UNIVERSITY OF KWAZULU-NATAL

DESIGN, SYNTHESIS, AND BIOLOGICAL EVALUATION OF NOVEL PENTACYCLO UNDECANE DERIVED PEPTIDES/PEPTOIDS AS POTENTIAL HIV-1 PROTEASE INHIBITORS.

DISSERTATION

2012

RAJSHEKHAR KARPOORMATH

DEDICATED

TO

MY BELOVED PARENTS

SHRI. V. S. KARPUR

AND

SMT. MANASA KARPUR

To raise new questions, new possibilities, to regard old problems from a new angle, requires creative imagination and marks real advance in science.

Albert Einstein

DESIGN, SYNTHESIS, AND BIOLOGICAL EVALUATION OF NOVEL PENTACUCLO UNDECANE DERIVED PEPTIDES/PEPTOIDS AS POTENTIAL HIV-1 PROTEASE INHIBITORS.

RAJSHEKHAR KARPOORMATH

2012

A thesis submitted to the School of Chemistry and Physics, College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Westville, for the degree of Doctor of Philosophy.

This thesis has been prepared according to **Format 4** as outlined in the guidelines from the Faculty of Science and Agriculture which states:

> This is a thesis in which the chapters are written as a set of discrete research papers, with an overall introduction and final discussion, where one or all of the chapters have already been published. Typically these chapters will have been published in internationally recognized, peer-reviewed journals.

As the Dean and HOD, School of Chemistry and Physics, I have/have not approved this thesis for submission.

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

Supervisor:

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

Co-Supervisor:

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

BIOGRAPHICAL SKETCH

Rajshekhar Karpoormath was born in Hubli (Karnataka, India) on 02 November 1975. He joined the faculty of Pharmaceutical Chemistry, University of KwaZulu-Natal in July 2010.

- After his secondary education, Rajshekhar was awarded Bachelor of Pharmacy degree from K.L.E. Society's College of Pharmacy, Belgaum affiliated with Karnataka University, Dharwad, India.

- His interest in Pharmaceutical Chemistry led him to register for Masters degree in Pharmaceutical sciences with major in Pharmaceutical Chemistry and Drug design. During this period, his research involved "*Synthesis and anti-tubercular activity of a series of 2-sulfonamido/ trifluoromethyl-6- substituted imidazo (2,1-b)-1,3,4-thiadiazole derivatives*." This work was then published in a peer reviewed journal (*Bioorganic Medicinal Chemistry, 2004, 12, 5651-5659*) for which he was awarded his Masters degree in 2004.

- In March 2004, he was appointed as lecturer in the Department of Pharmaceutical Chemistry in Dayanand Sagar College of Pharmacy, Bangalore, India.

- With an urge to further pursue his research interest, he joined National Chemical Laboratory, CSIR, Pune, India as a Senior Research Fellow (June 2005 - July 2006). At this institute his research was focused on "*structural and conformational studies of tRNA and the significance of modified nucleosides present in tRNA at various positions*." Outcomes of this research were presented in National and International Conferences.

- From Aug 2006, Rajshekhar joined the School of Chemistry and Physics at the University of KwaZulu-Natal for his Doctoral studies under the supervision of Prof. H. G. Kruger and Dr. Glenn Maguire. His research work titles, "*Design, Synthesis, and Biological Evaluation of Novel Pentacuclo Undecane derived Peptides/Peptoids as potential HIV-1 protease inhibitors*."

- Exposure in this research and academic environment enabled him to supervise both Honors and Masters students. In Nov 2011, he was awarded the Competitive

Research Grant, which heightened his interest to pursue independent research in University of KwaZulu-Natal's Department of Pharmaceutical Chemistry.

- Rajshekhar is a fun loving and sociable personality. In his spare time he enjoys sports and spending time with his family.

ABSTRACT

This study reports a series of promising and structurally diverse potential HIV-1 protease inhibitors.

Human Immunodeficiency Virus (HIV) is the causative agent of Acquired Immune Deficiency Syndrome (AIDS). HIV infection disrupts the immune system and makes the body susceptible to opportunistic infections. If untreated, AIDS is generally fatal. Today, AIDS has become a long lasting pandemic. According to the World Health Organization (WHO) and Joint United Nations Program (UNAIDS-2009) report, it is estimated that 33.3 million men, women and children worldwide are infected with HIV. This situation is steadily deteriorating in some parts of the world compared to the previous years. One of the major drawbacks associated with the currently FDA-approved anti-HIV drugs are severe side effects, toxicities, high dosage and high treatment cost. Thus, an urgent need for new drugs to combat HIV is apparent.

In the first part of the study, research efforts were focused to synthesize potent pentacycloundecane (PCU) derived peptide and peptoids as protease inhibitors. It is proposed that these inhibitors bind to wild type C-South African HIV protease (C-SA) catalytic site *via* a non-cleavable or non-hydrolysable cyclic ether bond for the first polycyclic cage compound and *via* a dihydroxylethelene type functional group for the second cage compound. The desired compounds were synthesized by coupling of the peptides and peptoids to the PCU derived cage.

Second part of the study involves, biological evaluation against wild type C-SA enzyme and characterization of the synthesized compounds by Nuclear Magnetic Resonances (NMR). All the synthesized novel compounds were evaluated against wild type C-SA enzyme for their ability to inhibit 50% of the enzyme's activity (IC_{50}). Some of the compounds reported herein showed promising activity by inhibiting the enzyme activity at concentrations of less than 0.6 nM. 2D NMR investigations employing a new Efficient Adiabatic Symmetrized Rotating Overhauser Effect Spectroscopy (ROESY / NOESY) technique enabled the attainment of vital information about the 3D structure of these small linear peptides and peptoids in solution. The activity could be related to conformations induced by the PCU moiety on the coupled peptide side chain. Further quantum mechanics/molecular mechanics/molecular dynamics (QM/MM/MD) simulations were carried out to confirm the observed NMR experimental results.

Docking studies were performed for the synthesized compounds. Binding energies obtained from the docking calculations were then used to further validate the experimental IC_{50} results. These experimental and theoretical methods provided valuable insight into the interaction mode of these cage peptide and peptoids inhibitors with the enzyme.

DECLARATIONS

DECLARATION 1 – PLAGIARISM

I, Rajshekhar Karpoormath, 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

DECLARATION 2 - PUBLICATIONS AND CONTRIBUTIONS

DETAILS OF CONTRIBUTION TO PUBLICATIONS that form part and/or include research presented in this thesis (include publications in preparation, submitted, *in press* and published and give details of the contributions of each author to the experimental work and writing of each publication)

Publication 1

Rajshekhar Karpoormath, Yasien Sayed, Thavendran Govender, Hendrik G. Kruger, Glenn E. M. Maguire,* and Mahmoud E. S. Soliman*; *Novel PCU cage diol peptides as potential targets against wild type CSA HIV-1 protease; Synthesis, biological screening and molecular modeling studies*; Medicinal Chemistry Research, **Accepted** (**In Press**).

Contributions: I have synthesized all the compounds, carried out the biological evaluations against HIV-1 protease enzyme and wrote the manuscript. All other authors are supervisors.

Publication 2

Rajshekhar Karpoormath, Oluseye K. Onajole, Tricia Naicker, Thavendran Govender, Glenn E. M. Maguire, and Hendrik G. Kruger*; *Synthesis and NMR elucidation of Novel Pentacycloundecane-Derived Peptides*; South African Journal of Chemistry, **2012**, **Accepted** (**In Press**).

Contributions: I have solely carried out the synthesis, 2D NMR experiments and wrote the entire manuscript. Dr. Oluseye K. Onajole assisted in 2D NMR elucidations of PCU diol diphenylalanine, rest all of the compounds were elucidated by me. Dr. Tricia Naicker proofread the manuscript. All other authors are supervisors.

Publication 3

Rajshekhar Karpoormath, Yasien Sayed, Patrick Govender, Thavendran Govender, Hendrik G. Kruger, Mahmoud E. S. Soliman* and Glenn E. M. Maguire*; *Pentacycloundecane derived hydroxy acid peptides: A new class of irreversible non-scissile ether bridged type isoster as potential HIV-1 wild type C-SA protease inhibitors*, Bioorganic Chemistry, **2012**, 40, 19-29.

Contributions: I contributed to the original idea, synthesised all compounds, characterised, carried out biological evaluations against HIV-1 protease enzyme, completed computational studies, wrote the introduction, results, discussion and experimental for the manuscript. All other authors are supervisors.

Publication 4

Rajshekhar Karpoormath, Fernando Albericio, Thavendran Govender, Glenn E. M. Maguire, and Hendrik G. Kruger*; *Synthesis and NMR elucidation of Pentacycloundecane derived hydroxy acid peptides as potential anti-HIV-1 agents*, Structural Chemistry, **Accepted** (**In Press**).

Contributions: Synthesis, 2D NMR experiments, structural elucidations of all compounds and the manuscript was written by me. All other authors are supervisors.

Other Publications

1. R. Karpoormath, T. Naicker, T. Govender, H.G. Kruger, G.E.M. Maguire; *Benzyl 5-hydroxy-4-oxapentacyclo[5.4.1.$0^{2,6}$.$0^{3,10}$.$0^{8,11}$]dodecane-3-carboxylate*, Acta Crystallographica Section E-Structure Reports Online, **2011**, 67, O877-U539.

2. R. Karpoormath, P. Govender, T. Govender, H.G. Kruger, G.E.M. Maguire; *endo-11-(Dibenzylamino)tetracyclo[5.4.0.0^{3,10}.$0^{5,9}$]undecane-8-one*, Acta Crystallographica Section E-Structure Reports Online, **2011**, 67, O619-U1716.

3. R. Karpoormath, T. Govender, P. Govender, H.G. Kruger, G.E.M. Maguire; *tert-Butyl N-[(11-exo-benzyloxycarbonyl-8-oxopentacyclo[5.4.0.0^{2,6}.$0^{3,10}$.$0^{5,9}$]undecane-11-endo-yloxy)carbonyl-methyl] carbamate*, Acta Crystallographica Section E-Structure Reports Online, **2010**, 66, O2607-U2900.

4. R. Karpoormath, P. Govender, H.G. Kruger, T. Govender, G.E.M. Maguire, *N, N '-[(8-endo,11-endo-Dihydroxypentacyclo[5.4.0.0^{2,6}.$0^{3,10}$.$0^{5,9}$]undec ane-8,11-diyl)bis(methylenecarbonyl)]di-L-phenylalanine*, Acta Crystallographica Section E-Structure Reports Online, **2010**, 66, O2537-U2268.

5. G.A. Boyle, T. Govender, R. Karpoormath, H.G. Kruger, *exo-8, exo-11-Divinylpentacyclo[5.4.0.0^{2,6}.$0^{3,10}$.$0^{5,6}$]undecane-endo -8, endo-11-diol*, Acta Crystallographica Section E-Structure Reports Online, **2007**, 63, O3977-U2352.

6. G.A. Boyle, T. Govender, R. Karpoormath, H.G. Kruger, *exo-8, exo-11-Diallylpentacyclo-[5.4.0.0^{2,6}.$0^{3,10}$.$0^{5,9}$]undecane-end o-8,endo-11-diol*, Acta Crystallographica Section E-Structure Reports Online, **2007**, 63, O4797-U5875.

Contributions: In the above publications **1**-**6**, I carried out synthesis, growing suitable crystal samples and wrote the manuscript. In publications **1**, **5** and **6**, Dr. T. Naicker and Dr. G.A. Boyle proofread the manuscripts. All others are supervisors.

ACKNOWLEDGEMENTS

Completion of this Doctoral thesis was possible with support of a few special people and I express here my sincere and heartfelt gratitude to them.

- No research would be possible without the resources and services the University of KwaZulu-Natal provides its students. I express my humble gratitude to this esteemed University.

- Prof. Andrew Kindness, Dean and HOD, School of Chemistry and Physics, for his valuable guidance and consistent encouragement throughout the completion of my PhD. I thank you Sir.

- Prof. Fernando Albericio for his compassionate and treasured guidance during my research. You have been a great inspiration, I thank you Sir.

- Prof. H.G.Kruger and Dr. Glenn Maguire, I thank you both for all your support rendered until the completion.

- I also would like to extend my gratitude to Dr. Patrick Govender (Dept. of Biochemistry, UKZN) and Dr. M.E.S. Soliman (Dept. of Pharmaceutical Chemistry, UKZN) for their valuable assistance during my research.

- My study would not have been successful without the help I received from the technical staff of the School of Chemistry and Physics. I wholeheartedly thank each and every one of you.

- My family has been my support system, Mr. V. S. Karpur (Father), Mrs. Manasa Karpur (Mother), Ms. Vijaya Karpur (Sister), Mrs. Shweta Hiremath (Sister), Mr. Saishankar Karpoormath (Brother), Mrs. Sharadha Karpoormath (Sister-in-law) Mr. Shanmukh Hiremath (Squandran Leader, Indian Air Force) and my in-laws Mr. & Mrs Ashwathnarayana. I will always be greatful to you for having stood by me throughout this period and for being my strength, courage and moral support during my research.

- My little adorable angels, Shradha Hiremath, Shubam Hiremath, Shashi Karpoormath, Shreyas Karpoormath, James Koorbanally, Rachel Pocock Giladi, Jonah Pocock Giladi, Gabrieal Marks, Jamie Marks and Mohammed Shaikh, you guys made me smile even at most stressful and difficult times. I can't thank enough for the love and affection.

- No words of appreciation can fulfill the status of my dear wife, Pruthvi. I should not belittle you by thanking, as you are so essential in my life. I must mention that my personal and professional life has changed being with you. I could not have completed my research without your love, respect, patience, appreciation, faith and help. They kept my spirits alive. It was you and only you who transformed my life drastically in every possible way.

- To my dear friends and collegues, Dr. Neil and Dr. Chantal Koorbanally, Dr. Brenda Moodley and family, Ms Jaymbal Govender, Prof. Monique Marks and family, Mr. Jean-Francois, Mr. Joe Grant, Mr. Richard Pocock, Mrs. Karen Giladi, Mr. Jason Nicholas Paraskevoaoulos, Dr. Sai Kumar Chakka, Dr. Mahidansha Mahiboob Shaikh and family Mr. J. P Oliver, Dr. Jane Oliver, Dr. Y. Naidoo and Dr. Alok Sen (NCL, Pune, India), I take this opportunity to whole heartedly thank them for making all the worst times more bearable and believing in me.

- I would like to extend my friendly & faithful appreciation to my lab mates Mrs. Monisola Ikhile, Mr. Michael Nivendran Pillay, Mr. Vicent Sithole, Mr. Vashen Moodely, Mr. Shirveen Sewpersad, and Mr. Fezile Potwana, for their support and encouragement throughtout. It was a pleasure working with you all.

- I would also like to take this opportunity to extend my gratitude to Prof. Sabiha Essack, Dean of Health Sciences, and my colleagues in the School of Pharmaceutical Sciences. I thank you one and all.

- Lastly I would like to thank everyone I have not mentioned who has helped me in any way to complete my work and more especially those who have motivated me to finish.

- Above all, I thank you God for deeming me fit and granting me the knowledge, will, strength and health to undertake this task and enabling me to its completion.

TABLE OF CONTENTS

LIST OF PUBLICATIONS

1. Rajshekhar Karpoormath, Yasien Sayed, Thavendran Govender, Hendrik G. Kruger, Glenn E. M. Maguire,[*] and Mahmoud E. S. Soliman[*]; *Novel PCU cage diol peptides as potential targets against wild type CSA HIV-1 protease; Synthesis, biological screening and molecular modeling studies*; Medicinal Chemistry Research, **Accepted** (**In Press**).

2. Rajshekhar Karpoormath, Oluseye K. Onajole, Tricia Naicker, Thavendran Govender, Glenn E. M. Maguire, and Hendrik G. Kruger[*]; *Synthesis and NMR elucidation of Novel Pentacycloundecane-Derived Peptides*; South African Journal of Chemistry, **2012**, **Accepted** (**In Press**).

3. Rajshekhar Karpoormath, Yasien Sayed, Patrick Govender, Thavendran Govender, Hendrik G. Kruger, Mahmoud E. S. Soliman[*] and Glenn E. M. Maguire[*]; *Pentacycloundecane derived hydroxy acid peptides: A new class of irreversible non-scissile ether bridged type isoster as potential HIV-1 wild type C-SA protease inhibitors*, Bioorganic Chemistry, **2012**, 40, 19-29.

4. Rajshekhar Karpoormath, Fernando Albericio, Thavendran Govender, Glenn E. M. Maguire, and Hendrik G. Kruger[*]; *Synthesis and NMR elucidation of Pentacycloundecane derived hydroxy acid peptides as potential anti-HIV-1 agents*, Structural Chemistry, **Accepted** (**In Press**).

OTHER PUBLICATIONS

1. R. Karpoormath, T. Naicker, T. Govender, H.G. Kruger, G.E.M. Maguire; *Benzyl 5-hydroxy-4-oxapentacyclo[5.4.1.$0^{2,6}$.$0^{3,10}$.$0^{8,11}$]dodecane-3-carboxylate*, Acta Crystallographica Section E-Structure Reports Online, **2011**, 67, O877-U539.

2. R. Karpoormath, P. Govender, T. Govender, H.G. Kruger, G.E.M. Maguire; *endo-11-(Dibenzylamino)tetracyclo[5.4.0.0^{3,10}.$0^{5,9}$]undecane-8-one*, Acta Crystallographica Section E-Structure Reports Online, **2011**, 67, O619-U1716.

3. R. Karpoormath, T. Govender, P. Govender, H.G. Kruger, G.E.M. Maguire; *tert-Butyl N-[(11-exo-benzyloxycarbonyl-8-oxopentacyclo[5.4.0.0^{2,6}.$0^{3,10}$.$0^{5,9}$]undecane-11-endo-yloxy)carbonyl-methyl] carbamate*, Acta Crystallographica Section E-Structure Reports Online, **2010**, 66, O2607-U2900.

4. R. Karpoormath, P. Govender, H.G. Kruger, T. Govender, G.E.M. Maguire, *N, N '-[(8-endo,11-endo-Dihydroxypentacyclo[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]undecane-8,11-diyl)bis(methylenecarbonyl)]di-L-phenylalanine*, Acta Crystallographica Section E-Structure Reports Online, **2010**, 66, O2537-U2268.

5. G.A. Boyle, T. Govender, R. Karpoormath, H.G. Kruger, *exo-8, exo-11-Divinylpentacyclo[5.4.0.0^{2,6}.0^{3,10}.0^{5,6}]undecane-endo-8,endo-11-diol*, Acta Crystallographica Section E-Structure Reports Online, **2007**, 63, O3977-U2352.

6. G.A. Boyle, T. Govender, R. Karpoormath, H.G. Kruger, *exo-8, exo-11-Diallylpentacyclo-[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]undecane-endo-8,endo-11-diol*, Acta Crystallographica Section E-Structure Reports Online, **2007**, 63, O4797-U5875.

CHAPTER 1

1 HIV/AIDS BACKGROUND

One of the greatest threats and challenge ever faced by human kind in the present day is AIDS (Acquired Immune Deficiency Syndrome) epidemic. The causative organism responsible for this disease in humans is thehuman immunodeficiency virus (HIV). The first cases of HIV/AIDS were reported in 1981 in Los Angeles, wherein the patients suffered from rare and unusual opportunistic infections, such as *Pneumocystis carinii* pneumonia and Kaposi's cancer.[1] These rare diseases were previously seen in patients on immunosuppressant therapy but now previously healthy homosexual men, intravenous drug abusers, hemophiliacs, infants and blood transfusion recipients were the main groups affected.[2] In 1983, two researchers (Luc Montagnier and Francoise Barre-Sinoussi) from the Institute Pasteur in Paris (France) successfully isolated and identified the infectious virus responsible for causing AIDS, for which they were awarded the Nobel-Price for Physiology or Medicine in 2008.[3] Initially the virus had been given various names like lymphadenopathy-associated virus (LAV),[4] human T-cell lymphotropic virus type III (HTLV-III),[5] and AIDS- associated retrovirus (ARV).[6] Later the International Committee on the Taxonomy of Viruses proposed the name presently used as Human Immunodeficiency Virus (HIV).[7]

One of the major problems in the early detection of HIV infection is that the patients do not show any symptoms of any opportunistic infection even though they might have contacted the virus years before. Such people can lead a normal life for an extended time.[8] The primary infection is characterized by flu-like symptoms (fever, sweats, headaches, sore throat and rashes) which may last for 3-14 days.[9,10] During acute infection the virus levels are extremely high.[11,12] A rapid fall in virus levels is observed when the immune response sets in. However, the virus levels soon reach a steady state that can persist for months or years.[13] During this stage the virus is still active, replicating and infecting new lymphocytes and macrophages,[14,15] but this activity is unnoticed due to the rapid turnover of virions and cells.[16,17] The estimated life span of an HIV-infected cell is about two days as compared with about eight hours for a viron.[18] The infection

progresses to AIDS as the number of lymphocytes decreases and the viral load increases past the critical level.[19,20] Today, AIDS has become a long lasting pandemic and is steadily increasing in some parts of the world compared to the previous years (Figure 1).[21]

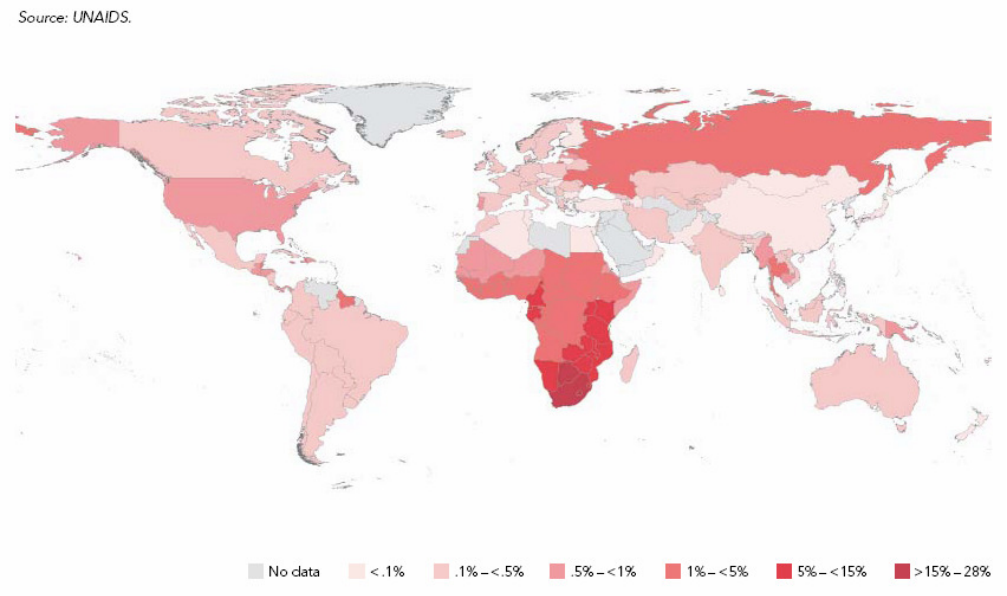


Figure 1. Global prevalence of HIV in 2010- Total 3.3 million.[21]

According to the World Health Organization (WHO) and Joint United Nations Program (UNAIDS) report, it is estimated that 33.3 million men, women and children are infected with HIV. Latest statistics on Global HIV/AIDS 2010 is shown in Table 1.

Table 1. Regional statistics for HIV and AIDS, end of 2009.[21]

Region	Adults & children living with HIV/AIDS	Adults & children newly infected	Adult prevalence*	AIDS-related deaths in adults & children (2009)
Sub-Saharan Africa	22.5 million	1.8 million	5.0%	1.3 million
North Africa & Middle East	460,000	75,000	0.2%	24,000
South and South-East Asia	4.1 million	270,000	0.3%	260,000
East Asia	770,000	82,000	<0.1%	36,000
Oceania	57,000	4,500	0.3%	1,400
Central & South America	1.4 million	92,000	0.5%	58,000
Caribbean	240,000	17,000	1.0%	12,000
Eastern Europe & Central Asia	1.4 million	130,000	0.8%	76,000
North America	1.5 million	70,000	0.5%	26,000
Western & Central Europe	820,000	31,000	0.2%	8,500
Global Total	**33.3 million**	**2.6 million**	**0.8%**	**1.8 million**

There is no known cure or vaccine for AIDS and HIV, though antiretroviral treatment reduces both the mortality and morbidity of HIV infection.[22] High cost and inaccessibility of the antiretroviral treatment in all developing countries was one of the major drawbacks in the fight against AIDS. In the past decade as a result of the considerable joint effort by world organization like WHO (World Health Organization), UNAIDS (Joint United Nations Programme on HIV/Acquired Immune Deficiency Syndrome), UNITAID (United States Agency International Development) and Clinton Foundation (CHAI- Clinton Health Access Initiative) has led to the drop in the number of people dying from AIDS-related causes in several countries.[21,23,24] According to UNAIDS 2012 annual report the number of people acquiring HIV infection in 2011 (2.5 million [2.2 million-2.8 million]) was 20% lower than in 2001. The sharpest decline in new HIV infections was observed in Caribbean (42%) and sub-Saharan

Africa (25%) since 2001.[24] This could be mainly attributed to three main factors: **a**) Accessibility of anti-retroviral therapy (ART) to the Low- and Middle-Income Countries (LMICs), **b**) Low cost of manufacturing anti-retroviral drugs (ARV) and **c**) Promoting greater awareness towards HIV/AIDS infection, therapy and prevention.

UNITAID in 2010 has spent about $12 billion USD on HIV/AIDS to make the anti-retroviral therapy (ART) accessible to the Low- and Middle-Income Countries (LMICs).[23] At the end of 2012 about 8.5 million people in LMICs will have access to ARTs and this account to approximately 50% of the patients in need. This was achieved with spend of about $1.5 billion USD in drug procurements.[23] According to Clinton Foundation, the first line ART is now as low as $75 USD/patient-year. But the second line and salvage therapies that contain protease inhibitors (PIs) are available for about $400/patient-year.[25] This could be attributed to the low-cost quality generic medicines manufactured by Indian pharmaceutical companies, which currently accounts to 80% of these medicines supplied to the majority of developing countries.[21,23,24] Due to the difficulty in treating the HIV infection, preventing infection by promoting safe sex and needle-exchange programs has become the key aim in controlling the AIDS pandemic, with the participation of Government / Non-government health organizations (NGO).

The international organizations (WHO, UNAIDS, UNITAID), NGO's together working with the local governments of the respective countries have been responsible for bringing down the HIV/AIDS drug prices considerably. But the present cost of the HIV/AIDS drugs currently marketed are still too high to support access of these drugs for everyone in need. Hence the aim of this study and the long term goals of our research group are:

- Building potent protease inhibitors (PI) by incorporating an inexpensive scaffold like PCU cage (Pentacycloundecane cage derivatives) into short inexpensive peptides or peptoids, which could be less expensive than making the current important PI drugs.
- It is also possible that the PCU cage moieties incorporated into short peptides or peptoids could have better pharmacokinetic profiles than the existing PI drugs due to its rigid and hydrophobic carbon skeletons.

- Lastly, prevention of disease progression by providing a very high barrier to genetic resistance is a key point in designing these new PI drugs.

2 THE HIV VIRUS

HIV is a virus that causes AIDS (acquired immune deficiency syndrome). HIV is a retroactive virus and belongs to *Retroviridae* family,[26] which means it uses single stranded RNA (ribonucleic acid) to store its genetic information and replicate.[27] AIDS is mainly caused by two types of human viruses, namely HIV-1[5-7,28,29] and HIV-2.[30,31] These two types of viruses belong to a group of retroviruses called lentiviruses.[32] Both these viruses are closely related to Simian Immunodeficiency Viruses (SIV's), and are distinguished based on their genomic organization and their evolutionary relationship to other lentiviruses.[33] Among the two human viruses causing AIDS, HIV-1 is the main causative organism responsible for AIDS pandemic worldwide compared to HIV-2 which is mostly prevalent in West Africa but has slowly spread to Asia and Europe.[34] The most likely origin of HIV-1 is due to cross species transmission (zoonosis) of a Simian Immunodeficiency Virus (SIV) from a subspecies of Chimpanzee (*Pan troglodytes troglodytes*)[35] but probably from different events for each group.[33,36,37] Whereas HIV-2 is thought to have originated from zoonosis events from sooty mangabeys (Cercocebusatys).[38] HIV primarily infects $CD4^+$ T lymphocytes which are part of our immune system.[39-41] The infection occurs in three phases as shown in Figure 2.

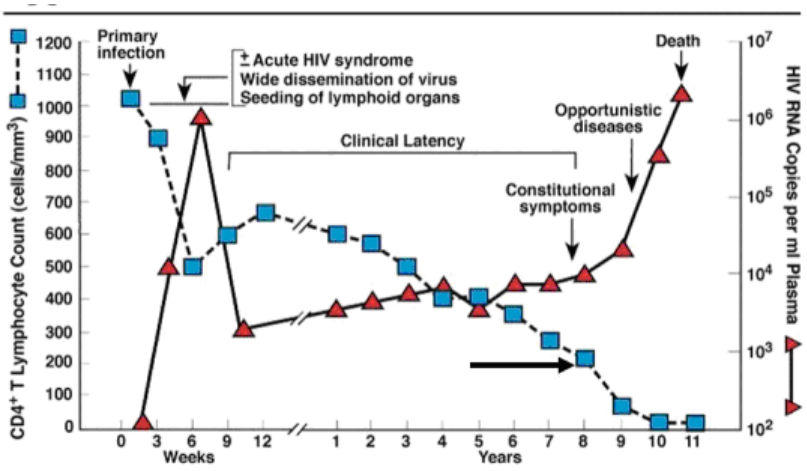


Figure 2. Graph showing HIV virus and CD4+ levels over the course of an untreated infection.[42]

Phase-I The infection progresses rapidly and exhibits mononucleosis-like symptoms within weeks.[43] High viron concentration is observed which may exceed million copies per ml of blood.[11]

Phase-II The second phase starts when the host immune system decreases the viron concentration in the blood. This long, asymptomatic period characterizes HIV as a lentivirus ("slow virus").[44] During this phase the viral replication is active and healthy cells are being infected.[16,17]

Phase-III This is the third and final phase in which the $CD4^+$ T count drops drastically as the CD8 cytotoxic lymphocytes recognizes the infected $CD4^+$ T and kills them. As a result the viral load exceeds enormously to produce clinical immunodeficiency.[45,46]

2.1 STRUCTURE OF HIV

The mature HIV virion is an essentially spherical particle with a radius of about 10 nm (Figure 3).

