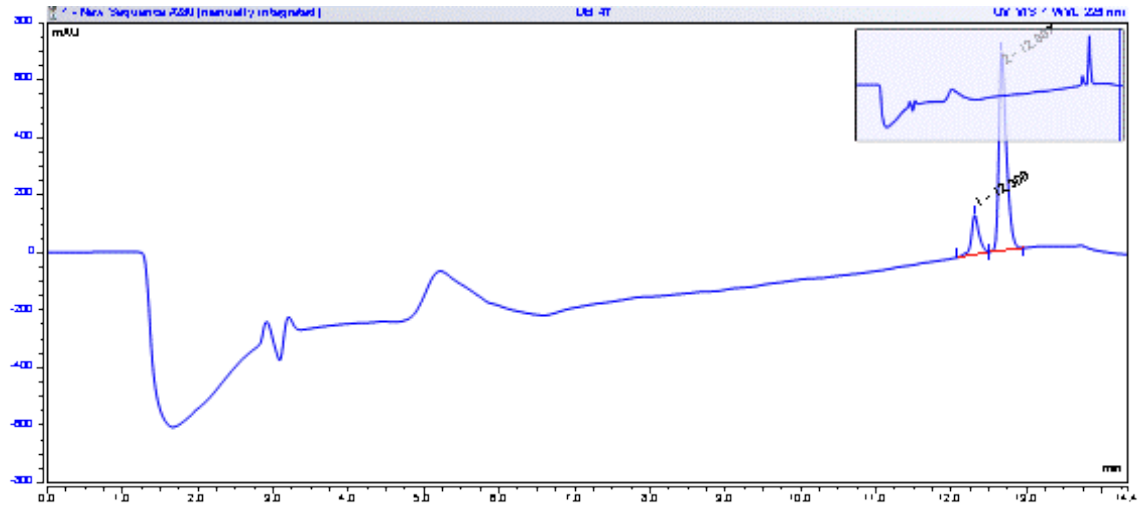
6.2 Synthesis of (S)-1-(((9H-fluoren-9-yl)methoxy)carbonyl)azetidine-2-carboxylic acid, 2:



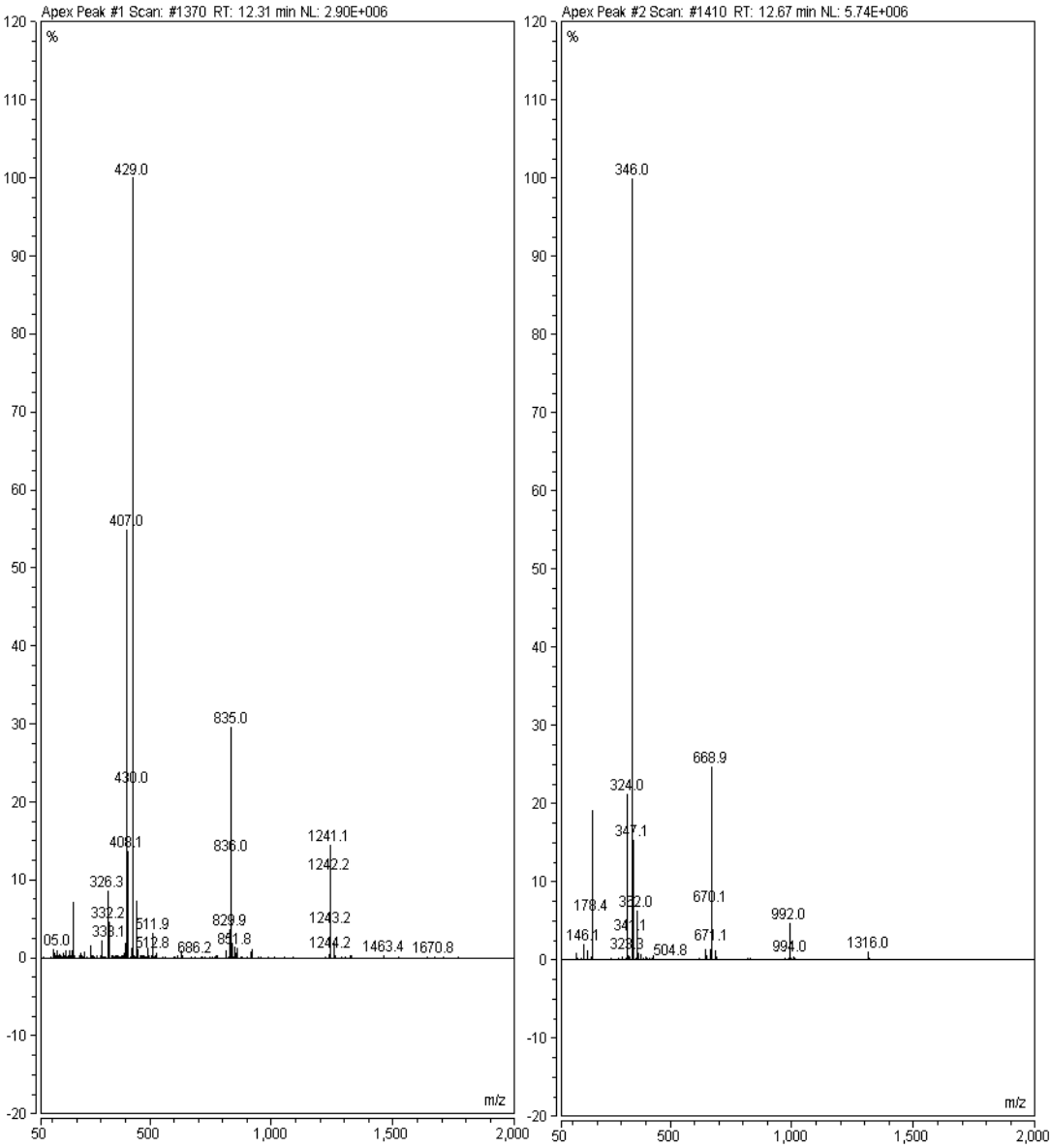
^{1}H NMR of compound 2 (400 MHz, $CDCl_3$)



^{13}C NMR of compound 2 (101 MHz, $CDCl_3$)

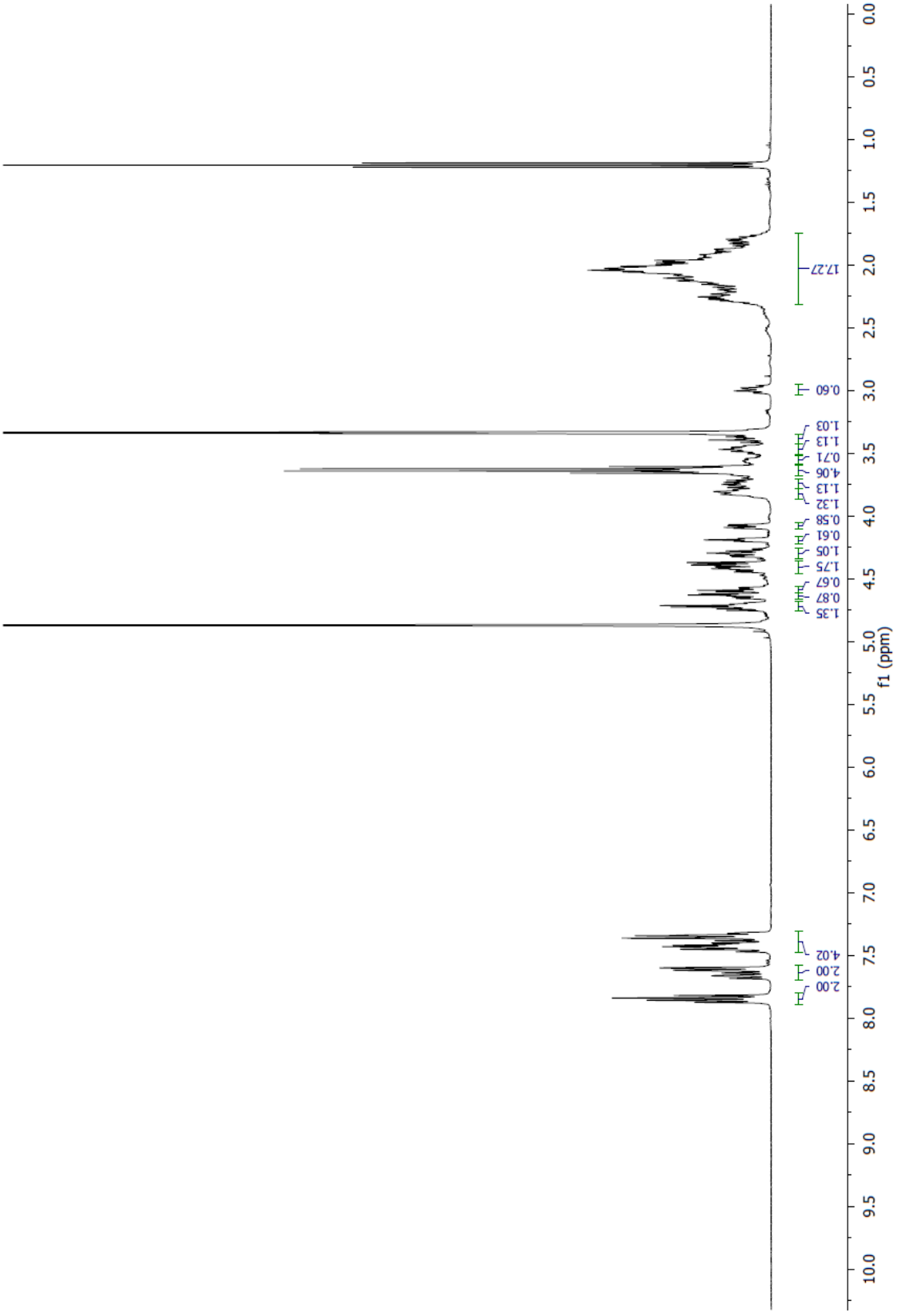


HPLC UV-Vis trace of compound 2, rt = 12.31 min, 12.67 min, 225 nm

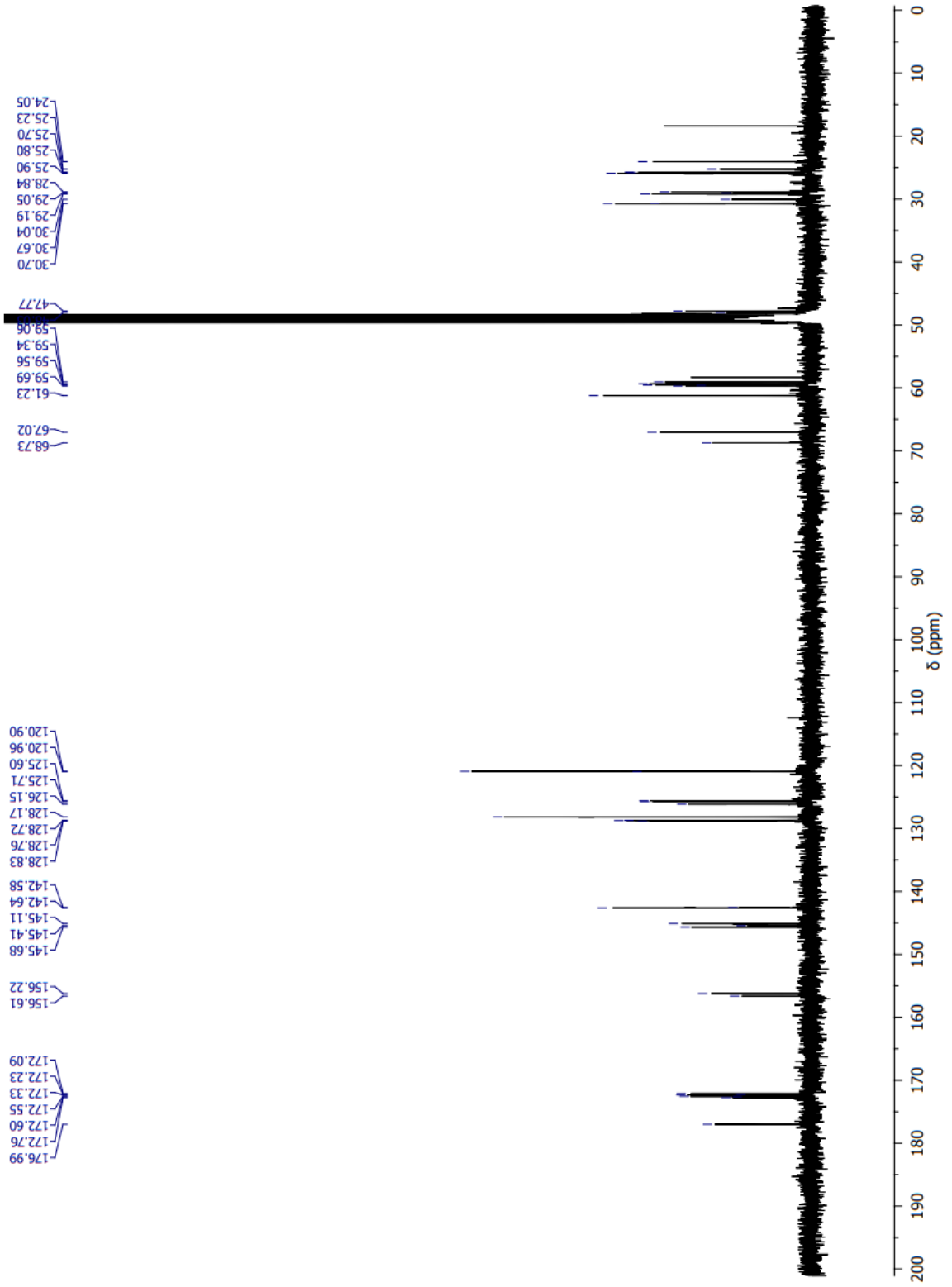


Mass spectrum (100-2000 m/z) of peaks at 12.31 (**2+Na_2CO_3**) and 12.67 min (**2**)

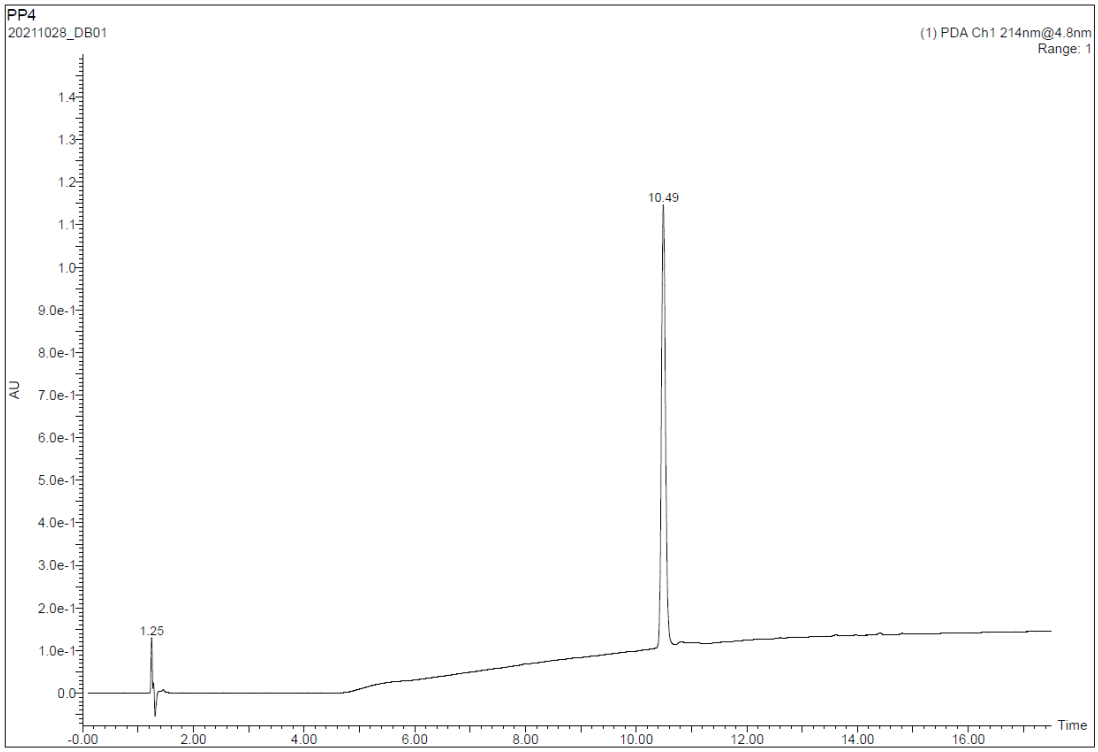
6.3 Synthesis of Fmoc-$(Pro)_4$-NH_2 (2):



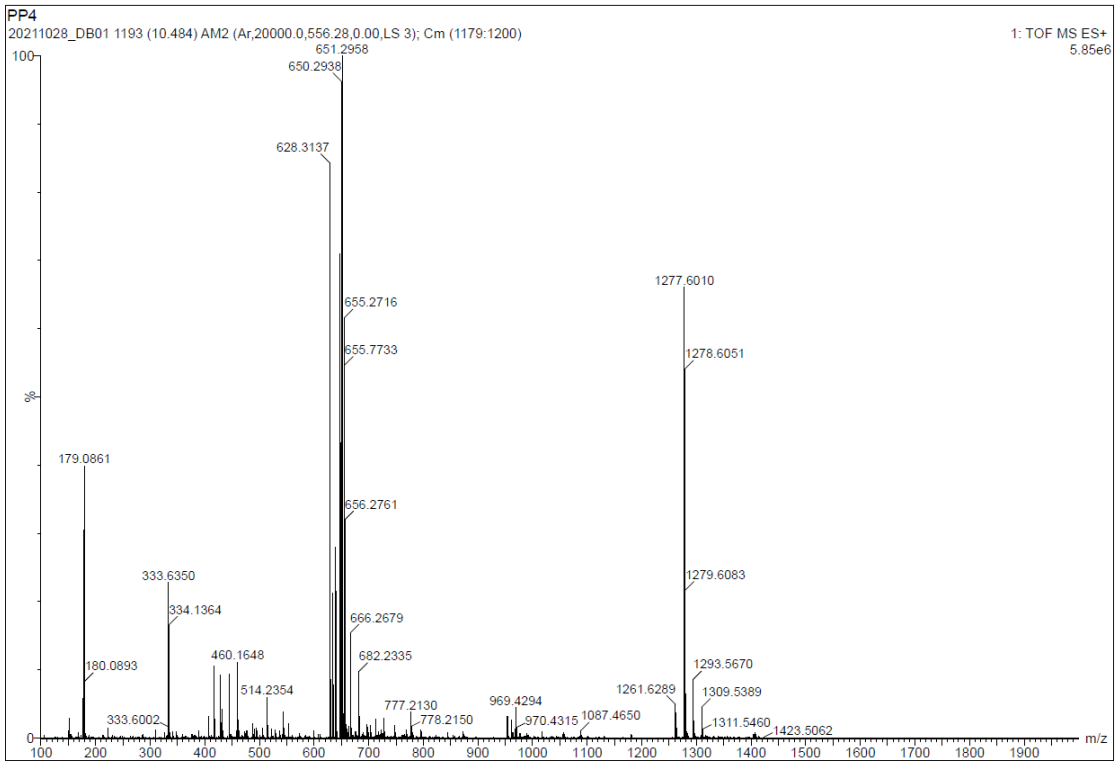
^{1}H NMR of **P_4 (400 MHz, MeOD)**



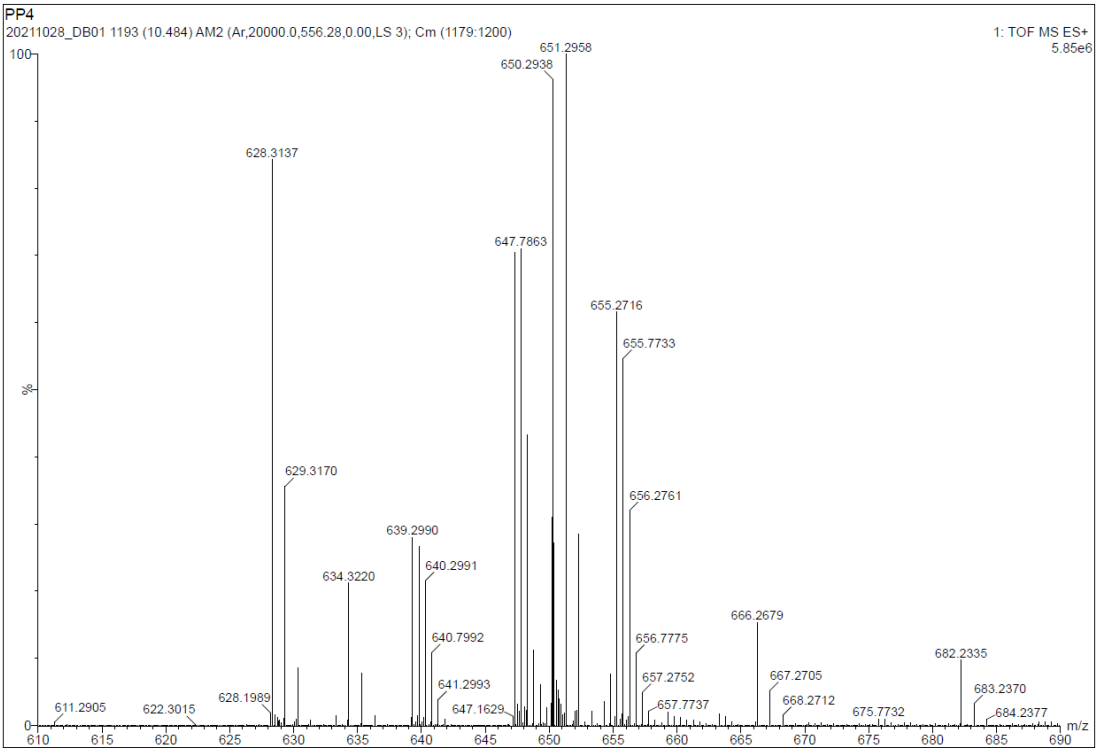
^{13}C NMR of **P_4 (101 MHz, MeOD)**



UHPLC UV-Vis trace of peptide **P_4**, rt = 10.49 min, 214 nm

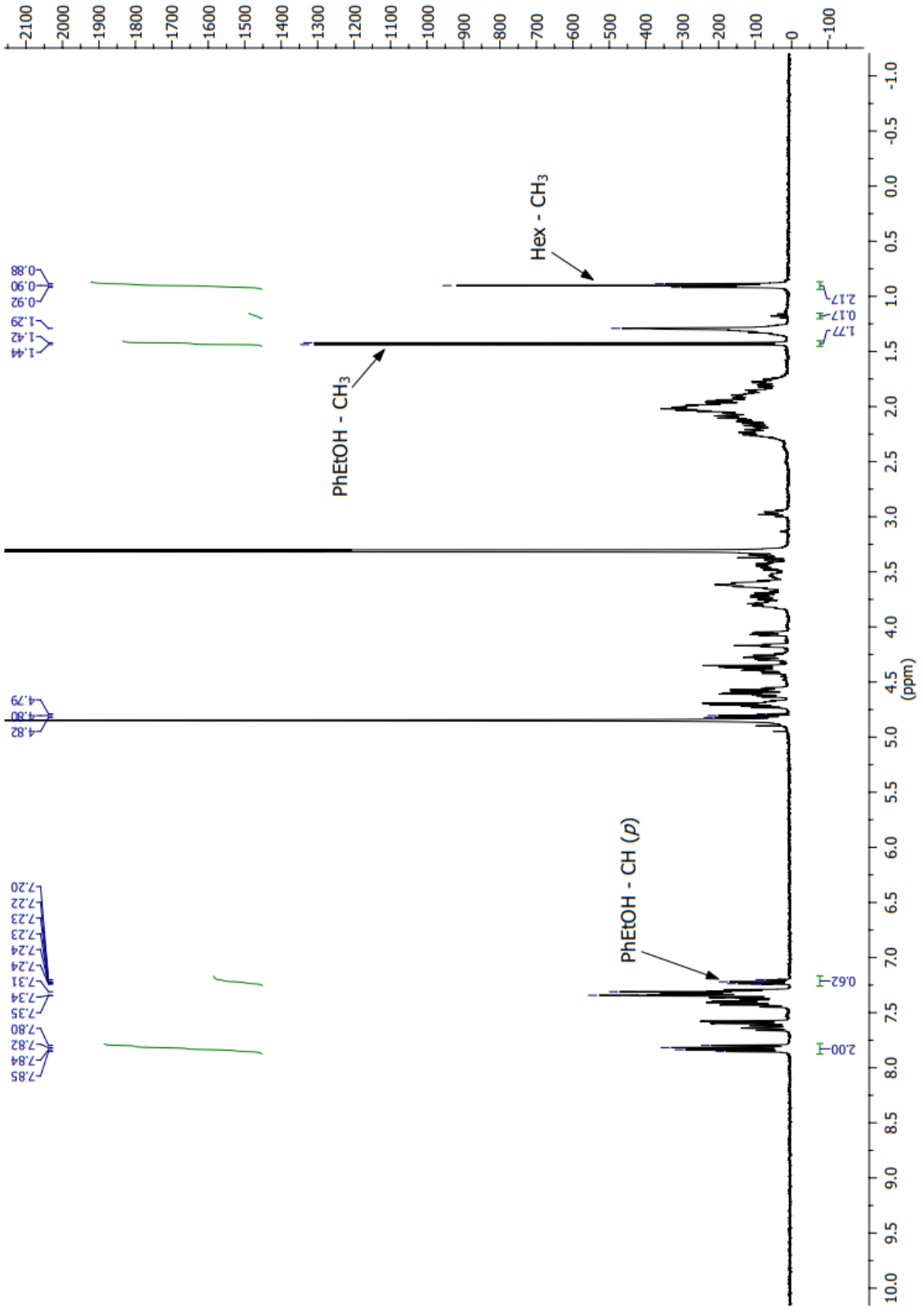


Mass spectrum (100-2000 m/z) of peak at 10.49 min, **P_4** UHPLC

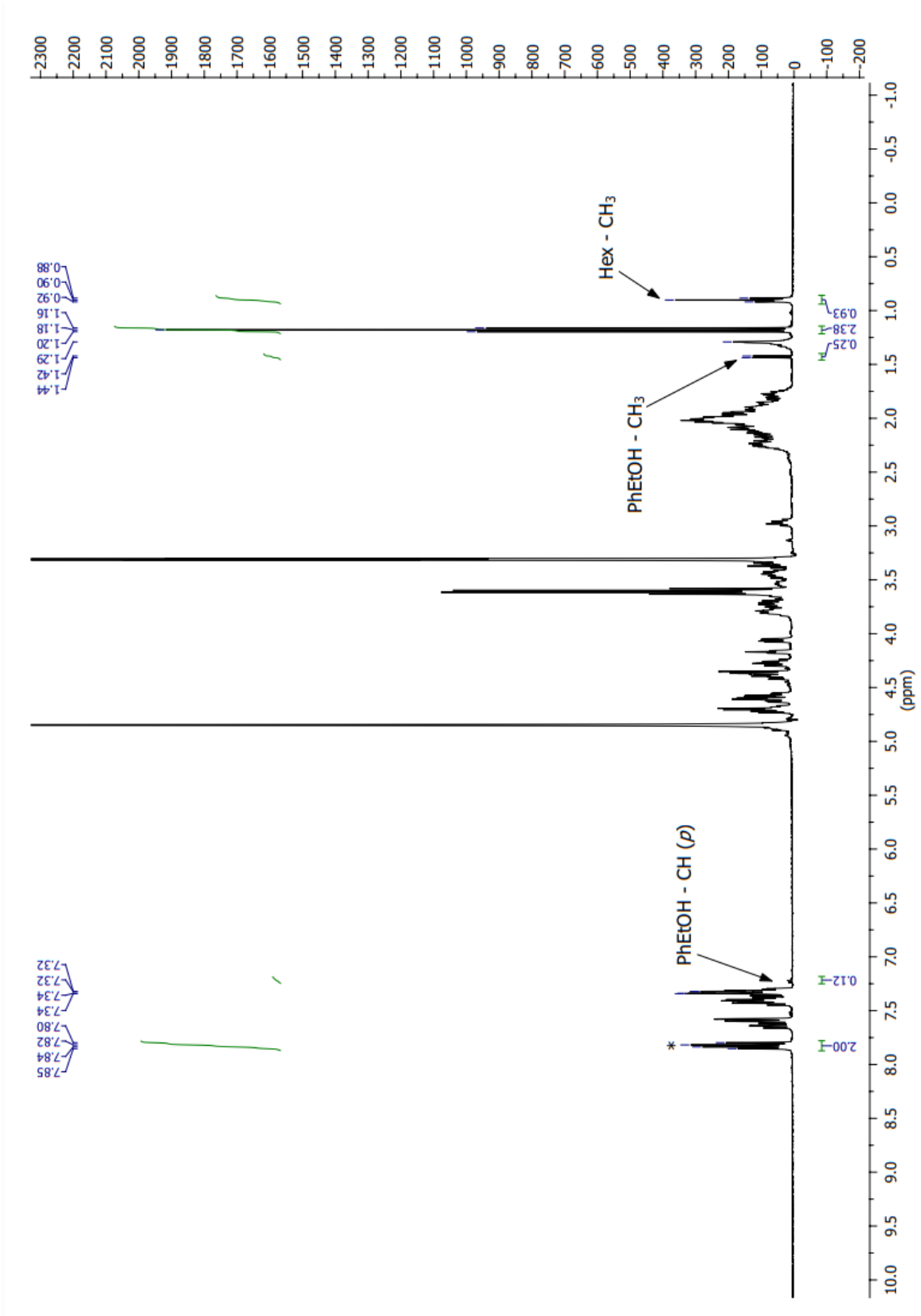


Mass spectrum (610-690 m/z) of peak at 10.49 min, **P_4** UHPLC

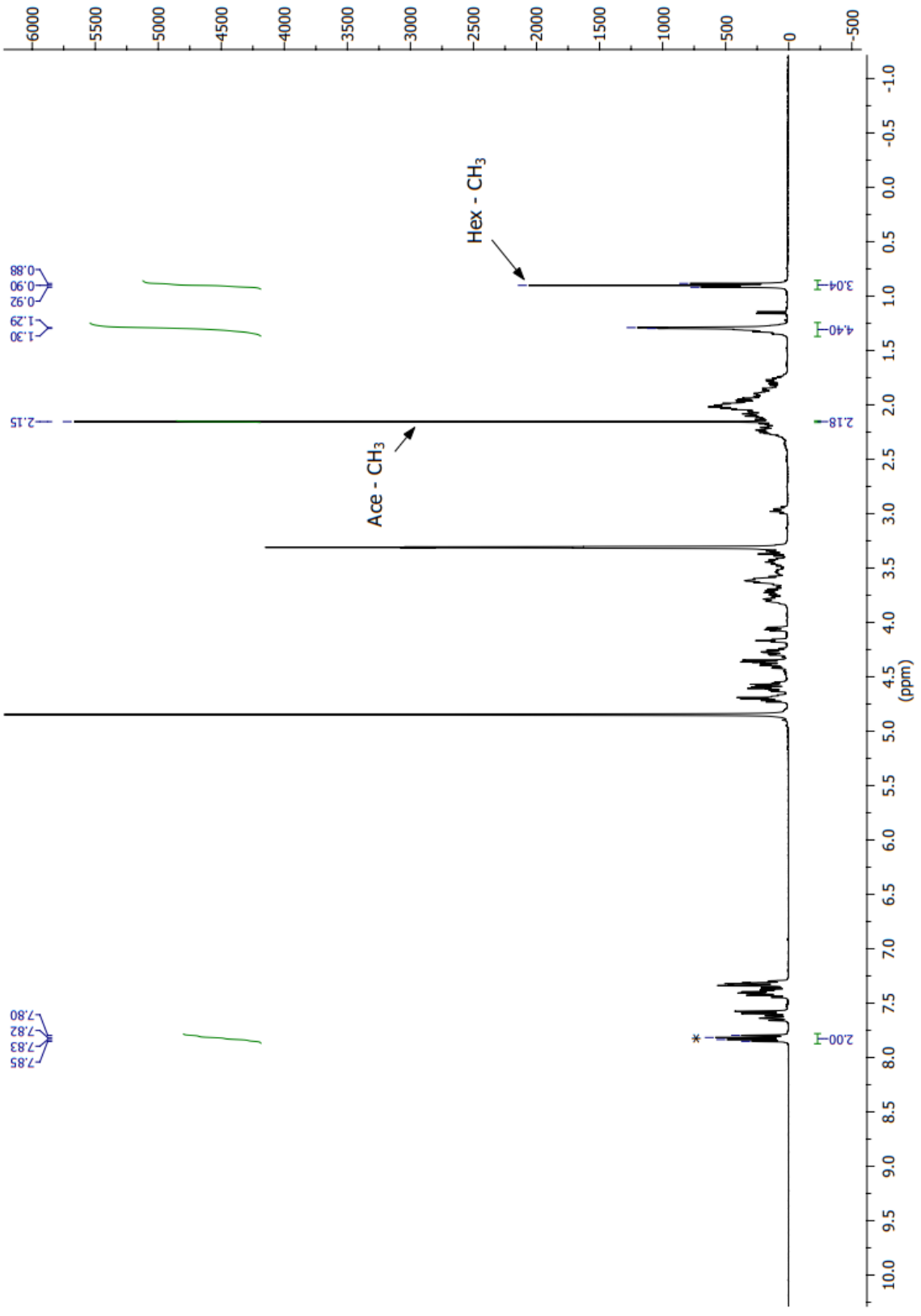
6.3.1 Experiments with Fmoc-$(Pro)_4$-NH_2 crystalline material (2):



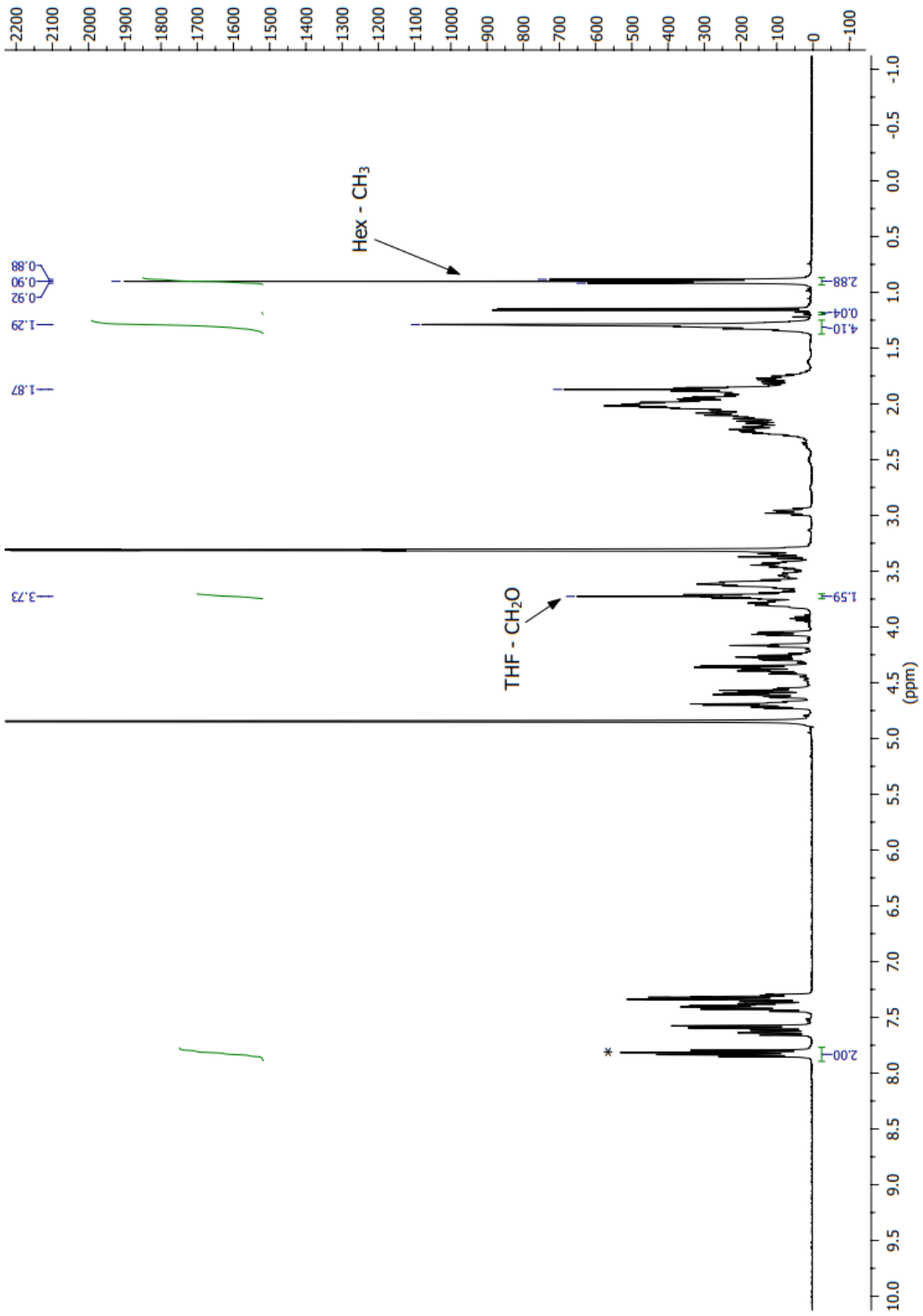
^{1}H NMR (400 MHz, MeOD) of **PP_4-SPF_{act}** after soaking in 1-phenylethanol (5%, 300 µL in hexane) for 16 h and washed with hexane (HPLC grade, x 5) before drying for 30 min at rt.



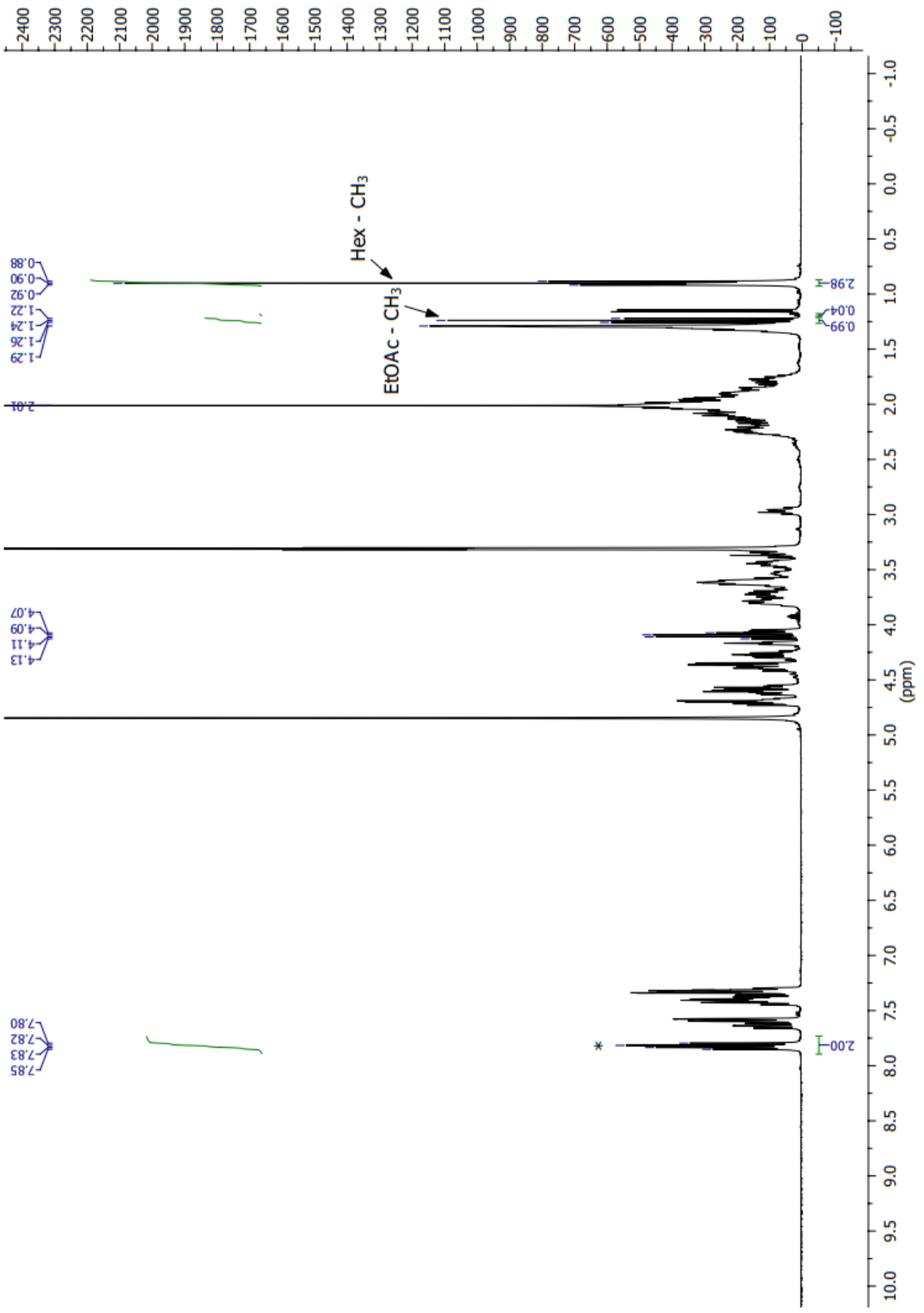
^{1}H NMR (400 MHz, MeOD) of **PP_4** after soaking in 1-phenylethanol (5%, 300 µL in hexane) for 1.5 h and washed with hexane (HPLC grade, x 5) before drying for 30 min at rt.



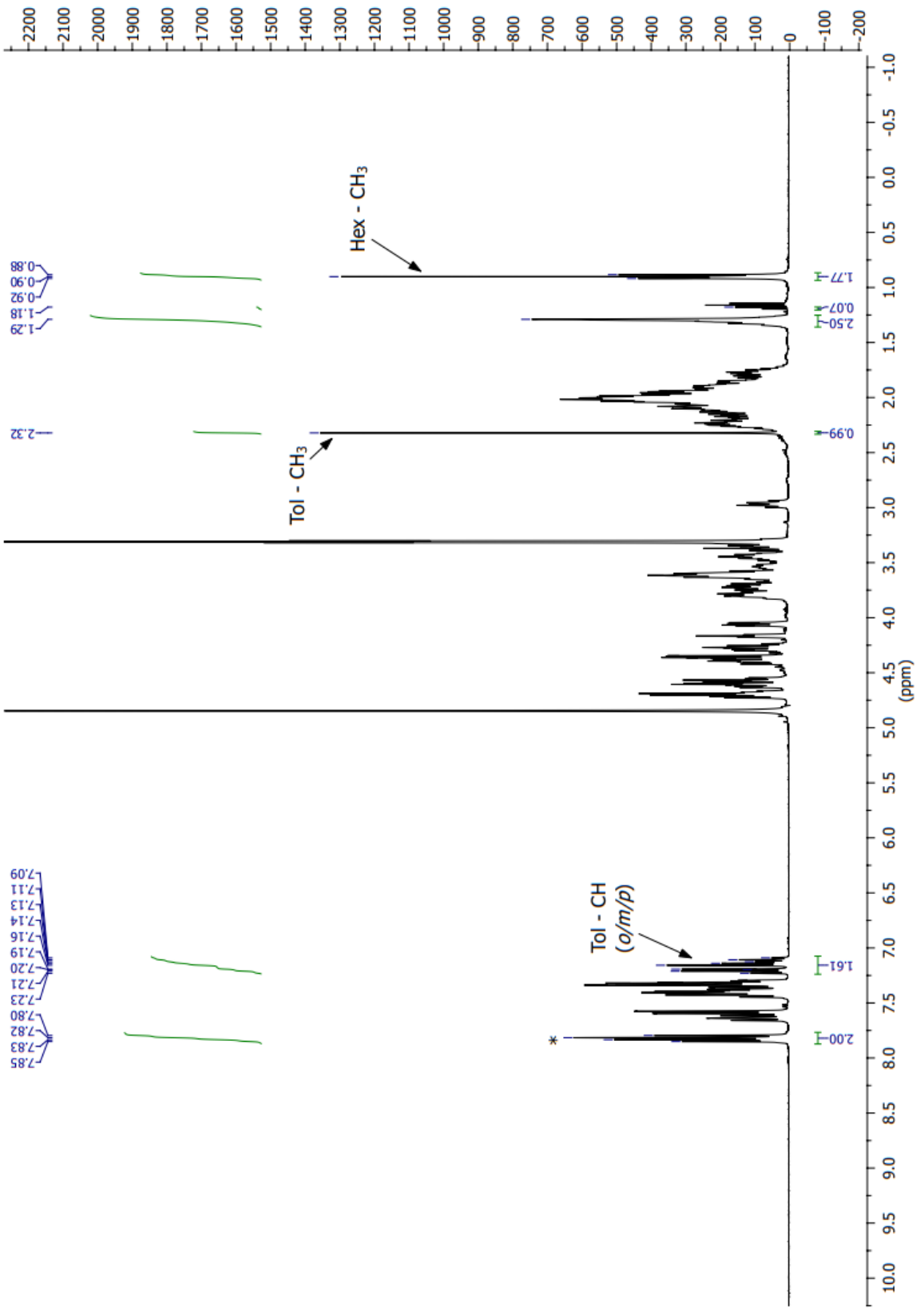
^{1}H NMR (400 MHz, MeOD) of **PP_4-SPF_{act}** after soaking in Acetone (5%, 300 µL in hexane) for 16 h and washed with hexane (HPLC grade, x 5) before drying for 30 min at rt.



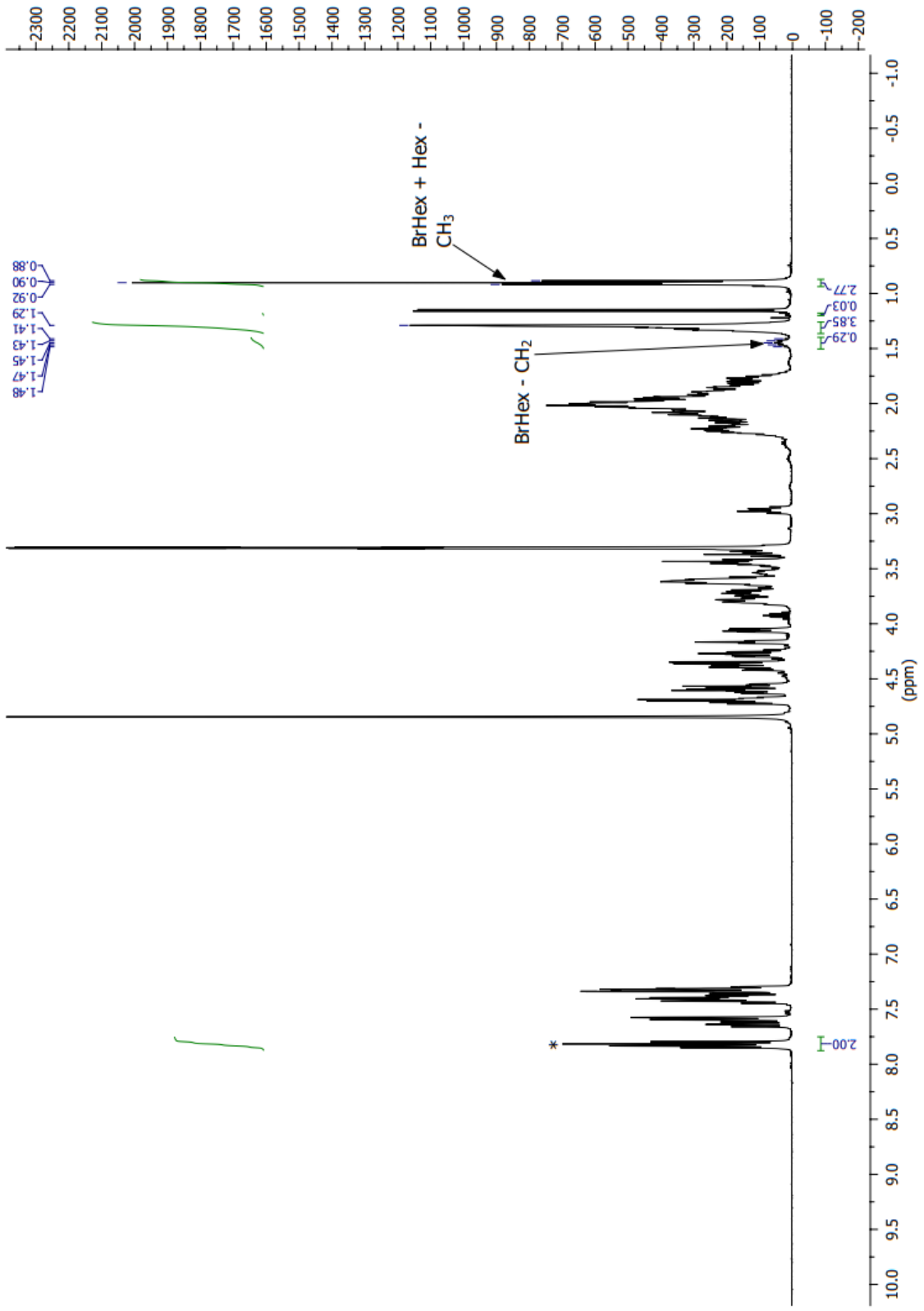
^{1}H NMR (400 MHz, MeOD) of **PP_4-SPF_{act}** after soaking in THF (5%, 300 µL in hexane) for 16 h and washed with hexane (HPLC grade, x 5) before drying for 30 min at rt.



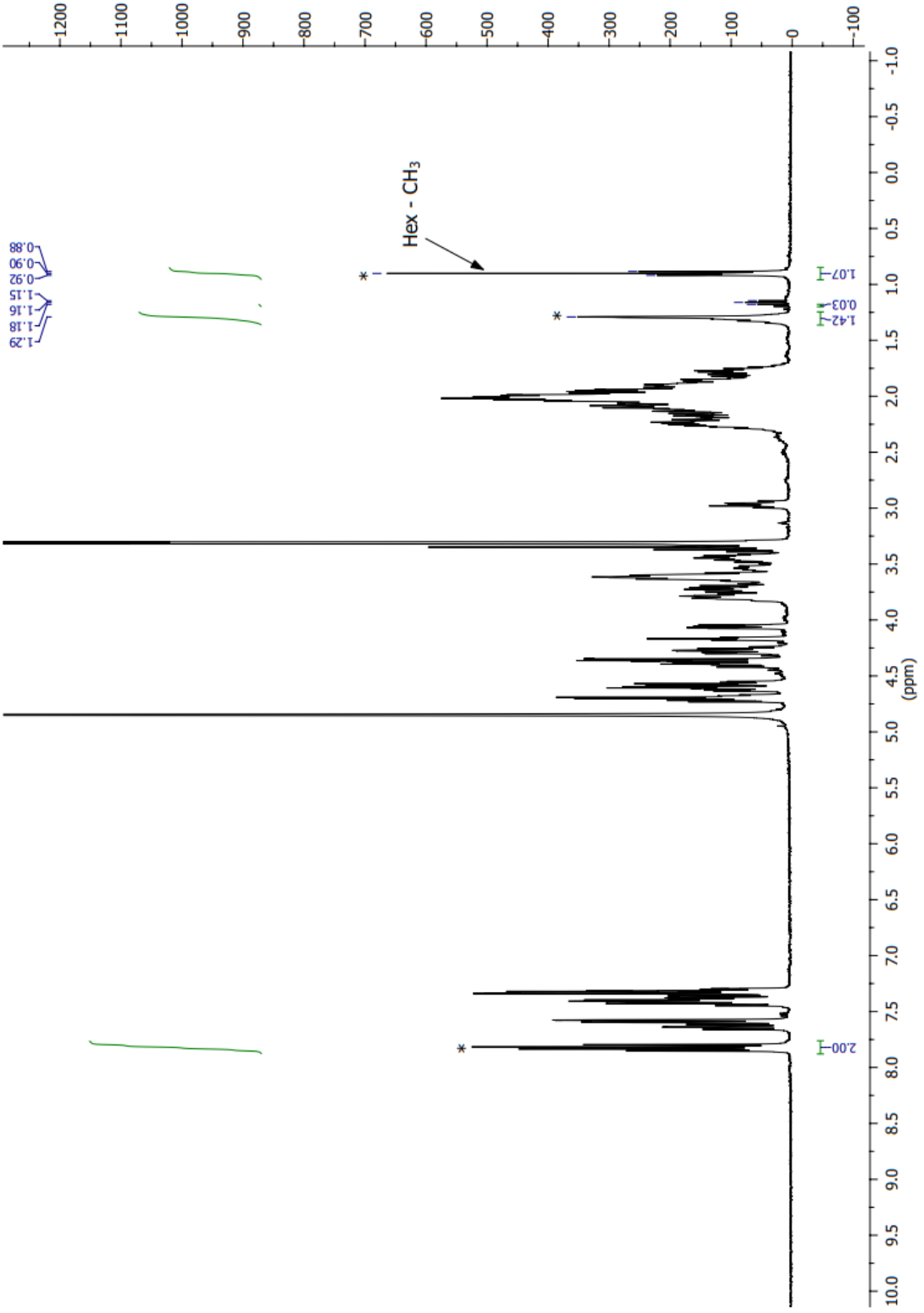
^{1}H NMR (400 MHz, MeOD) of **PP_4-SPF_{act}** after soaking in ethyl acetate (5%, 300 µL in hexane) for 16 h and washed with hexane (HPLC grade, x 5) before drying for 30 min at rt.



^{1}H NMR (400 MHz, MeOD) of **PP_4-SPF_{act}** after soaking in toluene (5%, 300 µL in hexane) for 16 h and washed with hexane (HPLC grade, x 5) before drying for 30 min at rt.

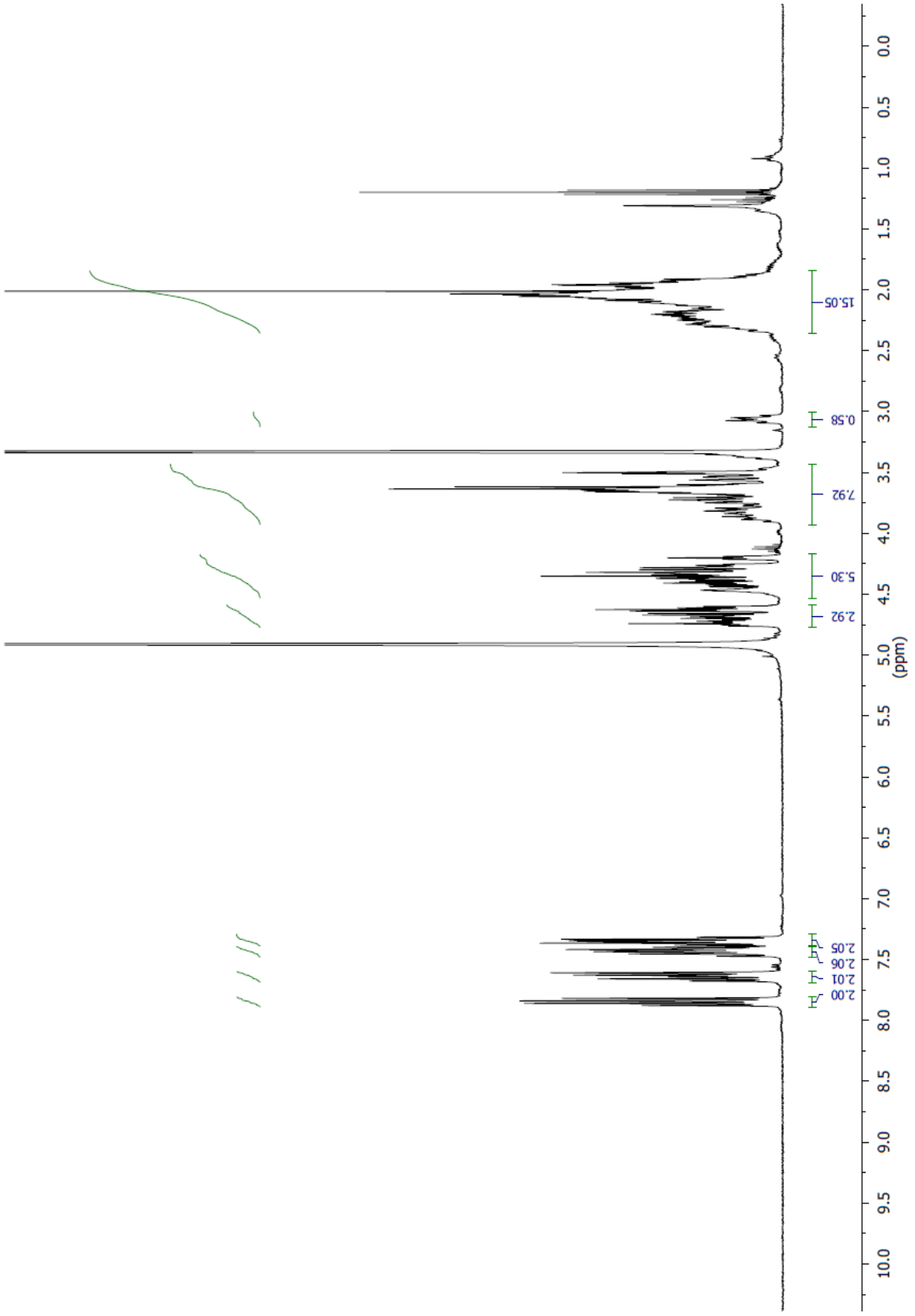


^{1}H NMR (400 MHz, MeOD) of **PP_4-SPF_{act}** after soaking in 1-bromohexane (5%, 300 µL in hexane) for 16 h and washed with hexane (HPLC grade, x 5) before drying for 30 min at rt.

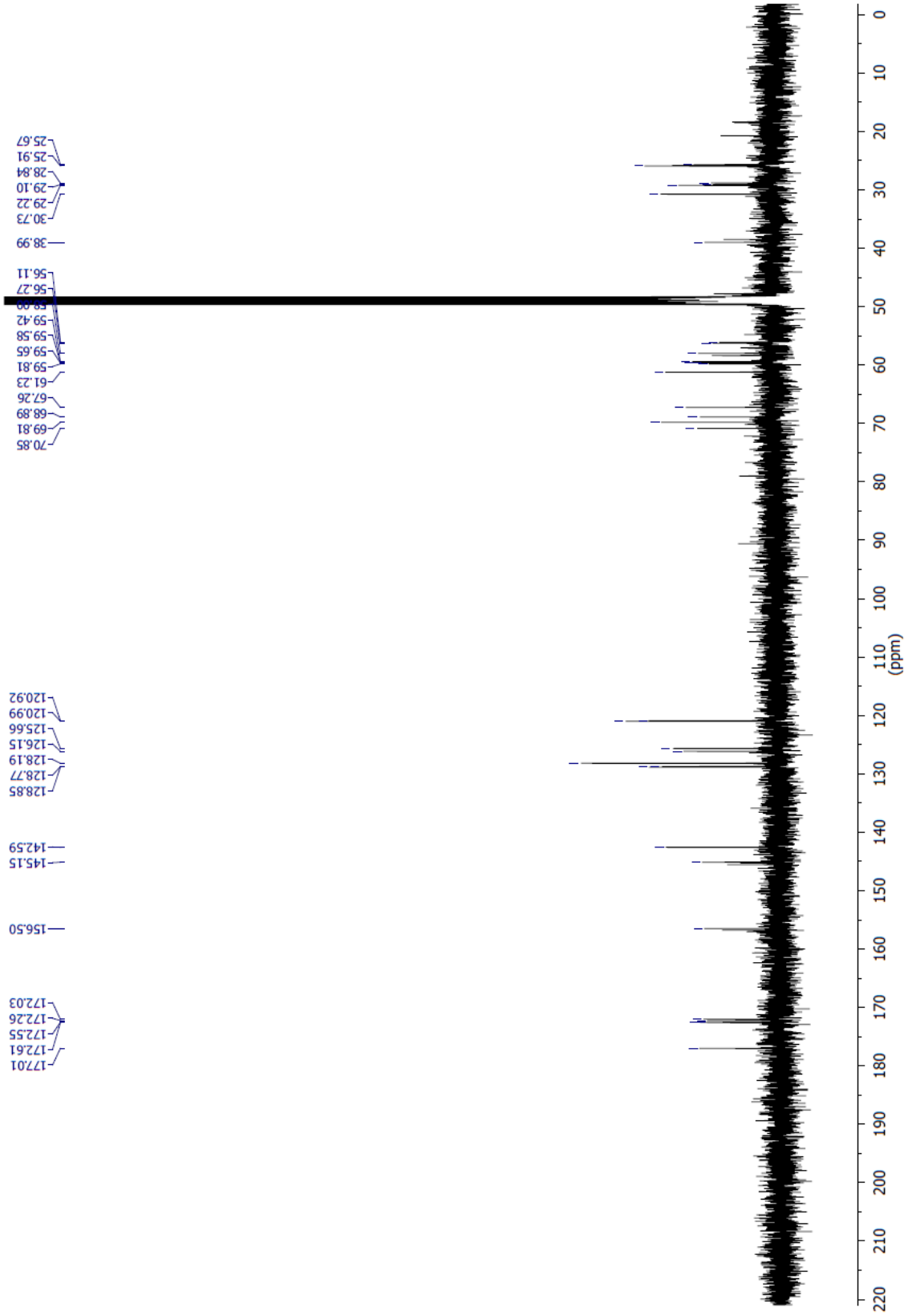


^{1}H NMR (400 MHz, MeOD) of **PP_4-SPF_{act}** after soaking in saturated iodine solution (300 µL in hexane) for 16 h and washed with hexane (HPLC grade, x 5) before drying for 30 min at rt.

