

UNIVERSITY OF KWAZULU-NATAL

**DESIGN, SYNTHESIS AND SCREENING OF NOVEL
PCU-PEPTIDE/PEPTOID DERIVED HIV PROTEASE
INHIBITORS**

2011

MAYA MELLISA MAKATINI

DESIGN, SYNTHESIS AND SCREENING OF NOVEL PCU- PEPTIDE/PEPTOID DERIVED HIV PROTEASE INHIBITORS

MAYA MELLISA MAKATINI

2011

A thesis submitted to the School of Pharmacy and Pharmacology, Faculty of Health Science, University of KwaZulu-Natal, Westville, for the degree of Doctor of Philosophy.

This is a thesis in which the chapters are written as a set of discrete research papers, with an overall Introduction and a final Discussion. These research papers are not published as yet, but one has already been submitted for publication.

As the candidate's supervisor, I have approved this thesis for submission

Supervisor:

Signed:----- Name:----- Date:-----

Co-Supervisor:

Signed:----- Name:----- Date:-----

Co-Supervisor:

Signed:----- Name:----- Date:-----

ABSTRACT

The AIDS epidemic in Africa has reached dramatic proportions. Of the 42 million people infected with HIV worldwide, 30 million are in Africa. Current available therapies have begun to transform this fatal disease into a chronic condition but there are still major obstacles that have resulted in a great demand for new and better drugs. The aim of this study was to synthesize novel and effective HIV protease inhibitors.

This work describes the first account of pentacycloundecane (PCU)-peptide and peptoid based protease inhibitors. These inhibitors are proposed to bind the wild type C-South African HIV protease (C-SA) catalytic site via the norstatine or dihydroxyethelene type functional group of the PCU. The desired compounds were synthesized by the coupling of the peptides and peptoids to the PCU cage which resulted in a series of promising and structurally diverse HIV-1 protease inhibitors. The inhibitors were characterized by Nuclear Magnetic Resonance (NMR) and evaluated against the wild type C-SA enzyme for its ability to inhibit 50 % of the enzyme's activity (IC_{50}). Two of the compounds reported herein, inhibited the enzyme activity at concentrations less than 80 nM.

NMR investigations indicated that the activity was related to the chirality of the PCU moiety and its ability to induce conformations of the coupled peptide side chain. Employing the new Efficient Adiabatic Symmetrized Rotating Overhauser Effect Spectroscopy (EASY-ROESY) technique enabled us to obtain vital information about the 3D structure of these small linear peptides and peptoids in solution. This technique is the first example describing the successful through space correlations of such small peptides. Furthermore, docking and a combined quantum mechanics/molecular mechanics (QM/MM) molecular dynamics MD simulation at the AM1 semi empirical level mirrored the observed NMR results and the experimental IC_{50} activity profile of the considered inhibitors. The combination of these experimental and theoretical methods provided a powerful insight into the interaction mode of these cage peptide and peptoid inhibitors with the enzyme.

DECLARATION

I,, declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
 - a. Their words have been re-written but the general information attributed to them has been referenced
 - b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed

ACKNOWLEDGEMENTS

I wish to thank God for keeping me safe and for his guidance.

I also wish to express my sincere gratitude to my supervisors Dr T. Govender, Prof H.G. Kruger, Dr P. Govender and Dr G.E.M Maguire for their guidance, patience and support during the course of this research.

My thanks also go out to:

The GGKM group.

Mr D. Jagjivan for his assistance with recording NMR spectra.

My friends Sabelo Kubheka, Ningie Langa, Nomfundo Langa, Jabulani Makhathini, Simphiwe Buthelezi, Bongani Shandu and Oluseye Onajole.

My family for all the years of sacrifice, love and moral support.

My angel nephew, Sabelo Kubheka for brightening my life.

The University of KwaZulu-Natal, Durban, the National Research Foundation and DAAD for their financial assistance during the project.

TABLE OF CONTENTS

Abstract	iii
Declaration	iv
Acknowledgements	v
Table of Contents	vi
Chapter 1	1
Introduction.....	1
1.1 Background.....	1
1.2 The HIV virus	2
1.2.1 Parts of the HIV virus	3
1.2.2 The HIV-1 life cycle	3
1.3 HIV Inhibitors.....	4
1.3.1 Entry inhibitors	4
1.3.2 Nucleoside reverse transcriptase inhibitors and Non-nucleoside reverse transcriptase inhibitors	5
1.3.3 HIV Integrase inhibitors.....	7
1.3.4 HIV protease inhibitors.....	7
1.4 HIV protease	10
1.4.1 Domains in the HIV protease dimer	10
1.4.2 Dimer stability	11
1.4.3 Substrate binding	11
1.4.4 HIV protease mechanism	13
1.5 Polycyclic cage compounds	14

1.6	Techniques employed in structure based drug discovery.....	15
1.6.1	NMR studies	15
1.6.2	Computational Studies	16
1.7	Aims of the present study.....	17
1.8	References.....	18
Chapter 2	21
Pentacycloundecane-based inhibitors of wild type C-South African HIV-protease		21
Introduction.....		21
Materials and Methods.....		22
Results and Discussion		24
Conclusion		31
References.....		32
Chapter 3	34
Pentacyclo-undecane lactam peptides and peptoids as potential HIV-1 wild type C-SA protease inhibitors		34
Introduction.....		34
Results and Discussion		34
Conclusion		40
References.....		41
Chapter 4	42
Design, synthesis and screening of pentacycloundecane-diol inhibitors targeting the HIV protease enzyme		42
Introduction.....		42

Results and Discussion	45
Conclusion	52
References.....	53
Chapter 5	54
Synthesis and screening of pentacycloundecane-peptoids as potent HIV protease inhibitors.....	54
Introduction.....	54
Results and discussion	58
Conclusion	62
References.....	63
Chapter 6	64
Summary and conclusion	64
PCU-lactam inhibitors	64
PCU-diol inhibitors.....	65
Conclusion	68
Appendices	69
Appendix 1	70
Appendix 2.....	169
Appendix 3.....	233
Appendix 4.....	306

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Acquired immunodeficiency syndrome (AIDS) was the first major epidemic caused by a previously unknown pathogen to appear during the 20th century.¹ To date more than 42 million people are infected worldwide.² This disease is caused by the human immunodeficiency virus (HIV), a member of the *lentivirus* subfamily of retroviruses.³ At the onset of the epidemic in the early 1980s, no existing drug was known to be useful against AIDS.¹ The understanding of the structure and life cycle of the virus led to the unprecedented development of drugs targeting the retroviral enzymes—reverse transcriptase (RT), integrase (IN) and protease (PR).¹ In the mid 1990s the success of the combination therapy (highly active antiretroviral therapy (HAART)) which consists of a protease inhibitor (PI), a nucleoside reverse transcriptase inhibitor (NRTI) and a non-nucleoside reverse transcriptase inhibitor (NNRTI) dramatically changed the situation for HIV/AIDS patients.⁴⁻⁶

HIV protease inhibitors reduce the virus proliferation and the spread of viral infection in susceptible cells and this success made the HIV aspartic protease the prime target for AIDS therapies.^{7,8,9} In spite of recent progress in the development of the PIs for the treatment of AIDS, a number of problems persist. The problems with most synthetic anti-retroviral peptidomimetic PIs are their low membrane permeability,¹⁰ protease degradation to form toxic byproducts and loss of activity after enzyme mutation.^{9,11-13} As a consequence of long term use of antiretroviral therapy, a high prevalence of side effects such as lipodystrophy, dyslipidaemia, diarrhoea, insulin resistance, an increase in cardiovascular risks and birth defects has been recorded in patients.^{11,5,14} Numerous drug resistant strains have also been identified and as a result there is still a need for new antiretroviral drugs.^{15,16}

At present, we face a situation where HIV/AIDS has developed into one of the most devastating pandemics the world has ever seen as the young adults die, orphans and the elderly are left behind.^{17,2,18} The lack of a healthy work force has also severely disturbed the economy in the worst-affected regions.¹⁸

In this study we have designed HIV protease inhibitors utilizing different pentacycloundecane (PCU) cage moieties. The PCU moieties have been previously reported to enhance the biological activity,^{19,20} membrane permeability and retard biodegradation when incorporated into biological active molecules.²¹⁻²⁶ A number of antiviral active cage compounds has also been reported.^{27-29,22} Herein we report the

synthesis, biological activity, NMR, quantum mechanics/molecular mechanics (QM/MM) based molecular dynamics (MD) simulations and docking results of PCU derived HIV protease inhibitors.

1.2 THE HIV VIRUS

The HIV is an enveloped *lentivirus* (a member of the retrovirus family) that exists in two forms: HIV-1 and HIV-2.^{30,31} The HIV-1 is the predominant form worldwide while the HIV-2 type is concentrated in West Africa and is rarely found elsewhere.³⁰ The genes of the HIV virus are composed of two single RNA strands which are later transcribed to DNA by viral transcriptase as it enters the human cell.³² HIV infects primarily vital cells in the human immune system such as CD4⁺ T-helper cells, macrophages, and dendritic cells.^{33,34} HIV infection leads to low levels of CD4⁺ T cells through three main mechanisms: First, direct viral killing of infected cells; second, increased rates of apoptosis in infected cells;³⁵ and third, killing of infected CD4⁺ T cells by CD8 cytotoxic lymphocytes that recognize infected cells.³⁶ When CD4⁺ T cell numbers decline below a critical level, cell-mediated immunity is lost, and the body becomes progressively more susceptible to opportunistic infections.³⁷

1.2.1 PARTS OF THE HIV VIRUS

A schematic drawing of the mature HIV virion is shown below in **Error! Reference source not found.**

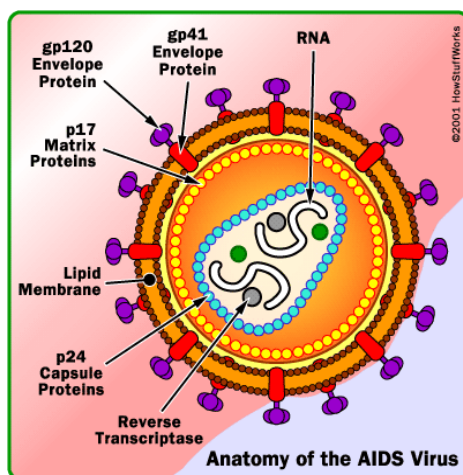


Figure 1: Schematic representation of the HIV virus.³⁸

- The viral envelope is the outer coat of the virus and it is composed of two layers of lipids with the proteins from the host cell embedded in the layers. There are also about 72 copies of the envelope proteins, which protrudes from the envelope surface. It consists of a cap made of three or four surface glycoproteins (gp120) and a stem consisting of three to four transmembrane protein (gp4) molecules.³⁹
- The p17 protein is the HIV matrix protein that lies between the envelope and the core. The viral core contains 2,000 copies of the viral protein (p24). These proteins surround two single strands of RNA, each containing a copy of the virus's nine genes. Three of these genes *gag*, *pol* and *env* contain information needed to make structural proteins for new virions.³⁹

1.2.2 THE HIV-1 LIFE CYCLE

The replication cycle is initiated by the binding of the virus to the cell surface, (A) Figure 2. The virus then fuses with the cell membrane (B) and empties its contents into the cell (C). The reverse transcriptase copies the viral genetic material from ribonucleic acid (RNA) into double-stranded deoxyribonucleic acid (ds-DNA) (D). This process is subject to a vast number of errors resulting in mutations for each viral generation produced. NRT and NNRT inhibitors have been developed to block the activity of the reverse transcriptase. As the viral DNA enters the nucleus of the cell (E), it gets integrated with the host's DNA (provirus) by the viral integrase (IN). Activity of this enzyme can be blocked by integrase inhibitors.

Upon activation of the infected host cell, the proviral DNA is transcribed into messenger RNA (mRNA) (F). mRNA is transported to the cytoplasm of the cell where it is translated into viral polyprotein precursors which includes the *env*, *gag* and *gag-pol* polyprotein precursors (G). The HIV protease, cleaves the new polyproteins (H), enabling them to join the RNA in the new viral particles that bud from the cell and infect others (I). Protease inhibitors can block the catalytic activity of the protease resulting in the production of non-infectious virions.⁴⁰

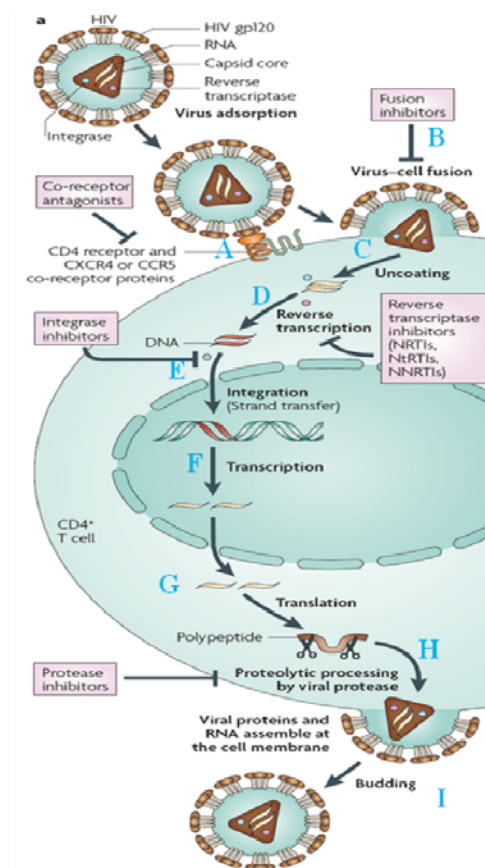


Figure 2. The replication cycle and targets for therapeutic intervention in the HIV life cycle.⁴¹

1.3 HIV INHIBITORS

Various inhibitors have proven to be effective in inhibiting the catalytic activity of the enzymes responsible for the replication of the HIV virus. This section is a brief description of these compounds.

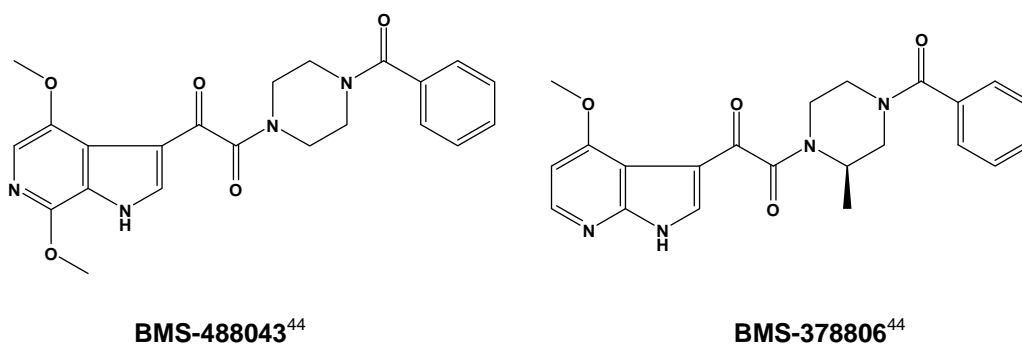
1.3.1 ENTRY INHIBITORS

Entry inhibitors block cell-to-cell transmission of HIV-1 and there are three broad classes of these inhibitors: attachment inhibitors, co-receptor antagonists, and fusion inhibitors.⁴²

Attachment inhibitors (AIs) represent a diverse class of compounds which disrupt the initial binding of the HIV-1 gp120 surface protein to the CD4 cell surface receptor by one of several possible mechanisms.⁴³ Examples of clinically effective AIs are BMS-488043 and BMS-378806 (Figure 3).⁴⁴

Co-receptor Antagonists (CRAs) block the HIV entry by preventing the virus from interacting with either of its two principal co-receptors: CCR5 and CXCR4.⁴³ The ability of co-receptor antagonists (CRAs) to inhibit HIV infection has been demonstrated in cell culture systems and animal models, and several compounds are advancing in clinical trials.⁴⁵ Lead candidates targeting the CCR5 co-receptor include the monoclonal antibody, Pro140 which blocks HIV-1 infection by occupying the gp120 binding site on CCR5.⁴⁴

Fusion Inhibitors (FIs) inhibit the final conformational change of gp41 required for membrane fusion and virus entry.⁴³ Enfuvirtide is the first and so far the only therapeutically used fusion inhibitor.⁴⁶ It is a linear 36-amino acid peptide (Figure 3) corresponding to residues 127-162 of the viral glycoprotein gp41.⁴¹



Ac-Tyr-Thr-Ser-Leu-Ile-His-Ser-Leu-Ile-Glu-Glu-Ser-Gln-Asn-Gln-Gln-Glu-Lys-Asn-Glu-Gln-Glu-Leu-Leu-Glu- Leu-Asp-Lys-Trp-Ala-Ser-Leu-Trp-Asn-Trp-Phe-NH₂⁴⁶

Enfuvirtide⁴¹

Figure 3. Structures of entry inhibitors.

1.3.2 NUCLEOSIDE REVERSE TRANSCRIPTASE INHIBITORS AND NON-NUCLEOSIDE REVERSE TRANSCRIPTASE INHIBITORS

NRTIs are analogues of the naturally occurring deoxynucleotides needed to synthesize the viral DNA and they compete with the natural deoxynucleotides for incorporation into the growing viral DNA chain. However, unlike the natural deoxynucleotides substrates, NRTIs lack a 3'-hydroxyl group on the

deoxyribose moiety. As a result, following incorporation of an NRTI the next incoming deoxynucleotide cannot form the subsequent 5'-3' phosphodiester bond needed to extend the DNA chain. Thus, when an NRTI is incorporated, viral DNA synthesis is halted, a process known as chain termination.^{47,48} All NRTIs are classified as competitive substrate inhibitors and there are currently eight drugs available on the market and a few examples are shown in Figure 4.

In contrast, NNRTIs have a completely different mode of action. NNRTIs block the reverse transcriptase by binding at a different site on the enzyme. They are not incorporated into the viral DNA but instead inhibit the movement of protein domains of reverse transcriptase that are needed to carry out the process of DNA synthesis. NNRTIs are therefore classified as non-competitive inhibitors of reverse transcriptase.^{49,50} The United State food and drug administration (FDA) approved NNRTIs are shown in Figure 5.

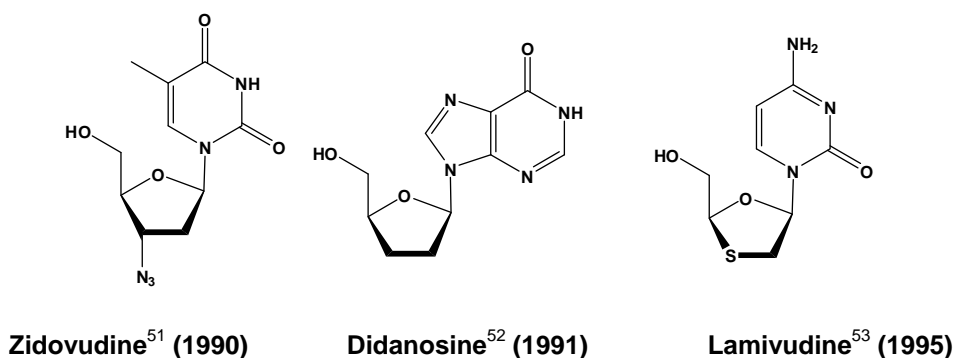


Figure 4. Nucleotide Reverse Transcriptase Inhibitors (NRTI) approved by the FDA.

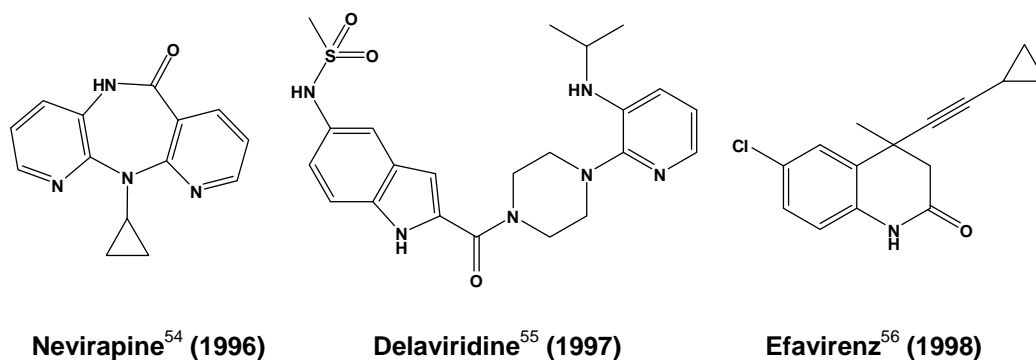


Figure 5. Non Nucleotide Reverse Transcription Inhibitors (NNRTI) approved by the FDA.

1.3.3 HIV INTEGRASE INHIBITORS

The HIV integrase enzyme is the last of the three essential viral enzymes to be targeted by antiretroviral drugs. Development of the HIV integrase inhibitor class of drugs has lagged behind because of the incomplete knowledge of the enzyme structure. To date all HIV integrase inhibitors are in the investigational testing phase and none have been approved by the FDA for treatment of HIV/AIDS.⁵⁷

1.3.4 HIV PROTEASE INHIBITORS

HIV protease inhibitors, inhibits the maturation of new virions by binding to the catalytic site of the protease and preventing the cleavage of *gag* and *gag-pol* into functional constituents. All commercially available (Figure 6, Figure 7) and some inhibitors that are currently being tested clinically have non-hydrolyzable transition state isosteres such as the hydroxyethylene, norstatine and many others moieties (Figure 8) as the basic core of the molecule. Basically, the core is a good isostere replacement at the scissile bond that is believed to mimic the tetrahedral transition state of the proteolytic reaction.^{8,1}

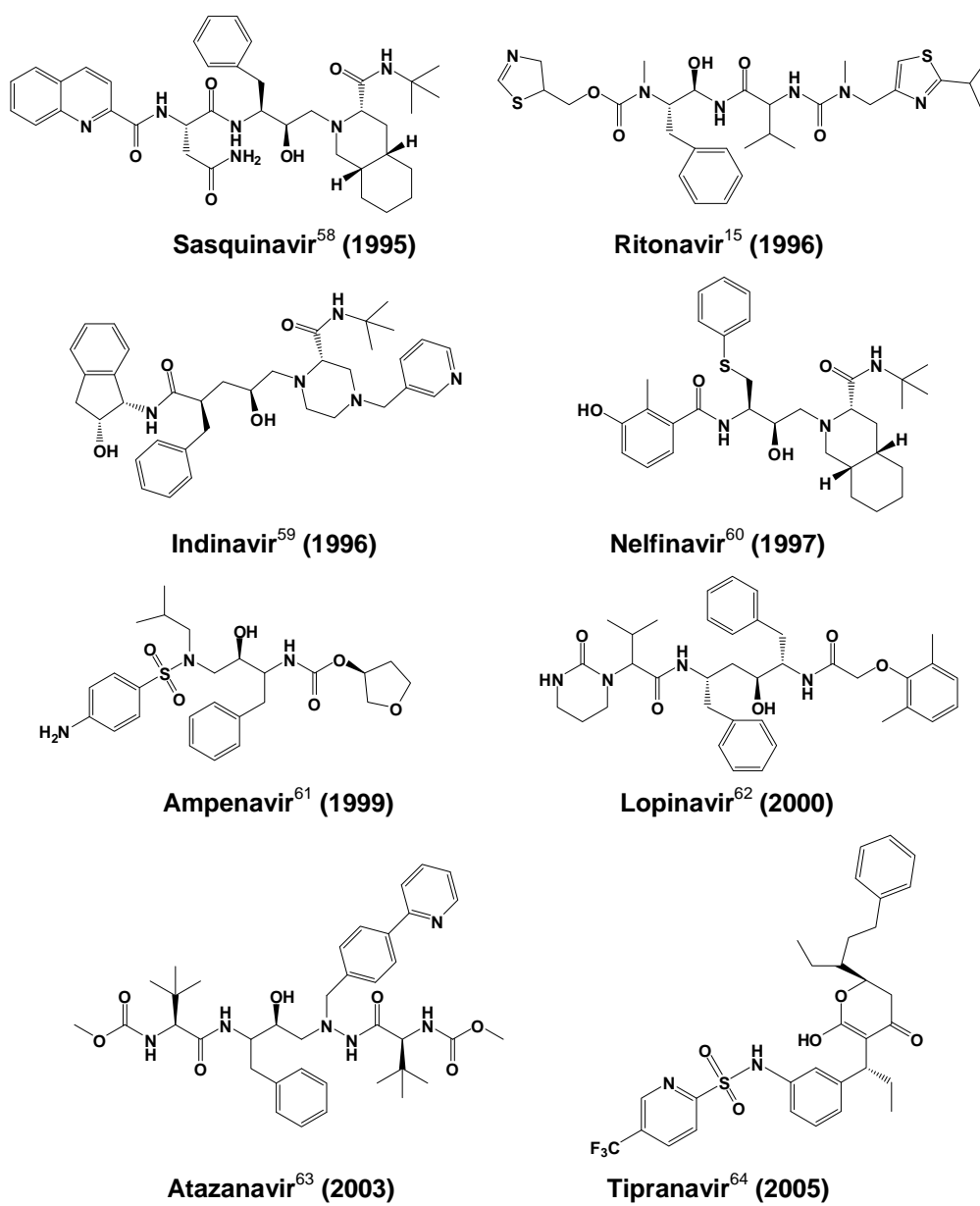
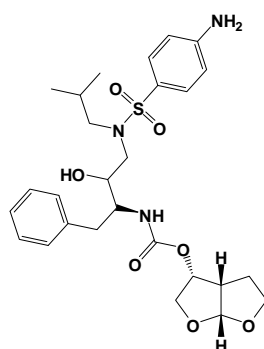


Figure 6. FDA approved HIV-1 protease inhibitors.



Darunavir⁶⁵ (2006)

Figure 7 continuation of Figure 6. FDA approved HIV-1 protease inhibitors.

Inhibitors derived from the non-cleavable transition state isostere (Figure 8) have been reported to bind to the active site of the enzyme with an affinity several orders of magnitude greater than the natural peptide substrate.⁶⁶

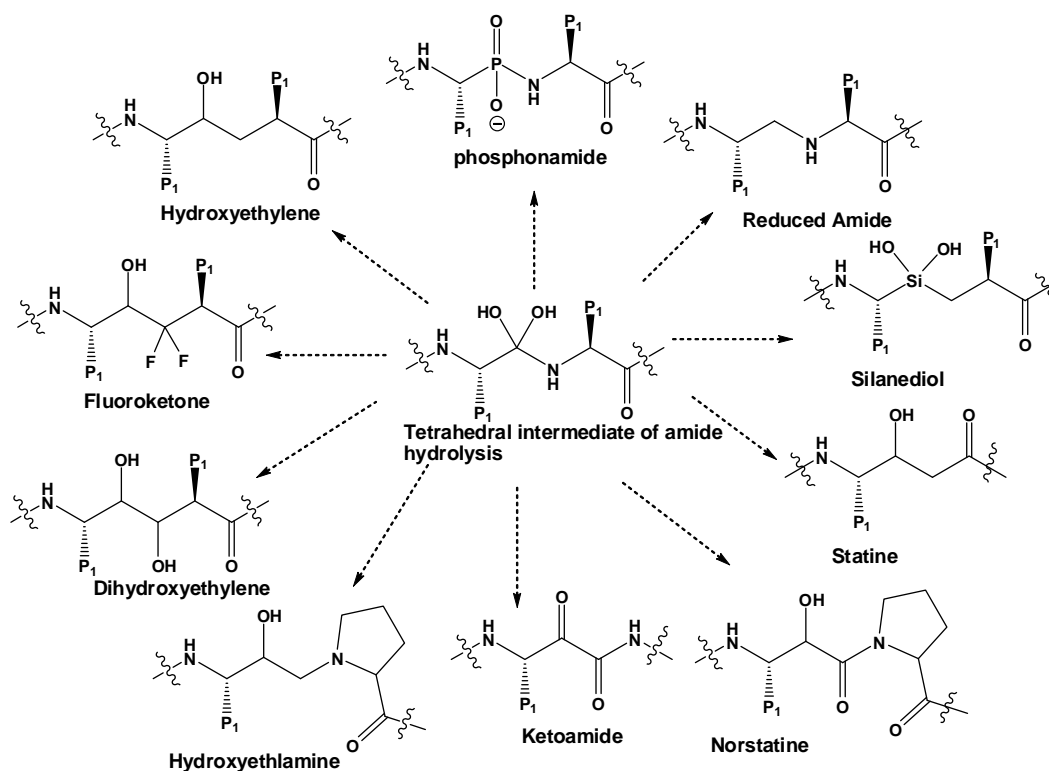


Figure 8. Non-cleavable transition state isostere developed for the synthesis of HIV PR inhibitors.⁸

Although there are nine anti-HIV PR inhibitors, and many other anti-reverse transcriptase drugs that are commercially available, their effectiveness has been hampered by the emergence of drug-resistant and cross-resistant mutants, rendering AIDS with no definitive cure.¹⁶

1.4 HIV PROTEASE

The HIV protease enzyme has been characterized biochemically and structurally as a member of the aspartyl protease family of enzymes.⁶⁷ It consists of two units, each composed of 99 amino acids, which join together to form the C_2 -symmetric active homo-dimer.⁶⁸ The C-termini of each dimer interdigitate to form a compact four-"stranded" interface. The interfacing four terminal strands align together like interlocking fingers (see Figure 9).⁶⁹

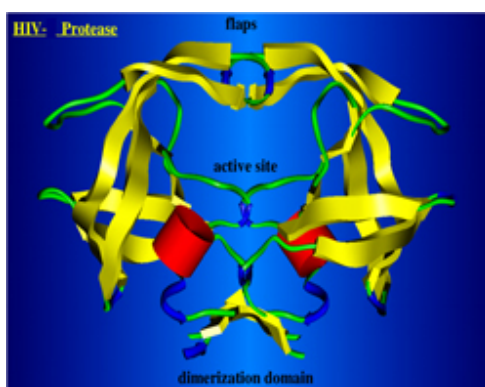


Figure 9. Structure of HIV protease.⁷⁰

On account of high error rates of the HIV reverse transcriptase, the nucleotide sequence of HIV protease changes with generations. However, residues in the catalytic triad (Asp-Thr-Gly) are well conserved both from the point of genetic variation and drug resistance.⁷¹

1.4.1 DOMAINS IN THE HIV PROTEASE DIMER

- a. Terminal domain or dimerization domain
 - b. The core domains
 - c. The flap domains
- a. Terminal or dimerization domain consists of the termini four-stranded beta-sheet residues 1-4, and 95-99 of each monomer, a turn encompassing residues 4-9, and the helix residues 86-94 of each monomer. This domain is quite crucial in dimer formation and stabilization of an active protease.^{46,47}
 - b. The core domains are primarily composed of four beta-strand structures. The domain sequence is made up of quite compact residues 10-32 and 63-85 from each monomer. The conserved Asp25-Thr26-Gly27 catalytic triad is situated at the interface of the core domains from the two monomers. This domain is quite useful in dimer stabilization, as well as the catalytic site stability. The interface between the core and terminal domains is composed primarily of small hydrophobic residues. The helix of the terminal domain packs against several beta strands of the core domain.^{46,47}

- c. The flap domains: This domain includes the most solvent exposed loop residues 33-43 preceding the beta hairpin containing the flaps residue 44-63. The flexible flap encloses the active site and it provides important ligand binding interactions.^{46,47}

1.4.2 DIMER STABILITY

The dimer is stabilized by many factors such as non-covalent interactions, hydrophobic packing of side chains and interactions involving the catalytic residues. The catalytic residues triad Asp25–Thr26–Gly27 form a network of hydrogen bonds referred to as a “fireman’s grip”, Figure 10.^{46,48}

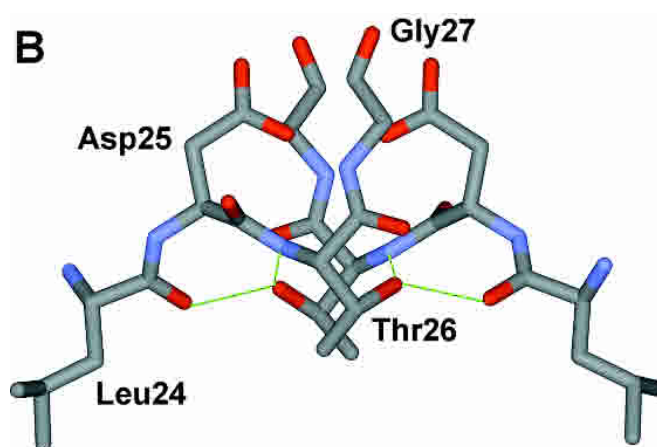


Figure 10. The catalytic residues triad Asp25–Thr26–Gly27 forming a network of hydrogen bonds.⁷²

1.4.3 SUBSTRATE BINDING

The pocket forming the active site is composed of a number of well defined sub sites which accommodate the side chains of the substrates. These distinct sub sites are named as S_1 to S_n and the corresponding side chain of the substrate fitting into these pockets are termed P_1 to P_n .

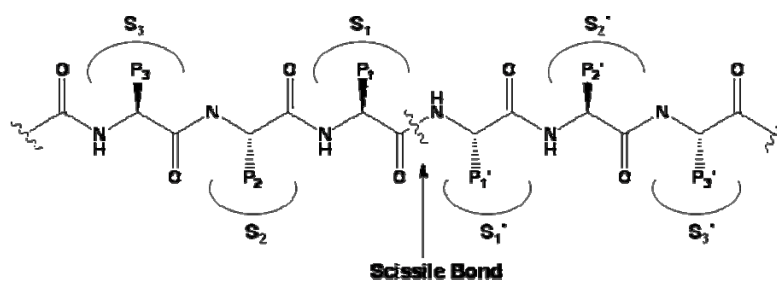


Figure 11. Standard nomenclature P_1 ___ P_n , P_1' ___ P_n' is used to designate amino acid residues of peptide substrates. The corresponding binding sites on the protease are referred to as S_1 ___ S_n , S_1' ___ S_n' subsites.⁶⁸

Since the protease is C_2 symmetric the corresponding sub sites binding the P_1' to P_n' terminal are termed as S_1' to S_n' .^{68,73}

1.4.4 HIV PROTEASE MECHANISM

The HIV protease catalyzes the hydrolysis of peptide bonds with high sequence selectivity (Figure 12, Table 1) and catalytic proficiency.

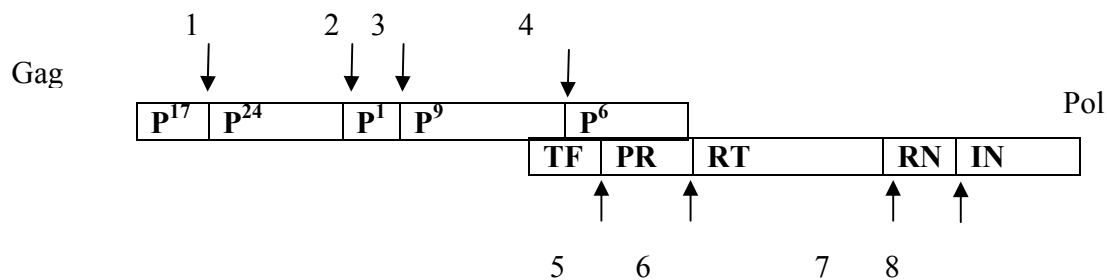


Figure 12. Cleavage sites within the gag and gag-pol precursor proteins.⁷³

Table 1. Amino acid sequences of the sites cleaved by HIV protease.⁷³

Site	Sequence
1	-Ser-Gln-Asn-Tyr*Pro-Ile-Val-Gln-
2	-Ala-Arg-Val-Leu*Ala-Glu-Ala-Met-
3	-Ala-Thr-Ile-Met*Met-Gln-Arg-Gly-
4	-Pro-Gly-Asn-Phe*Leu-Gln-Ser-Arg-
5	-Ser-Phe-Asn-Phe*Pro-Gln-Ile-Thr-
6	-Thr-Leu-Asn-Phe*Pro-Ile-Ser-Pro-
7	-Ala-Glu-Thr-Tyr*Phe-Val-Asp-Gly-
8	-Arg-Lys-Ile-Leu*Phe-Leu-Asp-Gly-

The asterisk* indicates to the position of the scissile bond

The HIV protease accomplishes the catalysis by using an activated water molecule to attack the amide bond carbonyl of the substrate's scissile bond. The activation of the water molecule is achieved by the two aspartyl β -carboxy groups at the active site. The nucleophilic water molecule held between the catalytic aspartates, after its activation by the negative aspartate side chain, attacks the carbonyl group in the substrate scissile bond to generate an oxyanion tetrahedral intermediate (Figure 13). Protonation of

the scissile amide N atom and rearrangement results in the breakdown of the tetrahedral intermediate to the hydrolysed products.^{8,73}

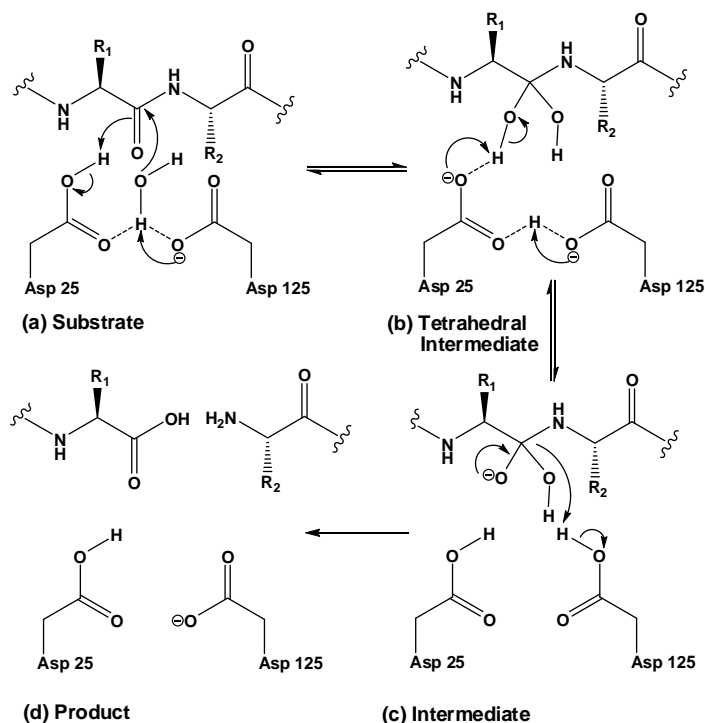


Figure 13. Cleavage mechanism of HIV-1 protease.⁸

The knowledge accumulated through the last two decades about the structure and mechanism of HIV PR has paved the way towards the development of new effective drugs for this enzyme through a structure based activity approach.⁸

1.5 POLYCYCLIC CAGE COMPOUNDS

Polycyclic cage compounds are useful scaffolds that yield drugs with a wide scope of applications.⁷⁴ These moieties can be used to modify and improve the pharmacokinetic and pharmacodynamic properties of current drugs.^{22,79,80} Interest in the pharmacology of polycyclic cage compounds was stimulated by early findings that the adamantane amine and 1-amino-adamantane or amantadine, exhibited broad range antiviral activity. These viruses include the influenza, hepatitis C and herpes zoster neuralgia.^{79,81,82} Further studies into related polycyclic cage compounds such as pentacycloundecane, homocubane, and trishomocubane derivatives have led to the discovery of more pharmaceutical applications for cage compounds.²² Derivatives of trishomocubane exhibit promising *in vivo* activity against Herpes simplex II and the Influenza A2/Taiwan virus.^{22,23,24,27} The activity of these derivatives is comparable to that of

acyclovir and amantadine.²² Cubane and homocubane derivatives have been reported by Hasegana *et al.* to have anti-ulcer activity with cycloprotective effect and histamine (H-2) receptor antagonism activity.²² Pentacycloundecylamines have been observed to influence the profile of KCl-induced membrane depolarization.⁷⁵

Furthermore, the highly rigid structure of cage molecules induces receptor site specificity for lipophilic regions on the receptor molecule.⁷⁶ Their hydrophobic cage skeleton improved transport of drugs across the cell membrane thus increasing their affinity for lipophilic regions in receptor molecules.^{19,20,22} and thus enabling the drug to cross the blood-brain barrier of the central nervous system (CNS).^{24,82}

1.6 TECHNIQUES EMPLOYED IN STRUCTURE BASED DRUG DISCOVERY

There are three common techniques that can be used to relate structure to activity (1) X-ray; (2) NMR; and (3) computational methods as depicted in Figure 14.⁷⁷

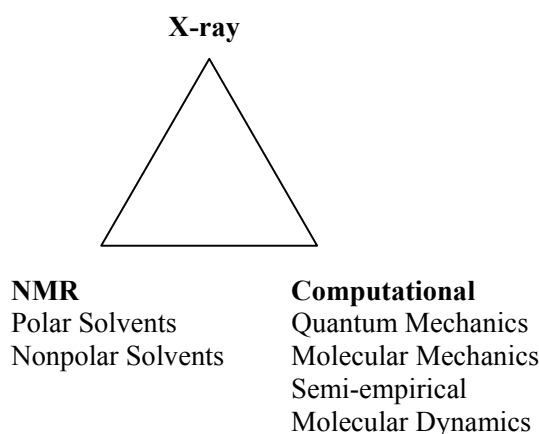


Figure 14. Technique employed to assess conformational details.⁷⁷

In this section we briefly describe the NMR and computational approaches since they enabled us to determine conformations of the synthesized inhibitors. X-ray diffraction could not be used because it requires suitable crystals which we were not able to obtain.

1.6.1 NMR STUDIES

The most important information needed for a 3-dimensional (3D) structure determination can be obtained from a Nuclear Overhauser Effect (NOE) spectroscopy (NOESY). The NOESY shows cross peaks between hydrogen atoms and these peaks are observed if the two protons are separated by a distance

shorter than 5.0 Å. Since the NOE depends on the through space correlation the two interacting protons can be as far apart as 100 residues or more.^{78,79}

Determination of 3D structures of small to medium sized organic and biomolecules are limited by the well known phenomenon that little or no NOE is observed for medium-sized compounds (MW = 1000 gmol⁻¹).⁸⁰ This problem can be solved by using rotating-frame Nuclear Overhauser effect spectroscopy (ROESY).⁸¹ However, several experimental problems namely direct cross-peaks due to J coupling, Hartmann-Hahn matching leading to, total correlation spectroscopy (TOCSY), cross-peaks and offset dependence have to be avoided. Recent design of an EASY ROESY in the year 2009 has made it possible to obtain ROESY spectra of medium sized peptide inhibitors without these artefacts. This enables the extraction of distance information in peptide inhibitors. These distances are then used to obtain valuable information about the 3D structure in solution phase.⁸⁰

1.6.2 COMPUTATIONAL STUDIES

The 3D structure of organic molecules can be determined by the combined QM/MM/MD approach. The fundamental idea behind the use of hybrid potentials is that a system is partitioned into several regions which are then modeled with different levels of approximation. The atoms from the residues or molecules that are participating in the reaction process, or atoms that are rich in electrons, such as the reacting groups of the enzyme, are treated quantum mechanically. The rest of the system is described by standard molecular mechanics force fields. Finally a boundary region is defined to account for the finite size of the system.^{82,83} In the combined method the QM part corresponds to what is to be studied in detail such as a molecule or several molecules, a fragment of a large molecule. Atoms in this part are explicitly expressed as electrons and nuclei. The MM part is the "environment" to the QM part. It is "non reactive" meaning there is no charge transfer or other chemical exchange between the QM and MM part.⁸⁴ In our case the environment was either water as a solvent, or the protease enzyme inside a cubic box of water. From the QM/MM calculations, information about position of and forces on the atoms can be obtained. This information is needed for the MD simulations. MD is a form of computer simulation in which atoms and molecules are allowed to interact for a period of time by approximations of known physics, giving a view of the motion of the particles due to a simulated temperature.⁸⁴ The hybrid calculations (QM/MM/MD) then afford low energy conformation structures of the inhibitors in solution.⁸⁵⁻⁸⁷

To obtain good starting structures for the inhibitors inside the protease enzyme, docking calculations can also be performed.⁸⁸ Molecular docking gives an approximation of the host-guest interaction. It is a method which predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex.⁸⁹ To perform a docking screen, the first requirement is a structure of the protein

of interest. Usually the structure has been determined using a biophysical technique such as x-ray crystallography, or less often, NMR spectroscopy. This protein structure and a database of potential ligands serve as inputs to a docking program. The success of a docking program depends on two components: the search algorithm (The search space in theory consists of all possible orientations and conformations of the protein paired with the ligand) and the scoring function (The scoring function takes a pose as input and returns a number indicating the likelihood that the pose represents a favorable binding interaction).⁸³ To obtain a better understanding of the interaction of the inhibitor inside the active pocket of the docked structure it can be subjected to a QM/MM/MD calculations.

1.7 AIMS OF THE PRESENT STUDY

This study is part of a research project aimed at the discovery of novel cage HIV-1 protease inhibitors, with the ultimate objective of developing new non hydrolysable transition state isosteres to be used in the simple synthesis of potent and bioavailable inhibitors. The specific objectives of this study have been:

- To design and synthesize PCU-peptide and peptoid derived HIV protease inhibitors.
- To determine the concentrations of the synthesized inhibitors that are needed to inhibit 50 % of the protease catalytic activity.
- To explore the PIs structural conformations *via* NMR and QM/MM/MD simulations.
- To understand the interactive mode of the inhibitor with the enzyme using docking.

1.8 REFERENCES

1. Wlodawer, A.; Vondrasek, J. *Annual Review of Biophysics and Biomolecular Structure* **1998**, *27*, 249-284.
2. Velaquez-Campoy, A.; Vega, S.; Fleming, E.; Bacha, U.; Sayed, Y.; Dirr, H. W.; Freire, E. *Aids Reviews* **2003**, *5*, 165-171.
3. Vacca, J. P.; Condra, J. H. *Drug Discovery Today* **1997**, *2*, 261-272.
4. Rodriguez-Barrios, F.; Gago, F. *Current Topics in Medicinal Chemistry* **2004**, *4*, 991-1007.
5. Lenhard, J. M.; Croom, D. K.; Weiel, J. E.; Winegar, D. A. *Arteriosclerosis Thrombosis and Vascular Biology* **2000**, *20*, 2625-2629.
6. Pomerantz, R. J.; Horn, D. L. *Nature Medicine* **2003**, *9*, 867-873.
7. Kempf, D. J.; Norbeck, D. W.; Codacovi, L. M.; Wang, X. C.; Kohlbrenner, W. E.; Wideburg, N. E.; Paul, D. A.; Knigge, M. F.; Vasavanonda, S.; Craiggennard, A.; Saldivar, A.; Rosenbrook, W.; Clement, J. J.; Plattner, J. J.; Erickson, J. *Journal of Medicinal Chemistry* **1990**, *33*, 2687-2689.
8. Brik, A.; Wong, C. H. *Organic & Biomolecular Chemistry* **2003**, *1*, 5-14.
9. Lee, T.; Laco, G. S.; Torbett, B. E.; Fox, H. S.; Lerner, D. L.; Elder, J. H.; Wong, C. H. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 939-944.
10. Zhang, X. Q.; Schooley, R. T.; Gerber, J. G. *Journal of Infectious Diseases* **1999**, *180*, 1833-1837.
11. Randolph, J. T.; DeGoey, D. A. *Current Topics in Medicinal Chemistry* **2004**, *4*, 1079-1095.
12. Mahato, R. I.; Narang, A. S.; Thoma, L.; Miller, D. D. *Critical Reviews in Therapeutic Drug Carrier Systems* **2003**, *20*, 153-214.
13. Mosebi, S.; Morris, L.; Dirr, H. W.; Sayed, Y. *Journal of Virology* **2008**, *82*, 11476-11479.
14. Chene, G.; Sterne, J. A.; May, M.; Costagliola, D.; Ledergerber, B.; Phillips, A. N.; Dabis, F.; Lundgren, J.; Monforte, A. D.; de Wolf, F.; Hogg, R.; Reiss, P.; Justice, A.; Leport, C.; Staszewski, S.; Gill, J.; Fatkenheuer, G.; Egger, M. E.; Collaboration, A. R. T. C. *Lancet* **2003**, *362*, 679-686.
15. Kempf, D. J.; Marsh, K. C.; Denissen, J. F.; McDonald, E.; Vasavanonda, S.; Flentge, C. A.; Green, B. E.; Fino, L.; Park, C. H.; Kong, X. P.; Wideburg, N. E.; Saldivar, A.; Ruiz, L.; Kati, W. M.; Sham, H. L.; Robins, T.; Stewart, K. D.; Hsu, A.; Plattner, J. J.; Leonard, J. M.; Norbeck, D. W. *Proceedings of the National Academy of Sciences of the United States of America* **1995**, *92*, 2484-2488.
16. Martinez-Cajas, J. L.; Wainberg, M. A. *Antiviral Research* **2007**, *76*, 203-221.
17. Cohen, J. *Science* **2000**, *288*, 2150-2153.
18. Bollinger, L.; Stover, J. *The Futures Group International* **1999**.
19. Schwab, R. S.; England, A. C.; Poskanze, D.; Young, R. R. *Journal of the American Medical Association* **1969**, *208*, 1168-1170.
20. Neumeyer, J. L. *Principles of Medicinal Chemistry*; Lea and Febiger, Philadelphia, Pa: London, **1989**, 223.
21. James, B.; Viji, S.; Mathew, S.; Nair, M. S.; Lakshmanan, D.; Kumar, R. A. *Tetrahedron Letters* **2007**, *48*, 6204-6208.
22. Geldenhuys, W. J.; Malan, S. F.; Bloomquist, J. R.; Marchand, A. P.; Van der Schyf, C. J. *Medicinal Research Reviews* **2005**, *25*, 21-48.
23. Oliver, D. W.; Malan, S. F. *Medicinal Chemistry Research* **2008**, *17*, 137-151.
24. Oliver, D. W.; Dekker, T. G.; Snyckers, F. O.; Fourie, T. G. *Journal of Medicinal Chemistry* **1991**, *34*, 851-854.
25. Bisetty, K.; Corcho, F. J.; Canto, J.; Kruger, H. G.; Perez, J. J. *Journal of Peptide Science* **2006**, *12*, 92-105.

26. Ito, F. M.; Petroni, J. M.; de Lima, D. P.; Beatriz, A.; Marques, M. R.; de Moraes, M. O.; Costa-Lotuflo, L. V.; Montenegro, R. C.; Magalhaes, H. I. F.; Pessoa, C. D. O. *Molecules* **2007**, *12*, 271-282.
27. Oliver, D. W.; Dekker, T. G.; Snyckers, F. O. *Arzneimittel-Forschung/Drug Research* **1991**, *41-1*, 549-552.
28. Inamoto, Y.; Aigami, K.; Kadono, T.; Nakayama, H.; Takatsuki, A.; Tamura, G. *Journal of Medicinal Chemistry* **1977**, *20*, 1371-1374.
29. Aigami, K.; Inamoto, Y.; Takaishi, N.; Fujikura, Y.; Takatsuki, A.; Tamura, G. *Journal of Medicinal Chemistry* **1976**, *19*, 536-540.
30. Clavel, F.; Guetard, D.; Brunvezinet, F.; Chamaret, S.; Rey, M. A.; Santosferreira, M. O.; Laurent, A. G.; Dauguet, C.; Katlama, C.; Rouzioux, C.; Klatzmann, D.; Champalimaud, J. L.; Montagnier, L. *Science* **1986**, *233*, 343-346.
31. Gallo, S. A.; Reeves, J. D.; Garg, H.; Foley, B.; Doms, R. W.; Blumenthal, R. *Retrovirology* **2006**, *3*, 90.
32. S. Kartikeyan, R. N. B., R. P. Tiwari *HIV and AIDS: Basic elements and priorities* Springer: Netherlands, **2007**, 39-51.
33. Renneberg, R. *Biotechnology for beginners* Springer: Germany, **2008**, 141-151.
34. Cunningham, A. L.; Donaghy, H.; Harman, A. N.; Kim, M.; Turville, S. G. *Current Opinion in Microbiology* **2010**, *13*, 524-529.
35. Ji, J. X.; Chen, J. J. Y.; Braciale, V. L.; Cloyd, M. W. *Journal of Leukocyte Biology* **2007**, *81*, 297-305.
36. Davenport, M. P.; Petravic, J. *Plos Pathogens* **2010**, *6*, e1000728.
37. Migueles, S. A.; Connors, M. *Jama-Journal of the American Medical Association* **2010**, *304*, 194-201.
38. Bonsor, K.; How AIDS Works:<http://health.howstuffworks.com/diseases-conditions/infectious/aids.htm>, **2010**.
39. Turner, B. G.; Summers, M. F. *Journal of Molecular Biology* **1999**, *285*, 1-32.
40. Monini, P.; Sgadari, C.; Toschi, E.; Barillari, G.; Ensoli, B. *Nature Reviews Cancer* **2004**, *4*, 861-875.
41. Matthews, T.; Salgo, M.; Greenberg, M.; Chung, J.; DeMasi, R.; Bolognesi, D. *Nature Reviews Drug Discovery* **2004**, *3*, 215-225.
42. Ayoub, A.; Cannou, C.; Nugeyre, M. T.; Barre-Sinoussi, F.; Menu, E. *Retrovirology* **2008**, *5*, 31.
43. Whitcomb, J. M.; Huang, W.; Fransen, S.; Limoli, K.; Toma, J.; Wrin, T.; Chappey, C.; Kiss, L. D. B.; Paxinos, E. E.; Petropoulos, C. J. *Antimicrobial Agents and Chemotherapy* **2007**, *51*, 566-575.
44. Ho, H. T.; Fan, L.; Nowicka-Sans, B.; McAuliffe, B.; Li, C. B.; Yamanaka, G.; Zhou, N. N.; Fang, H.; Dicker, I.; Dalterio, R.; Gong, Y. F.; Wang, T.; Yin, Z. W.; Ueda, Y.; Matiske, J.; Kadow, J.; Clapham, P.; Robinson, J.; Colonna, R.; Lin, P. F. *Journal of Virology* **2006**, *80*, 4017-4025.
45. Shaheen, F.; Collman, R. G. *Current Opinion in Infectious Diseases* **2004**, *17*, 7-16.
46. Dau, B.; Holodniy, M. *Drugs* **2009**, *69*, 31-50.
47. Painter, G. R.; Almond, M. R.; Mao, S. L.; Liotta, D. C. *Current Topics in Medicinal Chemistry* **2004**, *4*, 1035-1044.
48. De Clercq, E. *Reviews in Medical Virology* **2000**, *10*, 255-277.
49. De Clercq, E. *Antiviral Research* **1998**, *38*, 153-179.
50. Drake, S. M. *Journal of Antimicrobial Chemotherapy* **2000**, *45*, 417-420.
51. Mitsuya, H.; Weinhold, K. J.; Furman, P. A.; Stclair, M. H.; Lehrman, S. N.; Gallo, R. C.; Bolognesi, D.; Barry, D. W.; Broder, S. *Proceedings of the National Academy of Sciences of the United States of America* **1985**, *82*, 7096-7100.
52. Yarchoan, R.; Mitsuya, H.; Thomas, R. V.; Pluda, J. M.; Hartman, N. R.; Perno, C. F.; Marczyk, K. S.; Allain, J. P.; Johns, D. G.; Broder, S. *Science* **1989**, *245*, 412-415.

53. Soudeyans, H.; Yao, X. J.; Gao, Q.; Belleau, B.; Kraus, J. L.; Nghe, N. B.; Spira, B.; Wainberg, M. A. *Antimicrobial Agents and Chemotherapy* **1991**, *35*, 1386-1390.
54. Hargrave, K. D.; Proudfoot, J. R.; Grozinger, K. G.; Cullen, E.; Kapadia, S. R.; Patel, U. R.; Fuchs, V. U.; Mauldin, S. C.; Vitous, J.; Behnke, M. L.; Klunder, J. M.; Pal, K.; Skiles, J. W.; McNeil, D. W.; Rose, J. M.; Chow, G. C.; Skoog, M. T.; Wu, J. C.; Schmidt, G.; Engel, W. W.; Eberlein, W. G.; Saboe, T. D.; Campbell, S. J.; Rosenthal, A. S.; Adams, J. *Journal of Medicinal Chemistry* **1991**, *34*, 2231-2241.
55. Romero, D. L.; Morge, R. A.; Biles, C.; Berriospena, T. N.; May, P. D.; Palmer, J. R.; Johnson, P. D.; Smith, H. W.; Busso, M.; Tan, C. K.; Voorman, R. L.; Reusser, F.; Althaus, I. W.; Downey, K. M.; So, A. G.; Resnick, L.; Tarpley, W. G.; Aristoff, P. A. *Journal of Medicinal Chemistry* **1994**, *37*, 999-1014.
56. Young, S. D.; Britcher, S. F.; Tran, L. O.; Payne, L. S.; Lumma, W. C.; Lyle, T. A.; Huff, J. R.; Anderson, P. S.; Olsen, D. B.; Carroll, S. S.; Pettibone, D. J.; O'Brien, J. A.; Ball, R. G.; Balani, S. K.; Lin, J. H.; Chen, I. W.; Schleif, W. A.; Sardana, V. V.; Long, W. J.; Byrnes, V. W.; Emini, E. A. *Antimicrobial Agents and Chemotherapy* **1995**, *39*, 2602-2605.
57. Kane, B. M. *Drugs the straight facts: HIV/AIDS treatment drugs*; Chelsea House Books: United states of America, **2008**, 66-69.
58. Roberts, N. A.; Martin, J. A.; Kinchington, D.; Broadhurst, A. V.; Craig, J. C.; Duncan, I. B.; Galpin, S. A.; Handa, B. K.; Kay, J.; Krohn, A.; Lambert, R. W.; Merrett, J. H.; Mills, J. S.; Parkes, K. E. B.; Redshaw, S.; Ritchie, A. J.; Taylor, D. L.; Thomas, G. J.; Machin, P. J. *Science* **1990**, *248*, 358-361.
59. Vacca, J. P.; Dorsey, B. D.; Schleif, W. A.; Levin, R. B.; McDaniel, S. L.; Darke, P. L.; Zugay, J.; Quintero, J. C.; Blahy, O. M.; Roth, E.; Sardana, V. V.; Schlabach, A. J.; Graham, P. I.; Condra, J. H.; Gotlib, L.; Holloway, M. K.; Lin, J.; Chen, I. W.; Vastag, K.; Ostovic, D.; Anderson, P. S.; Emini, E. A.; Huff, J. R. *Proceedings of the National Academy of Sciences of the United States of America* **1994**, *91*, 4096-4100.
60. Kaldor, S. W.; Kalish, V. J.; Davies, J. F.; Shetty, B. V.; Fritz, J. E.; Appelt, K.; Burgess, J. A.; Campanale, K. M.; Chirgadze, N. Y.; Clawson, D. K.; Dressman, B. A.; Hatch, S. D.; Khalil, D. A.; Kosa, M. B.; Lubbehusen, P. P.; Muesing, M. A.; Patick, A. K.; Reich, S. H.; Su, K. S.; Tatlock, J. H. *Journal of Medicinal Chemistry* **1997**, *40*, 3979-3985.
61. Kim, E. E.; Baker, C. T.; Dwyer, M. D.; Murcko, M. A.; Rao, B. G.; Tung, R. D.; Navia, M. A. *Journal of the American Chemical Society* **1995**, *117*, 1181-1182.
62. Sham, H. L.; Kempf, D. J.; Molla, A.; Marsh, K. C.; Kumar, G. N.; Chen, C. M.; Kati, W.; Stewart, K.; Lal, R.; Hsu, A.; Betebenner, D.; Korneyeva, M.; Vasavanonda, S.; McDonald, E.; Saldivar, A.; Wideburg, N.; Chen, X. Q.; Niu, P.; Park, C.; Jayanti, V.; Grabowski, B.; Granneman, G. R.; Sun, E.; Japour, A. J.; Leonard, J. M.; Plattner, J. J.; Norbeck, D. W. *Antimicrobial Agents and Chemotherapy* **1998**, *42*, 3218-3224.
63. Robinson, B. S.; Riccardi, K. A.; Gong, Y. F.; Guo, Q.; Stock, D. A.; Blair, W. S.; Terry, B. J.; Deminie, C. A.; Djang, F.; Colonno, R. J.; Lin, P. F. *Antimicrobial Agents and Chemotherapy* **2000**, *44*, 2093-2099.
64. Doyon, L.; Tremblay, S.; Bourgon, L.; Wardrop, E.; Cordingley, M. G. *Antiviral Research* **2005**, *68*, 27-35.
65. Ghosh, A. K.; Dawson, Z. L.; Mitsuya, H. *Bioorganic & Medicinal Chemistry* **2007**, *15*, 7576-7580.
66. Urban, J.; Konvalinka, J.; Stehlikova, J.; Gregorova, E.; Majer, P.; Soucek, M.; Andreansky, M.; Fabry, M.; Strop, P. *Febs Letters* **1992**, *298*, 9-13.
67. Desolms, S. J.; Giuliani, E. A.; Guare, J. P.; Vacca, J. P.; Sanders, W. M.; Graham, S. L.; Wiggins, J. M.; Darke, P. L.; Sigal, I. S.; Zugay, J. A.; Emini, E. A.; Schleif, W. A.; Quintero, J. C.; Anderson, P. S.; Huff, J. R. *Journal of Medicinal Chemistry* **1991**, *34*, 2852-2857.
68. Reetz, M. T.; Merk, C.; Mehler, G. *Chemical Communications* **1998**, 2075-2076.
69. Pettit, S. C.; Gulnik, S.; Everitt, L.; Kaplan, A. H. *Journal of Virology* **2003**, *77*, 366-374.

70. scottremley. In http://www.scottremley.com/viral_images/HIV_Protease_dimerization_domain_ColorFullPage_small.gif, **2010**.
71. Camarasa, M. J.; Velazquez, S.; San-Felix, A.; Perez-Perez, M. J.; Gago, F. *Antiviral Research* **2006**, *71*, 260-267.
72. Ingr, M.; Uhlikova, T.; Strisovsky, K.; Majerova, E.; Konvalinka, J. *Protein Science* **2003**, *12*, 2173-2182.
73. Abdel-Rahman, H. M.; Al-karamany, G. S.; El-Koussi, N. A.; Youssef, A. F.; Kiso, Y. *Current Medicinal Chemistry* **2002**, *9*, 1905-1922.
74. Van der Schyf, C. J.; Geldenhuys, W. J. *Neurotherapeutics* **2009**, *6*, 175-186.
75. Grobler, E.; Grobler, A.; Van der Schyf, C. J.; Malan, S. F. *Bioorganic & Medicinal Chemistry* **2006**, *14*, 1176-1181.
76. Brookes, K. B.; Hickmott, P. W.; Jutle, K. K.; Schreyer, C. A. *South African Journal of Chemistry-Suid-Afrikaanse Tydskrif Vir Chemie* **1992**, *45*, 8-11.
77. Chorghade, M. S. *Drug discovery and development*; John Wiley & Sons, **2006**, *1*, 42-45.
78. Wuthrich, K. *Journal of Biological Chemistry* **1990**, *265*, 22059-22062.
79. Wüthrich, K. *NMR in structural biology*; World Scientific, **1995**, 11-15.
80. Thiele, C. M.; Petzold, K.; Schleucher, J. *Chemistry-a European Journal* **2009**, *15*, 585-588.
81. Bothnerby, A. A.; Stephens, R. L.; Lee, J. M.; Warren, C. D.; Jeanloz, R. W. *Journal of the American Chemical Society* **1984**, *106*, 811-813.
82. Bernstein, N.; Kermode, J. R.; Csanyi, G. *Reports on Progress in Physics* **2009**, *72*.
83. Burger, S. K.; Thompson, D. C.; Ayers, P. W. *Journal of Chemical Information and Modeling*, *51*, 93-101.
84. Balbuena, P. B.; Seminario, J. M. *Molecular Dynamics*; Elsevier, **1999**, *7*, 8-15.
85. Leach, A. R. *Molecular Modelling* Addison Wesley Longman Limited: Singapore, **1996**, 524-541.
86. Dittrich, M.; Yu, J.; Schulten, K. *Atomistic Approaches in Modern Biology: From Quantum Chemistry to Molecular Simulations* **2007**, *268*, 319-347.
87. Miguel A.L. Marques, C. A. U., Fernando Nogueira *Time-Dependent Density Functional Theory*; Springer: Netherlands, **2006**, 224-227.
88. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *Journal of Computational Chemistry* **1998**, *19*, 1639-1662.
89. Lengauer, T.; Rarey, M. *Current Opinion in Structural Biology* **1996**, *6*, 402-406.

CHAPTER 2

PENTACYCLOUNDECANE-BASED INHIBITORS OF WILD TYPE C-SOUTH AFRICAN HIV-PROTEASE

INTRODUCTION

In this chapter, the first account of pentacycloundecane (PCU) lactam-peptide based HIV protease inhibitors is presented.

The PCU lactam was selected for this study because it contains a hydroxyl carbonyl functional group, resembling a norstatine transition-state isostere (Figure 13, Chapter 1).¹ Furthermore, it consists of a very stable amide bond situated at the axial position of a rigid cyclohexane boat structure (see the figure with Table 1) which can potentially interact with the catalytic residues in the active site.² Various peptide sequences were coupled to the racemic PCU-cage (see Table 1). The first sequence was selected from a natural HIV-protease substrate, FEAIS,³ where F is replaced with the cage lactam, to potentially ensure high specificity of the proposed inhibitor to the active site. This peptide sequence was then systematically varied.

Although the PCU lactam has a non-conventional bulky structure,⁴ it has been reported that the S_1 and S_1' sub-binding site of the HIV-PR accommodates bulky hydrophobic substituents.¹ We therefore proposed that the PCU group of the inhibitor could potentially occupy these sub-sites. Even though the S_2 and S_2' pockets are of a hydrophobic nature too, both hydrophilic and hydrophobic side chains from the *gag* and *gag-pol* polyproteins can occupy these sites.⁵ Based on literature and the HIV substrate sequence, we envisaged that either glutamic acid or alanine, in the peptide chain, would be a preferred binder to the S_2 and S_2' sub-sites.

Presented herein is the first report of the synthesis, biological activity, NMR and computational investigation of seven PCU derived HIV-protease inhibitors. We introduce the PCU-lactam as a potent and accessible moiety for protease inhibition.

MATERIALS AND METHODS

PCU-Peptide Synthesis: Peptides were synthesized on a CEM automated microwave assisted peptide synthesizer on 2-chlorotrityl-chloride resin *via* the Fmoc strategy. The desired crude peptides, all (*S*)-amino acids, were cleaved from the resin and coupled to the PCU lactam in solution. Purification of the peptides was achieved *via* semi-preparative HPLC. Only one of the diastereomeric cage peptides could be separated (see Table 1). Detailed methods for the synthesis and purification are described in the Appendix 1.

Optical Rotation (OR) and Circular Dichroism (CD): Measurements were recorded on a Perkin-Elmer Polarimeter (Model 341) and Jasco J-810 CD spectrometer instruments, respectively, and all compounds were dissolved in MeOH. Optical rotation was measured at 589 nm and the CD spectra were recorded from 190 to 260 nm.

NMR: Samples were dissolved in 450 μ l of deuterated DMSO- d_6 and measured at 25 °C. All ^1H , ^{13}C , COSY, HMBC and HSQC NMR spectra, required for assignment were recorded on a Bruker AVANCE III 400 MHz spectrometer, BBO probe with z-gradient with standard parameters. The EASY-ROESY measurements required for structural analysis were recorded on a Bruker AVANCE III 600 MHz spectrometer using a BBO probe with z-gradient at a transmitter frequency of 600.1 MHz (spectral width, 8223.7 Hz; acquisition time, 0.1704436 s; 90° pulse width, 11.02 μ s; scans, 8; relaxation delay, 2.0 s, mixing time 0.12s).⁶ Processing and assignments were carried out using the Topspin 2.3 software from Bruker Karlsruhe. In order to determine if the same 3D structural conformation observed for the inhibitors in DMSO- d_6 are also valid for the HIV-PR activity studies, the NMR spectra for PCU-EAIS were recorded in a buffered D_2O solution. All spectra mentioned in the text are available in Appendix 1.

***In vitro* HIV-1 protease inhibition assay:** Synthesis and purification of the HIV protease type C were performed as described in Appendix 1.⁷⁻⁹ The catalytic activity of the protease was monitored following the hydrolysis of a chromogenic peptide substrate, Ala-Arg-Val-Nle-*p*-nitro-Phe-Glu-Ala-Nle-NH₂ at 300 nm using an Epcord 210 spectrophotometer. Activity was standardized using commercially available drugs Atazanavir and Lopinavir. All cage peptides were soluble in the aqueous buffer solution.

Computational methods: The structures of the selected cage diastereomeric compounds [PCU-AISa/b and PCU-EAISa/b] were constructed using the Avogadro software.¹⁰ A β -sheet conformation was used as starting structure for both the conformational and docking studies.

3D structure of the inhibitors in aqueous solution

A computational method, in which molecular dynamics using a hybrid quantum mechanics/molecular mechanics (QM/MM) molecular dynamics MD, was employed to determine the low energy conformations of the inhibitors in aqueous solution. Gas phase AM1 optimized structures of the inhibitors were enveloped in a cubic box of TIP3P water of side-length 31.4 Å.¹¹ The energy of the inhibitor in solution was minimized to a residual gradient of less than 0.001 kJ mol⁻¹ Å⁻¹ with a QM/MM method using the DYNAMO package.¹² The QM region (PCU-AIS, 70 atoms and PCU-EAIS, 86 atoms) was described by AM1 semi-empirical Hamiltonian.¹³ The whole system was then equilibrated without any constraints by MD for 50 ps at 300 K, using the NVT ensemble prior to 1 ns of separate production MD runs at 300 K, 400 K, 500 K and 600 K (1 fs time step, switched cut off radius of 16 Å applied to all interactions).¹⁴⁻¹⁶

Docking of the inhibitors with the C-SA PR

A model system for C-SA HIV-PR was constructed by substituting T12S, I15V, L19I, M36I, R41K, H69K, L89M and I93L in the crystal structure of the subtype B HIV-PR (PDB accession code 1HXW).^{7,17} Docking studies were performed using the Autodock software¹⁸ under physiological pH conditions, ensuring the correct protonation state of all ionizable groups.¹⁹ Both aspartates 25 (D25) were fully protonated and as a control one of the D25 acid groups was deprotonated in an independent run. The suitability of the constructed C-SA PR for docking experiments was tested by a control experiment in which Ritonavir was docked into active site of the enzyme. The cage-peptide inhibitors were docked into the active site pocket of the modeled C-SA.

QM/MM/MD simulation of the HIV-PR and cage peptide inhibitor complexes

The docked complexes with the lowest binding energy were energy minimized as described previously and were used as an input for the MD simulations of the inhibitor-enzyme complex.²⁰ The QM region comprised of the inhibitors and the two catalytic residues, ASP25 and ASP25', was described by AM1 semi-empirical Hamiltonian. The MM region contained the rest of the enzyme and was described by the OPLSAA potential.²¹ The inhibitor-enzyme complex was enveloped in a cubic box of TIP3P water of side-length 79.5 Å. QM link atoms were placed along the C β (QM)–C α (MM) bonds of Asp25 and Asp25'. A no-constraint-MD run for 300 ps at 300 K was performed, where hydrogen bond interaction analysis and inhibitor-enzyme interaction energies were monitored along the MD simulations.

All inhibitor-enzyme complex structures mentioned in the text are available in PDB format in Appendix 1.

RESULTS AND DISCUSSION

The PCU-peptides were obtained as diastereomers from the coupling of the enantiopure short peptide to the racemic PCU-lactam.^{22,23} The sequences of the different inhibitor PCU-peptides are presented in Table 1. All inhibitors were characterized by high-resolution mass spectrometry and NMR. The peptides were tested for HIV-PR inhibition activity and the results are also presented in Table 1.

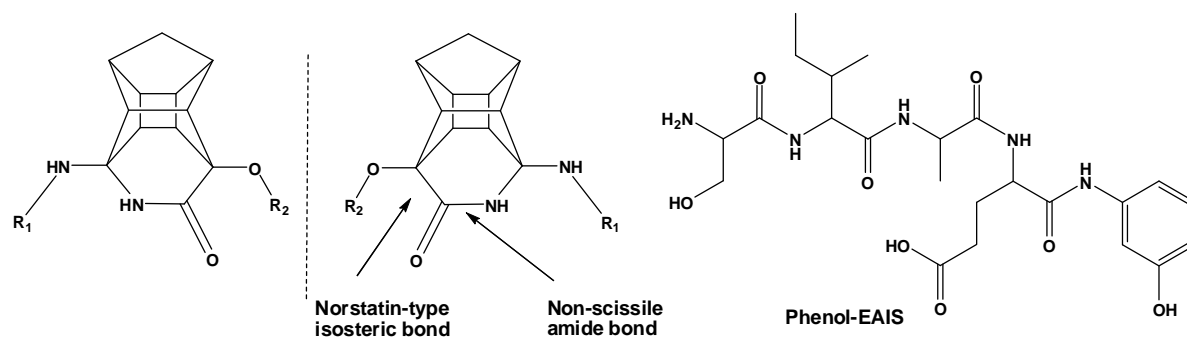


Table 1. IC₅₀ values for the inhibition of wild type C-SA HIV-1 protease by PCU lactam peptides and control compounds. Structural information from EASY-ROESY NMR and binding energies from docking experiments are also presented.

PCU-R ₁	R ₂	Yield (%)	Cage-peptide interaction (CP) ^d	Peptide-peptide interaction (PP) ^e	Docked binding energies /kcal mol ⁻¹	IC ₅₀ /μM
PCU-EAIS ^a	H	47	Yes	-	-10.23	0.078 ± 0.0035
Phenol-EAIS ^b		87	-	-	-	10.0 ± 3.53
PCU-EAIS ^a	Ac	27	-	-	-7.06	>10
PCU-EAI ^a	H	23	Yes	-	-	0.5 ± 0.035
PCU-EVIS ^a	H	33	Yes	-	-	2.0 ± 0.18
PCU-QAIS ^a	H	20	-	-	-	5.0 ± 0.71
PCU-AISa ^c	H	18	Yes	-	-9.41	0.5 ± 0.035
PCU-AISb ^c	H	17	-	Yes	-9.27	10.0 ± 1.06
Atazanavir						0.004 ± 0.00071
Lopinavir						0.025 ± 0.0014

a. Diastereomeric mixtures.

b. Enantiopure.

c. Diastereomer was separated with semi-preparative HPLC.

d. CP: ROE interaction of the Cage protons with a Peptide NH in solution was observed.

e. PP: ROE interaction of the same Peptide NH with the Peptide side chain in solution was observed.

Interestingly, the first cage peptide sequence (PCU-EAIS) that was chosen from the natural HIV-protease substrate (FEAIS³ where F is replaced with the cage lactam) gave the best inhibition activity ($IC_{50} = 78$ nM). To confirm the significance of the role of the cage in the inhibition of the HIV-PR active center it was replaced with hydroxyphenylamine to produce the control peptide (Phenol-EAIS), which had an $IC_{50} = 10.0$ μ M, a 120 fold decrease when compared to PCU-EAIS. Furthermore the hydroxyl-group at position C8 of the PCU-cage is essential, because after acetylation thereof, activity was completely lost (Ac-PCU-EAIS with $IC_{50} > 10.0$ μ M).

Systematic removal or substitution of the peptide amino acid residues resulted in a decrease in inhibitory activity. Removal of the serine segment resulted in a six fold decrease (PCU-EAI with $IC_{50} = 0.5$ μ M), the substitution of alanine with valine to PCU-EVIS caused a 25 fold decrease, substitution of the glutamic acid in PCU-EAIS with glutamine resulted in PCU-QAIS with a 64 fold decrease. Removal of the glutamic acid yielded diastereoisomers PCU-AISa ($IC_{50} = 0.5$ μ M) and PCU-AISb ($IC_{50} = 10.0$ μ M), which were separated by preparative HPLC. These diastereomers exhibited a 20-fold difference in activity and a 6 fold and 120 fold reduction in respective activities, compared to PCU-EAIS.

These results clearly confirm the significant bioactive contribution of the cage, with a non-cleavable norstatine-type bond and the importance of the appropriate peptide sequence for enhanced and specific binding to the active site.

NMR and computational chemistry techniques were employed to determine a potential correlation between the IC_{50} results and the 3D structure of the corresponding inhibitors in solution. We measured 1H and ^{13}C spectra, 2D $^1H, ^1H$ correlation and $^1H, ^{13}C$ correlation spectra to validate the chemical structure of all compounds (NMR assignments are presented within the Appendix 1).

The NMR elucidation of the mixture of PCU-EAIS diastereomers were extremely challenging. Firstly, the diastereomers could not be separated and secondly, it was clear from the data that each of the two diastereomers also exhibited different conformational preferences. It therefore made sense to assign the signals for the separated diastereomers PCU-AISa and b first. In addition, these diastereomers exhibited a 20-fold difference in IC_{50} , which may enable us to investigate the influence of the cage-chirality on the HIV PR inhibition by means of applied high-resolution heteronuclear NMR techniques. The 1H and ^{13}C NMR spectra of PCU-AISa and PCU-AISb were very similar except for carbons C-1 and C-10 of the PCU-cage.

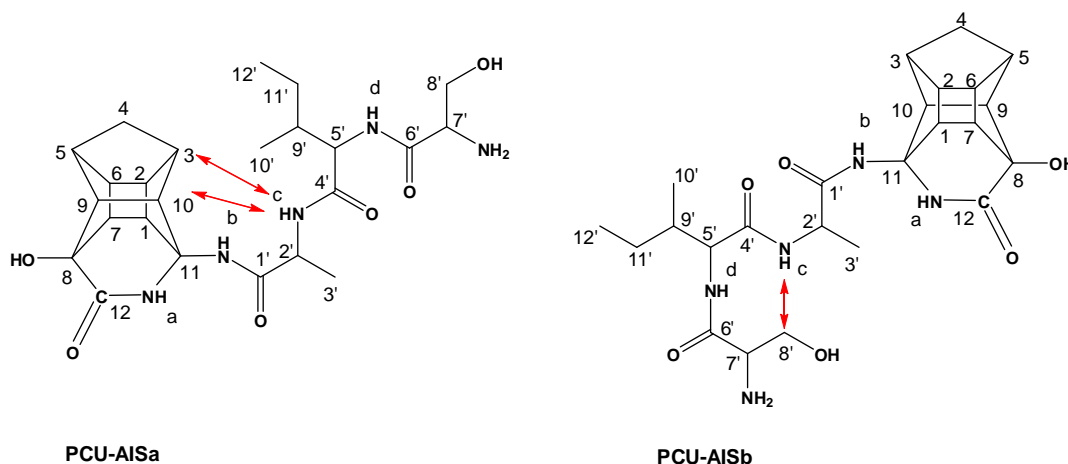


Figure 1. EASY-ROESY long-range correlations observed for the two cage peptide diastereomers.

The ^1H , ^{13}C -HSQC spectrum of PCU-AISa showed that carbons C-1 and C-10 of the cage were split, which was not observed in the spectrum of PCU-AISb. Upon heating, the splitting was reduced and disappeared completely at 333 K (60 °C) indicating that different inter-convertible conformations exist. The EASY-ROESY spectra enabled us to detect ^1H , ^1H through space ROESY interactions of the various inhibitors.⁶ PCU-AISa showed long-range correlations between the NHc amide protons of alanine with the PCU protons H-1, H-3 and H-10 and the side chain methyl protons of alanine (H-3') also showed a correlation with H-3 of the cage. This first conformation was named CP (long-range cage-peptide interaction). Therefore the splitting of the cage carbon signals C-10 and C-1 of PCU-AISa was attributed to the existence of at least two low energy conformations of the peptide side chain. In the dominant conformation (CP), a through space shielding of C-1 and C-10 existed as result of the interaction between one of the split alanine amide proton signals (NHc, 8.03 ppm) and the cage protons H-1 and H-10. We observed a second peptide conformation for PCU-AISa, in which the alanine NHc signal is at 8.07 ppm. These conformations are most certainly induced by the chiral cage structure and are remarkably stable in the light of the relatively high temperatures (~ 60 °C) required to obtain interconversion.

PCU-AISb showed only one set of C-1 and C-10 resonances in the ^1H , ^{13}C -HSQC, coinciding with the second conformation of PCU-AISa. It was clear from the EASY-ROESY spectra that CP interactions for PCU-AISb were absent. Interestingly, a ROESY interaction of the alanine-NHc (8.10 ppm) with the serine side chain (H-8' – see Figure 1) was observed for PCU-AISb; which was not observed for PCU-AISa.

This other long-range peptide-peptide interaction was named PP. We therefore concluded, that the chirality* of the cage induces different secondary structures on the peptide chains, either interacting with the cage (CP) or with the peptide itself (PP). The difference in HIV-PR inhibition activities of these two inhibitors (PCU-AISa is 20 times more active than PCU-AISb) seem to suggest that the peptide conformations induced by the cage enantiomers play an important role in the activity.

Applying the knowledge obtained from elucidating the NMR spectra of PCU-AIS, enabled us to solve the intriguing NMR spectra of the PCU-EAIS mixture. The EASY-ROESY spectrum of the most potent inhibitor, the diastereomeric mixture of PCU-EAIS, displayed the same dominant CP conformation, as present for PCU-AISa. We could not detect the PP conformation as present in PCU-AISb, but a second, most likely elongated, side chain conformation is present, indicated by the splitting of C-1 and C-10. To prove that the CP conformation is biologically relevant and present in the conditions of binding to the HIV-PR, NMR spectra of PCU-EAIS were also recorded in an aqueous buffer (D₂O) used for the HIV-PR inhibition assay; similar EASY-ROESY correlations and HSQC spectra were observed as with DMSO-d₆ (see Table 11 in Appendix 1).

The ROESY spectra of compounds PCU-EAI and PCU-EVIS also only revealed the CP conformation, but not the PP conformation. The ROESY spectra of PCU-QAIS and PCU-AISb showed no sign of CP interactions and these peptides were found to be the weakest inhibitors in the group. This chirality induced conformation of the cage with the peptide chain is crucial for better recognition and/or binding.

It was anticipated that a computational chemistry investigation could explain why the change of the peptide sequence reduced the inhibitory activity. We therefore embarked on a computational chemistry investigation involving three different approaches. First, a hybrid method of quantum mechanics, molecular mechanics and molecular dynamics (QM/MM/MD) was used to determine the prominent low energy conformations for selected cage peptide systems in aqueous solution. Second, the PCU-peptides were subjected to docking studies, and last QM/MM/MD simulations of the docked structures inside the HIV-PR enzyme were performed.

We have previously reported that cage peptides required much higher MD simulation temperatures than room temperature (298 K) for a reasonable exploration of the conformational profile.^{24,25} Our NMR results indicated that the two different conformations for PCU-AISa required heating to 333 K (60 °C) to overcome the energy barrier between them. The MD simulations for all peptides (PCU-EAISa/b and

* The chirality of the cage diastereomers were deduced from the computational work. The stereochemistry for the cage peptides are 8-(R)-11-(R)-PCU-AISa, 8-(S)-11-(S)-PCU-AISb.

PCU-AISa/b) were therefore executed at 300 K, 400 K, 500 K and 600 K. The QM/MM/MD results showed that the whole conformational space of each peptide could only be explored at 500 K. The Ramachandran plots and the 10 lowest energy structures for each of the diastereomeric peptides were analyzed (see Appendix 1). The Ramachandran plots at 500 K for PCU-EAISa revealed that only the first amino segment (Glu) favored a α -helical character (CP interaction is possible) while the other segments (Ala and Ile) both exhibited β -sheet conformations (CP interaction is not possible). The alanine segment of PCU-AISa adopted a α -helical conformation, while that of PCU-AISb exhibited a β -sheet conformation. The same trend holds for the isoleucine segment.

Analysis of the 10 lowest conformations of each of the cage peptides from the MD simulation (500 K) also confirmed the NMR observations. Nine out of the ten lowest energy structures from PCU-EAISa demonstrated cage peptide (CP) interactions (distances less than 5.0 Å) while none of the PCU-EAISb lowest energy structures showed CP interaction. Similar observations were made for PCU-AISa and b.

These MD results therefore confirmed the existence of the conformations observed with the EASY-ROESY experiments. It also suggests that the active conformation for the most active inhibitor (PCU-EAISa) is most likely twice as active as the mixture of diastereomers, since PCU-EAISb does not exhibit the same CP conformation as all the active inhibitors.

The PCU-peptide inhibitors were next subjected to a docking experiment with the C-SA PR (see Figure 2 and Appendix 1). To test the validity of the technique with the newly generated C-SA enzyme, a control docking experiment with Ritonavir was performed. The docked complex orientation gave a reasonable RMSD of 0.559 Å with the crystal structure of Ritonavir¹⁷ with HIV-PR type B. Upon docking it was found that for the most potent inhibitor PCU-EAISa, the hydroxyl-group of the PCU-lactam (the proposed norstatine moiety) was forming a hydrogen bond with Asp25 of the dimeric catalytic triad residues Asp25-Thr26-Gly27 (A/B chains), most noticeable Asp25 and Asp25' (see Figure 3). The docked structures of Ac-PCU-EAISa and PCU-EAISb failed to reveal any interaction between the cage lactam and the two Asp25 segments of the PR. The binding energies of the docked inhibitor enzyme complexes are reported in Table 1.

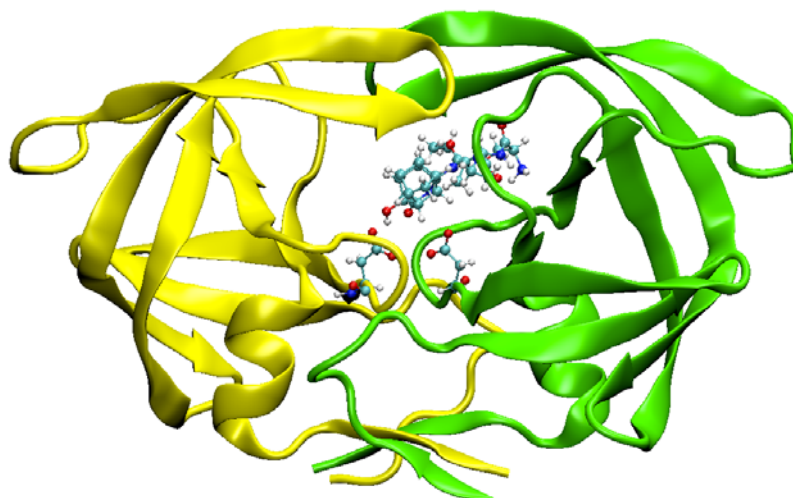


Figure 2. Lowest energy docked structure for PCU-AISa (CP) with C-SA HIV-PR. The two PR monomers are colored in yellow and green ribbon and the inhibitor and the two Asp25 residues are presented in ball and stick form. The 3D presentations for all docking results are available as PDB files in the Appendix 1.

The second lowest docked binding energy was obtained for PCU-AISa, which also exhibited the second best HIV-PR inhibition activity. Interestingly, the low energy docked cage peptide structures (PCU-EAISa and PCU-AISa) did not exhibit the same cage-peptide (CP) interaction observed with the NMR studies and the QM/MM/MD solution structures. This observation could suggest that the specific low energy CP conformation is not the active conformation.

The crude docked binding energies are largely in accordance with the *in vitro* HIV-PR inhibition results. Docking of the same inhibitors with C-SA PR where only one of the Asp25 residues was protonated, gave similar docking behavior to the unprotonated Asp25 residues.

The lowest energy docked complexes of the inhibitors (PCU-EAISa/b and PCU-AISa/b) with the C-SA PR were then subjected to unconstrained QM/MM/MD unconstrained simulations (300 ps) to obtain a better simulated picture where full flexibility of the inhibitor and the enzyme is taken into account. The RMSD values for each complex relative to the initial minimized structures were calculated. These values were $< 1.2 \text{ \AA}$ from the mean value, indicating that the stabilities of the dynamic equilibrium of all complexes were reliable. Very little movement of the cage lactam inside the enzyme pocket is observed during the course of the MD run, however significant degrees of movement for the peptide side chains were observed.

The success of an inhibitor is also dependant on the effectivity of hydrogen bonds that are formed with the enzyme pocket. Post-dynamic electrostatic and hydrogen bond interactions were plotted for the four complexes (Figure 3). Analysis of the hydrogen bond interactions along the MD trajectories clearly showed that PCU-EAISa and PCU-AISa were forming hydrogen bonds between the cage hydroxyl group and at least one of the two Asp25 groups. This interaction for the PCU-EAISb and PCU-AISb structures did not develop during the MD simulation as the low energy docked complexes gave a starting structure where the cage hydroxyl group was rotated away from the two Asp25 residues. Based on X-ray structures of inhibitor/HIV-PR (subtype B) complexes,^{17,26,27} it is clear that hydrogen bond interactions between the inhibitor and the ASP25 groups of the enzyme are crucial for effective for binding and hence effective inhibitory activity.

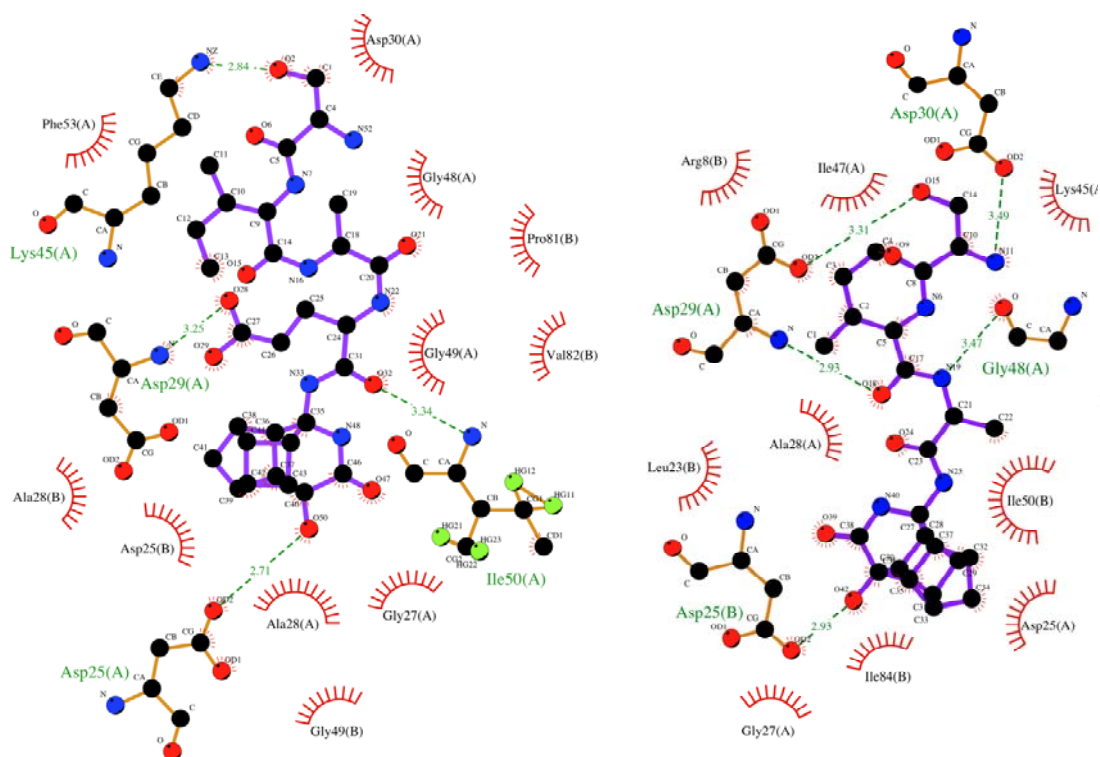


Figure 3. Selected electrostatic and hydrogen bond interactions for the CSA HIV-PR complexes with PCU-EAISa and PCU-AISa from the MD simulation. These plots were created with the Ligplot software.²⁸ (The plots for the electrostatic and hydrogen bond interactions of PCU-AISa and b and the corresponding 3D structures are provided in Appendix 1).

Unfortunately the QM/MM/MD results do not provide easily interpretable binding energy values for the enzyme-inhibitor interaction. MD results show that all inhibitors easily fit into the active enzyme pocket,

but do not provide an explanation why the cage with the natural peptide sequence (PCU-EAIS) is more active than the shorter version (PCU-AIS). The MD results confirm that the CP conformation, observed for the free inhibitor in solution, is not restored through movement of the side chain inside the enzyme pocket during the MD simulation. It therefore appears that the CP conformation is perhaps required as starting structure for an induced fit model.

CONCLUSION

The combination of the PCU-cage with HIV-PR substrate-peptides yields a potential lead compound for HIV-PR inhibition. A promising IC_{50} of 78 nM was obtained from a series of seven cage peptides. Variation of the amino acid sequence and protection of the cage hydroxyl group enabled us to determine a structure-activity relationship for this family of cage peptide inhibitors, showing that the cage with the natural HIV-PR substrate (PCU-EAIS) exhibits the best activity. These results indicate the importance of the cage for bioactivity: removal of the cage destroys activity, while activity is observed for cage-peptides where the natural peptide sequence is markedly varied. The EASY-ROESY technique has proven to be an extremely powerful method to provide vital information about the 3D solution structure of small peptides. All active inhibitors exhibited the same ROESY interaction between the cage protons and the peptide side chain. NMR and QM/MM/MD simulations were instrumental to show for the first time that the chirality of the PCU-cage induces different peptide conformations which are stable at temperatures up to 60 °C. The different cage enantiomers were shown to exhibit a 20-fold difference in their *in vitro* IC_{50} PR inhibition results. Protection of the cage hydroxyl group, which is proposed to be a norstatine type isosteric bond, resulted in a complete loss of inhibitory activity. This result was confirmed with a docking experiment where the acetyl protected cage peptide appeared to have lost the required interaction between the norstatine bond and the two enzymatic Asp25 residues. The docking results also confirmed that most active cage inhibitor (PCU-EAIS), gives the best binding energy. A MD simulation revealed that the most potent compounds exhibited interaction between the cage hydroxyl-group and one of the PR Asp25 groups. The computational results also suggest that the low energy CP conformation is most likely not the active conformation, as this conformation was not observed for the inhibitor/enzyme complexes. This interaction is likely to play a role in early host-guest recognition in a type of induced fit model. The combination of drug design with *in-vitro* assays, NMR and computational techniques has enabled us to rationalize the observed HIV-PR inhibition for the different cage peptides. Employment of these techniques, combined with a systematic variation of alternative cage analogues and peptide sequences should lead to the design of more active cage peptide inhibitors in future. The possibility of cage peptides to inhibit some of the many alternative disease related protease families will also be investigated.

Supporting Information Available in Appendix 1: Detailed procedures for the synthesis of inhibitors; over-expression, extraction, purification and *in-vitro* activity screening of the C-SA HIV-PR, Tables with NMR, mass, CLogP and microwave peptide synthesis conditions data; NMR, CD and HRMS spectra; PDB files of the enzyme-inhibitor complexes and Ramachandran plots

REFERENCES

1. Brik, A.; Wong, C. H. *Organic & Biomolecular Chemistry* **2003**, *1*, 5-14.
2. Kruger, H. G.; Martins, F. J. C.; Viljoen, A. M. *Journal of Organic Chemistry* **2004**, *69*, 4863-4866.
3. Lee, J.; Huh, M. S.; Kim, Y. C.; Hattori, M.; Otake, T. *Antiviral Research* **2010**, *85*, 425-428.
4. Kruger, H. G.; Mdluli, P.; Power, T. D.; Raasch, T.; Singh, A. *Journal of Molecular Structure-Theochem* **2006**, *771*, 165-170.
5. Wlodawer, A.; Vondrasek, J. *Annual Review of Biophysics and Biomolecular Structure* **1998**, *27*, 249-284.
6. Thiele, C. M.; Petzold, K.; Schleucher, J. *Chemistry-a European Journal* **2009**, *15*, 585-588.
7. Mosebi, S.; Morris, L.; Dirr, H. W.; Sayed, Y. *Journal of Virology* **2008**, *82*, 11476-11479.
8. Tie, Y. F.; Boross, P. I.; Wang, Y. F.; Gaddis, L.; Liu, F. L.; Chen, X. F.; Tozser, J.; Harrison, R. W.; Weber, I. T. *Febs Journal* **2005**, *272*, 5265-5277.
9. Klabe, R. M.; Bacheler, L. T.; Ala, P. J.; Erickson-Viitanen, S.; Meek, J. L. *Biochemistry* **1998**, *37*, 8735-8742.
10. Avogadro. **Version 1.0.0**<http://avogadro.openmolecules.net/>.
11. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *Journal of Chemical Physics* **1983**, *79*, 926-935.
12. Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. *Journal of Computational Chemistry* **2000**, *21*, 1088-1100.
13. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *Journal of the American Chemical Society* **1985**, *107*, 3902-3909.
14. Hamad, S.; Cristol, S.; Callow, C. R. A. *Journal of Physical Chemistry B* **2002**, *106*, 11002-11008.
15. Bisetty, K.; Perez, J. J. *Journal of Physical Chemistry B* **2009**, *113*, 5234-5238.
16. Zhang, H. Z.; Huang, F.; Gilbert, B.; Banfield, J. F. *Journal of Physical Chemistry B* **2003**, *107*, 13051-13060.
17. Kempf, D. J.; Marsh, K. C.; Denissen, J. F.; McDonald, E.; Vasavanonda, S.; Flentge, C. A.; Green, B. E.; Fino, L.; Park, C. H.; Kong, X. P.; Wideburg, N. E.; Saldivar, A.; Ruiz, L.; Kati, W. M.; Sham, H. L.; Robins, T.; Stewart, K. D.; Hsu, A.; Plattner, J. J.; Leonard, J. M.; Norbeck, D. W. *Proceedings of the National Academy of Sciences of the United States of America* **1995**, *92*, 2484-2488.
18. AutoDock. <http://www.scripps.edu/rpubrolson-web/docrautodock>.
19. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *Journal of Computational Chemistry* **1998**, *19*, 1639-1662.
20. Warshel, A.; Levitt, M. *Journal of Molecular Biology* **1976**, *106*, 421-437.
21. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *Journal of Physical Chemistry B* **2001**, *105*, 6474-6487.
22. Marchand, A. P. *Advances in theoretically interesting molecules*; JAI Press: Greenwich CT, 1989; Vol. 1.
23. Marchand, A. P.; Allen, R. W. *Journal of Organic Chemistry* **1974**, *39*, 1596-1596.

24. Bisetty, K.; Gomez-Catalan, J.; Aleman, C.; Giralt, E.; Kruger, H. G.; Perez, J. J. *Journal of Peptide Science* **2004**, *10*, 274-284.
25. Bisetty, K.; Govender, P.; Kruger, H. G. *Biopolymers* **2006**, *81*, 339-349.
26. Tyndall, J. D. A.; Reid, R. C.; Tyssen, D. P.; Jardine, D. K.; Todd, B.; Passmore, M.; March, D. R.; Pattenden, L. K.; Bergman, D. A.; Alewood, D.; Hu, S. H.; Alewood, P. F.; Birch, C. J.; Martin, J. L.; Fairlie, D. P. *Journal of Medicinal Chemistry* **2000**, *43*, 3495-3504.
27. Wu, X. Y.; Oehrngren, P.; Ekegren, J. K.; Unge, J.; Unge, T.; Wallberg, H.; Samuelsson, B.; Hallberg, A.; Larhed, M. *Journal of Medicinal Chemistry* **2008**, *51*, 1053-1057.
28. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. *Protein Engineering* **1995**, *8*, 127-134.

CHAPTER 3

PENTACYCLO-UNDECANE LACTAM PEPTIDES AND PEPTOIDS AS POTENTIAL HIV-1 WILD TYPE C-SA PROTEASE INHIBITORS

INTRODUCTION

Biodegradation is one of the major limiting factors of peptide drugs (Chapter 1). Considerable effort has gone into improving the pharmacological properties of peptides, mainly by increasing their enzymatic stability while preserving the chemical moieties required for functionality.^{1,2} Peptoids represent a class of oligomeric compounds that closely mimic the natural structure of peptides and possess increased enzymatic stability as compared with homologous peptides. Peptoids are polymeric compounds composed of *N*-substituted glycine (NSG) residues. The advances in the synthetic strategies used to prepare peptoid oligomers make them an attractive class of peptidomimetics.³⁻⁵

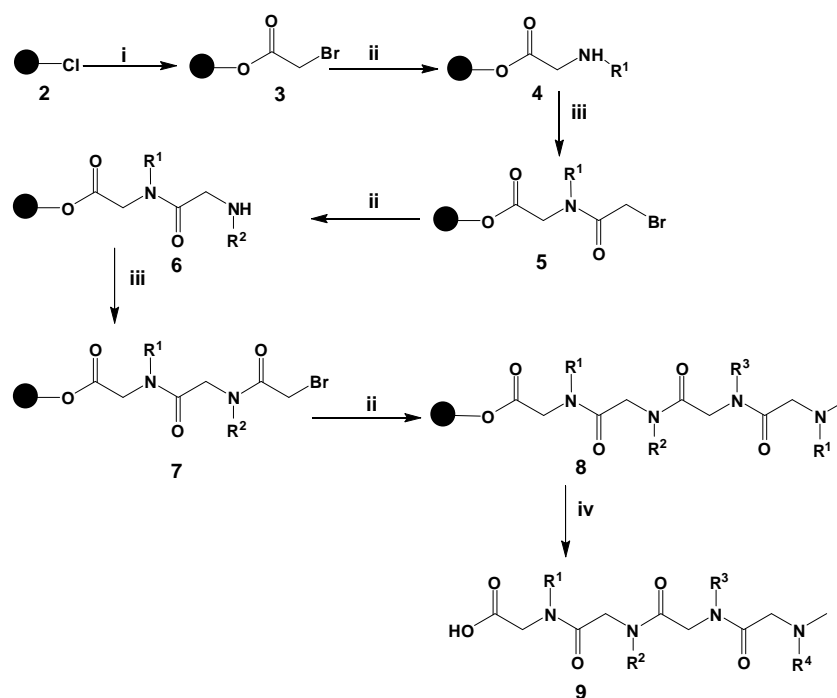
In this chapter, peptoid chains mimicking the peptide sequence phenylalanine-valine-alanine-carbobenzyloxy (FVACbz) were coupled to the PCU-lactam and tested against the wild type C-SA HIV protease enzyme. The FVACbz amino acid sequence have been proven to bind efficiently to the S₁-S₄ subsites of the HIV protease.⁶ The cage hydroxyl has been clearly shown in chapter 2 to form hydrogen bonds with at least one of the two Asp25 groups. This proved the PCU-lactam moiety's tendency to bind to the protease active site.

Herein, the synthesis, biological testing and NMR studies using the new Efficient Adiabatic SYmmetrized Rotating Overhauser Effect Spectroscopy (EASY-ROESY) technique for eight PCU-peptide and peptoid inhibitors are presented.

RESULTS AND DISCUSSION

The peptides were synthesized *via* solid phase peptide synthesis (SPPS) on a 2-chlorotriyl chloride resin. Upon activation of the resin, the first amino acid was loaded onto the resin in the presence of *N,N*-diisopropylethylamine (DIPEA). The remaining amino acids were coupled to the loaded resin

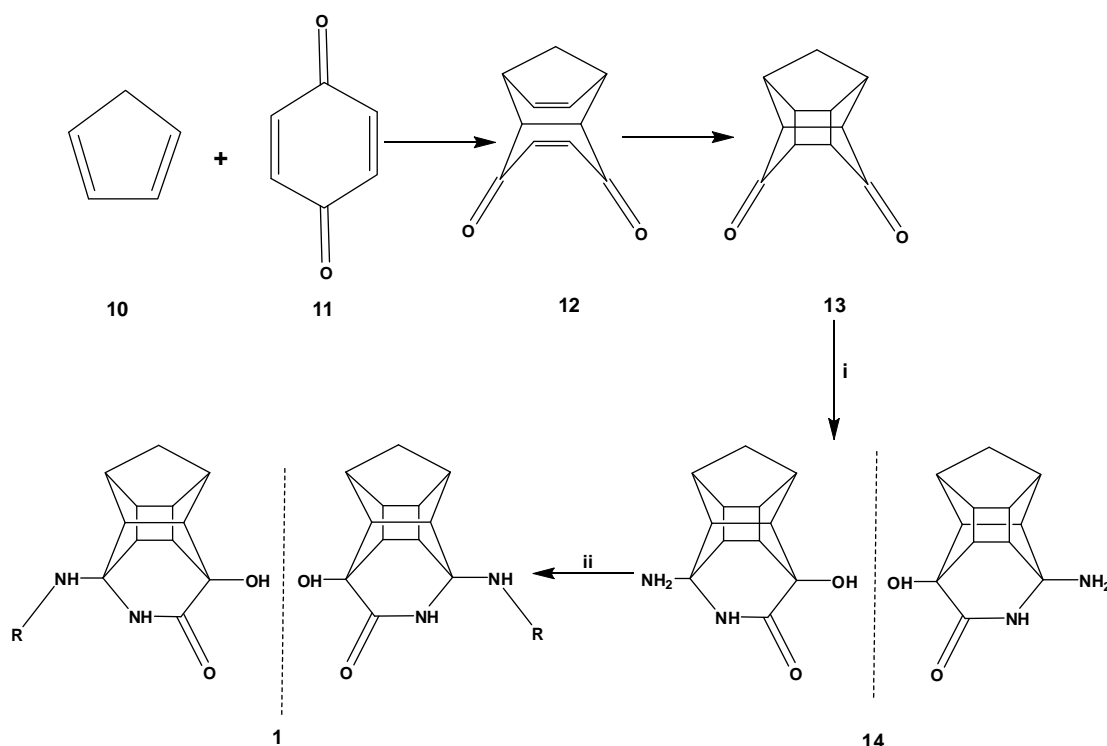
via microwave assistance on an automated peptide synthesizer. The completed peptides were cleaved from the resin with 5 % trifluoroacetic acid/dichloromethane (TFA/DCM) mixture. In most cases, the peptides/peptoids were of sufficient purity to be coupled to the lactam after lyophilisation. The peptoid oligomers were synthesized *via* solid phase synthesis on activated 2-chlorotrityl chloride resin **2** by attaching bromoacetic acid in the presence of DIPEA in dry DCM affording compound **3**. Upon filtration, the amine solution in DMF was added and mixed under a stream of nitrogen bubbles to give compound **4**. The chain was elongated by coupling bromoacetic acid in the presence of *N,N*-diisopropylcarbodiimide (DIC). Similarly the corresponding amine groups were coupled to the extended chain followed by bromoacetic acid to give compounds **5-9**. The completed peptoids were cleaved from the resin using 5 % TFA/DCM as illustrated in Scheme 1.



Scheme 1. General procedure for the synthesis of peptoids. i = bromoacetic acid, DIPEA, dry DCM ii = amine, DMF iii = bromoacetic acid, DIC, DMF iv = 5 % TFA/DCM.

The pentacycloundecane (PCU) lactam **14** was synthesized from pentacycloundecane-8, 11-dione **13** which can be easily obtained from photocyclisation of the Diels-Alder adduct **12** of cyclopentadiene **10** and *p*-benzoquinone **11**.⁷ When the dione **13** is treated with ammonia, ammonium chloride and a mixture of aqueous sodium cyanide it produces racemic amino-

hydroxylactam **14**. The peptides and peptoids were coupled to the amino lactam with 2-(7-aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) and DIPEA in dimethylformamide (DMF) at ambient temperature as shown in Scheme 2. The compounds were purified via semi-preparative HPLC and obtained as non-separable diastereomeric mixtures.



Scheme 2. Synthesis of the PCU-lactam peptide and peptoid analogues. i = NaCN, NH₄Cl, NH₃, 0 °C ii = HATU, DIPEA, DMF, RT R = peptides or peptoids.

The catalytic activities of the HIV-1 proteases were determined by monitoring the hydrolysis of the chromogenic peptide substrate His-Lys-Ala-Arg-Val-Leu-Phe(*p*-NO₂)-Glu-Ala-Nle-Ser. This substrate mimics the conserved KARVL/AEAM cleavage site between the capsid protein and nucleocapsid (CA-p2) in the *gag* polyprotein precursor and its use for the determination of the HIV-1 protease inhibitors efficacy has been established.^{8,9}

According to Table 1, it can be inferred that compound **16a** in which the phenylalanine was removed as the first amino acid from the peptide sequence showed a slightly higher activity than **15**. This could be attributed to the phenylalanine occupying the S₂ subsite instead of the S₁ subsite since the cage would already occupy this space. As mentioned before, the S₂ subsite prefers less bulky molecules such as valine thus phenylalanine would not have maximum interaction with this

subsite.¹⁰ It was therefore postulated that for **16a**, the cage norstatine moiety interacted with the active site while the cage occupied the S₁ subsite which can accommodate bulky groups. The valine and alanine would then fit into the S₂ and S₃ subsite respectively. The carbobenzyloxy (Cbz) group was removed from **16a** to increase aqueous solubility but this led to a decrease in activity (**16b** with an IC₅₀ = 0.75 μM). In an attempt to increase the aqueous solubility, the Cbz group was replaced with 2-pyrimidinyl-ethane-thioic acid and this led to an increased inhibitory activity (**16c** with an IC₅₀ = 0.5 μM).¹¹

The peptoid **17** consisting of an aromatic moiety showed the best inhibitory activity when compared to the other PCU-lactam peptoid analogs **18a-c**. This observation was attributed to the conformational difference reported by Fowler *et al.*¹² between peptoids consisting of an aromatic and/or negatively charged side chain and those with only alkyl chains. According to Fowler *et al.*¹² aromatic groups on the side chains cause charge-charge repulsion with backbone carbonyls that assist in peptoid folding. This folding can be advantageous since it gives the peptoid a well-defined secondary structure which can be lacking in most peptoids. Peptoids are highly flexible because of the lack of hydrogen bonding and the tertiary nature of the amide bonds that allows them to isomerize between *trans* and *cis* conformations far more readily than secondary amides in peptides.¹² The decrease in activity could be attributed to the lack of a defined structure for peptoids **18a-c**.

According to the NMR observations of **18a**, more than two conformations were observed for this inhibitor suggesting that the compound do not exhibit a well defined structure. Compound **18a** was designed to mimic the peptide inhibitor **16a** but the inhibitory activity of **18a** was observed to be approximately 15 fold less. NMR was then employed to justify this difference in activity.

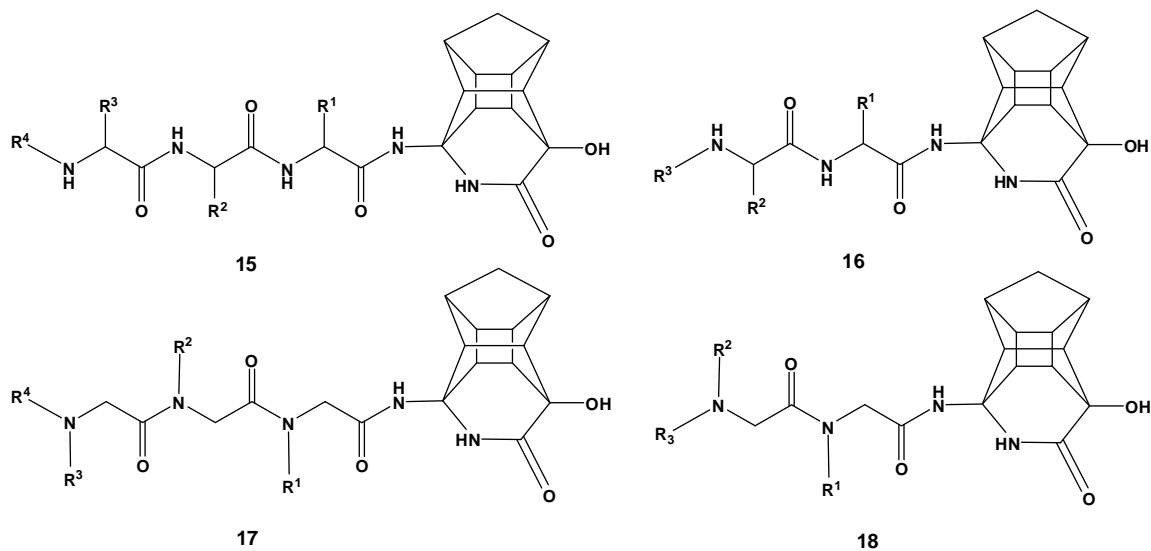


Table 1. Inhibition of wild type C-SA HIV-1 protease by PCU lactam peptide derivatives **15-16c** and peptoid derivatives **17-18c**.

Compounds	R ¹	R ²	R ³	R ⁴	IC ₅₀ (μM)	Yield (%)
15	benzyl	isopropyl	methyl	Cbz	0.75±0.035	52.5
16a	isopropyl	methyl	Cbz	-	0.6±0.071	76.2
16b	isopropyl	methyl	H	-	0.75±0.071	78.3
16c	isopropyl	methyl	2-pyrimidinyl-	-	0.5±0.035	43.2
17	benzyl	isopropyl	methyl	Cbz	0.5±0.035	46.2
18a	isopropyl	methyl	Cbz	-	10.0±0.71	65.3
18b	isopropyl	methyl	H	-	10.0±0.71	68.3
18c	isopropyl	methyl	2-pyrimidinyl-	-	5.0±1.06	38.3
	Atazanavir				0.004±0.0007	
	Lopinavir				0.025± 0.0014	

IC₅₀ = inhibitor concentration that inhibits 50 % of the protease activity, CBz = Benzyl oxy carbonyl.

The EASY-ROESY measurements required for structural analysis were recorded on a Bruker AVANCE III 600 MHz spectrometer using a BBO probe with z-gradient at a transmitter frequency of 600.1 MHz (spectral width, 8223.7 Hz; acquisition time, 0.1704436 s; 90° pulse width, 11.02 μ s; scans, 8; relaxation delay, 2.0 s, mixing time 0.12s).¹³ Processing and assignments were carried out using the Topspin 2.3 software from Bruker Karlsruhe. In order to investigate the effect of using DMSO-d₆ as a solvent had on the conformations, one of the compounds ROESY spectrum in D₂O was compared. These two spectra did not show any significant differences in the correlations and it was therefore decided to use DMSO-d₆ to dissolve all other NMR samples.

The ROESY spectra of compounds **16a** and **18a** confirmed the presence of more than one conformation. Based on the 2D spectra, it was suggested that **16a** had fewer conformations as compared to **18a**, see Appendix 2 Tables 1 and 2. The HSQC spectrum of compound **18a** revealed three different chemical shifts for each of the methylene protons (H-2' and H-6') of the peptoid chain. According to the ROESY spectra, the most prevalent conformations in both inhibitors are those with the valine methyl group's H-4' interacting with the cage atoms H-1 and H-3 (Figure 1). In the other conformation the H-4' atoms interacted with the cage amine NH_a (Figure 1). The weaker activity observed for **18a** was then attributed to the lack of a stable conformation that can interact with the catalytic site and the sub-site of the enzyme.

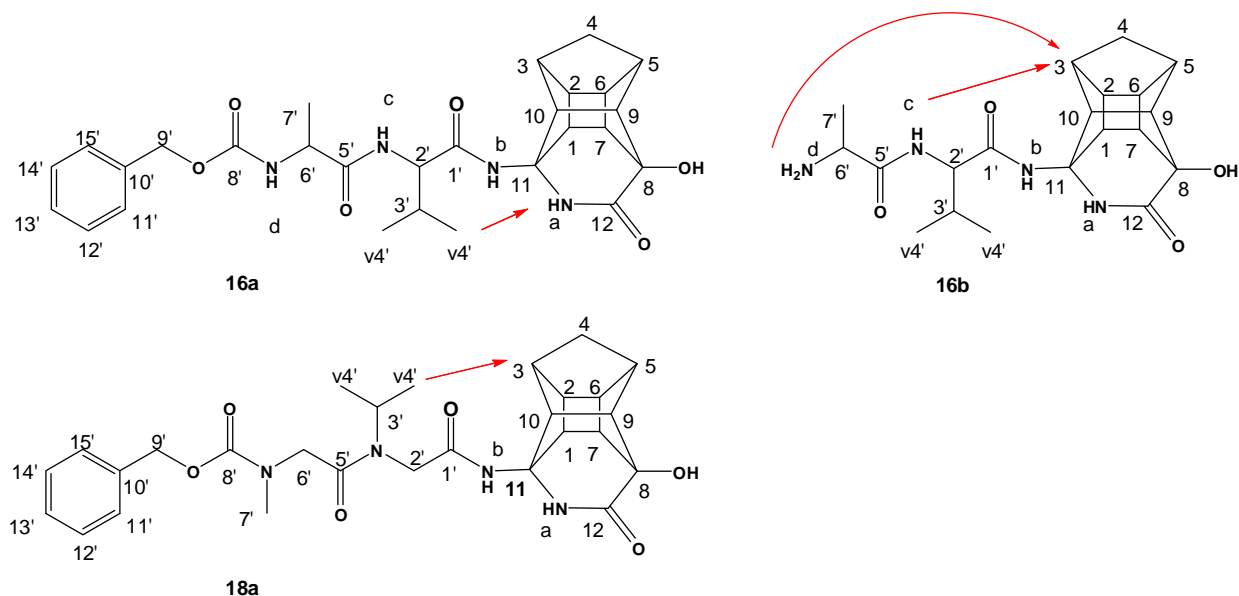


Figure 1. Inhibitor **16a**, **16b** and **18a** with arrows showing the EASY-ROESY correlations.

The ROESY spectrum of compound **16b** showed the interaction between the alanine amine NHd with the cage atoms H-1 and H-3 and the valine amine NHc with the H-3 (Figure 1). This suggests a cyclised conformation which allows only the cage and the valine to interact with the catalytic site and S₁ sub-site of the enzyme respectively, leaving the S₂ pocket unoccupied. Most of the clinically available inhibitors extend to either the S₂ or S₃ pocket of the enzyme for maximum binding.^{10,14} It was therefore concluded that for the PCU-lactam derived inhibitors, the P₃ substituent was necessary for efficient binding to the enzyme.

CONCLUSION

The coupling of the PCU-lactam to a series of peptides and peptoid chains resulted in compounds that inhibit 50 % of the HIV protease activity at good concentrations. The proper selection of residues for the peptoid was proven to be essential since the stability or formation of the well defined structure depends on the charge-charge repulsion caused by the residue. PCU-peptides/peptoid inhibitors that do not extend to either the P₂/P₃ subsite bind less efficiently to the enzyme. The EASY-ROESY technique enabled the determination of 3D structures of the inhibitors and correlate it to activity. Hence proper modifications of the residue in the inhibitors will be done in the future to obtain inhibitors with better activity. This is currently the only example of the use of a polycyclic cage framework to be employed as a transition state analogue and shows excellent promise to target other disease related proteases.

Supporting Information Available in Appendix 2: Detailed procedures for the synthesis of inhibitors; over-expression, extraction, purification and *in-vitro* activity screening of the C-SA HIV protease, Tables with NMR, mass, and microwave peptide synthesis conditions data; NMR and HRMS spectra.

REFERENCES

1. Adessi, C.; Soto, C. *Current Medicinal Chemistry* **2002**, *9*, 963-978.
2. Hruby, V. J.; Balse, P. M. *Current Medicinal Chemistry* **2000**, *7*, 945-970.
3. Miller, S. M.; Simon, R. J.; Ng, S.; Zuckermann, R. N.; Kerr, J. M.; Moos, W. H. *Bioorganic & Medicinal Chemistry Letters* **1994**, *4*, 2657-2662.
4. Miller, S. M.; Simon, R. J.; Ng, S.; Zuckermann, R. N.; Kerr, J. M.; Moos, W. H. *Drug Development Research* **1995**, *35*, 20-32.
5. Zuckermann, R. N.; Kerr, J. M.; Kent, S. B. H.; Moos, W. H. *Journal of the American Chemical Society* **1992**, *114*, 10646-10647.
6. Lee, T.; Laco, G. S.; Torbett, B. E.; Fox, H. S.; Lerner, D. L.; Elder, J. H.; Wong, C. H. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 939-944.
7. Marchand, A. P. *Advances in theoretically interesting molecules*; JAI Press: Greenwich CT, **1989**; Vol. 1.
8. Mosebi, S.; Morris, L.; Dirr, H. W.; Sayed, Y. *Journal of Virology* **2008**, *82*, 11476-11479.
9. Velaquez-Campoy, A.; Vega, S.; Fleming, E.; Bacha, U.; Sayed, Y.; Dirr, H. W.; Freire, E. *Aids Reviews* **2003**, *5*, 165-171.
10. Brik, A.; Wong, C. H. *Organic & Biomolecular Chemistry* **2003**, *1*, 5-14.
11. Huang, L.; Lee, A.; Ellman, J. A. *Journal of Medicinal Chemistry* **2002**, *45*, 676-684.
12. Fowler, S. A.; Blackwell, H. E. *Organic & Biomolecular Chemistry* **2009**, *7*, 1508-1524.
13. Thiele, C. M.; Petzold, K.; Schleucher, J. *Chemistry-A European Journal* **2009**, *15*, 585-588.
14. Wlodawer, A.; Vondrasek, J. *Annual Review of Biophysics and Biomolecular Structure* **1998**, *27*, 249-284.