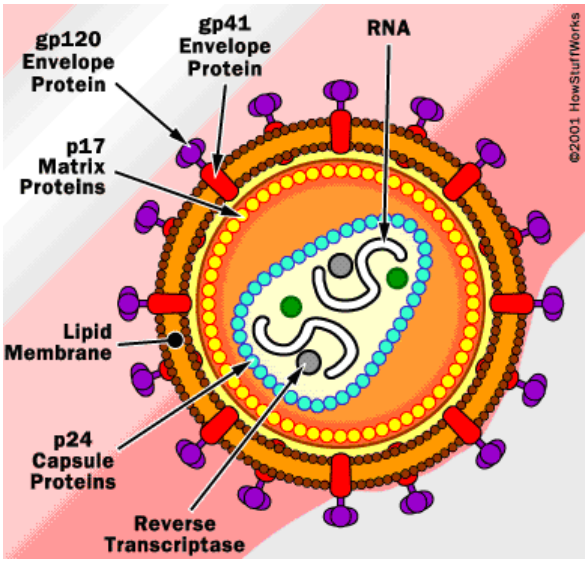


Figure 3. Schematic of the HIV viron.[47]

The basic structure of the virus is as follows:

- The viral envelope, the outer coat of the virus is a lipid bilayer. This outer lipid bilayer is embedded with different cellular membrane proteins such as glycoprotein gp120 which are adhered to the transmembrane gp41. This lipid layer is borrowed from the host cell during the budding process. In HIV virus, gp120 is essential for attaching itself to the host cell and g41 is crucial for the cell fusion process.[48,49]

- The HIV matrix proteins consisting of the p17 protein, which lie between the envelope and core.

- The viral core contains the viral capsule proteinp24which surrounds two single strands of HIV RNA and the enzymes needed for HIV replication, such as reverse transcriptase, protease, ribonuclease, and integrase. Out of the nine virus genes, there are three, namely *gag, pol* and *env* that contain the information needed to make structural proteins for new virus particles.[49]

2.2 THE LIFE CYCLE OF THEHIV VIRUS

The gp120 protein on the HIV virus binds to the $CD4^+$ T cells of the host. This initiates the replication cycle (Figure 4). After binding, the HIV viral envelope fuses with the cell membrane of the host cell and allows the viral RNA and viral enzymes toenter the cytoplasm of the host cell. This activates the reverse transcriptase (RT)enzyme which acts as a DNA polymerase to synthesize double stranded DNA from the viral RNA.

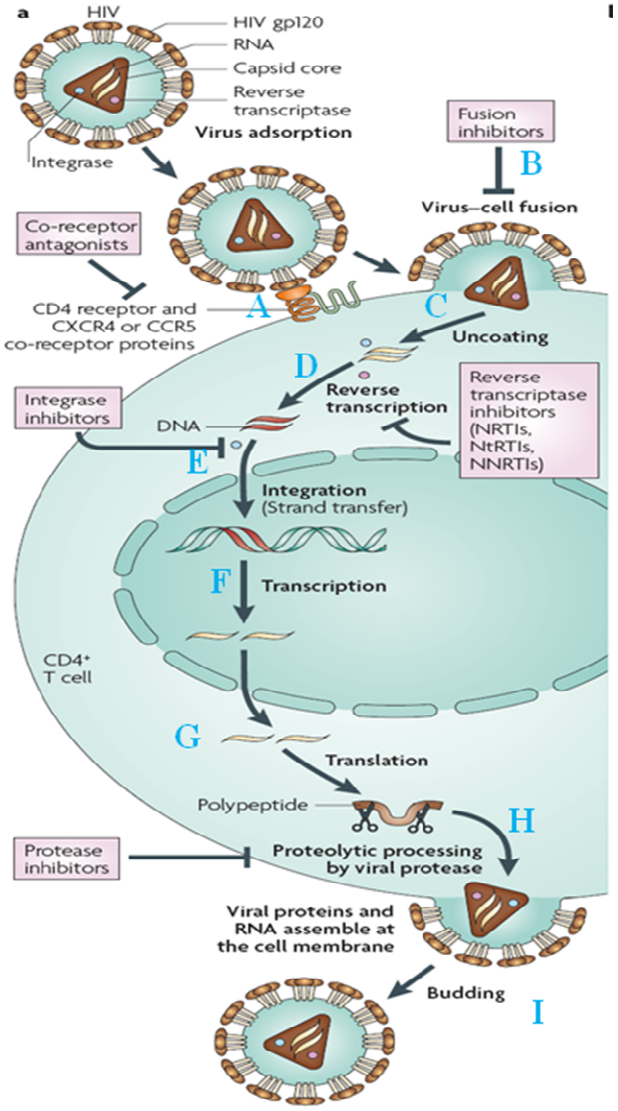


Figure 4. Life Cycle of the HIV virus.[50-53] **a)** attachment; **b)** fusion; **c)** uncoating; **d)** reverse transcription; **e)** integration; **f)** transcription; **g)** translation; **h)** post-translational processing and assembly; **i)** budding; **j)** maturation.

This process is responsible for mutations for each viral generation produced. The virally coded DNA enters the nucleus of the host cell and integrates with the genetic material of the host cell. This process is catalyzed by viral enzyme integrase (IN) which results in the synthesis of viral messenger RNA (mRNA). The viral mRNA leaves the nucleus and utilizes the host cellular mechanism and is translated into the virus proteins *gag* and *gag-pol* polyproteins, which are multi protein molecules that are further processed into smaller functional proteins during viral maturation. Polyproteins and the viral RNA are then transported to the cell membrane for assembly and viral budding. At the host cell membrane the virus buds from the cell surface and is released with all the viral proteins and RNA needed to form virions. New HIV particles from the host cell membrane at the budding stage activates the viral protease enzyme (PR) resulting in the proteolytic cleavage of the polyproteins into various subunits, which leads to the generation of the mature HIV virus. HIV PR enzyme is essential for the generation of infectious virions.[52,53]

3 ANTIRETROVIRAL THERAPY

During 1980s there were no known anti-HIV drugs. Patients treatment was restricted only to the opportunistic infections experienced due to AIDS. The following decade lead to intense research worldwide to understand the morphology and life cycle of the HIV virus. As a result various therapeutic targets were identified to combat the HIV 1 replication cycle. Today, based on the therapeutic targets the anti-HIV drugs are classified into four main classes, which will be discussed briefly in the following sections.[54,55]Anti-HIV agents acting on several new targets are also under investigation.[56]

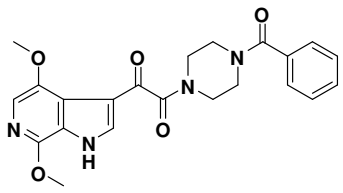
3.1 ENTRY INHIBITORS

Entry inhibitors inhibit the transmission of HIV-1 virus to a $CD4^+$ human cell.[57] HIV entry into human cells requires the following steps in sequence:

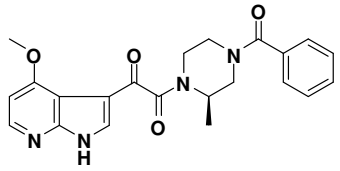
- The binding of HIV surface protein gp120 to the CD4 receptor.
- A conformational change in gp120, which both increases its affinity for a co-receptor and exposes gp41.
- The binding of gp120 to a co-receptoreitherCCR5 or CXCR4.

- The penetration of the cell membrane by gp41, which approximates the membrane of HIV and the T cell and promotes their fusion.
- The entry of the viral core into the cell.

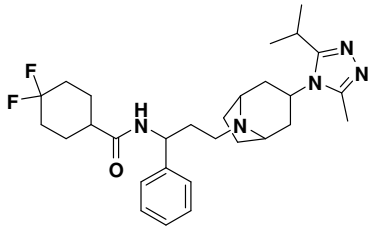
Entry inhibitors work by interfering with one aspect of this process.[58] These entry inhibitors (Figure 5).can be broadly classified into three categories.



BMS-488043



BMS-378806



Maraviroc

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-NH2

Enfuvirtide

Figure 5. Examples of entry inhibitors which inhibit the attachment of HIV-1 virus to the $CD4^+$ cell membrane proteins.

Attachment inhibitors (**AIs**), inhibit the initial binding of gp120 surface protein of HIV-1 virus to the $CD4^+$ cell receptor. Several possible mechanisms have been proposed for the AIs, but the exact mechanism is yet to be confirmed.[59] Some clinically active AIs are BMS-488043 and BMS-378806 (Figure 5).[60]

Co-receptor Antagonists(CRAs), bind to either of the two co-receptors, CCR5 and CXCR4 of the $CD4^+$ cell membrane and thus prevent HIV-1 from entering the immune cells.[61] Examples of some CCR5 co-receptor antagonist are themonoclonal antibody, Pro140 which inhibits HIV-1 infection by binding to the gp120 receptor of the cell and Maraviroc.[60,62-64]

Fusion Inhibitors (FIs), bind to the viral transmembrane protein g41, which undergoes conformational change that assist in the fusion of the viral membrane to the host cell membrane.[59] Thus disrupting the HIV-1 molecular machinery at the final stage of fusion. Enfuvirtide (FDA approved in 2003) is the only drug so far used as a fusion inhibitor.[65,66] It is a linear peptide consisting of 36 amino acids (Figure 5), which corresponds to residues 127-162 of the viral glycoprotein gp41.[67]

3.2 NUCLEOSIDE REVERSE TRANSCRIPTASE INHIBITORS (NRTIS)

Zidovudine (AZT) was the first US Food and Drug Administration (FDA) approved drug for HIV/AIDS. It is a nucleoside reverse transcriptase inhibitor (NRTI).[68] NRTI's that are currently marketed are zidovudine,[68] didanosine,[69] zalcitabine,[70] stavudine,[70] lamivudine,[71] abacavir,[70] emtricitabine,[70] and entecavir.[72] Apricitabine[73] was under phase-III evaluation and was discontinued due to high twice-daily dosing (800 mg). Tenofovir disoproxil fumarate and GS-7340 are commonly known as tenofovir prodrugs.[74,75] They are structurally different from the other NRTIS and is a nucleotide. Structures of few NRTIs are shown in Figure 6.

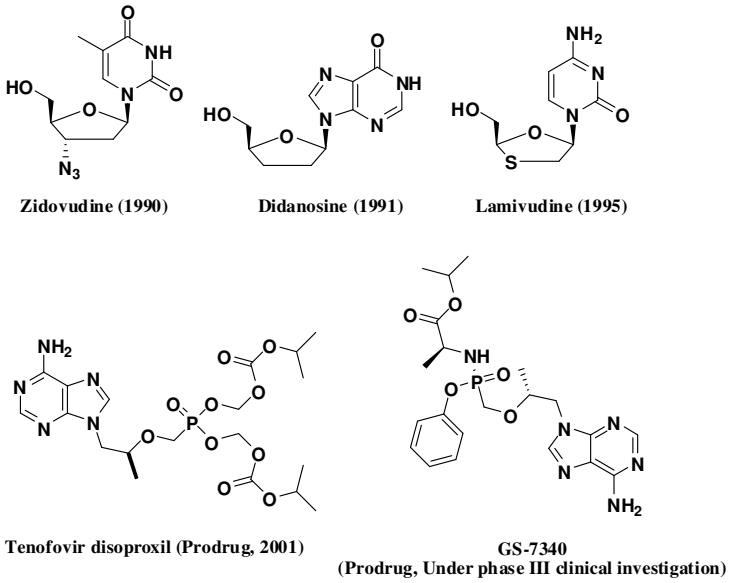


Figure 6. Examples of nucleoside reverse transcriptase inhibitors affecting the viral DNA synthesis.

This class of drugs are analogues to naturally occurring deoxynucleotides needed for viral DNA synthesis. These analogues lack the 3'-hydroxy group on the deoxyribose moiety, also NRTIs are phosphorylated in *vivo* and act as an alternative substrate for RT. Once the phosphorylated NRTIs are incorporated into the growth of the viral DNA, the chain is terminated as they prevent further attachment of the nucleosides. The result is that the viral reverse transcription process is prematurely terminated.[62,76,77]

3.3 NON-NUCLEOTIDE REVERSE TRANSCRIPTASE INHIBITORS (NNRTIs)

Non-nucleotide reverse transcriptase inhibitors are structurally diverse compounds and have different mechanism for inhibiting RT from the nucleoside analogs.[77-79] FDA approved NNRTIs (Figure 7) currently marketed are efavirenz,[70,80] nevirapine,[70,81] delavirdine (discontinued due to high toxicity profile),[70,82] and the latest addition in this class of drugs was etravirine (2008)[83,84] and rilpivirine (2011).[85,86]

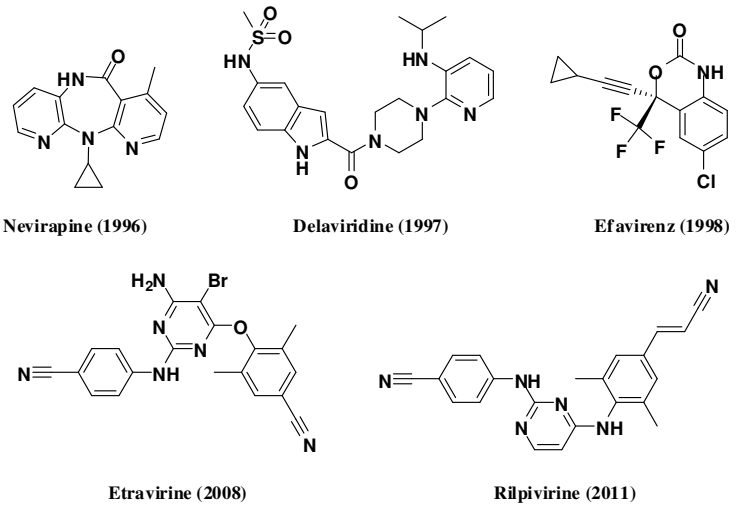


Figure 7. FDA approved non-nucleoside reverse transcriptase inhibitors (NNRTIs) which inhibit HIV by binding to an allosteric site on the reverse transcriptase enzyme (RT).

The NNRTIs cause conformational changes at the active site of the enzyme by binding to an allosteric site on the reverse transcriptase enzyme (RT) of the HIV virus, there by inhibiting its activity. Hence this class of drugs is known as non-competitive inhibitors of the RT enzyme.

3.4 HIV INTEGRASE INHIBITORS (INIs)

HIV integrase (IN) is a vital enzyme that catalyzes the insertion of reverse transcribed viral genome into host chromatin. This enzyme plays a crucial and unique role in viral replication cycle. In the life cycle of HIV, integrase is responsible for catalyzing two distinct steps, 3' processing and strand transfer. During 3' processing, integrase excises a dinucleotide from the 3' terminus of viral cDNA. During strand transfer this 3'-processed viral DNA then covalently binds to host DNA.[26,87,88] This unique process has always been considered a viable drug target, which several early studies attempted to exploit.[89] Early integrase inhibitors (INIs) included peptides,[90,91] nucleotides[92] and DNA complexes[93] as well as small molecules derived either from natural products[92] or by rational drug design strategies.[91,94] Major drawbacks of these early INIs were *in vivo* toxicity and non-specific off-target effects. In 2000 Merck investigators identified 4-aryl-2,4-diketobutanoic acid inhibitors containing a distinct diketo acid moiety (DKA) from a screening of 250,000 compounds.[95] The mode of action of these compounds containing the diketo acid moieties was confirmed by their antiviral activity in cell cultures, which revealed the development of resistance mutations in the IN protein. These compounds, exemplified by L-731988[96] (Figure 8), were found to inhibit strand transfer with much higher potency (IC_{50} = 80 nM) than 3' prime processing (6 μM)[95], and they were thus known as integrase strand transfer inhibitors (INSTIs). Like most nucleotidyltransferase enzymes, IN also requires two divalent cation bound at the active site for activity, for example Mg^{2+} is likely used *in vivo*, although Mn^{2+} is used in some *in vitro* assays.[97] Most INSTIs including DKA compounds inhibit IN by chelation of bound cations in a dose-dependent manner.[98]

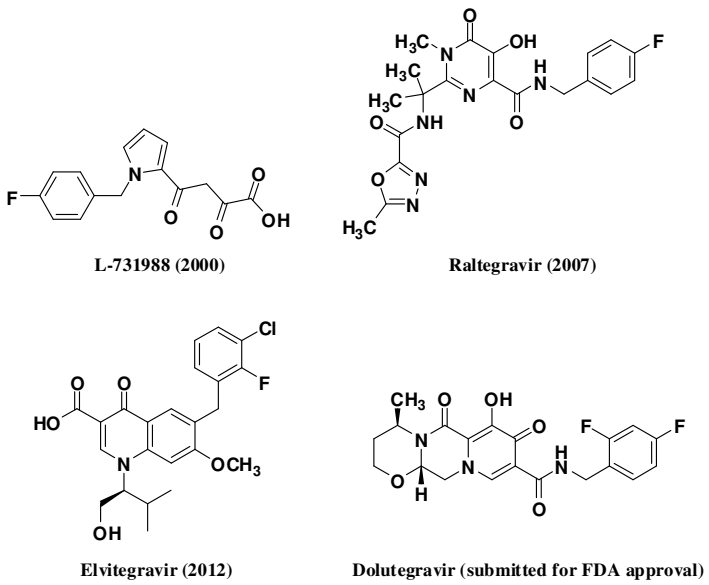


Figure 8. Structures of DKA based integrase inhibitors.

Currently marketed FDA approved DKA based integrase inhibitors as shown above in Figure 8 are raltegravir[99,100] (Merck & Co., October 2007) and elvitegravir[101,102] (Gilead Sciences, August 2012). Dolutegravir[103,104] (GlaxoSmithkline) is a new integrase inhibitor which has been submitted for FDA approval.

3.5 HIV PROTEASE INHIBITORS

Together with the greater understanding of the HIV-1 virus life cycle, new drug targets in addition to the first reverse transcriptase drus, soon emerged. In 1986, Kramer et al. suggested HIV protease as a potential target for AIDS therapy. It was shown that theHIV protease enzyme was crucial for the maturation of the virion particles to the infectious HIV virus. The mechanism involved was the frameshift mutation in the protease region of the *pol*-gene which prevented the cleavage of the *gag* polyprotein precursor, which is essential for the maturation of the HIV virus particles.[105] Thus inhibition of the HIV protease enzyme, would result in the formation of immature non-infectious virions.[106] In the past two decades compounds having the ability to inhibit the protease enzyme were studied intensively and numerous reports of HIV-1 protease inhibitors have been published.[105-112] The first FDA approved HIV-1 protease inhibitors was saquinavir in 1995.[113] Presently there are ten clinically approved inhibitors namely

saquinavir,[70,113] rotinavir,[70,114] indinavir,[70,115] nelfinavir,[70,116] ampenavir,[70,117] lopinavir,[70,118] atazanavir,[70,119] fosamprenavir,[120] tipranavir,[70,121] and darunavir[122] (Figure 9).

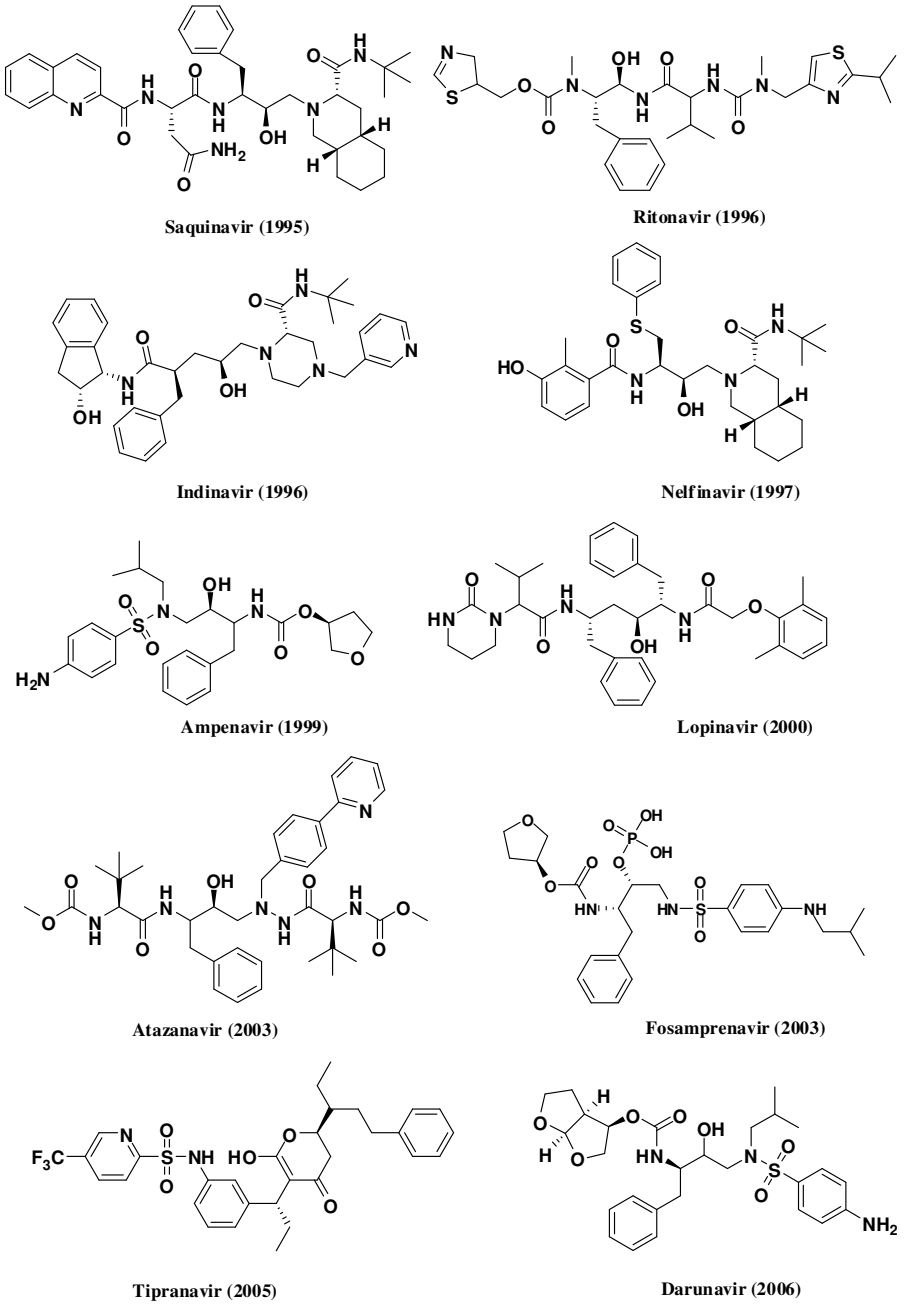


Figure 9. Examples of nine FDA approved HIV-1 protease inhibitors.

These inhibitors have the advantage of being highly selective but nonetheless have some major side effects such as lipodystrophy, hyperlipidaemia, and insulin resistance.[123-129] Hence there is

be a constant demand for new HIV protease inhibitors. In the past two decades many non-cleavable transition state isostere inhibitors (Figure 10) have been reported.

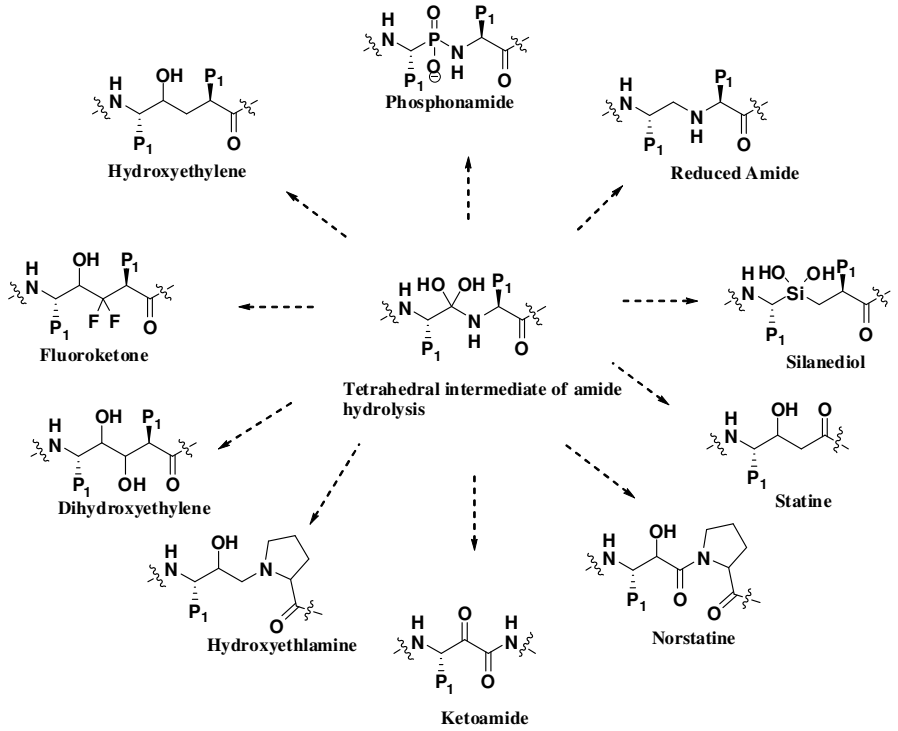


Figure 10. Some noncleavable transition-state isostere developed for the synthesis of HIV PR inhibitors.[130]

These isosteres bind to the active site of the enzyme with an affinity of greater magnitude of several orders than the natural peptide substrate.[131] The HIV-1 protease enzyme will be discussed in detail in later sections.

3.6 HIGHLY ACTIVE ANTIRETROVIRAL THERAPY (HAART)

Antiretroviral drugs (ARVs) are medications for the treatment of HIV Infection. When several such drugs, typically three or four, are taken in combination, the approach is known as Highly Active Antiretroviral Therapy or HAART. In early 1980s primarily AZTs were introduced for combating HIV infection. But within few years, it was observed that AZTs monotherapy showed limited effects against the HIV virus.[132] HIV strains with reduced sensitivity to zidovudine were isolated from the patients receiving prolonged therapy.[133] Zidovudine resistant strains were first reported in 1993.[134] In order to counter this HIV resistance, combination

therapy with two NRTIs was introduced, but strains resistance to this form of therapy soon emerged.[135]

With the introduction of first PR inhibitors (saquinavir and indinavir), triple therapy with zidovudine and zalcitabine [136] or lamivudine [135] was suggested. This combination therapy made it possible to target two steps in the HIV life cycle and was very effective. The patients receiving this highly active form of anti-retroviral therapy showed significantly low levels of HIV RNA.[135,136] Thus the hypothesis of triple therapy was proven correct and HAART was introduced. Mortality was reduced among patients on HAART, particularly if the therapy included protease inhibitors.[137] It was found that after few months of use of this therapy, patients had almost undetectable levels of HIV RNA. Hence it was speculated that the HIV virus could be completely eliminated from the body, if HAART was introduced early and continued for a long enough period, since the virus would be eradicated from the lymphoid tissues.[20,138] But this was proved wrong, as it was found that half life of $CD4^+$T cells was 44 months, which meant it would take approximately 70 years before the virus reservoirs were eradicated.[139] During the latent period of HIV virus infection, the virus simply exists as integrated DNA, unaffected by anti-retroviral drugs. The brain can also serve as a sanctuary site for the virus and can exist beyond the reach of anti-retroviral drugs.[138,140,141] Early introduction of HAART has benefits such as the preservation of the immune system, decreased risk of HIV transmission, early suppression of viral replication, and is also associated with severe side effects like lipodystropy syndrome,[127,142] resistance to insulin,[142] hyperlactemia[143] and coronary heart disease.[144]

Looking at the advantages and disadvantages of HAART therapy, it should be noted that this therapy is quite expensive and the vast majority of the patients worldwide, especially from the developing nations receive little or no treatment at all. Initiations have been taken by many nations to lower the cost of the drug by 90-99%, so that poor people could afford the treatment. In addition, many other issues involving logistics, education and complex social factors must be resolved before HAART can be a treatment for none but the patients of the wealthiest nations.[145,146]

4 THE HIV-1 PROTEASE

Proteases are a class of enzymes that are responsible for catalyzing the cleavage of specific amide bonds in peptides or proteins.[147] Depending on the mechanism of catalysis, the protease enzyme are classified into four categories: metalloproteases (e.g. carboxypeptidase A and angiotensin converting enzyme), serine proteases (e.g. trypsin and thrombin), cysteine proteases (e.g. papain) and aspartyl proteases (e.g. pepsin and renin).[148] Certain medical disorders are required to inhibit a specific protease enzyme as a treatment strategy.[28] One of the well known examples of this strategy is the treatment of hypertension wherein angiotensin converting enzyme is inhibited (e.g. captopril).[149]

The HIV-1 protease belongs to the aspartyl protease family of enzymes.[150] It is a C_2-symmetric dimer which is made up of two monomers (Figure 10). These monomers are non-covalently associated to form the C_2-symmetric active homo-dimer.[49,130,151,152]

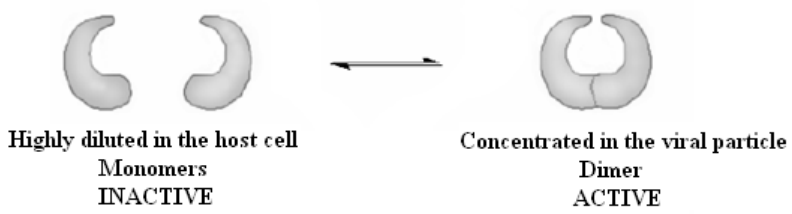


Figure 11. Dimerization of the HIV protease.[153]

The compact four- "stranded" interface is the result of the integration of the two C-terminals and N-terminals of each monomer. These four terminal strands align together like interlocking fingers.[154] Activation of the protease enzyme occurs with dimerization and it is a reversible process. In concentrated solution the protease is activated (in the budded viral particle), and in a highly diluted solution the protease is inactivated (in the host cell). This mechanism prevents premature breakdown of the polyproteins and to minimize damage to the host cellular proteins.[107,153]

4.1 STRUCTURE OF HIV-1 PROTEASE

The X-ray crystallographic structure of HIV-1 Protease enzyme was first reported in 1989,[155-160] and since then a large number of crystal structures with inhibitors have been reported. Some general features of the HIV-1 protease enzyme structure (Figure 11) can be described as below.