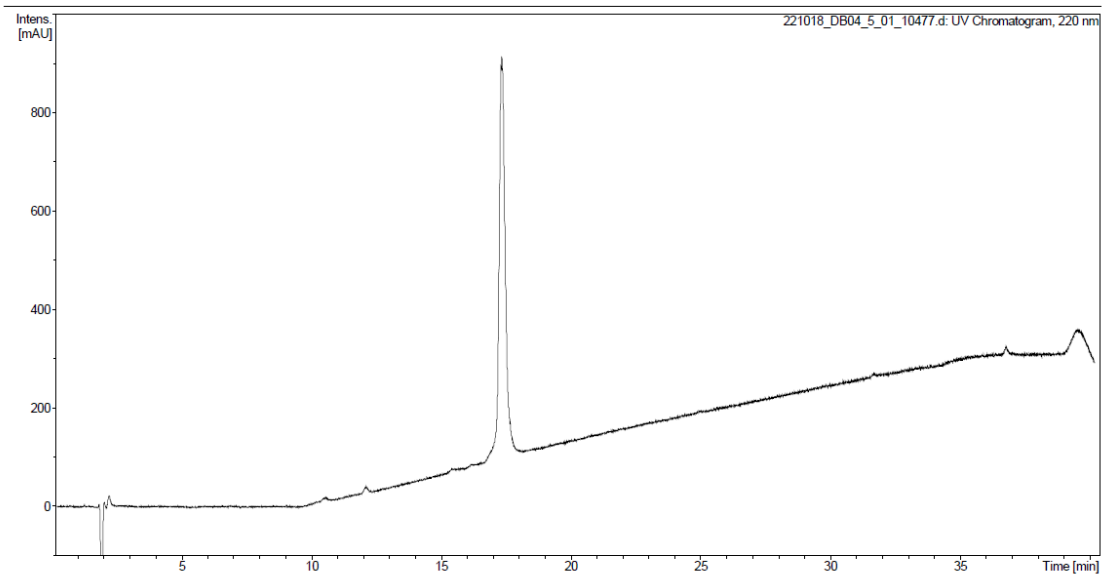
6.4 Synthesis of Fmoc-Hyp-$(Pro)_3$-NH_2 (3):



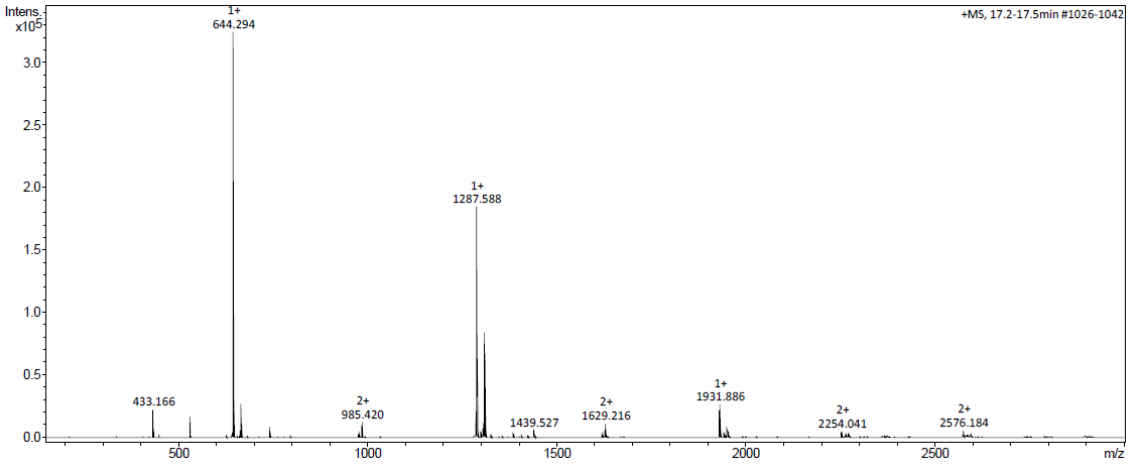
^{1}H NMR of **HP_3 (400 MHz, MeOD)**



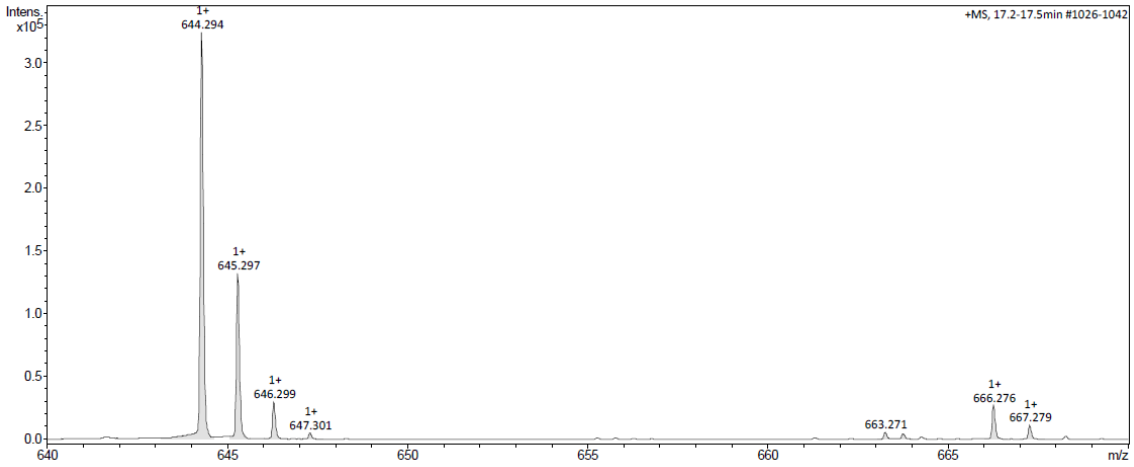
^{13}C NMR of **HP_3 (101 MHz, MeOD)**



HPLC UV-Vis trace of peptide **HP_3**, rt = 17.35 min, 220 nm

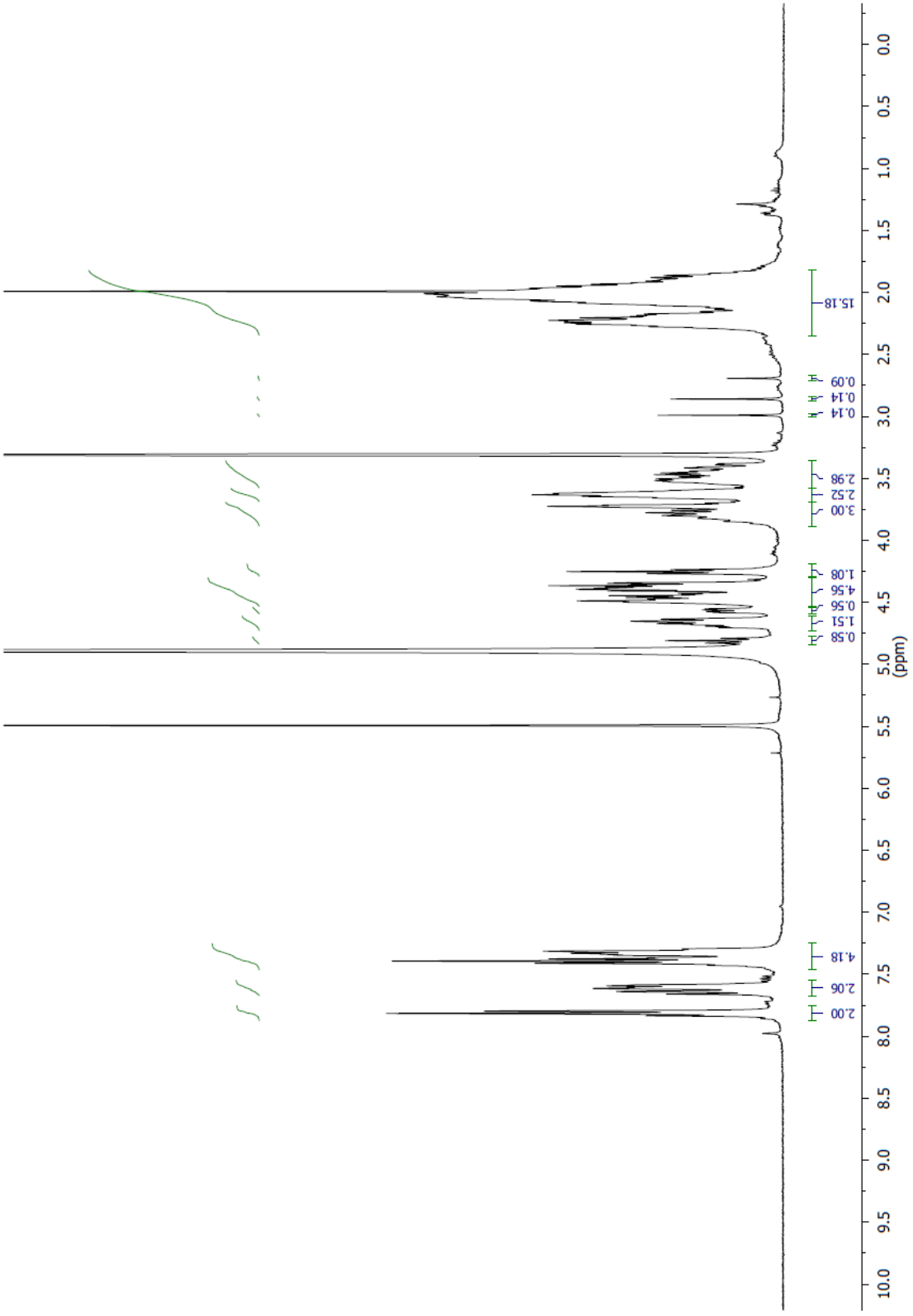


Mass spectrum (100-2000 m/z) of peak at 17.35 min, **HP_3** HPLC

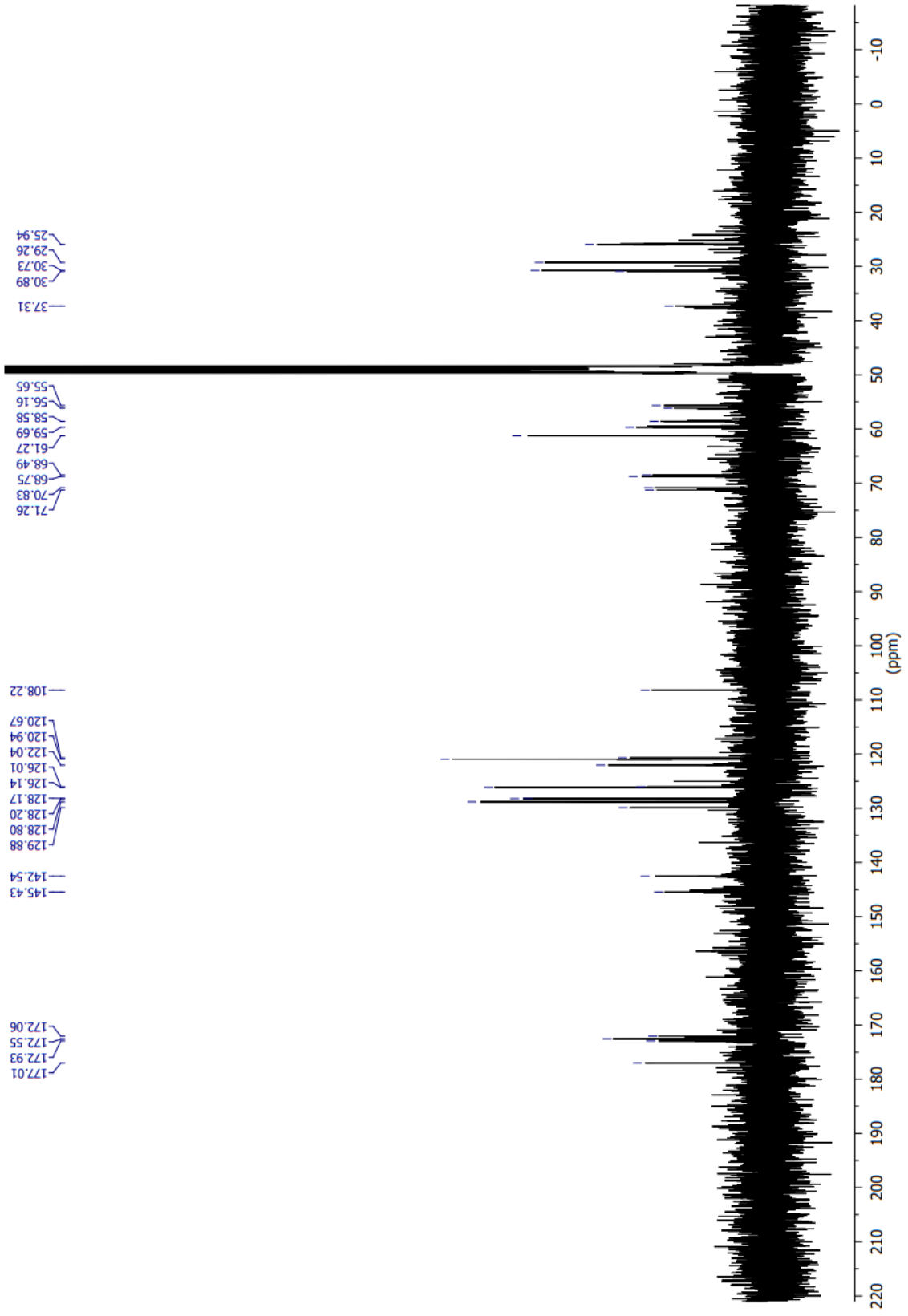


Isotopic distribution spectrum of peak at 17.35 min (m/z 644.294), **HP_3** HPLC

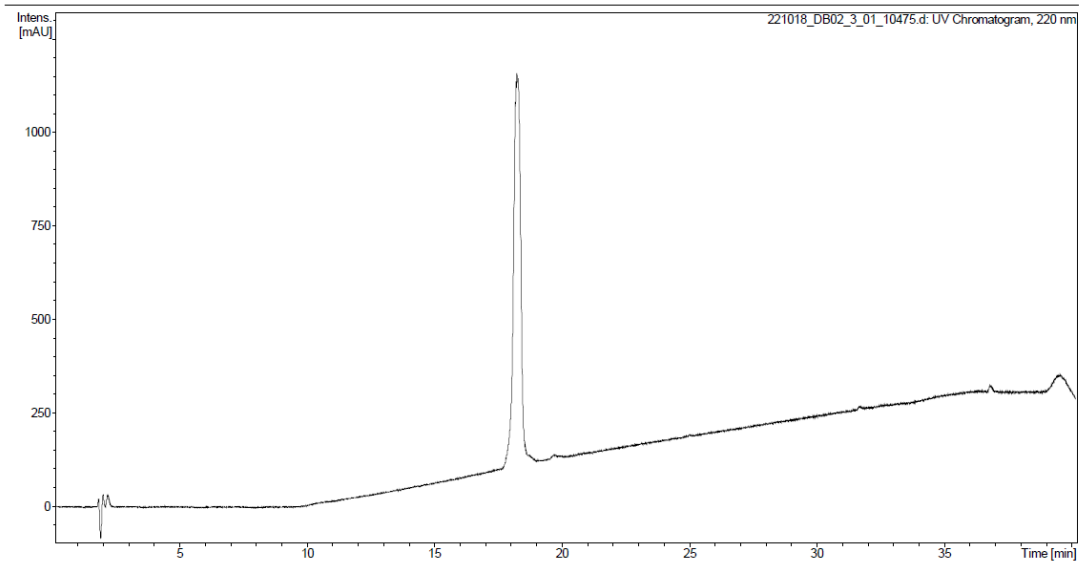
6.5 Synthesis of Fmoc-Pro-Hyp-$(Pro)_2$-NH_2 (4):



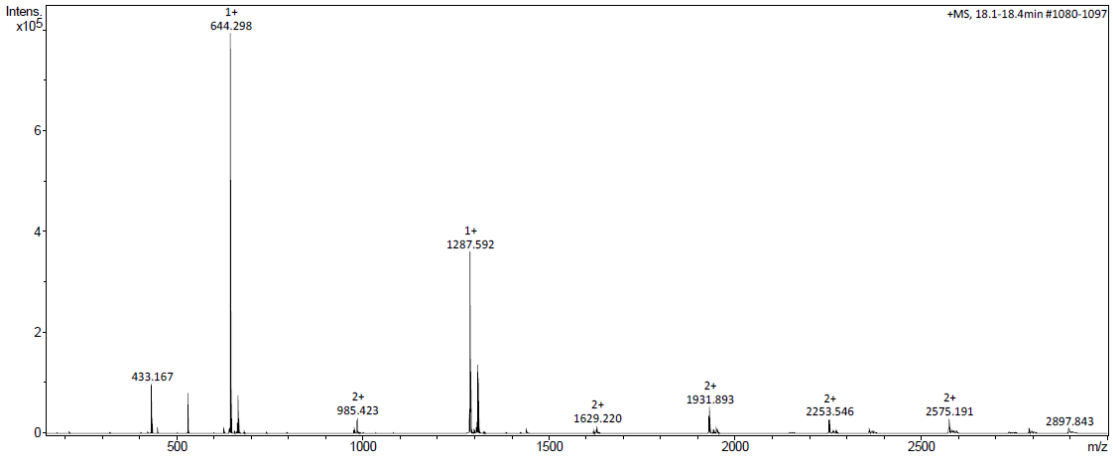
^{1}H NMR of **PHP_2 (400 MHz, MeOD)**



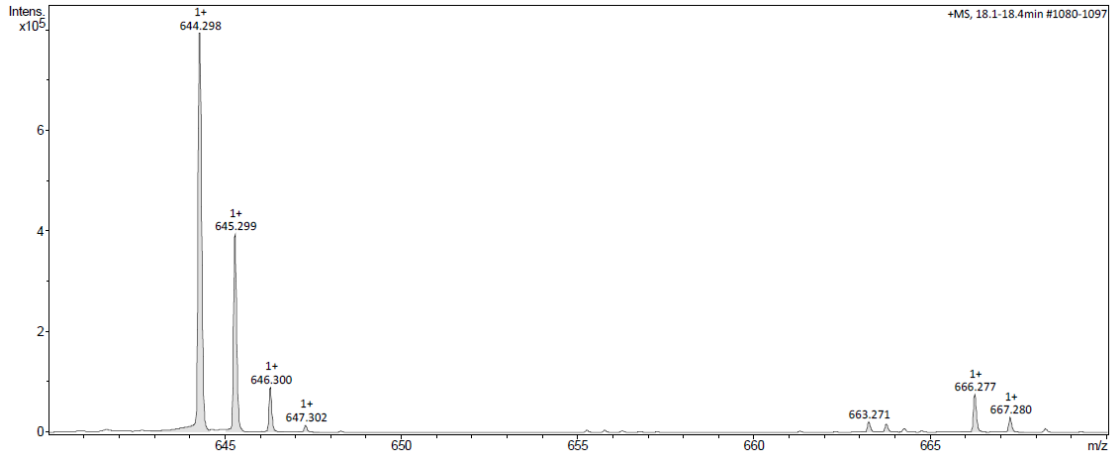
^{13}C NMR of **PHP_2 (101 MHz, MeOD)**



HPLC UV-Vis trace of peptide **PHP_2**, rt = 18.25 min, 220 nm

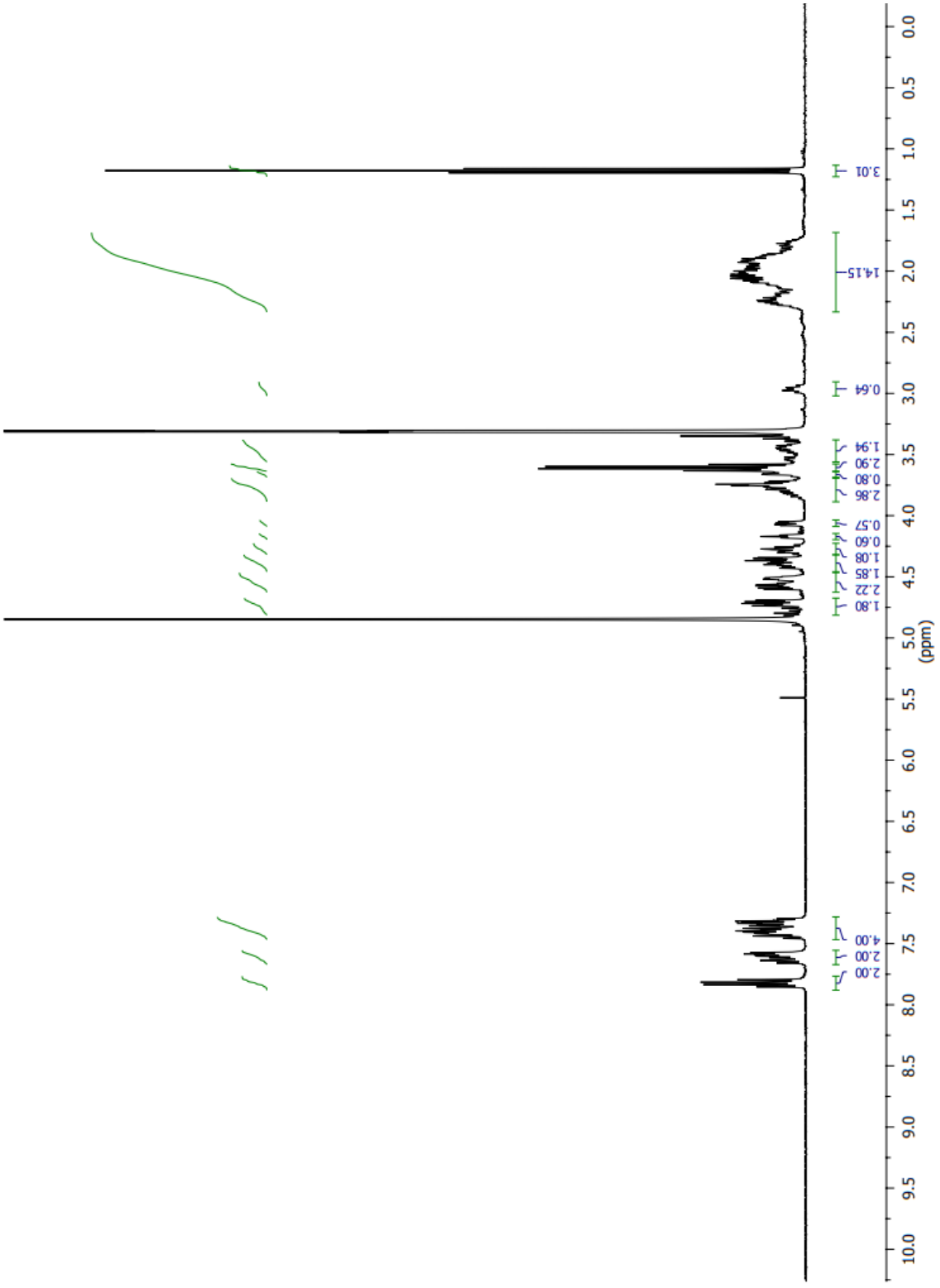


Mass spectrum (100-2000 m/z) of peak at 18.25 min, **PHP_2** HPLC

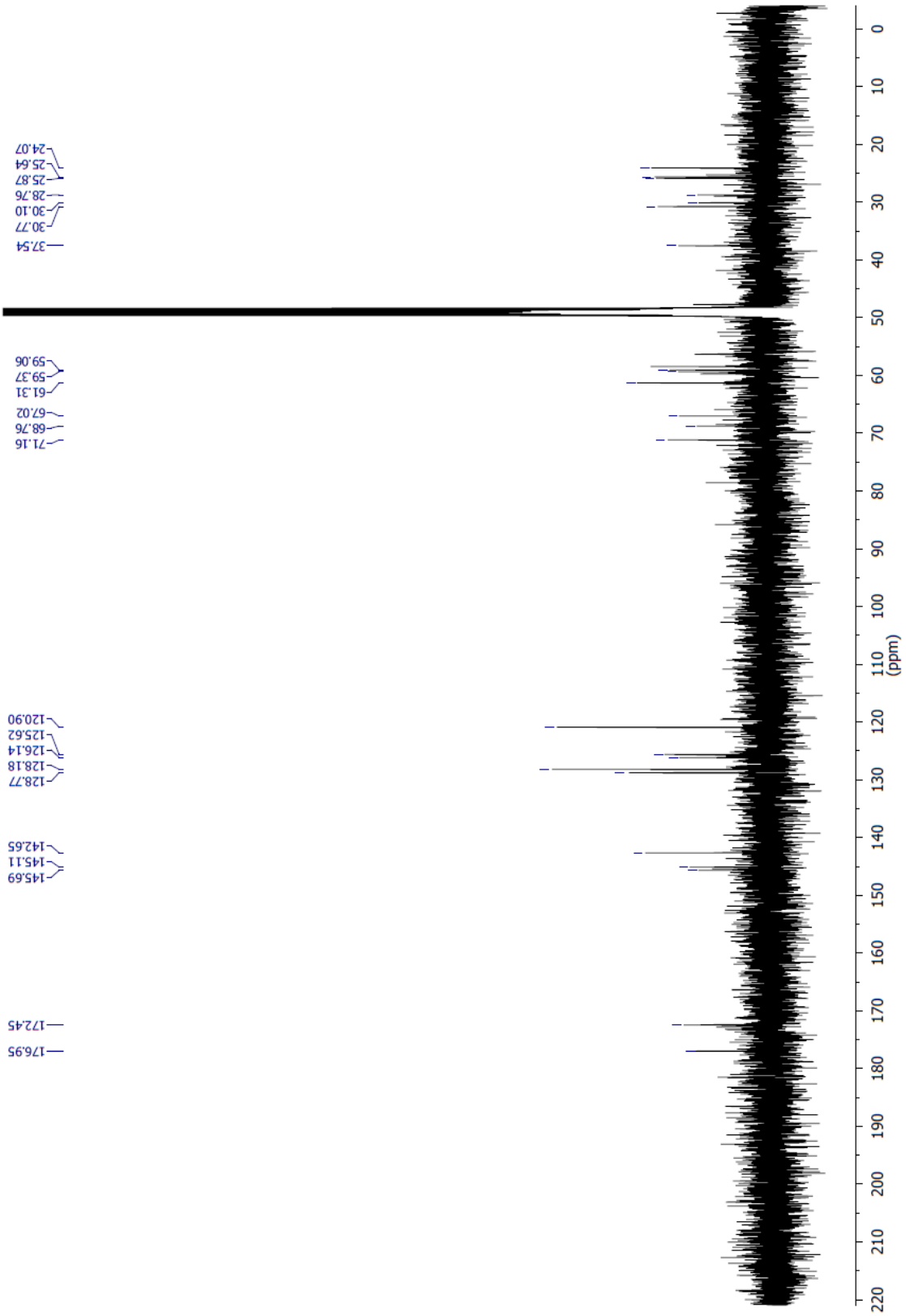


Isotopic distribution spectrum of peak at 18.25 min (m/z 644.298), **PHP_2** HPLC

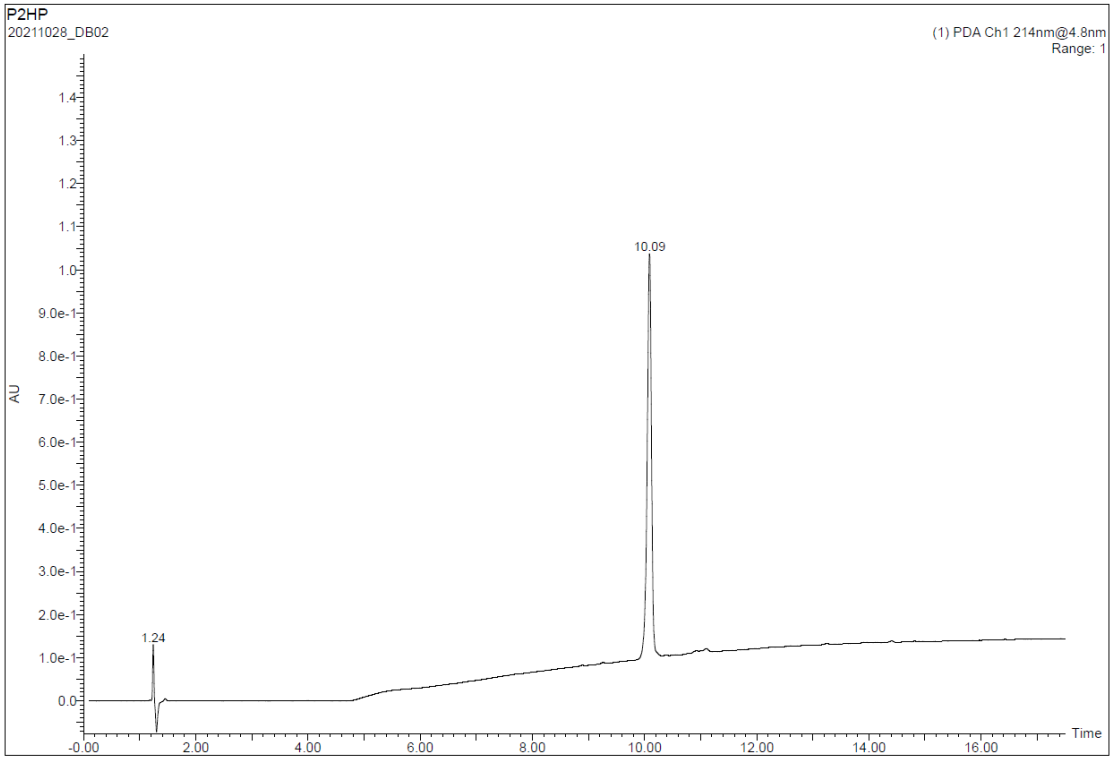
6.6 Synthesis of Fmoc-$(Pro)_2$-Hyp-Pro-NH_2 (5):



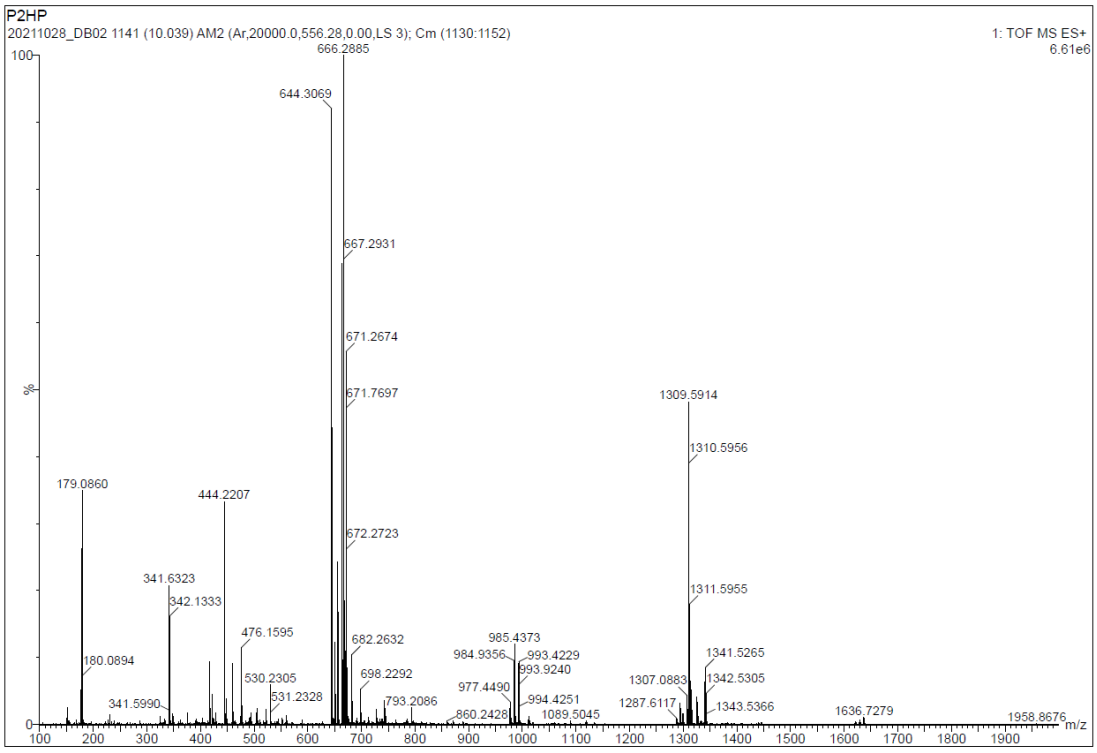
^{1}H NMR of **P_2HP (400 MHz, MeOD)**



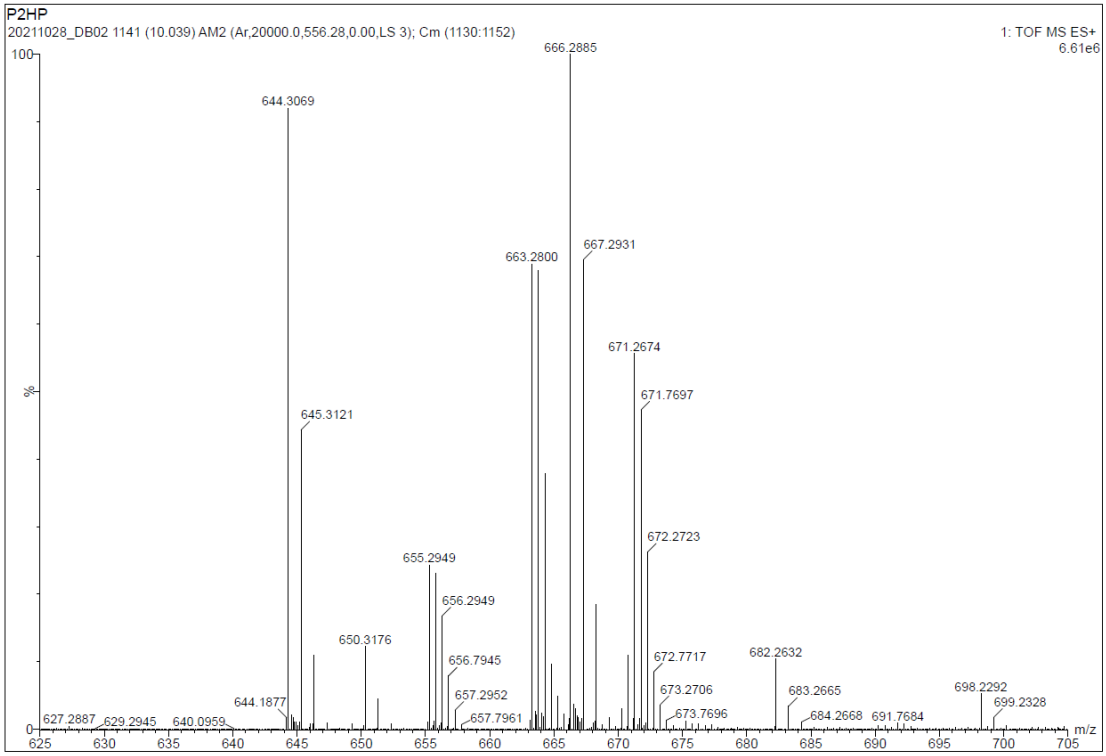
^{13}C NMR of **P_2HP (101 MHz, MeOD)**



UHPLC UV-Vis trace of peptide **P_2HP**, rt = 10.09 min, 214 nm

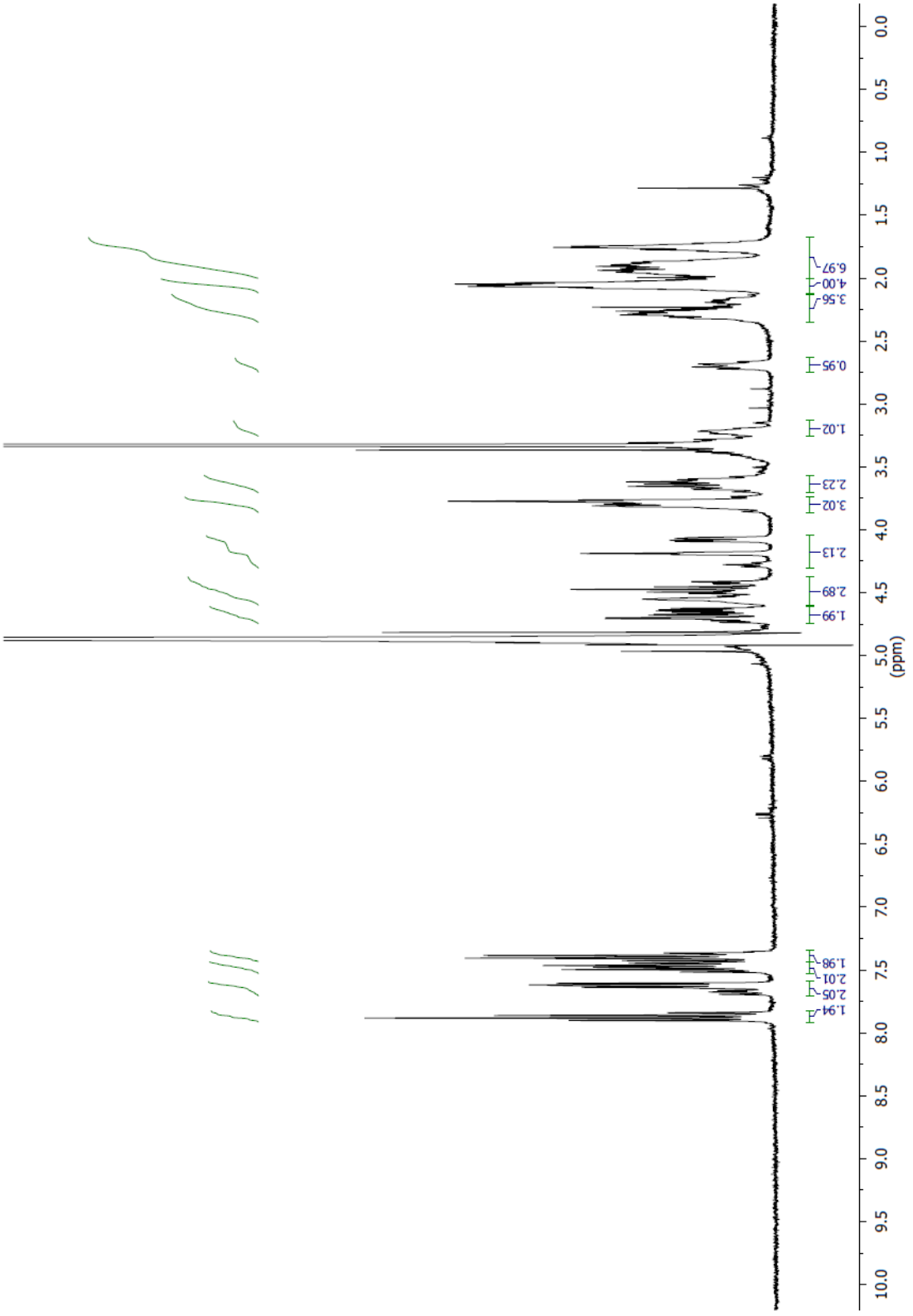


Mass spectrum (100-2000 m/z) of peak at 10.09 min, **P_2HP** UHPLC

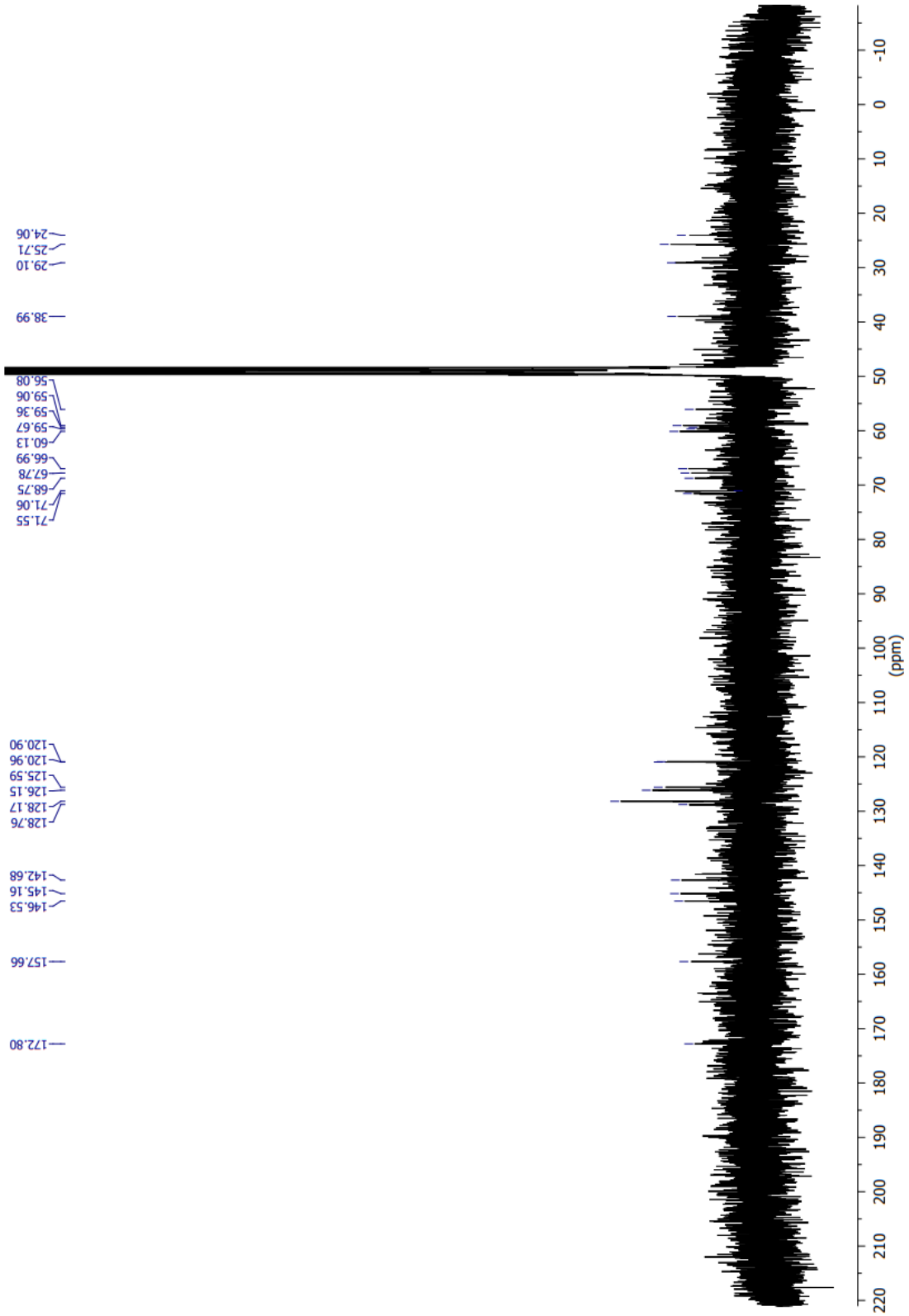


Mass spectrum (625-705 m/z) of peak at 10.09 min, **P_2HP** UHPLC

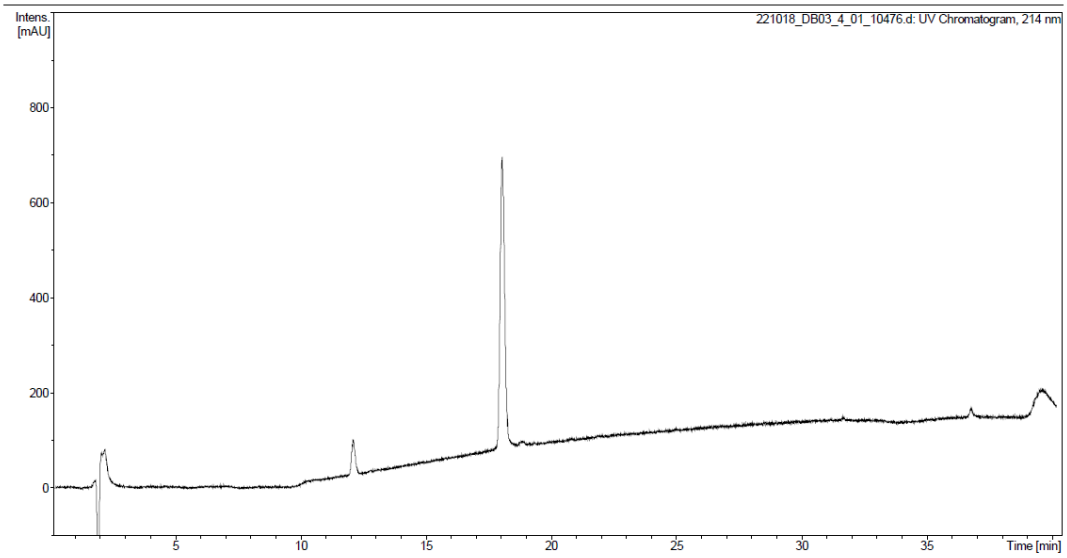
6.7 Synthesis of Fmoc-$(Pro)_3$-Hyp-NH_2 (6):



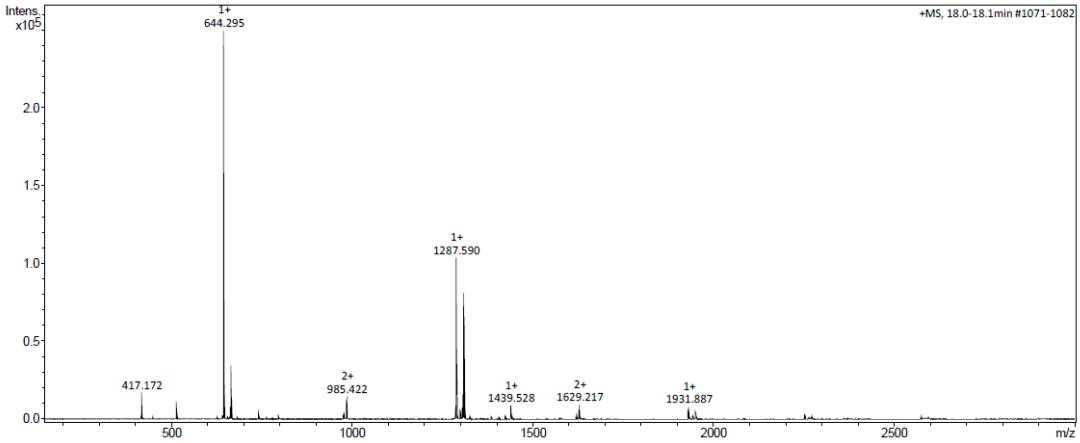
^{1}H NMR of **P_3H (400 MHz, MeOD)**



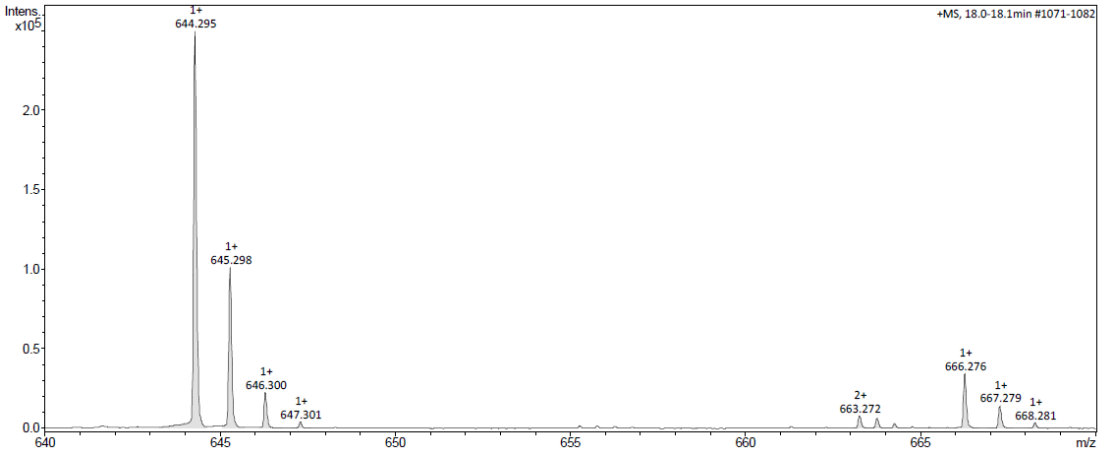
^{13}C NMR of **P_3H (101 MHz, MeOD)**



HPLC UV-Vis trace of peptide **P_3H**, rt = 18.05 min, 214 nm

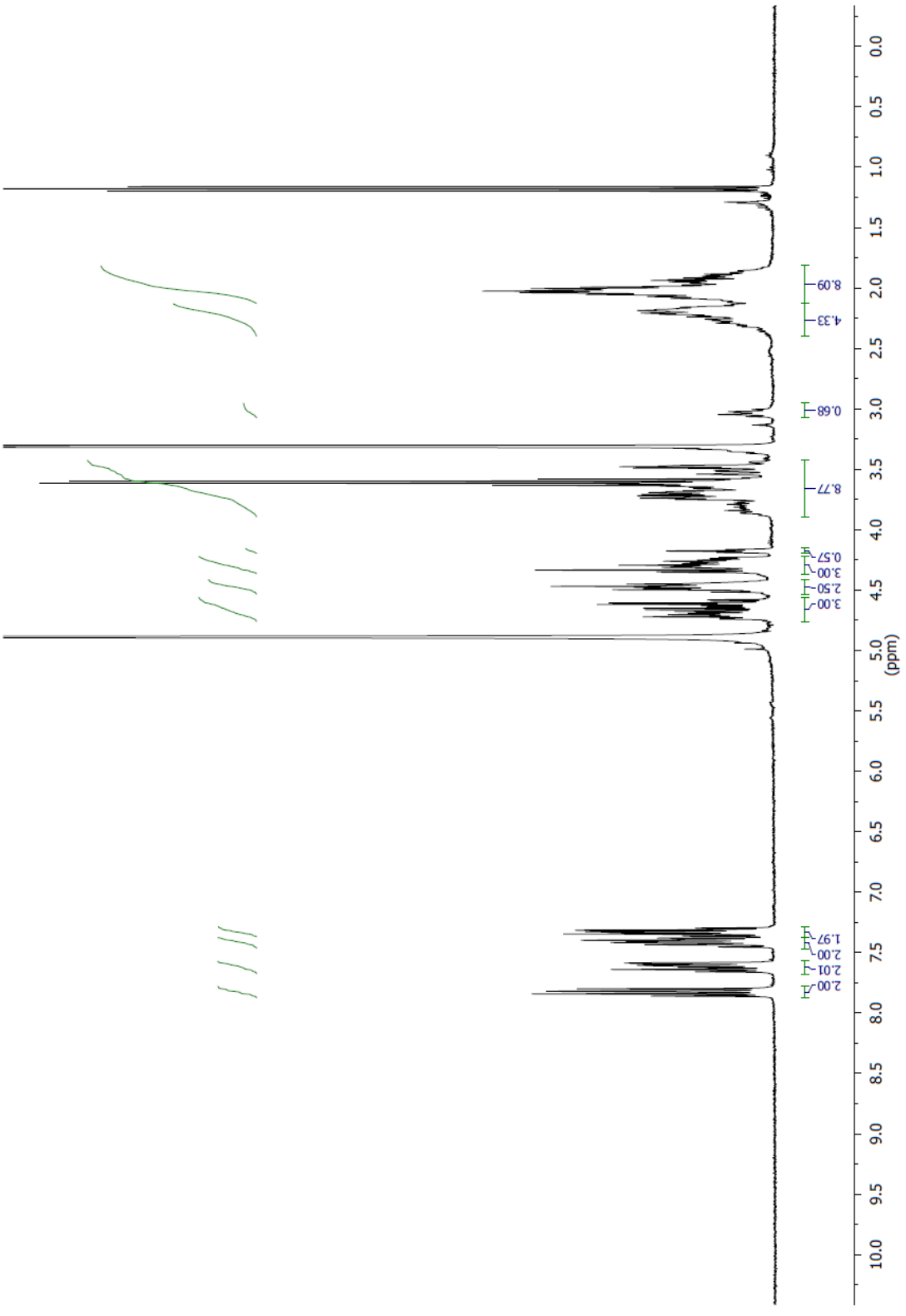


Mass spectrum (100-2000 m/z) of peak at 18.05 min, **P_3H** HPLC

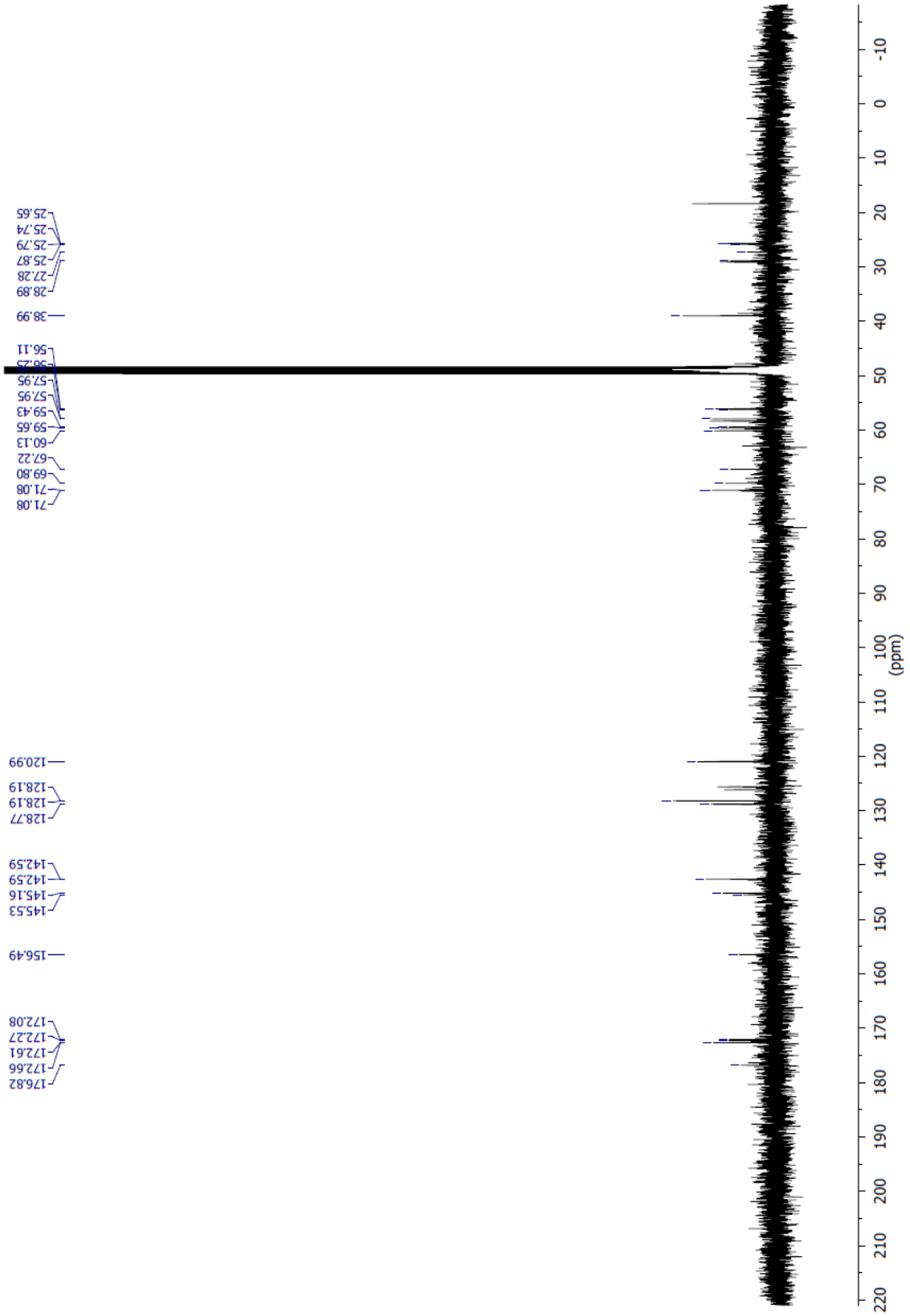


Isotopic distribution spectrum of peak at 18.05 min (m/z 644.295), **P_3H** HPLC

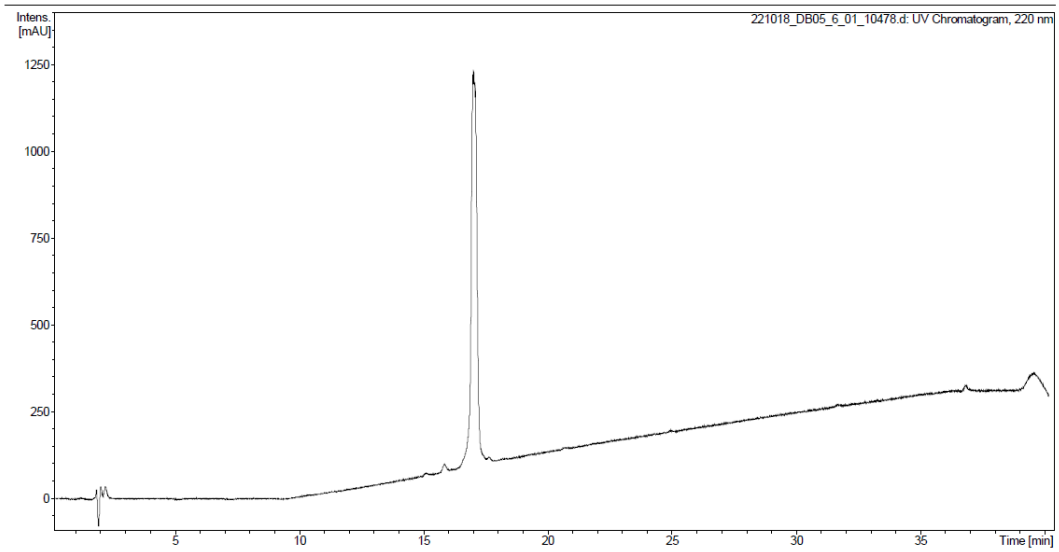
6.8 Synthesis of Fmoc-Hyp-$(Pro)_2$-Hyp-NH_2 (7):



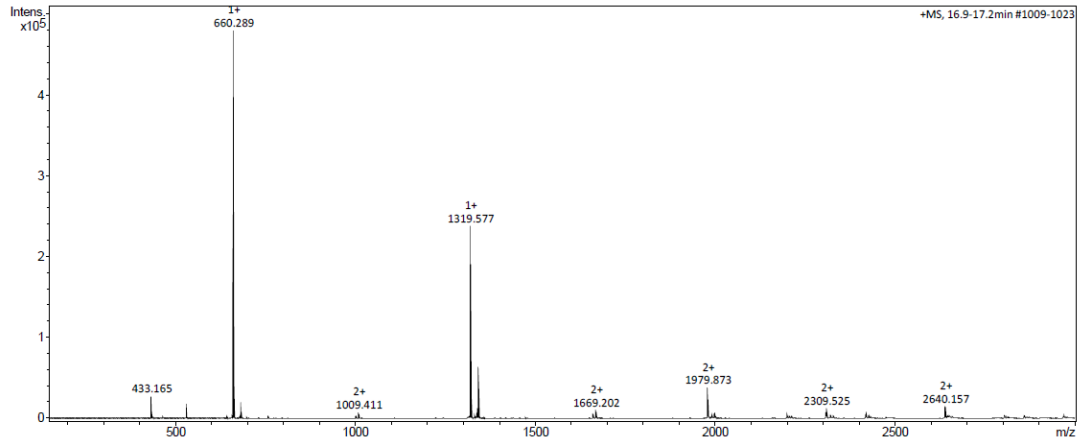
^{1}H NMR of **HP_2H (400 MHz, MeOD)**



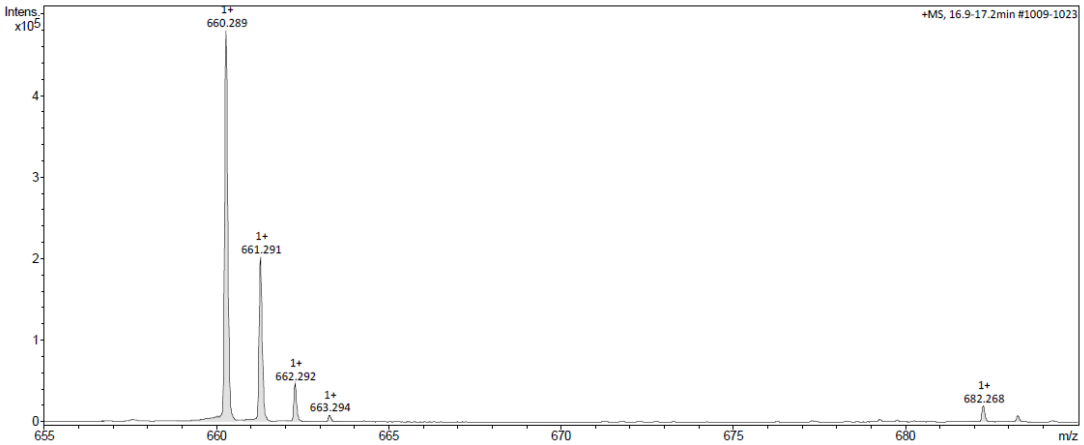
^{13}C NMR of **HP_2H (101 MHz, MeOD)**



HPLC UV-Vis trace of peptide **HP_2H**, rt = 17.05 min, 220 nm

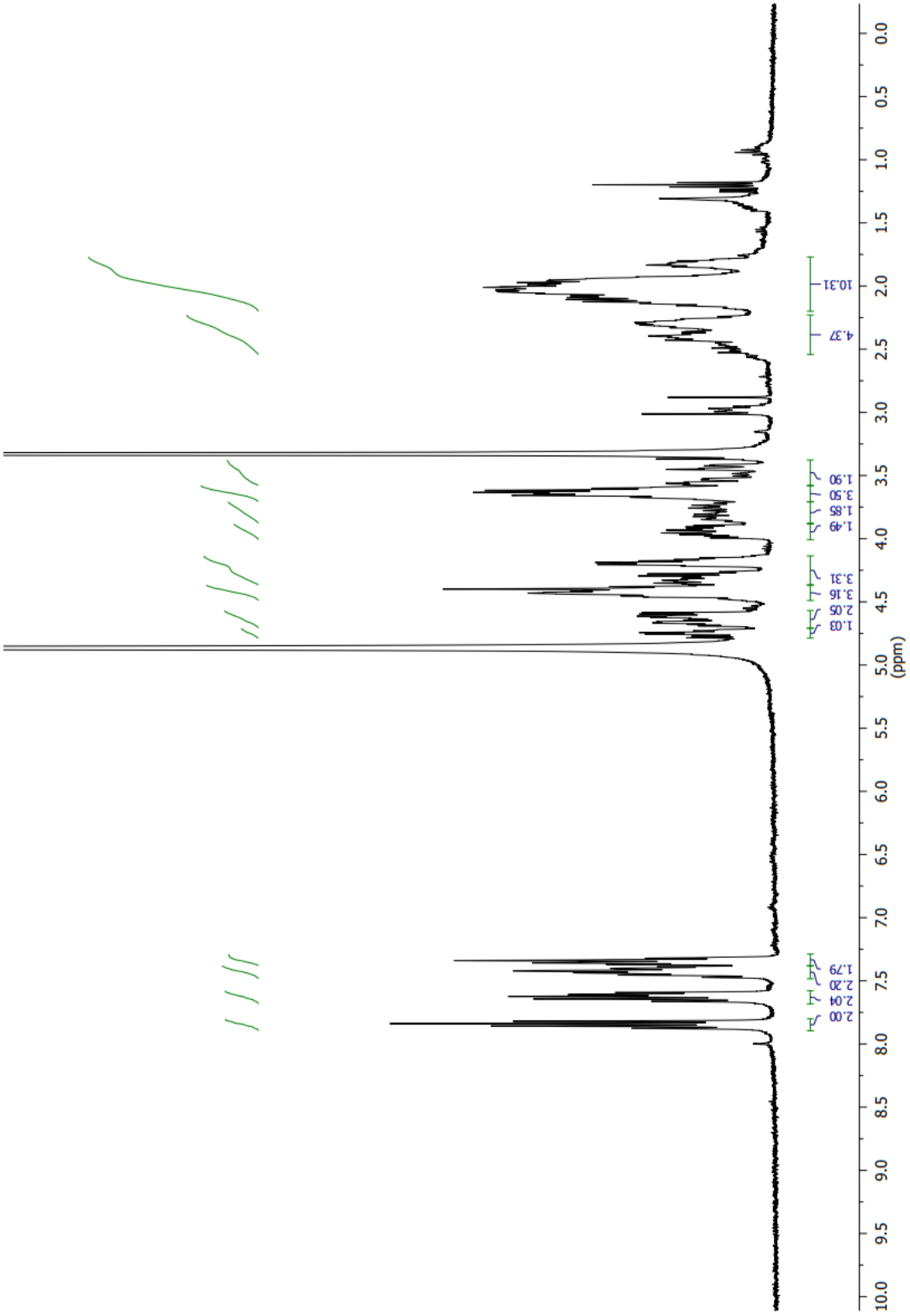


Mass spectrum (100-2000 m/z) of peak at 17.05 min, **HP_2H** HPLC

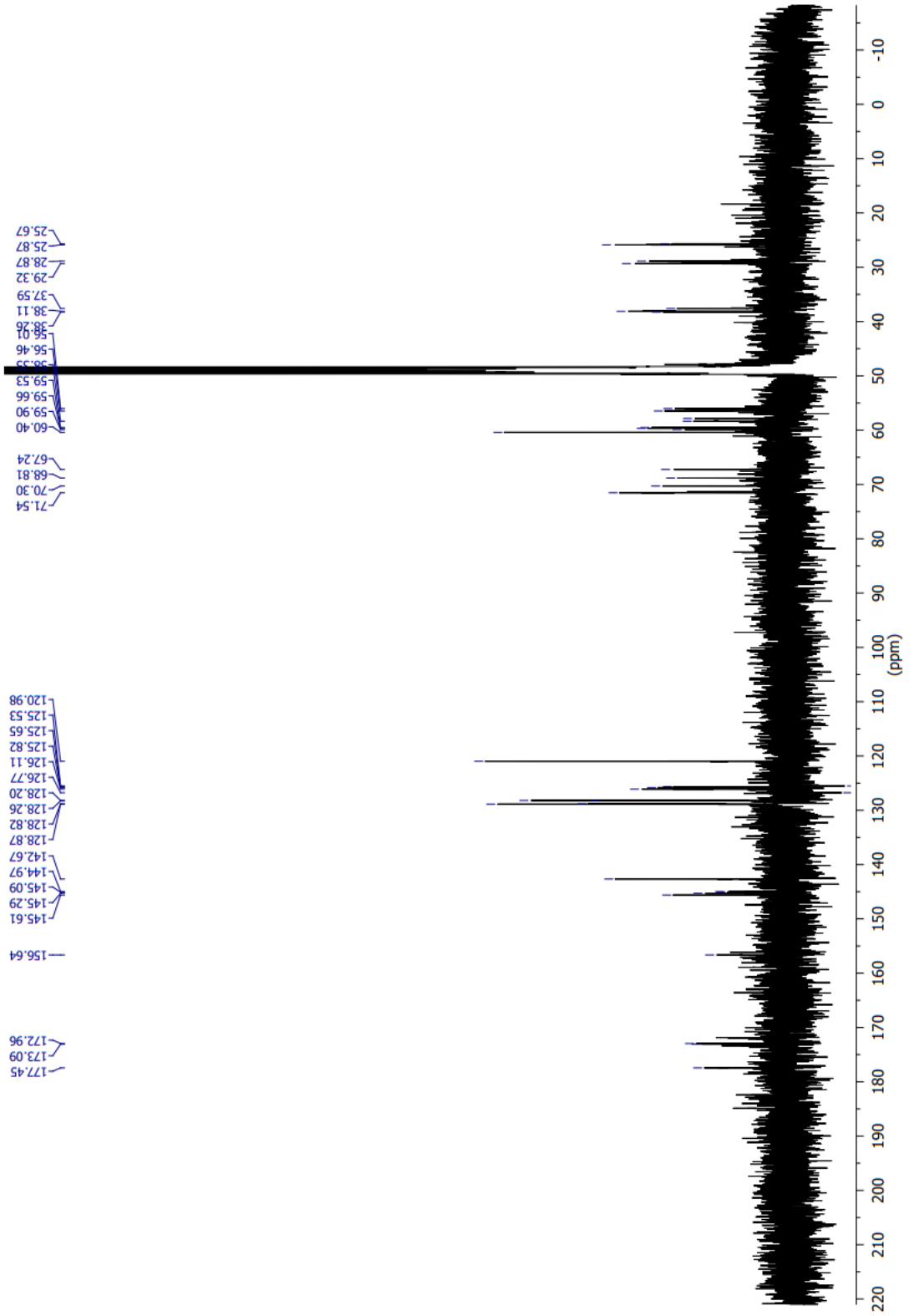


Isotopic distribution spectrum of peak at 17.05 min (m/z 660.289), **HP_2H** HPLC

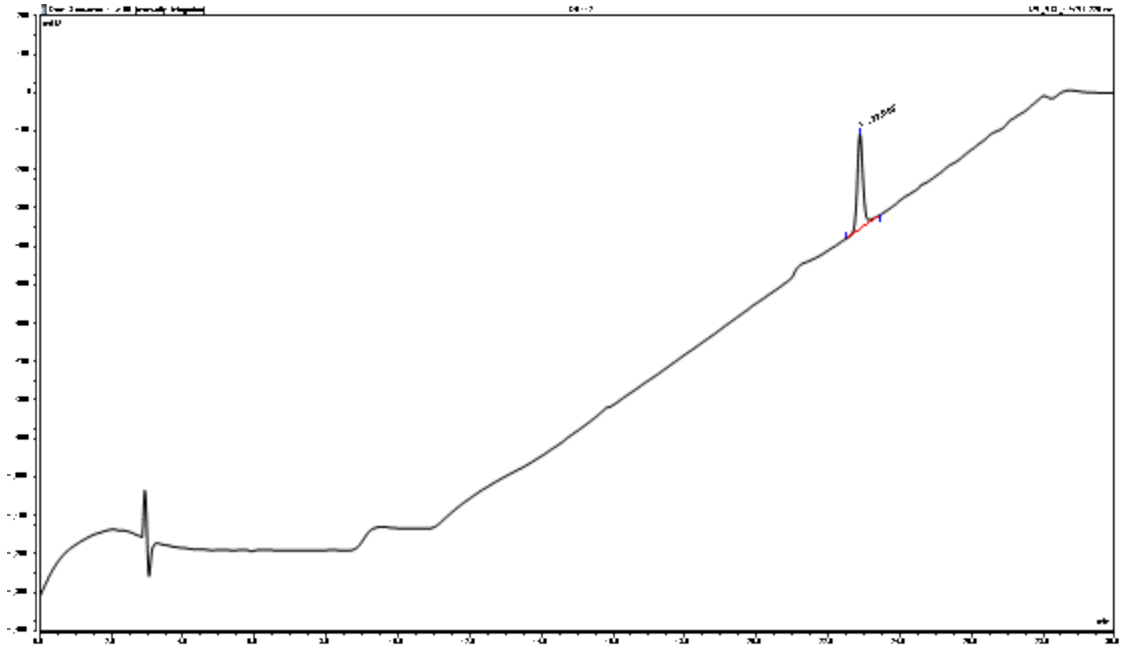
6.9 Synthesis of Fmoc-*cis*-Hyp-(Pro)$_2$-*cis*-Hyp-NH$_2$ (8):



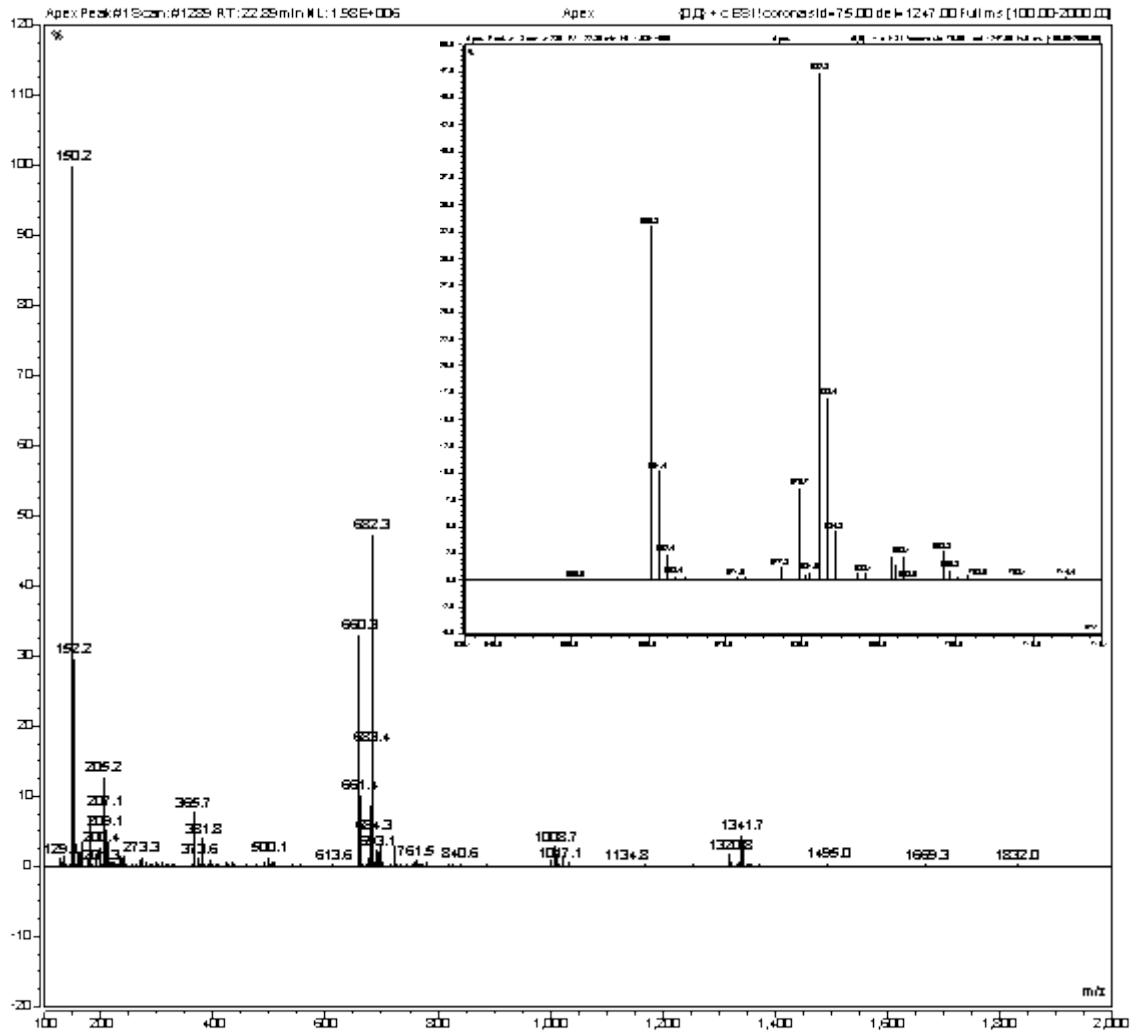
^{1}H NMR of ***cis*-HP$_2$H (400 MHz, MeOD)**



^{13}C NMR of ***cis*-HP$_2$H (101 MHz, MeOD)**

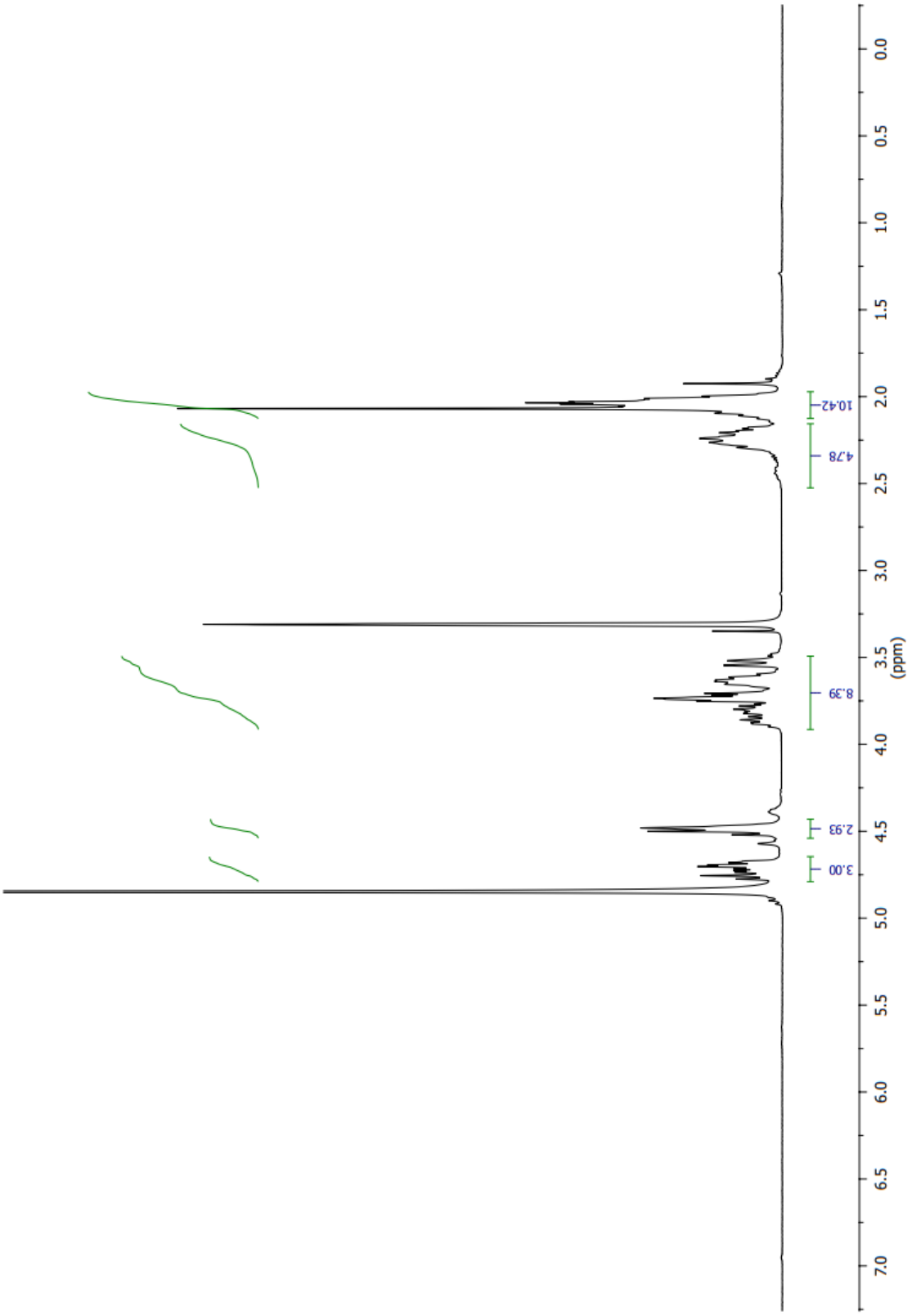


HPLC UV-Vis trace of peptide *cis*-**HP$_2$H**, rt = 22.887 min, 225 nm

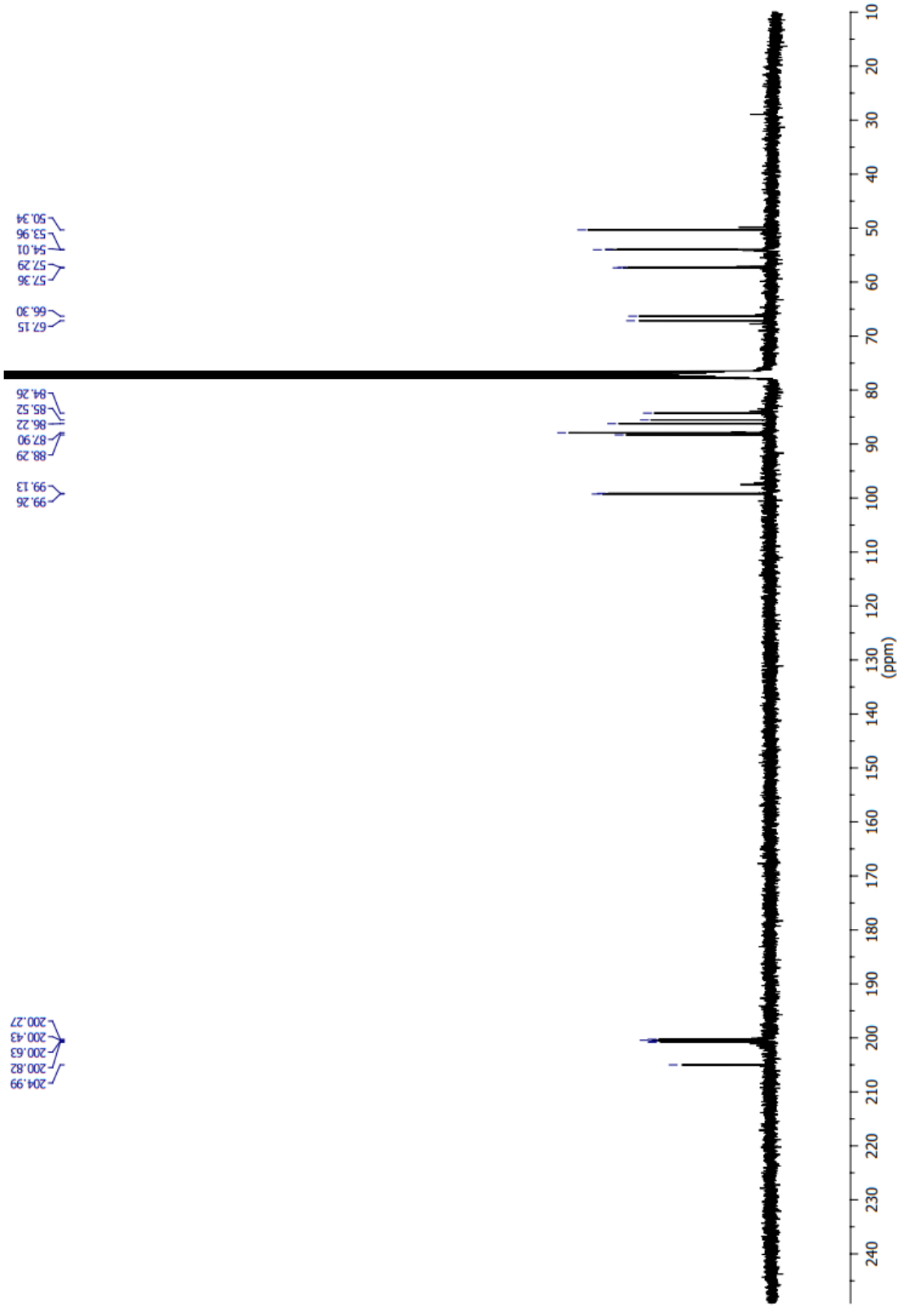


Mass spectrum (100-2000 m/z) of peak at 22.887 min, *cis*-**HP$_2$H** HPLC. M+ ion peak area enlarged.