CHAPTER 4

DESIGN, SYNTHESIS AND SCREENING OF PENTACYCLOUNDECANE-DIOL INHIBITORS TARGETING THE HIV PROTEASE ENZYME

INTRODUCTION

A strategy for inhibiting the HIV protease has been to prepare C_2 -symmetric inhibitors in the form of peptidomimetics. One family of active compounds are symmetric 1,2-diols flanked by short peptide units eg. A-75925 and DMP 450 (Figure 1).^{1,2} This family of inhibitors were designed after the hydroxyethylene isosteres provided the first highly potent inhibitors toward the HIV in a cell based assay.³⁻⁶ It was found that the contribution from the second hydroxyl group could be significant, as suggested by similar or better binding affinities of hydroxyethylene and dihydroxyethylene based transition state analogues.⁷ To date it has been proven using crystallography that diols form hydrogen bonds with the aspartic acid residues in the active site of the enzyme.^{8,9}

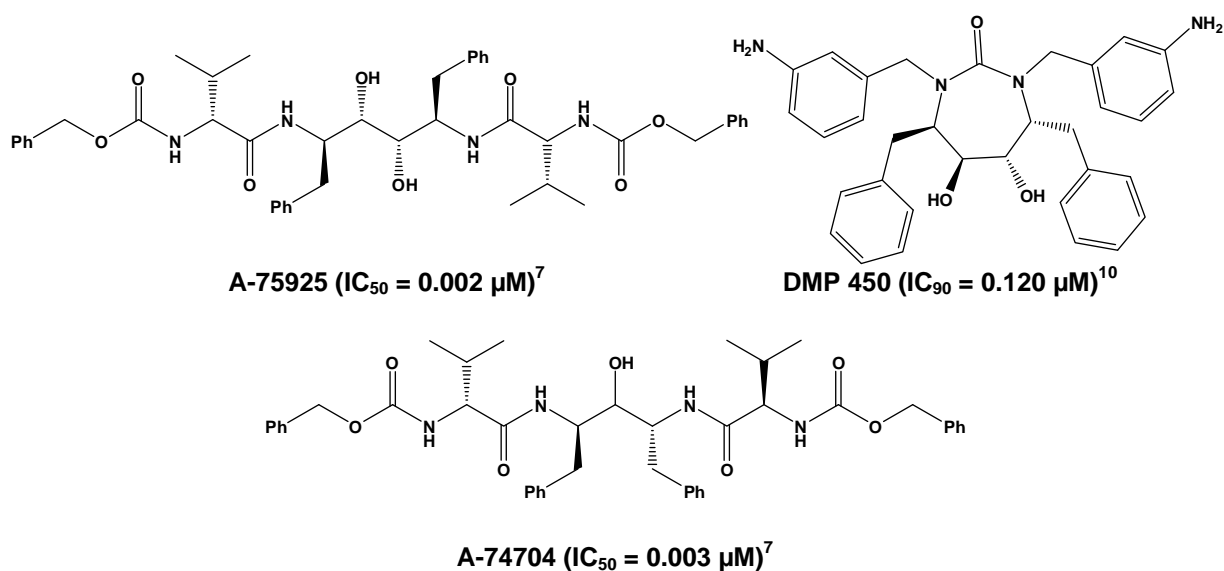


Figure 1. C_2 -symmetric 1,2-diol inhibitors.

In this study, the cage moiety was employed as a core molecule because of its 1,2-diol functional group (see Figure 2). Another advantage of using the PCU cage diol moiety would be to allow for the synthesis of symmetric inhibitors which are expected to exhibit higher specificity for the HIV protease over other aspartic proteases with less symmetric binding sites.

Rational for the design of these inhibitors

This study involved the reformulation of the ritonavir precursor A-75925 (Figure 1) by introducing well defined and systematic changes. These changes were introduced in view of first, substituting the core diols with cage diol moieties to increase stereospecificity and lipophilicity. The replacement of the P₃ position with groups that maintains the activity of the inhibitor after the protease had mutated. The binding affinity of the inhibitors to the mutated HIV protease enzyme have been reported to be affected by the inability of the P₃ group of the inhibitors to bind the enzyme S₃/S₃' subsite.¹¹ This was caused by a reduction in the size of the S₃ pocket in the mutated protease enzyme.¹¹ It was therefore crucial to design inhibitors with the appropriate P₃ groups. Taekyu Lee *et al.*¹¹ have reported alanine as an appropriate amino acid to use in the P₃ position. With this in mind all the inhibitors synthesized in this study have alanine in the P₃ position. Since the cage diol protease inhibitors were synthesized for the first time, the EASY ROESY was employed to elucidate the structural conformations and to get through space ¹H, ¹H distances/contacts.

Three different cage moieties were chosen for this investigation (Figure 2).

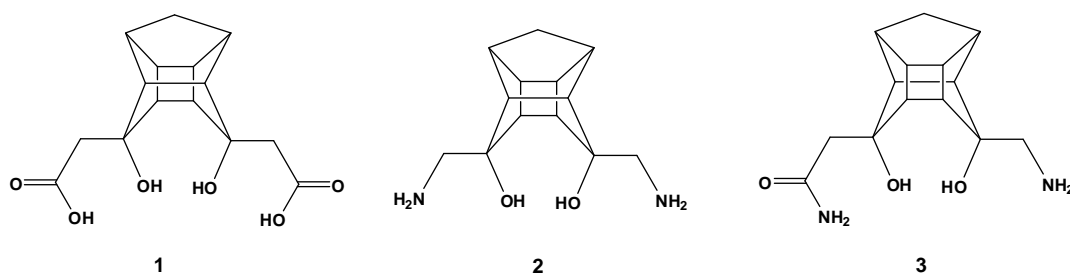


Figure 2. The structures of pentacycloundecane diol moieties.

Seven, inhibitors were designed namely **1a-1c**, **2a-2c** and **3a**. Inhibitors **1a-1c** consist of the cage diacid **1**¹² and inhibitors **2a-2c** and **3a** of the cage diamine **2** and cage amine **3** moieties, respectively. Cages **2** and **3** were synthesized for the first time in this study. All inhibitors except **1c** were designed such that the diol cage would potentially fit into the catalytic site of the enzyme while phenylalanine, valine, alanine, alanine/carbobenzyloxy parts occupy the S₁, S₂, S₃ and S₄ pockets, respectively (see Figure 3).

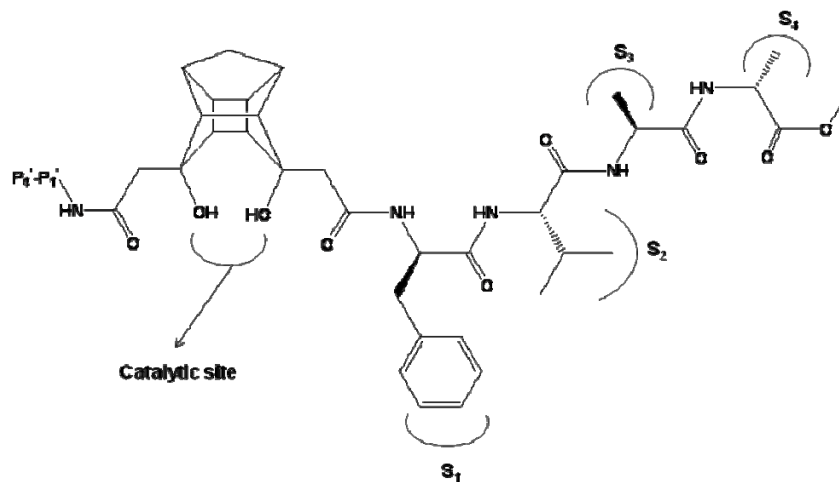


Figure 3. Schematic representation of how the cage diol inhibitor may bind to the HIV protease enzyme. [Standard nomenclature $P_1 \text{ --- } P_n$, $P_1' \text{ --- } P_n'$ is used to designate amino acid residues of peptide substrates. The corresponding binding sites on the protease are referred to as $S_1 \text{ --- } S_n$, $S_1' \text{ --- } S_n'$ subsites].¹³

Compound **1c** was designed to investigate the ability of the cage to fit into the S_1 pocket of the enzyme. The S_1 pocket has been observed to prefer binding to bulky groups.² In the absence of the phenylalanine compound **1c** will allow the investigation of the ability of the cage to bind to this site.

Cage **2** was synthesized to enable the coupling of the exact sequence (Phe-Val-Ala-Cbz) used by Lee *et al.*¹¹ and to investigate the effect of introducing the cage to this sequence. Inhibitor **2a** derived from cage **2** was designed to also allow the investigation of the effect that the positioning of the carbonyl next to the scissile bond would have on binding. On cage **1** the carbonyl is three bonds away from the diols while on cage **2** the carbonyl is four bonds away. The proposed mechanism for the binding of HIV protease to the inhibitors illustrated the significance of the efficient binding of the carbonyl after scissile bond to the flaps.

Pearl *et al.* have suggested that upon binding of the inhibitor to the catalytic site, the scissile bond is distorted. This reduces the double bond character in the C–N bond and polarizes the carbonyl, rendering it more electrophilic towards the incoming nucleophilic water molecule. The distortion of the amide bond is stabilized by interactions of the substrate with the extended binding cleft, mostly by hydrogen bond interactions between the interior side of the flap and the substrate carbonyls on either side of the scissile bond as shown in Figure 4.²

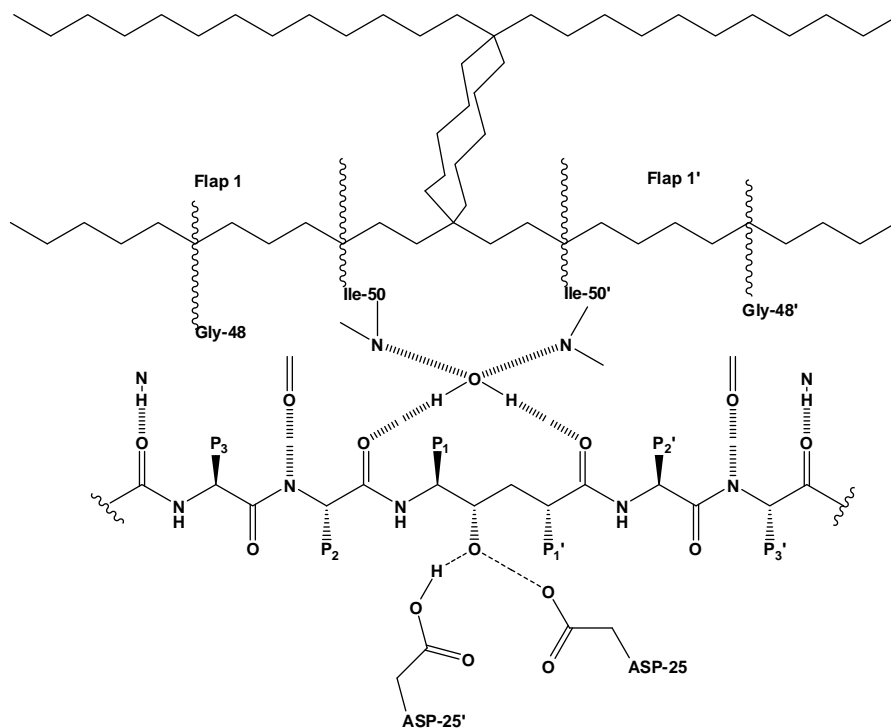
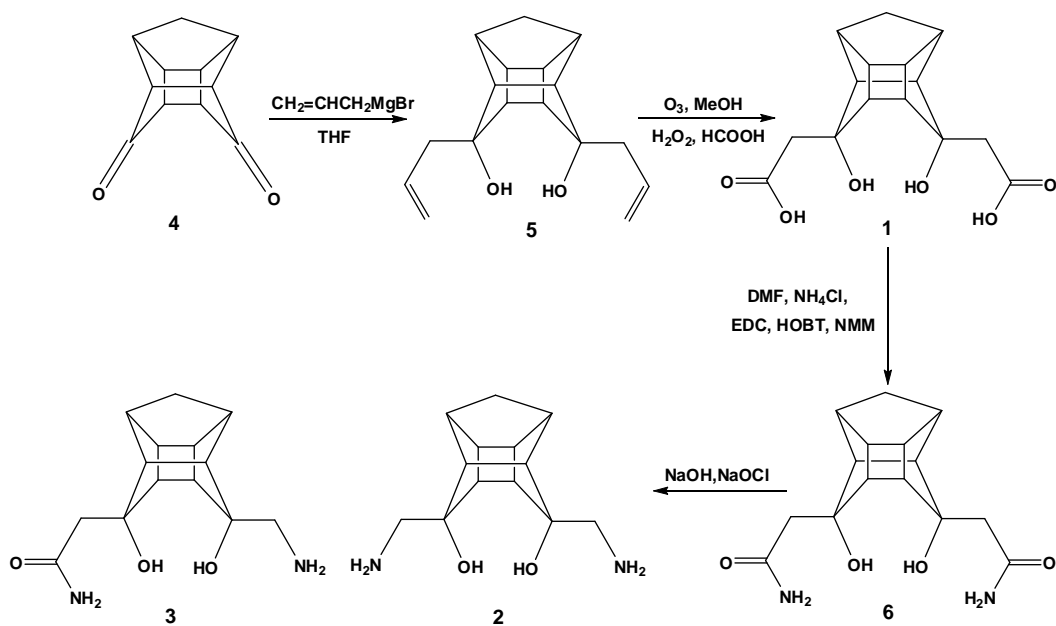


Figure 4. Schematic representation of two flaps of HIV PR and their hydrogen bonds with the water molecule.²

Inhibitor **3a** was designed for better exposure of the core diols for efficient binding to the catalytic site. An X-ray crystallography of the diacid coupled to one amino acid showed that one arm of the cage is positioned over the diols which can potentially cause steric hinderance and prevent efficient binding of the diols to the catalytic site.¹⁴

RESULTS AND DISCUSSION

The cage diacid **1** was the precursor in the synthesis of the novel cage moieties **2** and **3**. The cage acid was synthesized as illustrated in Appendix 3. Cage **2** and **3** were obtained from the Hofmann rearrangement of the diamide (Scheme 1) which was obtained by reacting the diacid with ammonium chloride (NH_4Cl), 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC), 1-hydroxybenzotriazole (HOBT) and N-methylmorpholine (NMM) in dimethylformamide (DMF) (see Appendix 3).



Scheme 1. Synthesis of cage diol from the cage dione.¹²

The peptides to fit into the S_1/S_1' - S_4/S_4' pockets were synthesized on an automated peptide synthesizer on 2-chlorotrityl chloride resin (see Appendix 3).

The peptides to be coupled to the cage **1** moiety were cleaved from the resin and the acid terminal converted to a methyl ester before coupling the amine to the cage diacid. Different coupling reagents were evaluated for the formation of the amide bond between the cage acid and the peptide amines. Firstly, the diacid was converted to an acid chloride and coupled with the amine terminal of the peptide in presence of a base. Secondly, the cage diacid and peptide were coupled in the presence of either *O*-benzotriazole-*N,N,N',N'*-tetramethyl-uronium-hexafluoro-phosphate/*N,N*-diisopropylethylamine (HBTU/DIPEA) or 2-(7-aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU)/DIPEA or EDC/NMM/HOBT. The acid chloride route afforded the product in low yield with too many impurities. The HBTU/DIPEA coupling reagents afforded the product with only one side coupled. The HATU/DIPEA afforded the product in low yield with impurities. The EDC/NMM/HOBT resulted in good yields of **1a** – **1c** with few impurities.

Coupling of the peptides to the cage diamine **2** and cage amine **3** to the peptide acid group were evaluated with three different coupling conditions that included HBTU/DIPEA, HATU/DIPEA and EDC/NMM/HOBT. The latter set gave good yields with fewer impurities. All the synthesized compounds were purified using semi-preparative HPLC carried out on a Shimadzu Instrument (Ace C18 150 mm x 21.2 mm x 5 microns) with a UV/VIS detector (215 nm). The peptides were characterized

using NMR Bruker AVANCE III 400 MHz and 600 MHz spectrometer and High Resolution Mass Spectroscopic analysis performed on a Bruker MicroTOF QII.

***In vitro* HIV-1 protease activity**

Synthesis and purification of the HIV protease type C were performed as described in the Appendix 3. The catalytic activity of the protease was monitored following the hydrolysis of a chromogenic peptide substrate, Ala-Arg-Val-Nle-*p*-nitro-Phe-Glu-Ala-Nle-NH₂ at 300 nm using an Fpecord 210 spectrophotometer.¹⁵⁻¹⁷ Activity was standardized using commercially available drugs Atazanavir and Lopinavir.

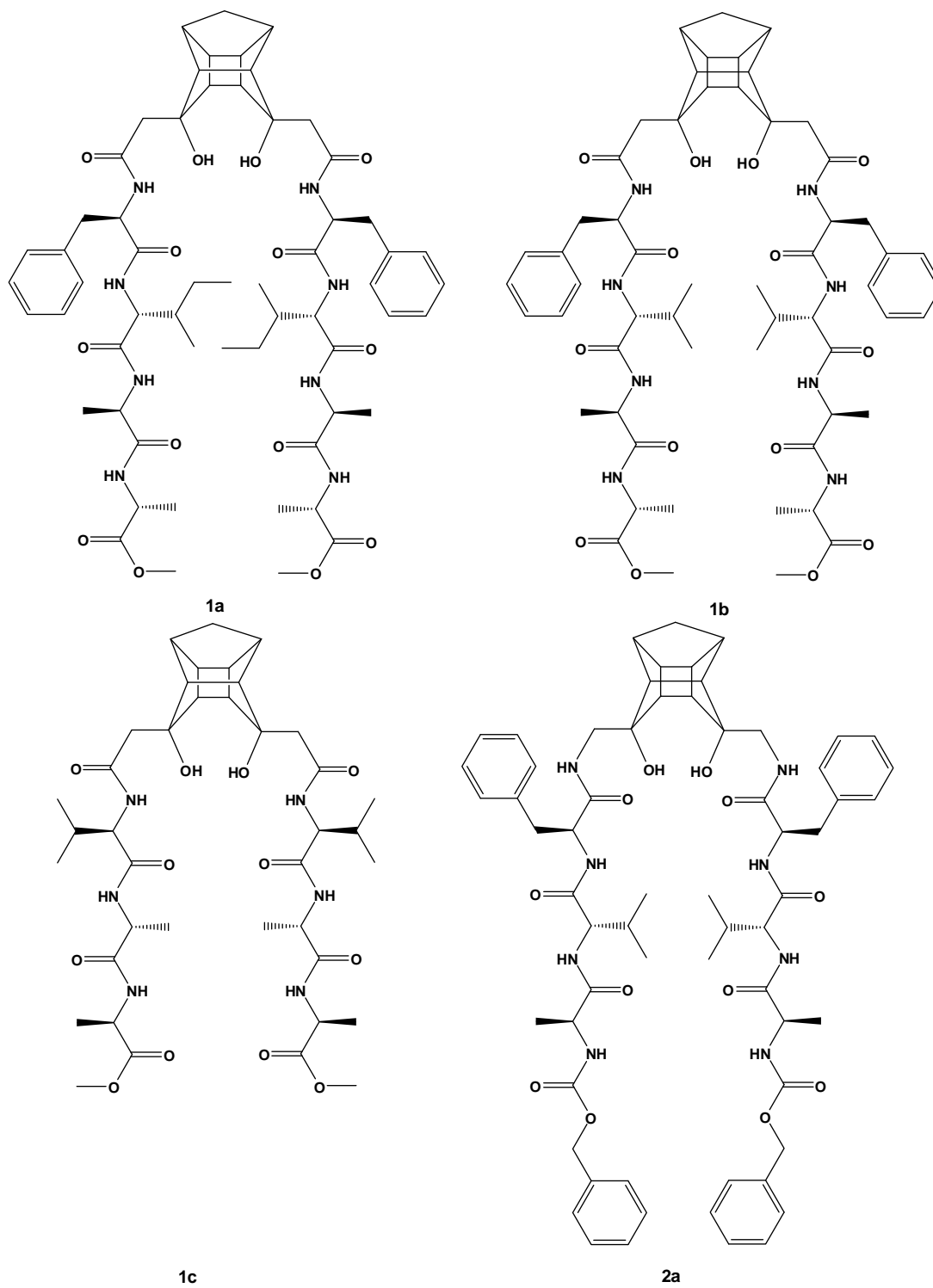


Figure 5. Structures of synthesized inhibitors.

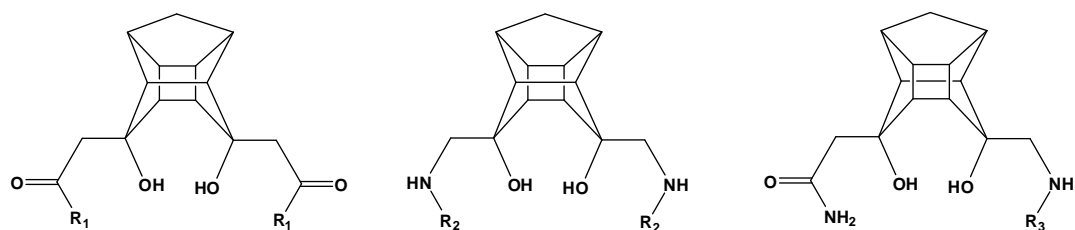


Table 1. Inhibition of wild type C-SA HIV protease by the PCU peptide derivatives.

Compound	R ₁	IC ₅₀ (μM)	Yield (%)	Log P
1a	Phe-Ile-Ala-Ala	5	38	2.59 ± 0.92
1b	Phe-Val-Ala-Ala	5	59	3.65±/ - 0.92
1c	Val-Ala-Ala	ND (> 10)	48	-0.56±/ - 0.85
	R ₂			
2a	Phe-Val-Ala-Cbz	7.5	19	6.47 ± 0.96
2b	Phe-Val-Ala-((2-pyrimidinylthio)acetic acid)	0.5	22	3.80±/ - 0.93
2c	Phe-Val-Ala-NH ₂	0.5	17	1.68 ± 0.89
	R ₃			
3a	Phe-Val-Ala-Cbz	0.6	44	1.67±/ - 0.85

All of the designed inhibitors were synthesized and evaluated for their activity against the wild type C-SA protease enzyme. According to the results in table 1, inhibitors **1a-1c** did not show considerable activity. Inhibitor **1c** showed the least activity and this was initially attributed to the absence of the phenylalanine ring which must interact with the S₁/S₁' subsites for efficient binding. Further investigations of the structural conformations of **1c** and **1b** using NMR enabled us to explain the activity observed. From the ROESY spectrum of **1c** it was noted that one of the valine methyl H-6' or H7' interacted with cage protons H-9/10 (labelled structure, Figure 7).

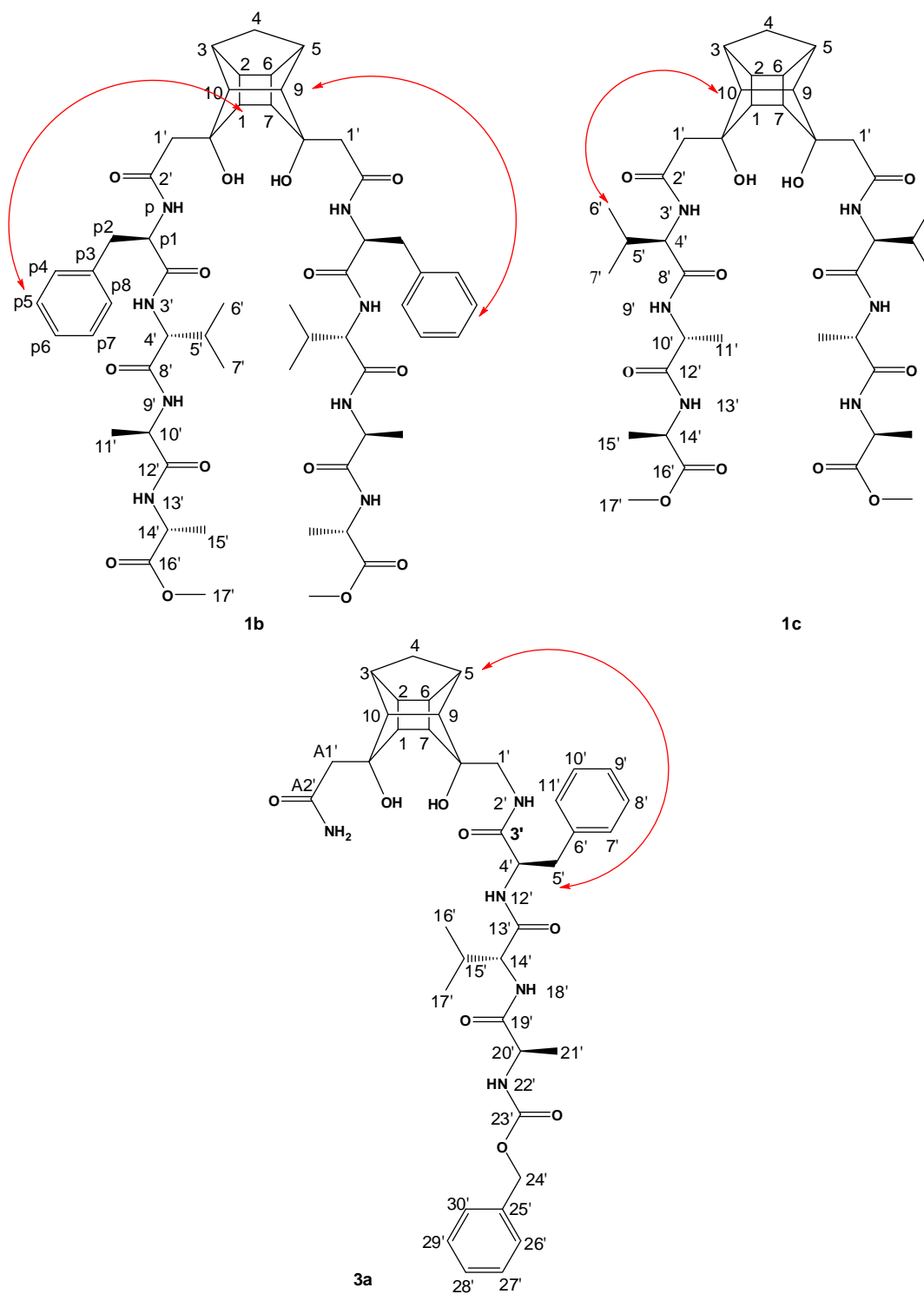


Figure 7. Numbered structures of inhibitor **1b**, **1c** and **3a** with arrows indicating the long range correlations observed in the EASY-ROESY spectra.

This arrangement of the cage to the peptide obstructs the diols, shielding them from the protease catalytic site. The ROESY spectrum of **1b** showed the presence of two conformations. In one conformation the phenyl ring (H-p4-H-p8) interacted with cage protons H-1/7 which are situated at the back of the cage. This conformation therefore allowed for the diols to be exposed to interact with the catalytic site. In the second conformation the phenyl ring (H-p4-H-p8) interacted with H-9/10 thus the peptide chain extended to the front and obstructed the diols. It was then concluded that the phenyl ring was important because it allowed the inhibitor to exist in a conformation that exposes the diols. All other inhibitors were then designed with the phenylalanine in the S₁ subsite.

The diamine inhibitor **2a** did not have good solubility in the pH 5 sodium acetate buffer that was used for the protease assay and thus the high IC₅₀ observed was attributed to low solubility. The benzyl group was then substituted with a (2-pyrimidinylthio) acetic acid moiety (**2b**). The IC₅₀ improved from 7.5 μM to 0.5 μM. It was then decided to remove the benzyl group completely to reduce the size and use the free amine group on the alanine to increase solubility. The IC₅₀ obtained for **2c** was the same as the one of **2b**. The diamine derivatives exhibited better activity when compared to the diacid inhibitors. Although inhibitor **3a** was synthesized with the benzyl group furthest away from the cage, like in **2a**, it showed good activity because of the free amide which increased its solubility. The EASY-ROESY spectrum of inhibitor **3a** showed a long range correlation between the methylene protons of phenylalanine H-5' and the cage protons H-5 (Figure 7). This arrangement of the PCU-peptide allowed the chain to extend away from the diols thereby exposing them for binding to the enzyme catalytic site.

CONCLUSION

Peptides derived from the cage diol moieties can inhibit the catalytic activity of the wild type C-SA HIV protease enzyme. The PCU-diol diacid derived inhibitors exhibited less activity when compared to the inhibitors derived from the PCU-diol diamine and PCU-diol amino amide moieties. The phenylalanine was observed to be an appropriate residue for the P₁ position of the inhibitors. Inhibitors with less bulky amino acid groups at the P₁ position were proven to be less active. Large inhibitors that were insoluble in the buffer exhibited low inhibitory activity when compared to more soluble compounds. The EASY-ROESY NMR technique enabled the determination of the 3D structures of the inhibitors. Using the predicted structural arrangement it was possible to explain the activity observed since it depended on the exposure of the diols which are crucial for binding to the catalytic site of the enzyme. The combination of drug design with *in-vitro* assays and NMR techniques has enabled the rationalization of the observed HIV-PR inhibition for the different cage peptides. Computational studies are currently being conducted to understand the interaction modes of these inhibitor with the enzyme.

Supporting Information Available in Appendix 3: Detailed procedures for the synthesis of inhibitors; over-expression, extraction, purification and *in-vitro* activity screening of the C-SA HIV protease, Tables with NMR, mass, and microwave peptide synthesis conditions data; NMR and HRMS spectra.

REFERENCES

1. Wlodawer, A.; Vondrasek, J. *Annual Review of Biophysics and Biomolecular Structure* **1998**, *27*, 249-284.
2. Brik, A.; Wong, C. H. *Organic & Biomolecular Chemistry* **2003**, *1*, 5-14.
3. Dreyer, G. B.; Metcalf, B. W.; Tomaszek, T. A.; Carr, T. J.; Chandler, A. C.; Hyland, L.; Fakhoury, S. A.; Magaard, V. W.; Moore, M. L.; Strickler, J. E.; Debouck, C.; Meek, T. D. *Proceedings of the National Academy of Sciences of the United States of America* **1989**, *86*, 9752-9756.
4. Meek, T. D.; Lambert, D. M.; Dreyer, G. B.; Carr, T. J.; Tomaszek, T. A.; Moore, M. L.; Strickler, J. E.; Debouck, C.; Hyland, L. J.; Matthews, T. J.; Metcalf, B. W.; Petteway, S. R. *Nature* **1990**, *343*, 90-92.
5. Dreyer, G. B.; Lambert, D. M.; Meek, T. D.; Carr, T. J.; Tomaszek, T. A.; Fernandez, A. V.; Bartus, H.; Cacciavillani, E.; Hassell, A. M.; Minnich, M.; Petteway, S. R.; Metcalf, B. W.; Lewis, M. *Biochemistry* **1992**, *31*, 6646-6659.
6. Kempf, D. J.; Codacovi, L.; Wang, X. C.; Kohlbrenner, W. E.; Wideburg, N. E.; Saldivar, A.; Vasavanonda, S.; Marsh, K. C.; Bryant, P.; Sham, H. L.; Green, B. E.; Betebenner, D. A.; Erickson, J.; Norbeck, D. W. *Journal of Medicinal Chemistry* **1993**, *36*, 320-330.
7. Kempf, D. J.; Sham, H. L. *Current Pharmaceutical Design* **1996**, *2*, 225-246.
8. Tomasselli, A. G.; Thaisrivongs, S.; Heinrikson, R. L. *Advances in Antiviral Drug Design* **1996**, 173-228.
9. Muhlman, A.; Lindberg, J.; Classon, B.; Unge, T.; Hallberg, A.; Samuelsson, B. *Journal of Medicinal Chemistry* **2001**, *44*, 3407-3416.
10. Jadhav, P. K.; Ala, P.; Woerner, F. J.; Chang, C. H.; Garber, S. S.; Anton, E. D.; Bacheler, L. T. *Journal of Medicinal Chemistry* **1997**, *40*, 181-191.
11. Lee, T.; Laco, G. S.; Torbett, B. E.; Fox, H. S.; Lerner, D. L.; Elder, J. H.; Wong, C. H. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 939-944.
12. Onajole, O. K.; Makatini, M. M.; Govender, P.; Govender, T.; Maguire, G. E. M.; Kruger, H. G. *Magnetic Resonance in Chemistry* **2010**, *48*, 249-255.
13. Reetz, M. T.; Merk, C.; Mehler, G. *Chemical Communications* **1998**, 2075-2076.
14. Boyle, G. A.; Govender, T.; Karpoornath, R.; Kruger, H. G. *Acta Crystallographica Section E-Structure Reports Online* **2007**, *63*, O4797-U5875.
15. Mosebi, S.; Morris, L.; Dirr, H. W.; Sayed, Y. *Journal of Virology* **2008**, *82*, 11476-11479.
16. Tie, Y. F.; Boross, P. I.; Wang, Y. F.; Gaddis, L.; Liu, F. L.; Chen, X. F.; Tozser, J.; Harrison, R. W.; Weber, I. T. *Febs Journal* **2005**, *272*, 5265-5277.
17. Klabe, R. M.; Bacheler, L. T.; Ala, P. J.; Erickson-Viitanen, S.; Meek, J. L. *Biochemistry* **1998**, *37*, 8735-8742.

CHAPTER 5

SYNTHESIS AND SCREENING OF PENTACYCLOUNDECANE-PEPTOIDS AS POTENT HIV PROTEASE INHIBITORS

INTRODUCTION

In the previous chapters we have reported PCU-peptide/peptoid derived HIV protease inhibitors. The activities of these inhibitors were related to their structural conformations with the aid of NMR and computational studies (Chapters 2, 3 and 4). The activity was also noted to be dependent on the chirality of the cage moiety (Chapter 2). For this study the cage diol moieties were coupled to peptoids. The peptoid side chains were selected to mimic the peptide, phenylalanine-valine-alanine-Cbz (FVA-Cbz), as in Chapter 3. This peptide sequence has been reported to have a high binding affinity for the HIV-protease subsites.^{1,2}

Herein, five novel pentacycloundecane-peptoid inhibitors that consist of the cage diol moieties coupled to a peptoid chain are presented.

Rational for the design of these inhibitors

Five inhibitors were designed, namely **1a**, **2a**, **2b**, **3a** and **3b** (Figure 1 and Figure 2). Inhibitor **1a** consists of the cage diacid **1**³ and inhibitors **2a**, **2b** and **3a**, **3b** of the cage diamine **2** and cage amine **3** moieties (see Figure 2 in Chapter 4 for cage moieties), respectively.

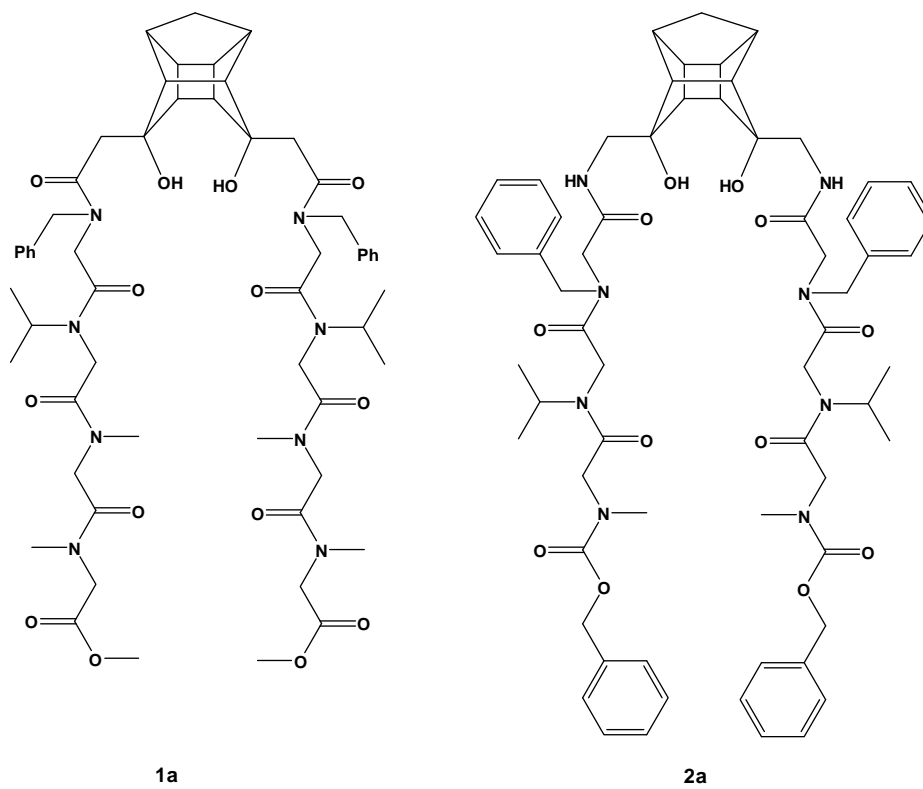


Figure 1. Structures of synthesized inhibitors.

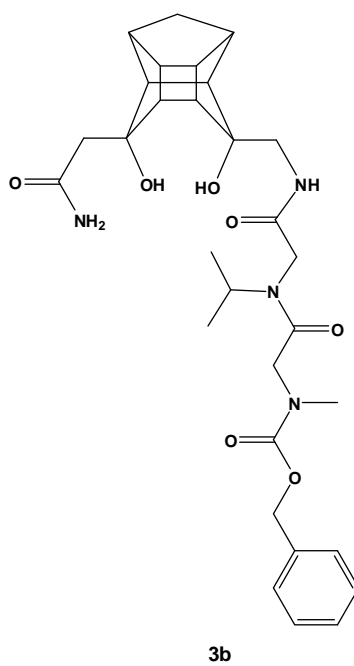
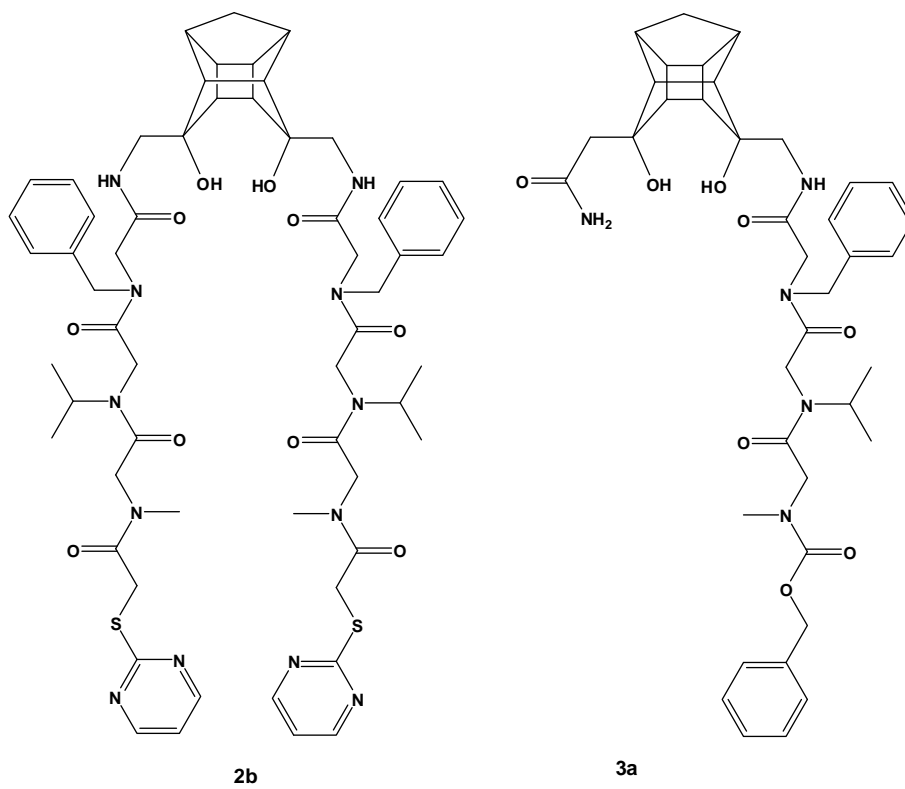


Figure 2 continuation of Figure 1. Structures of synthesized inhibitors.

Inhibitor **1a** was designed such that the diol cage groups would potentially fit into the catalytic site of the enzyme while the benzyl, isopropyl, methyl and methyl/ carbobenzyloxy groups occupied the S_1, S_2, S_3 and S_4 pockets, respectively (see Figure 3).

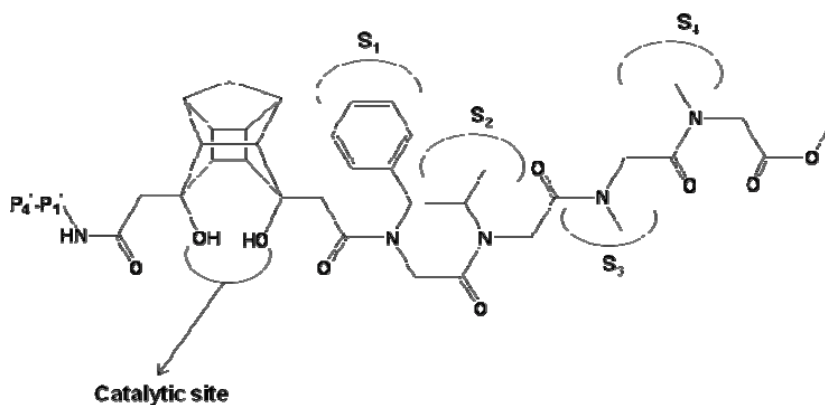


Figure 3. Schematic representation of the proposed interaction between the cage diol and the HIV protease enzyme. [Standard nomenclature P_1 ___ P_n , P_1' ___ P_n' is used to designate amino acid residues of peptide substrates. The corresponding binding sites on the protease are referred to as S_1 ___ S_n , S_1' ___ S_n' subsites according to literature].⁴

Cage **2** was synthesized to enable the coupling of the peptoid version of the sequence (FVA-Cbz) reported by Lee *et al.*¹ and to investigate the effect of introducing the cage to this sequence. Inhibitor **2a** derived from cage **2** enabled the investigation of the effect that the positioning of the carbonyl next to the scissile bond would have on binding. On cage **1** the carbonyl is three bonds away from the diols while on cage **2** the carbonyl is four bonds away. The proposed mechanism for the binding of HIV protease to inhibitors illustrated the importance of the positioning of the carbonyl after scissile for efficient binding with the enzyme flaps.⁵

The cage inhibitor **3a** was designed for better exposure of the core diols for efficient binding to the catalytic site. An X-ray crystallography of the diacid coupled to one amino acid showed that one arm of the cage is positioned over the diols and can cause steric hinderance and prevent efficient binding of the diols to the catalytic site, Figure 4.⁶

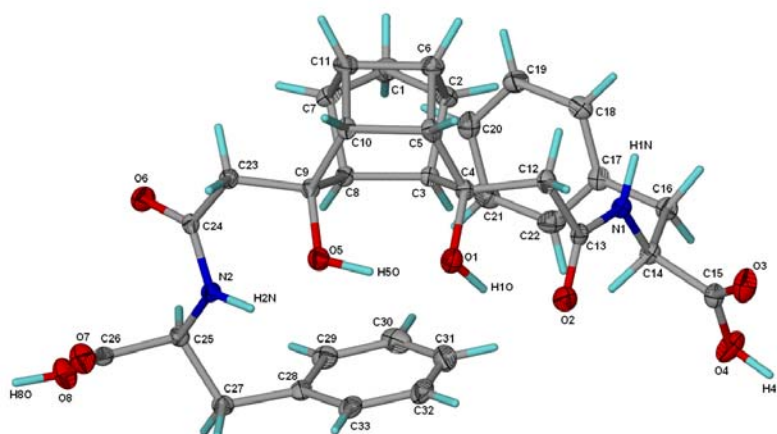


Figure 4. Crystal structure of the PCU diol diacid **1** coupled to phenylalanine.⁶

RESULTS AND DISCUSSION

All cage moieties were synthesized according to previously reported procedures (Chapter 4, Scheme 1, Appendix 4).

The peptoids to be coupled to the cage diacid **1** were synthesized according to procedures described in Chapter 3, cleaved from the resin and the acid terminal converted to a methyl ester before coupling the amine terminal to the cage. The peptoids synthesized from cages **2** and **3** were coupled directly after cleavage without any further modification using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC), 1-hydroxybenzotriazole (HOBT) and *N*-methylmorpholine (NMM) (see Appendix 4 and chapter 4).

All of the designed inhibitors (Figure 1 and Figure 2) were successfully synthesized and tested against the wild type C-SA protease enzyme, see Table 1.

Table 1. Inhibition of wild type C-SA HIV protease by the PCU diol peptoid derivatives, yields for the coupling of peptoids to the cage diols and the size as confirmed by high-resolution mass spectroscopy.

Compound	IC ₅₀ (μM)	Yield (%)	Mass calculated	Mass found
1a	6.5 ±0.35	19	1121.5530 C ₅₇ H ₇₈ N ₈ NaO ₁₄	1121.5530
2a	0.075±0.0035	30	1139.5812 C ₆₃ H ₇₉ N ₈ O ₁₂	1139.5780
2b	1.0±0.17	39	1175.5165 C ₅₉ H ₇₅ N ₁₂ O ₁₀ S ₂	1175.5159
3a	0.5±0.18	22	716.3654 C ₃₉ H ₅₀ N ₅ O ₈	716.3654
3b	0.75±0.071	33	591.2789 C ₃₀ H ₄₀ N ₄ NaO ₇	591.2787

According to the results in Table 1 inhibitor **1a** did not show good activity (IC₅₀ = 6.5 μM) when compared to the cage **2** and cage **3** peptoid derivatives. Thus other cage **1** derivatives were not synthesized. Inhibitor **2a** showed very good activity (IC₅₀ = 0.075 μM) and was derivatised by substituting the carbobenzyloxy group with the (2-pyrimidinylthio)acetic acid group in order to produce compound **2b** with an improved solubility. Inhibitor **2b** exhibited significantly less binding affinity to the enzyme (IC₅₀ = 1.0 μM). The same peptoid sequence used in **2a** was then coupled to cage **3** (**3a**) to investigate if the activity would increase when the diols were less sterically hindered. Inhibitor **3a** exhibited less activity than **2a**. This decrease was attributed to the presence of the free amide which could be binding to the catalytic site or form a hydrogen bond with one of the diols. If the primary amide was to interact with the catalytic site, the cage moiety would bind to the S₁ subsite which can accommodate bulky hydrophobic groups. The next substituent on the peptoid (benzyl group) will now have to fit into the S₂ subsite. This subsite does not accommodate bulky hydrophobic moieties therefore the benzyl group was removed to synthesize **3b** in order to investigate the assumption. Inhibitor **3b** showed a decrease in activity when compared to **3a**, prompting the full analysis of **2a** via NMR while further computational studies are being conducted to investigate the effect of hydrogen bonding between the amide and the diols and to also understand the interaction modes of the PCU-diol peptoids with the HIV protease. With this set of compounds it was not possible to suggest a structure activity relationship like in previous chapters. Peptoids have been reported to be highly flexible compounds (chapter 3) that can exist in many conformations.⁷ Thus relating their structure to activity would not be straight forward. It

requires assistance from computational studies. The good activity observed for inhibitor **2a** can be attributed to the presence of four benzyl groups. In peptoid aromatic groups on the side chains cause charge-charge repulsion with backbone carbonyls that assist in peptoid folding. This folding can be advantageous since it gives the peptoid a well-defined secondary structure. It was therefore suggested that **2a** does not have many conformations.

NMR studies

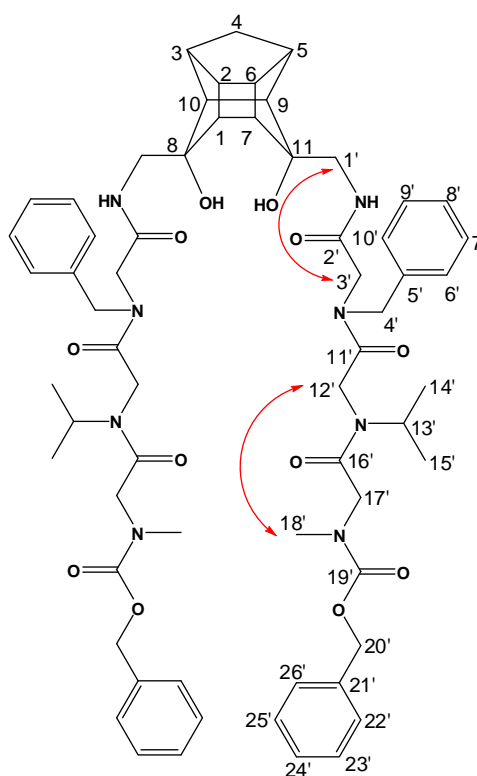


Figure 5. Numbered structure of inhibitor **2a** with arrows indicating the long range correlations observed in the EASY-ROESY spectrum.

Table 2. ^1H chemical shift and EASY-ROESY correlations of inhibitor 2a. The chemical shift was referenced to 2.50 ppm and the long-range interactions through space, measured by ROESY, are marked in red.

Atom	$\delta^1\text{H}^{\text{a,b}}$	ROESY ($\delta^1\text{H}^{\text{a,b}}$)
1/7	2.09-2.63	4a, 4s
2/6		
3/5		
9/10		
8/11		
4a	1.44	3/5, 4s
4s	1.78	3/5, 4a
1'	3.41	NH, 3'
NH	8.29	1', 3'
2'	-	-
3'	4.03	1', 4', 6'-10', 12', 14', 15', NH
4'	4.46	3', 6'-10'
5'	-	-
6'-10'	7.23-7.38	3', 4'
11'	-	-
12'	4.22	3', 13', 14', 15', 18'
13'	4.66	12', 14', 15'
14'	0.85	3', 12', 13'
15'	0.99	3', 12', 13'
16'	-	-
17'	4.66	
18'	2.83	12'
19'	-	-
20'	5.02, 5.08	22'-26'
21'	-	-
22'-26'	7.23-7.38	20'

a. Solvent $(\text{CD}_3)_2\text{SO}$.

b. 600 MHz for ^1H .

According to the EASY-ROESY correlations in Table 2, the peptoid chain did not exist in an extended conformation. The H-3' proton showed correlations with many groups including those at the end of the chain. Another interesting observation was noted in the HSQC spectrum, the C-4 signals of the cage diol are deshielded compared to the reported chemical shift of these carbons.³ The reported chemical shift ranged from 33.5 to 34.8 ppm and the observed chemical shift is 43.1 ppm. This downfield shift could be attributed to through space deshielding caused by electron rich atoms close to the C-4 as reported previously.^{3,8} The presence of electron rich groups around C-4 could not be proven using the ROESY spectra since the correlations observed were only between the peptoid side chain atoms and not between the cage and the peptoid. The HSQC spectrum of cage **2** coupled to *t*-BOC anhydride was recorded. The C-4 signal was observed at 33.8 ppm which proved that it was the presence of the peptoid chain that caused the deshielding. The other interesting point to note in Table 2 is that the cage and the peptoid protons did not show splitting which was observed from the PCU-lactam peptoids in chapter 3. This confirms that the four benzyl groups stabilized the secondary structure to a single conformation. Currently the hybrid method (quantum mechanics/molecular mechanics (QM/MM) molecular dynamics (MD) is being conducted to obtain the lowest energy conformations. This will enable us to fully predict the 3D structure of compound **2a**.

CONCLUSION

PCU-Peptoid derivatives can inhibit the catalytic activity of the wild type C-SA HIV protease enzyme. The best candidate inhibited 50 % of the enzyme activity at a concentration of 75 nM. Structural modification of this inhibitor with the best activity resulted in a decrease in activity. The EASY-ROESY technique provided vital information about the 3D structure of the inhibitor which was related to the biological activity. Aromatic residues in a peptoid chain stabilize the secondary structure of the inhibitor. Further studies involving computational are currently being conducted to determine the correct structural changes to be made to improve the activity.

Supporting Information Available in Appendix 4: Detailed procedures for the synthesis of inhibitors; over-expression, extraction, purification and *in-vitro* activity screening of the C-SA HIV protease, Tables with NMR, mass, and microwave peptide synthesis conditions data; NMR and HRMS spectra.

REFERENCES

1. Lee, T.; Laco, G. S.; Torbett, B. E.; Fox, H. S.; Lerner, D. L.; Elder, J. H.; Wong, C. H. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 939-944.
2. Wlodawer, A.; Vondrasek, J. *Annual Review of Biophysics and Biomolecular Structure* **1998**, *27*, 249-284.
3. Onajole, O. K.; Makatini, M. M.; Govender, P.; Govender, T.; Maguire, G. E. M.; Kruger, H. G. *Magnetic Resonance in Chemistry* **2010**, *48*, 249-255.
4. Reetz, M. T.; Merk, C.; Mehler, G. *Chemical Communications* **1998**, 2075-2076.
5. Brik, A.; Wong, C. H. *Organic & Biomolecular Chemistry* **2003**, *1*, 5-14.
6. Boyle, G. A.; Govender, T.; Karpoomath, R.; Kruger, H. G. *Acta Crystallographica Section E-Structure Reports Online* **2007**, *63*, O4797-U5875.
7. Fowler, S. A.; Blackwell, H. E. *Organic & Biomolecular Chemistry* **2009**, *7*, 1508-1524.
8. Kruger, H. G.; Ramdhani, R. *South African Journal of Chemistry* **2006**, *59*, 71-U28.

CHAPTER 6

SUMMARY AND CONCLUSION

To date there is an urgent demand for new drugs that can inhibit the catalytic activity of the HIV protease enzyme. Currently available HIV protease inhibitors suffer from low oral bioavailability, rapid metabolism, unspecific protein binding and increased drug-resistant as a result of mutations. Thus, in this study inhibitors that incorporate the cage moieties which can improve the pharmacokinetic and pharmacodynamic properties of molecules were designed. Polycyclic cage compounds are useful scaffolds that yield drugs with a wide scope of applications. Although cages have non-conventional structures their highly rigid nature induces receptor site specificity, especially for lipophilic regions. The hydrophobic cage skeleton improves transport of drugs across the cell membrane thus increasing their affinity for lipophilic regions in the receptor molecules and enabling the drug to cross the blood-brain barrier of the central nervous system (CNS). Four different cage molecules namely the PCU-lactam, PCU-diol diacid PCU-diol diamine and PCU-diol amino amide were chosen for this study. The cage moieties were coupled to peptide and peptoid chain sequences that fit the protease subsites and this resulted in a series of novel HIV inhibitors.

PCU-LACTAM INHIBITORS

The PCU-lactam was coupled to three different classes of sequences. The first class was made up of the EAIS amino acids, the second of the FVA-Cbz and the third of the peptoid chain that mimics the FVA-Cbz amino acids sequence. The EAIS sequence yielded the best inhibitor in this family with an IC_{50} value of 0.078 μ M. Substitution or removal of amino acids from this sequence resulted in a decrease in activity. The FVA-Cbz derived inhibitors exhibited moderate inhibition ($IC_{50} \leq 0.75\mu$ M). The peptoid compounds exhibited poor activity ($IC_{50} \geq 10\mu$ M) and this was attributed to a lack of stable conformations which was deduced from NMR (Figure 1).

Protection of the cage hydroxyl group, which is proposed to be a norstatine type isosteric bond and the results obtained from computational studies confirmed that these inhibitors bind to the catalytic site via this group. The inhibitory activity and structural conformations were proven to be dependent on the cage chirality. All of the active inhibitors from the EAIS class exhibited a correlation between the peptide chain alanine residue and the H-3, H-10 and H-1 cage protons (Figure 1). The combination of drug design with *in-vitro* assays, NMR and computational techniques enabled us to rationalize the observed HIV protease inhibition for the different PCU-lactam peptides/peptoids.

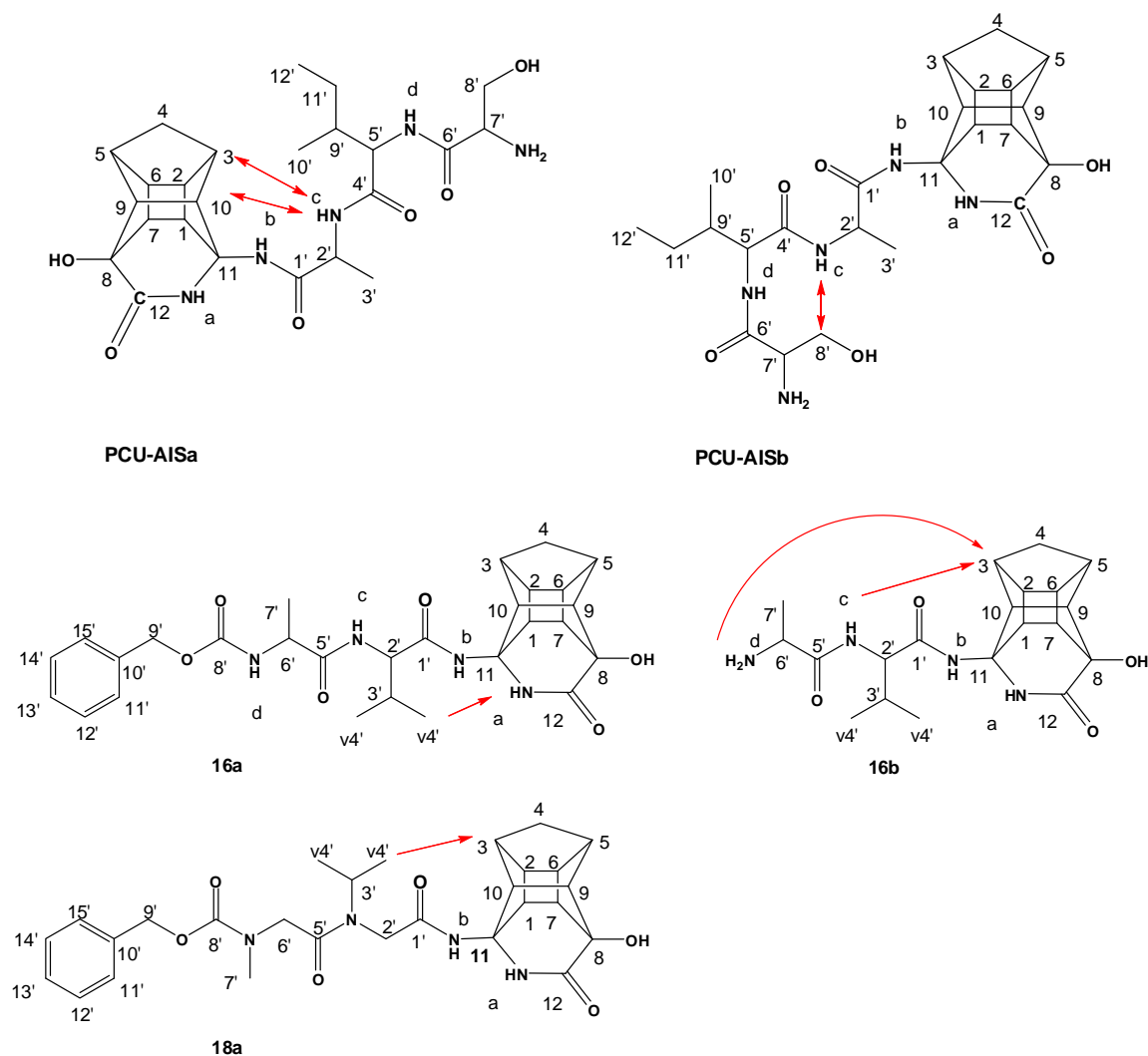


Figure 1. Inhibitor PCU-AISa, PCU-AISb, **16a**, **16b** and **18a** with arrows showing the EASY-ROESY correlations.

PCU-DIOL INHIBITORS

These inhibitors were synthesized by coupling the cage-diol moieties (PCU-diol diacid PCU-diol diamine and PCU-diol amino amide) to peptides (FVA-Cbz) and peptoids (FVA-Cbz mimics) chain sequences (Figure 2, and Figure 3)

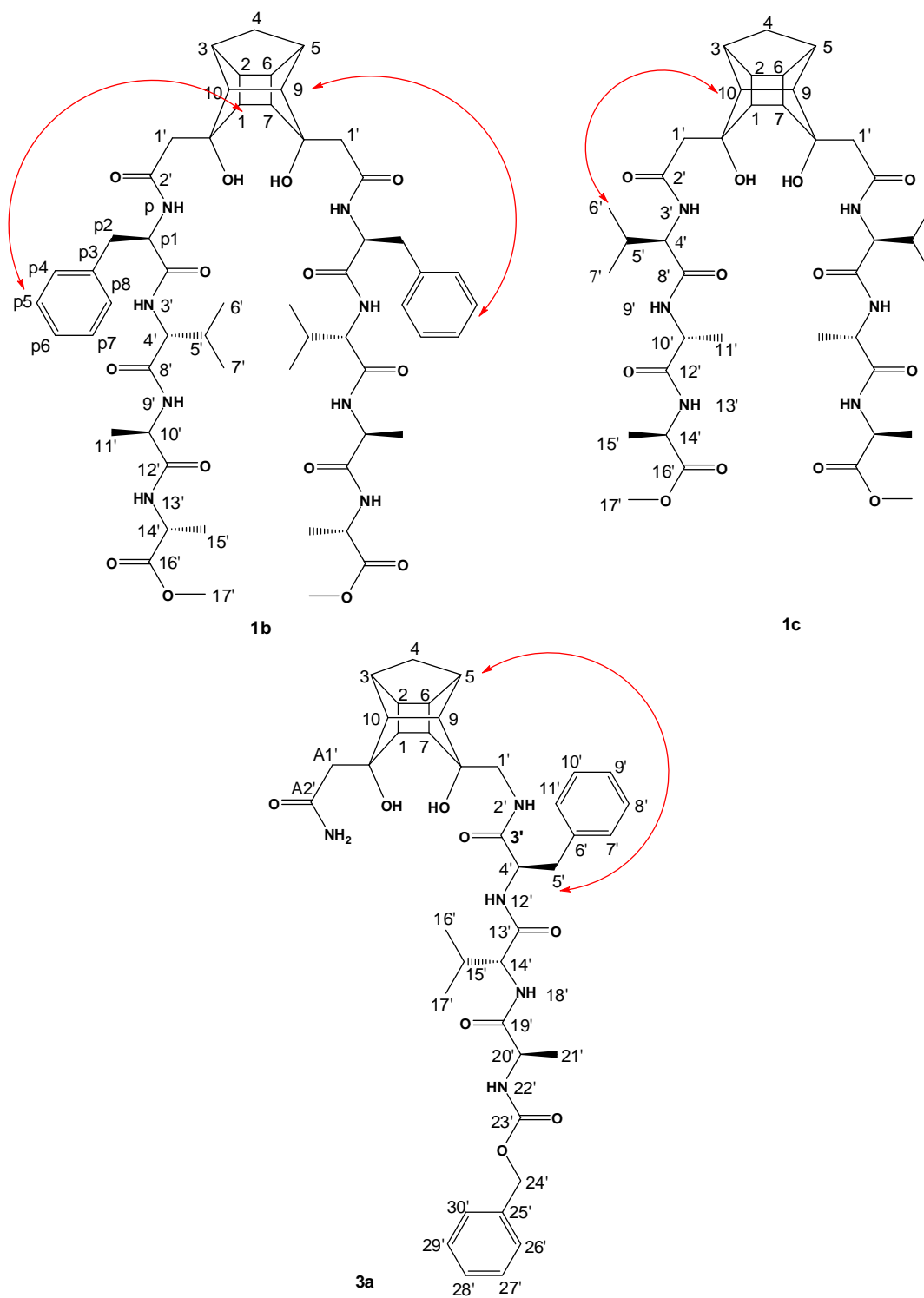


Figure 2. Numbered structures of inhibitor **1b**, **1c** and **3a** with arrows indicating the long range correlations observed in the EASY-ROESY spectra.

The different cage diol moieties were coupled to the peptide/peptoid sequences and yielded compounds with moderate to good activity. The peptoid coupled to the diamine resulted into the best inhibitor in this family with the IC_{50} of 75 nM (Figure 3).

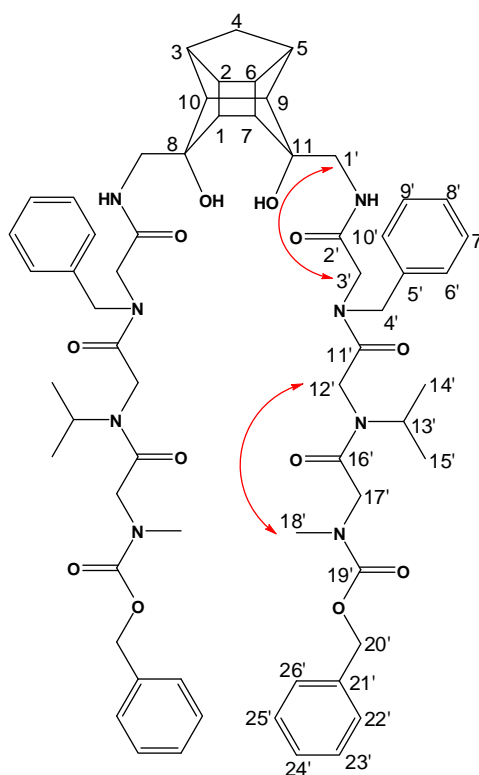


Figure 3. Numbered structure of inhibitor **2a** with arrows indicating the long range correlations observed in the EASY-ROESY spectrum.

The PCU-diol diacid derived inhibitors exhibited less activity when compared to the inhibitors derived from the PCU-diol diamine and PCU-diol amino amide moieties. Using the EASY-ROESY technique it was proven that the benzyl group is the appropriate P_1 substituent for this family of compounds. Large inhibitors that were insoluble in the buffer exhibited low inhibition activity when compared to large soluble compounds. The EASY-ROESY NMR technique enabled the determination of the 3D structures of these inhibitors.

CONCLUSION

The combination of the different cage moieties with peptides and peptoids yielded potential lead compounds for HIV-PR inhibition. Promising IC_{50} values of 75 nM and 78 nM were obtained from the peptoid and peptide series, respectively. Variation of the amino acid sequence, peptoid residues and protection of the cage hydroxyl (PCU-lactam) group enabled the determination of a structure-activity relationship for the cage peptide/peptoid inhibitors. Substitution of the cage moiety of the inhibitor PCU-EAIS ($IC_{50} = 78$ nM) with a phenol group destroyed the activity and this indicated the importance of the cage for bioactivity. The EASY-ROESY technique has proven to be an extremely powerful method to provide vital information about the 3D solution structure of small peptides and peptoid. NMR and computational techniques has enabled the rationalization of the observed HIV protease inhibition for the different cage peptides.

Employment of these techniques, combined with a systematic variation of alternative cage analogues and peptide/peptoids sequences lead to the design of active cage inhibitors. In the future the possibility of cage peptides/peptoids to inhibit some of the many alternative disease related protease families will also be investigated.

APPENDICES

Appendix 1: Chapter 2 supporting information

Appendix 2: Chapter 3 supporting information

Appendix 3: Chapter 4 supporting information

Appendix 4: Chapter 5 supporting information

APPENDIX 1

CHAPTER 2

SUPPORTING INFORMATION

MATERIALS AND METHODS

Analytical analysis was performed on an Agilent 1100 HPLC with a flowrate of 1 mL/min on a Waters Xbridge C18 (150 mm x 4.6 mm x 5 microns) coupled to a UV detector (215 nm) and an Agilent VL ion trap mass spectrophotometer in the positive mode. Semi-preparative HPLC was carried out on a Shimadzu semi-preparative instrument with a flowrate of 17 mL/min on a Ace C18 (150 mm x 21.2 mm x 5 microns) with a UV/VIS detector (215 nm) and an automated fraction collector. A two-buffer system was employed, utilizing formic acid as the ion-pairing agent. Buffer A consisted of 0.1 % formic acid/H₂O (v/v) and buffer B consisted of 0.1 % formic acid/acetonitrile (v/v). High resolution mass spectroscopic analysis was performed on a Bruker MicroTOF QII mass spectrometer in positive mode with an internal calibration. Peptides were synthesized on an automated CEM Liberty microwave peptide synthesizer (see Table 1 for conditions, not same as supplier). All ¹H, ¹³C, HSQC, COSY and HMBC NMR data were recorded on a Bruker AVANCE III 400 MHz spectrometer. The ROESY data was recorded on a Bruker AVANCE III 600 MHz spectrometer.

Supplementary Table 2. Microwave conditions for coupling and deprotection

	Microwave power (Watts)	Temperature (°C)	Time (sec)
Single coupling			
30 minute coupling	0	25	900
	35	73	900
Deprotection	40	73	180

SYNTHESIS OF AMINO LACTAM (6)

To an ice cold solution of 25 % ammonia (15 mL) the dione (1.00 g, 5.7 mmol) was added, followed by NH_4Cl (0.42 g, 9.1 mmol, 2.2 eq) and NaCN (0.45 g, 7.9 mmol, 2.3 eq), respectively. The mixture was stirred in the ice-bath for 2 hours resulting in a white thick solution that was filtered using a sintered glass funnel. The precipitate was washed with acetone (2 x 15 mL) and the product was recrystallized from methanol to yield **14** as white crystals (73 % yield).

GENERAL PROCEDURE FOR THE SYNTHESIS OF PEPTIDES USING MICROWAVE POWER

Stock solutions of all amino acids (0.2 mM), DIPEA (1 mM) and HBTU (2 mM) were prepared in DMF and peptides were synthesized on a 0.5 mmol scale. The modified coupling method described in Table 1 above was used for coupling the subsequent amino acids to the 2-chlorotrityl chloride resin (1 g) preloaded with the first amino acid.

Final cleavage : The resin bound peptide was removed from the peptide synthesizer and transferred to a manual peptide synthesis reaction vessel for the final cleavage. It was washed with DCM (3 x 10 mL) and a cleavage mixture of 5:95 % (v/v) TFA:DCM was added to the resin while nitrogen was bubbled through the solution for 10 minutes. The resin was washed three times with the cleavage mixture and the cleaved peptide was removed by filtration and collected in a flask containing water (100 mL). The filtrate was extracted several times with DCM so as to remove the peptide from the water layer. DCM was removed under reduced vacuum with at 40°C and the peptide remained as a white powder.

GENERAL PROCEDURE FOR COUPLING OF PEPTIDES TO THE PCU-LACTAM

The cleaved peptide (1.2 eq.) was dissolved in DCM (3 mL) followed by addition of PCU lactam **14** (1 eq), HATU (2.5 eq) in DMF (7 mL) and DIPEA (3 eq). The mixture was stirred at room temperature for 24 hours and thereafter the DMF was removed under reduced pressure using a Teflon pump at 40°C. The crude product was dissolved with EtOAc (200 mL) and the resulting solution was washed with 1M HCl, NaHCO_3 (saturated solution) and NaCl (saturated solution), dried over MgSO_4 , filtered and concentrated *in vacuo*. A cleavage mixture (10 mL) of 95:5 (v/v) TFA:DCM was added to the crude product and stirred for 24 hours at room temperature. Cold ether was then added and the product was obtained as yellow precipitate which was purified by reverse phase semi-preparative HPLC.

OVER-EXPRESSION, EXTRACTION AND PURIFICATION OF THE C-SA PROTEASE

Plasmid encoding HIV-1 subtype C protease (containing the mutation Q7K designed to reduce the hypersensitive autolytic site) is expressed as inclusion bodies (2) in *Escherichia coli* BL21 (DE3) pLysS cells. Briefly, *Escherichia coli* cells harboring the plasmid DNA were grown at 37 °C in LB medium supplemented with 100 µg/mL of ampicillin and 35 µg/mL of chloramphenicol. The overnight culture was diluted 100-fold into fresh 2 × YT medium supplemented with ampicillin (100 µg/mL) and chloramphenicol (35 µg/mL) and grown at 37 °C. When the optical density (OD₆₀₀) of the culture reached 0.4 to 0.5, over-expression of the HIV-1 C-SA protease was induced by adding IPTG. IPTG was added to final concentrations of 0.4 mM. Over-expression of the protease was allowed to continue for four hours.

The cells were pelleted after growth and resuspended in ice-cold extraction buffer [10 mM Tris, 1 mM EDTA, and 1 mM PMSF (added only fresh before use), pH 8] and disrupted using an ultra-sonicator. Following the addition of MgCl₂ and DNase I to final concentrations of 10 mM and 10 µg/mL, respectively, the culture medium was stirred on ice until the viscosity of the mixture decreased. The cells were then ruptured by sonication and centrifuged at 15 000 × g for 30 minutes at 4 °C. The pellet was resuspended in ice-cold extraction buffer containing 1 % (v/v) of Triton X-100. Cell debris and protease-containing inclusion bodies were pelleted by centrifugation at 15 000 × g for 30 minutes at 4 °C. The pellet was then resuspended in a freshly prepared solubilization buffer containing 10 mM Tris, 2 mM DTT, 8 M urea, pH 8.0, at room temperature, and centrifuged at 15 000 × g for 30 minutes at 20 °C.

The protease, in the supernatant, was purified by passing it through an anion exchange (DEAE) column previously equilibrated with solubilization buffer. Upon elution from the column, the protease was acidified by adding formic acid to a final concentration of 25 mM. Precipitation of significant amount of contaminating proteins occurred upon acidification. Following an overnight incubation, the precipitated contaminants were removed by centrifugation at 15 000 × g for 30 minutes at 4 °C.

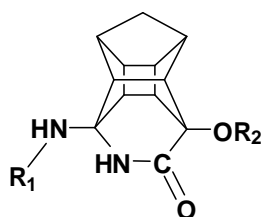
HIV-1 protease was refolded by dialysis into 10 mM formic acid at 4 °C. Subsequently, the protease was dialyzed into storage buffer containing 10 mM sodium acetate, 1 mM NaCl and 1 mM DTT, pH 5.0. The folded protease was concentrated to a final volume of ~5 mL and stored at -20 °C.

***IN VITRO* HIV-1 PROTEASE ACTIVITY**

The catalytic activity of the HIV-1 protease was monitored following the hydrolysis of the chromogenic peptide substrate Lys-Ala-Arg-Val-Nle-*p*-nitro-Phe-Glu-Ala-Nle-NH₂. This substrate mimics the conserved KARVL/AEAM cleavage site between the capsid protein and nucleocapsid (CA-p2) in the Gag polyprotein precursor.

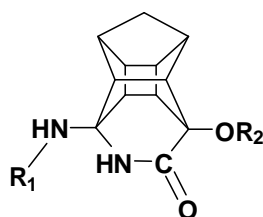
For this study the chromogenic substrate was synthesized using the Discovery CEM Liberty microwave peptide synthesizer on Rink amide resin. The substrate was cleaved from the resin and deprotected using 95 %:5 % (v/v) TFA:TIS for 3 hours. It was then precipitated using cold ether and purified *via* reverse phase semi-preparative HPLC on a Shimadzu instrument and characterized using the Bruker microTOF-Q II instrument (Table 2).

To determine the concentration of the inhibitors that resulted in 50 % inhibition (IC₅₀) of HIV-1 protease enzyme activity, the protein (100 nM) and chromogenic substrate (50 μM) were added into a 120 μL microcuvette containing increasing concentrations of inhibitor in a pH 5.0 buffer (50 mM sodium acetate and 0.1 M NaCl). Protease hydrolytic activity was measured by monitoring the relative decrease in absorbance at 300 nm using Fpecord 210 spectrophotometer.



Supplementary Table 3. The inhibition of wild type C-SA HIV-1 protease by PCU lactam peptide derivatives compared with the commercial available drugs, the optical rotation, the membrane permeability, Clog*P* values and stability

Compound	R ₁	R ₂	IC ₅₀ (μM)	CLog <i>P</i>
PCU-EAIS	-Glu-Ala-Ile-Ser	H	0.078 ± 0.0035	-3.45 ± 0.92
PCU-EAI	-Glu-Ala-Ile	H	0.5 ± 0.035	-2.53 ± 0.90
PCU-EVIS	-Glu-Val-Ile-Ser	H	2.0 ± 0.18	-2.57 ± 0.89
PCU-QAIS	-Gln-Ala-Ile-Ser	H	5.0 ± 0.71	-3.84 ± 0.84
PCU-AISa	-Ala-Ile-Ser	H	0.5 ± 0.035	-2.65 ± 0.75
PCU-AISa	-Ala-Ile-Ser	H	10.0 ± 1.06	-2.65 ± 0.75
Ac-PCU-EAIS	-Glu-Ala-Ile-Ser	Ac	>10	-2.55 ± 0.92
Phenol-EAIS			10.0 ± 3.53	-0.06 ± 0.92
Atazanavir			0.004 ± 0.00071	
Lopinavir			0.025 ± 0.0014	

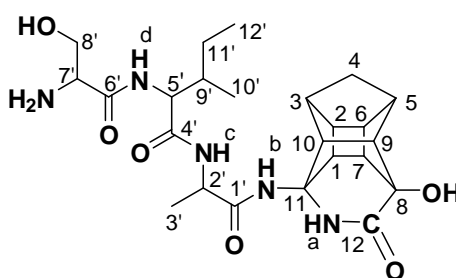


Supplementary Table 3. Yields for coupling of peptides to the lactam and the size as confirmed by high-resolution mass spectroscopy

Compound	R ₁	R ₂	Yields (%)	Calculated mass	Found mass	Molecular ion
PCU-EAIS	-Glu-Ala-Ile-Ser	H	47	619.2880 C ₂₉ H ₄₃ N ₆ O ₉	619.2878	(M+H ⁺)
PCU-EAI	-Glu-Ala-Ile	H	23	532.2766 C ₂₆ H ₃₈ N ₅ O ₇	532.2776	(M+H ⁺)
PCU-EVIS	-Glu-Val-Ile-Ser	H	33	647.3399 C ₃₁ H ₄₇ N ₆ O ₉	647.3389	(M+H ⁺)
PCU-QAIS	-Gln-Ala-Ile-Ser	H	20	618.3255 C ₂₉ H ₄₄ N ₇ O ₈	618.3253	(M+H ⁺)
PCU-AISa	-Ala-Ile-Ser	H	18	490.2460 C ₂₄ H ₃₆ N ₅ O ₆	490.2464	(M+H ⁺)
PCU-AISb	-Ala-Ile-Ser	H	17	490.2660 C ₂₄ H ₃₆ N ₅ O ₆	490.2665	(M+H ⁺)
Ac-PCU-EAIS	-Glu-Ala-Ile-Ser	Ac	27	661.3192 C ₃₁ H ₄₅ N ₆ O ₉	661.3191	(M+H ⁺)

Phenol-EAIS	87	510.2572	510.2573	(M+H ⁺)
			C ₂₃ H ₃₅ N ₅ O ₈	
Atazanavir		-		
Lopinavir		-		

NMR elucidation of PCU-AISa



Supplementary Table 4. ¹H and ¹³C chemical shift of the PCU-AISa as measured by HSQC, COSY, HMBC, ROESY, ¹H and ¹³C spectra. The chemical shift was referenced to 2.50 ppm for COSY and ROESY the correlation that aided in assignment are marked in black (sequential correlations) and the long-range interactions through space, measured by ROESY, are marked in red

Atom	$\delta^1\text{H}^{a,b}$	$\delta^{13}\text{C}^{a,d}$	COSY ($\delta^1\text{H}^{a,b}$)	HMBC	ROESY ($\delta^1\text{H}^{a,c}$)
1x	2.82	43.4	1y, 2, 7		3, 2, 7, 9, NHaa, NHbb
1y	2.97	42.9	1x, 2, 7		3, 7, NHa, NHc, NHcc
2	2.46	39.6	1x, 1y, 6, 7	3	1x, 4a, NHaa, NHbb
3	2.42	43.1	4a, 4s, 10x, 10y	2, 10x	1x, 1y, 3', 6, 4a, 4s, 10y, NHc, NHcc
4a	1.35 (d. $J=10.2$ Hz)	36.9	3, 4s, 5	9	2, 3, 4s, 5, 6, 10x
4s	1.68 (d. $J=10.5$ Hz)	36.9	3, 4a, 5	2, 6	3, 5, 4a, 6, 9, 10x

	Hz)				
5	2.66	45.2	4a, 4s, 9	2, 9	4a, 4s, 7, 9
6	2.75	41.6	2, 7	9	3, 4a, 4s 7, 9
7	2.34	44.3	1x, 1y, 2, 6	8, 9, 11	1x, 1y, 5, 6, 10x
8		79.7			
9	2.17	51.9	5, 10x, 10y	5, 8	1x, 4s, 5, 6,
10x	2.72	50.8	3, 9, 10y	3, 9	4a, 4s, 7, NHa, NHc, NHcc
10y	2.88	50.2	3, 9, 10x	3	3, NHa, NHaa
11		72.9			
12		173.6			
1'		172.4			
2'	4.21	48.5	*	1', 3'	3', NHbb, NHc, NHcc
3'	1.18	18.0	*	1', 2'	2', 3
4'		170.1			
5'	4.25	56.9	9'	4', 6', 9', 11'	9', 10', 11a', 11s', 12'
6'		167.4			
7'	3.88	54.3		6', 8a', 8s'	8a', 8s', NHd
8a'	3.61	60.9		6', 7', 8s'	7', 8s, NHd
8s'	3.72	60.9		6', 7', 8a	7', 8a
9'	1.74	37.0	5', 10', 11a, 11s	10'	5', 10', 11a', 11s', 12' NHd
10'	0.85	15.3	9', 11a, 11s	5', 11a', 11s'	5', 9', 11a', 11s', NHd

11a'	1.08	24.2	9', 11s', 12'		5', 9', 10', 11s', 12', NHd
11s'	1.43	24.2	9' 11a', 12'		5', 9', 10, 11a', 12'
12'	0.82	11.3	11a', 11s'	9', 11'	5', 9', 11a', 11s'
NHa	7.94		*	1, 8, 10, 11, 12	1y, 10x, 10y
NHaa	7.99		*	8, 10, 11, 12	1x, 2, 10y
NHb	8.09		*	1', 1x, 10a, 10s, 11	
NHbb	8.11		*	1', 1x, 10a, 10s 11	1x, 2, 2'
NHc	8.03		*	2', 3', 4'	1y, 2', 3, 10x
NHcc	8.07		*	2', 3', 4'	1s, 2', 3, 10x
NHd	8.42		*	5, 6'	7', 8a', 9', 10', 11a

a Solvent (CD₃)₂SO.

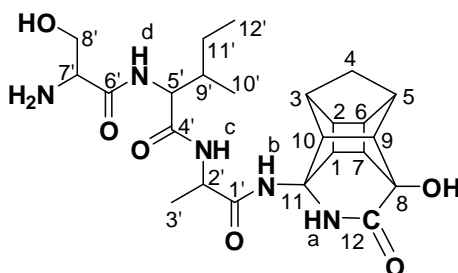
b 400 MHz for ¹H.

c 600 MHz for ¹H.

d 100 MHz for ¹³C.

* Overlapping signals

NMR elucidation of PCU-AISb



Supplementary Table 5. ^1H and ^{13}C chemical shift of PCU-AISb as measured by HSQC, COSY, HMBC, ROESY, ^1H and ^{13}C spectra. The chemical shift was referenced to 2.50 ppm for COSY and ROESY the correlation aided in assignment are marked in black (sequential correlations) and the long-range interactions through space, measured by ROESY, are marked in red. The carbon spectrum was also compared with the carbon spectrum of PCU-AISa

Atom	$\delta^1\text{H}^{\text{a,b}}$ (PCU-AISb)	$\delta^{13}\text{C}^{\text{a,d}}$ (PCU-AISb)	$\delta^{13}\text{C}^{\text{a,d}}$ (PCU-AISa)	ROESY ($\delta^1\text{H}^{\text{a,c}}$) (PCU-AISb)
1x			43.4	
1y	2.92	42.9	42.9	2, 7, NHa, NHb
2	2.47	39.3	39.6	1x, 1y, 5, 6 NHa, NHb
3	2.42	43.1	43.1	4a, 4s, 10x, 10y, NHa, NHb
4a	1.35 (d. $J = 10.1$ Hz)	36.9	36.9	3, 4s, 5, 6
4s	1.68 (d. $J = 10.4$ Hz)	36.9	36.9	3, 4a, 5, 6, 9, 10x
5	2.66	45.2	45.2	2, 4a, 4s, 7, 6, 9
6	2.75	41.6	41.6	2, 4a, 4s, 5, 7, 9, 10s
7	2.34	44.3	44.3	1x, 1y, 6, 5, NHa, NHaa
8		79.7	79.7	

9	2.17	51.9	51.9	4s, 5, 6, 10x, 10y, NHa, NHaa
10x	2.83	50.6	50.8	3, 4s, 9, NHa, NHb
10y			50.2	
11		72.9	72.9	
12		173.7	173.6	
1'		172.3	172.4	
2'	4.29	48.5	48.5	3', NHb, NHc
3'	1.18	18.0	18.0	2', NHa, NHb, NHc, NHcc
4'		170.1	170.1	
5'	4.26	57.1	56.9	9', 10', NHd
6'		166.9	167.4	
7'	3.95	54.1	54.3	8a', 8s', NHcc, NHd
8a'	3.62	60.6	60.9	7', 8s', NHcc, NHd
8s'	3.72	60.6	60.9	7', 8a', NHcc, NHd
9'	1.74	37.0	37.0	5', 10', 11a', 11s', 12' NHc, NHd
10'	0.85	15.3	15.3	5', 9, 11s'
11a'	1.09	24.2	24.2	9', 11s', NHd
11s'	1.43	24.2	24.2	9', 10', 11a', 12' NHd, NHc
12'	0.81	11.3	11.3	9', 11s', NHc, NHcc, NHd
NHa	7.99			1s, 2, 3, 3', 7, 9, 10x
NHaa	8.01			1a, 7, 9, 10y

NHb	8.15			1s, 2, 2'/5', 3, 3', 10x, 10y
NHbb				
NHc	8.09			2'/5', 3', 9', 11s', 12'
NHcc	8.10			3', 7', 8a', 8s', 9', 12'
NHd	8.44			2'/5', 7', 8a', 8s', 9', 11a', 11s', 12'

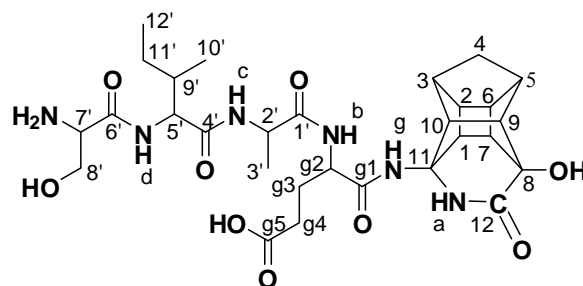
a Solvent (CD₃)₂SO.

b 400 MHz for ¹H.

c 600 MHz for ¹H.

d 100 MHz for ¹³C.

NMR elucidation of PCU-EAIS



Supplementary Table 6. ^1H and ^{13}C chemical shift of PCU-EAIS as measured by HSQC, COSY, HMBC, ROESY, ^1H and ^{13}C spectra. The chemical shift was referenced to 2.50 ppm for COSY and ROESY the correlation aided in assignment are marked in black (sequential correlations) and the long-range interactions through space, measured by ROESY, are marked in red. The carbon spectrum was also compared with the carbon spectrum of PCU-AISa

Atom	$\delta^1\text{H}^{\text{a,b}}$ (PCU-EAIS)	$\delta^{13}\text{C}^{\text{a,d}}$ (PCU-EAIS)	$\delta^{13}\text{C}^{\text{a,d}}$ (PCU-AISa)	ROESY($\delta^1\text{H}^{\text{a,c}}$)(PCU-EAIS)
1x	2.89	43.2	43.4	2, 3,6, 7, NHc
1y	2.97	42.9	42.9	2, 3, 6, 7, 10, NHa, NHg
2	2.45	39.6	39.6	1x, 1y, NHa, NHb
3	2.41	43.1	43.1	1x, 1y, 3', 4a, 4s, 6, 10, 10', NHa, NHb, NHc, NHg
4a	1.35 (d. $J = 10.6$ Hz)	36.9	36.9	3, 4s, 5, 6
4s	1.68	36.9	36.9	3, 4a, 5, 6, 9, 10
5	2.66	45.2	45.2	4a, 4s, 6, 9
6	2.75	41.6	41.6	1x, 1y, 2, 3, 4a, 4s, 5, 7, 9
7	2.34	44.4	44.3	1x, 1y, 6
8		79.6	79.7	

9	2.17	51.9	51.9	4s, 5, 6, 10
10x			50.8	
10y	2.77	50.5	50.2	1y, 3, 4s, 9, 10, NHa, NHb, NHc
11		72.9	72.9	
12		171.8	173.6	
g1'		171.2		
g2'	4.22	51.9		g3a', g3s, g4, NHa
g3a'	1.72	27.4		g2', g3s, g4, NHa, NHg
g3s'	1.89	27.4		g2', g3a, g4, NHa, NHg
g4'	2.21	29.8		3', g2', g3a', g3s', 10', NHa
g5'		174.0		
1'		172.1	172.4	
2'	4.27	48.6	48.5	3', NHb, NHc, NHcc
3'	1.19	17.9	18.0	2', 3, g4', NHb, NHc, NHcc, NHg
4'		170.3	170.1	
5'	4.25	56.9	56.9	9', 10', 11a, 11s, NHd
6'		166.7	167.4	
7'	3.93	54.1	54.3	8a', 8s', 10', 12', NHd
8a'	3.62	60.5	60.9	7', 8s', NHd
8s'	3.72	60.5	60.9	7', 8a', NHd
9'	1.73	37.0	37.0	5', 10', 11a', 11s', 12', NHc,

				NHcc NHd
10'	0.86	15.3	15.3	3, g4', 5', 7, 9', 11a', NHc, NHcc
11a'	1.05	24.2	24.2	5', 9', 10, 11s', 12', NHd
11s'	1.43	24.2	24.2	5', 7, 9', 11a', 12' NHd, NHc
12'	0.82	11.2	11.3	5', 9', 11a', 11s', NHcc, NHd
NHa	8.08			1x, 1y, 2, 3, 10, g2', g3a', g3s, g4'
NHaa				
NHg	8.10			1y, 3, 3', g3a', g3s
NHb	7.80			2', 2, 3, 3', g2', g3a', g3s, g4'
NHbb	7.94			2', 2, 3, 3', g2', g3a', g3s, g4'
NHc	8.16			1x, 2', 3', 3, 6, 9', 10', 11s'
NHcc	8.20			2', 3', 9', 10', 12'
NHd	8.40			2'/5', 7', 8a', 8s', 9', 11a', 11s', 12'

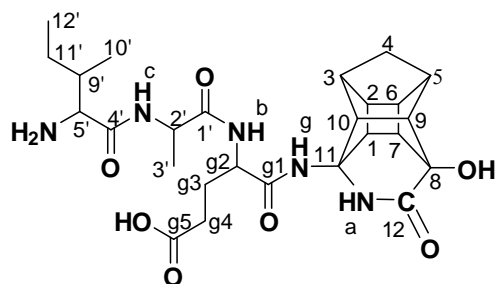
a Solvent (CD₃)₂SO.

b 400 MHz for ¹H.

c 600 MHz for ¹H.

d 100 MHz for ¹³C

NMR elucidation of PCU-EAI



Supplementary Table 7. ^1H and ^{13}C chemical shift of PCU-EAI as measured by HSQC, COSY, HMBC, ROESY, ^1H and ^{13}C spectra. The chemical shift was referenced to 2.50 ppm for COSY and ROESY the correlation aided in assignment are marked in black (sequential correlations) and the long-range interactions through space, measured by ROESY, are marked in red. The carbon spectrum was also compared with the carbon spectrum of PCU-EAIS

Atom	$\delta^1\text{H}^{\text{a,b}}$ (PCU-EAI)	$\delta^{13}\text{C}^{\text{a,d}}$ (PCU-EAI)	$\delta^{13}\text{C}^{\text{a,d}}$ (PCU-EAIS)	ROESY($\delta^1\text{H}^{\text{a,c}}$) (PCU-EAI)
1x	2.93	43.6	43.2	2, 5, 6, 7, NHa, NHb, NHg, NHgg
1y	3.01	42.9	42.9	2, 5, 6, 7 NHa, NHbb
2	2.47	39.6	39.6	1x, 1y, 4a, 4s, 6, 10y, NHa, NHb, NHg, NHgg
3	2.44	43.0	43.1	3', 10x
4a	1.35 (d. $J = 10.1$ Hz)	36.9	36.9	2, 4s, 5, 6
4s	1.68	36.9	36.9	2, 4a, 5, 9, 10x
5	2.66	45.1	45.2	1x, 1y, 4a, 4s, 6, 7, 9
6	2.75	41.5	41.6	1x, 1y, 2, 4a, 4s, 5, 7, 10y
7	2.34	44.3	44.4	1x, 1y, 5, 6, NHa, NHb

8		79.6	79.6	
9	2.17	52.0	51.9	4s, 5, 10x, 10y, NHa
10x	2.79	50.7		3, 4s, 9, 10y, NHa, NHbb, NHg, NHgg
10y	2.87	50.5	50.5	2, 6, 9, 10x, NHb
11		72.9	72.9	
12		173.6	171.8	
g1'		171.5	171.2	
g2'	4.14	53.0	51.9	g3a', g3s, g4, NHa, NHb, NHg, NHgg
g3a'	1.78	27.2	27.4	g2', g3s, NHb, NHbb, NHc, NHcc
g3s'	1.88	27.2	27.4	g2', g3a, NHb, NHbb, NHc, NHcc
g4'	2.18	31.0	29.8	3', g2', NHb, NHbb, NHc, NHcc
g5'		175.5	174.0	
1'		171.6	172.1	
2'x	4.18	48.6	48.6	3', 10', NHb, NHbb, NHc
2'y	4.25	48.4		3', 10', NHcc
3'	1.25	17.9	17.9	3, 2'x, 2'y, g4', NHc, NHcc, NHb
4'		170.8	170.3	
5'	3.37	57.8	56.9	9', 10'

6'			166.7	
7'			54.1	
8a'			60.5	
8s'			60.5	
9'	1.74	37.1	37.0	5', 10', 11a', 11s', 12'
10'	0.86	15.1	15.3	2'x, 2'y, g4', 5', 7, 9', 11a'
11a'	1.12	24.0	24.2	9', 10', 11s', 12'
11s'	1.46	24.0	24.2	9', 11a', 12'
12'	0.82	11.3	11.2	9', 11a', 11s'
NHa	7.96			1x, 1y, 2, g2', 7, 9, 10x
NHg	8.12			1x, 2, g2', 10x
NHgg	8.19			1x, 2, g2, 10x
NHb	7.95			1x, 2'x, 3', g2', g3a', g3s g4', 7, 10y
NHbb	8.06			2'x, 1y, 10x, g3a', g3s, g4',
NHc	8.14			2'x, 2'y, 3', g3a', g3s, g4'
NHcc	8.31			2'y, 3', 5', g3a', g3s, g4'

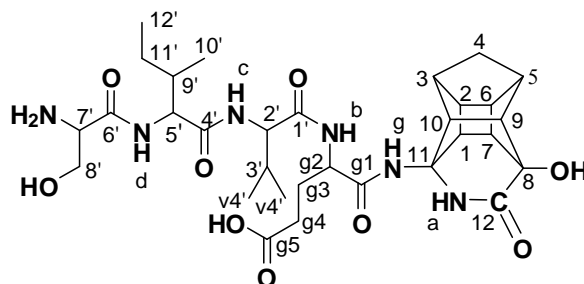
a Solvent (CD₃)₂SO.

b 400 MHz for ¹H.

c 600 MHz for ¹H.

d 100 MHz for ¹³C

NMR elucidation of PCU-EVIS



Supplementary Table 8. ^1H and ^{13}C chemical shift of PCU-EVIS as measured by HSQC, COSY, HMBC, ROESY, ^1H and ^{13}C spectra. The chemical shift was referenced to 2.50 ppm for COSY and ROESY the correlation aided in assignment are marked in black (sequential correlations) and the long-range interactions through space, measured by ROESY, are marked in red. The carbon spectrum was also compared with the carbon spectrum of PCU-EAIS

Atom	$\delta^1\text{H}^{\text{a,b}}$ (PCU-EVIS)	$\delta^{13}\text{C}^{\text{a,d}}$ (PCU-EVIS)	$\delta^{13}\text{C}^{\text{a,d}}$ (PCU-EAIS)	ROESY($\delta^1\text{H}^{\text{a,c}}$) (PCU-EVIS)
1x	2.84	43.1	43.2	3, 4s, 7, 9, 10', NHg, NHgg
1y	2.95	42.9	42.9	2, 4a, 7, 10', NHa, NHg
2	2.46	39.6	39.6	1y, 4a, 5, 10x, NHbb
3	2.43	43.1	43.1	1x, 4a, 4s, 6, 10', 10x, 10y, NHb, NHg, NHgg
4a	1.35 (d. $J = 10.1$ Hz)	36.9	36.9	1y, 2, 3, 4s, 5, 10x
4s	1.67 (d. $J = 10.8$ Hz)	36.9	36.9	1x, 3, 4a, 5, 9
5	2.66	45.1	45.2	2, 4a, 4s, 6, 7, 9, 10x
6	2.75	41.6	41.6	3, 5, g4'
7	2.34	44.4	44.4	1x, 1y, 5, 10x, 10y, 12'

8		79.6	79.6	
9	2.17	51.9	51.9	1x, 4s, 5
10x	2.73	50.7		2, 3, 4a, 5, 6, 7
10y	2.81	50.5	50.5	3, 7, g4'
11		72.9	72.9	
12		173.6	171.8	
g1'		171.4	171.2	
g2'	4.21	52.1	51.9	g3a', g3s, g4, NHb, NHbb, NHg, NHgg
g3a'	1.73	27.4	27.4	g2', g3s
g3s'	1.93	27.4	27.4	g2', g3a,
g4'	2.20	30.2	29.8	g2', 6, V4', 10y
g5'		174.3	174.0	
1'		170.8	172.1	
2'	4.09	58.5	48.6	3', v4', NHb, NHc, NHbb
3'	1.94	30.2	17.9	2', V4', NHb, NHc, NHbb
v4'	0.84	19.1		2', 3', g4', V4', NHb, NHc, NHbb
v4'	0.84	19.1		2', 3', g4', NHb, NHc, NHbb
4'		171.2	170.3	
5'	4.27	57.0	56.9	9', 10', 11a, 11s
6'		173.2	166.7	
7'	3.54	55.6	54.1	-

8a'	3.51	62.9	60.5	-
8s'	3.59	62.9	60.5	-
9'	1.74	37.0	37.0	5', 10', 11a', 11s', NHb, NHc,
10'	0.82	15.3	15.3	1x, 1y, 5', 9'
11a'	1.06	24.2	24.2	5', 9', 11s', 12'
11s'	1.41	24.2	24.2	5', 9', 11a', 12'
12'	0.79	11.2	11.2	11a', 11s',
NHa	8.04			1y
NHaa				
NHg	8.15			1x, 1y, 2, 3, g2'
NHgg	8.21			1x, 3, g2'
NHb	7.93			2', 3', g2', g3'a, g3's, V4'
NHbb				1y, 2', 2, 3', 3, g2', V4'
NHc	8.02			2', 3', 5', 9', V4'
NHcc				
NHd	8.24			-

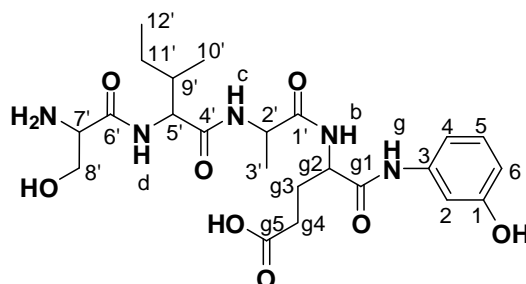
a Solvent (CD₃)₂SO.

b 400 MHz for ¹H.

c 600 MHz for ¹H.

d 100 MHz for ¹³C

NMR elucidation of Phenol-EAIS



Aromatic protons 1-6 were observed at 6.94 ppm -7.23 ppm. The HSQC spectrum showed the same pattern for the EAIS side chain as the one observed for PCU-EAIS. The ROESY spectrum did not show long range correlations.

Supplementary Table 9. ^1H and ^{13}C chemical shift of Phenol-EAIS as measured by HSQC, COSY, HMBC, ROESY, ^1H and ^{13}C spectra. The chemical shift was referenced to 2.50 ppm for COSY and ROESY the correlation aided in assignment are marked in black (sequential correlations) and the long-range interactions through space, measured by ROESY, are marked in red. The carbon spectrum was also compared with the carbon spectrum of PCU-EAIS

Atom	$\delta^1\text{H}^{\text{a,b}}$ (Phenol-EAIS)	$\delta^{13}\text{C}^{\text{a,d}}$ (Phenol-EAIS)	$\delta^{13}\text{C}^{\text{a,d}}$ (PCU-EAIS)	ROESY($\delta^1\text{H}^{\text{a,c}}$) (Phenol-EAIS)
1-6	6.94-7.23	106.5-157.6		
g1'			171.2	
g2'	4.35	52.6	51.9	g2', g3a', g3s', NHg
g3a'	1.85	27.4	27.4	g2', g3s', NHg
g3s'	2.02	27.4	27.4	g2', g3a', NHg
g4'	2.33	29.1	29.8	g2', NHg
g5'		173.0	174.0	
1'		172.2	172.1	
2'	4.26	48.7	48.6	3', NHb, NHg

3'	1.23	17.6	17.9	2', NHb, NHc, NHd
4'			170.3	
5'	4.23	56.5	56.9	9', 12', NHd
6'		169.6	166.7	
7'	3.52	55.7	54.1	-
8a'	3.51	62.8	60.5	-
8s'	3.59	62.8	60.5	-
9'	1.75	37.2	37.0	5', 10', 11a', 11s', 12', NHc, NHd
10'	0.83	15.2	15.3	9', NHc
11a'	1.08	24.1	24.2	9', 11s', 12'
11s'	1.43	24.1	24.2	9', 11a', 12'
12'	0.80	11.1	11.2	5', 9', 11a', 11s'
NHg	8.20			2', g2', g3a', g3s', g4'
NHb	8.00			2', 3'
NHc	8.16			3', 5', 9', 10'
NHd	8.31			3', 5', 9', 12'

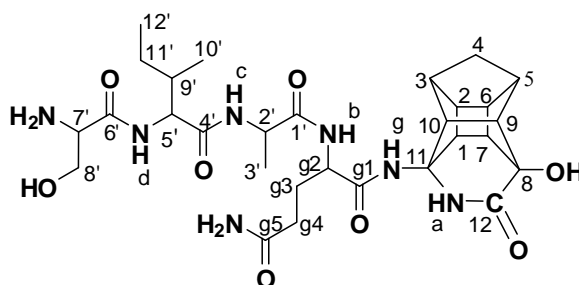
a Solvent (CD₃)₂SO.

b 400 MHz for ¹H.

c 600 MHz for ¹H.

d 100 MHz for ¹³C

NMR elucidation of PCU-QAIS



Supplementary Table 10. ^1H and ^{13}C chemical shift of PCU-QAIS as measured by HSQC, COSY, HMBC, ROESY, ^1H and ^{13}C spectra. The chemical shift was referenced to 2.50 ppm for COSY and ROESY the correlation aided in assignment are marked in black (sequential correlations) and the long-range interactions through space, measured by ROESY, are marked in red. The carbon spectrum was also compared with the carbon spectrum of PCU-EAIS

Atom	$\delta^1\text{H}^{\text{a,b}}$ (PCU-QAIS)	$\delta^{13}\text{C}^{\text{a,d}}$ (PCU-QAIS)	$\delta^{13}\text{C}^{\text{a,d}}$ (PCU-EAIS)	ROESY($\delta^1\text{H}^{\text{a,c}}$)PCU-QAIS
1x	2.91	43.1	43.2	7, NHaa
1y	2.98	42.7	42.9	3, NHg,
2	2.45	39.6	39.6	4a, 4s, 10, NHaa, NHb, NHbb, NHg, NHgg
3	2.41	43.0	43.1	1y
4a	1.35 (d. $J = 10.5$ Hz)	36.9	36.9	2, 4s, 5, 6
4s	1.68	36.9	36.9	2, 4a, 5, 9, 10
5	2.66	45.1	45.2	4a, 4s, 6, 9
6	2.75	41.6	41.6	4a, 5, 7
7	2.34	44.4	44.4	1x, 6
8		79.6	79.6	

9	2.17	51.9	51.9	4s, 5, 10
10x				
10y	2.78	50.7	50.5	2, 4s, 9, NHa, NHaa, NHg, NHgg
11		72.9	72.9	
12		173.6	171.8	
g1'		171.4	171.2	
g2'	4.18	52.3	51.9	g3a', g3s, g4, NHb, NHbb, NHg, NHgg
g3a'	1.66	27.6	27.4	g2', g3s, g4, NHb, NHbb, NHg, NHgg
g3s'	1.88	27.6	27.4	g2', g3a, g4', NHb, NHbb, NHg, NHgg
g4'	2.05	31.1	29.8	g2', g3s', NHb, NHbb
g5'		173.7	174.0	
1'		172.0	172.1	
2'	4.22	56.3	48.6	3', NHb, NHbb, NHc, NHcc
3'	1.17	17.9	17.9	2', NHb, NHbb, NHc, NHcc, NHd
4'		170.7	170.3	
5'	4.23	48.6	56.9	9', 10', 11a, 11s, 12', NHc
6'		-	166.7	
7'	3.37	56.3	54.1	-
8a'	3.46	-	60.5	-

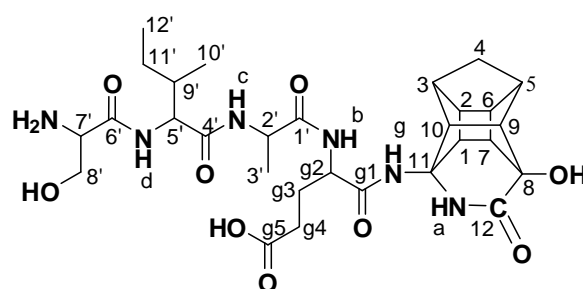
8s'	3.53	-	60.5	-
9'	1.70	37.2	37.0	5', 10', 11a', 11s', 12', NHcc
10'	0.82	15.2	15.3	5', 9', 11s'
11a'	1.03	24.1	24.2	5', 9', 11s', 12'
11s'	1.40	24.1	24.2	5', 9', 10', 11a', 12'
12'	0.79	11.2	11.2	5', 9', 11a', 11s'
NHa	8.00			10
NHaa	8.14			1x, 2, 10y
NHg	8.05			1y, 2, g2', g3a' g3s', 10
NHgg	8.08			2, g2', g3a' g3s', 10
NHb	7.93			2', 2, 3', g2', g3a', g3s', g4'
NHbb	7.95			2', 2, 3', g2', g3a', g3s', g4'
NHc	8.16			2', 3', 5'
NHcc	8.23			2', 3', 9'
NHd	8.30			3'

a Solvent (CD₃)₂SO.

b 400 MHz for ¹H.

c 600 MHz for ¹H.

d 100 MHz for ¹³C



Supplementary Table 11. EASY ROESY correlations observed in the spectra of PCU-EAIS recorded in DMSO- d_6 and the D_2O as the solvent. The chemical shift was referenced to 2.50 ppm

Atom	1H PCU-EAIS	ROESY(δ^1H^c)(PCU-EAIS) (DMSO- d_6)	ROESY(δ^1H^c)(PCU-EAIS (D_2O))
1x	2.89	2, 3, 6, 7, NHc	2, 3, 7, NHa, NHc, NHcc
1y	2.97	2, 3, 6, 7, 10, NHa, NHg	2, 3, 7, NHa
2	2.45	1a, 1s, NHa, NHb	1a, 1s
3	2.41	1a, 1s, 3', 4a, 4s, 6, 10, 10', NHa, NHb, NHc, NHg	1a, 1s, 4a, 4s, 10, NHa, NHc, NHcc
4a	1.35 (d. $J = 10.6$ Hz)	3, 4s, 5, 6	3, 4s, 5, 6
4s	1.68	3, 4a, 5, 6, 9, 10	3, 4a, 5, 6, 9, 10
5	2.66	4a, 4s, 6, 9	4a, 4s, 6
6	2.75	1a, 1s, 2, 3, 4a, 4s, 5, 7, 9	4a, 4s, 5, 7, 9
7	2.34	1a, 1s, 6	1a, 1s, 6
8			
9	2.17	4s, 5, 6, 10	4s, 6, 10
10	2.77	1s, 3, 4s, 9, 10, NHa, NHb, NHc	3, 4s, 9, NHa, NHc

11			
12			
g1'			
g2'	4.22	g3a', g3s, g4, NHa	g3a', g3s, g4, NHa, NHg
g3a'	1.72	g2', g3s, g4, NHa, NHg	g2', g3s, g4, NHg, NHbb,
g3s'	1.89	g2', g3a, g4, NHa, NHg	g2', g3a, g4, NHg
g4'	2.21	3', g2', g3a', g3s', 10', NHa	g2', g3a', g3s', NHg
g5'			
1'			
2'	4.27	3', NHb, NHc, NHcc	3', NHbb, NHc, NHcc
3'	1.19	2', 3, g4', NHb, NHc, NHcc, NHg	2', NHbb, NHc, NHcc
4'			
5'	4.25	9', 10', 11a, 11s, NHd	9', 11a', 11s', 12'
6'			
7'	3.93	8a', 8s', 10', 12', NHd	8a', 8s', NHd
8a'	3.62	7', 8s', NHd	7', 8s', NHd
8s'	3.72	7', 8a', NHd	7', 8a', NHd
9'	1.73	5', 10', 11a', 11s', 12', NHc, NHcc, NHd	5', 10', 11a', 11s', 12', NHc, NHcc, NHd
10'	0.86	3, g4', 5', 7, 9', 11a', NHc, NHcc	9', 11a', NHcc
11a'	1.05	5', 9', 10, 11s', 12', NHd	5', 9', 10, 11s', NHd
11s'	1.43	5', 7, 9', 11a', 12' NHd, NHc	5', 9', 11a', 12'

12'	0.82	5', 9', 11a', 11s', NHcc, NHd	5', 9', 11a', 11s'
NHa	8.08	1a, 1s, 2, 3, 10, g2', g3a', g3s, g4'	1a, 1s, 2, 3, 10, g2'
NHaa			
NHg	8.10	1s, 3, 3', g3a', g3s	g2', g3a', g3s, g4'
NHb	7.80	2', 2, 3, 3', g2', g3a', g3s, g4'	
NHbb	7.94	2', 2, 3, 3', g2', g3a', g3s, g4'	2', 3', g3a'
NHc	8.16	1a, 2', 3', 3, 6, 9', 10', 11s'	1a, 2', 3', 3, 9', 10'
NHcc	8.20	2', 3', 9', 10', 12'	1a, 2', 3, 3', 3, 9', 10'
NHd	8.40	2'/5', 7', 8a', 8s', 9', 11a', 11s', 12'	7', 8a', 8s', 9', 11a'

c 600 MHz for ^1H .

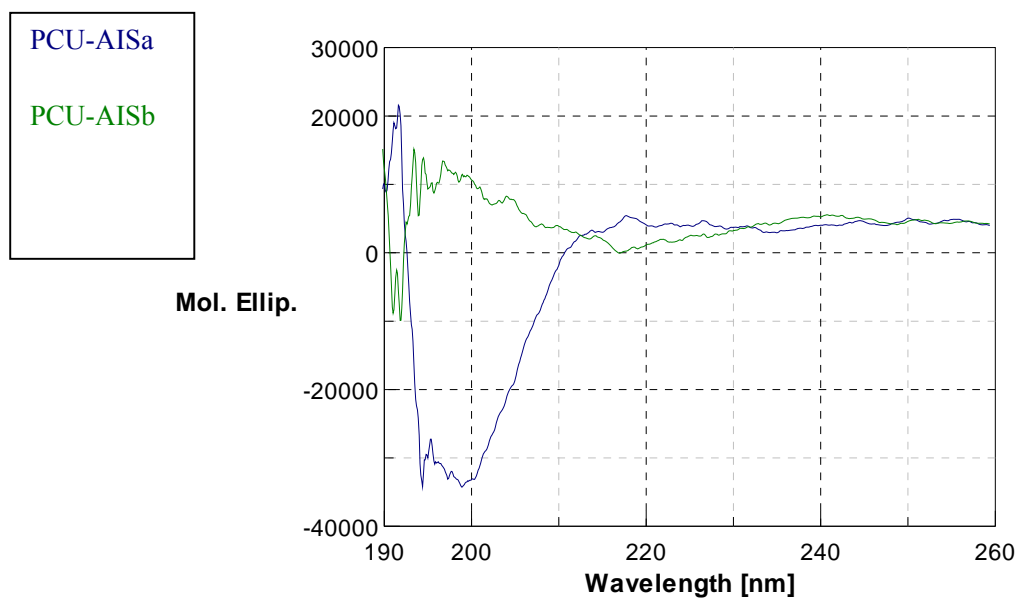


Figure 1. Circular dichroism spectra of compound PCU-AISa in blue and PCU-AISb in green merged

INPUT/SCRIPT FILES

MD Minimization PCU-EAISb 400k.f90

```
PROGRAM QMMM_RELAX_SYSTEM
```

```
USE DYNAMO
```

```
IMPLICIT NONE
```

```
integer          :: i, cnt, A1, A2, A3, A4, A5
```

```
real( kind=dp )   :: d1, d2, d3, d4
```

```
REAL ( KIND=DP ), DIMENSION (1:2) :: COEFF
```

```
CHARACTER ( LEN = 10 )      :: si
```

```
character ( len = 10 ) :: scnt
```

```
logical, dimension(:), allocatable :: quant, flg
```

```
call dynamo_header
```

```
call random_initialize ( 314159 )
```

```
CALL MM_SYSTEM_READ ( 'inh_ext_box.sys_bin' )
```

```
CALL COORDINATES_READ ( 'inh_ext_box.crd' )
```

```
!A1 = ATOM_NUMBER ( ATOM_NAME = 'H41', RESIDUE_NUMBER = 1 , SUBSYSTEM =
'LAALSA' )
```

```
!A2 = ATOM_NUMBER ( ATOM_NAME = 'H7', RESIDUE_NUMBER = 1 , SUBSYSTEM =
'LAALSA' )
```

```
!A3 = ATOM_NUMBER ( ATOM_NAME = 'H37', RESIDUE_NUMBER = 1 , SUBSYSTEM =
'LAALSA' )
```

```
!A4 = ATOM_NUMBER ( ATOM_NAME = 'H12', RESIDUE_NUMBER = 1 , SUBSYSTEM =
'LAALSA' )
```

```
allocate (quant(1:natoms))
```

```
quant = atom_selection( &
```

```
    subsystem = (/ "INH_EXT" /) )
```

```
! quant = quant .or. atom_selection( &
```

```
!     atom_name = (/ "CB", "HB1", "HB2", "HG2", "HG1", "CG", "CD", "OE1", "OE2" /), &
```

```
!     residue_number = (/ 78 /), &
```

```
!     subsystem = (/ "PROTEIN" /) )
```

```
! quant = quant .or. atom_selection( &
```

```
!     atom_name = (/ "CB", "HB1", "HB2", "HG2", "HG1", "CG", "CD", "OE1", "OE2"/), &
```

```
!     residue_number = (/ 172 /), &
```

```
!     subsystem = (/ "PROTEIN" /) )
```

```
! quant = quant .or. atom_selection( &
```

```
!      residue_number = (/ 2275 /), &
!      subsystem      = (/ "WAT" /))

! keep the following atom fixed

!      call atoms_fix ( atom_selection ( subsystem = (/ 'ACS' /)))
!      call atoms_fix ( atom_selection ( subsystem = (/ 'PROTEIN' /)))
!      call atoms_fix ( atom_selection ( subsystem = (/ 'WAT' /)))

call mopac_setup ( &
    method      = 'AM1', &
    charge      = 0, &
    selection    = quant )
!      link_atom_distances = (/ 1.1_dp, 1.1_dp /))

CALL ENERGY_NON_BONDING_OPTIONS ( &
    LIST_CUTOFF = 15.5_dp, &
    OUTER_CUTOFF = 13.5_dp, &
```

```
INNER_CUTOFF = 11.5_dp, &
```

```
MINIMUM_IMAGE = .true. )
```

```
CALL ENERGY_INITIALIZE
```

```
CALL GRADIENT
```

```
!call constraint_initialize
```

```
!      call constraint_point_define( &
```

```
!          atom_selection( atom_number = (/ A1 /) ) )
```

```
!      call constraint_point_define( &
```

```
!          atom_selection( atom_number = (/ A2 /) ) )
```

```
!      d1 = 3.400_dp
```

```
!      call constraint_define( &
```

```
!          type = 'DISTANCE', &
```

```
!          fc = 10000._dp, &
```

```
!          eq = d1 )
```

```
!
```

```
!      call constraint_point_define( &
```

```
!          atom_selection( atom_number = (/ A3 /) ) )
```

```
!      call constraint_point_define( &
```

```
!          atom_selection( atom_number = (/ A4 /) ) )
```

```
!      d2 = 3.400_dp
```

```
!      call constraint_define( &
```



```
!      type = 'DISTANCE', &  
!  
!      fc = 10000._dp, &  
!  
!      eq = d2 )
```

```
CALL OPTIMIZE_CONJUGATE_GRADIENT ( &
```

```
    STEP_SIZE    = 0.001_dp, &
```

```
    PRINT_FREQUENCY = 100, &
```

```
    STEP_NUMBER   = 30000 )
```

```
CALL COORDINATES_WRITE ('opt_LaEAIb_from_ext.crd')
```

```
CALL PDB_WRITE      ('opt_LaEAIb_from_ext.pdb')
```

```
call velocity_assign ( tneeded = 400._dp, remove_translation = .false. )
```

```
    call dynamics_options ( &
```

```
        time_step    = 0.001_dp, &
```

```
        steps        = 1000000, &
```

```
        save_frequency = 100 , &
```

```
        coordinate_file = '400K_LaEAIb_ext.trj.dcd', &
```

```
        print_frequency = 100 )
```

```
call constraint_writing_start
```

```
call langevin_verlet_dynamics ( tbath = 400._dp, gamma = 100._dp )
```

```
call constraint_writing_stop
```

```
deallocate( quant )
```

```
CALL COORDINATES_WRITE ('400K_rel_LaEASb_from_ext.crd')
```

```
CALL PDB_WRITE      ('400K_rel_LaEASb_from_ext.pdb')
```

```
END
```

QM-energy PCU-EAISb 400k.f90

```
PROGRAM QM
```

```
USE DYNAMO
```

```
IMPLICIT NONE
```

```
TYPE (DCD_TYPE) :: TRJ
```

```
logical, dimension(:), allocatable :: flg, quant
```

```
INTEGER      :: NF, I
```

```
OPEN (UNIT=70, FILE="qm.energy", STATUS="UNKNOWN")
```

```
call dynamo_header
```

```
CALL MM_SYSTEM_READ ( "inh_ext_box.sys_bin" )
```

```
allocate (quant(1:natoms))
```

```
quant = atom_selection( &
subsystem = (/ "INH_EXT" /))

call mopac_setup ( &
method      = 'AM1', &
charge      = 0, &
selection   = quant )
! link_atom_distances = (/ 1.1_dp, 1.1_dp /))

CALL ENERGY_NON_BONDING_OPTIONS ( &
LIST_CUTOFF = 15.5_dp, &
OUTER_CUTOFF = 13.5_dp, &
INNER_CUTOFF = 11.5_dp, &
MINIMUM_IMAGE = .true. )

CALL DCD_INITIALIZE ( TRJ )
! Put the trajectory file
CALL DCD_ACTIVATE_READ ( "400K_LaEAIb_ext.trj.dcd", TRJ )
```

```
NF = TRJ%NFRAMES
```

```
CALL SYMMETRY_CUBIC_BOX ( 34.5_DP )
```

```
DO I = 1, NF
```

```
    CALL DCD_READ (DCD = TRJ, BOXDAT = BOXL, ATMDAT = ATMCRD )
```

```
    CALL ENERGY
```

```
    write(70,*) eqm
```

```
END DO
```

```
END
```

Dihedral PCU-EAISb 400k.f90

Last login: Fri Mar 11 17:04:18 on ttys000

```
/Volumes/SABELO/Maye\ refree/dihedral-PCU-EAISb-400k-f90 ; exit;
```

```
Administrators-Mac-mini:~ Mac$ /Volumes/SABELO/Maye\ refree/dihedral-PCU-EAISb-400k-f90 ;  
exit;
```

```
/Volumes/SABELO/Maye refree/dihedral-PCU-EAISb-400k-f90: line 1: PROGRAM: command not  
found
```

```
/Volumes/SABELO/Maye refree/dihedral-PCU-EAISb-400k-f90: line 2: USE: command not found
```

```
/Volumes/SABELO/Maye refree/dihedral-PCU-EAISb-400k-f90: line 3: IMPLICIT: command not found
```

```
/Volumes/SABELO/Maye refree/dihedral-PCU-EAISb-400k-f90: line 4: syntax error near unexpected  
token `DCD_TYPE'
```

```
/Volumes/SABELO/Maye refree/dihedral-PCU-EAISb-400k-f90: line 4: `TYPE (DCD_TYPE) :: TRJ'
```

```
logout
```

```
[Process completed]
```

PDB FILES OF THE PCU-PEPTIDE ENZYME COMPLEXES

See the attached folder

ELECTROSTATIC AND HYDROGEN BOND INTERACTIONS OF HIV-PR COMPLEXES FROM THE MD SIMULATION.