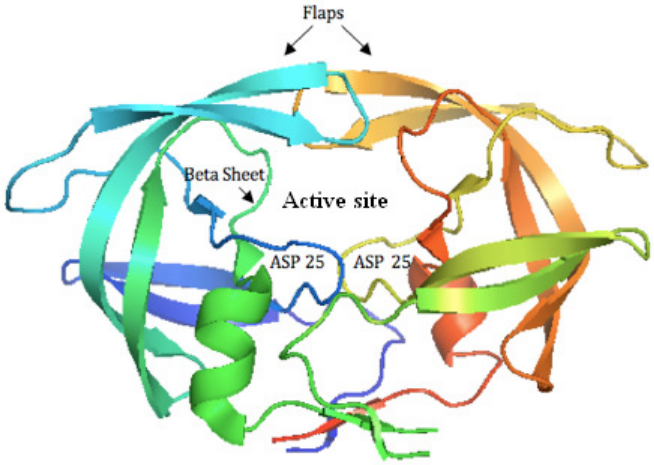


Figure 12. Structure of HIV protease.[161]

The HIV-1 Protease enzyme consist of two identical monomers consisting of 99 amino acid residues each and form a C_2-symmetrical elliptical-shaped enzyme.

The dimer is held together by N and C- termini of each monomer and is juxtaposed in a four-stranded β-sheet (the dimer interface). In the sequence of each monomer, amino acids 1-4 are β-sheet residues, 4-9 are turn encompassing residues and 86-94 are helix residues. This domain is quite crucial in dimer formation and stabilization of an active protease.[162,67]

The active site has a characteristic catalytic triad sequence Asp25-Thr26-Gly27 common to aspartic proteases which is situated between the identical monomer subunits. Apart from the four stranded beta-sheet structure it is made up of quite compact residues at positions 10-32 and 63-85 of each monomer. Each of these monomers consists of a hydrophobic core consisting of the two loops, one of which includes the active site aspartic acid Asp25.[162,67]

The dimers come together to create an extended substrate-binding cleft capable of interacting with a minimum of seven consecutive amino acids in the substrate. Each monomer contributes a

flexible flap that folds down to make important contacts with the bound substrate. The two flaps (residues 44-63) enclose the active site and provide important ligand binding interactions.[162,67,163]

The 3D structures of the enzyme was taken from the reported X-ray data of subtype B HIV-PR (PDB accession code 1HXW) as well as the C-SA wild type HIV PR strain used in this study are available in PDB format as part of the supplementary material to this thesis.

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 (Figure 12).

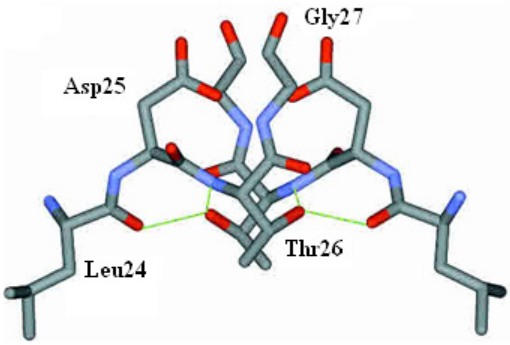


Figure 13. The catalytic residues triad Asp25–Thr26–Gly27 forming a network of hydrogen bonds.[67]

The catalytic residues triad Asp25–Thr26–Gly27 form a network of hydrogen bonds referred to as a "firearm's grip.[50,162]

4.3 SUBSTRATE BINDING

The active site of HIV-1 protease contains a number of well defined subsites where the substrate or inhibitor side chains can be accommodated on binding. Starting from the aspartates, in the active site, there are distinct subsites named S_1, S_2, S_3, and S_4 in the first monomer with corresponding S_1', S_2', S_3' and S_4' subsites in the other monomer. These subsites were coined by Schecheter and Berger.[164] The corresponding side-chains of the substrate or of the inhibitor are

named P_1 to P_n outwards from the scissile peptide bond towards the amino terminus and P_1' to P_n' towards carboxyl terminus (Figure 13).

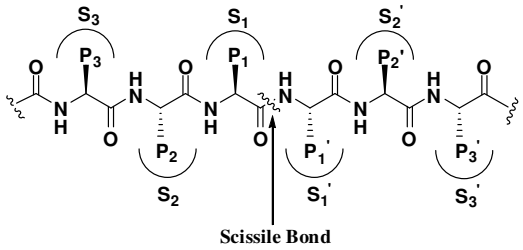
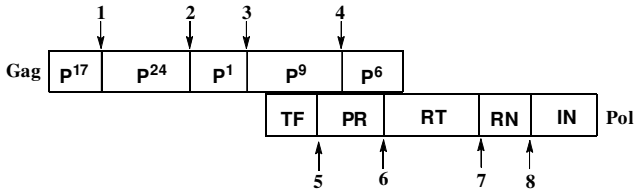


Figure14. Standard nomenclature P_1 _ _ _ P_n, P_1'_ _ _ _ P_n'_ is used to designate amino acid residues of peptide substrates. The corresponding binding sites on the protease are referred to as S_1_ _ _ S_n, S_1'_ _ _ _ S_n'_ subsites.[50,151]

All subsites in the HIV-1 protease are bound by mostly aliphatic side chains and have hydrophobic character, with the exception of S_4/S_4', which are exposed to water.[153,151,165]

4.4 CATALYTIC MECHANISM

The HIV- protease recognizes and cleaves the amino acid sequences within *gag* and *gag-pol* polyproteins at specific cleavage site as shown in Figure 14.[112,166-170]



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-

Figure 15. Cleavage sites within the gag and gag-pol precursor proteins. The asterisk* indicates to the position of the scissile bond.

HIV-1 protease cleaves the *gag* and *gag-pol* polyproteins by general acid-base mechanism (Figure 15).[130] The protonation states for the catalytic aspartic acids have been investigated using different methods.[171-176] The water molecule located between the two aspartic residues in the active site is activated by the deprotonated aspartic acid. This activated water molecule acts as a nucleophile and attacks at the carbonyl carbon in the substrate (a). The protonated aspartic acid donates a proton to the nitrogen and forms a oxyanion tetrahedral intermediate (b and c) that collapses into two hydrolyzed products (d): the acid (*N*-terminal cleavage product) and the amine (*C*-terminal cleavage product).[28,50,130,177,178]

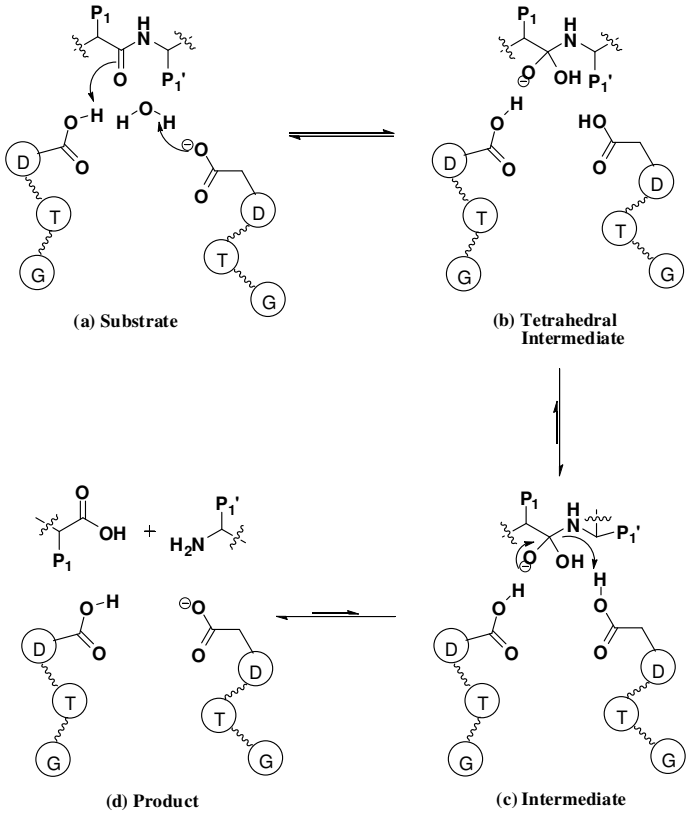


Figure 16. Cleavage mechanism of HIV-1 protease enzyme.[130]

The knowledge accumulated through the last two decades about the structure and the mechanism of HIV PR has paved the way towards the development of new effective drugs to inhibit this enzyme.[130]

5 TOOLS EMPLOYED FOR CONFORMATIONAL ANALYSIS AND PROTEIN-DRUG INTERACTIONS IN MEDICINAL CHEMISTRY

This section we will briefly describe the common techniques used for determining the conformational analysis of the substrates and will involve X-ray crystallography, NMR spectroscopy, and computational chemistry methods.

5.1 X-RAY CRYSTALLOGRAPHY

Single-crystal X-ray diffraction is a non-destructive analytical technique which provides detailed information about the internal lattice of crystalline substances, including unit cell dimensions, bond-lengths, bond-angles, and details of site-ordering.[179-181] Directly related is single-crystal refinement, where the data generated from the X-ray analysis is interpreted and refined to obtain the crystal structure. In this method a mounted crystal is bombarded with a beam of X-rays, resulting in diffraction of the X-rays into many specific directions.[181] From the angles and intensities of the diffracted beams a three dimensional picture of the density of electrons is generated, which is then used to calculate the mean position of the atoms in the crystal. Thus a detailed 3D structure is obtained.

X-ray crystallography has become an important tool and plays particularly important role in the structure-assisted drug design and discovery process.[182] The protease targeted drugs against AIDS was possible with the determination in early 1989 of the structures of retroviral proteases, first from RSV[183]and subsequently from HIV-1 PR.[184] Many experimental crystal structures of native HIV-1 PR for the recombinant and synthetic enzymes were reported by several laboratories.[155,157,184-186] The first fully correct structure of HIV-1 PR and the cocrystal structure of its complex with an inhibitor (a substrate-derived, reduced isostere hexapeptide MVT-101) were determined using chemically synthesized protein.[187] These studies confirmed that the HIV-PR is a homodimer and the active site closely resembles the active sites of pepsin like proteases. In the past two decades many crystal structures of HIV-1 PR complexed with known inhibitors like saquinavir,[188-190] ritonavir,[114,191,192] indinavir,[193,194] nelfinavir,[194,195] ampenavir,[196] lopinavir,[197,198] atazanavir,[199,200] tipranavir,[201,202] and darunavir[203] have been reported. These X-ray structures of protein-inhibitor complex gave detailed morphology of the

HIV-1 PR enzyme and its interactions with the drug inside the active site, which further allowed for the calibration of computational methods and the design of more potent inhibitors.

5.2 NMR STUDIES

Two-dimensional nuclear magnetic resonance (2D NMR) technique has emerged as a powerful tool in determining the structure of a molecule.[204] 2D NMR spectra provide more information about a molecule than one-dimensional NMR spectra and are especially useful in determining the structure of a molecule that is too complicated to solve with one dimensional NMR spectra. Three dimensional (3D) structural determinations can be achieved from Nuclear Overhauser Enhancement (NOE) spectroscopy (NOESY).[205,206] The NOESY spectrum 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.[207,208] These NOE interactions are typically observed for protons which are in relative close proximity (< 5 Å) relative stable conformations are required to observe good NOE interactions.[209-211]

One limitation of NOESY spectrum is that small and medium sized flexible molecules (MW = 1000 $gmol^{-1}$),[212] show little or no NOE phenomenon as a result determining the 3D structures of such molecules with this method becomes difficult. To resolve this problem rotating-frame frame Nuclear Overhauser effect spectroscopy (ROESY) is employed.[213] However, several experimental problems namely direct cross-peaks due to *J* coupling, Hartmann-Hahn matching (leading to TOCSY, total correlation spectroscopy, cross-peaks) and offset dependence have to be avoided.[214-217] Recently an EASY ROESY technique was developed to obtain a ROESY spectrum for medium sized flexible molecules without these artifacts. These distances are then used to obtain valuable information about the 3D structure in solution phase.[212] This method was used in later chapters.

5.3 COMPUTATIONAL STUDIES

Computational chemistry is a discipline using mathematical methods for the calculations of molecular properties or for the simulations of molecular behavior.[218] This branch of chemistry

has become a very important tool to design new molecules of new reactions which are later investigated experimentally. In other cases it is used to supplement experimental studies by providing data which are hard to probe experimentally (for example, transition state structures, energies, protein drug interactions or binding energies). Since its modest beginnings in the 1950's and 1960's, advances in theoretical techniques and in computer power have dramatically increased the usefulness and importance of computational chemistry.

The Schrödinger equation ($\mathbf{H\psi = E\psi}$) is the basis for the electronic based (quantum mechanical - QM) computational chemistry methods. Due to the size of the mathematical calculations involved with QM methods, only small to medium sized molecules can currently be investigated. Larger systems such as proteins and enzymes are mostly subjected to molecular mechanical (MM) methods. Typical chemical entities that can be calculated are:[219]

- electronic structure determinations
- geometry optimizations
- frequency calculations
- transition state structures
- protein calculations, i.e. docking and molecular dynamics (MD)
- electron and charge distributions
- potential energy surfaces (PES)
- rate constants for chemical reactions (kinetics)
- thermodynamic calculations-heat of reactions, energy of activation

The computational chemistry is divided into two main branches:

Classical Mechanics also called Molecular Mechanics (MM) or "force-field method. All molecular mechanics methods are *empirical* that is the parameters in the model are obtained by fitting to known experimental data.[220] The specifics of each computational method are given briefly in Table 2.

- Quantum Mechanics (QM) which can be further classified as *ab initio* or *semi-empirical*[221]

- Density functional theory (DFT)

Table 2. Shows the specifics of each of the methods.[222]

Method Type	Basis	Considerations	Best for	
Molecular Mechanics	uses classical physics relies on force-field with embedded empirical parameters	computationally least intensive - fast and useful with limited computer resources can be used for molecules as large as enzymes	particular force field applicable only for a limited class of molecules does not calculate electronic properties requires experimental data (or data from *ab initio*) for parameters	large systems (thousands of atoms) systems or processes with no breaking or forming of bonds
Semi-empirical	uses quantum physics uses experimentally derived empirical parameters uses approximation extensively	less demanding computationally than *ab initio* methods capable of calculating transition states and excited states	requires experimental data (or data from *ab initio*) for parameters less rigorous than *ab initio* methods	medium-sized systems (hundreds of atoms) systems involving electronic transitions
Ab Initio and DFT	uses quantum physics mathematically rigorous, no empirical parameters uses approximation extensively	useful for a broad range of systems does not depend on experimental data capable of calculating transition states and excited states	computationally expensive	small systems (tens of atoms) systems involving electronic transitions molecules or systems without available experimental data ("new" chemistry) systems requiring rigorous accuracy

In Quantum mechanics, *ab initio*, means "from the beginning" and implies an approach which contains no *empirical parameters*. This category includes Hartree-Fock (HF), configuration interaction (CI), many-body perturbation theory (MBPT), coupled-cluster (CC) theory, and other approaches. The second category, *semi-empirical*, includes methods which make large approximations to the quantum mechanical laws and then employ a few empirical parameters to assist with correcting the result. These methods include the modified neglect of differential

overlap (MNDO), Austin Model 1 (AM1), Parameterized Model number 3(PM3), Monte Carlo method and many others.[219,221,223-225] Due to the nature of the empirical parameters involved, a specific method will work better for a certain family of molecules, while it will give unreliable results for a different family of molecules.

Density functional theory (DFT) methods are quantum mechanical modeling method used in physics and chemistry to investigate the electronic structure (principally ground state) of many body systems, in particular atoms, molecules and the condensed phases.[219,226] Some DFT methods are free from empirical parameters, while others rely heavily on calibration with experiment.[227] DFT was originally developed by Hohenberg and Kohn[228] and later developed by Kohn and Sham.[229] For the past 30 years density functional theory has been the dominant method for the quantum mechanical simulation of periodic systems. In recent years it has also been adopted by quantum chemists and is now very widely used for the simulation of energy surfaces in small molecules.[230] A grand challenge in computational chemistry and biology is the accurate quantum mechanical calculation of interaction energies for biological molecules such as proteins. Due to a larger number of atoms, standard full quantum mechanical or *ab initio* calculation of protein interaction energy with the molecules is beyond computational reach.[231,232] Due to the size of systems involved in this study, MM and MD calculations will be used.

5.3.1 MOLECULAR MECHANICS (MM)

Molecular mechanics is the more popular of the methods used to obtain molecular models of very large systems as it requires considerably less computing resources and time. Newtonian mechanics is applied to model molecular systems. The potential energy of all systems in molecular mechanics is calculated using force fields and can be expressed mathematically as shown in Figure 16. MM describes molecules in terms of "bonded atoms" which have been distorted from some idealized geometry due to non-bonded van der Waals and Coulombic interactions. The force constant inherently represents the ability of the "spring" connecting to atoms to strech, bend and rotate. Each family of bond type (alkane, alkene, amide etc.) will have its characteristics incorporated into the force constant.[219,233]

Force constants are developed for specific families of molecules (peptides, sugars, metal complexes etc.). X-ray data and high level *ab initio* (or DFT) results are used to calibrate the forcefield for a specific family of molecules.[234,235] The success of molecular mechanic model depends on a high degree on the transferability of geometrical parameters from one molecule to another, as well as predictable dependence of the parameters on atomic hybridization.[235] Therefore researchers should ensure that a specific force field is well calibrated for the molecule to be investigated.

$$E - E_{bond} + E_{angle} + E_{dihedral} + E_{vdw} + E_{elec}$$

$$= \frac{1}{2}\sum_{bonds} K_b\,(b - b_0)^2$$

$$+ \frac{1}{2}\sum_{angles} K_\theta\,(\theta - \theta_0)^2$$

$$+ \frac{1}{2}\sum_{dihedrals} K_\phi\,[1 + \cos(n\phi -)]^2$$

$$+ \frac{1}{2}\sum_{non-bonds} \left\{ 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + \frac{q_1 q_2}{Dr} \right\}$$

(Short range)

(Long range)

Figure 17. General form of the Molecular Mechanics equations.[234,236]

The total potential energy (E_{Total}) of a molecule is given by the sum of all the energies of the attractive and repulsive forces between atoms in the structure[219,237,238] Figure 16. Even though advanced optimization algorithms have been develop to find the lowest energy geometry molecules, when highly flexible molecules such as peptides and proteins are involved, it is clear that another approach will have to be followed in order to ensure that a representative ensemble of low energy conformations are indeed found.

5.3.2 MOLECULAR DYNAMICS

The picture of a molecular system obtained from a molecular mechanics calculation is far from the true image of the molecule or structure. Apart from the fact that many more low energy conformations may exist, molecules experience continuous movement. In biological macromolecules the atoms and groups of molecules are in constant motion and these movements

are concerted and may be essential for biological function. NMR and X-ray techniques have been used to observe this behavior of biological macromolecules.[219,234,239,240]

Thermodynamic properties of molecules, especially when they form a complex or are immersed in a solvent, cannot be derived from harmonic approximation, which inherently assumes small amplitude motions around a symmetric minimum. Molecular dynamics simulations are used to present researchers with a dynamic picture of the movements of atoms and molecules at a certain temperature.[241,242] Molecular mechanics and molecular dynamics share the same potential energy function but they differ conceptually. In molecular mechanics, the information is derived from a single geometry of the molecule where as in molecular dynamics simulation methods thousands to millions of geometries are explored to produce meaningful results.[219,240,243,244] MD is classified as a deterministic approach.[245-247] In this thesis we have only applied deterministic approach so stochastic approach will not be discussed.

Molecular Dynamics (MD) simulates the time evolution of the molecular system and provides us with the actual trajectory of the system. The information generated from simulation methods can be in principal used to fully characterize the thermodynamic state of the system. Classical molecular dynamics (MD) can be used to propagate in time the nuclear coordinates of molecular system using the classical equations of motion.[219,239,248]

$$-\frac{dV}{dq} = m\frac{d^2q}{dt^2} \qquad \textbf{(2)}$$

Equation 2 must be solved numerically propagating a trajectory at small time-steps, while a typical time-step is about one femtosecond (10^{-15}second). Most of chemical interesting events take place at time scales several orders of magnitude higher (micro or millisecond). Therefore the MD equations should be propagated until 10^9-10^{12} steps to observe a reactive event (rare event).[249] Despite of the recent acceleration techniques this task is yet too expensive to be performed with current the computational resources. In this thesis MD will be employed as a technique that may be used to compute equilibrium as well as kinetic properties[237,249,250] of a many-body system.

The QM/MM energy and forces can in principle be used within any molecular dynamics (MD) or Monte Carlo (MC) scheme. The objective of such simulations is the sampling of configurational space to calculate statistical thermodynamics ensemble averages (Example: Free energy differences such as reaction, activation or salvation free energies). The computational demands for these simulations are extremely high as the amount of sampling necessary to obtain converged averages is considerable. To reduce the computational effort, approximate treatments have therefore been developed by taking advantage of the QM-MM partitioning. This is achieved by avoiding the demanding direct sampling of the QM contribution while fully sampling the MM configurations.[251,252] An early example of this approach[253] kept the QM region fixed while sampling the MM region and used ESP (electrostatic potential) derived charges for the QM atoms to evaluate the electrostatic QM/MM interactions during the MD run. This was shown to be successful in the context of a QM/MM free energy perturbation treatment in which the entropic contributions from the QM region are estimated separately.[253-255] There are a number of recent other activities to improve the available QM/MM MD technology.[251,256,257]

5.3.3 DOCKING

In molecular modeling, docking is a method, which predicts the preferred orientation of the guest (drug molecule) to a host (biological system) to form a stable complex. This knowledge of preferred orientation can be used to predict the strength of association or binding affinity between the ligand (drug) and the receptor (enzyme).[258,259] In the process of docking the conformations of the ligand and specified residues of the protein are adjusted to achieve an overall "best fit." This kind of conformational adjustments result the number of overall preferred binding modes which is referred to as an "induced fit" model.[260,261]

The aim of molecular docking is therefore to achieve an optimized conformation for both protein and ligand, and relative orientation with respect to each other so that the free energy of the overall system is minimized. The low energy conformations found in this way should at least hypothetically, come close to the actual binding conformations of the ligand to the active site of the enzyme. In order to obtain meaningful results, it is essential to have a basic understanding of the mechanism of action in the active site. This can be obtained from a X-ray structure of a

known inhibitor inside the active pocket of the enzyme, or from computational results on known systems.[234] It may also be required to refine the crude docking results by doing a MM/MD or even a QM/MM/MD simulation study on the best docked inhibitor/enzyme structure. Since Kuntz and colleagues' pioneering work[262] significant progress has been made in docking research to improve the computational speed and accuracy. Among them protein-ligand docking is a particularly vibrant research area because of its importance to structure-based drug design.[263]

Two essential components of protein-ligand binding are sampling and scoring.[234,264] Sampling refers to the generation of putative ligand binding orientations/conformations near a binding site of protein and can be further divided into two aspects, ligand sampling and protein flexibility. Scoring is the prediction of the binding tightness for individual ligand orientation/conformation with a physical or empirical energy function. The top orientation/conformation, one with the lowest energy score, is predicted as the binding mode.[263,265-267] The general classification of the methods for protein-ligand docking is shown in Figure 17.

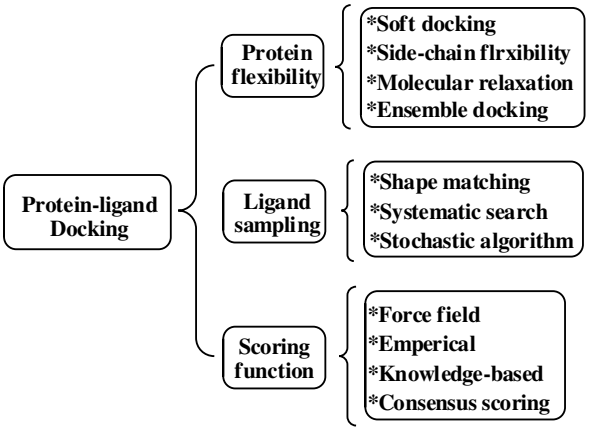


Figure 18. Classification of the methods for protein-ligand docking.[263]

In the past two decades, as many as 60 different docking programs have been proposed for both commercial and academic use. Some of the most popular docking programs currently in use are DOCK,[268] AutoDock,[269] FlexX,[270] Surflex,[271] GOLD,[264] ICM,[272] Glide,[273] Cdocker,[274] LigandFit,[275] and MCDock.[276] All of these docking software's exploit different strategies in the ligand placement. These docking software's can be categorized into four broad categories: stochastic Monte Carlo, fragment-based, evolutionary-based, and the shape complementary

methods.[277] None of these programs use a systematical search to fully explore all degrees of freedom in the ligand molecule because of the enormous computational cost of such a procedure.[278]

As mentioned earlier most of the docking programs are composed of two main components, search/sampling and scoring.[234,264] Out of the two components, scoring function is most widely preferred approach to predict the strength of binding the ligand to the protein target. The advantage of this approach is its speed, due to relative simplicity of the functions.[277] However, when comparing the experimentally derived score (in vitro measured activity of the ligand), weak correlations were usually obtained.[265,279] It is also observed that scoring functions cannot recreate ligands true binding affinities and cannot be ranked according to the RMSD (root mean square deviation, usually 2Å) to the native structure.[265,280,281] This is because all the docking programs belong to the same class of empirical based scoring functions as they explore similar parameters, like van der Waals repulsive hydrogen bonds and electrostactic terms.[277]

Literature reports suggest that, presently there is no single docking software that consistently outperforms all others. Nevertheless, popular docking programs give 60-70% docking accuracy in terms of pose prediction, which is a reasonable number. But in case of docking score and *in vitro* binding affinities there is a poor correlation. Hence it is concluded form the literature reviews that docking program performance is both target- and ligand-dependent and should always be evaluated and validated for the problem at hand.[263,265,277]

6 NEED FOR NEW HIV-1 PR INHIBITORS

In recent years there is an increase in the number of HIV-1 PR mutant strains, which are showing resistance to many marketed drugs.[282] Currently marketed, FDA approved HIV-1 PR drugs especially when prescribed through HAART, are associated with severe side effects, toxicities, high dosage and high treatment cost.[283,284] These factors have urged researchers worldwide to search for new HIV-1 PR drugs which would demonstrate high potency, high genetic barrier, lack of cross-resistance towards the resistant mutant strains, good tolerability, low toxicity and low cost of production. In this thesis, due to the long standing interest of our research group in

polycyclic cage compounds[285-294] we have decided to investigate some PCU diol diacid and PCU hydroxy acid (5-hydroxy-4- oxahexacyclo[5.4.1.$0^{2,6}$.$0^{3,10}$.$0^{5,9}$.$0^{8,11}$]dodecane)[295]derived peptides and peptoids as potential HIV-1 PR inhibitors. This follows the successful study by Ndlovu[296] and Makatini[297,298] where a PCU lactam was employed as a HIV PR new transition state mimic.

6.1 POLYCYCLIC CAGE COMPOUNDS

Polycyclic cages[285-287,299] are an interesting and highly promising group of compounds that are receiving intense scrutiny as potential chemical scaffolds for the development of new drugs.[288,300,301] It is a well know fact that hydrocarbon cage moieties promote the transport of drugs attached to them across cell membranes including the blood brain barrier (BBB) and central nervous system (CNS).[289,290] Incorporation of polycyclic "cage" compounds have shown to increase their affinity for lipophilic regions in the receptor molecules.[302-305] The rigid carbon frameworks of the polycyclic "cage" compounds, incorporated into drugs should also have the added advantage of retarding the metabolic degradation due to the inherent stability and steric bulk of the cage skeleton, thus prolonging the activity and reducing the frequency of drug administration.[288,289,291] Some examples of polycyclic cage compounds are given below in Figure 18.

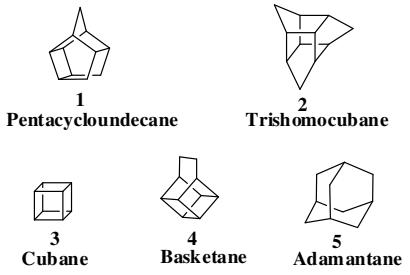


Figure 19. Examples of some polycyclic compounds.

The potential of polycyclic "cage" compounds as biologically active agents was first realized with the discovery by Davis and co-workers[306-308] that 1-amino-adamantane **(6)**, better known as amantadine, did exhibit antiviral activity and was an effective drug against Parkinson's disease. Over the years adamantane became the most widely explored polycyclic "cage" compound.

Various derivatives of adamantane were synthesized in several laboratories and their pharmacological activities were evaluated. Some adamantane derivatives with their corresponding pharmacological profiles are presented in Figure 19.

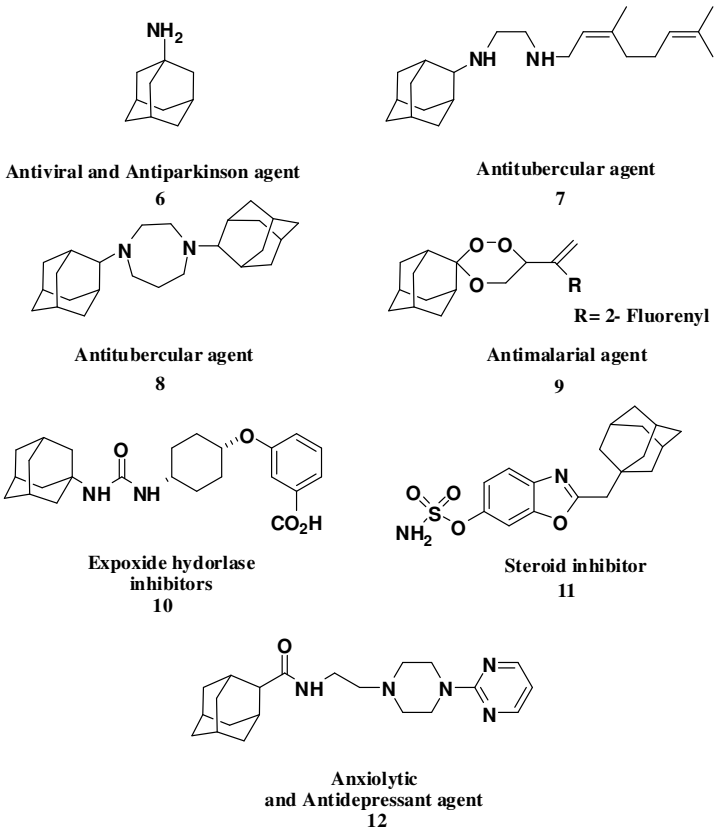


Figure 20. Examples of biological active adamantane compounds.[288]

The introduction of the adamantane moiety into medicinal drugs has shown to result in longevity of drug action, increased drug potency, speed of action and receptor specificity. These were observed in areas such as antibacterial activity,[309,310] anabolic action,[311] analgesic activity,[312] sedative activity,[313] anti-tubercular activity,[314-316] antimalarial activity,[317] antidepressant activity,[318] anti HIV activity,[317] epoxide hydrolase inhibitor,[319] steroid inhibitor and antiviral activity.[317,320]

6.2 PENTACYCLOUNDECANE COMPOUNDS

Since the introduction of amantadine, considerable synthetic research has been carried out on polycyclic compounds.[285-287,299] However pharmacological evaluation and synthesis of

potential target molecules has been neglected to a large extent. The success of amantadine as a potential biologically active molecule, led to the investigation for finding other polycyclic lead compounds. One such polycyclic compound whose derivatives showed promising therapeutic value was pentacyclic "cage" compounds. Their potential pharmacological applications have been shown recently in many case studies. These pentacyclic "cage" compounds showed promising properties as anti-prakinsonian,[321,322] antiviral,[293,298,323-326] anti-tubercular,[327-329] anti-HIV, calcium channel antagonist,[330,331] neuroprotective agent,[332] anti-bacterial,[333] anti-fungal[333] and anti-inflammatory[334] activities which compared favourable with that of adamantane derivatives. Some examples of pentacyclic cage compounds and their pharmacological profiles are given below in Figure 20.