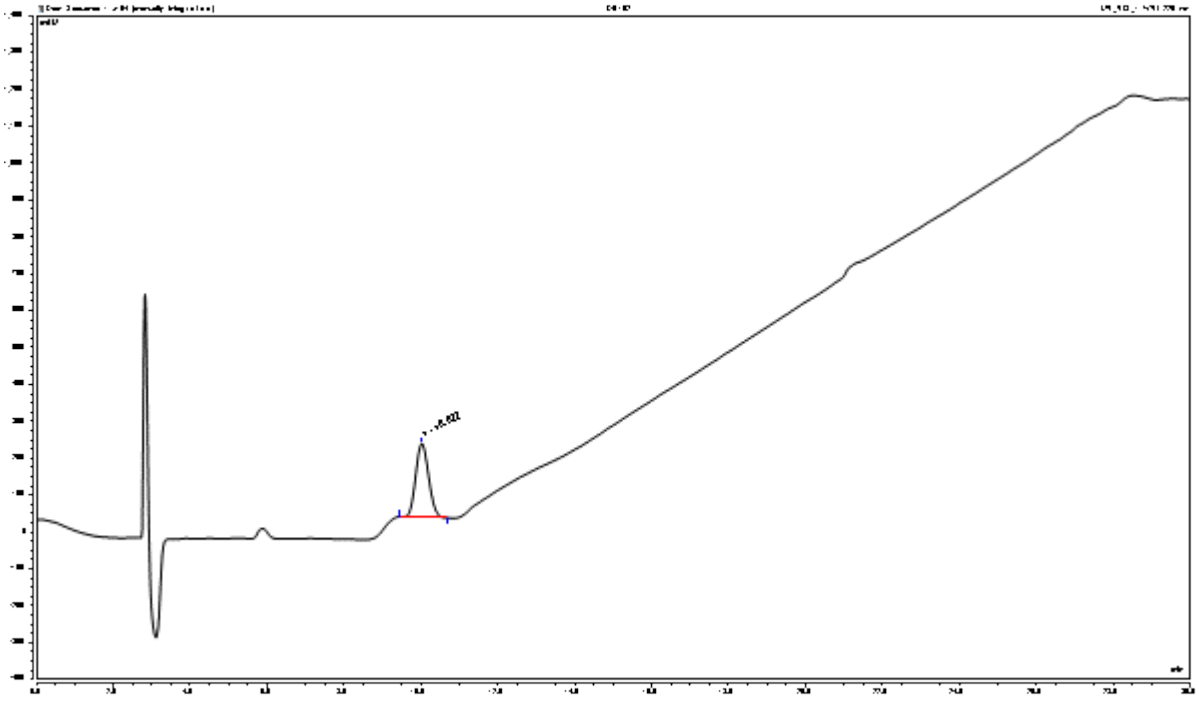
6.10 Synthesis of Ac-Hyp-$(Pro)_2$-Hyp-NH_2 (9):



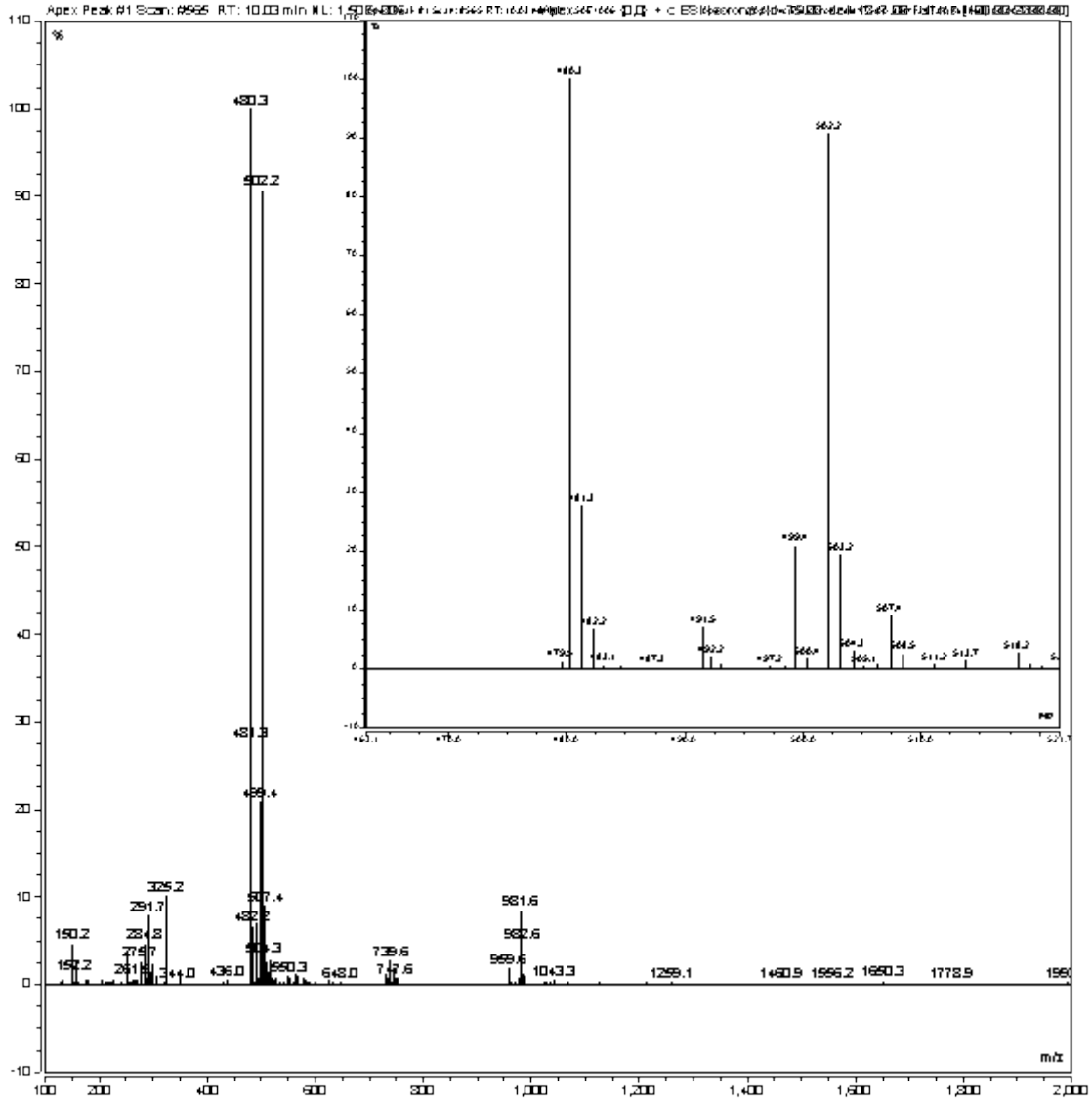
^{1}H NMR of **$AcHP_2H$ (400 MHz, MeOD)**



^{13}C NMR of **AcHP$_2$H (101 MHz, MeOD)**

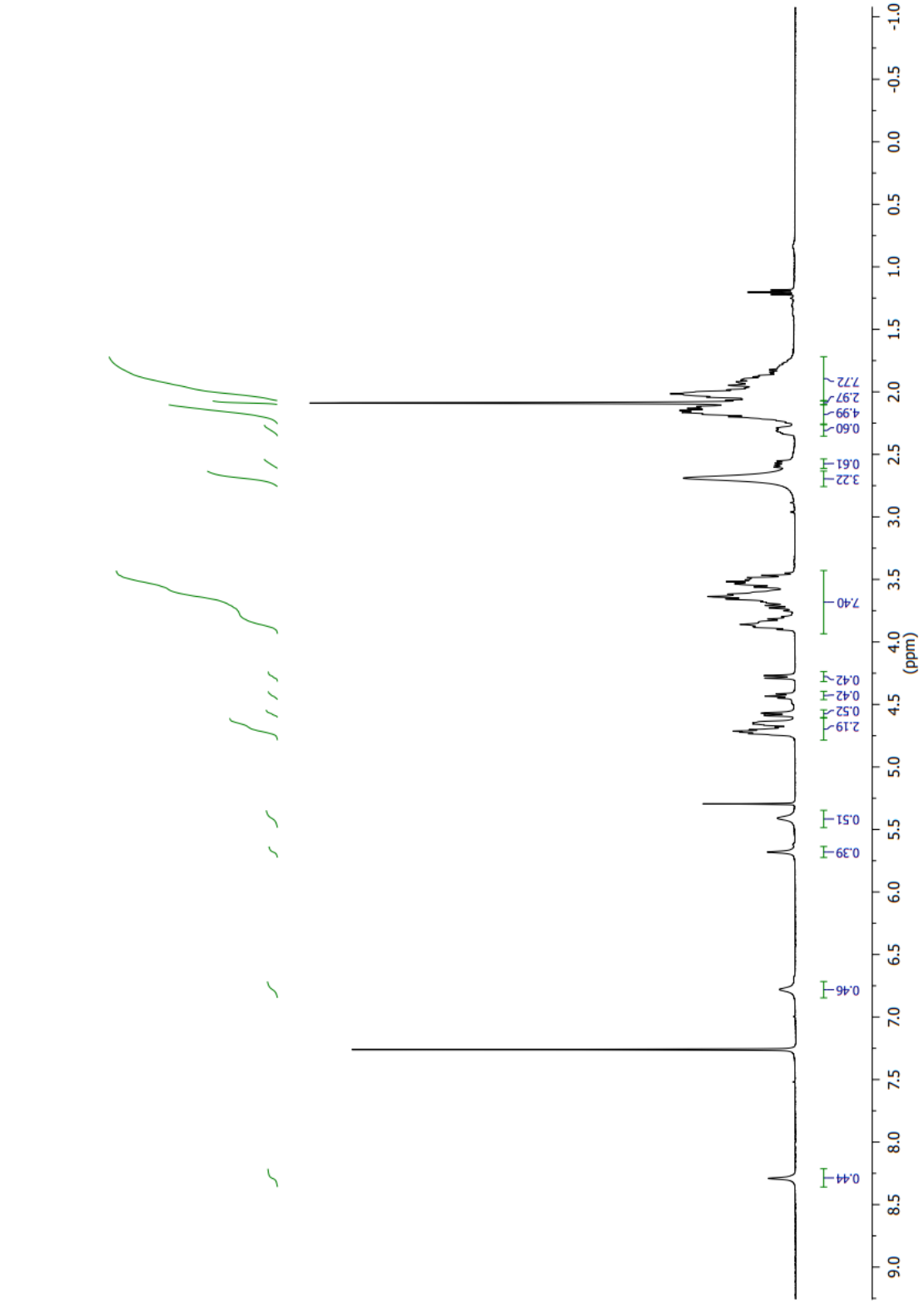


HPLC UV-Vis trace of peptide **AcHP$_2$H**, rt = 10.022 min, 225 nm

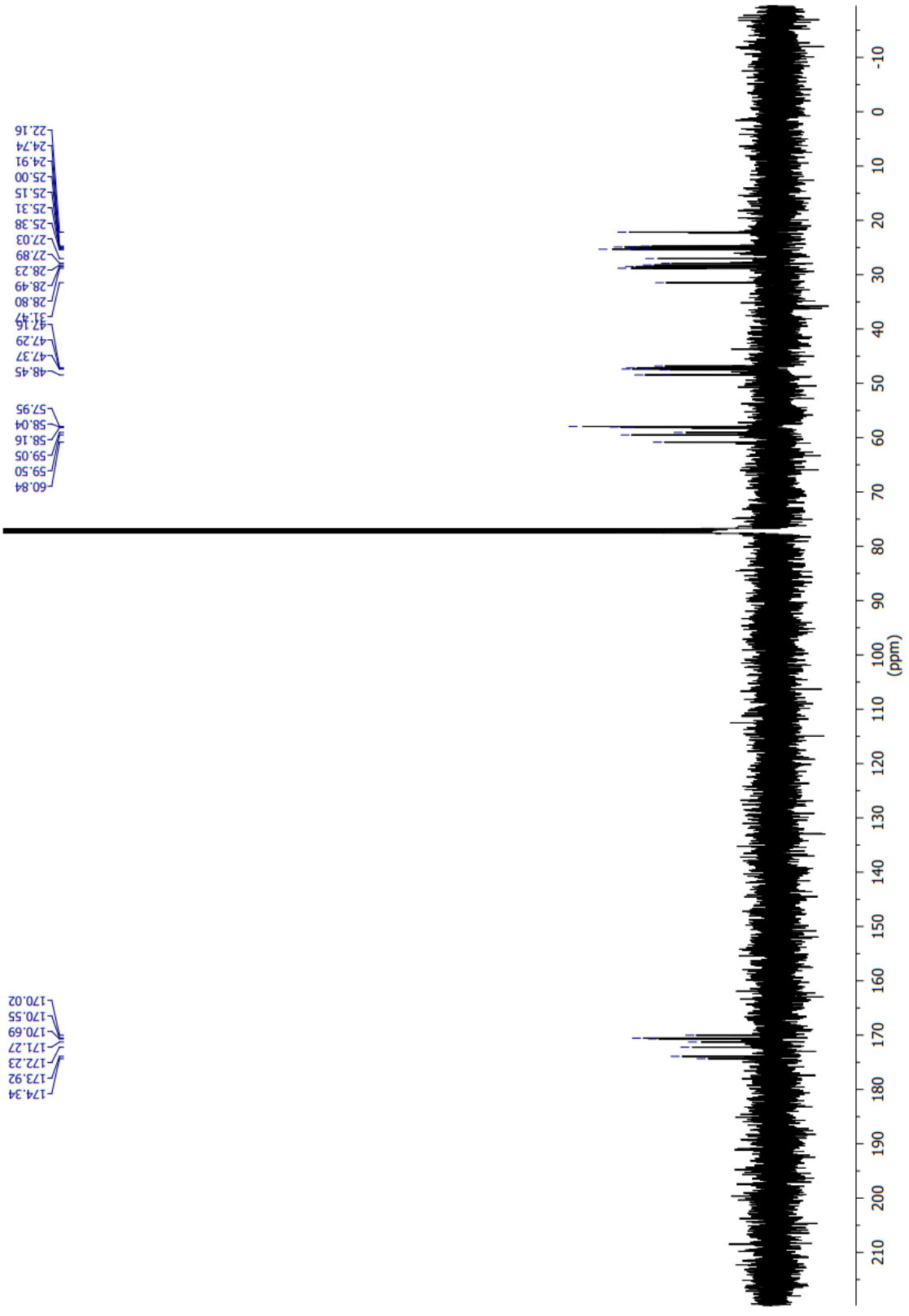


Mass spectrum (100-2000 m/z) of peak at 10.022 min, **AcHP$_2$H** HPLC. M+ ion peak area enlarged.

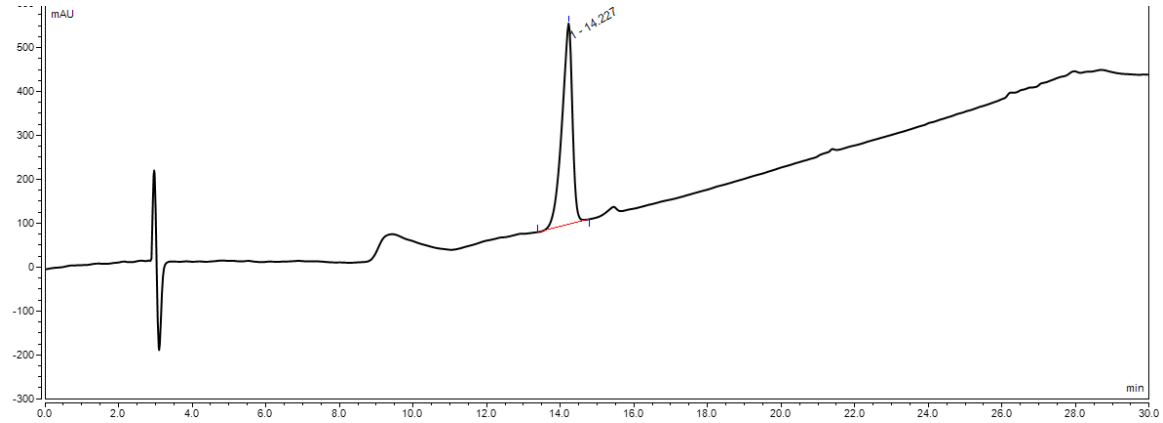
6.11 Synthesis of Ac-Hyp-$(Pro)_2$-Hyp-NH_2, (AcP_4):



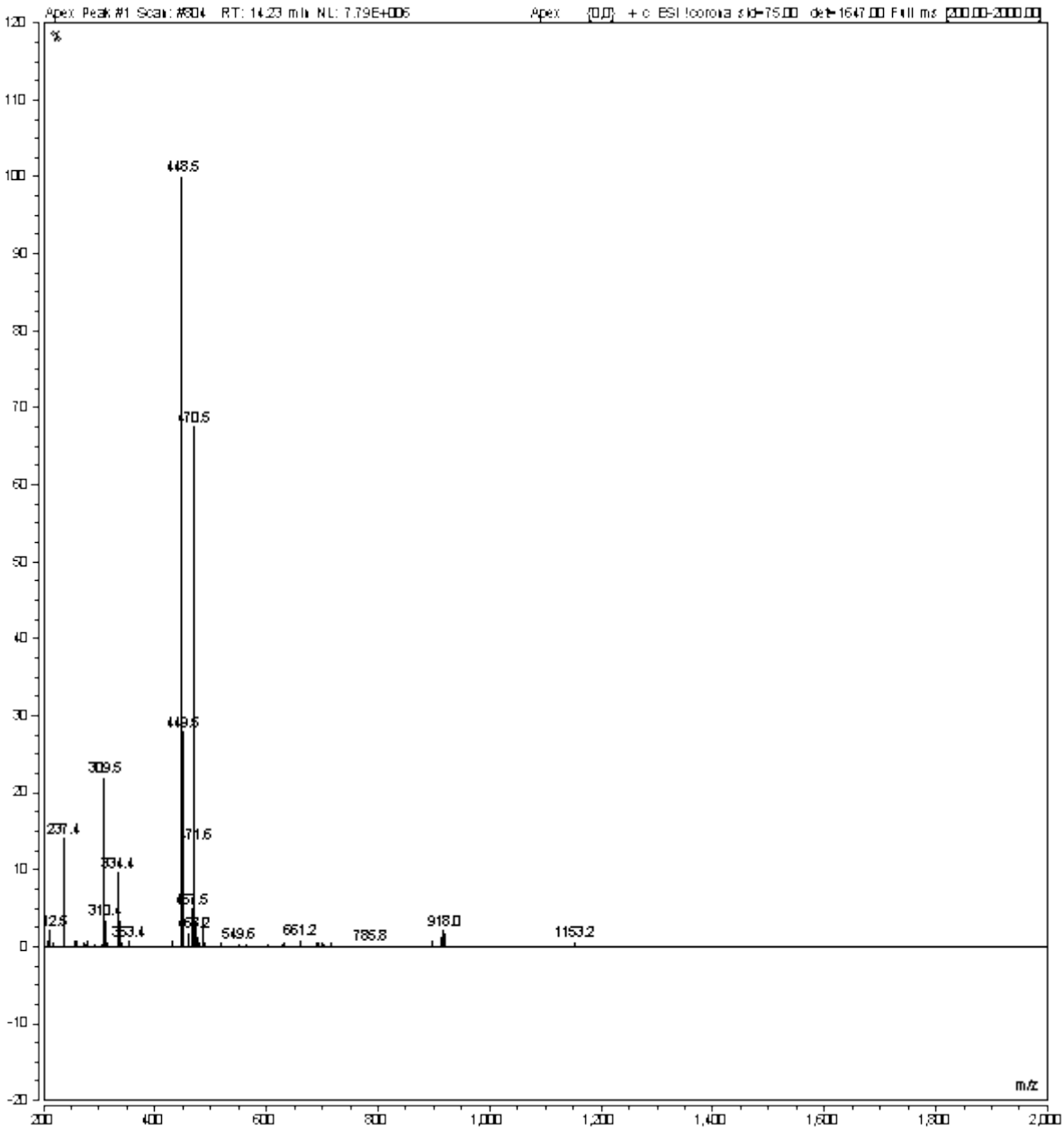
^{1}H NMR of **AcP_4 (400 MHz, $CDCl_3$)**



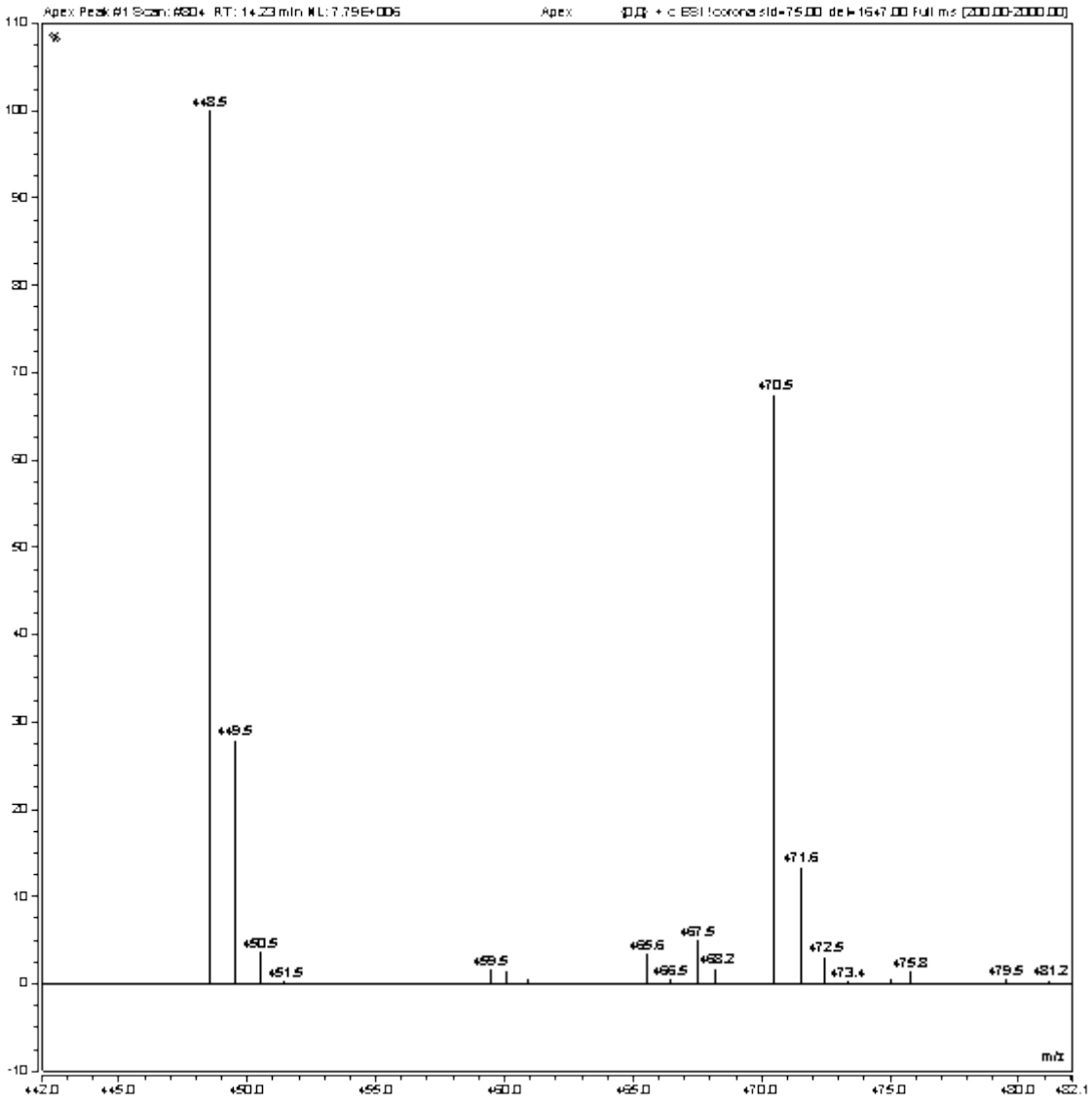
^{13}C NMR of **AcP_4 ($CDCl_3$, 101 MHz)**



HPLC UV-Vis trace of peptide **AcP$_4$**, rt = 14.227 min, 225 nm

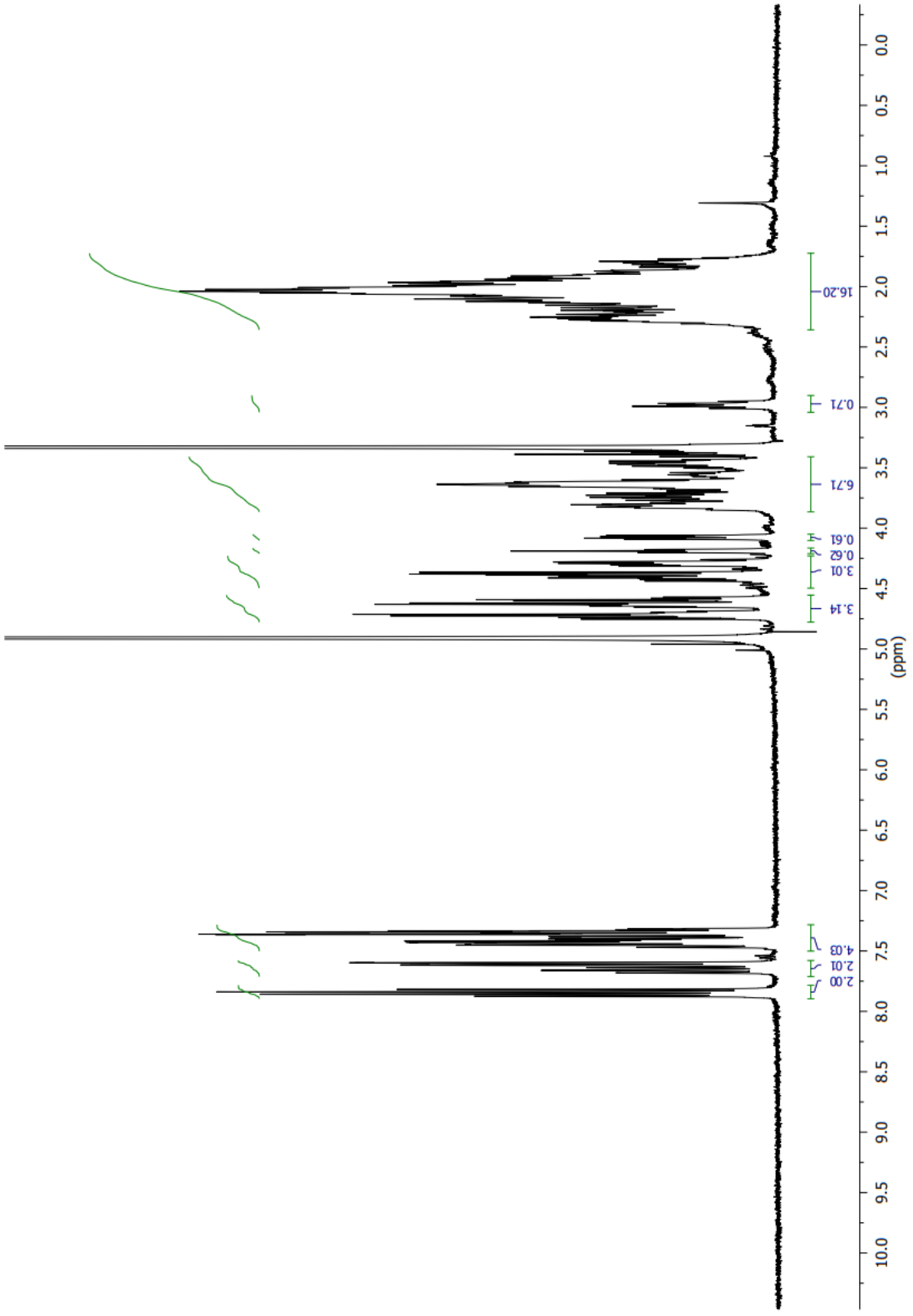


Mass spectrum (200-2000 m/z) of peak at 14.23 min, **AcP$_4$**.

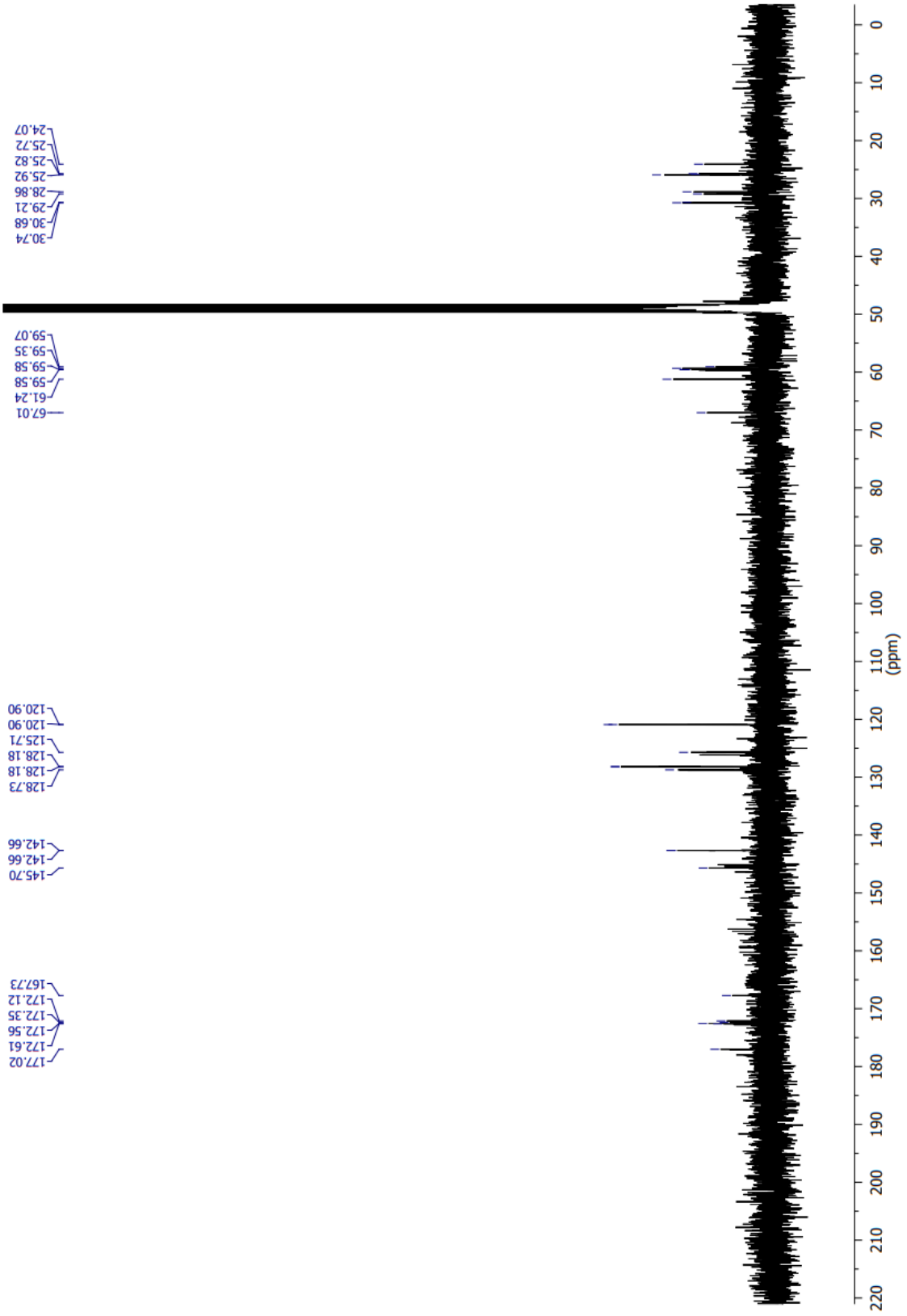


M+ ion peak area enlarged region of MS, peak at 14.23 min, **AcP$_4$**.

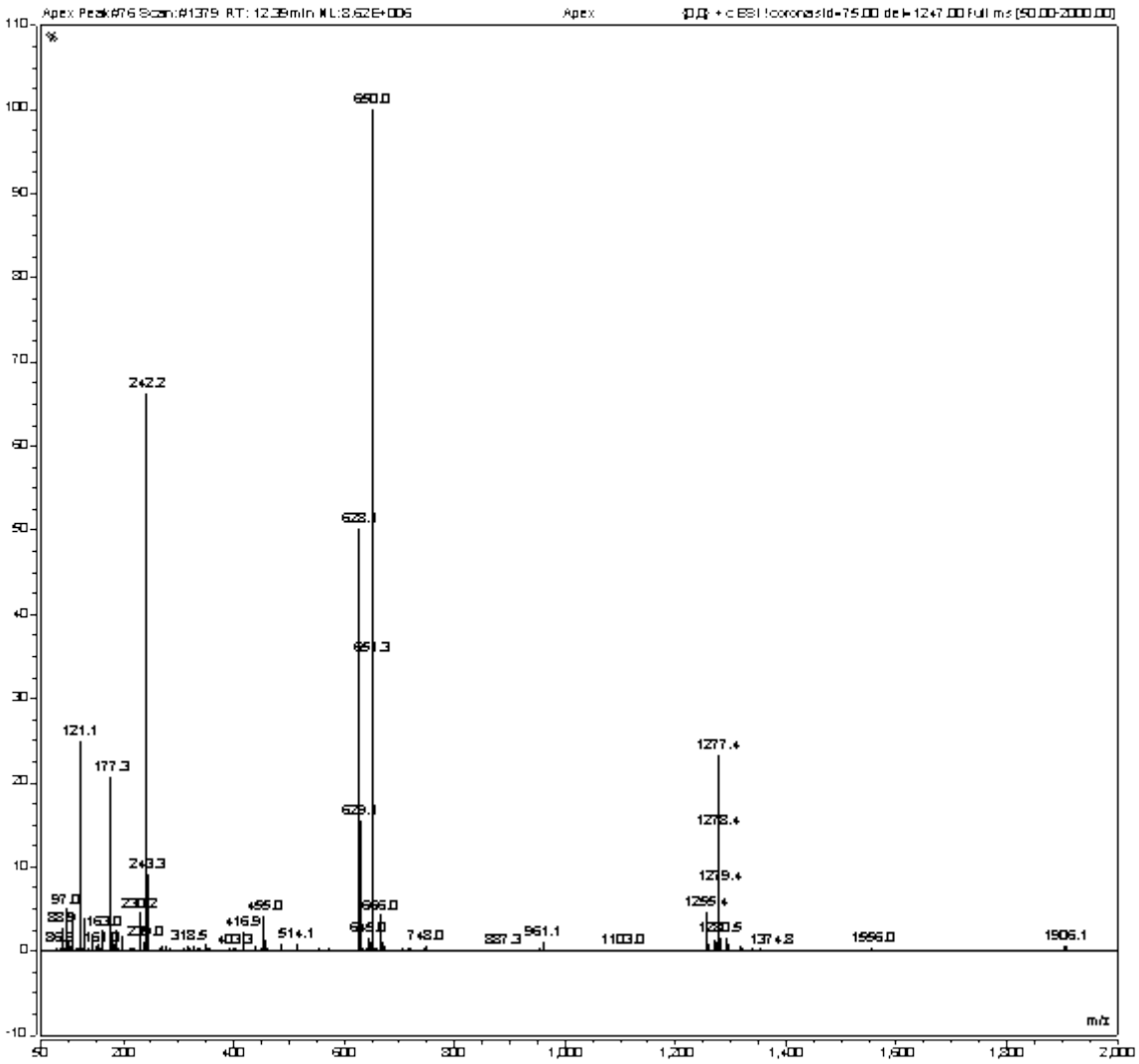
6.12 Synthesis of Fmoc-$(pro)_4$-NH_2 (13):



^{1}H NMR of ***D*-P_4 (400 MHz, MeOD)**

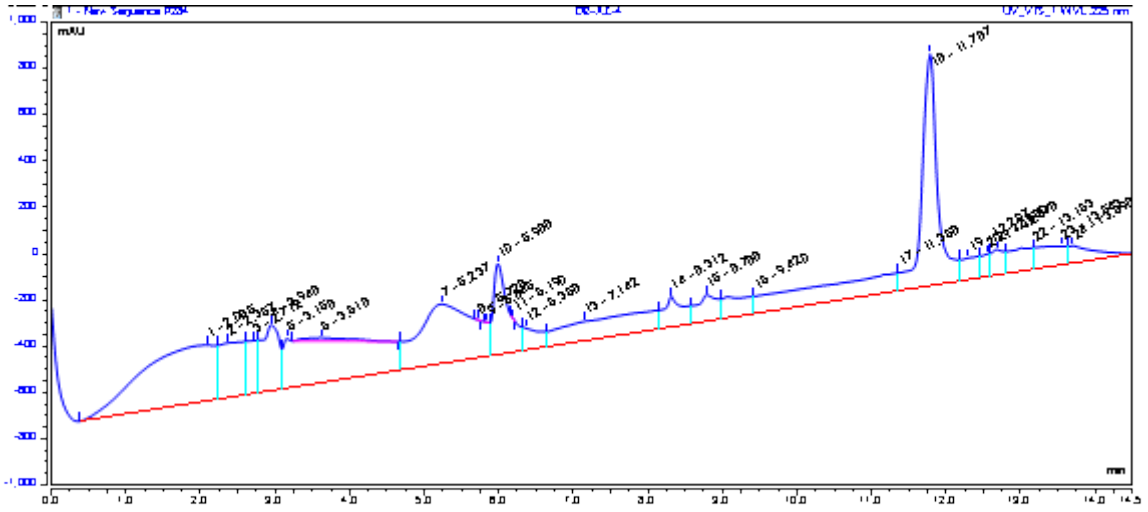


^{13}C NMR of ***D*-P$_4$ (101 MHz, MeOD)**

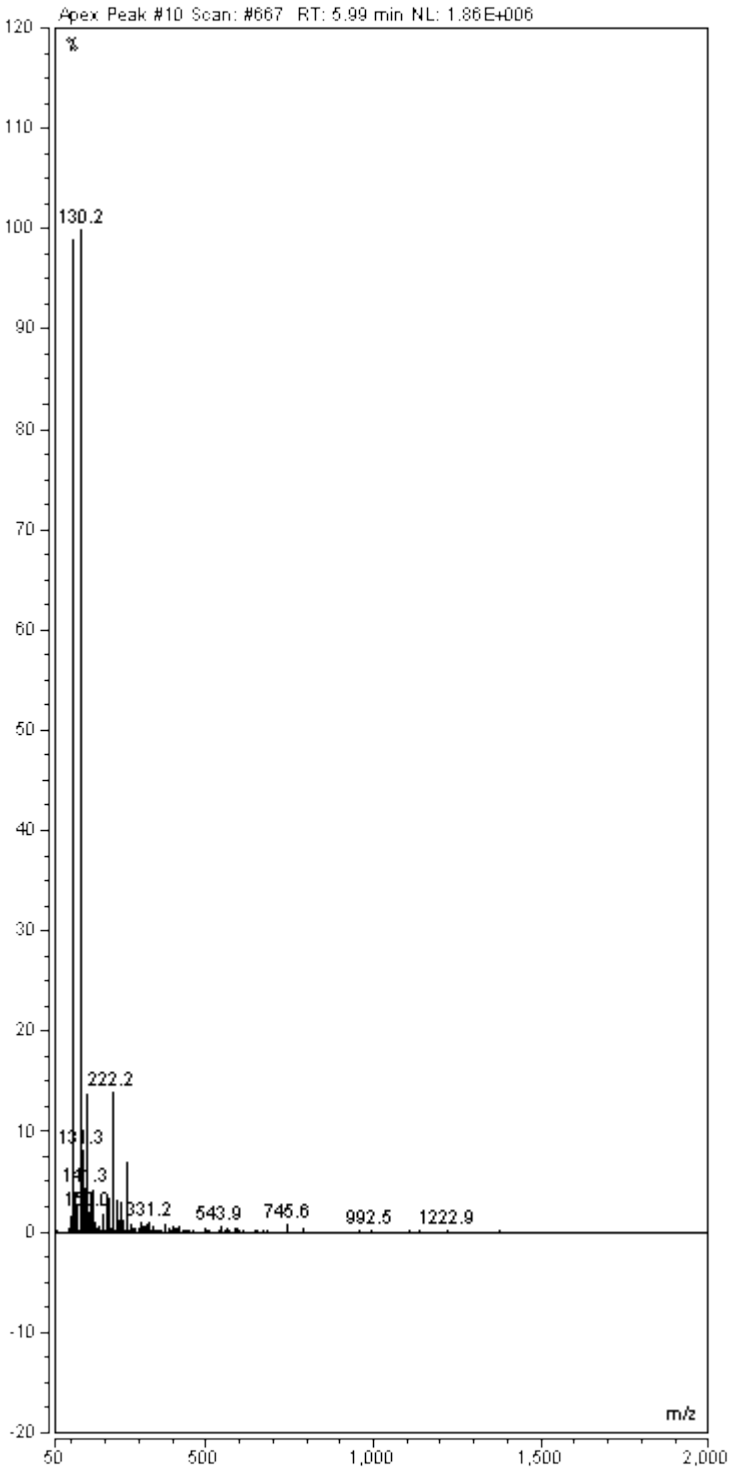


Mass spectrum (50-2000 m/z), ***D*-P$_4$**, m/z found: $[M+H]^+$ 628.1, $[M+Na]^+$ 650.0, $[2M+Na]^+$ 1277.4

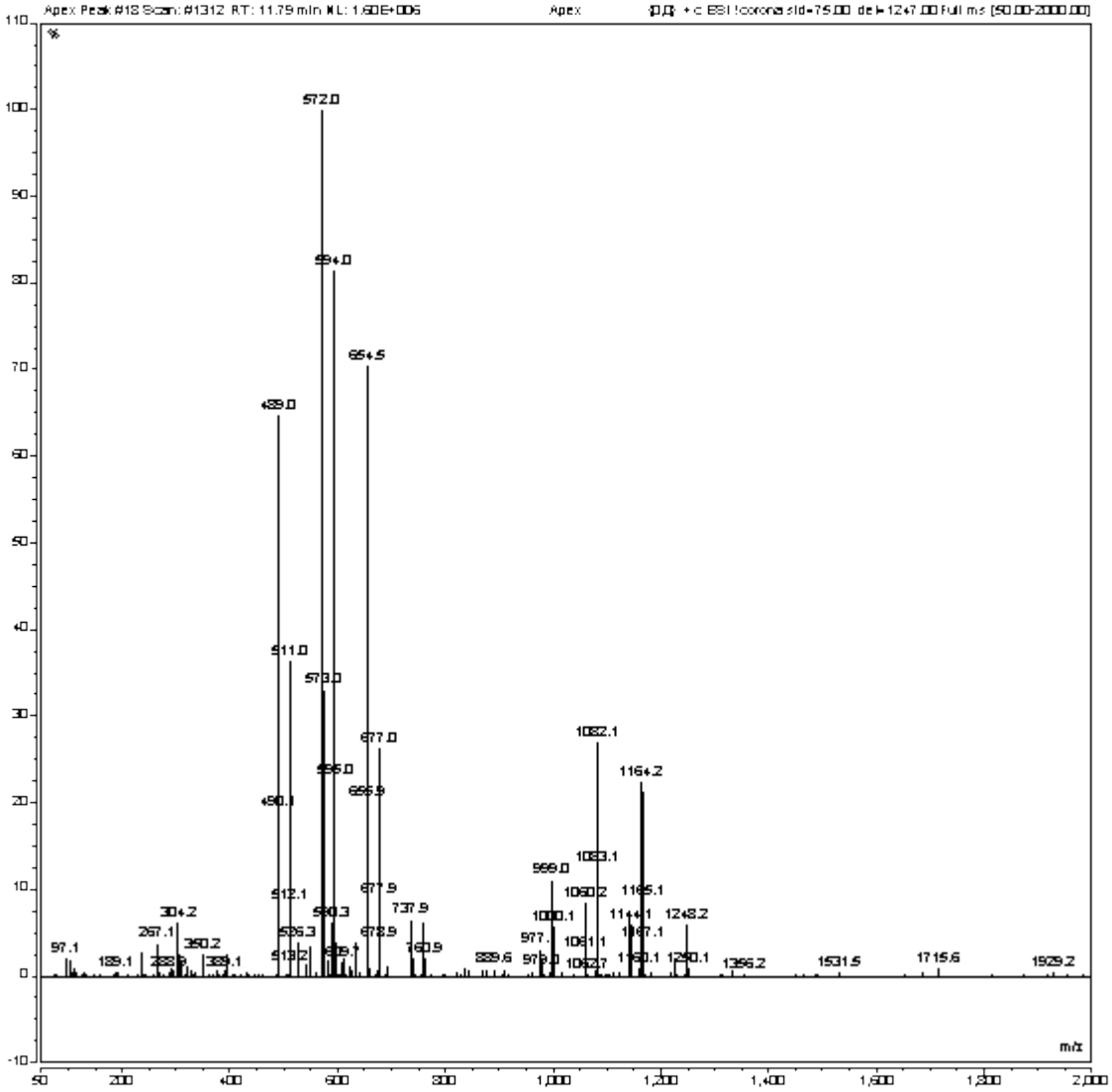
6.13 Synthesis of Fmoc-(Az)$_4$-NH$_2$ (14):



HPLC UV-Vis trace of peptide **Fmoc-(Az)$_4$-NH$_2$**, **Az$_4$**, rt = 11.787 min, 225 nm

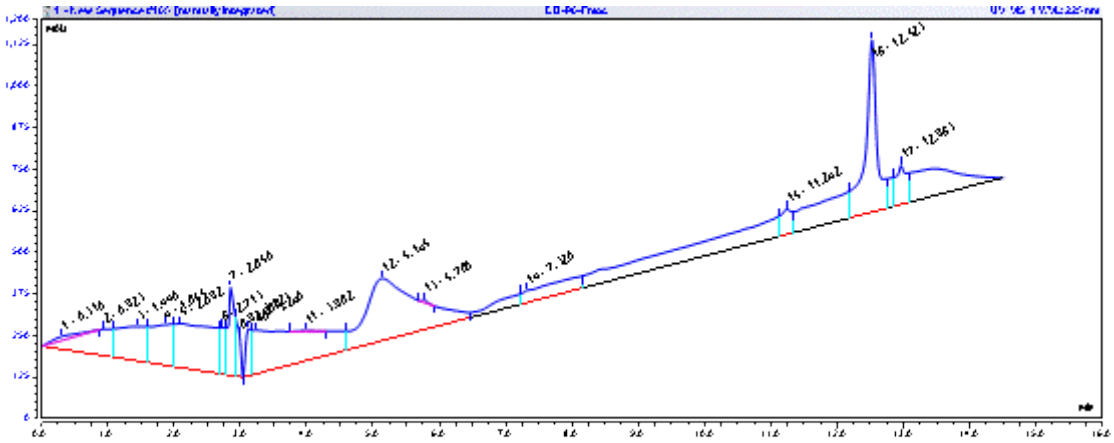


Mass spectrum (50-2000 m/z) of peak at 11.79 min, **Az_4**(without Fmoc), m/z calcd. for $[M+H]^+$ $C_{16}H_{24}N_5O_4^+$: 350.1823; found: $[M+H]^+$ not found

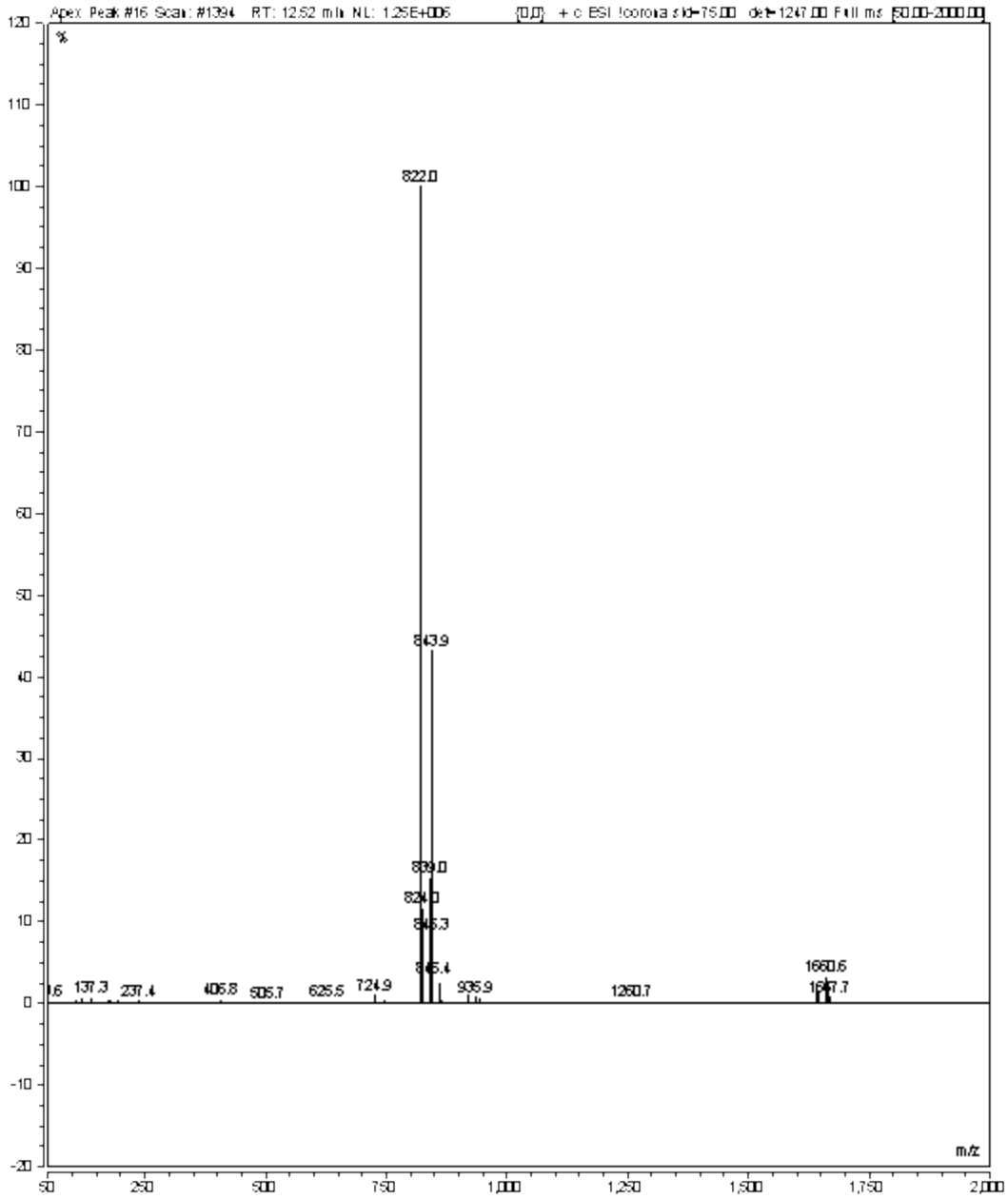


Mass spectrum (50-2000 m/z) of peak at 11.79 min, **Az$_4$**, m/z calcd. for $[M+H]^+$ $C_{31}H_{34}N_5O_6^+$: 572.2504; found: $[M+H]^+$ 572.0, $[M+Na]^+$ 594.0

6.14 Synthesis of Fmoc-(Pro)$_6$-NH$_2$ (15):

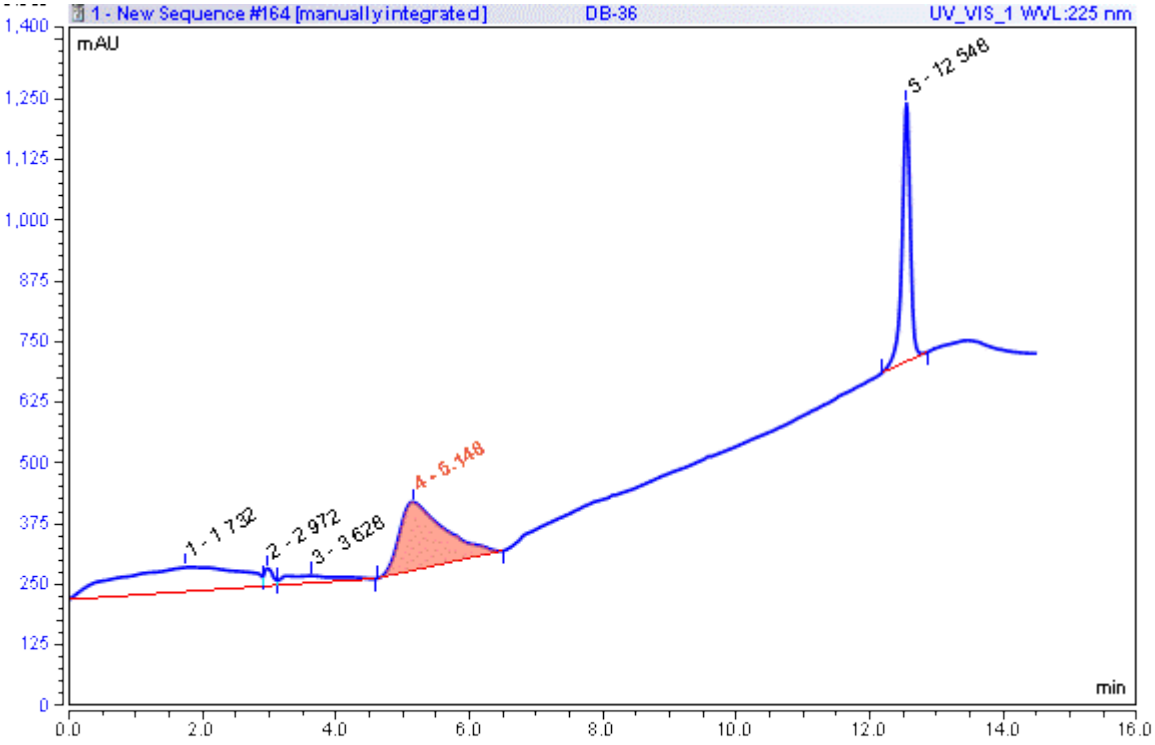


HPLC UV-Vis trace of peptide **Fmoc-(Pro)$_6$-NH$_2$**, **P$_6$**, rt = 12.523 min, 225 nm

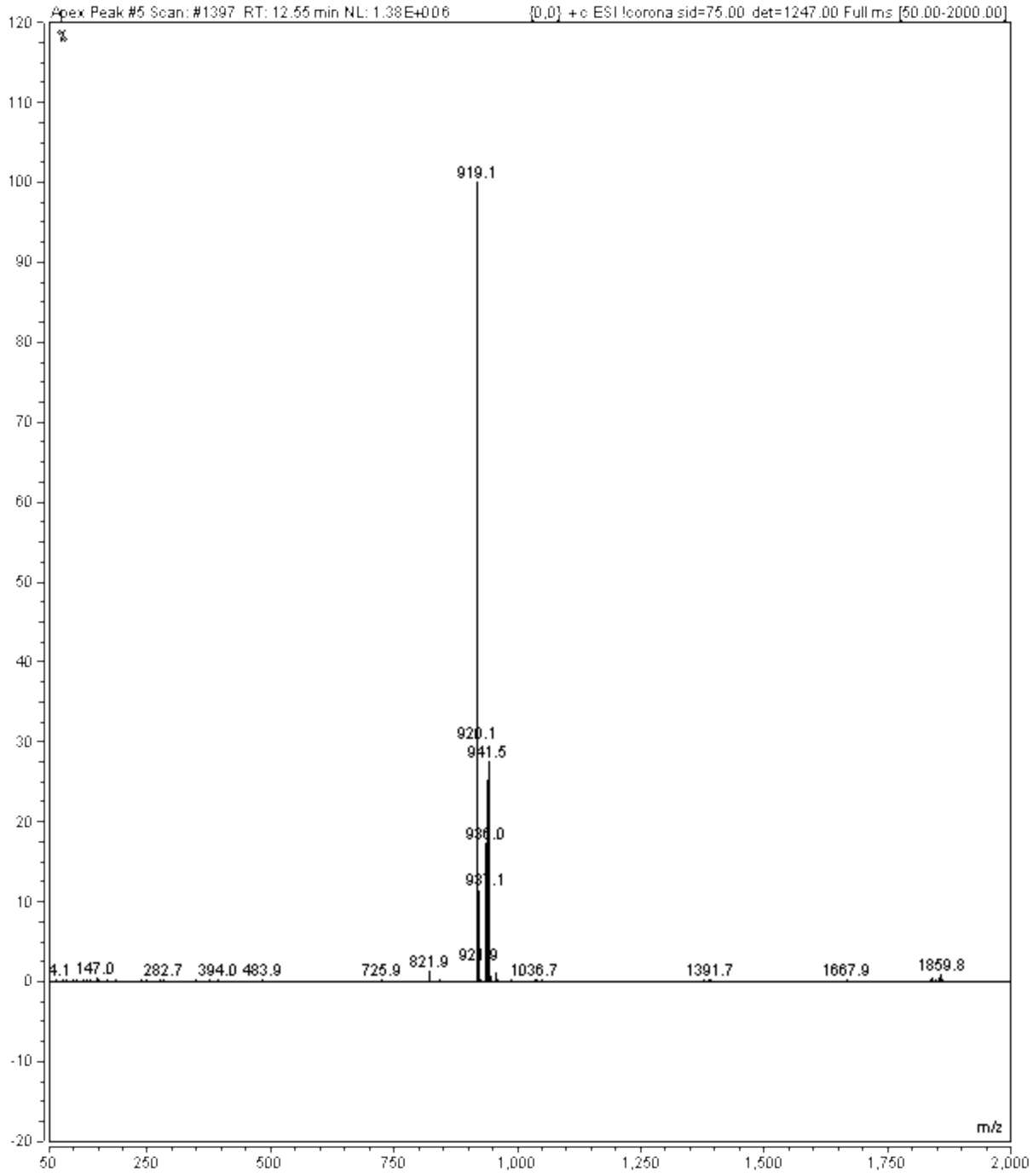


Mass spectrum (50-2000 m/z) of peak at 12.55 min, **P_6**, m/z calcd. for $[M+H]^+$ $C_{45}H_{56}N_7O_8^+$: 822.4185; found: $[M+H]^+$ 822.8, $[M+Na]^+$ 843.9

6.15 Synthesis of Fmoc-$(Pro)_7$-NH_2 (16):

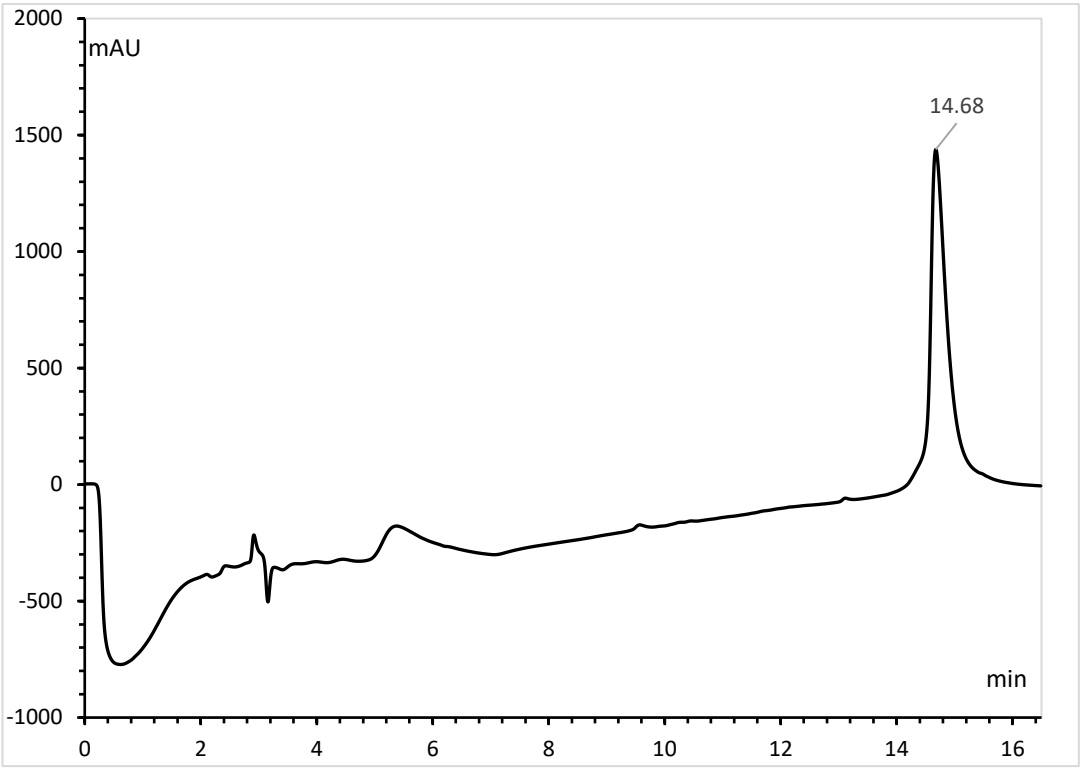


HPLC UV-Vis trace of peptide **Fmoc-$(Pro)_7$-NH_2**, **P_7**, rt = 12.548 min, 225 nm



Mass spectrum (50-2000 m/z) of peak at 12.55 min, **P_7,** m/z calcd. for $[M+H]^+$ $C_{50}H_{63}N_8O_9^+$: 919.4713; found: $[M+H]^+$ 919.1, $[M+Na]^+$ 941.5

6.16 Synthesis of Fmoc-$(Pro)_{13}$-NH_2 (18):



HPLC UV-Vis trace of peptide **Fmoc-$(Pro)_{13}$-NH_{13}, P_7**, rt = 14.68 min, 225 nm

Appendices: Chapter 3

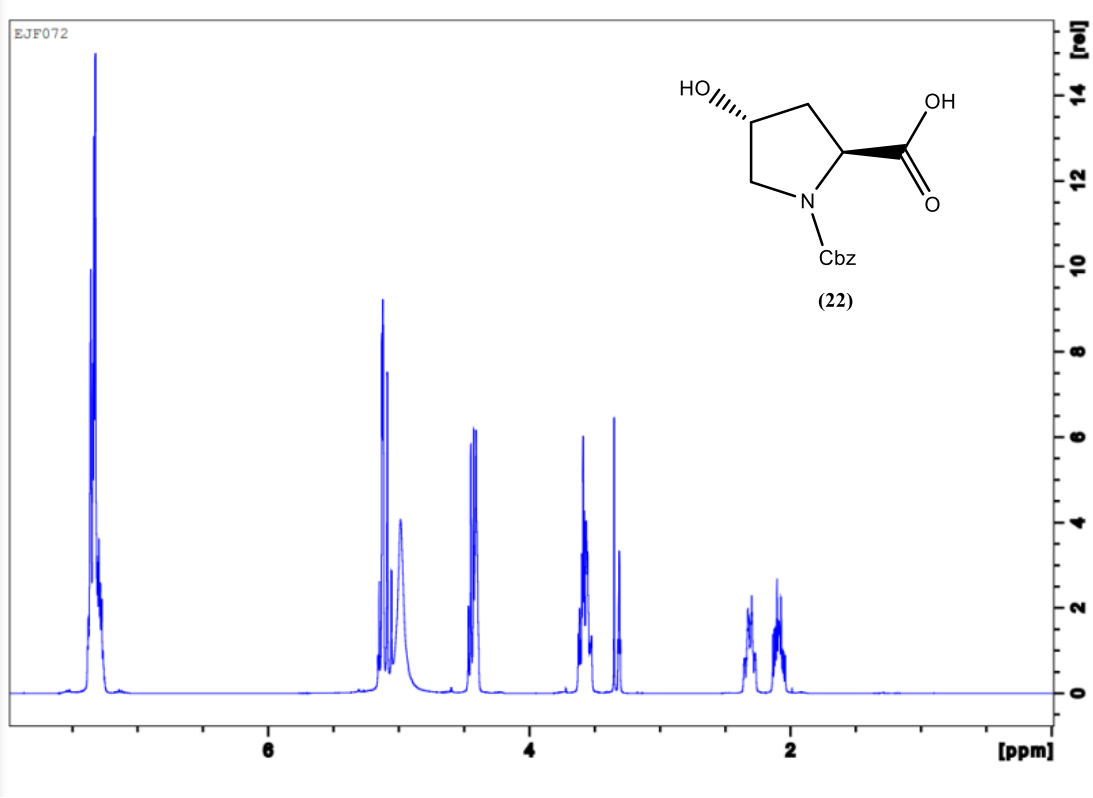
NMR spectroscopy. NMR spectra were collected on Bruker Avance II 400 MHz spectrometer, using Topspin software. 2D NMR experiments were measured on the Bruker Avance III 400 MHz at 298 K, using Topspin software. Chemical shifts are with reference to the residual solvent peak, with J values in Hz. For multiplicity of the peaks, the abbreviations used are (s) singlet, (d) doublet, (t) triplet, and (q) quartet.