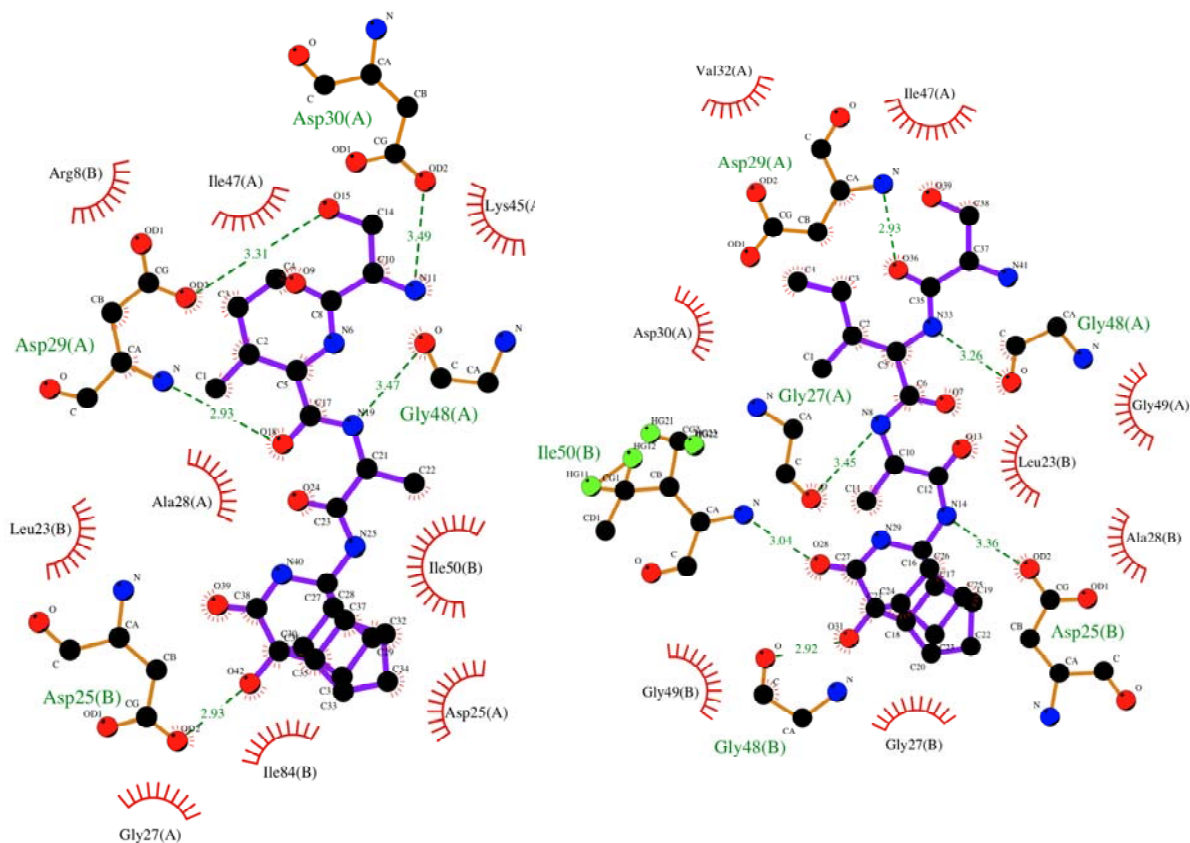


Figure 2. Electrostatic and hydrogen bond interactions for the PCU-AISa (CP) and PCU-AISa (PP) CSA HIV-PR complexes from the MD simulation. These plots were created with the Ligplot software.

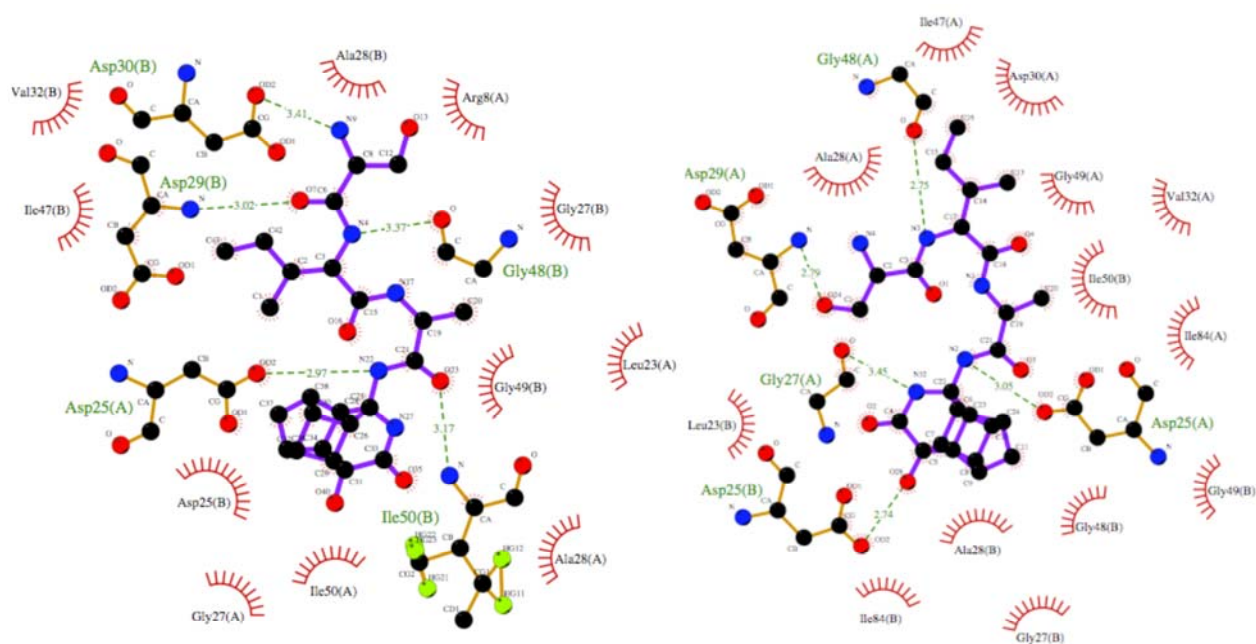


Figure 3. Electrostatic and hydrogen bond interactions for the PCU-AISb (CP) and PCU-AISb (PP) CSA HIV-PR complexes from the MD simulation. These plots were created with the Ligplot software.

RAMACHANDRAN PLOTS

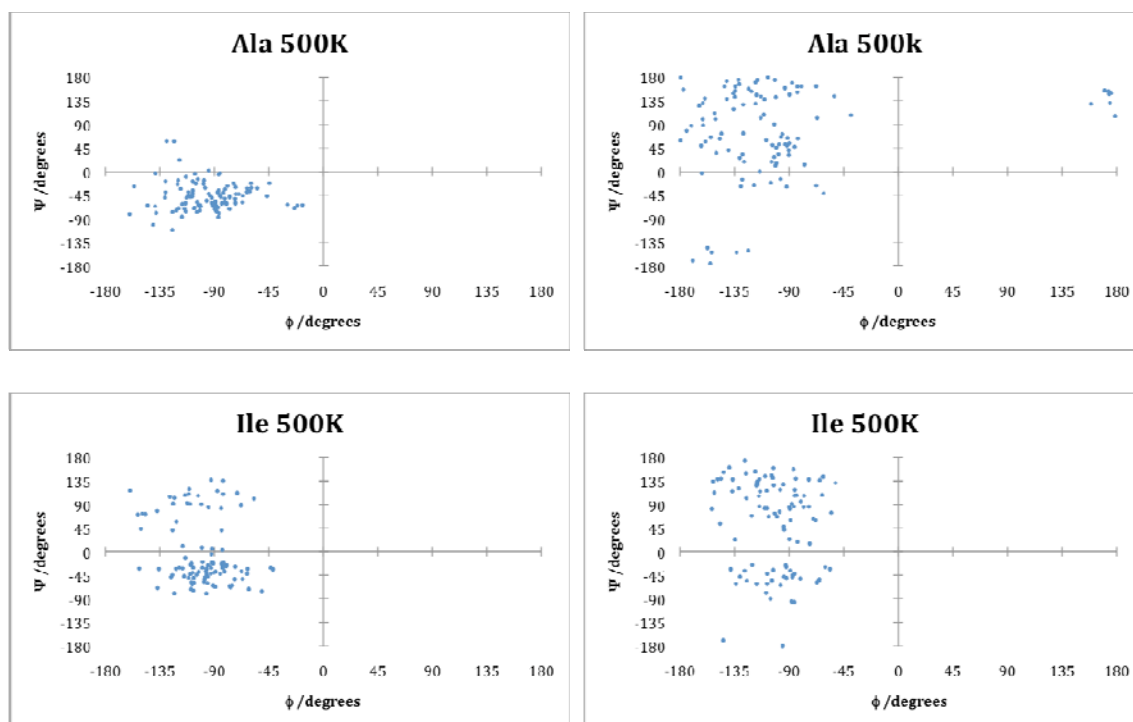
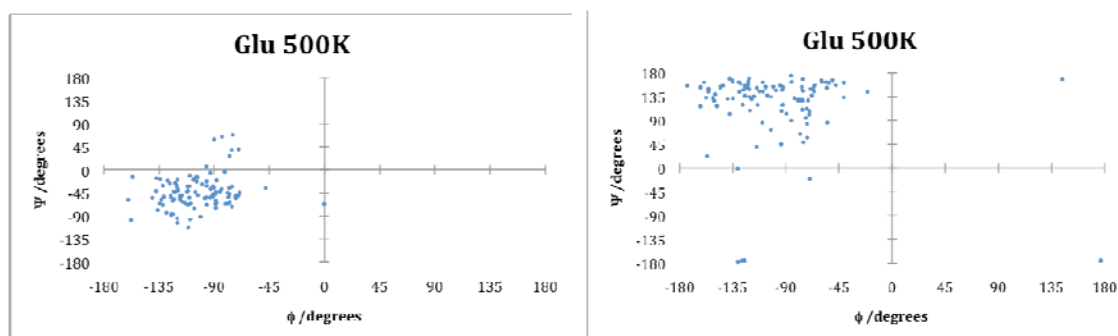


Figure 4. Ramachandran plots for PCU-AISa (left) and PCU-AISb (right) at 500 K. The 100 lowest energy structures were plotted.



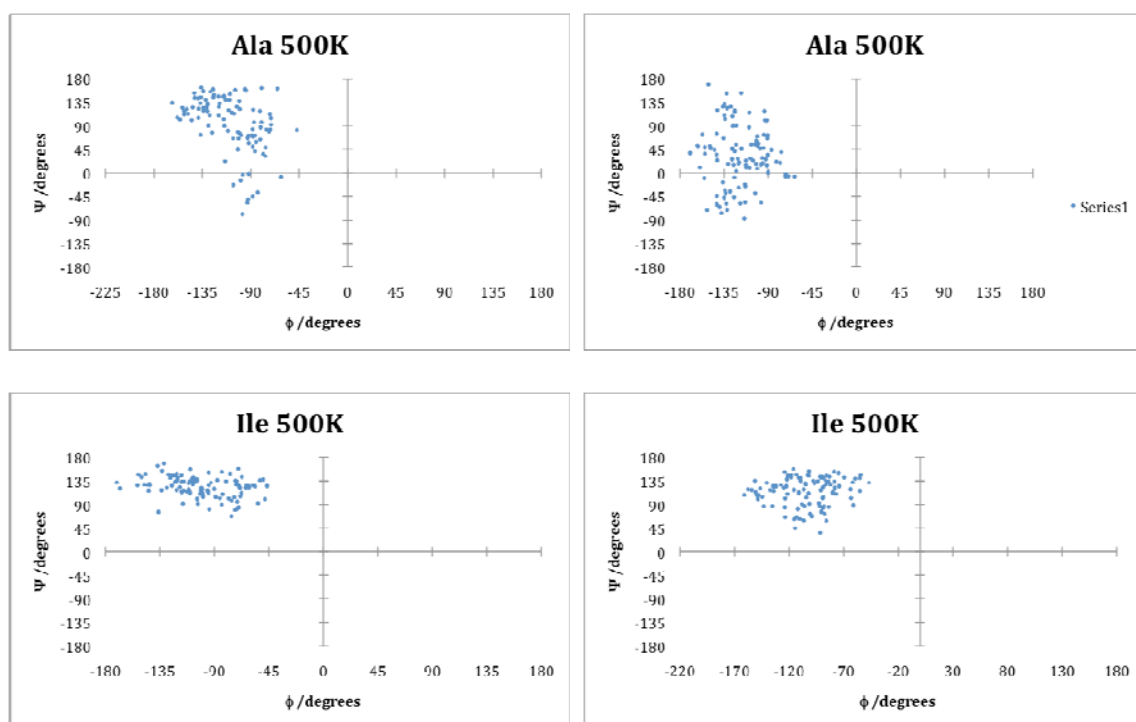


Figure 5. Ramachandran plots for PCU-EAISa (left) and PCU-EAISb (right) at 500 K. The 100 lowest energy structures were plotted.

Display Report

Analysis Info

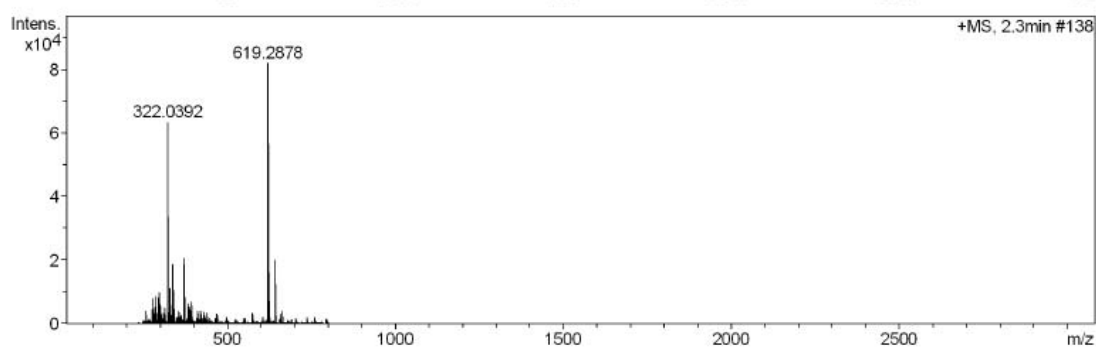
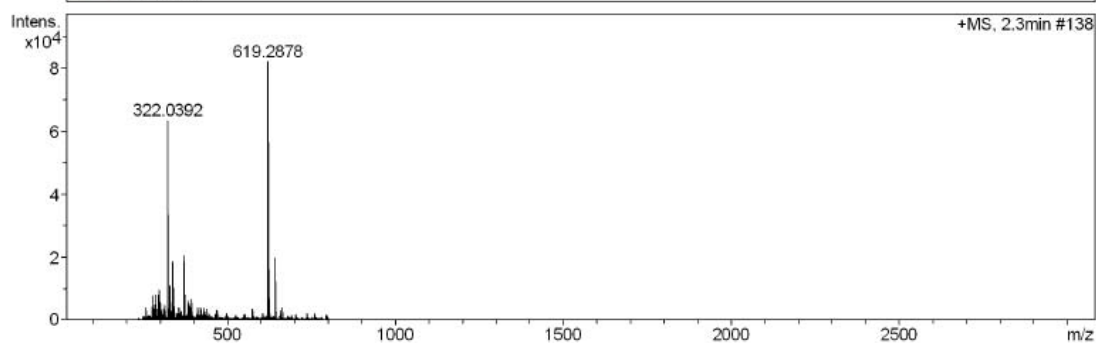
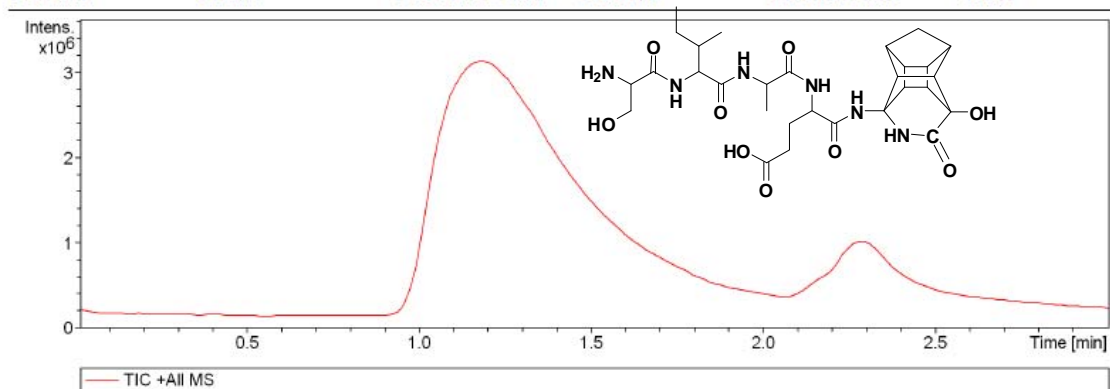
Analysis Name D:\Data\Sri\lactam-glu-ala-Ile-Ser2.d
Method tune_wide_expert.m
Sample Name lactam-glu-ala-Ile-Ser2
Comment

Acquisition Date 12/24/2009 11:23:23 AM

Operator BDAL@DE
Instrument micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	800 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



Display Report

Analysis Info

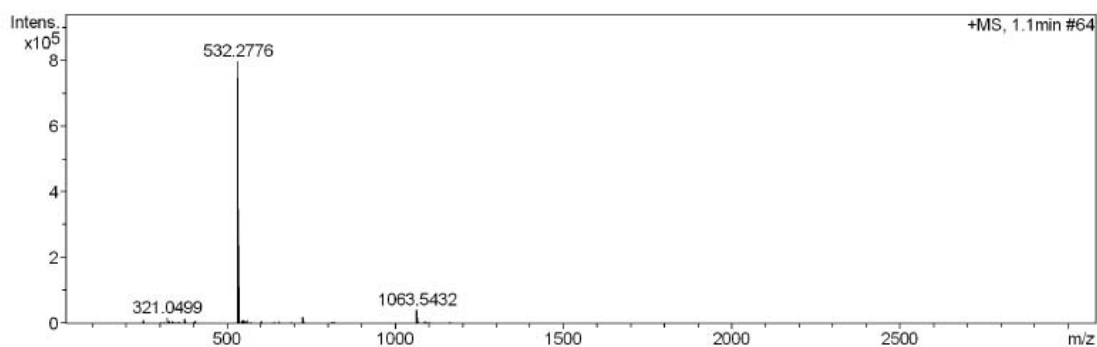
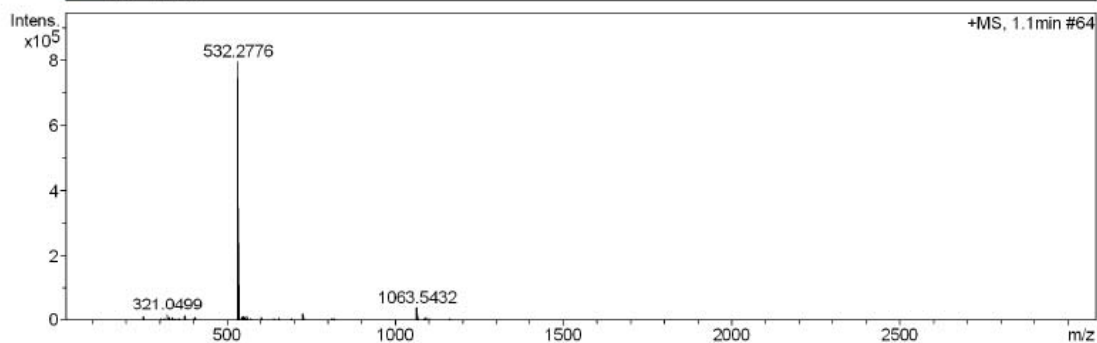
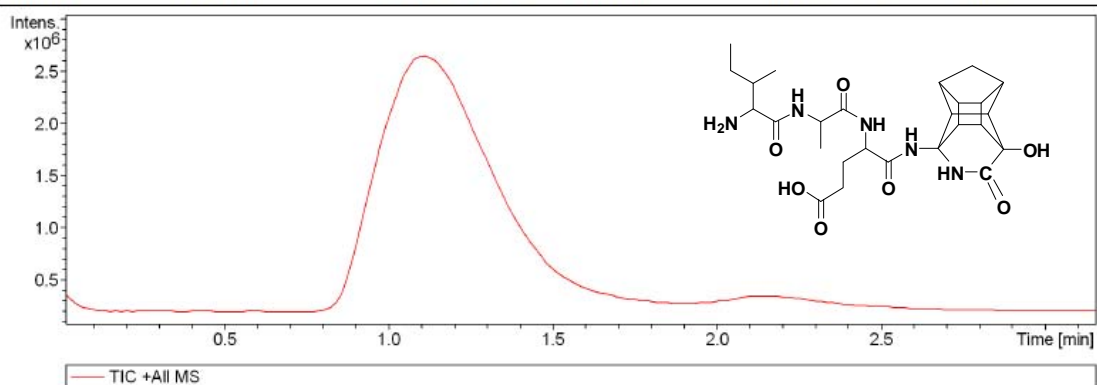
Analysis Name D:\Data\Sri\lactam-Glu-Ala-Ile.d
Method tune_wide_expert.m
Sample Name lactam-Gln-Ala-Ile
Comment

Acquisition Date 12/23/2009 3:21:48 PM

Operator BDAL@DE
Instrument micrOTOF-Q 10139

Acquisition Parameter

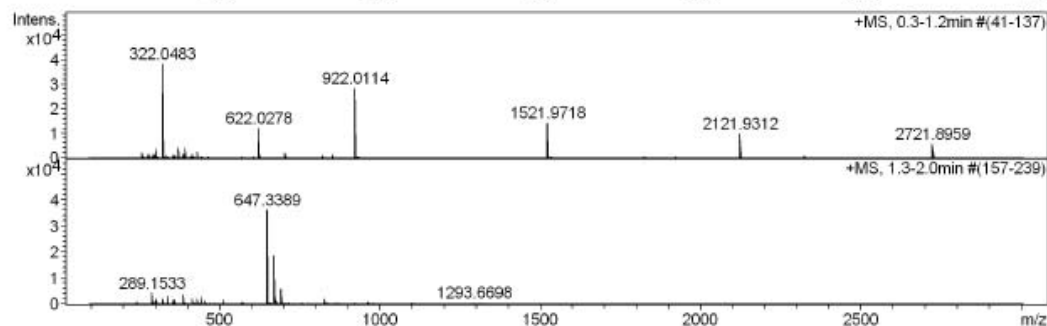
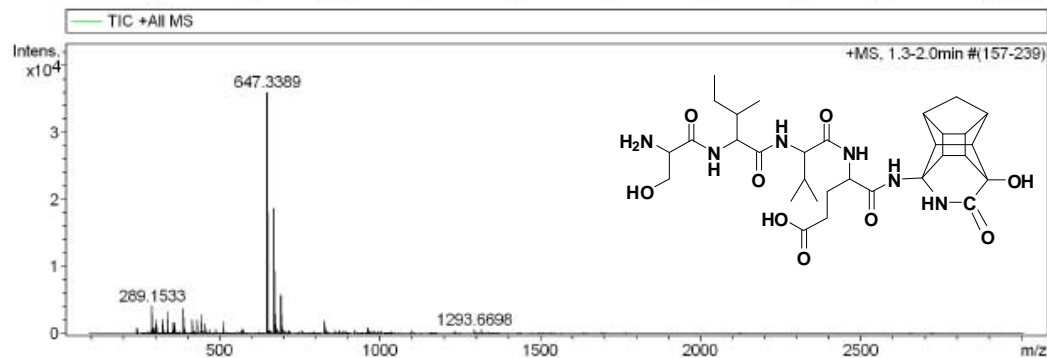
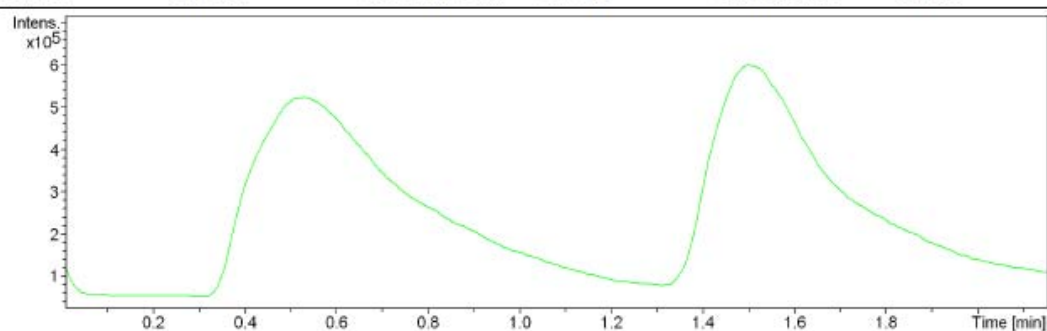
Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



Display Report

Analysis Info		Acquisition Date	5/24/2010 8:33:18 PM
Analysis Name	D:\Data\maya\LGVIS.d	Operator	BDAL@DE
Method	tune_wide_expert.m	Instrument	micrOTOF-Q 10139
Sample Name	LGVIS		
Comment			

Acquisition Parameter					
Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



Display Report

Analysis Info

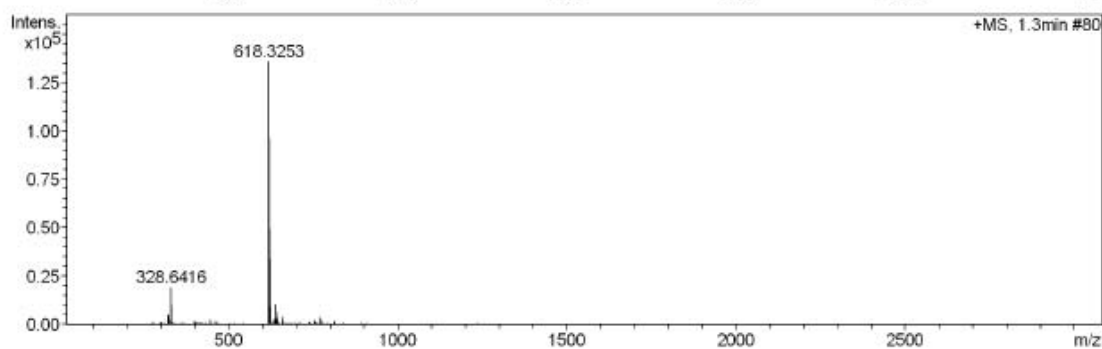
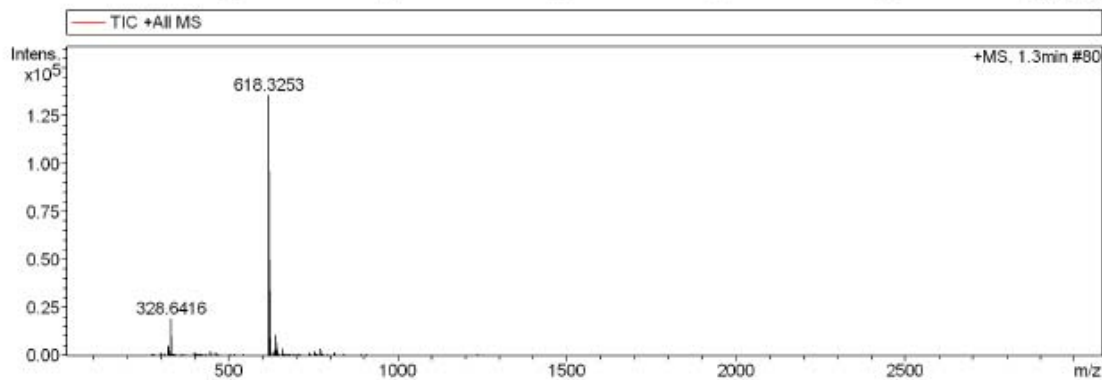
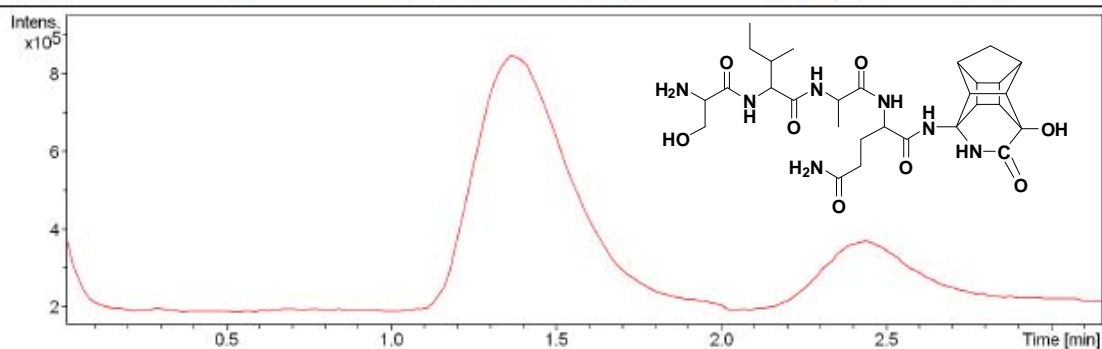
Analysis Name D:\Data\Srnlactam-Gln-Ala-Ile-Ser.d
Method tune_wide_expert.m
Sample Name lactam-Gln-Ala-Ile-Ser
Comment

Acquisition Date 12/23/2009 3:11:45 PM

Operator BDAL@DE
Instrument micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



Display Report

Analysis Info

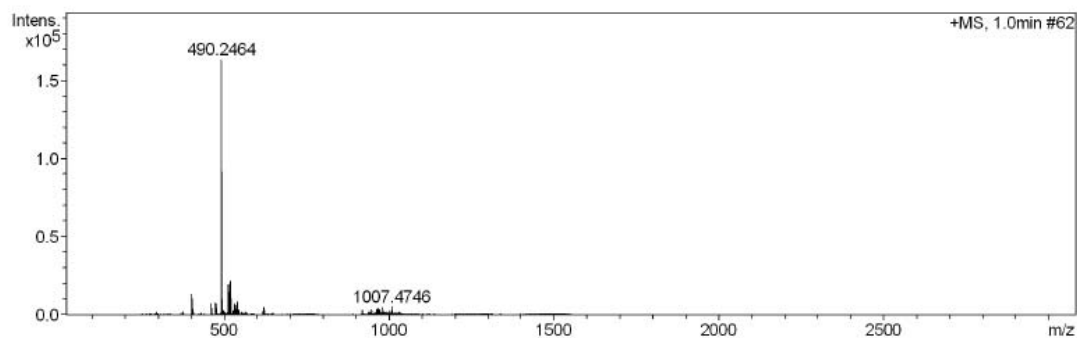
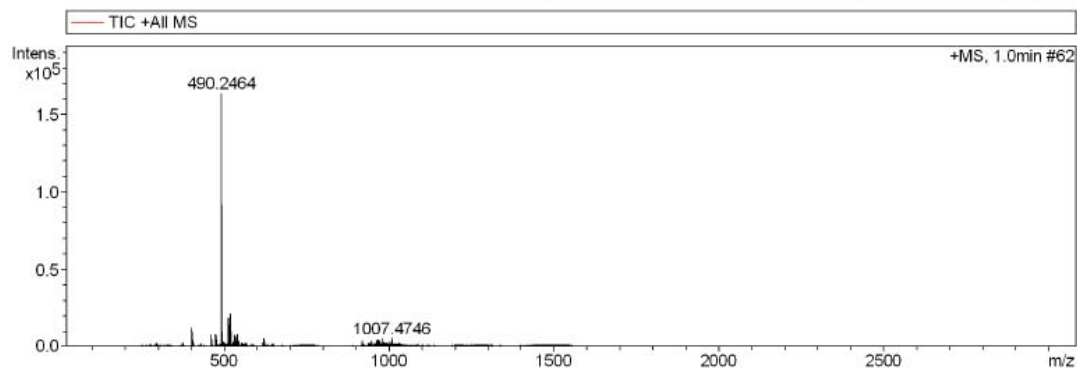
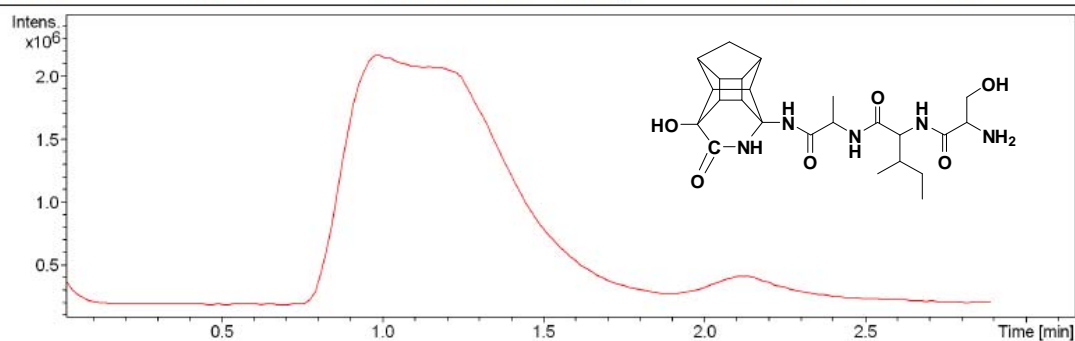
Analysis Name D:\Data\Sri\lactam-Ala-Ile-Ser-A.d
Method tune_wide_expert.m
Sample Name lactam-Ala-Ile-Ser-A
Comment

Acquisition Date 12/23/2009 3:32:56 PM

Operator BDAL@DE
Instrument micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



Analysis Info

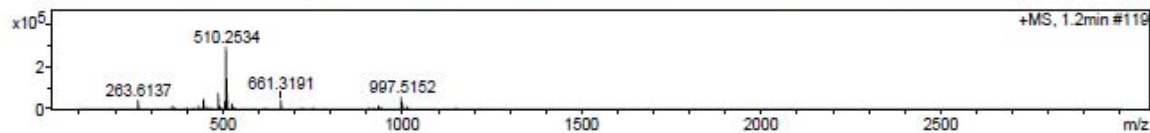
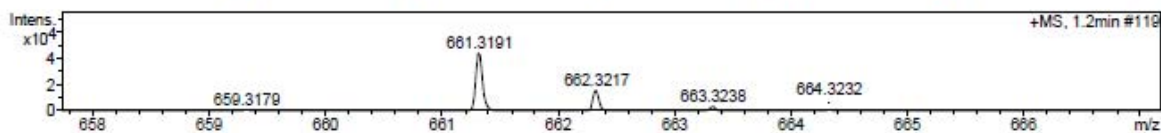
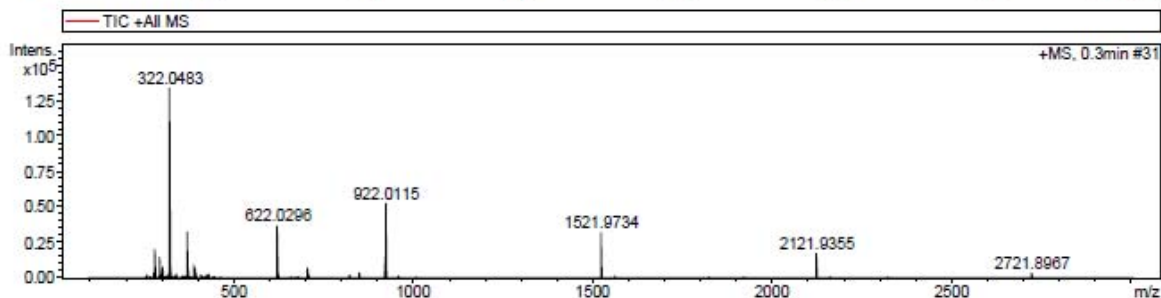
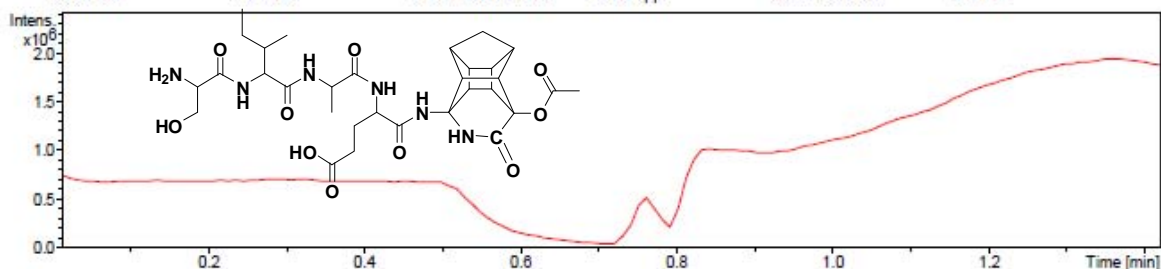
Analysis Name D:\Data\maya\Ac-PCU-EAIS.d
 Method tune_wide_expert.m
 Sample Name Ac-PCU-EAIS
 Comment

Acquisition Date 10/9/2010 4:01:10 PM

Operator BDAL@DE
 Instrument / Ser# micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	14.5 psi
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



Meas. #	Formula	m/z	err [ppm]	Mean err [ppm]	rdB	N-Rule	e ⁻ Conf	mSigma	Std I	Std Mean m/z	Std I VarNo	Std I m/z Diff	Std I Comb Dev
661.3191													
1	C ₂₇ H ₄₁ N ₁₂ O ₈	661.3165	-3.9	-3.5	13.5	ok	even	6.66	0.0116	0.0024	0.0040	0.0012	0.7973
2	C ₃₀ H ₄₉ N ₂ O ₁₄	661.3178	-1.9	-1.4	7.5	ok	even	7.14	0.0099	0.0011	0.0043	0.0012	0.7012
3	C ₂₈ H ₃₇ N ₁₆ O ₄	661.3178	-1.9	-1.4	18.5	ok	even	7.52	0.0101	0.0012	0.0045	0.0012	0.7194

HRMS spectrum of Ac-PCU-EAIS

Meas. m/z	#	Formula	m/z	err [pp m]	Mea n err [pp m]	rdb	N- Ru le	e ⁻ Conf	mSig ma	Std l	Std Mean m/z	Std l VarNo rm	Std m/z Diff	Std Comb Dev
	4	C 31 H 45 N 6 O 10	661.3192	0.1	0.6	12.5	ok	even	8.39	0.0133	0.0007	0.0047	0.0012	0.6602
	5	C 26 H 45 N 8 O 12	661.3151	-5.9	-5.5	8.5	ok	even	16.60	0.0261	0.0037	0.0086	0.0012	0.9086
	6	C 32 H 41 N 10 O 6	661.3205	2.2	2.6	17.5	ok	even	18.05	0.0260	0.0019	0.0084	0.0012	0.8401
	7	C 29 H 33 N 20	661.3192	0.1	0.6	23.5	ok	even	19.64	0.0259	0.0008	0.0083	0.0012	0.7705
	8	C 35 H 49 O 12	661.3219	4.2	4.7	11.5	ok	even	21.16	0.0337	0.0032	0.0112	0.0013	0.9144
	9	C 33 H 37 N 14 O 2	661.3218	4.2	4.6	22.5	ok	even	30.47	0.0423	0.0032	0.0135	0.0012	0.9274
	10	C 42 H 45 O 7	661.3160	-4.7	-4.1	20.5	ok	even	57.25	0.0821	0.0028	0.0264	0.0013	0.9645
	11	C 43 H 41 N 4 O 3	661.3173	-2.7	-2.1	25.5	ok	even	69.62	0.0966	0.0016	0.0306	0.0012	0.9527
	12	C 48 H 41 N 2 O	661.3213	3.4	4.0	29.5	ok	even	94.55	0.1265	0.0027	0.0408	0.0013	0.9830

HRMS spectrum of Ac-PCU-EAIS

Analysis Info

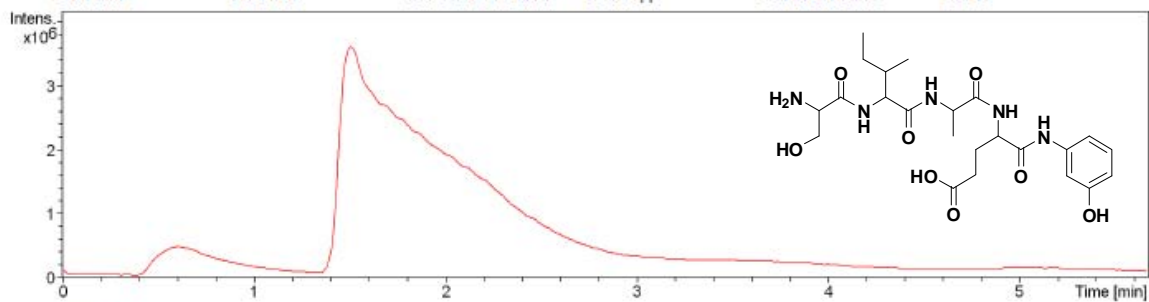
Analysis Name D:\Data\maya\PHe-GAIS2.d
 Method tune_wide_expert.m
 Sample Name PHE-GAIS2
 Comment

Acquisition Date 5/24/2010 8:54:35 PM

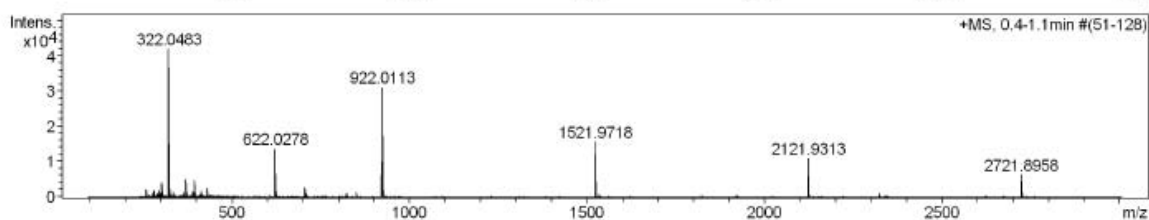
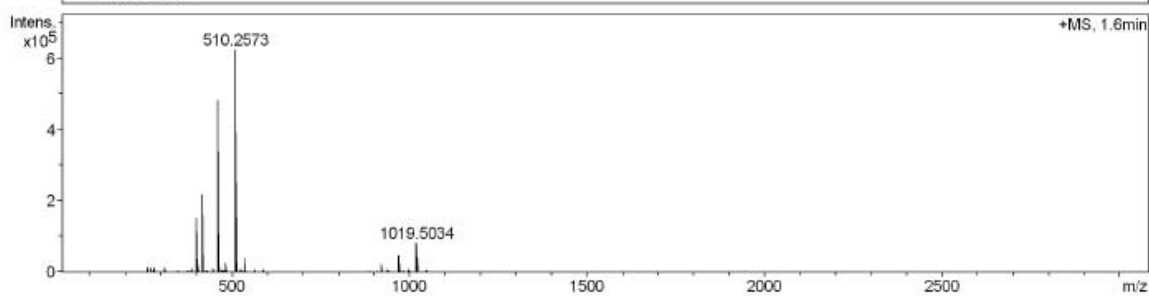
Operator BDAL@DE
 Instrument / Ser# micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



TIC +All MS

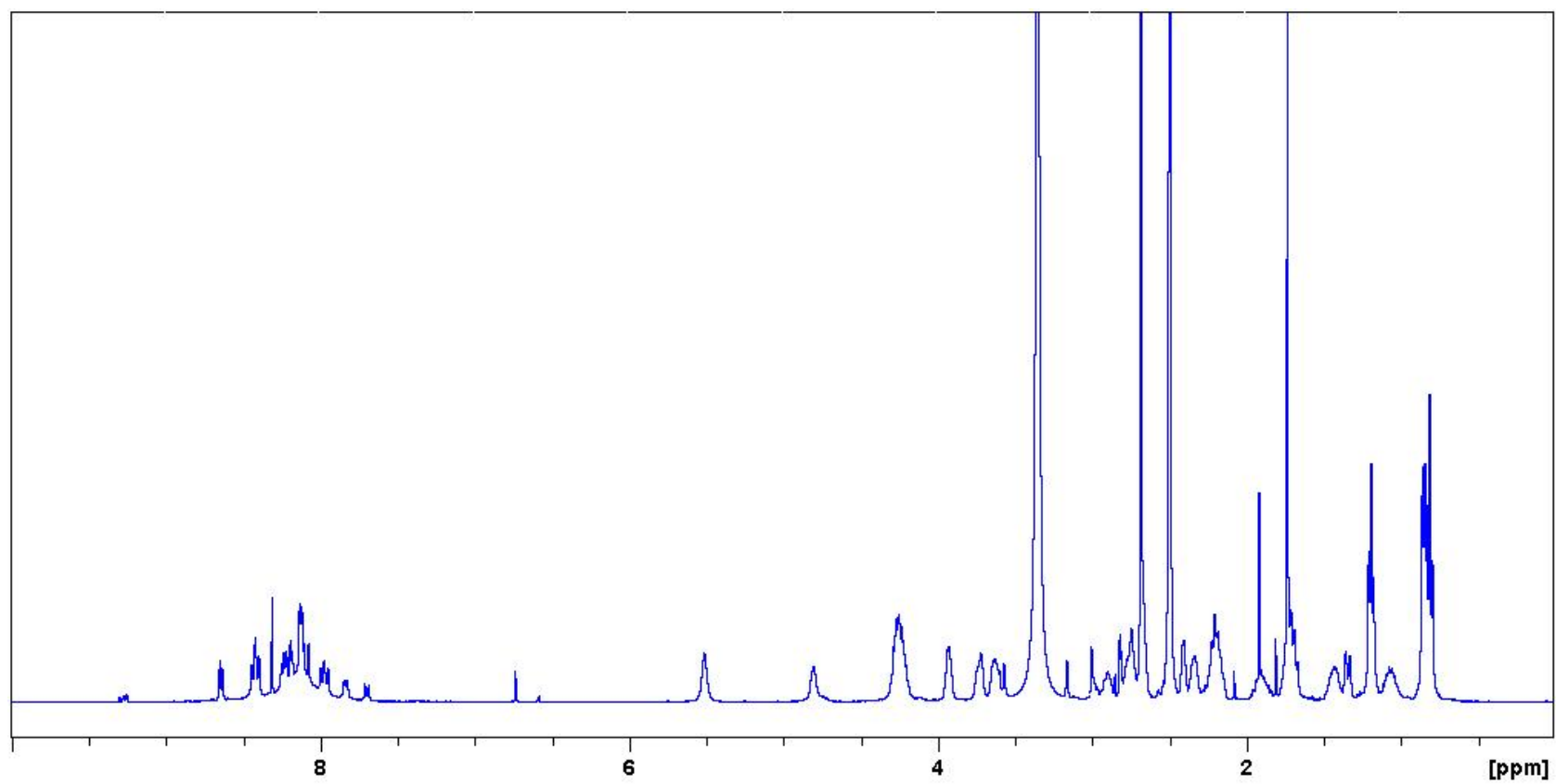


Meas. m/z	#	Formula	m/z	err [ppm]	Mean err [ppm]	rdB	N-Rule	e ⁻ Conf	mSig ma	Std I	Std Mean m/z	Std I VarNo rm	Std m/z Diff	Std Comb Dev
-----------	---	---------	-----	-----------	----------------	-----	--------	---------------------	---------	-------	--------------	----------------	--------------	--------------

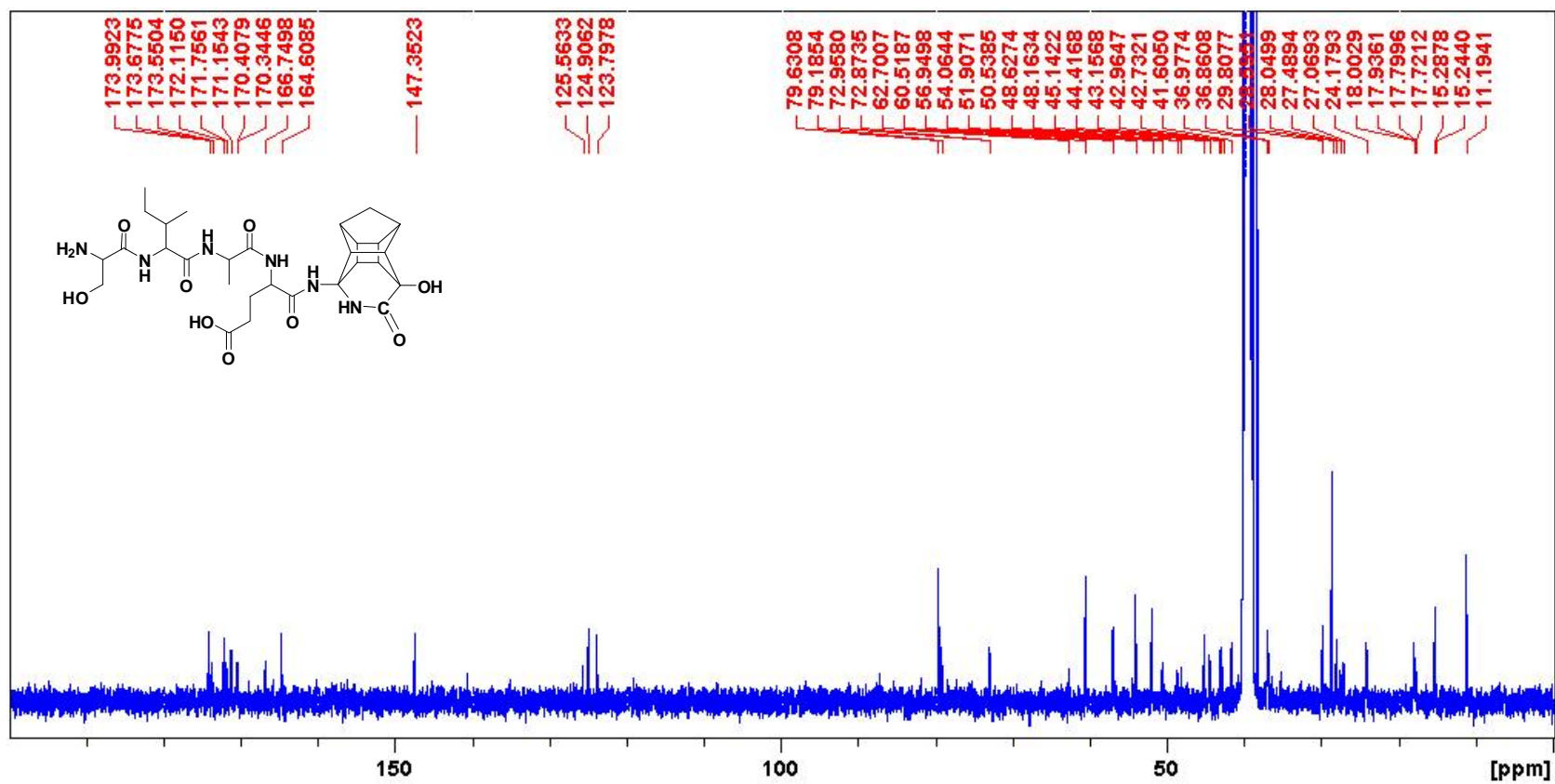
HRMS spectrum of Phenol-EAIS

REFERENCES

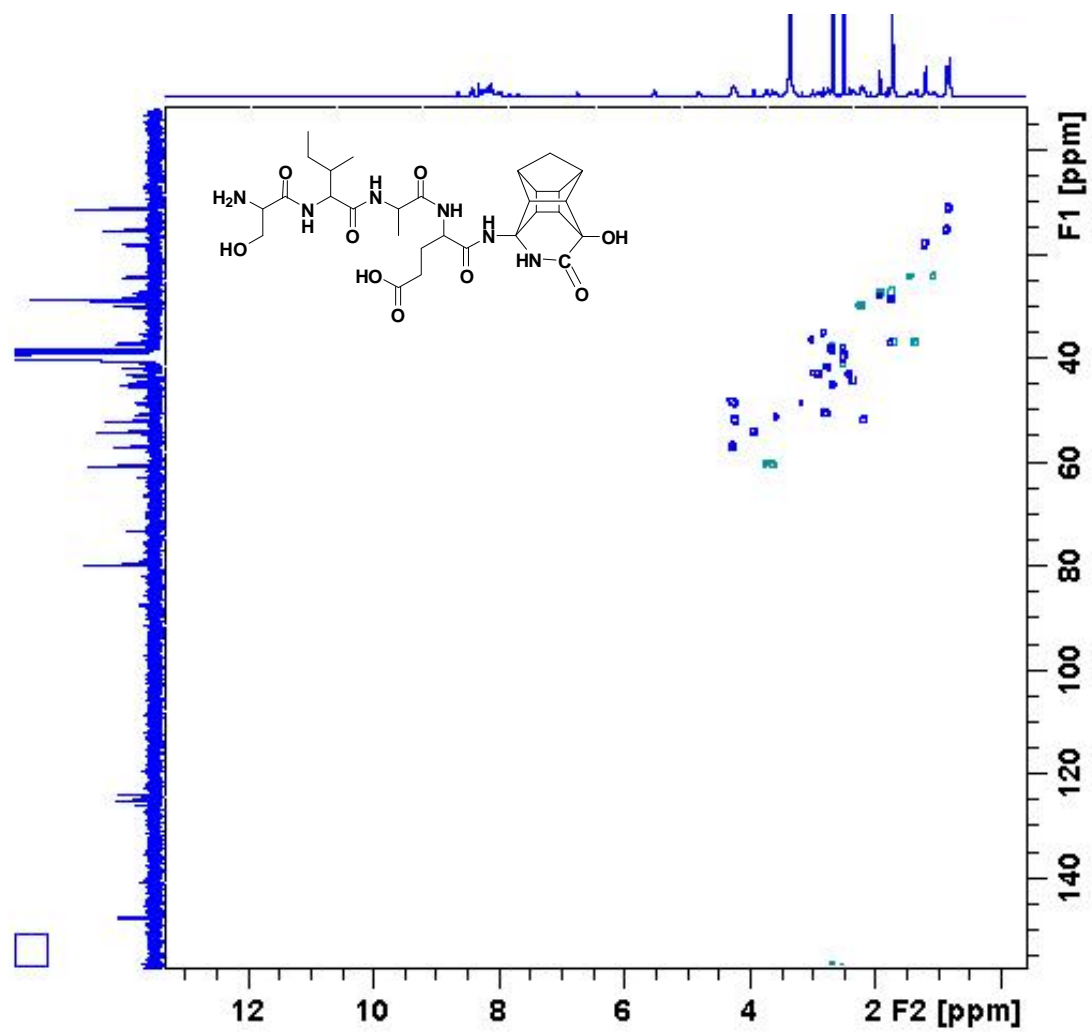
1. White WCCaPD (2004) *Fmoc solid phase peptide synthesis: A practical approach* (Oxford University Press, New York).
2. Ido E, Han HP, Kezdy FJ, & Tang J (1991) *Journal of Biological Chemistry* **266**, 24359-24366.

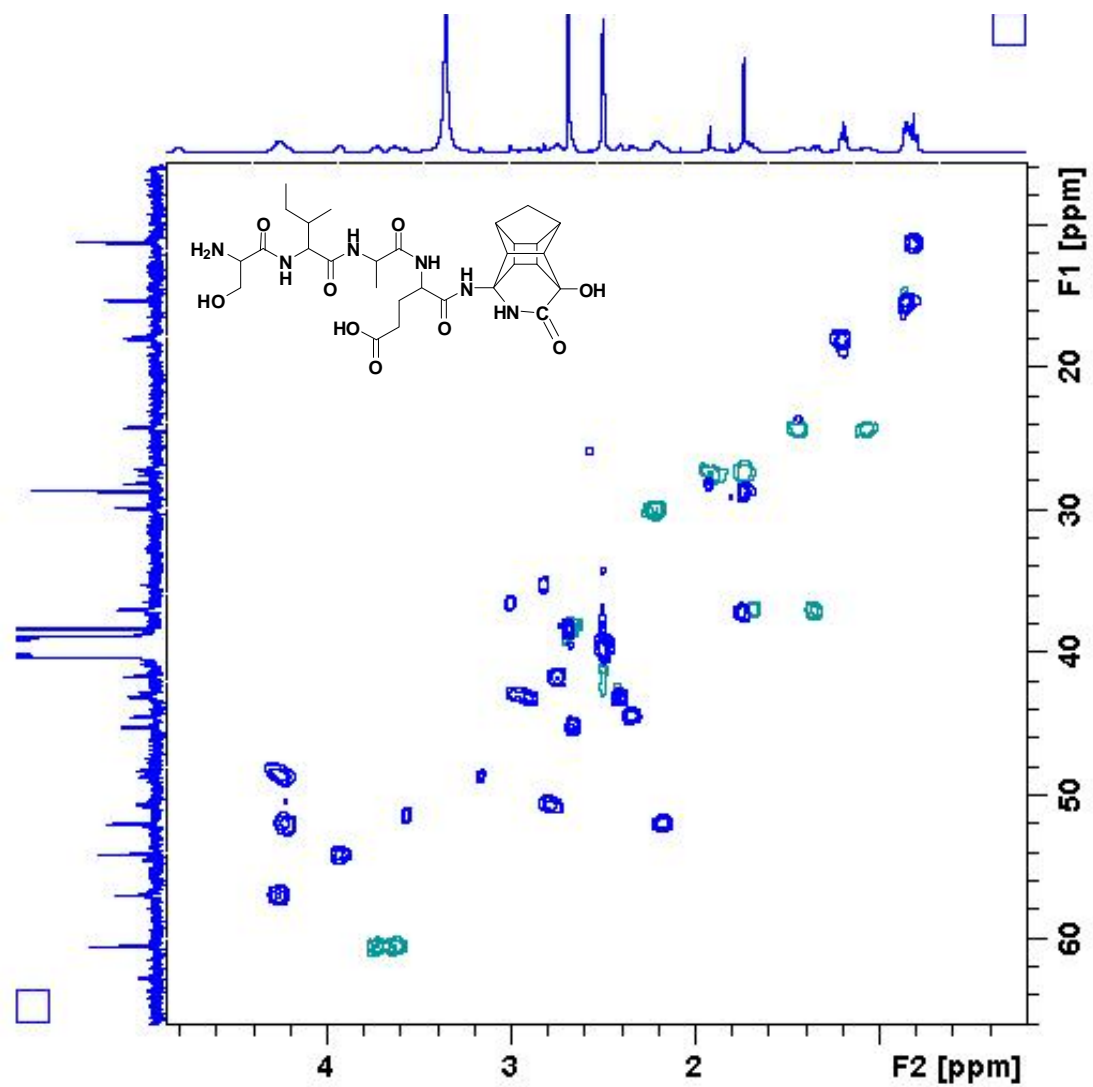


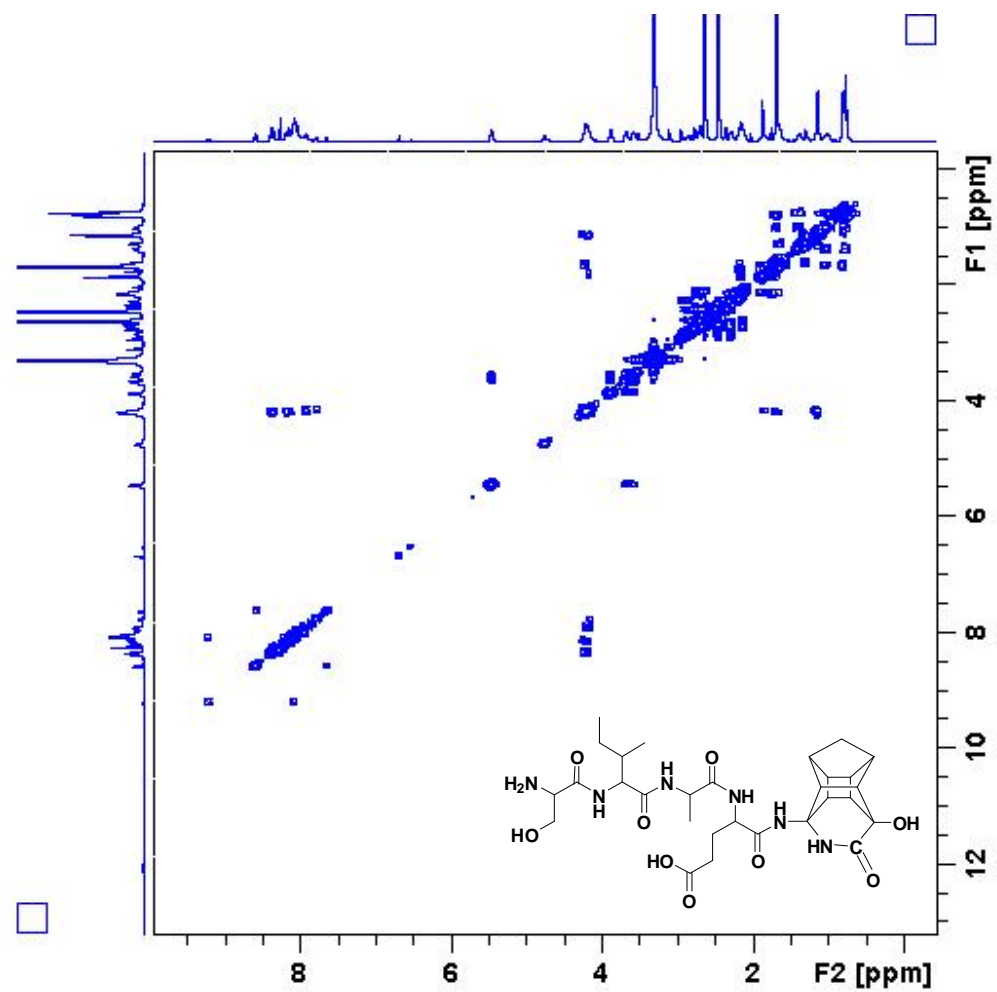
^1H NMR spectrum of PCU-EAIS in $(\text{CD}_3)_2\text{SO}$

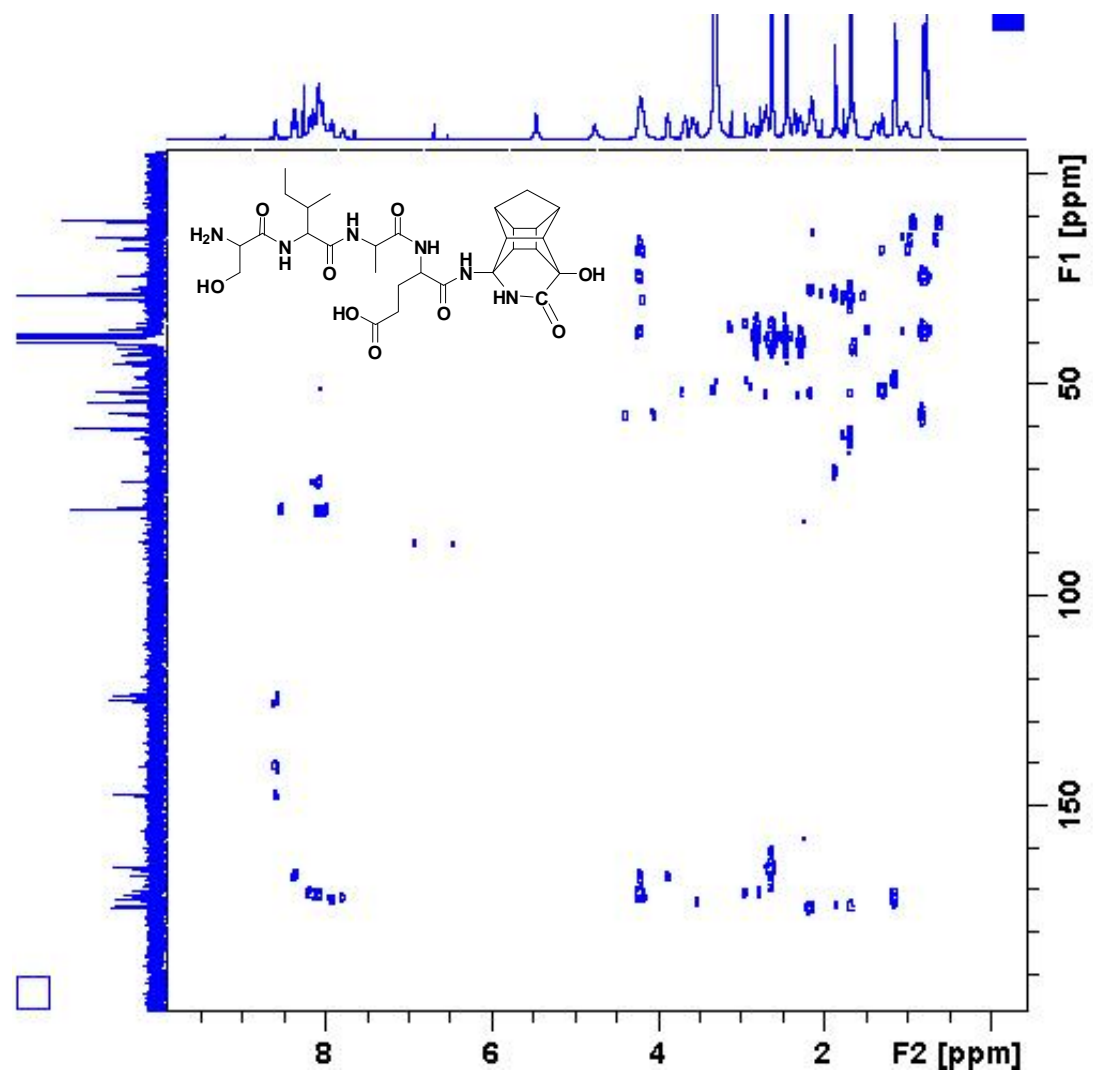


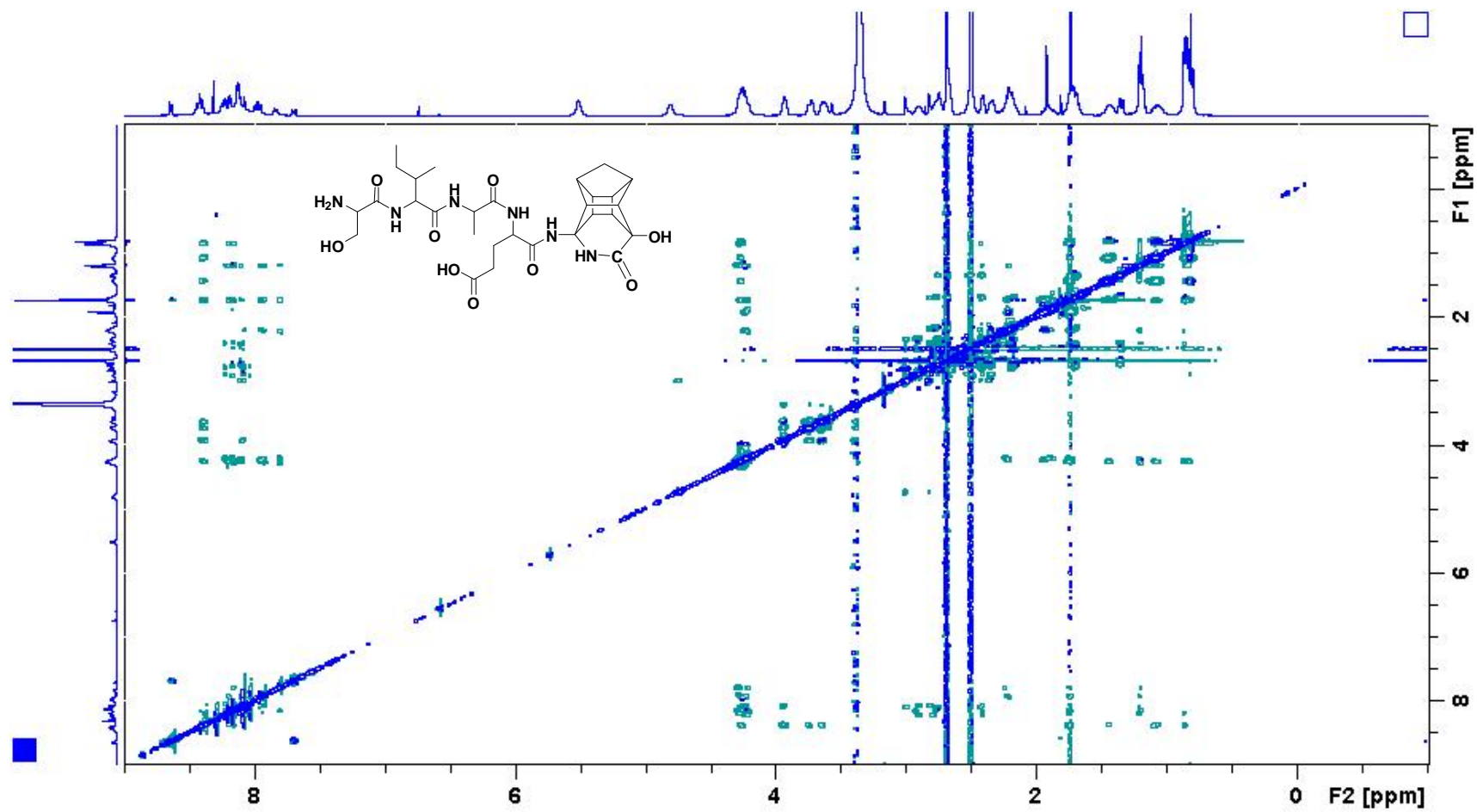
^{13}C NMR spectrum of PCU-EAIS in $(\text{CD}_3)_2\text{SO}$

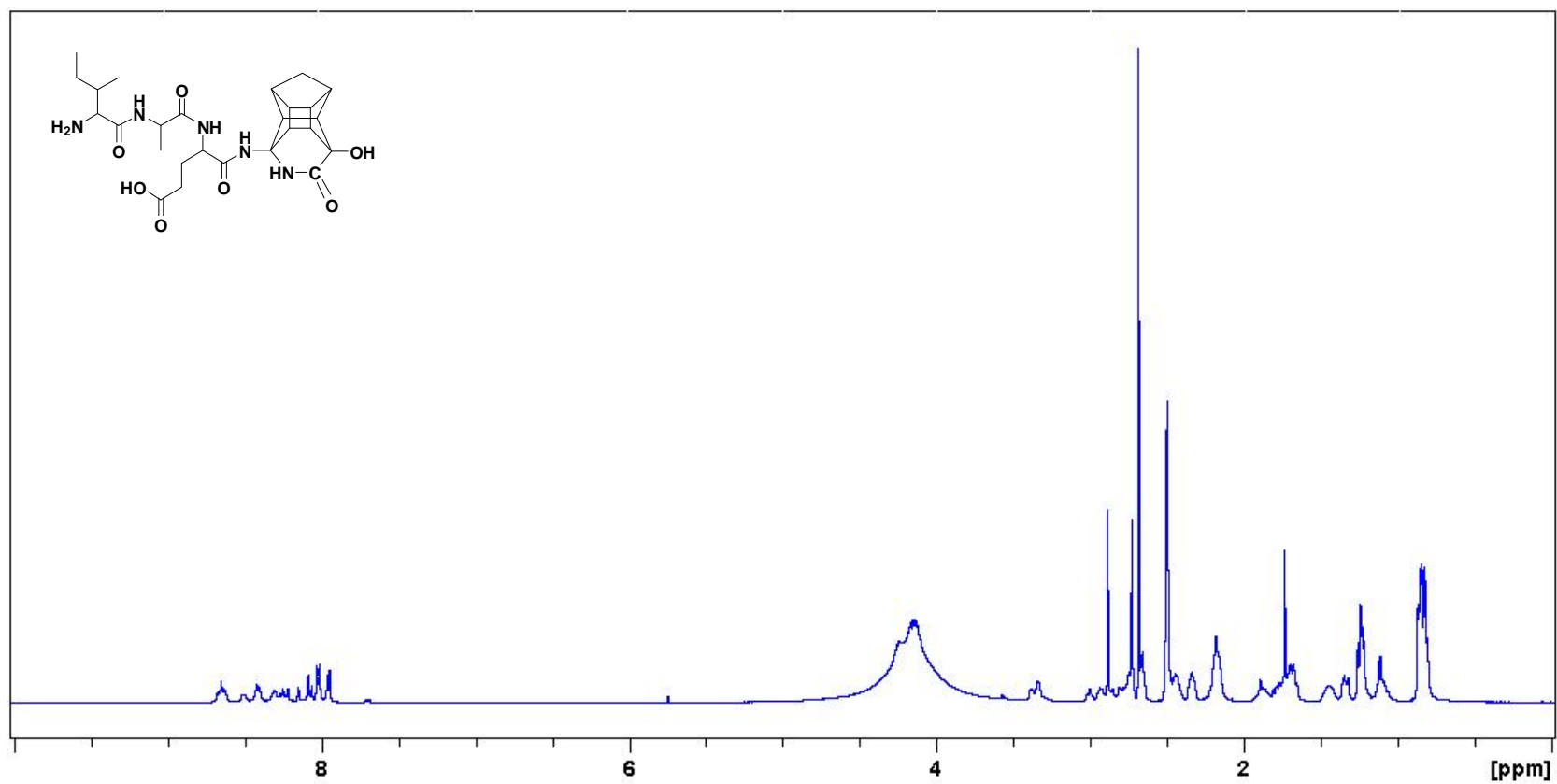




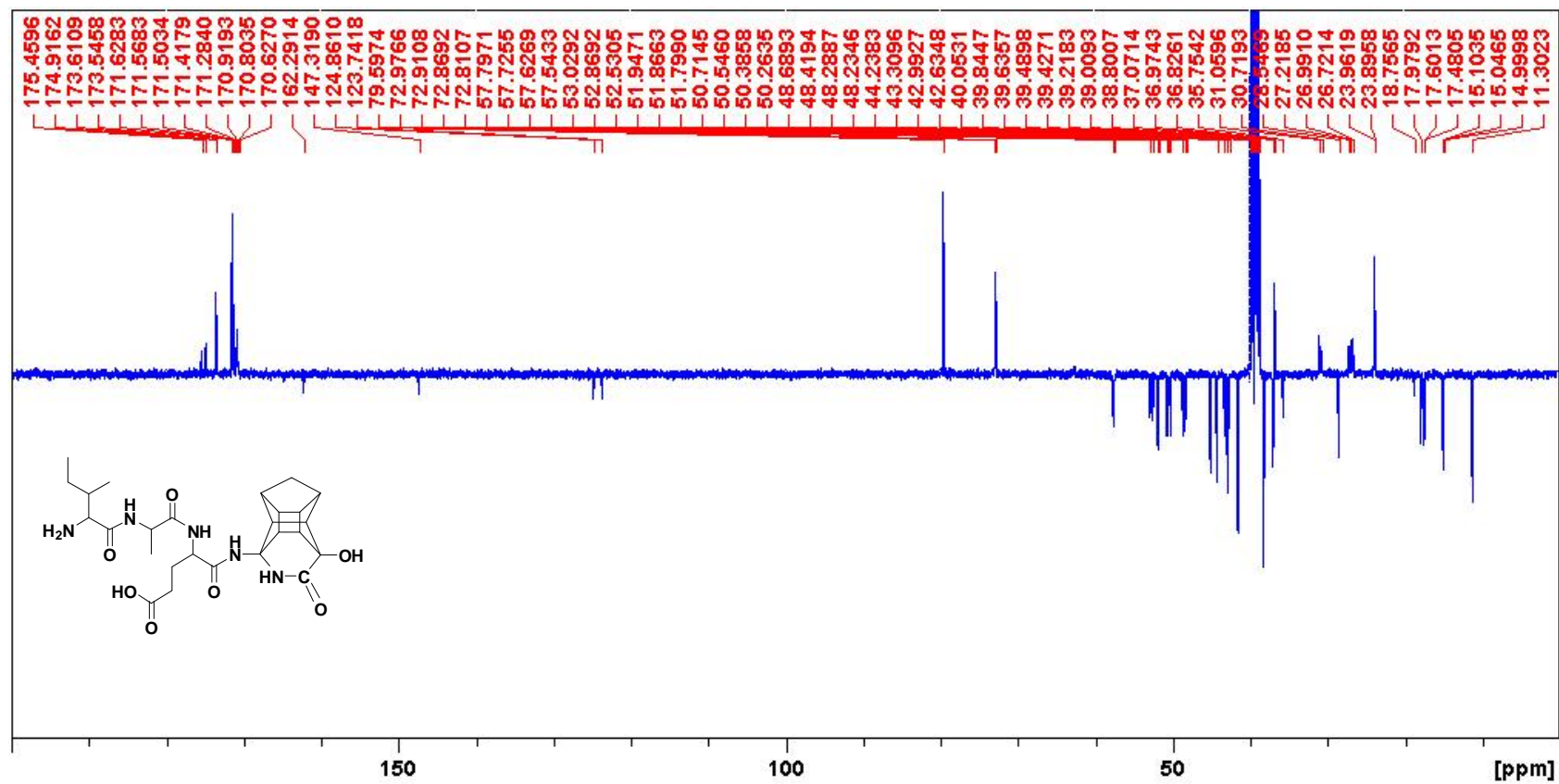
COSY spectrum of PCU-EAIS in $(\text{CD}_3)_2\text{SO}$



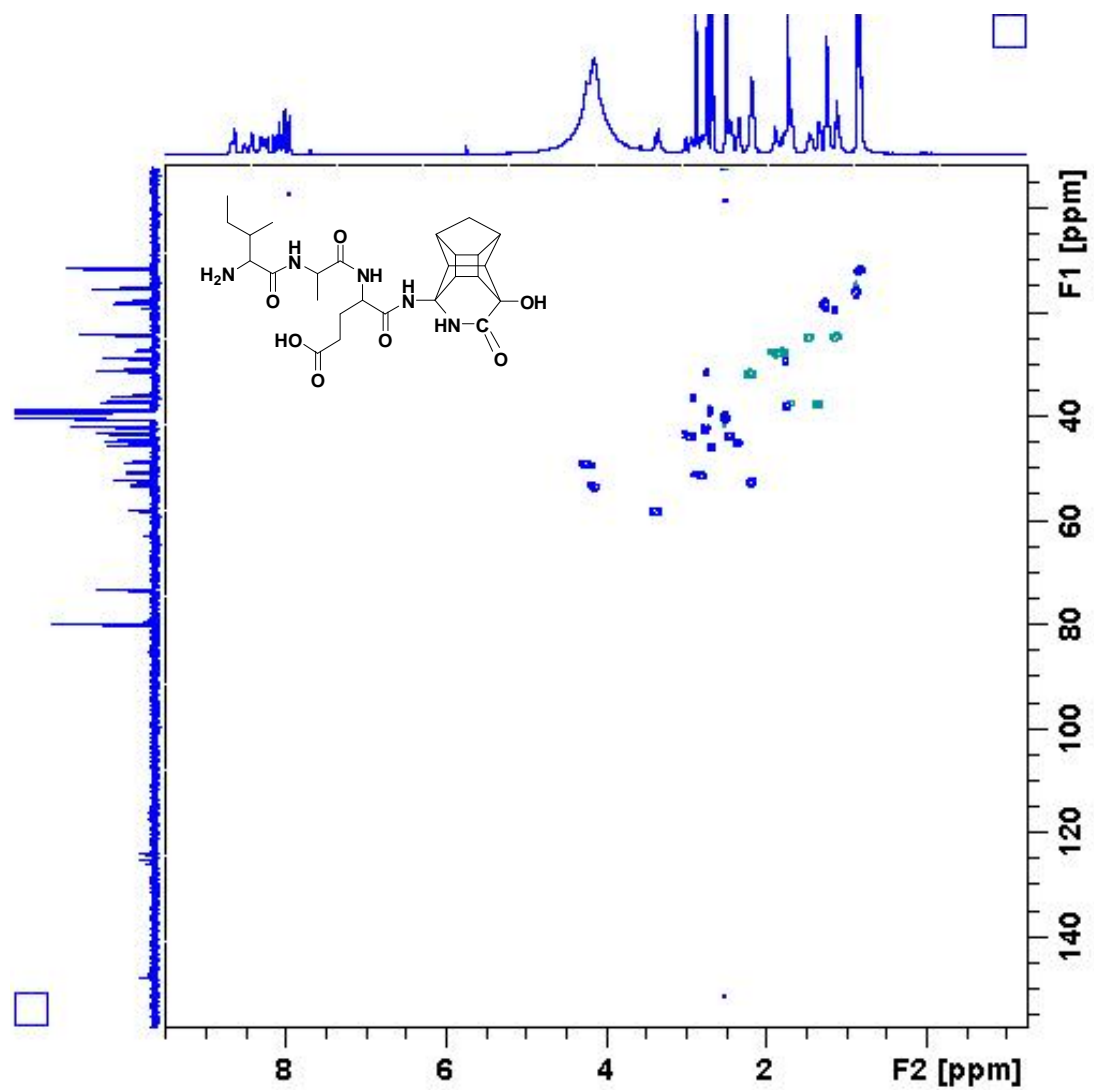
ROESY spectrum of PCU-EAIS in $(\text{CD}_3)_2\text{SO}$

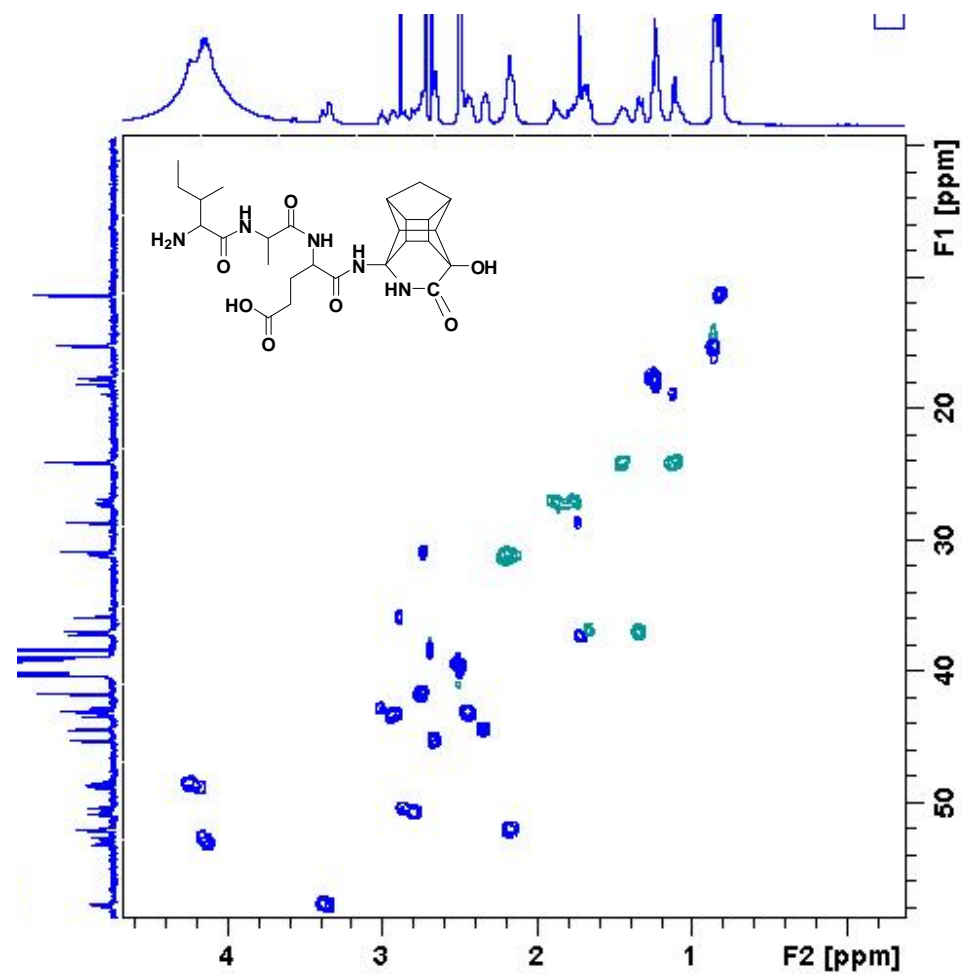


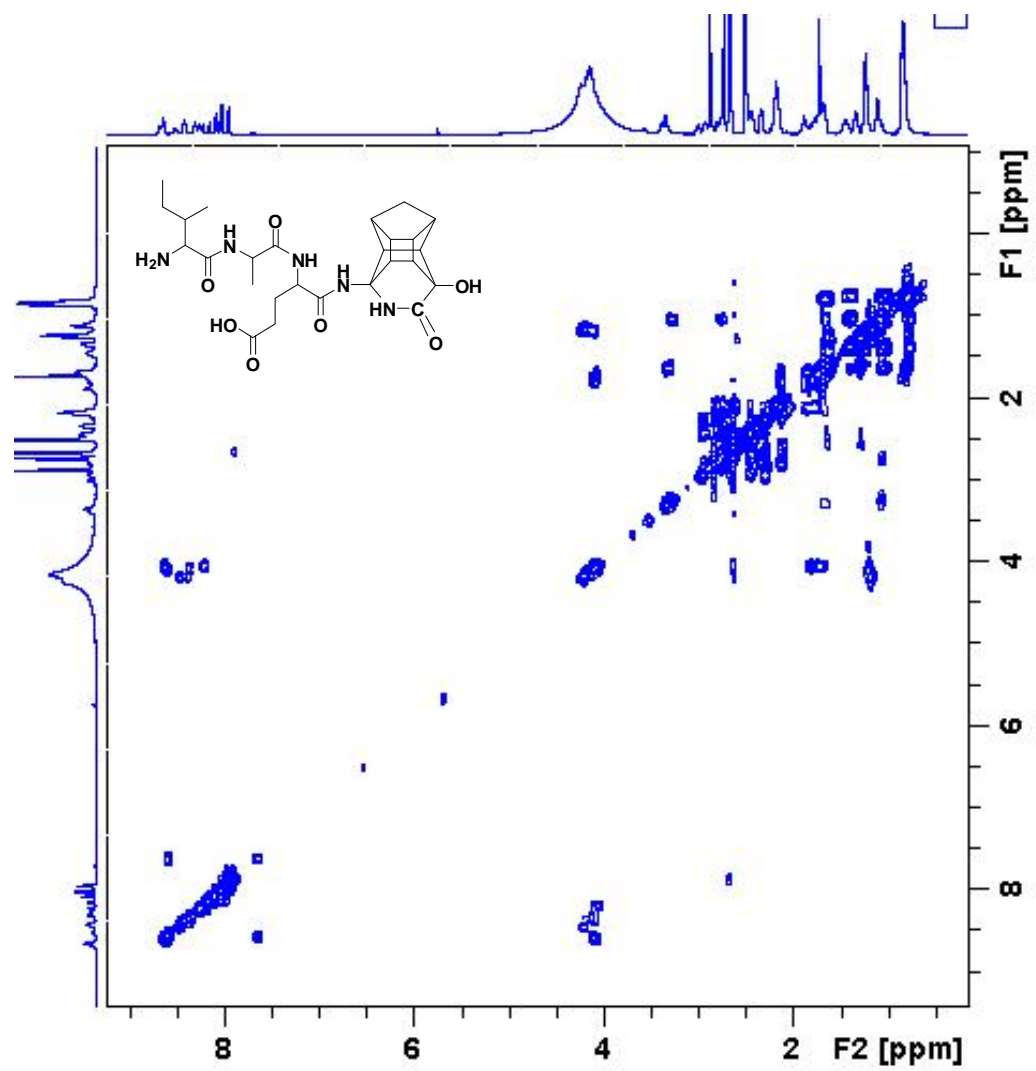
^1H NMR spectrum of PCU-EAI in $(\text{CD}_3)_2\text{SO}$

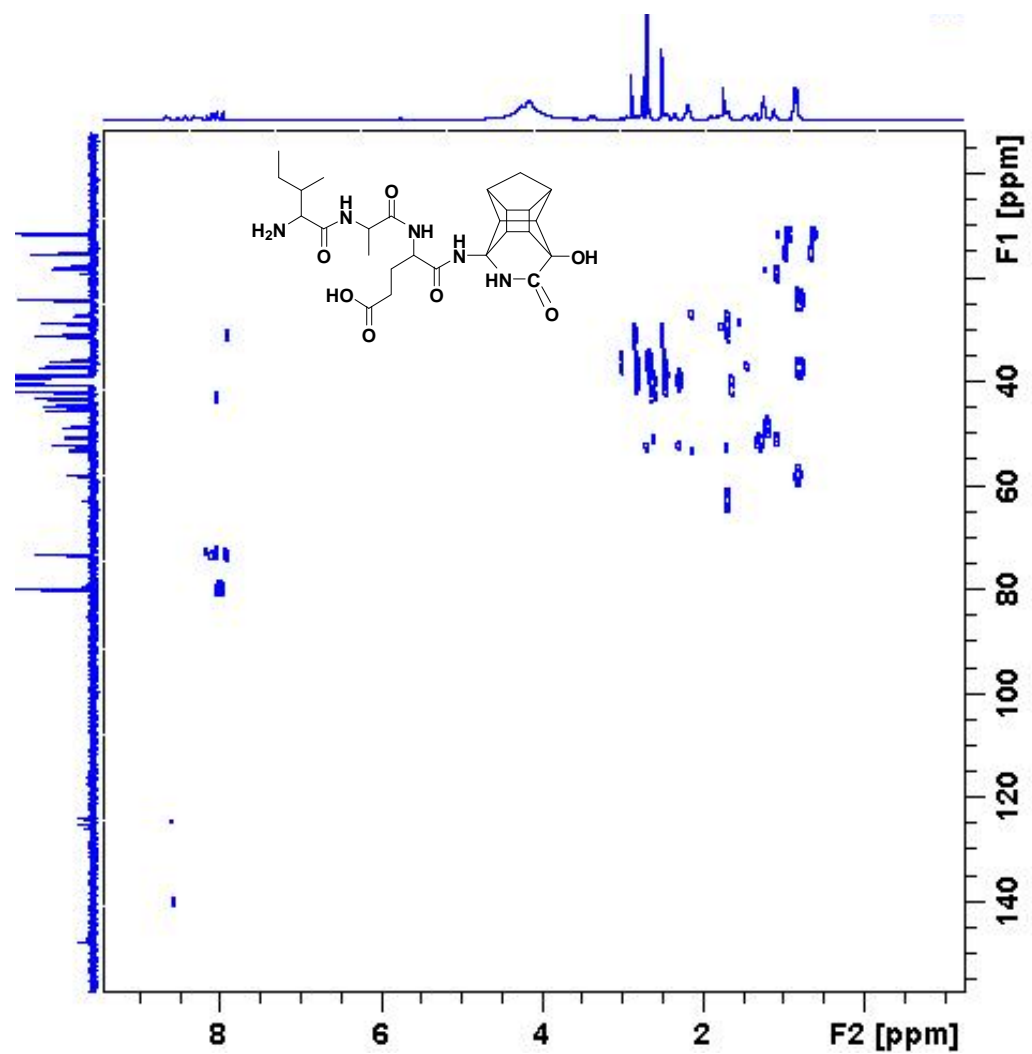


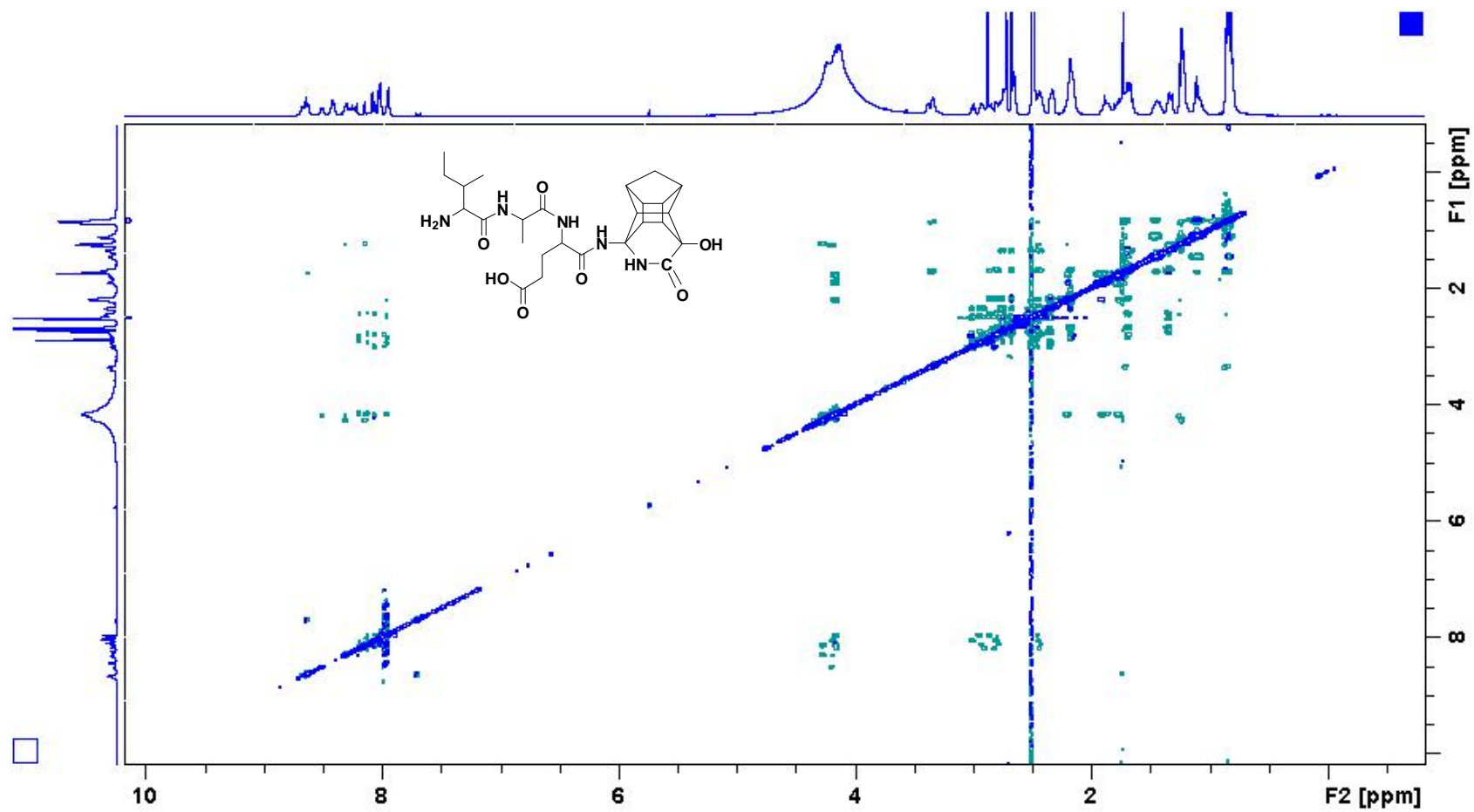
^{13}C NMR spectrum of PCU-EAI in $(\text{CD}_3)_2\text{SO}$

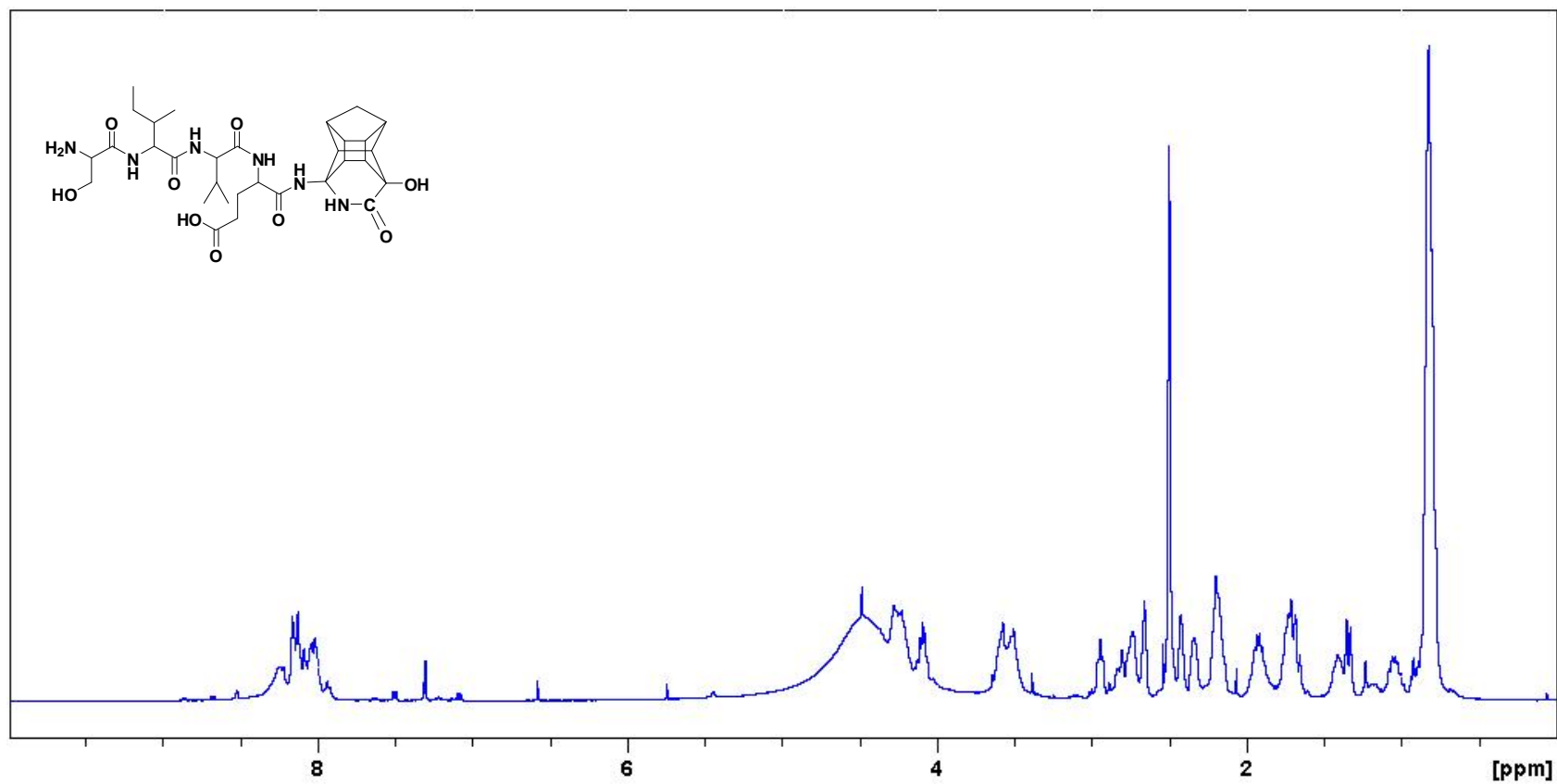


Expanded HSQC spectrum of PCU-EAI in $(\text{CD}_3)_2\text{SO}$

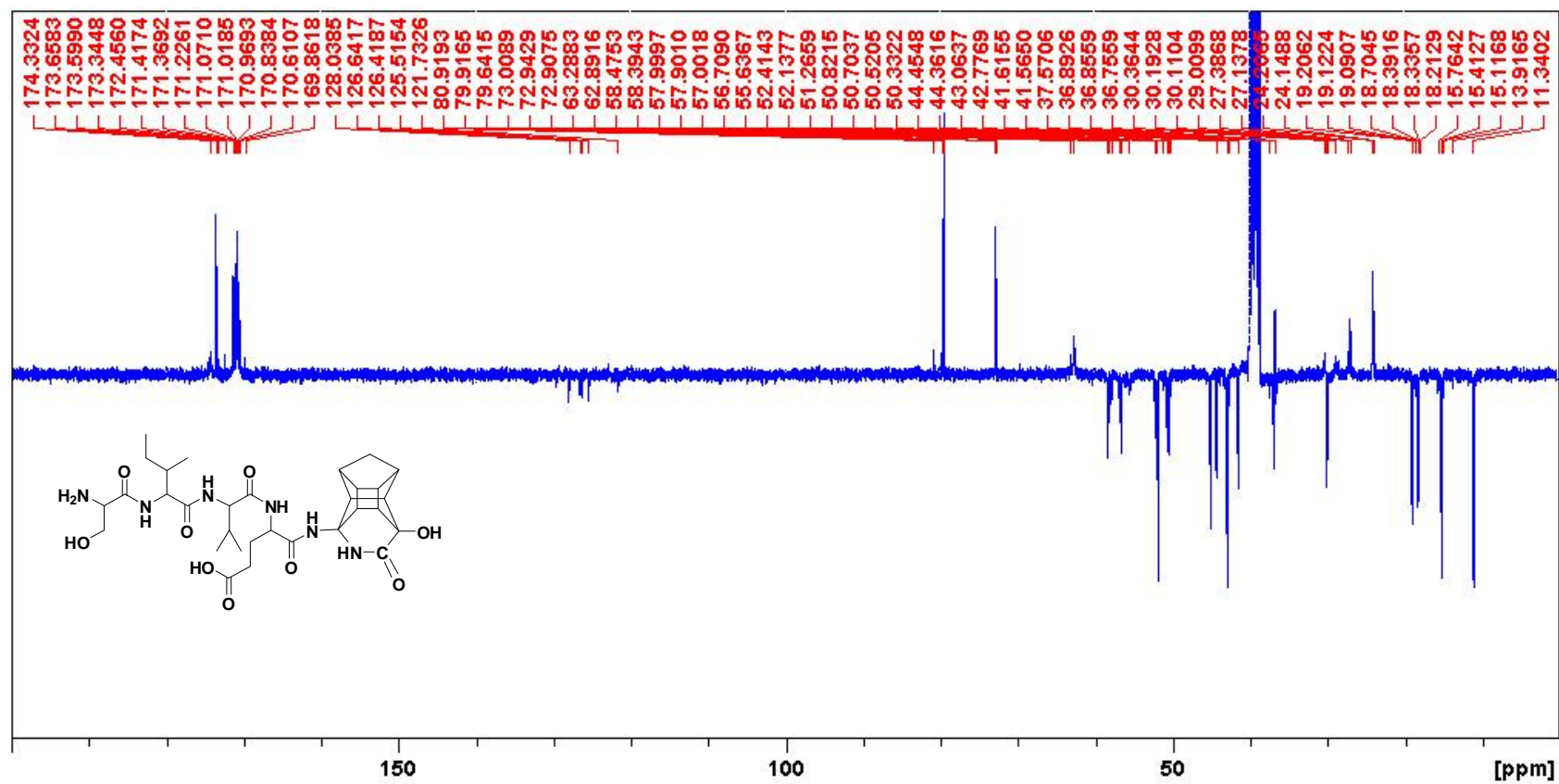
COSY spectrum of PCU-EAI in $(\text{CD}_3)_2\text{SO}$

HMBC spectrum of PCU-EAI in $(\text{CD}_3)_2\text{SO}$

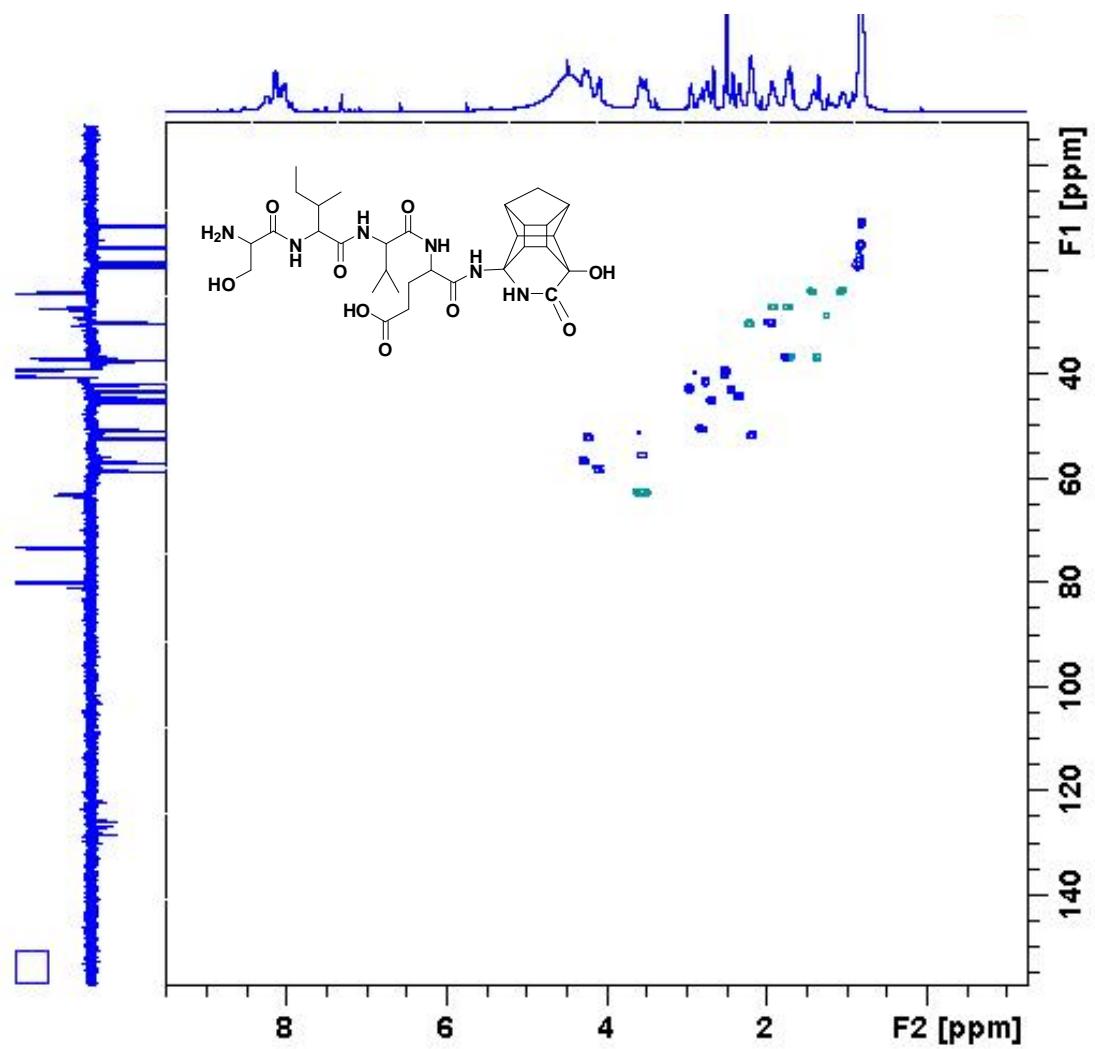
ROESY spectrum of PCU-EAI in $(\text{CD}_3)_2\text{SO}$

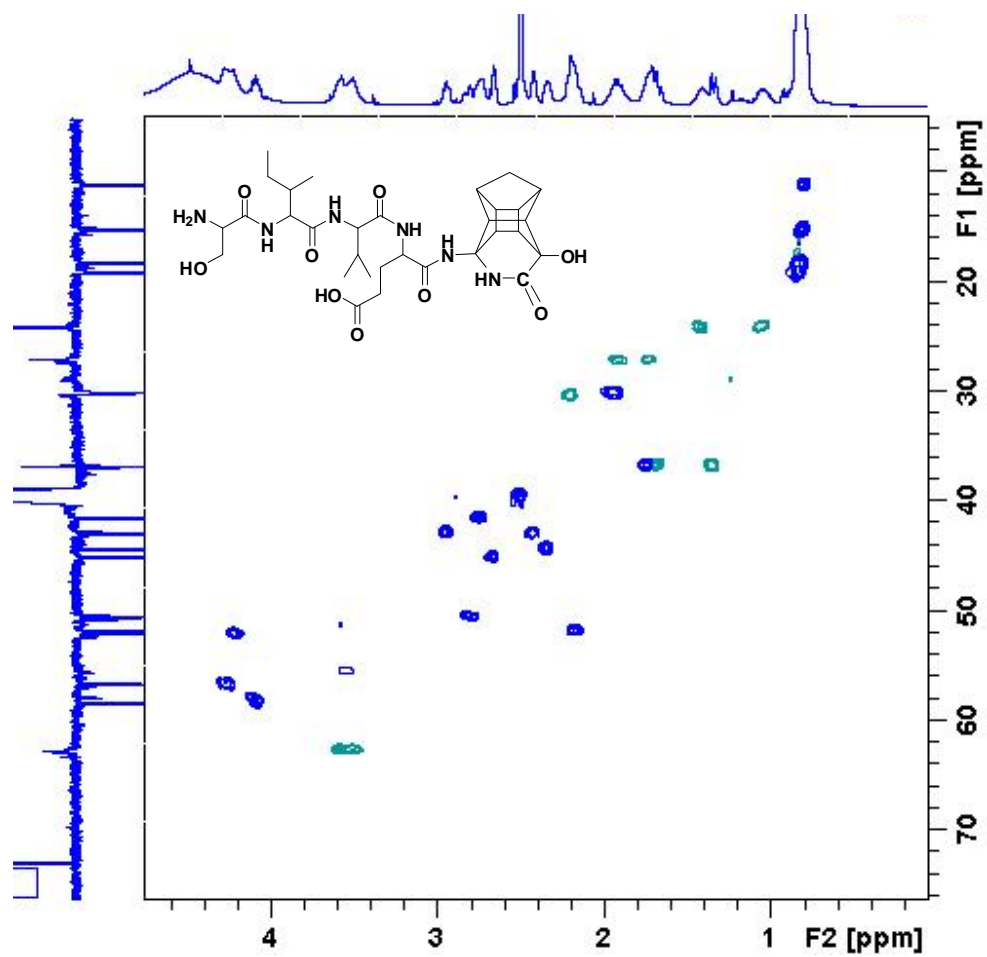


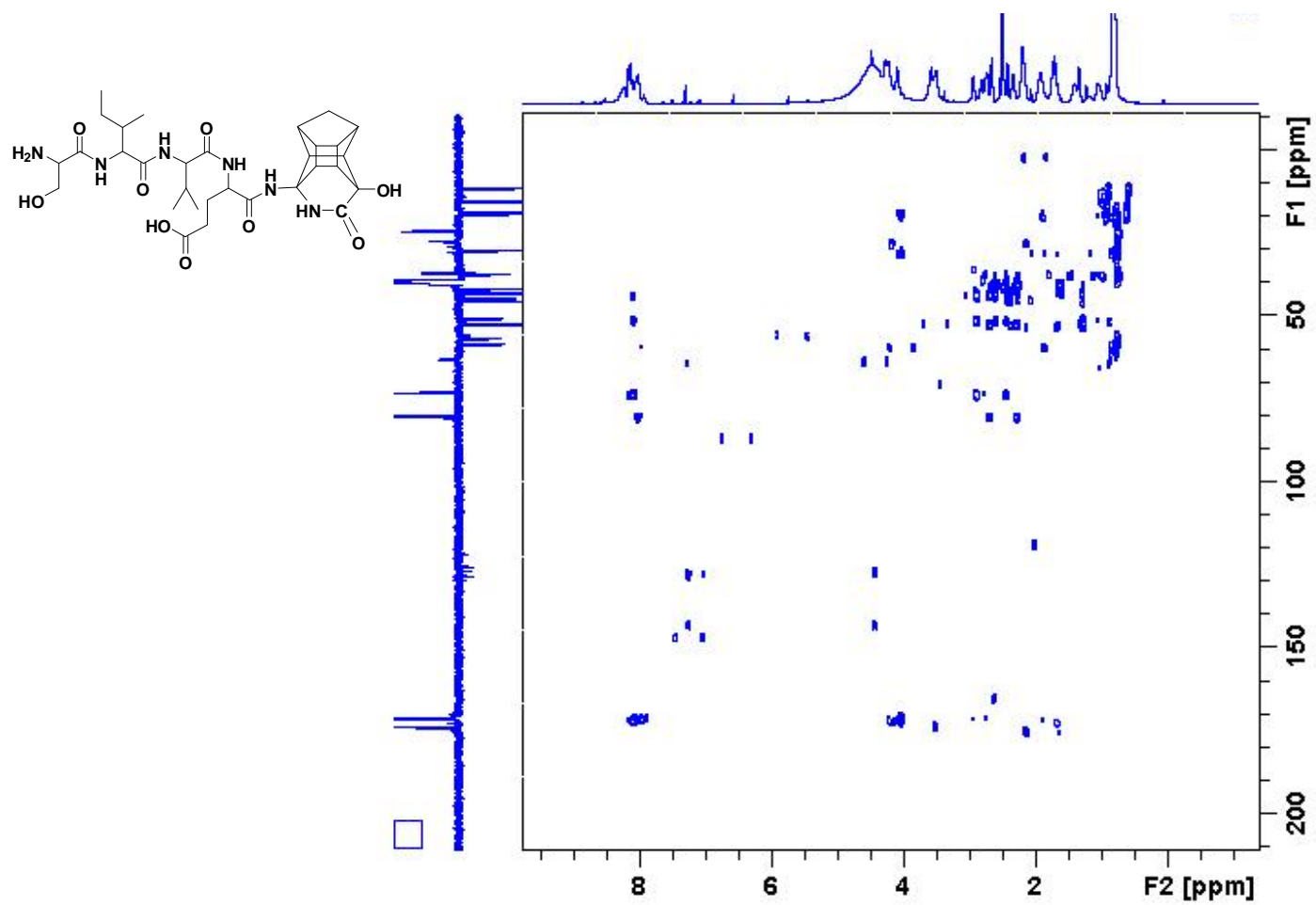
^1H NMR spectrum of PCU-EVIS in $(\text{CD}_3)_2\text{SO}$

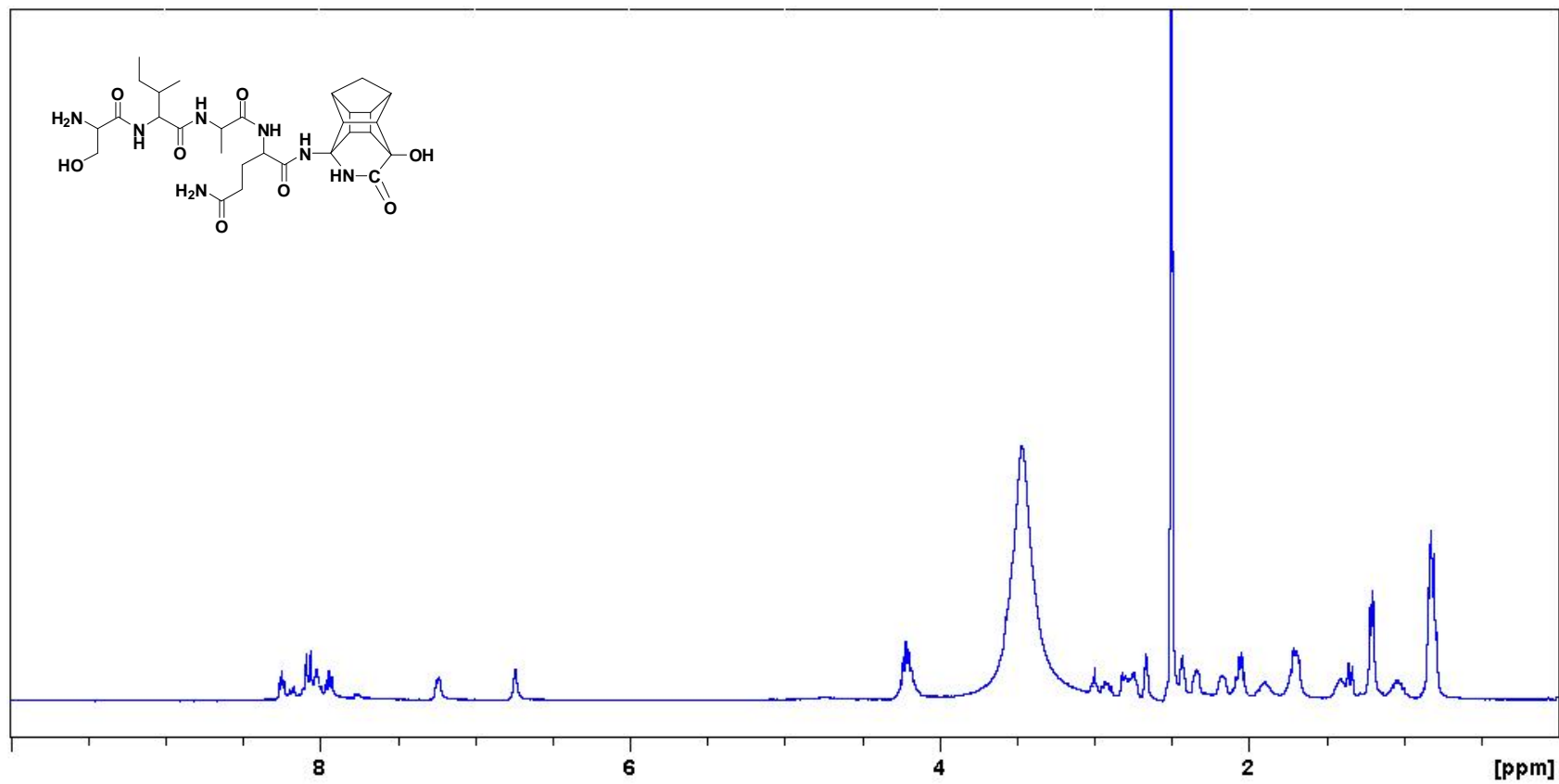


^{13}C NMR spectrum of PCU-EVIS in $(\text{CD}_3)_2\text{SO}$

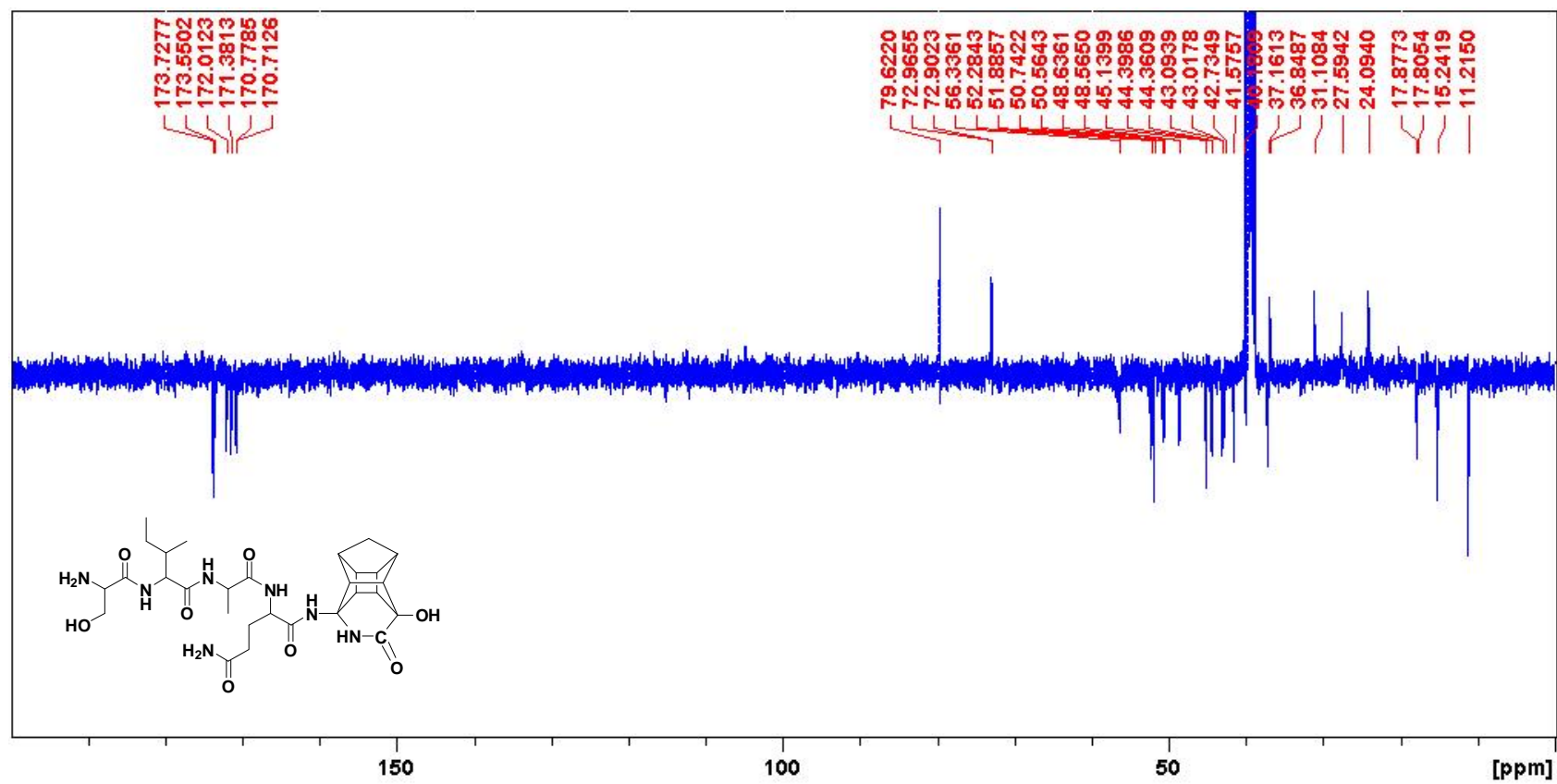
HSQC spectrum of PCU-EVIS in $(\text{CD}_3)_2\text{SO}$