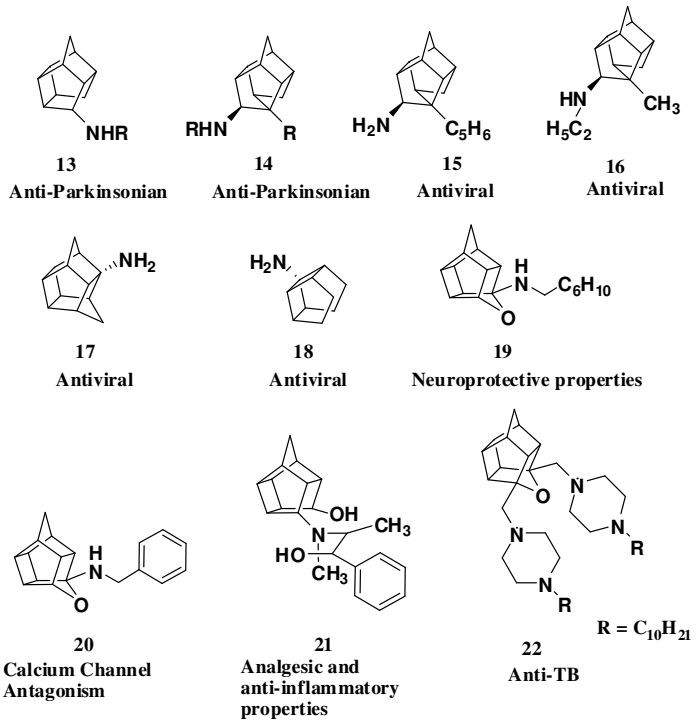


Figure 21. Examples of biological active pentacyclic compounds.[288]

With pentacycloundecane cage derivatives showing a wide range of pharmacological activities against various pathogenic organisms, it has become a major interest for many researchers, especially in the field of medicinal chemistry and drug discovery. The rest of this dissertation

will focus extensively on the incorporation of the PCU cage (**1**, Figure 18) into peptides and peptoids as potential anti HIV-I protease inhibitors.

7 AIMS OF THE PRESENT STUDY

This investigation is part of a research project aimed to develop new non-hydrolysable transition state isosteres to be used in the simple synthesis of potent and bioavailable PCU incorporated peptide and peptoid HIV-1 protease inhibitors. The specific objectives of this study have been:

- To design and synthesise new PCU derived peptide and peptoid compounds as potential HIV PR 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.

8. REFERENCES

1. Gottlieb, M. S.; Schroff, R.; Schanker, H. M.; Weisman, J. D.; Fan, P. T.; Wolf, R. A.; Saxon, A. *N. Engl. J. Med.* **1981,** *305*, 1425-1431.
2. Prusiner, S. B. *Science* **2002,** *298*, 1726-1727.
3. Barresinoussi, F.; Chermann, J. C.; Rey, F.; Nugeyre, M. T.; Chamaret, S.; Gruest, J.; Dauguet, C.; Axlerblin, C.; Vezinetbrun, F.; Rouzioux, C.; Rozenbaum, W.; Montagnier, L. *Science* **1983,** *220*, 868-871.
4. Brunvezinet, F.; Barresinoussi, F.; Saimot, A. G.; Christol, D.; Montagnier, L.; Rouzioux, C.; Klatzmann, D.; Rozenbaum, W.; Gluckmann, J. C.; Chermann, J. C. *Lancet* **1984,** *1*, 1253-1256.
5. Gallo, R. C.; Salahuddin, S. Z.; Popovic, M.; Shearer, G. M.; Kaplan, M.; Haynes, B. F.; Palker, T. J.; Redfield, R.; Oleske, J.; Safai, B.; White, G.; Foster, P.; Markham, P. D. *Science* **1984,** *224*, 500-503.
6. Levy, J. A.; Hoffman, A. D.; Kramer, S. M.; Landis, J. A.; Shimabukuro, J. M. *Science* **1984,** *225*, 840-842.
7. Coffin, J.; Haase, A.; Levy, J. A.; Montagnier, L.; Oroszlan, S.; Teich, N.; Temin, H.; Toyoshima, K.; Varmus, H.; Vogt, P.; Weiss, R. *Science* **1986,** *232*, 697-697.
8. Jolles, S.; DeLoes, S. K.; Johnson, M. A.; Janossy, G. *Brit. Med. J.* **1996,** *312*, 1243-1244.
9. Tindall, B.; Barker, S.; Donovan, B.; Barnes, T.; Roberts, J.; Kronenberg, C.; Gold, J.; Penny, R.; Cooper, D. *Arch. Intern. Med.* **1988,** *148*, 945-949.

10. Cooper, D. A.; Maclean, P.; Finlayson, R.; Michelmore, H. M.; Gold, J.; Donovan, B.; Barnes, T. G.; Brooke, P.; Penny, R. *Lancet* **1985,** *1*, 537-540.
11. Daar, E. S.; Moudgil, T.; Meyer, R. D.; Ho, D. D. *N. Engl. J. Med.* **1991,** *324*, 961-964.
12. Clark, S. J.; Saag, M. S.; Decker, W. D.; Campbell-Hill, S.; Roberson, J. L.; Veldkamp, P. J.; Kappes, J. C.; Hahn, B. H.; Shaw, G. M. *N. Engl. J. Med.* **1991,** *324*, 954-960.
13. Lindback, S.; Karlsson, A. C.; Mittler, J.; Blaxhult, A.; Carlsson, M.; Briheim, G.; Sonnerborg, A.; Gaines, H.; HIV, K. I. P. *AIDS* **2000,** *14*, 2283-2291.
14. Pantaleo, G.; Graziosi, C.; Demarest, J. F.; Butini, L.; Montroni, M.; Fox, C. H.; Orenstein, J. M.; Kotler, D. P.; Fauci, A. S. *Nature* **1993,** *362*, 355-358.
15. Embretson, J.; Zupancic, M.; Ribas, J. L.; Burke, A.; Racz, P.; Tennerracz, K.; Haase, A. T. *Nature* **1993,** *362*, 359-362.
16. Wei, X. P.; Ghosh, S. K.; Taylor, M. E.; Johnson, V. A.; Emini, E. A.; Deutsch, P.; Lifson, J. D.; Bonhoeffer, S.; Nowak, M. A.; Hahn, B. H.; Saag, M. S.; Shaw, G. M. *Nature* **1995,** *373*, 117-122.
17. Ho, D. D.; Neumann, A. U.; Perelson, A. S.; Chen, W.; Leonard, J. M.; Markowitz, M. *Nature* **1995,** *373*, 123-126.
18. Perelson, A. S.; Neumann, A. U.; Markowitz, M.; Leonard, J. M.; Ho, D. D. *Science* **1996,** *271*, 1582-1586.
19. Coffin, J. M. *Science* **1995,** *267*, 483-489.
20. Coffin, J. M. *AIDS* **1996,** *10 Suppl 3*, S75-84.
21. UNAIDS/WHO. In *UNAIDS Report on The Global AIDS Epidemic*; United Nations: Geneva, **2010**; p. 102.
22. Palella, F. J.; Delaney, K. M.; Moorman, A. C.; Loveless, M. O.; Fuhrer, J.; Satten, G. A.; Aschman, D. J.; Holmberg, S. D. *N. Engl. J. Med.* **1998,** *338*, 853-860.
23. UNITAID/WHO. In *UNITAID Annual Report*; World Health Organization: Geneva, **2010**; p. 88.
24. UNAIDS/WHO. In *UNAIDS Report on The Global AIDS Epidemic*; United Nations: Geneva, **2012**; p. 107.
25. In *Clinton Health Access Initiative, CHAI, HIV/AIDS,* **2012** Clinton Foundation.
26. Coffin J. M; Hughes S. H; Varmus H. E, Eds. *Retroviruses*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), **1997**.
27. Lodish H; Berk A; Zipursky S. L; Matsudaira P; Baltimore D; Darnell, J. *Molecular Cell Biology*, 4th ed.; W. H. Freeman: New York, **2000**.
28. Leung, D.; Abbenante, G.; Fairlie, D. P. *J. Med. Chem.* **2000,** *43*, 305-341.
29. Hyland, L. J.; Tomaszek, T. A.; Meek, T. D. *Biochemistry* **1991,** *30*, 8454-8463.
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*, -.
32. [Anon]. *Nature* **2000,** *406*, 15-16.
33. Hahn, B. H.; Shaw, G. M.; De Cock, K. M.; Sharp, P. M. *Science* **2000,** *287*, 607-614.
34. Popper, S. J.; Sarr, A. D.; Travers, K. U.; Gueye-Ndiaye, A.; Mboup, S.; Essex, M. E.; Kanki, P. J. *J. Infect. Dis.* **1999,** *180*, 1116-1121.

35. Gao, F.; Bailes, E.; Robertson, D. L.; Chen, Y. L.; Rodenburg, C. M.; Michael, S. F.; Cummins, L. B.; Arthur, L. O.; Peeters, M.; Shaw, G. M.; Sharp, P. M.; Hahn, B. H. *Nature* **1999,** *397*, 436-441.
36. Simon, F.; Mauclere, P.; Roques, P.; Loussert-Ajaka, I.; Muller-Trutwin, M. C.; Saragosti, S.; Georges-Courbot, M. C.; Barre-Sinoussi, F.; Brun-Vezinet, F. *Nat. Med.* **1998,** *4*, 1032-7.
37. Corbet, S.; Muller-Trutwin, M. C.; Versmisse, P.; Delarue, S.; Ayouba, A.; Lewis, J.; Brunak, S.; Martin, P.; Brun-Vezinet, F.; Simon, F.; Barre-Sinoussi, F.; Mauclere, P. *J. Virol.* **2000,** *74*, 529-534.
38. Chen, Z.; Luckay, A.; Sodora, D. L.; Telfer, P.; Reed, P.; Gettie, A.; Kanu, J. M.; Sadek, R. F.; Yee, J.; Ho, D. D.; Zhang, L.; Marx, P. A. *J. Virol.* **1997,** *71*, 3953-3960.
39. Haase, A. T. *Annu. Rev. Immunol.* **1999,** *17*, 625-656.
40. Renneberg, R. In *Biotechnology for beginners;* Demain, A. L. Ed.; Academic Press: Berlin, Germany, **2008**.
41. Cunningham, A. L.; Donaghy, H.; Harman, A. N.; Kim, M.; Turville, S. G. *Curr. Opin. Microbiol.* **2010,** *13*, 524-529.
42. Fauci, A. S.; Pantaleo, G.; Stanley, S.; Weissman, D. *Ann. Intern. Med.* **1996,** *124*, 654-663.
43. Quinn, T. C. *JAMA-J. Am. Med. Assoc.* **1997,** *278*, 58-62.
44. Ratner, L. *Perspect. Drug Discovery Des.* **1991,** *1*, 3-22.
45. Ji, J. X.; Chen, J. J. Y.; Braciale, V. L.; Cloyd, M. W. *J. Leukocyte Biol.* **2007,** *81*, 297-305.
46. Davenport, M. P.; Petravic, J. *PLoS Pathog.* **2010,** *6*, e1000728.
47. Bonsor, K. In *www.health.howstuffworks.com;* Discovery Health, **2001**.
48. Arthur, L. O.; Bess, J. W.; Sowder, R. C.; Benveniste, R. E.; Mann, D. L.; Chermann, J. C.; Henderson, L. E. *Science* **1992,** *258*, 1935-1938.
49. Turner, B. G.; Summers, M. F. *J. Mol. Biol.* **1999,** *285*, 1-32.
50. Abdel-Rahman, H. M.; Al-karamany, G. S.; El-Koussi, N. A.; Youssef, A. F.; Kiso, Y. *Curr. Med. Chem.* **2002,** *9*, 1905-1922.
51. Unger, J. M. Licentiate Thesis, Stockholm University, 2006.
52. Monini, P.; Sgadari, C.; Toschi, E.; Barillari, G.; Ensoli, B. *Nat. Rev. Cancer* **2004,** *4*, 861-875.
53. Matthews, T.; Salgo, M.; Greenberg, M.; Chung, J.; DeMasi, R.; Bolognesi, D. *Nat. Rev. Drug Discovery* **2004,** *3*, 215-225.
54. Pomerantz, R. J.; Horn, D. L. *Nat. Med.* **2003,** *9*, 867-73.
55. Richman, D. D. *Nature* **2001,** *410*, 995-1001.
56. De Clercq, E. *J. Med. Chem.* **2005,** *48*, 1297-1313.
57. Ayouba, A.; Cannou, C.; Nugeyre, M. T.; Barre-Sinoussi, F.; Menu, E. *Retrovirology* **2008,** *5*, 31.
58. Biswas, P.; Tambussi, G.; Lazzarin3, A. *Expert Opin. Pharmacother.* **2007,** *8*, 923-933.
59. 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. *Antimicrob. Agents Ch.* **2007,** *51*, 566-575.
60. 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.;

Matiskella, J.; Kadow, J.; Clapham, P.; Robinson, J.; Colonno, R.; Lin, P. F. *J. Virol.* **2006,** *80*, 4017-4025.
61. Briz, V.; Poveda, E.; Soriano, V. *J. Antimicrob. Chemother.* **2006,** *57*, 619-627.
62. Painter, G. R.; Almond, M. R.; Mao, S.; Liotta, D. C. *Curr. Top. Med. Chem.* **2004,** *4*, 1035-44.
63. Levy, J. A. *AIDS* **2009,** *23*, 147-160.
64. Abel, S.; Russell, D.; Whitlock, L. A.; Ridgway, C. E.; Nedderman, A. N. R.; Walker, D. K. *Brit. J. Clin. Pharmaco.* **2008,** *65*, 60-67.
65. Lalezari, J. P.; Eron, J. J.; Carlson, M. *AIDS* **2003,** *17*, 691-698.
66. Dau, B.; Holodniy, M. *Drugs* **2009,** *69*, 31-50.
67. Ingr, M.; Uhlikova, T.; Strisovsky, K.; Majerova, E.; Konvalinka, J. *Protein Sci.* **2003,** *12*, 2173-2182.
68. Mitsuya, H.; Weinhold, K. J.; Furman, P. A.; St Clair, M. H.; Lehrman, S. N.; Gallo, R. C.; Bolognesi, D.; Barry, D. W.; Broder, S. *Proc. Natl. Acad. Sci. U.S.A.* **1985,** *82*, 7096-7100.
69. 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.
70. Wlodawer, A. *Annu. Rev. Med.* **2002,** *53*, 595-614.
71. Soudeyns, H.; Yao, X. J.; Gao, Q.; Belleau, B.; Kraus, J. L.; Nghe, N. B.; Spira, B.; Wainberg, M. A. *Antimicrob. Agents Ch.* **1991,** *35*, 1386-1390.
72. Levine, S.; Hernandez, D.; Yamanaka, G.; Zhang, S.; Rose, R.; Weinheimer, S.; Colonno, R. J. *Antimicrob Agents Chemother* **2002,** *46*, 2525-2532.
73. Cahn, P.; Cassetti, I.; Wood, R.; Phanuphak, P.; Shiveley, L.; Bethell, R. C.; Sawyer, J. *AIDS* **2006,** *20*, 1261-1268.
74. Emau, P.; Jiang, Y.; Agy, M. B.; Tian, B.; Bekele, G.; Tsai, C. C. *AIDS Res. Ther.* **2006,** *3*, 29.
75. Eisenberg, E. J.; He, G. X.; Lee, W. A. *Nucleosides Nucleotides Nucleic Acids* **2001,** *20*, 1091-1098.
76. De Clercq, E. *Rev. Med. Virol.* **2000,** *10*, 255-277.
77. De Clercq, E. *J. Clin. Virol.* **2004,** *30*, 115-133.
78. De Clercq, E. *Antivir. Res.* **1998,** *38*, 153-179.
79. Drake, S. M. *J. Antimicrob. Chemother.* **2000,** *45*, 417-420.
80. 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.; Obrien, 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. *Antimicrob. Agents Ch.* **1995,** *39*, 2602-2605.
81. 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. *J. Med. Chem.* **1991,** *34*, 2231-2241.
82. 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. *J. Med. Chem.* **1994,** *37*, 999-1014.
83. Stellbrink, H. J. *Eur. J. Med. Res.* **2007,** *12*, 483-495.

84. Das, K.; Clark, A. D.; Lewi, P. J.; Heeres, J.; de Jonge, M. R.; Koymans, L. M. H.; Vinkers, H. M.; Daeyaert, F.; Ludovici, D. W.; Kukla, M. J.; De Corte, B.; Kavash, R. W.; Ho, C. Y.; Ye, H.; Lichtenstein, M. A.; Andries, K.; Pauwels, R.; de Béthune, M.-P.; Boyer, P. L.; Clark, P.; Hughes, S. H.; Janssen, P. A. J.; Arnold, E. *J. Med. Chem.* **2004,** *47*, 2550-2560.
85. Stellbrink, H.-J. *Eur. J. Med. Res.* **2007,** *12*, 483-495.
86. Goebel, F.; Yakovlev, A.; Pozniak, A. L.; Vinogradova, E.; Boogaerts, G.; Hoetelmans, R.; de Bethune, M.-P. P.; Peeters, M.; Woodfall, B. *AIDS* **2006,** *20*.
87. Pommier, Y.; Johnson, A. A.; Marchand, C. *Nat. Rev. Drug Discovery* **2005,** *4*, 236-248.
88. Quashie, P. K.; Sloan, R. D.; Wainberg, M. A. *Bmc Medicine* **2012,** *10*.
89. Gupta, S. P.; Nagappa, A. *Curr. Med. Chem.* **2003,** *10*, 1779-1794.
90. Sluis-Cremer, N.; Tachedjian, G. *Eur. J. Biochem.* **2002,** *269*, 5103-5111.
91. Maurin, C.; Bailly, F.; Cotelle, P. *Curr. Med. Chem.* **2003,** *10*, 1795-1810.
92. Singh, S. B.; Jayasuriya, H.; Salituro, G. M.; Zink, D. L.; Shafiee, A.; Heimbuch, B.; Silverman, K. C.; Lingham, R. B.; Genilloud, O.; Teran, A.; Vilella, D.; Felock, P.; Hazuda, D. *J. Nat. Prod.* **2001,** *64*, 874-882.
93. Jing, N.; Xu, X. *Curr. Drug Targets Infect. Disord.* **2001,** *1*, 79-90.
94. Brigo, A.; Mustata, G. I.; Briggs, J. M.; Moro, S. *Med. chem. (Shariqah (United Arab Emirates))* **2005,** *1*, 263-75.
95. Hazuda, D. J.; Felock, P.; Witmer, M.; Wolfe, A.; Stillmock, K.; Grobler, J. A.; Espeseth, A.; Gabryelski, L.; Schleif, W.; Blau, C.; Miller, M. D. *Science* **2000,** *287*, 646-650.
96. Reinke, R.; Lee, D. J.; Robinson, W. E. *Antimicrob. Agents Chemother.* **2002,** *46*, 3301-3303.
97. Hazuda, D.; Iwamoto, M.; Wenning, L. In *Annu. Rev. Pharmacol. Toxicol.*, 2009; pp. 377-394.
98. Marchand, C.; Zhang, X. C.; Pais, G. C. G.; Cowansage, K.; Neamati, N.; Burke, T. R.; Pommier, Y. *J. Biol. Chem.* **2002,** *277*, 12596-12603.
99. Savarino, A. *Expert Opin. Invest. Drugs* **2006,** *15*, 1507-1522.
100. FDA; notifications. *AIDS alert* **2009,** *24*, 106-7.
101. Marchand, C.; Maddali, K.; Metifiot, M.; Pommier, Y. *Curr. Top. Med. Chem.* **2009,** *9*, 1016-1037.
102. Shimura, K.; Kodama, E.; Sakagami, Y.; Matsuzaki, Y.; Watanabe, W.; Yamataka, K.; Watanabe, Y.; Ohata, Y.; Doi, S.; Sato, M.; Kano, M.; Ikeda, S.; Matsuoka, M. *J. Virol.* **2008,** *82*, 764-774.
103. Hare, S.; Smith, S. J.; Metifiot, M.; Jaxa-Chamiec, A.; Pommier, Y.; Hughes, S. H.; Cherepanov, P. *Mol. Pharmacol.* **2011,** *80*, 565-572.
104. Wainberg, M. A.; Quashie, P. K.; Mesplede, T. *Drugs Future* **2012,** *37*, 697-707.
105. Kramer, R. A.; Schaber, M. D.; Skalka, A. M.; Ganguly, K.; Wongstaal, F.; Reddy, E. P. *Science* **1986,** *231*, 1580-1584.
106. Kohl, N. E.; Emini, E. A.; Schleif, W. A.; Davis, L. J.; Heimbach, J. C.; Dixon, R. A.; Scolnick, E. M.; Sigal, I. S. *Proc. Natl. Acad. Sci. U.S.A.* **1988,** *85*, 4686-4690.
107. Tomasselli, A. G.; Heinrikson, R. L. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **2000,** *1477*, 189-214.
108. DeLucca, G. V.; EricksonViitanen, S.; Lam, P. Y. S. *Drug Discovery Today* **1997,** *2*, 6-18.
109. Kempf, D. J.; Sham, H. L. *Curr. Pharm. Des.* **1996,** *2*, 225-246.

110. Kempf, D. J. *Methods. Enzymol.* **1994,** *241*, 334-354.
111. Erickson, J. W. *Perspect. Drug Discovery Des.* **1993,** *1*, 109-128.
112. Moore, M. L.; Dreyer, G. B. *Perspect. Drug Discovery Des.* **1993,** *1*, 85-108.
113. 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.
114. 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. *Proc. Natl. Acad. Sci. U.S.A.* **1995,** *92*, 2484-2488.
115. 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. *Proc. Natl. Acad. Sci. U.S.A.* **1994,** *91*, 4096-4100.
116. Kaldor, S. W.; Kalish, V. J.; Davies, J. F., 2nd; 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. *J. Med. Chem.* **1997,** *40*, 3979-3985.
117. Kim, E. E.; Baker, C. T.; Dwyer, M. D.; Murcko, M. A.; Rao, B. G.; Tung, R. D.; Navia, M. A. *J. Am. Chem. Soc.* **1995,** *117*, 1181-1182.
118. 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. *Antimicrob. Agents Ch.* **1998,** *42*, 3218-3224.
119. 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. *Antimicrob. Agents Ch.* **2000,** *44*, 2093-2099.
120. Eron, J., Jr.; Yeni, P.; Gathe, J., Jr.; Estrada, V.; DeJesus, E.; Staszewski, S.; Lackey, P.; Katlama, C.; Young, B.; Yau, L.; Sutherland-Phillips, D.; Wannamaker, P.; Vavro, C.; Patel, L.; Yeo, J.; Shaefer, M. *Lancet.* **2006,** *368*, 476-82.
121. Doyon, L.; Tremblay, S.; Bourgon, L.; Wardrop, E.; Cordingley, M. G. *Antivir. Res.* **2005,** *68*, 27-35.
122. Ghosh, A. K.; Dawson, Z. L.; Mitsuya, H. *Bioorg. Med. Chem.* **2007,** *15*, 7576-7580.
123. Flexner, C. *N. Engl. J. Med.* **1998,** *338*, 1281-1292.
124. Walli, R.; Segerer, S.; Herfort, O.; Michl, G.; Loch, O.; Dieterle, C.; Muhlbayer, D.; Bogner, J. R.; Landgraf, R.; Demant, T.; Goebel, F. D. *AIDS* **1998,** *12*, S52-S52.
125. Lo, J. C.; Mulligan, K.; Tai, V. W.; Algren, H.; Schambelan, M. *Lancet* **1998,** *351*, 867-870.
126. Miller, K. D.; Jones, E.; Yanovski, J. A.; Shankar, R.; Feuerstein, I.; Falloon, J. *Lancet* **1998,** *351*, 871-875.
127. Carr, A.; Samaras, K.; Burton, S.; Law, M.; Freund, J.; Chisholm, D. J.; Cooper, D. A. *AIDS* **1998,** *12*, F51-F58.
128. Viraben, R.; Aquilina, C. *AIDS* **1998,** *12*, F37-F39.

129. Martinez, E.; Casamitjana, R.; Conget, I.; Gatell, J. M. *AIDS* **1998,** *12*, 2077-2079.
130. Brik, A.; Wong, C. H. *Org. Biomol. Chem.* **2003,** *1*, 5-14.
131. Urban, J.; Konvalinka, J.; Stehlikova, J.; Gregorova, E.; Majer, P.; Soucek, M.; Andreansky, M.; Fabry, M.; Strop, P. *FEBS Lett.* **1992,** *298*, 9-13.
132. Larder, B. A.; Darby, G.; Richman, D. D. *Science* **1989,** *243*, 1731-1734.
133. Larder, B. A.; Kemp, S. D. *Science* **1989,** *246*, 1155-1158.
134. Erice, A.; Balfour, H. H. *N. Engl. J. Med.* **1993,** *329*, 1124-1124.
135. Gulick, R. M.; Mellors, J. W.; Havlir, D.; Eron, J. J.; Gonzalez, C.; McMahon, D.; Richman, D. D.; Valentine, F. T.; Jonas, L.; Meibohm, A.; Emini, E. A.; Chodakewitz, J. A. *N. Engl. J. Med.* **1997,** *337*, 734-739.
136. Collier, A. C.; Coombs, R. W.; Schoenfeld, D. A.; Bassett, R. L.; Timpone, J.; Baruch, A.; Jones, M.; Facey, K.; Whitacre, C.; McAuliffe, V. J.; Friedman, H. M.; Merigan, T. C.; Reichman, R. C.; Hooper, C.; Corey, L. *N. Engl. J. Med.* **1996,** *334*, 1011-1017.
137. Mocroft, A.; Vella, S.; Benfield, T. L.; Chiesi, A.; Miller, V.; Gargalianos, P.; Monforte, A. D.; Yust, I.; Bruun, J. N.; Phillips, A. N.; Lundgren, J. D. *Lancet* **1998,** *352*, 1725-1730.
138. Stephenson, J. *JAMA-J. Am. Med. Assoc.* **1997,** *277*, 614-616.
139. Siliciano, J. D.; Siliciano, R. F. *J. Antimicrob. Chemother.* **2004,** *54*, 6-9.
140. Chun, T. W.; Stuyver, L.; Mizell, S. B.; Ehler, L. A.; Mican, J. A. M.; Baseler, M.; Lloyd, A. L.; Nowak, M. A.; Fauci, A. S. *Proc. Natl. Acad. Sci. U.S.A.* **1997,** *94*, 13193-13197.
141. Finzi, D.; Hermankova, M.; Pierson, T.; Carruth, L. M.; Buck, C.; Chaisson, R. E.; Quinn, T. C.; Chadwick, K.; Margolick, J.; Brookmeyer, R.; Gallant, J.; Markowitz, M.; Ho, D. D.; Richman, D. D.; Siliciano, R. F. *Science* **1997,** *278*, 1295-1300.
142. Kinlaw, W. B.; Marsh, B. *Endocrinology* **2004,** *145*, 484-486.
143. Gerard, Y.; Maulin, L.; Yazdanpanah, Y.; De la Tribonniere, X.; Amiel, C.; Maurage, C. A.; Robin, S.; Sablonniere, B.; Dhennain, C.; Mouton, Y. *AIDS* **2000,** *14*, 2723-2730.
144. Saves, M.; Chene, G.; Ducimetiere, P.; Leport, C.; Le Moal, G.; Amouyel, P.; Arveiler, D.; Ruidavets, J. B.; Reynes, J.; Bingham, A.; Raffi, F. *Clin. Infect. Dis.* **2003,** *37*, 292-298.
145. Attaran, A.; Sachs, J. *Lancet* **2001,** *357*, 57-61.
146. Steinbrook, R.; Drazen, J. M. *N. Engl. J. Med.* **2001,** *344*, 1781-1782.
147. Babine, R. E.; Bender, S. L. *Chem. Rev.* **1997,** *97*, 1359-1472.
148. Ripka, A. S.; Rich, D. H. *Curr. Opin. Chem. Biol.* **1998,** *2*, 441-452.
149. Brown, N. J.; Vaughan, D. E. *Circulation* **1998,** *97*, 1411-1420.
150. 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.; et al. *J. Med. Chem.* **1991,** *34*, 2852-2857.
151. Reetz, M. T.; Merk, C.; Mehler, G. *Chem. Commun.* **1998**, 2075-2076.
152. Huff, J. R. *J. Med. Chem.* **1991,** *34*, 2305-2314.
153. Aletrman, M., Uppsala University, 2001.
154. Pettit, S. C.; Gulnik, S.; Everitt, L.; Kaplan, A. H. *J. Virol.* **2003,** *77*, 366-374.
155. Navia, M. A.; Fitzgerald, P. M. D.; Mckeever, B. M.; Leu, C. T.; Heimbach, J. C.; Herber, W. K.; Sigal, I. S.; Darke, P. L.; Springer, J. P. *Nature* **1989,** *337*, 615-620.
156. Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B. H.; Wlodawer, A. *Science* **1989,** *246*, 1149-1152.