LCMS: Compounds were separated *via* RP-HPLC using a HiChrom KR100 5C18 5263 column at 40 ^{o}C on a Dionex UltiMate 3000. Gradient: 5% B for 5 minutes then from 5% B to 100% B over 20 minutes, and held at 100% B for 5 minutes. Where A is Water (0.1% formic acid) and B is methanol (0.1% formic acid). Flow rate is 1.0 mL/min. Wavelength: 225 nm. The flow was directed into directed into the electrospray source of a Thermo Scientific MSQ Plus Mass Detector, operating in positive ion mode, at 75 kV and mass spectra recorded from 100-2000 m/z.

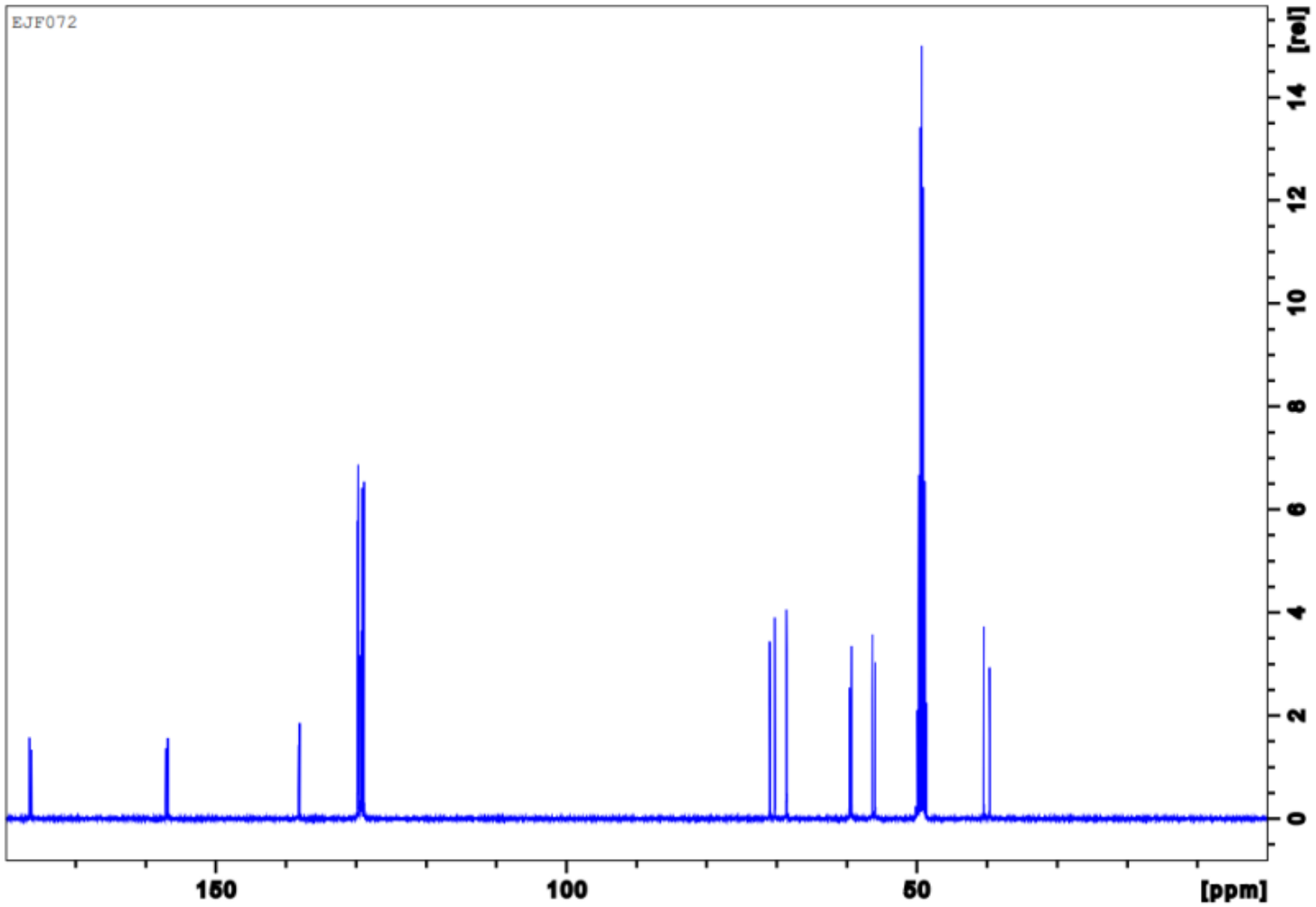
7.1 Monomers

7.1.1 – Synthesis of 4-position carboxylic acid functionalised *trans*-fmoc-proline (25):

Synthesis of benzyl-1-(2S,4R)-4-hydroxypyrrolidine-1-carboxylate-2-carboxylic acid (22):

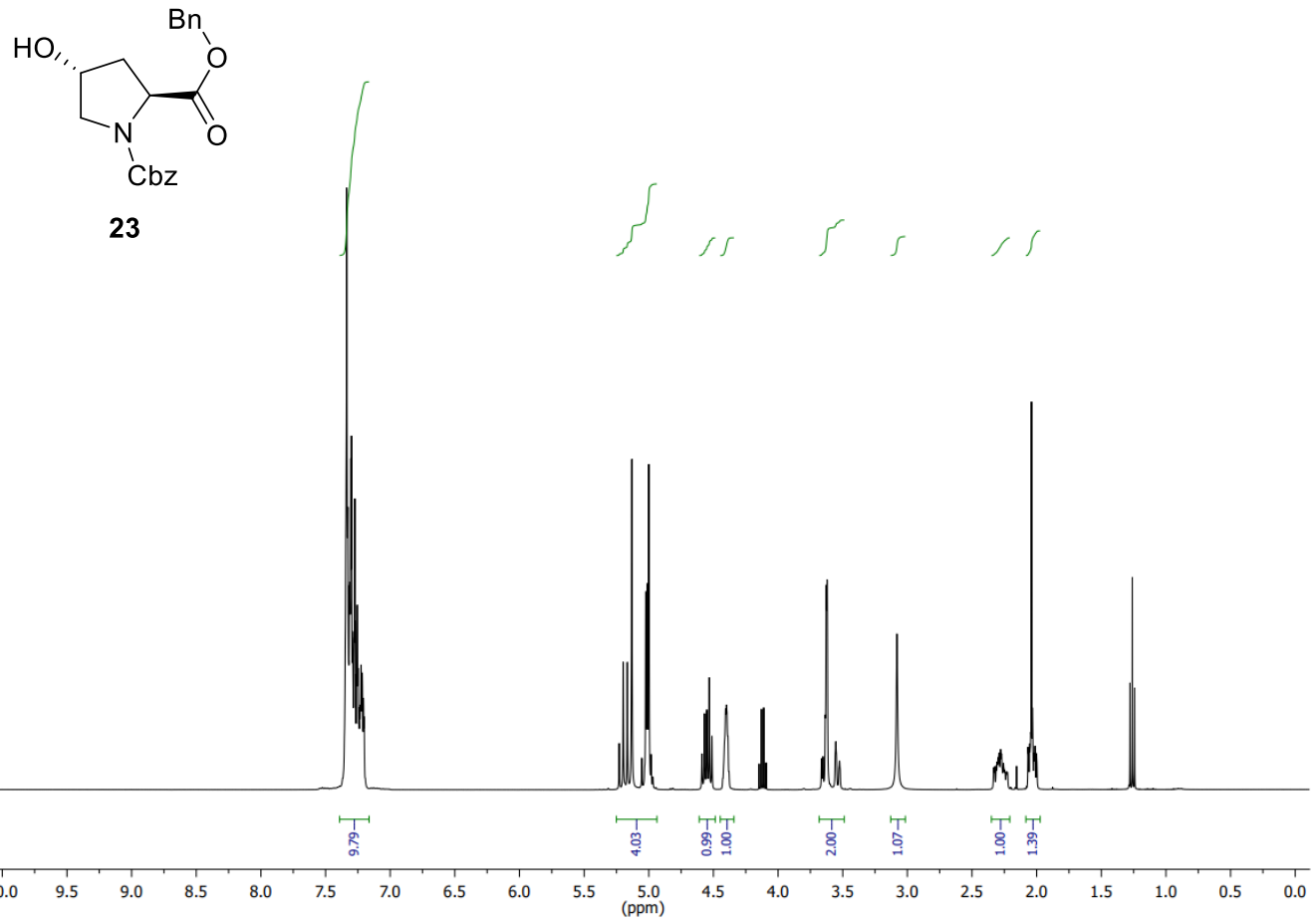


^{1}H NMR of **22 (400 MHz, MeOD)**

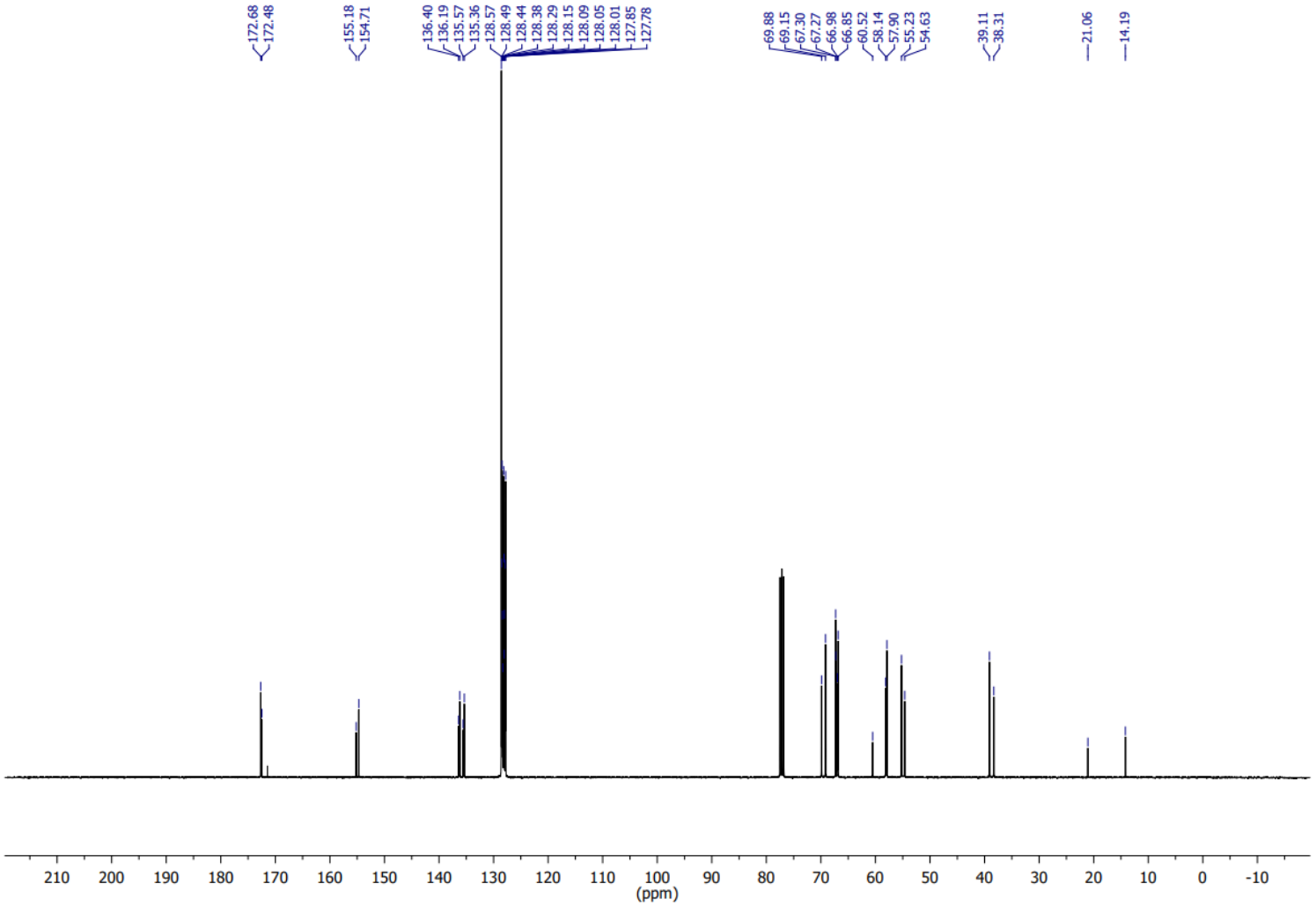


^{13}C NMR of **22 (101 MHz, MeOD)**

Synthesis of dibenzyl (2S,4R)-4-hydroxypyrrolidine-1,2-dicarboxylate (23):

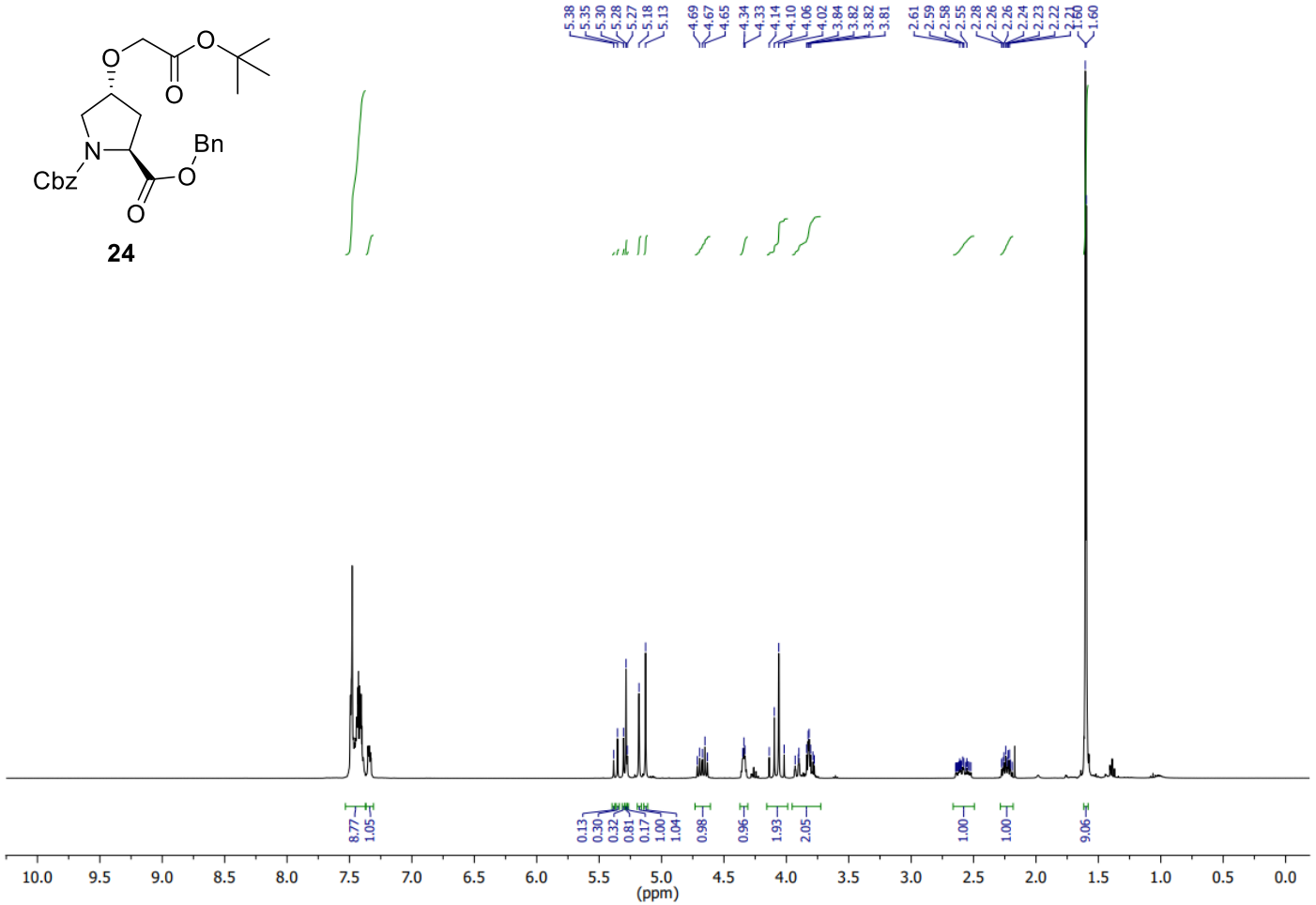


^{1}H NMR of **23 (400 MHz, $CDCl_3$)**

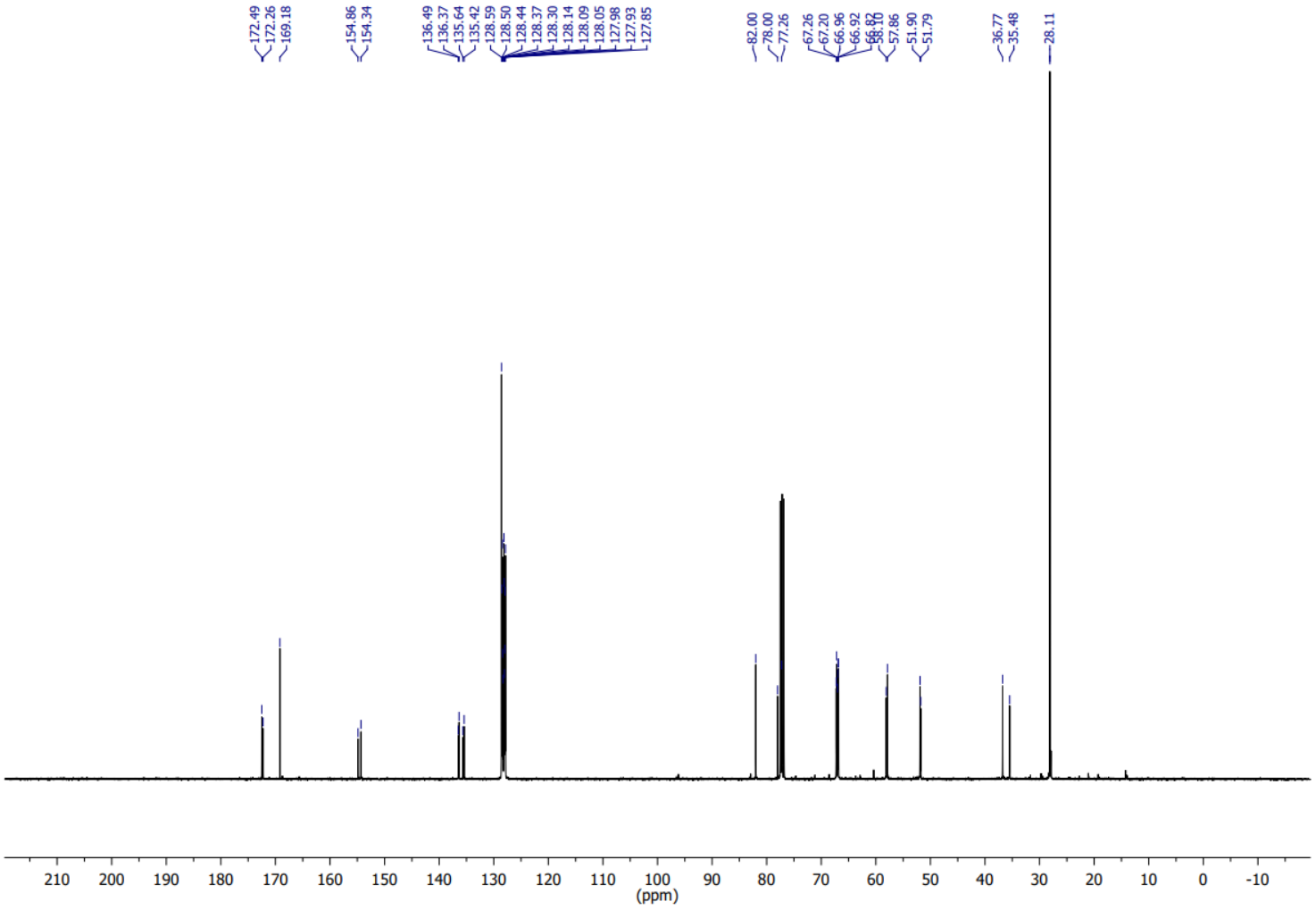


^{13}C NMR of **23 (101 MHz, $CDCl_3$)**

Synthesis of dibenzyl (2S,4R)-4-(2-(tert-butoxy)-2-oxoethoxy)pyrrolidine-1,2-dicarboxylate (24):



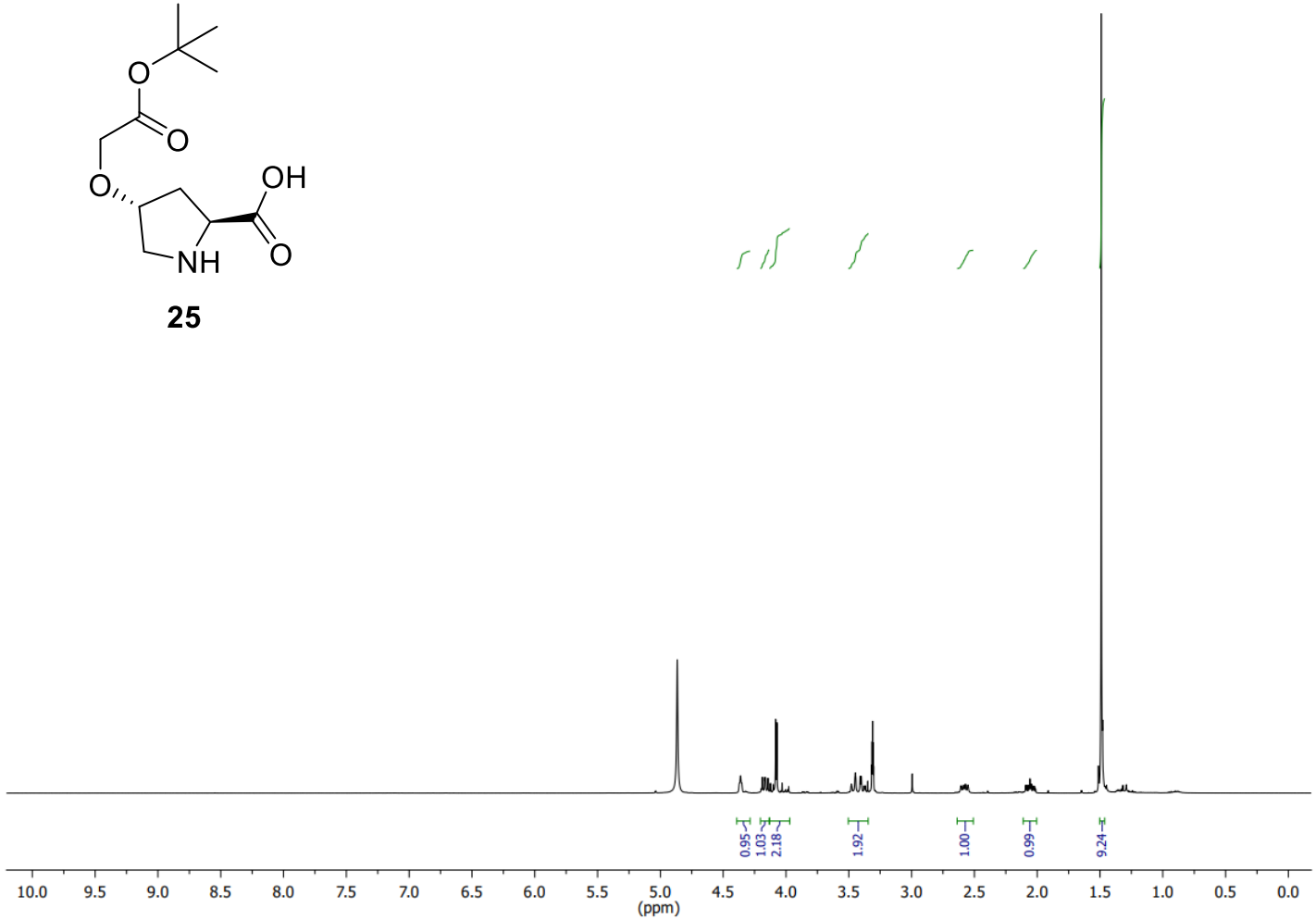
^{1}H NMR of **24 (400 MHz, $CDCl_3$)**



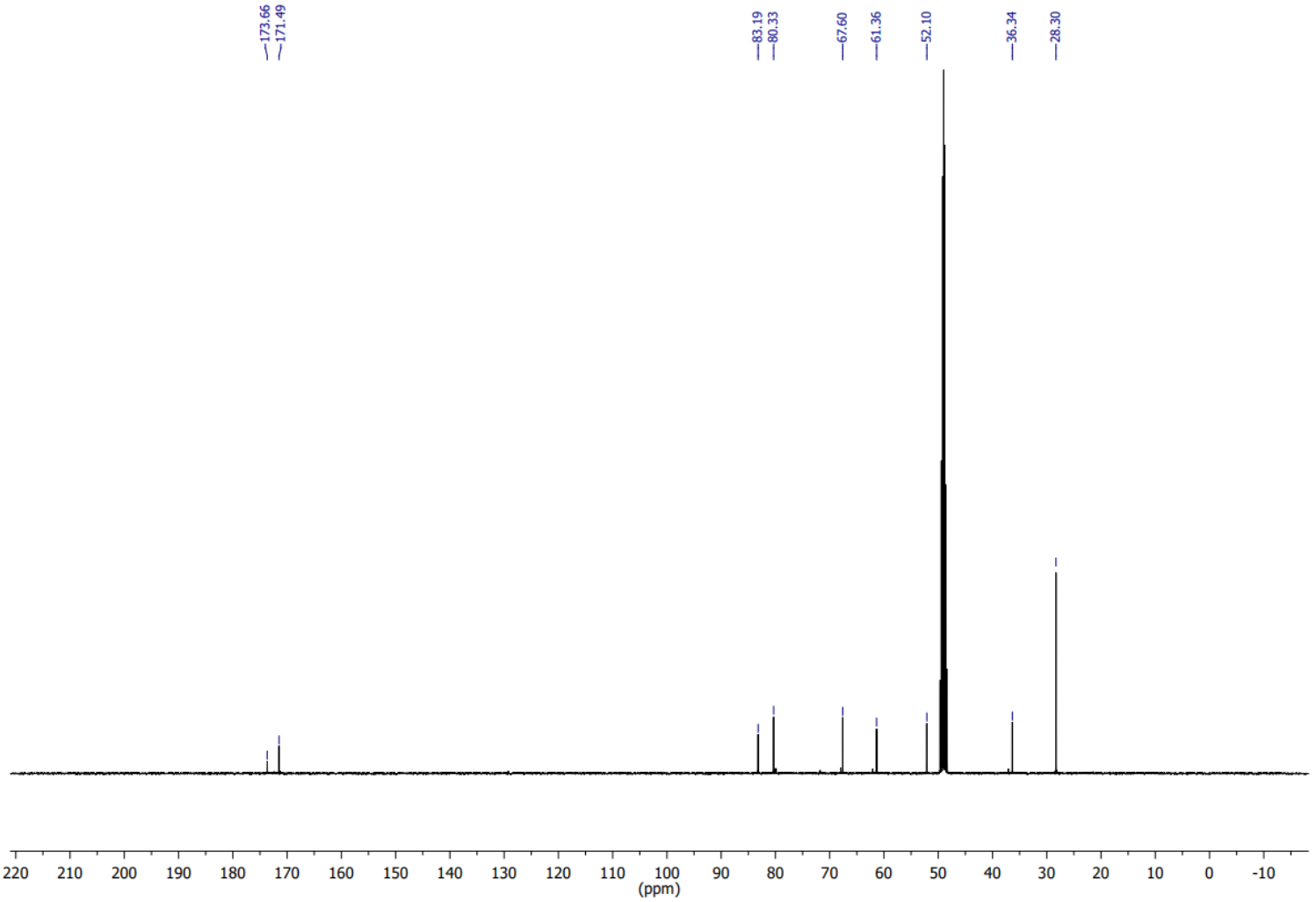
^{13}C NMR of **24 (101 MHz, $CDCl_3$)**

Synthesis of (2S,4R)-4-(2-(tert-butoxy)-2-oxoethoxy)pyrrolidine-2-carboxylic acid (25):

25

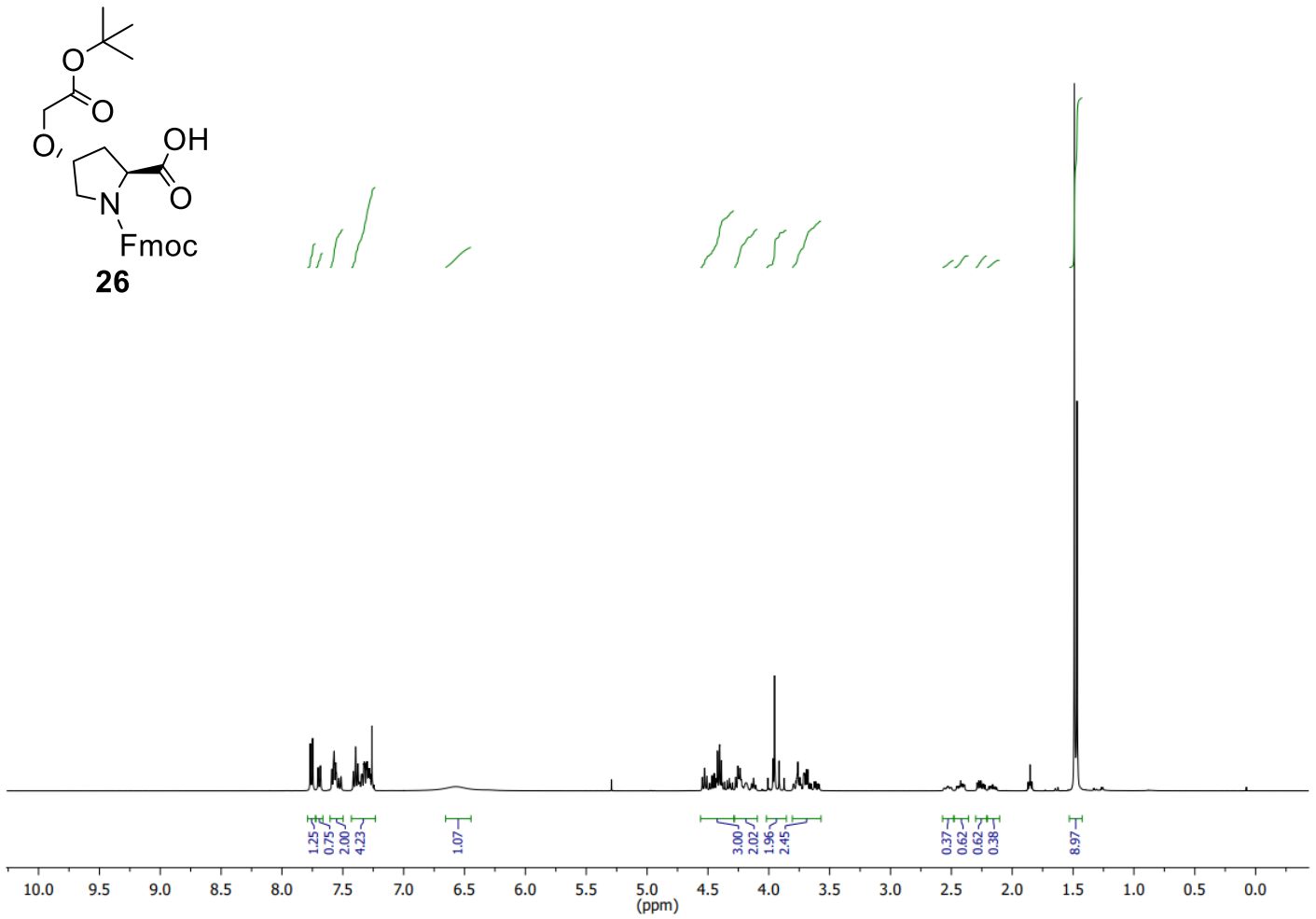


^{1}H NMR of **25 (400 MHz, MeOD)**

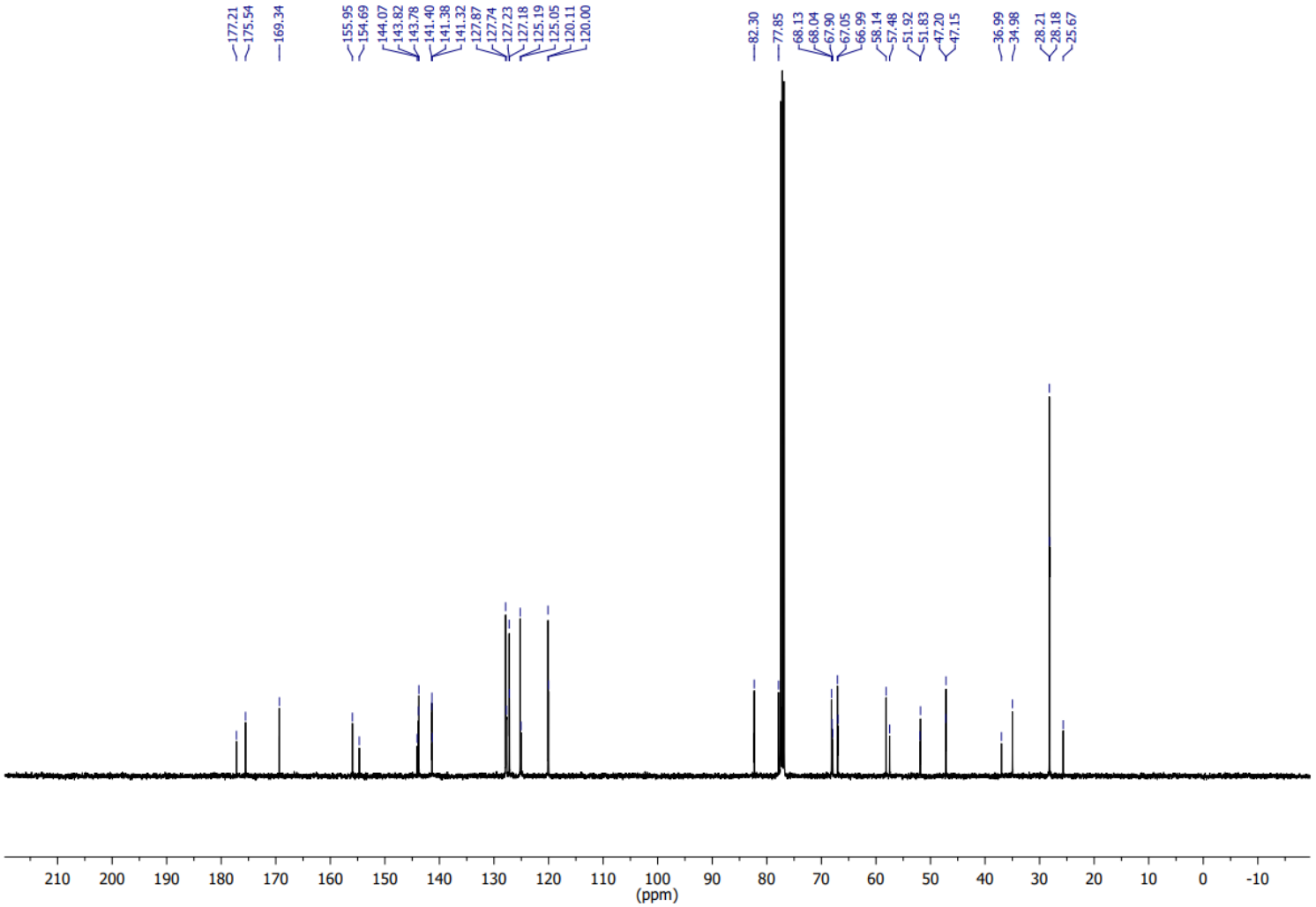


^{13}C NMR of **25 (101 MHz, MeOD)**

Synthesis of (2S,4R)-1-(((9H-fluoren-9-yl)methoxy)carbonyl)-4-(2-(tert-butoxy)-2-oxoethoxy)pyrrolidine-2-carboxylic acid (26):



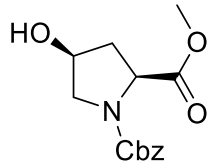
^{1}H NMR of **26 (400 MHz, $CDCl_3$)**



^{13}C NMR of **26 (101 MHz, $CDCl_3$)**

7.1.2 Synthesis of (*2S,4R*)-1-(((9H-fluoren-9-yl)methoxy)carbonyl)-4-(4-(*tert*-butoxycarbonyl)phenoxy)pyrrolidine-2-carboxylic acid (35):

Synthesis of 1-benzyl 2-methyl (2S,4S)-4-hydroxypyrrolidine-1,2-dicarboxylate (29):

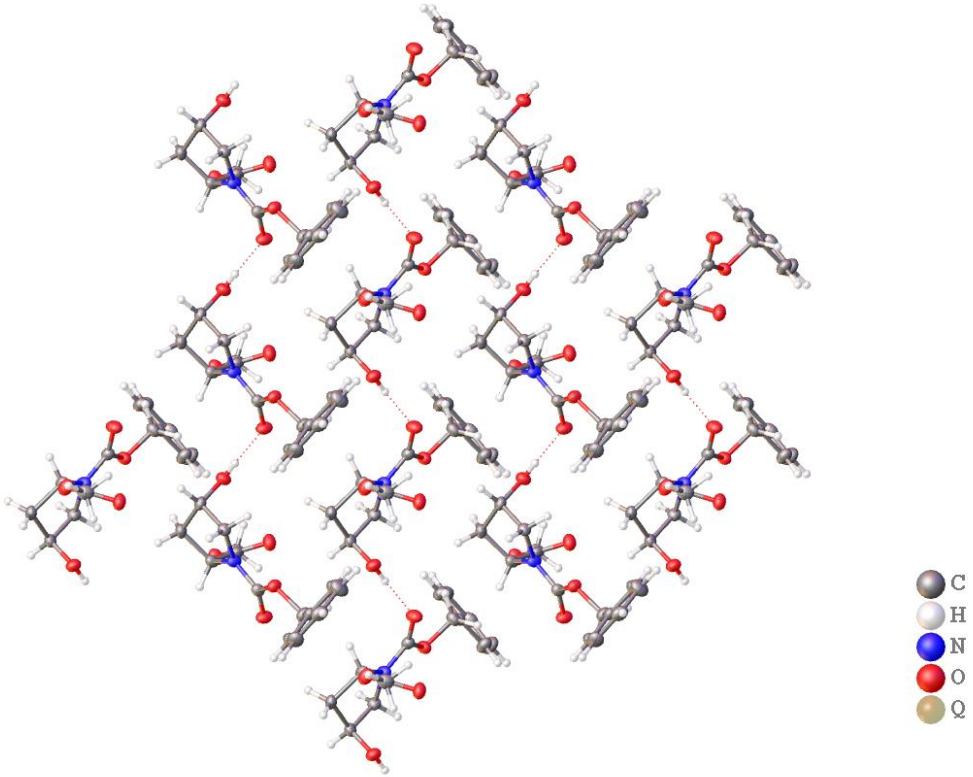


29

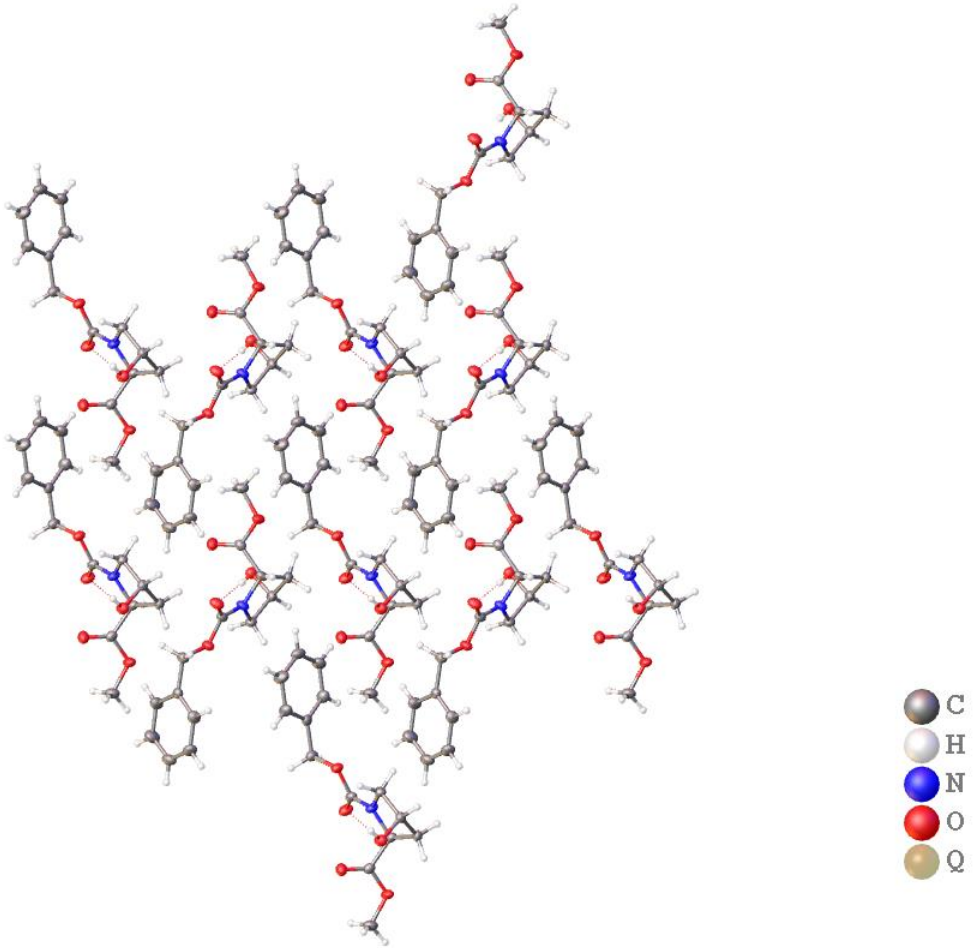
Table 1 Crystal data and structure refinement for *cis*-Cbz-Hyp-OMe.

Identification code	cis-Cbz-Hyp-OMe
Empirical formula	$C_{14}H_{17}NO_5$
Formula weight	279.28
Temperature/K	150.00(10)
Crystal system	monoclinic
Space group	$P2_1$
a/Å	6.7483(3)
b/Å	10.5173(4)
c/Å	9.6302(4)
α/°	90
β/°	100.951(4)
γ/°	90
Volume/Å^3	671.05(5)
Z	2
ρ_{calc}g/cm^3	1.382
μ/mm^{-1}	0.105
F(000)	296.0
Crystal size/mm^3	0.35 × 0.15 × 0.1
Radiation	Mo Kα (λ = 0.71073)

2Θ range for data collection/°	6.806 to 59.48
Index ranges	$-9 \leq h \leq 8$, $-13 \leq k \leq 14$, $-13 \leq l \leq 13$
Reflections collected	15236
Independent reflections	3440 [R_{int} = 0.0346, R_{sigma} = 0.0314]
Data/restraints/parameters	3440/1/183
Goodness-of-fit on F^2	1.041
Final R indexes [I>=2σ (I)]	R_1 = 0.0362, wR_2 = 0.0801
Final R indexes [all data]	R_1 = 0.0413, wR_2 = 0.0840
Largest diff. peak/hole / e $Å^{-3}$	0.17/-0.19
Flack parameter	0.0(4)
CCDC No.	2290760

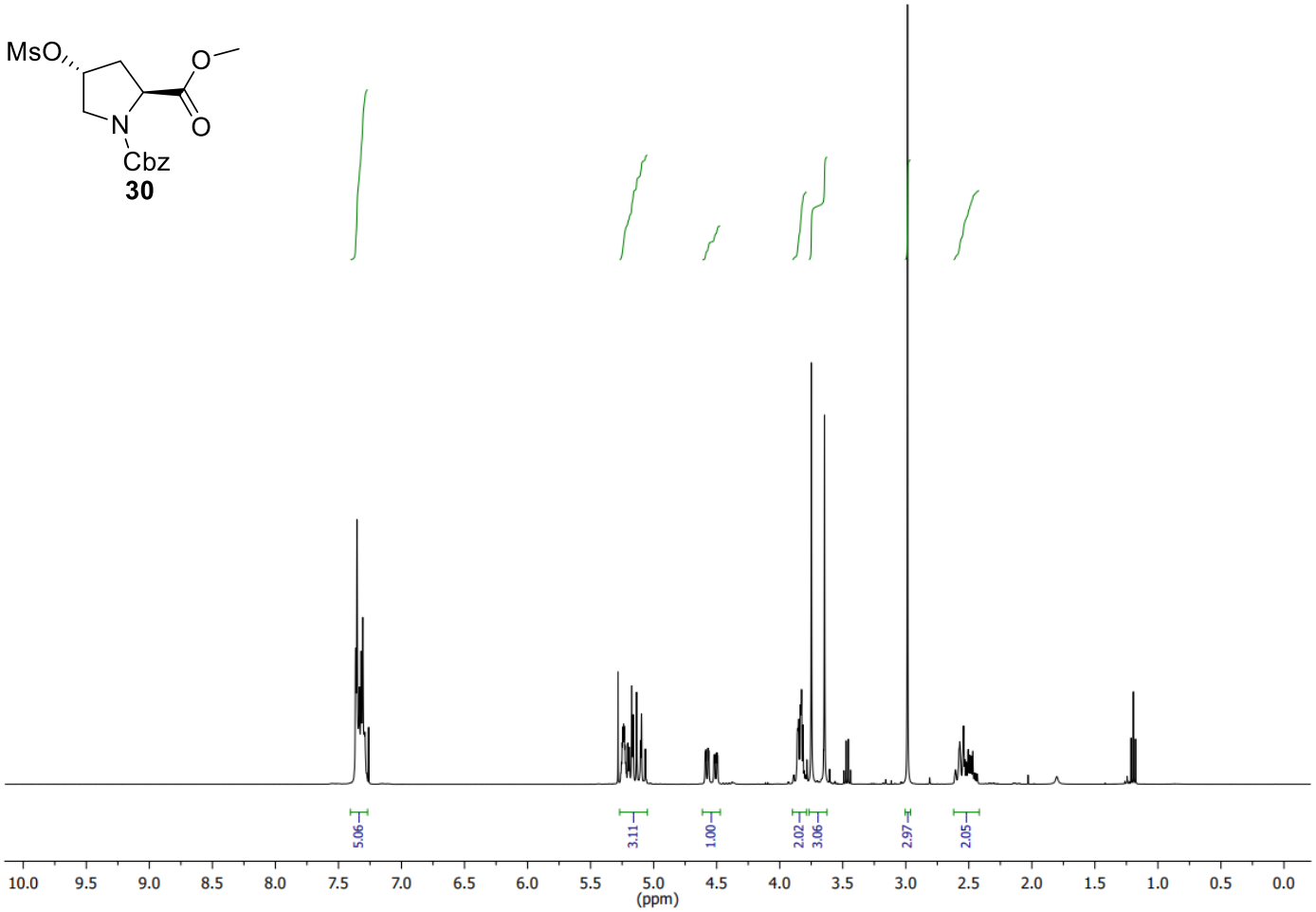


Crystal structure of compound **29**, viewed along the c-axis, 50% ellipsoids (Olex 2)



Crystal structure of compound **29**, viewed along the a-axis, 50% ellipsoids (Olex 2)

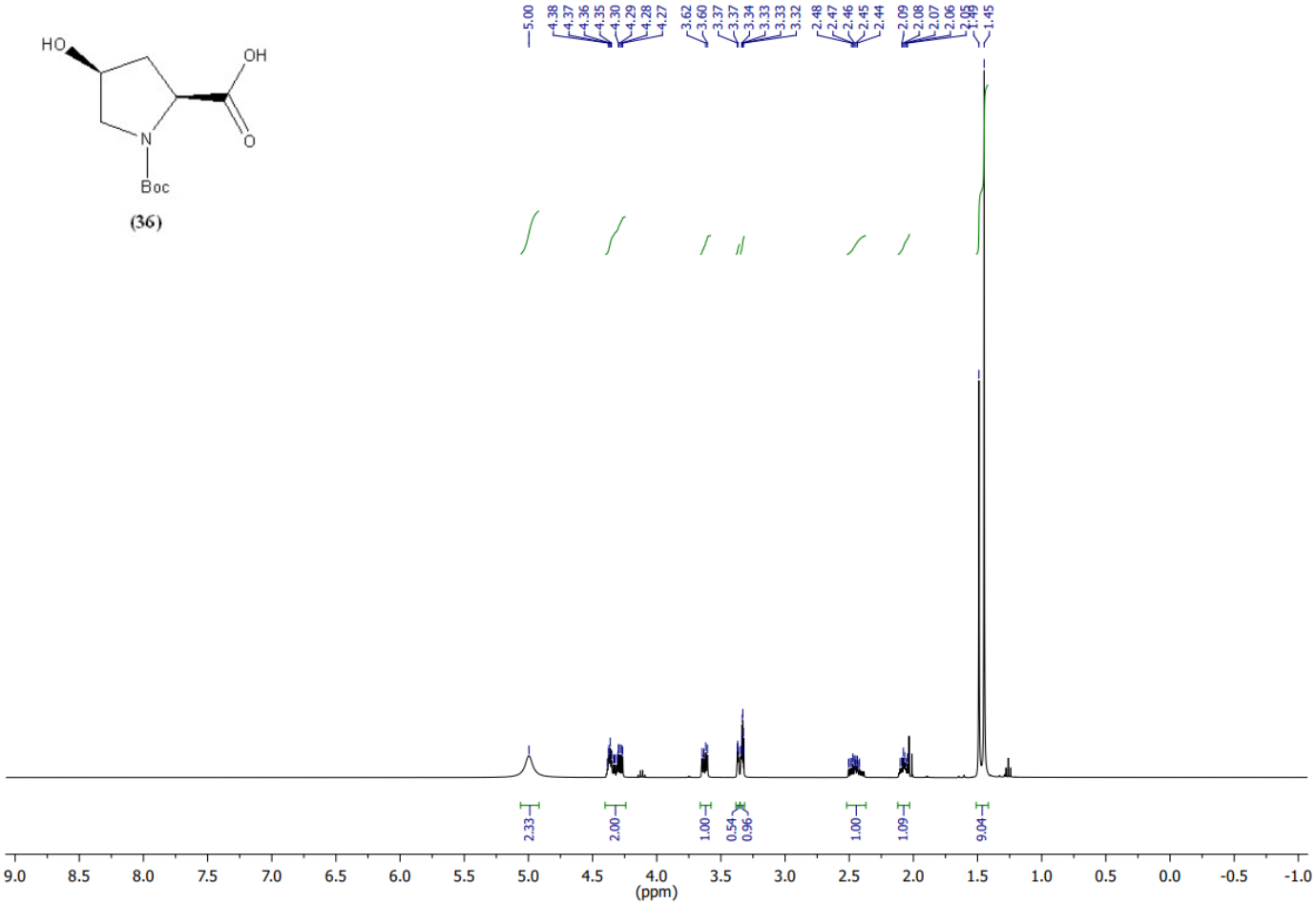
Synthesis of 1-benzyl 2-methyl (*2S,4R*)-4-((methylsulfonyl)oxy)pyrrolidine-1,2-dicarboxylate (30):



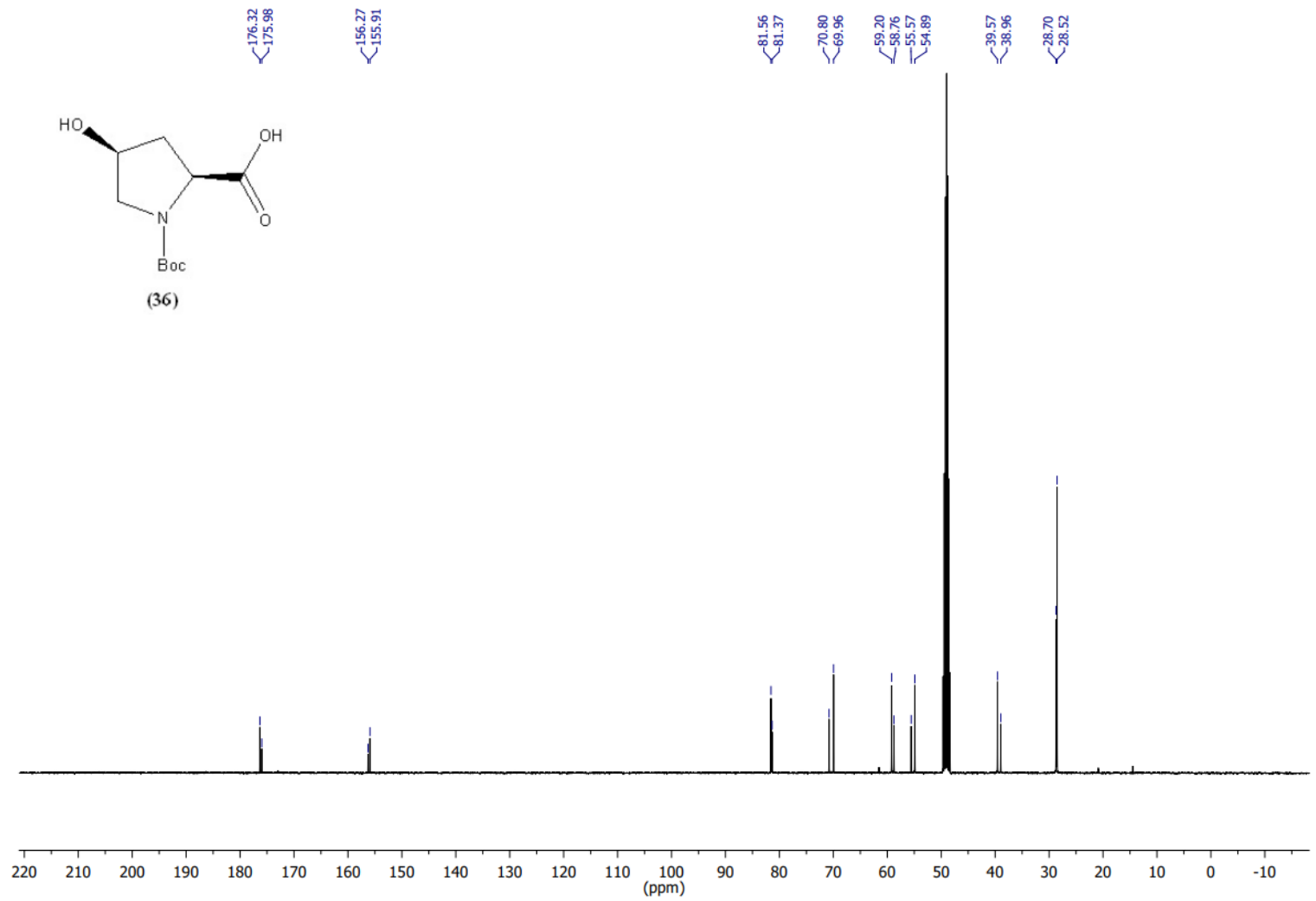
^{1}H NMR of **30** (400 MHz, $CDCl_3$)

7.1.3 Synthesis of (*2S,4R*)-1-(((9H-fluoren-9-yl)methoxy)carbonyl)-4-(4-(methoxycarbonyl)phenoxy)pyrrolidine-2-carboxylic acid (42):

Synthesis of (2S,4S)-1-(*tert*-butoxycarbonyl)-4-hydroxypyrrolidine-2-carboxylic acid (36):

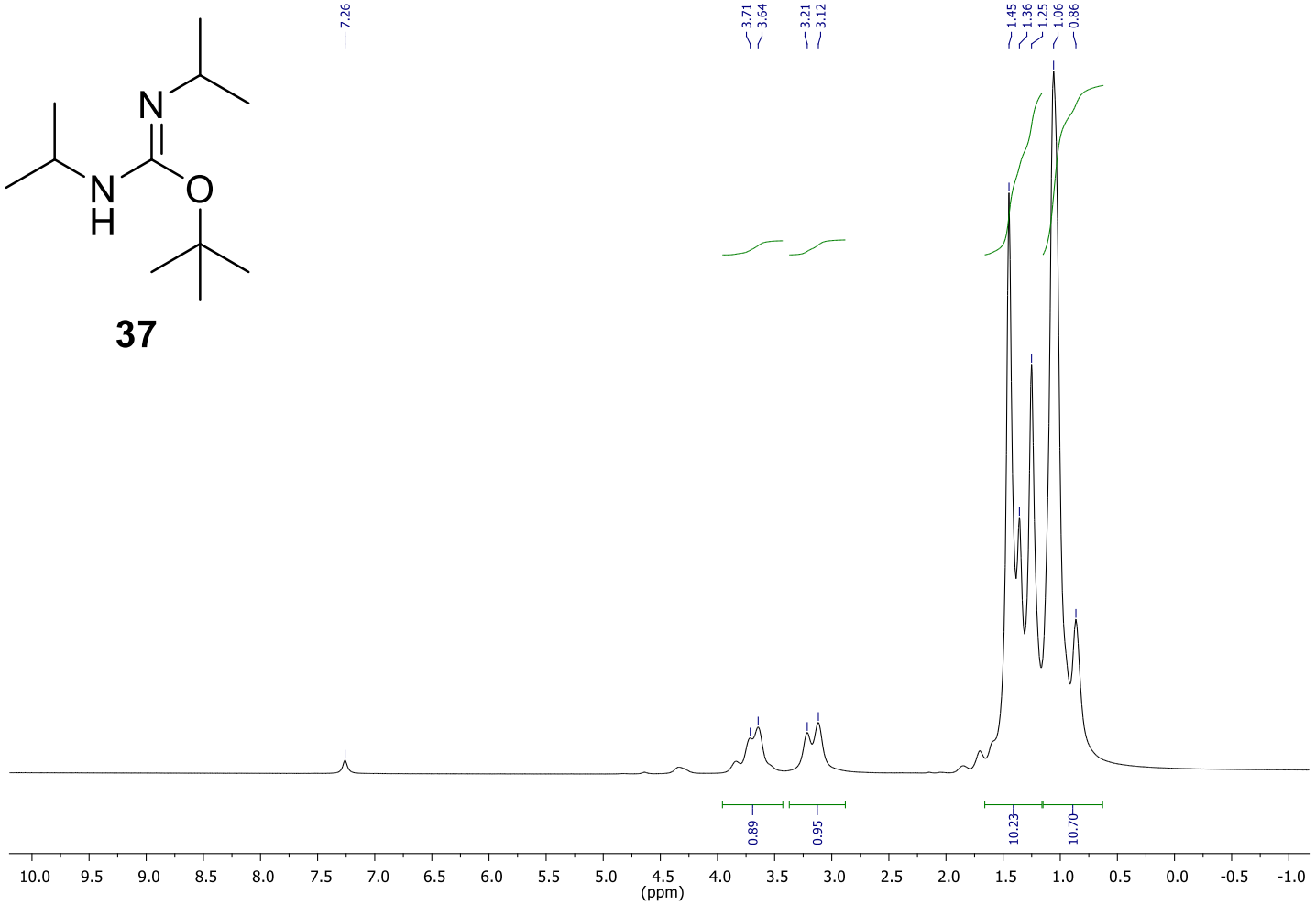


^{1}H NMR of **36** (400 MHz, MeOD)



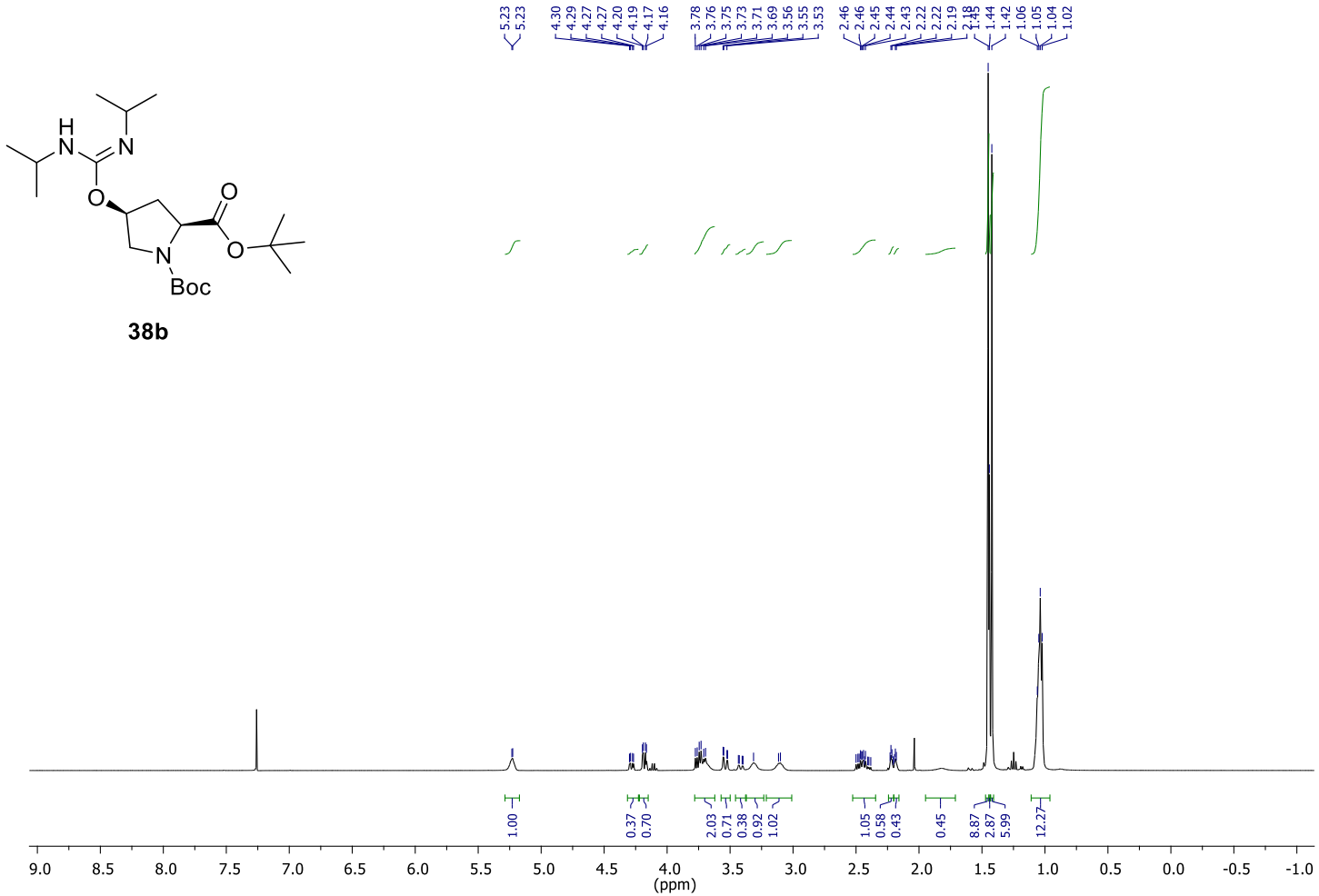
^{13}C NMR of **36** (101 MHz, MeOD)