Expanded HSQC spectrum of PCU-EVIS in $(\text{CD}_3)_2\text{SO}$

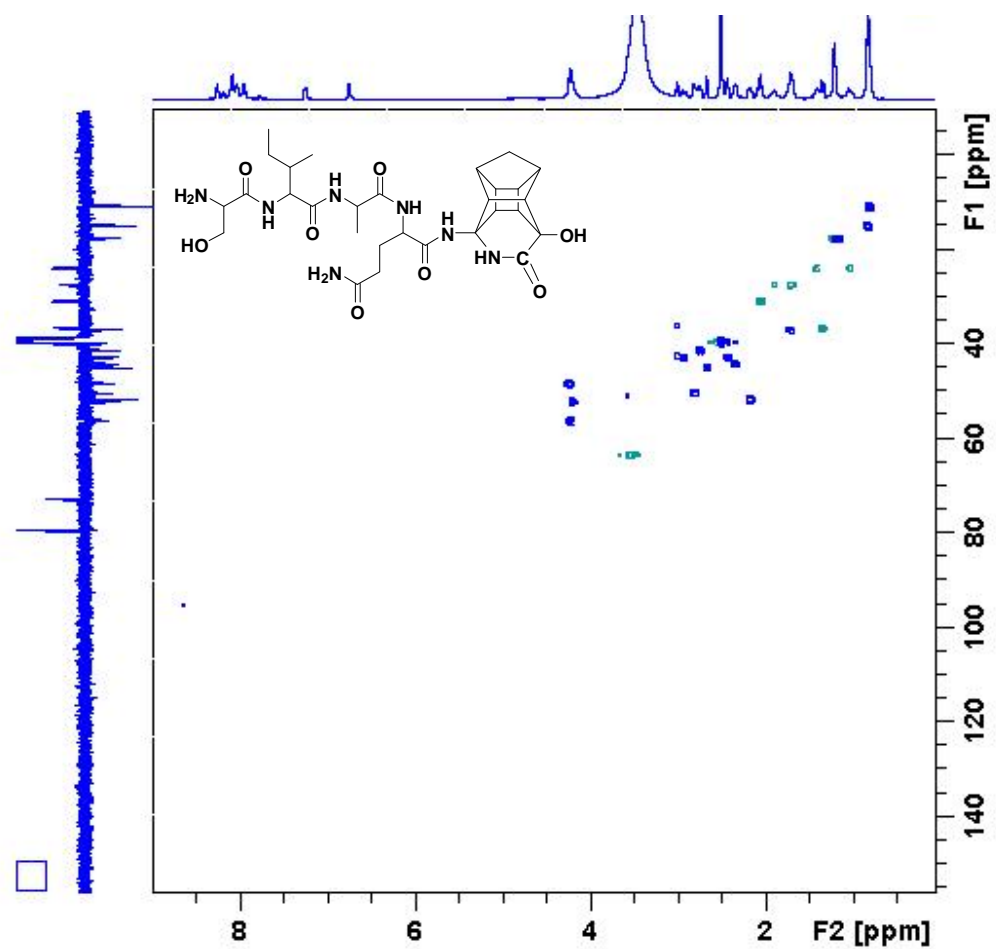
HMBC spectrum of PCU-EVIS in $(\text{CD}_3)_2\text{SO}$

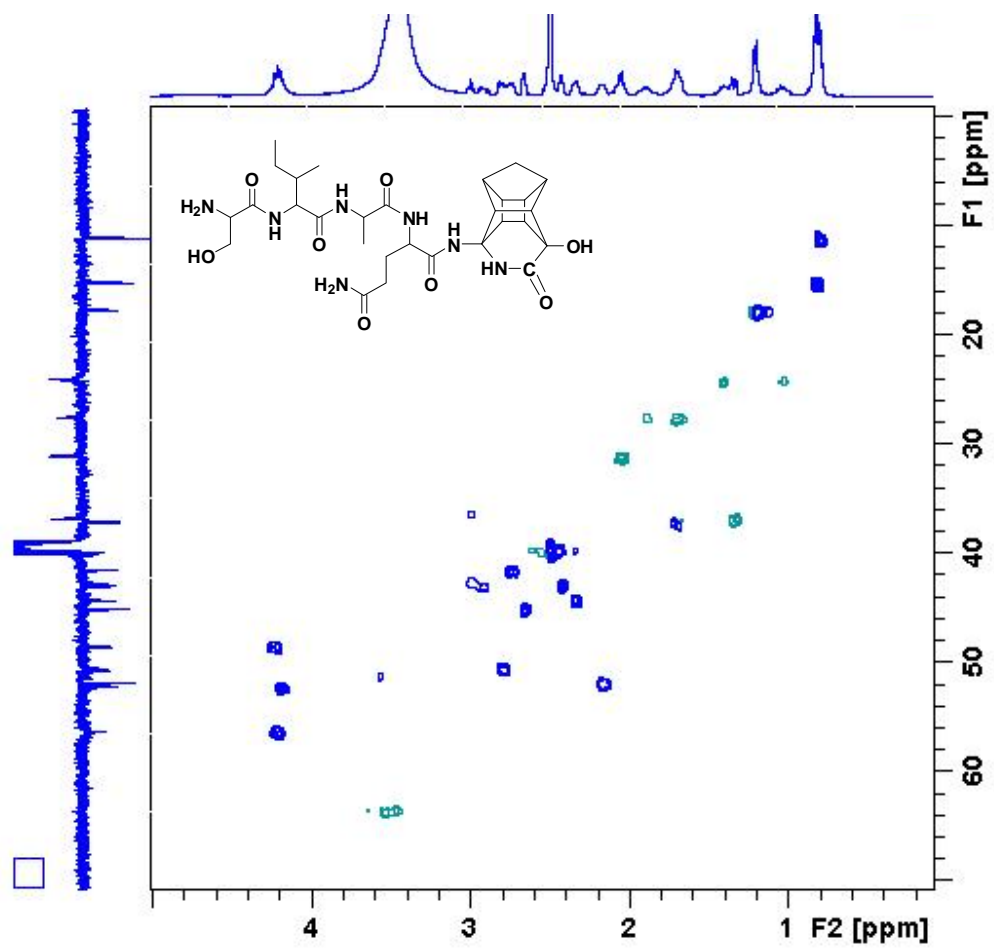


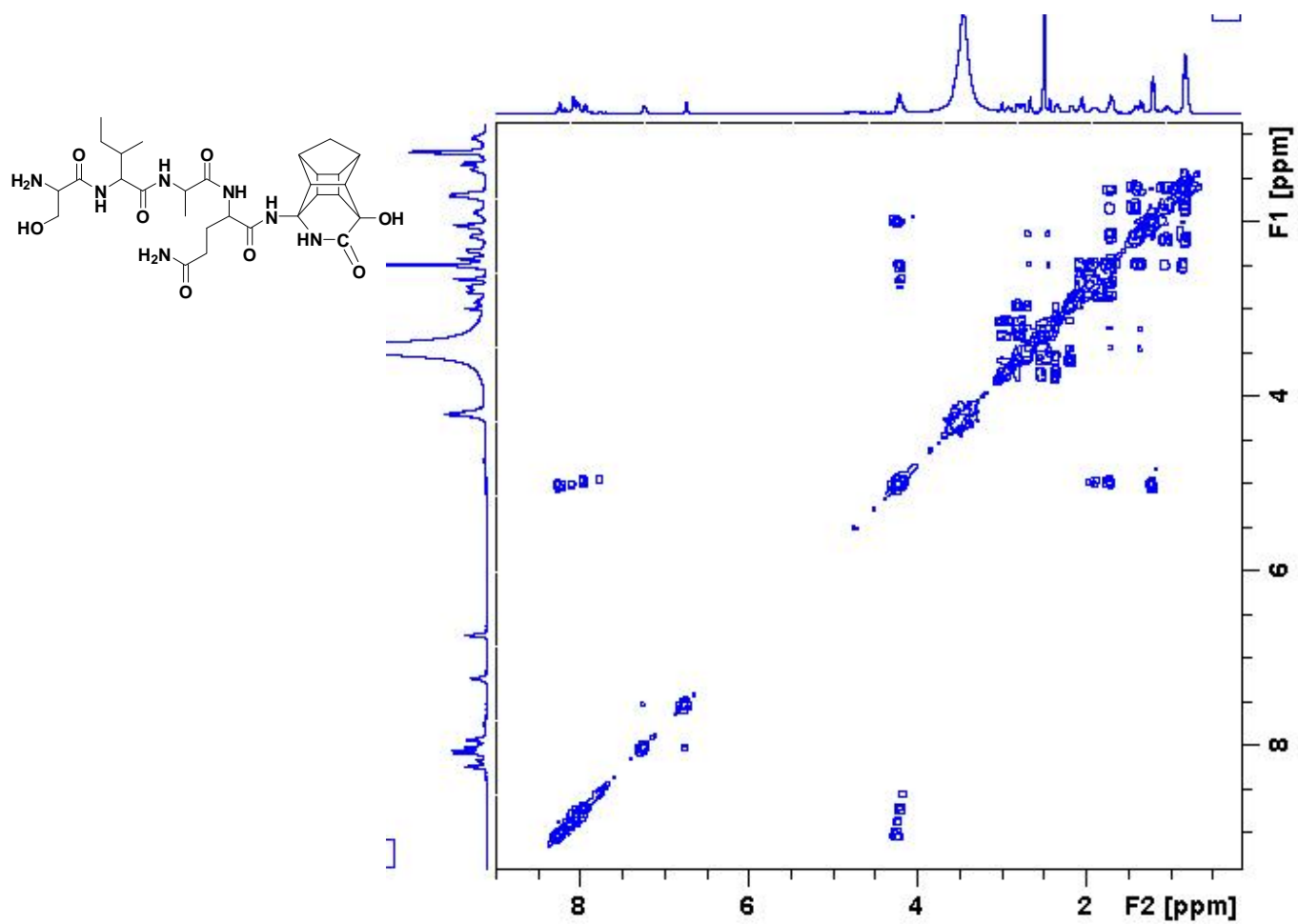
^1H NMR spectrum of PCU-QAIS in $(\text{CD}_3)_2\text{SO}$

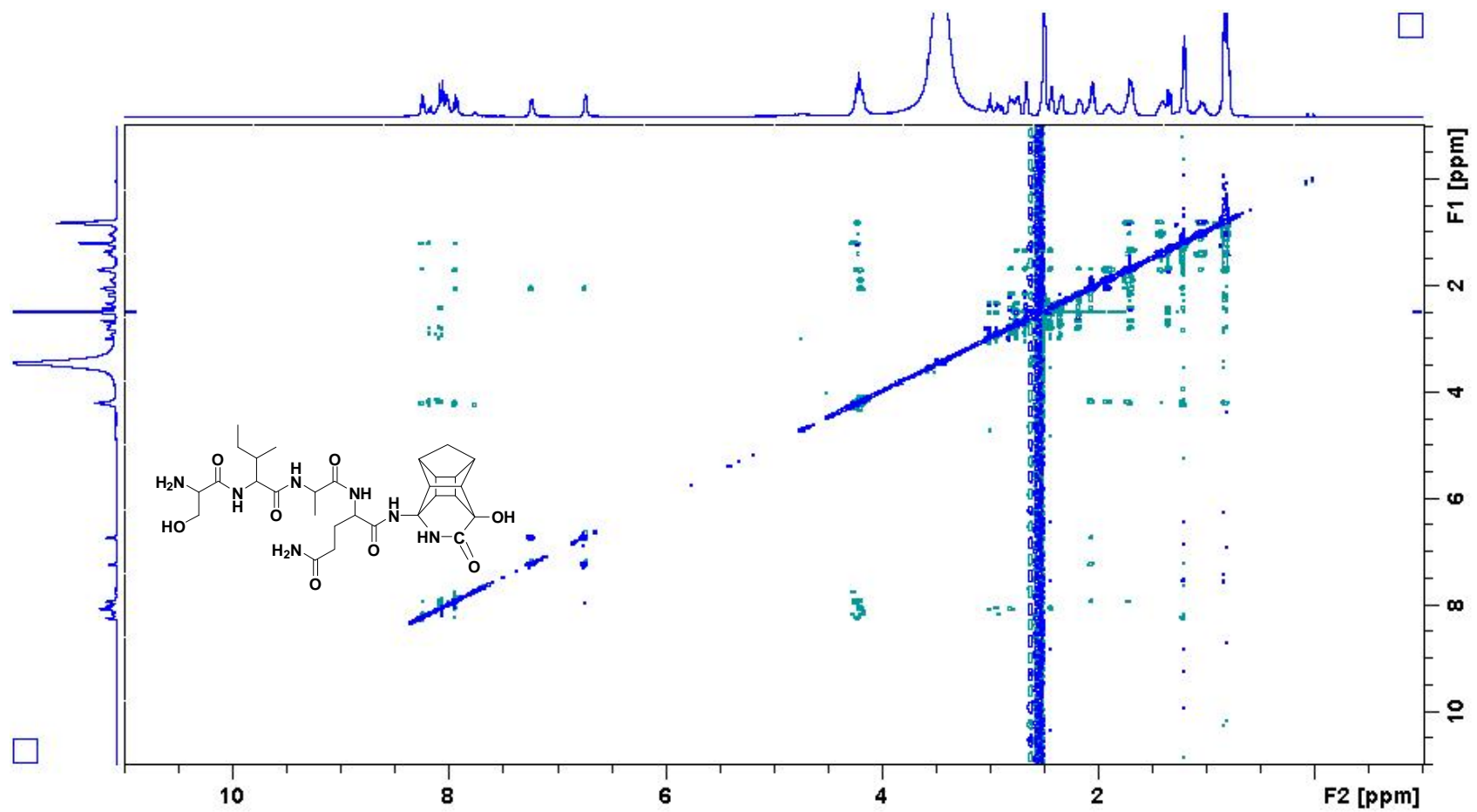


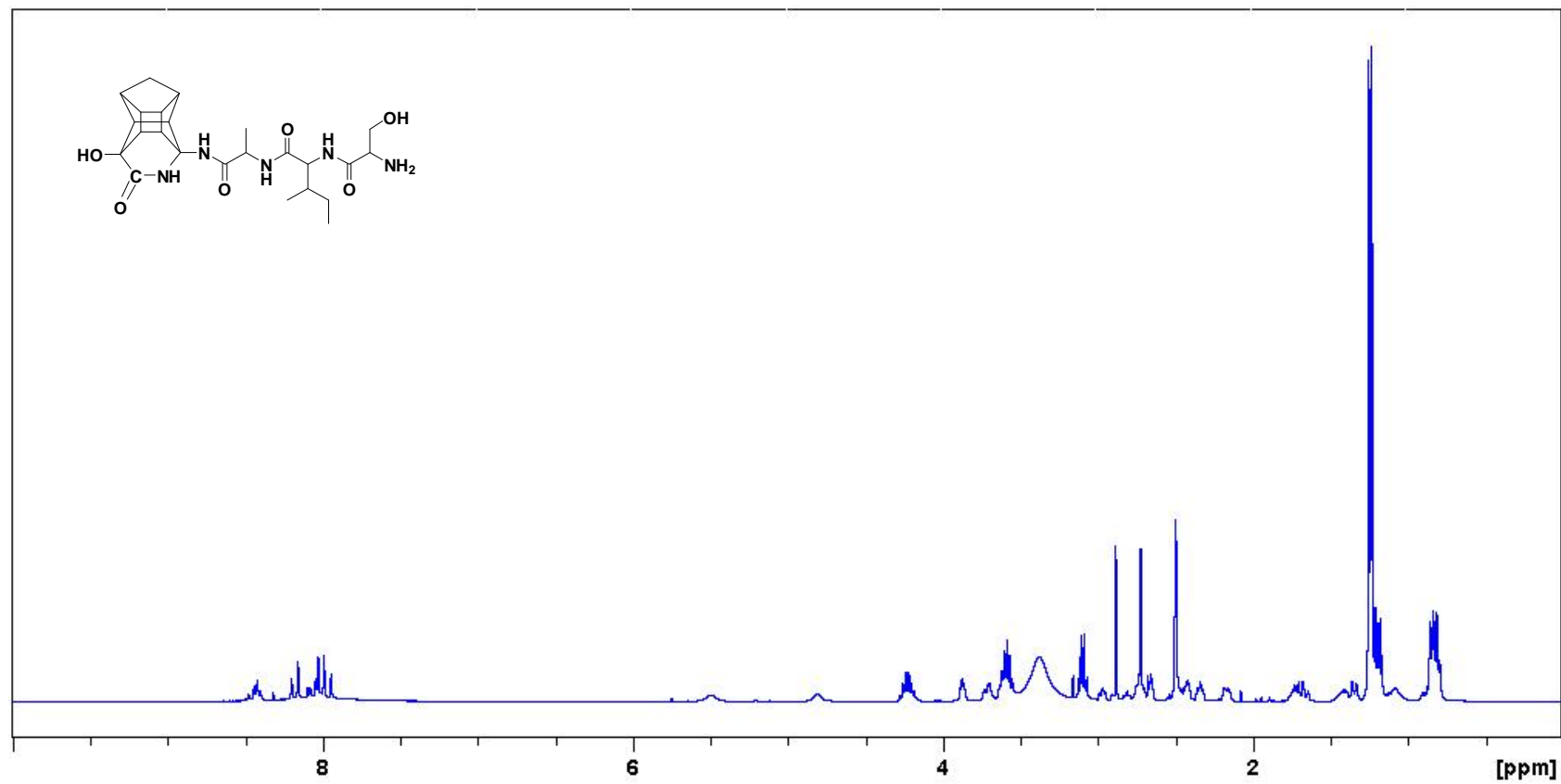
^{13}C NMR spectrum of PCU-QAIS in $(\text{CD}_3)_2\text{SO}$

HSQC spectrum of PCU-QAIS in $(\text{CD}_3)_2\text{SO}$

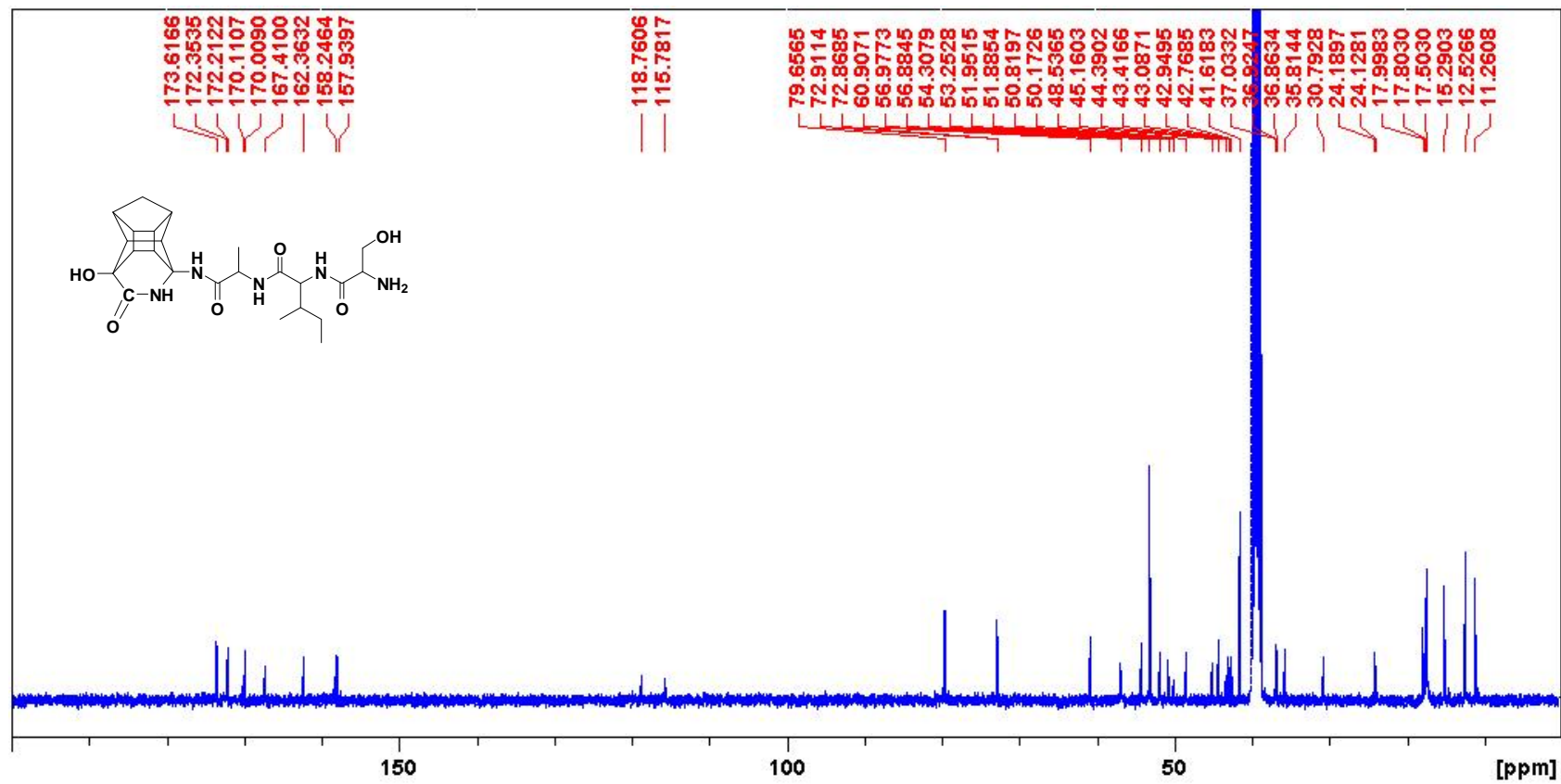
Expanded HSQC spectrum of PCU-QAIS in $(\text{CD}_3)_2\text{SO}$

COSY spectrum of PCU-QAIS in $(\text{CD}_3)_2\text{SO}$

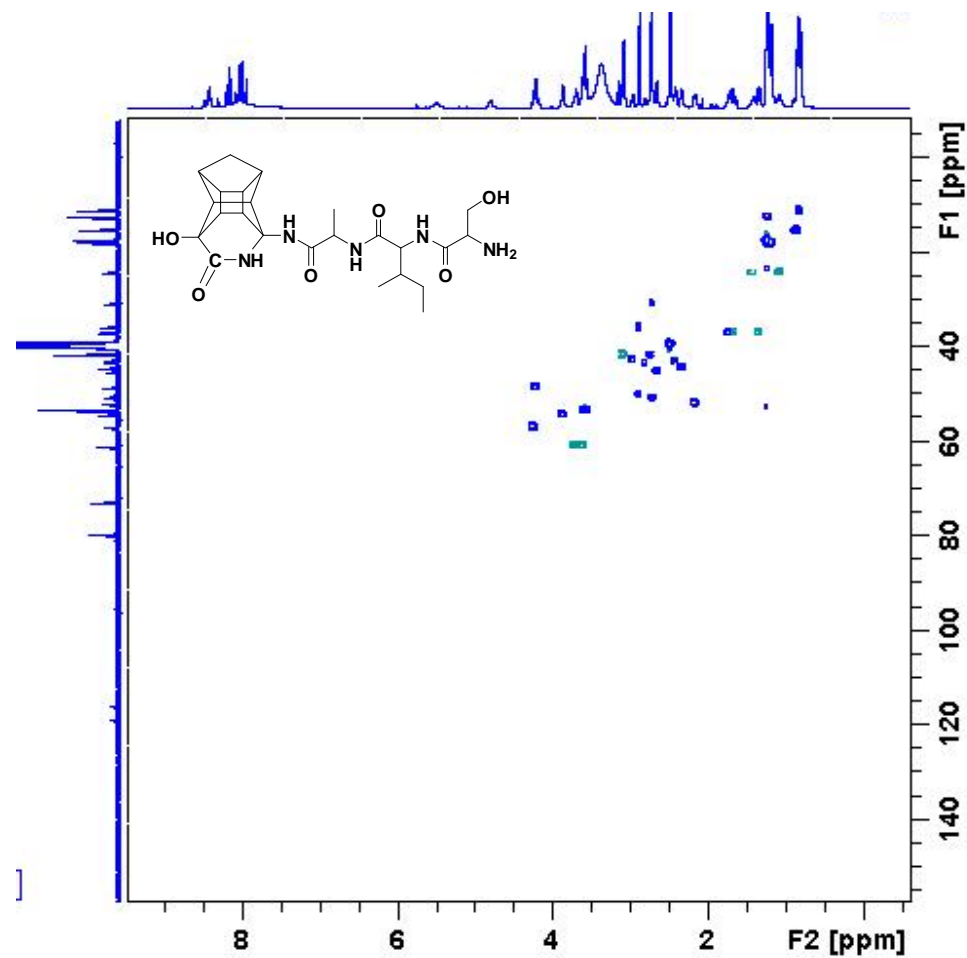
ROESY spectrum of PCU-QAIS in $(\text{CD}_3)_2\text{SO}$

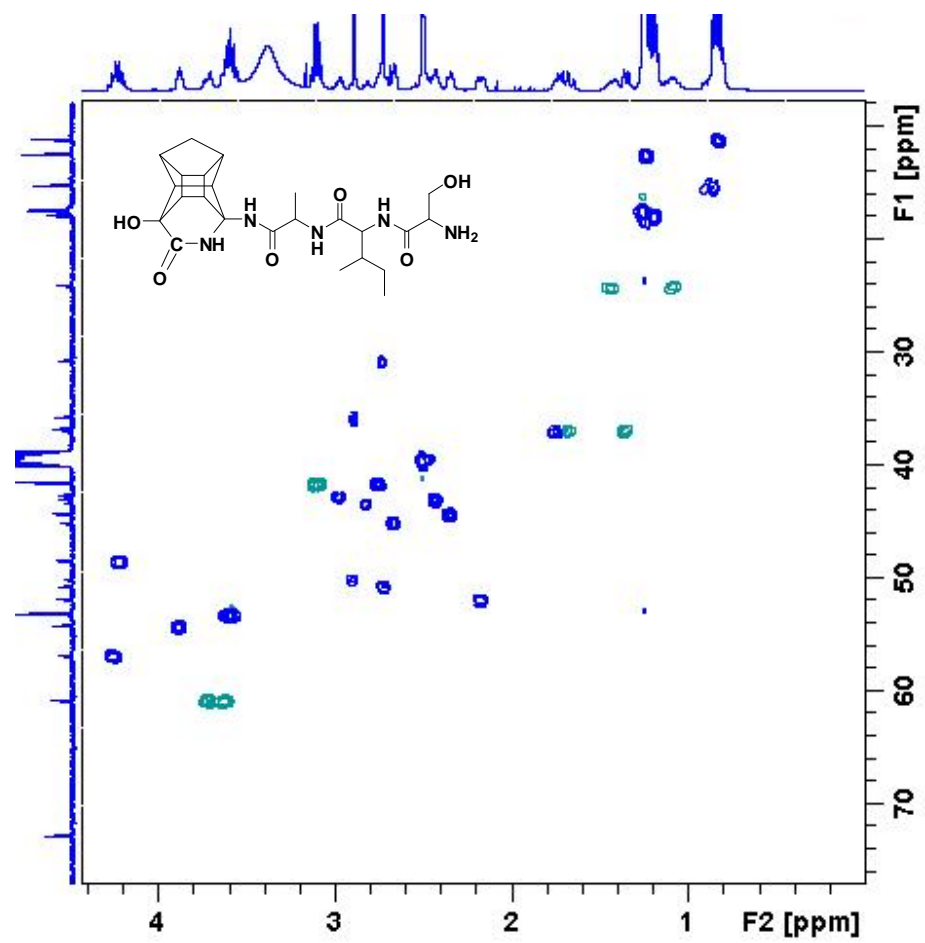


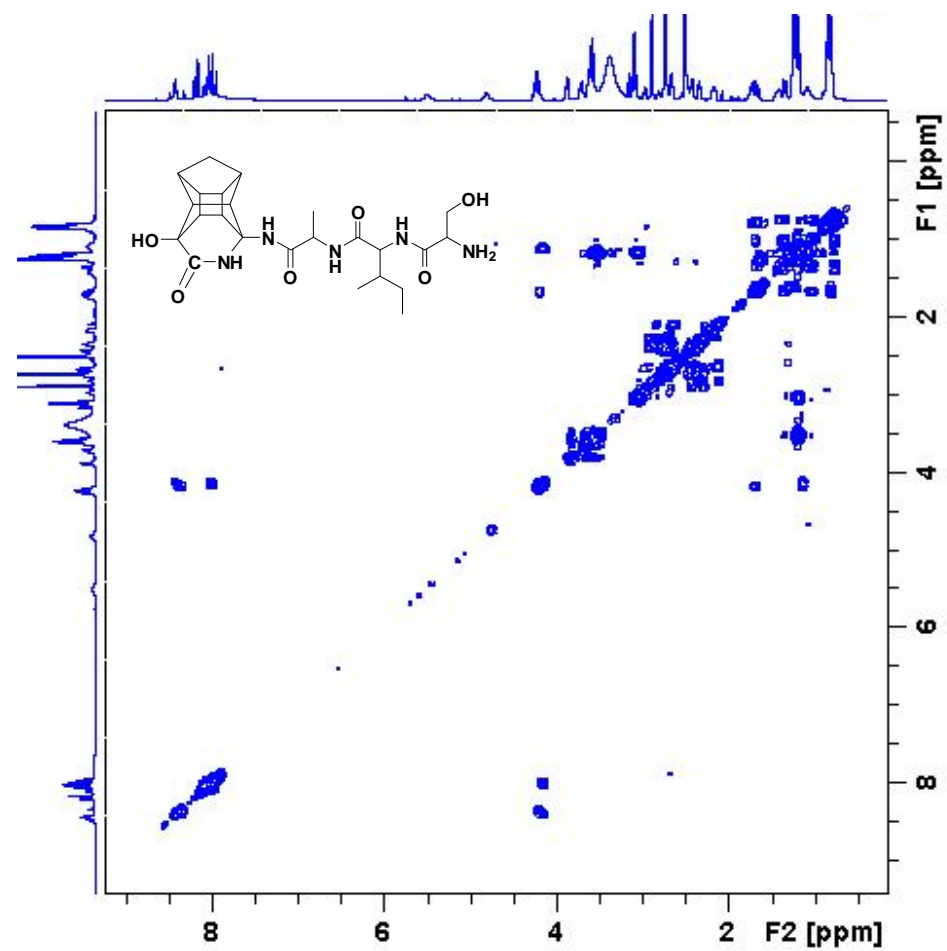
^1H NMR spectrum of PCU-AISa in $(\text{CD}_3)_2\text{SO}$

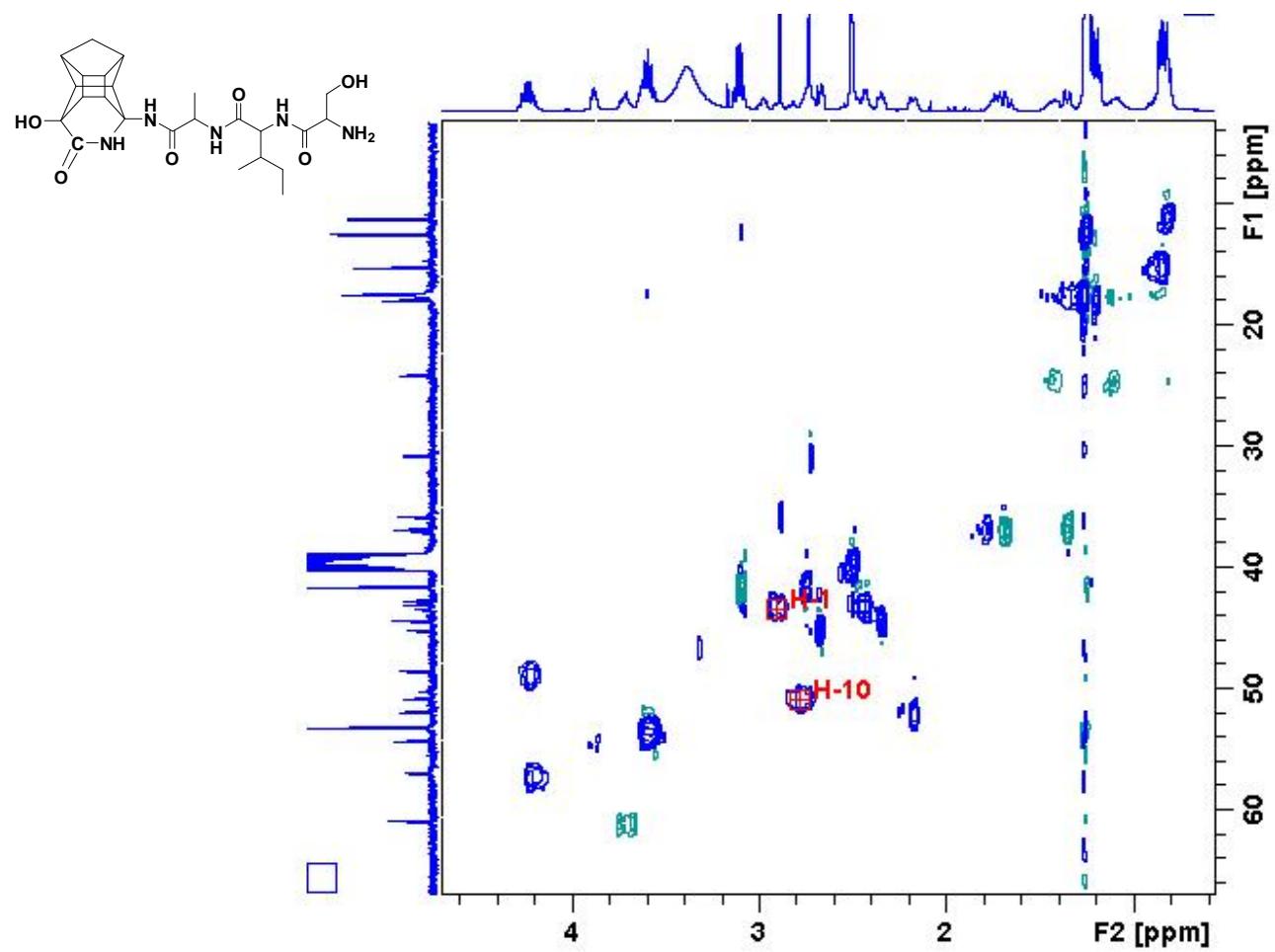


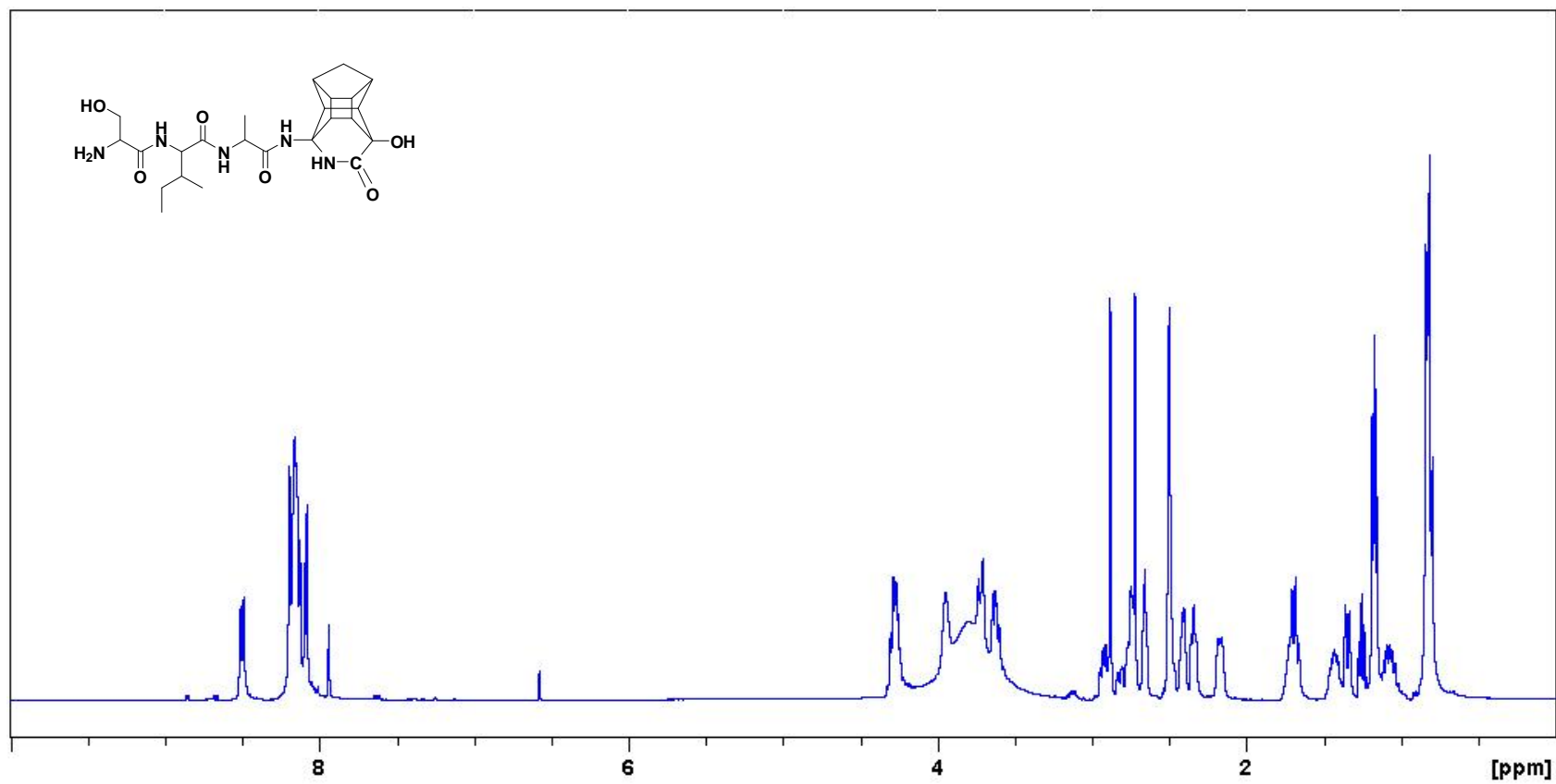
^{13}C NMR spectrum of PCU-AISa in $(\text{CD}_3)_2\text{SO}$

HSQC spectrum of PCU-AISa in $(\text{CD}_3)_2\text{SO}$

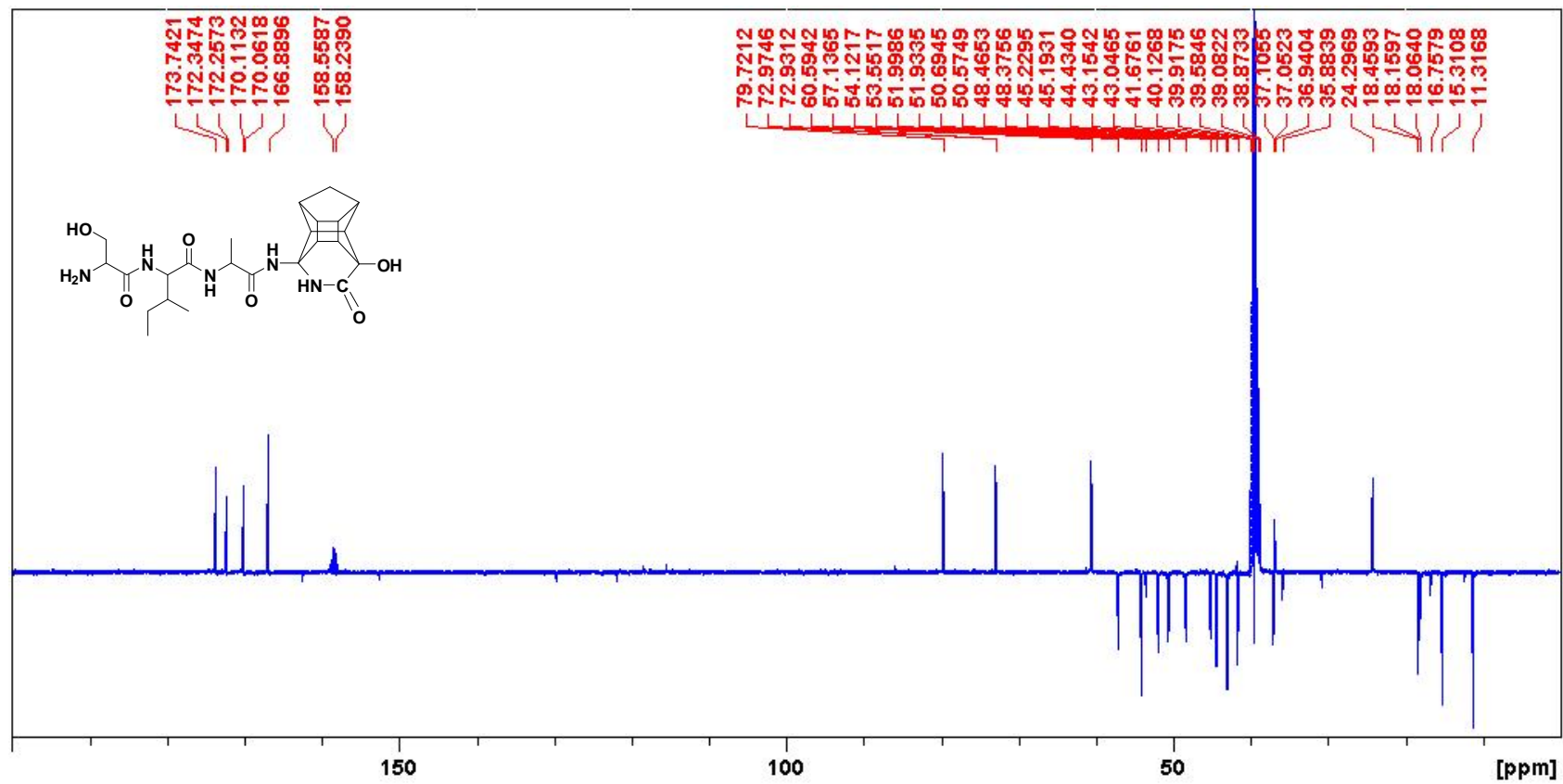
Expanded HSQC spectrum of PCU-AISa in $(\text{CD}_3)_2\text{SO}$

COSY spectrum of PCU-AISa in $(\text{CD}_3)_2\text{SO}$

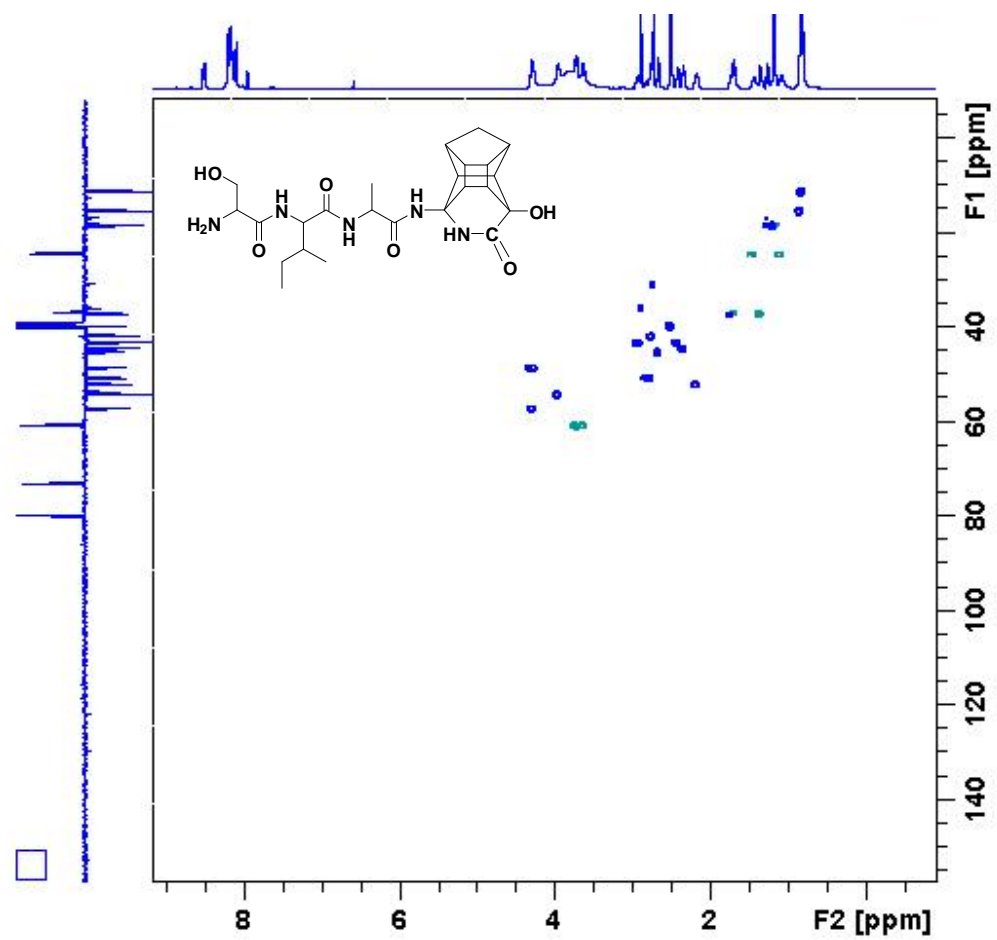
Expanded HSQC spectrum of PCU-AISa in $(\text{CD}_3)_2\text{SO}$ at 333K

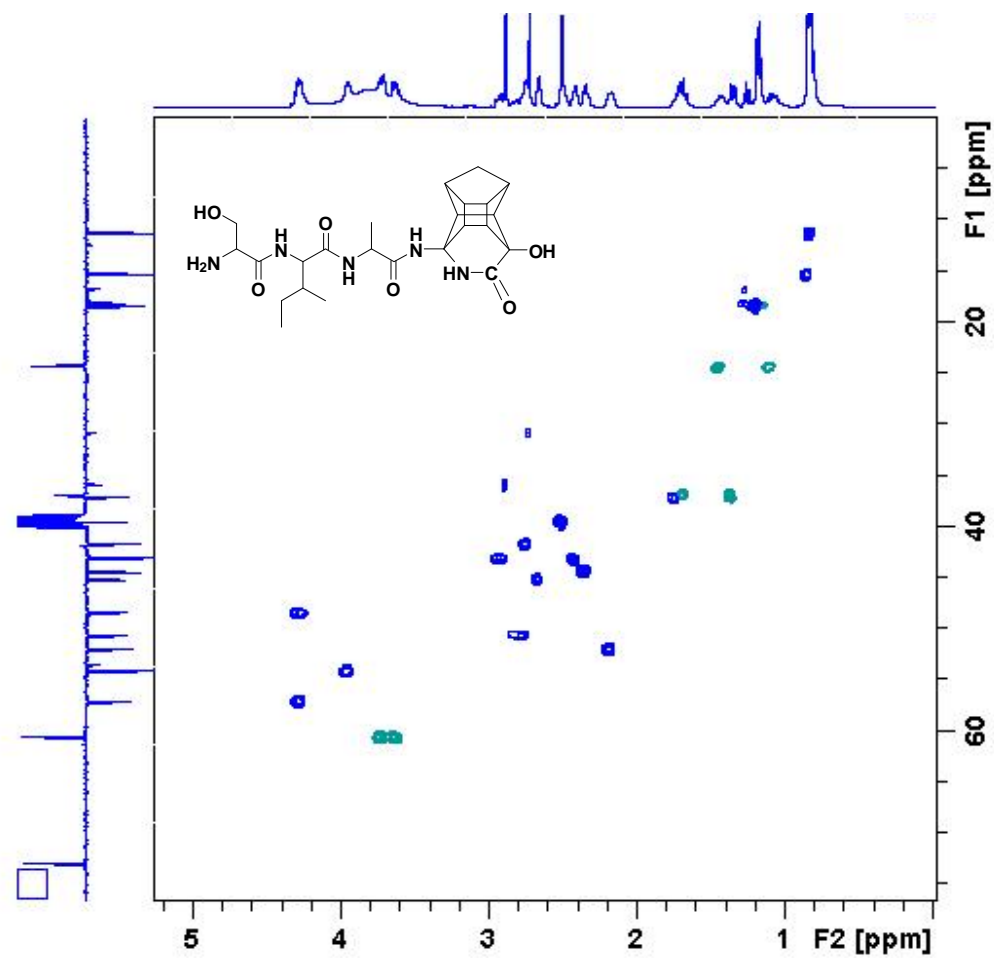


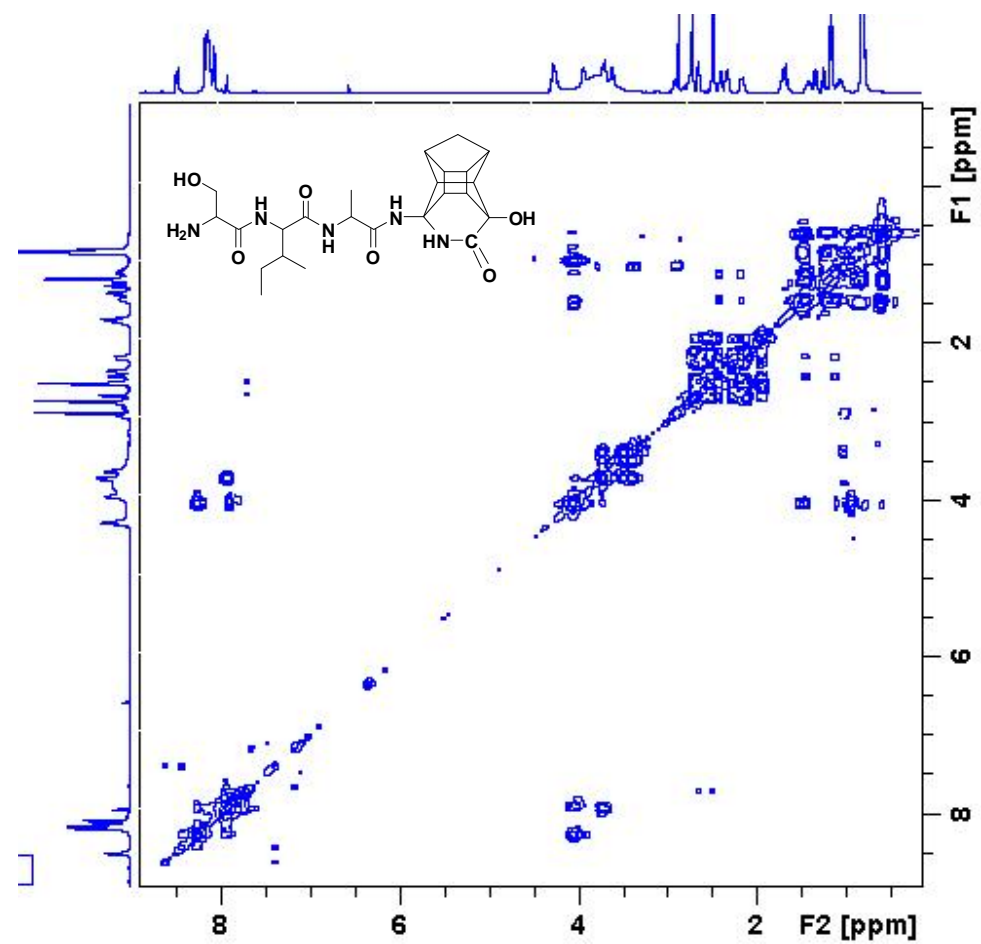
^1H NMR spectrum of PCU-AISb in $(\text{CD}_3)_2\text{SO}$

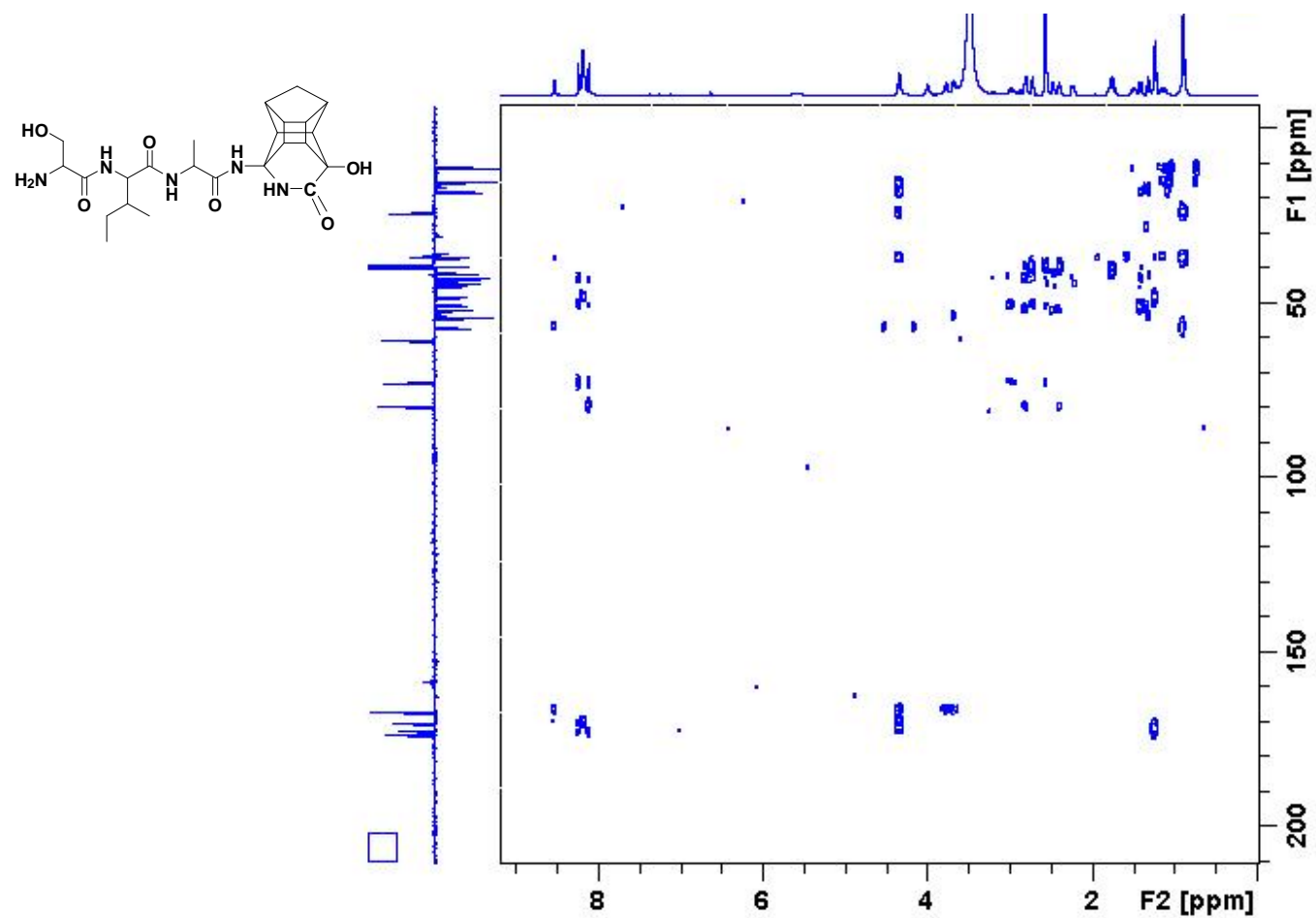


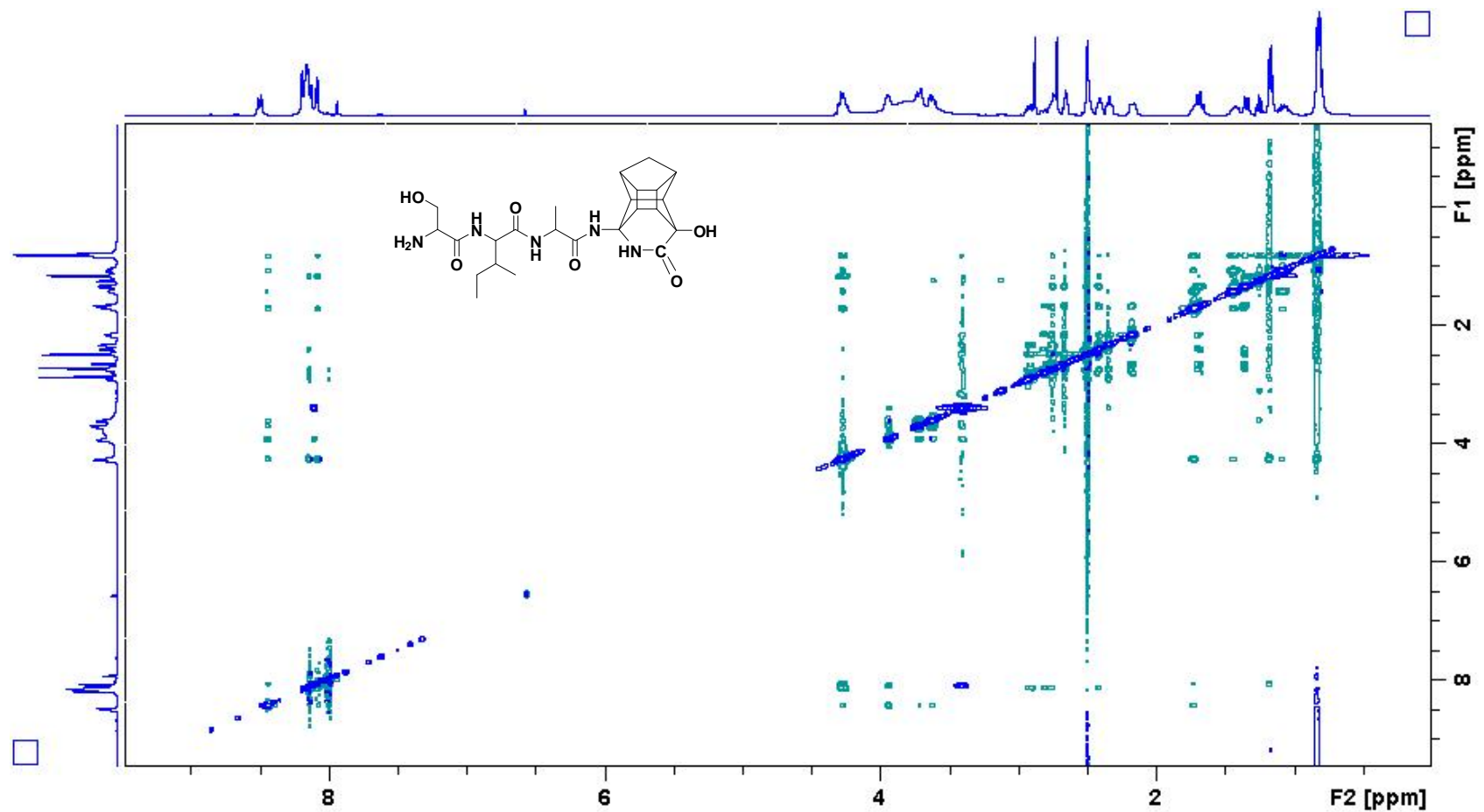
^{13}C NMR spectrum of PCU-AISb in $(\text{CD}_3)_2\text{SO}$

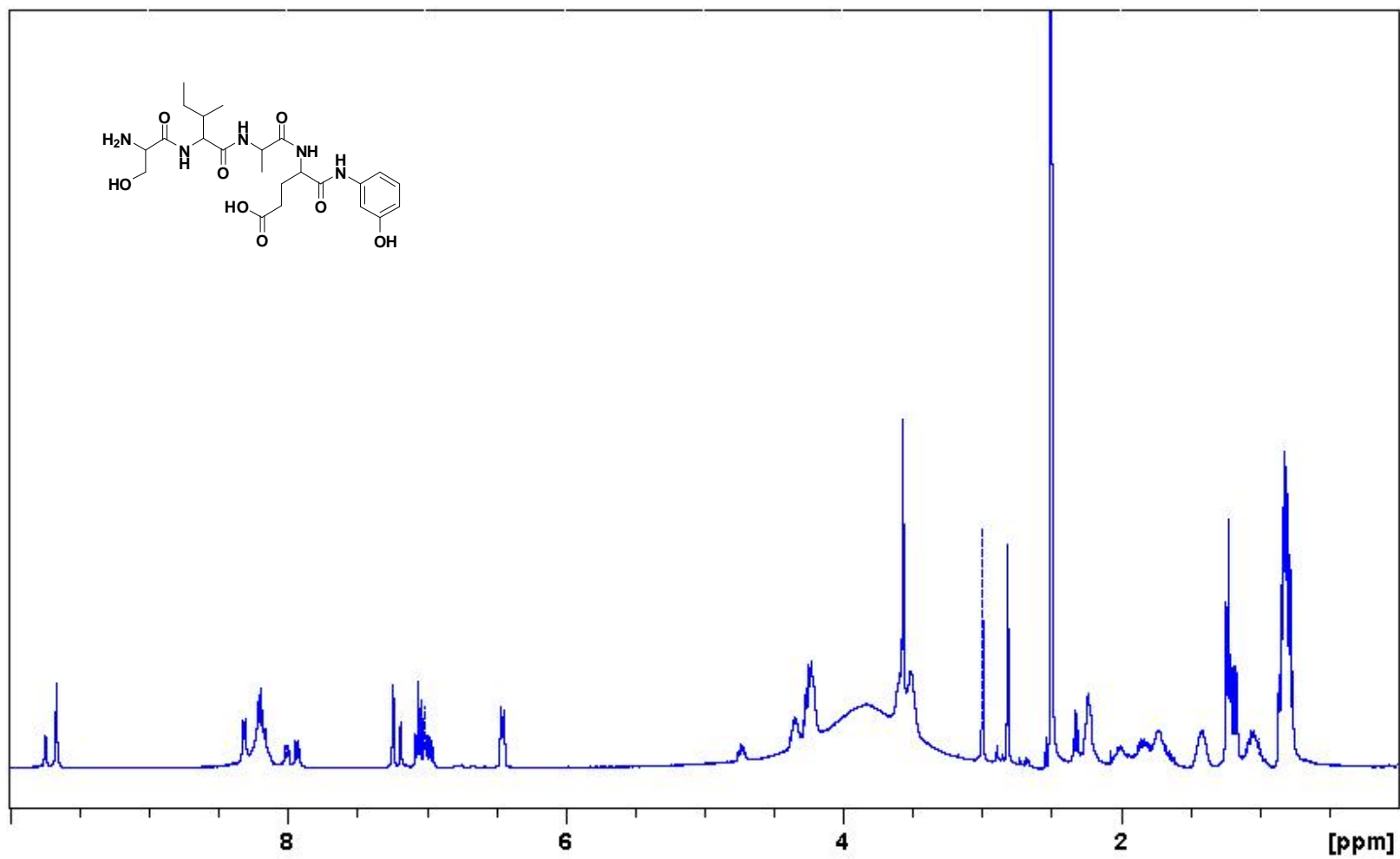


Expanded HSQC spectrum of PCU-AISb in $(\text{CD}_3)_2\text{SO}$

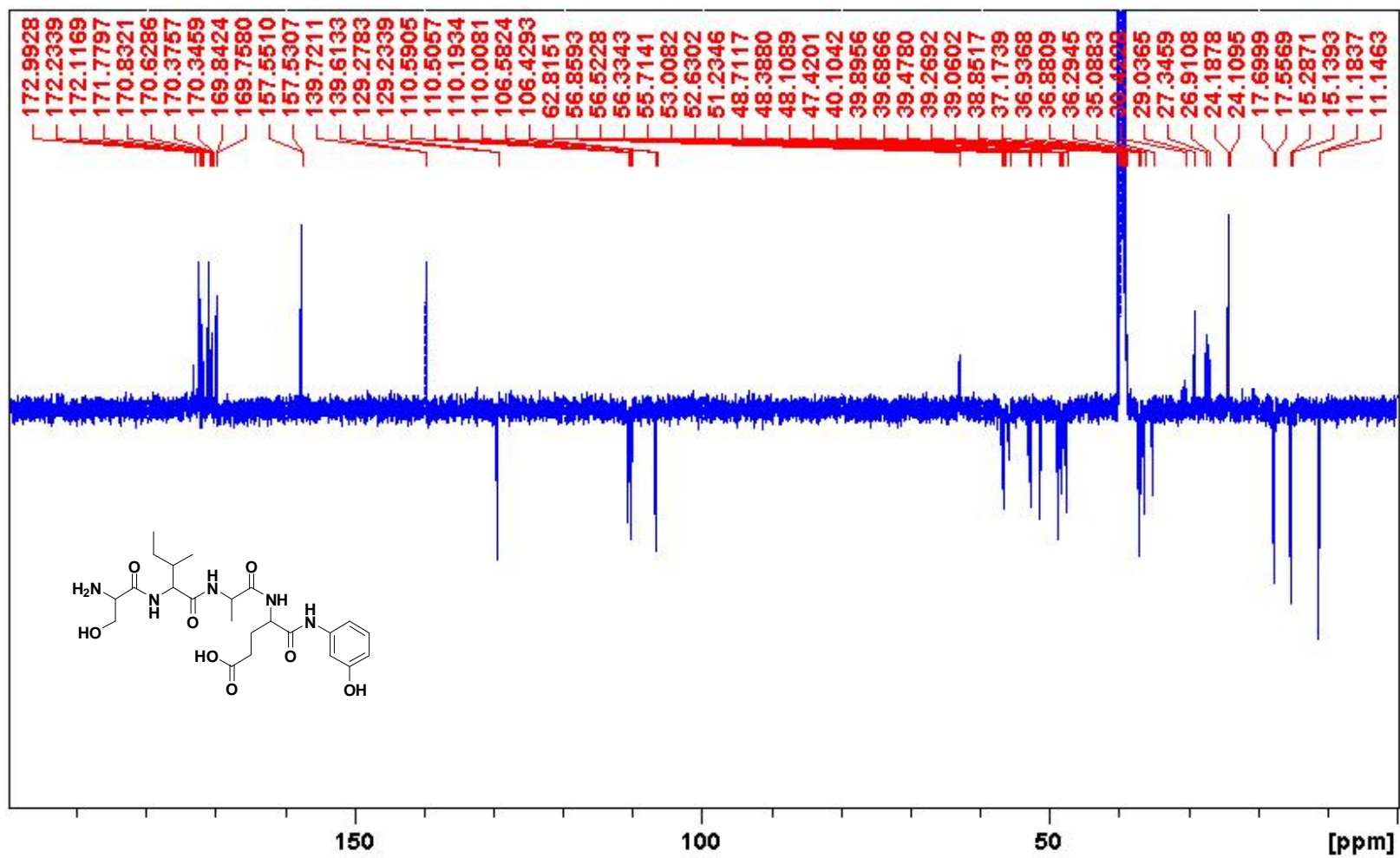
COSY spectrum of PCU-AISb in $(\text{CD}_3)_2\text{SO}$

HMBC spectrum of PCU-AISb in $(\text{CD}_3)_2\text{SO}$

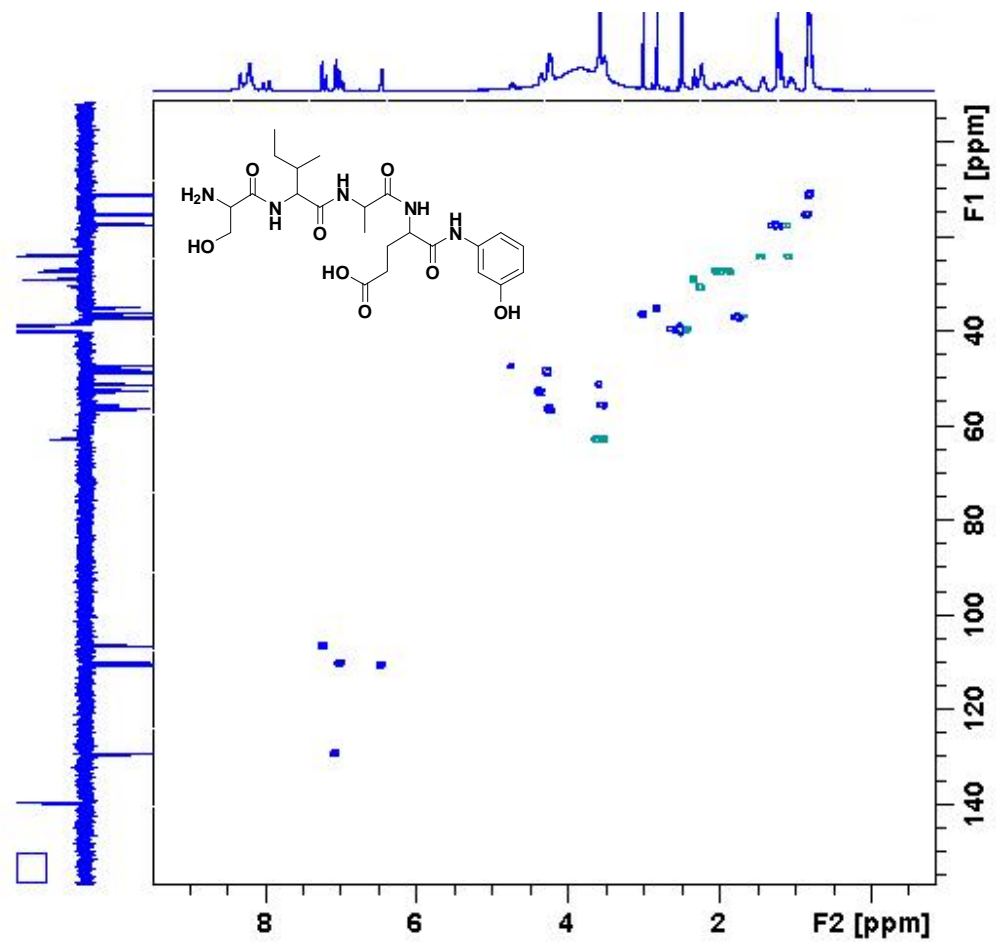
ROESY spectrum of PCU-AISb in $(\text{CD}_3)_2\text{SO}$

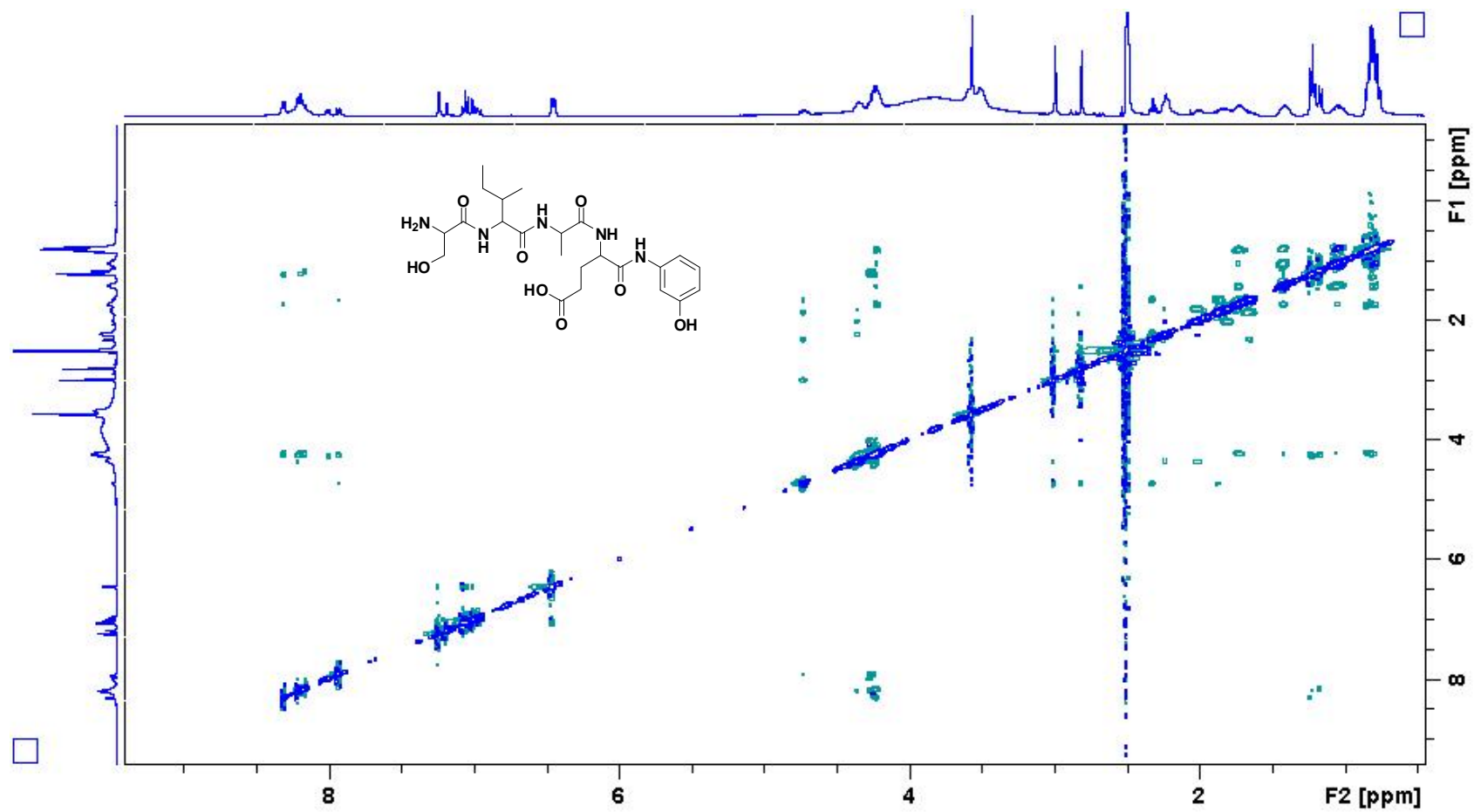


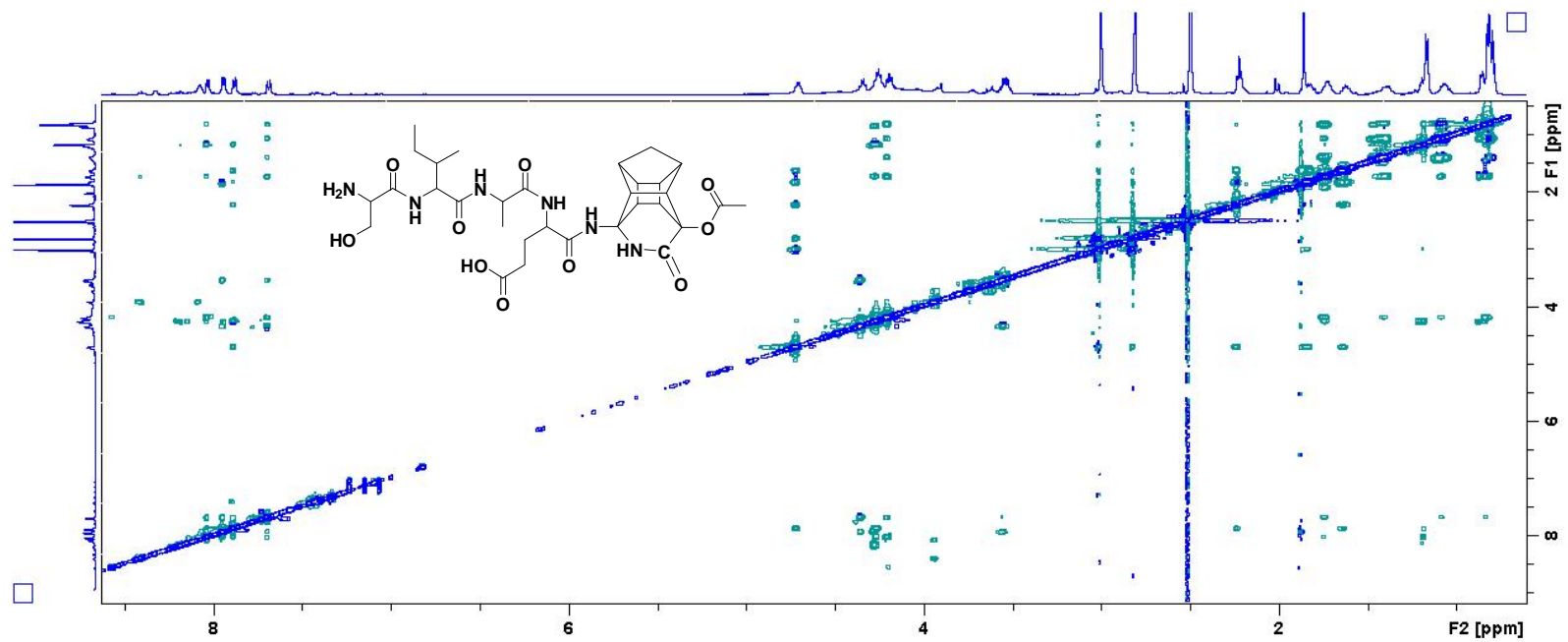
^1H NMR spectrum of Phenol-EAIS in $(\text{CD}_3)_2\text{SO}$



^{13}C NMR spectrum of Phenol-EAIS in $(\text{CD}_3)_2\text{SO}$

HSQC spectrum of Phenol-EAIS in $(\text{CD}_3)_2\text{SO}$

ROESY spectrum of Phenol-EAIS in $(\text{CD}_3)_2\text{SO}$

ROESY spectrum of Ac-PCU-EAIS in $(\text{CD}_3)_2\text{SO}$

APPENDIX 2**CHAPTER 3****SUPPORTING INFORMATION****MATERIALS AND METHODS**

Analytical analysis was performed on an Agilent 1100 HPLC with a flow rate of 1 mL/min on a Waters Xbridge C18 (150 mm x 4.6 mm x 5 microns) coupled to a UV detector (215 nm) and an Agilent VL ion trap mass spectrophotometer in the positive mode. Semi-preparative HPLC was carried out on a Shimadzu semi-preparative instrument with a flow rate of 17 mL/min on a Ace C18 (150 mm x 21.2 mm x 5 microns) with a UV/VIS detector (215 nm) and an automated fraction collector. A two-buffer system was employed, utilizing formic acid as the ion-pairing agent. Buffer A consisted of 0.1 % formic acid/H₂O (v/v) and buffer B consisted of 0.1 % formic acid/acetonitrile (v/v). High resolution mass spectroscopic analysis was performed on a Bruker MicroTOF QII mass spectrometer in positive mode with an internal calibration. Peptides were synthesized on an automated CEM Liberty microwave peptide synthesizer (see Table 1 for conditions, not same as supplier). All ¹H, ¹³C, HSQC, COSY and HMBC NMR data were recorded on a Bruker AVANCE III 400 MHz spectrometer. The ROESY data was recorded on a Bruker AVANCE III 600 MHz spectrometer.

Supplementary Table 4. Microwave conditions for coupling and deprotection

	Microwave power (Watts)	Temperature (°C)	Time (sec)
Single coupling			
	0	25	900
30 minute coupling	35	73	900
Deprotection	40	73	180

SYNTHESIS OF AMINO LACTAM

To an ice cold solution of 25 % ammonia (15 mL) the dione (1.00 g, 5.7 mmol) was added, followed by NH_4Cl (0.42 g, 9.1 mmol, 2.2 eq) and NaCN (0.45 g, 7.9 mmol, 2.3 eq), respectively. The mixture was stirred in the ice-bath for 2 hours resulting in a white thick solution that was filtered using a sintered glass funnel. The precipitate was washed with acetone (2 x 15 mL) and the product was recrystallized from methanol to yield **14** as white crystals (73 % yield).

GENERAL PROCEDURE FOR LOADING OF FIRST AMINO ACID TO THE 2-CHLOROTRITYL CHLORIDE RESIN

Activated 2-chlorotrityl chloride resin (1 g, 1.33 mmol) was swelled in dry DCM (10 mL) for 10 minutes in a sintered glass reaction vessel. Fmoc-amino acid-OH (3.99 mmol) was dissolved in DCM (10 mL) and DIPEA (6.65 mmol) and then added to the reaction vessel. The mixture was bubbled for 2 hours using nitrogen. The solvent was removed by filtration and the resin was washed with DCM (3 x 10 mL). A few resin beads (~5 mg) were removed from the reaction vessel and left to dry in the vacuum oven for an hour. A solution of 20 % piperidine in DMF was added to the dried resin beads (1 mg) and left to stand for 20 min. General washing procedure for manual SPPS: after each step, the resin was washed with DCM (3 x 10 mL), DMF (3 x 10 mL) then DCM (3 x 10 mL).

GENERAL PROCEDURE FOR THE SYNTHESIS OF PEPTIDES USING MICROWAVE POWER

Stock solutions of all amino acids (0.2 mM), DIPEA (1 mM) and HBTU (2 mM) were prepared in DMF and peptides were synthesized on a 0.5 mmol scale. The modified coupling method described in Table 1 above was used for coupling the subsequent amino acids to the 2-chlorotrityl chloride resin (1 g) preloaded with the first amino acid.

Final cleavage : The resin bound peptide was removed from the peptide synthesizer and transferred to a manual peptide synthesis reaction vessel for the final cleavage. It was washed with DCM (3 x 10 mL) and a cleavage mixture of 5:95 % (v/v) TFA:DCM was added to the resin while nitrogen was bubbled through the solution for 10 minutes. The resin was washed three times with the cleavage mixture and the cleaved peptide was removed by filtration and collected in a flask containing water (100 mL). The filtrate was extracted several times with DCM so as to remove the peptide from the water layer. DCM was removed under reduced vacuum with at 40°C and the peptide remained as a white powder.

GENERAL PROCEDURE FOR THE SYNTHESIS OF THE PEPTOIDS.

Activated 2-chlorotrityl chloride resin (1 g, 1.33 mmol) was swelled in dry DCM (10 mL) for 10 minutes in a sintered glass reaction vessel. Bromoacetic acid (1.6 g) in DMF (final concentration of 1.2 M) and DIPEA (2 mL) was added to the reaction vessel and the mixture was bubbled for 2 hours with nitrogen. The solvent was removed by filtration and the resin washed with DCM (3 x 10 mL) and DMF (3 x 10 mL). A solution of the amine (1 M) in DMF was added to the reaction vessel and bubbled for 1 hr using nitrogen. The solvent was removed by filtration and the resin was washed with DCM (3 x 10 mL) and DMF (3 x 10 mL). Bromoacetic acid (1.6 g) in DMF (final concentration of 1.2 M) and DIC (2 mL) was added to the reaction vessel and bubbled for 2 hours using nitrogen. These steps were repeated for the desired sequences. Final cleavage of the peptoid from the resin was carried out similarly as that of peptides described above.

GENERAL PROCEDURE FOR COUPLING OF PEPTIDES AND PEPTOIDS TO PCU-LACTAM

The cleaved peptide (1.2 eq.) was dissolved in DCM (3 mL) followed by addition of PCU lactam **14** (1 eq), HATU (2.5 eq) in DMF (7 mL) and DIPEA (3 eq). The mixture was stirred at room temperature for 24 hours and thereafter evaporated to dryness under reduced pressure using a Teflon pump at 40°C. A cleavage mixture (10 mL) of 95:5 (v/v) TFA:DCM was added to the crude product and stirred for 24 hours at room temperature. The TFA was removed from the mixture by bubbling air through the solution and the remaining DCM was removed under reduced pressure at 30°C. The product was obtained as yellow oil which was purified by reverse phase semi-preparative HPLC.

OVER-EXPRESSION, EXTRACTION AND PURIFICATION OF THE C-SA PROTEASE

Plasmid encoding HIV-1 subtype C protease (containing the mutation Q7K designed to reduce the hypersensitive autolytic site) is expressed as inclusion bodies (2) in *Escherichia coli* BL21 (DE3) pLysS cells. Briefly, *Escherichia coli* cells harboring the plasmid DNA were grown at 37 °C in LB medium supplemented with 100 µg/mL of ampicillin and 35 µg/mL of chloramphenicol. The overnight culture was diluted 100-fold into fresh 2 × YT medium supplemented with ampicillin (100 µg/mL) and chloramphenicol (35 µg/mL) and grown at 37 °C. When the optical density (OD₆₀₀) of the culture

reached 0.4 to 0.5, over-expression of the HIV-1 C-SA protease was induced by adding IPTG. IPTG was added to final concentrations of 0.4 mM. Over-expression of the protease was allowed to continue for four hours.

The cells were pelleted after growth and resuspended in ice-cold extraction buffer [10 mM Tris, 1 mM EDTA, and 1 mM PMSF (added only fresh before use), pH 8] and disrupted using an ultra-sonicator. Following the addition of MgCl₂ and DNase I to final concentrations of 10 mM and 10 µg/mL, respectively, the culture medium was stirred on ice until the viscosity of the mixture decreased. The cells were then ruptured by sonication and centrifuged at 15 000 × g for 30 minutes at 4 °C. The pellet was resuspended in ice-cold extraction buffer containing 1 % (v/v) of Triton X-100. Cell debris and protease-containing inclusion bodies were pelleted by centrifugation at 15 000 × g for 30 minutes at 4 °C. The pellet was then resuspended in a freshly prepared solubilization buffer containing 10 mM Tris, 2 mM DTT, 8 M urea, pH 8.0, at room temperature, and centrifuged at 15 000 × g for 30 minutes at 20 °C.

The protease, in the supernatant, was purified by passing it through an anion exchange (DEAE) column previously equilibrated with solubilization buffer. Upon elution from the column, the protease was acidified by adding formic acid to a final concentration of 25 mM. Precipitation of significant amount of contaminating proteins occurred upon acidification. Following an overnight incubation, the precipitated contaminants were removed by centrifugation at 15 000 × g for 30 minutes at 4 °C.

HIV-1 protease was refolded by dialysis into 10 mM formic acid at 4 °C. Subsequently, the protease was dialyzed into storage buffer containing 10 mM sodium acetate, 1 mM NaCl and 1 mM DTT, pH 5.0. The folded protease was concentrated to a final volume of ~5 mL and stored at -20 °C.

***IN VITRO* HIV-1 PROTEASE ACTIVITY**

The catalytic activity of the HIV-1 protease was monitored following the hydrolysis of the chromogenic peptide substrate Lys-Ala-Arg-Val-Nle-*p*-nitro-Phe-Glu-Ala-Nle-NH₂. This substrate mimics the conserved KARVL/AEAM cleavage site between the capsid protein and nucleocapsid (CA-p2) in the Gag polyprotein precursor.

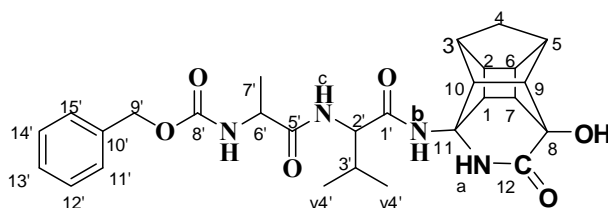
For this study the chromogenic substrate was synthesized using the Discovery CEM Liberty microwave peptide synthesizer on Rink amide resin. The substrate was cleaved from the resin and deprotected using 95 %:5 % (v/v) TFA:TIS for 3 hours. It was then precipitated using cold ether and purified *via* reverse phase semi-preparative HPLC on a Shimadzu instrument and characterized using the Bruker microTOF-Q II instrument.

To determine the concentration of the inhibitors that resulted in 50 % inhibition (IC_{50}) of HIV-1 protease enzyme activity, the protein (100 nM) and chromogenic substrate (50 μ M) were added into a 120 μ L microcuvette containing increasing concentrations of inhibitor in a pH 5.0 buffer (50 mM sodium acetate and 0.1 M NaCl). Protease hydrolytic activity was measured by monitoring the relative decrease in absorbance at 300 nm using Fpecord 210 spectrophotometer.

Table 2. High Resolution Mass Spectral data and percentage yields

Compounds	Molecular Formula	Found mass	Calculated mass	Molecular ion	IC_{50} (μ M)	% Yield
15	$C_{37}H_{43}N_5O_7$	670.3228	670.3235	(M+H ⁺)	0.75 \pm 0.035	52.5
16a	$C_{28}H_{34}N_4O_6$	523.2554	523.2551	(M+H ⁺)	0.6 \pm 0.071	76.2
16b	$C_{20}H_{28}N_4O_4$	389.2186	389.2183	(M+H ⁺)	0.75 \pm 0.071	78.3
16c	$C_{26}H_{32}N_6O_4S$	546.2943	546.2928	(M+Na ⁺)	0.5 \pm 0.035	43.2
17	$C_{37}H_{43}N_5O_7$	670.3255	670.3249	(M+H ⁺)	0.5 \pm 0.035	46.2
18a	$C_{28}H_{34}N_4O_6$	545.2375	545.2371	(M+Na ⁺)	10.0 \pm 0.71	65.3
18b	$C_{20}H_{28}N_4O_4$	389.2185	389.2183	(M+H ⁺)	10.0 \pm 0.71	68.3
18c	$C_{26}H_{32}N_6O_4S$	546.2947	546.2948	(M+Na ⁺)	5.0 \pm 1.06	38.3

NMR elucidation of PCU-VAZ, 16a



Supplementary Table 1. ^1H and ^{13}C chemical shift of PCU-VAZ as measured by HSQC, COSY, HMBC, ROESY and ^1H and ^{13}C spectra. The chemical shift was referenced to 2.50 ppm for COSY and ROESY the correlation aided in assignment are marked in black (sequential correlations) and the long-range interactions through space, measured by ROESY, are marked in red

Atom	$\delta^1\text{H}^{\text{a,b}}$	$\delta^{13}\text{C}^{\text{a,d}}$	COSY($\delta^1\text{H}^{\text{a,b}}$)	HMBC	ROESY ($\delta^1\text{H}^{\text{a,c}}$)
1x	2.79	43.5	7		7, NHa, NHb, NHbb
1y	3.11	41.9	2, 7		2, 7, v4', NHaa, NHaaa, NHbb, NHbbb
2	2.46	39.6	1y, 6	3, 5, 7	1y, 4a
3x	2.38	43.1	4a, 4s, 10y		4a, 4s, v4', 10y, NHb, NHbb
3y	2.45	43.1	4a, 4s		4a, 4s, v4', NHbb, NHbbb
4a	1.34 (d. J = 10.3 Hz)	36.9	3x, 3y, 4s, 5	10x	2, 3x, 3y, 4s, 5, 6
4s	1.68 (d. J = 10.5 Hz)	36.9	3x, 3y, 4a, 5		3x, 3y, 4a, 5, 9, 10x, 10y
5	2.66	45.2	4a, 4s, 9		4a, 4s, 9
6	2.74	41.7	2, 7		4a, 7
7	2.34	44.4	1x, 1y, 6,	2	1x, 1y, 6

8		79.9			
9	2.17 (d. J = 10.8 Hz)	51.9	4s, 5, 10x, 10y		4s, 5, 10x, 10y
10x	2.58	51.4	9		4s, 9, NHaa, NHaaa, NHbb, NHbbb
10y	2.91	50.1	3x, 9		3x, 4s, 9, NHa, NHb, NHbb
11		72.9			
12		173.5			
1'		171.1			
2'x	4.19	57.3	3'	1', 3', v4'	3', v4', NHbb, NHbbb
2'y	4.25	57.0	3'	1', 3', v4'	3', v4', NHb, NHbb
3'	1.95	30.9	2'x, 2'y, v4'		2'x, 2'y, v4', NHb, NHbb, NHbbb, NHc, NHcc, NHccc
v4'	0.78	17.5	3'	2', 3'	1y, 2'x, 2'y, 3', 3x, 3y, NHa, NHaa, NHb, NHbb, NHbbb, NHc, NHcc, NHccc
v4'	0.82	19.4	3'	2', 3'	1y, 2'x, 2'y, 3', 3x, 3y, NHa, NHaa, NHb, NHbb, NHbbb, NHc, NHcc, NHccc
5'		172.2			
6'x	4.11	50.2	7'		7', NHc, NHccc
6'y	4.16	50.1	7'		7', NHcc, NHd
7'	1.20	18.5	6'x, 6'y	5', 6'	6', 11'-15', NHc, NHcc, NHccc, NHd

8'		155.6			11'-15'
9'	5.06	65.4		11'-15'	11'-15'
10'		137.0			
11'-15'	7.30-7.36	127.6-128.3		9'	7', 9'
NHa	8.04				1x, v4', 10y
NHaa	8.06			8	v4', 10x, 1y
NHaaa	8.09			8	1y, 10x
NHb	8.17				1, 2'y, 3', 3, v4', 10y
NHbb	8.24			1', 11	1x, 1y, 2'x, 2'y, 3', 3x, 3y, v4', 10x, 10y
NHbbb	8.32			1', 11	1y, 2'x, 3', 3y, v4', 10x,
NHc	7.60				3', v4', 6'x, 7'
NHcc	7.79		2'x, 2'y		3', v4', 6'y, 7'
NHccc	7.94		2'y		3', v4', 6'x, 7'
NHd	7.47		6'		6'y, 7'

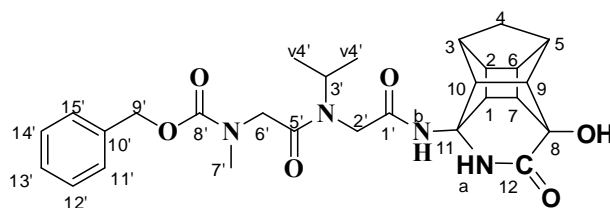
a Solvent (CD₃)₂SO.

b 400 MHz for ¹H.

c 600 MHz for ¹H.

d 100 MHz for ¹³C.

NMR elucidation of PCU-VAZ peptoid, 18a



Supplementary Table 2: ^1H and ^{13}C chemical shift of PCU-VAZ peptoid as measured by HSQC, COSY, HMBC, ROESY and ^1H and ^{13}C spectra. The chemical shift was referenced to 2.50 ppm for COSY and ROESY the correlation aided in assignment are marked in black (sequential correlations) and the long-range interactions through space, measured by ROESY, are marked in red

Atom	$\delta^1\text{H}^{\text{a,b}}$	$\delta^{13}\text{C}^{\text{a,d}}$	COSY ($\delta^1\text{H}^{\text{a,b}}$)	HMBC	ROESY ($\delta^1\text{H}^{\text{a,c}}$)
1	2.89	43.1	2, 7		2, v4'x, 7, NHa, NHb, NHbb, NHbbb, NHbbbb
2	2.49	39.6	1, 6		1, 4a, v4'x, NHbbbb
3	2.43	43.1	4a, 4s		4a, 4s, v4'x, 10, NHb, NHbb, NHbbbb
4a(d. $J = 10.6$ Hz)	1.35	36.9	3, 4s, 5		2, 3, 4s, 5, 6
4s	1.68	36.9	3, 4a, 5		3, 4a, 5, 9, 10
5	2.66	45.1	4a, 4s, 9		4a, 4s, 9
6	2.74	41.7	2, 7		4a, 7
7	2.34	44.4	1, 6,	2, 5	1, 6, NHbbb
8		79.5			
9	2.17	51.9	5, 10		4s, 5, 10, NHbbb
10	2.79	50.6	9		3, 4s, v4'x, 9, NHa, NHb, NHbb, NHbbb, NHbbbb

11		72.9			
12		173.4			
1'		169.2			
2'x	3.74	43.8	v4'	1', 3'	3', v4'x, v4'y, v4'z, NHb, NHbb, NHbbbb
2'y	3.79	44.2	v4'	1', 3'	3', v4'x, v4'y, v4'z, NHb, NHbb, NHbbbb
2'z	3.82	44.2	v4'	1', 3'	3', v4'x, v4'y, v4'z, NHb, NHbb, NHbbbb, NHbbbb
3'	4.61	44.5			2'x, 2'y, 2'z, v4'
v4'x	0.96	19.5	3'	3'	1, 2, 2'x, 2'y, 2'z, 3', 3, 10, NHa, NHb, NHbb, NHbbb, NHbbbb
v4'y	1.05	20.6	3'	3'	2'x, 2'y, 2'z, 3'
v4'z	1.13	20.6	3'	3'	2'x, 2'y, 2'z, 3'
5'		168.3			
6'x	4.05	50.5		5', 8'	7'x, 7'y
6'y	4.09	50.2		5', 8'	7'x, 7'y
6'z	4.18	50.2		5', 8'	7'x, 7'y
7'x	2.82	35.8		8'	6'x, 6'y, 6'z, 11'-15'
7'y	2.86	35.1		8'	6'x, 6'y, 6'z, 11'-15'
8'		156.1			
9'x	4.99	66.1		8, 11'-15'	11'-15'
9'y	5.04	65.8		8, 11'-15'	11'-15'

9'z	5.06	66.1		8, 11'-15'	11'-15'
10'		137.1			
11'-15'	7.25-7.32	126.8-128.3		9'	7', 9'
NHa	8.16			8, 12	1, v4', 10
NHb	7.95				1, 2'x, 2'y, 2'z, 3, v4', 10
NHbb	8.02				1, 2'x, 2'y, 2'z, 3, v4', 10
NHbbb	8.23				1, 2'z, 7, 9, v4', 10
NHbbbb	8.36			1', 1, 10, 11	1, 2, 2'x, 2'y, 2'z, 3, v4', 10

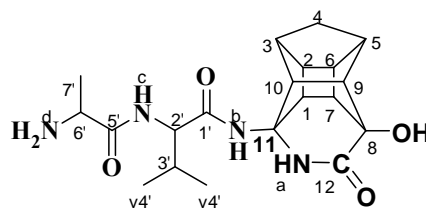
a Solvent (CD₃)₂SO.

b 400 MHz for ¹H.

c 600 MHz for ¹H.

d 100 MHz for ¹³C.

NMR elucidation of PCU-VANH₂, 16b



Supplementary Table 3. ¹H and ¹³C chemical shift of PCU-VANH₂ as measured by HSQC, COSY, HMBC, ROESY and ¹H and ¹³C spectra. The chemical shift was referenced to 2.50 ppm for COSY and ROESY the correlation aided in assignment are marked in black (sequential correlations) and the long-range interactions through space, measured by ROESY, are marked in red

Atom	$\delta^1\text{H}^{a,b}$	$\delta^{13}\text{C}^{a,d}$	$\delta^{13}\text{C}^{a,d}$	ROESY ($\delta^1\text{H}^{a,c}$)
	PCU-VANH ₂	PCU-VANH ₂	PCU-VAZ	PCU-VANH ₂
1x	2.77	43.5	43.5	7, NHa, NHc
1y	3.10	41.6	41.9	2, 7, NHb, NHd
2	2.49	39.7	39.6	1y
3x	2.39	43.5	43.1	4a, 4s, v4', 10y, NHc, NHd
3y	2.45	43.2	43.1	4a, 4s, v4'
4a	1.36	36.9	36.9	3x, 3y, 4s, 5, 6
4s(d. <i>J</i> = 10.4 Hz)	1.70	36.9	36.9	3x, 3y, 4a, 5, 9, 10y
5	2.67	45.2	45.2	4a, 4s, 6, 9
6	2.75	41.6	41.7	4a, 5, 7
7	2.36	44.3	44.4	1x, 1y, 6
8		79.5	79.9	

9	2.18	51.9	51.9	4s, 5, 10x, 10y
10x	2.58	51.3	51.4	9, NHb
10y	2.92	50.0	50.1	3x, 4s, 9, NHa, NHc
11		72.9	72.9	
12		173.6	173.5	
1'		170.8	171.1	
2'x	4.23	57.8	57.3	3', v4', NHb, NHc
2'y	4.34	57.7	57.0	3', v4', NHbb, NHcc
3'	1.95	30.9	30.9	2'x, 2'y, v4', NHb, NHbb, NHc, NHcc
v4'	0.85	18.0	17.5	2'x, 2'y, 3', 3x, 3y, NHb, NHbb, NHc, NHcc
v4'	0.87	19.4	19.4	2'x, 2'y, 3', 3x, 3y, NHb, NHbb, NHc, NHcc
5'		169.3	172.2	
6'	3.94	50.0	50.2	7'x, 7'y, NHc, NHcc, NHd
7'x	1.33	17.6	18.5	6', NHc, NHd
7'y	1.36	18.0		6', NHcc
8'			155.6	
9'			65.4	
10'			137.0	
11'-15'			127.6-128.3	
NHa	8.04			1x, 10y

NHaa				
NHaaa				
NHb	8.41			1y, 2'x, 3', v4', 10x
NHbb	8.44			1x, 1y, 2'y, 3', v4', 10x, 10y
NHbbb				
NHc	8.36			1x, 2'x, 3', 3x, v4', 6', 7'x, 10y,
NHcc	8.47			2'y, 3', v4', 6', 7'y
NHccc				
NHd	8.07			1y, 3x, 6', 7'x

a Solvent (CD₃)₂SO.

b 400 MHz for ¹H.

c 600 MHz for ¹H.

d 100 MHz for ¹³C.

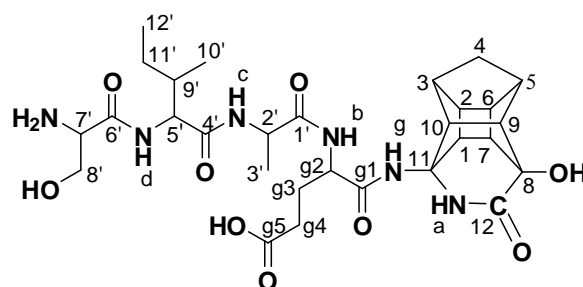
Supplementary Table 4. ^1H chemical shifts of **15**, **16c**, **17**, **18b** and **18c**. The chemical shift was referenced to 2.50 ppm

Atom	$\delta^1\text{H}^{\text{a,b}}$	$\delta^1\text{H}^{\text{a,b}}$	$\delta^1\text{H}^{\text{a,b}}$	$\delta^1\text{H}^{\text{a,b}}$	$\delta^1\text{H}^{\text{a,b}}$
	15	16c	17	18b	18c
1	3.03	2.80	2.91	2.91	2.69
2	2.55	2.55	2.50	2.49	2.50
3	2.40	2.45	2.41	2.43	2.61
4a	1.34	1.47 (d. $J = 10.5$ Hz)	1.36 (d. $J = 10.3$ Hz)	1.36 (d. $J = 10.3$ Hz)	1.40 (d. $J = 10.3$ Hz)
4s	1.70	1.77 (d. $J = 10.6$ Hz)	1.70 (d. $J = 10.3$ Hz)	1.69 (d. $J = 10.3$ Hz)	1.74 (d. $J = 10.3$ Hz)
5	2.69	2.66	2.67	2.67	2.70
6	2.78	2.75	2.75	2.75	2.62
7	2.37	2.33	2.34	2.35	2.43
8	-	-	-	-	-
9	2.21	2.18	2.18	2.18	2.28
10	2.89	2.89	2.77	2.82	2.40
11	-	-	-	-	-
12, 1', 5', 8'	-	-	-	-	-
P1'	-	-	-	-	-
P2'	4.64	-	4.01-4.47	-	-
P3'a	3.11	-	3.89	-	-
P3's	3.14	-	3.91	-	-

P4'-P9'	7.16-7.45	-	7.21-7.41	-	-
2'	4.47-4.90	4.12-4.41	4.01-4.47	3.69-3.82	3.00-3.81
3'	2.07	1.98	4.61	4.60	3.54
v4'	0.85 (d. $J = 6.8$ Hz)	1.14	0.99	0.96	0.99
v4'	0.90 (d. $J = 6.8$ Hz)	1.16	1.01	1.08	0.99
6'x	4.11	4.12-4.41	4.01-4.47	3.44-3.59	3.00-3.81
6'y	4.16				
7'	1.20	1.20	2.76-2.86	2.35	2.54
9'	5.08	3.90	5.04	-	2.70
10'	-	-	-	-	-
11'-15'	7.16-7.45	7.21-7.23 8.60-8.63	7.21-7.49	-	7.40-7.97
NHa- NHd	8.01-8.89				
NHa- NHb		8.04-8.63	8.14-8.71	8.16-8.47	8.61-8.67

a Solvent (CD₃)₂SO.

b 400 MHz for ¹H.



Supplementary Table 5. EASY ROESY correlations observed in the spectra of PCU-EAIS recorded in DMSO- d_6 and the D_2O as the solvent. The chemical shift was referenced to 2.50 ppm

Atom	1H PCU-EAIS	ROESY(δ^1H^c)(PCU-EAIS) (DMSO- d_6)	ROESY(δ^1H^c)(PCU-EAIS (D_2O))
1x	2.89	2, 3,6, 7, NHc	2, 3, 7,NHa, NHc, NHcc
1y	2.97	2, 3, 6, 7, 10, NHa, NHg	2, 3, 7, NHa
2	2.45	1a, 1s, NHa, NHb	1a, 1s
3	2.41	1a, 1s, 3', 4a, 4s, 6, 10, 10', NHa, NHb, NHc, NHg	1a, 1s, 4a, 4s, 10, NHa, NHc, NHcc
4a	1.35 (d. $J = 10.6$ Hz)	3, 4s, 5, 6	3, 4s, 5, 6
4s	1.68	3, 4a, 5, 6, 9, 10	3, 4a, 5, 6, 9, 10
5	2.66	4a, 4s, 6, 9	4a, 4s, 6
6	2.75	1a, 1s, 2, 3, 4a, 4s, 5, 7, 9	4a, 4s, 5, 7, 9
7	2.34	1a, 1s, 6	1a, 1s, 6
8			
9	2.17	4s, 5, 6, 10	4s, 6, 10
10	2.77	1s, 3, 4s, 9, 10, NHa, NHb, NHc	3, 4s, 9, NHa, NHc

11			
12			
g1'			
g2'	4.22	g3a', g3s, g4, NHa	g3a', g3s, g4, NHa, NHg
g3a'	1.72	g2', g3s, g4, NHa, NHg	g2', g3s, g4, NHg, NHbb,
g3s'	1.89	g2', g3a, g4, NHa, NHg	g2', g3a, g4, NHg
g4'	2.21	3', g2', g3a', g3s', 10', NHa	g2', g3a', g3s', NHg
g5'			
1'			
2'	4.27	3', NHb, NHc, NHcc	3', NHbb, NHc, NHcc
3'	1.19	2', 3, g4', NHb, NHc, NHcc, NHg	2', NHbb, NHc, NHcc
4'			
5'	4.25	9', 10', 11a, 11s, NHd	9', 11a', 11s', 12'
6'			
7'	3.93	8a', 8s', 10', 12', NHd	8a', 8s', NHd
8a'	3.62	7', 8s', NHd	7', 8s', NHd
8s'	3.72	7', 8a', NHd	7', 8a', NHd
9'	1.73	5', 10', 11a', 11s', 12', NHc, NHcc, NHd	5', 10', 11a', 11s', 12', NHc, NHcc, NHd
10'	0.86	3, g4', 5', 7, 9', 11a', NHc, NHcc	9', 11a', NHcc
11a'	1.05	5', 9', 10, 11s', 12', NHd	5', 9', 10, 11s', NHd
11s'	1.43	5', 7, 9', 11a', 12' NHd, NHc	5', 9', 11a', 12'

12'	0.82	5', 9', 11a', 11s', NHcc, NHd	5', 9', 11a', 11s'
NHa	8.08	1a, 1s, 2, 3, 10, g2', g3a', g3s, g4'	1a, 1s, 2, 3, 10, g2'
NHaa			
NHg	8.10	1s, 3, 3', g3a', g3s	g2', g3a', g3s, g4'
NHb	7.80	2', 2, 3, 3', g2', g3a', g3s, g4'	
NHbb	7.94	2', 2, 3, 3', g2', g3a', g3s, g4'	2', 3', g3a'
NHc	8.16	1a, 2', 3', 3, 6, 9', 10', 11s'	1a, 2', 3', 3, 9', 10'
NHcc	8.20	2', 3', 9', 10', 12'	1a, 2', 3, 3', 3, 9', 10'
NHd	8.40	2'/5', 7', 8a', 8s', 9', 11a', 11s', 12'	7', 8a', 8s', 9', 11a'

c 600 MHz for ^1H .