157. Wlodawer, A.; Miller, M.; Jaskolski, M.; Sathyanarayana, B. K.; Baldwin, E.; Weber, I. T.; Selk, L. M.; Clawson, L.; Schneider, J.; Kent, S. B. H. *Science* **1989,** *245*, 616-621.
158. Lapatto, R.; Blundell, T.; Hemmings, A.; Overington, J.; Wilderspin, A.; Wood, S.; Merson, J. R.; Whittle, P. J.; Danley, D. E.; Geoghegan, K. F.; Hawrylik, S. J.; Lee, S. E.; Scheld, K. G.; Hobart, P. M. *Nature* **1989,** *342*, 299-302.
159. Swain, A. L.; Miller, M. M.; Green, J.; Rich, D. H.; Schneider, J.; Kent, S. B.; Wlodawer, A. *Proc. Natl. Acad. Sci. U.S.A.* **1990,** *87*, 8805-9.
160. Fitzgerald, P. M.; McKeever, B. M.; VanMiddlesworth, J. F.; Springer, J. P.; Heimbach, J. C.; Leu, C. T.; Herber, W. K.; Dixon, R. A.; Darke, P. L. *J. Biol. Chem.* **1990,** *265*, 14209-14219.
161. Sriphanlop. In *Google Images;* Sriphanlop_fig1.png Ed.; people.mbi.ucla.edu: Los Angeles, 2009.
162. Camarasa, M. J.; Velazquez, S.; San-Felix, A.; Perez-Perez, M. J.; Gago, F. *Antivir. Res.* **2006,** *71*, 260-267.
163. Pettit, S. C.; Michael, S. F.; Swanstrom, R. *Perspect. Drug Discovery Des.* **1993,** *1*, 69-83.
164. Schechte.I; Berger, A. *Biochem. Biophys. Res. Commun.* **1967,** *27*, 157-162.
165. Appelt, K. *Perspect. Drug Discovery Des.* **1993,** *1*, 23-48.
166. Fitzgerald, P. M. D.; Springer, J. P. *Annu. Rev. Bioph. Biom.* **1991,** *20*, 299-320.
167. Prabu-Jeyabalan, M.; Nalivaika, E.; Schiffer, C. A. *Structure* **2002,** *10*, 369-381.
168. Prabu-Jeyabalan, M.; Nalivaika, E.; Schiffer, C. A. *J. Mol. Biol.* **2000,** *301*, 1207-1220.
169. Ridky, T. W.; Cameron, C. E.; Cameron, J.; Leis, J.; Copeland, T.; Wlodawer, A.; Weber, I. T.; Harrison, R. W. *J. Biol. Chem.* **1996,** *271*, 4709-4717.
170. Tozser, J.; Bagossi, P.; Weber, I. T.; Louis, J. M.; Copeland, T. D.; Oroszlan, S. *J. Biol. Chem.* **1997,** *272*, 16807-16814.
171. Smith, R.; Brereton, I. M.; Chai, R. Y.; Kent, S. B. *Nat Struct Biol* **1996,** *3*, 946-50.
172. Northrop, D. B. *Acc. Chem. Res.* **2001,** *34*, 790-797.
173. Hyland, L. J.; Tomaszek, T. A.; Roberts, G. D.; Carr, S. A.; Magaard, V. W.; Bryan, H. L.; Fakhoury, S. A.; Moore, M. L.; Minnich, M. D.; Culp, J. S.; Desjarlais, R. L.; Meek, T. D. *Biochemistry* **1991,** *30*, 8441-8453.
174. Piana, S.; Sebastiani, D.; Carloni, P.; Parrinello, M. *J. Am. Chem. Soc.* **2001,** *123*, 8730-8737.
175. Yamazaki, T.; Nicholson, L. K. e. a. *J. Am. Chem. Soc.* **1994,** *116*, 10791-10792.
176. Wang, Y. X.; Freedberg, D. I.; Yamazaki, T.; Wingfield, P. T.; Stahl, S. J.; Kaufman, J. D.; Kiso, Y.; Torchia, D. A. *Biochemistry* **1996,** *35*, 9945-9950.
177. Jaskolski, M.; Tomasselli, A. G.; Sawyer, T. K.; Staples, D. G.; Heinrikson, R. L.; Schneider, J.; Kent, S. B. H.; Wlodawer, A. *Biochemistry* **1991,** *30*, 1600-1609.
178. Davies, D. R. *Annu. Rev. Bioph. Biom.* **1990,** *19*, 189-215.
179. Giacovazzo, C.; Monaco, H. L.; Artioli, G.; Viterbo, D.; Ferraris, G.; Gilli, G.; Zanoti, G.; Catti, M. *Fundamentals of Crystallography*; Oxford University Press: Oxford, **2002**.
180. Bennett, D. W. *Undertsanding single-crystal X-ray crystallography*, First ed.; Wiley-VCH Verlag GmbH & Co. : Berlin, Germany, **2010**.
181. Carter, C. W. J.; Sweet, R. M. *Methods in Enzymology*; Academic Press: New York, **1997**; Vol. 276.
182. Appelt, K.; Bacquet, R. J.; Bartlett, C. A.; Booth, C. L. J.; Freer, S. T.; Fuhry, M. A. M.; Gehring, M. R.; Herrmann, S. M.; Howland, E. F.; Janson, C. A.; Jones, T. R.; Kan, C.

C.; Kathardekar, V.; Lewis, K. K.; Marzoni, G. P.; Matthews, D. A.; Mohr, C.; Moomaw, E. W.; Morse, C. A.; Oatley, S. J.; Ogden, R. C.; Reddy, M. R.; Reich, S. H.; Schoettlin, W. S.; Smith, W. W.; Varney, M. D.; Villafranca, J. E.; Ward, R. W.; Webber, S.; Webber, S. E.; Welsh, K. M.; White, J. *J. Med. Chem.* **1991,** *34*, 1925-1934.
183. Miller, M.; Jaskolski, M.; Rao, J. K. M.; Leis, J.; Wlodawer, A. *Nature* **1989,** *337*, 576-579.
184. Wlodawer, A.; Vondrasek, J. *Annu. Rev. Bioph. Biom.* **1998,** *27*, 249-284.
185. Lapatto, R.; Blundell, T.; Hemmings, A.; Overington, J.; Wilderspin, A.; Wood, S.; Merson, J. R.; Whittle, P. J.; Danley, D. E.; Geoghegan, K. F.; Hawrylik, S. J.; Lee, S. E.; Scheld, K. G.; Hobart, P. M. *Nature* **1989,** *342*, 299-302.
186. Spinelli, S.; Liu, Q. Z.; Alzari, P. M.; Hirel, P. H.; Poljak, R. J. *Biochimie* **1991,** *73*, 1391-1396.
187. Miller, M. *Biopolymers* **2010,** *94*, 521-529.
188. Daopin, S.; Piez, K. A.; Ogawa, Y.; Davies, D. R. *Science* **1992,** *257*, 369-373.
189. Prashar, V.; Bihani, S. C.; Das, A.; Rao, D. R.; Hosur, M. V. *Biochem. Biophys. Res. Commun.* **2010,** *396*, 1018-1023.
190. Foulkes, J. E.; Prabu-Jeyabalan, M.; Cooper, D.; Henderson, G. J.; Harris, J.; Swanstrom, R.; Schiffer, C. A. *J. Virol.* **2006,** *80*, 6906-6916.
191. 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. *J. Med. Chem.* **2000,** *43*, 3495-3504.
192. Wu, X. Y.; Ohrngren, P.; Ekegren, J. K.; Unge, J.; Unge, T.; Wallberg, H.; Samuelsson, B.; Hallberg, A.; Larhed, M. *J. Med. Chem.* **2008,** *51*, 2586-2586.
193. Liu, F. L.; Boross, P. I.; Wang, Y. F.; Tozser, J.; Louis, J. M.; Harrison, R. W.; Weber, I. T. *J. Mol. Biol.* **2005,** *354*, 789-800.
194. Coman, R. M.; Robbins, A. H.; Fernandez, M. A.; Gilliland, C. T.; Sochet, A. A.; Goodenow, M. M.; McKenna, R.; Dunn, B. M. *Biochemistry* **2008,** *47*, 731-743.
195. Kozisek, M.; Bray, J.; Rezacova, P.; Saskova, K.; Brynda, J.; Pokorna, J.; Mammano, F.; Rulisek, L.; Konvalinka, J. *J. Mol. Biol.* **2007,** *374*, 1005-1016.
196. Shen, C. H.; Wang, Y. F.; Kovalevsky, A. Y.; Harrison, R. W.; Weber, I. T. *FEBS J.* **2010,** *277*, 3699-3714.
197. Stoll, V.; Qin, W. Y.; Stewart, K. D.; Jakob, C.; Park, C.; Walter, K.; Simmer, R. L.; Helfrich, R.; Bussiere, D.; Kao, J.; Kempf, D.; Sham, H. L.; Norbeck, D. W. *Bioorg. Med. Chem.* **2002,** *10*, 2803-2806.
198. Reddy, G.; Ali, A.; Nalam, M. N. L.; Anjum, S. G.; Cao, H.; Nathans, R. S.; Schiffer, C. A.; Rana, T. M. *J. Med. Chem.* **2007,** *50*, 4316-4328.
199. Klei, H. E.; Kish, K.; Lin, P. F. M.; Guo, Q.; Friborg, J.; Rose, R. E.; Zhang, Y.; Goldfarb, V.; Langley, D. R.; Wittekind, M.; Sheriff, S. *J. Virol.* **2007,** *81*, 9525-9535.
200. Clemente, J. C.; Coman, R. M.; Thiaville, M. M.; Janka, L. K.; Jeung, J. A.; Nukoolkarn, S.; Govindasamy, L.; Agbandje-McKenna, M.; McKenna, R.; Leelamanit, W.; Goodenow, M. M.; Dunn, B. M. *Biochemistry* **2006,** *45*, 5468-5477.
201. Thaisrivongs, S.; Skulnick, H. I.; Turner, S. R.; Strohbach, J. W.; Tommasi, R. A.; Johnson, P. D.; Aristoff, P. A.; Judge, T. M.; Gammill, R. B.; Morris, J. K.; Romines, K. R.; Chrusciel, R. A.; Hinshaw, R. R.; Chong, K. T.; Tarpley, W. G.; Poppe, S. M.; Slade, D. E.; Lynn, J. C.; Horng, M. M.; Tomich, P. K.; Seest, E. P.; Dolak, L. A.; Howe, W. J.; Howard, G. M.; Schwende, F. J.; Toth, L. N.; Padbury, G. E.; Wilson, G. J.; Shiou, L. H.;

Zipp, G. L.; Wilkinson, K. F.; Rush, B. D.; Ruwart, M. J.; Koeplinger, K. A.; Zhao, Z. Y.; Cole, S.; Zaya, R. M.; Kakuk, T. J.; Janakiraman, M. N.; Watenpaugh, K. D. *J. Med. Chem.* **1996,** *39*, 4349-4353.
202. Muzammil, S.; Armstrong, A. A.; Kang, L. W.; Jakalian, A.; Bonneau, P. R.; Schmelmer, V.; Amzel, L. M.; Freire, E. *J. Virol.* **2007,** *81*, 5144-5154.
203. Bandaranayake, R. M.; Kolli, M.; King, N. M.; Nalivaika, E. A.; Heroux, A.; Kakizawa, J.; Sugiura, W.; Schiffer, C. A. *J. Virol.* **2010,** *84*, 9995-10003.
204. Kumar, M. R.; Hosur, R. V.; Roy, K. B.; Miles, H. T.; Govil, G. *Biochemistry* **1985,** *24*, 7703-7711.
205. Wuthrich, K. *J. Biol. Chem.* **1990,** *265*, 22059-22062.
206. Simon, B.; Madl, T.; Mackereth, C. D.; Nilges, M.; Sattler, M. *Angew. Chem. Int. Ed.* **2010,** *49*, 1967-1970.
207. Wuthrich, K. *J. Biol. Chem.* **1990,** *265*, 22059-22062.
208. Wuthrich, K. In *NMR in structural biology: A Collection of Papers by Kurt Wuthrich* Wuthrich, K. Ed.; World Scientific Publising Co. Pte. Ltd.: Singapore, 1995; pp. 11-15.
209. Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979,** *71*, 4546-4553.
210. Wuethrich, K. In *[Nova Acta Leopoldina; Nuclear magnetic resonance in biology and medicine];* Oelssner, W.; Stieve, F. E. Eds., 1992; pp. 47-56.
211. Rosen, O.; Sharon, M.; Quadt-Akabayov, S. R.; Anglister, J. *Proc. Natl. Acad. Sci. U.S.A.* **2006,** *103*, 13950-13955.
212. Thiele, C. M.; Petzold, K.; Schleucher, J. *Chem. Eur. J.* **2009,** *15*, 585-588.
213. Bothnerby, A. A.; Stephens, R. L.; Lee, J. M.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984,** *106*, 811-813.
214. Tisné, C.; Hantz, E.; Hartmann, B.; Delepierre, M. *J. Mol. Biol.* **1998,** *279*, 127-142.
215. Konrat, R.; Burghardt, I.; Bodenhausen, G. *J. Am. Chem. Soc.* **1991,** *113*, 9135-9140.
216. Konrat, R.; Zieger, G.; Sterk, H. *Chem. Phys. Lett.* **1993,** *212*, 78-80.
217. Vincent, S. J. F.; Zwahlen, C.; Bodenhausen, G. *J. Am. Chem. Soc.* **1993,** *115*, 9202-9209.
218. van de waterbeemd, H.; Carter, R. E.; Grassy, G.; Kubinyi, H.; Martin, Y. C.; Tute, M. S.; Willett, P. *Academic P: San Diego* **1998**.
219. Leach, A. R. *Molecular Modelling: principles and applications*; Longman: Harlow, **1996**.
220. Allinger, B. U. a. N. L. *Molecular Mechanics*; ACS monographs,American Chemical Society: Washington, D.C, **1982**; Vol. volume 177.
221. A. Szabo and N. S. Ostlund *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; McGraw-Hill: New York, **1989**.
222. The Shodor Education Foundation, Inc., **2000**.
223. Stewart, J. J. P. *J. Comput. Chem.* **1989,** *10*, 209-220.
224. Stewart, J. J. P. *J. Comput. Chem.* **1989,** *10*, 221-264.
225. Kalos, M. H.; Whitlock, P. A. *Monte Carlo Methods*, Second ed.; Wiley-VCH, **2008**.
226. Parr, R. G. *Annu. Rev. Phys. Chem.* **1983,** *34*, 631-656.
227. R. G. Parr and W. Yang *Density-Functional Theory of Atoms and Molecules*; Oxford: New York, **1989**; Vol. 16.
228. Hohenberg, P.; Kohn, W. *Phys. Rev. B* **1964,** *136*, B864-B871.
229. Kohn, W.; Sham, H. L. *Phys. Rev. A* **1965,** *140*, A1133-A1138.
230. Harrison, N. M.: University of California, **2008**; pp. 1-26.

231. Zhang, D. W.; Zhang, J. Z. H. *J. Chem. Phys.* **2003,** *119*, 3599-3605.
232. Kaschner, R.; Hohl, D. *J. Phys. Chem. A* **1998,** *102*, 5111-5116.
233. Schlick, T. *Molecular modeling and simulation: an interdisciplinary guide*, Second ed.; Springer: Berlin, **2002**.
234. Barril, X.; Soliva, R. *Mol. Biosyst.* **2006,** *2*, 660-681.
235. Rangarajan, E. S.; Proteau, A.; Cui, Q. Z.; Logan, S. M.; Potetinova, Z.; Whitfield, D.; Purisima, E. O.; Cygler, M.; Matte, A.; Sulea, T.; Schoenhofen, I. C. *J. Biol. Chem.* **2009,** *284*, 20989-21000.
236. Tur, D.; CESCA: New york, **2010**.
237. Mahmoud, E. S. S., University of Bath, **2009**.
238. Thomas, G. *Medicinal Chemistry : An Introduction*, 2nd ed.; John Wiley & sons Inc.: London, **2007**.
239. Showalter, S. A.; Bruschweiler, R. *J. Chem. Theory Comput.* **2007,** *3*, 961-975.
240. Karplus, M.; McCammon, J. A. *Nat. Struct. Biol.* **2002,** *9*, 788-788.
241. Stella, L.; Melchionna, S. *J. Chem. Phys.* **1998,** *109*, 10115-10117.
242. Stocker, U.; Juchli, D.; van Gunsteren, W. F. *Mol. Simul.* **2003,** *29*, 123-138.
243. Karplus, M. *Acc. Chem. Res.* **2002,** *35*, 321-323.
244. Aqvist, J.; Wennerstrom, P.; Nervall, M.; Bjelic, S.; Brandsdal, B. O. *Chem. Phys. Lett.* **2004,** *384*, 288-294.
245. Goss, P. J. E.; Peccoud, J. *Proc. Natl. Acad. Sci. U.S.A.* **1998,** *95*, 6750-6755.
246. Schilstra, M. J.; Martin, S. R. *Methods Enzymol.: Computer Methods, Part B* **2009,** *467*, 381-409.
247. Bender, C.; Henjes, F.; Frohlich, H.; Wiemann, S.; Korf, U.; Beissbarth, T. *Bioinformatics* **2010,** *26*, i596-i602.
248. Sagui, C.; Darden, T. A. *Annu. Rev. Bioph. Biom.* **1999,** *28*, 155-179.
249. Elber, R.; Ghosh, A.; Cardenas, A. *Acc. Chem. Res.* **2002,** *35*, 396-403.
250. *Acc. Chem. Res.* **2002,** *35*, 321-489.
251. Senn, H. M.; Thiel, W. *Angew. Chem. Int. Ed.* **2009,** *48*, 1198-1229.
252. Senn, H. M.; Thiel, S.; Thiel, W. *J. Chem. Theory Comput.* **2005,** *1*, 494-505.
253. Zhang, Y.; Liu, H.; Yang, W. *J. Chem. Phys.* **2000,** *112*, 3483-3492.
254. Kastner, J.; Senn, H. M.; Thiel, S.; Otte, N.; Thiel, W. *J. Chem. Theory Comput.* **2006,** *2*, 452-461.
255. Hofer, T. S.; Pribil, A. B.; Randolf, B. R. *Pure Appl. Chem.* **2008,** *80*, 1195-1210.
256. Hu, H.; Yang, W. T. *Annu. Rev. Phys. Chem.* **2008,** *59*, 573-601.
257. Thiel, W. In *Multiscale Simulation Methods in Molecular Sciences,*Grotendorst, J.; Attig, N.; Bluegel, S.; Marx, D. Eds.; Julich Supercomputing Centre, **2009**; pp. 203-214.
258. Lengauer, T.; Rarey, M. *Curr. Opin. Struct. Biol.* **1996,** *6*, 402-406.
259. Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. *Nat. Rev. Drug Discovery* **2004,** *3*, 935-949.
260. Jorgensen, W. L. *Science* **1991,** *254*, 954-955.
261. Wei, B. Q.; Weaver, L. H.; Ferrari, A. M.; Matthews, B. W.; Shoichet, B. K. *J. Mol. Biol.* **2004,** *337*, 1161-1182.
262. Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. *J. Mol. Biol.* **1982,** *161*, 269-288.
263. Huang, S. Y.; Zou, X. Q. *Int. J. Mol. Sci.* **2010,** *11*, 3016-3034.

264. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997,** *267*, 727-748.
265. Yuriev, E.; Agostino, M.; Ramsland, P. A. *J. Mol. Recognit.* **2011,** *24*, 149-164.
266. Waszkowycz, B.; Clark, D. E.; Gancia, E. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011,** *1*, 229-259.
267. Spyrakis, F.; Cozzini, P.; Kellogg, G. E.; Abraham, D. J. *Docking and Scoring in Drug Discovery*; John Wiley & Sons, Inc., 2003.
268. Ewing, T. J. A.; Makino, S.; Skillman, A. G.; Kuntz, I. D. *J. Comput.-Aided Mol. Des.* **2001,** *15*, 411-428.
269. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998,** *19*, 1639-1662.
270. Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.* **1996,** *261*, 470-489.
271. Jain, A. N. *J. Med. Chem.* **2003,** *46*, 499- 511.
272. Abagyan, R.; Totrov, M.; Kuznetsov, D. *J. Comput. Chem.* **1994,** *15*, 488-506.
273. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004,** *47*, 1739-1749.
274. Wu, G.; Robertson, D. H.; Brooks, C. L.; Vieth, M. *J. Comput. Chem.* **2003,** *24*, 1549-1562.
275. Venkatachalam, C. M.; Jiang, X.; Oldfield, T.; Waldman, M. *J. Mol. Graphics Modell.* **2003,** *21*, 289-307.
276. Liu, M.; Wang, S. M. *J. Comput.-Aided Mol. Des.* **1999,** *13*, 435-451.
277. Plewczynski, D.; Łaźniewski, M.; Augustyniak, R.; Ginalski, K. *J. Comput. Chem.* **2011,** *32*, 742-755.
278. Moitessier, N.; Englebienne, P.; Lee, D.; Lawandi, J.; Corbeil, C. R. *Brit. J. Pharmacol.* **2008,** *153*, S7-S26.
279. Wang, R.; Lu, Y.; Fang, X.; Wang, S. *J. Chem. Inf. Comput. Sci.* **2004,** *44*, 2114-2125.
280. Kolb, P.; Irwin, J. J. *Curr. Top. Med. Chem.* **2009,** *9*, 755-770.
281. Baber, J. C.; Thompson, D. C.; Cross, J. B.; Humblet, C. *J. Chem. Inf. Model.* **2009,** *49*, 1889-1900.
282. Deeks, S. G. *Lancet* **2003,** *362*, 2002-2011.
283. Waters, L.; Nelson, M. *Int. J. Clin. Pract.* **2007,** *61*, 983-990.
284. Waters, L.; Nelson, M. *Int. J. Clin. Pract.* **2007,** *61*, 999-1014.
285. Griffin, G. W.; Marchand, A. P. *Chem. Rev.* **1989,** *89*, 997-1010.
286. Marchand, A. P. In *Advances in theoretically interesting molecules.;* Thumemel, R. P. Ed.; JAI: Greenwhich, CT, 1989; pp. 375-397.
287. Marchand, A. P. *Chem. Rev.* **1989,** *89*, 1011-1033.
288. Geldenhuys, W. J.; Malan, S. F.; Bloomquist, J. R.; Marchand, A. P.; Van der Schyf, C. J. *Med. Res. Rev* **2005,** *25*, 21-48.
289. Brookes, K. B.; Hickmott, P. W.; Jutle, K. K.; Schreyer, C. A. *S. Afr. J. Chem.* **1992,** *45*, 8-11.
290. Kitagawa, K.; Mizobuchi, N.; Hama, T.; Hibi, T.; Konishi, R.; Futaki, S. *Chem. Pharm. Bull.* **1997,** *45*, 1782-1787.
291. Ito, F. M.; Petroni, J. M.; de Lima, D. P.; Beatriz, A.; Marques, M. R.; de Moraes, M. O.; Costa-Lotufo, L. V.; Montnegro, R. C.; Magalhaes, H. I. F.; Pessoa, C. D. O. *Molecules* **2007,** *12*, 271-282.

292. Oliver, D. W.; Dekker, T. G.; Snyckers, F. O. *Eur. J. Med. Chem.* **1991,** *26*, 375-379.
293. Oliver, D. W.; Dekker, T. G.; Snyckers, F. O.; Fourie, T. G. *J. Med. Chem.* **1991,** *34*, 851-854.
294. Van der Schyf, C. J.; Dekker, T. G.; Snyckers, F. O. *Arch. Pharm.* **1986,** *319*, 409-415.
295. Martins, F. J. C.; Viljoen, A. M.; Kruger, H. G.; Wessels, P. L. *Tetrahedron* **1993,** *49*, 6527-6532.
296. Ndlovu, N. I., University of KwaZulu-Natal, **2009**.
297. Makatini, M. M., University of KwaZulu-Natal, **2010**.
298. Makatini, M. M.; Petzold, K.; Sriharsha, S. N.; Soliman, M. E. S.; Honarparvar, B.; Arvidsson, P. I.; Sayed, Y.; Govender, P.; Maguire, G. E. M.; Kruger, H. G.; Govender, T. *Bioorg. Med. Chem. Lett.* **2011,** *21*, 2274-2277.
299. Oliver, D.; Malan, S. *Med. Chem. Res.* **2008,** *17*, 137-151.
300. Geldenhuys, W. J.; Malan, S. F.; Bloomquist, J. R.; Marchand, A. P.; Van der Schyf, C. J. *Medicinal Research Reviews* **2005,** *25*, 21-48.
301. Geldenhuys, W. J.; Malan, S. F.; Bloomquist, J. R.; Van der Schyf, C. J. *Bioorganic & Medicinal Chemistry* **2007,** *15*, 1525-1532.
302. Nagasawa, H. T.; Elberling, J. A.; Shirota, F. N. *Journal of Medicinal Chemistry* **1975,** *18*, 826-830.
303. Nagasawa, H. T.; Elberlin.Ja; Shirota, F. N. *Journal of Medicinal Chemistry* **1973,** *16*, 823-826.
304. Zah, J.; Terre'Blanche, G.; Erasmus, E.; Malan, S. F. *Bioorganic & Medicinal Chemistry* **2003,** *11*, 3569-3578.
305. Joubert, J.; Geldenhuys, W. J.; Van der Schyf, C. J.; Oliver, D. W.; Kruger, H. G.; Govender, T.; Malan, S. F. *Chem. Med. Chem.* **2012,** *7*, 375-384.
306. Davies, W. L.; Grunert, R. R.; Haff, R. F.; McGahen, J. W.; Neumayer, E. M.; Paulshock, M.; Watts, J. C.; Wood, T. R.; Hermann, E. C.; Hoffmann, C. E. *Science (Washington, DC, U. S.)* **1964,** *144*, 862-863.
307. Davies, W. L.; Grunert, R. R.; Hoffmann, C. E. *J. Immunol.* **1965,** *95*, 1090-1094.
308. Grunert, R. R.; McGahen, J. W.; Davies, W. L. *Virology* **1965,** *26*, 262-9.
309. Gerzon, K.; Kau, D. *J. Med. Chem.* **1967,** *10*, 189-199.
310. Onajole, O. K.; Belewa, X. V.; Coovadia, Y.; Govender, T.; Kruger, H. G.; Maguire, G. E. M.; Naidu, D.; Somai, B.; Singh, N.; Govender, P. *Medicinal Chemistry Research* **2010**, No pp yet given.
311. Rapala, R. T.; Kraay, R. J.; Gerzon, K. *J. Med. Chem.* **1965,** *8*, 580-583.
312. Voldeng, A. N.; Bradley, C. A.; Kee, R. D.; King, E. L.; Melder, F. L. *J. Pharm. Sci.* **1968,** *57*, 1053-1055.
313. Gerzon, K.; Tobias, D. J.; Holmes, R. E.; Rathbun, R. E.; Kattau, R. W. *J. Med. Chem.* **1967,** *10*, 603-606.
314. Lee, R. E.; Protopopova, M.; Crooks, E.; Slayden, R. A.; Terrot, M.; Barry, C. E. *Journal of Combinatorial Chemistry* **2003,** *5*, 172-187.
315. Bogatcheva, E.; Hanrahan, C.; Nikonenko, B.; Samala, R.; Chen, P.; Gearhart, J.; Barbosa, F.; Einck, L.; Nacy, C. A.; Protopopova, M. *Journal of Medicinal Chemistry* **2006,** *49*, 3045-3048.
316. Onajole, O. K.; Govender, P.; van Helden, P. D.; Kruger, H. G.; Maguire, G. E. M.; Wiid, I.; Govender, T. *Eur. J. Med. Chem.* **2010,** *45*, 2075-2079.

317. Balzarini, J.; Orzeszko, B.; Maurin, J. K.; Orzeszko, A. *European Journal of Medicinal Chemistry* **2007,** *42*, 993-1003.
318. Hwang, H. S. T., H; Liu, J.; Morisseau, C.; Hammock, B.D.;. *Journal of Medicinal Chemistry* **2007,** *50*, 3825-3840.
319. Wennekes, T.; van den Berg, R.; Donker, W.; van der Marel, G. A.; Donker, W.; van der Marel, G. A.; Strijland, A.; Aerts, J.; Overkleeft, H. S. *Journal of Organic Chemistry* **2007,** *72*, 1088-1097.
320. Tsuzuki, N.; Hama, T.; Kawada, M.; Hasui, A.; Konishi, R.; Shiwa, S.; Ochi, Y.; Futaki, S.; Kitagawa, K. *Journal of Pharmaceutical Sciences* **1994,** *83*, 481-484.
321. Oliver, D. W.; Dekker, T. G.; Snyckers, F. O.; Fourie, T. G. *Journal of Medicinal Chemistry* **1991,** *34*, 851-854.
322. Oliver, D. W.; Dekker, T. G.; Snyckers, F. O. *European Journal of Medicinal Chemistry* **1991,** *26*, 375-379.
323. Oliver, D. W.; Dekker, D. G.; Snyckers, F. O. *Arzneimittel-Forsch.* **1991,** *41*, 549-552.
324. Aigami, K.; Inamoto, Y.; Takaishi, N.; Fujikura, Y.; Takatsuki, A.; Tamura, G. *J. Med. Chem.* **1976,** *19*, 536-540.
325. Inamoto, Y.; Aigami, K.; Kadono, T.; Nakayama, H.; Takatsuki, A.; Tamura, G. *J. Med. Chem.* **1977,** *20*, 1371-1374.
326. Inamoto, Y.; Aigami, K.; Takaishi, N.; Fujikura, Y.; Ohsugi, M.; Ikeda, H.; Tsuchihashi, K.; Takatsuki, A.; Tamura, G. *J. Med. Chem.* **1979,** *22*, 1206-1214.
327. Onajole, O. K.; Govender, K.; Govender, P.; van Helden, P. D.; Kruger, H. G.; Maguire, G. E. M.; Muthusamy, K.; Pillay, M.; Wiid, I.; Govender, T. *Eur. J. Med. Chem.* **2009,** *44*, 4297-4305.
328. Govender, T.; Kruger, H. G.; Makatini, M.; Onajole, O. K. *Struct. Chem.* **2008,** *19*, 719-726.
329. Onajole, O. K.; Govender, T.; Makatini, M.; Kruger, H. G. *Magn. Reson. Chem.* **2008,** *46*, 1007-1014.
330. J.C.A. Boeyens, L. M. C., G.N. Nelson and T.G. Fourie. *South African Journal of Chemistry-Suid-Afrikaanse Tydskrif Vir Chemie* **1994,** *47*, 72.
331. Van der Schyf, C. J.; Squier, G. J.; Coetzee, W. A. *Pharmacological Research* **1986,** *18*, 407-417.
332. Geldenhuys, W. J.; Terre'Blanche, G.; Van der Schyf, C. J.; Malan, S. F. *European Journal of Pharmacology* **2003,** *458*, 73-79.
333. Onajole, O. K.; Coovadia, Y.; Govender, T.; Kruger, H. G.; Maguire, G. E. M.; Naidu, D.; Singh, N.; Govender, P. *Chem. Biol. Drug Des.* **2011,** *77*, 295-299.
334. Van der Schyf, C. J.; Dekker, T. G.; Snyckers, F. O. *Archiv der Pharmazie* **1986,** *319*, 409-415.