7.1.4 Synthesis of *tert*-butyl *N,N*'-diisopropylcarbamimidate (37):

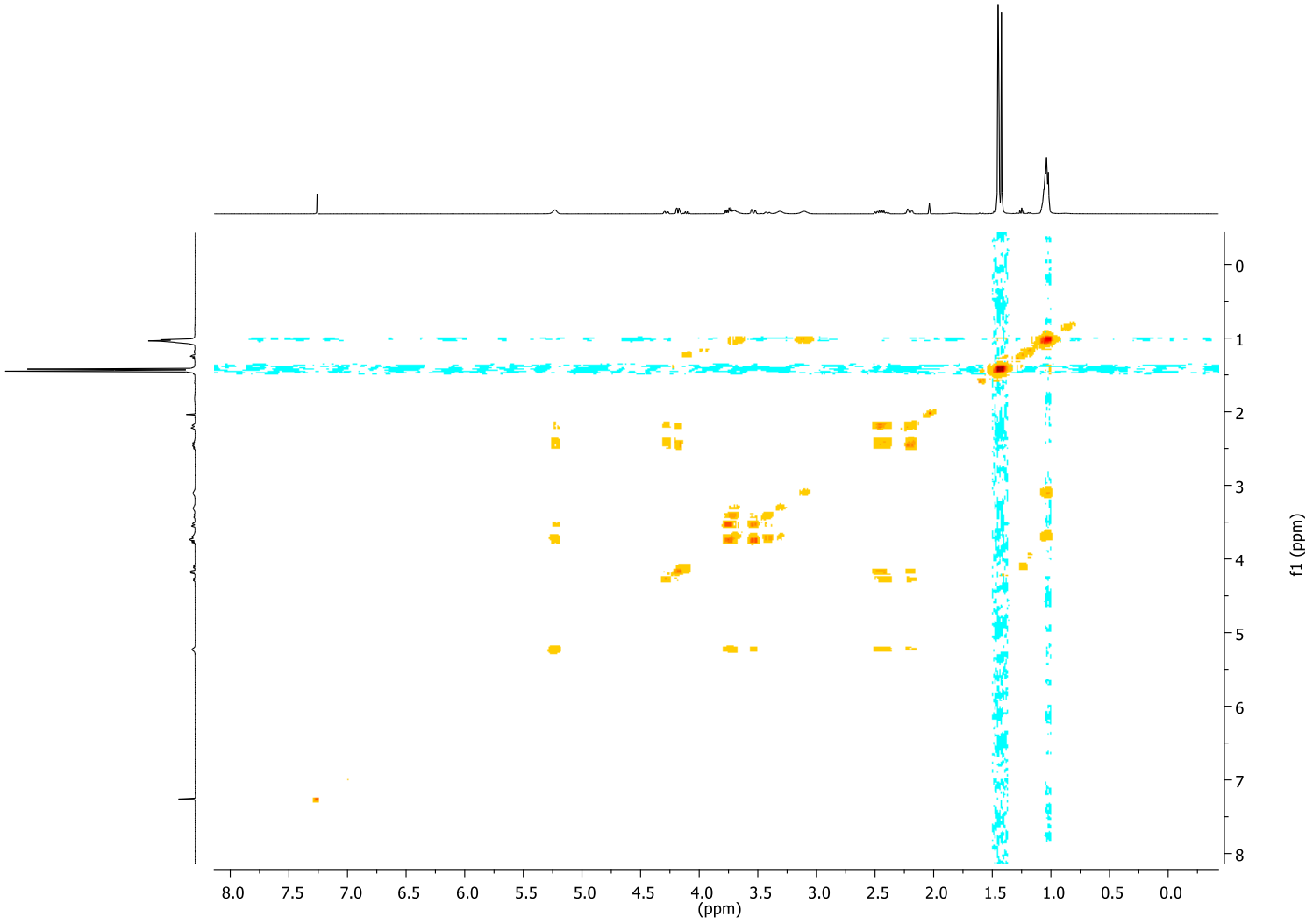


^{1}H NMR spectrum of **37** (400 MHz, $CDCl_3$)

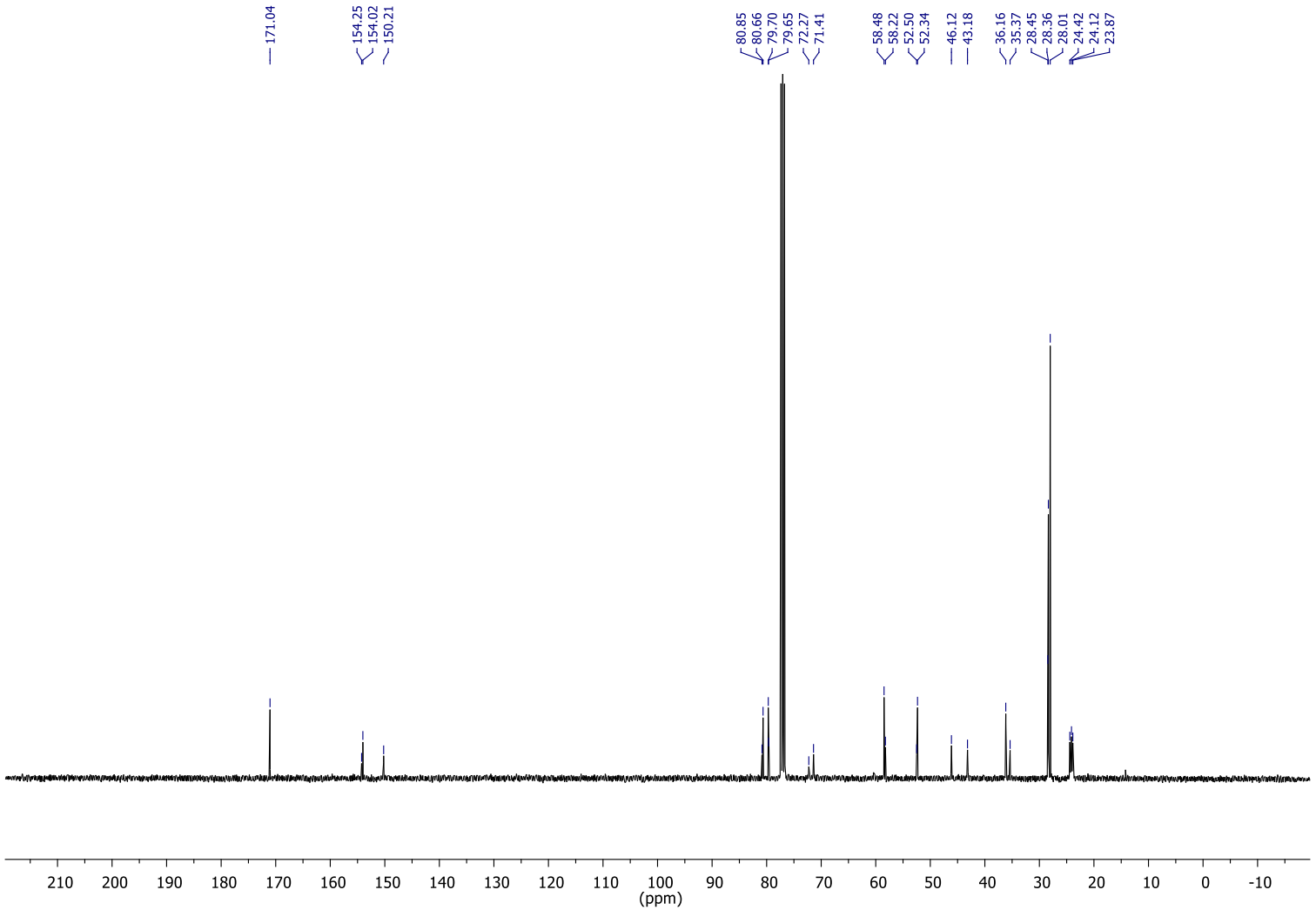
7.1.5 Synthesis of di-*tert*-butyl (2S,4S)-4-(((E)-*N,N*'-diisopropylcarbamimidoyl)oxy)pyrrolidine-1,2-dicarboxylate (38b):



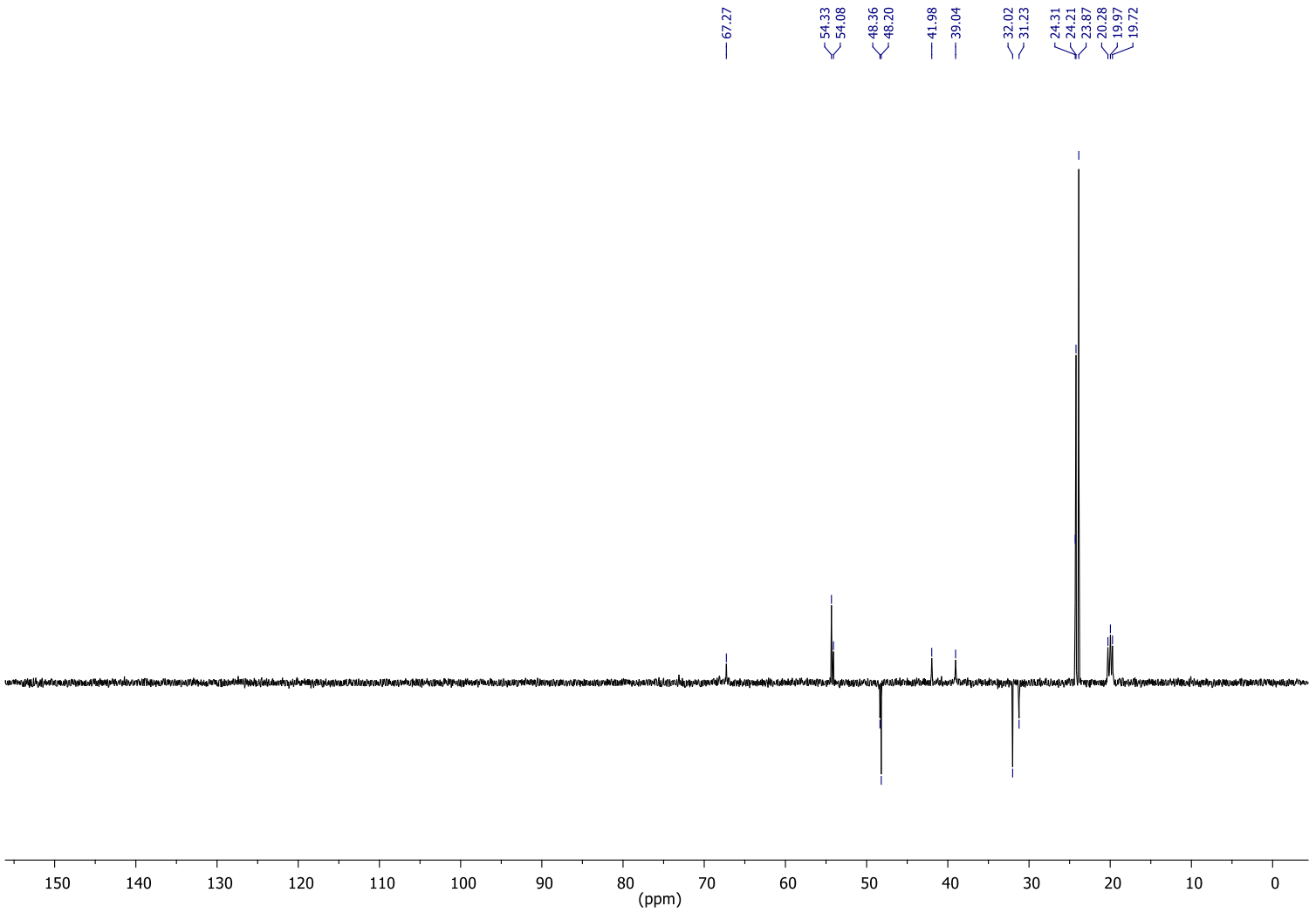
^{1}H NMR of **38b** (400 MHz, $CDCl_3$)



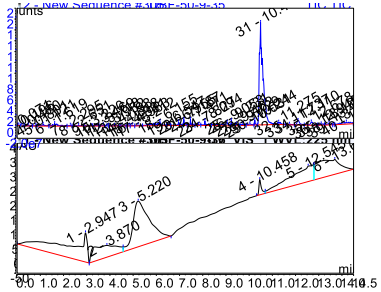
^{1}H COSY NMR of **38b** (400 MHz, $CDCl_3$)



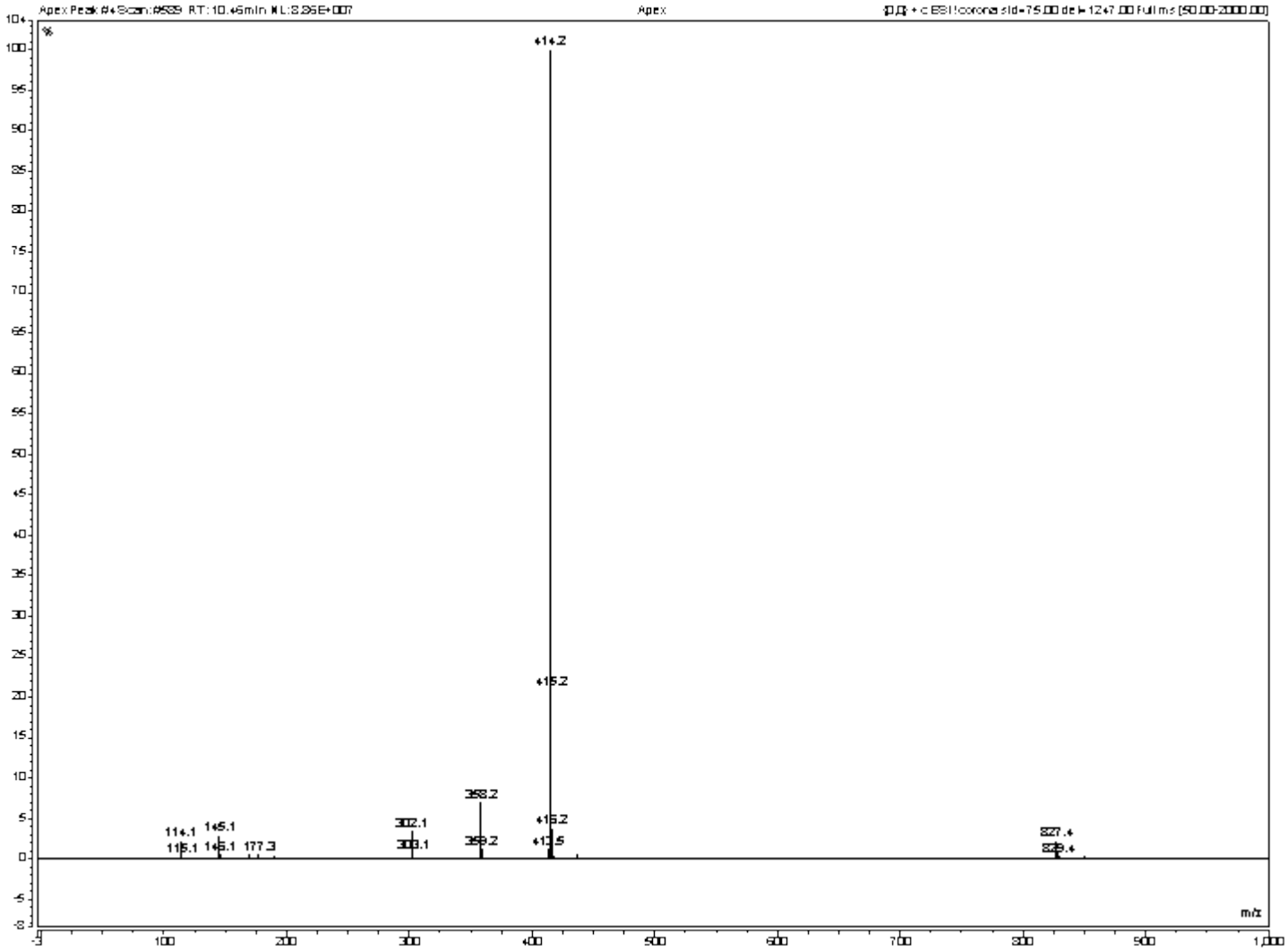
^{13}C NMR of **38b** (101 MHz, $CDCl_3$)



^{13}C DEPT NMR of **38b** (101 MHz, $CDCl_3$)

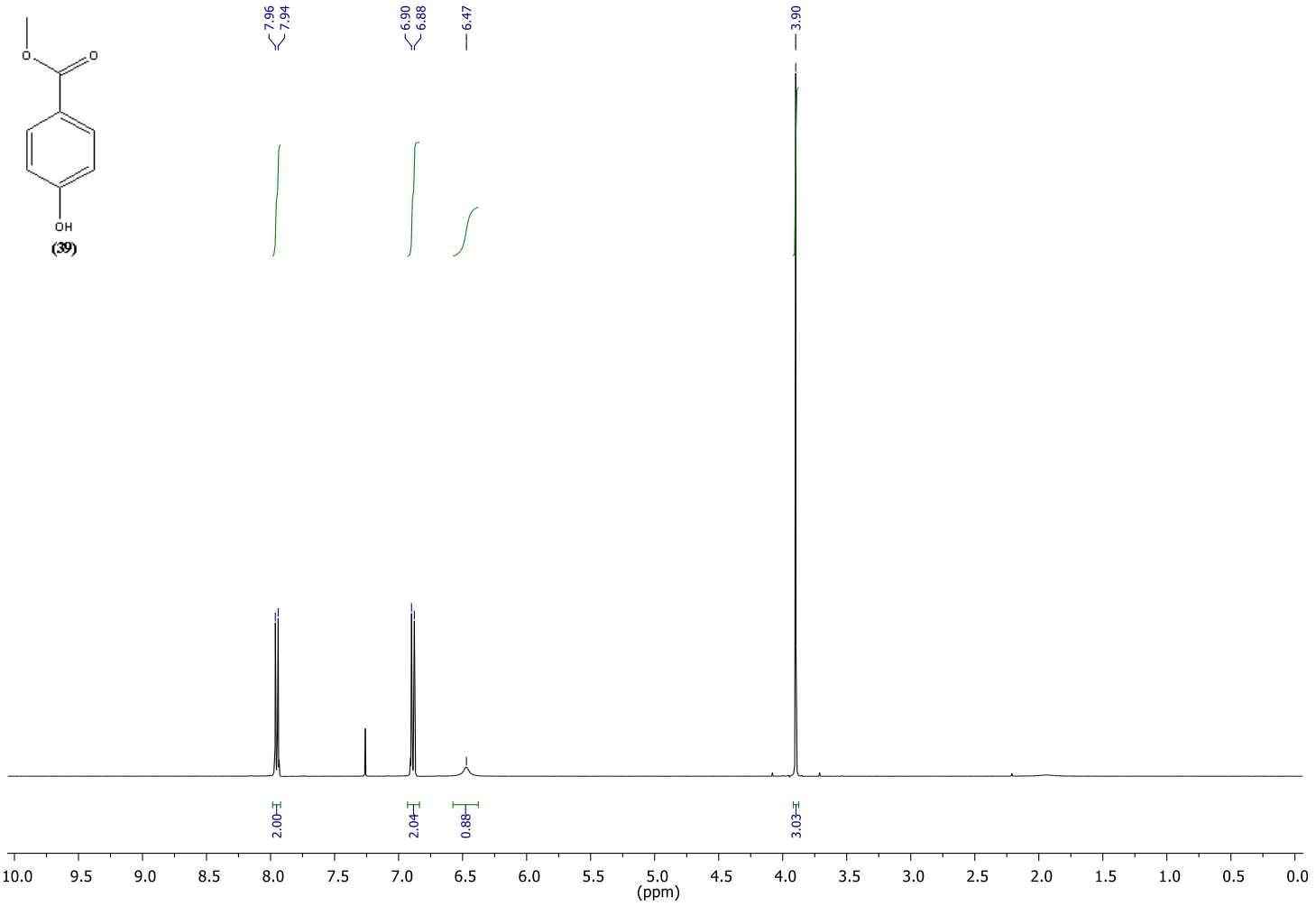


HPLC UV-Vis spectrum of **38**, rt = 10.468, 225 nm

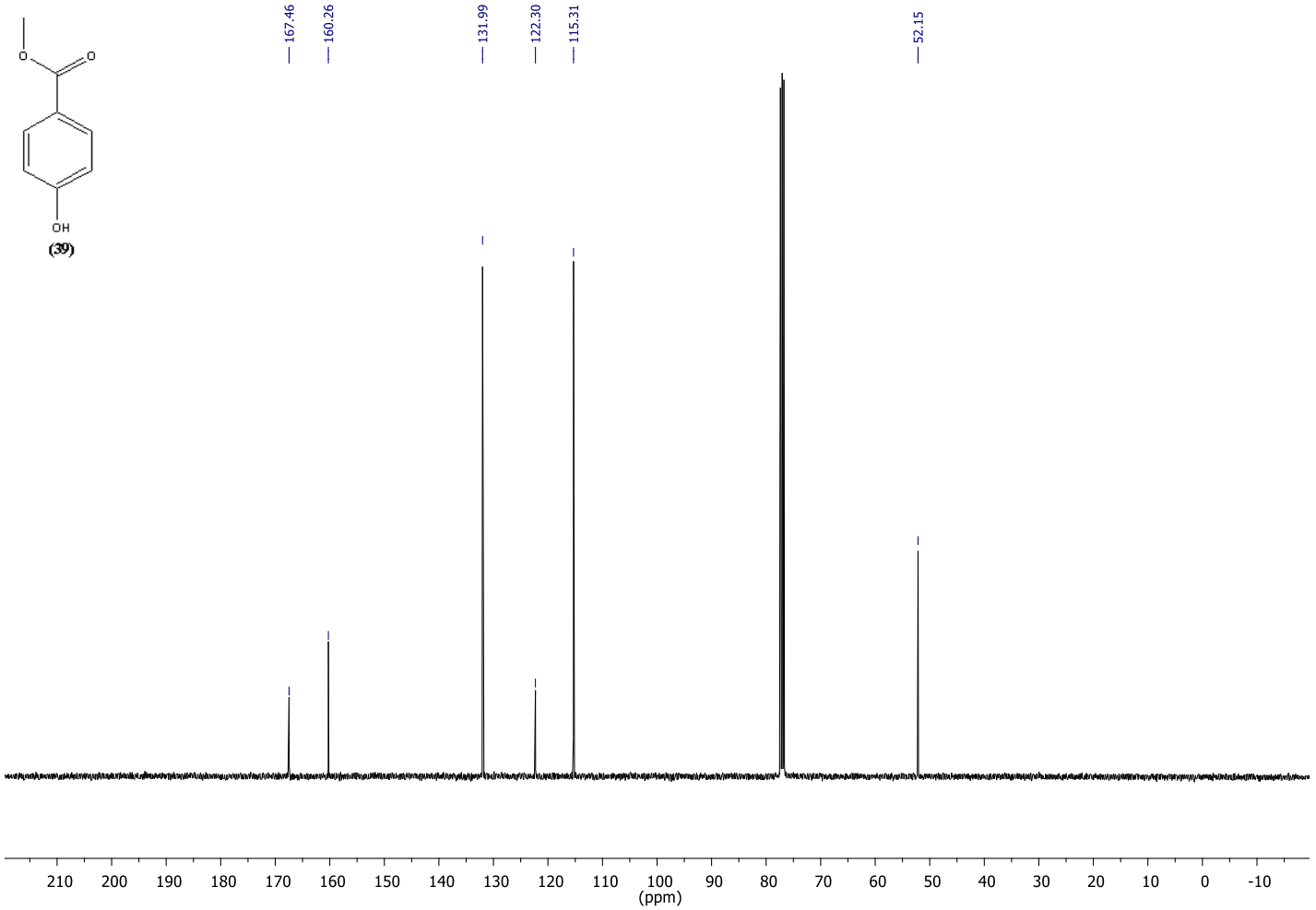


Mass spectrum (0-1000 m/z) of peak at 10.46 min, **38b**, (m/z):[M+H]$^+$calcd. for $C_{21}H_{40}N_3O_5^+$, 414.2963; found, [M+H]$^+$, 414.2; [2M+H]$^+$, 827.4

Synthesis of methyl 4-hydroxybenzoate (39):



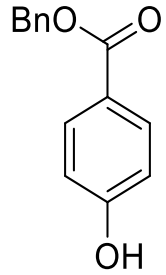
^{1}H NMR of **39** (400 MHz, $CDCl_3$)



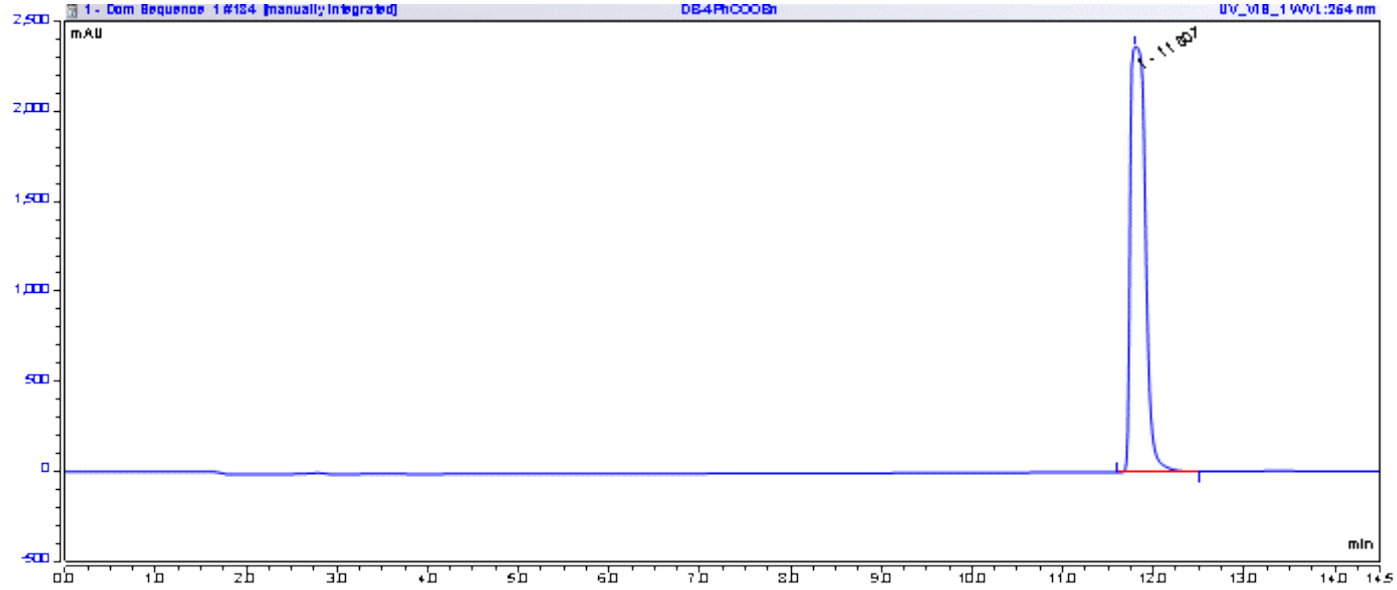
^{1}H NMR of **39** (400 MHz, $CDCl_3$)

7.1.6 Synthesis of (*2S,4R*)-1-(((9H-fluoren-9-yl)methoxy)carbonyl)-4-(4-((benzyloxy)carbonyl)phenoxy)pyrrolidine-2-carboxylic acid (56):

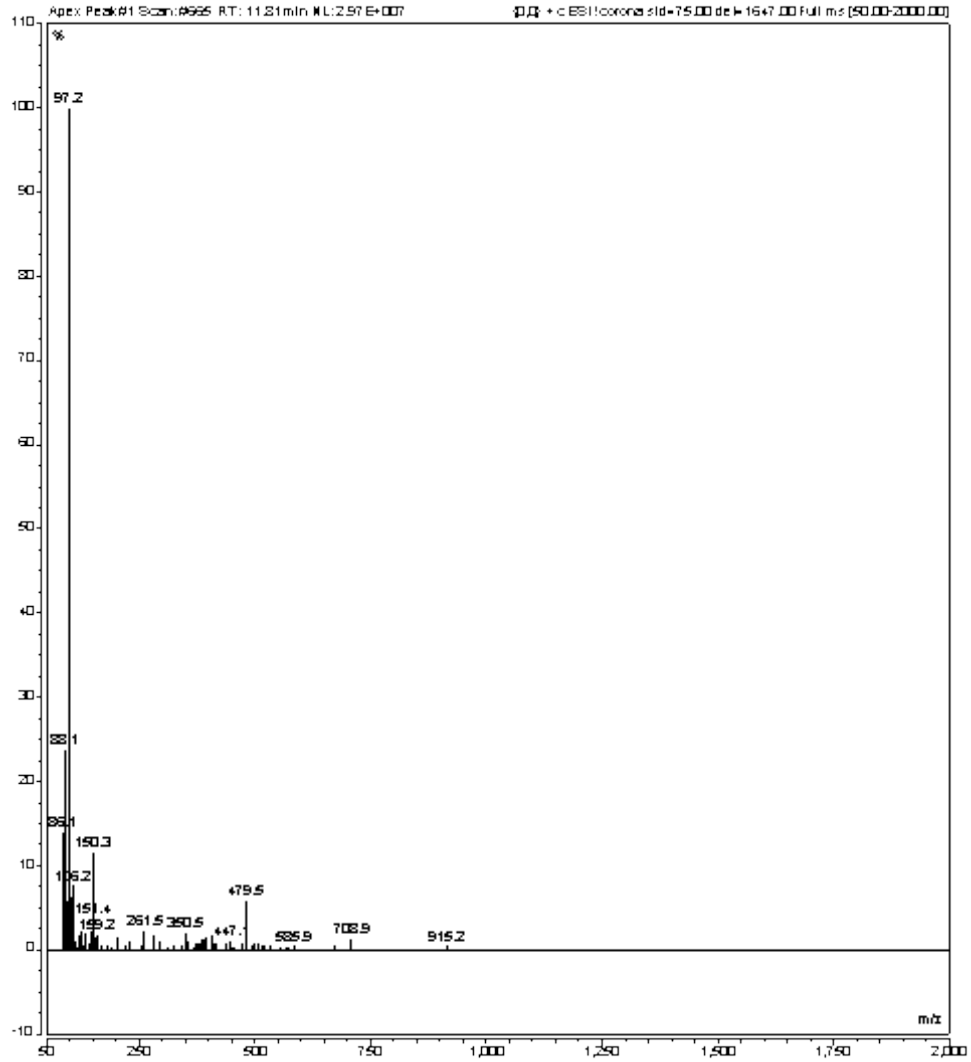
Synthesis of benzyl 4-hydroxybenzoate (52):



52

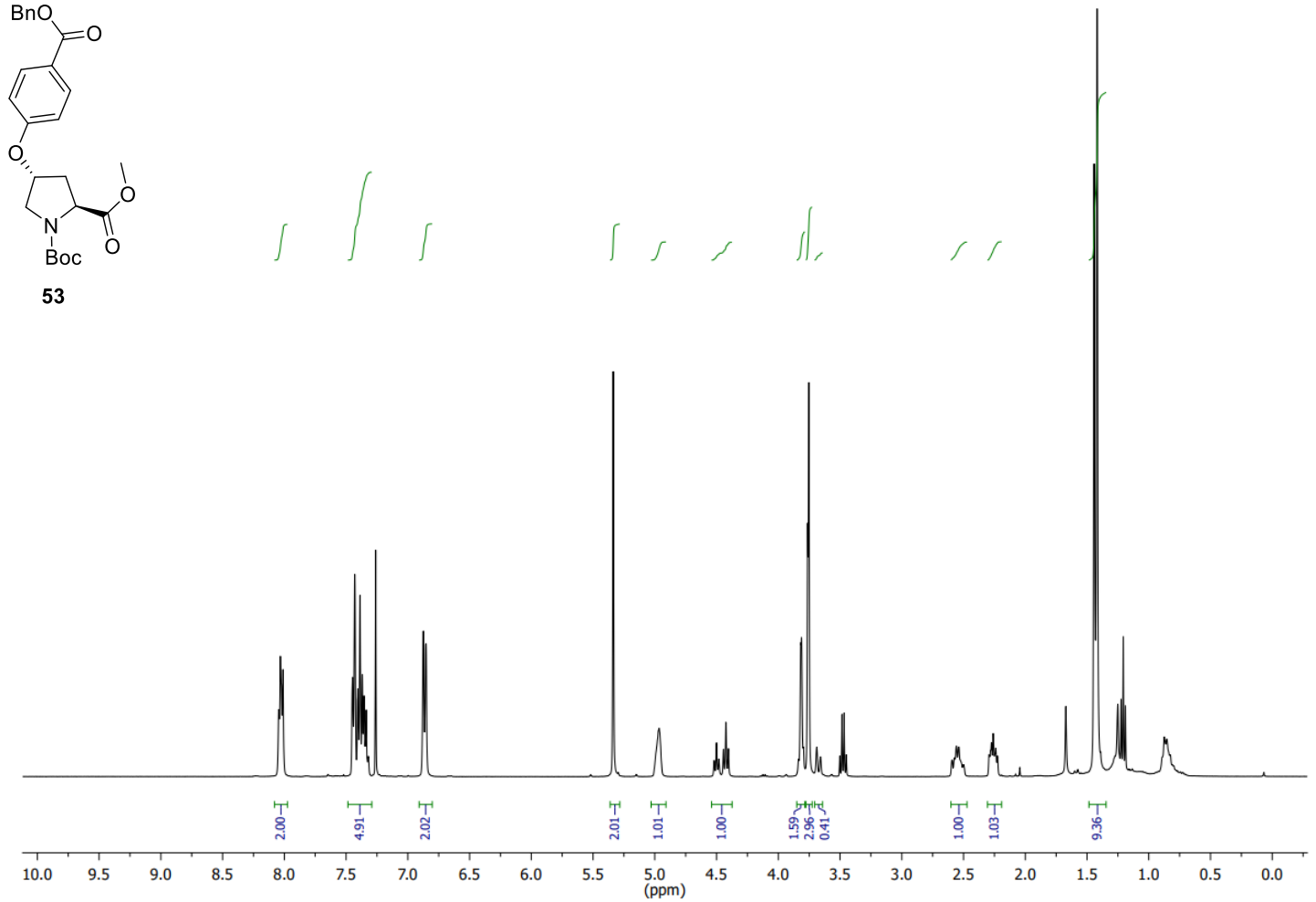


HPLC UV-Vis trace of compound **52**, rt = 11.807 min, 254 nm

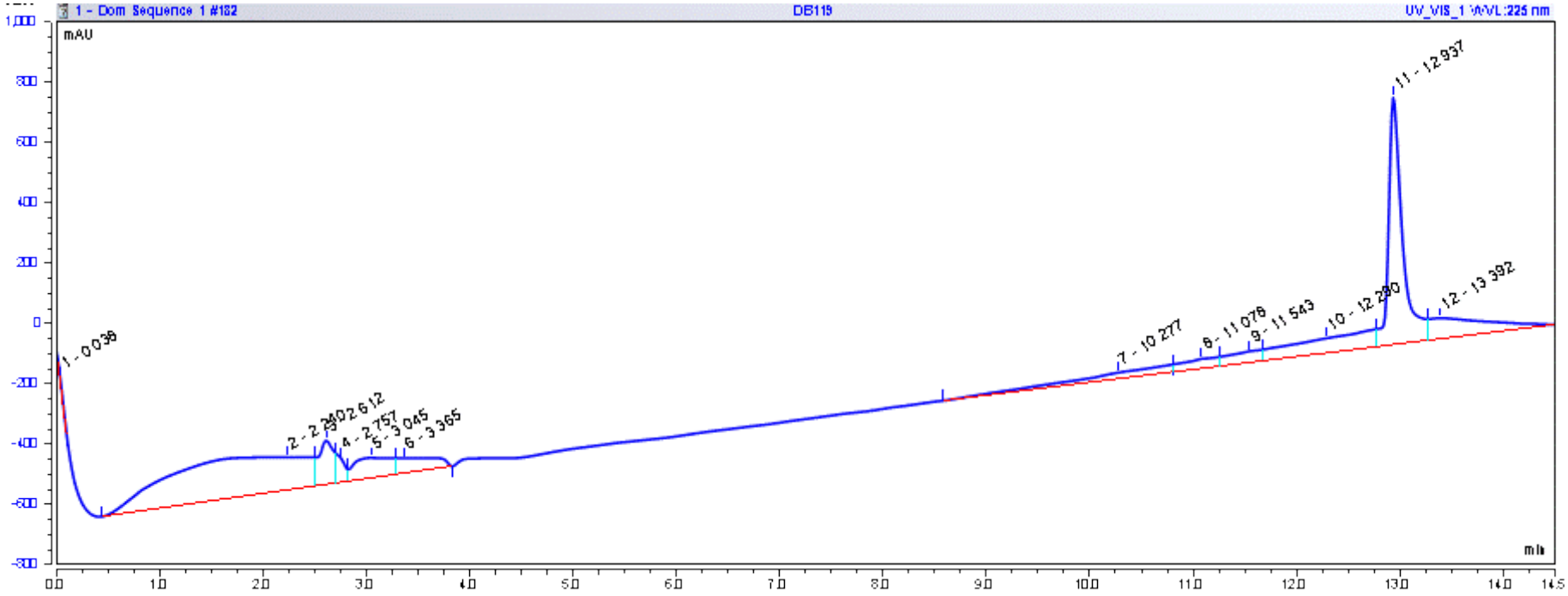


Mass spectrum (50-2000 m/z) of peak at 11.807 min, **52**

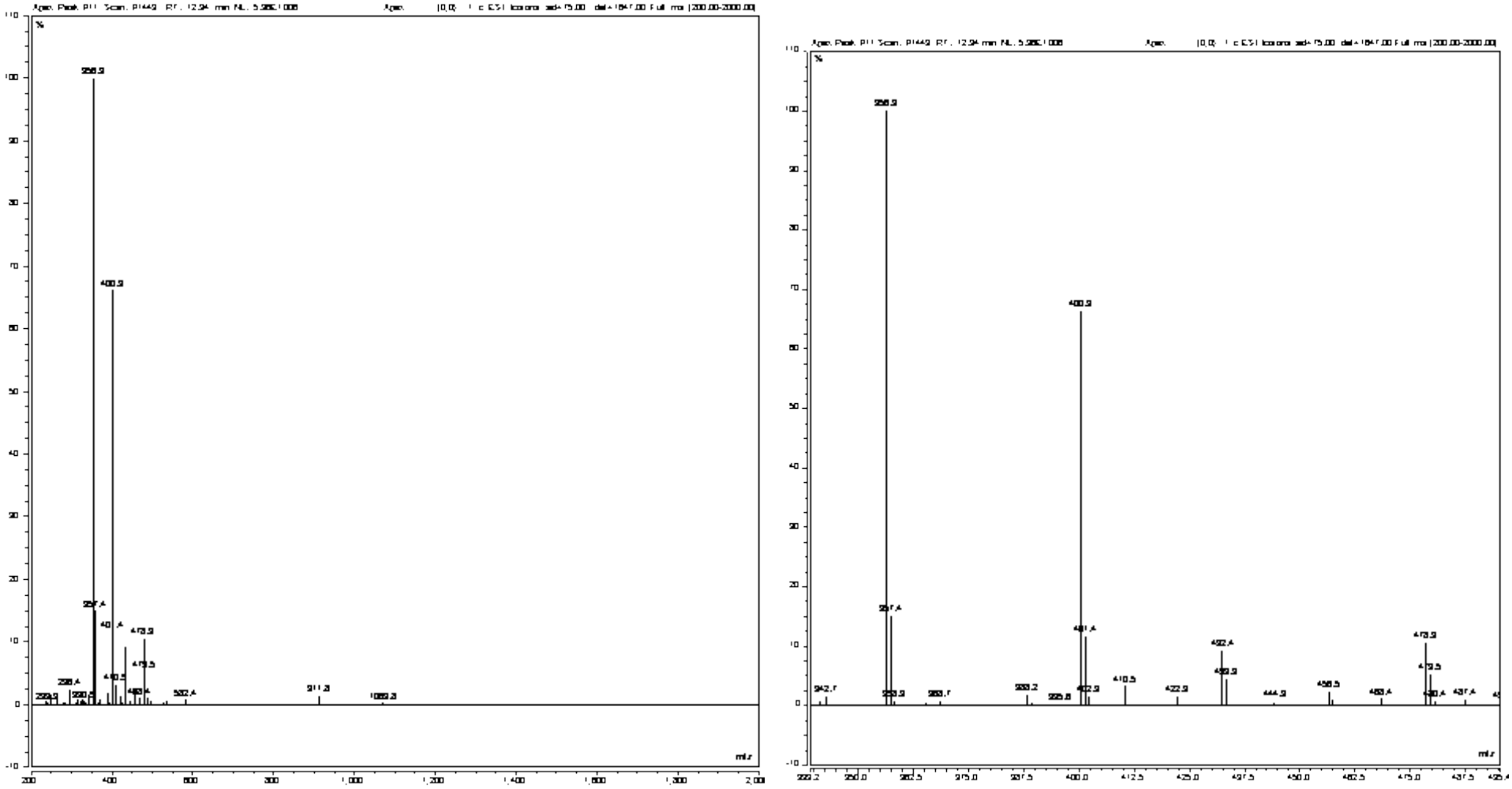
Synthesis of 1-(*tert*-butyl) 2-methyl (*2S,4R*)-4-(4-((benzyloxy)carbonyl)phenoxy)pyrrolidine-1,2-dicarboxylate (53):



^{1}H NMR of **53** (400 MHz, $CDCl_3$)

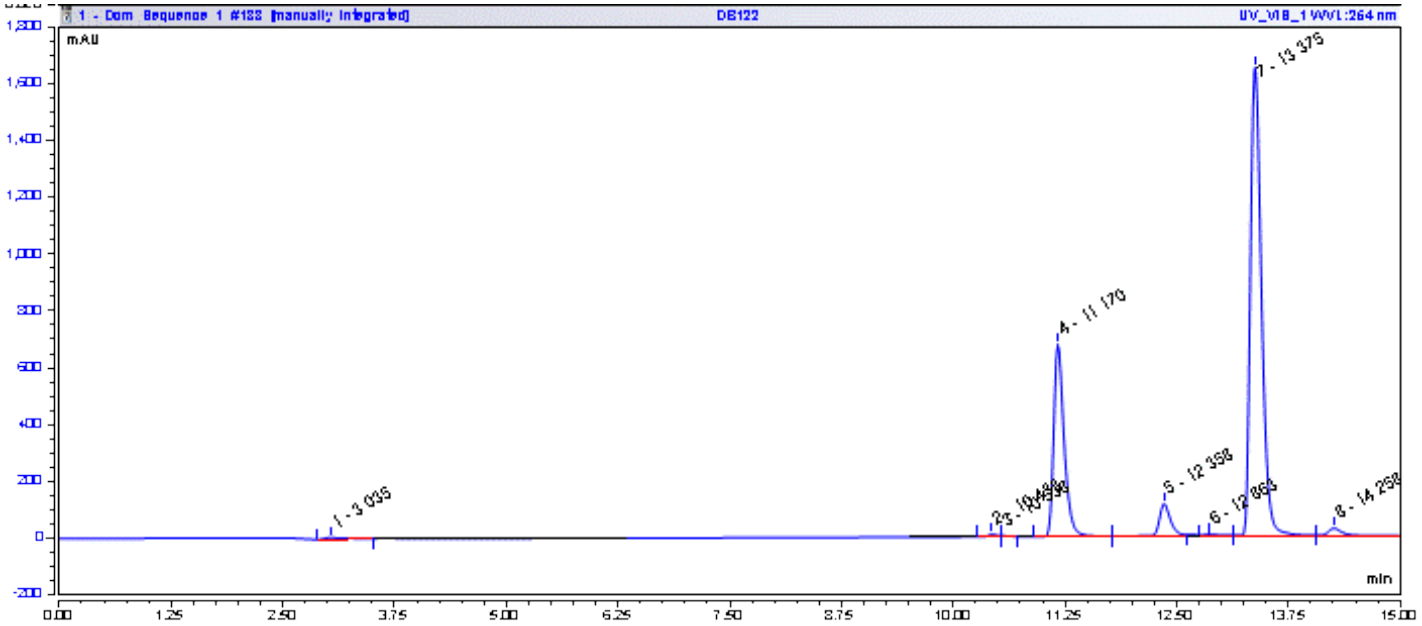


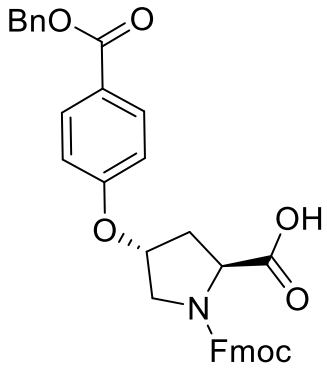
HPLC UV-Vis trace of compound **53**, rt = 12.937 min, 225 nm



Mass spectrum, 53, (200-2000 m/z) of peak at 12.937 min, and enlarged view of M^+ ion, section (340-495 m/z)

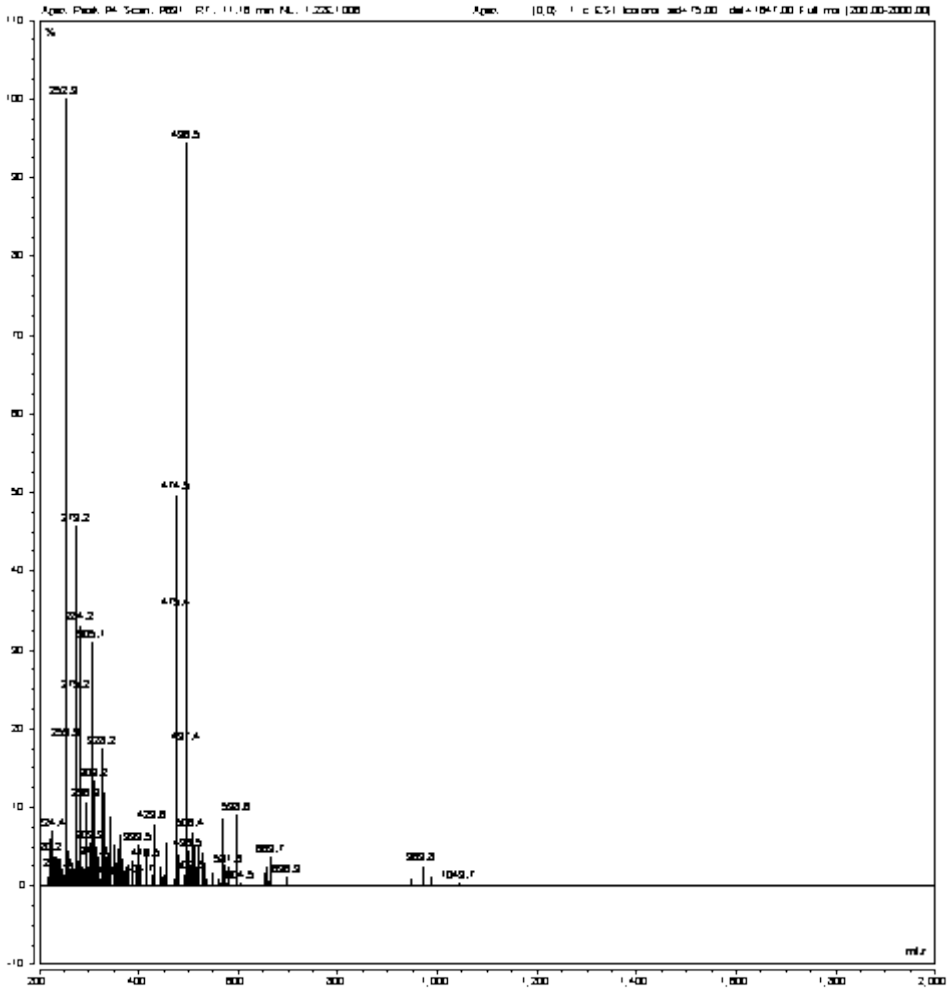
Synthesis of (*2S,4R*)-1-(((9H-fluoren-9-yl)methoxy)carbonyl)-4-(4-((benzyloxy)carbonyl)phenoxy)pyrrolidine-2-carboxylic acid (56):



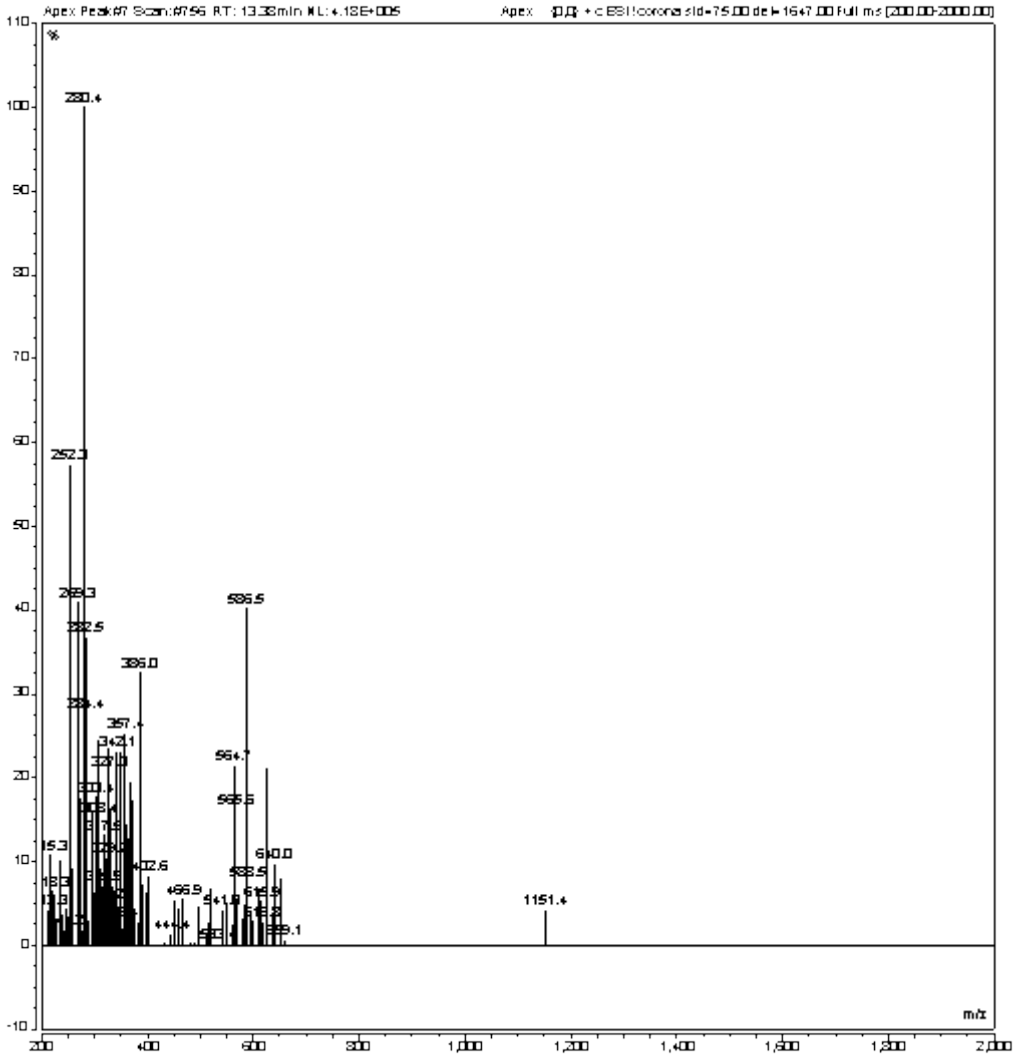


56

HPLC UV-Vis of **56** before FCC, rt (**56 minus Bn**) = 11.170 min, rt **56** = 13.375 min



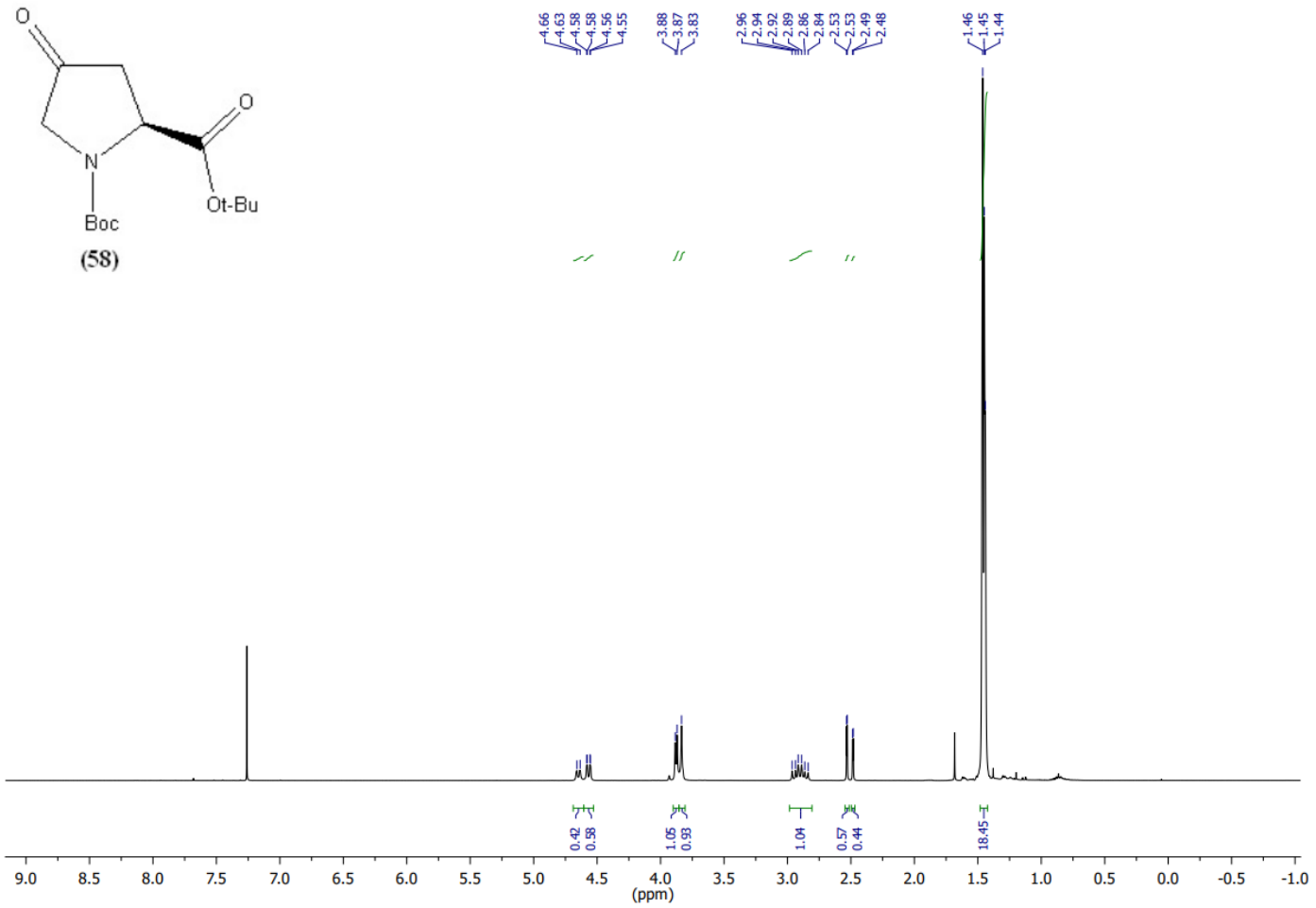
Mass spectrum (200-2000 m/z) of peak at 11.16 min, **56 minus Bn**, (m/z):$[M+H]^+$calcd. for $C_{27}H_{24}NO_7^+$, 474.1548; found, $[M+H]^+$, 474.5; $[M+Na]^+$, 496.5; $[2M+Na]^+$, 969.8



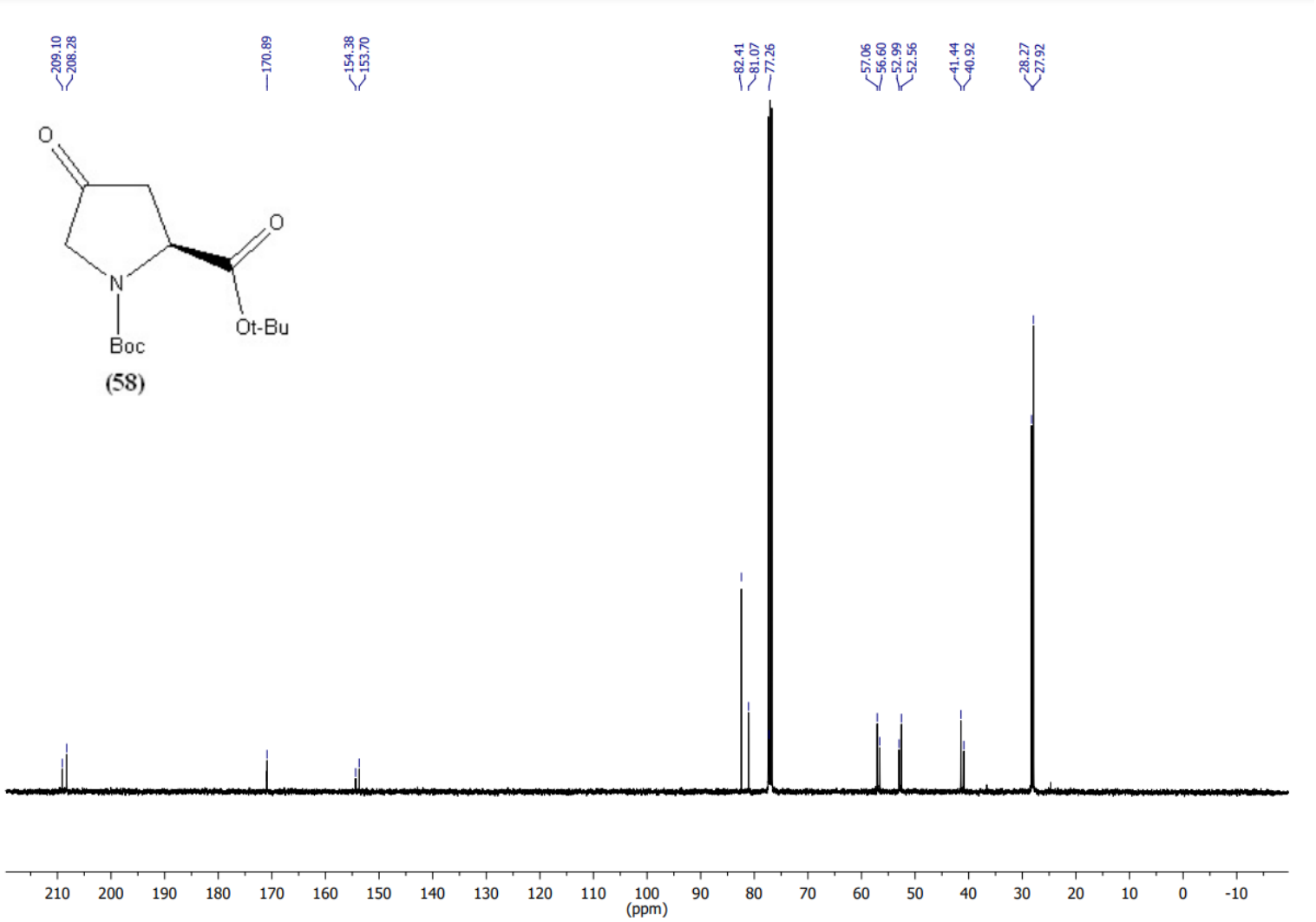
Mass spectrum (200-2000 m/z) of peak at 13.38 min, **56**, (m/z):[M+H]$^+$calcd. for $C_{27}H_{24}NO_7^+$, 564.2017; found, [M+H]$^+$, 564.7; [M+Na]$^+$, 586.5; [2M+Na]$^+$, 1151.4

7.1.7 Synthesis of (S)-1-(((9H-fluoren-9-yl)methoxy)carbonyl)-4-(4-((benzyloxy)carbonyl)phenoxy)-2,5-dihydro-1H-pyrrole-2-carboxylic acid, (**62**):

Synthesis of di-*tert*-butyl (S)-4-oxopyrrolidine-1,2-dicarboxylate, (58):

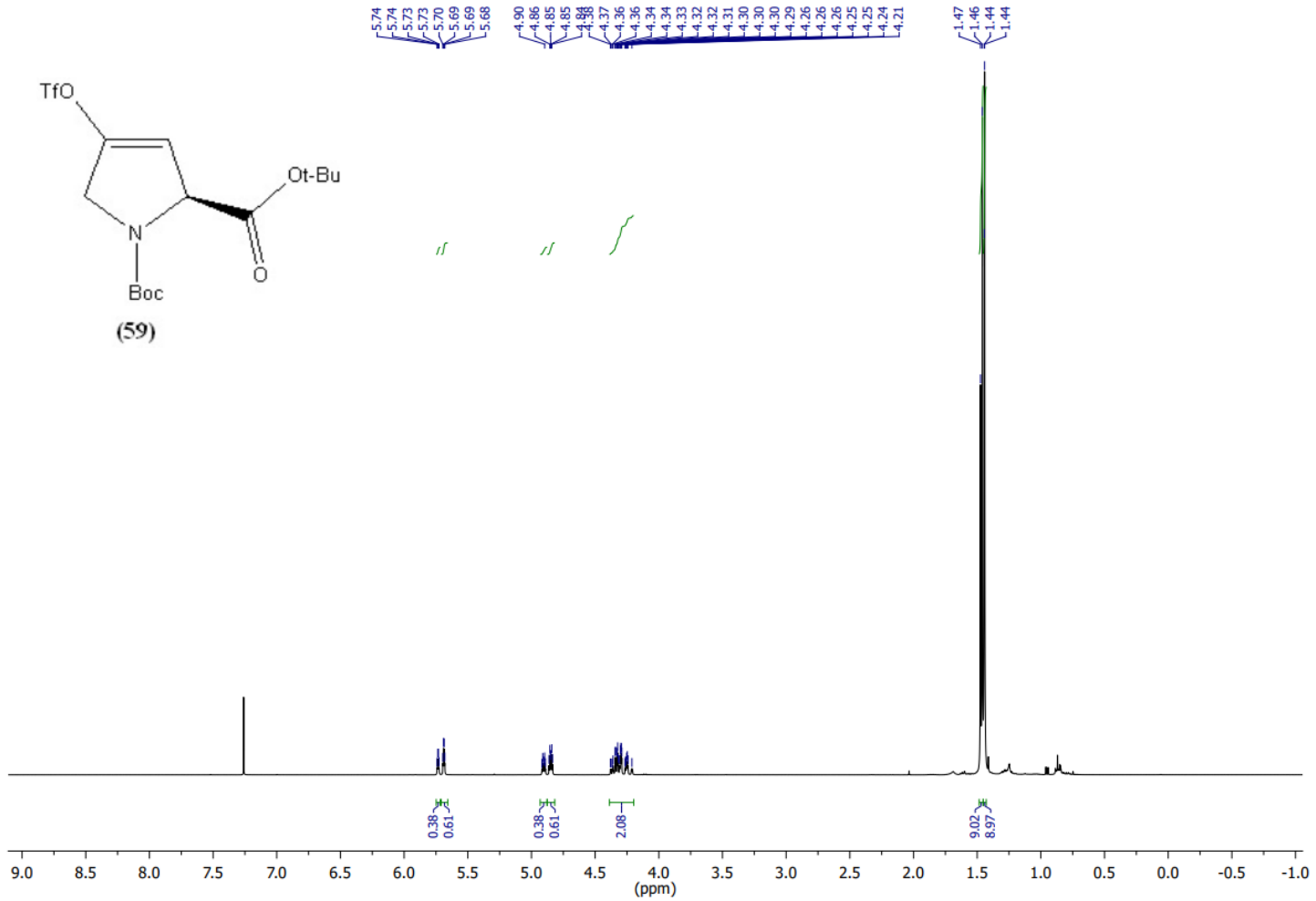


^{1}H NMR of **58** (400 MHz, $CDCl_3$)

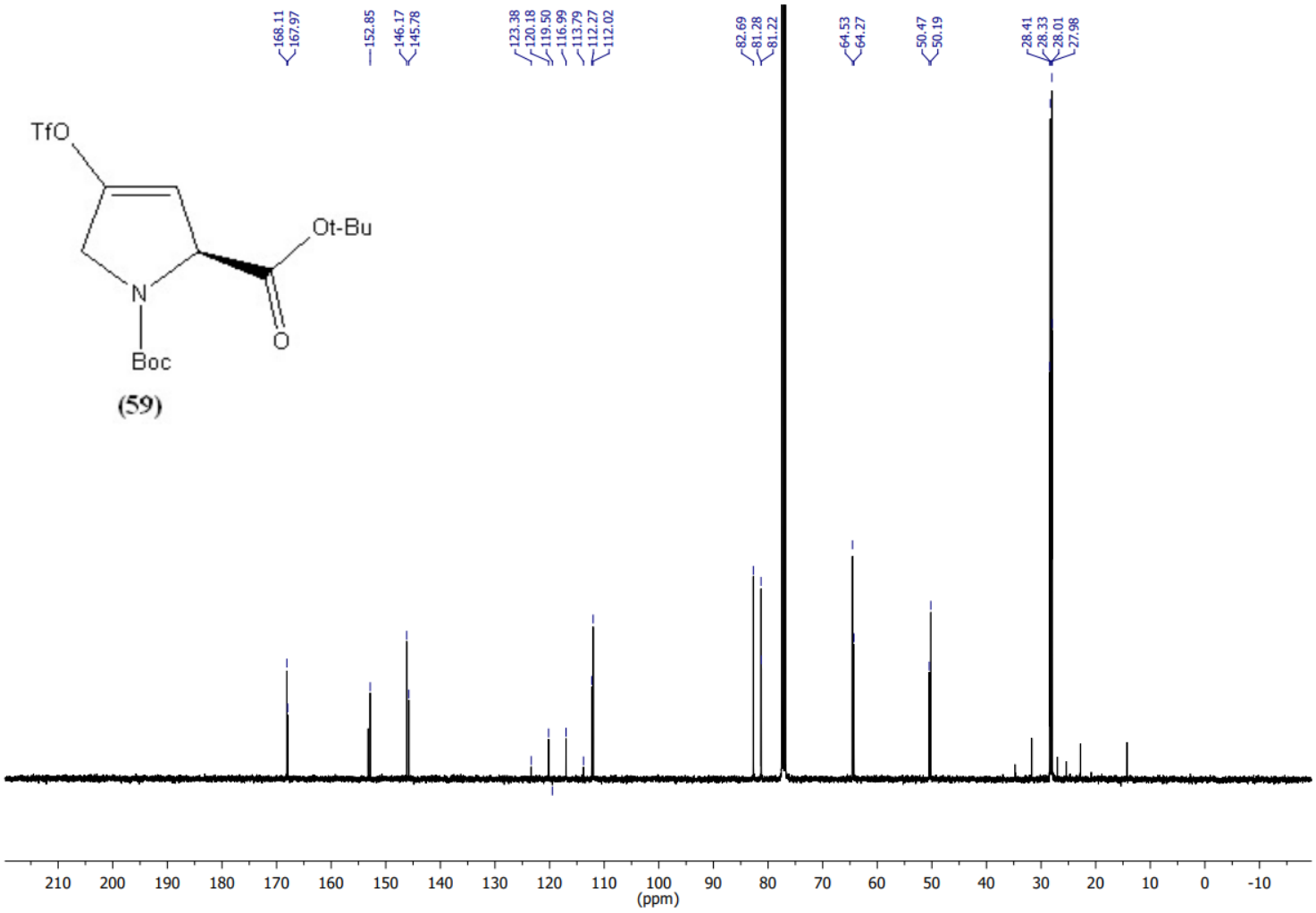


^{13}C NMR of **58** (101 MHz, $CDCl_3$)

Synthesis of di-*tert*-butyl (S)-4-(((trifluoromethyl)sulfonyl)oxy)-2,5-dihydro-1H-pyrrole-1,2-dicarboxylate, (59):