Display Report

Analysis Info

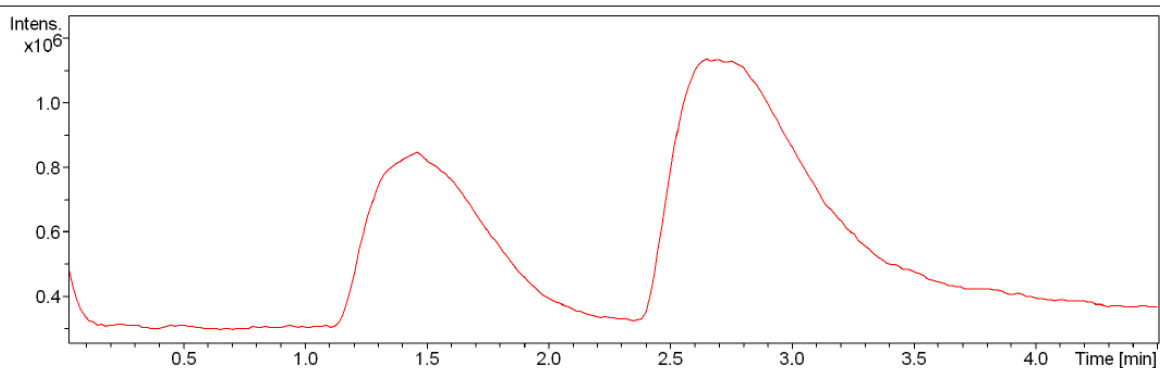
Analysis Name D:\Data\Sri\Lactam-Phe-Val-Ala-CBz1.d
 Method tune_wide_expert.m
 Sample Name Lactam-Phe-Val-Ala-CBz1
 Comment

Acquisition Date 12/19/2009 6:33:01 PM

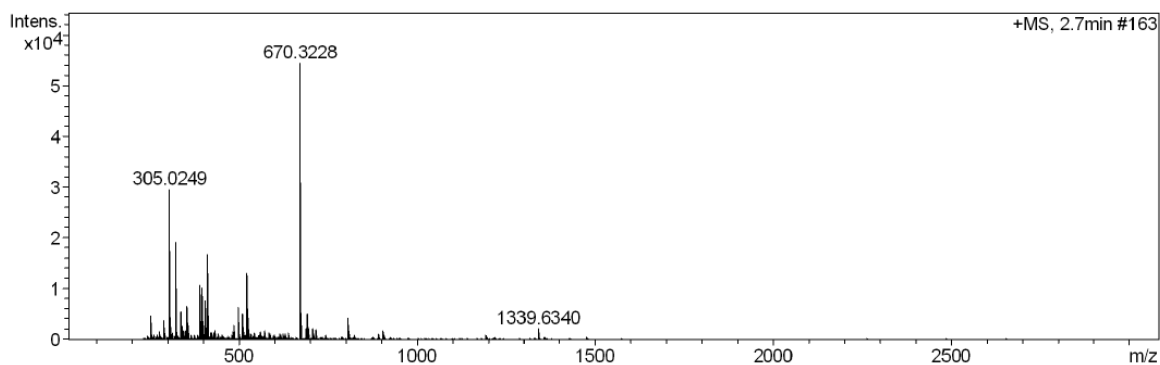
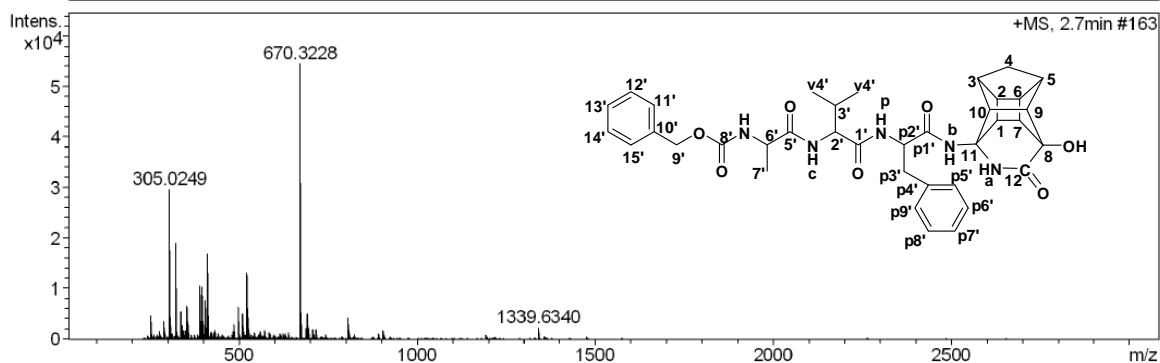
Operator BDAL@DE
 Instrument micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



TIC +All MS



Display Report

Analysis Info

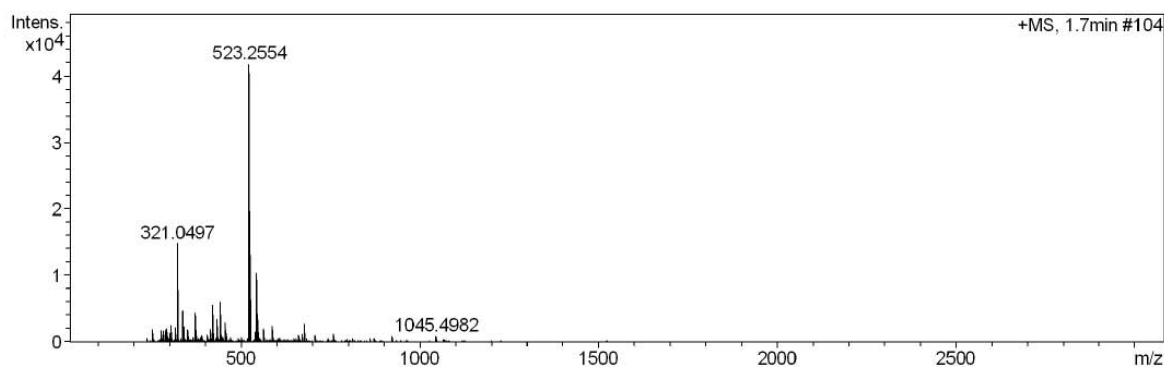
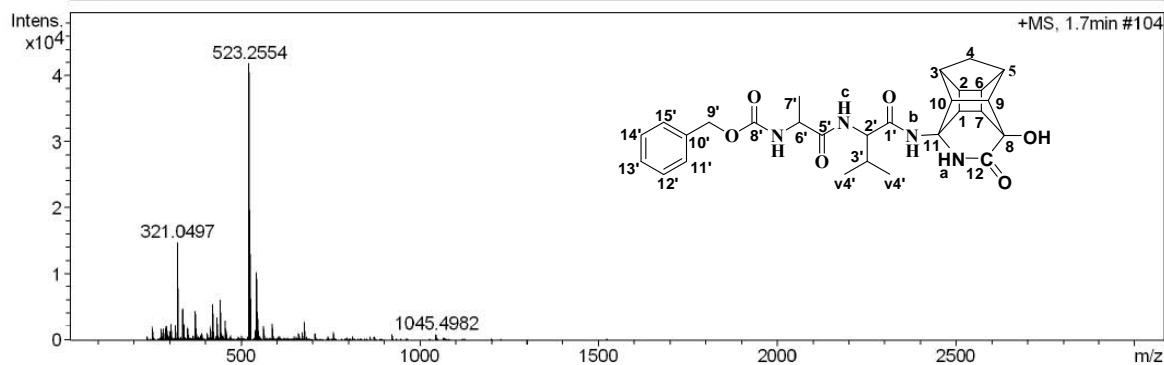
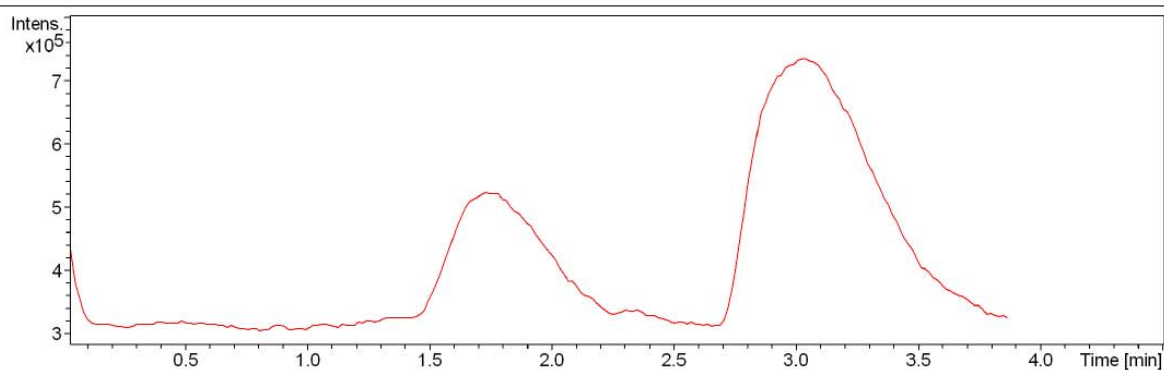
Analysis Name D:\Data\Sri\Lactam-Val-Ala-CBz2.d
 Method tune_wide_expert.m
 Sample Name Lactam--Val-Ala-CBz2
 Comment

Acquisition Date 12/19/2009 6:41:05 PM

Operator BDAL@DE
 Instrument micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



Display Report

Analysis Info

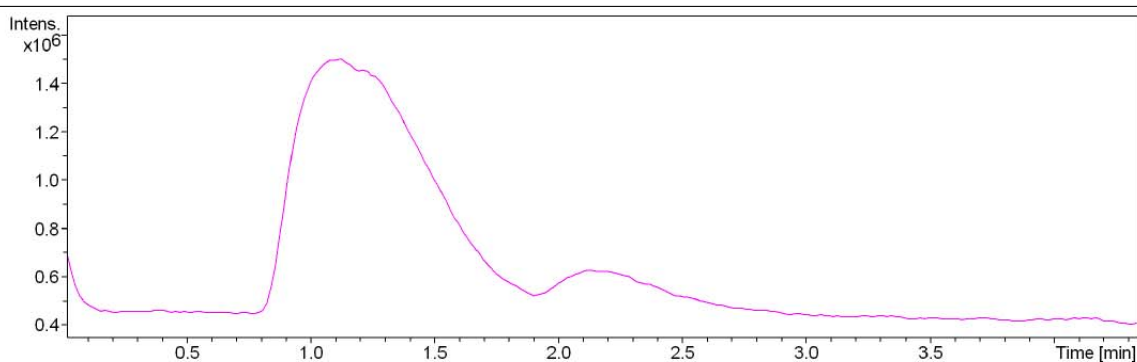
Analysis Name D:\Data\Sri\lactam-Val-Ala-NH2.d
 Method tune_wide_expert.m
 Sample Name lactam-Val-Ala-NH2
 Comment

Acquisition Date 12/19/2009 5:35:45 PM

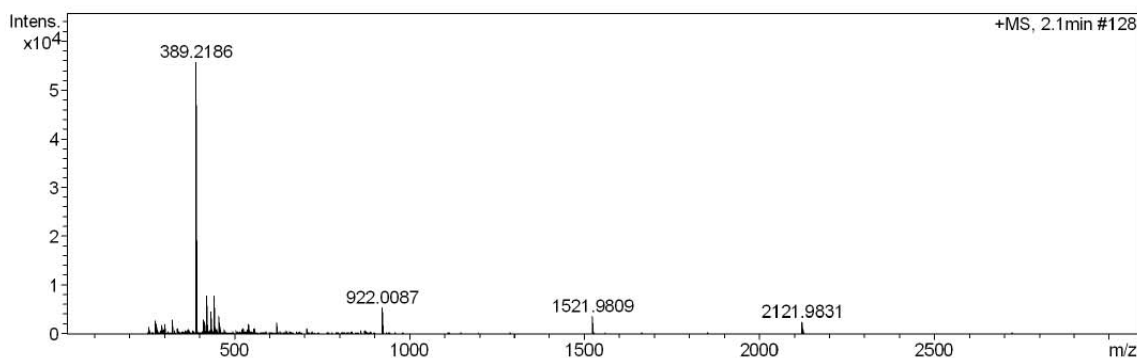
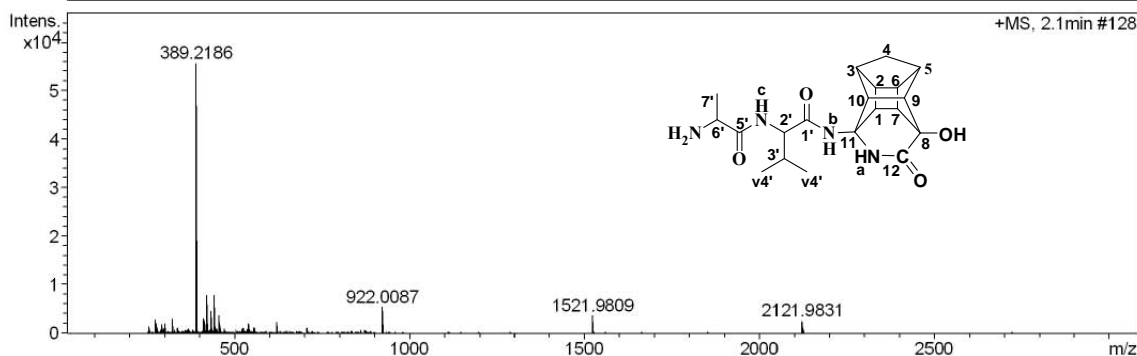
Operator BDAL@DE
 Instrument micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



— TIC +All MS



Display Report

Analysis Info

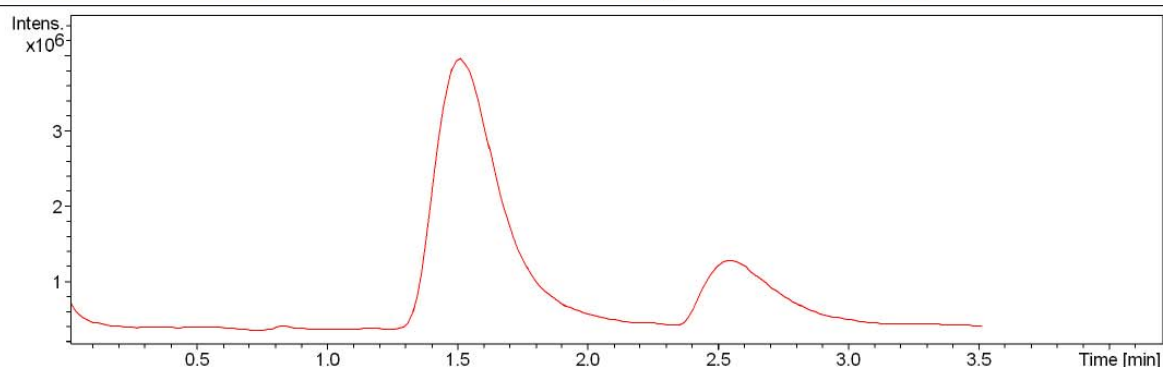
Analysis Name D:\Data\Sri\L-Val-Ala-2-Pyrimidyl-ethanethioicacid.d
 Method tune_wide_expert.m
 Sample Name L-Val-Ala-2-Pyrimidyl-ethanethioicacid
 Comment

Acquisition Date 12/25/2009 11:35:02 AM

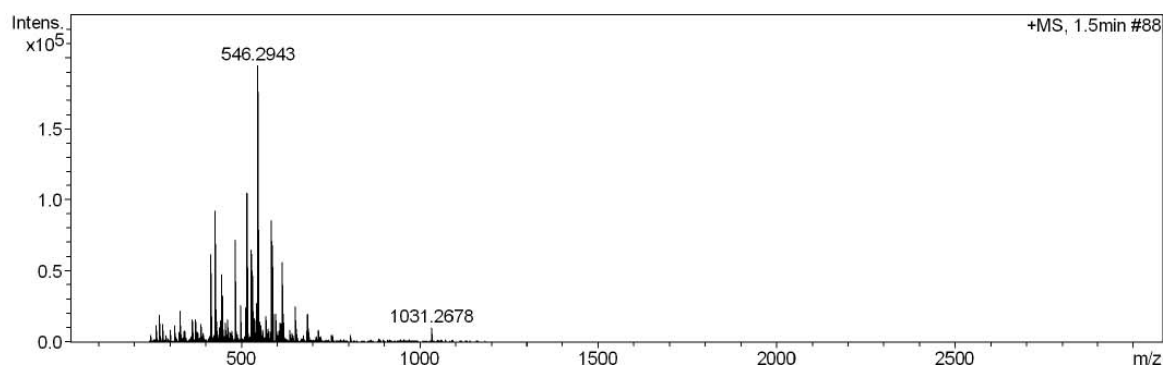
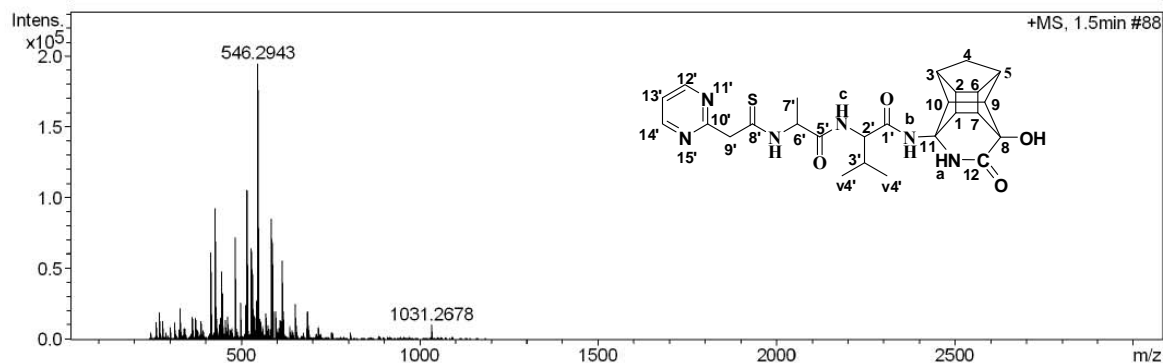
Operator BDAL@DE
 Instrument micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



— TIC +All MS



Display Report

Analysis Info

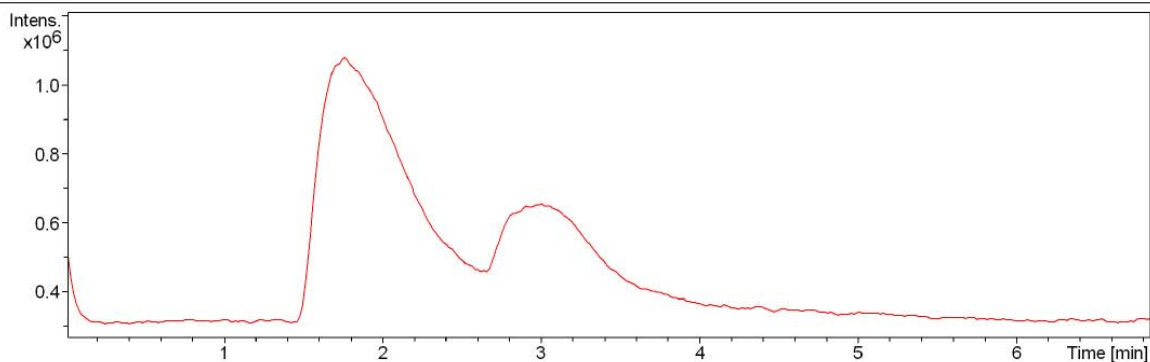
Analysis Name D:\Data\Sri\Lactam-Phe-Val-Ala-CBz-PEPTOID1.d
 Method tune_wide_expert.m
 Sample Name Lactam-Phe-Val-Ala-CBz-PEPTOID1
 Comment

Acquisition Date 12/19/2009 6:23:18 PM

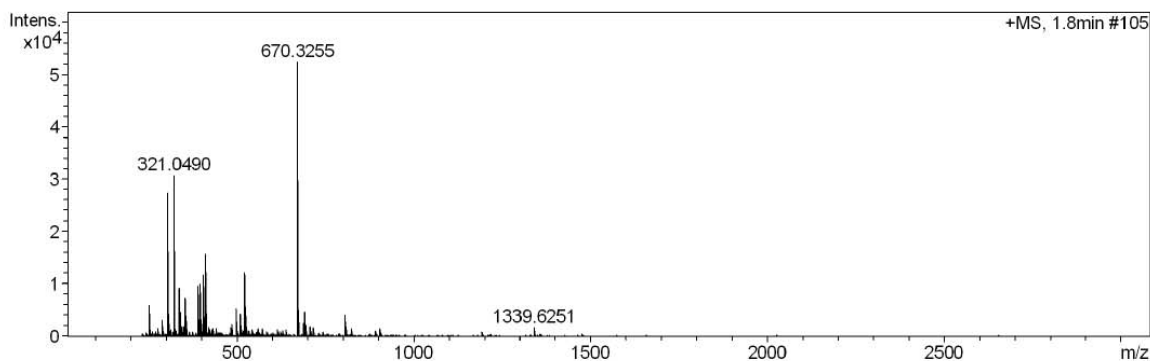
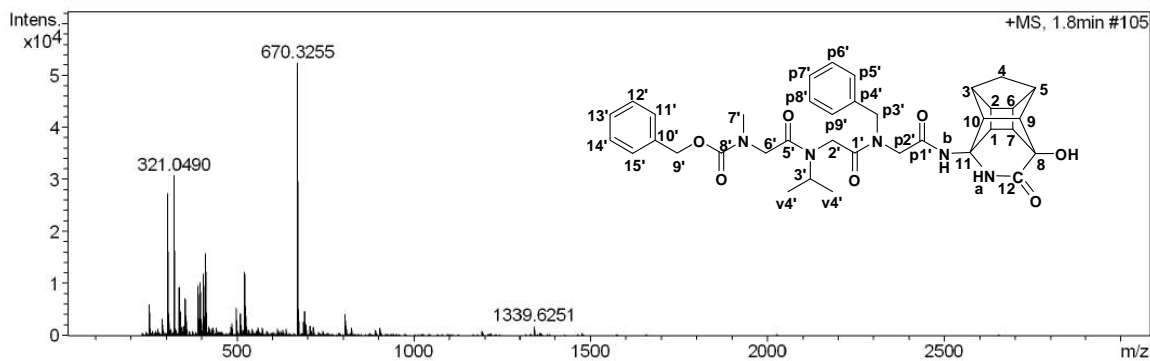
Operator BDAL@DE
 Instrument micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



— TIC +All MS



Display Report

Analysis Info

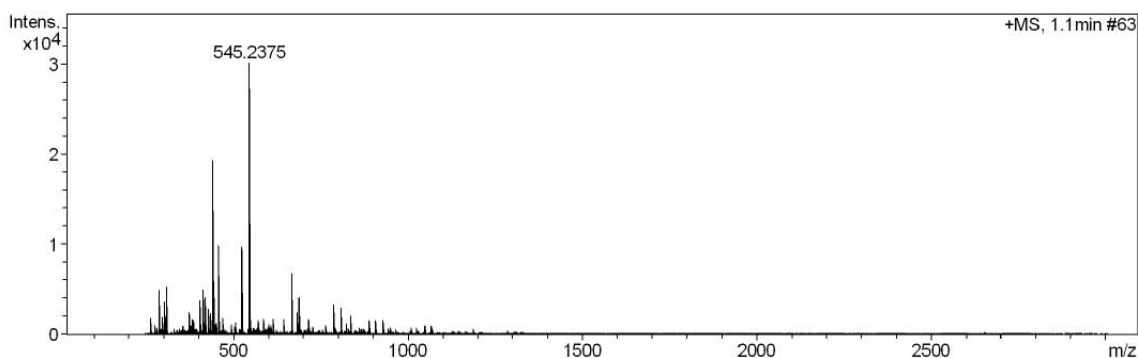
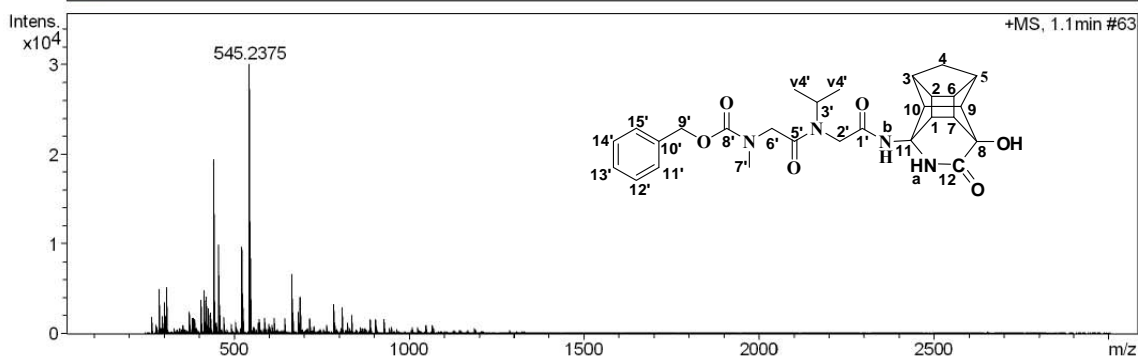
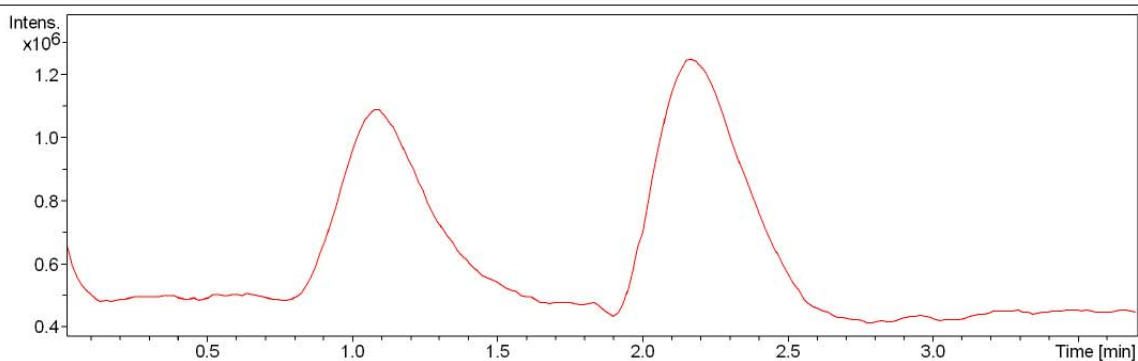
Analysis Name D:\Data\Sri\lactam-Val-Ala-CBz-peptoid.d
 Method tune_wide_expert.m
 Sample Name lactam-Val-Ala-CBz-peptoid
 Comment

Acquisition Date 12/19/2009 5:20:40 PM

Operator BDAL@DE
 Instrument micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



Display Report

Analysis Info

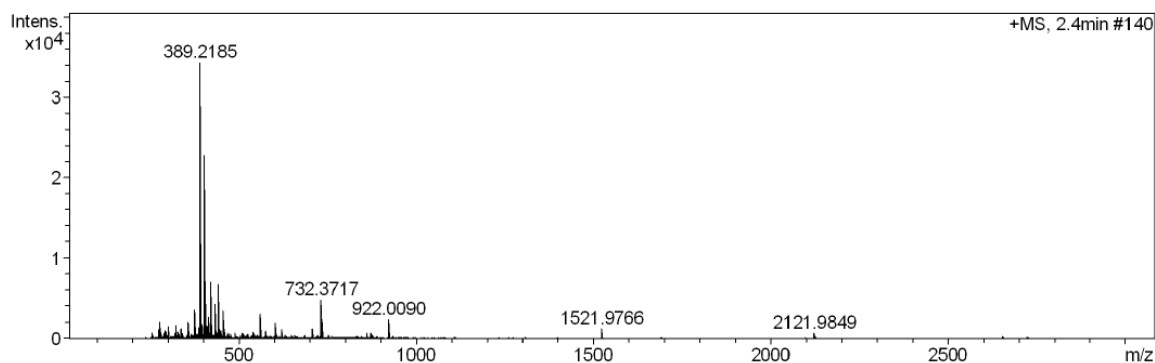
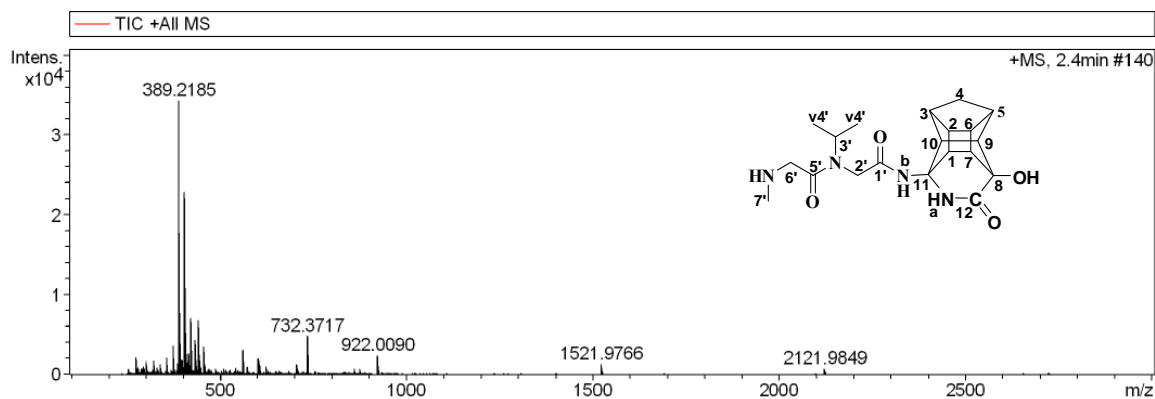
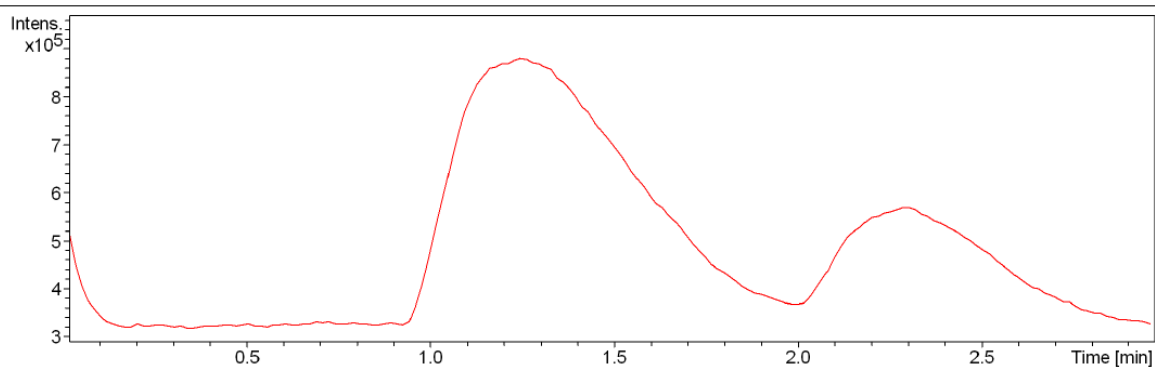
Analysis Name D:\Data\Sri\lactam-Val-Ala-NH2-PEPTOID1.d
Method tune_wide_expert.m
Sample Name lactam-Val-Ala-NH2-PEPTOID1
Comment

Acquisition Date 12/19/2009 6:16:39 PM

Operator BDAL@DE
Instrument micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



Display Report

Analysis Info

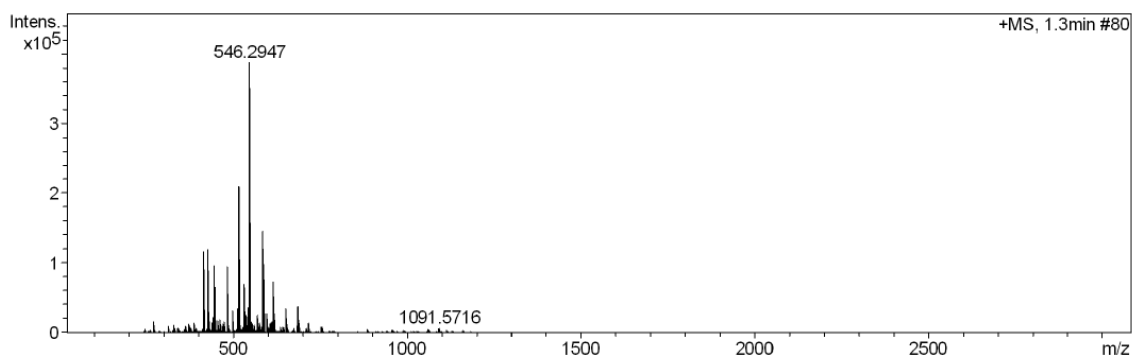
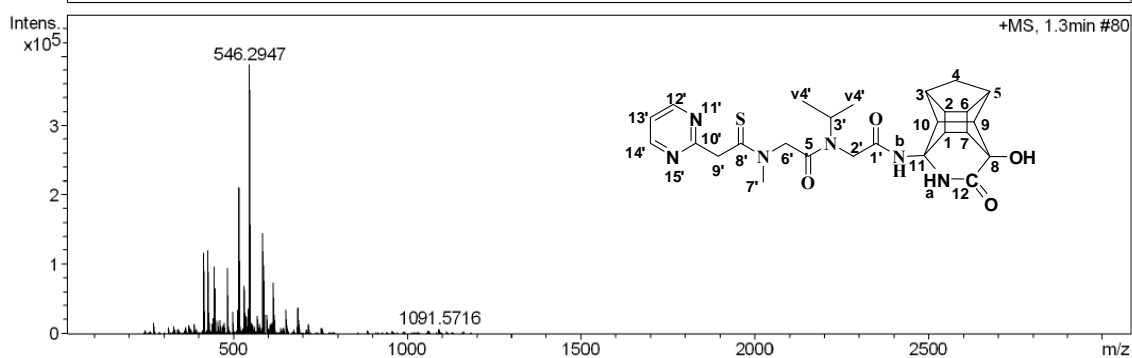
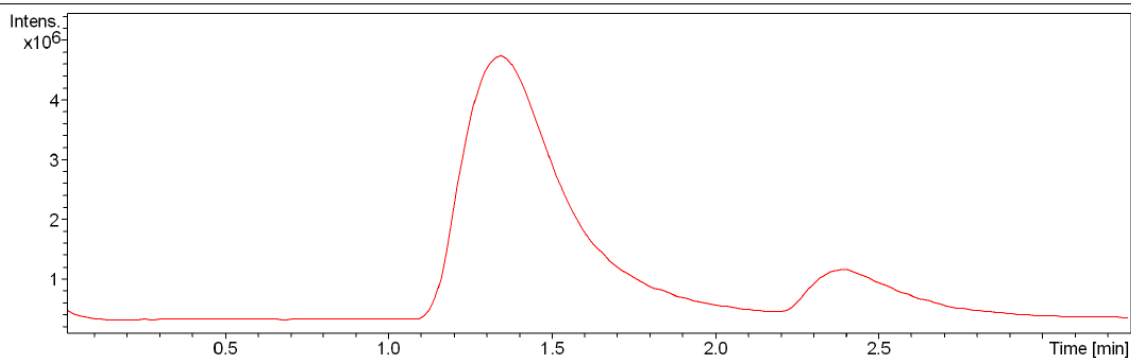
Analysis Name D:\Data\Sri\L-Val-Ala-2-Pyrimidyl-ethanethioicacid-peptoid2.d
 Method tune_wide_expert.m
 Sample Name L-Val-Ala-2-Pyrimidyl-ethanethioicacid-peptoid2
 Comment

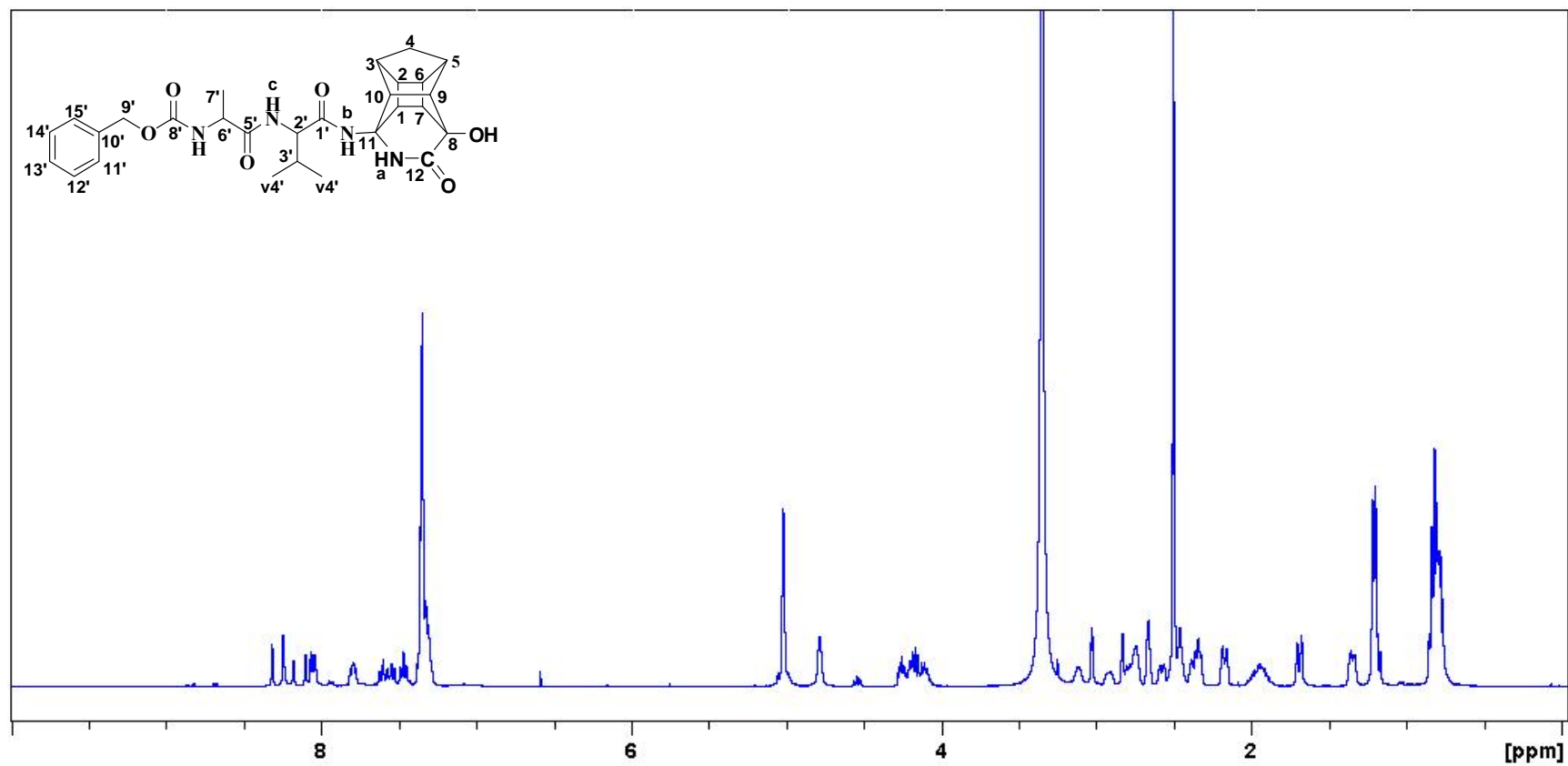
Acquisition Date 12/25/2009 11:59:59 AM

Operator BDAL@DE
 Instrument micrOTOF-Q 10139

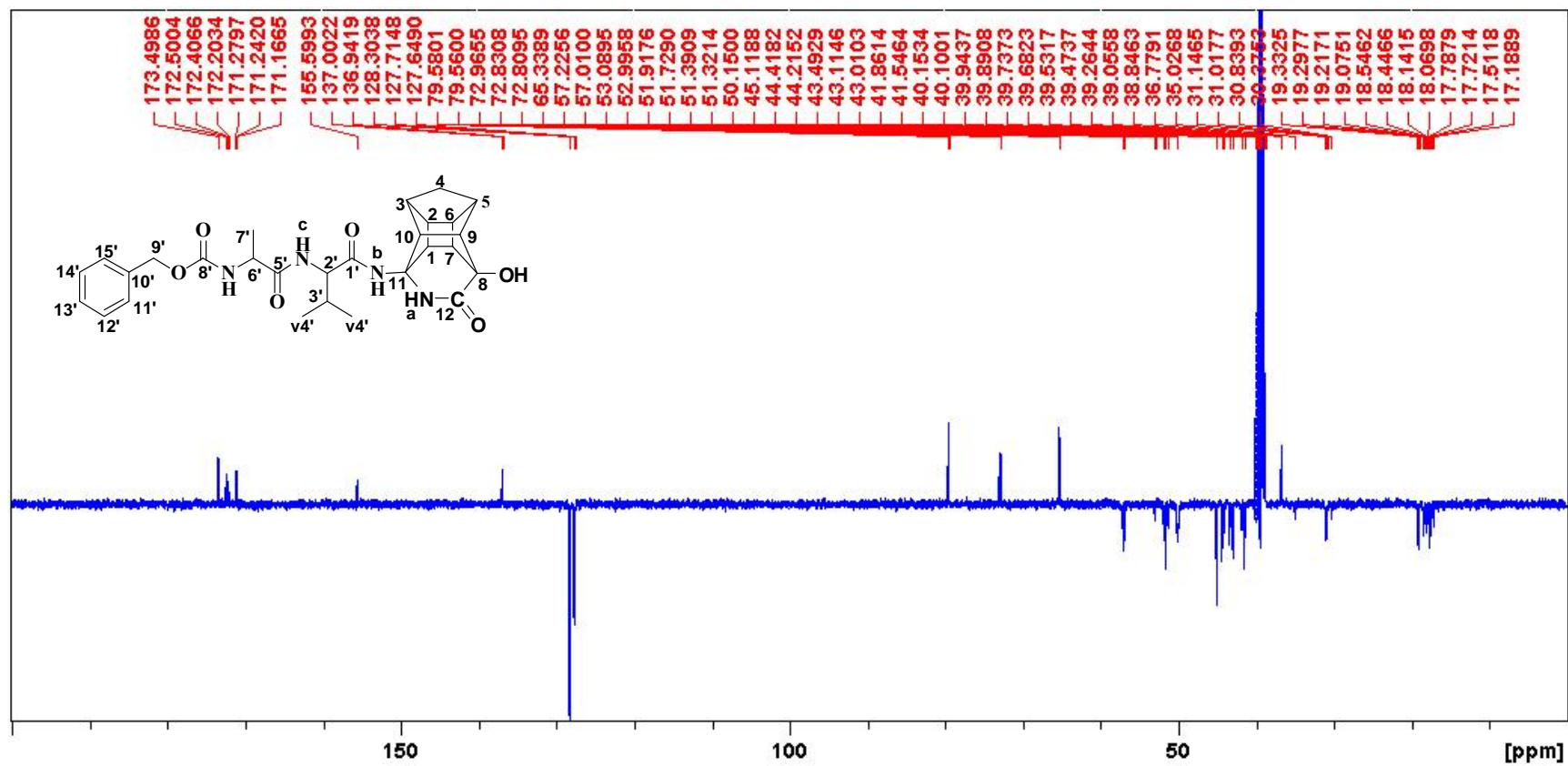
Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source

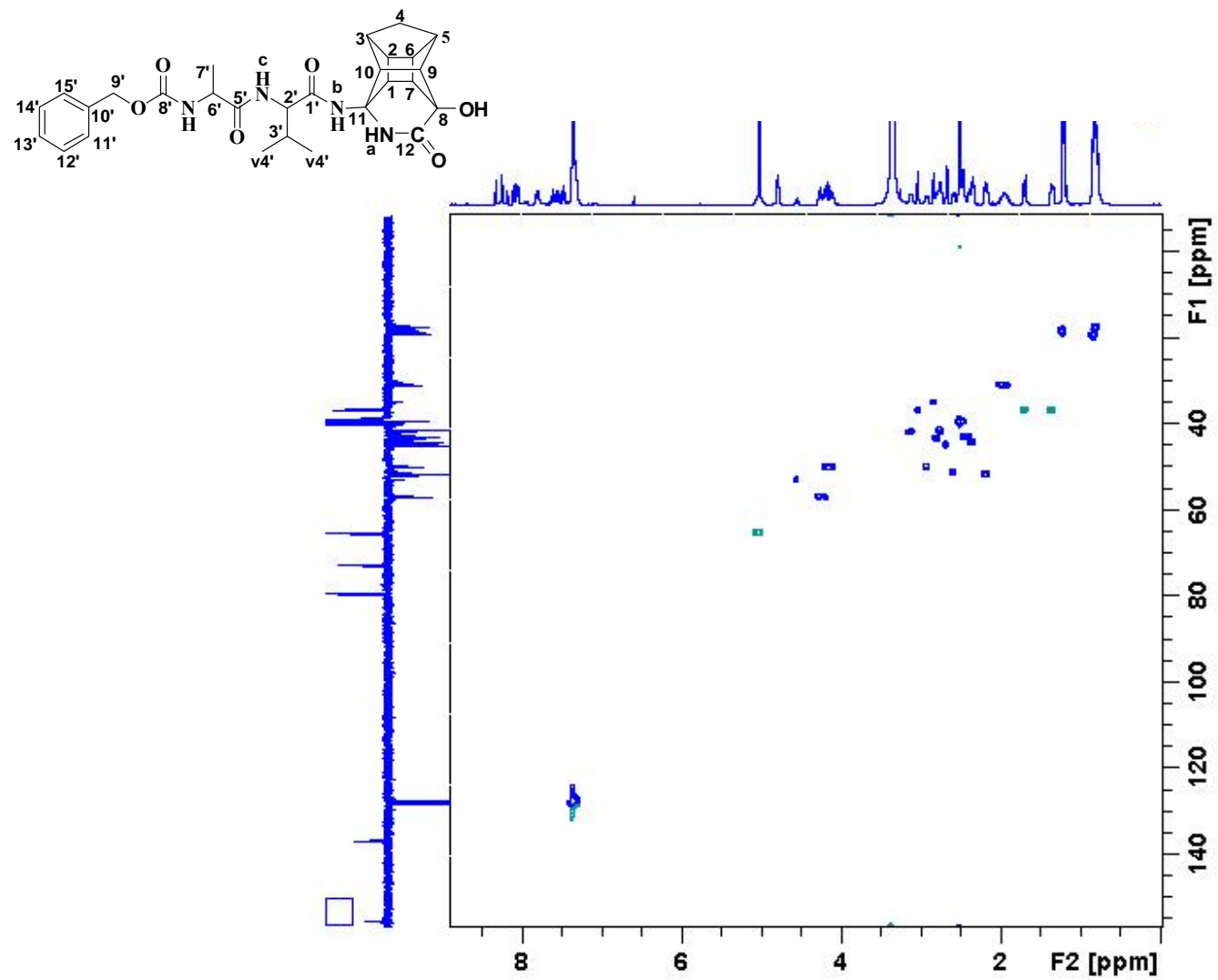




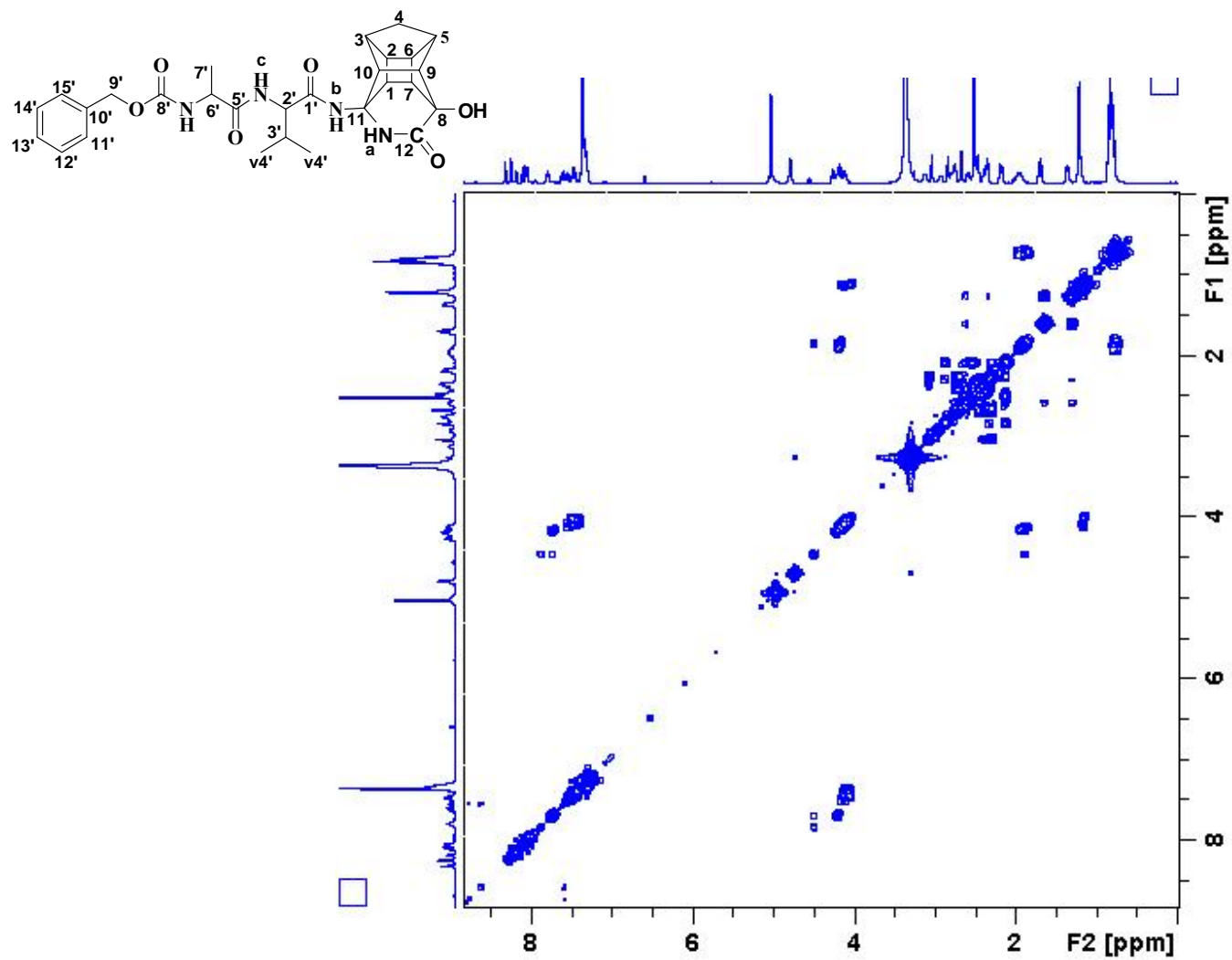
^1H NMR spectrum of PCU-VAZ peptide, 16a



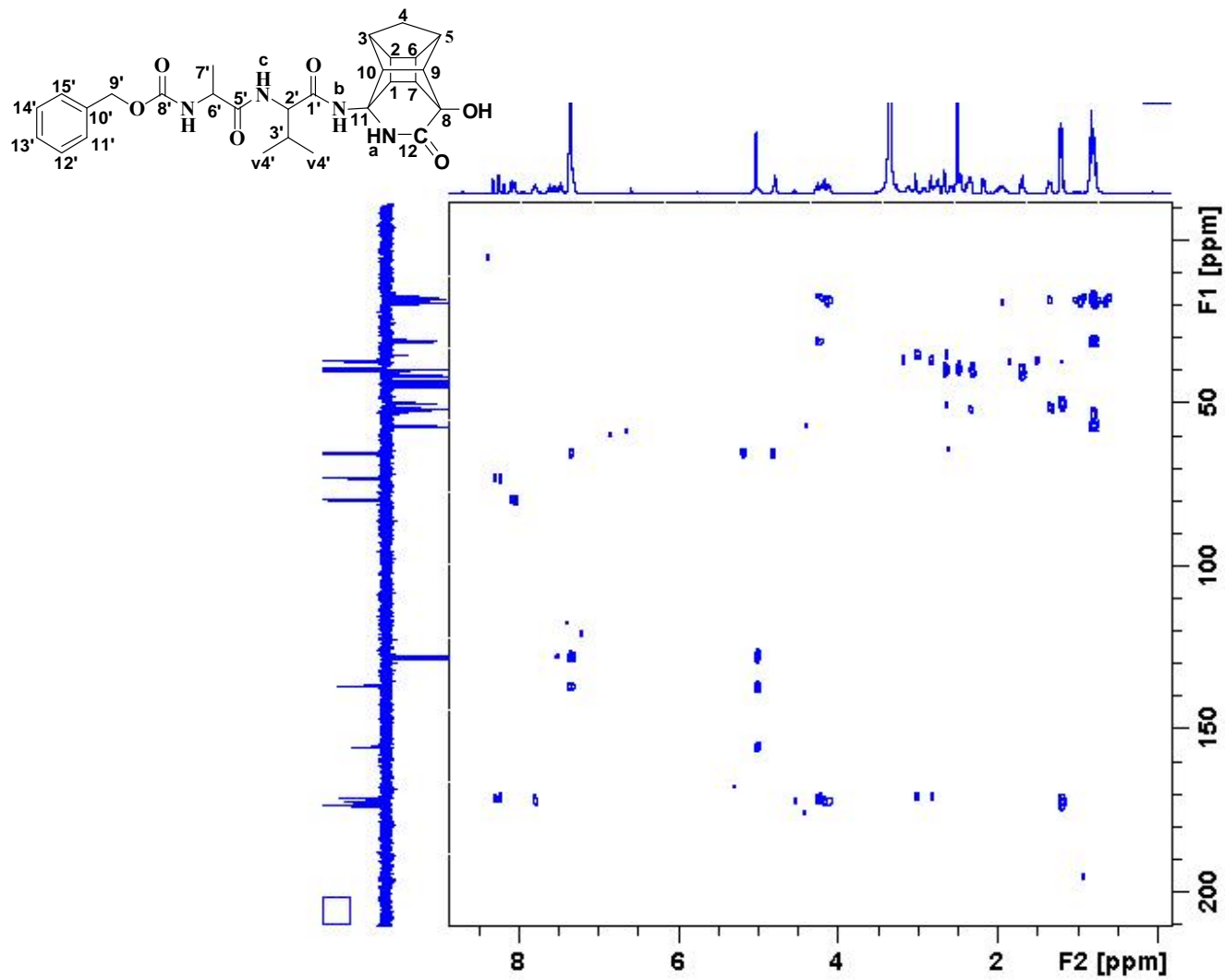
^{13}C NMR spectrum of PCU-VAZ peptide, 16a



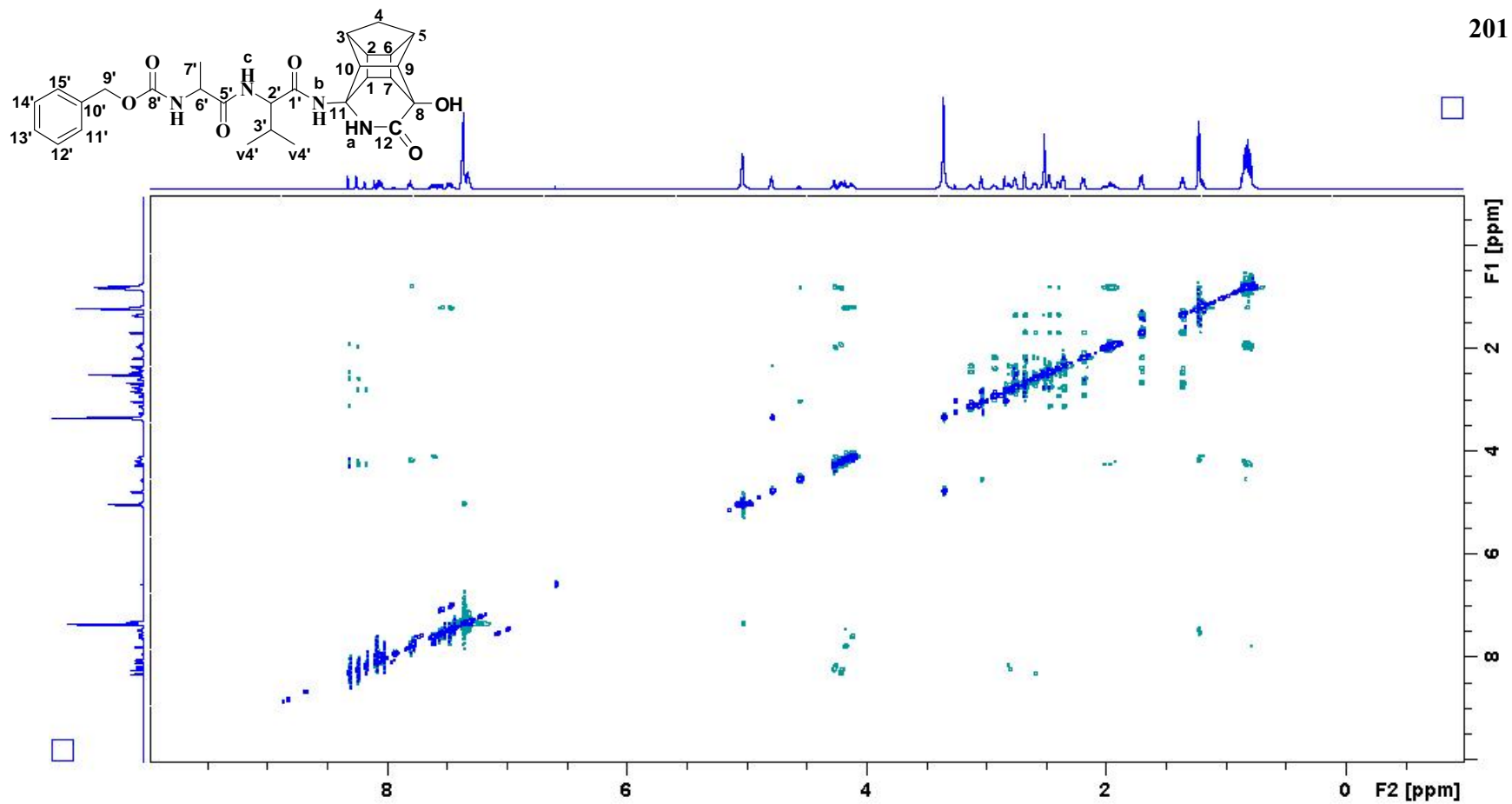
HSQC spectrum of PCU-VAZ peptide, 16a



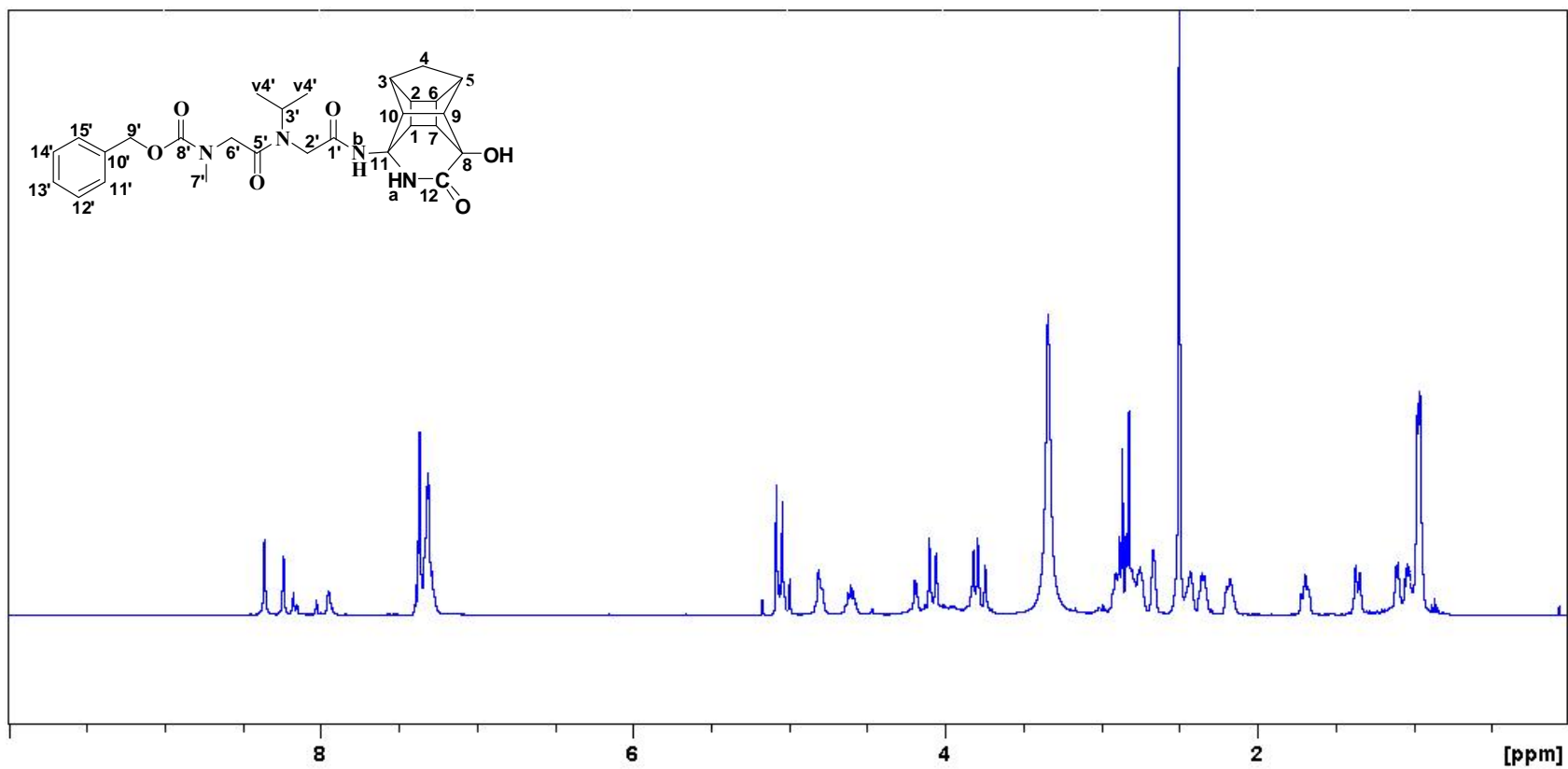
COSY spectrum of PCU-VAZ peptide, 16a



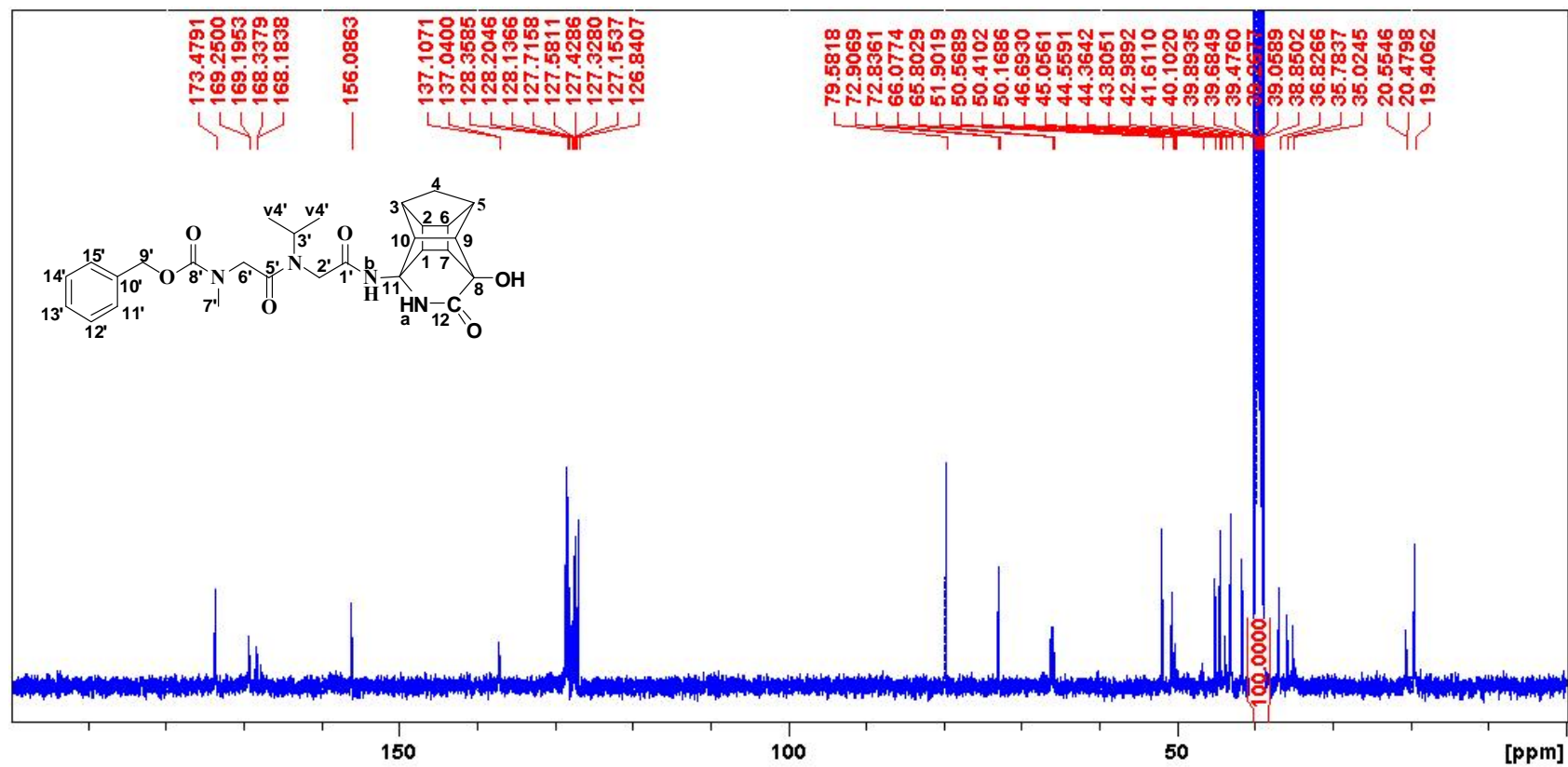
HMBC spectrum of PCU-VAZ peptide, 16a



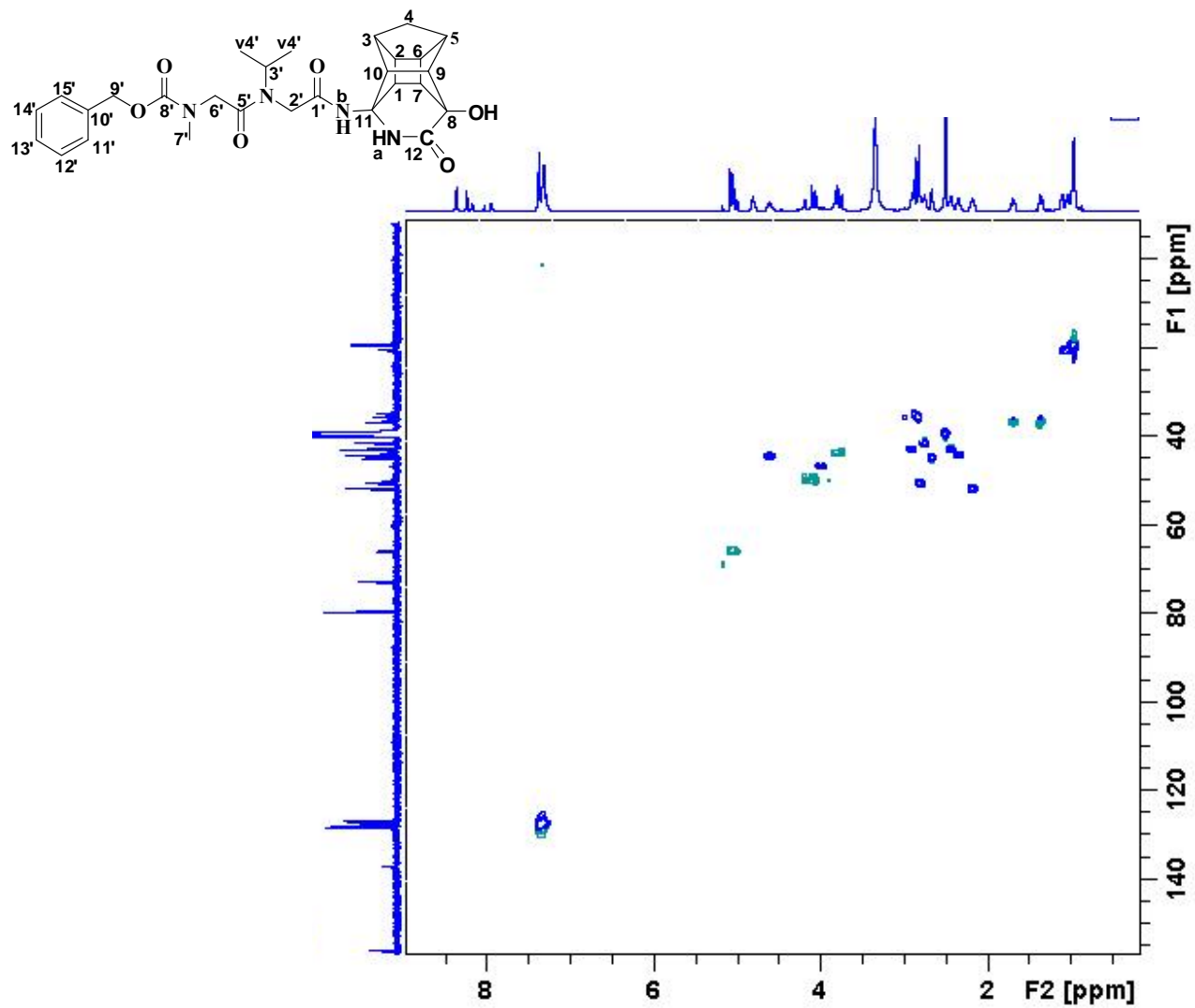
ROESY spectrum of PCU-VAZ peptide, 16a



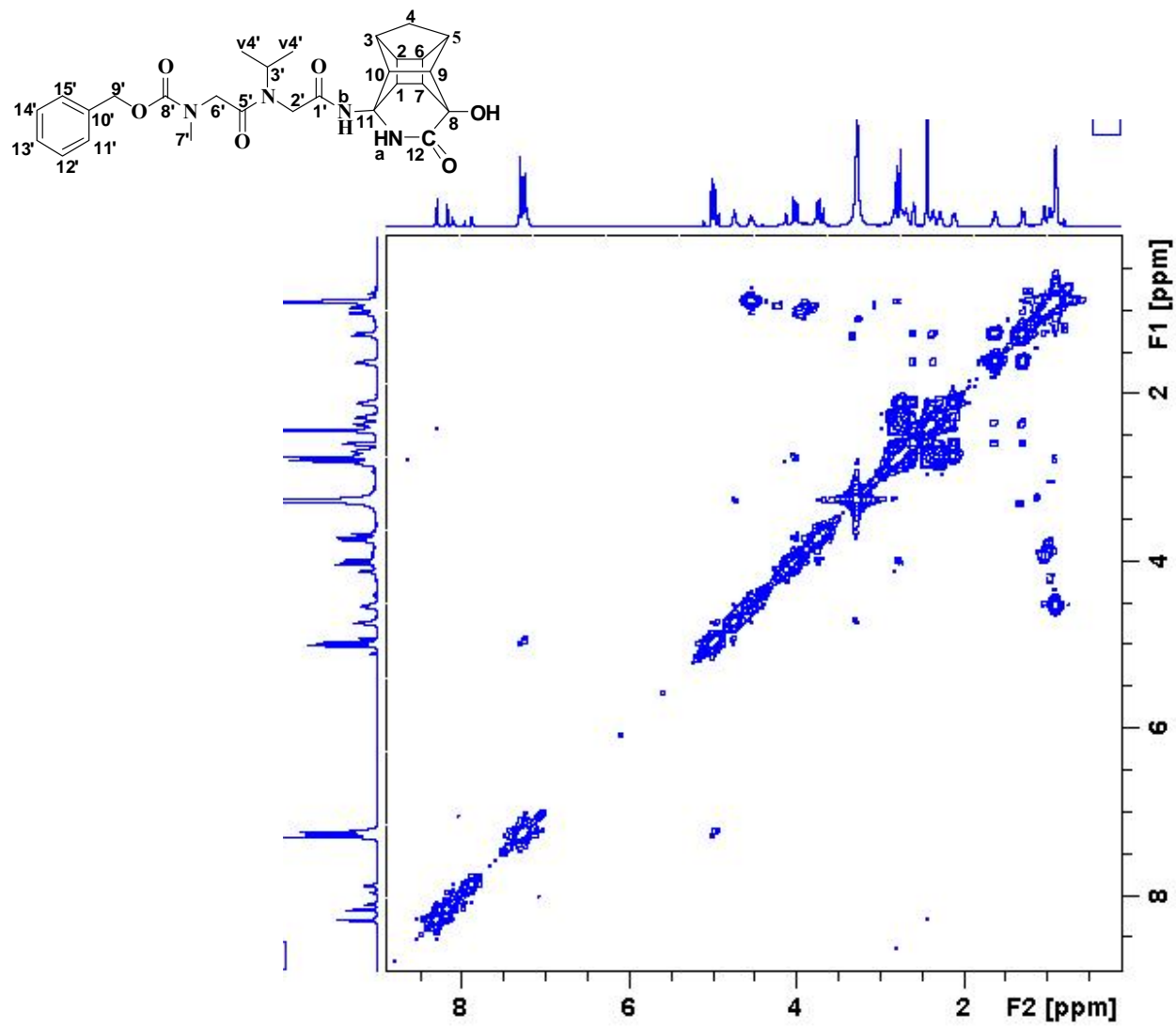
^1H NMR spectrum of PCU-VAZ peptoid, 18a



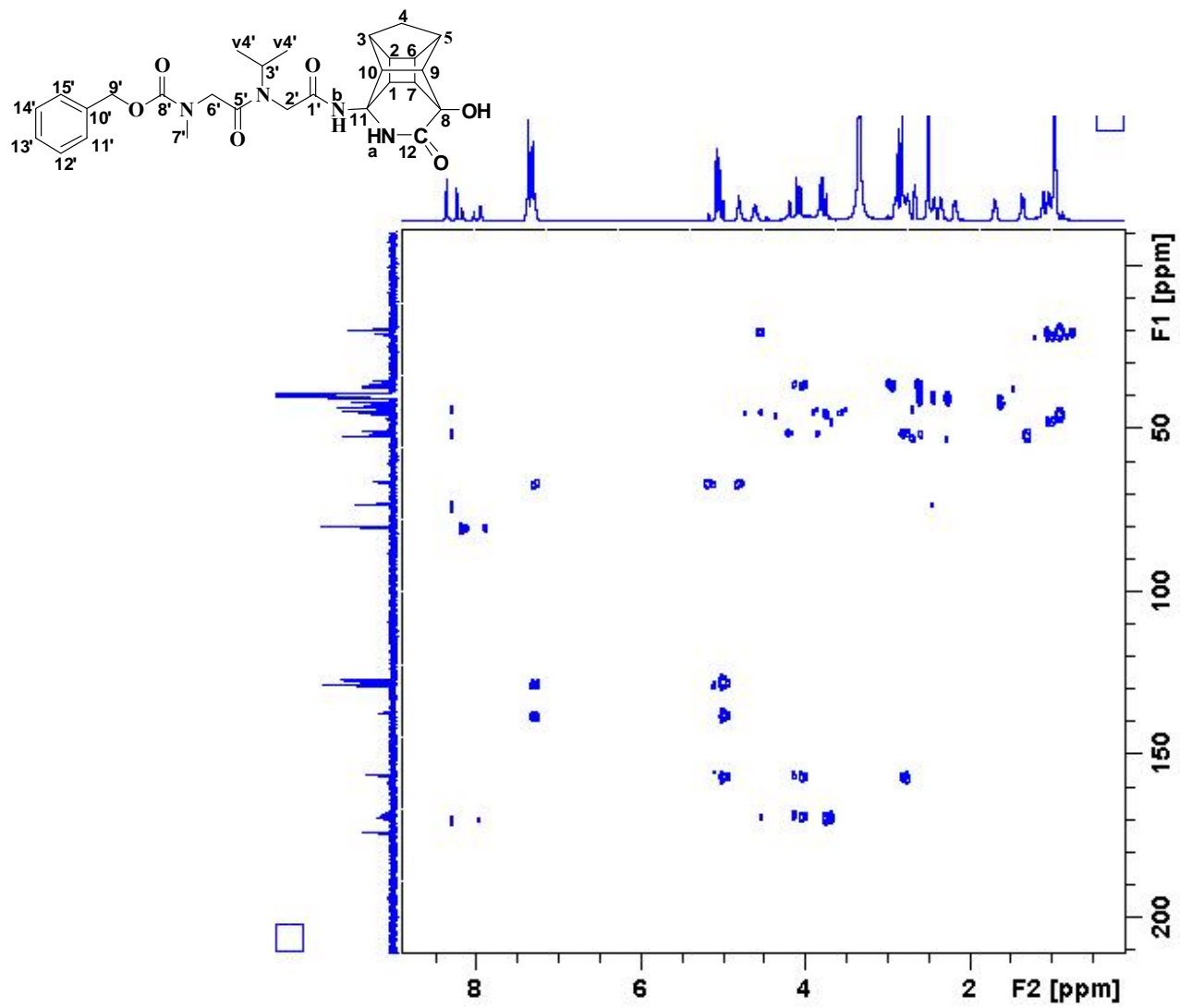
^{13}C NMR spectrum of PCU-VAZ peptoid, 18a



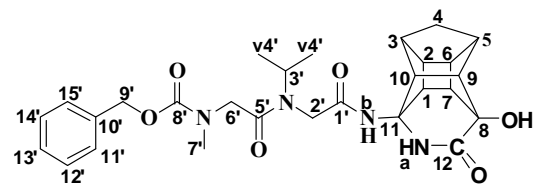
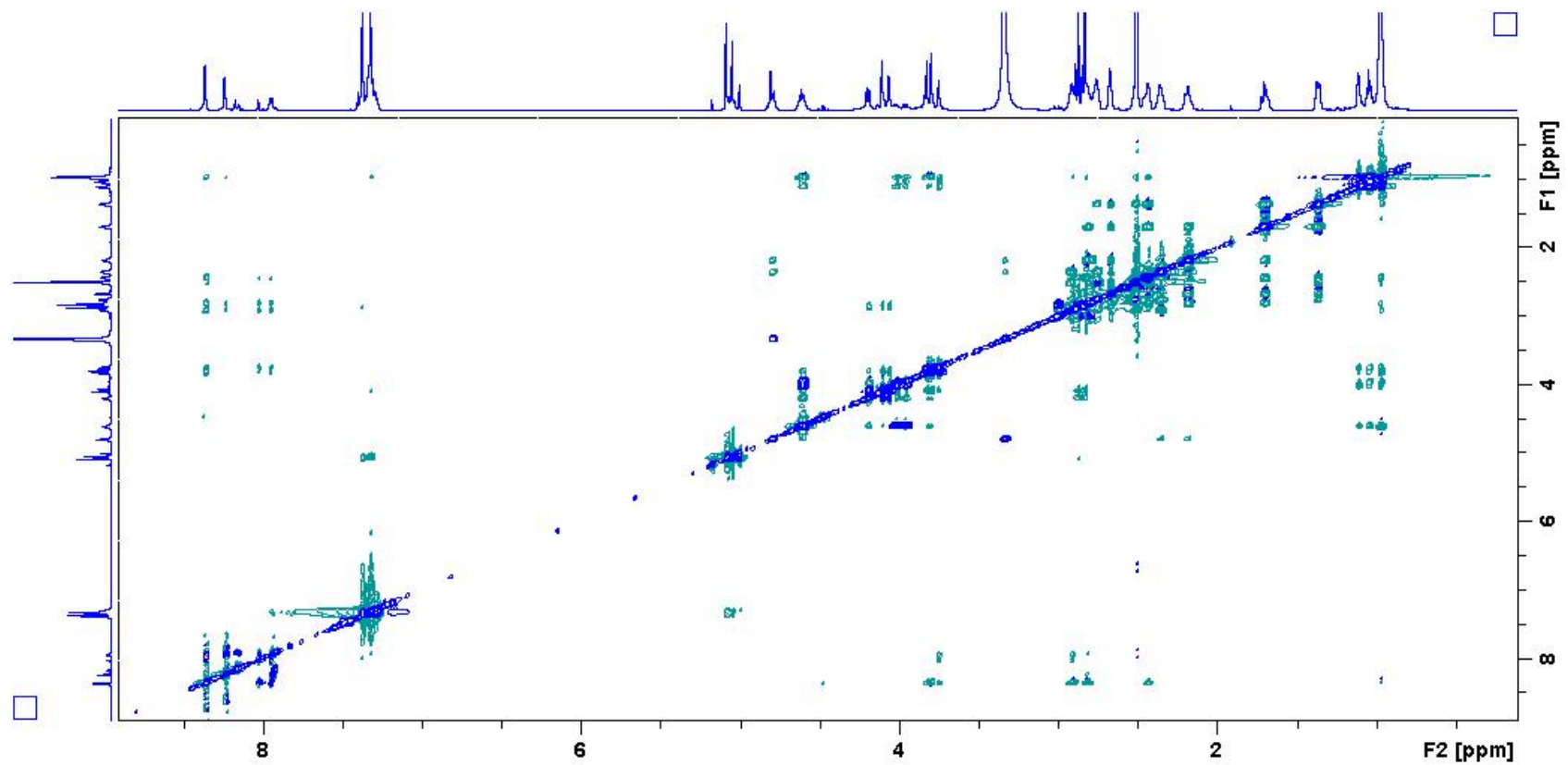
HSQC spectrum of PCU-VAZ peptoid, 18a



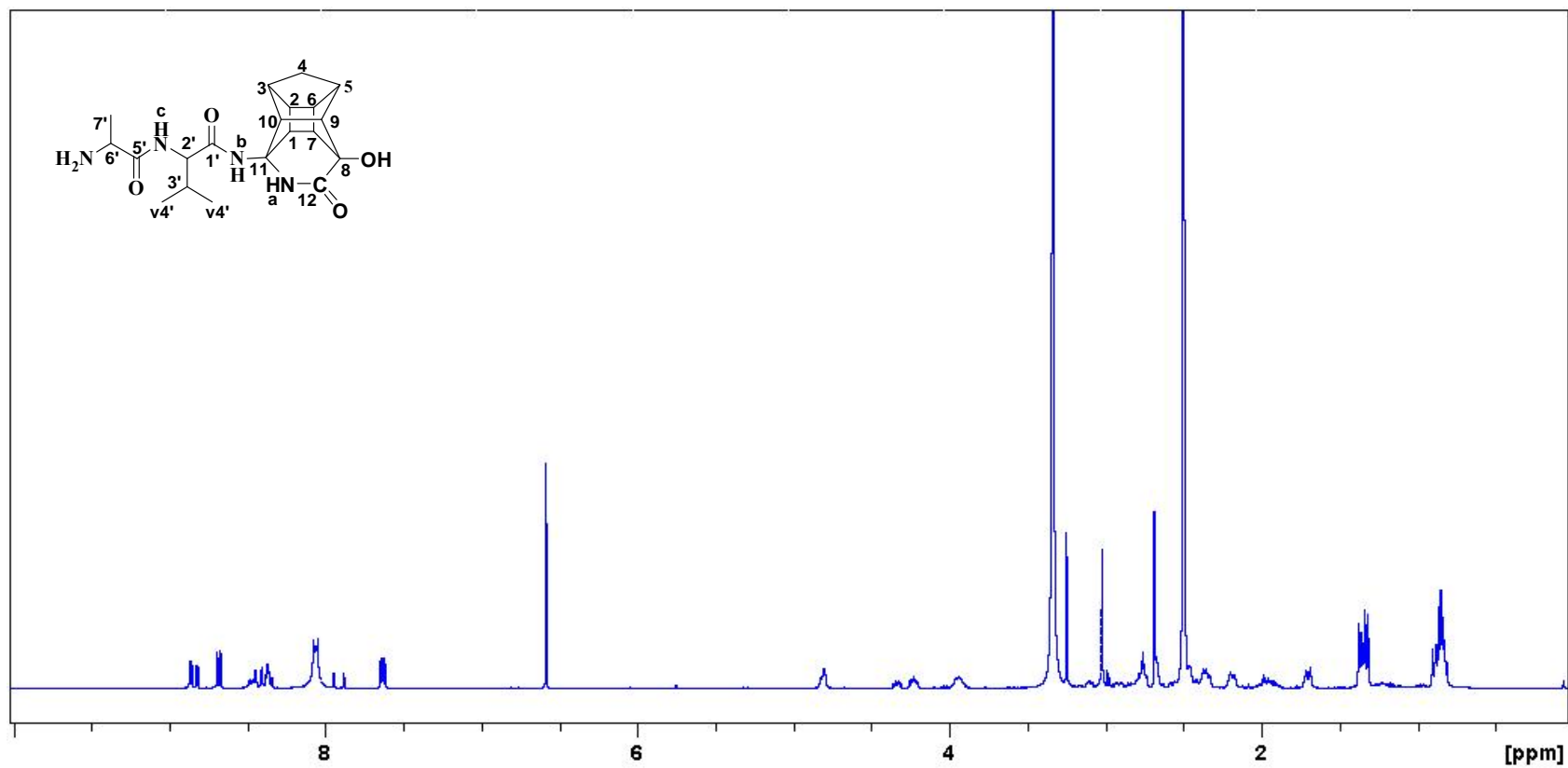
COSY spectrum of PCU-VAZ peptoid, 18a



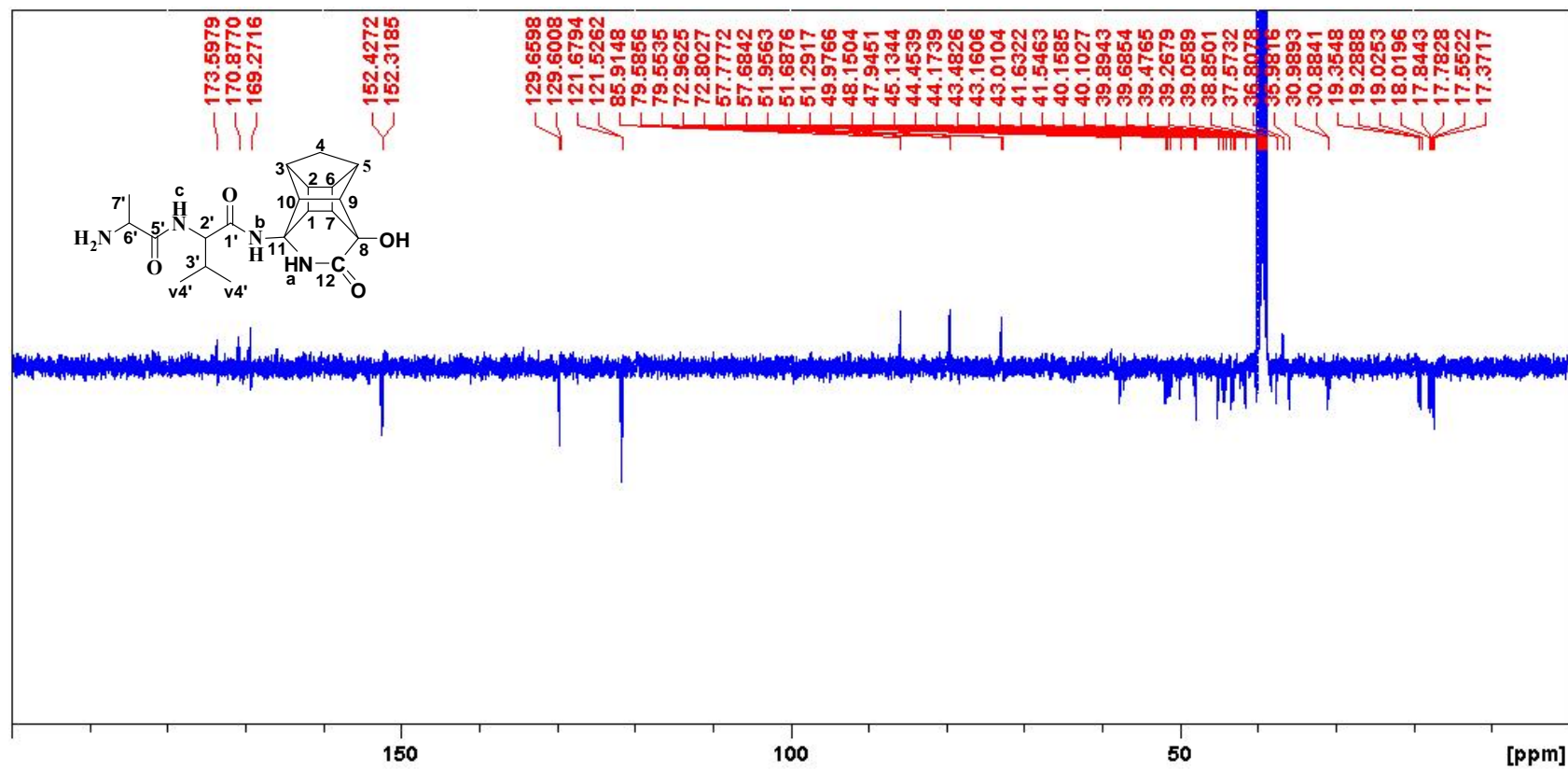
HMBC spectrum of PCU-VAZ peptoid, 18a



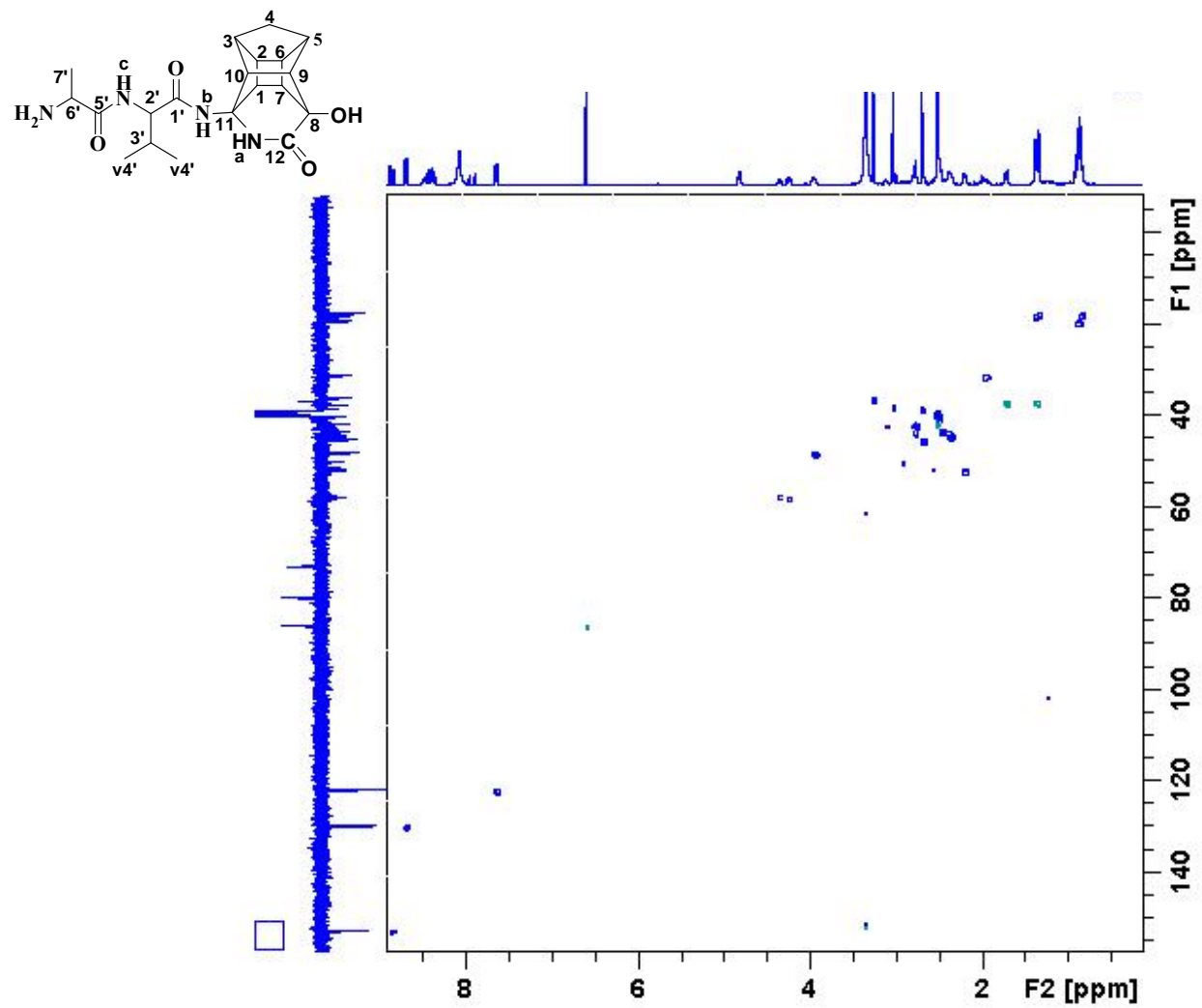
ROESY spectrum of PCU-VAZ peptoid, 18a

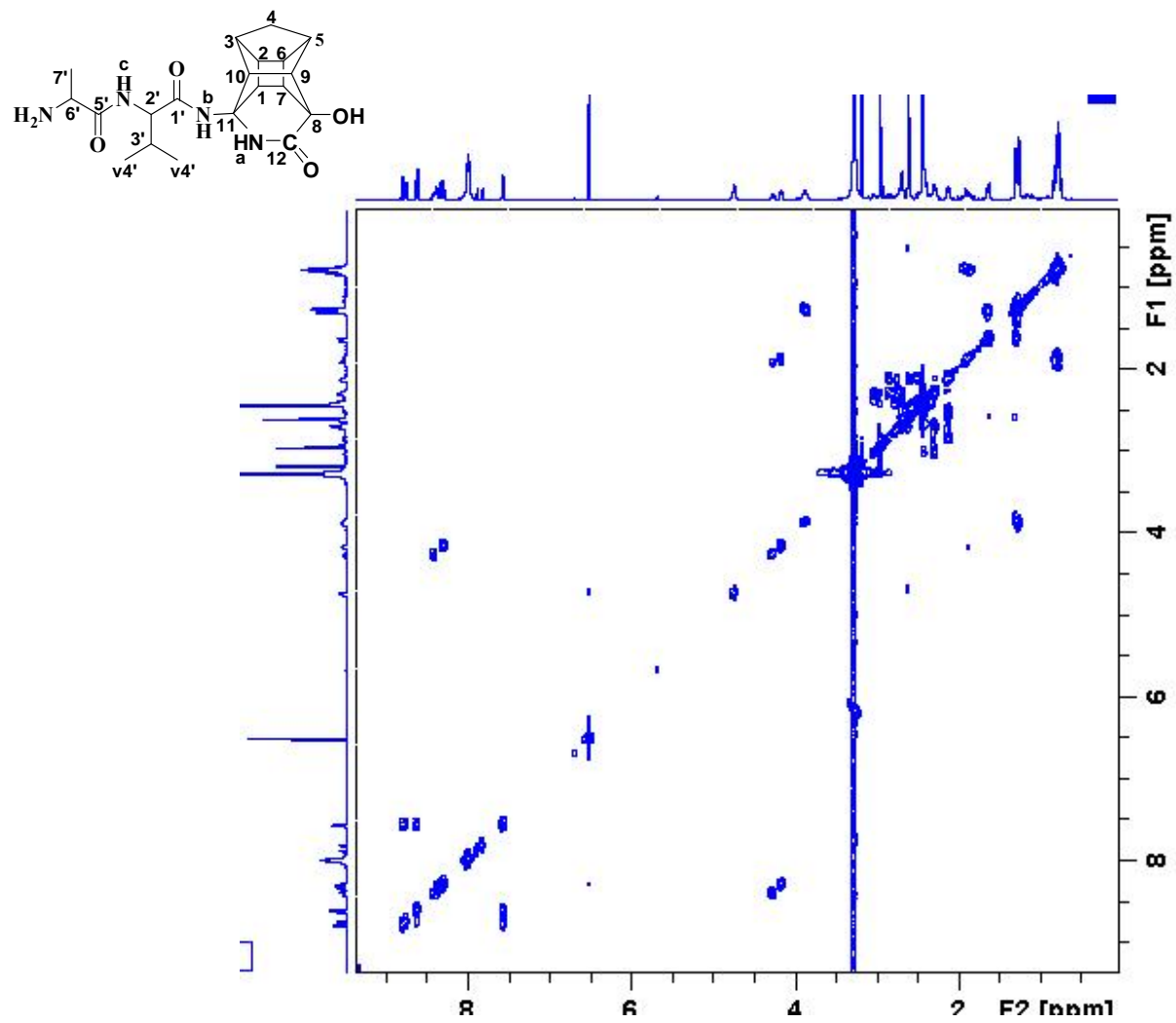


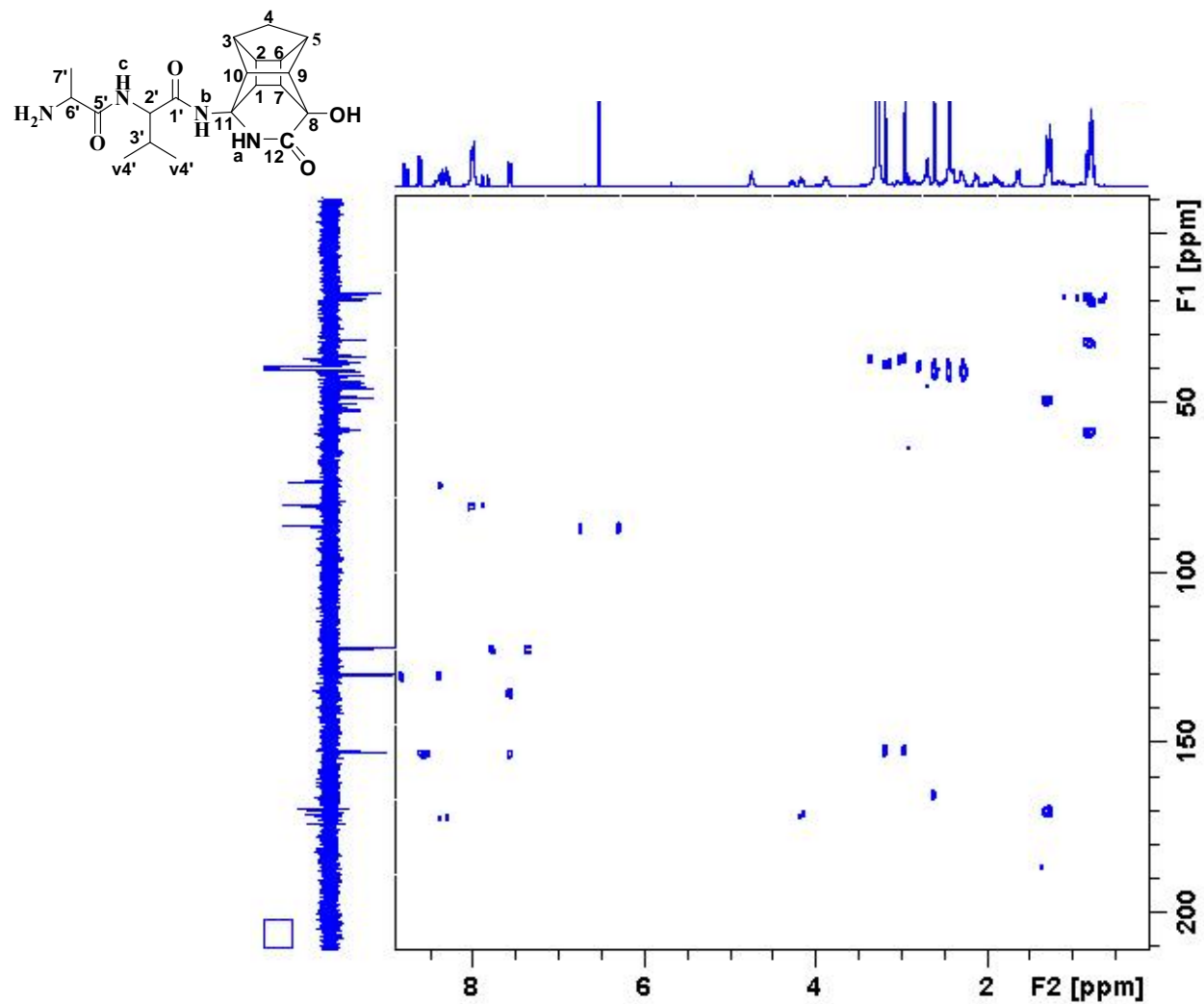
^1H NMR spectrum of PCU-VANH₂ peptide, 16b

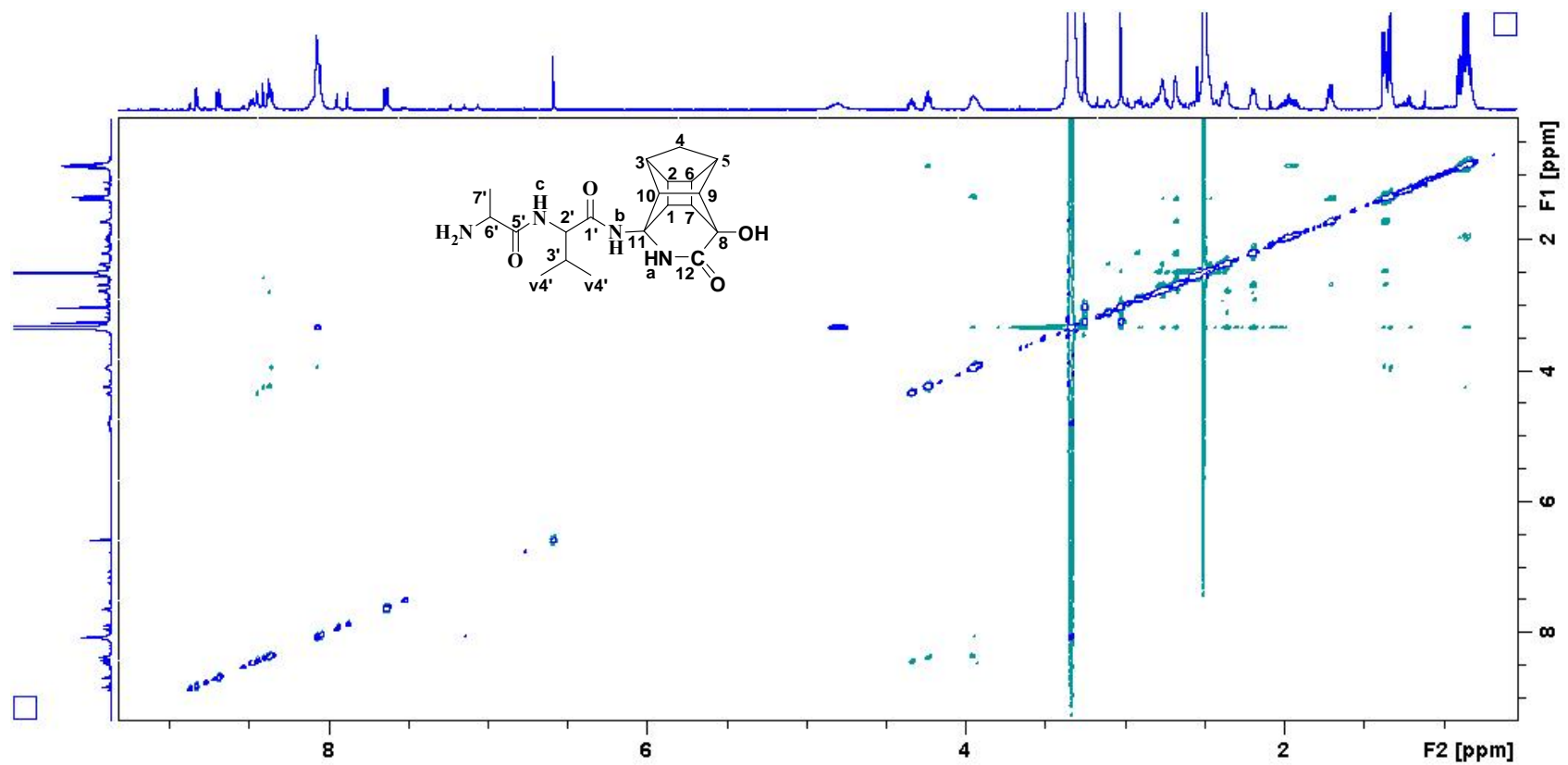


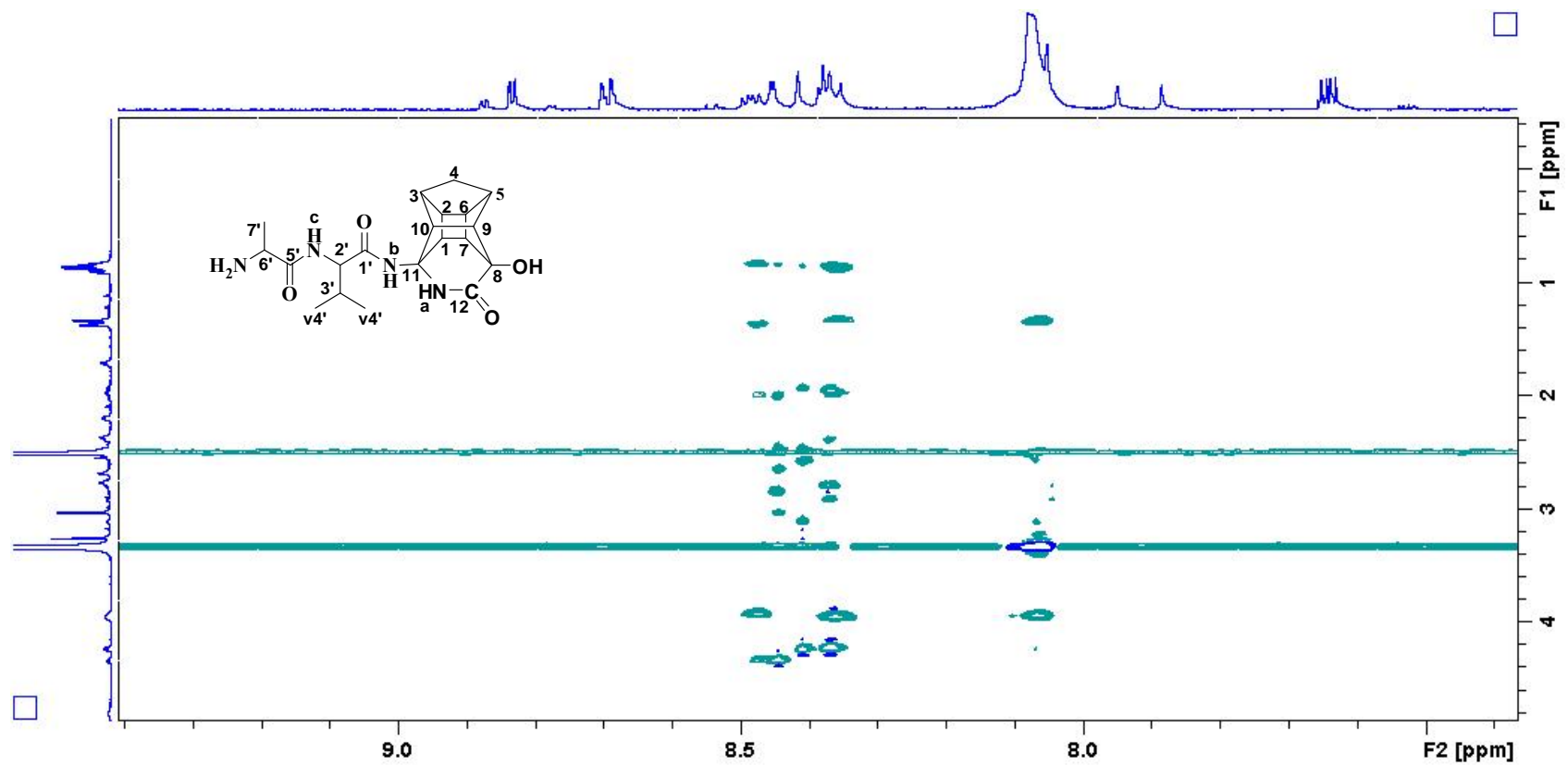
^1H NMR spectrum of PCU-VANH₂ peptide, 16b

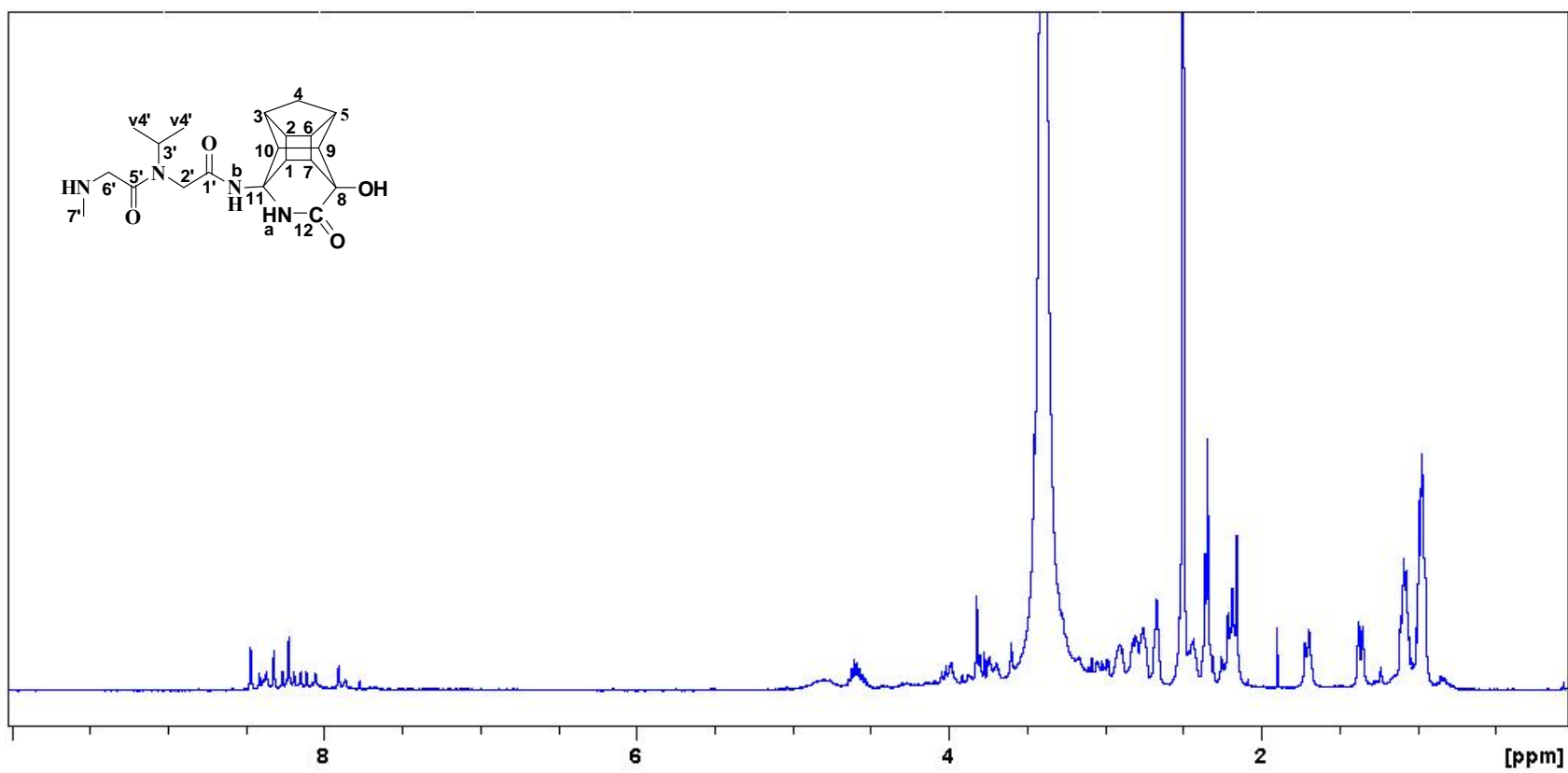
HSQC spectrum of PCU-VANH₂ peptide, 16b

COSY spectrum of PCU-VANH₂ peptide, 16b

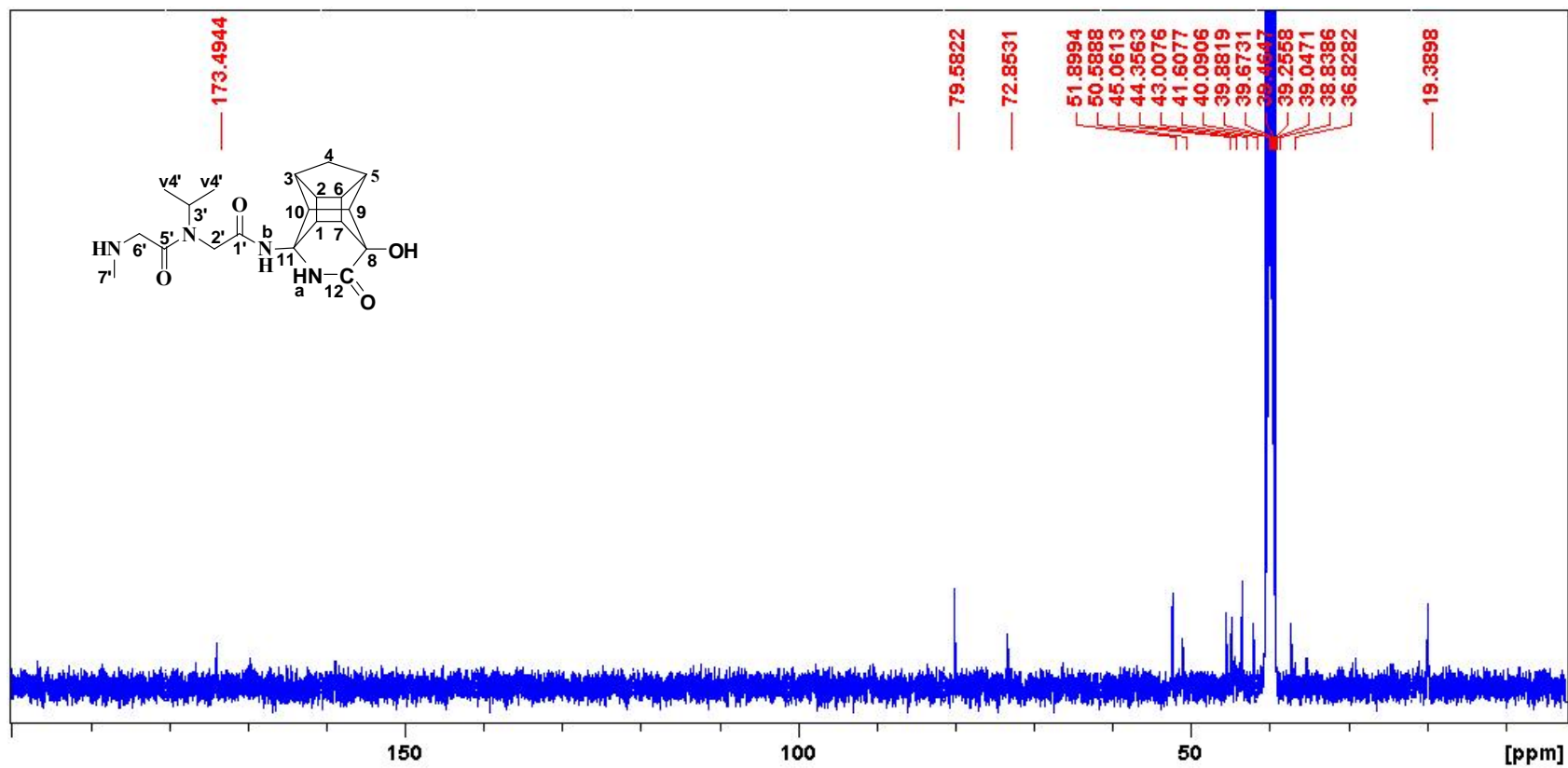
HMBC spectrum of PCU-VANH₂ peptide, 16b

ROESY spectrum of PCU-VANH₂ peptide, 16b

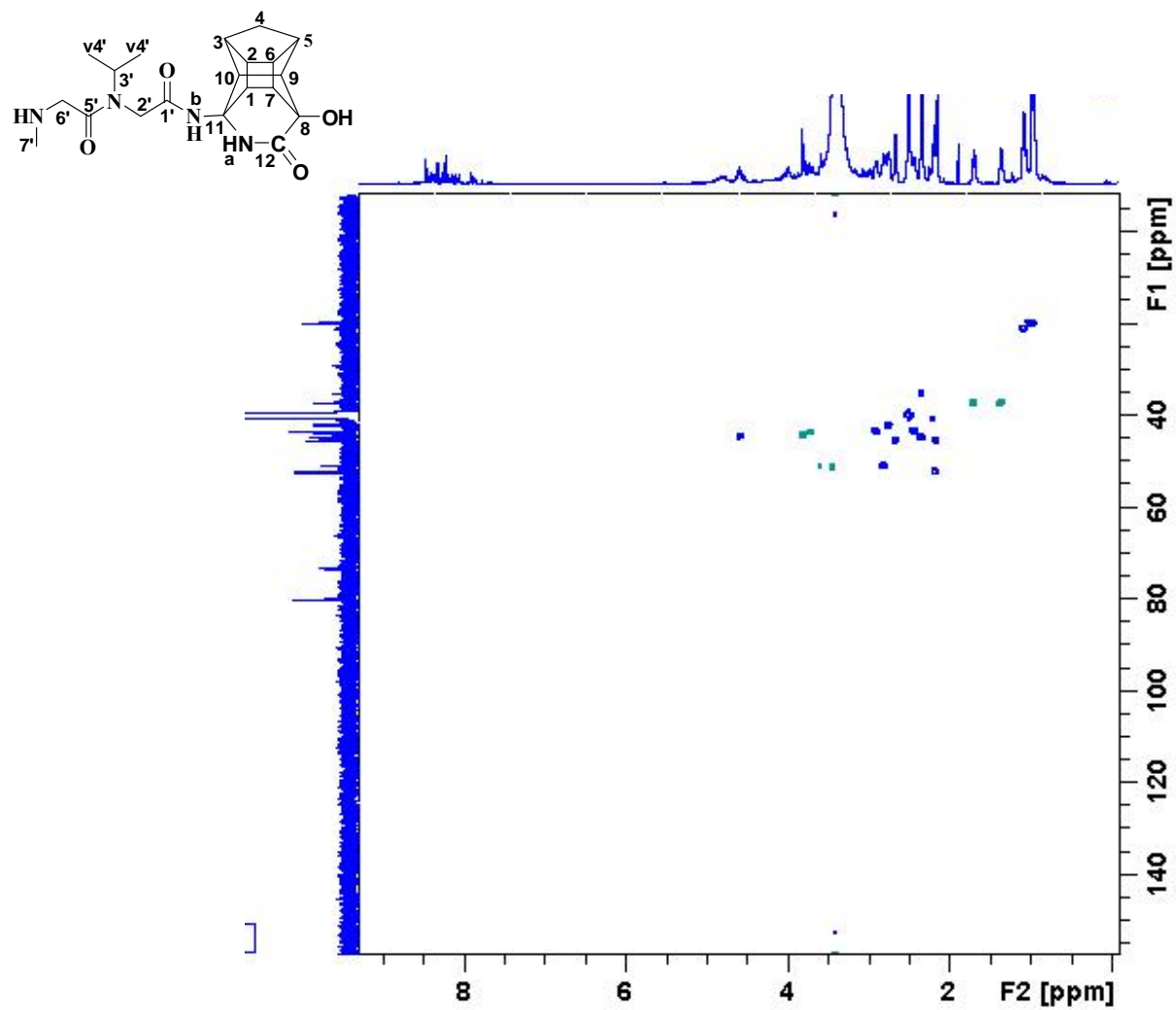
ROESY spectrum of PCU-VANH₂ peptide expanded, 16b

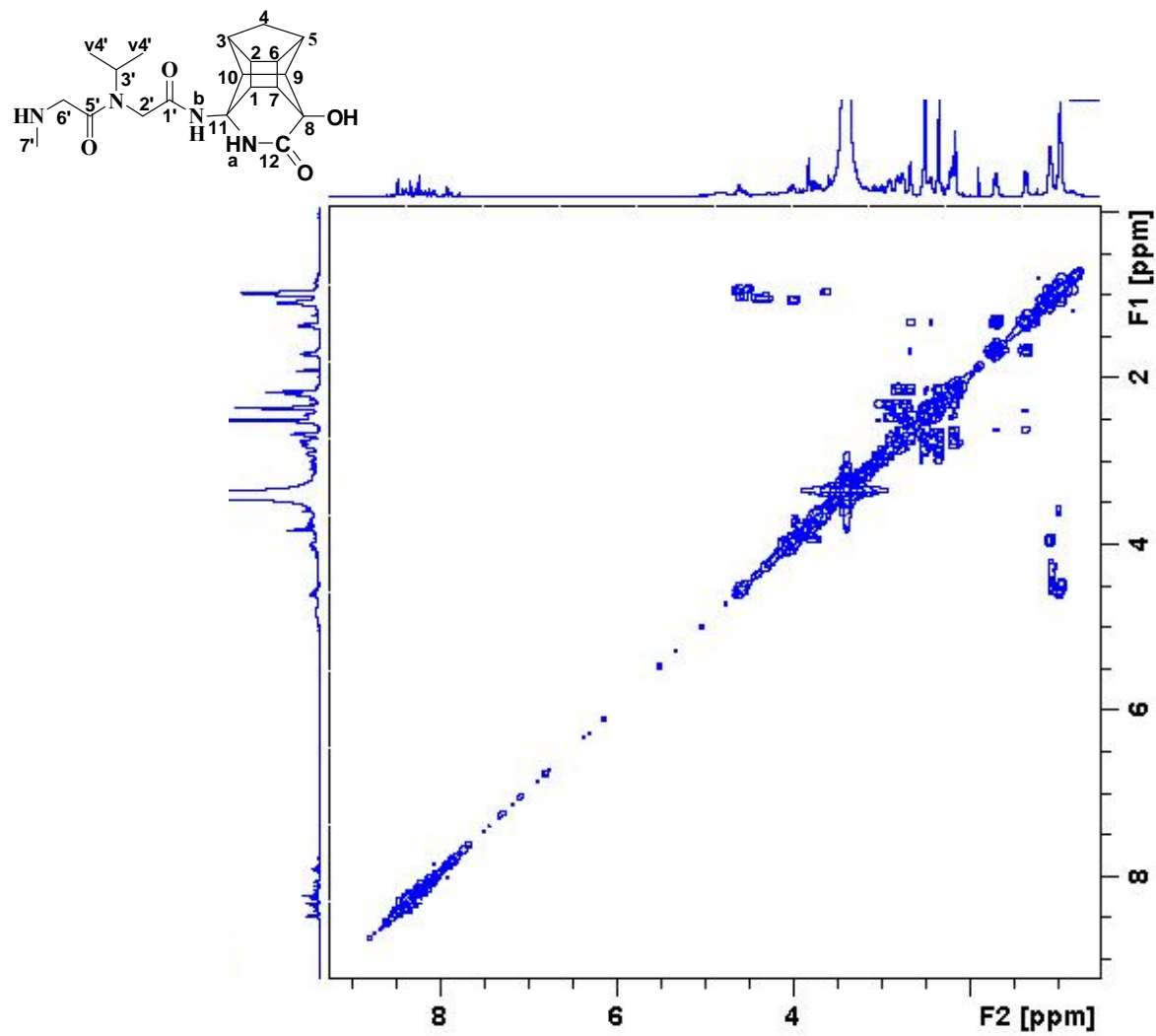


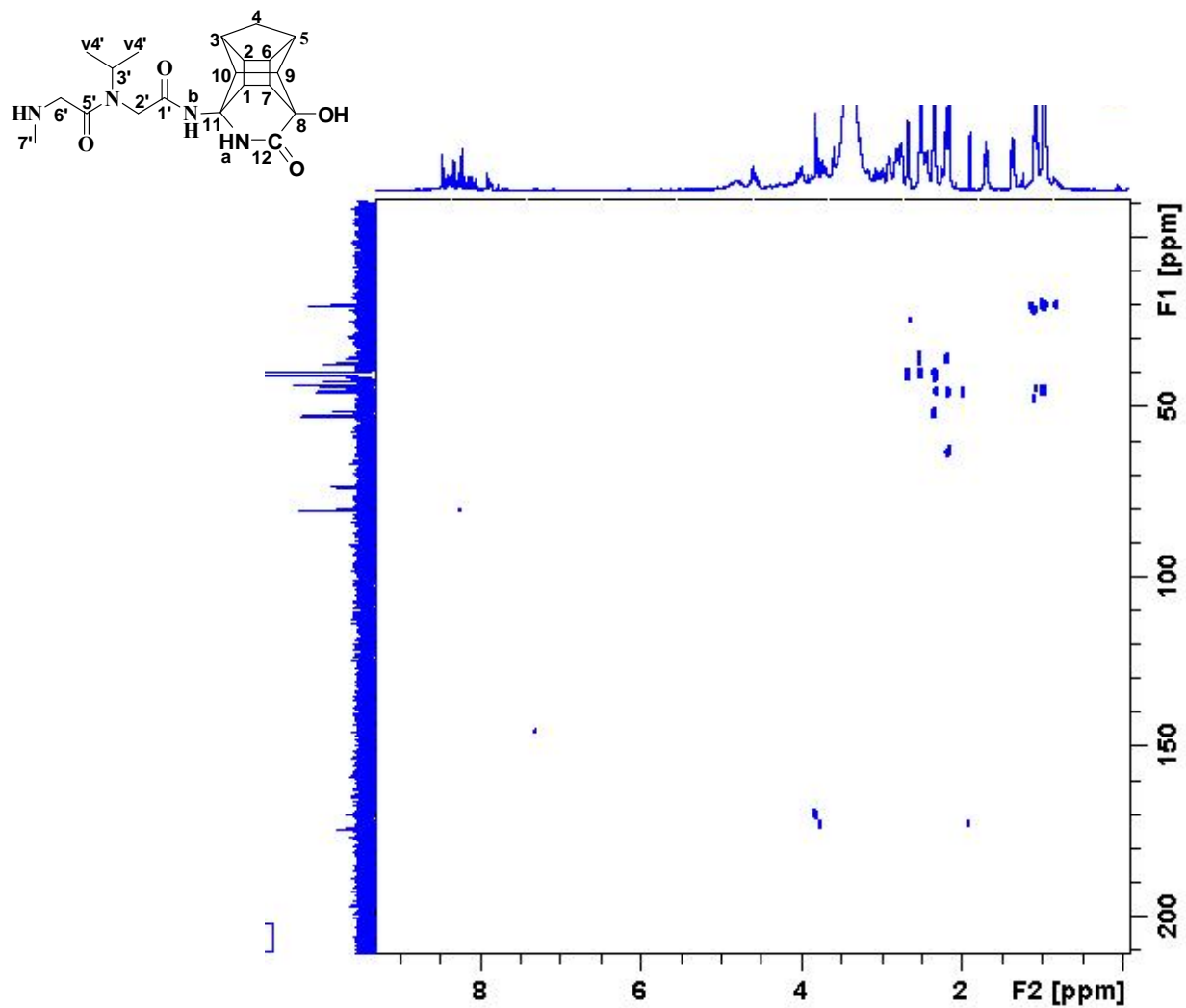
^1H NMR spectrum of PCU-VANH₂ peptoid, 18b