CHAPTER 2

NOVEL PCU CAGE DIOL PEPTIDES AS POTENTIAL TARGETS AGAINST WILD TYPE CSA HIV-1 PROTEASE; SYNTHESIS, BIOLOGICAL SCREENING AND MOLECULAR MODELING STUDIES.

Rajshekhar Karpoormath,[a,c] Yasien Sayed,[b] Thavendran Govender,[c] Hendrik G. Kruger,[a] Glenn E. M. Maguire,[a*] and Mahmoud E. S. Soliman[c,d*]

[a]School of Chemistry, University of KwaZulu Natal, Durban 4001, South Africa
[b]Protein Structure-Function Research Unit, School of Molecular and Cell Biology, University of the Witwatersrand, Wits 2050, South Africa
[c]School of Pharmacy and Pharmacology, University of KwaZulu Natal, Durban 4001, South Africa
[d]Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Zagazig University, Zagazig 44519, Egypt
*** Corresponding authors: email: kruger@ukzn.ac.za, Fax: +27-31-2603091**

ABSTRACT

*We have synthesized a series of novel pentacycloundecane cage diol diacid (PCU diol diacid) incorporated C_2-symmetric peptides. Their activity against the resistance-prone wild type C-South African (C-SA) HIV-protease is reported. These compounds were obtained in moderate yields of 42-72%. Amongst the nine compounds reported herein only compound **6**, **7**, **10** and **11** showed moderate IC_{50} values of 5 to 10 µM. Peptides **7** and **10** contain two phenylglycine and two tryptophan (Figure 3) amino acids respectively as side arms to the cage diol. Phenylglycine is a non-natural amino acid. Docking and molecular dynamics (MD) studies were carried out in order to understand the binding mode of the PCU moiety at the active site of the HIV protease enzyme. The computational results show that the cage diol peptide fits quite comfortably inside the active site of the enzyme. Not much movement is observed during the MD simulation and the hydrogen bondings that develop between the inhibitor and the enzyme pocket suggests that the inhibitor/HIV-PR complex is stable at room temperature.*

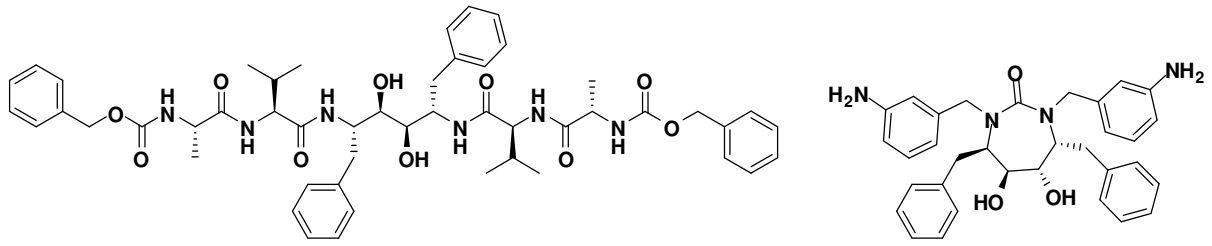
***Keywords*:** Transition state analogs, PCU cage diol diacid peptides, HIV-1 wild type C-SA protease, PCU derived peptides, HIV protease inhibitor, inhibitory concentration (IC_{50}), Docking.

1. INTRODUCTION

Acquired Immune Deficiency Syndrome (AIDS) pandemic has become a global health problem. Causative organism responsible for this disease in humans is the human immunodeficiency virus type 1 (HIV-1).[1-5] According to the World Health Organization (WHO) and Joint United Nations Program report (UNAIDS, 2010), it is estimated that 33.3 million men, women and children are infected with HIV and this number is steadily increasing in some parts of the world compared to the previous years.[6] The HIV-1 protease belongs to an aspartyl protease family of enzymes.[7] It is a C_2-symmetric dimer which is made up of two monomers. Each monomer consists of 99 amino acid residues, which are non-covalently associated to form a C_2-symmetric active homo-dimer.[8-11] The viral polyproteins p55 *gag* and p160 *gag- pol* of the HIV protease enzyme (HIV PR) play an essential role in the assembly and maturation of the virus.[12] Inhibition of HIV protease enzyme leads to the formation of non-infectious virions; hence this enzyme has become a major target for intervention in HIV infection.[13-16] Presently there are nine clinically approved HIV PR inhibitors, saquinavir,[17] ritonavir,[18] indinavir,[19] nelfinavir,[20] ampenavir,[21] lopinavir,[22] atazanavir,[23] tipranavir,[24] and darunavir.[25] These inhibitors have the advantage of being highly active against HIV PR but nonetheless exhibit major side effects such as lipodystrophy, hyperlipidaemia, insulin resistance and toxicity.[26-32] Long term use of these drugs has led to the occurrence of HIV PR inhibitor resistant mutants.[33] High dosage and high treatment cost of clinically available HIV PR inhibitors has become another major drawback to the wide spread implementation of AIDS therapies, especially in developing countries. Thus it has become very important to develop drugs, which are effective against resistant viruses and at the same time are available at low cost.[34-36]

One of the strategies explored for inhibiting the HIV protease has been to synthesize *C2*-symmetric, diol-containing inhibitors in the form of peptidomimetics, which mimics the transition state of amide hydrolysis. This concept has been proved to be efficient for the design

of potent compounds with desired pharmacological properties. One such family of active compounds being *C*2-symmetric 1,2-diol flanked by short peptide units for example Wong's inhibitor (**1**) developed in 1998[37] and DMP 450 (Mozinavir, **2**) in 1996[38,39] have shown potent anti-HIV PR activity (Figure 1).



1
Wong's Inhibitor (K_i=1.5 nM)

2
Mozinavir (DMP450, K_i=0.3 nM)

Figure 1. Examples of two known diol containing HIV PR inhibitors

Mozenavir (DMP 450) was a potent C_2-symmetric cyclic urea designed to inhibit HIV protease enzyme and its human trial ended in 2000 due to safetly concerns.[38] DMP450 has pseudo-diequatorial hydroxyl groups which are 2.8 angstroms apart and forms a hydrogen bond with the Asp 25/25' residues. The rigid cyclic urea scaffold allows for optimal interactions of the P1/P1' and P2/P2' residues with their corresponding pockets in the active site.[40]

Our group has previously demonstrated the potential of a cage lactam attached to a segment of the natural HIV-PR substrate to inhibit HIV-PR.[41] We have also reported the HIV-PR activity of a family of cage ether peptides.[42] Both these cage peptides had one peptide side arm and the best cage lactam (PCU-lactam-EAIS, IC_{50} = 0.078 μM) exhibited better HIV-PR activity than the best cage ether (PCU-ether-VA, IC_{50} = 0.6 μM) as shown below in Figure 2.

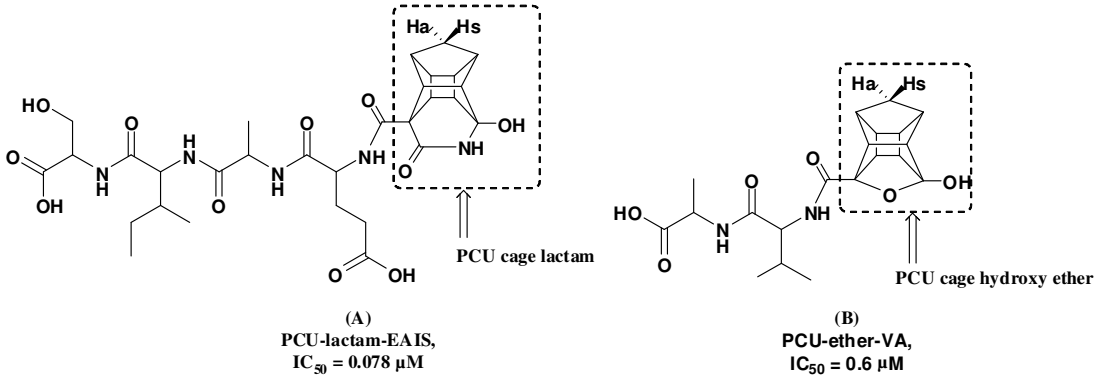


Figure 2. Structures of PCU cage lactam peptide[41] (**A**) and PCU cage hydroxy ether peptide[42] (**B**)

The rationale behind using a PCU cage diol peptide is that the diols could perhaps bind tightly to the catalytic ASP25/ASP25' in the active site of HIV PR enzyme and serve as a diol transition state mimic, since it has structural resemblance to a dihydroxyethylene type isostere[9,43] (Figure 3, structures **3** and **4**). Another potential important aspect of PCU cage diol is the distance between the diols (2.3 Å) which is quite close to the calculated diol distance (2.3 Å) of the transition state (R1-NH-C$(OH)_2$-CH-NH-R2) resulting from the hydrolysis of the natural HIV-PR protein substrate. The length of the link between the PCU cage moiety and the peptide side arms can be easily changed.[44-46]

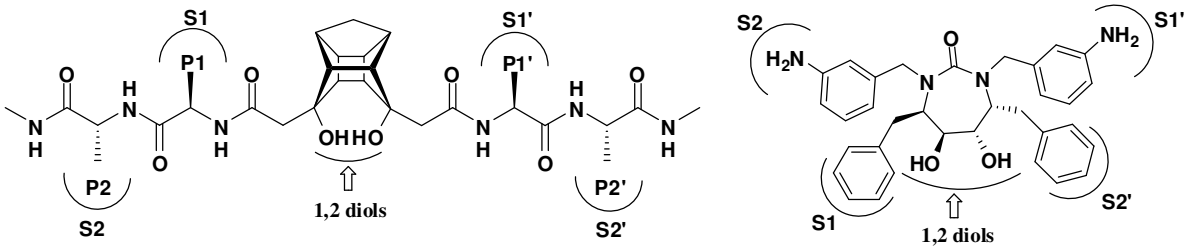


Figure 3. Structures depicting resemblance to a dihydroxy type isosteres (1,2 diols) with the potential mode of interaction in the active site of HIV PR enzyme.

This closeness of PCU hydroxyl groups [45-47] could play an important role in mimicking the transition state of HIV protease natural substrate.

Incorporation of cage compounds to medicinal drugs introduces a variety of advantages in terms of drug action and pharmacology.[48] The lipophilic nature of cage compounds makes them good candidates for drug delivery. The bulky hydrocarbon framework serves as a transport aid in carrying drugs attached to them across the cellular membranes including the blood brain barrier (BBB) and the central nervous system (CNS).[48-56] Cage compounds covalently attached to pharmaceutically active molecules also reduce drug bio-degradation which helps to increase the half life of such compounds in the biological system, thus reducing the dosage frequency.[57,52] This has led to an evolution in drug discovery with the emergence of different classes of polycyclic "cage" based skeletons being synthesized and tested against a series of diseases.[58,59]

Keeping in mind the concept of hydroxyl groups and the rigidity of Mozenavir (DMP 450) we have synthesized a series of pentacycloundecane (PCU) *endo-, endo*-diol-*exo-, exo*-amino acids as potential anti-HIV protease inhibitors (Figure 4).

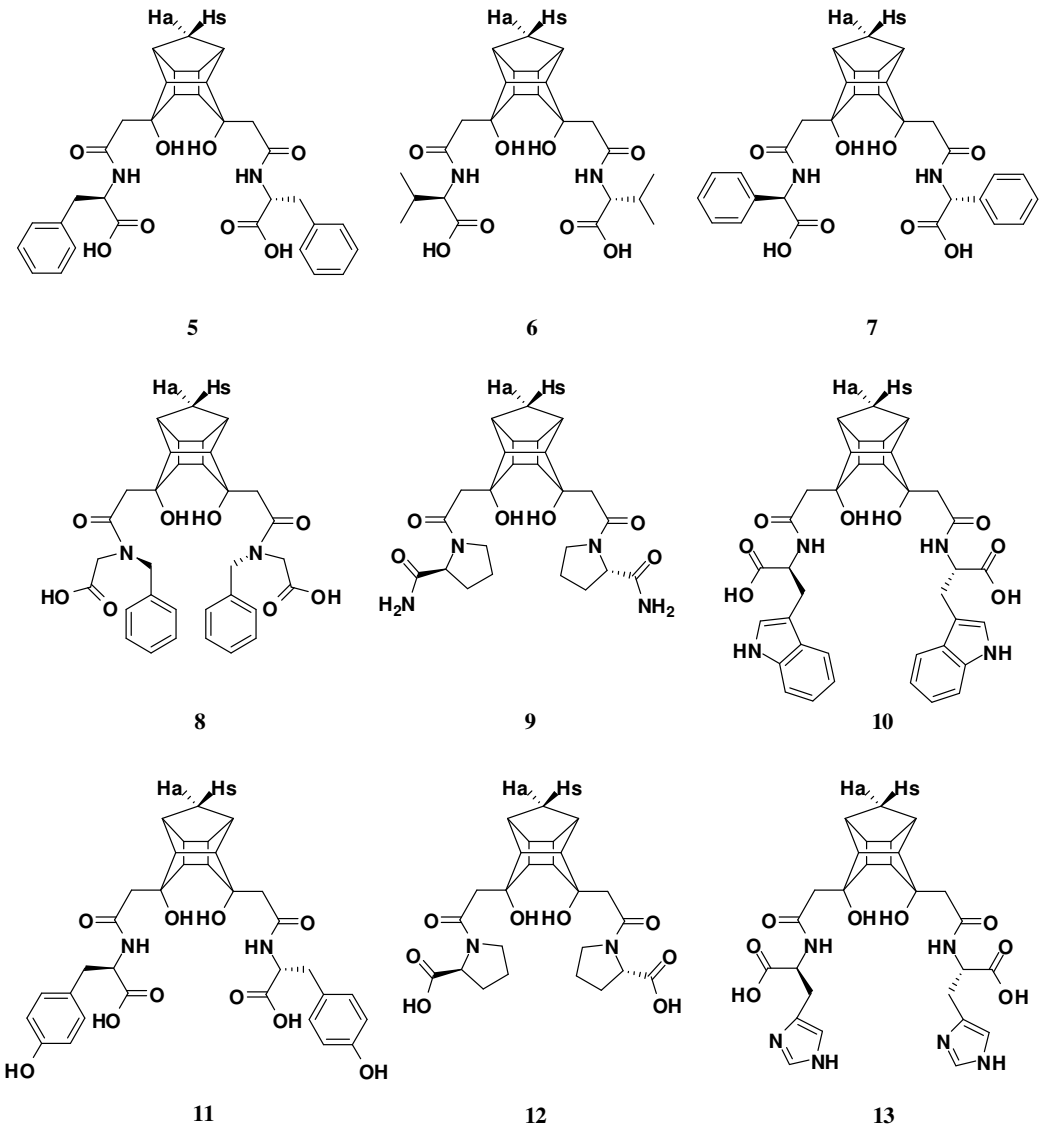


Figure 4. PCU-peptide and PCU-peptoid based inhibitors reported in this work.

The amino acids chosen for attachment to the cage diol acid were phenylalanine (F), valine (V), tryptophan (W), tyrosine (Y), proline (P), and histidine (H) for peptides **5**, **6** and **10-13** respectively. The non-natural amino acids phenylglycine and prolinamide (Figure 4 compounds **7** and **9**) were also incorporated into cage diol acid. The peptoid analogue of **5**, the cage diol di-*N*-benzylglycine derivative **8** was also synthesized.

2. MATERIAL AND METHODS

2.1. EXPERIMENTAL

High Resolution Electron Spray Ionization Mass Spectroscopic analysis was performed for all compounds on a Bruker MicroTOF QII mass spectrometer in positive mode with an internal calibration with the exception for compounds **10-14,** where data was collated in negative mode with an internal calibration. Microwave assisted synthesis was performed using CEM Discover SP microwave oven. Melting point analysis (uncorrected) was performed on a Stuart Scientific digital melting point apparatus SMP3. IR spectra were obtained from a Perkin Elmer Spectrum 100 instrument with an attenuated total reflectance attachment. Tetrahydrofuran (THF) was freshly distilled using sodium wire/benzophenone under nitrogen (N_2) atmosphere. Dichloromethane (DCM) was dried using phosphorus pentoxide as drying agent. The NMR data were recorded on Bruker AVANCE III 400 and 600 MHz instruments; the chemical shifts were referenced to the solvent peak, namely δ = 7.24 ppm for $CDCl_3$, δ = 2.50 ppm for $(CD_3)_2SO$, and δ = 3.34 ppm for CD_3OD at ambient temperature. Analytical analysis was performed on an Agilent 1100 HPLC (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 8A instrument (Ace C18 150 mm x 21.2 mm x 5 microns) with a UV/VIS detector (215 nm) and automated fraction collector. A two-mobile phase system was employed, utilizing formic acid as the ion-pairing agent. Mobile phase A consisted of 0.1% formic acid/H_2O (v/v) and mobile phase B consisted of 0.1% formic acid/acetonitrile (v/v). Optical rotations were measured at room temperature in dry methanol using a Perkin Elmer Polarimeter-Model 341. All the amino acids and coupling reagents are commercially available and were purchased from GL Biochem (Shanghai) Ltd. Analytical grade solvents and reagents like allyl magnesium bromide, *tert*-butyl acetate, DIPEA (*N,N*-Diisopropylethylamine), prolinamide for synthesis were procured from Sigma-Aldrich (*South Africa).*

Synthesis of the PCU diol diacid **(14)**

A solution of diene **16**[60] (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 ºC). Ozone was bubbled into

the reaction mixture until a blue-purple colour persisted indicating the presence of excess of ozone in the system and hence the completion of the reaction. Excess ozone gas was purged with nitrogen and the solvent (CH_3OH) 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 cooled stirred reaction mixture. The reaction was left to attain ambient temperature for 1 h and then heated to reflux gently for 12 h. The resulting mixture was concentrated in *vacuo* to yield crude product. The crude product was purified with column chromatography on silica gel using CH_3COOH: MeOH: EtOAc: hexane (1:2:20:77, R_f = 0.3, 92% yield) to obtain compound **14** as viscous pale yellow oil, which solidified on standing at room temperature to give white solid. Melting point: 155 – 157 ºC, IR ν_{max}: 3102, 2985, 2864, 1698, 1639, 1451, 1292, 1280, 1181, 1165, 1120, 1068, 1045, 923, 862, 799, 736, 652, 575 and 477 cm^{-1}. ^1H NMR [$(CD_3)_2SO$, 400 MHz]: δ_H 1.02 (AB, J_{AB} = 10.4 H_Z, 1H), 1.49 (AB, J_{AB} = 10.4 H_Z, 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), 7.16 (s, 2H), 11.92 (s, 2H). ^{13}C NMR [$(CD_3)_2SO$, 100 MHz]: δ_C 33.4, 38.8, 42.2, 43.8, 44.0, 49.9, 76.4, 172.3. HR ESI *m/z*: calculated for $C_{15}H_{18}O_6$ $[M-H]^-$: 293.1031 found 293.1030.

General procedure for the synthesis of PCU cage diol diamino acids (5-8 and 10-12)

A solution of compound **14** (1 eq, 0.5 g) in 15 mL of dry DCM was stirred at room temperature for 5 min. To this mixture was added *tert*-butyl amino acid esters/amino acid methyl esters (4 eq, 1.5 g) while the reaction mixture was cooled in an ice water bath and stirred for 5 min. To the above cooled mixture was added HATU (5 eq, 3.24 g) followed by DIPEA (8 eq, 1.76 g, 2.4 mL) as a base. The solution was then slowly brought to room temperature and stirred for 6 h. The crude reaction mixtures was washed with 100 mL water and then with 100 mL of 10% HCl. The organic layer was dried over anhydrous Na_2SO_4 and filtered. The crude product was evaporated to dryness under vacuum at 40 °C to obtain thick yellow oil. The crude *tert*-butyl esters of PCU cage diol diamino acids (**17** and **18**) were further dissolved in 1:1 (v/v) DCM and TFA solvent mixture and stirred overnight. TFA was removed from the mixture by bubbling air through the peptide and the remaining DCM was removed under vacuum at 30 °C, to obtain crude PCU cage diol diamino acids (**5** and **6**) as pale yellow oils.

The methyl esters of PCU cage diol diamino acids (**19-23**) were dissolved in methanol and excess of 1 N KOH and stirred overnight. It is known from literature[10,61,62] that peptides are not racemized under such basic conditions (1 N KOH) even upon treatment over extended times. Glacial acetic acid was added until the mixture became neutral and then diluted with chloroform. The organic solution is extracted once with water and brine. The organic layer is dried over anhydrous $MgSO_4$ and concentrated *in vacuo* to give crude oily compounds **7, 8** and **10**-**12**. The crude compounds were then purified by preparative HPLC and solid phase extraction to give pure PCU cage diol diamino acids (**5**-**8** and **10**-**12**).

General procedure for the synthesis of *tert*-butyl amino acid esters

Compounds **33** and **34** were synthesized according to the reported method.[63]

General procedure for the synthesis of amino acid methyl esters

Compounds **35**-**39** were synthesized according to the reported method.[64,65]

***PCU cage diol diphenylalanine* (5)**

A white solid (76%). Retention time: 15.23 min (HPLC), Melting point: 231 – 232 ºC, $[\alpha]^{20}_D$:+14.29 (*c* 0.14 in MeOD). IR ν_{max}: 3261, 2957, 1758, 1638, 1522, 1191, 758, 701 and 489 cm^{-1}.^{1}H NMR [CD_3OD, 400 MHz]: δ_H1.04 (1H, d, *J* = 10.6 Hz), 1.39 (1H, d, *J* = 10.5 Hz), 1.67 (1H, d, *J* = 8.8 Hz), 1.90 (1H, d, *J* = 8.8 Hz), 2.19-2.34 (7H, m), 2.45-2.47 (2H, m), 2.91-3.01 (2H, m), 3.24-3.31 (2H, m), 4.69-4.73 (1H, q, *J* = 4.7 Hz), 4.77-4.81 (1H, q, *J* = 4.6 Hz), 7.13-7.17 (1H, m), 7.22-7.32 (9H, m). ^{13}C NMR [CD_3OD, 100 MHz]: δ_C 33.9, 37.5, 37.6, 39.4, 39.7, 43.2, 43.8, 43.8, 44.5, 44.8, 49.1, 49.4, 53.8, 53.9, 77.2, 77.8, 127.1, 127.3, 128.8, 128.9, 129.5, 137.2, 137.5, 173.1, 173.9. HR ESI *m/z*: calculated for $C_{33}H_{36}N_2O_8$ $[M+H]^+$: 589.2544 found 589.2541.

***PCU cage diol divaline* (6)**

A colourless oil (72%). Retention time: 14.61 min (HPLC), $[\alpha]^{20}_D$:+8.69 (*c* 0.23 in MeOD). IR ν_{max}: 2960, 2929, 1721, 1621, 1537, 1464, 1392, 1259, 1209, 1064, 986 and 544 cm^{-1}. ^{1}H NMR [CD_3OD, 400 MHz]: δ_H1.18 (1H, d, *J* = 10.8 Hz), 1.60 (1H, d, *J* = 10.7 Hz), 0.94-1.00 (12H,

m), 2.19 (2H, q, *J* = 4.6 Hz), 2.20-2.62 (12H, m), 4.36 (2H, t, 4.8). ^{13}C NMR [CD_3OD, 100 MHz]: δ_C 34.8, 40.6, 40.7, 44.3, 45.2, 45.3, 45.7, 50.7, 50.9, 78.5, 78.6, 174.3, 174.6. HR ESI *m/z*: calculated for $C_{25}H_{36}N_2O_8$ $[M+H]^+$: 493.2531 found 493.2544.

PCU cage diol di-2-phenylglycine (**7**)

A colourless oil (58%). Retention time: 15.10 min (HPLC), $[\alpha]^{20}_D$:-30.00 (*c* 0.10 in MeOD). IR ν_{max}: 3060, 2994, 2863, 1723, 1629, 1524, 1452, 1256, 1179, 1133, 1063, 722, 696, 596 and 513 cm^{-1}. ^{1}H NMR [CD_3OD, 400 MHz]: δ_H1.10 (1H, d, *J* = 10.8 Hz), 1.50 (1H, d, *J* = 10.7 Hz), 2.15-2.58 (12H, m) 5.43 (2H, s), 7.30-7.43 (10H, m). ^{13}C NMR [CD_3OD, 100 MHz]: δ_C 34.8, 40.6, 44.4, 45.3, 45.7, 50.8, 58.2, 78.7, 128.7, 129.4, 129.8, 173.8. HR ESI *m/z*: calculated for $C_{31}H_{32}N_2O_8$ $[M+H]^+$: 561.2213 found 561.2231.

Synthesis of PCU cage diol di -N-benzylglycine **(8)**

A pale yellow oil (42%). Retention time: 16.06 min (HPLC), $[\alpha]^{20}_D$:+7.14 (*c* 0.14 in MeOD). IR ν_{max}: 3054, 2981, 1667, 1615, 1475, 1450, 1293, 1182, 1130, 1033, 957, 837, 799, 721 and 598 cm^{-1}. ^{1}H NMR [CD_3OD, 600 MHz]: δ_H1.07 (1H, d, *J* = 10.6 Hz), 1.56 (1H, d, *J* = 10.6 Hz), 2.26-2.69 (12H, m), 4.05-4.21 (4H, m), 4.63-4.75 (4H, m), 7.28-7.37 (10H, m). ^{13}C NMR [CD_3OD, 100 MHz]: δ_C 34.8, 40.6, 44.3, 45.7, 45.8, 48.0, 50.9, 52.6, 58.2, 79.4, 85.2, 101.50, 117.7, 128.8, 129.7, 130.0, 138.1, 169.8. HR ESI *m/z*: calculated for $C_{33}H_{36}N_2O_8$ $[M+H]^+$: 589.2535 found 589.2544.

Synthesis of PCU cage diol diprolinamide (**9**)

A solution of compound **14** (1 eq, 0.5 g) in 15 mL of dry DCM was stirred at room temperature for 5 min. To this mixture was added prolinamide (4 eq, 0.8 g). The solution was cooled in an ice water bath and stirred for 5 min. To the above cooled mixture was added HATU (5 eq, 3.24 g) followed by DIPEA (8 eq, 1.76 g, 2.4 mL) as a base. The solution was then slowly brought to room temperature and stirred for 6 h. The crude reaction mixtures was diluted with DCM and filtered. The crude product was evaporated to dryness under vacuum using a teflon pump at 40 °C to obtain thick yellow oil which was purified by preparative HPLC and solid phase extraction to give pure PCU cage diol diprolinamide **9** as pale yellow solid (65%). Melting point: 83 – 85

ºC. Retention time: 13.42 min (HPLC), $[\alpha]^{20}_{D}$:-37.50 (*c* 0.16 in MeOD). IR ν_{max}: 3179, 2957, 2872, 1669, 1603, 1448, 1417, 1288, 1191, 1168, 1071, 987, 908, 653, 539 and 421 cm^{-1}. ^{1}H NMR [$(CD_3)_2SO$, 400 MHz]: δ_H1.04 (1H, t, *J* = 10.6 Hz), 1.48 (1H, t, *J* = 10.5 Hz), 1.74-2.72 (14H, m), 3.50-3.68 (2H, m), 4.16-4.22 (2H, m), 6.92-6.95 (2H, m), 7.08-7.22 (2H, m). ^{13}C NMR [$(CD_3)_2SO$, 100 MHz]: δ_C 24.03, 29.3, 33.5, 38.5, 41.3, 42.0, 42.2, 43.9, 47.3, 50.3, 59.3, 77.8, 170.6, 173.8. HR ESI *m/z*: calculated for $C_{25}H_{34}N_4O_6$ $[M+H]^+$: 487.2542 found 487.2551.

Synthesis of PCU cage diol ditryptophan **(10)**

A pale white solid (62%). Retention time: 15.58 min (HPLC), Melting point: 168 – 170 ºC, $[\alpha]^{20}_{D}$:-30.77 (*c* 0.13 in MeOD). IR ν_{max}: 3307, 2924, 2855 1720, 1624, 1532, 1457, 1429, 1343,1218, 1190, 1099, 1064, 1010, 910, 821, 740, 536 and 424 cm^{-1}. ^{1}H NMR [DMSO, 600 MHz]: δ_H0.96 (1H, d, *J* = 10.3 Hz), 1.23 (1H,s), 1.32 (1H, d, *J* = 10.1 Hz), 1.88 (1H, d, *J* = 8.9 Hz), 1.95 (1H, d, *J* = 8.8 Hz), 2.05-2.12 (2H, m), 2.19-2.22 (3H, m), 2.28-2.29 (3H, m), 2.37-2.43 (2H, m), 3.03-3.08 (2H,m), 3.16-3.28 (2H,m), 4.52-4.60 (2H, m), 6.94-7.06 (4H, m,), 7.12 (2H, q, *J* = 2.1 Hz), 7.32 (2H, t, *J* = 7.5 Hz), 7.42 (2H, s), 7.52 (2H, q, *J* = 4.1 Hz), 8.08 (1H, d, *J* = 7.6 Hz), 8.19 (1H, d, *J* = 7.6 Hz), 10.81 (1H, d, *J* = 1.3 Hz), 10.86 (1H, d, *J* = 1.8 Hz), 12.69 (1H,s). ^{13}C NMR [CD_3OD, 100 MHz]: δ_C 27.1, 33.3, 38.7, 42.3, 43.6, 43.9, 44.1, 48.7, 48.8, 52.5, 64.8, 76.6, 76.7, 109.3, 111.3, 118.1, 118.3, 120.8, 123.5,123.7, 127.1, 127.2, 136.0, 171.1, 171.2, 173.1. HR ESI *m/z*: calculated for $C_{37}H_{38}N_4O_8$ $[M-H]^-$: 665.2617 found 665.2497.