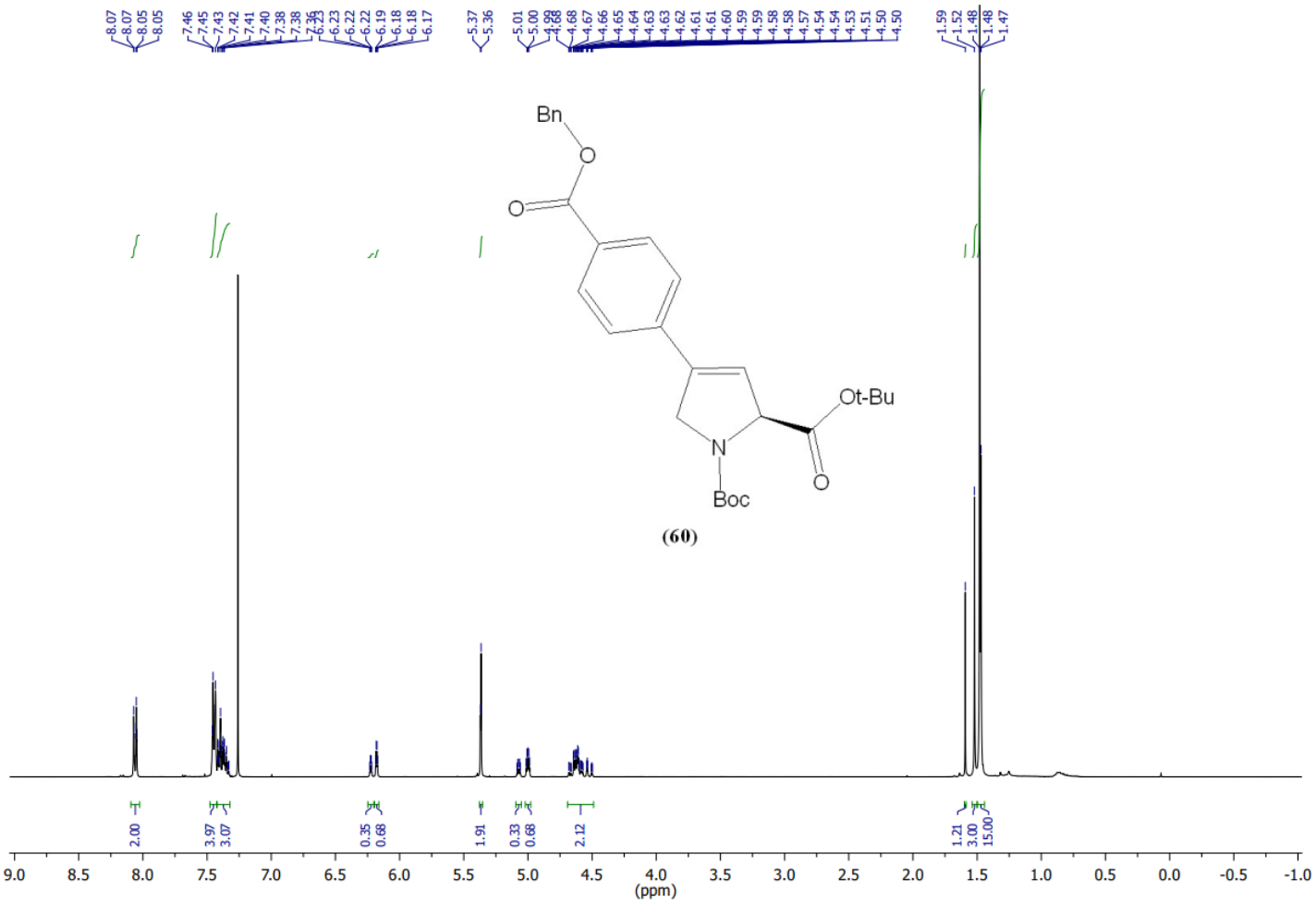


^{1}H NMR of **59** (400 MHz, $CDCl_3$)

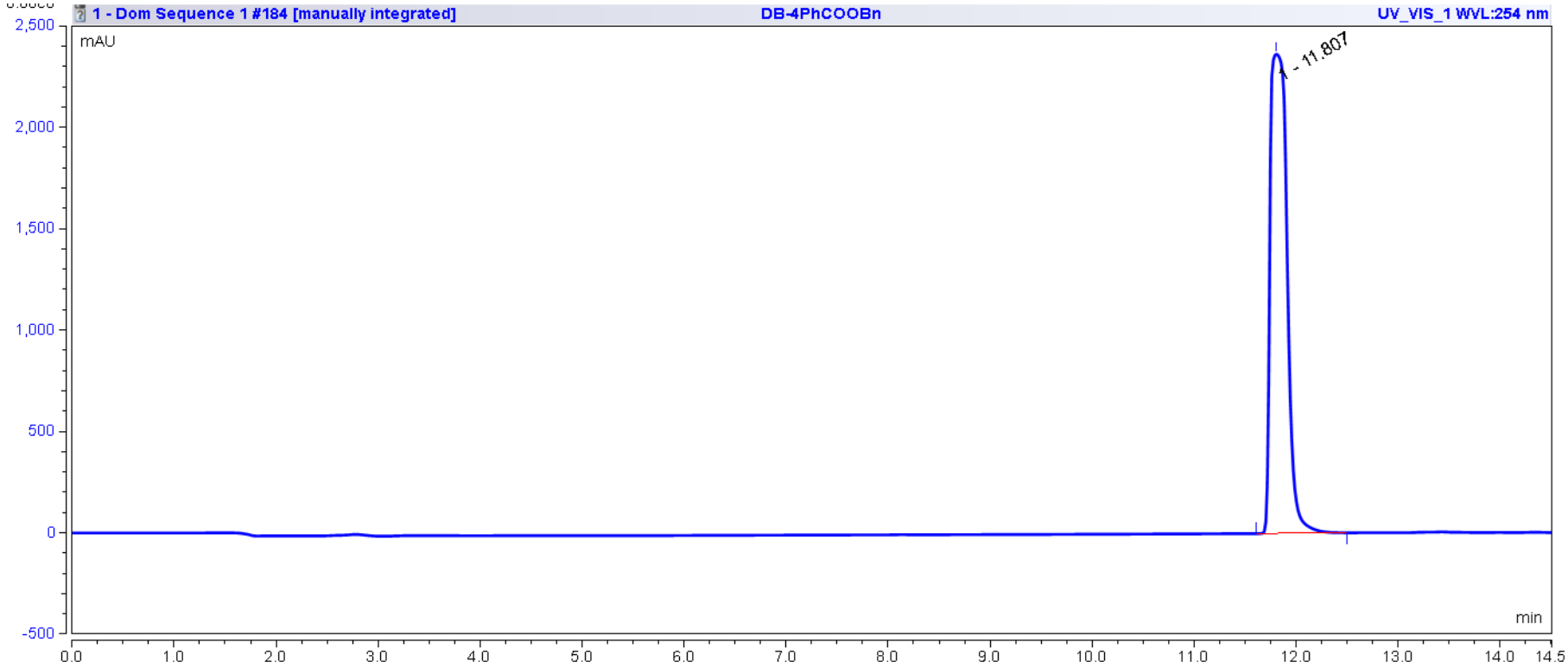


^{13}C NMR of **59** (101 MHz, $CDCl_3$)

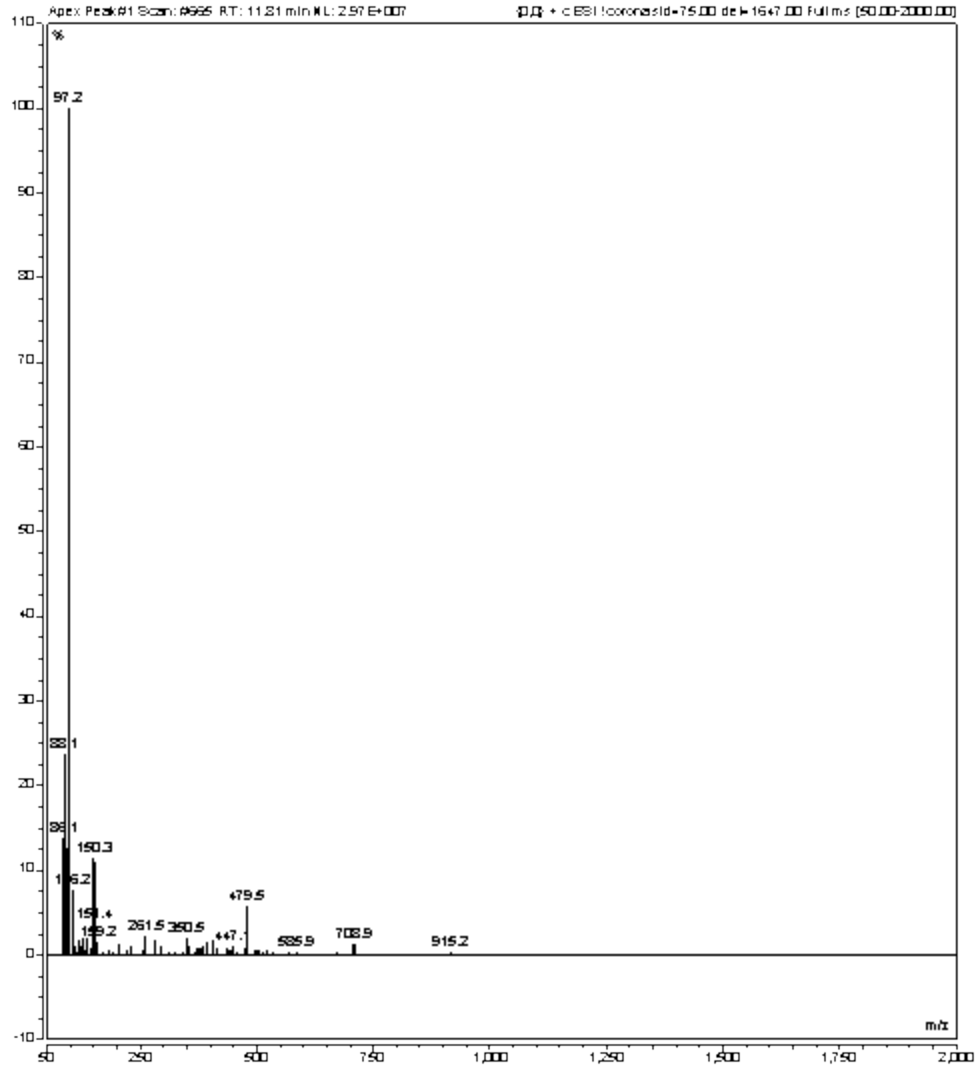
Synthesis of di-*tert*-butyl (S)-4-(4-((benzyloxy)carbonyl)phenyl)-2,5-dihydro-1H-pyrrole-1,2-dicarboxylate, (60):



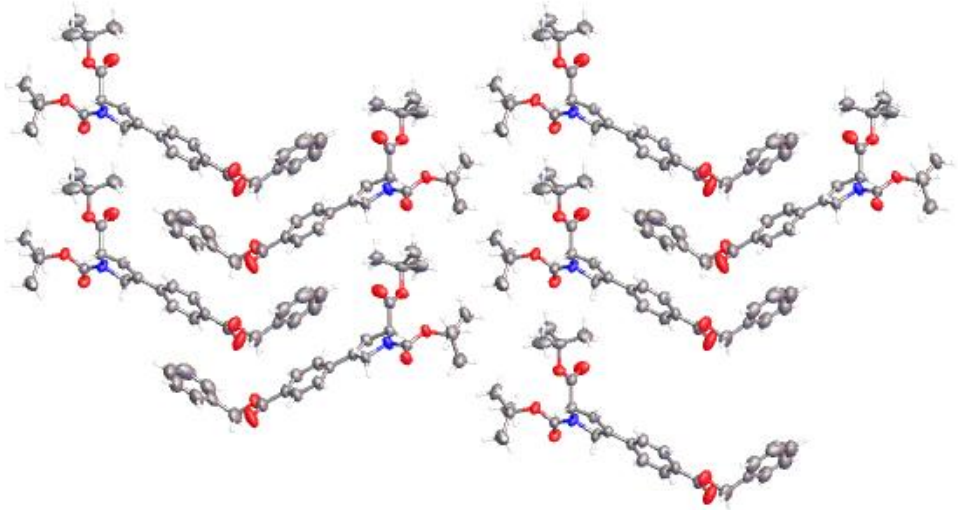
^{1}H NMR of **60** (400 MHz, $CDCl_3$)



HPLC UV-Vis of **60**, rt = 11.807 min, 254 nm



Mass spectrum (50-2000 m/z) of peak at 11.81 min, **60**, (m/z):[M+H]$^+$calcd. for $C_{28}H_{34}NO_6^+$, 480.2381; found, [M+H]$^+$, 479.5



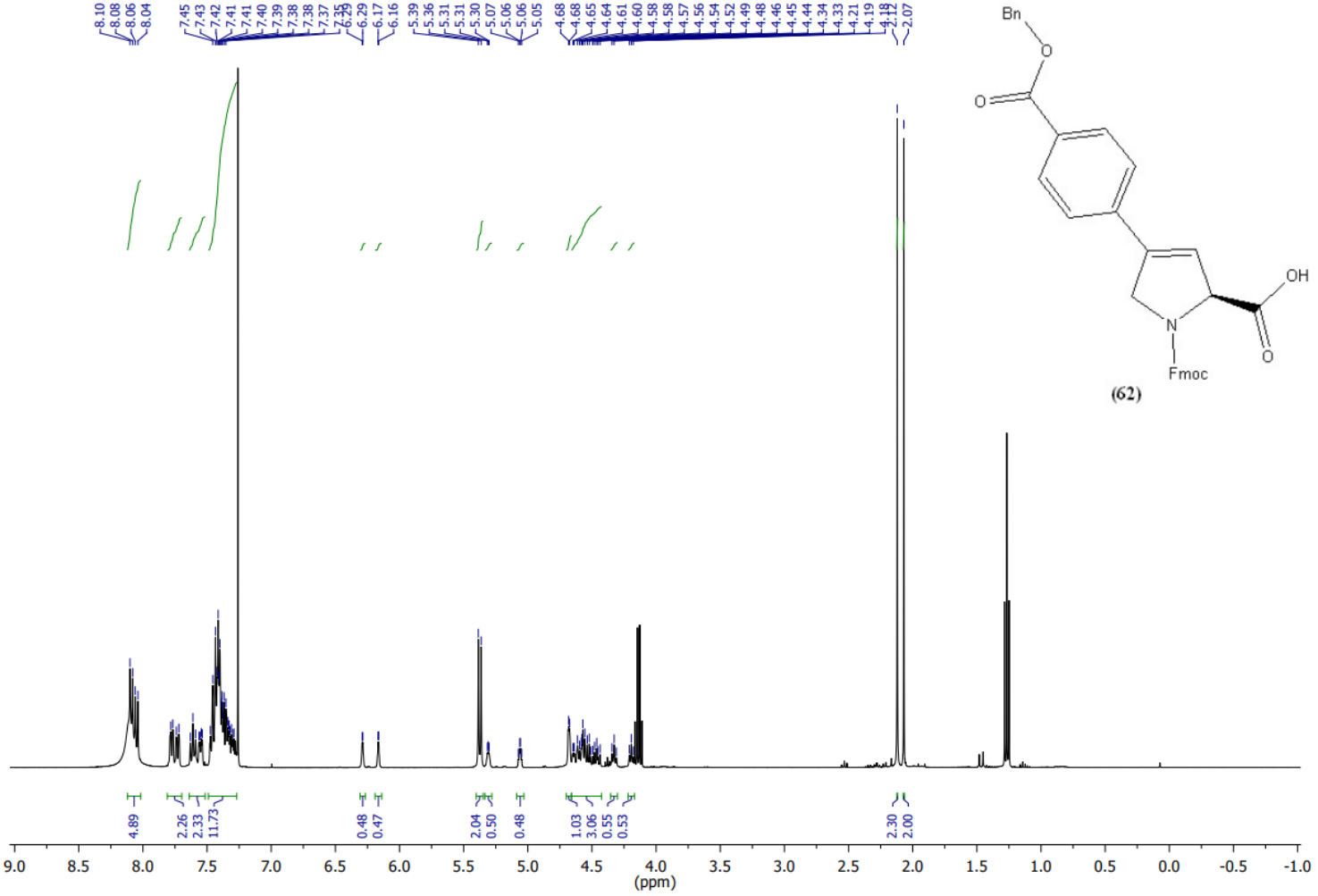
Single crystal structure *of compound 60, 50% ellipsoids, view along the a axis*

Table 2 Crystal data and structure refinement for compound 60.

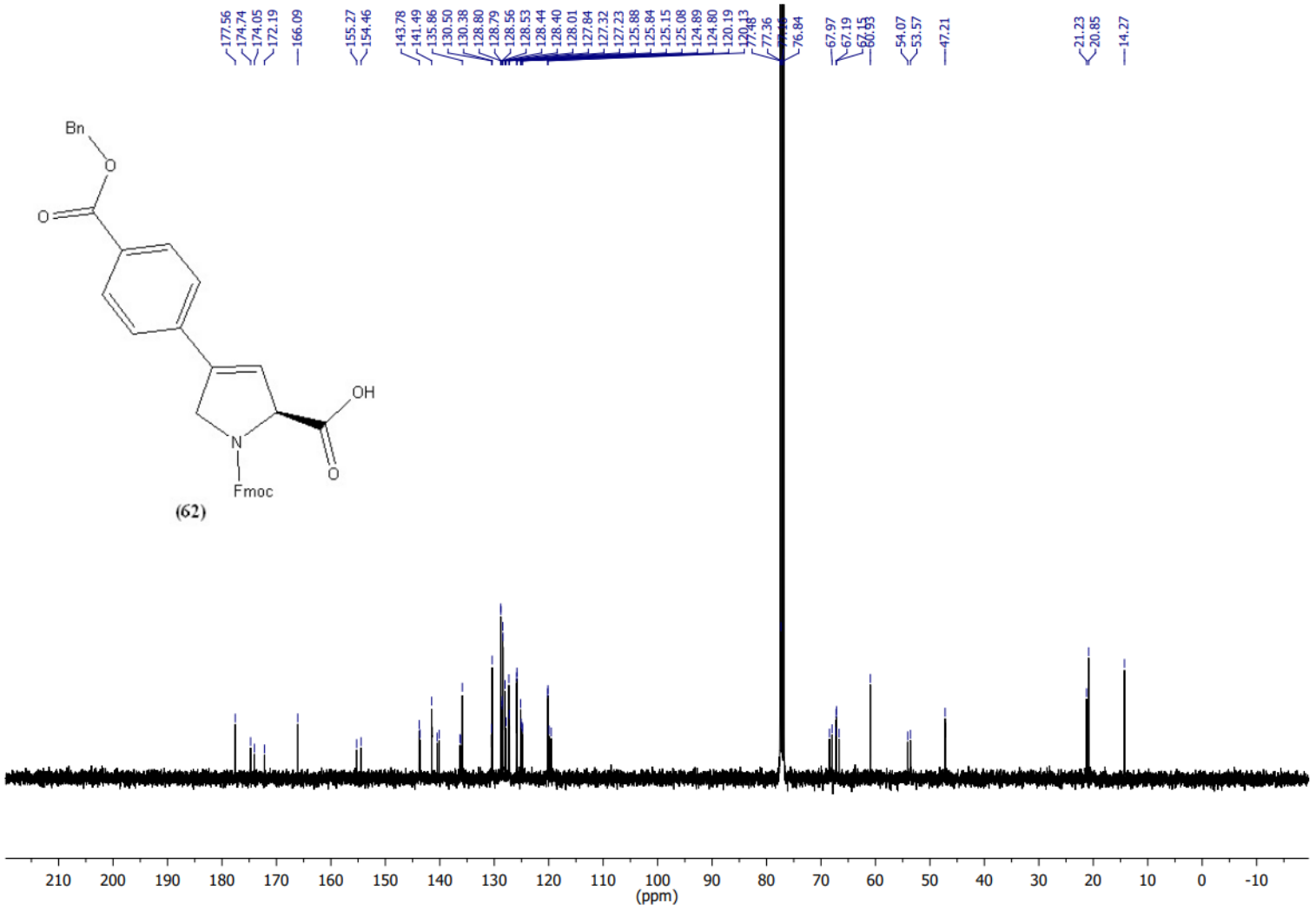
Identification code	2178751
Empirical formula	$C_{28}H_{33}NO_6$
Formula weight	479.55
Temperature/K	293
Crystal system	monoclinic
Space group	$P2_1$
a/Å	6.14630(10)
b/Å	8.31350(10)
c/Å	25.4908(3)
α/°	90

β/°	93.1660(10)
γ/°	90
Volume/Å^3	1300.52(3)
Z	2
ρ_{calc}g/cm^3	1.225
μ/mm^{-1}	0.698
F(000)	512.0
Crystal size/mm^3	0.96 × 0.31 × 0.07
Radiation	Cu Kα (λ = 1.54184)
2Θ range for data collection/°	6.946 to 136.5
Index ranges	$-7 \leq h \leq 7$, $-10 \leq k \leq 10$, $-30 \leq l \leq 30$
Reflections collected	28430
Independent reflections	4772 [R_{int} = 0.0426, R_{sigma} = 0.0197]
Data/restraints/parameters	4772/1/322
Goodness-of-fit on F^2	1.123
Final R indexes [I>=2σ (I)]	R_1 = 0.0615, wR_2 = 0.1808
Final R indexes [all data]	R_1 = 0.0621, wR_2 = 0.1815
Largest diff. peak/hole / e Å^{-3}	0.35/-0.26
Flack parameter	-0.02(9)

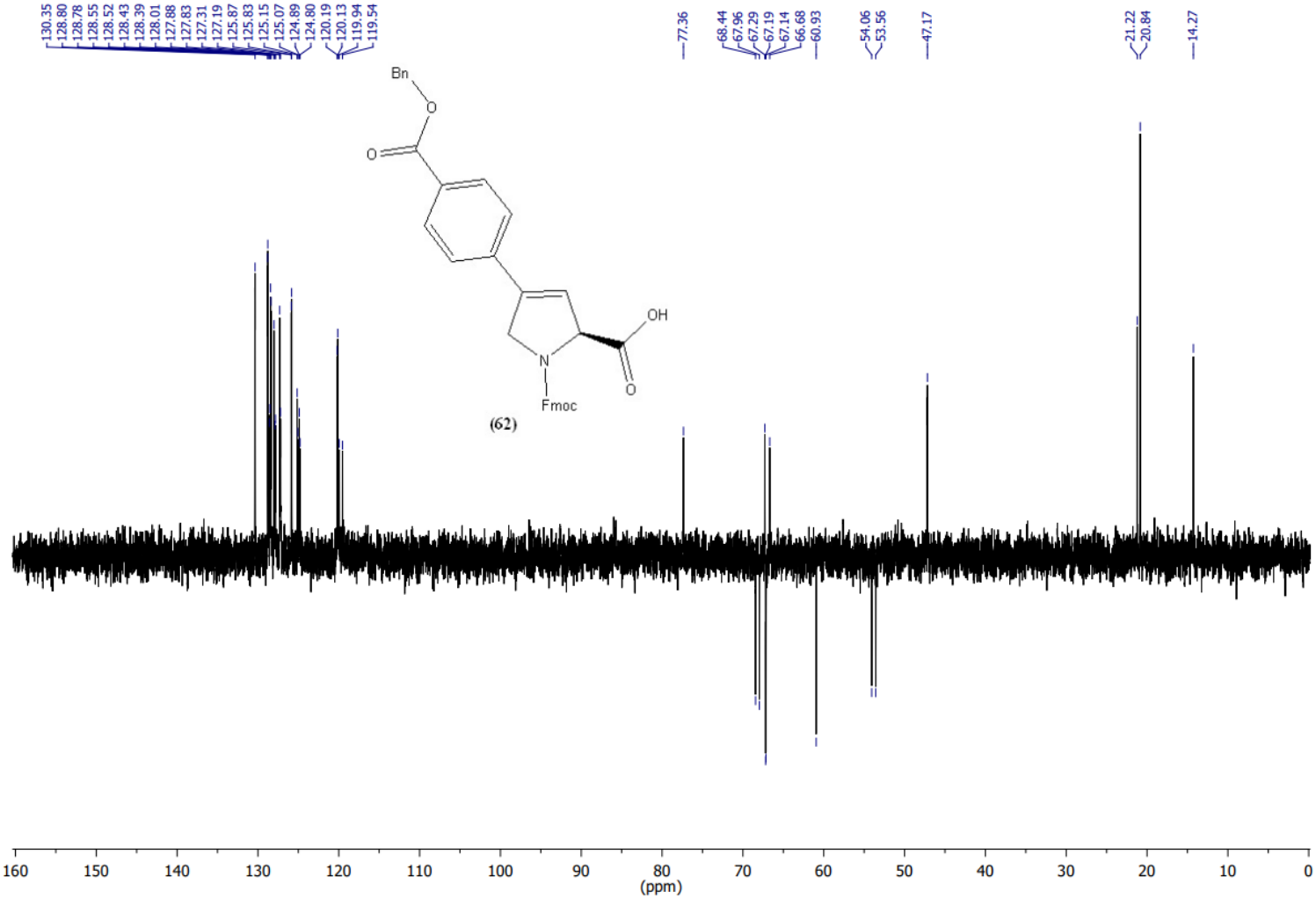
Synthesis of (S)-1-(((9H-fluoren-9-yl)methoxy)carbonyl)-4-(4-((benzyloxy)carbonyl)phenyl)-2,5-dihydro-1H-pyrrole-2-carboxylic acid, (62):



^{1}H NMR of **62** (400 MHz, $CDCl_3$)



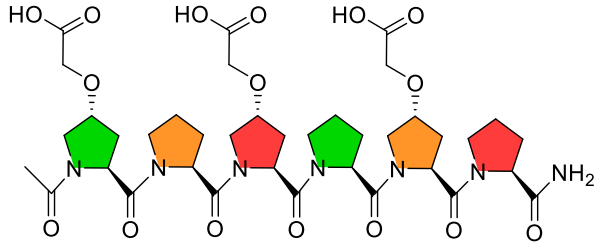
^{13}C NMR of **62** (101 MHz, $CDCl_3$)



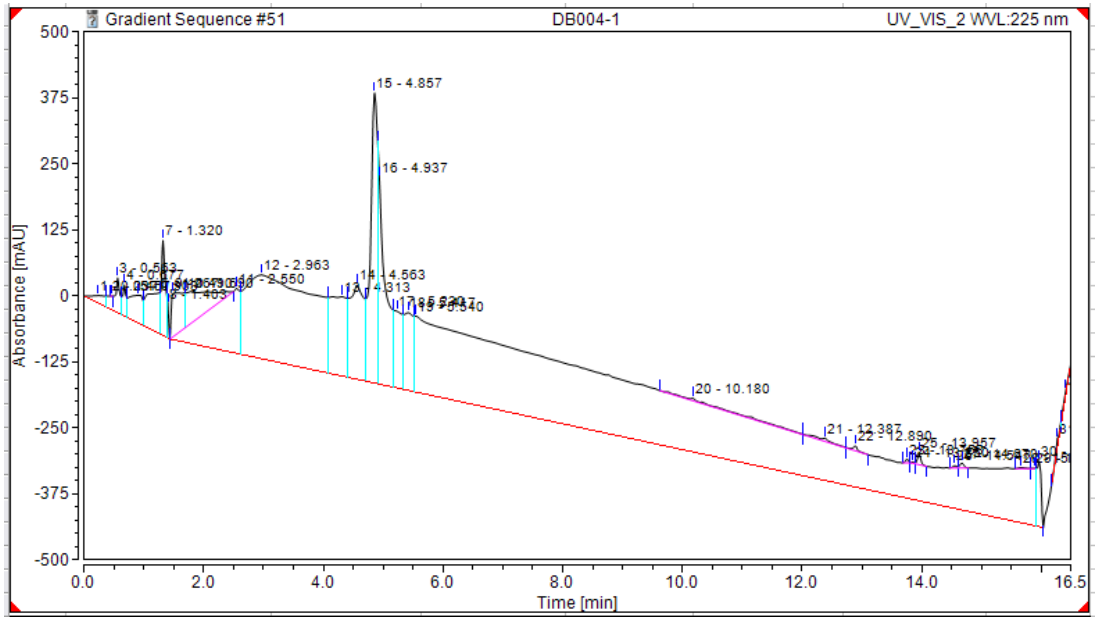
DEPT-135 ^{13}C NMR of **62** (101 MHz, $CDCl_3$)

7.2 Peptides

7.2.1 Synthesis of Ac-Pro_6-NH_2 (-$OCH_2COOH)_3$(i: 1, 3, 5) (67):

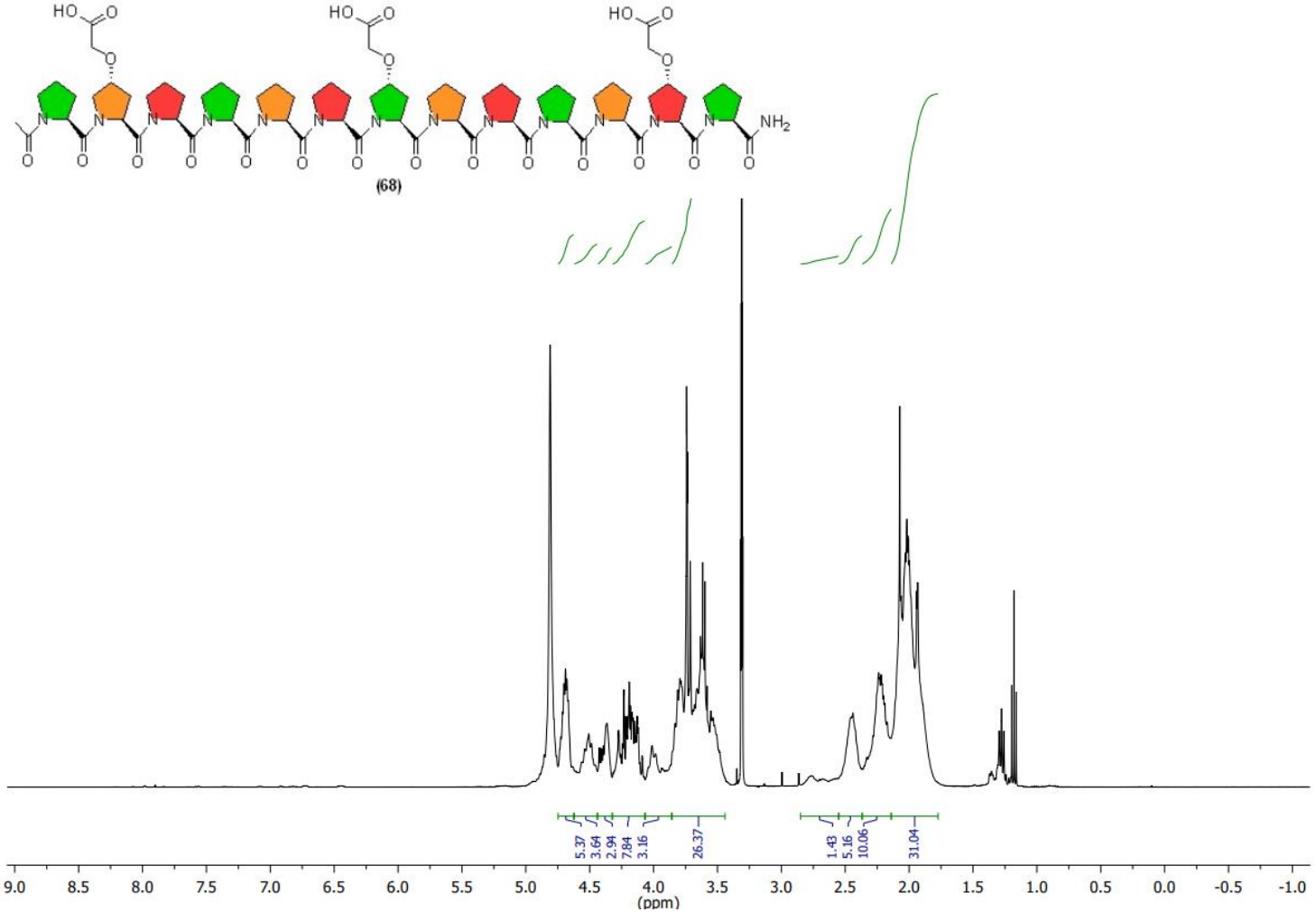


(67)

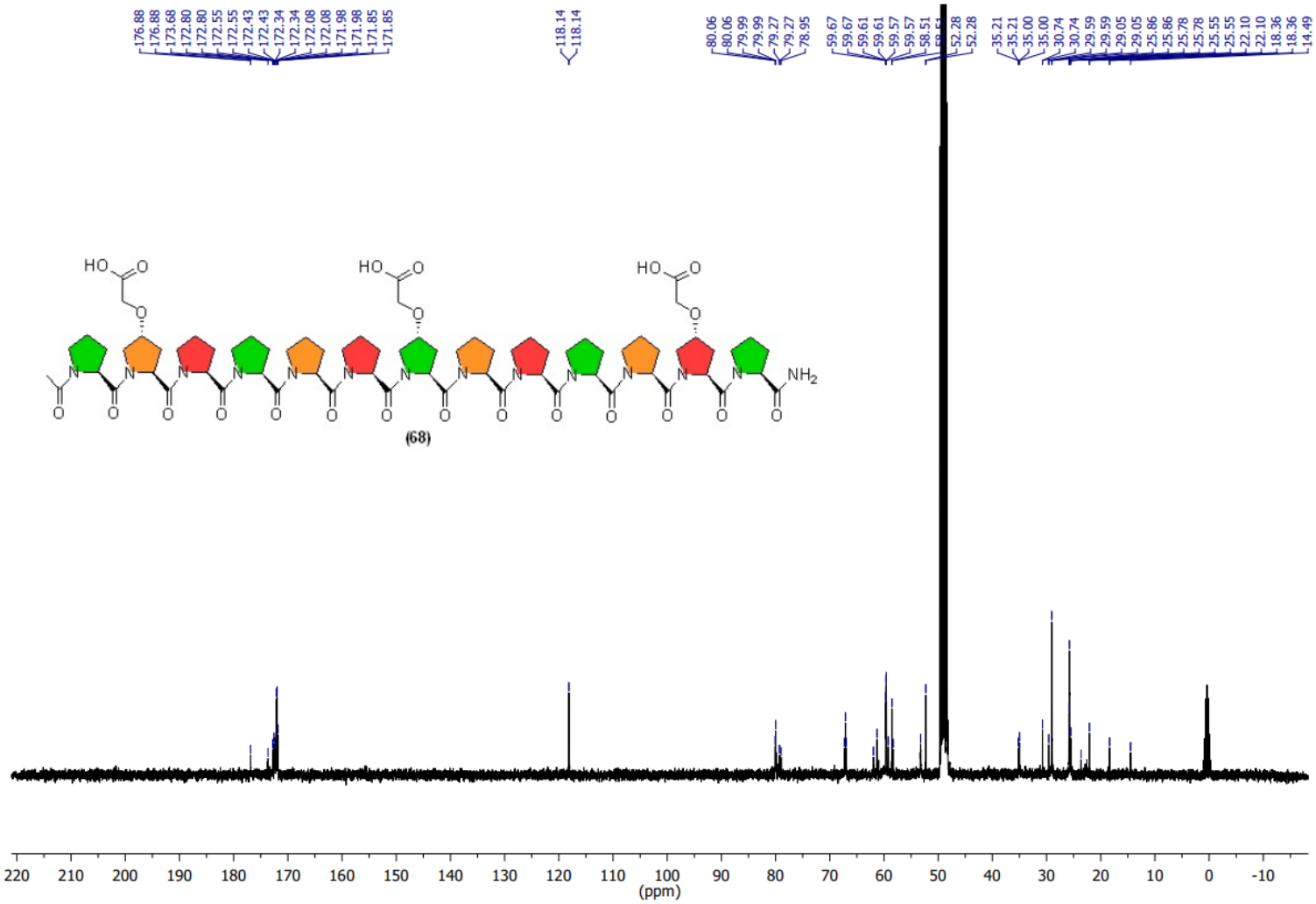


HPLC UV-Vis trace for peptide **67**, rt = 4.857 min, 225 nm

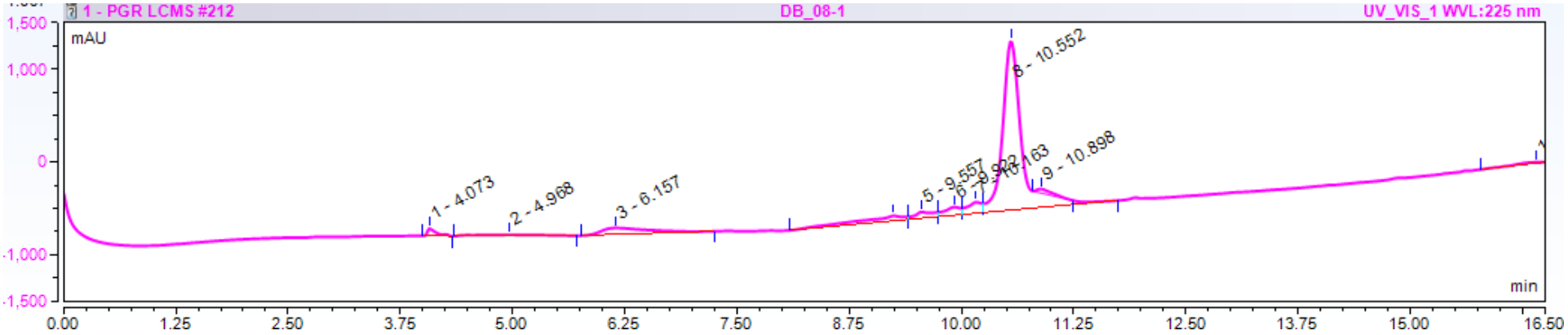
7.2.2 Synthesis of Ac-Pro_{13}-NH_2 (-$OCH_2COOH)_3$(i: 2, 7, 12) (68):



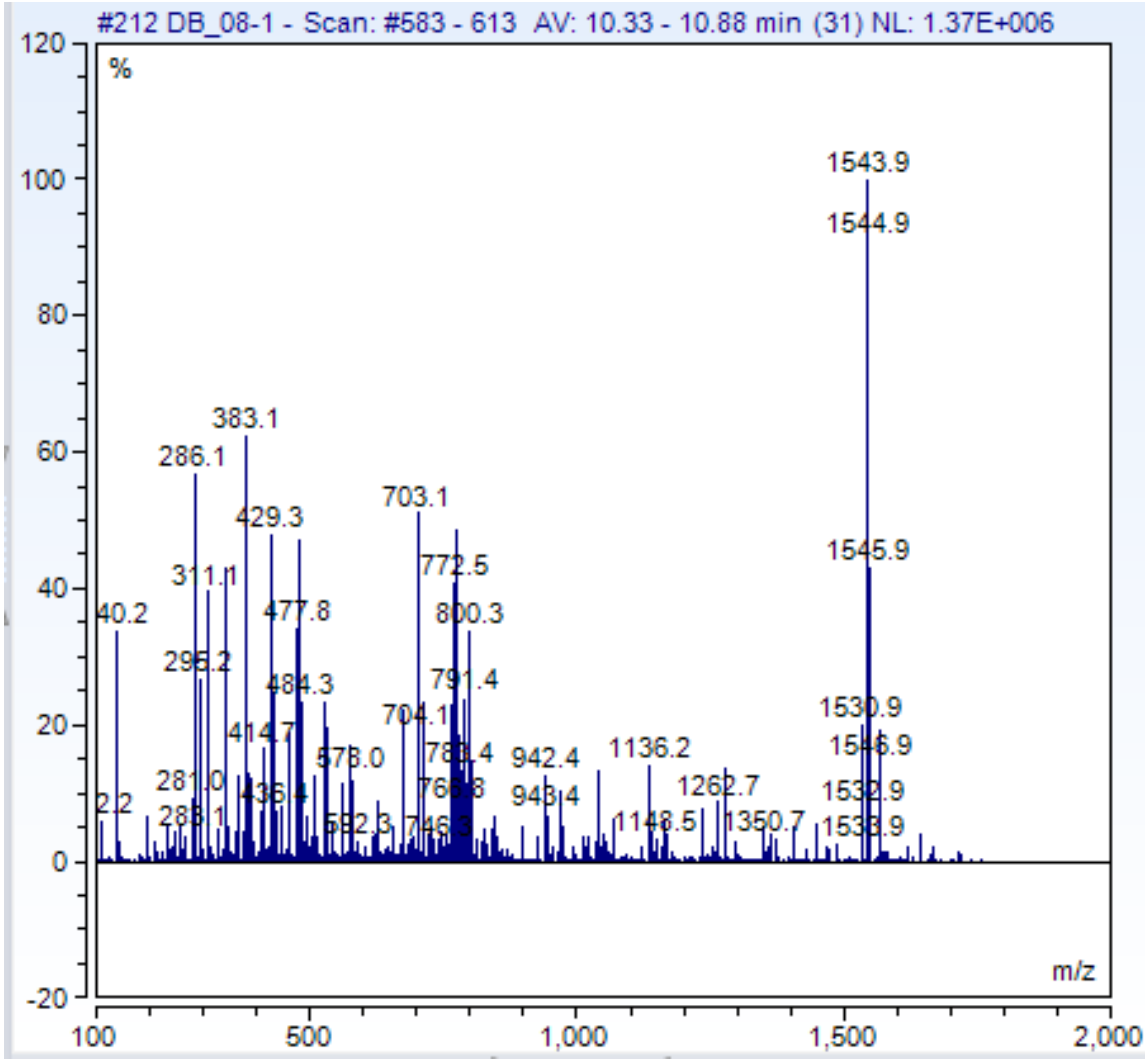
^{1}H NMR of **68** (400 MHz, MeOD)



^{13}C NMR of **68** (101 MHz, MeOD)

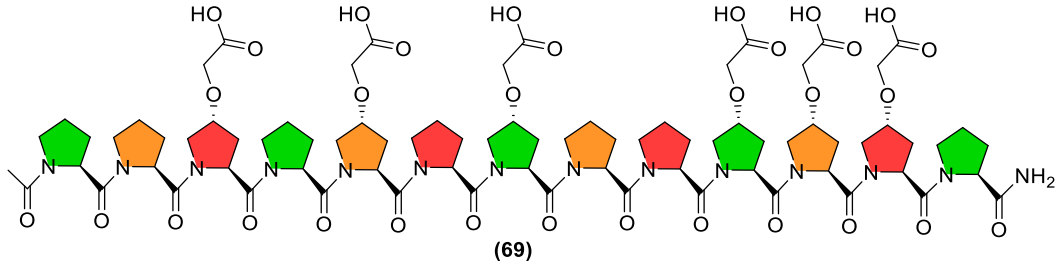


HPLC UV-Vis Spectrum of peptide **68**, rt = 10.552 min, 225 nm

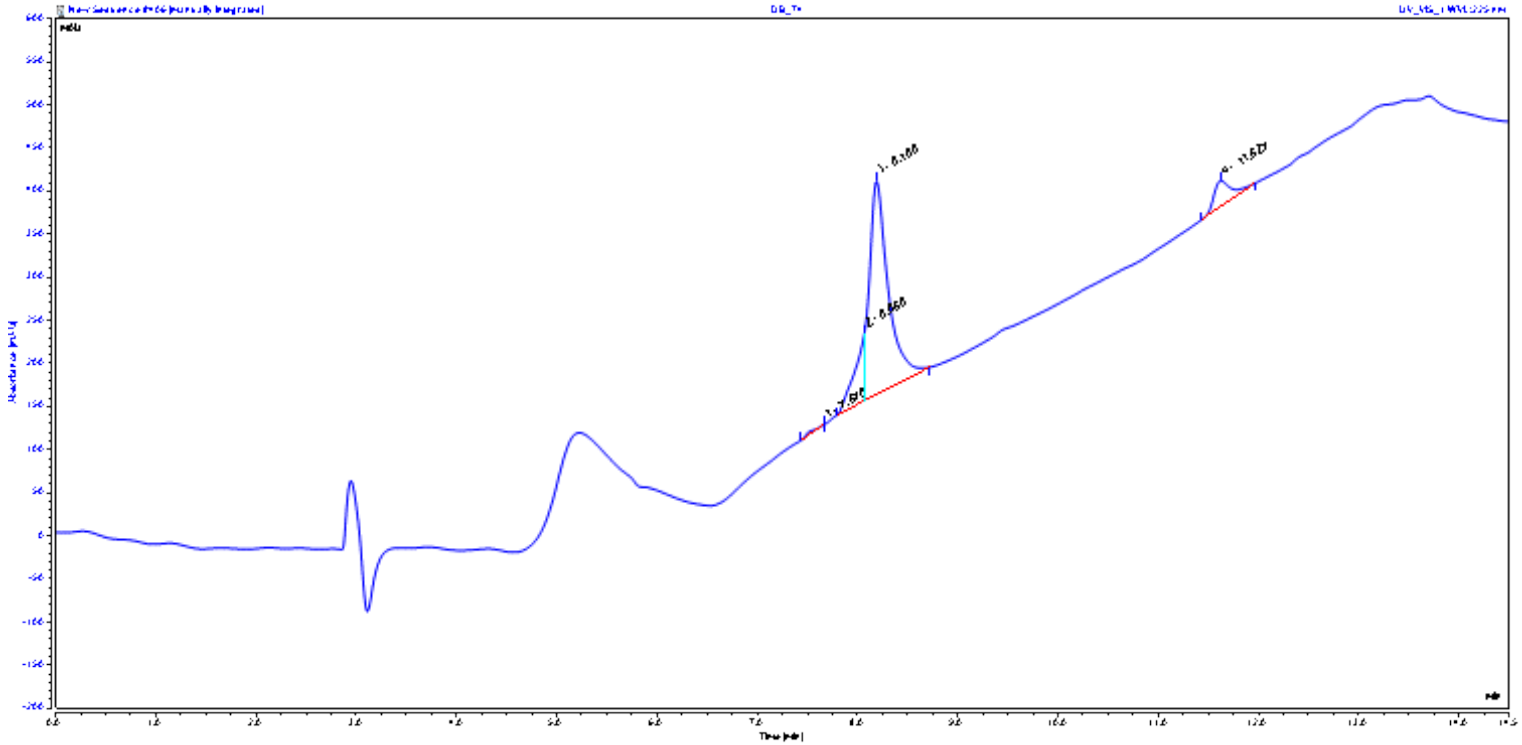


Mass spectrum (50-2000 m/z) of peak at 10.88 min for peptide **68**, m/z calcd for $[M+H]^+$: $C_{73}H_{103}N_{14}O_{23}^+$, 1543.7315; found; $[M+H]^+$ 1543.7315

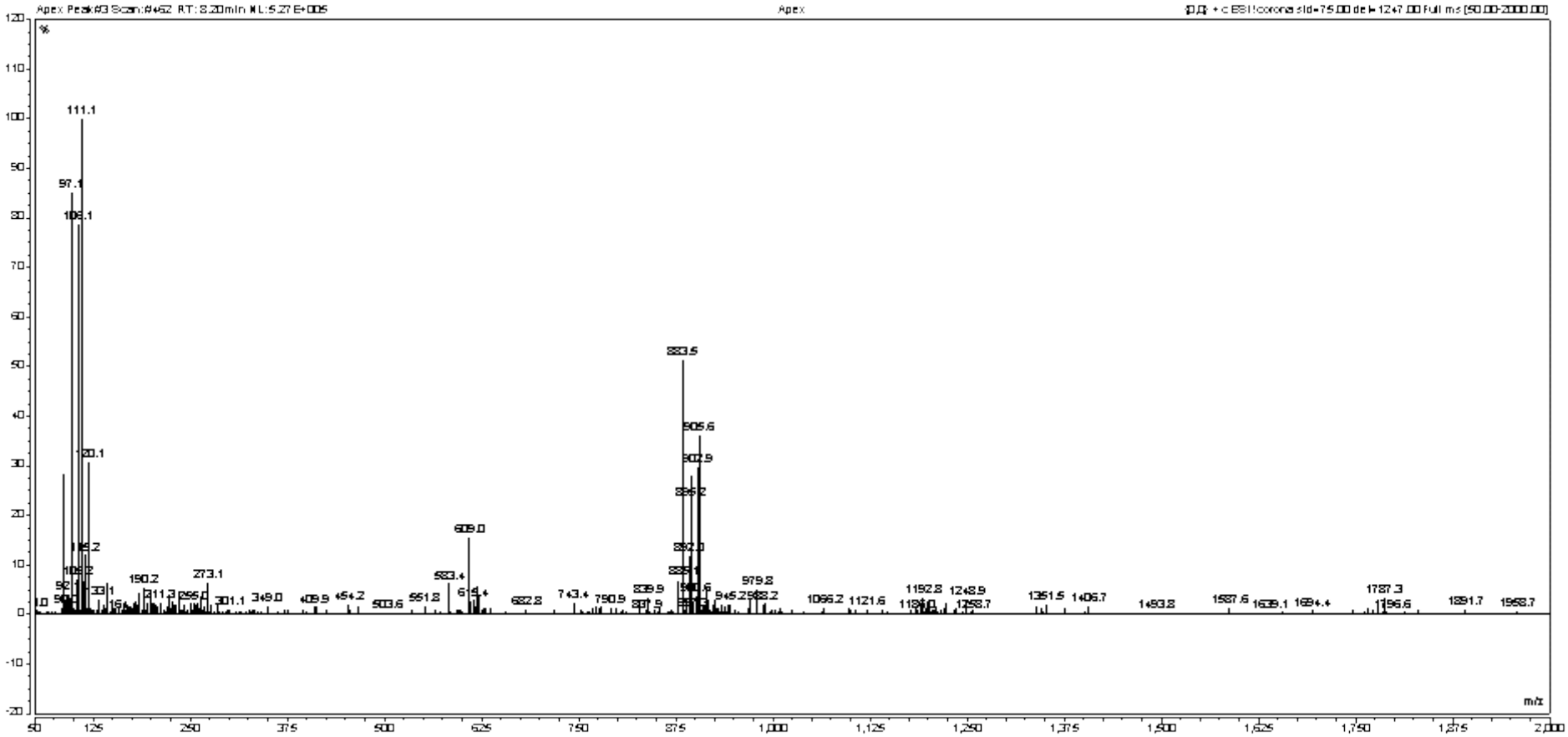
7.2.3 Synthesis of Ac-Pro_{13}-NH_2 (-$OCH_2COOH)_6$(i: 3, 5, 7, 10, 11, 12) (69):



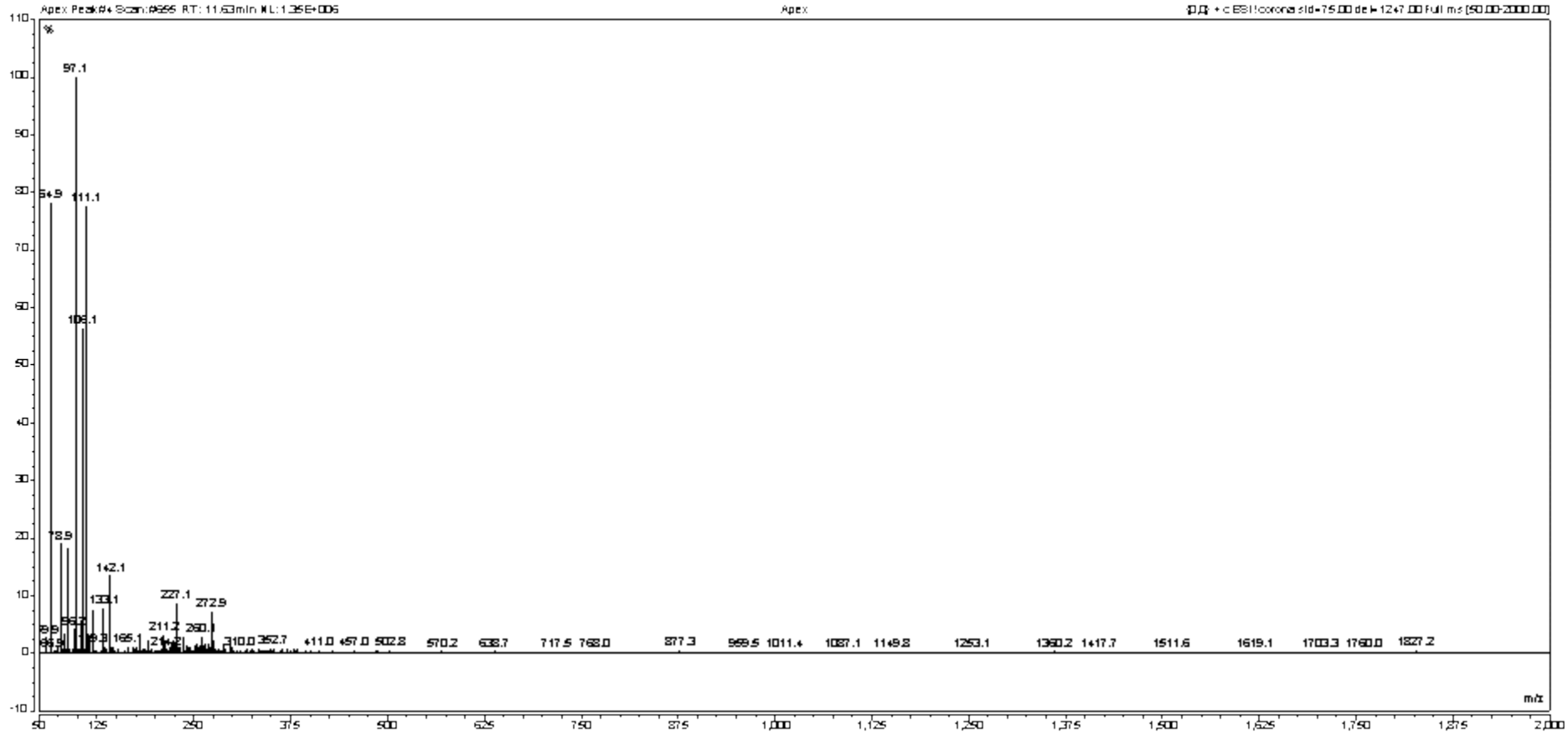
(69)



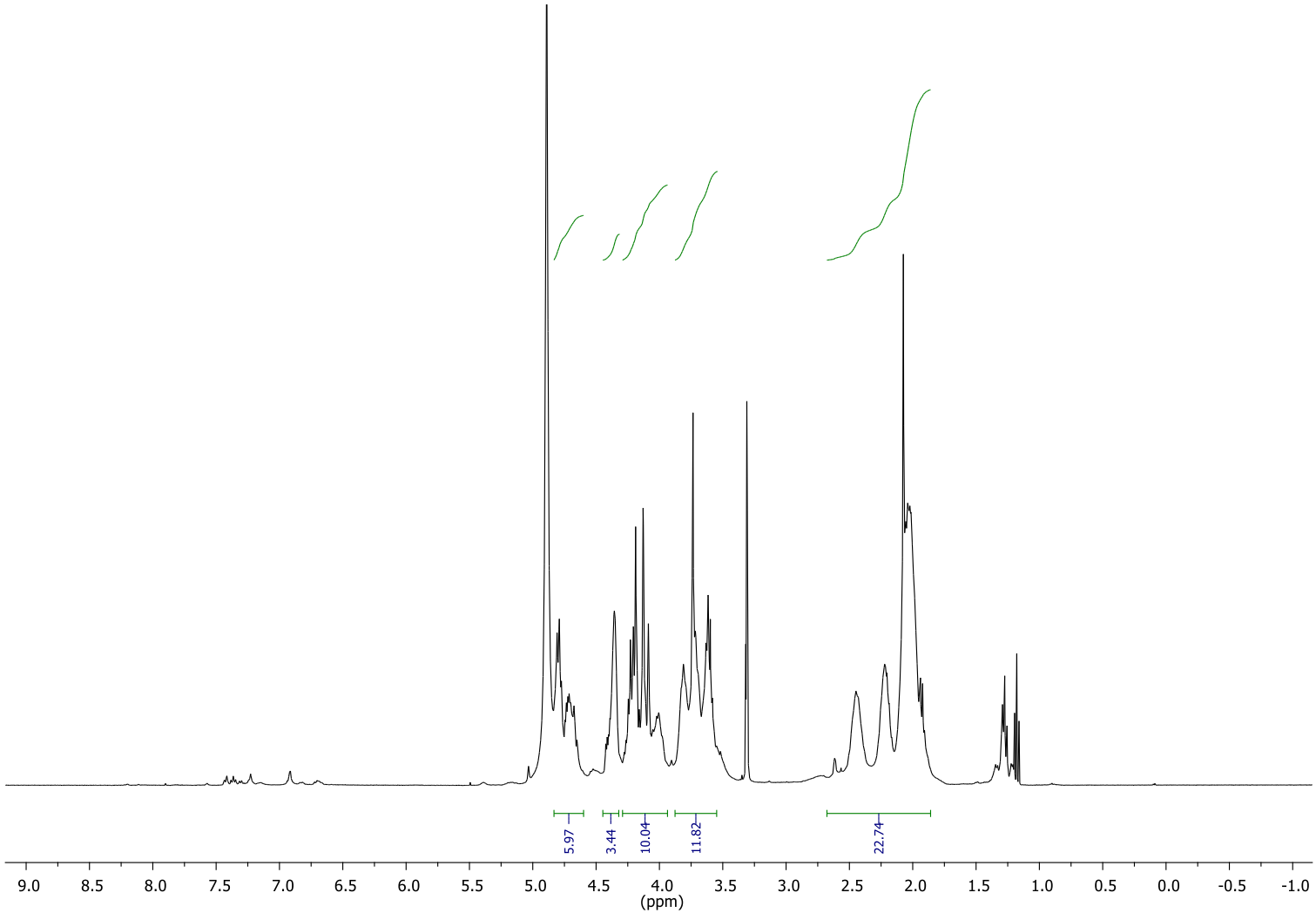
HPLC UV-Vis Spectrum of peptide **69**, rt = 8.188 min, 225 nm



Mass spectrum (50-2000 m/z) of peak at 8.20 min for peptide **69**, m/z calcd for $[M+H]^+$: $C_{79}H_{109}N_{14}O_{32}^+$, 1765.7327; found; $[M+Na]^+$ 1787.3; $[M+2H]^{2+}$ 883.5

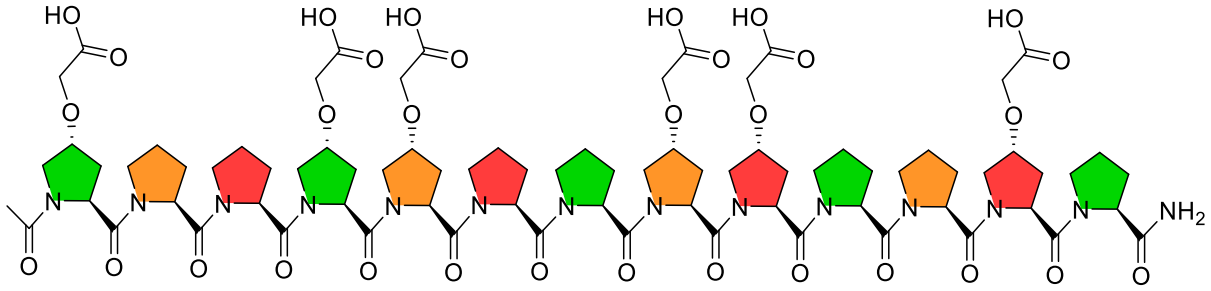


Mass spectrum (50-2000 m/z) of peak at 11.63 min for peptide **69**,

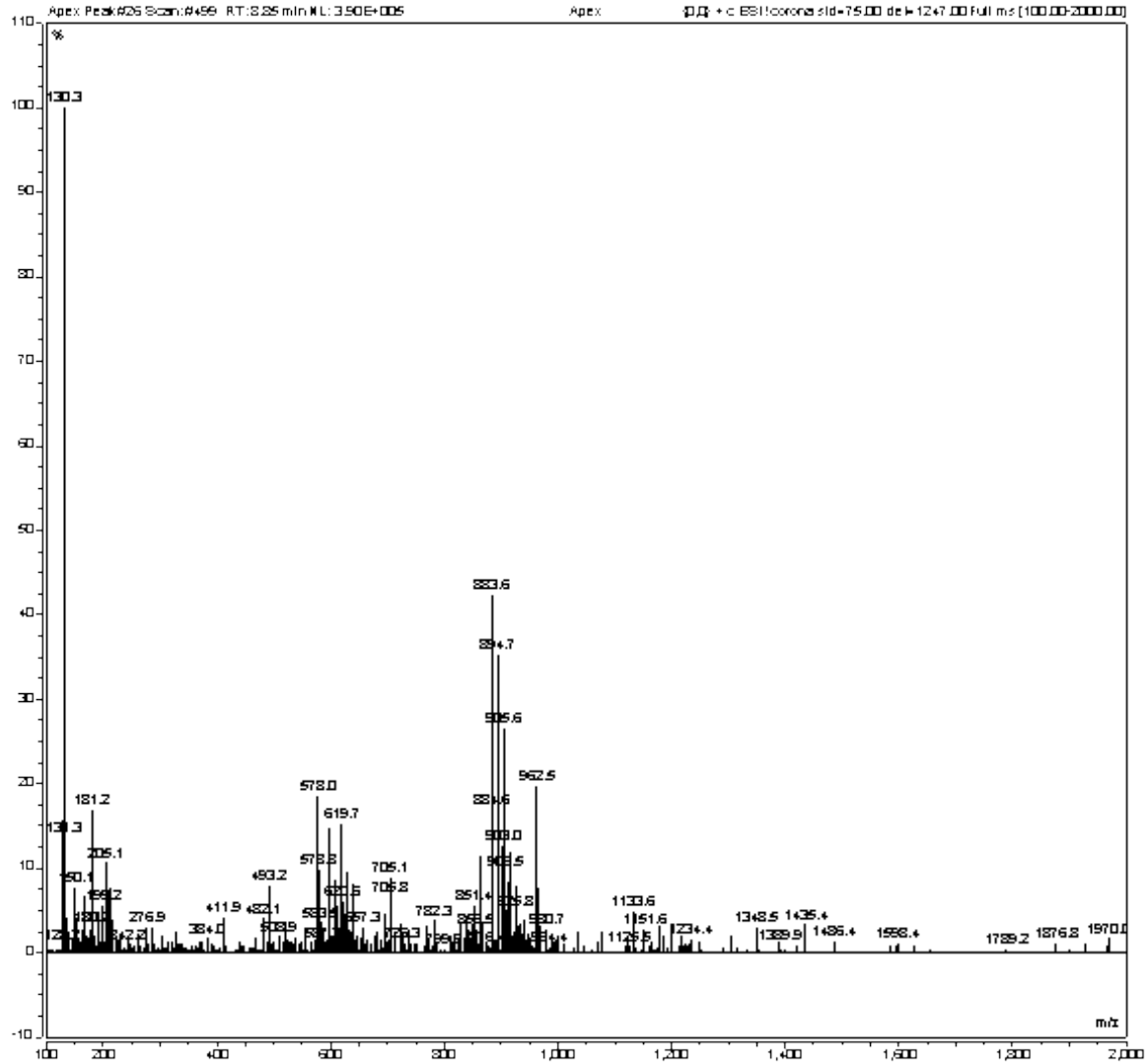


^{1}H NMR of **69** (400 MHz, MeOD)

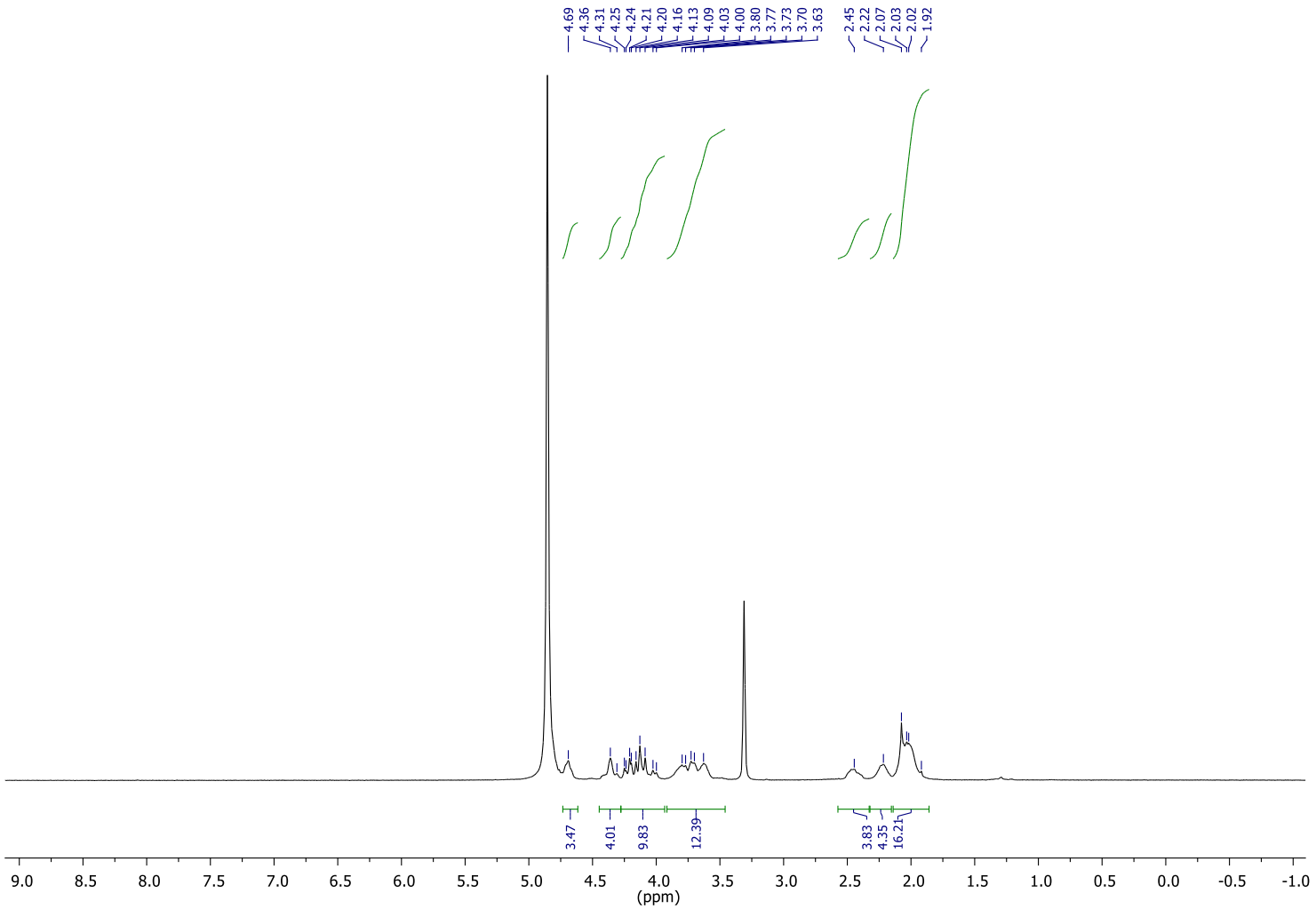
7.2.4 Synthesis of Ac-Pro13-NH2 (-OCH2COOH)6(i: 1, 4, 5, 8, 9, 12) (70):



(70)

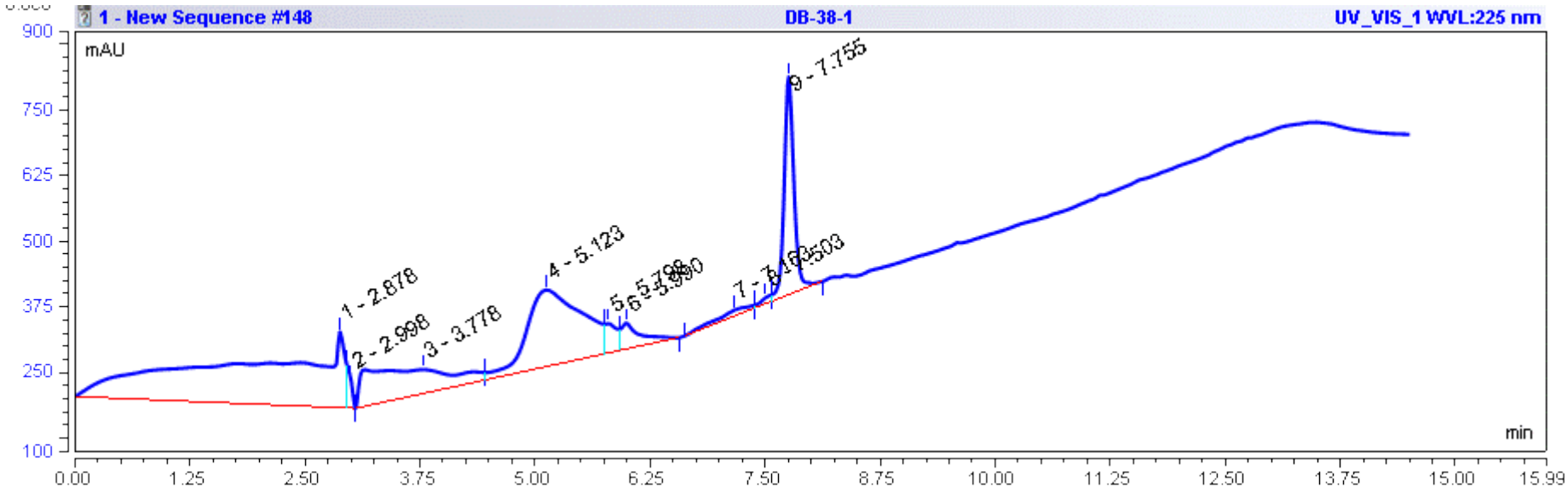


Mass spectrum (50-2000 m/z) of peak at 8.20 min for peptide **70**, m/z calcd for $[M+H]^+$: $C_{79}H_{109}N_{14}O_{32}^+$, 1765.7327; found; $[M+2H]^{2+}$ 883.6