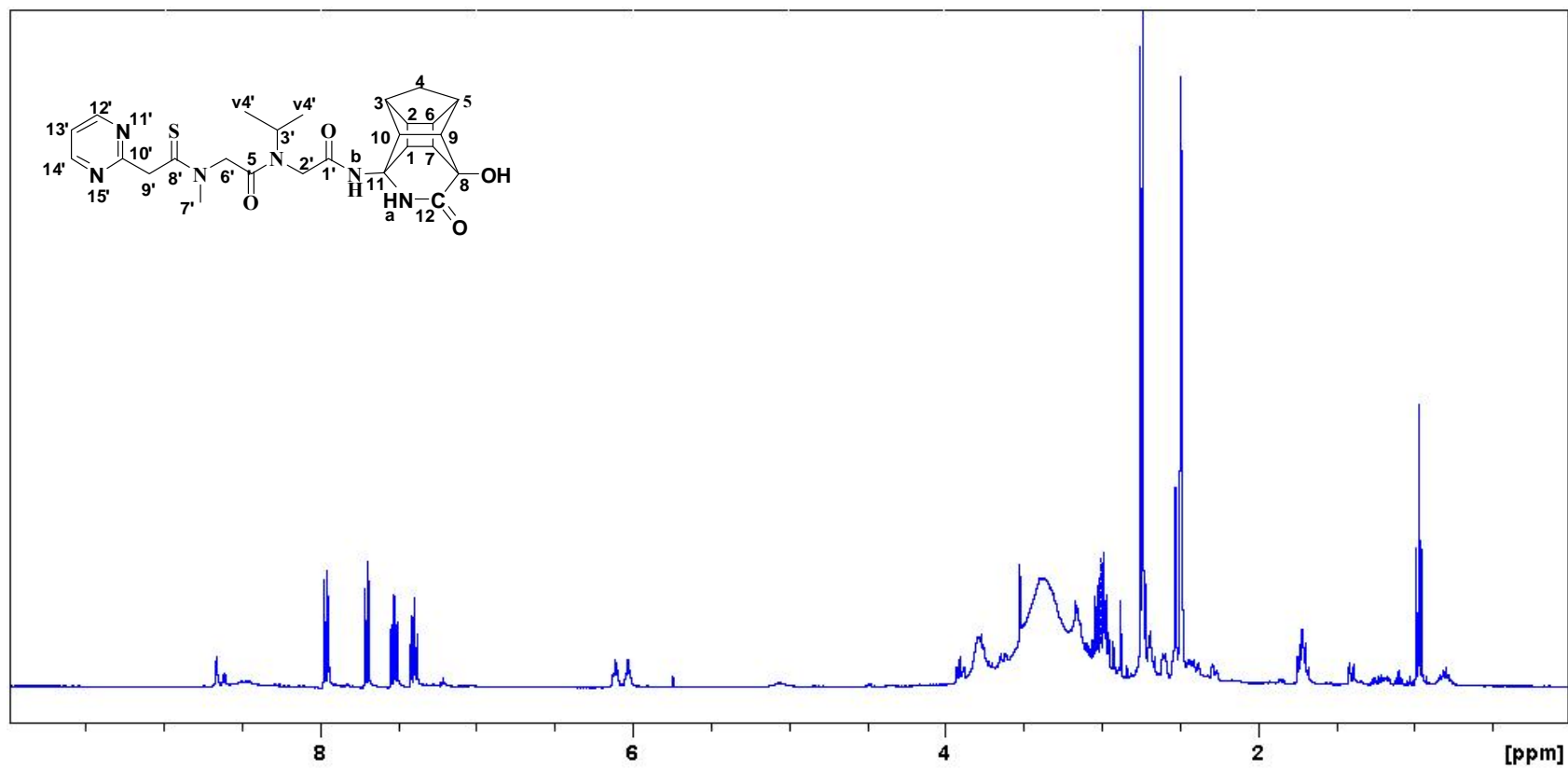


^{13}C NMR spectrum of PCU-VANH₂ peptoid, 18b

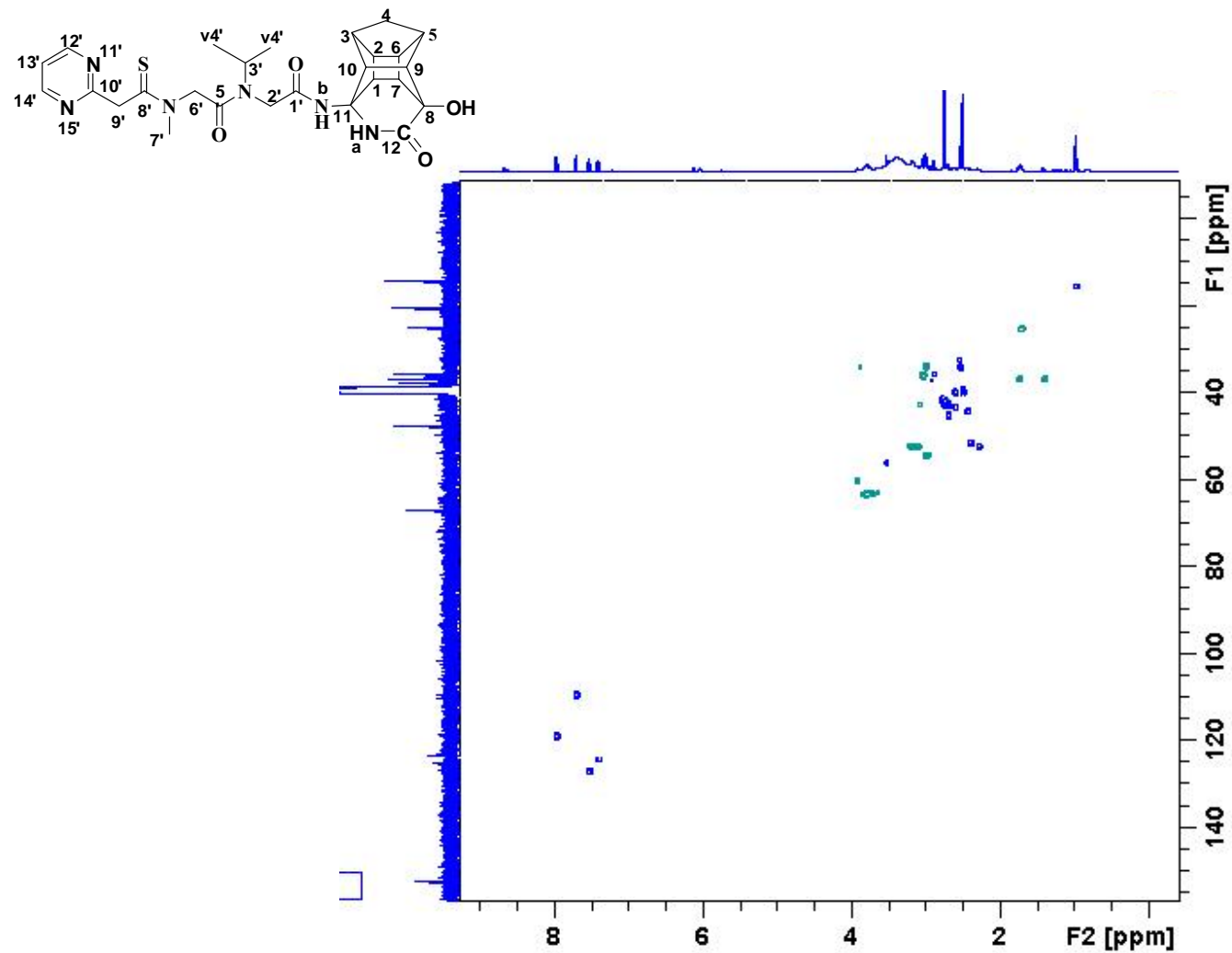
HSQC spectrum of PCU-VANH₂ peptoid, 18b

COSY spectrum of PCU-VANH₂ peptoid, 18b

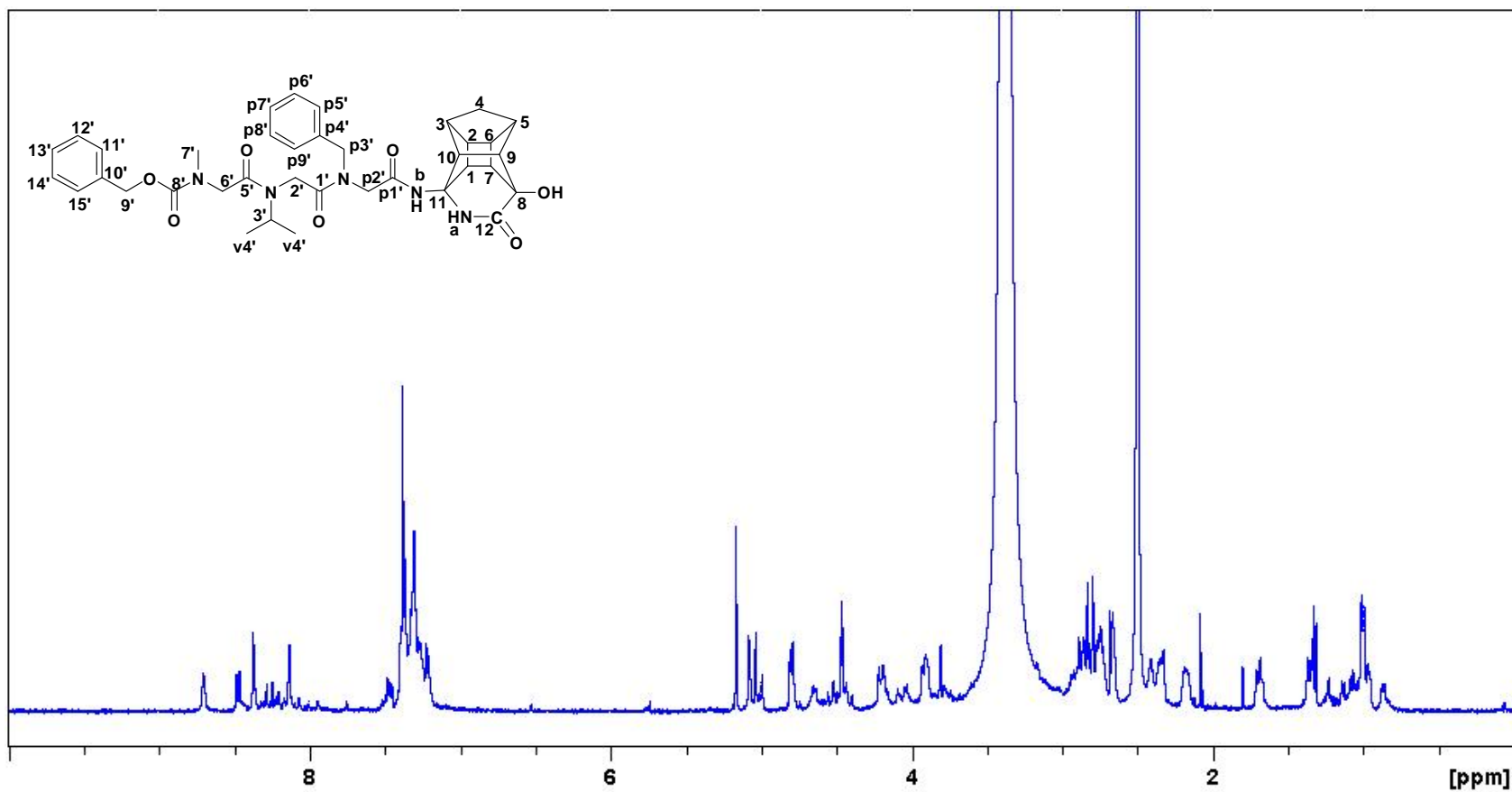
HMBC spectrum of PCU-VANH₂ peptoid, 18b



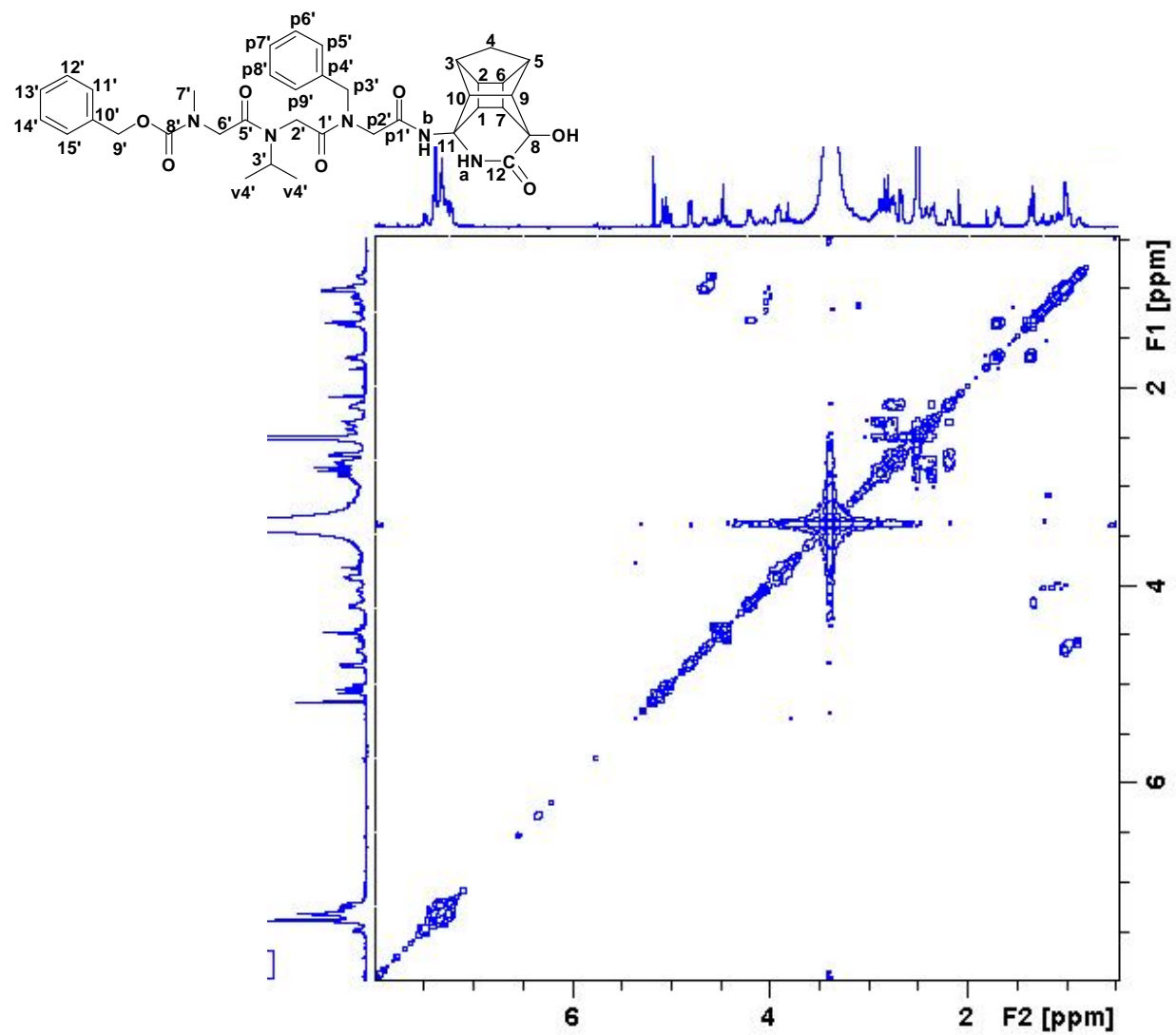
^1H NMR spectrum of PCU-VAT peptoid, 18c



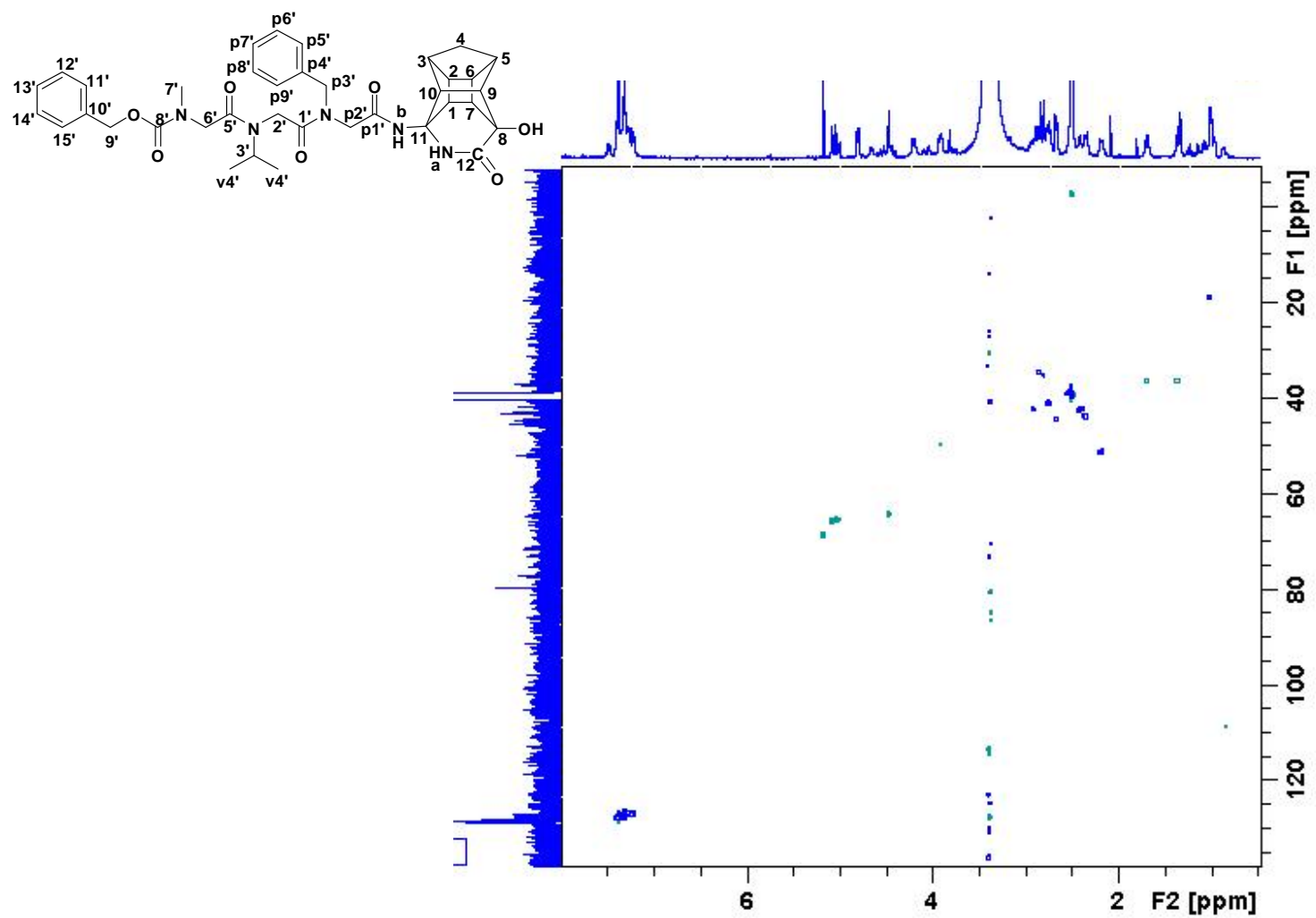
HSQC spectrum of PCU-VAT peptoid, 18c



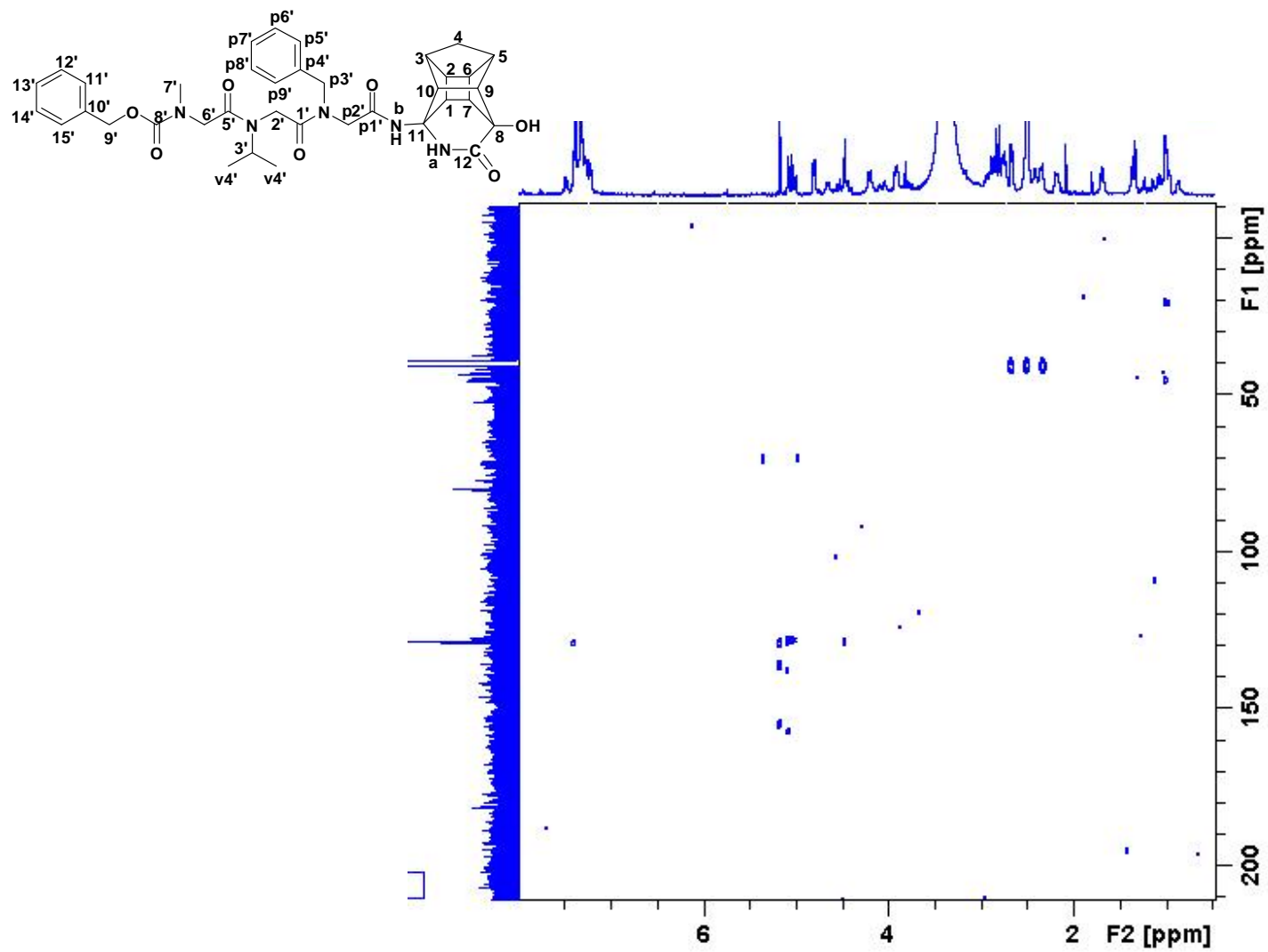
^1H NMR spectrum of PCU-PVAZ peptoid, 17



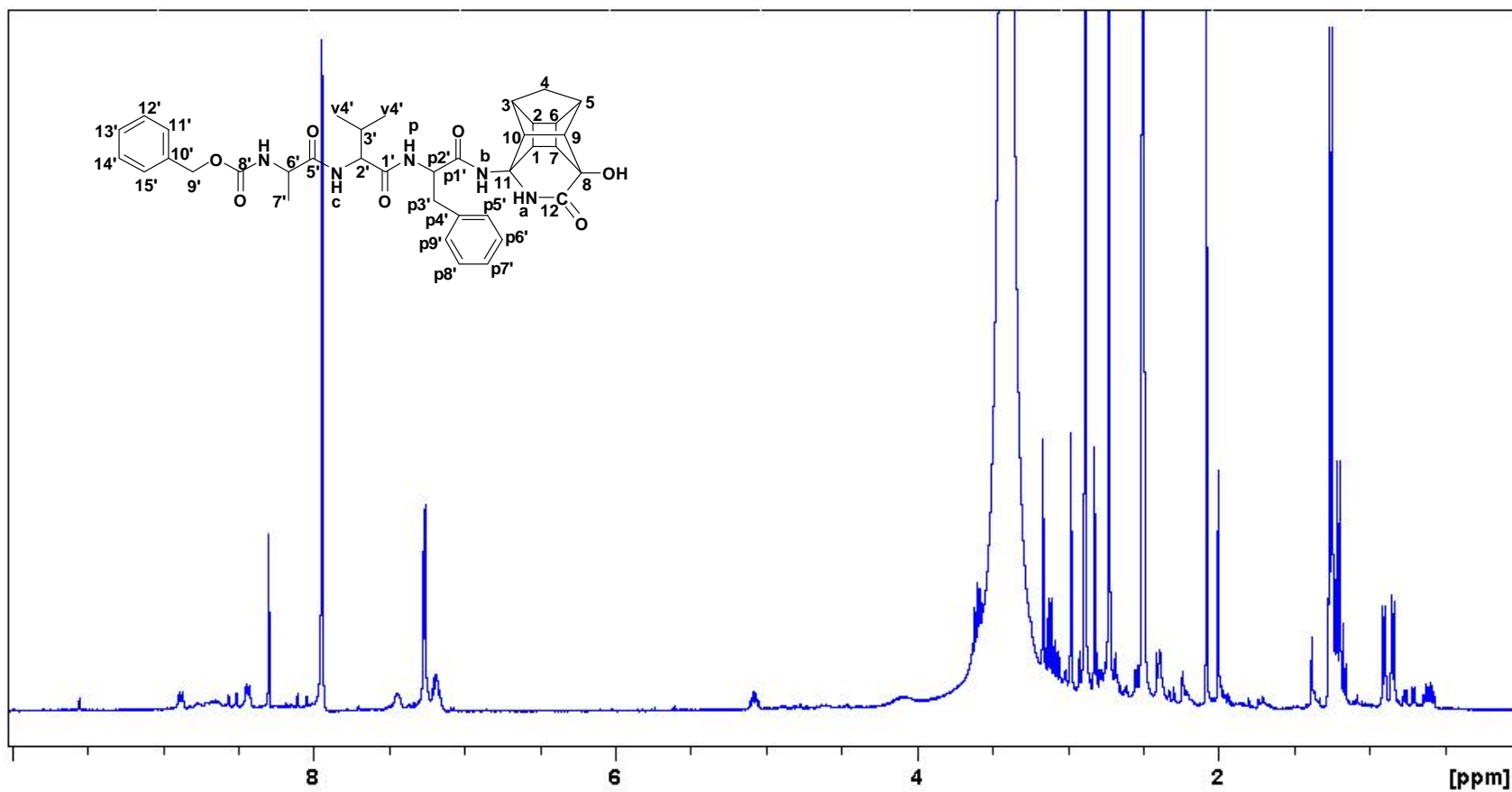
COSY spectrum of PCU-PVAZ peptoid,17



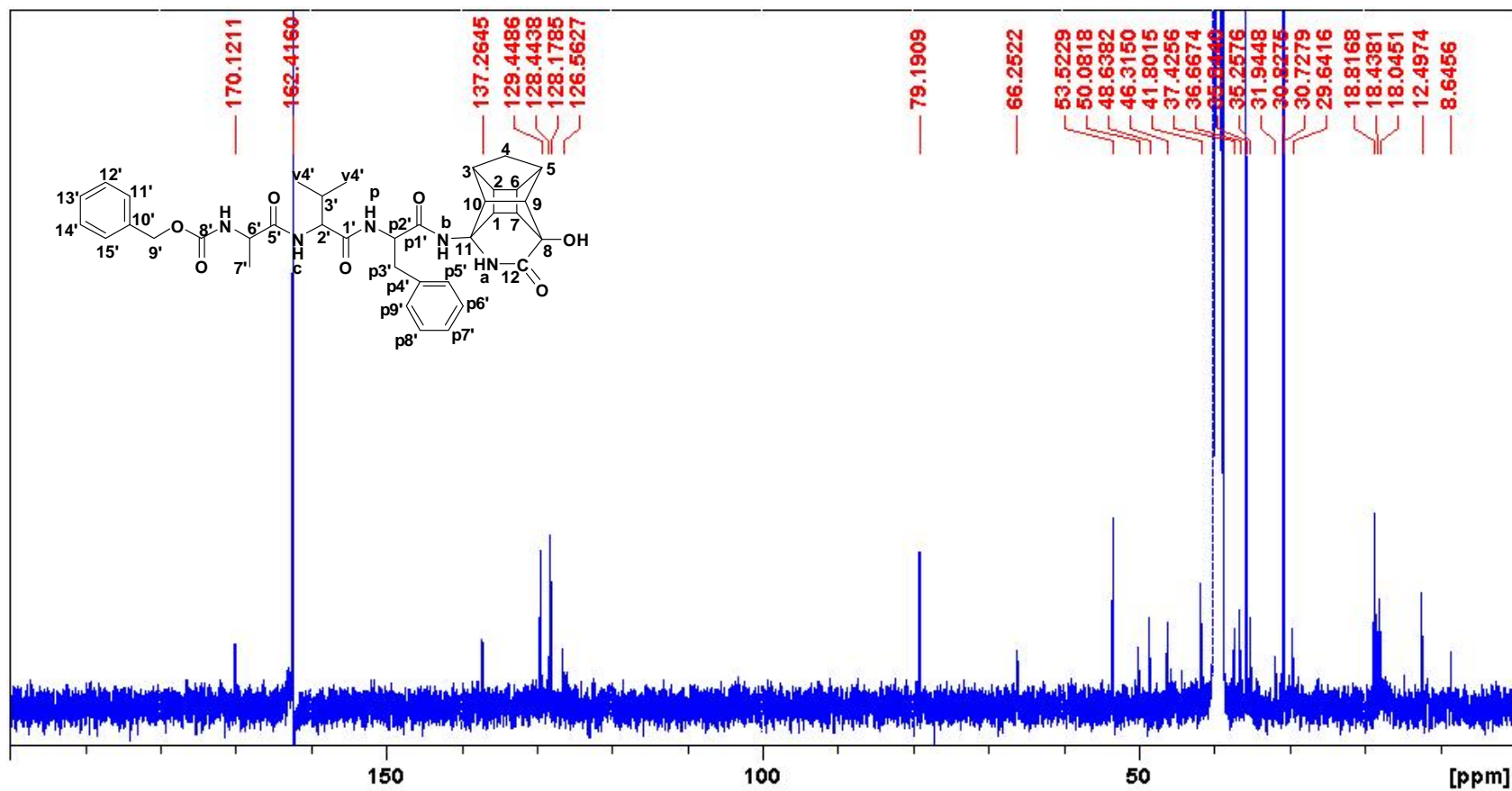
HSQC spectrum of PCU-PVAZ peptoid,17



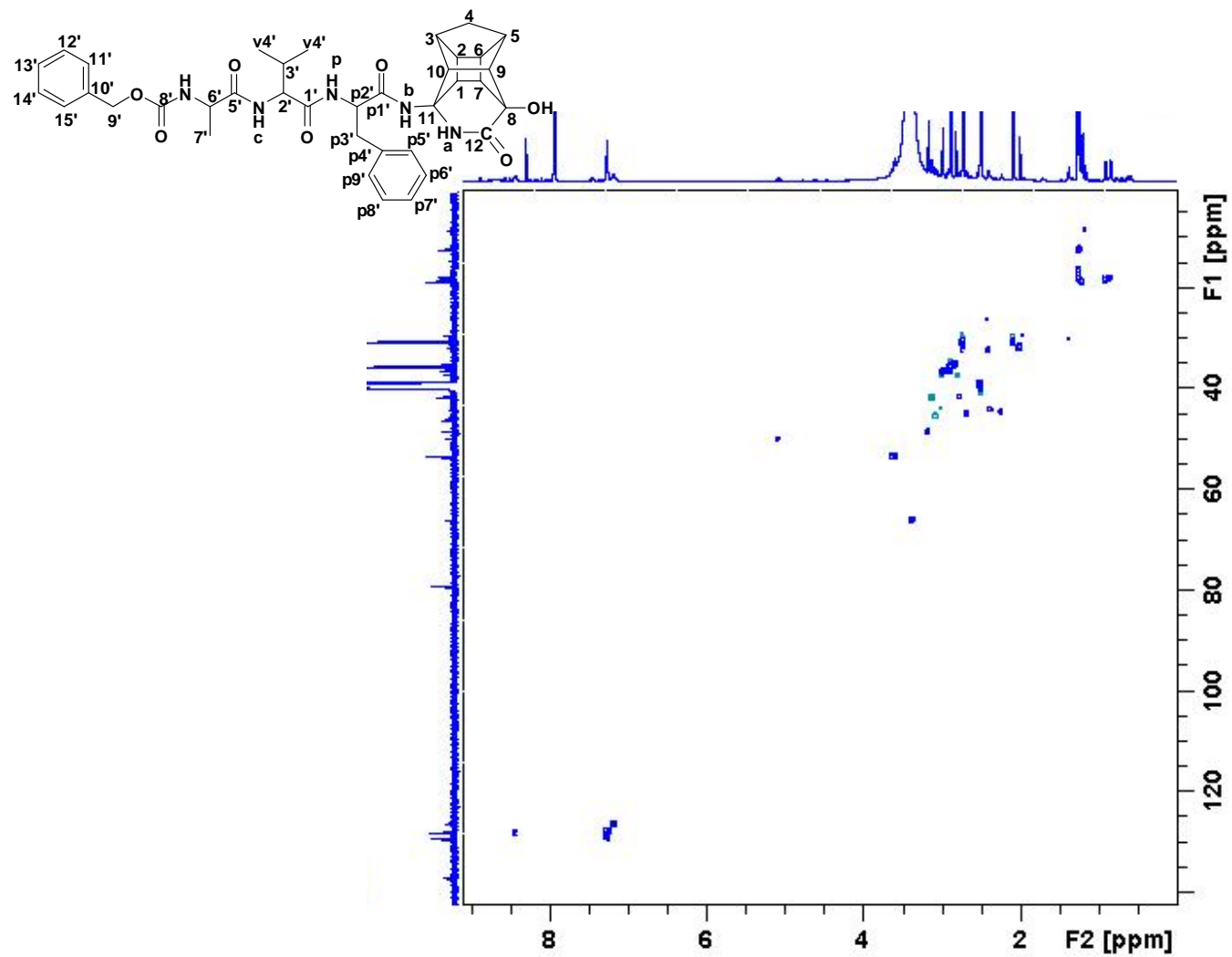
HMBC spectrum of PCU-PVAZ peptoid,17



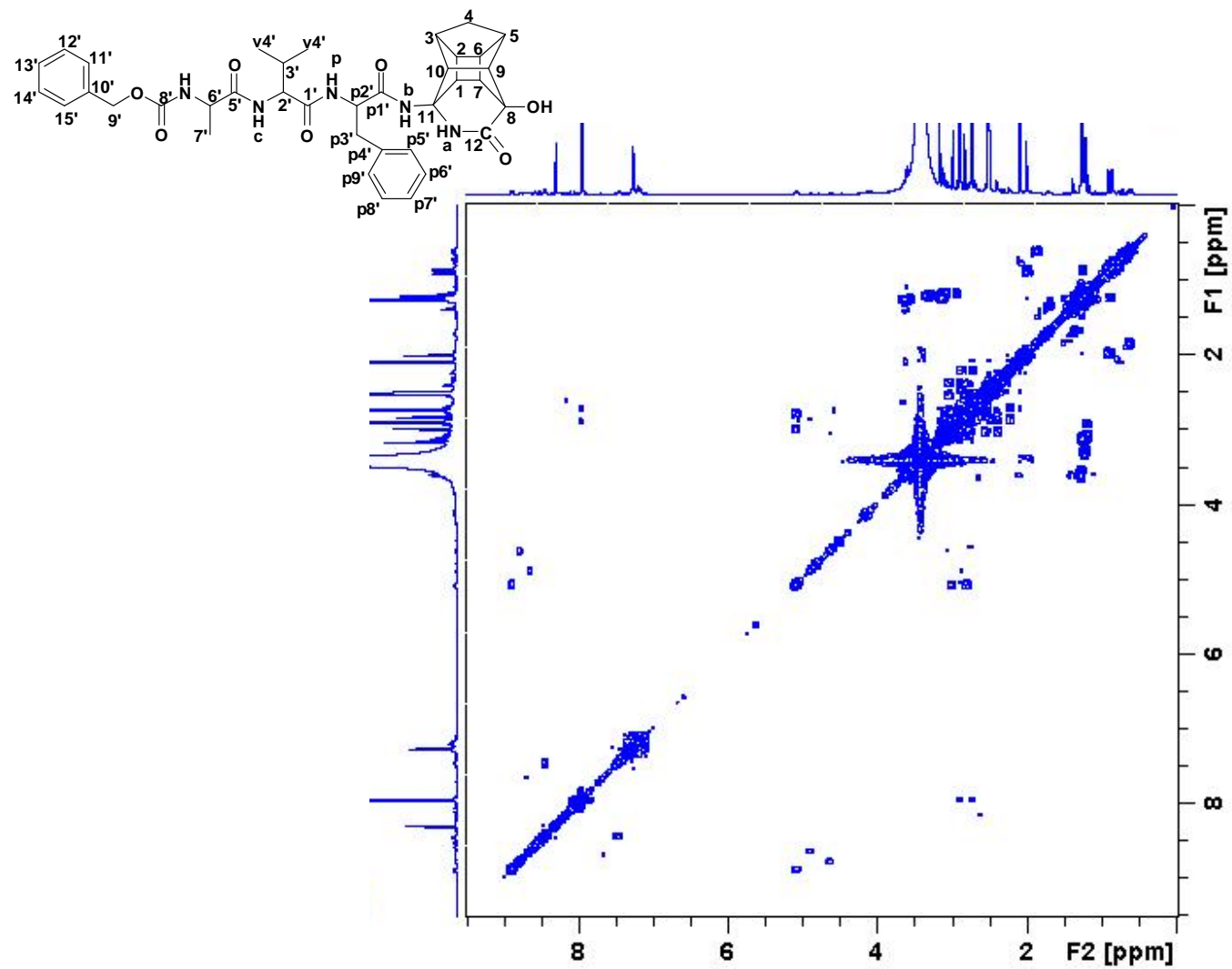
^1H NMR spectrum of PCU-PVAZ peptide, 15



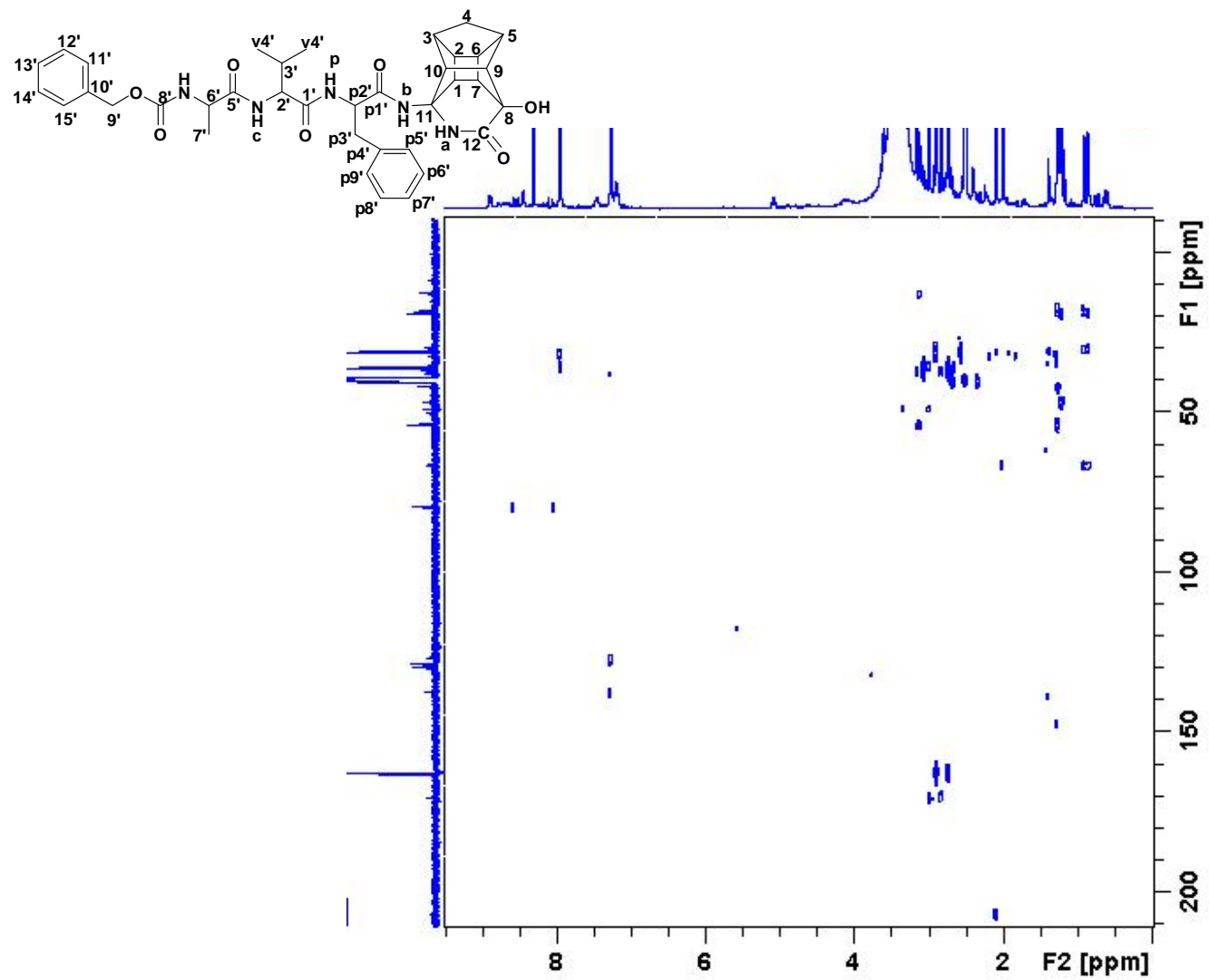
^{13}C NMR spectrum of PCU-PVAZ peptide, 15



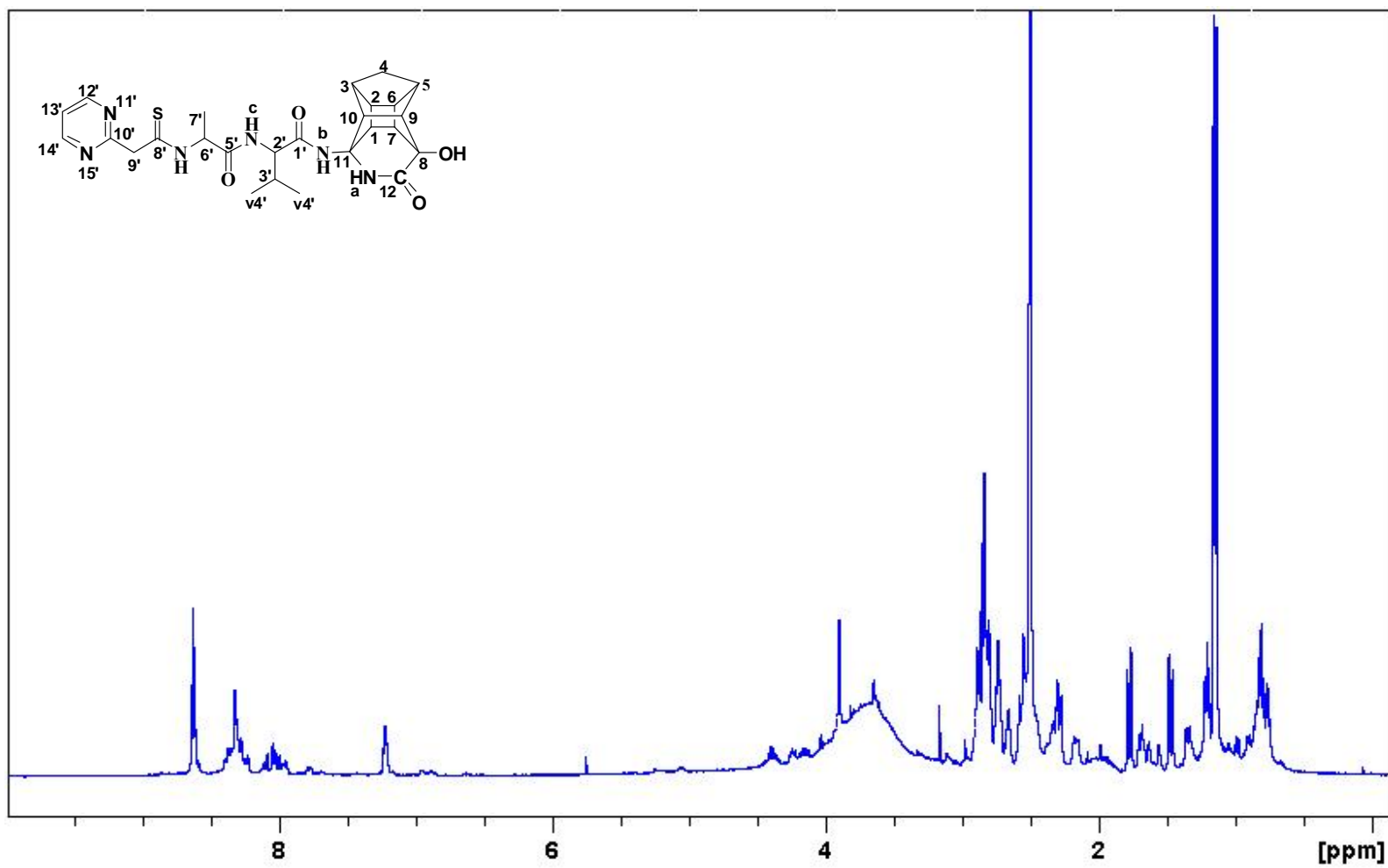
HSQC spectrum of PCU-PVAZ peptide,15



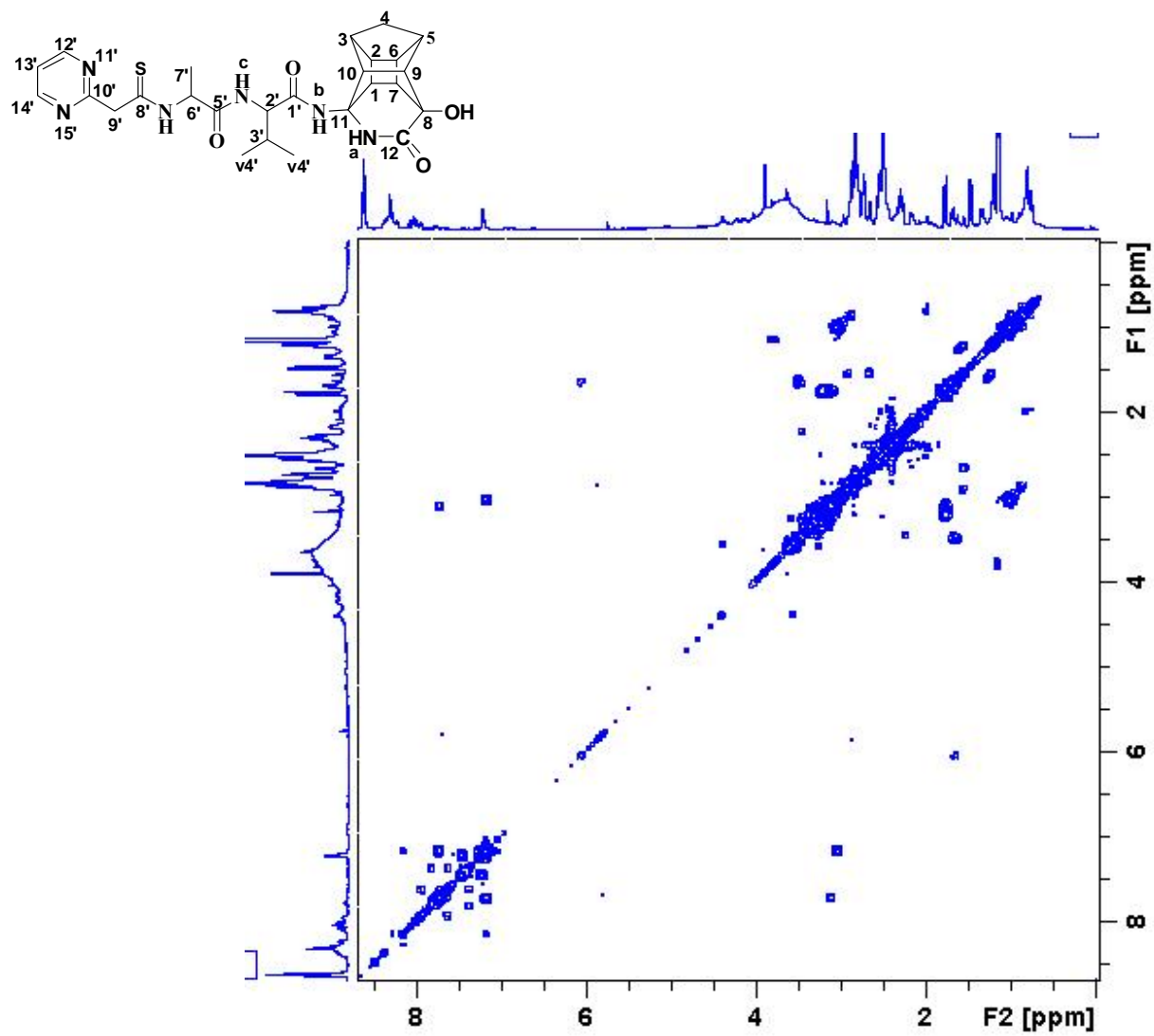
COSY spectrum of PCU-PVAZ peptide,15



HMBC spectrum of PCU-PVAZ peptide, 15



^1H NMR spectrum of PCU-VAT peptide, 16c



COSY spectrum of PCU-VAT peptide, 16c

APPENDIX 3

CHAPTER 4

SUPPORTING INFORMATION

MATERIALS AND METHODS

Analytical analysis was performed on an Agilent 1100 HPLC with a flowrate of 1 mL/min on a Waters Xbridge C18 (150 mm x 4.6 mm x 5 microns) coupled to a UV detector (215 nm) and an Agilent VL ion trap mass spectrophotometer in the positive mode. Semi-preparative HPLC was carried out on a Shimadzu semi-preparative instrument with a flowrate of 17 mL/min on a Ace C18 (150 mm x 21.2 mm x 5 microns) with a UV/VIS detector (215 nm) and an automated fraction collector. A two-buffer system was employed, utilizing formic acid as the ion-pairing agent. Buffer A consisted of 0.1 % formic acid/H₂O (v/v) and buffer B consisted of 0.1 % formic acid/acetonitrile (v/v). High resolution mass spectroscopic analysis was performed on a Bruker MicroTOF QII mass spectrometer in positive mode with an internal calibration. Peptides were synthesized on an automated CEM Liberty microwave peptide synthesizer (see Table 1 for conditions, not same as supplier). All ¹H, ¹³C, HSQC, COSY and HMBC NMR data were recorded on a Bruker AVANCE III 400 MHz spectrometer. The ROESY data was recorded on a Bruker AVANCE III 600 MHz spectrometer.

Supplementary Table 1. Microwave conditions for coupling and deprotection

	Microwave power (Watts)	Temperature (°C)	Time (sec)
Single coupling			
30 minute coupling	0	25	900
	35	73	900
Deprotection	40	73	180

NMR

Samples were dissolved in 450 µl of deuterated DMSO-d₆ and measured at 25 °C. All ¹H, ¹³C, COSY, HMBC and HSQC NMR spectra, required for assignment were recorded on a Bruker AVANCE III 400 MHz spectrometer, BBO probe with z-gradient with standard parameters. The EASY-ROESY⁽²⁸⁾ measurements required for structural analysis were recorded on a Bruker AVANCE III 600 MHz spectrometer using a BBO probe with z-gradient at a transmitter frequency of 600.1 MHz (spectral width,

10.0143 Hz; acquisition time, 0.1704436 s; 90° pulse width, 11.02 μ s; scans, 8; relaxation delay, 2.0 s, mixing time 0.12s, 2048 points in F2 dimensions). Processing and assignments were carried out using the Topspin 2.3 software from Bruker Karlsruhe. In order to determine if the same 3D structural conformation observed for the inhibitors in DMSO- d_6 are also valid for the HIV-PR activity studies, the NMR spectra for PCU-EAIS were recorded in a buffered D₂O solution (see chapter 1).

SYNTHESIS OF PCU DIACID 1

A solution of the diene **5** (31) (15.0 g, 58.0 mmol) in dry methanol (100 mL) was purged with nitrogen for 6 h while cooling in a dry ice–isopropanol bath (–78 °C). Ozone was bubbled into the reaction mixture until a blue–purple color persisted indicating the presence of excess ozone in the system and hence the completion of the reaction. The excess ozone gas was purged with nitrogen and the solvent (MeOH) removed *in vacuo*. Formic acid (100 mL) was added to the ozonide and the mixture was cooled in an ice bath with stirring. Hydrogen peroxide (150 mL, 30 %) was then added drop-wise to the stirring cooled reaction mixture. The reaction was left to attain ambient temperature for 1 h and refluxed gently for 12 h, the resulting mixture was concentrated *in vacuo* to yield the diol diacid **1** (15.6 g, 92 % yield) as a white solid. mp: 153–157°C, ¹H NMR [(CD₃)₂SO, 400 MHz]: δ H 1.03 (AB, J_{AB} = 10.4 Hz, 1H), 1.51 (AB, J_{AB} = 10.4 Hz, 1H), 2.23 (s, 3H), 2.26 (s, 2H), 2.37 (s, 2H), 2.48–2.50 (m, 3H), 2.57–2.59 (m, 2H). ¹³C NMR [(CD₃)₂SO, 100 MHz]: δ C 33.6 (CH₂), 38.9 (CH), 42.3 (CH), 43.9 (CH), 44.2 (CH₂), 50.0 (CH), 76.6 (C), 172.5 (C).

SYNTHESIS OF PCU DIAMIDE 6

The PCU diol diacid **1** (5.0 g, 17.0 mmol), NH₄Cl (2.4 g, 44.2 mmol), HOBt (5.7 g, 37.4 mmol), EDC (7.1 g, 37.4 mmol), and NMM (4.9 mL, 44.2 mmol) was dissolved in dimethylformamide (DMF) (30 mL) and stirred at room temperature for 36 h. The reaction mixture was concentrated *in vacuo*. The crude product was dissolved in ethylacetate and washed successively with 1 M HCl (50 mL), 10 % NaHCO₃ (50 mL), and brine (50 mL). The organic solution was dried over Na₂SO₄ and concentrated *in vacuo*. The crude product was purified *via* column chromatography on silica gel using CHCl₃:MeOH:25 % NH₄OH (88 : 10 : 2, R_f = 0.3, 72 % yield) to obtain a white solid. mp; 207–209 °C, IR ν_{\max} : 3172, 2961, 1656, 1408, 1287, 728, 634, and 436 cm⁻¹. Mass spectrometry (time of flight) [MS (TOF)] calculated for C₁₅H₂₁N₂O₄ (M⁺ H⁺) 293.1496, found 293.1484. ¹H NMR [(CD₃)₂SO, 400 MHz]: δ H 1.03 (AB, J_{AB} = 10.5 Hz, 1H), 1.48 (AB, J_{AB} = 10.5 Hz, 1H), 2.09 (AB, J_{AB} = 14.4 Hz, 1H), 2.19 (AB, J_{AB} = 14.45 Hz, 1H) 2.19 (s, 2H), 2.40 (s, 2H), 2.41 (s, 2H), 2.48 (s, 2H). ¹³C NMR [(CD₃)₂SO, 100 MHz]: δ C 33.5 (CH₂), 38.9 (CH), 42.5 (CH), 43.8 (CH), 43.9 (CH₂), 49.3 (CH), 76.6 (C), 173.8 (C).

SYNTHESIS OF PCU DIAMINE 2

Sodium hypochlorite (90 mL) and 4 M NaOH (48 mL) were added to 7 g of the diamide. The reaction was stirred for an hour at 80°C and then left to stir at room temperature for 24 hours. HCl (1M) was then added to adjust the pH to 7 before the reaction was concentrated in *vacuo*. The crude product was then purified using column chromatography CHCl₃:MeOH:25 % NH₄OH (88 : 10 : 2) (2.9 g, 53 % yield). The diamine (2.5g, 10.6 mmol) was dissolved in 15 mL of water and 3 equivalence (eq) of NaHCO₃ was added while stirring. The resulting solution was cooled to 5°C and *t*-BOC anhydride (2.4 eq) was added slowly as a solution in cooled THF (15 mL). The mixture was stirred at 0°C for 1 hour and allowed to warm to room temperature overnight. Water was then added and extracted twice with DCM. The organic layer was back extracted with saturated NaHCO₃. The aqueous layer was acidified to pH1 and then extracted with DCM twice. All organic layers were combined, dried over MgSO₄ and concentrated under reduced pressure. The crude *t*-BOC-PCU diamine product was purified with column chromatography and stored below 4°C. (4.1 g, 75 % yield) ¹H NMR [(CD₃)₂SO, 400 MHz]: δH 1.02 (*AB*, *J*_{AB} = 10.5 Hz, 1H), 1.44 (*AB*, *J*_{AB} = 10.5 Hz, 1H), 1.38 (s, 9H), 2.09 (s, 2H), 2.34 (s, 2H), 2.42-2.45 (m, 4H), 2.86 (*AB*, *J*_{AB} = 6.18 Hz, 1H), 2.90 (*AB*, *J*_{AB} = 6.18 Hz, 1H), 2.94 (*AB*, *J*_{AB} = 6.18 Hz, 1H), 2.98 (*AB*, *J*_{AB} = 6.18 Hz, 1H) 6.28 (s, 1H), 7.02 (s, 1H). ¹³C NMR [(CD₃)₂SO, 100 MHz]: δ_C 28.2 (CH₃), 33.7 (CH₂), 39.2 (CH), 40.4 (CH), 43.0 (CH), 43.0 (CH), 47.9 (CH), 48.1 (CH₂), 77.1 (C), 77.7 (C), 156.2 (C)

The removal of *t*-BOC from the diamine was achieved by stirring the compound in methanol (30 mL) and conc HCl (12M, 5 mL) overnight.

SYNTHESIS OF THE CAGE AMINE AMIDE 3

Sodium hypochlorite (45 mL) and 4M NaOH (24 mL) were added to 7 g of the diamide. The reaction was stirred for an hour at 80°C and then left to stir at room temperature for 24 hours. HCl (1 M) was then added to adjust the PH to 7 before the reaction was concentrated in *vacuo*. The crude product was then purified using column chromatography CHCl₃:MeOH:25 % NH₄OH (88 : 10 : 2) (1.6 g, 28 % yield). The amine amide (2.5 g, 10.6 mmol) was dissolved in 15 mL of water and 1.5 equivalence (eq) of NaHCO₃ was added while stirring. The resulting solution was cooled to 5°C and *t*-BOC anhydride (1.2 eq) was added slowly as a solution in cooled THF (15 mL). The mixture was stirred at 0°C for 1 hour and allowed to warm to room temperature overnight. Water was then added and extracted four times with DCM, dried over MgSO₄ and concentrated under reduced pressure. The crude *t*-BOC-PCU amine amide product was purified with column chromatography and stored below 4°C. (1.9 g, 86 % yield) ¹H NMR [(CD₃)₂SO, 400 MHz]: δH 1.02 (*AB*, *J*_{AB} = 10.5 Hz, 1H), 1.47 (*AB*, *J*_{AB} = 10.5 Hz, 1H), 1.37 (s, 9H), 2.09 (s, 2H), 2.33 (s, 2H), 2.42-2.45 (m, 4H), 2.92 (m, 2H), 3.12 (m, 1H), 3.16 (m, 1H), 6.20 (s, 1H), 8.32 (s,

¹H). ¹³C NMR [(CD₃)₂SO, 100 MHz]: δ_C 28.0 (CH₃), 33.7 (CH₂), 36.4 (CH₂), 40.4 (CH), 43.1 (CH), 43.5 (CH), 45.3 (CH₂), 48.0 (CH), 48.1 (CH₂), 76.7 (C), 76.9 (C), 77.6 (C), 78.4 (C), 153.3 (C), 156.2 (C), 172.1 (C), 174.1 (C),

The removal of *t*-BOC from the amine was achieved by stirring the compound in methanol (30 mL) and conc HCl (12M, 5 mL) overnight

GENERAL PROCEDURE FOR LOADING OF FIRST AMINO ACID TO THE 2-CHLOROTRITYL CHLORIDE RESIN

Activated 2-chlorotrityl chloride resin (1 g, 1.33 mmol) was swelled in dry DCM (10 mL) for 10 minutes in a sintered glass reaction vessel. Fmoc-amino acid-OH (3.99 mmol) was dissolved in DCM (10 mL) and DIPEA (6.65 mmol) and then added to the reaction vessel. The mixture was bubbled for 2 hours using nitrogen. The solvent was removed by filtration and the resin was washed with DCM (3 x 10 mL). A few resin beads (~5 mg) were removed from the reaction vessel and left to dry in the vacuum oven for an hour. A solution of 20 % piperidine in DMF was added to the dried resin beads (1 mg) and left to stand for 20 min. General washing procedure for manual SPPS: after each step, the resin was washed with DCM (3 x 10 mL), DMF (3 x 10 mL) and DCM (3 x 10 mL).

GENERAL PROCEDURE FOR THE SYNTHESIS OF PEPTIDES USING MICROWAVE POWER

Stock solutions of all amino acids (0.2 mM), DIPEA (1 mM) and HBTU (2 mM) were prepared in DMF and peptides were synthesized on a 0.5 mmol scale. The modified coupling method described in Table 1 above was used for coupling the subsequent amino acids to the 2-chlorotrityl chloride resin (1 g) preloaded with the first amino acid.

Final cleavage: The resin bound peptide was removed from the peptide synthesizer and transferred to a manual peptide synthesis reaction vessel for the final cleavage. It was washed with DCM (3 x 10 mL) and a cleavage mixture of 5:95 % (v/v) TFA:DCM was added to the resin while nitrogen was bubbled through the solution for 10 minutes. The resin was washed three times with the cleavage mixture and the cleaved peptide was removed by filtration and collected in a flask containing water (100 mL). The filtrate was extracted several times with DCM so as to remove the peptide from the water layer. DCM was removed under reduced vacuum with at 40°C and the peptide remained as a white powder.

GENERAL PROCEDURE FOR COUPLING OF PEPTIDES TO THE PCU-DIOL

The cleaved peptide (2.6 eq.) was dissolved in DMF (3 mL) followed by addition of PCU-diol (1 eq), EDC (2.4 eq), HOBT (2.4 eq), in DMF (7 mL) and NMM (2.6 eq). The mixture was stirred at room temperature for 24 hours and thereafter evaporated to dryness under reduced pressure using a Teflon pump at 40°C. The crude product was diluted with EtOAc (200 mL) and the resulting solution was washed with 1M HCl, sat. aq. NaHCO₃ and sat. aq. NaCl, dried over MgSO₄, filtered and concentrated *in vacuo*. The product was then purified by reverse phase semi-preparative HPLC.

OVER-EXPRESSION, EXTRACTION AND PURIFICATION OF THE C-SA PROTEASE

Plasmid encoding HIV-1 subtype C protease (containing the mutation Q7K designed to reduce the hypersensitive autolytic site) is expressed as inclusion bodies (Ido *et al.*, 1991) in *Escherichia coli* BL21 (DE3) pLysS cells. Briefly, *Escherichia coli* cells harboring the plasmid DNA were grown at 37 °C in LB medium supplemented with 100 µg/mL of ampicillin and 35 µg/mL of chloramphenicol. The overnight culture was diluted 100-fold into fresh 2 × YT medium supplemented with ampicillin (100 µg/mL) and chloramphenicol (35 µg/mL) and grown at 37 °C. When the optical density (OD₆₀₀) of the culture reached 0.4 to 0.5, over-expression of the HIV-1 C-SA protease was induced by adding IPTG. IPTG was added to final concentrations of 0.4 mM. Over-expression of the protease was allowed to continue for four hours.

The cells were pelleted after growth and resuspended in ice-cold extraction buffer [10 mM Tris, 1 mM EDTA, and 1 mM PMSF (added only fresh before use), pH 8] and disrupted using an ultra-sonicator. Following the addition of MgCl₂ and DNase I to final concentrations of 10 mM and 10 U/µL, respectively, the culture medium was stirred on ice until the viscosity of the mixture decreased. The cells were then ruptured by sonication and centrifuged at 15 000 × g for 30 minutes at 4 °C. The pellet was resuspended in ice-cold extraction buffer containing 1 % (v/v) of Triton X-100. Cell debris and protease-containing inclusion bodies were pelleted by centrifugation at 15 000 × g for 30 minutes at 4 °C. The pellet was then resuspended in a freshly prepared solubilization buffer containing 10 mM Tris, 2 mM DTT, 8 M urea, pH 8.0, at room temperature, and centrifuged at 15 000 × g for 30 minutes at 20 °C.

The protease, in the supernatant, was purified by passing through an anion exchange (DEAE) column previously equilibrated with solubilization buffer. Upon elution from the column, the protease was acidified by adding formic acid to a final concentration of 25 mM. Precipitation of significant amount of

contaminating proteins occurred upon acidification. Following an overnight incubation, the precipitated contaminants were removed by centrifugation at $15\,000 \times g$ for 30 minutes at $4\text{ }^{\circ}\text{C}$.

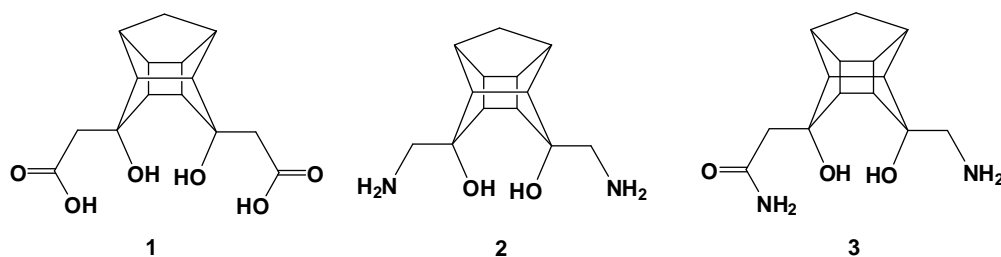
HIV-1 protease was refolded by dialysis into 10 mM formic acid at $4\text{ }^{\circ}\text{C}$. Subsequently, the protease was dialyzed into storage buffer containing 10 mM sodium acetate, 1 mM NaCl and 1 mM DTT, pH 5.0. The folded protease was concentrated to a final volume of $\sim 5\text{ mL}$ and stored at $-20\text{ }^{\circ}\text{C}$.

***IN VITRO* HIV-1 PROTEASE ACTIVITY**

The catalytic activity of the HIV-1 protease was monitored following the hydrolysis of the chromogenic peptide substrate Lys-Ala-Arg-Val-Nle-*p*-nitro-Phe-Glu-Ala-Nle-NH₂. This substrate mimics the conserved KARVL/AEAM cleavage site between the capsid protein and nucleocapsid (CA-p2) in the Gag polyprotein precursor.

For this study the chromogenic substrate was synthesized using the Discovery CEM Liberty microwave peptide synthesizer on the Rink amide resin. The substrate was cleaved from the resin and deprotected using 95 %:5 % (v/v) TFA:TIS for 3 hours. It was then precipitated using cold ether and purified *via* reverse phase semi-preparative HPLC on a Shimadzu instrument and characterized using the Bruker microTOF-Q II instrument (Table 2).

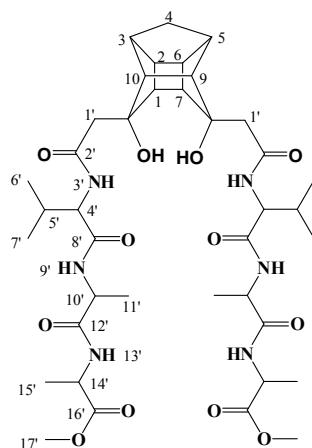
To determine the concentration of the inhibitors that resulted in 50 % inhibition (IC₅₀) of HIV-1 protease enzyme activity, the protein (100 nM) and chromogenic substrate (50 μM) were added into a 120 μL microcuvette containing increasing concentrations of inhibitor in a pH 5.0 buffer (50 mM sodium acetate and 0.1 M NaCl). Protease hydrolytic activity was measured by monitoring the relative decrease in absorbance at 300 nm using Fpecord 210 spectrophotometer



Supplementary Table 2. Percentage yields for the coupling of peptides to the cage diols and the size as confirmed by High-resolution mass spectroscopy,

Compound	R	Yields (%)	Calculated mass	Found mass	Molecular ion
1a	-Phe-Ile-Ala-Ala(OMe)	38	1127.6023 C ₅₉ H ₈₃ N ₈ O ₁₄	1127.5986	(M+H ⁺)
1b	-Phe-Val-Ala-Ala(OMe)	59	1099.5710 C ₅₇ H ₇₉ N ₈ O ₁₄	1099.5705	(M+H ⁺)
1c	-Val-Ala-Ala(OMe)	48	805.4342 C ₃₉ H ₆₁ N ₆ O ₁₂	805.4337	(M+H ⁺)
2a	-Phe-Val-Ala-Cbz	19	1139.5812 C ₆₃ H ₇₉ N ₈ O ₁₂	1139.5794	(M+H ⁺)
2b	-Phe-Val-Ala-(2-pyrimidinylthioacetic acid)	22	1175.5115 C ₅₉ H ₇₅ N ₁₂ O ₁₀ S ₂	1175.5160	(M+H ⁺)
2c	-Phe-Val-Ala(NH ₂)	17	871.5076 C ₄₇ H ₆₇ N ₈ O ₈	871.5076	(M+H ⁺)
3a	-Phe-Val-Ala-Cbz	44	716.3654 C ₃₉ H ₅₀ N ₅ O ₈	716.3652	(M+H ⁺)

NMR elucidation of diacid PCU-VAA



Supplementary Table 3: ^1H and ^{13}C chemical shift of PCU-VAA as measured by HSQC, COSY, HMBC, ROESY and ^1H and ^{13}C spectra. The chemical shift was referenced to 2.50 ppm for COSY and ROESY the correlation aided in assignment are marked in black (sequential correlations) and the long-range interactions through space, measured by ROESY, are marked in red.

Atom	$\delta^1\text{H}^{\text{a,b}}$	$\delta^{13}\text{C}^{\text{a,d}}$	COSY ($\delta^1\text{H}^{\text{a,b}}$)	HMBC	ROESY ($\delta^1\text{H}^{\text{a,c}}$)
1/7	2.43	42.4			
2/6	2.51	42.4			
3/5	2.40	43.8	4a, 4s, 9/10		4a, 4s, 9/10
4a	1.04	33.6	3/5, 4s	9/10	4s, 3/5
4s	1.45	33.6	3/5, 4a		4a, 3/5, 9/10
8/11x	-	76.9	-		
8/11y	-	70.0	-		
9/10	2.17	49.9	3/5		3/5, 4s, 6'
1'a	2.24	44.4		1/7, 2', 8/11	3'x, 3'y
1's	2.30	43.8			3'x, 3'y

2'	-	171.5	-		
3'x(NH)	7.85	-	4'	2'	1a', 1s', 4', 5', 6'
3'y(NH)	7.91	-	4'	2'	1a', 1s', 4', 5', 6'
4'	4.24	57.3	3'x, 3'y, 5'	2', 5', 6', 7', 8'	3'x, 3'y, 5', 6', 7', 9'
5'	2.00	30.4	4', 6', 7'	4', 6', 7'	3'x, 3'y, 4', 6', 7', 9'
6'	0.84	17.8	5'	4', 5', 7'	3'x, 3'y, 4', 5', 9', 9/10
7'	0.89	19.2	5'	4', 5', 6'	4', 5'
8'	-	170.4	-		
9'(NH)	8.06	-	10'	8', 10', 11'	4', 5', 6', 11' (1', 2/6)
10'	4.25	47.9	9', 11'	11'	11'
11'	1.20	17.8	10'	10', 12'	9', 10'
12'	-	172.1	-		
13'x(NH)	8.17	-	14'	12', 14', 15'	14', 15'
13'y(NH)	8.41	-	14'	12', 14', 15'	14', 15'
14'	4.24	47.6	13'x, 13'y, 15'		13'x, 13'y
15'	1.27	16.9	14'	14', 16'	13'x, 13'y, 14'
16'	-	172.9	-		
17'	3.61	51.8	-	16'	

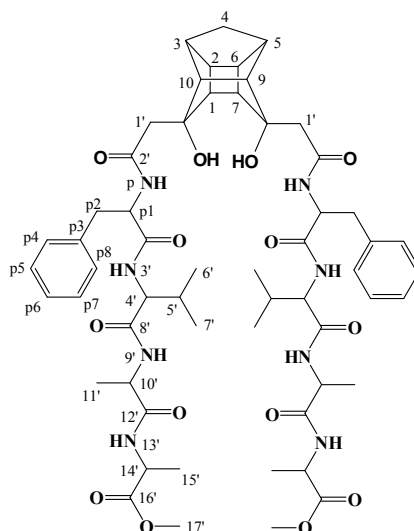
a Solvent (CD₃)₂SO.

b 400 MHz for ¹H.

c 600 MHz for ¹H.

d 100 MHz for ¹³C.

NMR elucidation of diacid PCU-PVAA



Supplementary Table 4. ^1H and ^{13}C chemical shift of PCU-PVAA as measured by HSQC, COSY, HMBC, ROESY and ^1H and ^{13}C spectra. The chemical shift was referenced to 2.50 ppm for COSY and ROESY the correlation aided in assignment are marked in black (sequential correlations) and the long-range interactions through space, measured by ROESY, are marked in red.

Atom	$\delta^1\text{H}^{\text{a,b}}$	$\delta^{13}\text{C}^{\text{a,d}}$	COSY ($\delta^1\text{H}^{\text{a,b}}$)	HMBC	ROESY ($\delta^1\text{H}^{\text{a,c}}$)
1/7	2.20	42.3	2/6		1'a, 2/6, p4-p8
2/6	2.37	38.8	1/7, 3/5		1/7, 4a
3/5	2.20	43.7	2/6, 4a, 4s, 9/10x, 9/10y		4a, 4s, 9/10x, 9/10y
4a	0.94	33.4	3/5		4s, 9/10x, 9/10y, 3/5, 2/6
4s	1.25	33.4	3/5		4a, 3/5, (p4-p8)
8/11x		76.3			
8/11y		77.0			
9/10x	1.53	48.8	3/5, 9/10y		9/10y, 3/5, 4a, p4-p8, NHpx

9/10y	1.66	48.8	3/5, 9/10x		9/10x, 3/5, 4a, NHpy, p4-p8
1'a	1.96	44.7	1's	1/7, 2', 8/11, 9/10	1/7, 1's, p4-p8, NHpx
1's	2.15	44.7	1'a		1'a, p4-p8, NHpx, NHpy
2'		170.8			
NHpx	7.97		p1x, p1y	p1x, p1y, 2'	1's, 1a, p1x, p2'a, 9/10x
NHpy	8.36		p1x, p1y	p1x, p1y, 2'	1's, p1x, p2'a, 9/10y
p1x	4.60	53.9	p2a, p2s, NHpx, NHpy	p2a, p2s, p3, p9	p2a, p2s, 3'x, NHpx, NHpy, p4-p8
p1y	4.76	53.5	p2a, p2s, NHpx, NHpy	p2a, p2s, p3, p9	p2a, p2s, 3'y, p4-p8
p2a	2.78	37.4	p1x, p1y, p2s	p1x, p1y, p3, p9	p1x, p1y, 3'x, 3'y, NHpx, NHpy, p4-p8
P2s	3.05	37.4	p1x, p1y, p2a	p1x, p1y, p3, p9	p1x, p1y, 3'x, 3'y, NHpx, NHpy, p4-p8
p3		138.3			
p4-p8	7.06-7.32	126.2-129.2	p4-p8	p2a, p2s	1'a, 1's, 1/7, p1x, p1y, p2a, p2s, 4s, 6', 7', 9/10x, 9/10y
3'x(NH)	7.91		4'	4', p9	p1x, p2a, p2s, 4', 5', 6', 7'
3'y(NH)	8.04		4'	4', p9	p1y, p2a, p2s, 4', 5', 6', 7'
4'	4.18	57.5	3'x, 3'x	5', 8'	5', 6', 7', 3'x, 3'y, 9'y
5'	1.97	30.6	6', 7'	4', 6', 7', 8'	4', 6', 7', 3'x, 3'y
6'	0.81	18.1	5'	4', 5', 7'	4', 5', 10', 14', p4-p8, 3'x, 3'y, 13'x, 13'y, 9'x

7'	0.86	19.1	5'	4', 5', 6'	4', 5', 10', 14', p4-p8, 3'x, 3'y, 13'x, 13'y, 9'x
8'		170.3			
9'x(NH)	7.98		10'	10', 11', 12'	4', 6', 7', 10', 11'
9'y(NH)	8.03		10'	10', 11', 12'	4', 10', 11'
10'	4.23	47.8	9', 11', 9'x, 9'y	11', 12'	6', 7', 11', 9'x, 9'y, 13'x, 13'y
11'	1.23	18.1	10'	10', 12'	9'x, 9'y, 10'
12'		172.0			
13'x(NH)	8.20		14'	14', 15'	6', 7', 10', 15'
13'y(NH)	8.23		14'	14', 15'	6', 7', 10', 15'
14'	4.25	47.5	13'x, 13'x, 15'	15', 16'	6', 7', 15', 17'
15'	1.26	16.9	14'	14', 16'	13'x, 13'y, 14'
16'		172.9			
17'	3.61	51.8		16'	14'

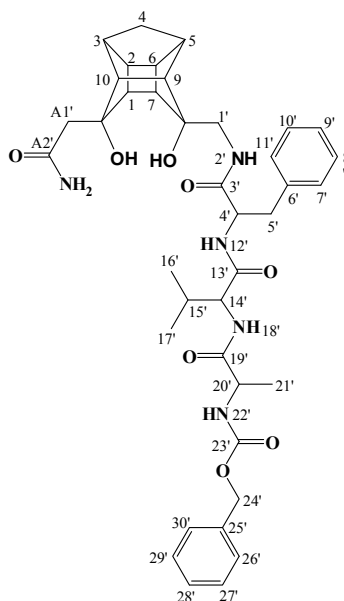
a Solvent (CD₃)₂SO.

b 400 MHz for ¹H.

c 600 MHz for ¹H.

d 100 MHz for ¹³C.

NMR elucidation of amine amide PCU-PVAZ



Supplementary Table 5: ^1H and ^{13}C chemical shift of PCU-PVAZ as measured by HSQC, COSY, HMBC, ROESY and ^1H and ^{13}C spectra. The chemical shift was referenced to 2.50 ppm for COSY and ROESY the correlation aided in assignment are marked in black (sequential correlations) and the long-range interactions through space, measured by ROESY, are marked in red.

Atom	$\delta^1\text{H}^{\text{a,b}}$	$\delta^{13}\text{C}^{\text{a,d}}$	COSY ($\delta^1\text{H}^{\text{a,b}}$)	HMBC	ROESY ($\delta^1\text{H}^{\text{a,c}}$)
1	2.40	42.3	2, 7		
2	2.46	36.5	1, 3		4a
3	2.40	39.0	2, 4a, 4s, 10		4a, 4s, 10
4a	1.05	43.1	3		2, 3, 4s, 5, 6
4s	1.45	43.8	3		3, 4a, 5, 6, 10
5	2.33	43.2	4a, 4s, 6, 9		4a, 4s, 5's, 6, 9
6	2.45	33.4	5, 7		4a, 4s, 5
7	2.24	33.4	1, 6		

8/11x		76.5			
8/11y		77.1			
9	1.92	47.7	5, 10		5, 10
10	2.13	49.2	3, 9		3, 4s, 9
A1'a	3.26	45.5			4', 5's
A1's	3.41	45.5		A2'	4', 5's, 17'
A2'		169.2			
1'	3.36	41.9			15', 17', 21'
2' (NH)	8.31	8.31		3', 4'	4', 5'a, 5's, 14', 15', 17', 21'
3'		170.3			
4'	4.91	49.8	5'a, 5's, 12'x	5', 6', 13'	5'a, 5;s, A1'a, A1's, 2', 7'- 11', 12', 20'
5'a	2.83	37.5	4'	3', 4', 6', 7'- 11'	2', 4', 7'-11', 14'
5's	2.93	37.5	4'	3', 4', 6', 7'- 11'	A1'a, A1's, 2', 4', 5, 7'-11', 12', 14', 20'
6'		137.3			
7'-11'	7.19-7.25	126.4-129.4		6'	4', 5'a, 5;s, 16', 21'
12'x(NH)	8.21		4'	4', 13'	4', 5's, 15', 17', 20', 21'
12'y(NH)					
13'		170.4			
14'	4.19	57.2	15', 18'	13', 15', 16'	2', 5'a, 5;s, 15', 16', 18'x, 18'y

15'	1.91	30.9	14', 16', 17'	13', 15', 16'	2', 12', 14', 15', 16', 17', 18'x, 18'y
16'	0.77	17.9	15'	14', 15', 17'	14', 7'-14', 15', 18'x, 18'y
17'	0.79	19.2	15'	14', 15', 16'	1', 2', 12', 15', 18'x, 18'y, 20', 22'
18'x(NH)	7.64		14'	14', 19'	14', 15', 16', 17', 20', 21'
18'y(NH)	7.71				14', 15', 16', 17', 20', 21'
19'		172.3			
20'	4.11	50.1	21', 22'	19', 21', 23'	5'a, 5's, 12', 17', 18'x, 18'y, 21', 22'
21'	1.19	18.1	20'	19', 20'	1', 2', 7-11', 12', 18'x, 18'y, 20', 22'
22'(NH)	7.94		20'	20', 21', 23'	17', 20', 21'
23'		155.6			
24'	5.16	65.6		25'	26'-30'
25'		137.0			
26'-30'	7.30-7.36	126.4-19.4		25'	24'

a Solvent (CD₃)₂SO.

b 400 MHz for ¹H.

c 600 MHz for ¹H.

d 100 MHz for ¹³C.

Analysis Info

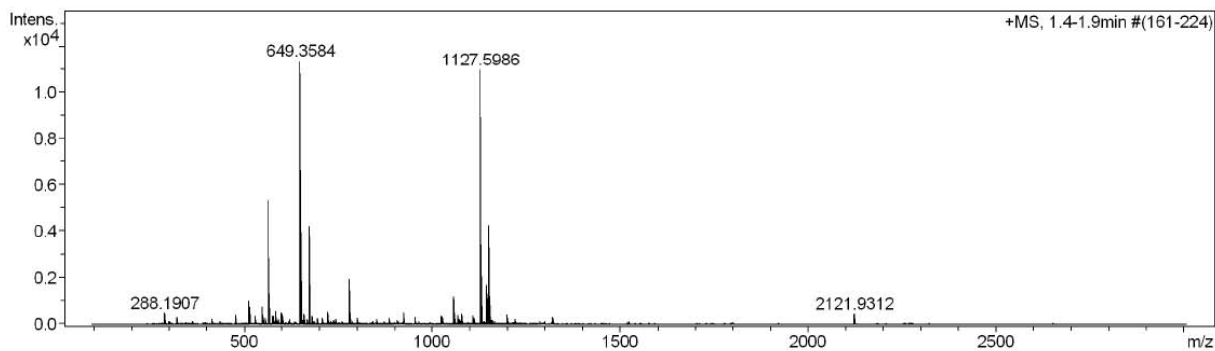
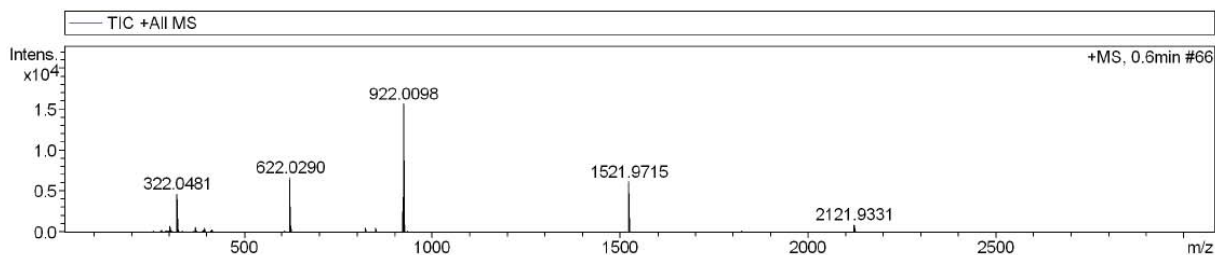
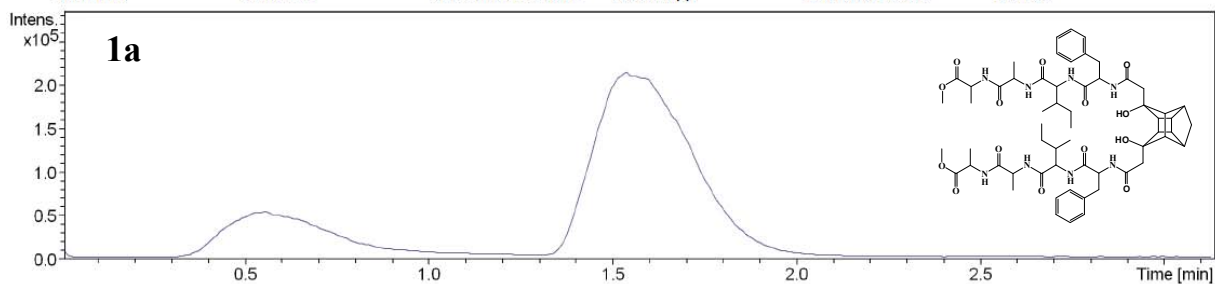
Analysis Name D:\Data\maya\PCUPIAA000001.d
 Method tune_wide_expert.m
 Sample Name PCUPIAA
 Comment 1126

Acquisition Date 7/30/2010 11:04:19 PM

Operator BDAL@DE
 Instrument / Ser# micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



Meas. m/z	#	Formula	m/z	err [ppm]	Me an err [ppm]	rdb	N-R ul e	e ⁻ Conf	mSig ma	Std I	Std Mean m/z	Std I VarNo rm	Std m/z Diff	Std Comb Dev
322.0481			322.0481											
622.0290			622.0290											
922.0098			922.0098											
1521.9715			1521.9715											
2121.9331			2121.9331											
288.1907			288.1907											
649.3584			649.3584											
1127.5986			1127.5986											
2121.9312			2121.9312											

HRMS spectrum of compound 1a

Analysis Info

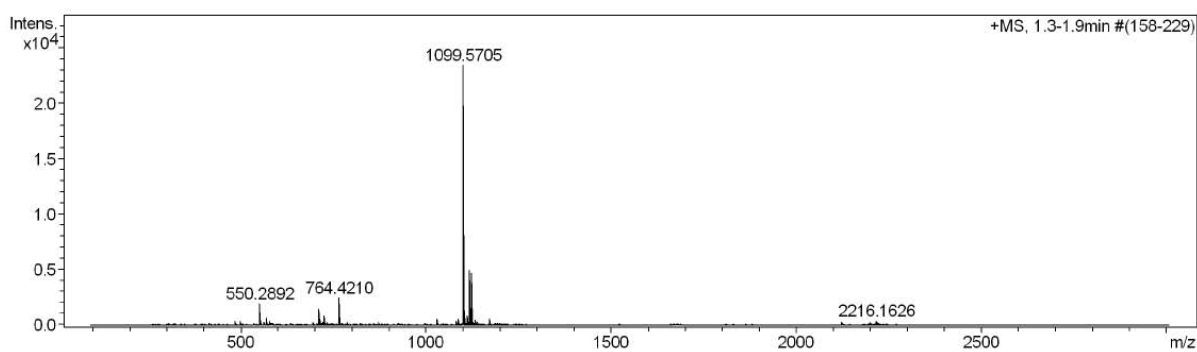
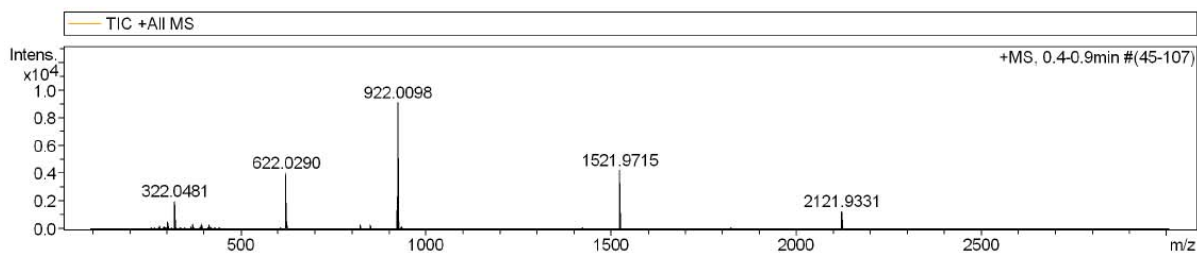
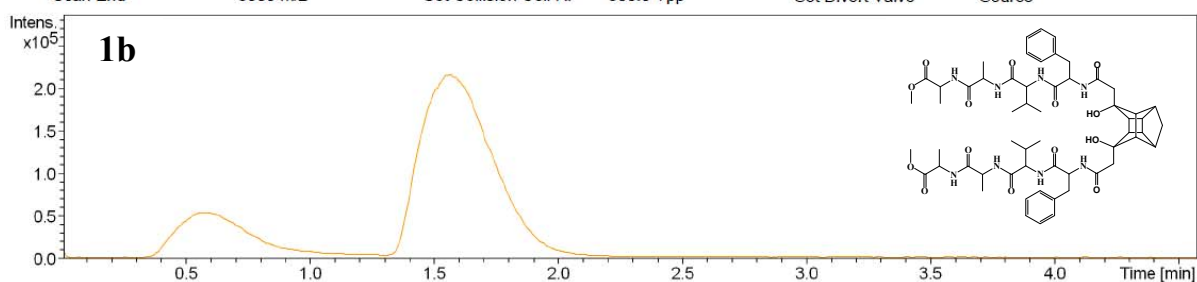
Analysis Name D:\Data\maya\PCUP\AA34000001.d
Method tune_wide_expert.m
Sample Name PCUP\AA4
Comment 1099

Acquisition Date 7/31/2010 12:34:03 AM

Operator BDAL@DE
Instrument / Ser# microTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source

HRMS spectrum of compound **1b**

Analysis Info

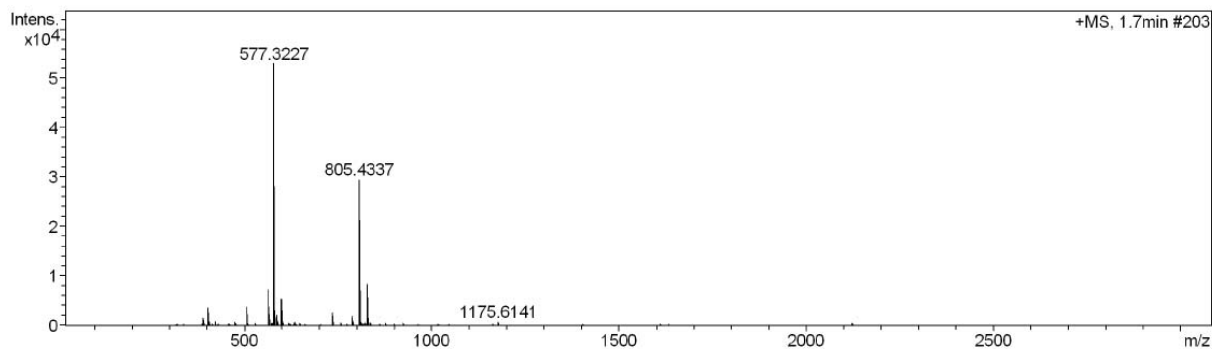
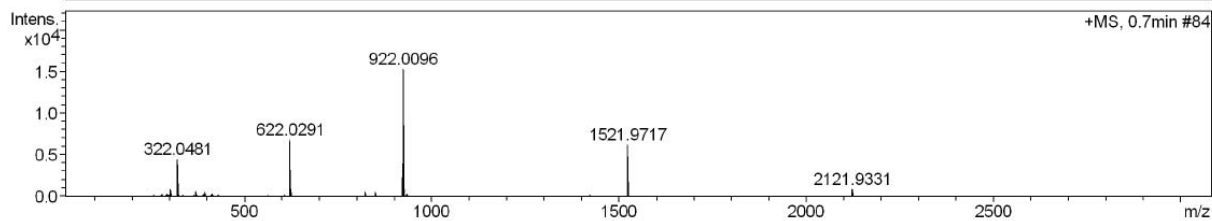
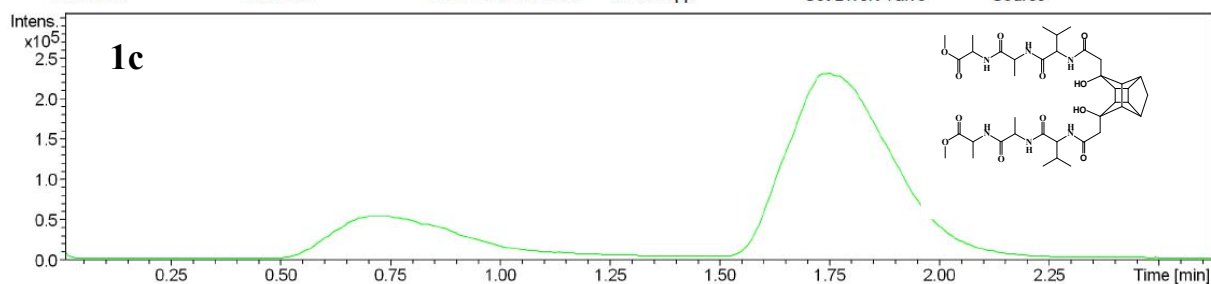
Analysis Name D:\Data\maya\PCUVAA 000001.d
 Method tune_wide_expert.m
 Sample Name PCUVAA
 Comment 1126

Acquisition Date 7/30/2010 11:13:23 PM

Operator BDAL@DE
 Instrument / Ser# micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



HRMS spectrum of compound **1c**

Analysis Info

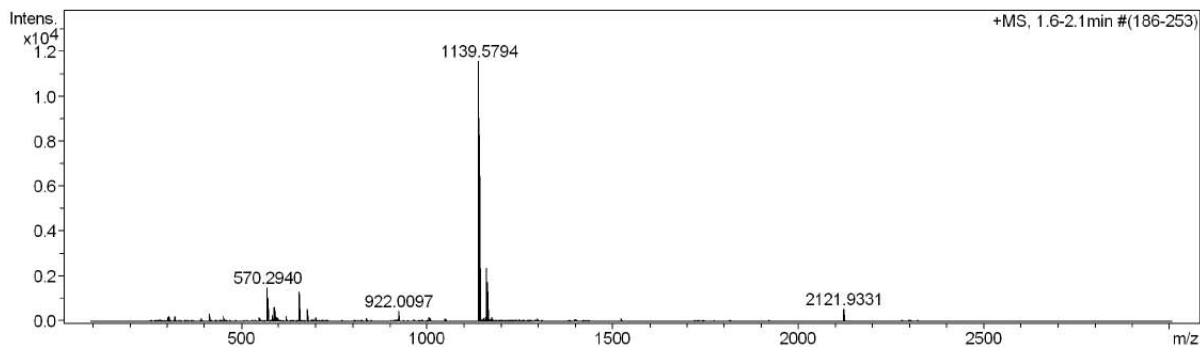
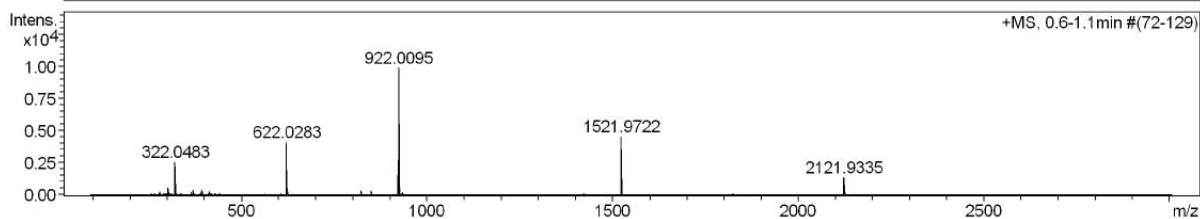
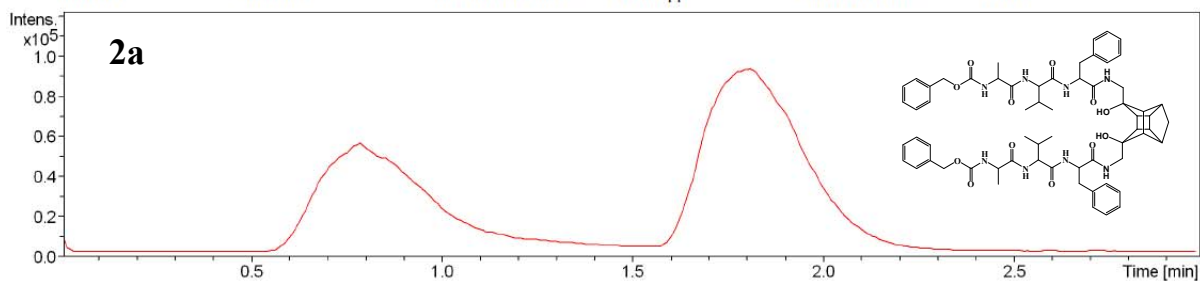
Analysis Name D:\Data\maya\PCUPVAZ 000001.d
 Method tune_wide_expert.m
 Sample Name PCUPVAZ
 Comment 1139

Acquisition Date 7/30/2010 11:31:40 PM

Operator BDAL@DE
 Instrument / Ser# micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



HRMS spectrum of compound **2a**

Analysis Info

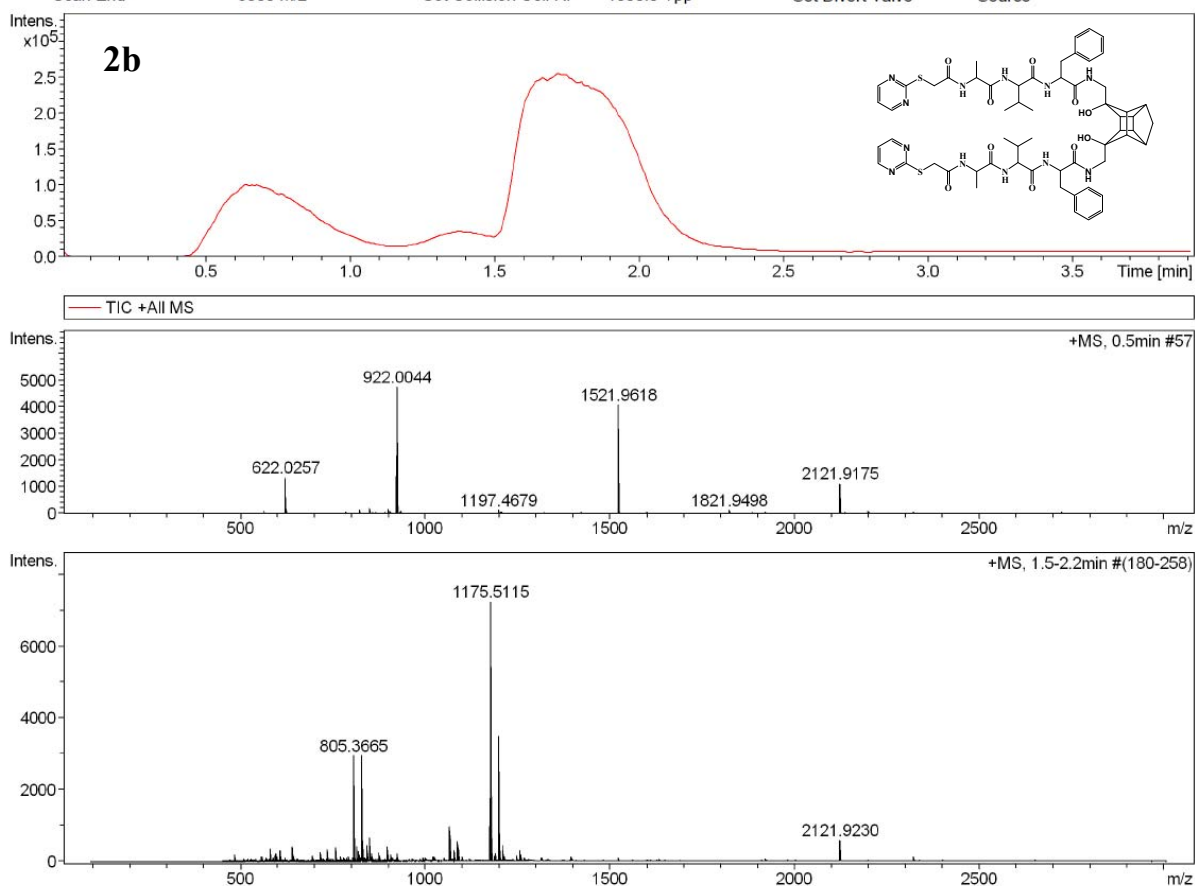
Analysis Name D:\Data\maya\cpvata peptide000001.d
Method tune_wide_expert.m
Sample Name cpvata
Comment 540

Acquisition Date 8/2/2010 10:42:51 PM

Operator BDAL@DE
Instrument / Ser# micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	1000.0 Vpp	Set Divert Valve	Source

HRMS spectrum of compound **2b**

Analysis Info

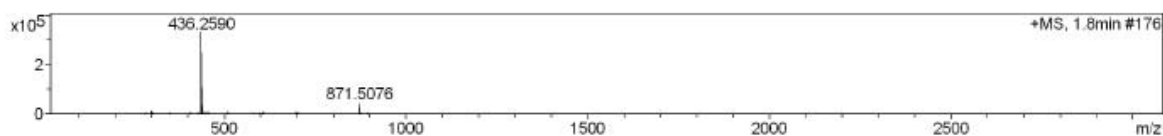
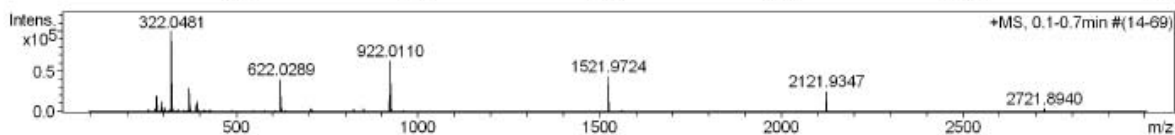
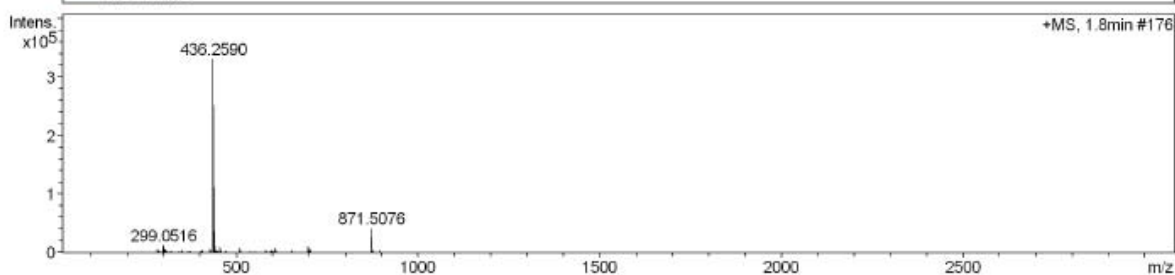
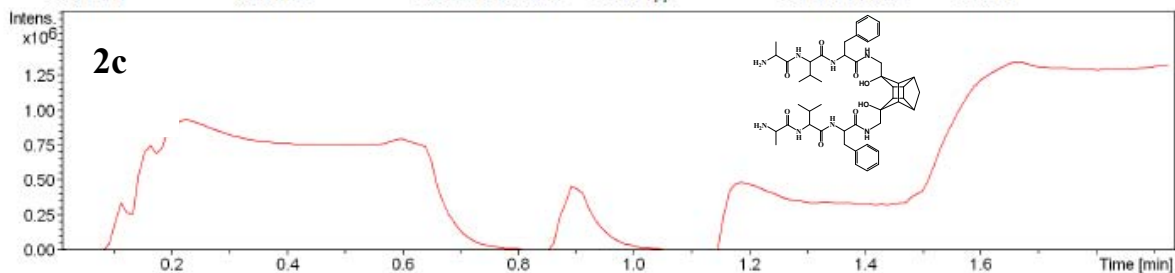
Analysis Name D:\Data\maya\PCU-PVANH2 peptide.d
 Method tune_wide_expert.m
 Sample Name PCU-PVANH2 peptide
 Comment

Acquisition Date 10/9/2010 4:14:56 PM

Operator BDAL@DE
 Instrument / Ser# micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	14.5 psi
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



Meas. m/z	#	Formula	m/z	err [ppm]	Me an err [ppm]	rdB	N-Ru le	e ⁻ Conf	mSig ma	Std I	Std Mean m/z	Std I VarNo	Std m/z Diff	Std Comb Dev
871.5078	1	C ₄₀ H ₅₅ N ₂₄	871.5036	-4.5	-4.2	25.5	ok	even	4.75	0.0051	0.0037	0.0021	0.0007	0.7814
	2	C ₄₃ H ₆₃ N ₁₄ O ₆	871.5050	-3.0	-2.6	19.5	ok	even	6.55	0.0092	0.0023	0.0037	0.0007	0.7495
	3	C ₄₂ H ₆₇ N ₁₀ O ₁₀	871.5036	-4.5	-4.2	14.5	ok	even	9.72	0.0113	0.0037	0.0048	0.0008	0.8338

HRMS spectrum of compound **2c**

Meas. m/z	#	Formula	m/z	err [ppm]	Me an err [ppm]	rdb	N-Rule	e ⁻ Conf	mSigma	Std I	Std Mean m/z	Std I VarNorm	Std m/z Diff	Std Comb Dev
	4	C ₄₅ H ₇₅ O ₁₆	871.5050	-3.0	-2.6	8.5	ok	even	13.50	0.0196	0.0023	0.0071	0.0008	0.8073
	5	C ₄₆ H ₇₁ N ₄ O ₁₂	871.5063	-1.4	-1.0	13.5	ok	even	14.04	0.0204	0.0010	0.0088	0.0008	0.7441
	6	C ₄₄ H ₅₉ N ₁₈ O ₂	871.5063	-1.5	-1.1	24.5	ok	even	17.12	0.0200	0.0011	0.0068	0.0007	0.7139
	7	C ₄₇ H ₆₇ N ₈ O ₈	871.5076	0.1	0.5	18.5	ok	even	19.80	0.0260	0.0008	0.0100	0.0008	0.7134
	8	C ₄₈ H ₆₃ N ₁₂ O ₄	871.5090	1.6	2.0	23.5	ok	even	30.99	0.0369	0.0018	0.0132	0.0008	0.8493
	9	C ₅₁ H ₇₁ N ₂ O ₁₀	871.5103	3.2	3.6	17.5	ok	even	35.68	0.0447	0.0032	0.0172	0.0008	0.9214
	10	C ₄₉ H ₅₉ N ₁₆	871.5103	3.2	3.5	28.5	ok	even	43.20	0.0487	0.0031	0.0169	0.0007	0.9182
	11	C ₅₈ H ₆₇ N ₂ O ₅	871.5044	-3.6	-3.2	26.5	ok	even	66.46	0.0832	0.0028	0.0234	0.0008	0.9380
	12	C ₅₉ H ₆₃ N ₆ O	871.5058	-2.0	-1.6	31.5	ok	even	77.21	0.0931	0.0015	0.0257	0.0008	0.9125
	13	C ₆₃ H ₆₇ O ₃	871.5085	1.0	1.5	30.5	ok	even	90.74	0.1084	0.0014	0.0305	0.0008	0.9254

HRMS spectrum of compound **2c**

Analysis Info

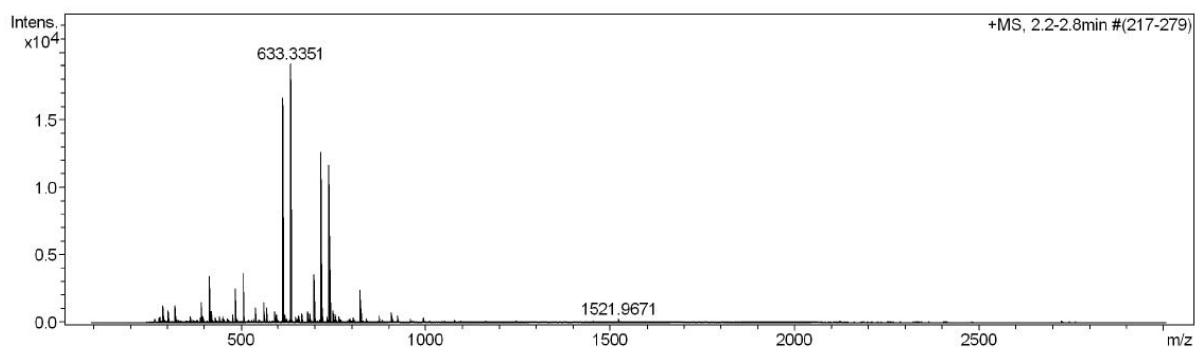
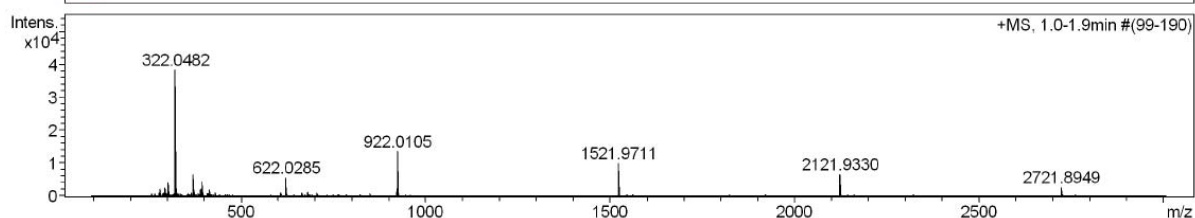
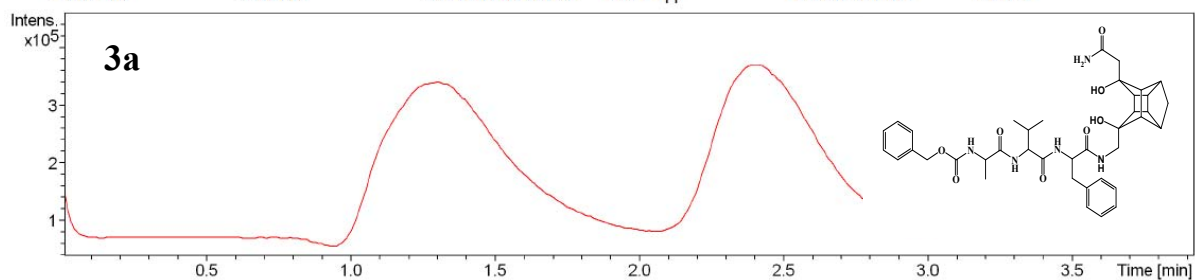
Analysis Name D:\Data\maya\CPUOOPVAZB PEPTIDE.d
 Method tune_wide_expert.m
 Sample Name CPUOOPVAZB PEPTIDE
 Comment

Acquisition Date 9/4/2010 9:01:49 PM

Operator BDAL@DE
 Instrument / Ser# micrOTOF-Q 10139

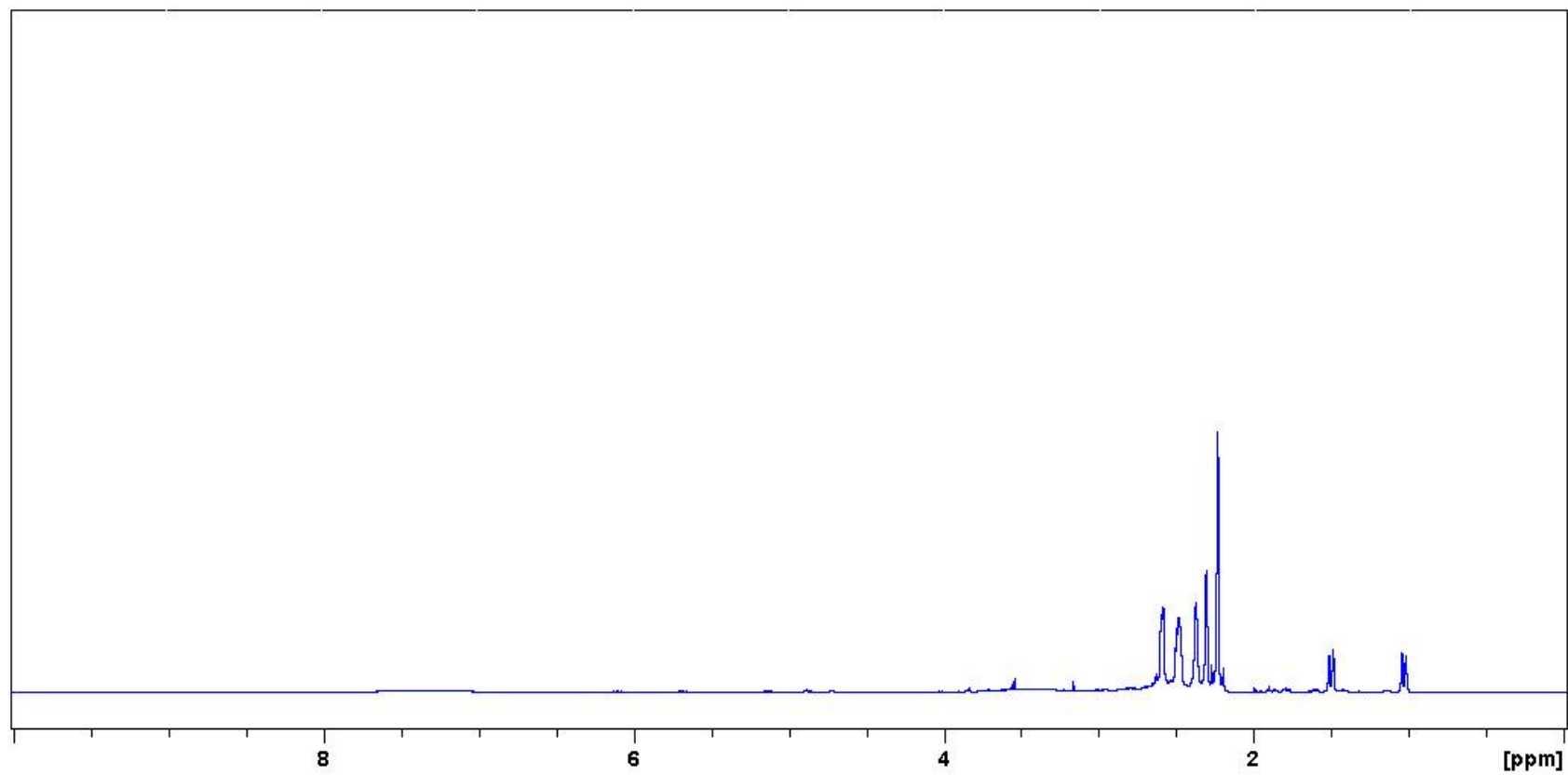
Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	1.0 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source

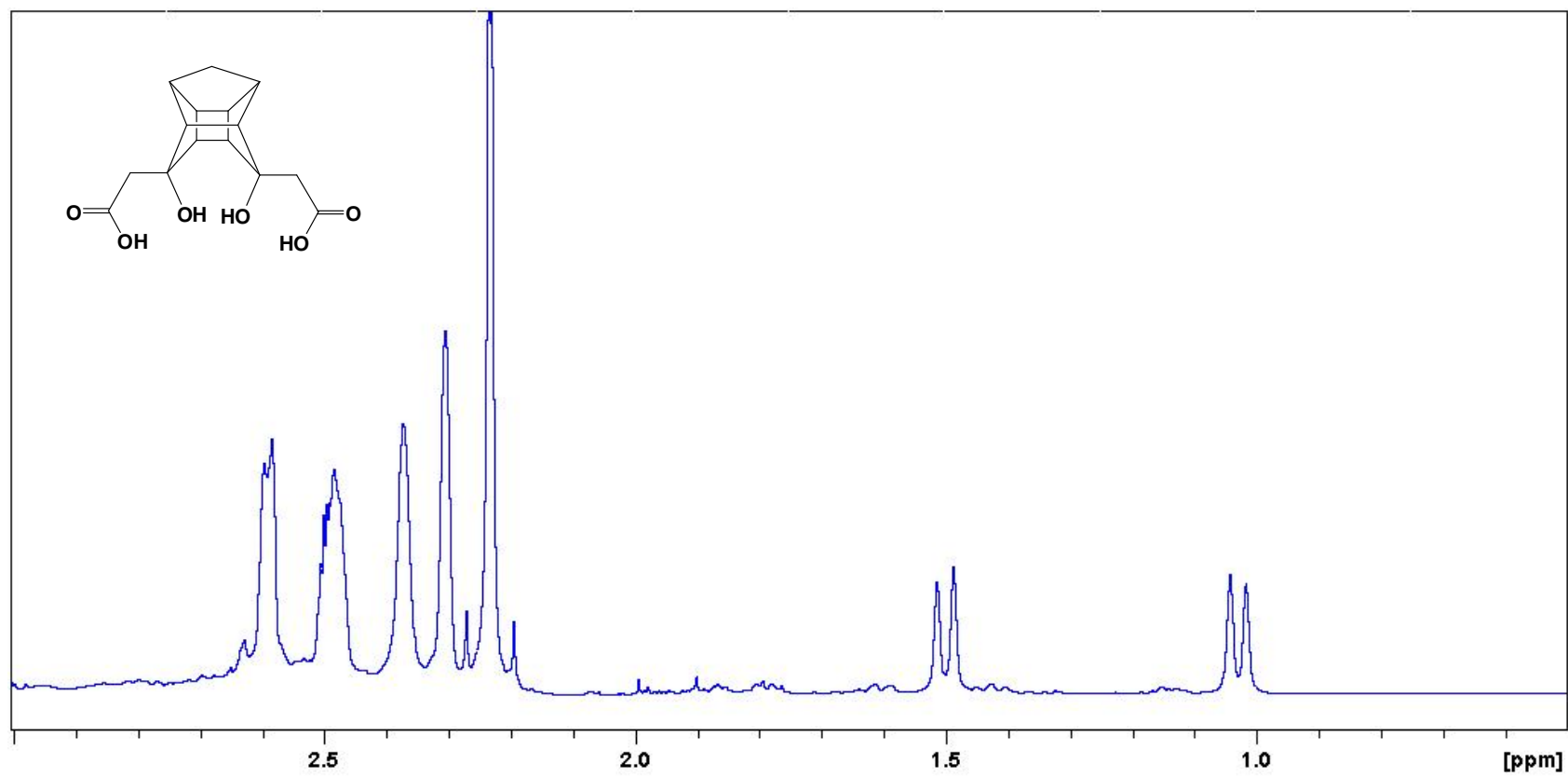


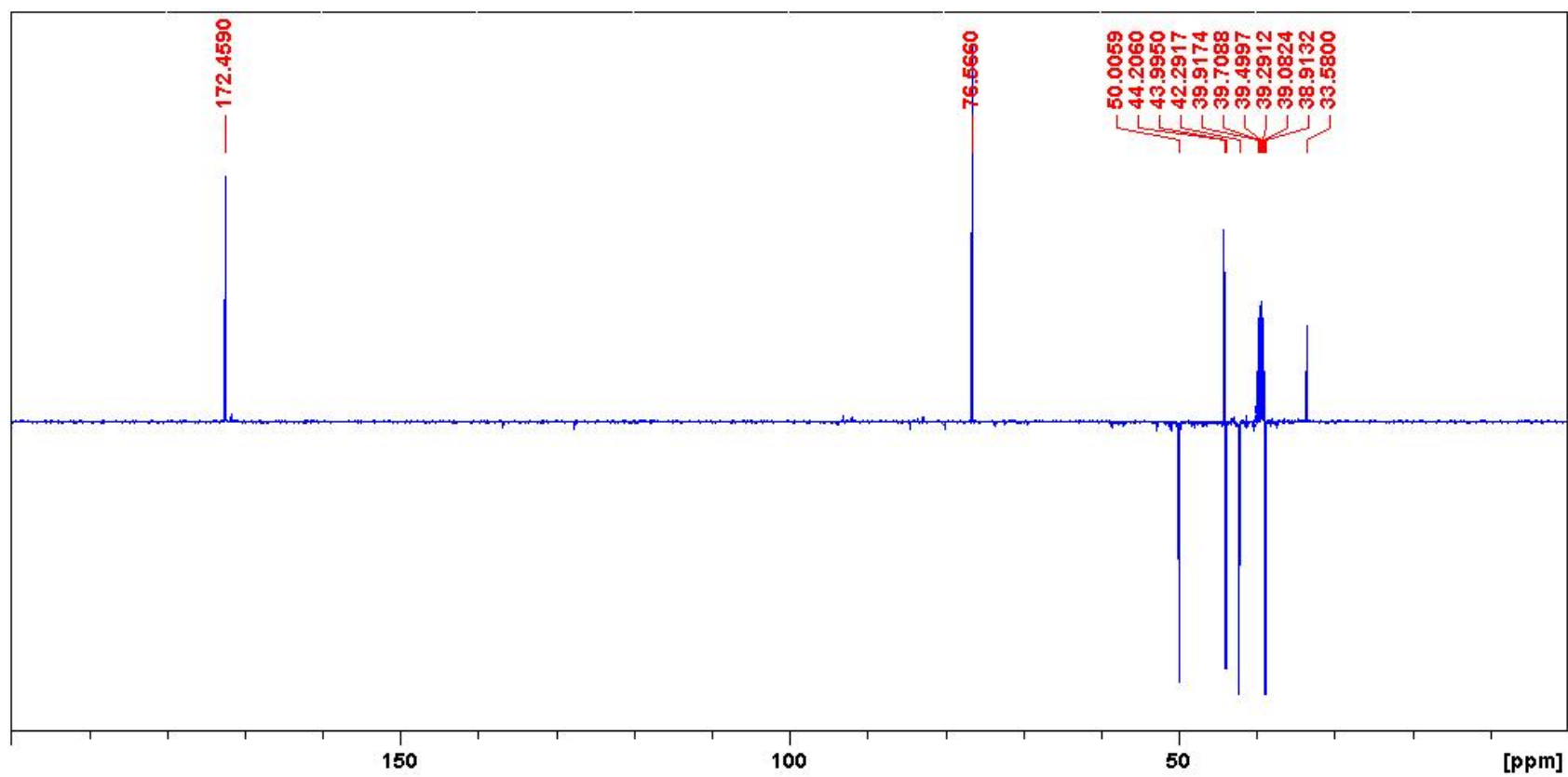
Meas. #	Formula	m/z	err [ppm]	Mean err [ppm]	rdb	N-Rule	e ⁻ Conf	mSigma	Std I	Std Mean m/z	Std I VarNo	Std m/z Diff	Std Comb Dev
		322.0482											
		622.0285											
		922.0105											
		1521.9711											
		2121.9330											
		2721.8949											
		633.3351											
		1521.9671											