Synthesis of PCU cage diol ditryrosine **(11)**

A white solid (32%). Retention time: 14.23 min (HPLC), Melting point: 199 – 201 ºC, $[\alpha]^{20}_{D}$:+23.08 (*c* 0.13 in MeOD). IR ν_{max}: 3232, 2954, 2865 1721, 1614, 1536, 1514, 1441, 1364, 1221, 1105, 1064, 986, 910, 825, 530 and 428 cm^{-1}. ^{1}H NMR [CD_3OD, 600 MHz]: δ_H1.05 (1H, d, *J* = 10.6 Hz), 1.47 (1H, d, *J* = 10.5 Hz), 1.85 (1H, d, *J* = 8.4 Hz), 2.02 (1H, d, *J* = 8.7 Hz), 2.10-2.35 (8H, m), 2.49 (2H, s), 2.86-2.92 (2H, m), 3.14-3.18 (2H, m), 3.71 (1H, s) 4.65-4.67 (1H, q, *J* = 4.5 Hz), 4.72-4.74 (1H, q, , *J* = 4.5 Hz), 6.69-6.73 (4H, m), 7.01-7.06 (4H, m). ^{13}C NMR [CD_3OD, 100 MHz]: δ_C 34.8, 37.8, 40.5, 40.7, 44.2, 44.5, 45.0, 45.6, 45.7, 45.8, 50.3, 50.5, 52.8, 55.0, 78.3, 78.7, 116.4, 116.5, 128.9, 131.5, 131.6, 157.4, 157.5, 174.0, 174.2. HR ESI *m/z*: calculated for $C_{33}H_{36}N_2O_{10}$ $[M-H]^-$: 619.2297 found 619.2045.

***Synthesis of PCU cage diol diproline* (12)**

A white solid (35%). Retention time: 14.34 min (HPLC), Melting point: 170 – 172 ºC, $[\alpha]^{20}_{D}$:-63.64 (*c* 0.11 in MeOD). IR ν_{max}: 3164, 2957, 2875 1720, 1598, 1448, 1423, 1288, 1223, 1188, 1072, 978, 908, 873, 654, 544, and 429 cm^{-1}. ^{1}H NMR [CD_3OD, 600 MHz]: δ_H1.16 (1H, d, *J* = 10.7 Hz), 1.62 (1H, d, *J* = 11.2 Hz), 1.98-2.01 (5H, m), 2.18-2.31 (5H, m), 2.36-2.40 (1H, m), 2.43-2.62 (8H, m), 2.75-2.83 (1H, m), 3.48-3.56 (1H, m), 3.63-3.73 (3H, m), 4.41-4.44 (1H, m), 4.71-4.77 (1H, m). ^{13}C NMR [CD_3OD, 100 MHz]: δ_C 23.6, 23.7, 30.5, 30.6, 32.1, 34.9, 35.0, 40.6, 40.7, 40.8, 40.9, 41.0, 43.1, 43.2, 43.5, 43.7, 44.7, 44.9, 45.8, 45.9, 46.0, 47.5, 51.1, 51.4, 51.6, 52.9, 53.0, 53.4, 60.3, 61.5, 79.4, 79.5, 79.6, 79.7, 173.4, 175.8, 175.9. HR ESI *m/z*: calculated for $C_{25}H_{32}N_2O_8$ $[M-H]^-$: 487.2086 found 487.2025.

***Synthesis of PCU cage diol dihistidine* (13)**

A solution of compound **14** (1 eq, 0.5 g) in 2 mL in dry DMF was stirred in a microwave reactor vessel. To this above solution was added H-His(trt)-OMe.HCL (2.5 eq, 0.8 g) followed by HATU (2.5 eq, 3.24 g) and DIPEA (8 eq, 1.76 g, 2.4 mL) as a base. The solution mixture was then subjected to microwave irradiation at 200 Watt for 20 min at 55 ºC and then 300 Watt for 20 min at 60 ºC.[66] The resulting solution was diluted excess of water (50 mL) and extracted twice with ethyl acetate (2 x 50 mL). The organic layer was then washed with brine (2 x 50 mL), dried over anhydrous Na_2SO_4 and filtered. The crude product **24** was concentrated *in vacuo* to obtain thick colourless oil. This crude product **24** was dissolved in a solution of TFA and water (3:1) and stirred at 60 ºC for 8h. Excess of TFA and water from the crude reaction mixture was removed *in vacuo* to give pale yellow oil, which was purified by preparative HPLC to give pure PCU cage diol dihistidine **13** as colourless oil (58%). Retention time: 09.80 min (HPLC), $[\alpha]^{20}_{D}$: = + 6.25 (*c* 0.12 in MeOD). IR ν_{max}: 3132, 2956, 2862, 1965, 1624, 1577, 1434, 1389, 1289, 1258, 1187, 1133, 1066, 984, 831, 797, 627, 542, and 435 cm^{-1}. ^{1}H NMR [$(CD_3)_2SO$, 600 MHz]: δ_H0.97 (1H, d, *J* = 10.0 Hz), 1.38 (1H, d, *J* = 10.1 Hz), 1.90 (1H, d, *J* = 08.1 Hz), 2.02-2.04 (2H, m) 2.09-2.11 (1H, m), 2.21-2.26, (2H, m), 2.30-2.37 (4H, m), 2.44 (2H, s), 2.85-2.93 (2H, m), 3.01-3.06 (2H, m), 4.47 (2H, s), 6.97 (2H, d, *J* = 36.9 Hz), 7.92 (2H, d, *J* = 39.3 Hz), 8.03 (1H, d, *J* = 07.3 Hz), 8.15 (2H, s). ^{13}C NMR [$(CD_3)_2SO$, 100 MHz]: δ_C 28.0, 28.3, 33.4, 38.8, 42.5, 42.9, 43.8, 44.6, 44.9, 48.5, 48.6, 48.8, 51.8, 52.4, 64.9, 76.5, 76.8, 92.9, 117.0, 117.1,

131.6, 131.9, 132.3, 134.2, 134.4, 157.9, 163.2, 170.4, 170.9, 172.8, 173.4. HR ESI *m/z*: calculated for $C_{27}H_{32}N_6O_8$ $[M-H]^-$: 567.2209 found 567.2180.

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-NH_2.[67-69] The substrate (H-1048) can be obtained from Bachem. 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 a 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 (see details provided in Experimental section 2.1) on a Shimadzu instrument and characterized using the Bruker microTOF-Q II instrument (see Table 1 in supplementary information).

To determine the concentration of the compounds that resulted in 50% inhibition (IC_{50}) 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 an Analytik Jena Specord 210 spectrophotometer. Activity was standardized using commercially available drugs Atazanavir and Lopinavir. All cage peptides were soluble in the aqueous buffer solution. (pH = 5, 50 mM sodium acetate and 0.1 M NaCl)

2.2. COMPUTATIONAL SIMULATIONS

Preparation of PCU derived peptide inhibitors

The structures of the cage-peptide/peptoid, **5-13**, were constructed and the geometry was optimized using the MMFF94 force field implemented in the Avogadro software (Avogadro 1.0.0, 2010).[70] The minimized structures were then subjected to docking studies.

C-SA PR enzyme model system

The X-ray structure of the South African HIV-1 protease subtype C (C-SA) has not yet been reported. The 3D structure of the enzyme was generated as described before.[41]

Docking of the inhibitors into the C-SA PR model

The energy minimized inhibitors, **5-13** were used in docking simulations. The effect of ionization states of docked compounds to the binding scores has been discussed in the literature.[71-73] We have previously reported that no significant difference was observed for docking of the cage lactam peptides with the different protonation states of the enzyme.[41] Docking of ritonavir (co-crystallized inhibitor) with subtype B HIV-PR (PDB accession code 1HXW)[18] was previously performed to evaluate the efficacy of Autodock for its use in docking experiments with this subtype as a target. Acceptable results were obtained when ritonavir was docked with the constructed C-SA HIV PR.[74] The interacting amino acids in the C-SA starting structure were conserved in the conformation of the enzyme after docking and the variation in orientation of these amino acids in the binding cavity was unnoticeable.

Docking studies were performed using the Autodock software.[75] Geisteger charges were computed and the Autodock atom types were defined using the Autodock Tools graphical user interface supplied by MGL Tools.[76] The Lamarckian Genetic Algorithm (LGA), which is considered one of the best docking methods available in Autodock[75,77] was employed. This algorithm yields superior docking performance compared to simulated annealing or the simple genetic algorithm and the other search algorithms available in Autodock. The docked conformations of each of the inhibitors were ranked into clusters based on the binding energy. Top ranked conformations were visually analyzed.

MD simulations of inhibitor-enzyme complex:

MD simulations were performed for the inhibitor- enzyme complex of the most active compound **7**. The diastereomer with the best docked binding energy was used. Partial charges and the force field parameters for the inhibitor were generated using the Antechamber program[78] in the

Amber10 package.[79] These were described by the general Amber force field (GAFF).[80,81] All hydrogen atoms of the proteins were added using the Leap module in Amber10. The standard AMBER force field for bioorganic systems (ff03)[82] was used to describe the HIV PR enzyme parameters. Counter ions were added to neutralize the complex. Then, the system was solvated using atomistic TIP3P water[83] in a cubic box with 8.0 Å distance around the complex.

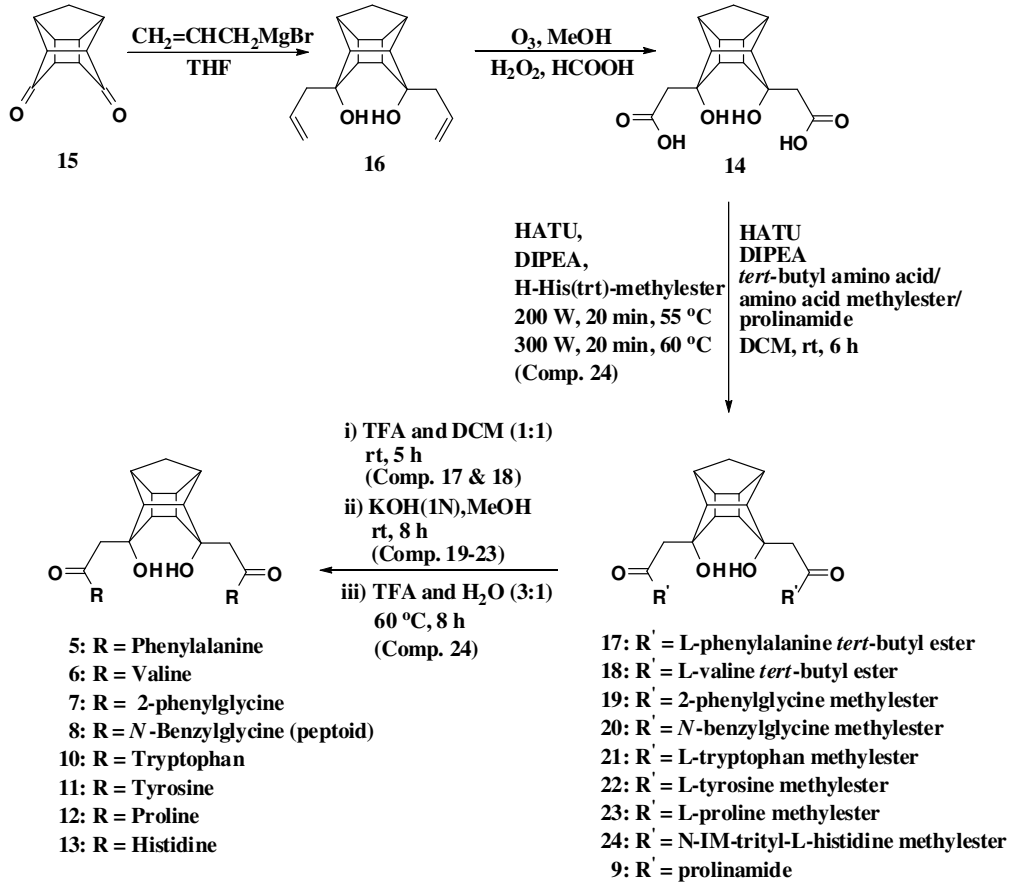
The molecular dynamics package Amber10[79] was used for the minimization and equilibration protocols. Cubic periodic boundary conditions were imposed and the long-range electrostatic interactions were treated with the particle-mesh Ewald method[84] implemented in Amber10 with a non-bonding cut-off distance of 10.0 Å. The energy minimization was first conducted using the steepest descent method in Amber10 for 1000 iterations switched to conjugate gradient for 2000 steps with a restraint potential of 2 kcal/mol $Å^2$ applied to the solute. Then the total system was freely minimized for 1000 iterations. For the equilibration and the subsequent production run, the SHAKE algorithm[85] was employed on all atoms covalently bonded to a hydrogen atom, allowing for an integration time step of 2 fs. Harmonic restraints with force constants 2.0 kcal/mol $Å^2$ were applied to all solute atoms. A canonical ensemble (NVT) MD was carried out for 70 ps, during which the system was gradually annealed from 0 to 300 K using a Langevin thermostat with a coupling coefficient of 1.0/ps. Subsequently, the system was equilibrated at constant volume and temperature (300 K) with a 2 fs time step for 100 ps while maintaining the force constants on the restrained solute. With no restraints imposed, a production run was performed for 2 ns in isothermal isobaric (NPT) ensemble using a Berendsen barostat[86] with a target pressure of 1 bar and a pressure coupling constant of 2 ps. The coordinate file was saved every 1 ps and the trajectory was analyzed at every 1 ps using the Ptraj module implemented in Amber10.

3. RESULTS AND DISCUSSION

3.1. CHEMISTRY AND BIOLOGICAL ACTIVITY

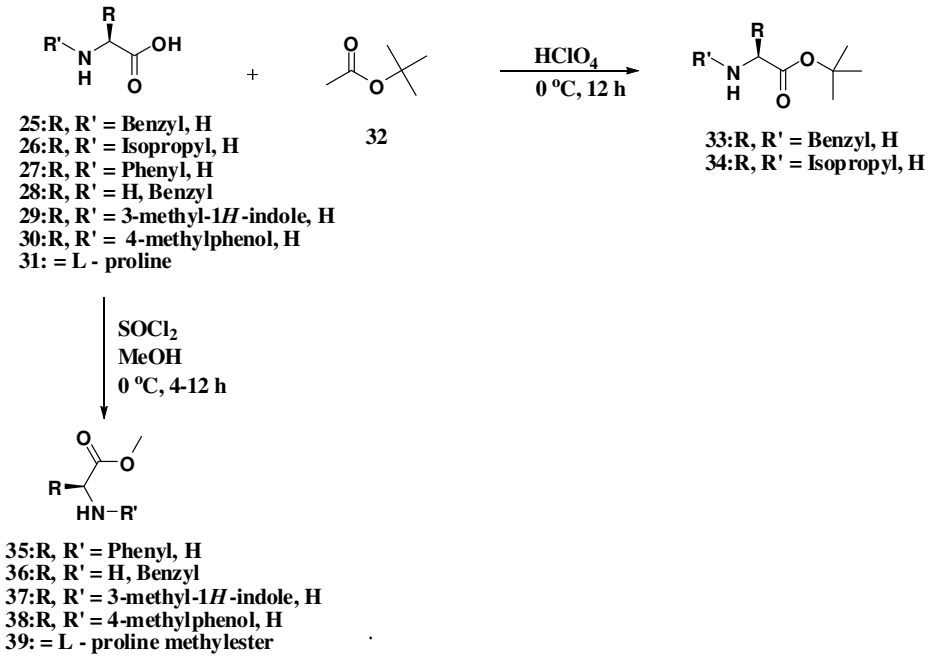
Synthesis of PCU cage peptides and peptoid was carried out as described in the experimental section. The *tert*-butyl amino acid esters and amino acid methyl esters were synthesized of sufficient purity to be coupled to the PCU cage diol dicarboxylic acid (**14**). The PCU cage diol

dicarboxylic acid **14**[87] was synthesized from Cookson's dione **15**,[88,89] which can be easily obtained from the photocyclisation of the Diels-Alder adduct[88,89] of cyclopentadiene and *p*-benzoquinone. The PCU dione **15** was reacted with freshly prepared allyl magnesium bromide (Grignard reaction) to obtain *endo*-8, *endo*-11 diol **16**. Further ozonolysis of the diol **16** followed by an oxidative workup yielded the PCU diol dicarboxylic acid **14** in good yield as described in Scheme 1.



Scheme 1. Synthesis of PCU diol dicarboxylic acid and its derivatives (peptides and peptoid).

The *tert*-butyl amino acid esters (**33** and **34)** were synthesized by reacting L-amino acids (**25** and **26**) with *tert*-butyl acetate (**32**) and $HClO_4$.[63] Treating L-amino acids with thionyl chloride and methanol gave the corresponding methyl esters (**35** to **39**).[64,65] Synthesis of *tert*-butyl and methyl esters is shown in Scheme 2.



Scheme 2. Synthesis of *tert*-butyl and methyl ester derivatives of L-amino acids.

The starting material for the synthesis of compounds **5** to **13** is the PCU diol dicarboxylic acid **14**. Compound **14** in dry DCM was reacted with *tert*-butyl amino acid esters **33** and **34** or the amino acid methyl esters **35** to **39** in the presence of the coupling reagent HATU (2-(1*H*-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyl uronium hexafluorophosphate methanaminium) and DIPEA (*N,N*-Diisopropylethylamine) as a base to give the corresponding compounds **17** to **23** and **9** except in the case of compound **24** microwave synthesizer was used for the synthesis as shown in Scheme 1.[66] Deprotection of PCU diol *tert*-butyl amino acid esters **17** and **18** was carried out using TFA (trifluoroacetic acid) in DCM (1:1) to yield the corresponding cage amino acids **5** and **6** while the deprotection of compound **19** to **23** was achieved using diluted KOH in methanol to give compounds **7** to **12** (Scheme 1). Compound **24** was treated with a solution mixture of TFA and water (3:1) and stirred at 60 °C resulting in the removal of both the protecting group's (trityl and methyl ester) to give compound **13**. Crystallization of compound **5** was carried out by dissolving the pure compound in 3 mL solvent mixture of DCM and TFA (1:1) and stored at 20 °C to obtained pale white crystals[47] (Figure 4).

Isolation and purification of compounds **5** to **13** (Figure 3) were carried out as described in the experimental section. The correctness of the cage peptides and peptoids were confirmed with

standard NMR and MS techniques. The purity of the peptides was determined by HPLC to be > 98%.

The catalytic activity of the HIV-1 protease was monitored by following the hydrolysis of the chromogenic peptide substrate His-Lys-Ala-Arg-Val-Leu-Phe(*p*-NO_2)-Glu-Ala-Nle-Ser as previously described.[67] The use of chromogenic substrates has been established for the determination of the HIV-1 protease inhibitors efficacy. The method is used to measure the protease activity by recording an increase in absorbance as a result of the hydrolysis of a chromogenic substrate by the protease.[68,69] The inhibitory activities of the synthesized compounds were tested using an UV spectrophotometric assay in order to determine the IC_{50} values for each compound. Biological results of the synthesized substrates are presented in Table 1.

Table 1. Inhibition of wild type C-SA HIV-1 protease by PCU derived peptides (**5**-**13** and **9**) and peptoid (**8**).

Compound	**Compound (R)**	**Formula**	**% Yield**	**LogP**[a]	**Binding Energy**[b] **Kcal/mol**	**IC_{50} (µM)**[c]
5	Phenylalanine	$C_{33}H_{36}N_2O_8$	76	1.25	-7.24	>60
6	Valine	$C_{25}H_{36}N_2O_8$	72	0.27	-6.78	10
7	2-Phenylglycine	$C_{31}H_{32}N_2O_8$	52	0.83	-8.39	5
8 (Peptoid)	R = Benzyl	$C_{33}H_{36}N_2O_8$	42	2.26	-7.02	>60
9	Prolinamide	$C_{25}H_{34}N_4O_6$	65	-2.37	-7.12	>60
10	Tryptophan	$C_{37}H_{38}N_4O_8$	62	0.75	-7.18	5
11	Tyrosine	$C_{33}H_{36}N_2O_{10}$	32	0.57	-6.49	10
12	Proline	$C_{25}H_{32}N_2O_8$	35	-2.19	-6.20	>60
13	Histidine	$C_{27}H_{32}N_6O_8$	58	-3.85	-6.94	>60
Lopinavir						**0.025**
Atazanavir						**0.004**

[a] Calculated with ACD/labs software v11.0.
[b] Binding energy (kcal/mol), calculated with Autodock. The lowest binding energy of the compounds are reported here (interestingly, all the lowest binding energies correspond to the conformations that showed interactions with the Asp residues). The 3D structures of these docked complexes are available as supplementary material.
[c] IC_{50} = 50% inhibition constant. Standard deviations (SD) for all the compounds were ≤ 2% of the reported IC_{50}values.

It was reported in literature that phenylalanine is the preferred substituent (P^1) at the S^1 subsite and it is found in 40% of *Gag-Pol* gene sequences of the virus.[90] Based on the literature and the HIV substrate sequence, it was envisaged that different amino acids having aromatic side chain would be the best starting point in the peptide for binding at the $S^1/S^{1'}$ subsite.[16] This motivated the commencement of the synthesis of a short cage dipeptide, with phenylalanine as the first amino acid (P^1 substituent), attached to the PCU cage diol.

It was observed that compound **5** where phenylalanine was coupled to PCU cage diol as $P^1/P^{1'}$ substituent showed very little inhibition activity (IC_{50}> 60 μM). This could perhaps be attributed to the benzyl side chain of the phenylalanine shielding the interaction of the PCU cage diols with the catalytic site (Asp 25/25') of HIV protease enzyme. Interestingly this arrangement of the side arms in the crystal structure of compound **5** in Figure 5 was observed.[47] An alternative reason is that the specific sequence does not fit closely into the active site.

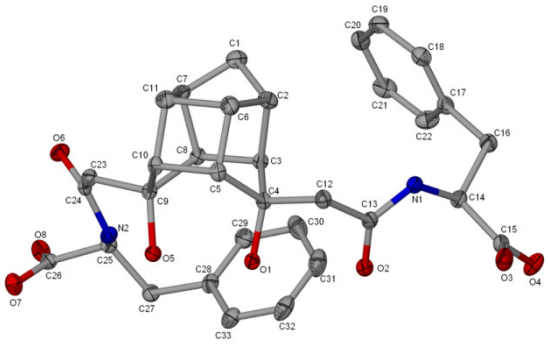


Figure 5. Crystal structure of compound **5** (CCDC 797748).[91] Hydrogen atoms were ommitted for clarity.

To validate this observation, valine was incorporated as the $P^1/P^{1'}$ substituent in compound **6** which resulted in six fold increase in inhibitory activity (IC_{50} = 10 μM). The increase in inhibition activity could be due to the shorter (isopropyl) side chain of valine, which could be altering the conformation of the two arms of the PCU cage diol in such a way that the diols are free to interact with the catalytic active site of the HIV protease enzyme. Since $S^1/S^{1'}$ subsite prefers more hydrophobic aromatic amino acids, the incorporation of 2-phenylglycine as the $P^1/P^{1'}$ substituent in compound **7**, resulted in a two fold increase of the inhibition activity (IC_{50} = 5 μM) in comparison to compound **6**.

Oligomers of *N*-substituted glycines are known to be resistant to proteolysis.[92-94] Keeping this in mind the peptoid **8**, which mimicked the dipeptide sequence of compound **5** was synthesised. Compound **8** failed to show any inhibitory activity (IC_{50} > 60 μM). No inhibition activity was observed for compound **9** when compared to compound **7**.

In order to further investigate, $P^1/P^{1'}$ substituents were replaced with amino acids having heteroaromatic and substituted aromatic groups as their side chains. It was observed that compound **10** with tryptophan as the potential $P^1/P^{1'}$ substituent showed similar inhibitory activity (IC_{50} = 5 μM) when compared to compound **7**. Inhibitory activity for compound **11** (IC_{50} = 10 μM) decreases by twofold when compared to compounds **7** and **10**. Compounds **12** and **13** failed to show any inhibitory activity (IC_{50} > 60 μM).

The LogP values for the most active peptides **6, 7**, **10** and **11** were 0.3, 0.8, 0.7 and -0.5 respectively. This is lower than the preferred range[95] of between 1.3 and 4.1 for the majority of orally available drugs.

The relative binding energies (-7.24 and -6.49 Kcal/mol) of compounds **5** and **11** could be attributed to the electrostatic interactions, hydrophobicity of the side arms and conserved hydrogen bonding even though structurally these compounds are similar. Similar explanation could be given for the relative binding energies (-7.18 and -6.94 Kcal/mol) for compounds **10** and **13**. It is also know that binding energies calculated by docking methods are not accurate and are just approximation of the orientation of the molecules within the active site of the enzyme.

From the results in Table 1 the best inhibition activity was obtained for compound **7** and **10** (IC_{50} = 5 μM), which shows potential activity when compared with Lopinavir but almost three orders of magnitude less active than Atazanavir. To further validate these experimental results, docking studies were performed. MD simulations were also performed for the most active compound.

3.2. COMPUTATIONAL SIMULATIONS

Docking and MD simulations

First, the PCU-peptides were subjected to docking studies to explore their binding pattern; second MD simulations of the docked inhibitor enzyme complex were performed to provide a

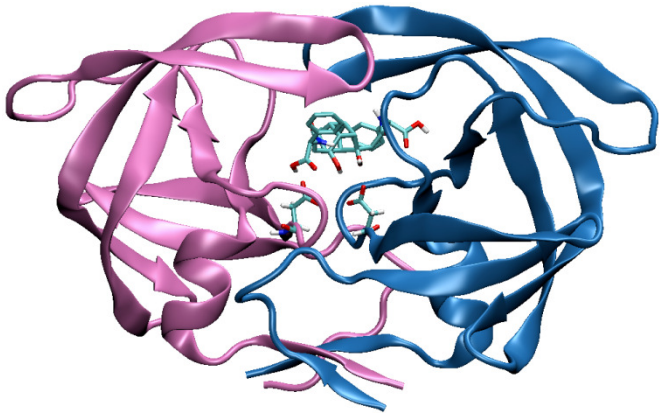
molecular view of the dynamic behaviour of such inhibitors inside the enzyme active site over a suitable time scale. Docking studies are used at different stages in drug discovery such as in the prediction of ligand-receptor complex structures and also to rank the ligand molecules based upon the binding energies of the corresponding ligand-enzyme complexes. Docking protocols aid in elucidation of the most energetically favourable binding mode of the ligand to the receptor. The objective of our docking study was to elucidate the potential interaction mode of the PCU-peptide derivatives with C-SA HIV PR and to see if a correlation between the binding energies and the observed IC_{50} results exists.

Docking routines of the new PCU-peptide inhibitors with the C-SA enzyme were performed. The binding energies of the docked PCU-peptide inhibitors have been tabulated (Table 1). The docking results are in reasonable agreement with the measured IC_{50} values.

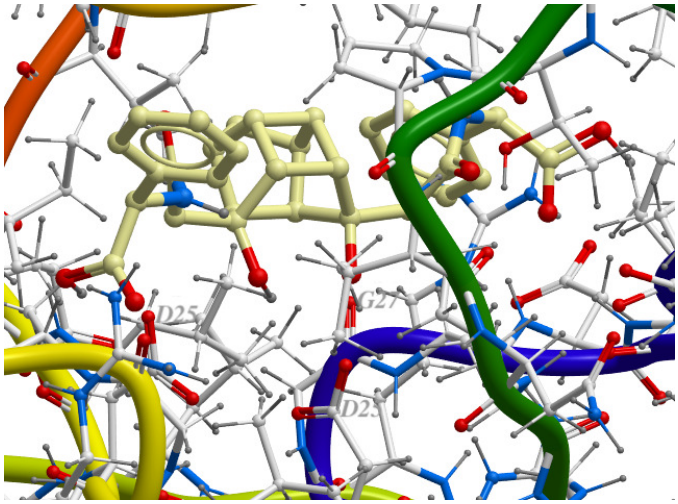
The docked binding energies previously calculated for the lactam peptides[74] and the cage ether peptides[42] correlate reasonably well to the binding energies calculated in this study for cage diol peptides with similar IC_{50} values. The most active lactam peptide (PCU-lactam-EAISa, IC_{50} = 0.075 μM) gave a stronger calculated binding energy (-10.23 kcal/mol) than all the peptides in this study. This is expected since the observed HIV-PR activities for the peptides in this study were at least two orders of magnitude weaker than that of PCU-EAISa. The most active cage ether peptide (PCU-ether-VA, IC_{50} = 0.6 μM) gave a docked binding energy of -9 kcal/mol, which is better than the best cage diol peptide **7** in this study (Fg-PCU-diol-Fg, IC_{50} = 5 μM) with a calculated binding energy of ~ -8.4 kcal/mol. In all our docking studies so far, the cage heteroatoms form strong hydrogen bonding interaction with the enzymatic aspartic acid (Asp25/Asp25') residues.

As previously reported,[41] it was demonstrated that the chirality of the PCU cage is important for enhanced inhibition of the diastereomeric cage peptides. In the present study no diastereomeric mixtures were obtained. Docking studies have clearly showed that conserved hydrogen bonds are formed between the cage hydroxyl groups and one of the triad residues, Asp25-Thr26-Gly27 (A/B chains) as shown in Figure 6. Such hydrogen bonds anchor the cage skeleton to the S1/S1' subsite.

Even though docking calculations cannot provide insight on the dynamic inhibitor-enzyme interactions, a general picture was provided of the most energetically favourable binding orientation of inhibitors to the enzyme. To obtain further insight into the dynamic changes of the docked inhibitors within the enzyme active site pocket over time, the lowest energy docked complex of the inhibitor **7** was subjected to unconstrained MD simulations (2 ns).



A



B

Figure 6. **(A)** Lowest energy docked structure for compound **7** with C-SA HIV-PR. The two PR monomers are coloured as the blue and mauve ribbon. The inhibitor and the two Asp25 residues presented as coloured sticks. **(B)** A closer view showing the binding mode of compound **7** inside the C-SA HIV-PR active site. (The PDB file from which this presentation was made is available as supplementary material).

The MD simulations clearly showed that the inhibitor easily fits into the active enzyme pocket. To assess the quality of these MD simulations, energetic and structural properties are monitored along the entire 2 ns MD trajectory of the complex. Figure 7A presents the plot of the potential energy of the system as a function of time. The fluctuation of potential energy is less than 1000 kcal mol^{-1} over the course of the reaction.

The 2 ns averaged backbone RMSDs for the **7** enzyme complex is 1.96 Å and this is an indication that the generated MD trajectory of the complex is quite stable. The overall orientation of the complex is preserved and very little movement of the PCU cage hydroxyl groups inside the enzyme pocket is observed during the course of the MD run. However, slight changes in the orientation in the cage inhibitor peptide side chains were observed.