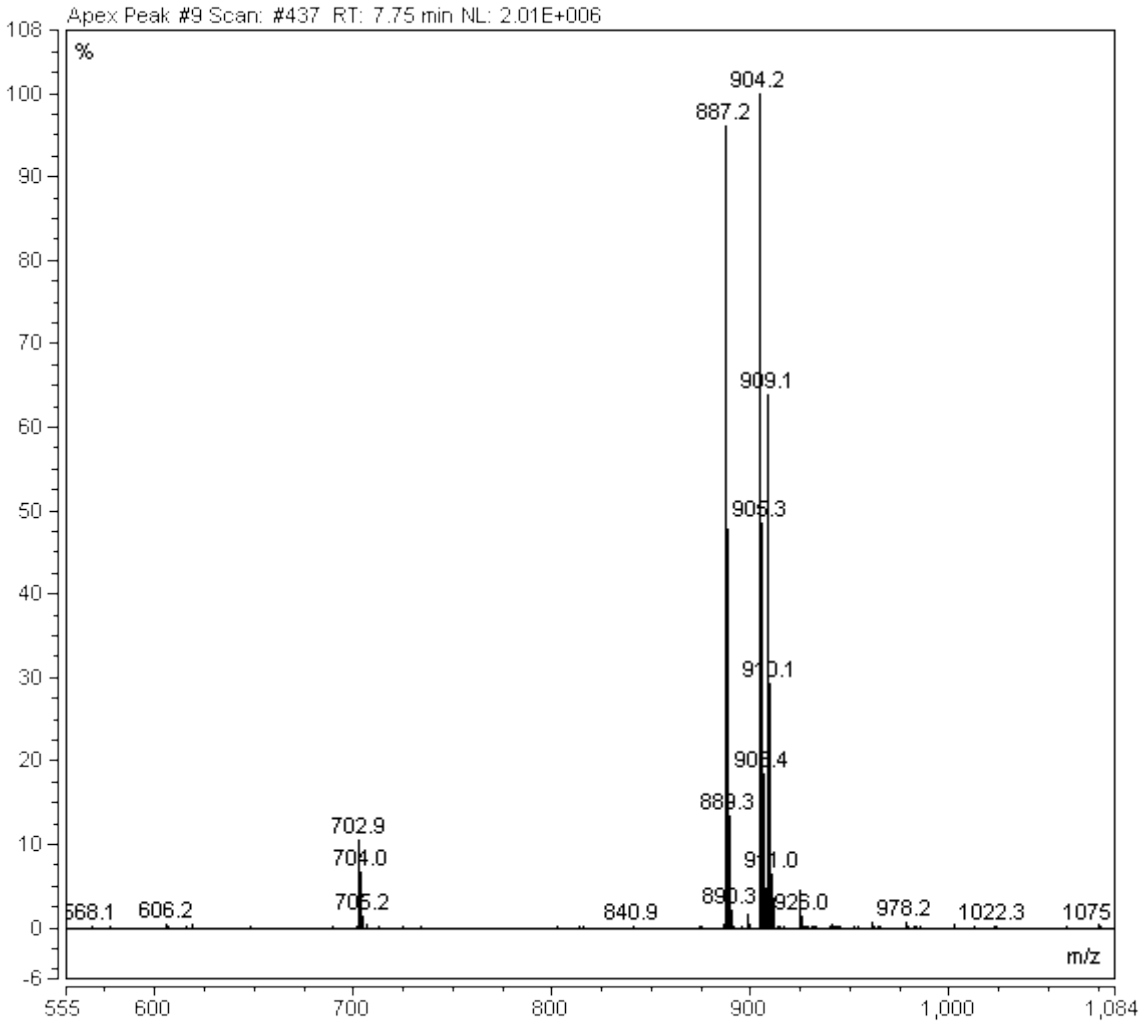


^{1}H NMR of peptide **70** (400 MHz, MeOD)

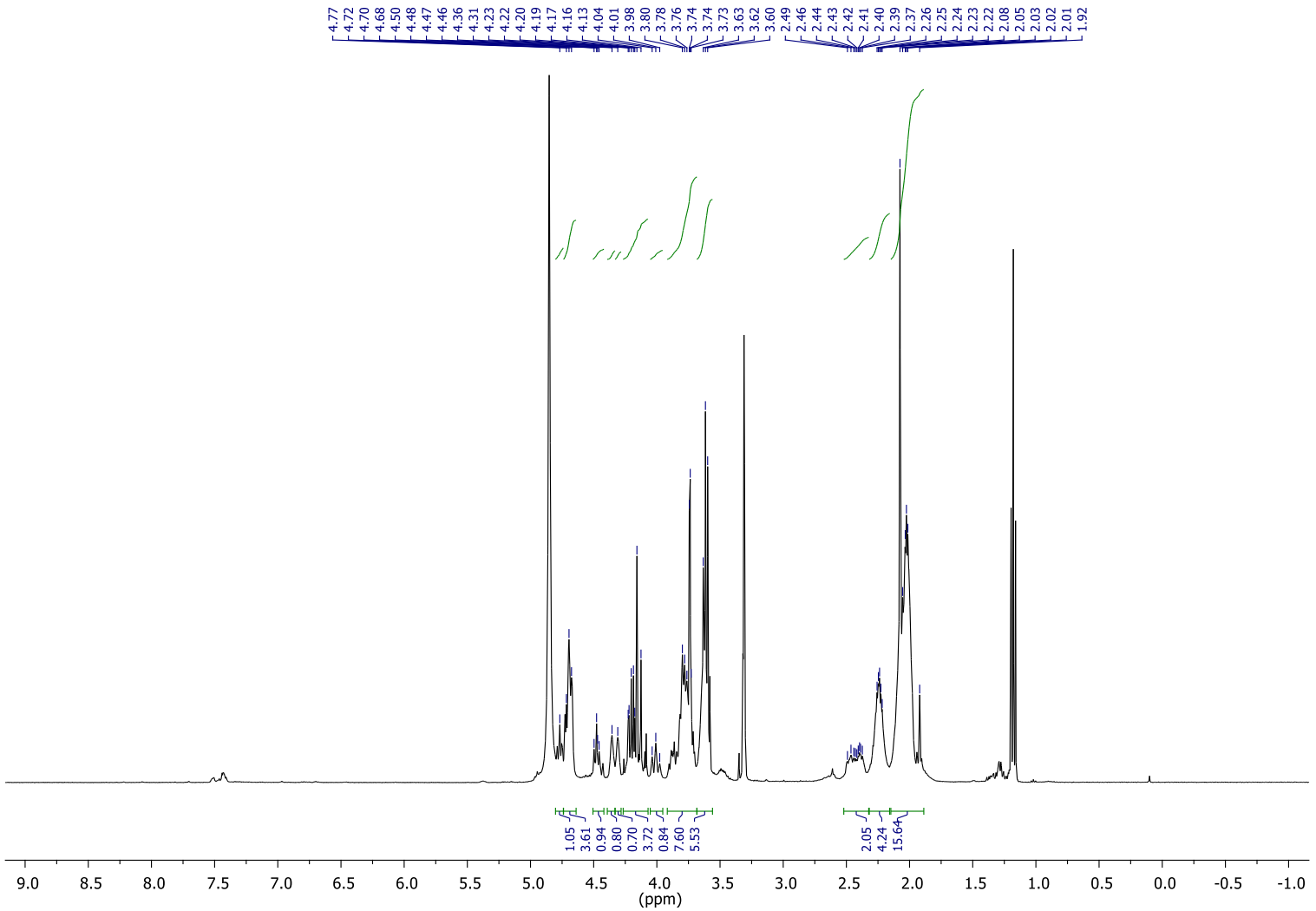
7.2.5 Synthesis of Ac-Pro$_7$-NH$_2$ (-OCH$_2$COOH)$_2$(i: 1, 7) (71):



HPLC UV-Vis Spectrum of peptide **71**, rt = 7.755 min, 225 nm

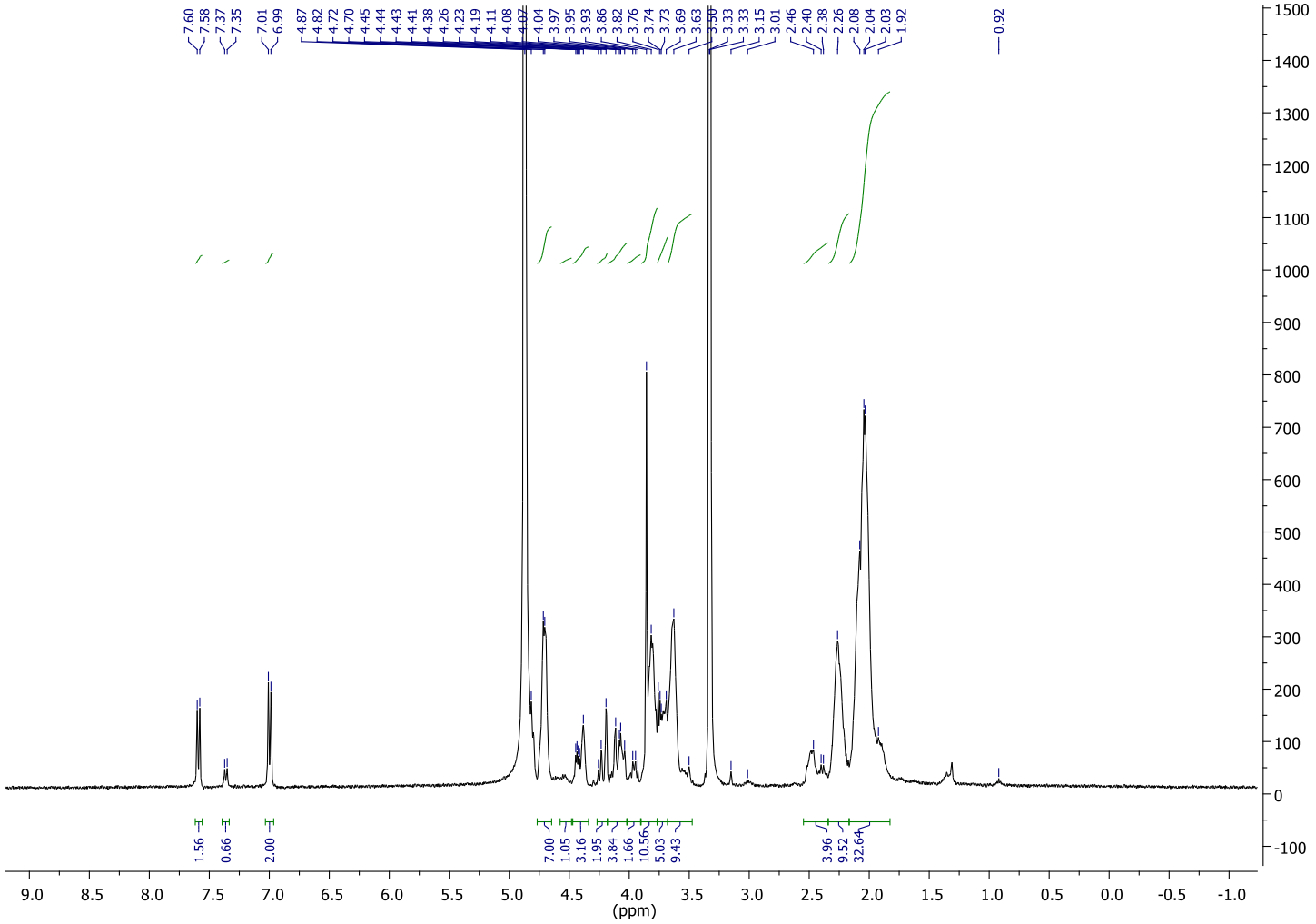


Mass spectrum (555-1084 m/z) of peak at 7.755 min for peptide **71**, m/z calcd for $[M+H]^+$: $C_{41}H_{59}N_8O_{14}^+$, 887.4145; found; $[M+H]^+$ 887.2; $[M+Na]^+$ 909.1

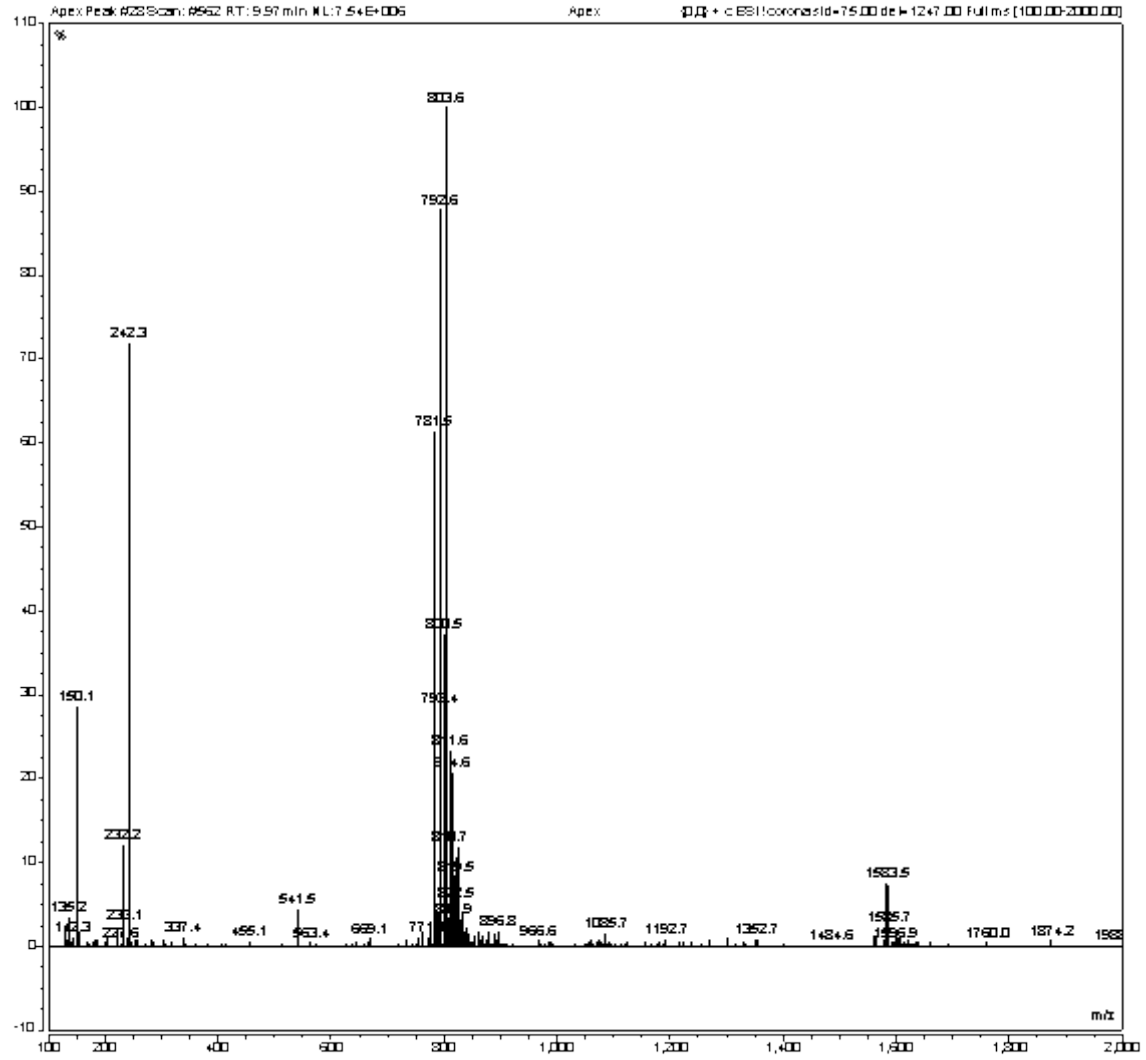


^{1}H NMR of peptide **71** (400 MHz, MeOD)

7.2.6 Synthesis of CH_3-C_6H_4-CO-Pro_{13}-NH_2 (-$OCH_2COOH)_2$(i: 3, 12) (72):

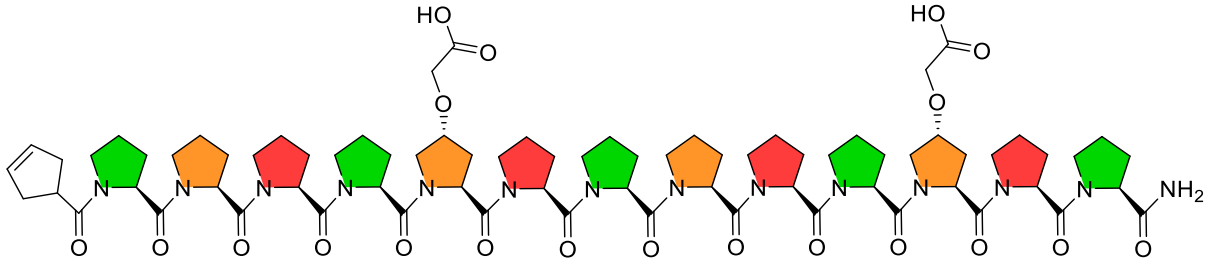


^{1}H NMR of peptide **72** (400 MHz, MeOD)



Mass spectrum (100-2000 m/z) of peak at 9.997 min for peptide **72**, m/z calcd for $[M+H]^+$: $C_{77}H_{105}N_{14}O_{21}^+$, 1561.7573; found; $[M+Na]^+$ 1583.5; $[M+2H]^{2+}$ 781.5

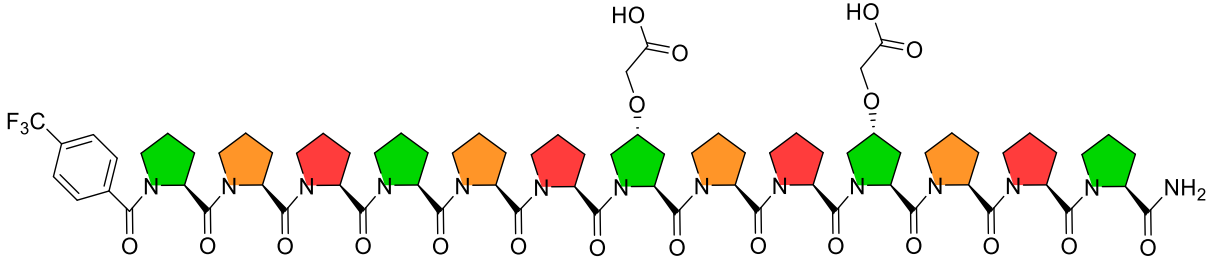
7.2.7 Synthesis of Cyclopent-3-ene-Pro_{13}-NH_2 $(-OCH_2COOH)_2$(i: 5, 11) (73):



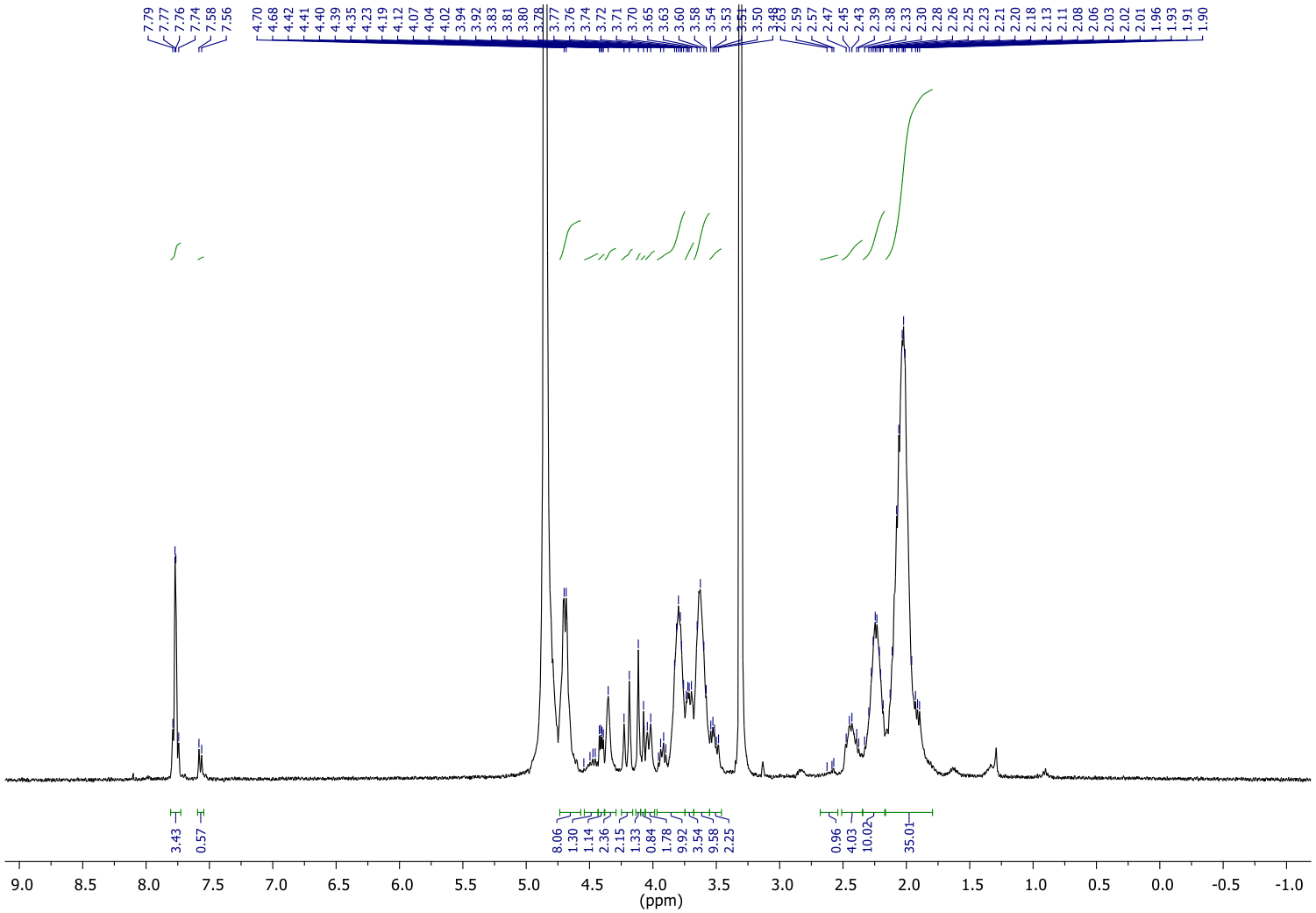
(73)

Impure product after SPPS from HPLC analysis, as such this peptide was not carried forward for further studies pending results from peptides 72 and 74.

7.2.8 Synthesis of F_3C-C_6H_4-CO-Pro_{13}-NH_2 $(-OCH_2COOH)_2$(i: 7, 10) (74):

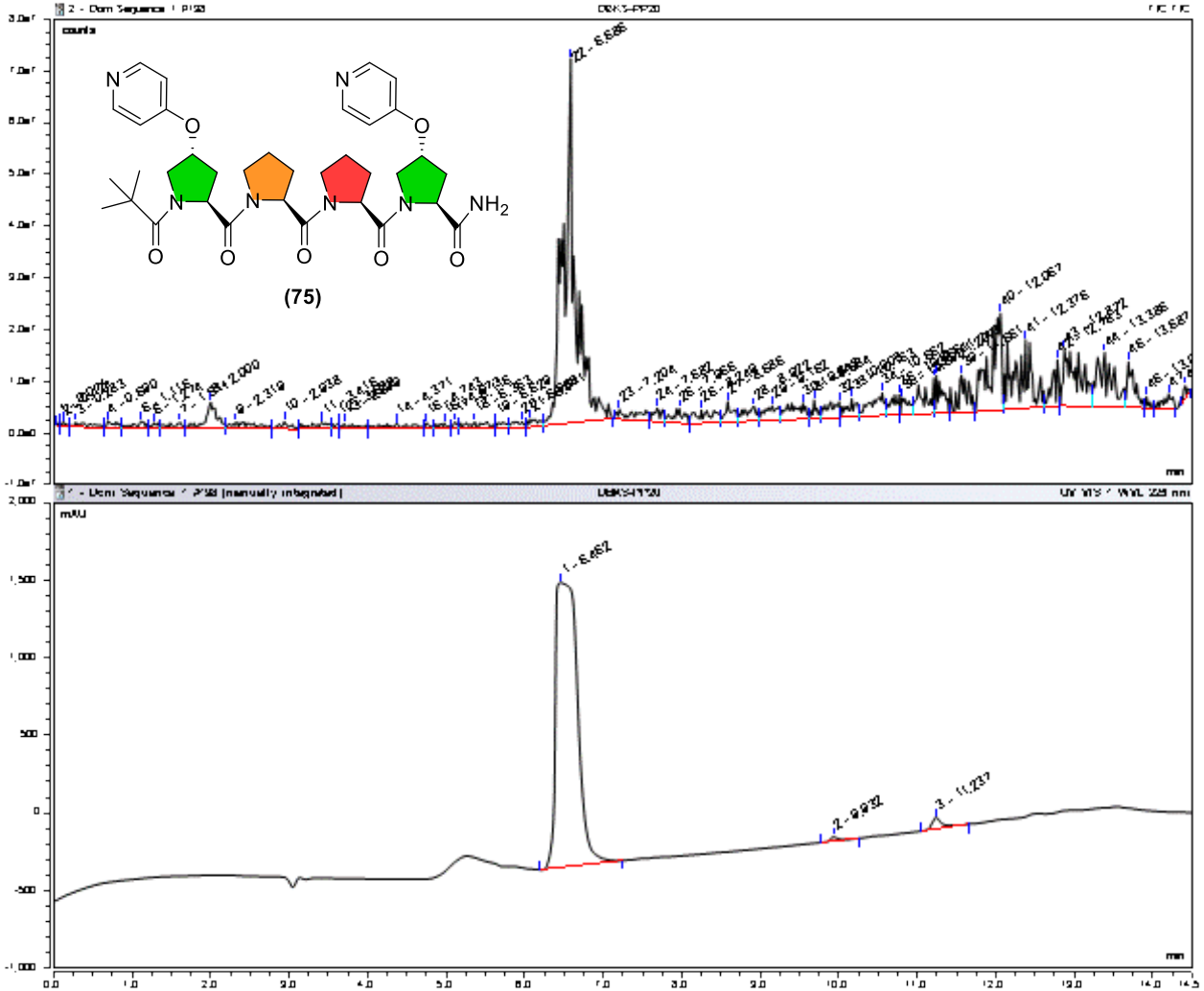


(74)

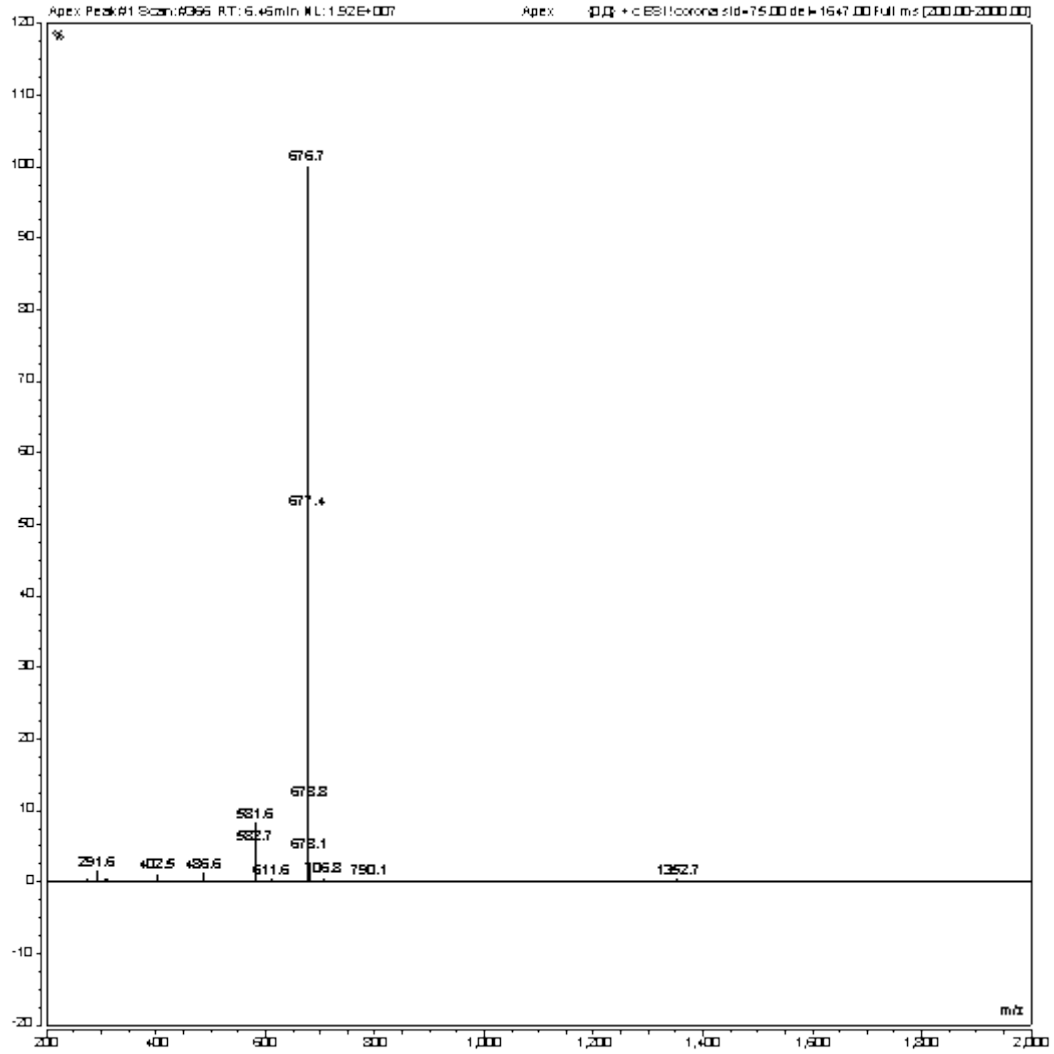


^{1}H NMR of peptide **74** (400 MHz, MeOD)

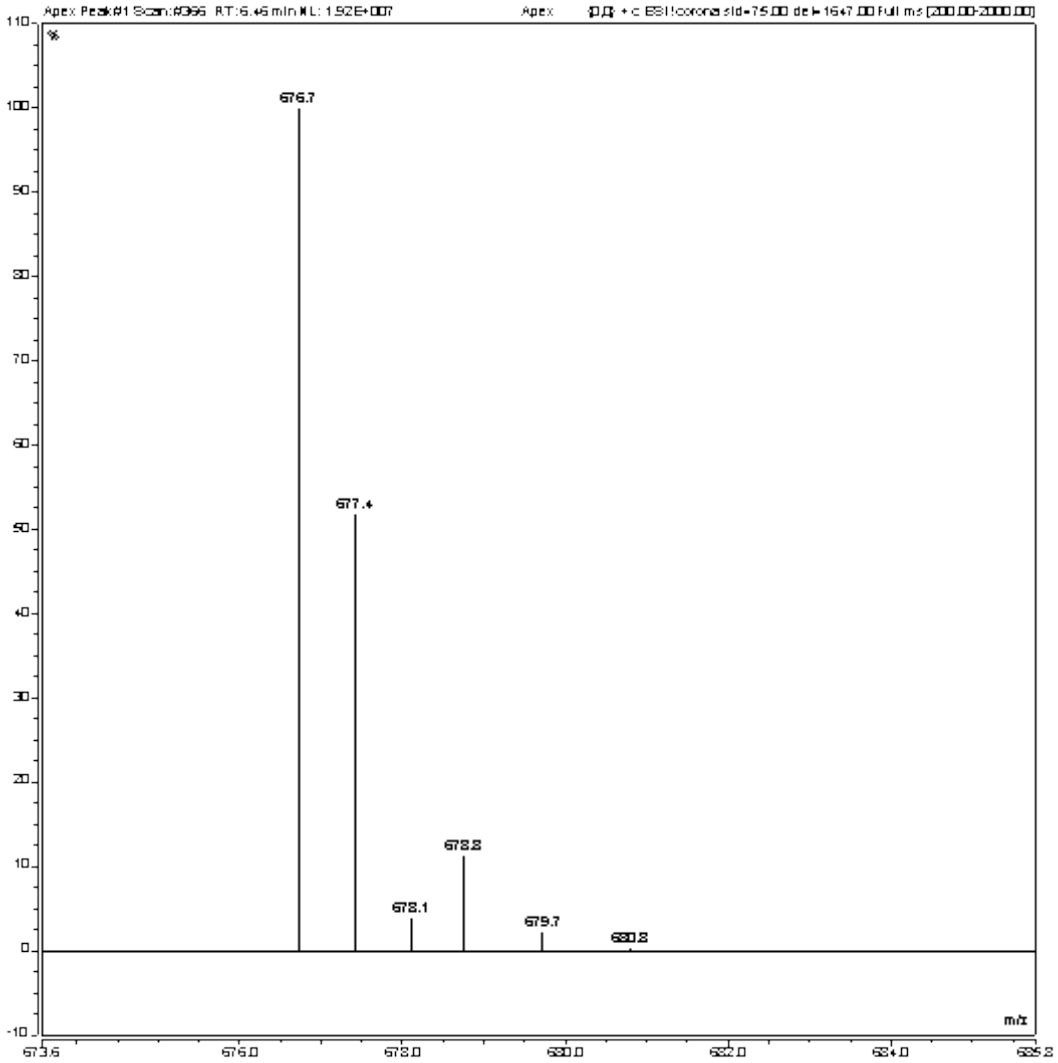
7.2.9 Synthesis of Piv-Pro$_4$-NH$_2$ (-OC$_5$N)$_2$(i: 1, 4) (75):



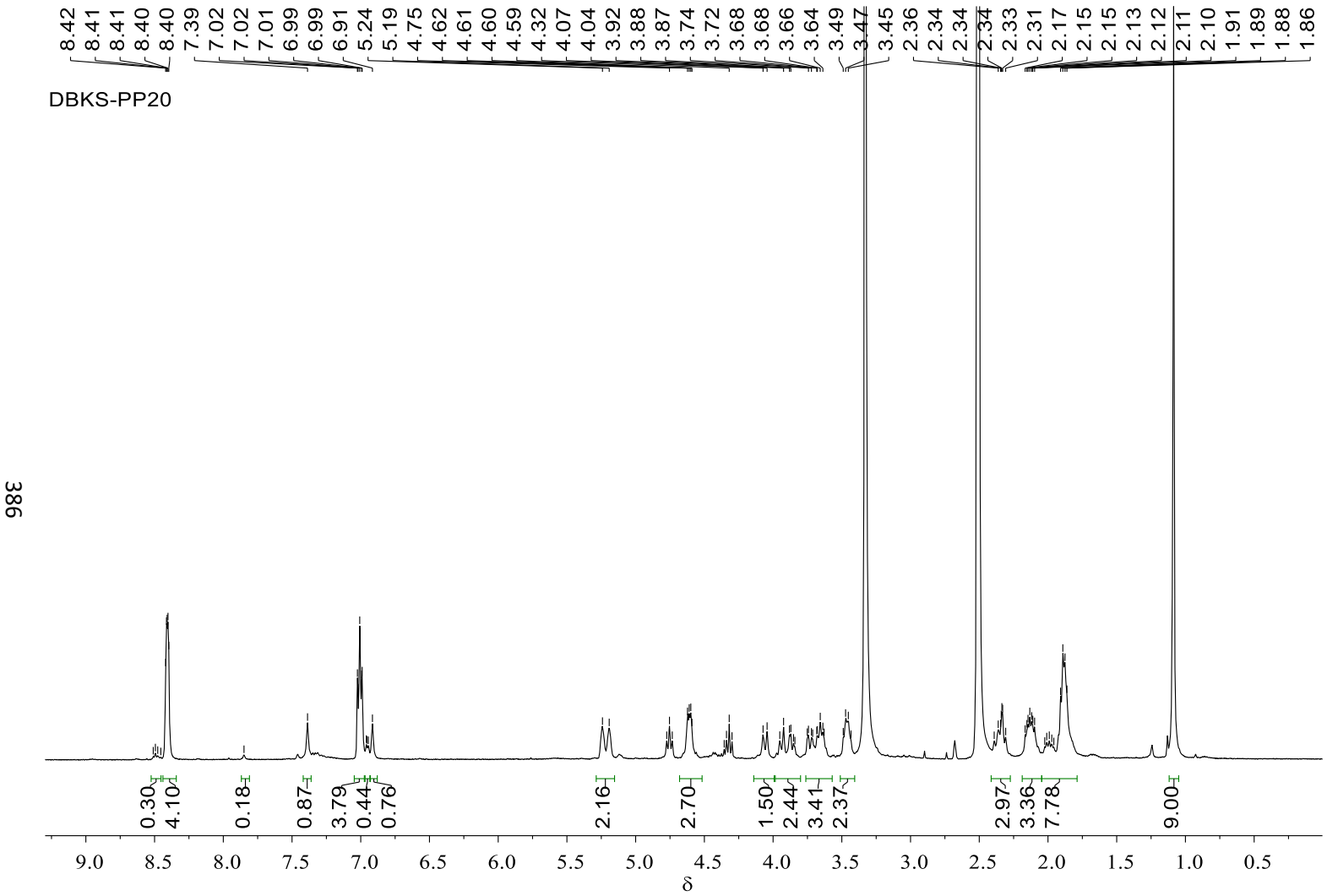
HPLC UV-Vis Spectrum of peptide **75**, rt = 6.462 min, 225 nm



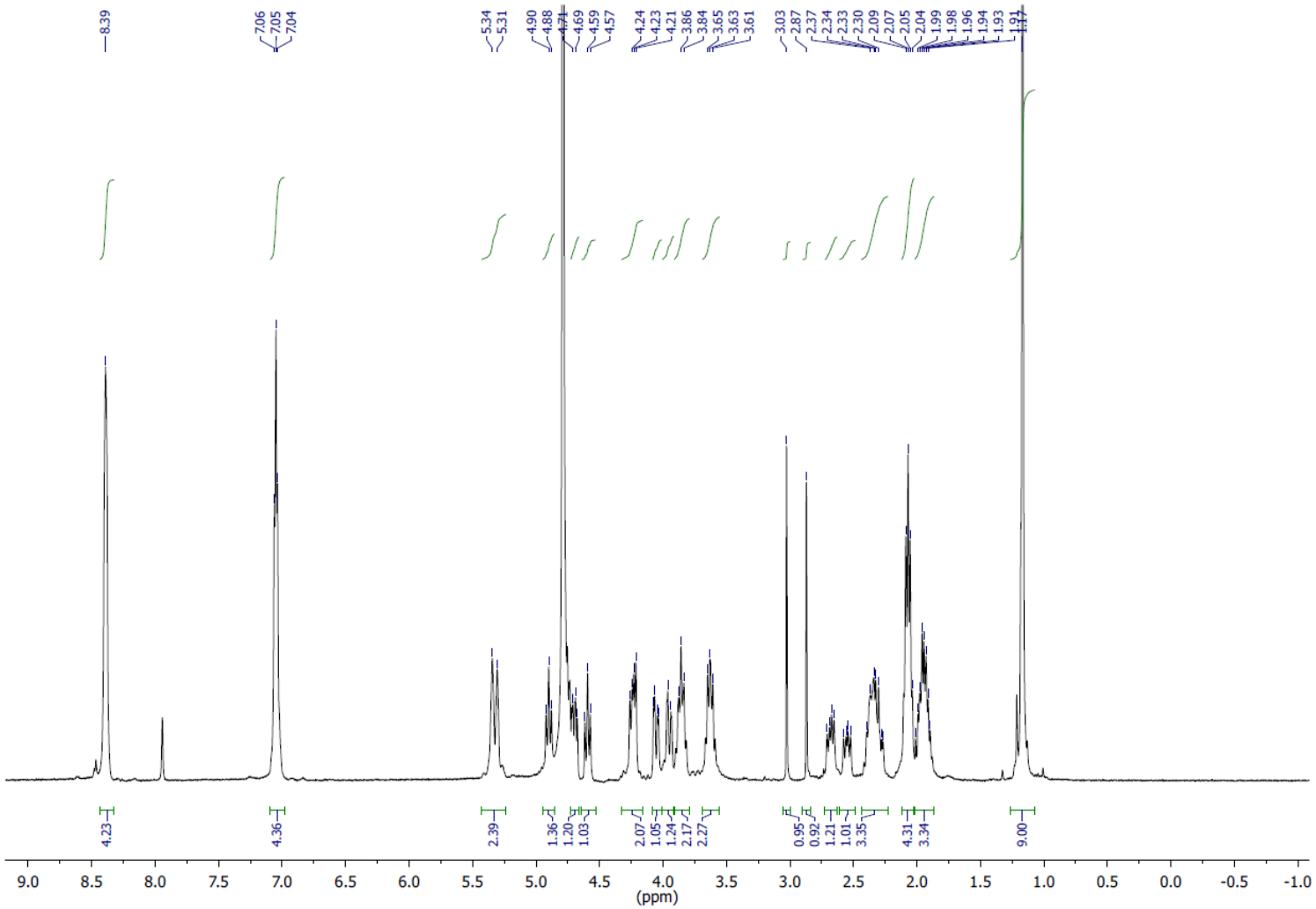
Mass spectrum (50-2000 m/z) of peak at 6.46 min for peptide **75**, m/z calcd for $[M+H]^+$: $C_{35}H_{46}N_7O_7^+$, 676.3453; found; $[M+H]^+$ 676.7; $[2M+H]^+$ 1352.7



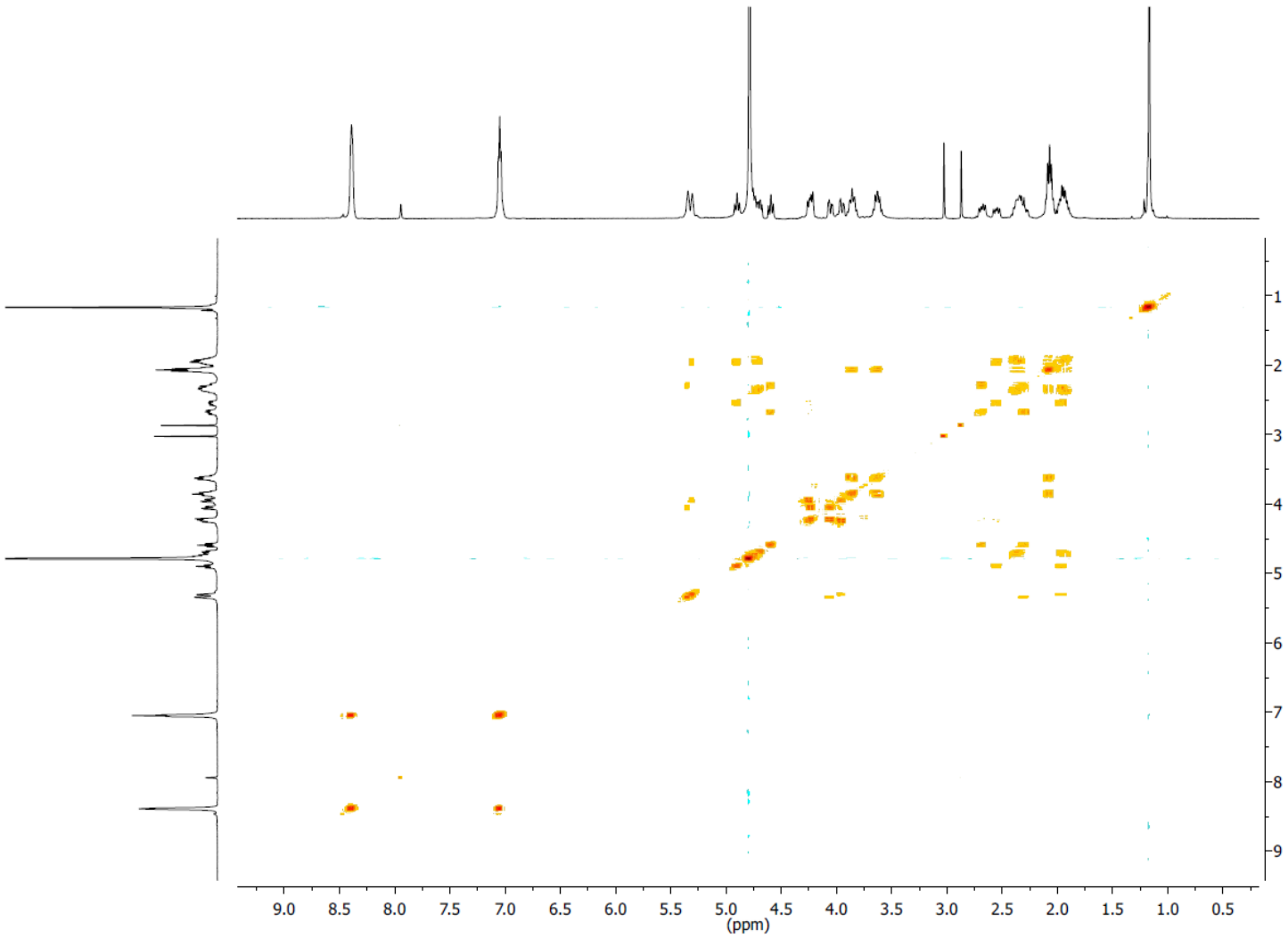
Mass spectrum (670-685 m/z) of peak at 6.46 min for peptide **75**,



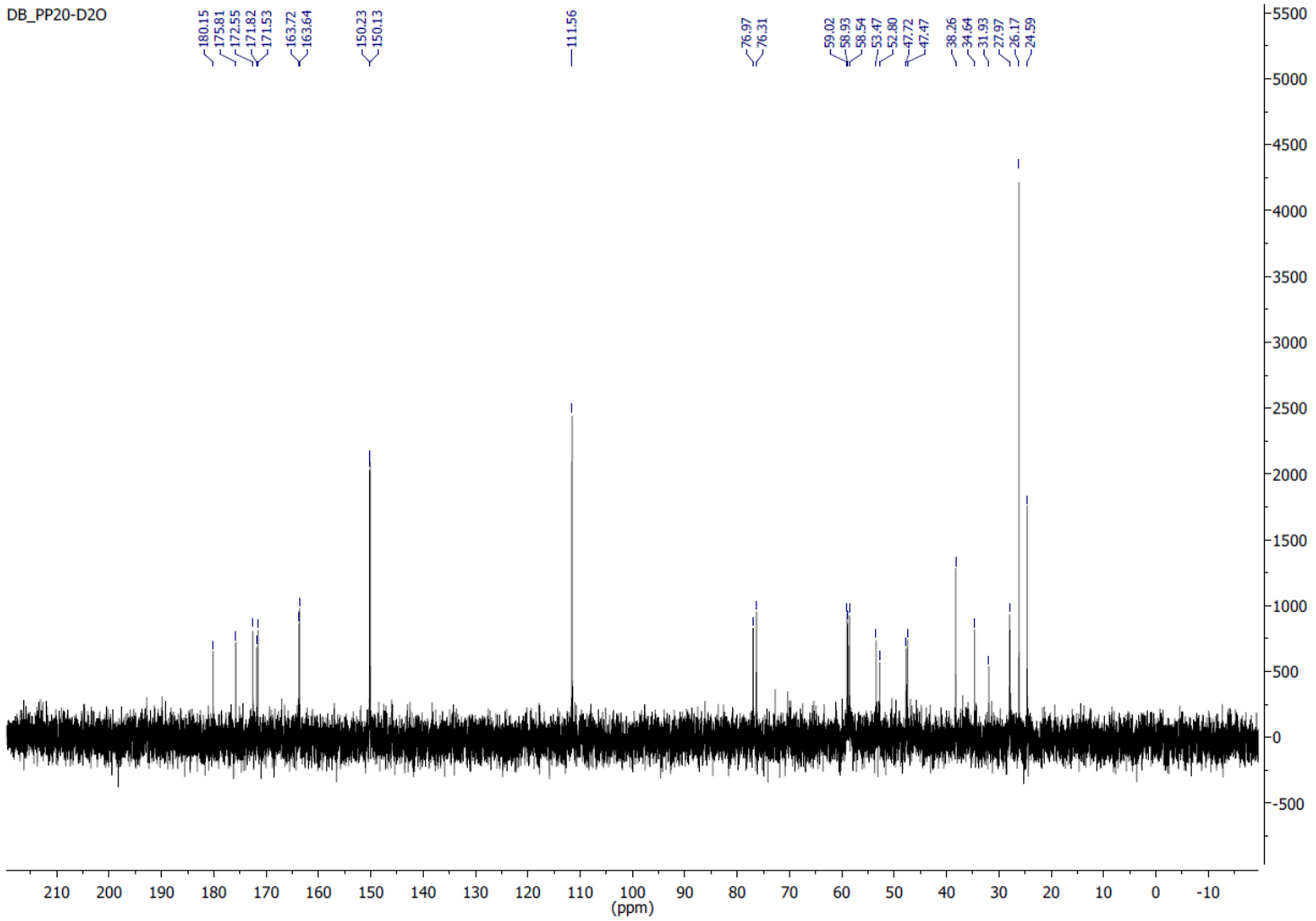
^{1}H NMR spectrum of **75** (400 MHz, DMSO-d_6, 298 K).



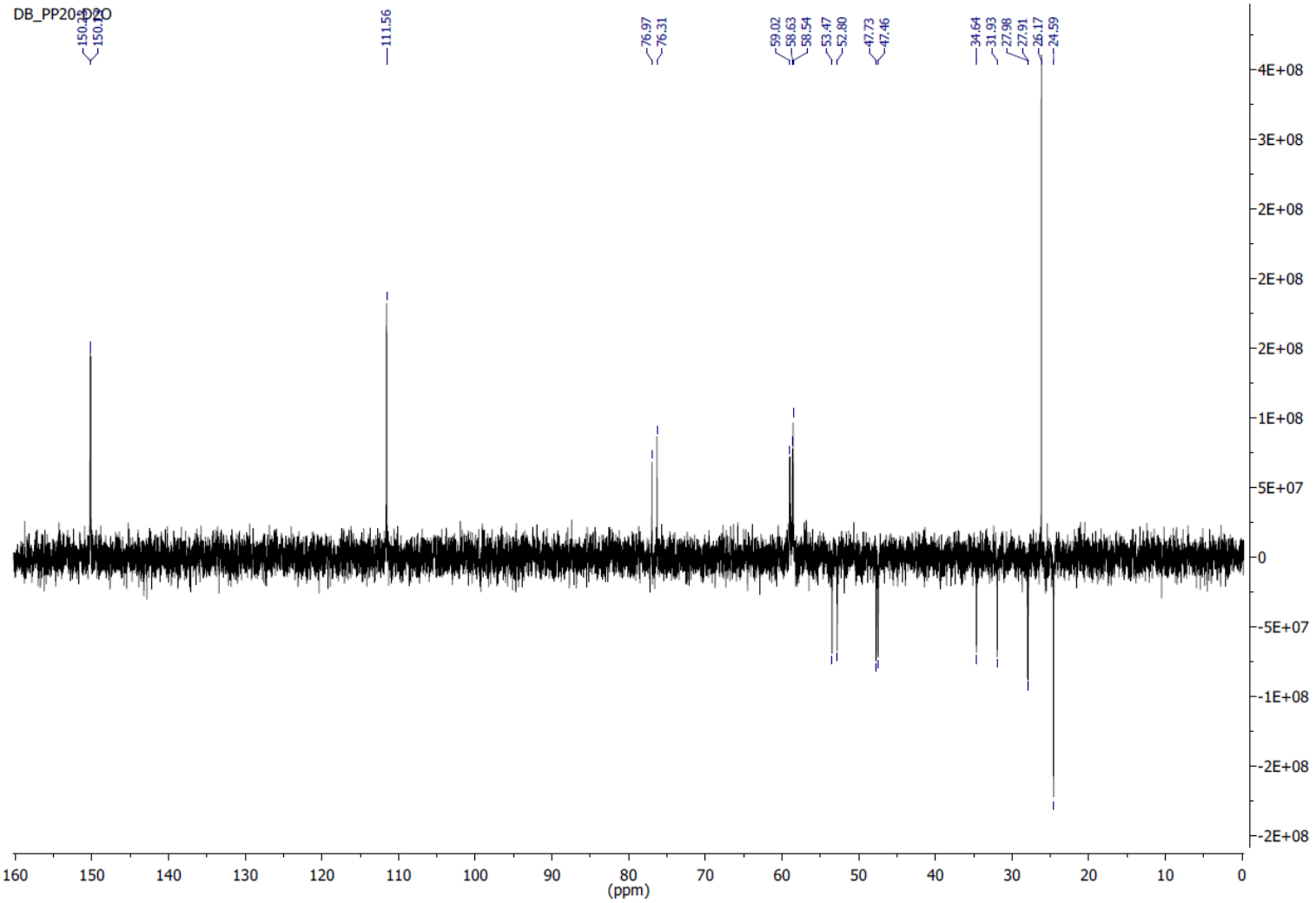
^{1}H NMR spectrum of **75** (400 MHz, D_2O).



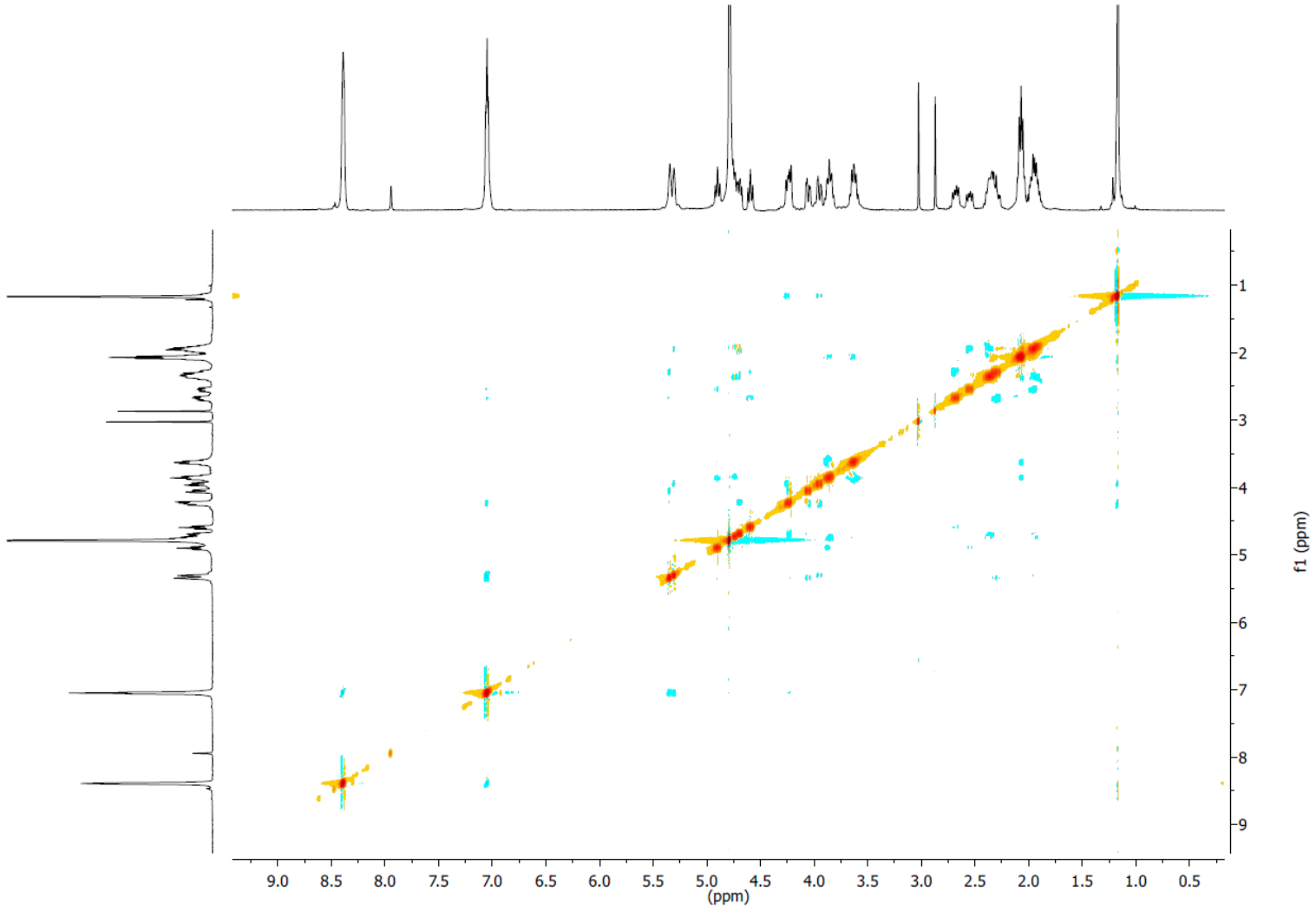
^{1}H-^{1}H COSY NMR spectrum of **75** (400 MHz, D_2O).



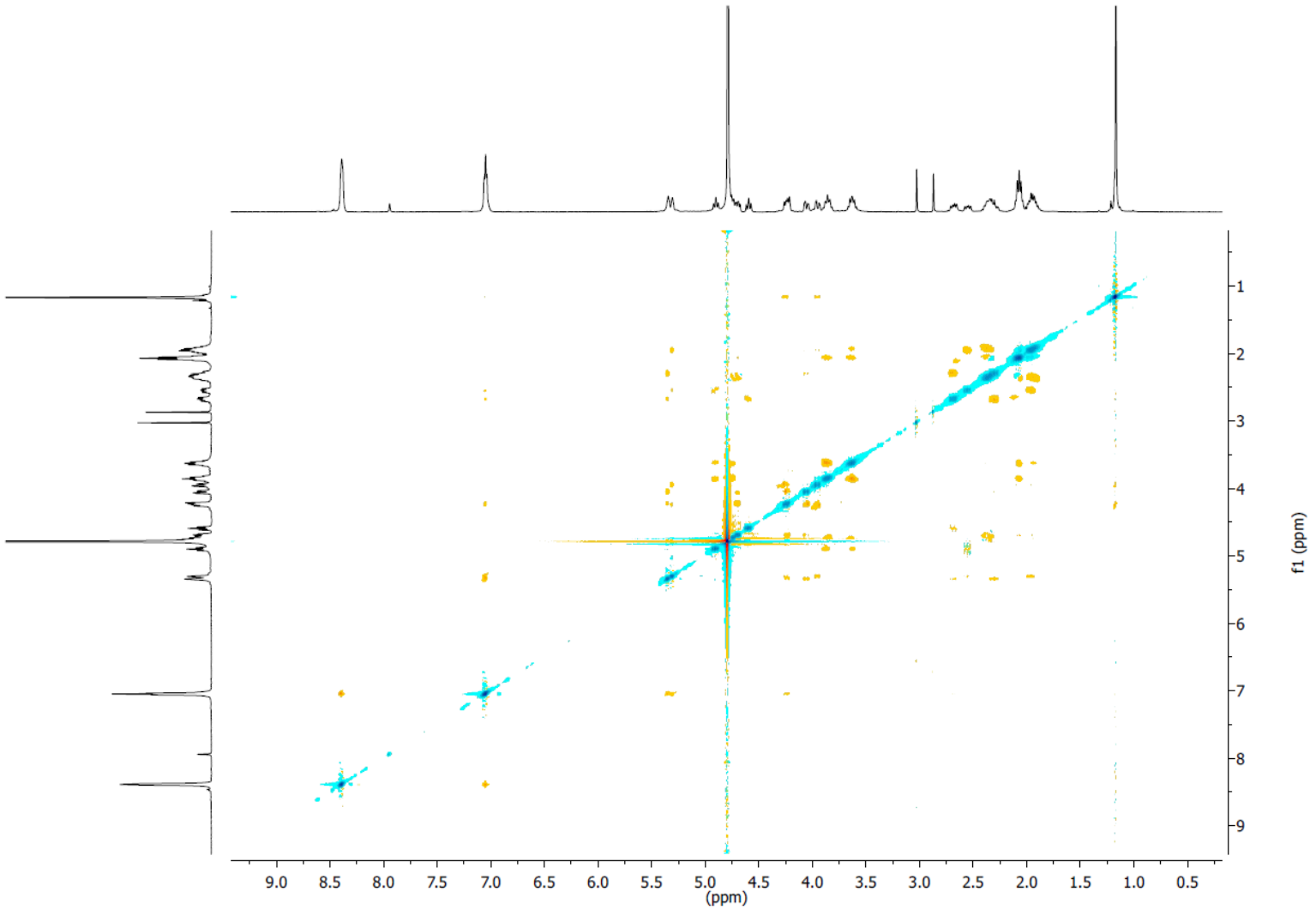
^{13}C NMR spectrum of **75** (101 MHz, D_2O).



^{13}C DEPT NMR spectrum of **75** (101 MHz, D_2O).

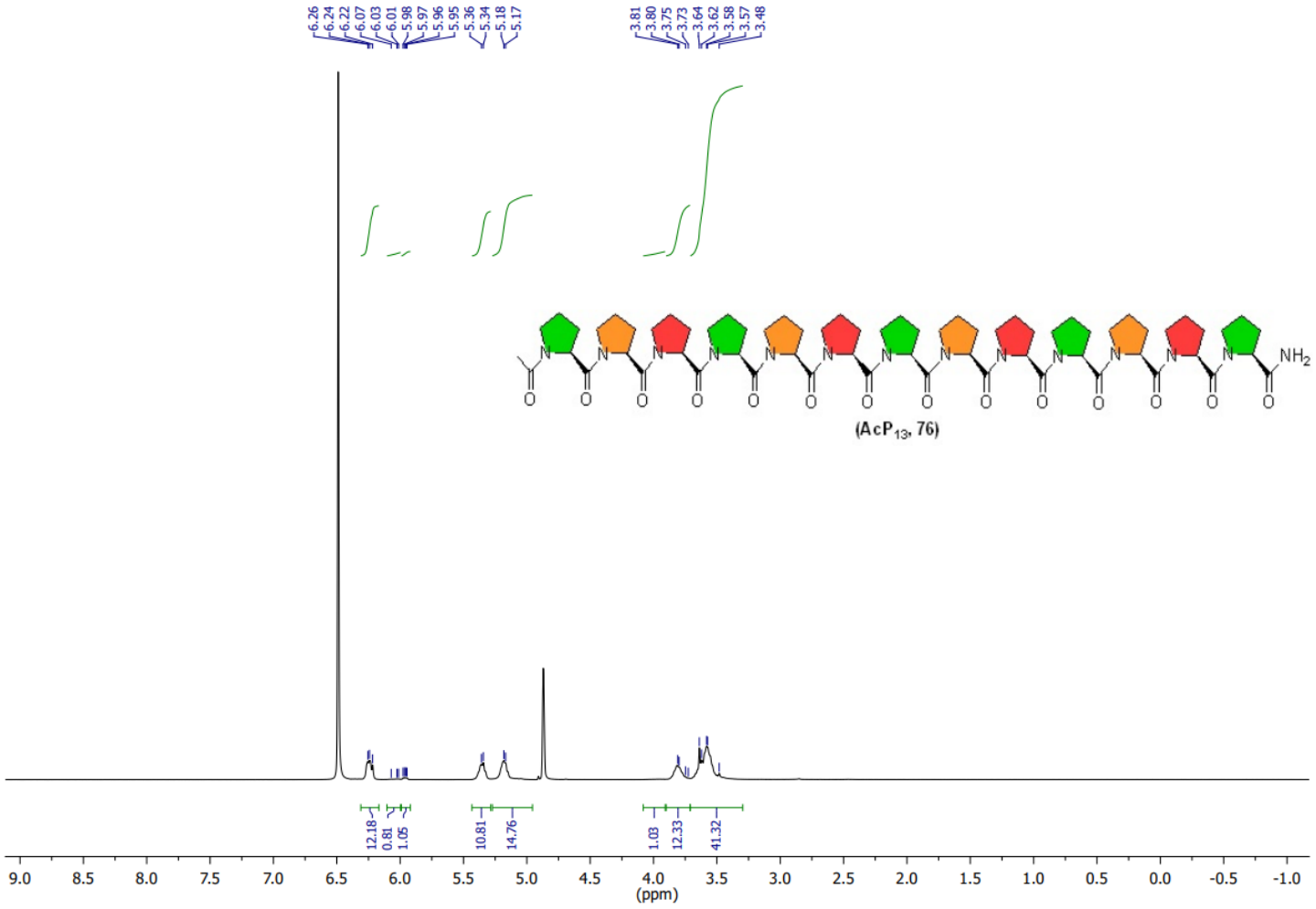


^{1}H NOESY NMR spectrum of **75** (400 MHz, D_2O).

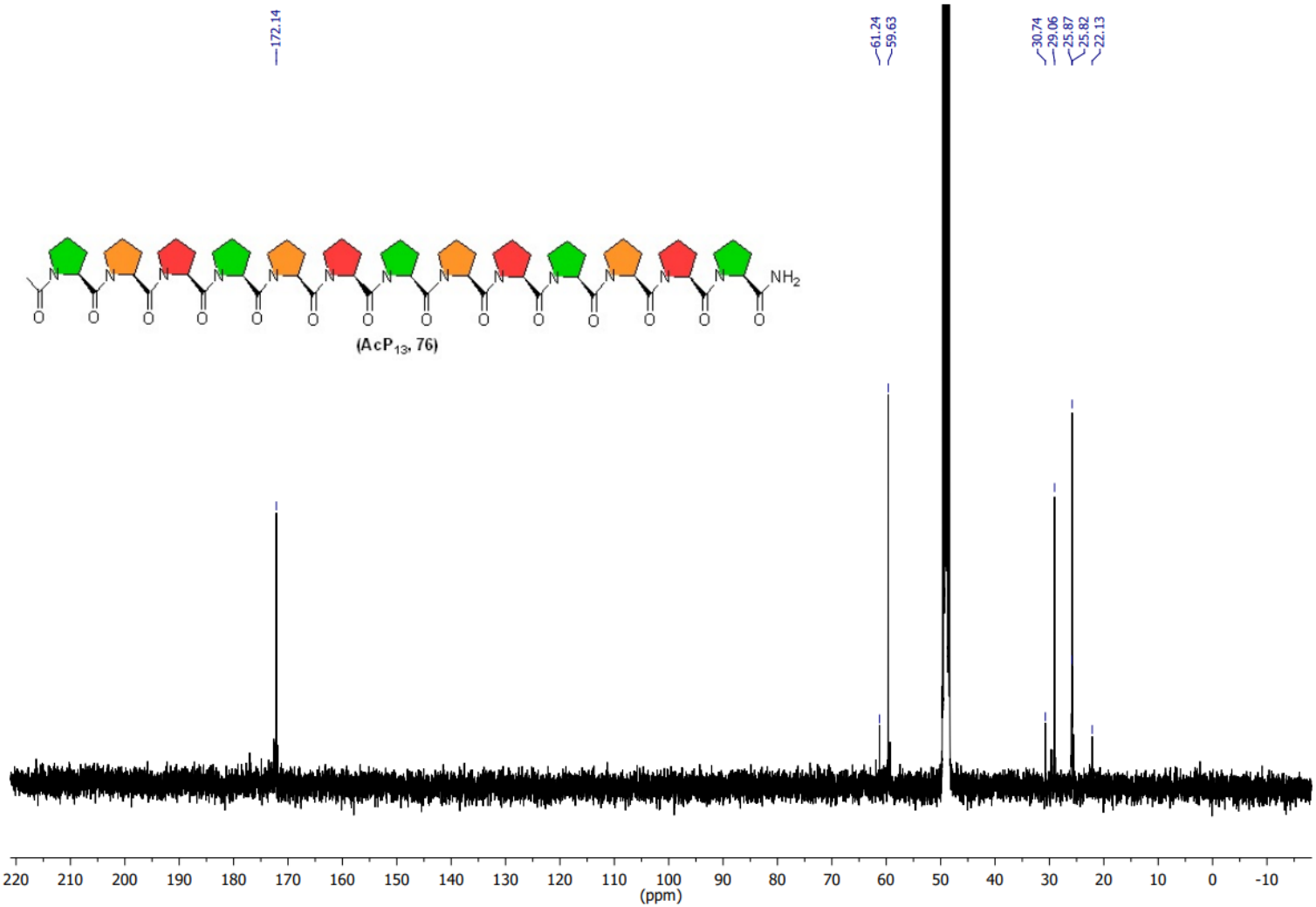


^{1}H ROESY NMR spectrum of **75** (400 MHz, D_2O).