HRMS spectrum of compound **3a**

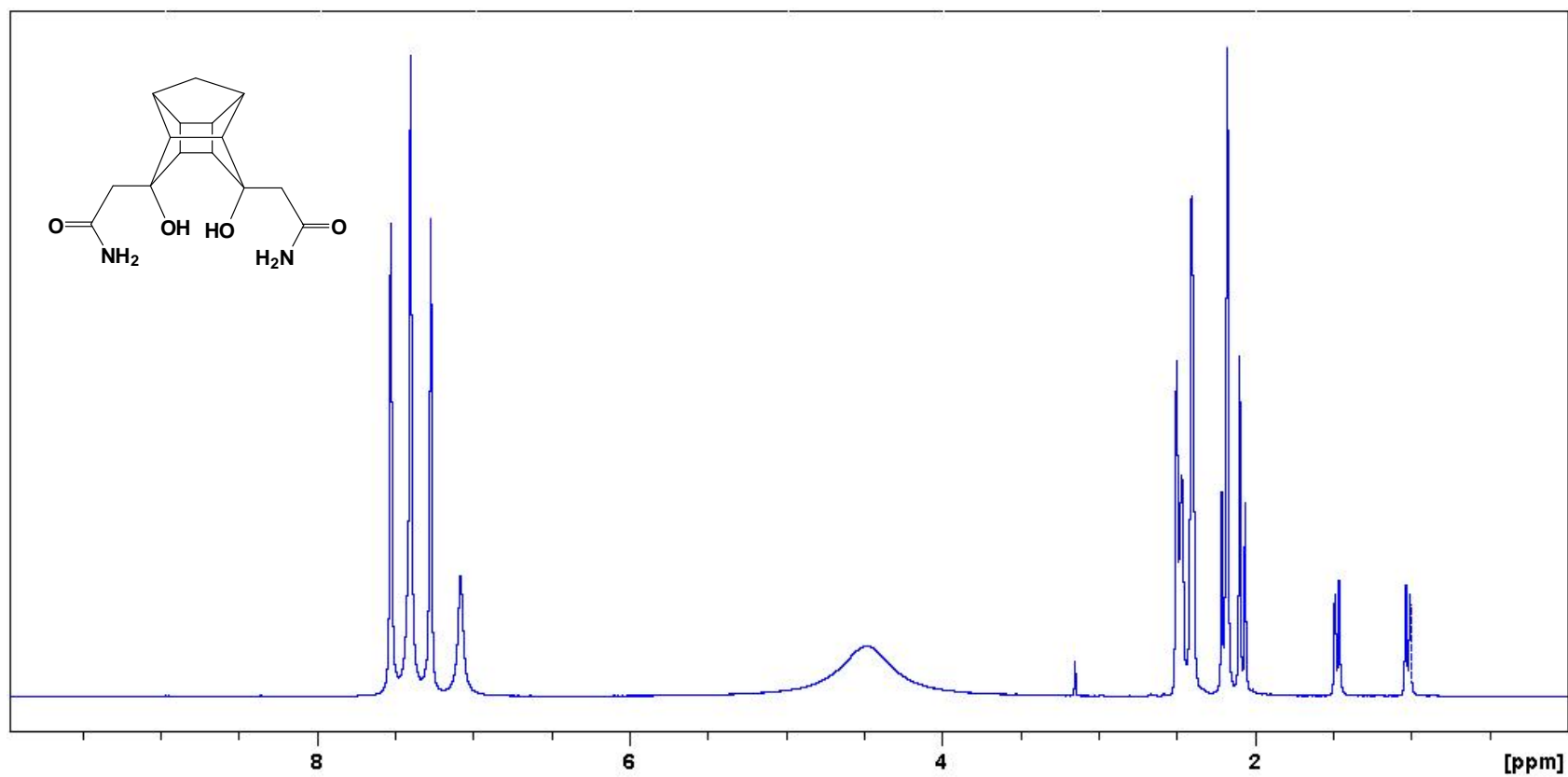


^1H NMR spectrum of Cage diol diacid 1

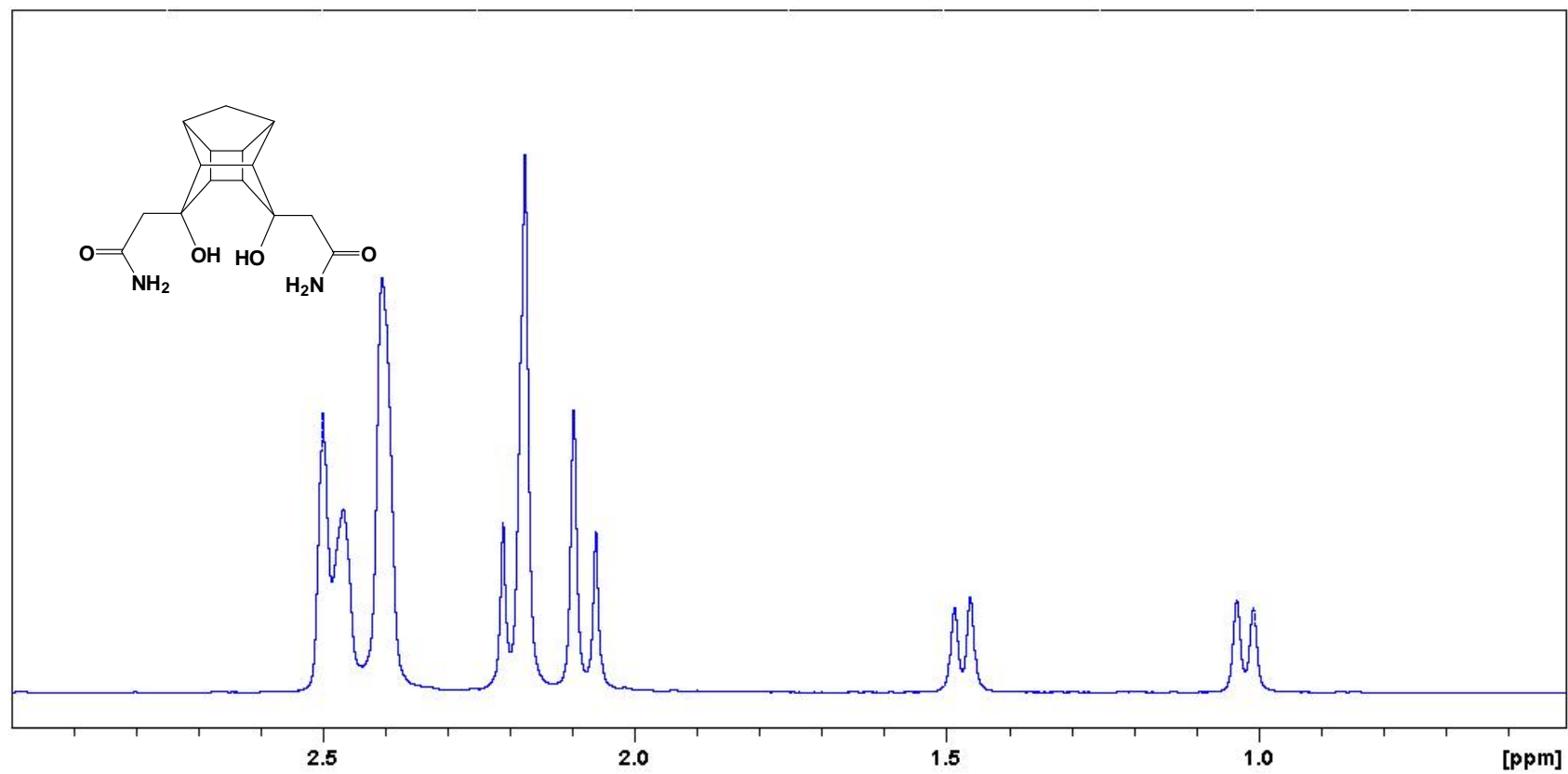
Expanded ^1H NMR spectrum of Cage diol diacid **1**

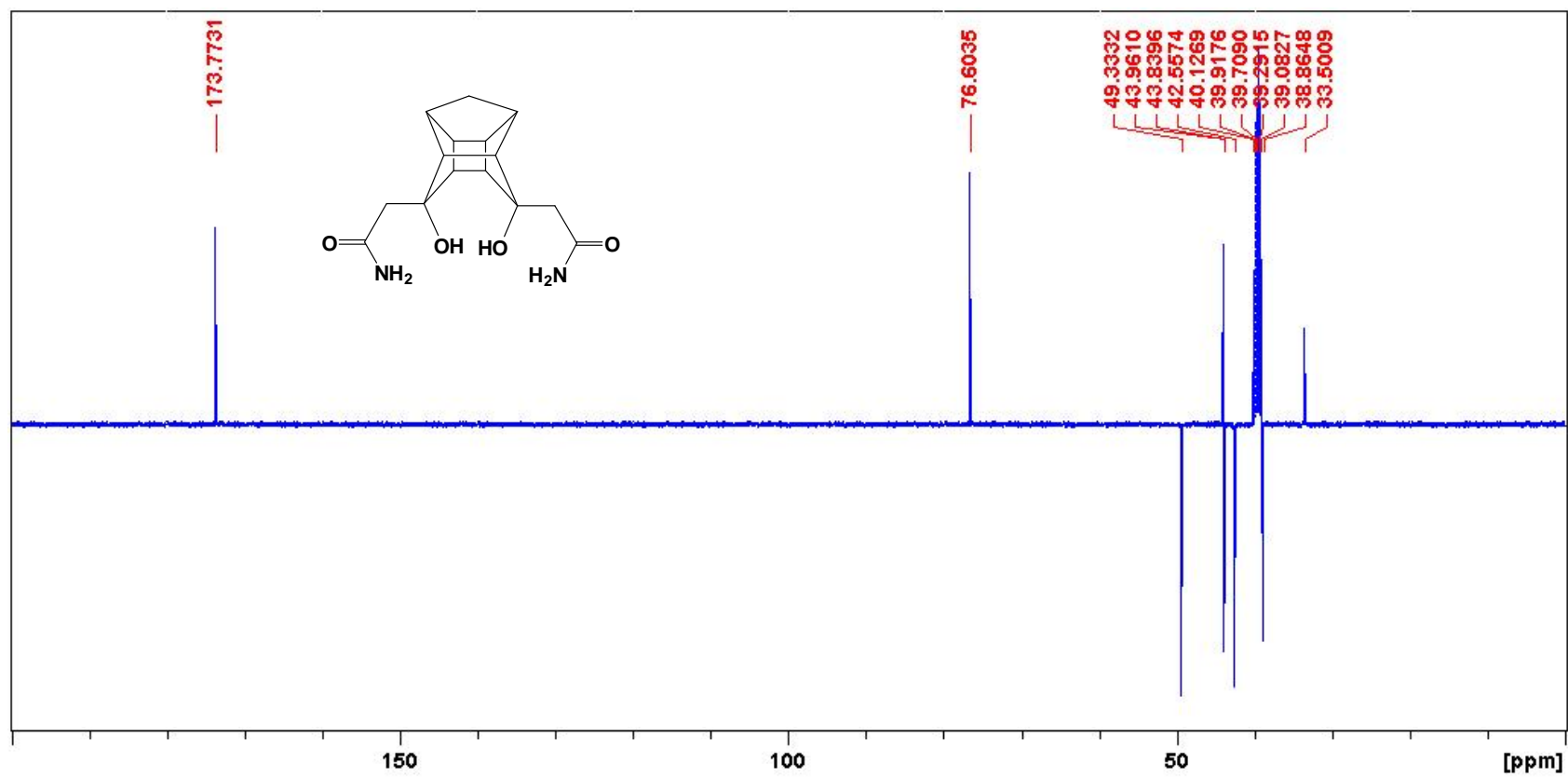


¹³C NMR spectrum of Cage diol diacid 1

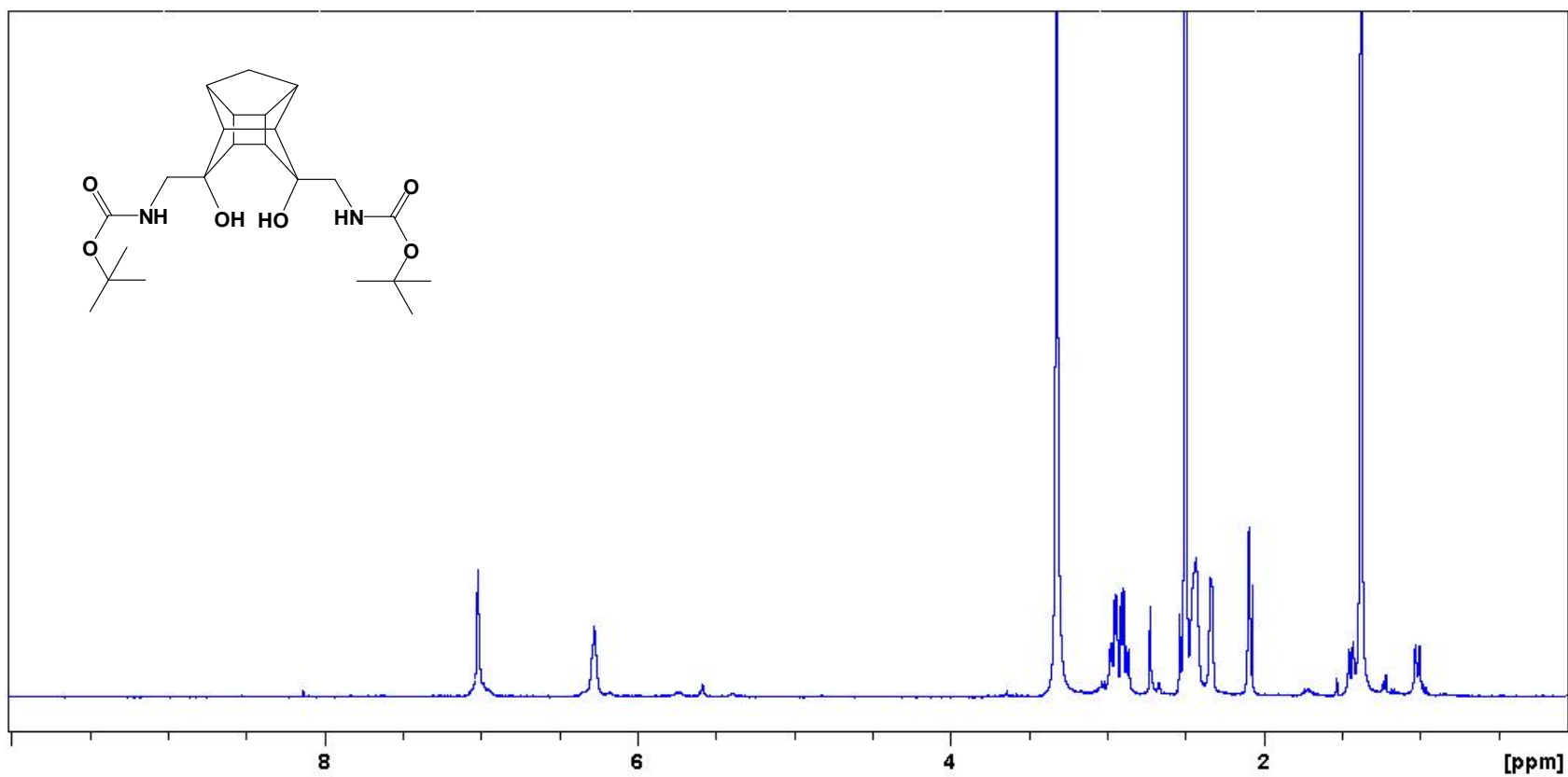


^1H NMR spectrum of Cage diol diamide

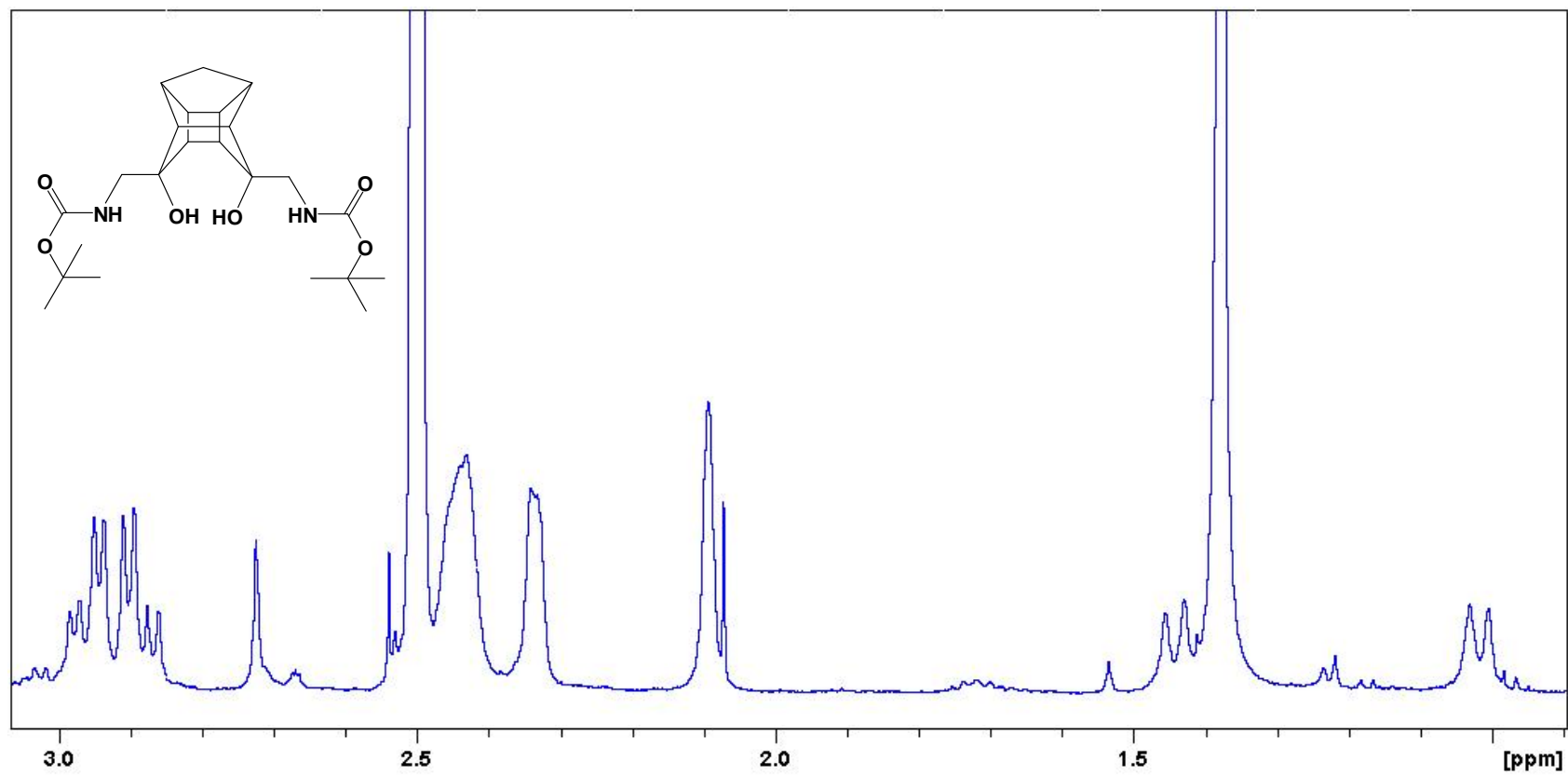
Expanded ^1H NMR spectrum of Cage diol diamide

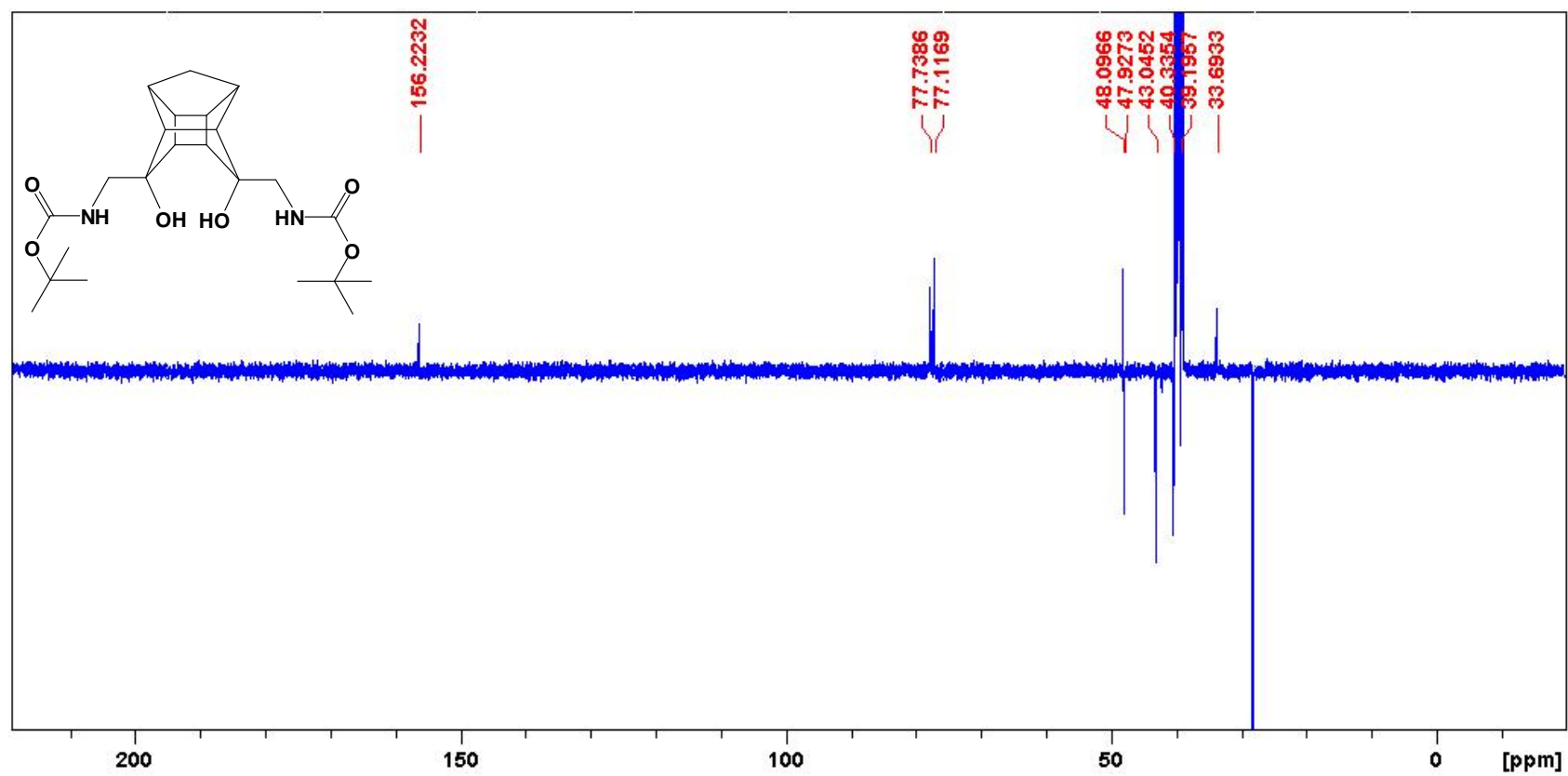


^{13}C NMR spectrum of Cage diol diamide

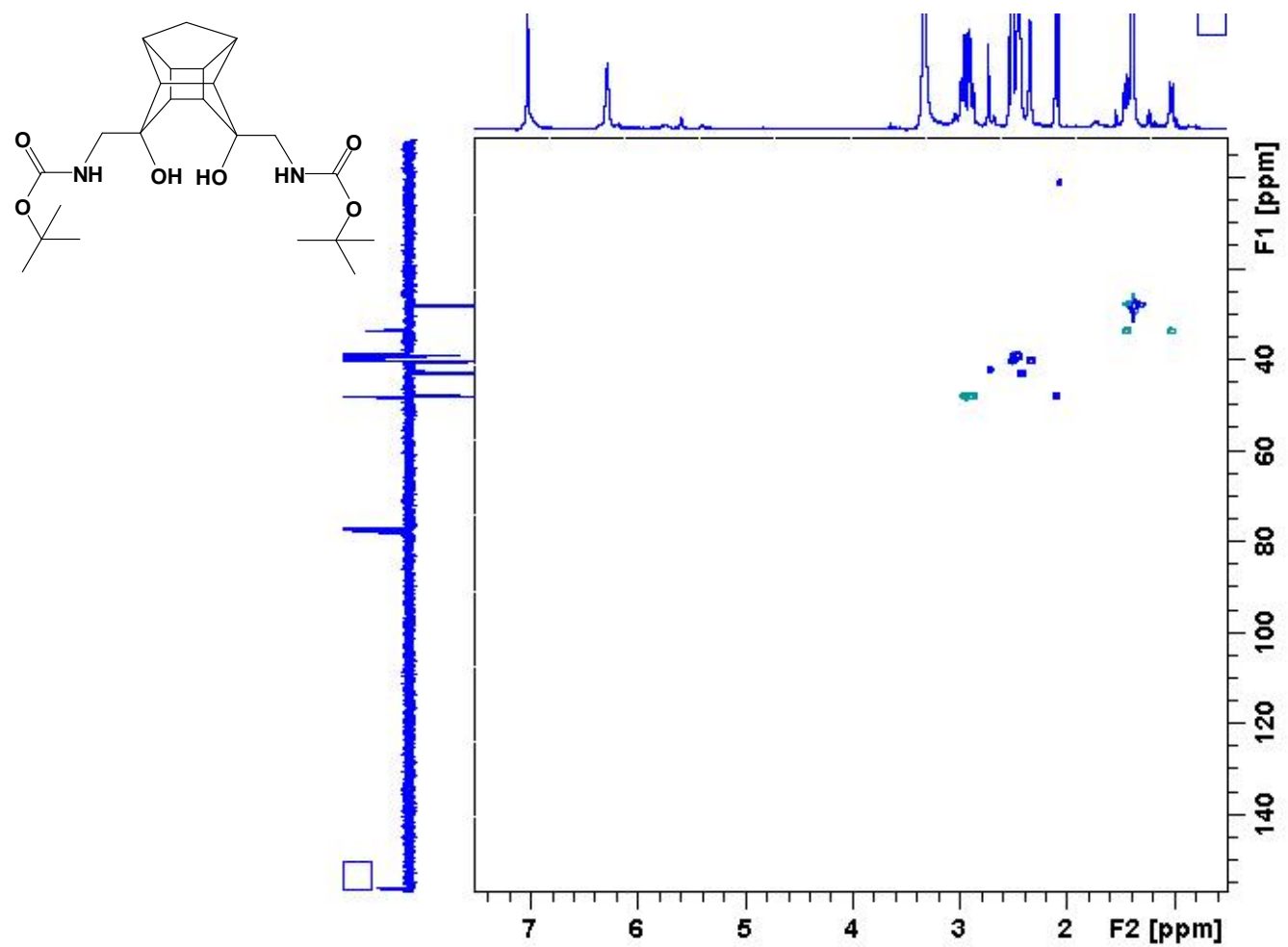


¹H NMR spectrum of BOC-cage diol diamine

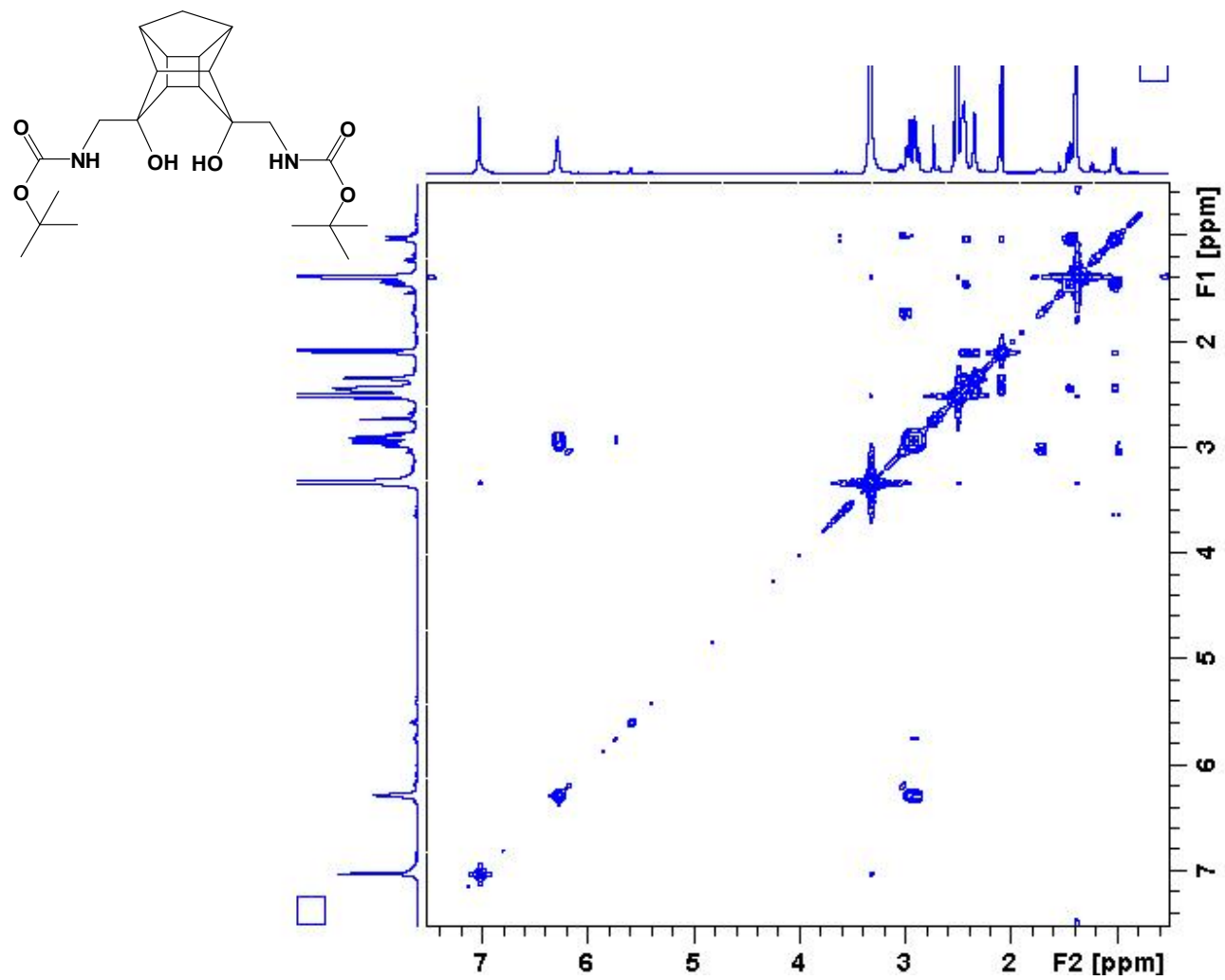
Expanded ^1H NMR spectrum of BOC-cage diol diamine



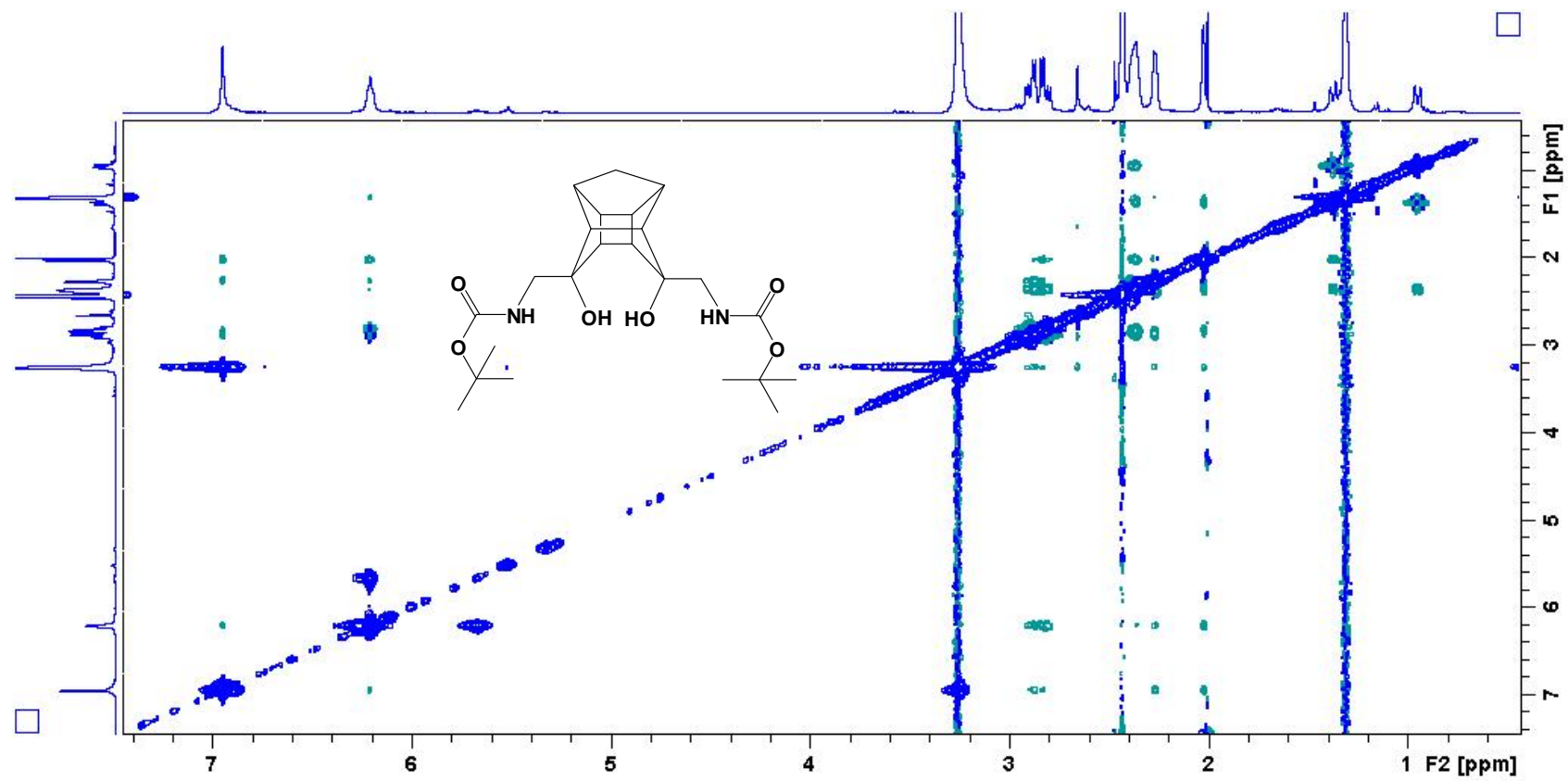
^{13}C NMR spectrum of BOC-cage diol diamine



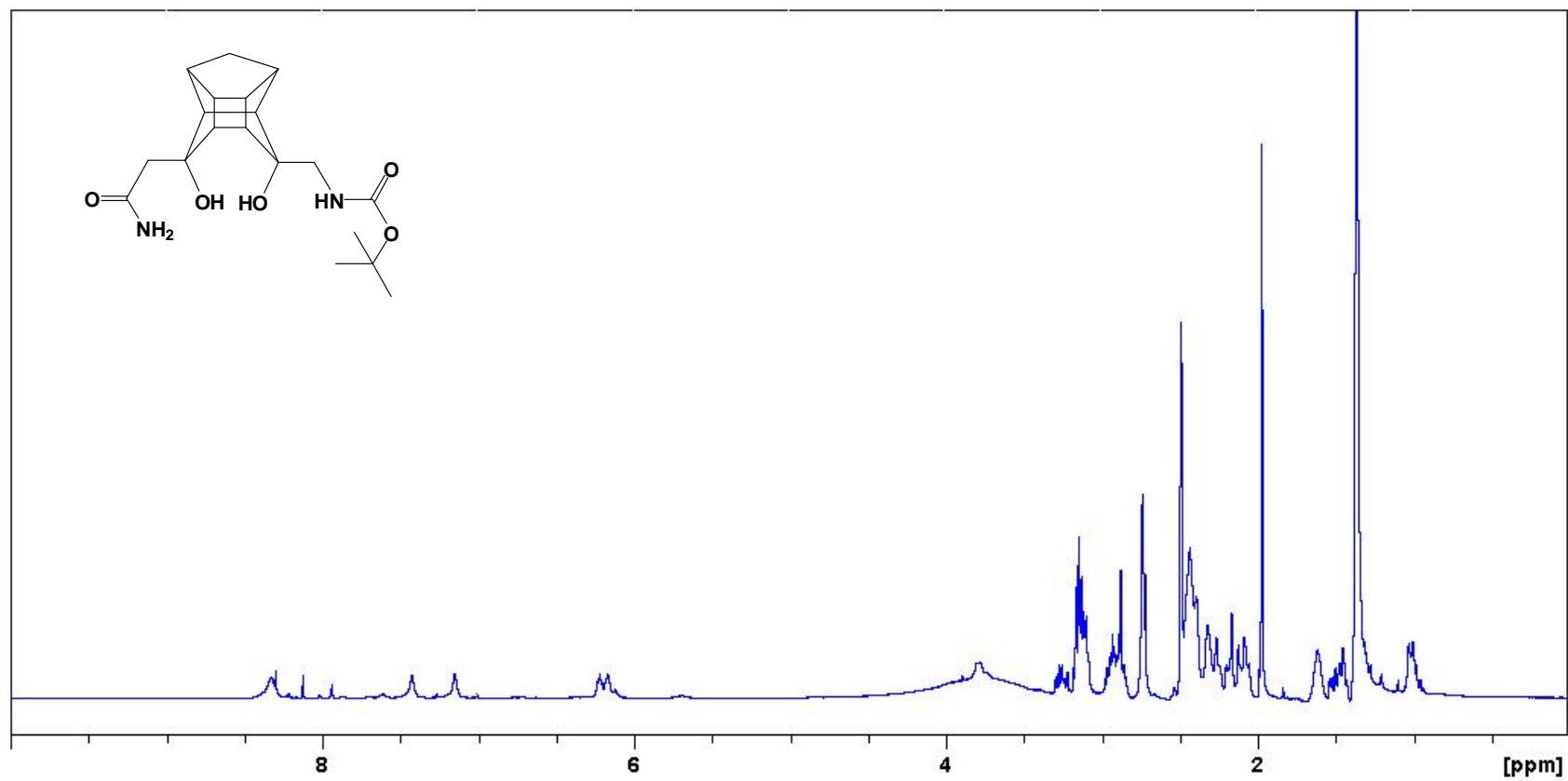
HSQC spectrum of BOC-cage diol diamine



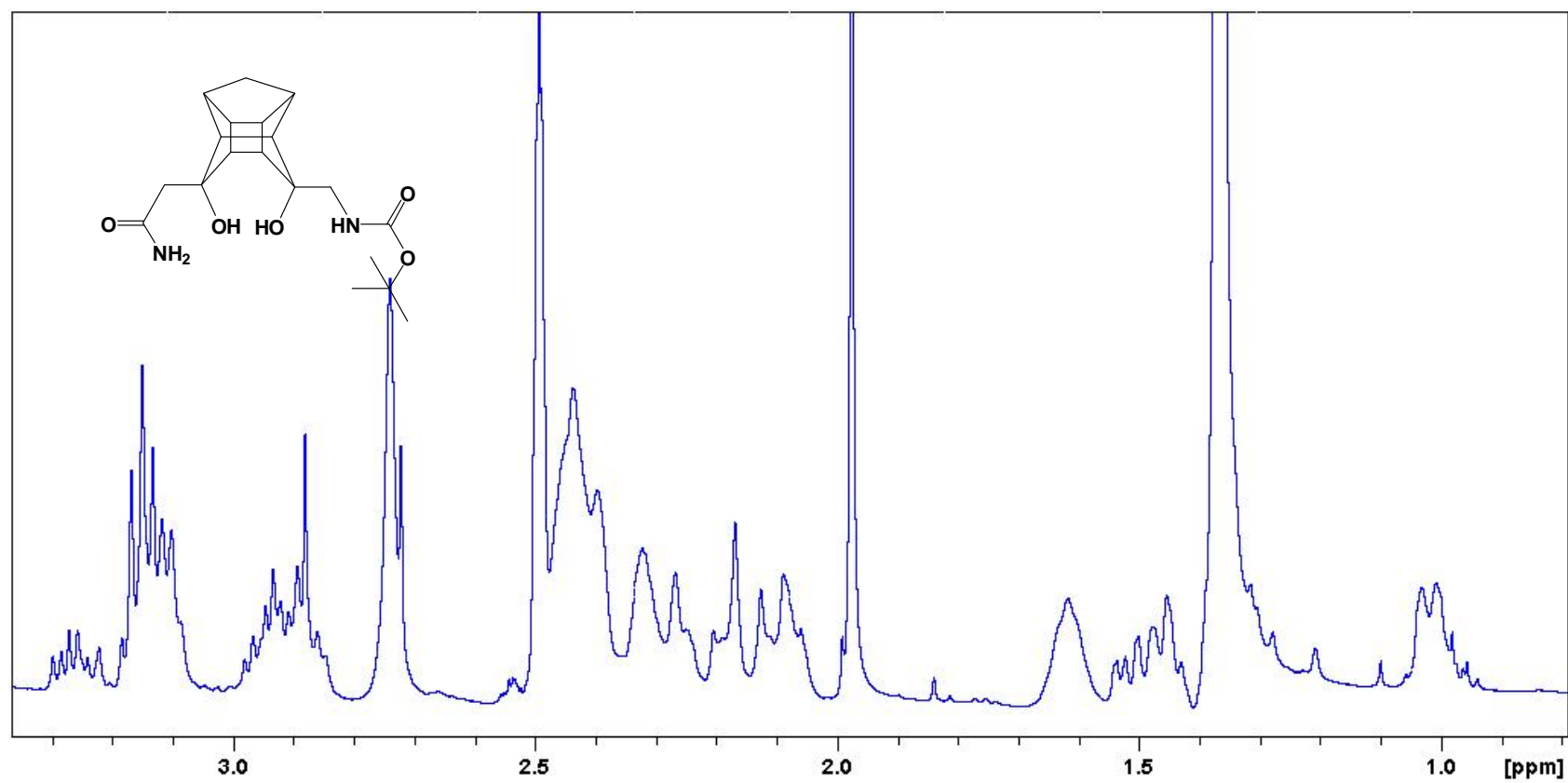
COSY spectrum of BOC-cage diol diamine

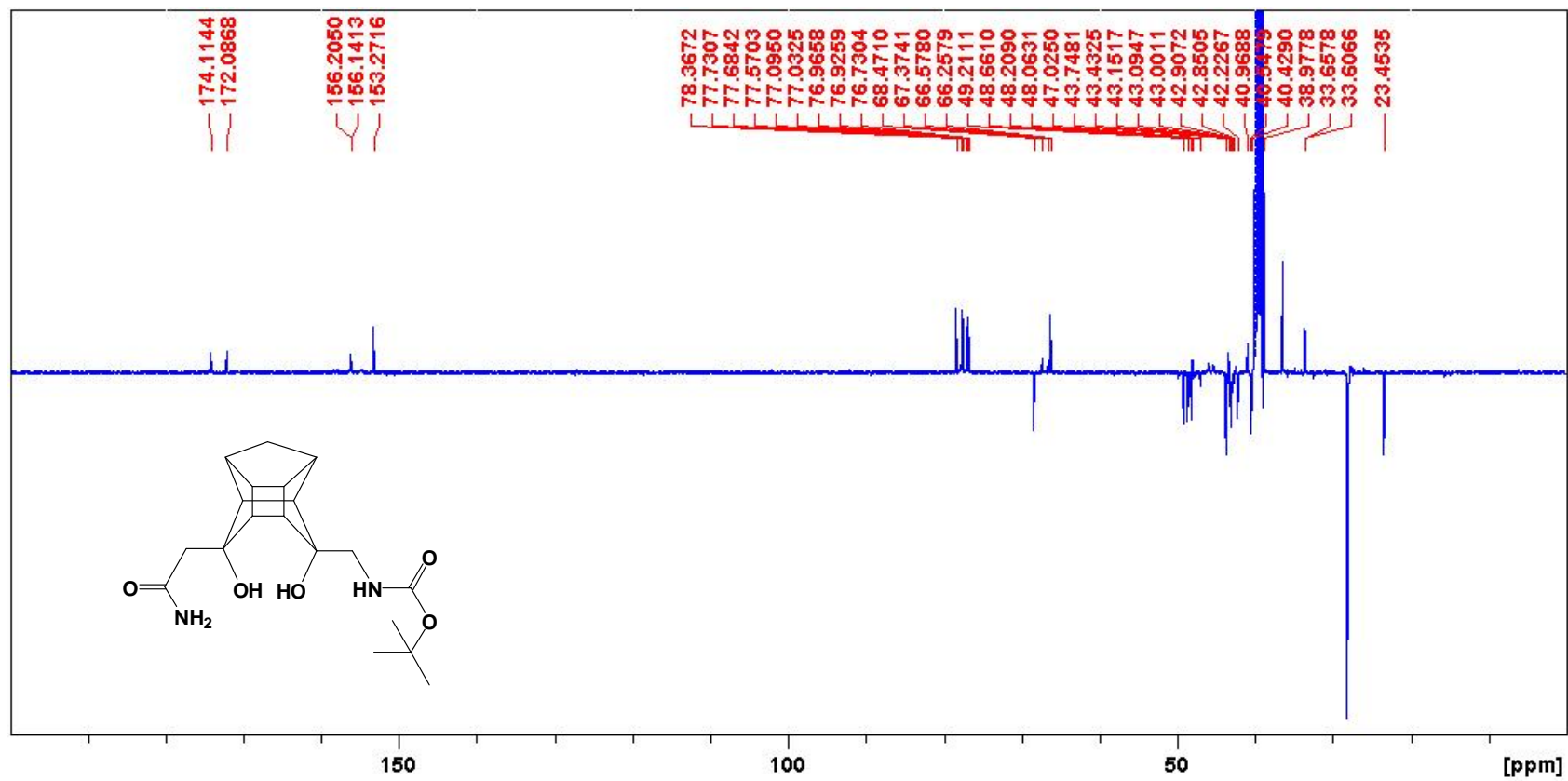


NOESY spectrum of BOC-cage diol diamine

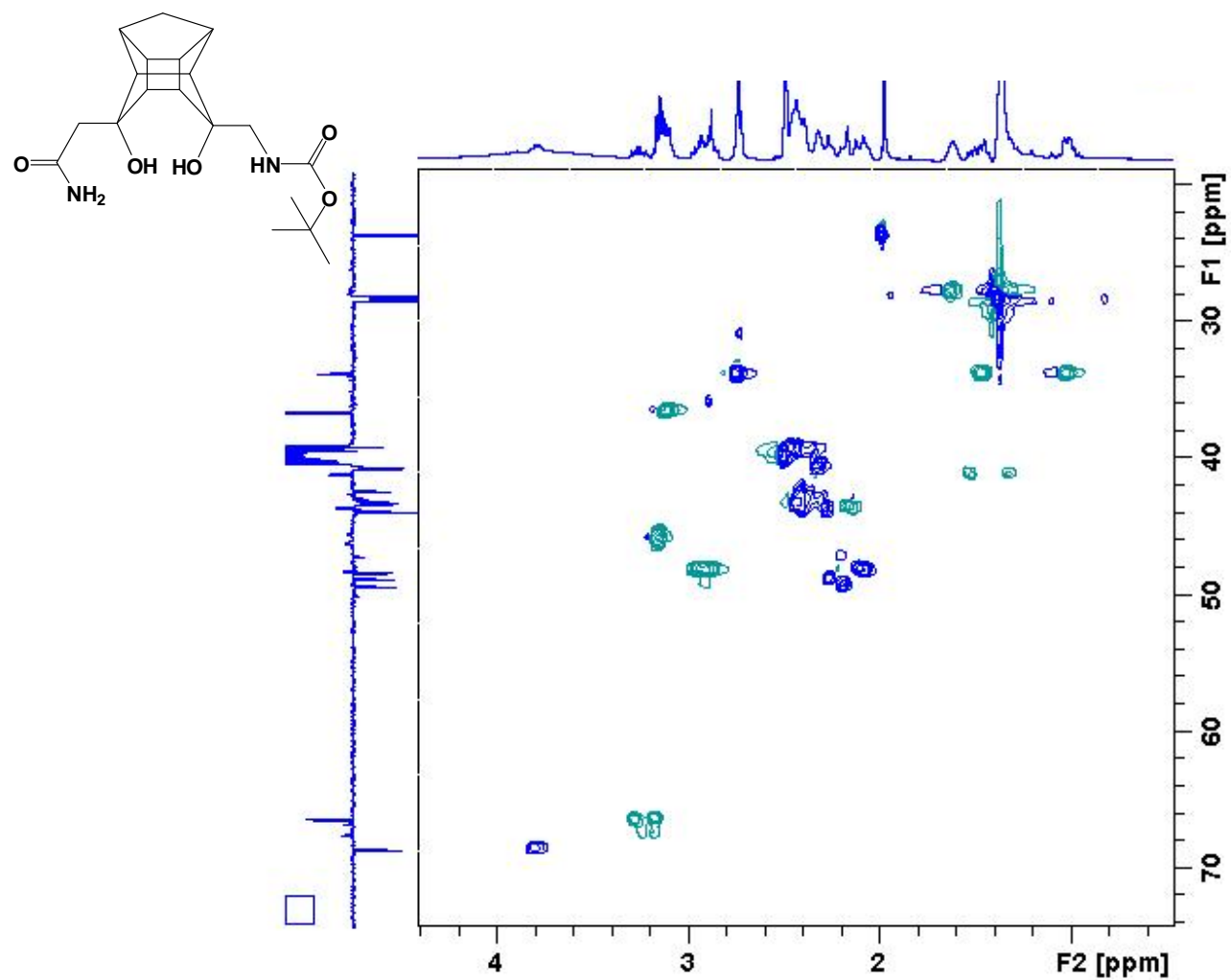


¹H NMR spectrum of BOC-cage diol amine amide

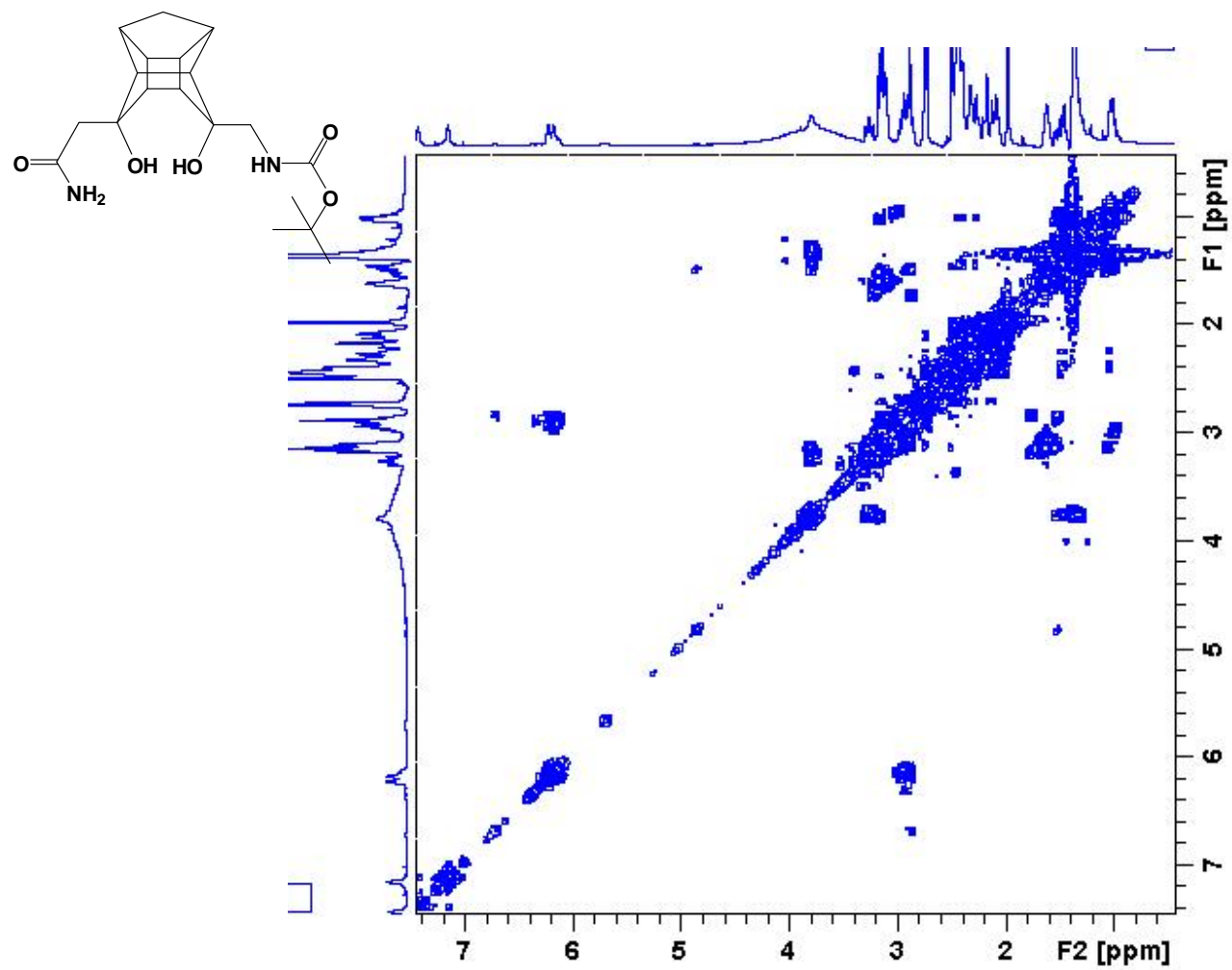




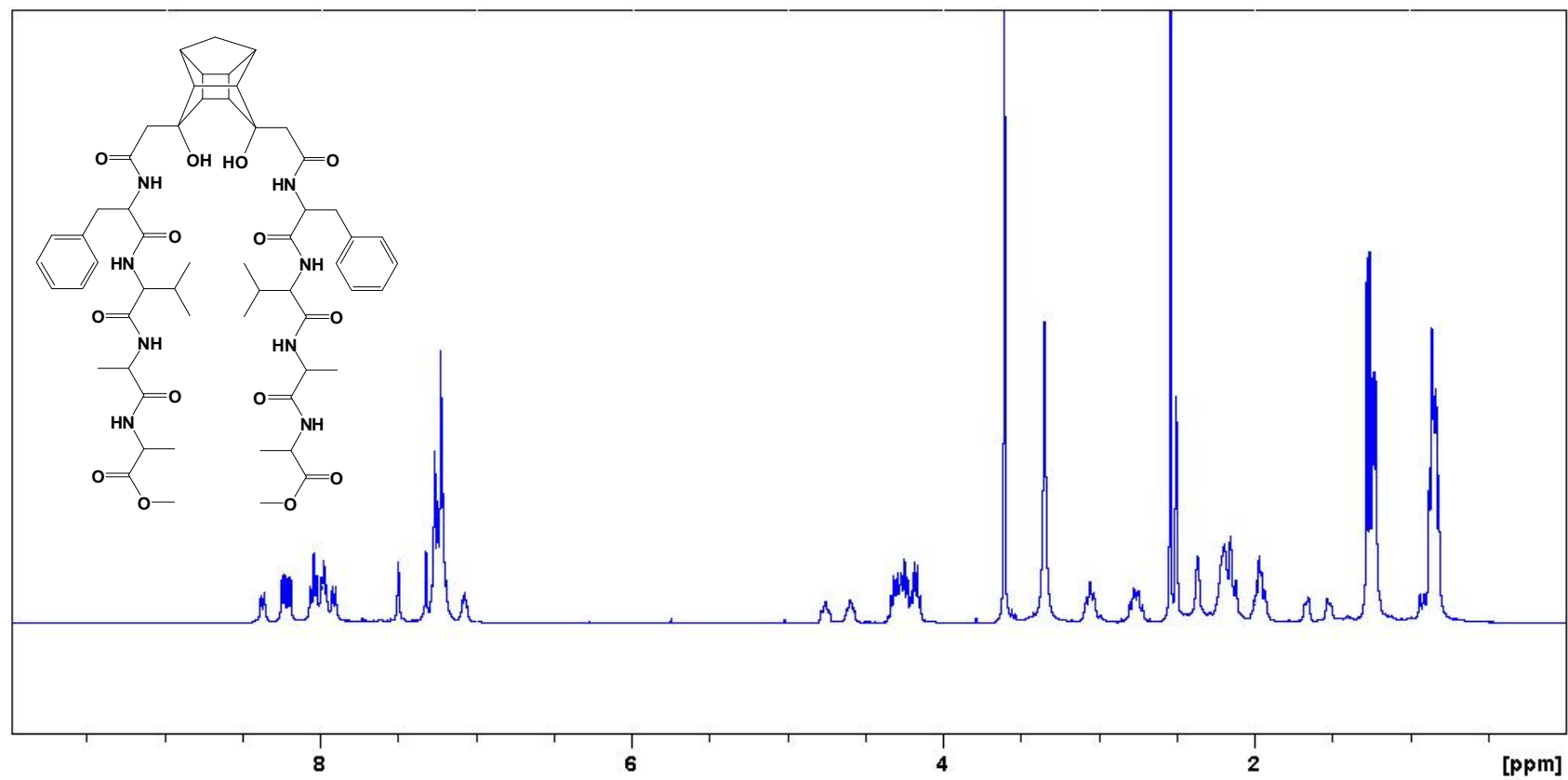
^{13}C NMR spectrum of BOC-cage diol amine amide



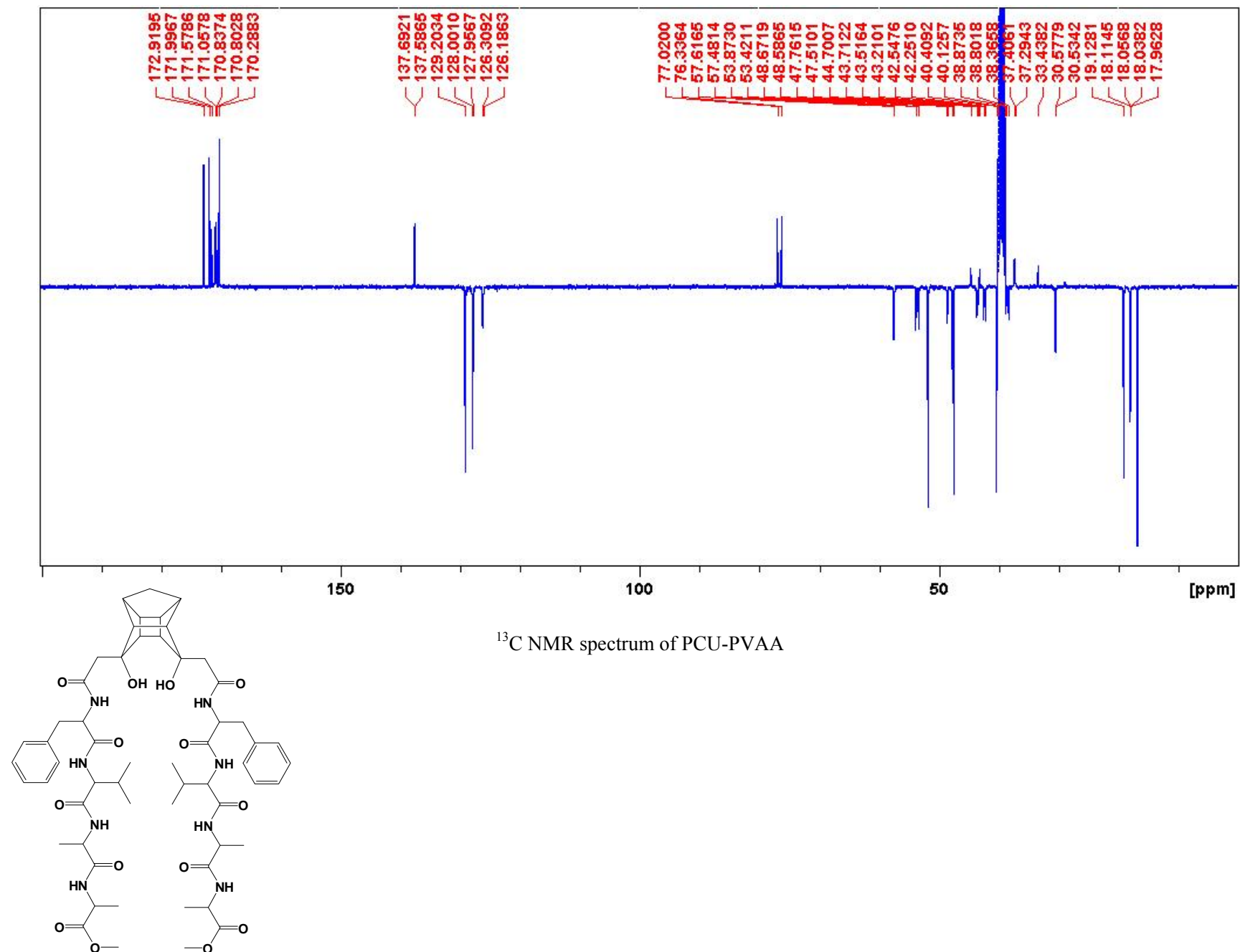
HSQC spectrum of BOC-cage diol amine amide

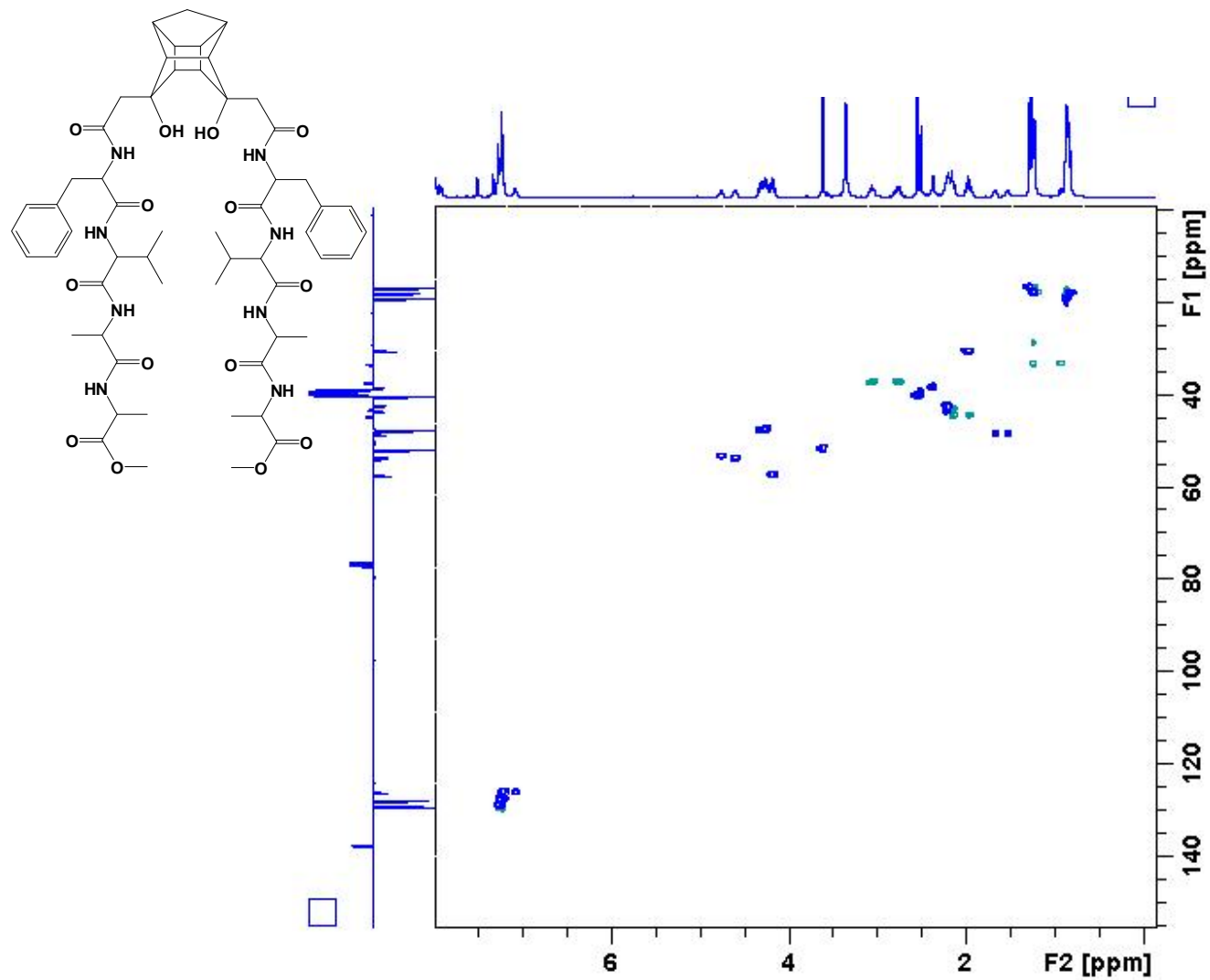


COSY spectrum of BOC-cage diol amine amide

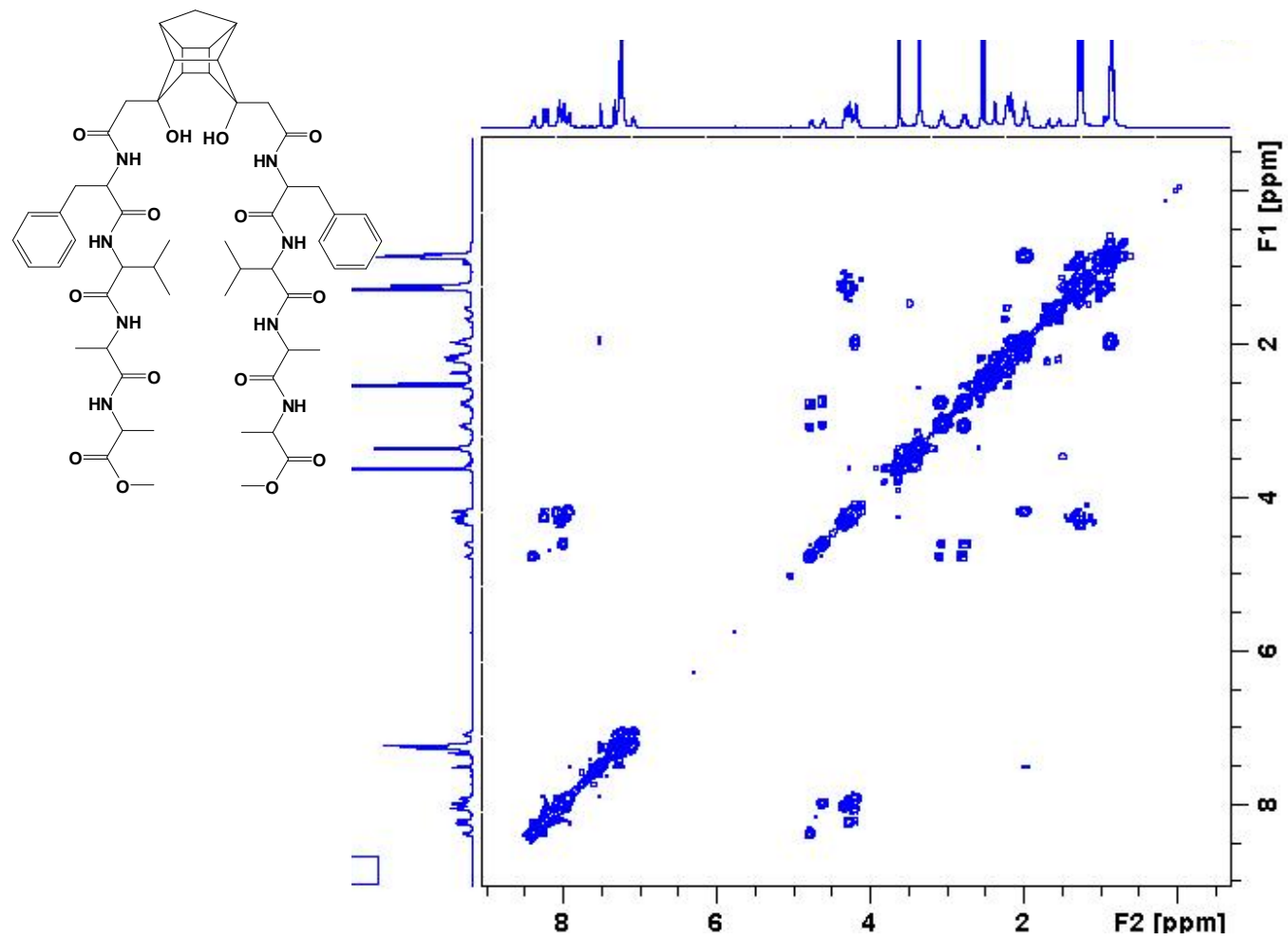


^1H NMR spectrum of PCU-PVAA

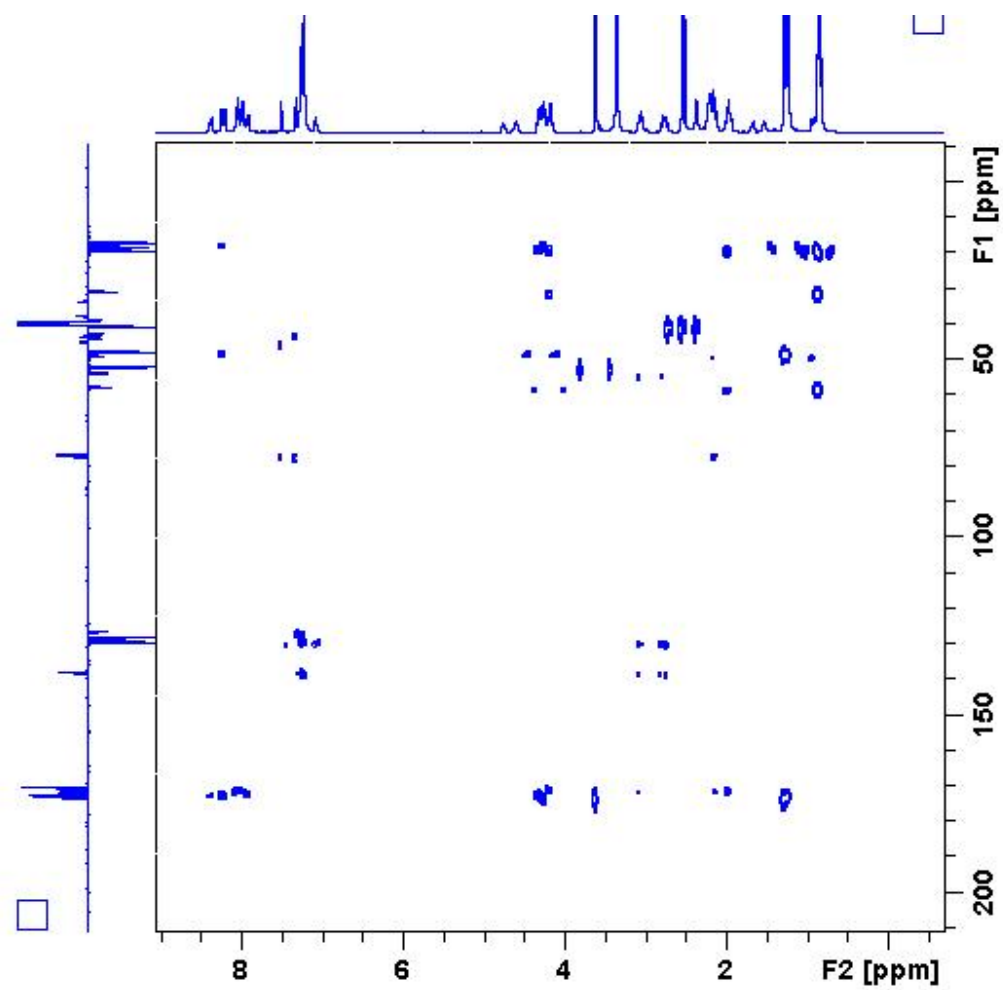




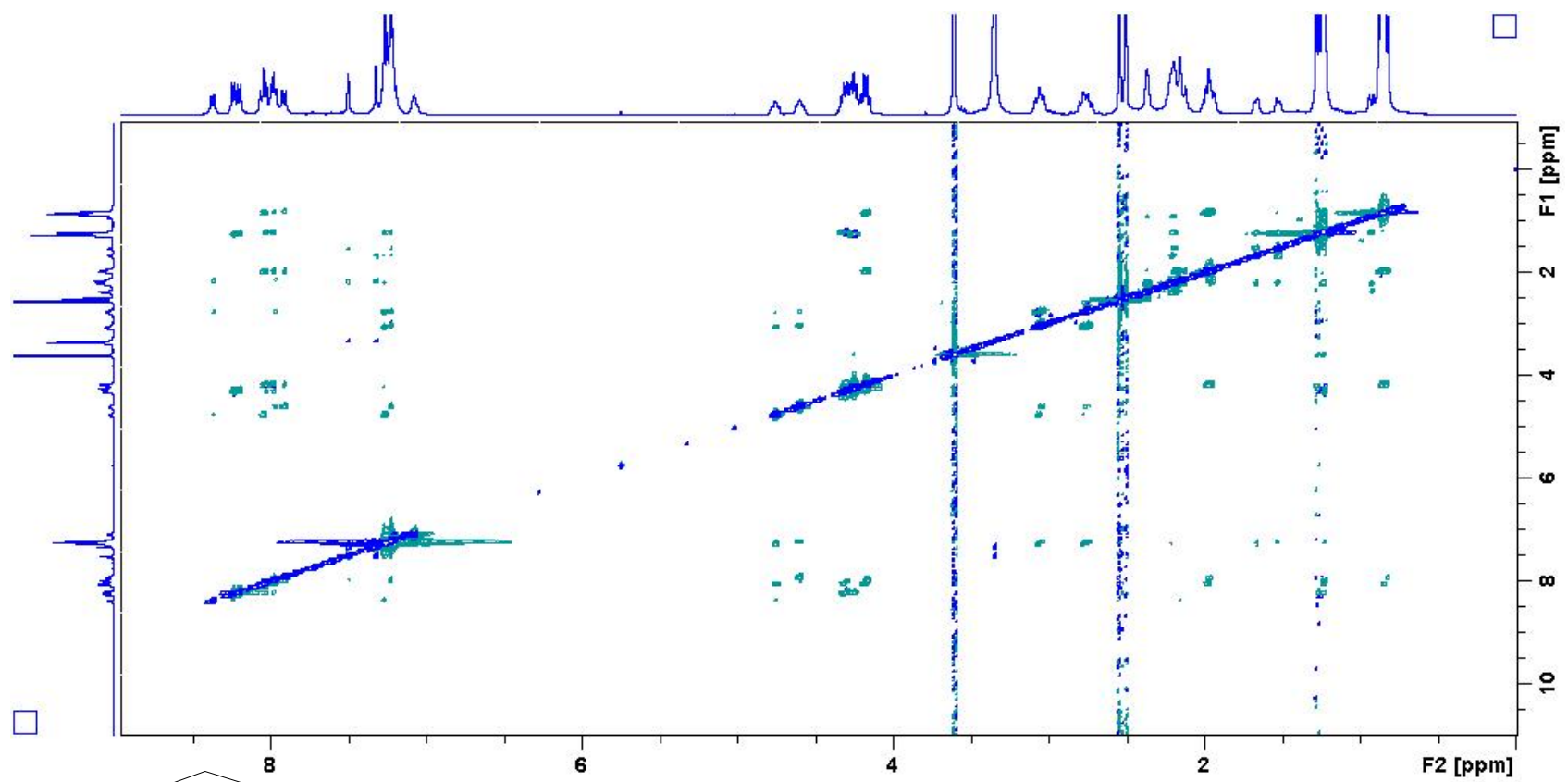
HSQC spectrum of PCU-PVAA



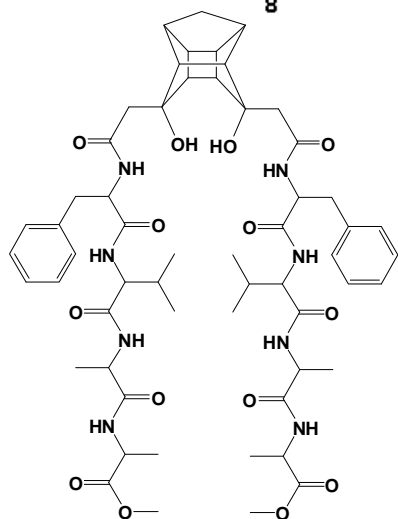
COSY spectrum of PCU-PVAA

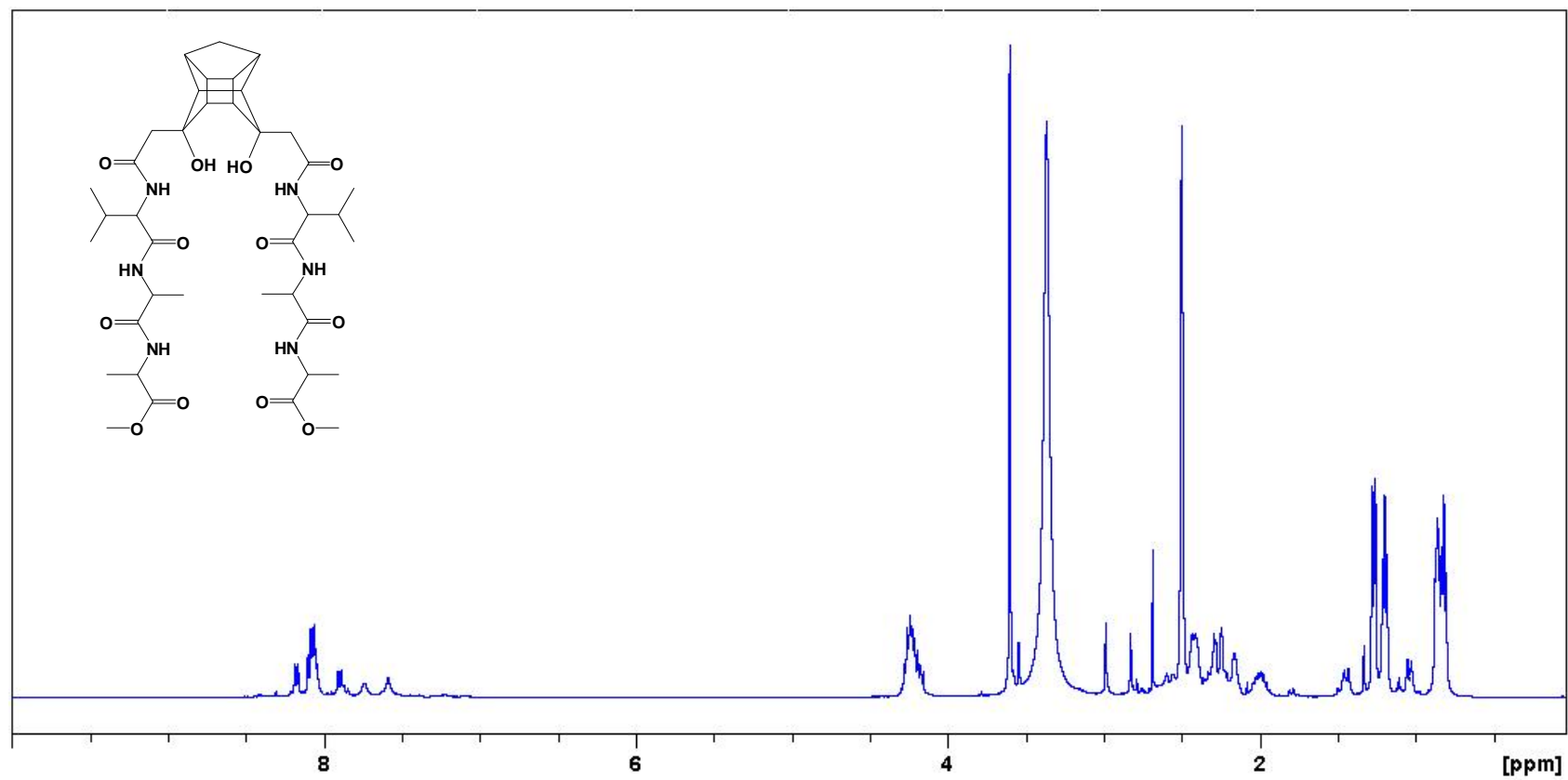


HMBC spectrum of PCU-PVAA

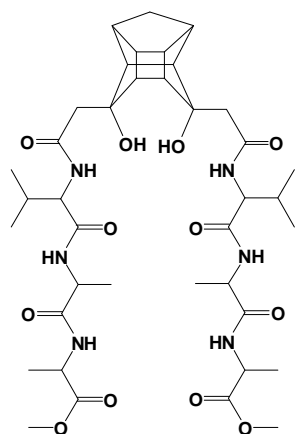
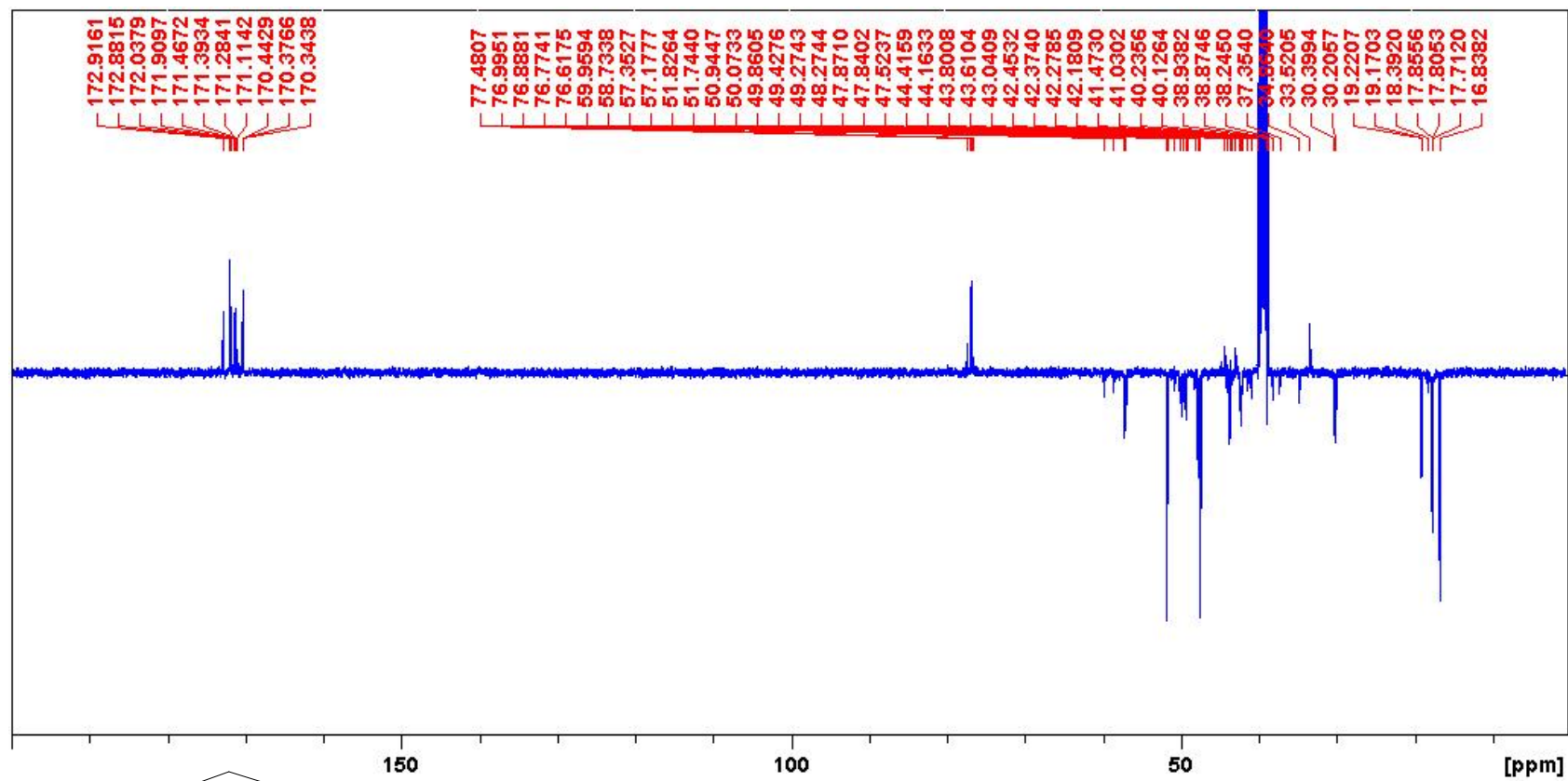


ROESY spectrum of PCU-PVAA

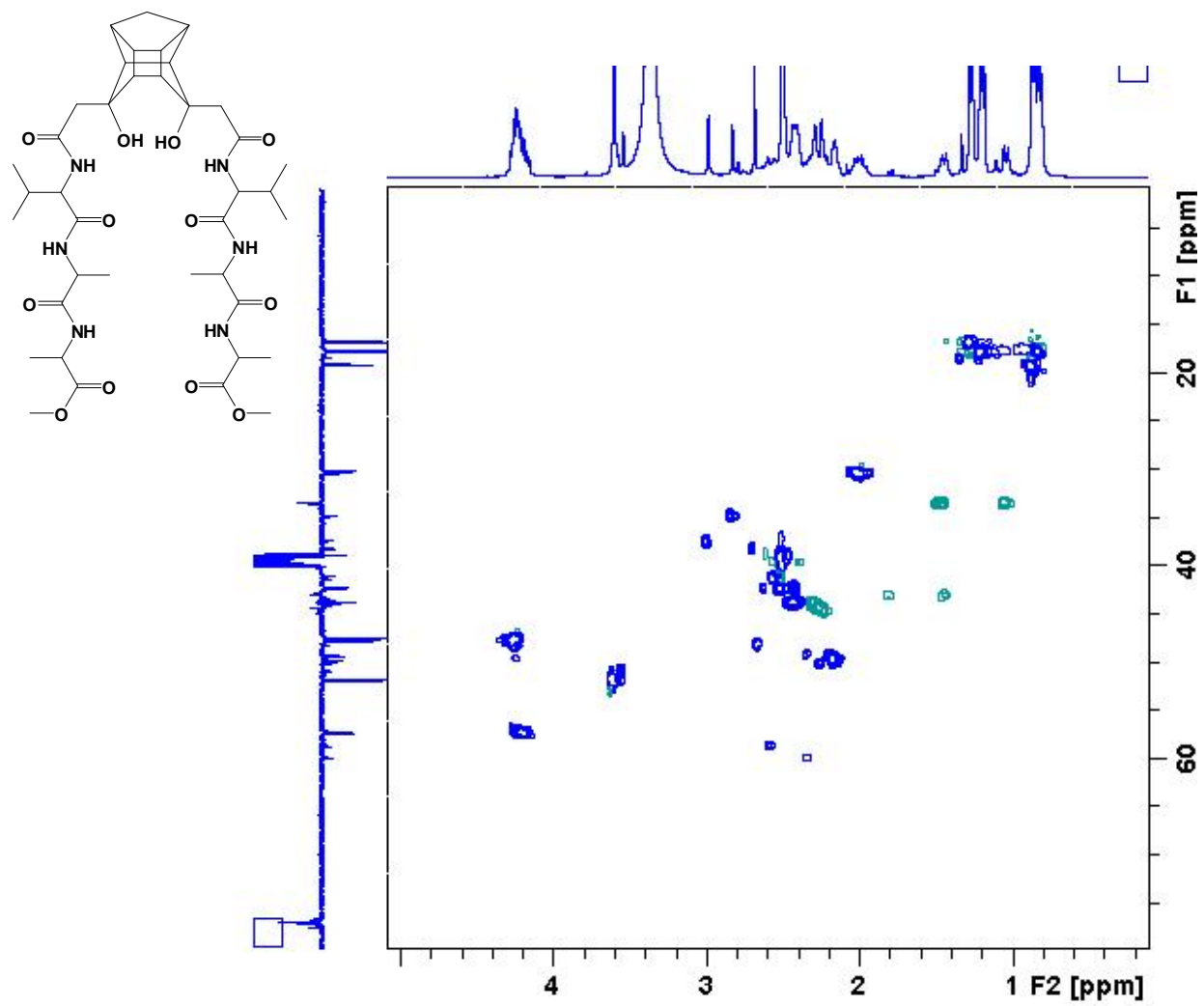




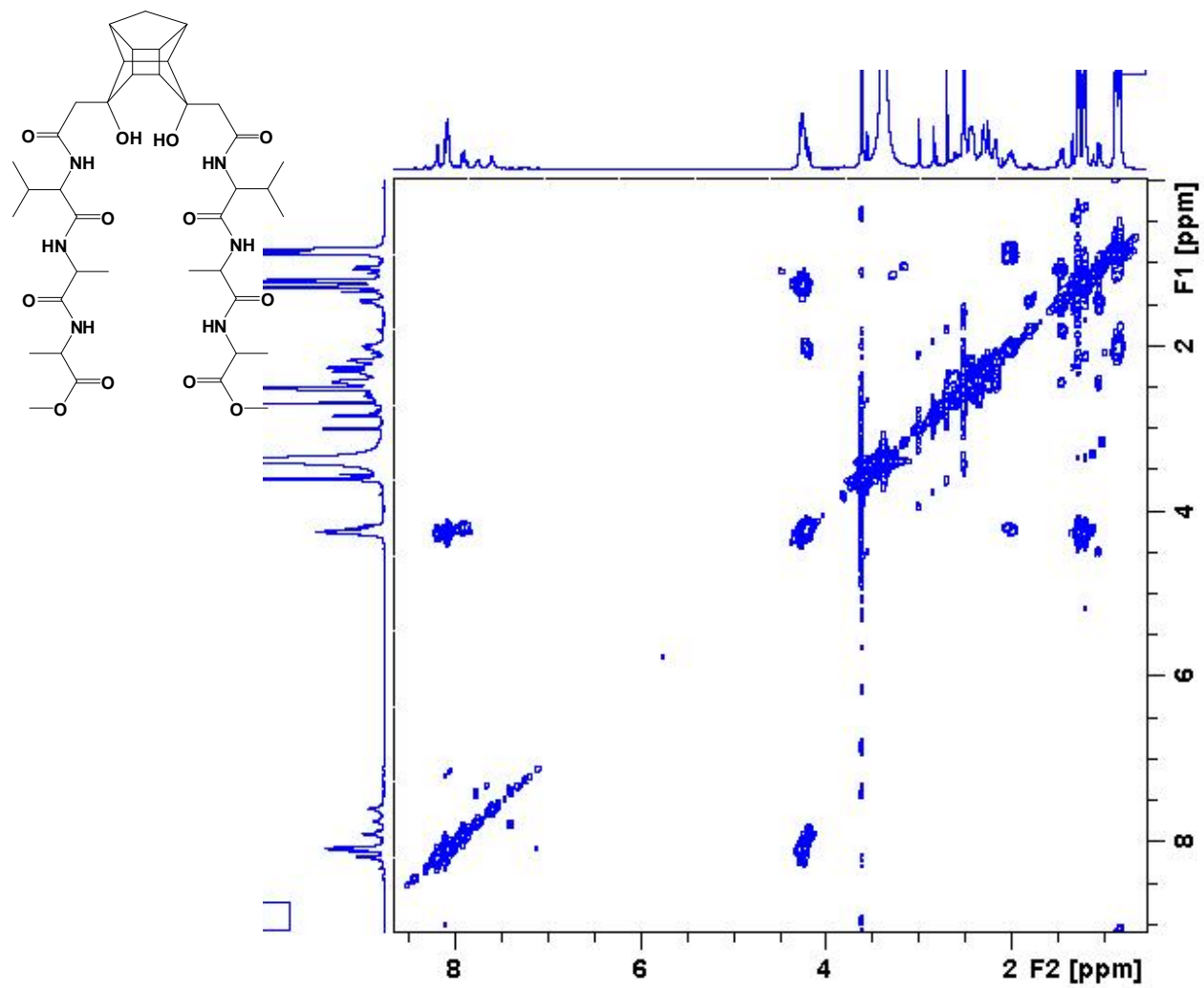
^1H NMR spectrum of PCU-VAA



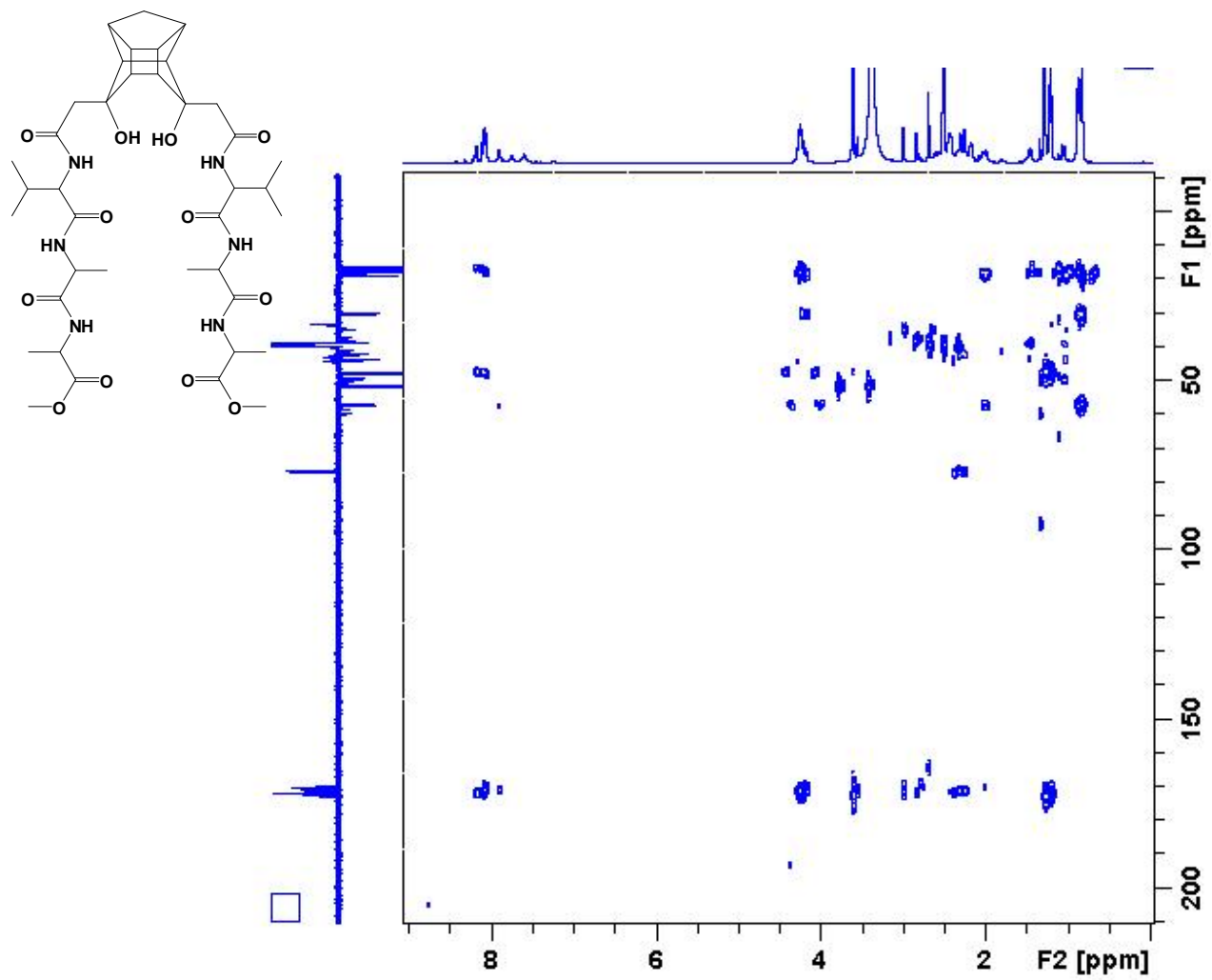
^{13}C NMR spectrum of PCU-VAA



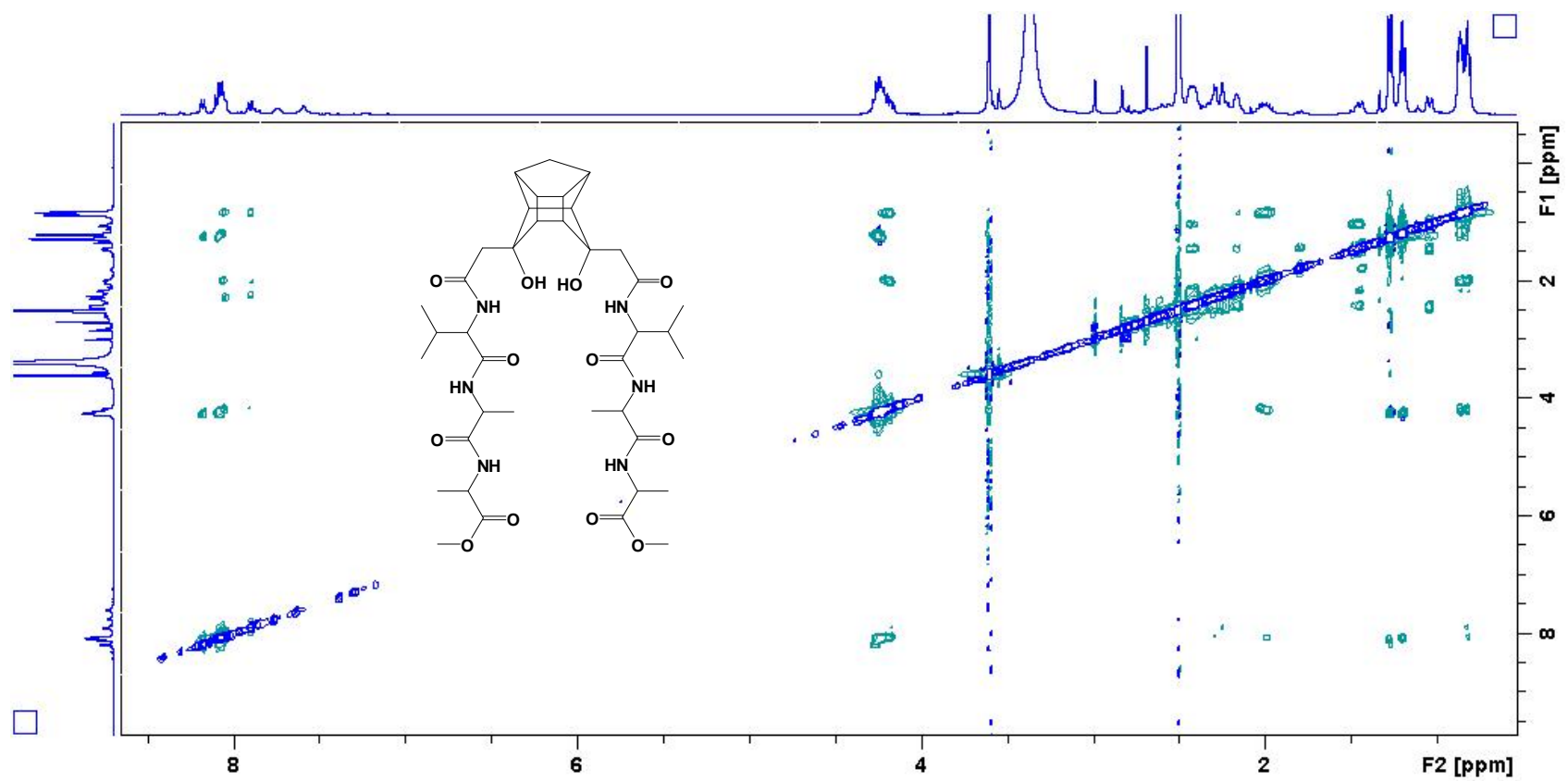
HSQC spectrum of PCU-VAA



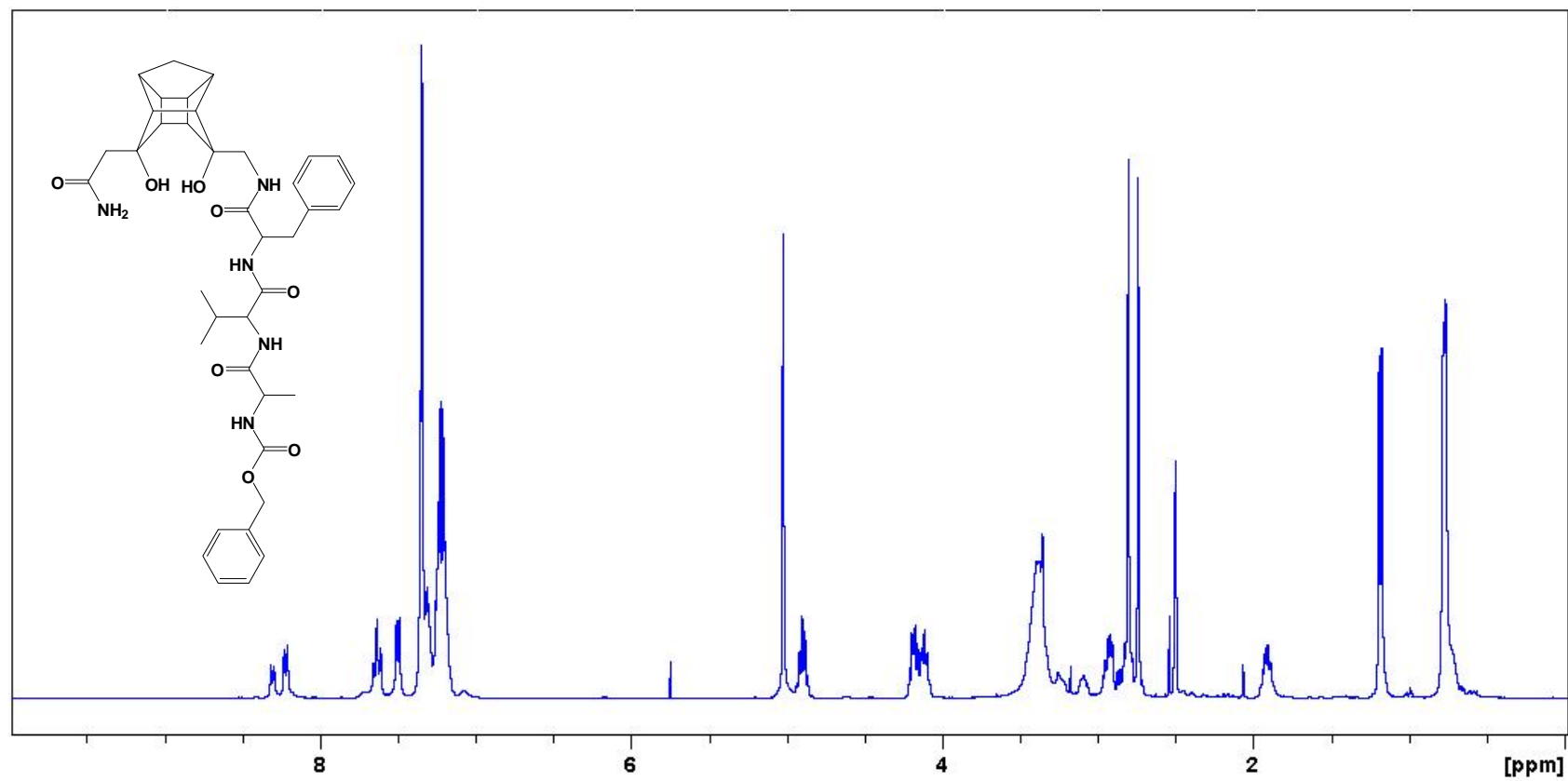
COSY spectrum of PCU-VAA



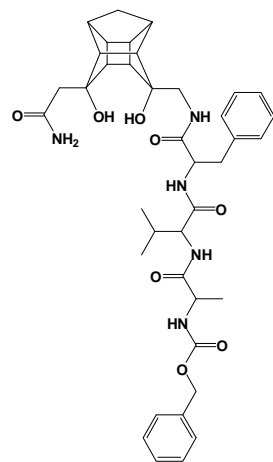
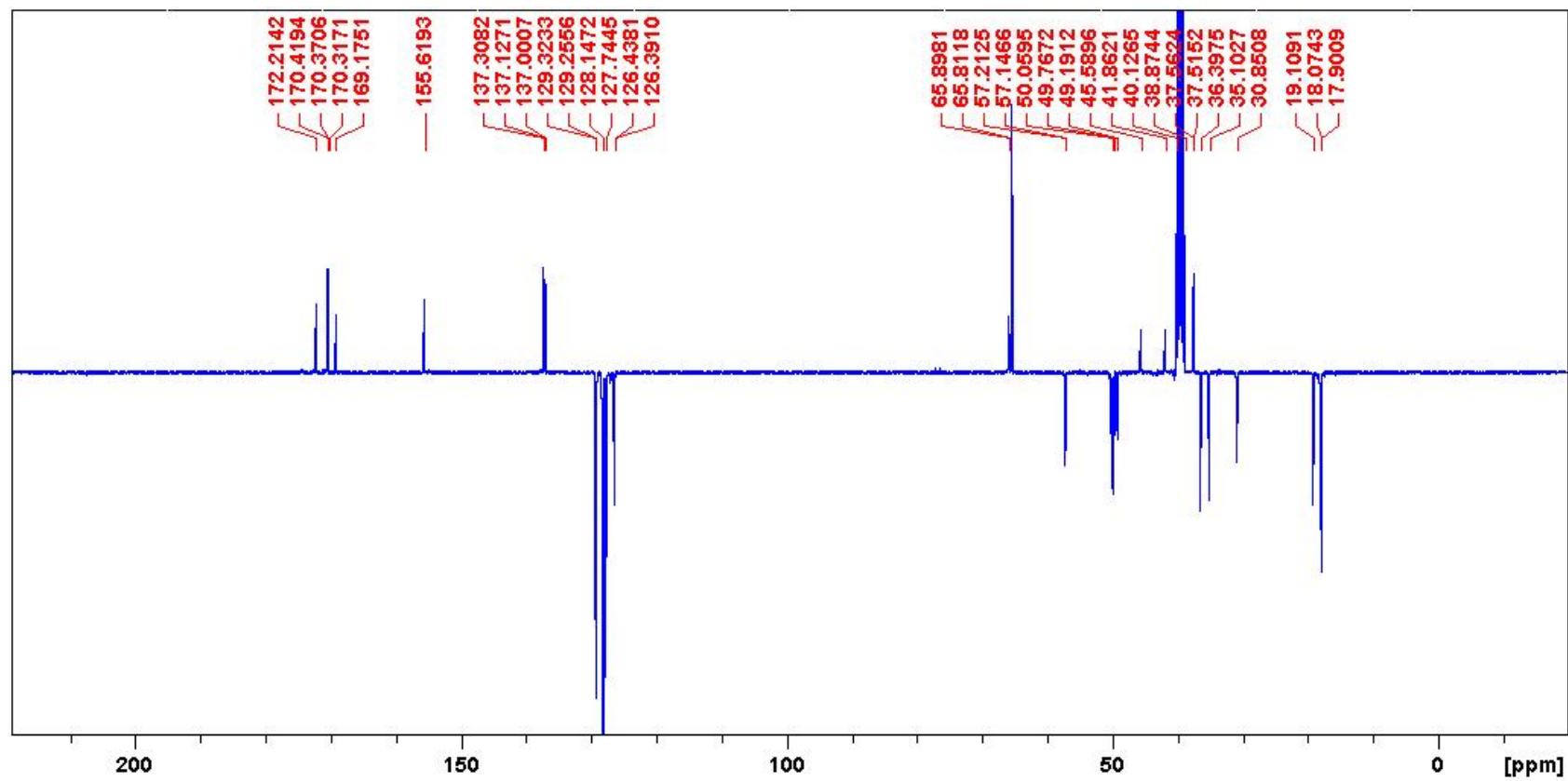
HMBC spectrum of PCU-VAA



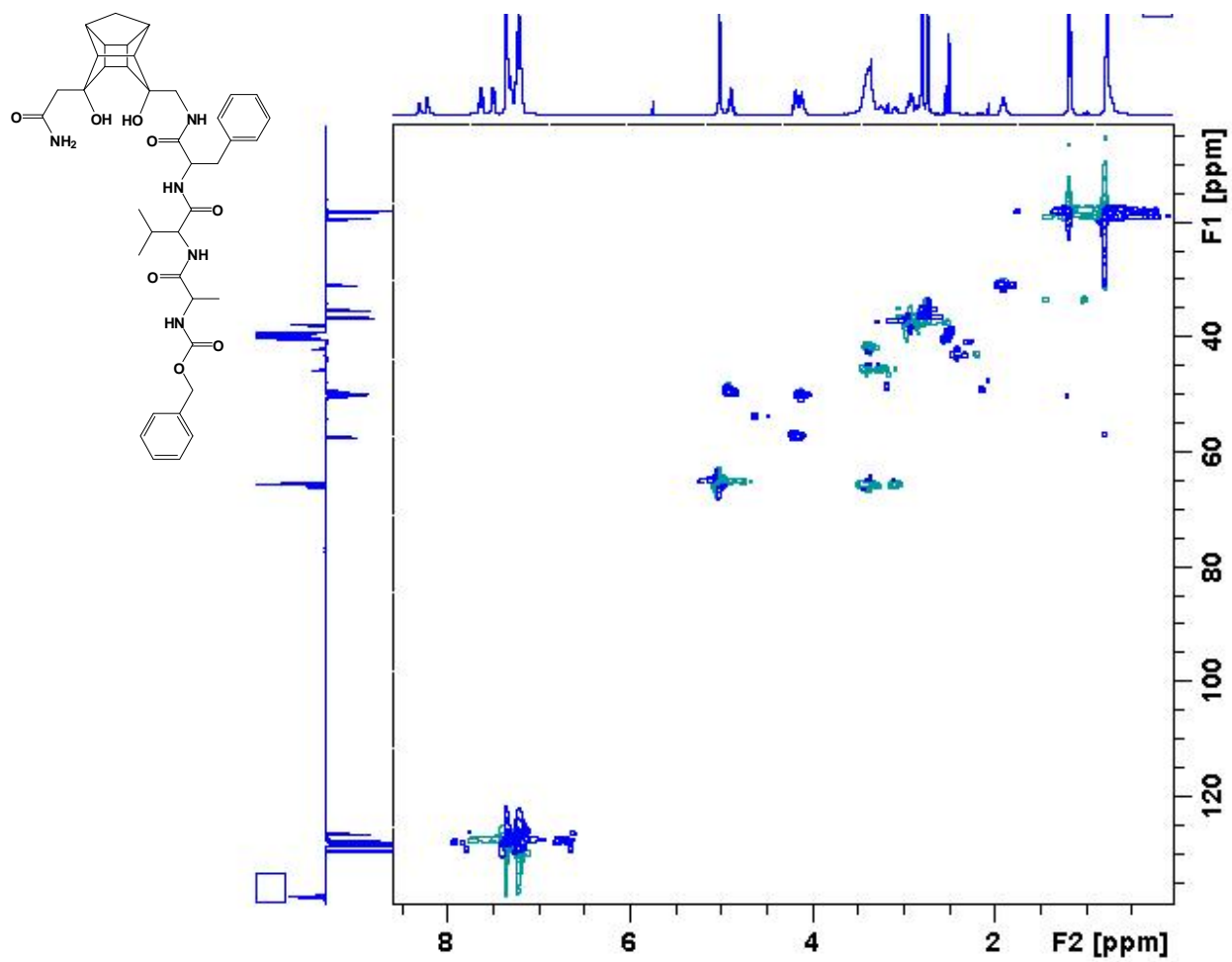
ROESY spectrum of PCU-VAA



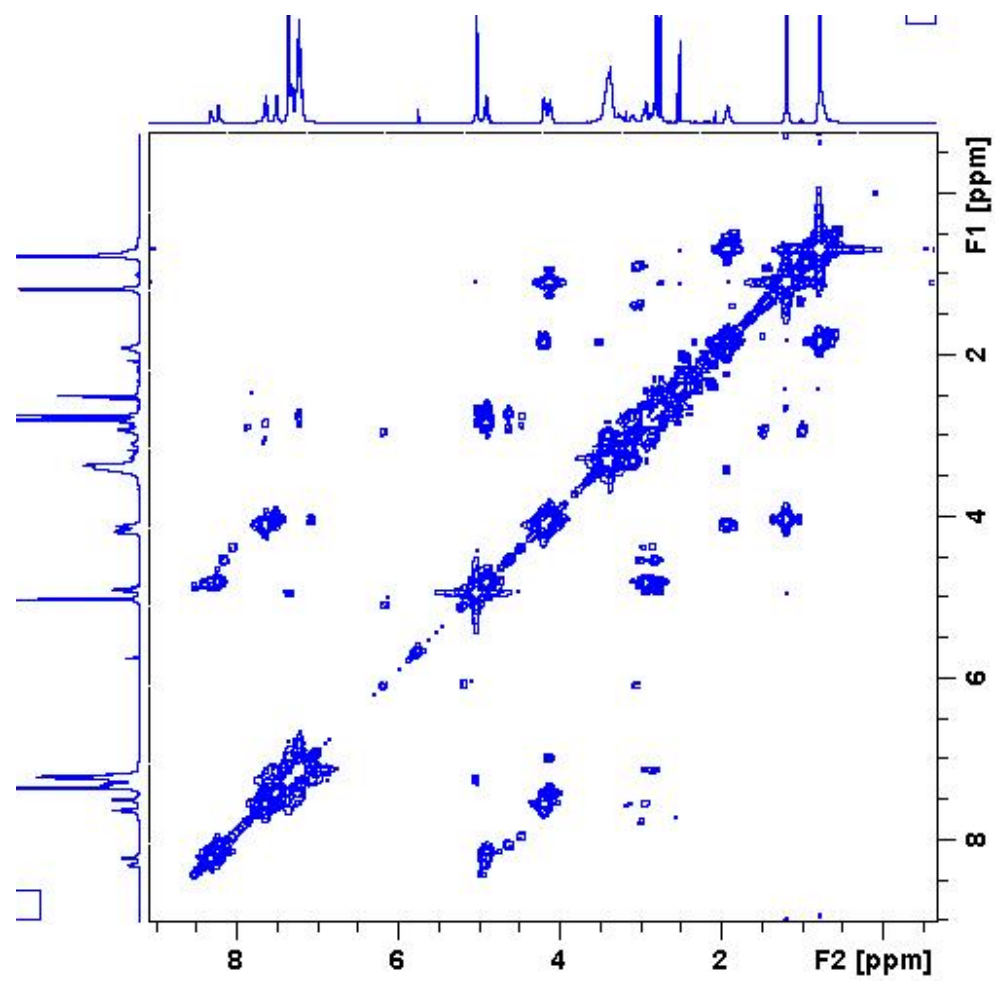
^1H NMR spectrum of PCU-OPVAZ, 3a



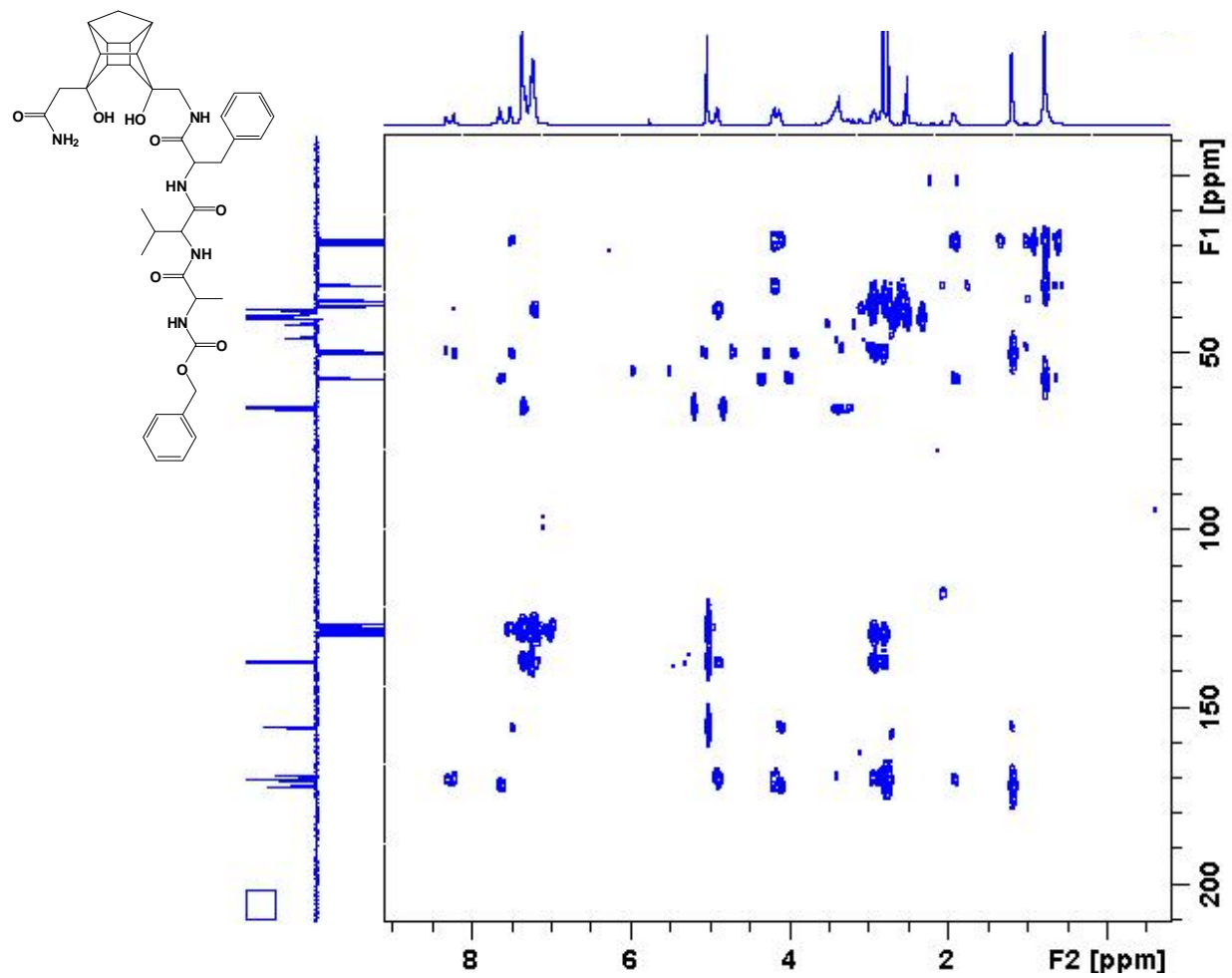
¹³C NMR spectrum of PCU-OPVAZ, 3a



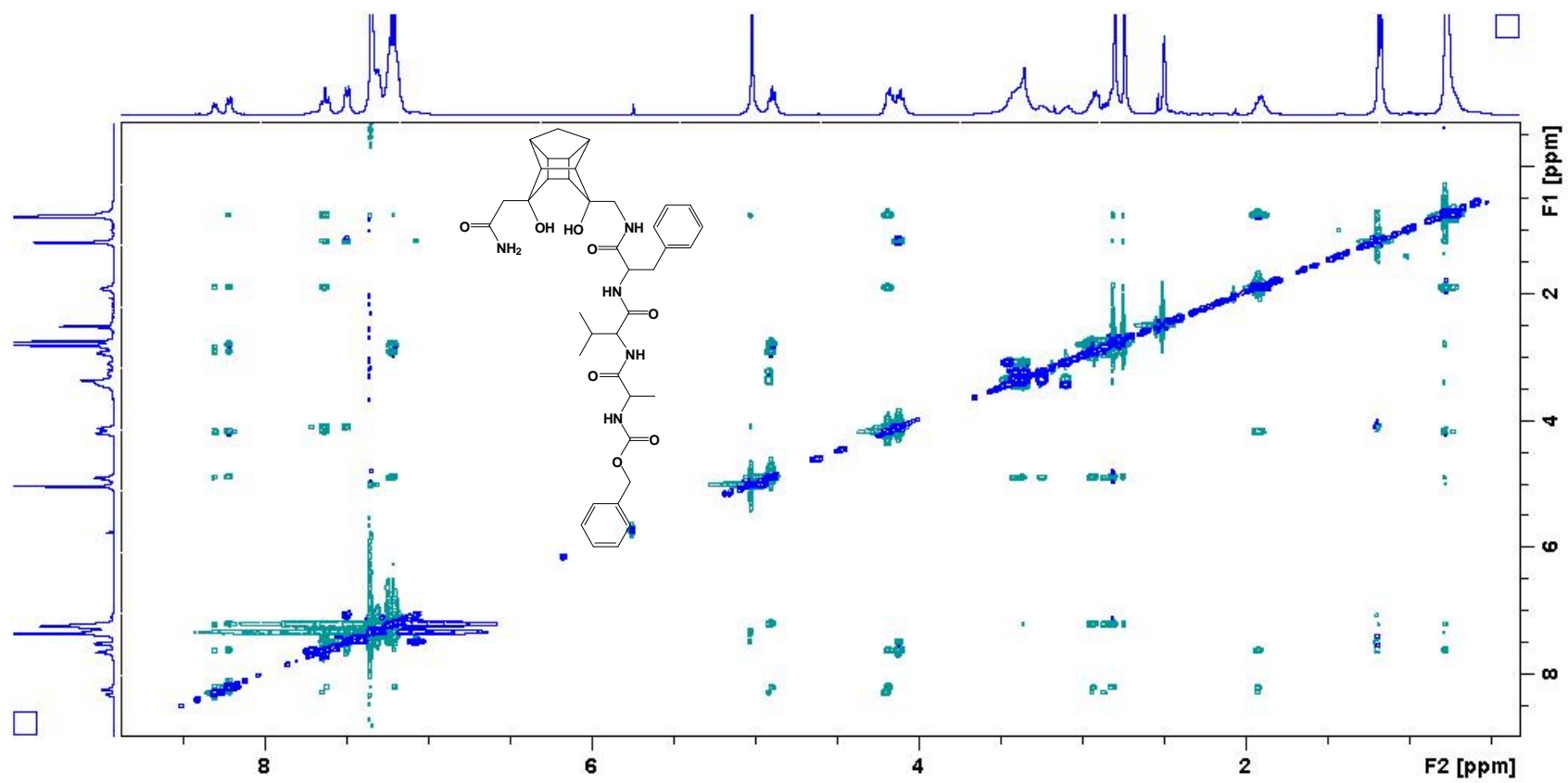
HSQC spectrum of PCU-OPVAZ, 3a



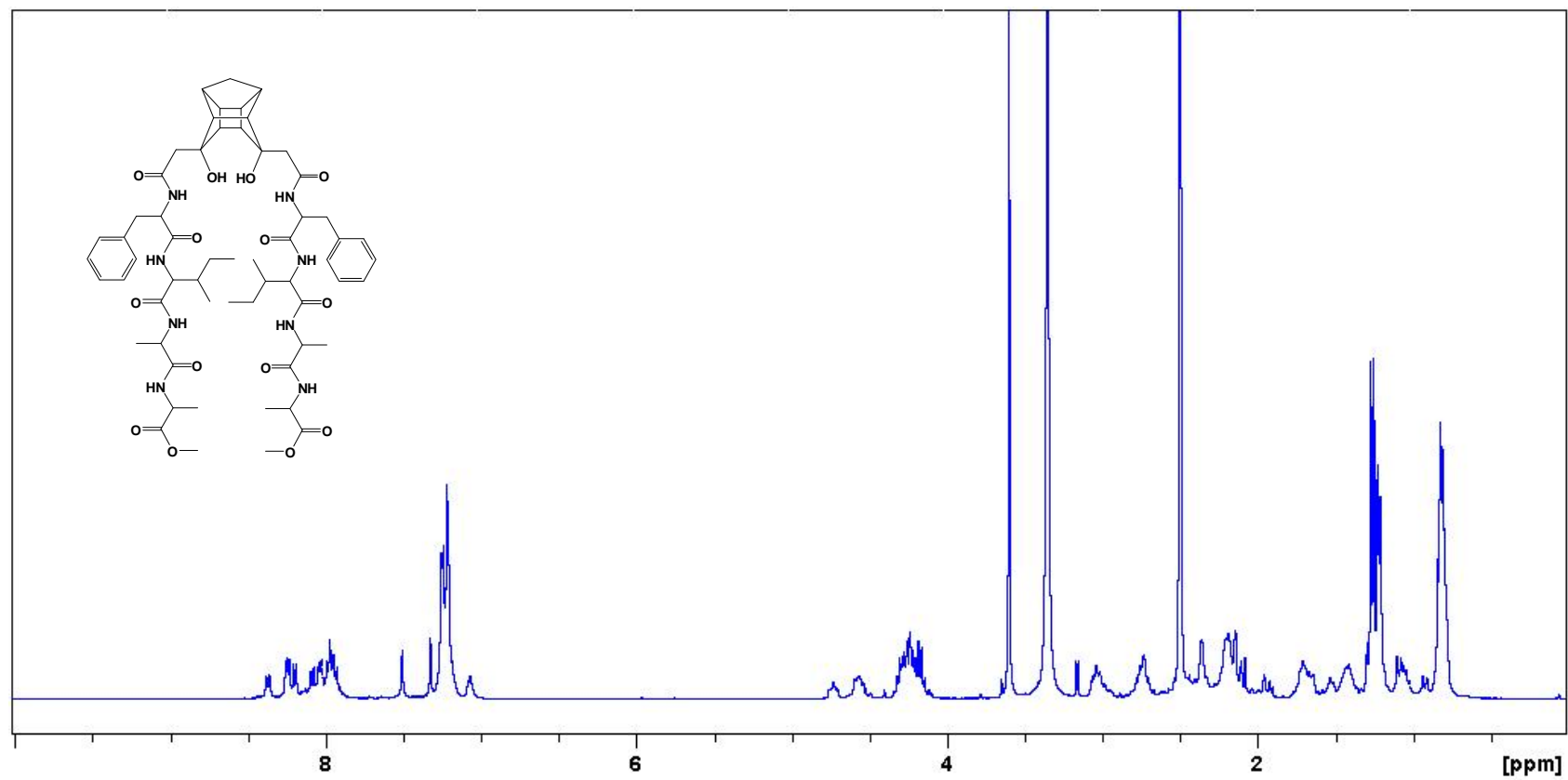
COSY spectrum of PCU-OPVAZ, 3a



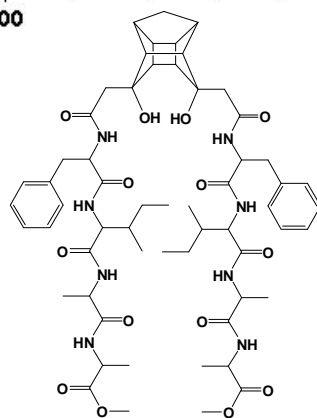
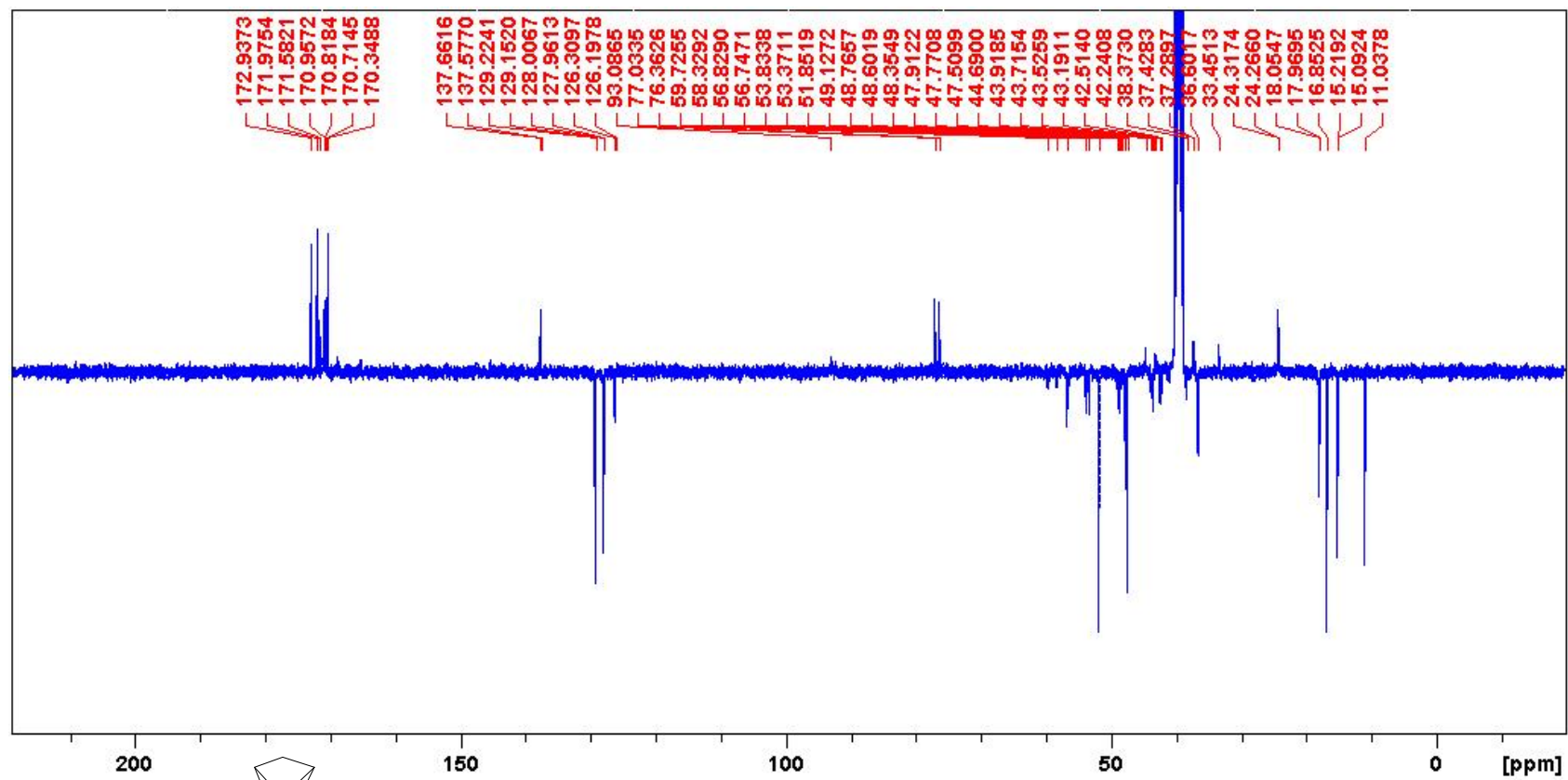
HMBC spectrum of PCU-OPVAZ, 3a