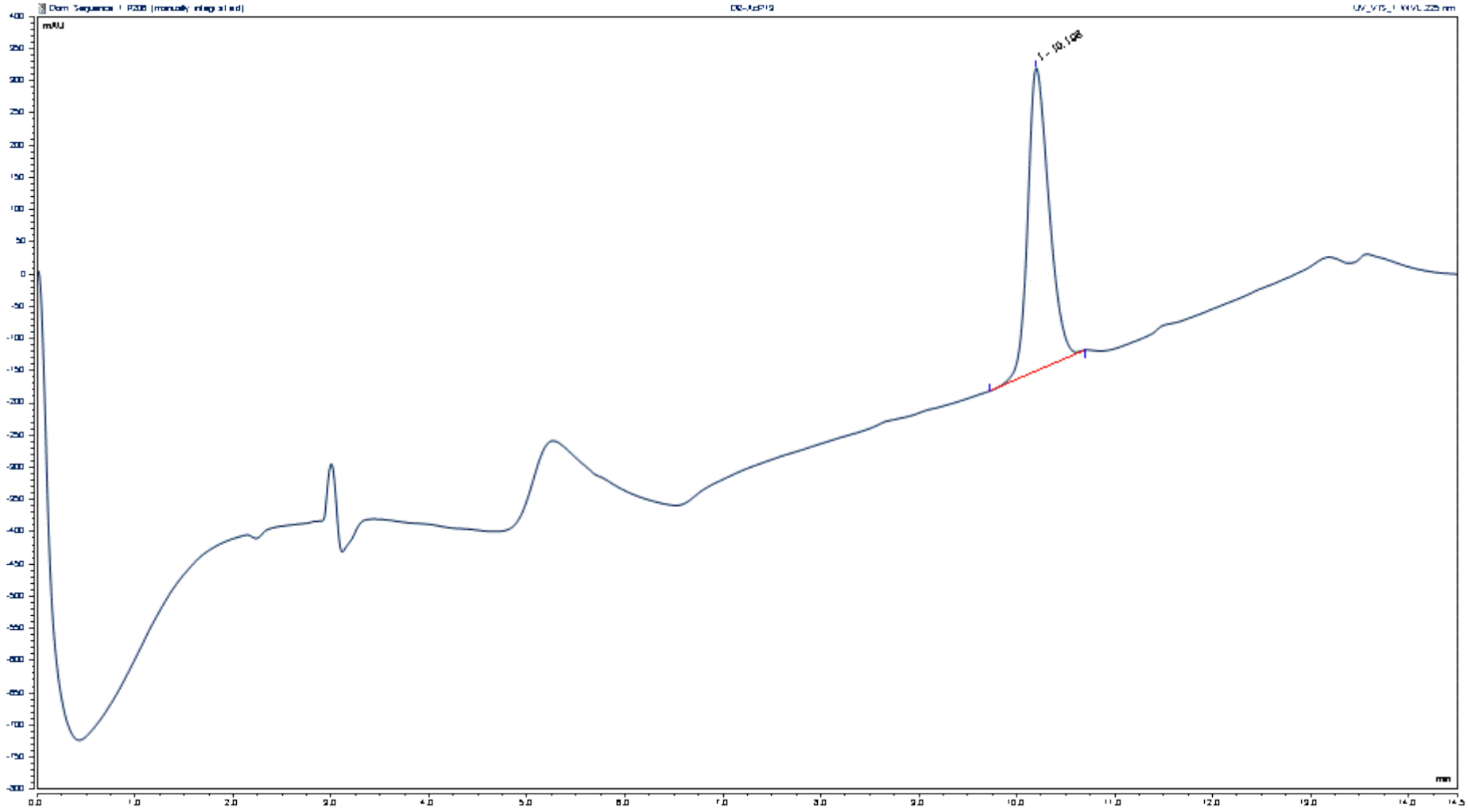
7.2.10 Synthesis of Ac-Pro_{13}-NH_2 (76):



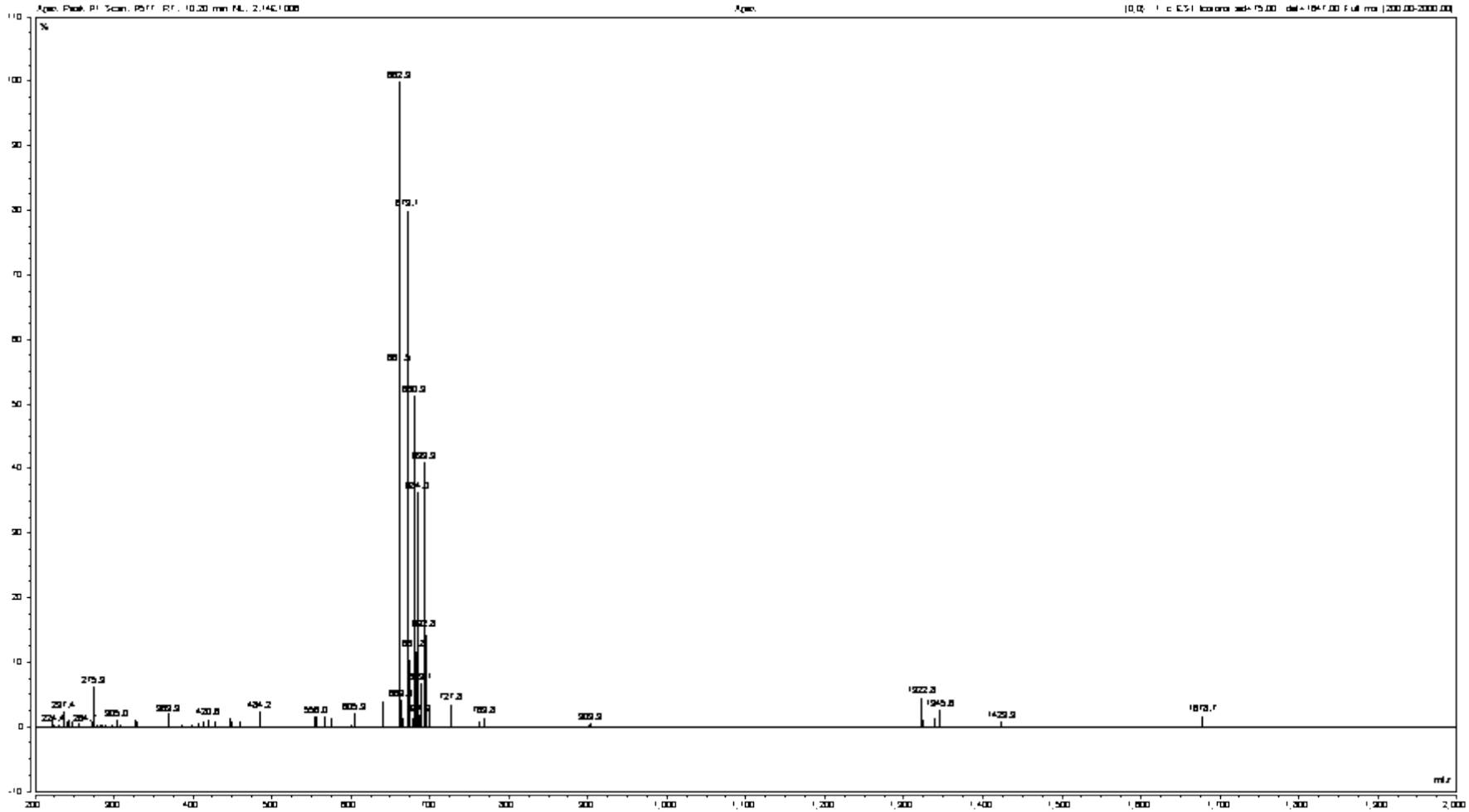
^{1}H NMR of **76** (400 MHz, MeOD)



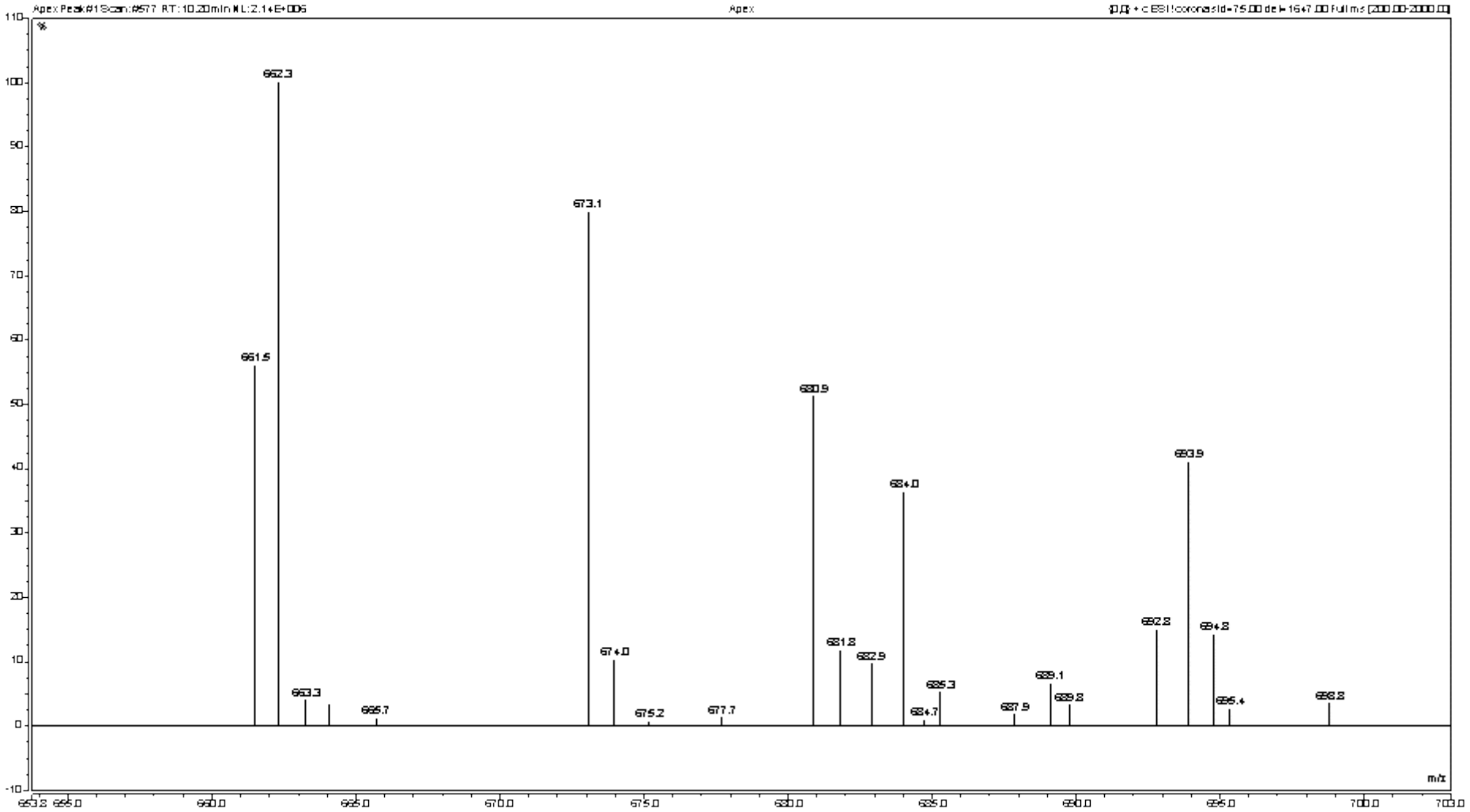
^{13}C NMR of **76** (101 MHz, MeOD)



HPLC UV-Vis Spectrum of peptide **76**, rt = 10.198 min, 225 nm

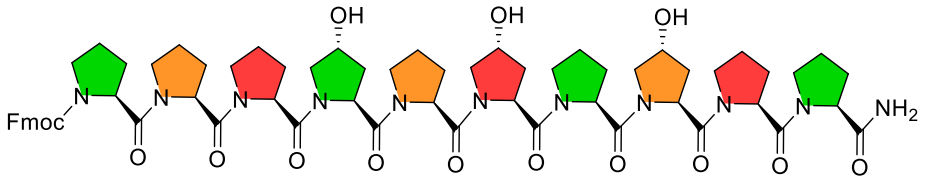


Mass spectrum (200-2000 m/z) of peak at 10.20 min for peptide **76**, m/z calcd for $[M+H]^+$: $C_{67}H_{97}N_{14}O_{14}^+$, 1321.7303; found; $[M+H]^+$ 1322.8; $[M+Na]^+$ 1345.6; $[M+2H]^{2+}$ 662.3

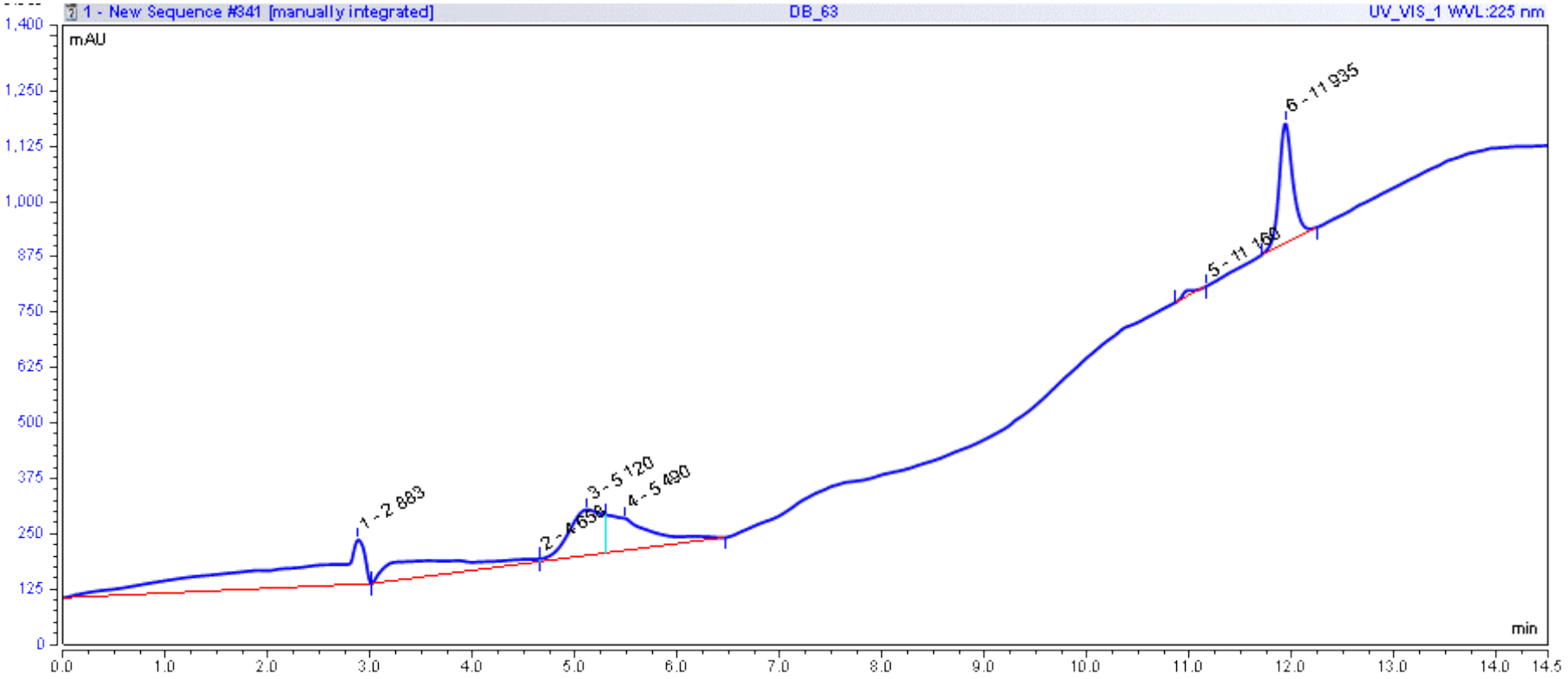


Mass spectrum (695-700 m/z) of peak at 10.20 min for peptide **76**, m/z calcd for $[M+H]^+$: $C_{67}H_{97}N_{14}O_{14}^+$, 1321.7303; found; $[M+H]^+$ 1322.8; $[M+2H]^{2+}$ 661.5; $[M+Na+H]^{2+}$ 673.1

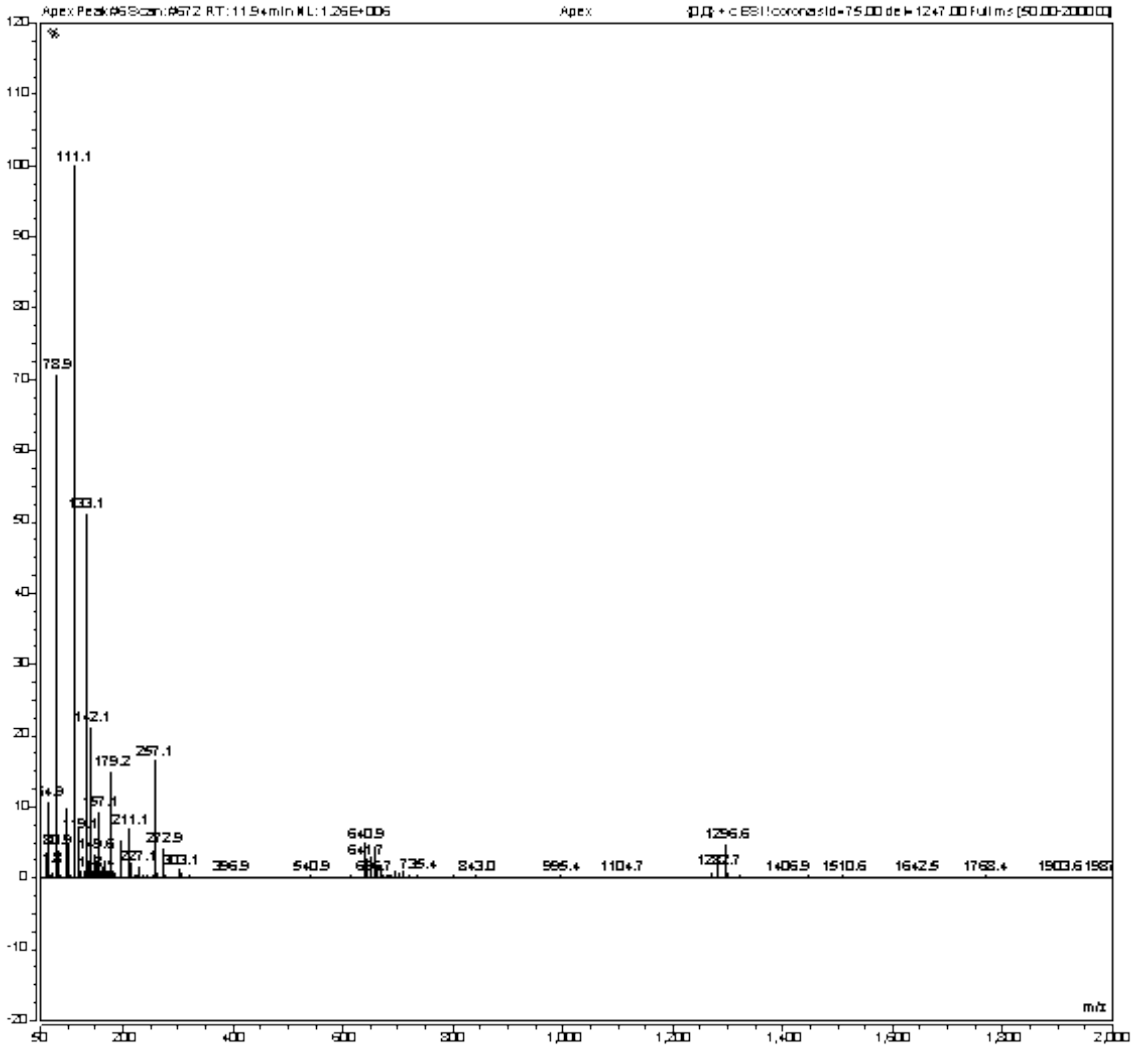
7.2.11 Synthesis of Fmoc-Pro_{10}-NH_2 $(-OH)_3$(i: 4, 6, 8) (77):



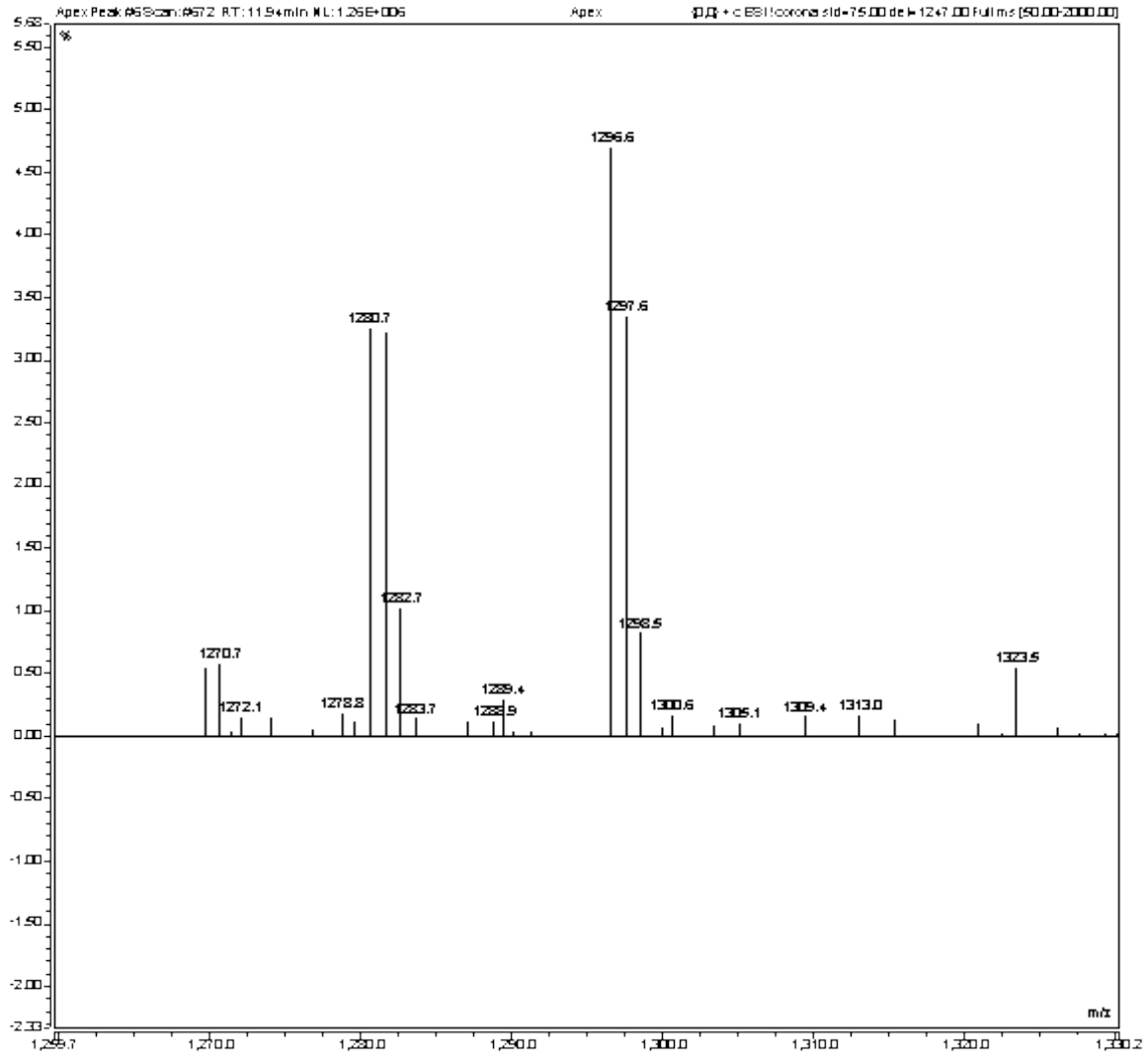
(77)



HPLC UV-Vis Spectrum of peptide **77**, rt = 11.935 min, 225 nm

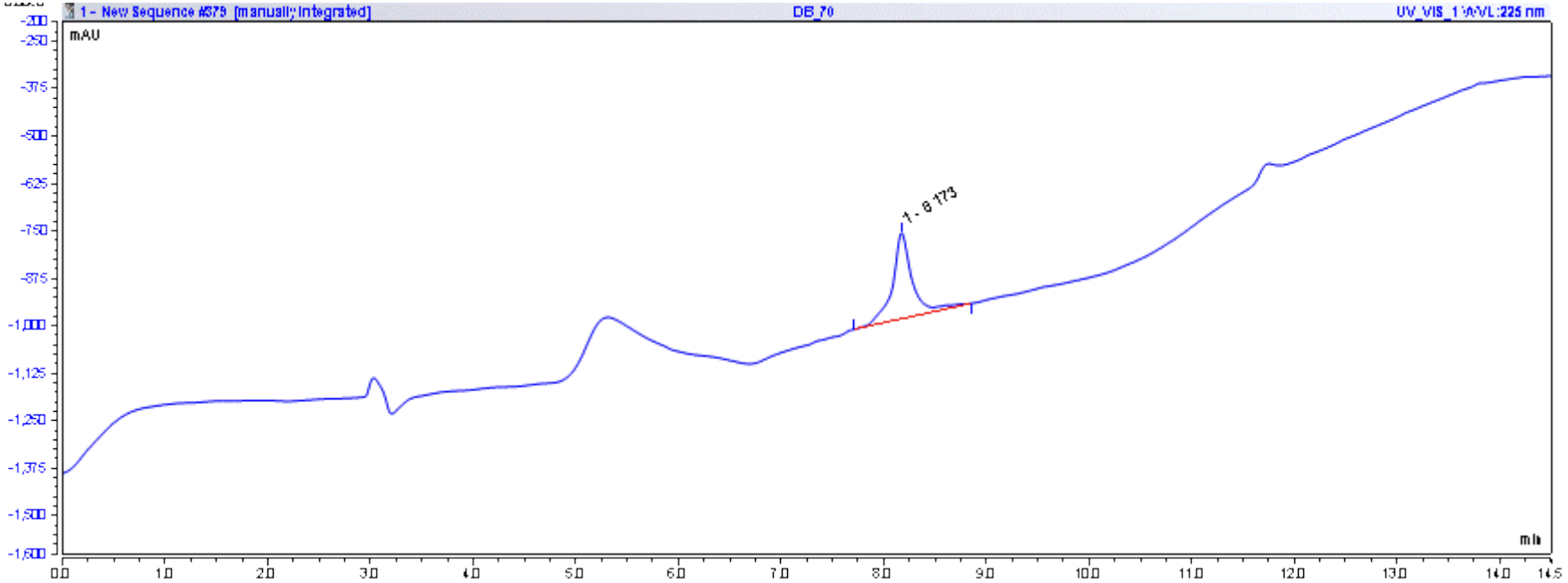


Mass spectrum (200-2000 m/z) of peak at 11.94 min for peptide **77**, m/z calcd for $[M+H]^+$: $C_{65}H_{84}N_{11}O_{15}^+$, 1258.6143; found; $[M+Na]^+$ 1280.7; $[M+H+Na]^{2+}$ 640.9

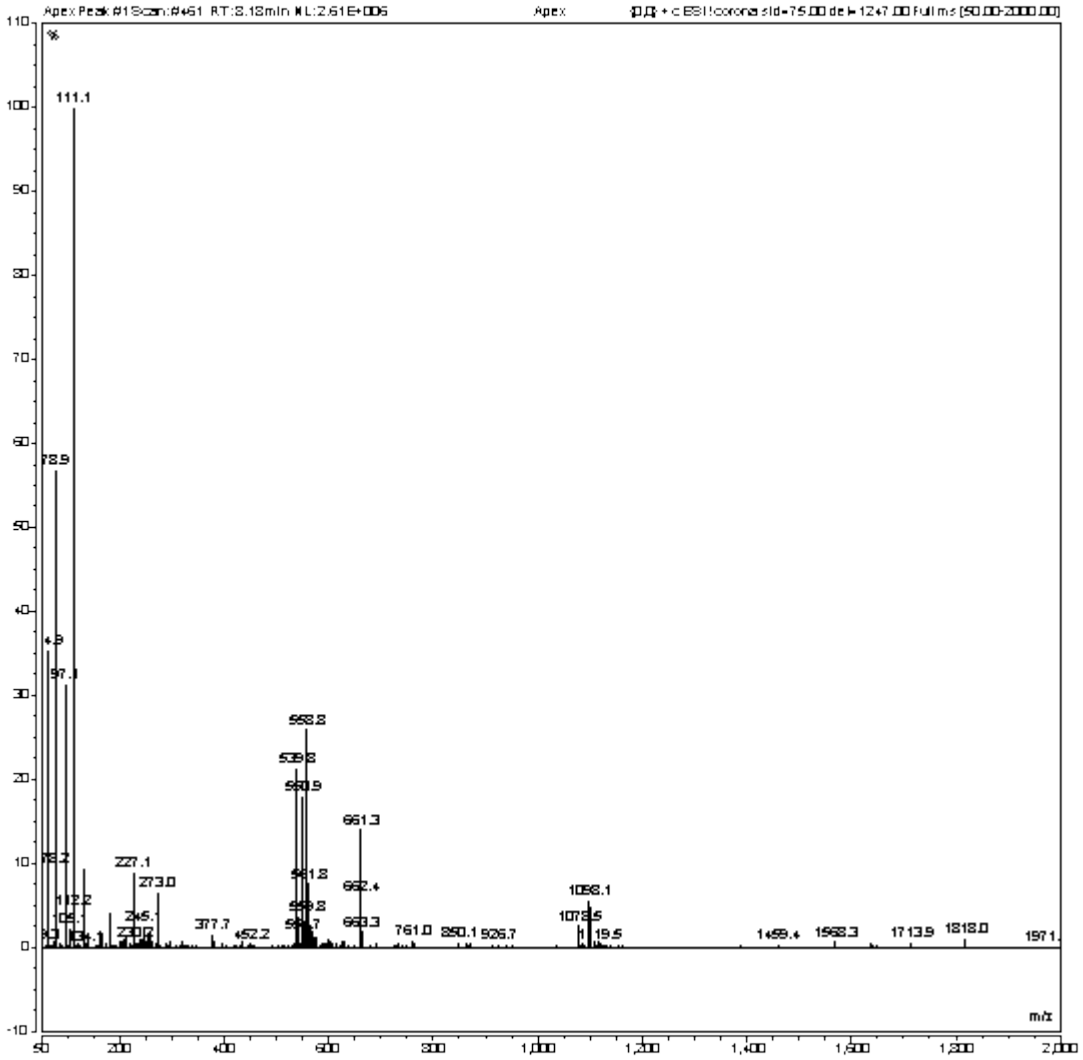


Mass spectrum (200-2000 m/z) of peak at 11.94 min for peptide **77**, m/z calcd for $[M+H]^+$: $C_{65}H_{84}N_{11}O_{15}^+$, 1258.6143; found; $[M+Na]^+$ 1280.7;

7.2.12 Synthesis of Ac-Pro_{10}-NH_2 (-OH)$_3$(i: 4, 6, 8) (78):

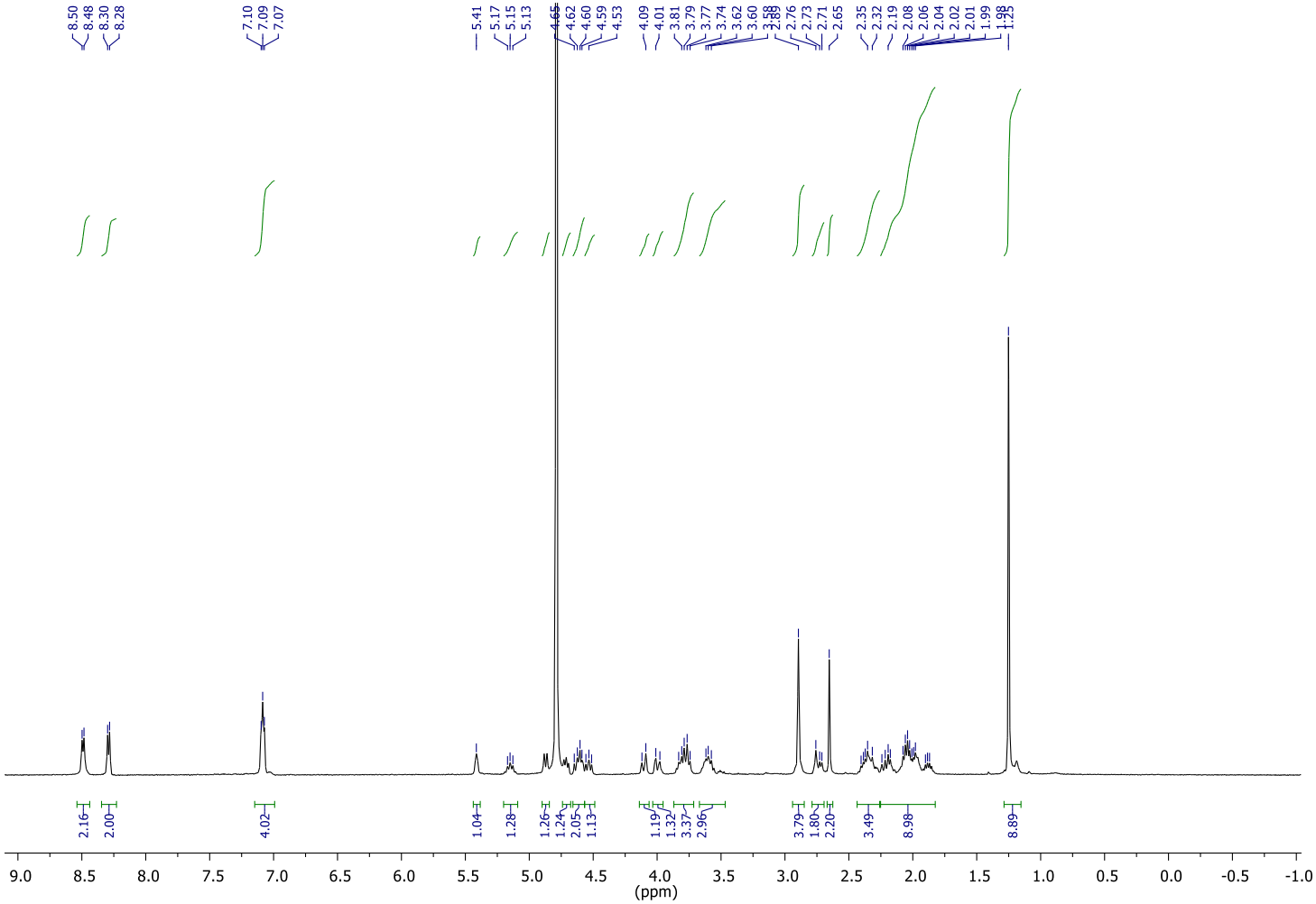


HPLC UV-Vis Spectrum of peptide **78**, rt = 8.173 min, 225 nm

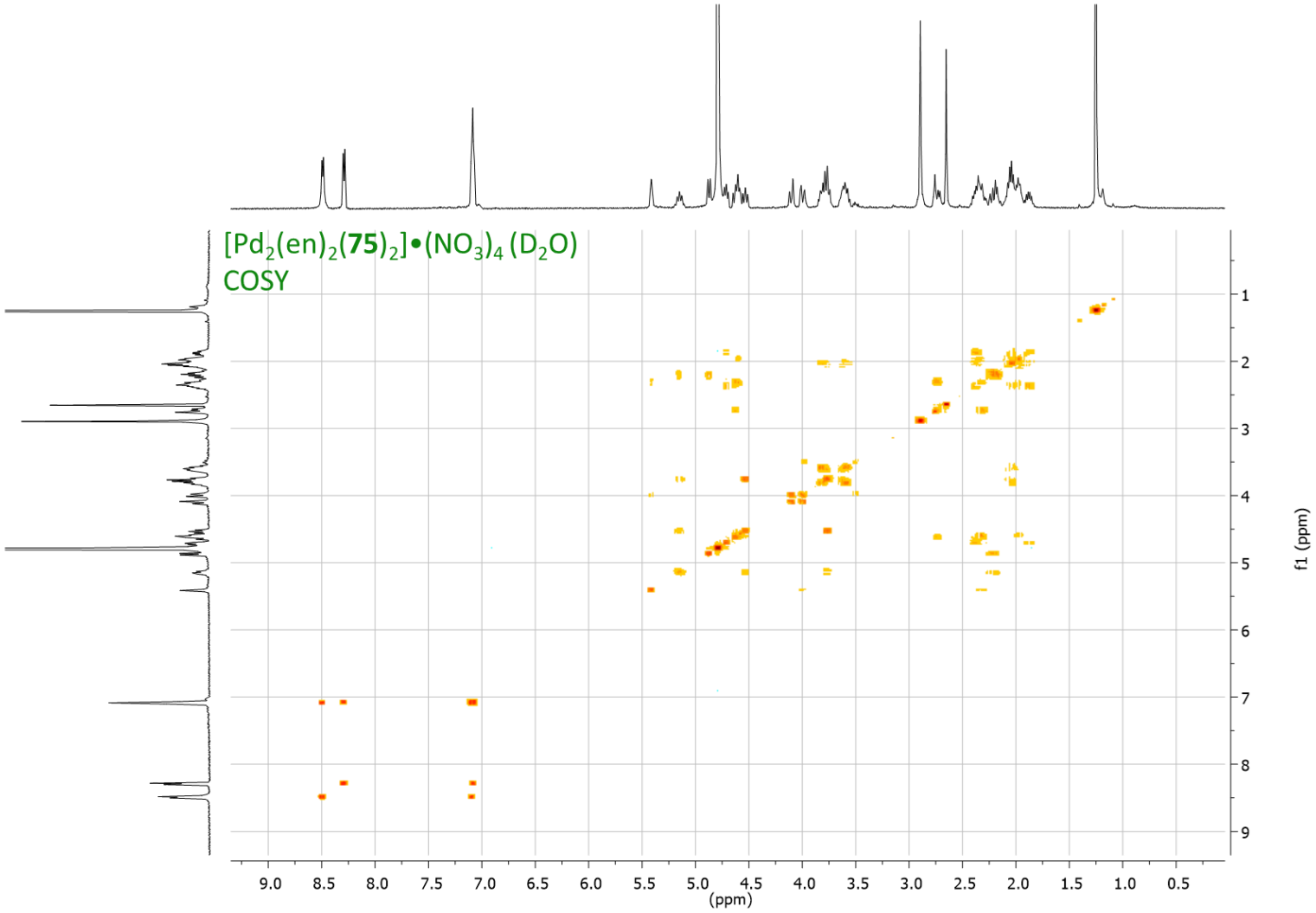


Mass spectrum (50-2000 m/z) of peak at 8.2 min for peptide **78**, m/z calcd for $[M+H]^+$: $C_{52}H_{76}N_{11}O_{14}^+$,1078.5568; found; $[M+H]^+$ 1078.5, $[M+2H]^{2+}$ 539.3

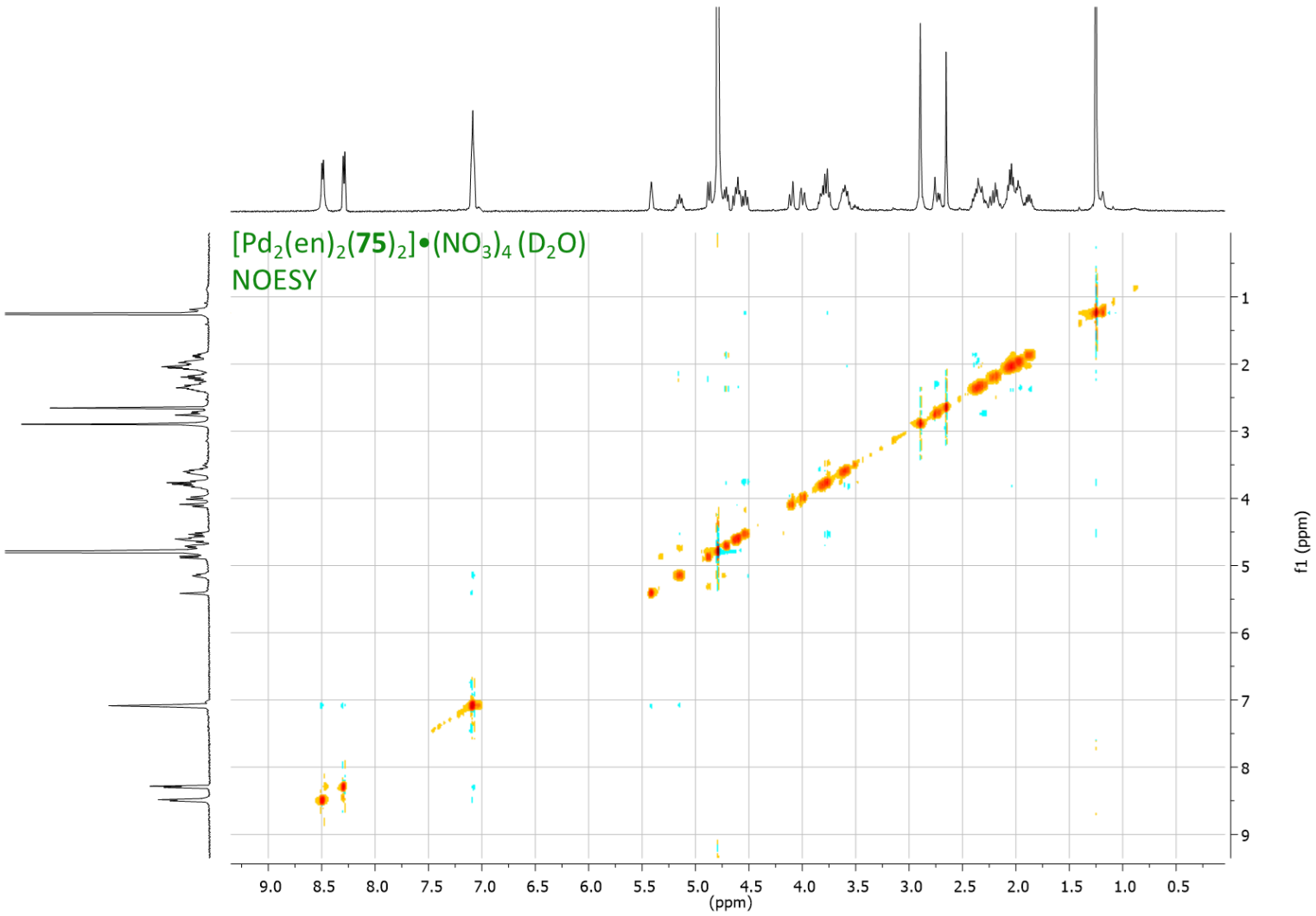
7.2.13 Synthesis of [Pd (en) (**75**)]•$(NO_3)_2$:



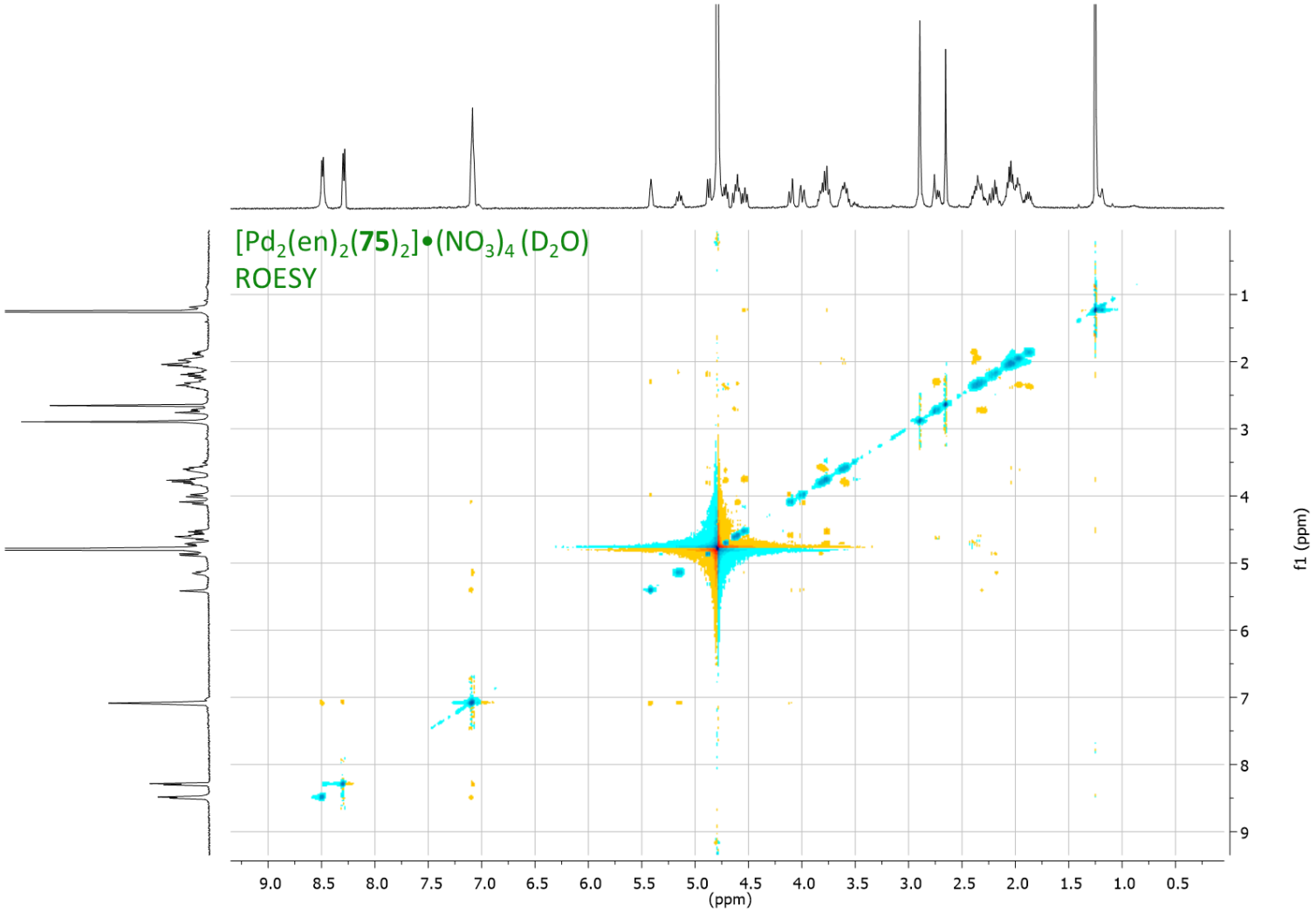
^{1}H NMR Spectrum of complex [Pd(en)(**75**)]•$(NO_3)_2$ (400 MHz, D_2O)



^{1}H-^{1}H COSY NMR Spectrum of complex [Pd(en)(**75**)]•$(NO_3)_2$ (400 MHz, D_2O)

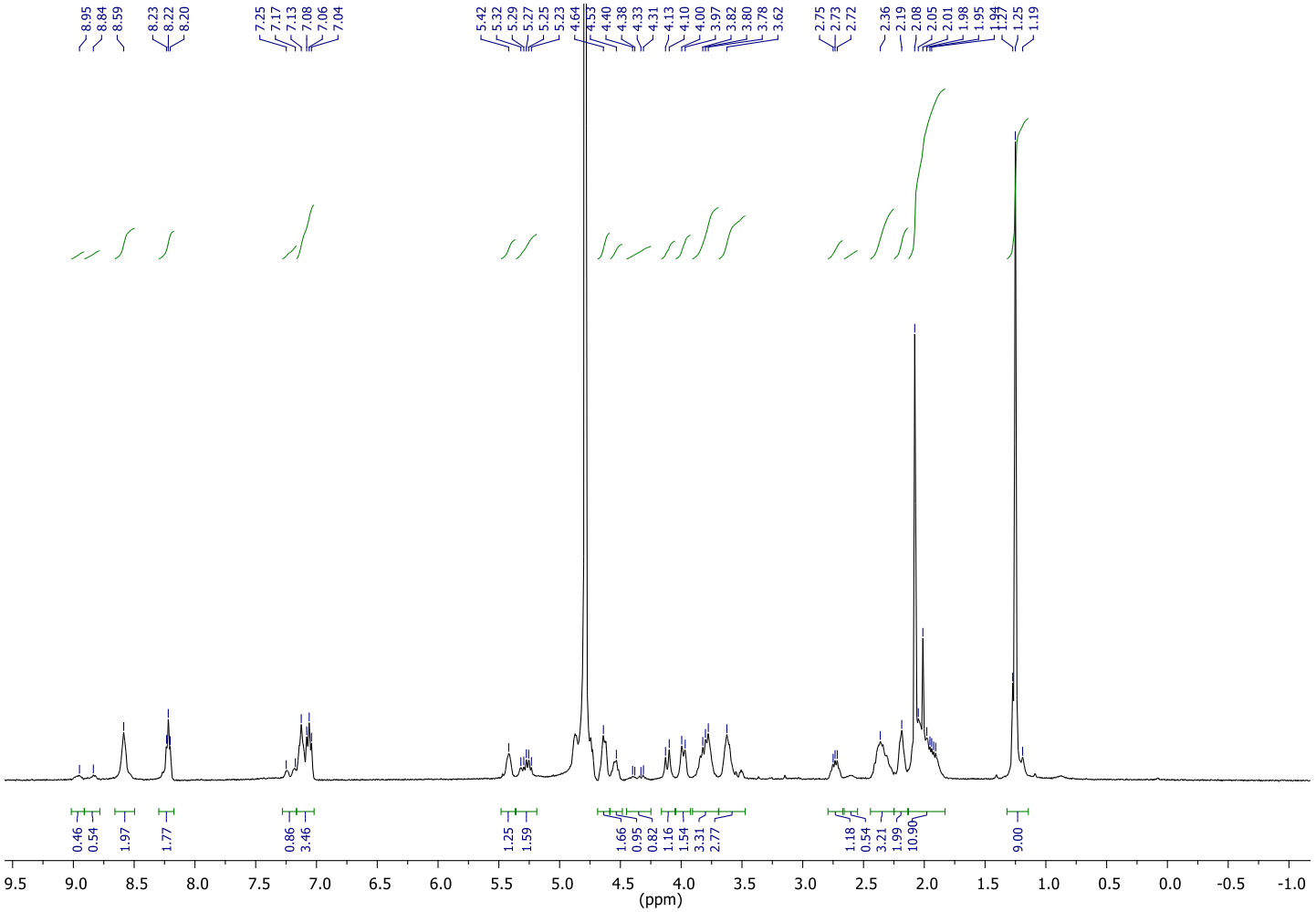


^{1}H-^{1}H NOESY NMR Spectrum of complex [Pd(en)(**75**)]•$(NO_3)_2$ (400 MHz, D_2O)

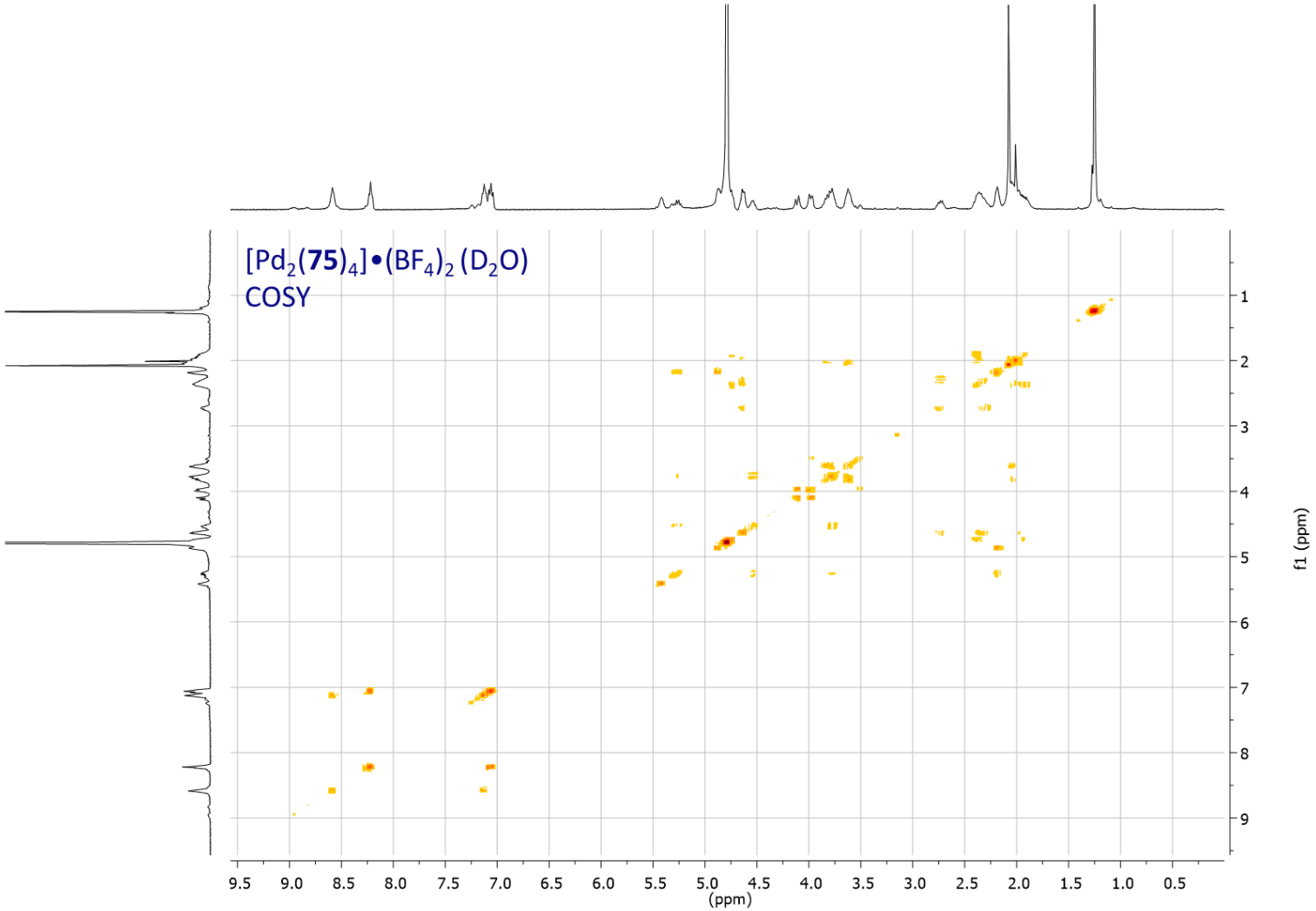


^{1}H-^{1}H ROESY NMR Spectrum of complex [Pd (en)(**75**)]•$(NO_3)_2$ (400 MHz, D_2O)

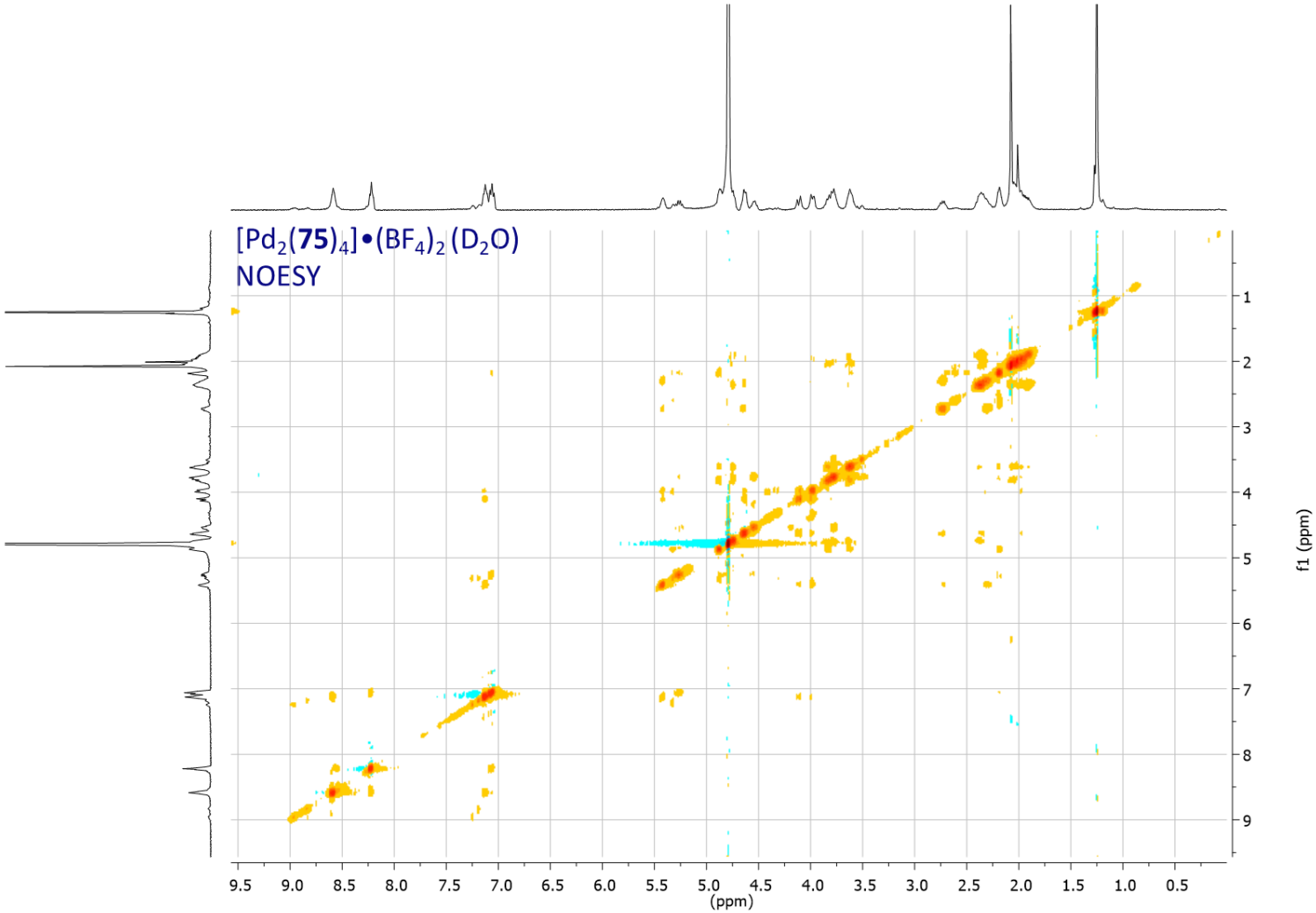
7.2.14 Synthesis of [Pd(**75**)$_2$]•(BF$_4$)$_2$:



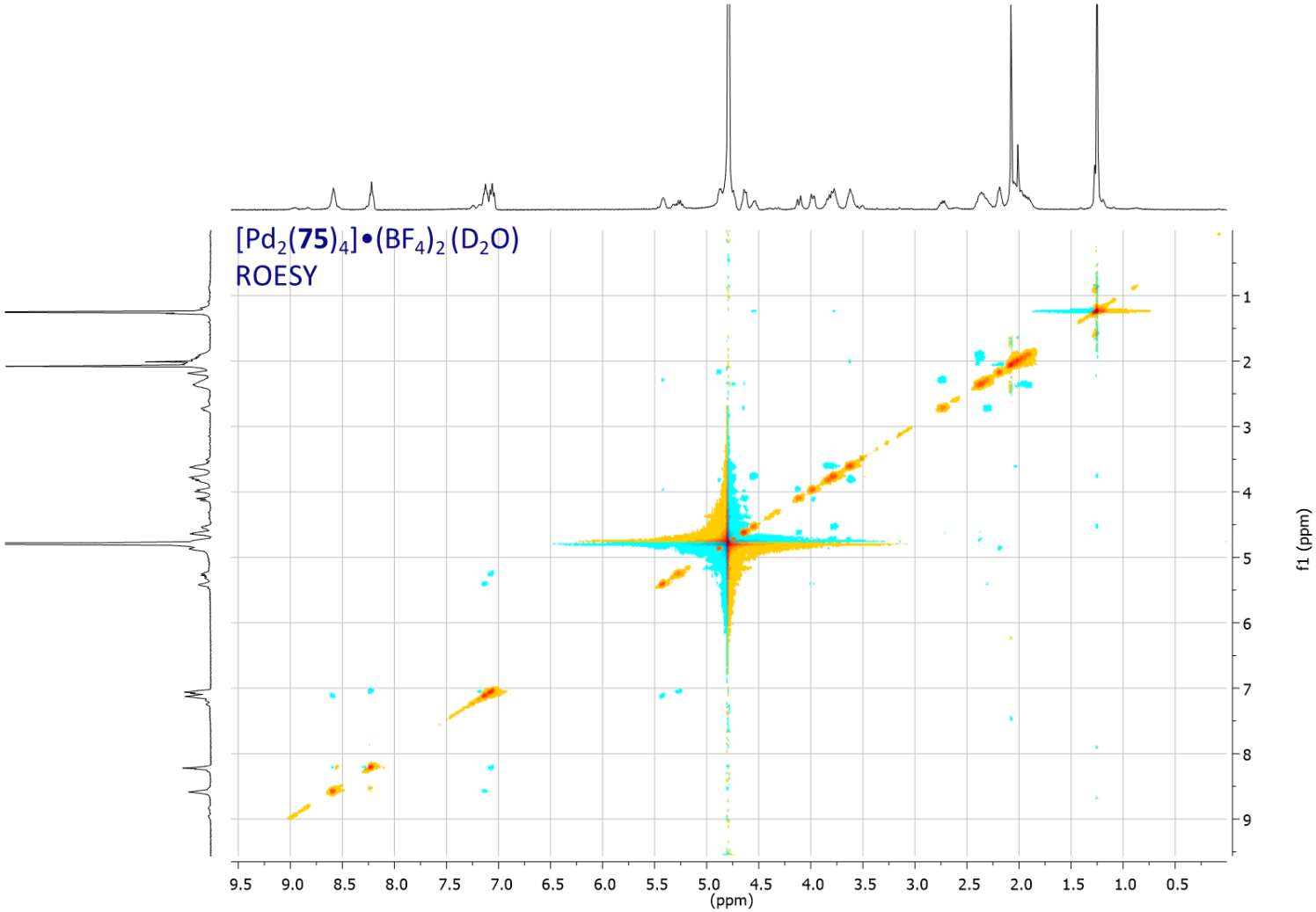
^{1}H NMR Spectrum of complex [Pd (**75**)$_2$]•(BF$_4$)$_2$ (400 MHz, D$_2$O)



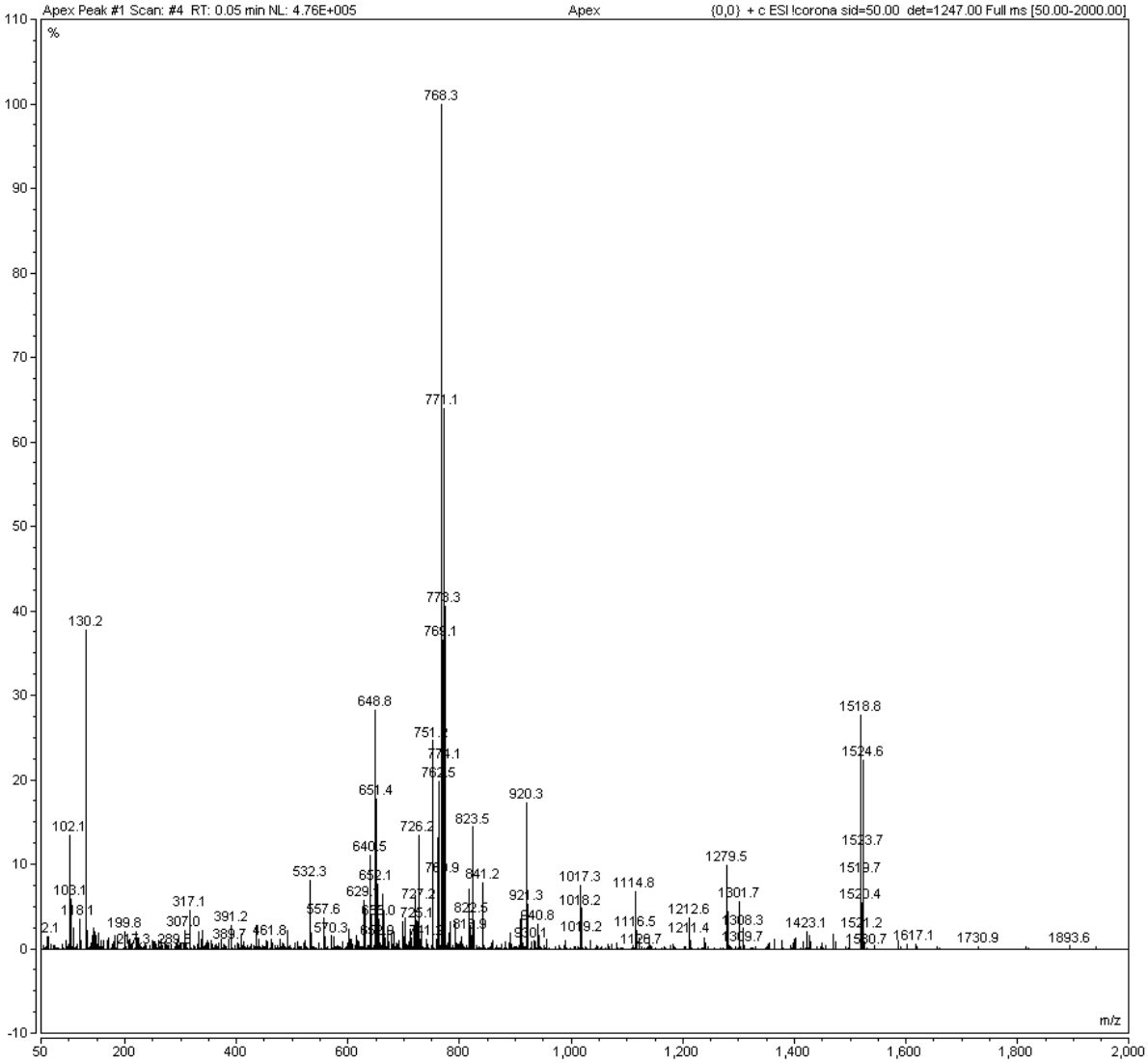
^{1}H-^{1}H COSY NMR Spectrum of complex [Pd (**75**)$_2$]•$(BF_4)_2$ (400 MHz, D_2O)



^{1}H-^{1}H NOESY NMR Spectrum of complex [Pd (**75**)$_2$]•$(BF_4)_2$ (400 MHz, D_2O)



^{1}H-^{1}H ROESY NMR Spectrum of complex [Pd (**75**)$_2$]•$(BF_4)_2$ (400 MHz, D_2O)



Mass spectrum (50-2000 m/z) of, **P_{13},** m/z calc. for $C_{80}H_{105}N_{14}O_{15}^{+}$: 1501.78.78 ,found: $[2M+H]^{2+}$ 751.2, $[2M+K]^{2+}$ 768.3, $[M+Na]^{+}$ 1524.6