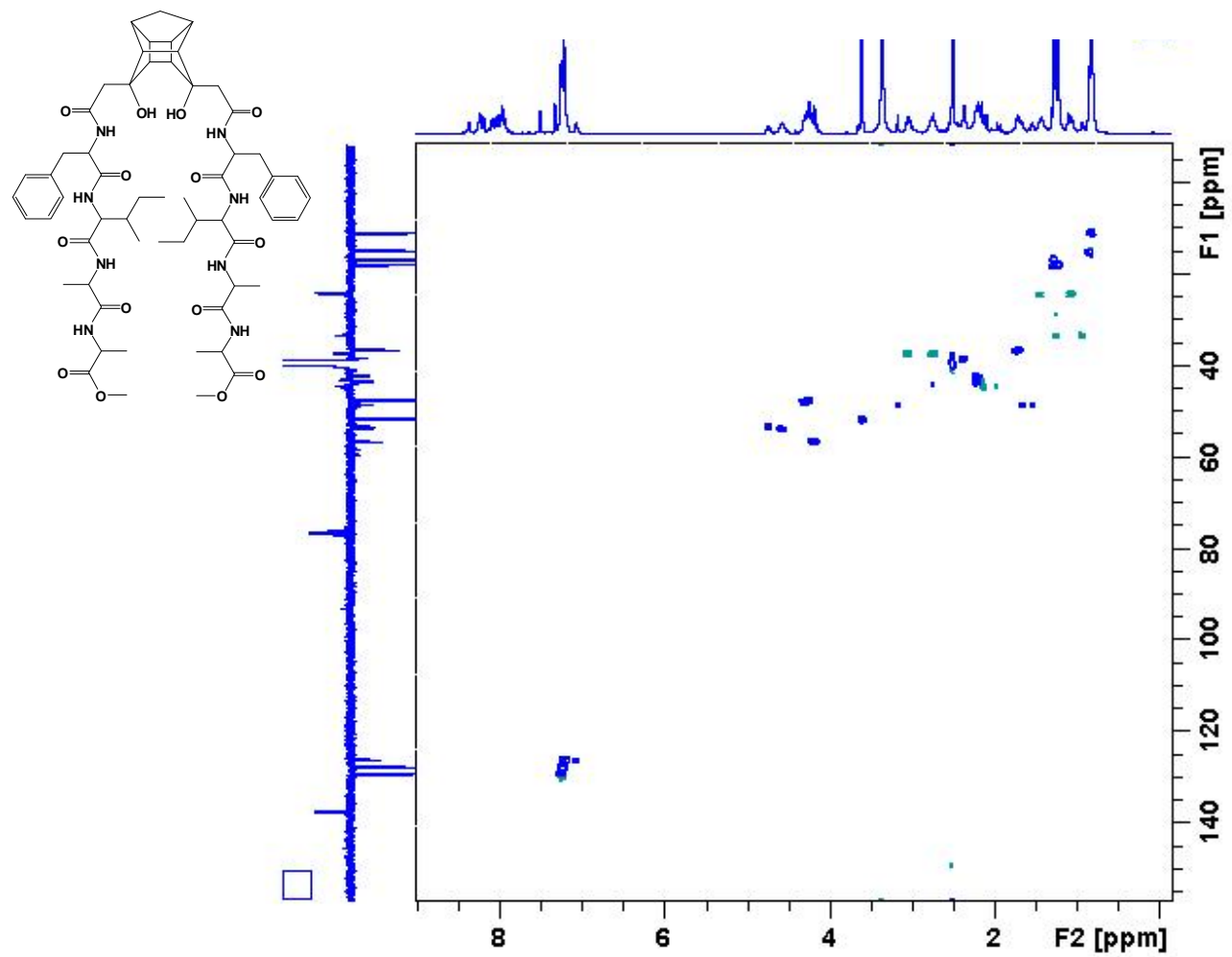
ROESY spectrum of PCU-OPVAZ, 3a



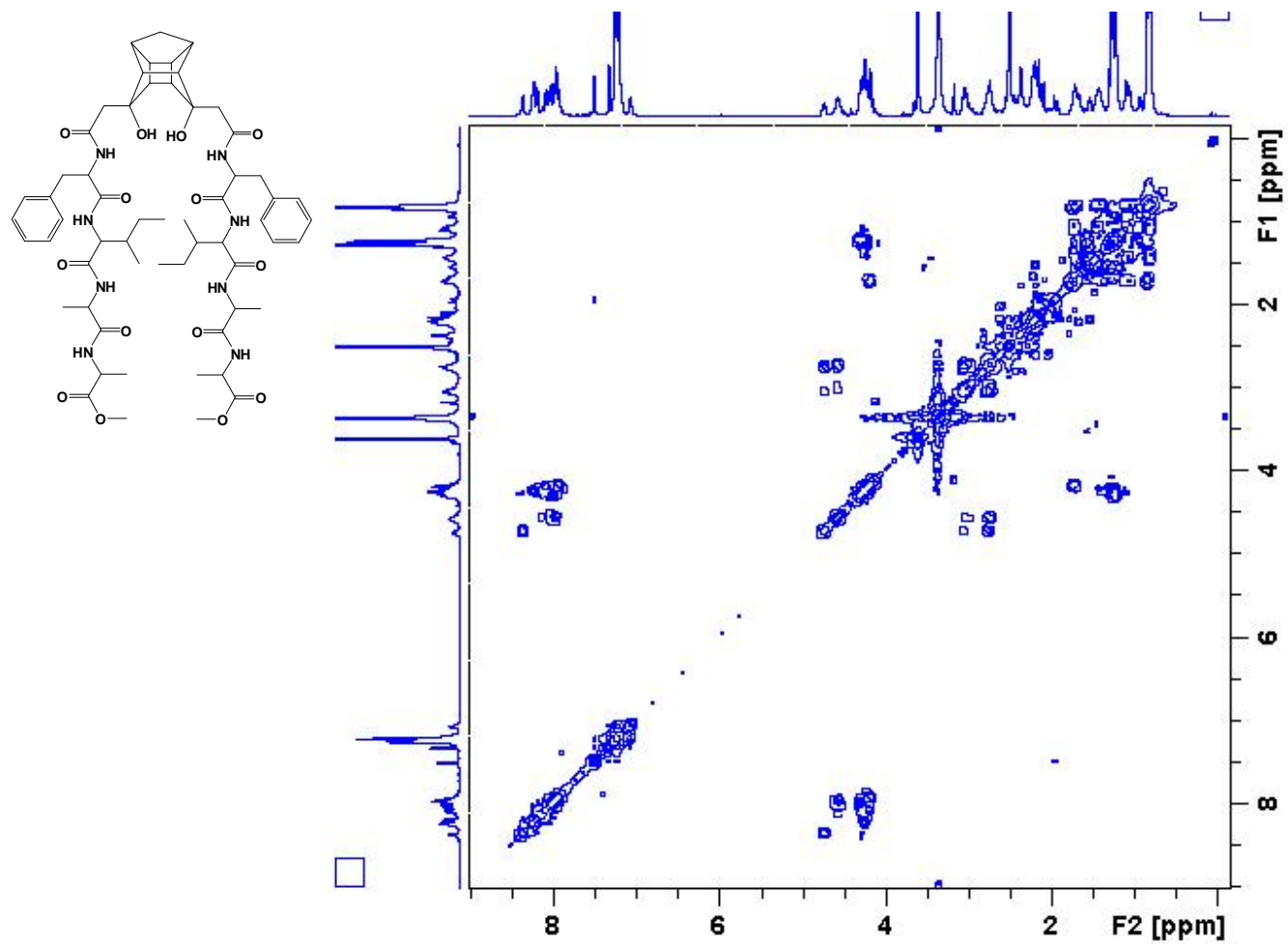
^1H NMR spectrum of PCU-PIAA



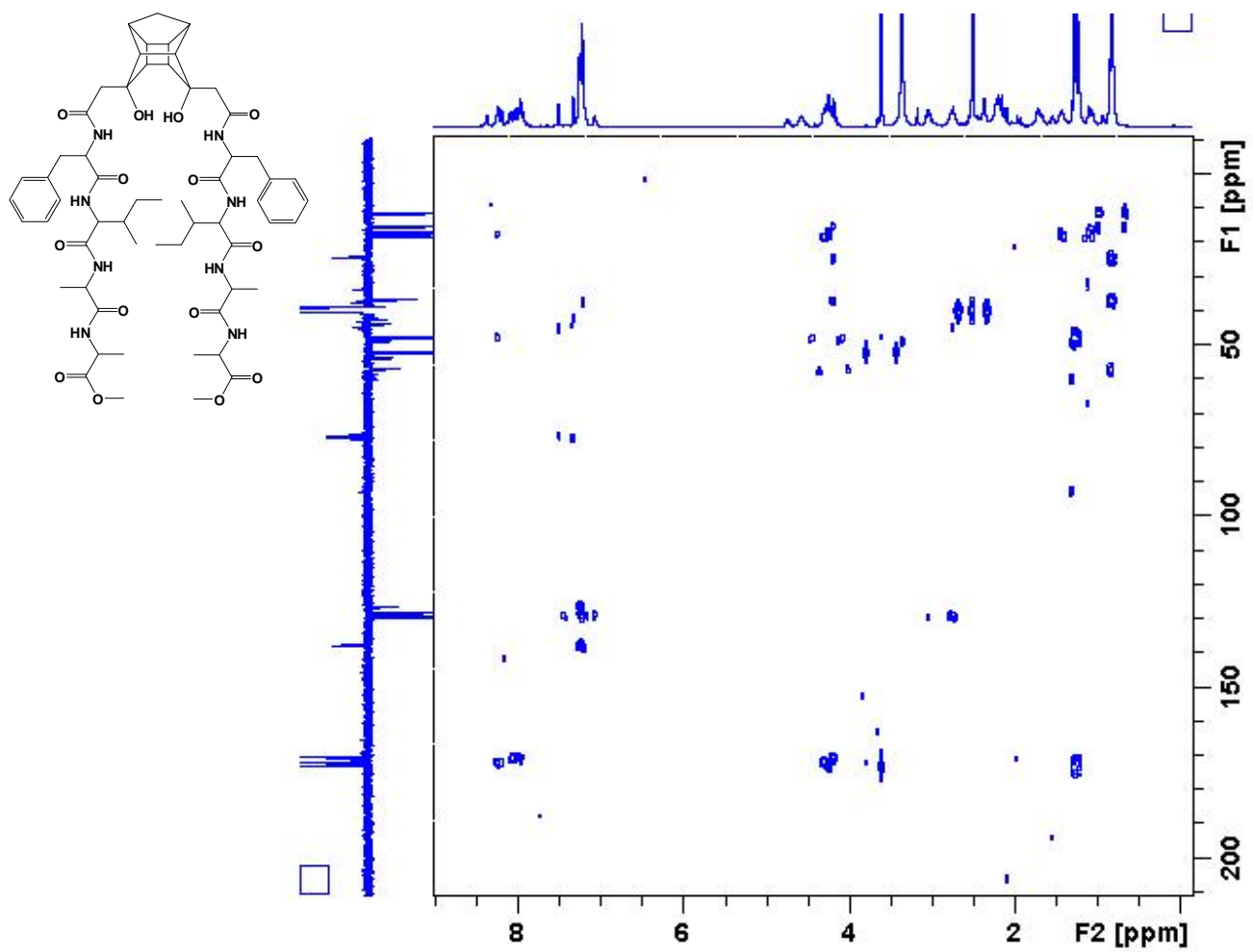
^{13}C NMR spectrum of PCU-PIAA



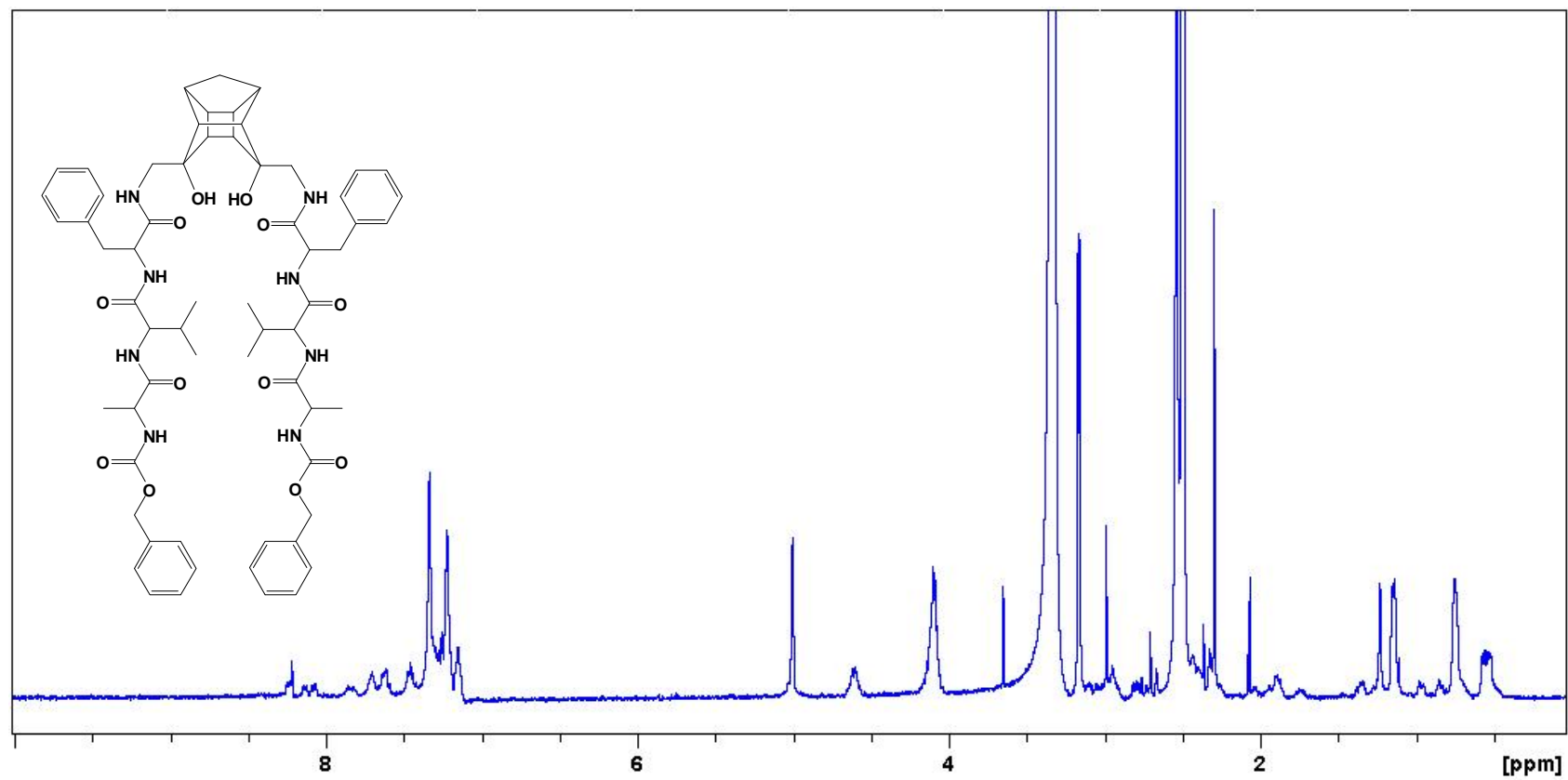
HSQC spectrum of PCU-PIAA



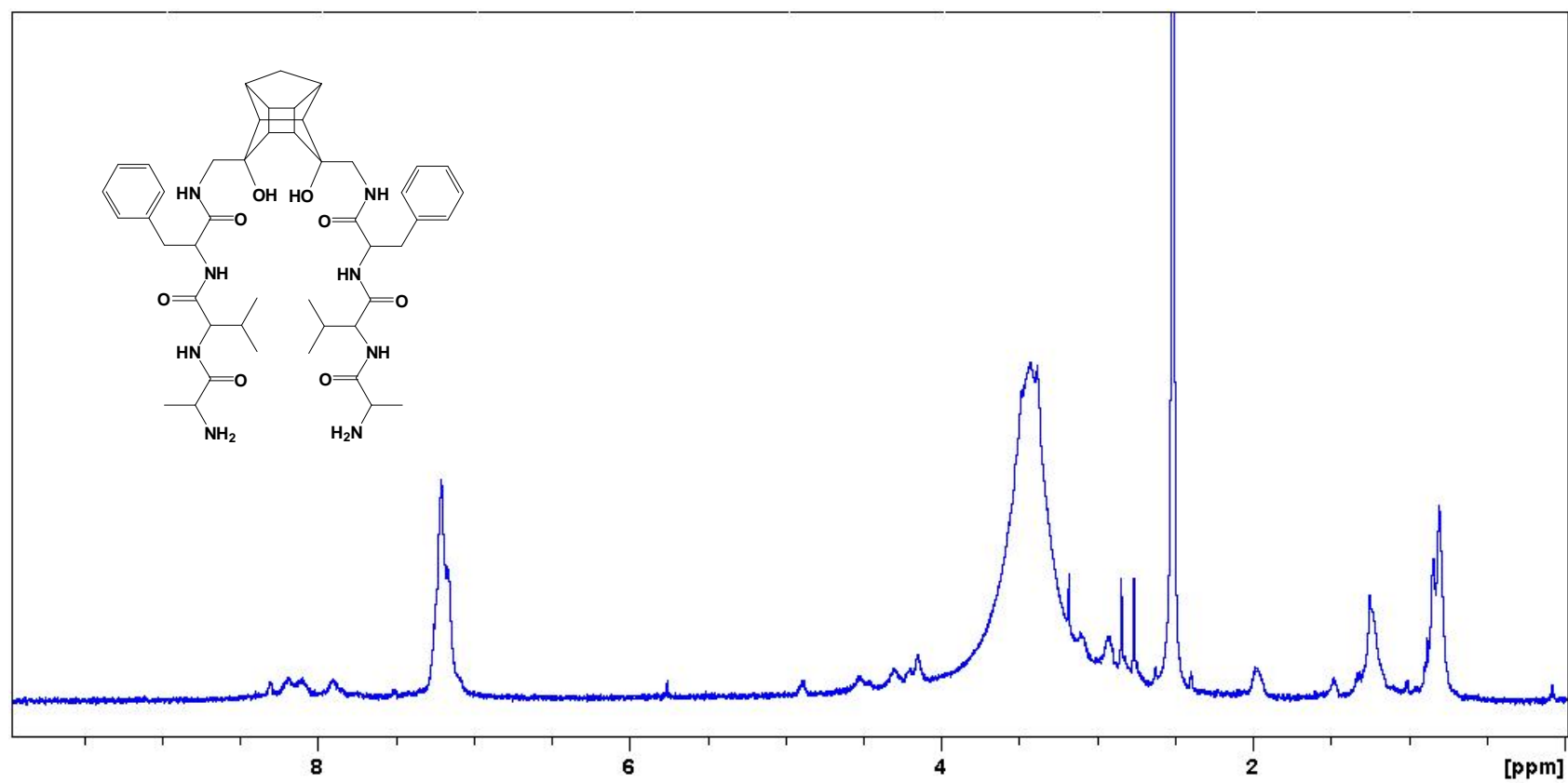
COSY spectrum of PCU-PIAA



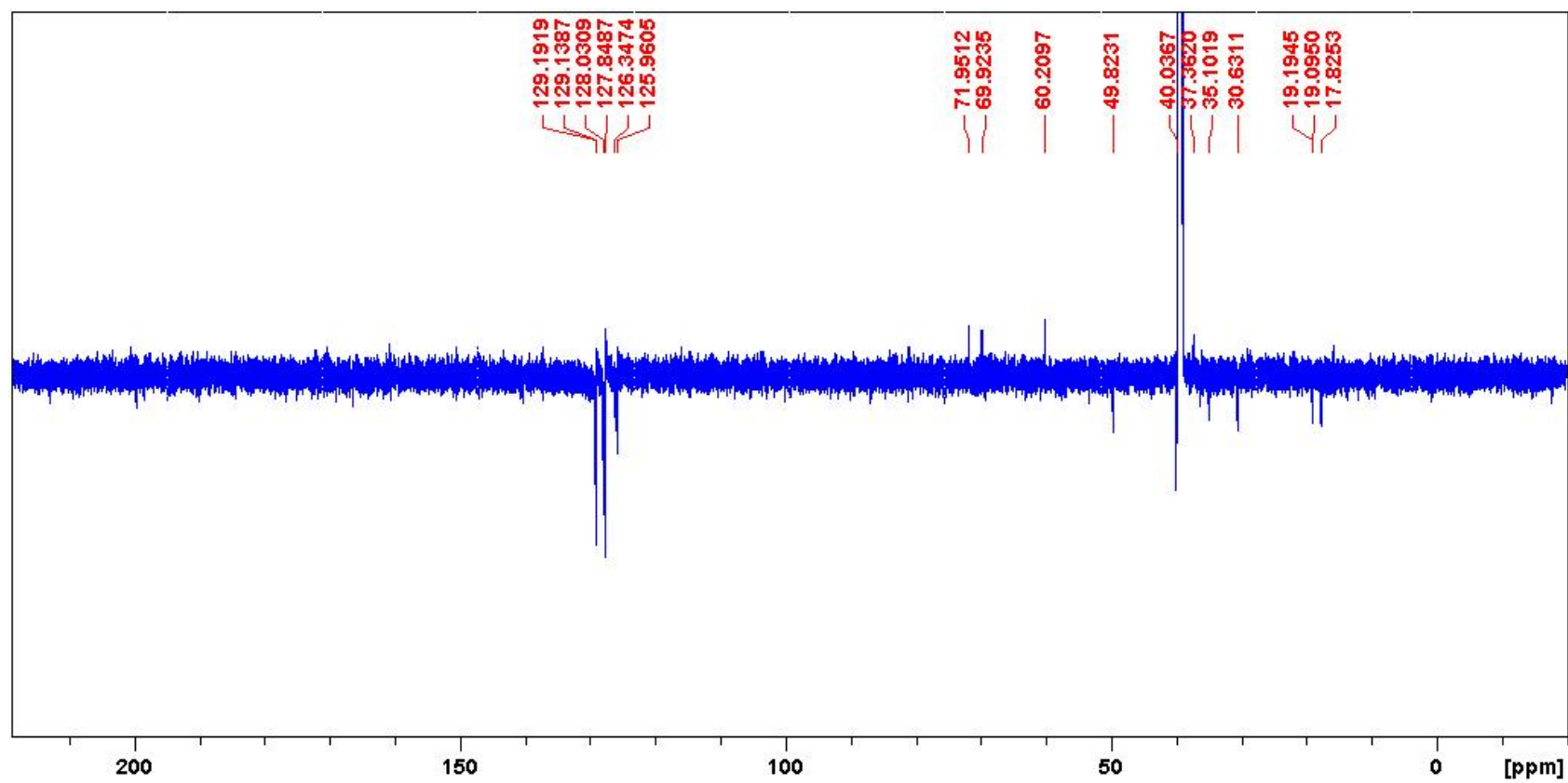
HMBC spectrum of PCU-PIAA



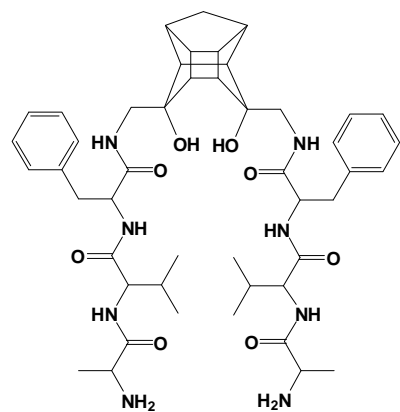
^1H NMR spectrum of PCU-PVAZ

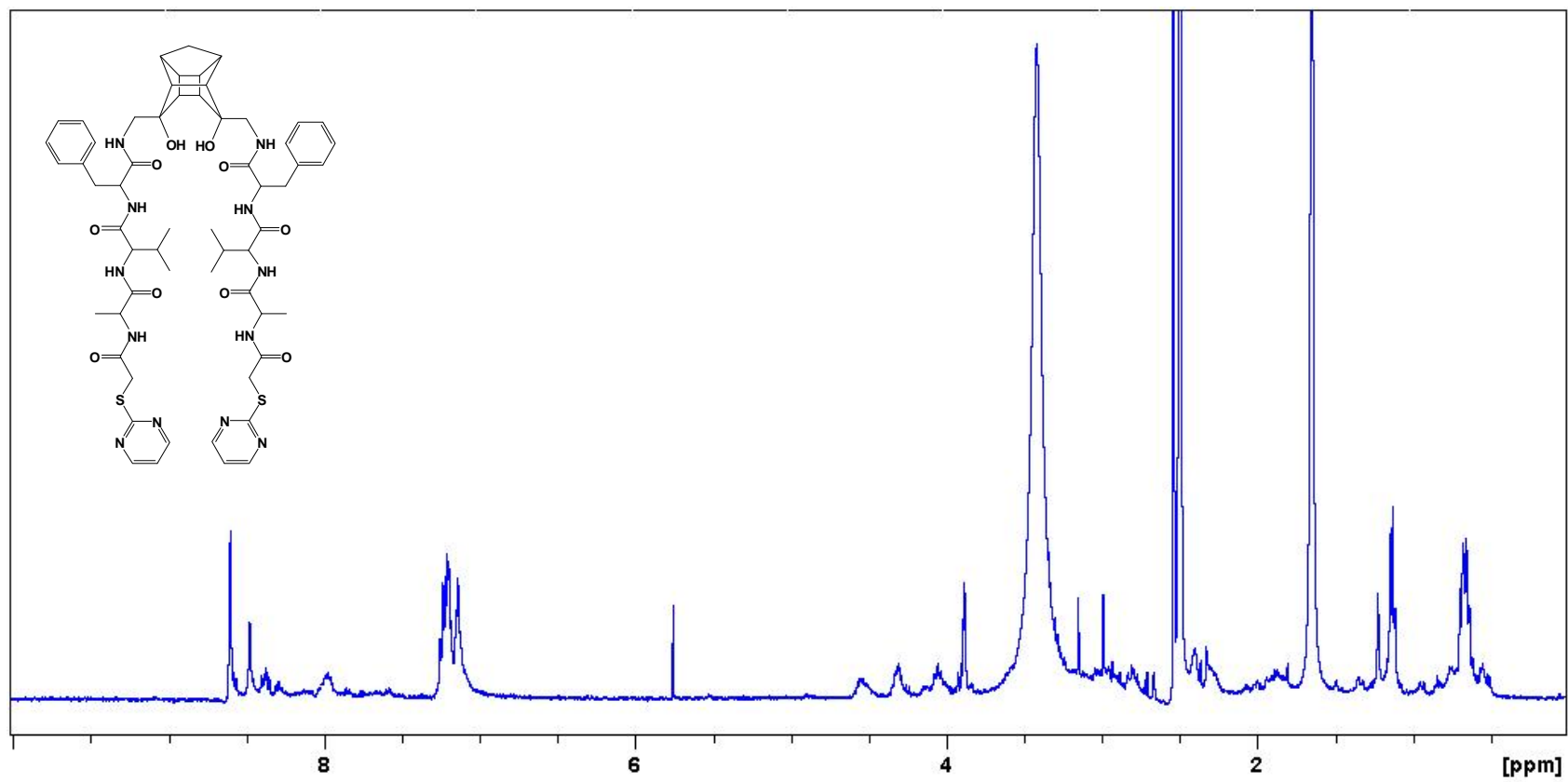


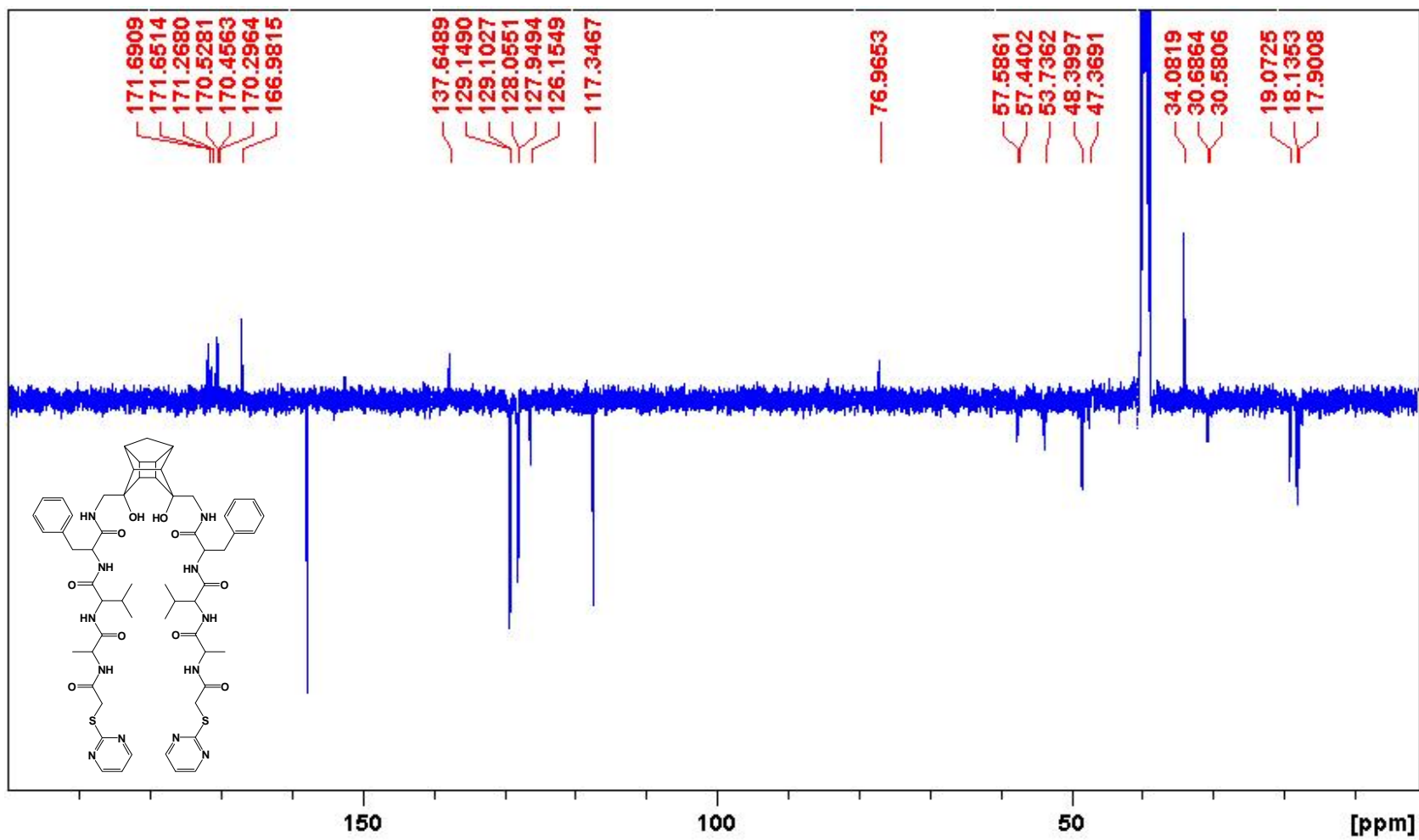
^1H NMR spectrum of PCU-PVANH₂

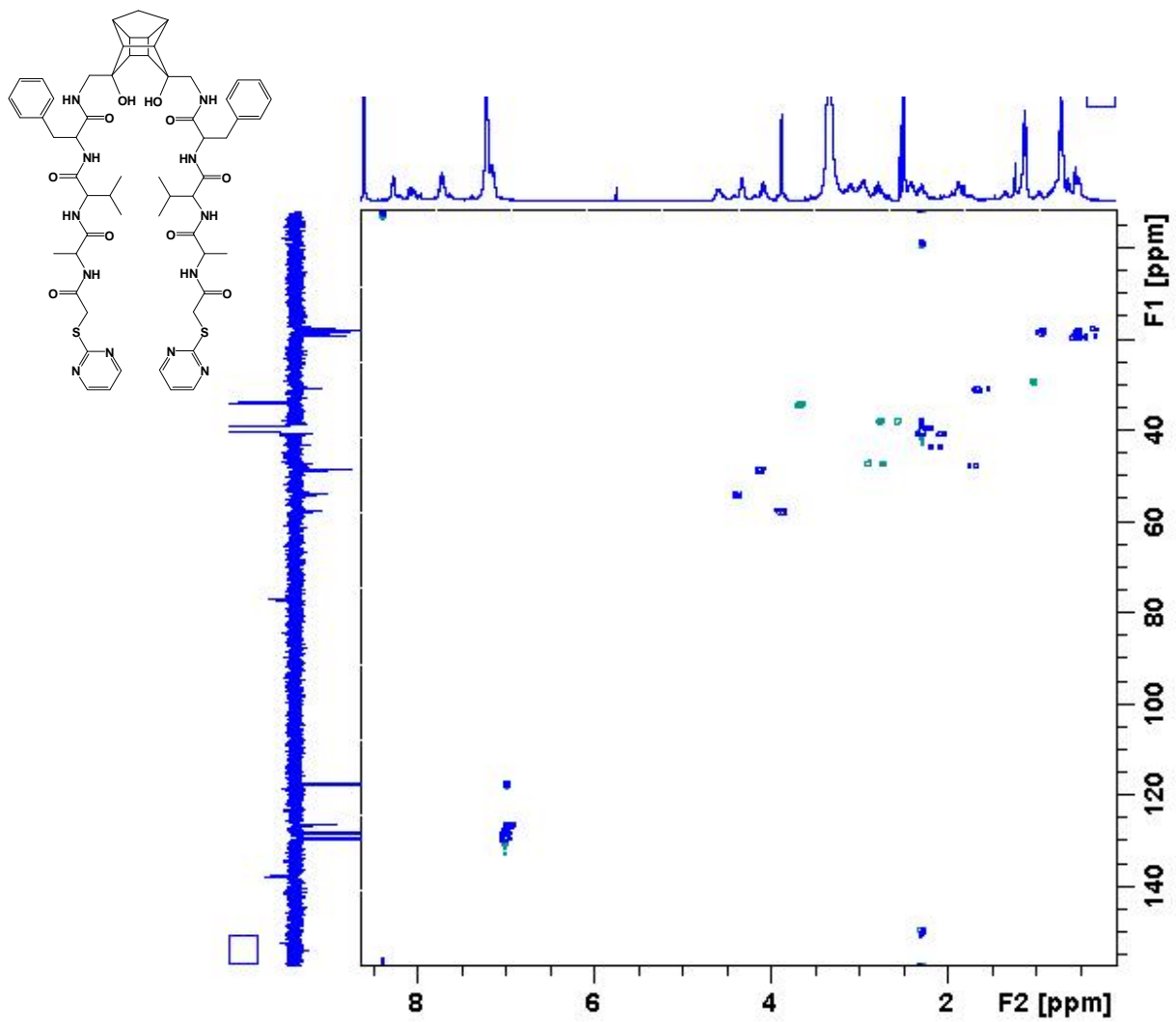


^{13}C NMR spectrum of PCU-PVANH₂

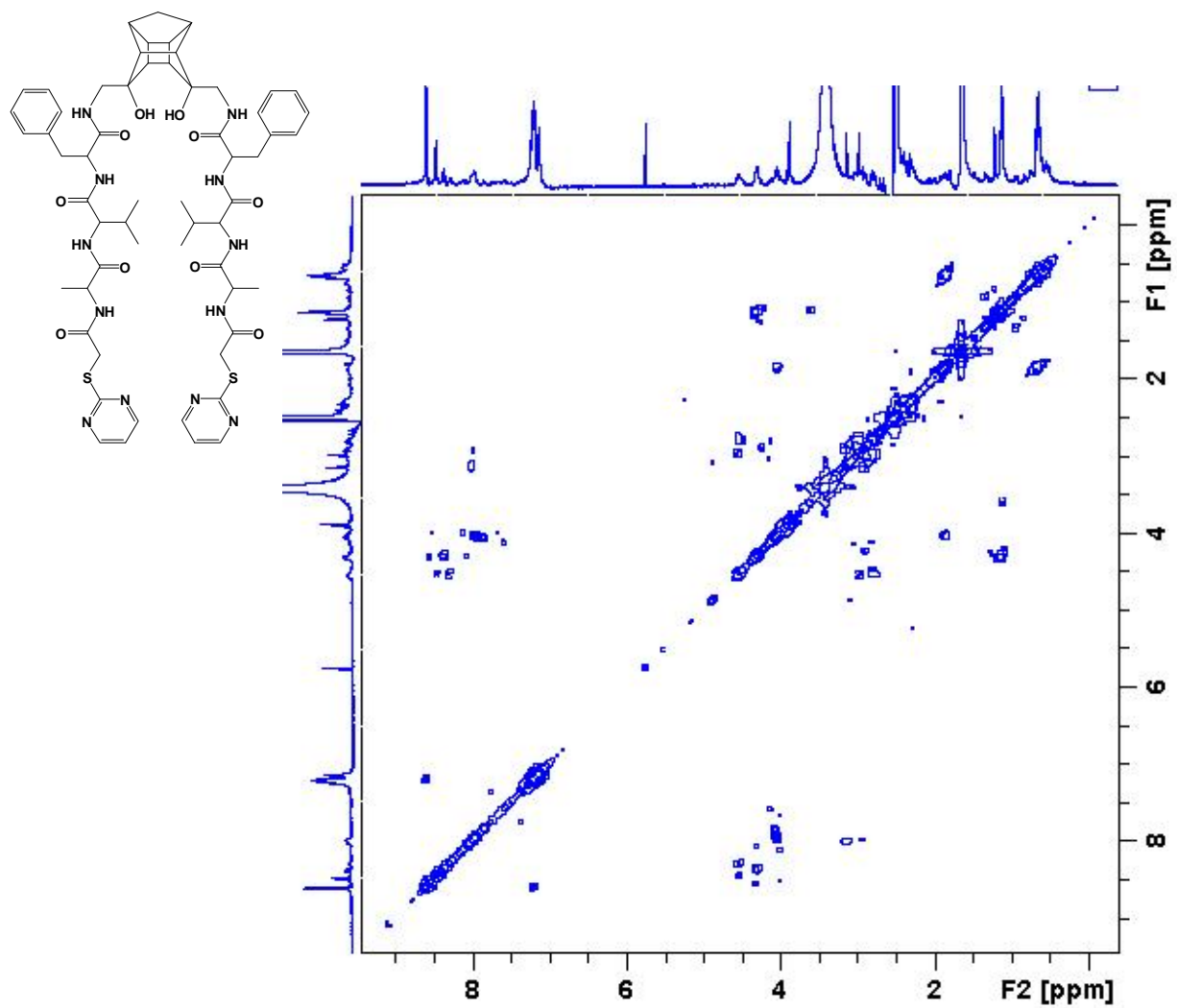


 ^1H NMR spectrum of PCU-PVAT

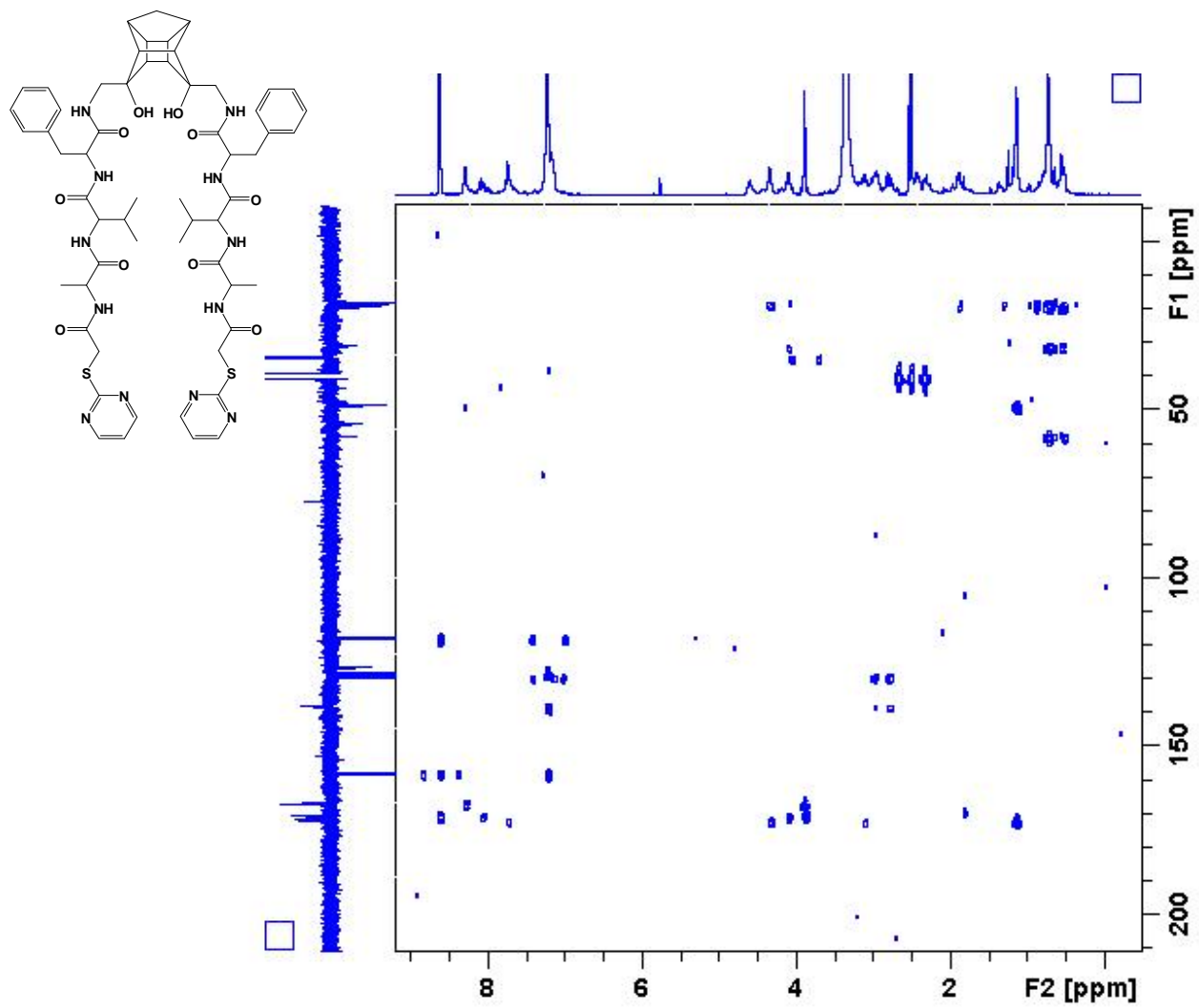
 ^{13}C NMR spectrum of PCU-PVAT



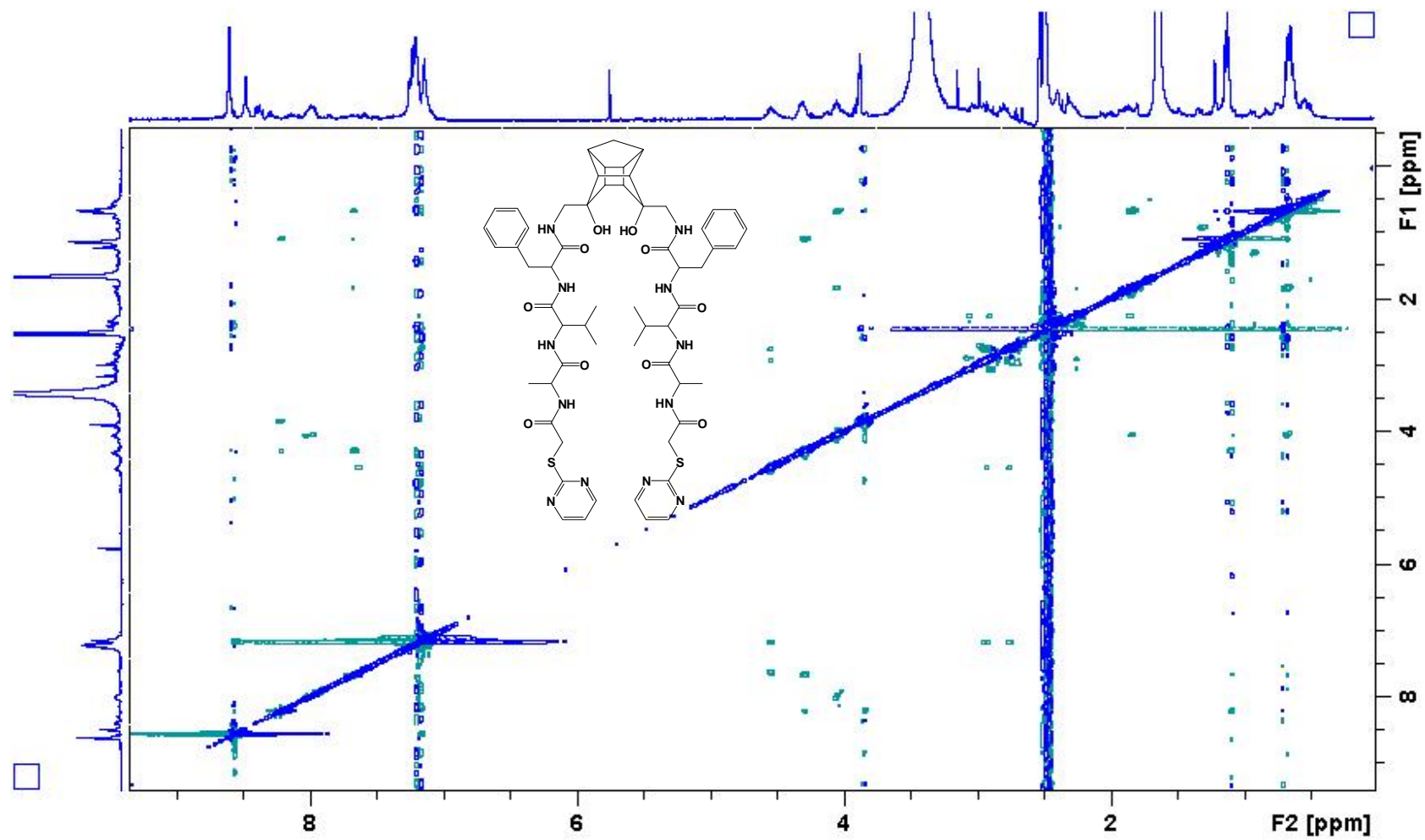
HSQC spectrum of PCU-PVAT



COSY spectrum of PCU-PVAT



HMBC spectrum of PCU-PVAT



ROESY spectrum of PCU-PVAT

APPENDIX 4

CHAPTER 5

SUPPORTING INFORMATION

MATERIALS AND METHODS

Analytical analysis was performed on an Agilent 1100 HPLC with a flow rate of 1 mL/min on a Waters Xbridge C18 (150 mm x 4.6 mm x 5 microns) coupled to a UV detector (215 nm) and an Agilent VL ion trap mass spectrophotometer in the positive mode. Semi-preparative HPLC was carried out on a Shimadzu semi-preparative instrument with a flow rate of 17 mL/min on a Ace C18 (150 mm x 21.2 mm x 5 microns) with a UV/VIS detector (215 nm) and an automated fraction collector. A two-buffer system was employed, utilizing formic acid as the ion-pairing agent. Buffer A consisted of 0.1 % formic acid/H₂O (v/v) and buffer B consisted of 0.1 % formic acid/acetonitrile (v/v). High resolution mass spectroscopic analysis was performed on a Bruker MicroTOF QII mass spectrometer in positive mode with an internal calibration. Peptides were synthesized on an automated CEM Liberty microwave peptide synthesizer (see Table 1 for conditions, not same as supplier). All ¹H, ¹³C, HSQC, COSY and HMBC NMR data were recorded on a Bruker AVANCE III 400 MHz spectrometer. The ROESY data was recorded on a Bruker AVANCE III 600 MHz spectrometer.

SYNTHESIS OF PCU DIACID 1

A solution of the diene **5** (31) (15.0 g, 58.0 mmol) in dry methanol (100 mL) was purged with nitrogen for 6 h while cooling in a dry ice–isopropanol bath (–78 °C). Ozone was bubbled into the reaction mixture until a blue–purple color persisted indicating the presence of excess ozone in the system and hence the completion of the reaction. The excess ozone gas was purge with nitrogen and the solvent (MeOH) removed *in vacuo*. Formic acid (100 mL) was added to the ozonide and the mixture was cooled in an ice bath with stirring. Hydrogen peroxide (150 mL, 30 %) was then added drop-wise to the stirring cooled reaction mixture. The reaction was left to attain ambient temperature for 1 h and refluxed gently for 12 h, the resulting mixture was concentrated *in vacuo* to yield the diol diacid **1** (15.6 g, 92 % yield) as a white solid. mp: 153–157°C, ¹H NMR [(CD₃)₂SO, 400 MHz]: δH 1.03 (AB, *J*_{AB} = 10.4 Hz, 1H), 1.51 (AB, *J*_{AB} = 10.4 Hz, 1H), 2.23 (s, 3H), 2.26 (s, 2H), 2.37 (s, 2H), 2.48–2.50 (m, 3H), 2.57–2.59 (m, 2H). ¹³C NMR [(CD₃)₂SO, 100 MHz]: δ_C 33.6 (CH₂), 38.9 (CH), 42.3 (CH), 43.9 (CH), 44.2 (CH₂), 50.0 (CH), 76.6 (C), 172.5 (C).

SYNTHESIS OF PCU DIAMIDE 6

The PCU diol diacid **1** (5.0 g, 17.0 mmol), NH₄Cl (2.4 g, 44.2 mmol), HOBt (5.7 g, 37.4 mmol), EDC (7.1 g, 37.4 mmol), and NMM (4.9 mL, 44.2 mmol) was dissolved in dimethylformamide (DMF) (30 mL) and

stirred at room temperature for 36 h. The reaction mixture was concentrated *in vacuo*. The crude product was dissolved in ethylacetate and washed successively with 1 M HCl (50 mL), 10 % NaHCO₃ (50 mL), and brine (50 mL). The organic solution was dried over Na₂SO₄ and concentrated *in vacuo*. The crude product was purified *via* column chromatography on silica gel using CHCl₃:MeOH:25 % NH₄OH (88 : 10 : 2, *R*_f = 0.3, 72 % yield) to obtain a white solid. mp; 207–209 °C, IR ν_{max} : 3172, 2961, 1656, 1408, 1287, 728, 634, and 436 cm⁻¹. Mass spectrometry (time of flight) [MS (TOF)] calculated for C₁₅H₂₁N₂O₄ (M⁺ H⁺) 293.1496, found 293.1484. ¹H NMR [(CD₃)₂SO, 400 MHz]: δ H 1.03 (AB, *J*_{AB} = 10.5 Hz, 1H), 1.48 (AB, *J*_{AB} = 10.5 Hz, 1H), 2.09 (AB, *J*_{AB} = 14.4 Hz, 1H), 2.19 (AB, *J*_{AB} = 14.45 Hz, 1H) 2.19 (s, 2H), 2.40 (s, 2H), 2.41 (s, 2H), 2.48 (s, 2H). ¹³C NMR [(CD₃)₂SO, 100 MHz]: δ _C 33.5 (CH₂), 38.9 (CH), 42.5 (CH), 43.8 (CH), 43.9 (CH₂), 49.3 (CH), 76.6 (C), 173.8 (C).

SYNTHESIS OF PCU DIAMINE 2

Sodium hypochlorite (90 mL) and 4 M NaOH (48 mL) were added to 7 g of the diamide. The reaction was stirred for an hour at 80°C and then left to stir at room temperature for 24 hours. HCl (1M) was then added to adjust the pH to 7 before the reaction was concentrated *in vacuo*. The crude product was then purified using column chromatography CHCl₃:MeOH:25 % NH₄OH (88 : 10 : 2) (2.9 g, 53 % yield). The diamine (2.5 g, 10.6 mmol) was dissolved in 15 mL of water and 3 equivalence (eq) of NaHCO₃ was added while stirring. The resulting solution was cooled to 5°C and *t*-BOC anhydride (2.4 eq) was added slowly as a solution in cooled THF (15 mL). The mixture was stirred at 0°C for 1 hour and allowed to warm to room temperature overnight. Water was then added and extracted twice with DCM. The organic layer was back extracted with saturated NaHCO₃. The aqueous layer was acidified to pH1 and then extracted with DCM twice. All organic layers were combined, dried over MgSO₄ and concentrated under reduced pressure. The crude *t*-BOC-PCU diamine product was purified with column chromatography and stored below 4°C. (4.1 g, 75 % yield) ¹H NMR [(CD₃)₂SO, 400 MHz]: δ H 1.02 (AB, *J*_{AB} = 10.5 Hz, 1H), 1.44 (AB, *J*_{AB} = 10.5 Hz, 1H), 1.38 (s, 9H), 2.09 (s, 2H), 2.34 (s, 2H), 2.42-2.45 (m, 4H), 2.86 (AB, *J*_{AB} = 6.18 Hz, 1H), 2.90 (AB, *J*_{AB} = 6.18 Hz, 1H), 2.94 (AB, *J*_{AB} = 6.18 Hz, 1H), 2.98 (AB, *J*_{AB} = 6.18 Hz, 1H) 6.28 (s, 1H), 7.02 (s, 1H). ¹³C NMR [(CD₃)₂SO, 100 MHz]: δ _C 28.2 (CH₃), 33.7 (CH₂), 39.2 (CH), 40.4 (CH), 43.0 (CH), 43.0 (CH), 47.9 (CH), 48.1 (CH₂), 77.1 (C), 77.7 (C), 156.2 (C)

The removal of *t*-BOC from the diamine was achieved by stirring the compound in methanol (30 mL) and conc HCl (12M, 5 mL) overnight.

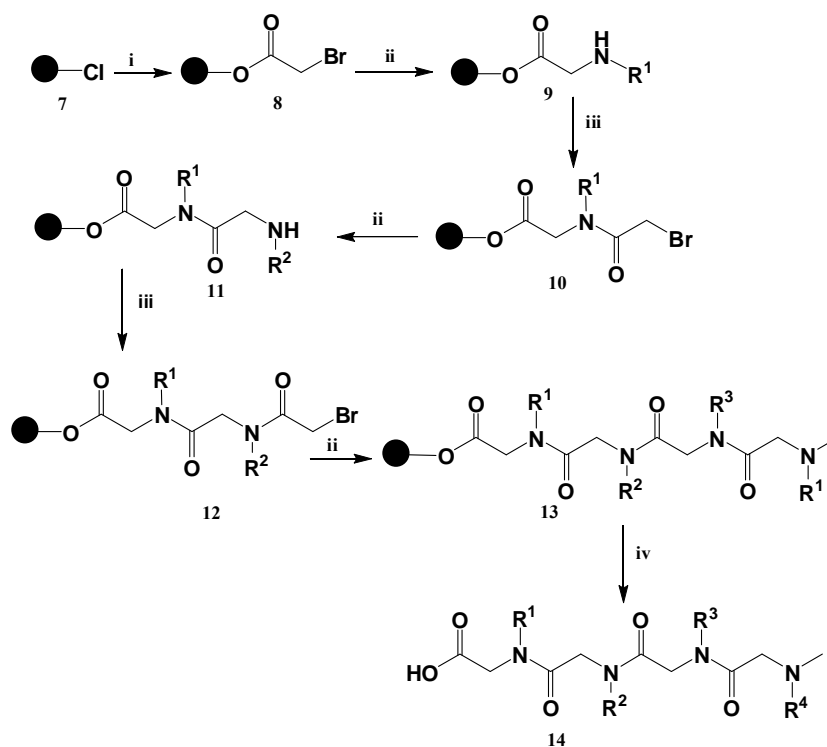
SYNTHESIS OF THE CAGE AMINE AMIDE 3

Sodium hypochlorite (45 mL) and 4M NaOH (24 mL) were added to 7 g of the diamide. The reaction was stirred for an hour at 80°C and then left to stir at room temperature for 24 hours. HCl (1 M) was then added to adjust the PH to 7 before the reaction was concentrated in *vacuo*. The crude product was then purified using column chromatography CHCl₃:MeOH:25 % NH₄OH (88 : 10 : 2) (1.6 g, 28 % yield). The amine amide (2.5 g, 10.6 mmol) was dissolved in 15 mL of water and 1.5 equivalence (eq) of NaHCO₃ was added while stirring. The resulting solution was cooled to 5°C and *t*-BOC anhydride (1.2 eq) was added slowly as a solution in cooled THF (15 mL). The mixture was stirred at 0°C for 1 hour and allowed to warm to room temperature overnight. Water was then added and extracted four times with DCM, dried over MgSO₄ and concentrated under reduced pressure. The crude *t*-BOC-PCU amine amide product was purified with column chromatography and stored below 4°C. (1.9 g, 86 % yield) ¹H NMR [(CD₃)₂SO, 400 MHz]: δH 1.02 (AB, *J*_{AB} = 10.5 Hz, 1H), 1.47 (AB, *J*_{AB} = 10.5 Hz, 1H), 1.37 (s, 9H), 2.09 (s, 2H), 2.33 (s, 2H), 2.42-2.45 (m, 4H), 2.92 (m, 2H), 3.12 (m, 1H), 3.16 (m, 1H), 6.20 (s, 1H), 8.32 (s, 1H). ¹³C NMR [(CD₃)₂SO, 100 MHz]: δ_C 28.0 (CH₃), 33.7 (CH₂), 36.4 (CH₂), 40.4 (CH), 43.1 (CH), 43.5 (CH), 45.3 (CH₂), 48.0 (CH), 48.1 (CH₂), 76.7 (C), 76.9 (C), 77.6 (C), 78.4 (C), 153.3 (C), 156.2 (C), 172.1 (C), 174.1 (C),

The removal of *t*-BOC from the amine was achieved by stirring the compound in methanol (30 mL) and conc HCl (12M, 5 mL) overnight

GENERAL PROCEDURE FOR THE SYNTHESIS OF PEPTOID

The peptoid oligomers to fit into the S1/S1'-S4/S4' subsite of the HIV protease were synthesized via solid phase synthesis on activated 2-chlorotrityl chloride resin **7**, scheme 2. Bromoacetic acid was coupled to the resin **7** in the presence of DIPEA in dry dichloromethane (DCM) for 2 hours and this affording compound **8**, scheme 2. Upon filtration, the amine solution in DMF was added and mixed under a stream of nitrogen bubbles to give compound **9**. The chain was elongated by coupling bromoacetic acid in the presence of N, N'-diisopropylcarbodiimide (DIC). Similarly the corresponding amines were coupled to the extended chain followed by bromoacetic acid to give compounds **10-14**. The completed peptoids were cleaved from the resin using 5 % trifluoroacetic acid (TFA) in DCM as illustrated in **Scheme 2**.



Scheme 2. General Synthesis of peptoids

GENERAL PROCEDURE FOR COUPLING OF PEPTOID TO THE PCU-DIOL

The cleaved peptide (2.6 eq.) was dissolved in DMF (3 mL) followed by addition of PCU-diol (1 eq.), EDC (2.4 eq), HOBT (2.4 eq.), in DMF (7 mL) and NMM (2.6 eq.). The mixture was stirred at room temperature for 24 hours and thereafter evaporated to dryness under reduced pressure using a Teflon pump at 40°C. The crude product was diluted with EtOAc (200 mL) and the resulting solution was washed with 1M HCl, sat. aq. NaHCO₃ and sat. aq. NaCl, dried over MgSO₄, filtered and concentrated *in vacuo*. The product was then purified by reverse phase semi-preparative HPLC carried out on a Shimadzu Instrument (Ace C18 150 mm x 21.2 mm x 5 microns) with a UV/VIS detector (215 nm). There were characterized using NMR Bruker AVANCE III 400 MHz and 600 MHz spectrometer and High Resolution Mass Spectroscopic analysis performed on a Bruker MicroTOF QII

OVER-EXPRESSION, EXTRACTION AND PURIFICATION OF THE C-SA PROTEASE

Plasmid encoding HIV-1 subtype C protease (containing the mutation Q7K designed to reduce the hypersensitive autolytic site) is expressed as inclusion bodies (Ido *et al.*, 1991) in *Escherichia coli* BL21 (DE3) pLysS cells. Briefly, *Escherichia coli* cells harboring the plasmid DNA were grown at 37 °C in LB medium supplemented with 100 µg/mL of ampicillin and 35 µg/mL of chloramphenicol. The overnight culture was diluted 100-fold into fresh 2 × YT medium supplemented with ampicillin (100 µg/mL) and chloramphenicol (35 µg/mL) and grown at 37 °C. When the optical density (OD₆₀₀) of the culture reached 0.4 to 0.5, over-expression of the HIV-1 C-SA protease was induced by adding IPTG. IPTG was added to final concentrations of 0.4 mM. Over-expression of the protease was allowed to continue for four hours.

The cells were pelleted after growth and resuspended in ice-cold extraction buffer [10 mM Tris, 1 mM EDTA, and 1 mM PMSF (added only fresh before use), pH 8] and disrupted using an ultra-sonicator. Following the addition of MgCl₂ and DNase I to final concentrations of 10 mM and 10 U/µL, respectively, the culture medium was stirred on ice until the viscosity of the mixture decreased. The cells were then ruptured by sonication and centrifuged at 15 000 × g for 30 minutes at 4 °C. The pellet was resuspended in ice-cold extraction buffer containing 1 % (v/v) of Triton X-100. Cell debris and protease-containing inclusion bodies were pelleted by centrifugation at 15 000 × g for 30 minutes at 4 °C. The pellet was then resuspended in a freshly prepared solubilization buffer containing 10 mM Tris, 2 mM DTT, 8 M urea, pH 8.0, at room temperature, and centrifuged at 15 000 × g for 30 minutes at 20 °C.

The protease, in the supernatant, was purified by passing through an anion exchange (DEAE) column previously equilibrated with solubilization buffer. Upon elution from the column, the protease was acidified by adding formic acid to a final concentration of 25 mM. Precipitation of significant amount of contaminating proteins occurred upon acidification. Following an overnight incubation, the precipitated contaminants were removed by centrifugation at 15 000 × g for 30 minutes at 4 °C.

HIV-1 protease was refolded by dialysis into 10 mM formic acid at 4 °C. Subsequently, the protease was dialyzed into storage buffer containing 10 mM sodium acetate, 1 mM NaCl and 1 mM DTT, pH 5.0. The folded protease was concentrated to a final volume of ~5 mL and stored at -20 °C.

***IN VITRO* HIV-1 PROTEASE ACTIVITY**

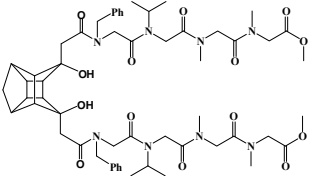
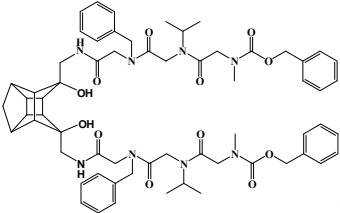
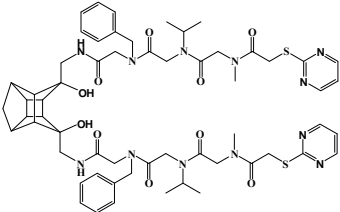
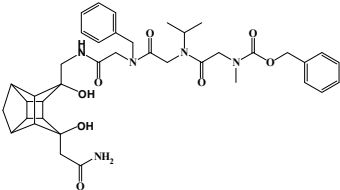
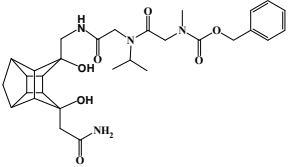
The catalytic activity of the HIV-1 protease was monitored following the hydrolysis of the chromogenic peptide substrate Lys-Ala-Arg-Val-Nle-*p*-nitro-Phe-Glu-Ala-Nle-NH₂. This substrate mimics the

conserved KARVL/AEAM cleavage site between the capsid protein and nucleocapsid (CA-p2) in the Gag polyprotein precursor.

For this study the chromogenic substrate was synthesized using the Discovery CEM Liberty microwave peptide synthesizer on the Rink amide resin. The substrate was cleaved from the resin and deprotected using 95 %:5 % (v/v) TFA:TIS for 3 hours. It was then precipitated using cold ether and purified *via* reverse phase semi-preparative HPLC on a Shimadzu instrument and characterized using the Bruker microTOF-Q II instrument (Table 2).

To determine the concentration of the inhibitors that resulted in 50 % inhibition (IC_{50}) of HIV-1 protease enzyme activity, the protein (100 nM) and chromogenic substrate (50 μ M) were added into a 120 μ L microcuvette containing increasing concentrations of inhibitor in a pH 5.0 buffer (50 mM sodium acetate and 0.1 M NaCl). Protease hydrolytic activity was measured by monitoring the relative decrease in absorbance at 300 nm using Fpecord 210 spectrophotometer.

Supplementary Table 1. The inhibition of wild type C-SA HIV protease by the PCU diol peptoid derivatives, yields for the coupling of peptoids to the cage diols and the size as confirmed by High-resolution mass spectroscopy

Compound	IC ₅₀ (μM)	Yield (%)	Mass calculated	Mass found	Molecular ion
 1a	6.5 ± 0.35	19	1121.5530 C ₅₇ H ₇₈ N ₈ NaO ₁₄	1121.5530	(M+Na ⁺)
 2a	0.075 ± 0.0035	30	1139.5812 C ₆₃ H ₇₉ N ₈ O ₁₂	1139.5780	(M+H ⁺)
 2b	1.0 ± 0.17	39	1175.5165 C ₅₉ H ₇₅ N ₁₂ O ₁₀ S ₂	1175.5159	(M+H ⁺)
 3a	0.5 ± 0.18	22	716.3654 C ₃₉ H ₅₀ N ₅ O ₈	716.3654	(M+H ⁺)
 3b	0.75 ± 0.071	33	591.2789 C ₃₀ H ₄₀ N ₄ NaO ₇	591.2787	(M+Na ⁺)

NMR elucidation of inhibitor 2a

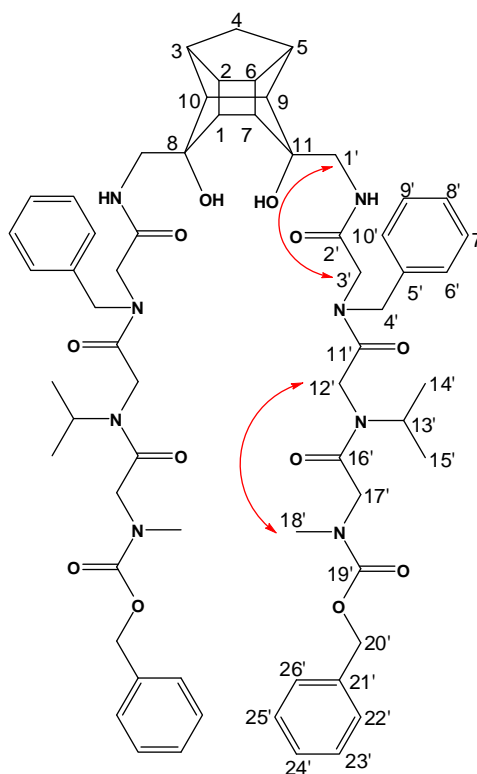


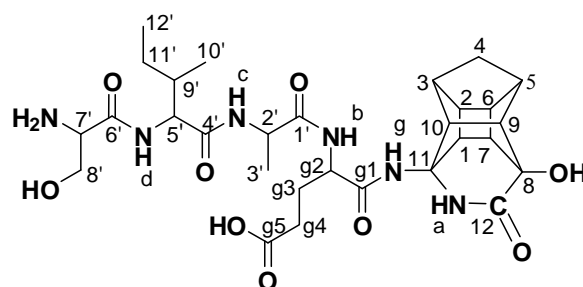
Table 1. ^1H chemical shift and EASY-ROESY correlations of inhibitor 2a. The chemical shift was referenced to 2.50 ppm and the long-range interactions through space, measured by ROESY, are marked in red.

Atom	$\delta^1\text{H}^{\text{a,b}}$	ROESY ($\delta^1\text{H}^{\text{a,b}}$)
1/7	2.09-2.63	4a, 4s
2/6		
3/5		
9/10		
8/11		
4a	1.44	3/5, 4s
4s	1.78	3/5, 4a
1'	3.41	NH, 3'
NH	8.29	1', 3'
2'	-	-

3'	4.03	1', 4', 6'-10', 12', 14', 15', NH
4'	4.46	3', 6'-10'
5'	-	-
6'-10'	7.23-7.38	3', 4'
11'	-	-
12'	4.22	3', 13', 14', 15', 18'
13'	4.66	12', 14', 15'
14'	0.85	3', 12', 13'
15'	0.99	3', 12', 13'
16'	-	-
17'	4.66	
18'	2.83	12'
19'	-	-
20'	5.02, 5.08	22'-26'
21'	-	-
22'-26'	7.23-7.38	20'

a Solvent (CD₃)₂SO.

b 600 MHz for ¹H.



Supplementary Table 2. EASY ROESY correlations observed in the spectra of PCU-EAIS recorded in DMSO- d_6 and the D_2O as the solvent. The chemical shift was referenced to 2.50 ppm

Atom	1H PCU-EAIS	ROESY(δ^1H^c)(PCU-EAIS) (DMSO- d_6)	ROESY(δ^1H^c)(PCU-EAIS (D_2O))
1x	2.89	2, 3, 6, 7, NHc	2, 3, 7, NHa, NHc, NHcc
1y	2.97	2, 3, 6, 7, 10, NHa, NHg	2, 3, 7, NHa
2	2.45	1a, 1s, NHa, NHb	1a, 1s
3	2.41	1a, 1s, 3', 4a, 4s, 6, 10, 10', NHa, NHb, NHc, NHg	1a, 1s, 4a, 4s, 10, NHa, NHc, NHcc
4a	1.35 (d. $J = 10.6$ Hz)	3, 4s, 5, 6	3, 4s, 5, 6
4s	1.68	3, 4a, 5, 6, 9, 10	3, 4a, 5, 6, 9, 10
5	2.66	4a, 4s, 6, 9	4a, 4s, 6
6	2.75	1a, 1s, 2, 3, 4a, 4s, 5, 7, 9	4a, 4s, 5, 7, 9
7	2.34	1a, 1s, 6	1a, 1s, 6
8			
9	2.17	4s, 5, 6, 10	4s, 6, 10
10	2.77	1s, 3, 4s, 9, 10, NHa, NHb, NHc	3, 4s, 9, NHa, NHc

11			
12			
g1'			
g2'	4.22	g3a', g3s, g4, NHa	g3a', g3s, g4, NHa, NHg
g3a'	1.72	g2', g3s, g4, NHa, NHg	g2', g3s, g4, NHg, NHbb,
g3s'	1.89	g2', g3a, g4, NHa, NHg	g2', g3a, g4, NHg
g4'	2.21	3', g2', g3a', g3s', 10', NHa	g2', g3a', g3s', NHg
g5'			
1'			
2'	4.27	3', NHb, NHc, NHcc	3', NHbb, NHc, NHcc
3'	1.19	2', 3, g4', NHb, NHc, NHcc, NHg	2', NHbb, NHc, NHcc
4'			
5'	4.25	9', 10', 11a, 11s, NHd	9', 11a', 11s', 12'
6'			
7'	3.93	8a', 8s', 10', 12', NHd	8a', 8s', NHd
8a'	3.62	7', 8s', NHd	7', 8s', NHd
8s'	3.72	7', 8a', NHd	7', 8a', NHd
9'	1.73	5', 10', 11a', 11s', 12', NHc, NHcc, NHd	5', 10', 11a', 11s', 12', NHc, NHcc, NHd
10'	0.86	3, g4', 5', 7, 9', 11a', NHc, NHcc	9', 11a', NHcc
11a'	1.05	5', 9', 10, 11s', 12', NHd	5', 9', 10, 11s', NHd
11s'	1.43	5', 7, 9', 11a', 12' NHd, NHc	5', 9', 11a', 12'

12'	0.82	5', 9', 11a', 11s', NHcc, NHd	5', 9', 11a', 11s'
NHa	8.08	1a, 1s, 2, 3, 10, g2', g3a', g3s, g4'	1a, 1s, 2, 3, 10, g2'
NHaa			
NHg	8.10	1s, 3, 3', g3a', g3s	g2', g3a', g3s, g4'
NHb	7.80	2', 2, 3, 3', g2', g3a', g3s, g4'	
NHbb	7.94	2', 2, 3, 3', g2', g3a', g3s, g4'	2', 3', g3a'
NHc	8.16	1a, 2', 3', 3, 6, 9', 10', 11s'	1a, 2', 3', 3, 9', 10'
NHcc	8.20	2', 3', 9', 10', 12'	1a, 2', 3, 3', 3, 9', 10'
NHd	8.40	2'/5', 7', 8a', 8s', 9', 11a', 11s', 12'	7', 8a', 8s', 9', 11a'

c 600 MHz for ^1H .

Analysis Info

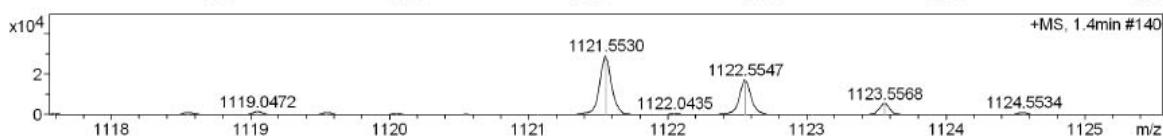
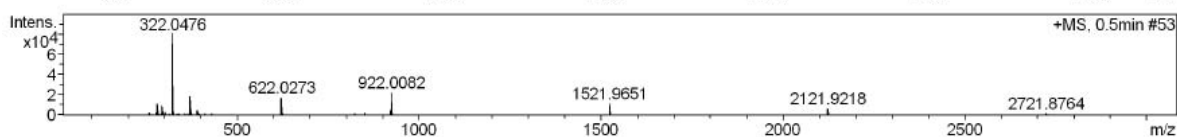
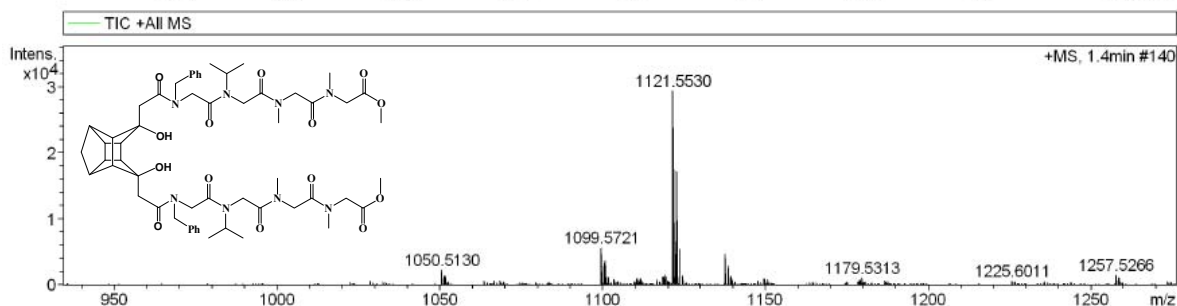
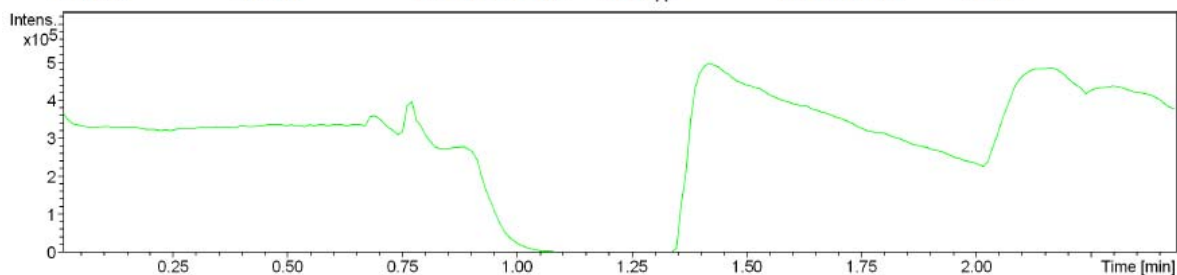
Analysis Name D:\Data\maya\PCU-PVAA peptoid.d
 Method tune_wide_expert.m
 Sample Name PCU-PVAA peptoid
 Comment

Acquisition Date 10/9/2010 4:36:59 PM

Operator BDAL@DE
 Instrument / Ser# micrOTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	14.5 psi
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



Meas. #	Formula	m/z	err [ppm]	Me an err [ppm]	rdB	N - R u l e	e ⁻ Con f	mSig ma	Std I	Std Mean m/z	Std I VarN orm	Std m/z Diff	Std Comb Dev
1121.5530		1121.5530											

HRMS spectrum of compound **1a**

Analysis Info

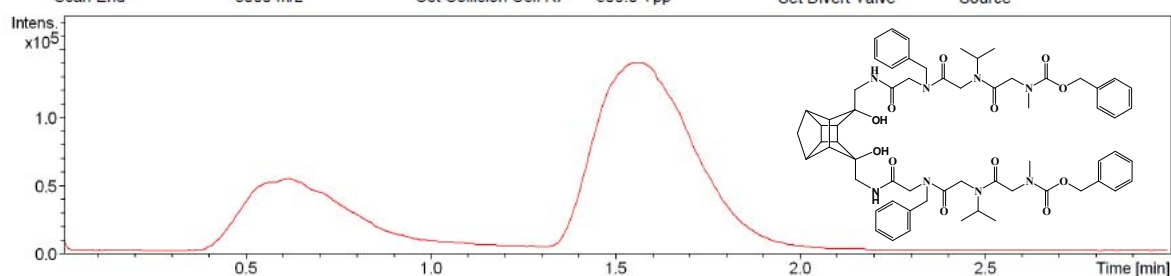
Analysis Name D:\Data\maya\PCUPVAZ PEPTOID000001.d
 Method tune_wide_expert.m
 Sample Name PCUPVAZ PEPTOID
 Comment 1139

Acquisition Date 7/30/2010 11:27:17 PM

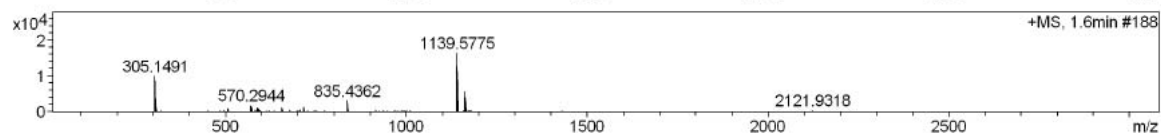
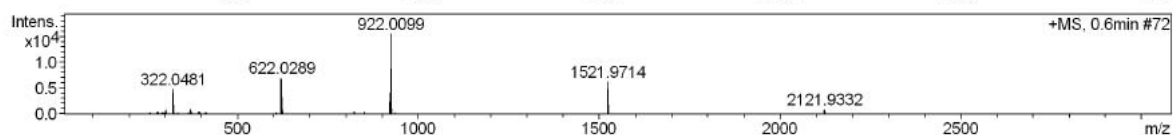
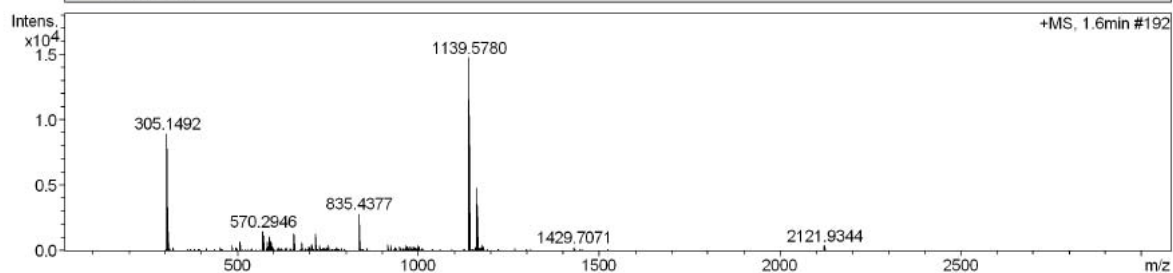
Operator BDAL@DE
 Instrument / Ser# microTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



— TIC +All MS



Meas. m/z	#	Formula	m/z	err [ppm]	Mean err [ppm]	rdb	N-Rule	e ⁻ Conf	mSig ma	Std I	Std Mean m/z	Std I VarNo rm	Std m/z Diff	Std Comb Dev
-----------	---	---------	-----	-----------	----------------	-----	--------	---------------------	---------	-------	--------------	----------------	--------------	--------------

HRMS spectrum of compound 2a

Analysis Info

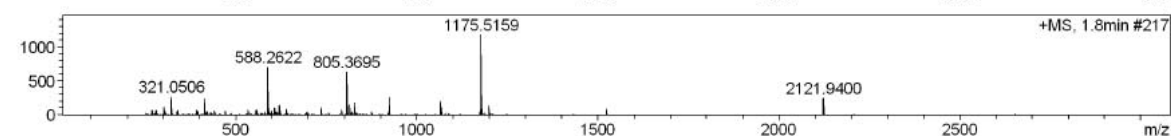
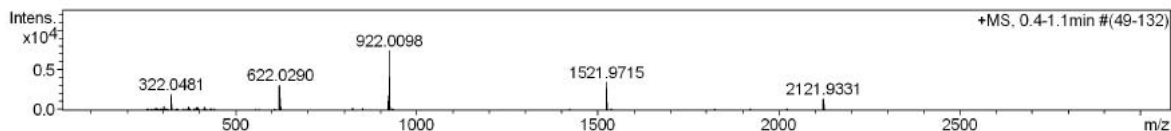
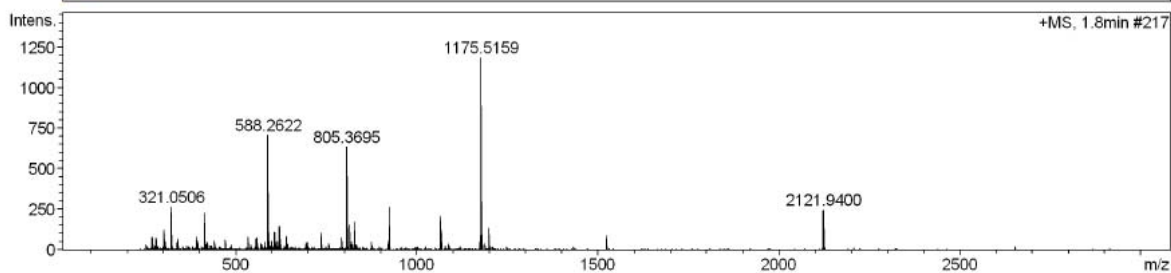
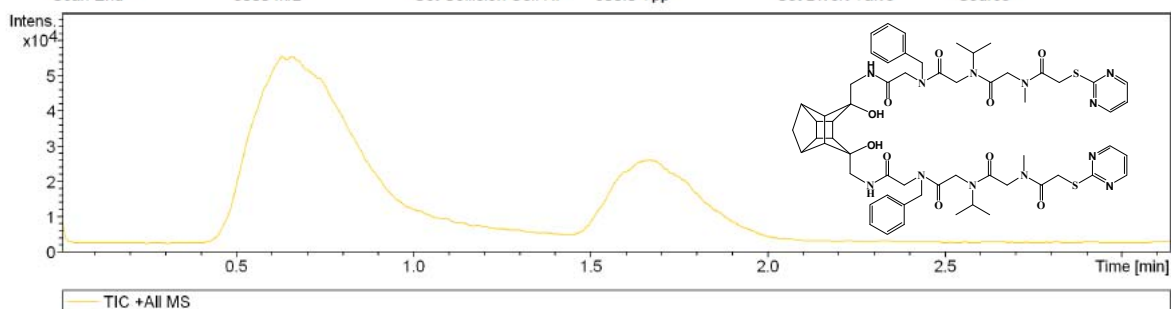
Analysis Name D:\Data\maya\PCUPVATA000001.d
 Method tune_wide_expert.m
 Sample Name PCUPVATA
 Comment 1194

Acquisition Date 7/30/2010 11:40:49 PM

Operator BDAL@DE
 Instrument / Ser# microTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source

HRMS spectrum of compound **2b**

Analysis Info

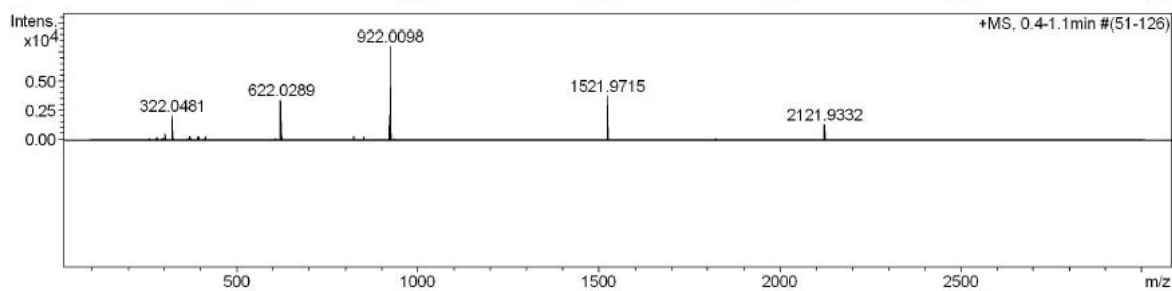
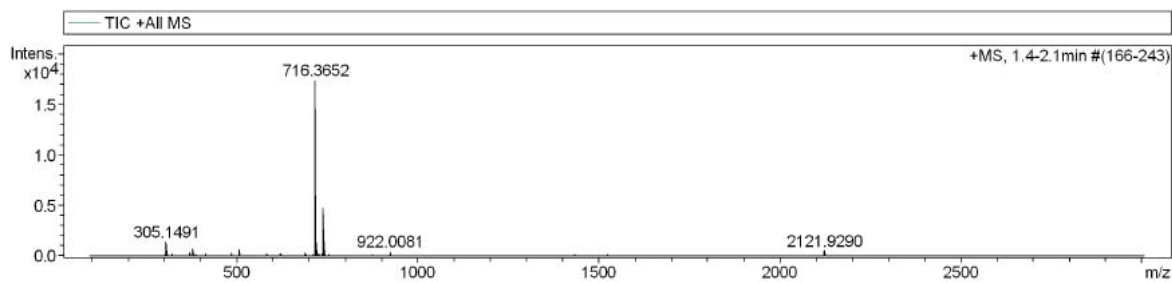
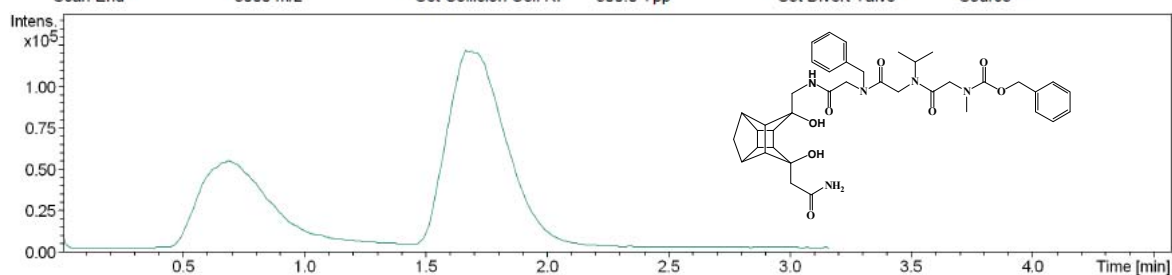
Analysis Name D:\Data\maya\PCUOPVAZ PEPTOID000001.d
 Method tune_wide_expert.m
 Sample Name PCUOPVAZ PEPTOID
 Comment 716

Acquisition Date 7/30/2010 11:22:33 PM

Operator BDAL@DE
 Instrument / Ser# microTOF-Q 10139

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



Meas. m/z	#	Formula	m/z	err [ppm]	Mean err [ppm]	rdb	N-Rule	e ⁻ Conf	mSig ma	Std I	Std Mean m/z	Std I VarNo rm	Std m/z Diff	Std Comb Dev
-----------	---	---------	-----	-----------	----------------	-----	--------	---------------------	---------	-------	--------------	----------------	--------------	--------------

HRMS spectrum of compound 3a

Analysis Info

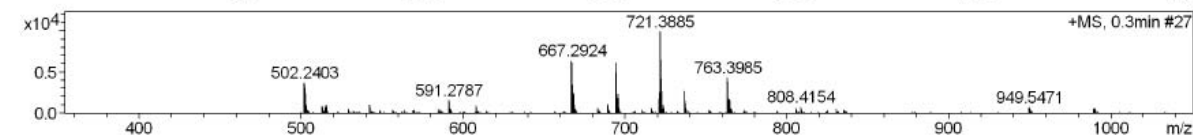
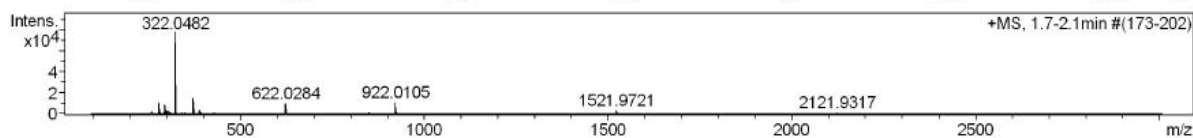
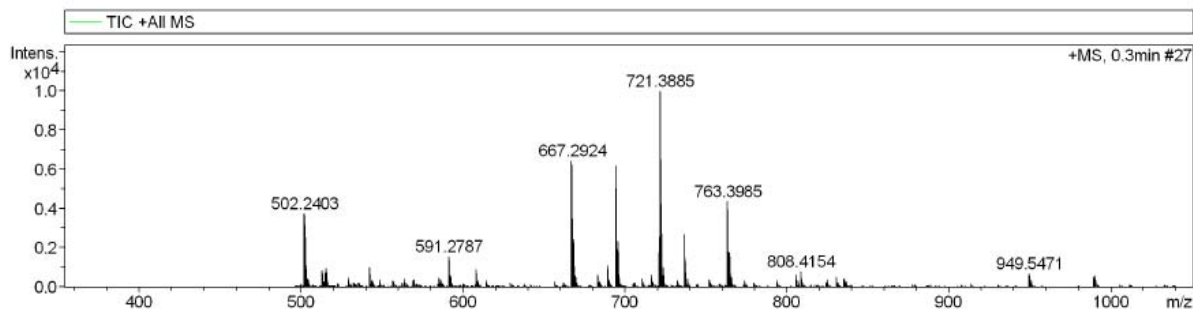
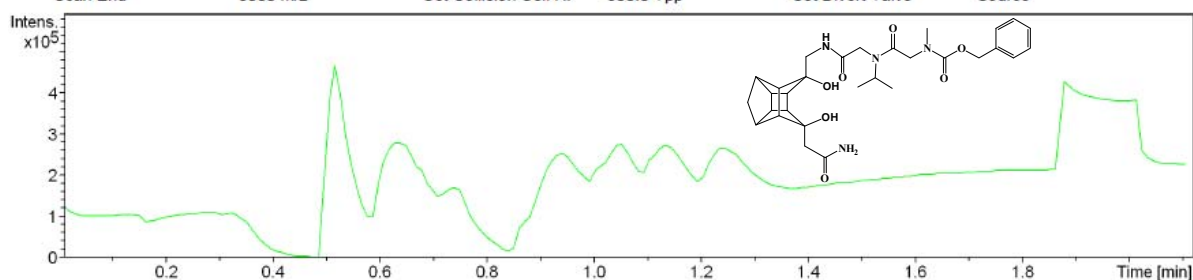
Analysis Name D:\Data\maya\OVAZ peptoid.d
 Method tune_wide_expert.m
 Sample Name OVAZ peptoid
 Comment

Acquisition Date 10/9/2010 5:38:07 PM

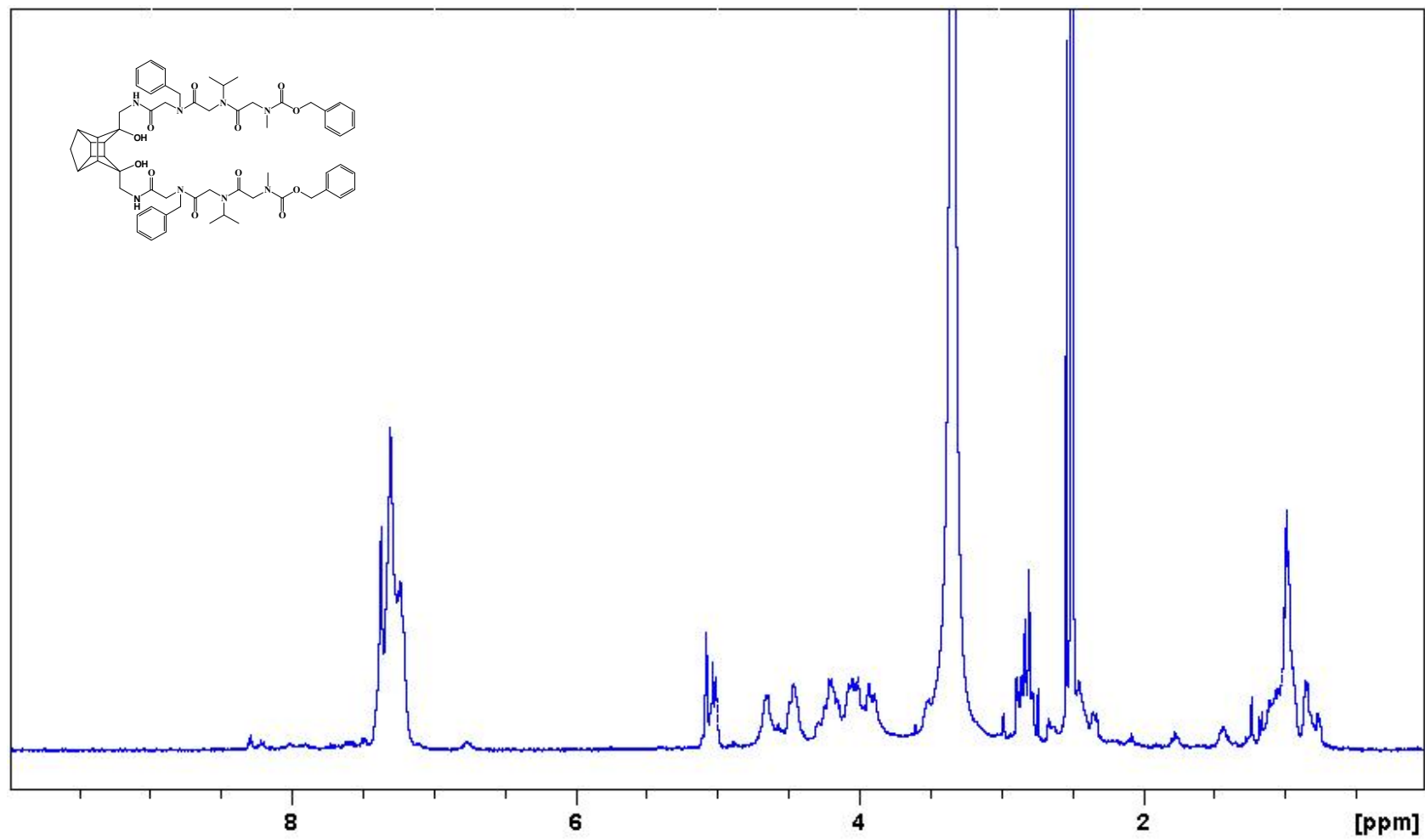
Operator BDAL@DE
 Instrument / Ser# microTOF-Q 10139

Acquisition Parameter

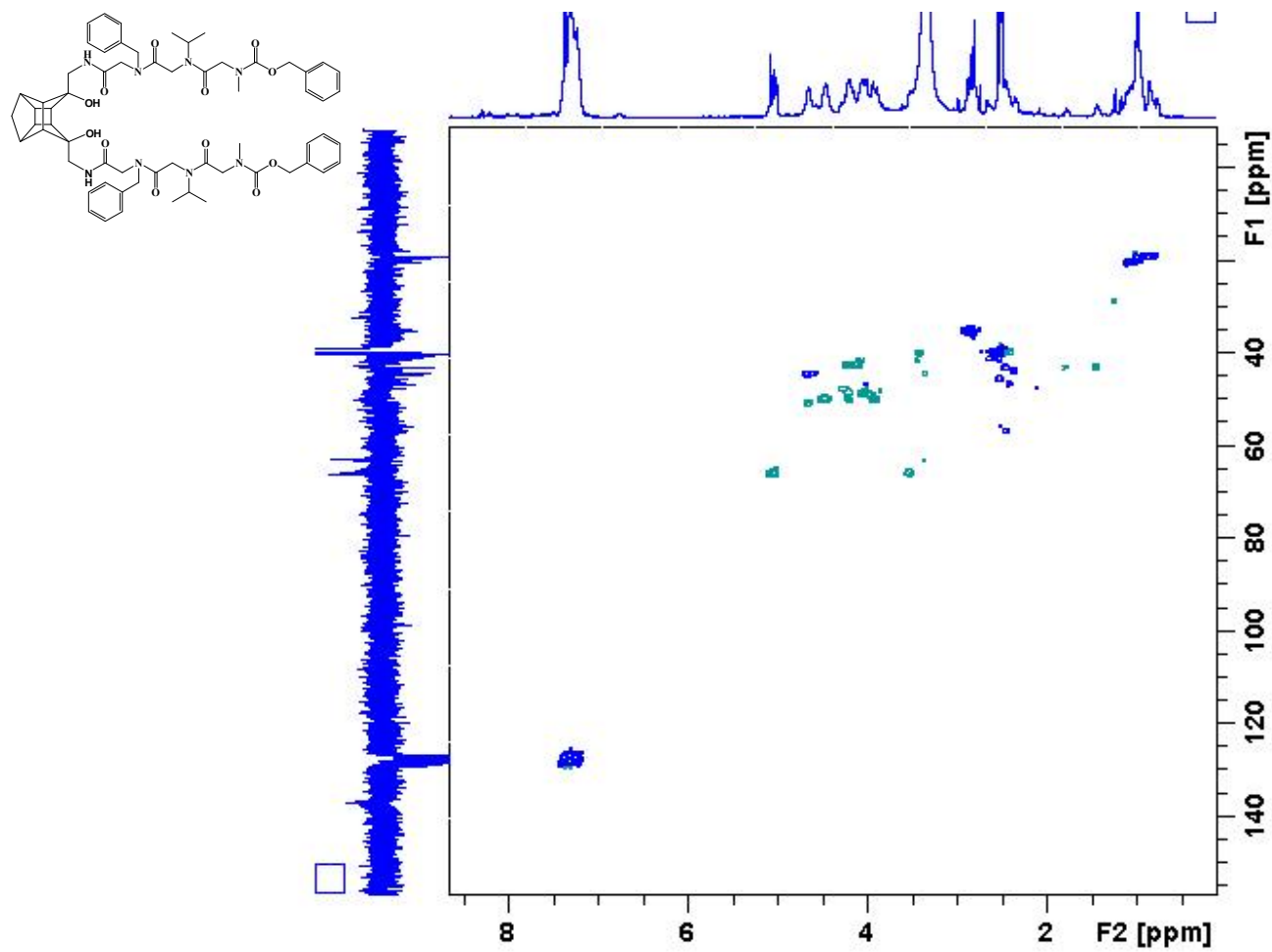
Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	14.5 psi
Focus	Not active	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan Begin	500 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



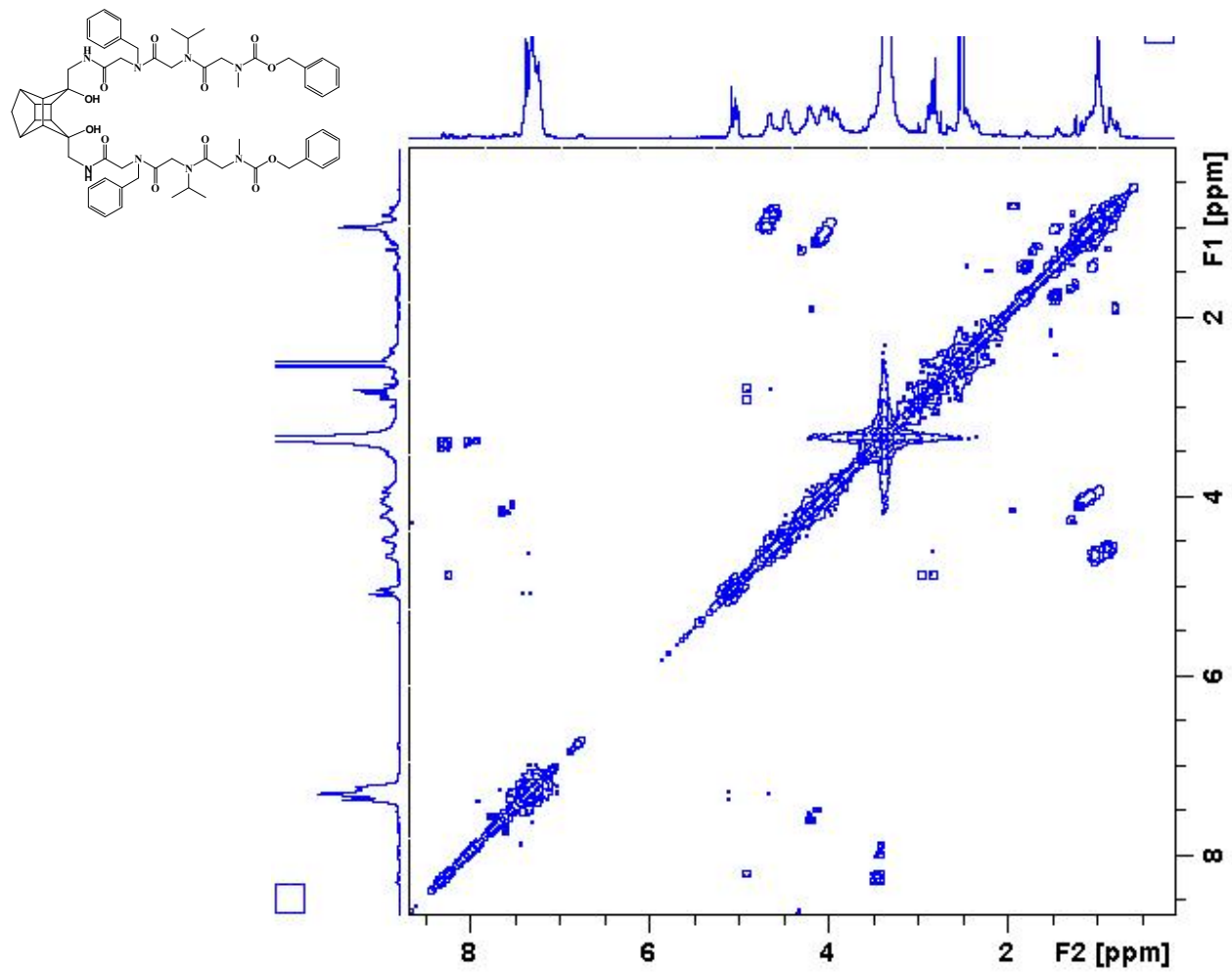
Meas. #	m/z	Formula	m/z	err [ppm]	Mean err [ppm]	rdb	N-R	e ⁻ Conf	mSig	Std I	Std Mean	Std I Var	Std m/z	Std Comb
591.2787	1	C ₃₃ H ₄₁ N ₂ Na ₂ O ₅	591.2805	3.1	2.2	13.5	ok	even	21.94	0.0325	0.0017	0.0157	0.0025	0.7766
	2	C ₃₀ H ₄₀ N ₄ Na ₂ O ₇	591.2789	0.4	-0.6	12.5	ok	even	22.91	0.0384	0.0011	0.0171	0.0025	0.7467
	3	C ₃₁ H ₃₆ N ₈ Na ₂ O ₃	591.2803	2.7	1.6	17.5	ok	even	23.73	0.0357	0.0015	0.0171	0.0026	0.7728



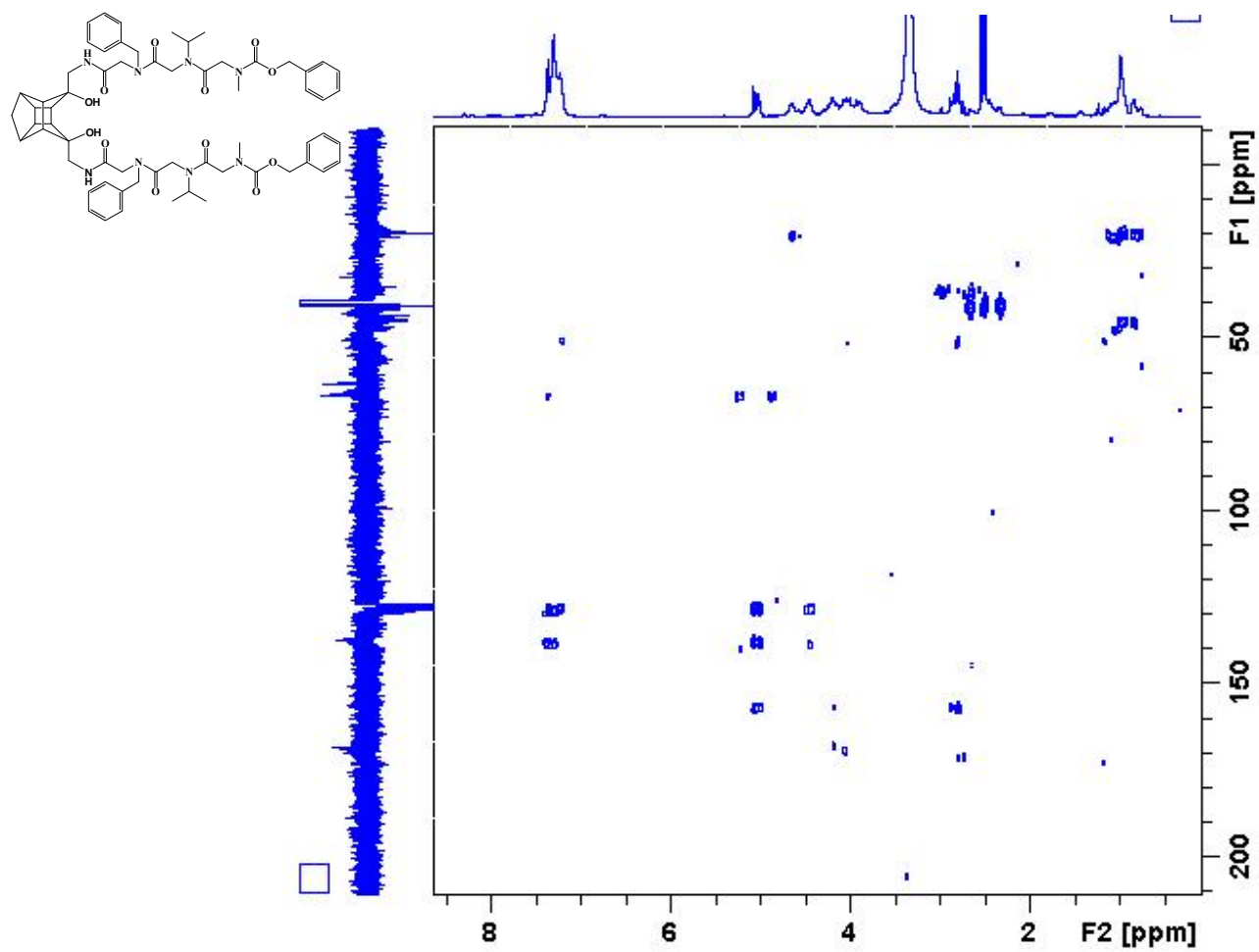
^1H NMR spectrum of PCU-PVAZ peptoid



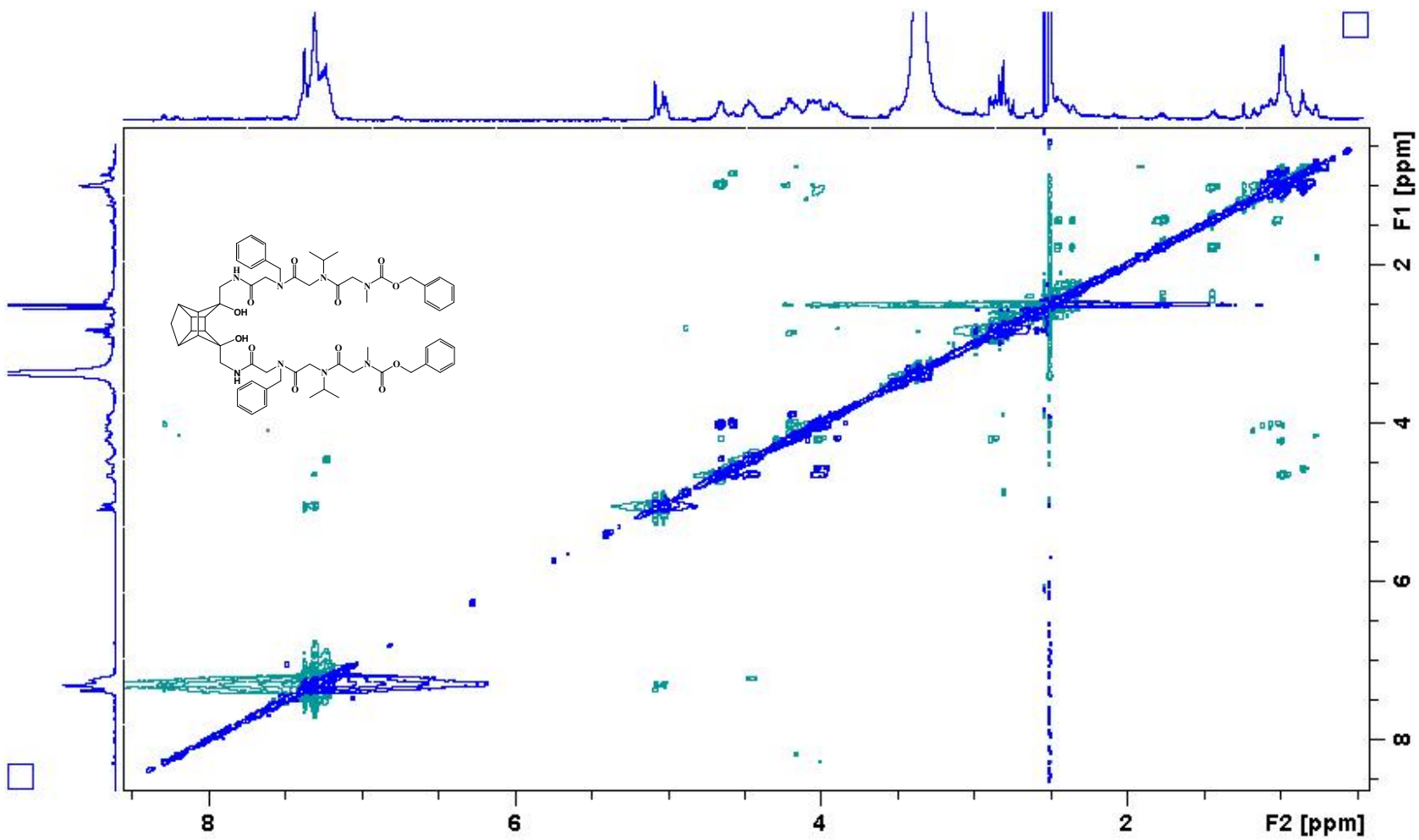
HSQC spectrum of PCU-PVAZ peptoid



COSY spectrum of PCU-PVAZ peptoid



HMBC spectrum of PCU-PVAZ peptoid



ROESY spectrum of PCU-PVAZ peptoid