As evident from the analysis of the hydrogen bonding interactions along the MD trajectories, the inhibitor is forming hydrogen bonds between the cage hydroxyl groups and at least one of the two Asp25 carboxyl groups (Figure 7 B and C). The post dynamic monitoring of the overall hydrogen bond networks and the electrostatic interaction between the PCU diol peptide inhibitor and nearby residues clearly revealed that these inhibitors settle comfortably inside the C-SA HIV-PR active site (Figure 8).

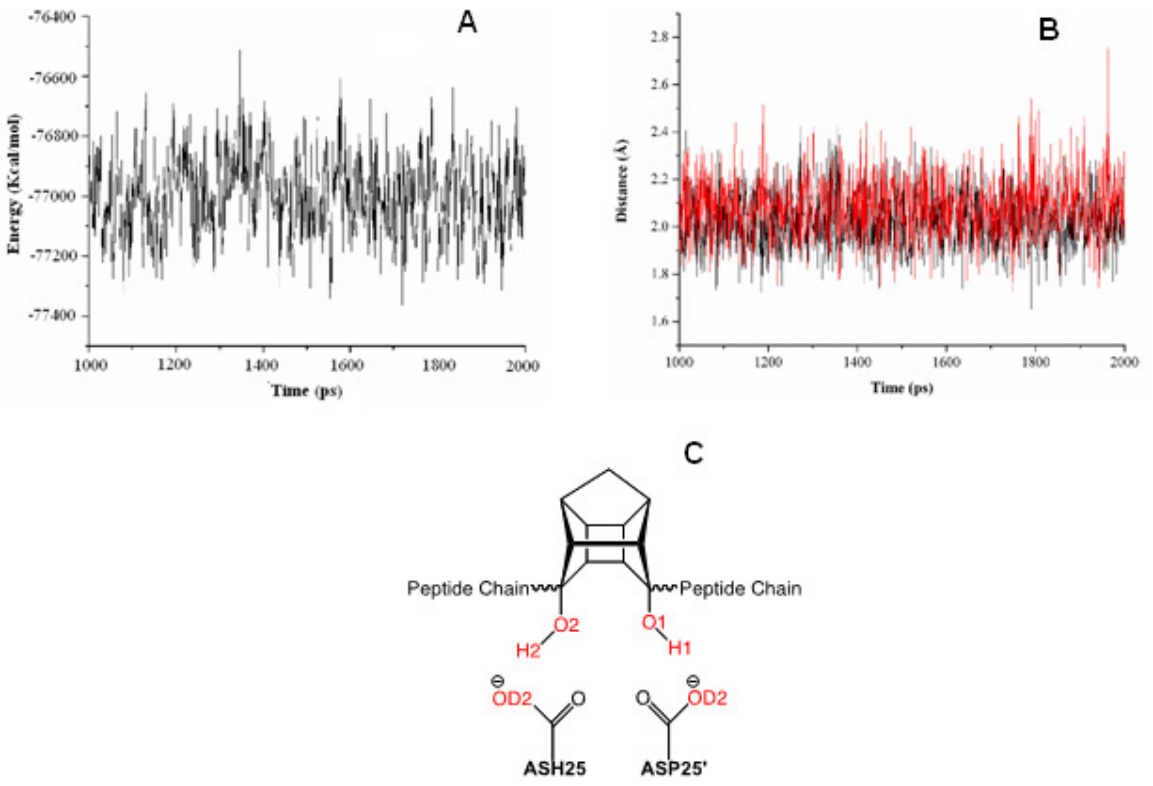


Figure 7. **(A)** The potential energy of **7** C-SA HIV-PR complex observed in MD simulation as a function of time. **(B)** Selected distances obtained from MD simulations: Black line, cage (H1)-Asp25'(OD2), Red line, cage (H2)-Asp25 (OD2) and (OD2). **(C)** Labelling of some atoms involved in hydrogen bond formation.

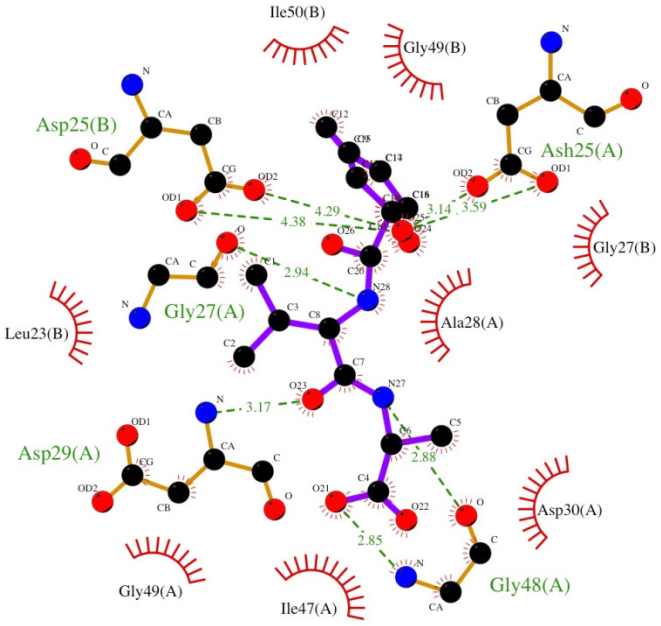


Figure 8. Modelled interaction of compound **7** with C-SA HIV-PR. Selected electrostatic and hydrogen bond interactions for the C-SA HIV-PR complex with compound **7** from the MD simulation (some of the atoms of the molecule have been deleted for clarity). These plots were created with the Ligplot software.[96] The 3D structure of this extraction from the MD simulation is available as supplementary material.

It is clear from the hydrogen bond interactions that the inhibitor would be fitting quite stable in the active site. Each hydrogen bond interaction contributes between 3-6 kcal/mol stabilization of the resulting complex.

4. CONCLUSION

In summary, this study designed and synthesised new HIV-1 protease inhibitors based on the pentacycloundecane derived diol diacid unit to which amino acids and oligomers of *N*-substituted glycine were attached. From the docking and MD simulations studies it was observed that the HIV-PR enzyme easily accommodates the most active cage diol peptides. Better IC_{50} values in the case of PCU cage diol peptides can potentially be achieved by changing the length of the link between the cage moiety and the side chain peptide. These possibilities currently form the basis of an extended theoretical investigation in our laboratories. Promising theoretical results will be scrutinised with a similar experimental approach. Although the

respective activities are much lower than those of the most potent drugs (Atzanavir and Lopinavir), the discovery of this new lead structure (the PCU cage diol diacid) opens exciting possibilities to serve as potential inhibitors to some of the wide variety of protease related diseases.

5. SUPPORTING INFORMATION

All the spectra mentioned in the text are available as supportive information (Appendix 1).

6. ACKNOWLEDGMENT

This research was supported by NRF (SA) TG (GUN: 66319), KP (GUN: 69728), HGK and PIA (SA-Sweden bilateral grant). Thanks to Aspen Pharmacare, University of KwaZulu-Natal and the NRF for financial support and the CHPC (www.chpc.ac.za) for computational resources. Mr. Dilip Jagjivan (UKZN, South Africa) is also thanked for his assistance with the NMR experiments.

7. REFERENCES

1. Gallo, R. C.; Sarin, P. S.; Gelmann, E. P.; Robertguroff, M.; Richardson, E.; Kalyanaraman, V. S.; Mann, D.; Sidhu, G. D.; Stahl, R. E.; Zollapazner, S.; Leibowitch, J.; Popovic, M. *Science* **1983**, *220*, 865-867.
2. Barresinoussi, F.; Chermann, J. C.; Rey, F.; Nugeyre, M. T.; Chamaret, S.; Gruest, J.; Dauguet, C.; Axlerblin, C.; Vezinetbrun, F.; Rouzioux, C.; Rozenbaum, W.; Montagnier, L. *Science* **1983**, *220*, 868-871.
3. Gallo, R. C.; Salahuddin, S. Z.; Popovic, M.; Shearer, G. M.; Kaplan, M.; Haynes, B. F.; Palker, T. J.; Redfield, R.; Oleske, J.; Safai, B.; White, G.; Foster, P.; Markham, P. D. *Science* **1984**, *224*, 500-503.
4. Levy, J. A.; Hoffman, A. D.; Kramer, S. M.; Landis, J. A.; Shimabukuro, J. M. *Science* **1984**, *225*, 840-842.
5. 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.
6. UNAIDS/WHO. In *UNAIDS*; UNAIDS: geneva, **2010**.
7. 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. *J. Med. Chem.* **1991**, *34*, 2852-2857.
8. Turner, B. G.; Summers, M. F. *J. Mol. Biol.* **1999**, *285*, 1-32.
9. Brik, A.; Wong, C. H. *Org. Biomol. Chem.* **2003**, *1*, 5-14.
10. Reetz, M. T.; Merk, C.; Mehler, G. *Chem. Commun.* **1998**, 2075-2076.
11. Huff, J. R. *J. Med. Chem.* **1991**, *34*, 2305-2314.
12. Freed, E. O. *Virology* **1998**, *251*, 1-15.

13. Meek, T. D. *J. Enzyme. Inhib.* **1992**, *6*, 65-98.
14. Meek, T. D.; Lambert, D. M.; Metcalf, B. W.; Petteway, S. R.; Dreyer, G. B. In *Design of Anti-AIDS Drugs;* Clereq, E. D. Ed.; Elservier: Amsterdam, **1990**; pp. 225-265.
15. Mastrolorenzo, A.; Rusconi, S.; Scozzafava, A.; Barbaro, G.; Supuran, C. T. *Curr. Med. Chem.* **2007**, *14*, 2734-2748.
16. Wlodawer, A.; Vondrasek, J. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 249-284.
17. 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.
18. 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. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2484-2488.
19. 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. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4096-4100.
20. Kaldor, S. W.; Kalish, V. J.; Davies, J. F., 2nd; 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. *J. Med. Chem.* **1997**, *40*, 3979-3985.
21. Kim, E. E.; Baker, C. T.; Dwyer, M. D.; Murcko, M. A.; Rao, B. G.; Tung, R. D.; Navia, M. A. *J. Am. Chem. Soc.* **1995**, *117*, 1181-1182.
22. 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. *Antimicrob. Agents Chemother.* **1998**, *42*, 3218-3224.
23. 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. *Antimicrob. Agents Chemother.* **2000**, *44*, 2093-2099.
24. Doyon, L.; Tremblay, S.; Bourgon, L.; Wardrop, E.; Cordingley, M. G. *Antivir. Res.* **2005**, *68*, 27-35.
25. Ghosh, A. K.; Dawson, Z. L.; Mitsuya, H. *Bioorg. Med. Chem.* **2007**, *15*, 7576-7580.
26. Carr, A.; Samaras, K.; Burton, S.; Law, M.; Freund, J.; Chisholm, D. J.; Cooper, D. A. *AIDS* **1998**, *12*, F51-F58.
27. Flexner, C. *N. Engl. J. Med.* **1998**, *338*, 1281-1292.
28. Lo, J. C.; Mulligan, K.; Tai, V. W.; Algren, H.; Schambelan, M. *Lancet* **1998**, *351*, 867-870.
29. Martinez, E.; Casamitjana, R.; Conget, I.; Gatell, J. M. *AIDS* **1998**, *12*, 2077-2079.
30. Miller, K. D.; Jones, E.; Yanovski, J. A.; Shankar, R.; Feuerstein, I.; Falloon, J. *Lancet* **1998**, *351*, 871-875.
31. Viraben, R.; Aquilina, C. *AIDS* **1998**, *12*, F37-F39.

32. Walli, R.; Herfort, O.; Michl, G. M.; Demant, T.; Jager, H.; Dieterle, C.; Bogner, J. R.; Landgraf, R.; Goebel, F. D. *AIDS* **1998**, *12*, F167-F173.
33. Erickson, J. W.; Burt, S. K. *Annu. Rev. Pharmacol. Toxicol.* **1996**, *36*, 545-571.
34. Waters, L.; Nelson, M. *Int. J. Clin. Pract.* **2007**, *61*, 983-990.
35. Waters, L.; Nelson, M. *Int. J. Clin. Pract.* **2007**, *61*, 999-1014.
36. Pecanha, E. P.; Figueiredo, L. J. O.; Brindeiro, R. M.; Tanuri, A.; Calazans, A. R.; Antunes, O. A. C. *Farmaco* **2003**, *58*, 149-157.
37. Lee, T.; Laco, G. S.; Torbett, B. E.; Fox, H. S.; Lerner, D. L.; Elder, J. H.; Wong, C. H. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 939-944.
38. Wlodawer, A. *Annu. Rev. Med.* **2002**, *53*, 595-614.
39. Hodge, C. N.; Aldrich, P. E.; Bacheler, L. T.; Chang, C. H.; Eyermann, C. J.; Garber, S.; Grubb, M.; Jackson, D. A.; Jadhav, P. K.; Korant, B.; Lam, P. Y.; Maurin, M. B.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Reid, C.; Sharpe, T. R.; Shum, L.; Winslow, D. L.; Erickson-Viitanen, S. *Chem. Biol.* **1996**, *3*, 301-314.
40. Kaltenbach, R. F., 3rd; Nugiel, D. A.; Lam, P. Y.; Klabe, R. M.; Seitz, S. P. *J. Med. Chem.* **1998**, *41*, 5113-5117.
41. Makatini, M. M.; Petzold, K.; Sriharsha, S. N.; Soliman, M. E. S.; Honarparvar, B.; Arvidsson, P. I.; Sayed, Y.; Govender, P.; Maguire, G. E. M.; Kruger, H. G.; Govender, T. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 2274-2277.
42. Karpoormath, R.; Sayed, Y.; Govender, P.; Govender, T.; Kruger, H. G.; Soliman, M. E. S.; Maguire, G. E. M. *Bioorg. Chem.* **2012**, *40*, 19-29.
43. Baker, W. R.; Condon, S. L. *J. Org. Chem.* **1993**, *58*, 3277-3284.
44. Levitskaia, T. G.; Moyer, B. A.; Bonnesen, P. V.; Marchand, A. P.; Krishnudu, K.; Chen, Z. B.; Huang, Z. L.; Kruger, H. G.; McKim, A. S. *J. Am. Chem. Soc.* **2001**, *123*, 12099-12100.
45. Boyle, G. A.; Govender, T.; Karpoormath, R.; Kruger, H. G. *Acta Crystallogr., Sect. E: Struct. Rep. Online* **2007**, *E63*, o4797-o4798.
46. Boyle, G. A.; Govender, T.; Karpoormath, R.; Kruger, H. G. *Acta Crystallogr., Sect. E: Struct. Rep. Online* **2007**, *E63*, So3977-So3978.
47. Karpoormath, R.; Govender, P.; Kruger, H. G.; Govender, T.; Maguire, G. E. M. *Acta Crystallogr., Sect. E: Struct. Rep. Online* **2010**, *E66*, o2537-o2538.
48. Geldenhuys, W. J.; Malan, S. F.; Bloomquist, J. R.; Marchand, A. P.; Van der Schyf, C. J. *Med. Res. Rev.* **2005**, *25*, 21-48.
49. Gerzon, K.; Kau, D. *J. Med. Chem.* **1967**, *10*, 189-199.
50. Gerzon, K.; Tobias, D. J., Sr.; Holmes, R. E.; Rathbun, R. E.; Kattau, R. W. *J. Med. Chem.* **1967**, *10*, 603-606.
51. Grobler, E.; Grobler, A.; Van der Schyf, C. J.; Malan, S. F. *Bioorg. Med. Chem.* **2006**, *14*, 1176-1181.
52. Ito, F. M.; Petroni, J. M.; de Lima, D. P.; Beatriz, A.; Marques, M. R.; de Moraes, M. O.; Costa-Lotufo, L. V.; Montnegro, R. C.; Magalhaes, H. I. F.; Pessoa, C. D. O. *Molecules* **2007**, *12*, 271-282.
53. Oliver, D. W.; Dekker, T. G.; Snyckers, F. O.; Fourie, T. G. *J. Med. Chem.* **1991**, *34*, 851-854.
54. Stamatiou, G.; Kolocouris, A.; Kolocouris, N.; Fytas, G.; Foscolos, G. B.; Neyts, J.; De Clercq, E. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2137-2142.

55. Zah, J.; Terre'Blanche, G.; Erasmus, E.; Malan, S. F. *Bioorg. Med. Chem.* **2003**, *11*, 3569-3578.
56. Joubert, J.; Geldenhuys, W. J.; Van der Schyf, C. J.; Oliver, D. W.; Kruger, H. G.; Govender, T.; Malan, S. F. *Chem. Med. Chem.* **2012**, *7*, 375-384.
57. Brookes, K. B.; Hickmott, P. W.; Jutle, K. K.; Schreyer, C. A. *S. Afr. J. Chem.* **1992**, *45*, 8-11.
58. Onajole, O. K.; Govender, K.; Govender, P.; van Helden, P. D.; Kruger, H. G.; Maguire, G. E. M.; Muthusamy, K.; Pillay, M.; Wiid, I.; Govender, T. *Eur. J. Med. Chem.* **2009**, *44*, 4297-4305.
59. Onajole, O. K.; Govender, P.; van Helden, P. D.; Kruger, H. G.; Maguire, G. E. M.; Wiid, I.; Govender, T. *Eur. J. Med. Chem.* **2010**, *45*, 2075-2079.
60. Govender, T.; Hariprakasha, H. K.; Kruger, H. G.; Marchand, A. P. *Tetrahedron-Asymmetr.* **2003**, *14*, 1553-1557.
61. Lassiani, L.; Pavan, M. V.; Berti, F.; Kokotos, G.; Markidis, T.; Mennuni, L.; Makovec, F.; Varnavas, A. *Bioorg. Med. Chem.* **2009**, *17*, 2336-2350.
62. Li, Y. G.; Zhang, J.; Xu, W. F.; Zhu, H. W.; Li, X. *Bioorg. Med. Chem.* **2010**, *18*, 1516-1525.
63. Chen, H.; Feng, Y.; Xu, Z.; Ye, T. *Tetrahedron* **2005**, *61*, 11132-11140.
64. Holmes, C. P.; Jones, D. G. *J. Org. Chem.* **1995**, *60*, 2318-2319.
65. Rutjes, F. P. J. T.; Veerman, J. J. N.; Meester, W. J. N.; Hiemstra, H.; Schoemaker, H. E. *Eur. J. Org. Chem.* **1999**, 1127-1135.
66. Santagada, V.; Fiorino, F.; Perissutti, E.; Severino, B.; De Filippis, V.; Vivenzio, B.; Caliendo, G. *Tetrahedron Letters* **2001**, *42*, 5171-5173.
67. Mosebi, S.; Morris, L.; Dirr, H. W.; Sayed, Y. *J. Virol.* **2008**, *82*, 11476-11479.
68. Gutierrez, O. A.; Salas, E.; Hernandez, Y.; Lissi, E. A.; Castrillo, G.; Reyes, O.; Garay, H.; Aguilar, A.; Garcia, B.; Otero, A.; Chavez, M. A.; Duarte, C. A. *Anal. Biochem.* **2002**, *307*, 18-24.
69. Simon, R. J.; Kania, R. S.; Zuckermann, R. N.; Huebner, V. D.; Jewell, D. A.; Banville, S.; Ng, S.; Wang, L.; Rosenberg, S.; Marlowe, C. K.; Spellmeyer, D. C.; Tan, R. Y.; Frankel, A. D.; Santi, D. V.; Cohen, F. E.; Bartlett, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9367-9371.
70. Avogadro: An open-source molecular builder and visualization tool. Version 1.0.0. Http://avogadro.Openmolecules.Net/
71. Frimurer, T. M.; Peters, G. H.; Iversen, L. F.; Andersen, H. S.; Moller, N. P. H.; Olsen, O. H. *Biophys. J.* **2003**, *84*, 2273-2281.
72. Locuson, C. W.; Gannett, P. M.; Ayscue, R.; Tracy, T. S. *J. Med. Chem.* **2007**, *50*, 1158-1165.
73. Soares, T. A.; Goodsell, D. S.; Ferreira, R.; Olson, A. J.; Briggs, J. M. *J. Mol. Recognit.* **2000**, *13*, 146-156.
74. Makatini, M. M.; Petzold, K.; Sriharsha, S. N.; Ndlovu, N.; Soliman, M. E. S.; Honarparvar, B.; Parboosing, R.; Naidoo, A.; Arvidsson, P. I.; Sayed, Y.; Govender, P.; Maguire, G. E. M.; Kruger, H. G.; Govender, T. *Eur. J. Med. Chem.* **2011**, *46*, 3976-3985.
75. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639-1662.
76. Sanner, M. F. *J. Mol. Graphics Modell.* **1999**, *17*, 57-61.

77. Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. *J. Comput. Chem.* **2007**, *28*, 1145-1152.
78. Wang, J. M.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Modell.* **2006**, *25*, 247-260.
79. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668-1688.
80. Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157-1174.
81. Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2005**, *26*, 114-114.
82. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M.; Kollman, P. A. *J. Comput. Chem.* **2003**, *24*, 1999–2012.
83. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
84. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577-8593.
85. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327-341.
86. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684-3690.
87. Onajole, O. K.; Makatini, M. M.; Govender, P.; Govender, T.; Maguire, G. E. M.; Kruger, H. G. *Magn. Reson. Chem.* **2010**, *48*, 249-255.
88. Cookson, R. C.; Crundwell, E.; Hill, R. R.; Hudec, J. *J. Chem. Soc. (Resumed)* **1964**, 3062-3075.
89. Marchand, A. P.; Allen, R. W. *J. Org. Chem.* **1974**, *39*, 1596-1596.
90. Clemente, J. C.; Coman, R. M.; Thiaville, M. M.; Janka, L. K.; Jeung, J. A.; Nukoolkarn, S.; Govindasamy, L.; Agbandje-McKenna, M.; McKenna, R.; Leelamanit, W.; Goodenow, M. M.; Dunn, B. M. *Biochemistry* **2006**, *45*, 5468-5477.
91. Karpoormath, R.; Govender, P.; Kruger, H. G.; Govender, T.; Maguire, G. E. M. *Acta Crystallogr., Sect. E: Struct. Rep. Online (CCDC 797748)* **2010**, *E66*, o2537-o2538.
92. Miller, S. M.; Simon, R. J.; Ng, S.; Zuckermann, R. N.; Kerr, J. M.; Moos, W. H. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2657-2662.
93. Miller, S. M.; Simon, R. J.; Ng, S.; Zuckermann, R. N.; Kerr, J. M.; Moos, W. H. *Drug Dev. Res.* **1995**, *35*, 20-32.
94. Zuckermann, R. N.; Kerr, J. M.; Kent, S. B. H.; Moos, W. H. *J. Am. Chem. Soc.* **1992**, *114*, 10646-10647.
95. Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. *J. Comb. Chem.* **1998**, *1*, 55-68.
96. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. *Protein Eng.* **1995**, *8*, 127-134.

CHAPTER 3

SYNTHESIS AND NMR ELUCIDATION OF NOVEL PENTACYCLOUNDECANE-DERIVED PEPTIDES

Rajshekhar Karpoormath,[a] Oluseye K. Onajole,[a] Tricia Naicker,[b] Thavendran Govender,[b] Glenn E. M. Maguire,[a] and Hendrik G. Kruger[a*]

[a]School of Chemistry, University of KwaZulu-Natal, Durban 4001, South Africa
[b]School of Pharmacy and Pharmacology, University of KwaZulu-Natal, Durban 4001, South Africa
Corresponding author: *Email: kruger@ukzn.ac.za, Fax: +27-31-2603091

ABSTRACT

Herein we report the synthesis and NMR elucidation of five novel pentacycloundecane (PCU) derived short peptides as potential HIV protease inhibitors. ^{1}H and ^{13}C spectral analysis show major overlapping of methine resonance of the PCU 'cage' thereby making it extremely difficult to assign the NMR signals. Attachment of short peptides to the cage at position C-8/C-11 results in conformational differences of the peptide side chains due to diastereomeric interactions between the cage skeleton and the chiral side chains. The use of two-dimensional NMR techniques proved to be highly effective in the elucidation of such systems.

***Keywords*:** ^{1}H NMR, ^{13}C NMR, 2D NMR, PCU diol diaminoacid, HIV protease inhibitors.

1. INTRODUCTION

The chemistry of polycyclic "cage" compounds such as adamantane,[1-8] pentacycloundecane (PCU),[9-16] trishomocubane[11,17-34] have been extensively studied by organic chemists over the years. A number of South African scientists have focused particularly on the chemistry and applications of PCU and trishomocubane polycyclic compounds.[8,10,14-18,24,25,27-30,35-52] Our research group has reported NMR studies in a bid to understand how the 'cage' skeleton interacts/relates to its side 'arms'.[15,16,53-63] As a part of ongoing research in this field, the NMR assignments of five PCU derivatives (**1-5**) are reported (Figure 1). Compounds **1** and **2** have

been previously reported,[63,75] however have never been fully characterized. Compounds **3-5** are novel PCU derivatives.

Several authors have commented on the challenges encountered with the NMR elucidation of cage compounds,[38,41,64-66] but the availability of 2D NMR techniques has helped to overcome these former difficulties. The unique features of the cage molecules such as their rigidity and long-range proton-proton interactions result in broad overlapping resonance and geminal/vicinal proton-proton couplings. This further complicates the structural elucidation of these compounds.

The incorporation of the cage moiety in medicinal compounds introduces a variety of advantages in terms of activity and pharmacology.[4,11] We have recently reported a family of PCU lactam peptides as potential South African human immunodeficiency virus type 1 subtype C protease (CSA-HIV-1 PR) inhibitors, which showed promising *in vitro* inhibition activity (IC_{50} = 0.78 μM) against the CSA-HIV-1 protease enzyme.[15,16] An extension of this study led to the design of compounds **1-5** (Figure 1).

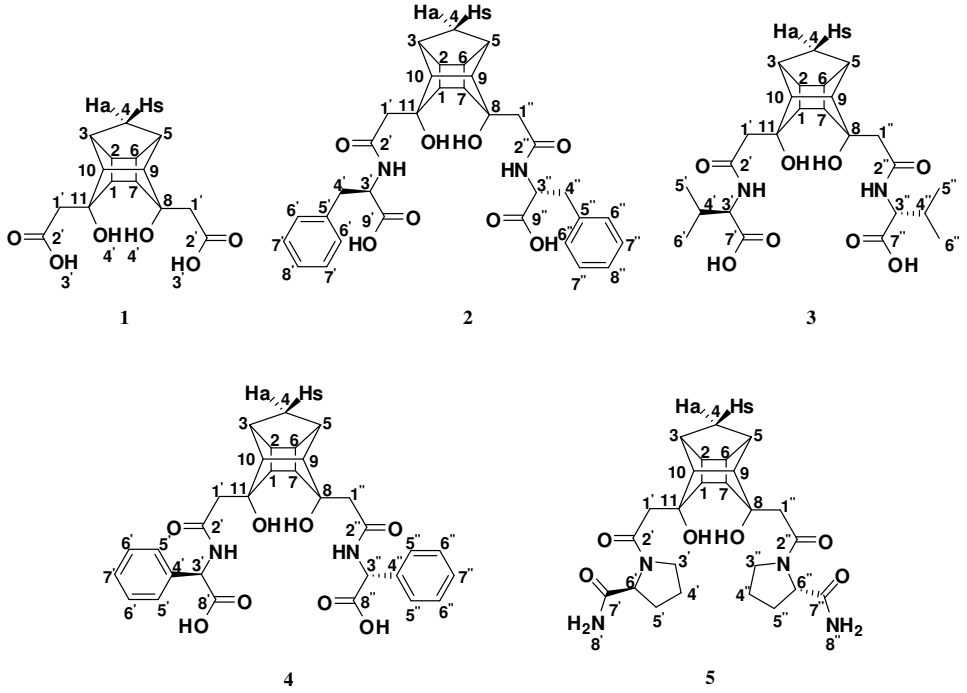
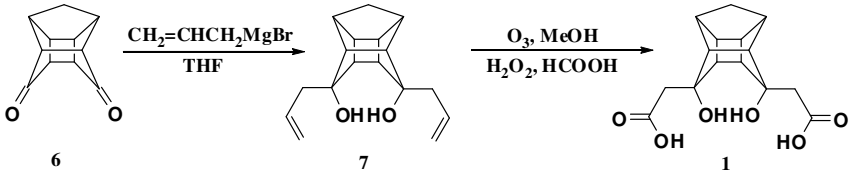


Figure 2. Structures of compounds **1-5**

The diols in the cage molecule are envisaged to interact with the catalytic ASP25/ASP25' residues in the active site of the HIV PR enzyme. It could also serve as a transition state mimic, due to the structural resemblance to a dihydroxyethylene type isostere.[67,68]

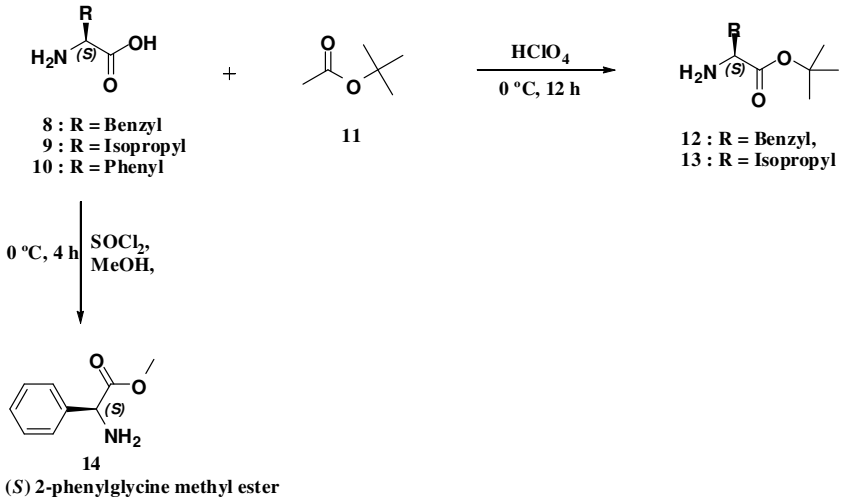
2. SYNTHESIS

Cookson's dione **6**[69], is reacted with freshly prepared allyl magnesium bromide (Grignard reaction) to obtain the *endo*-8, *endo*-11 diol **7**.[46,70] Ozonolysis of the diol **7** followed by an oxidative workup yielded the PCU diol dicarboxylic acid **1** in good yields[63] (Scheme 1).



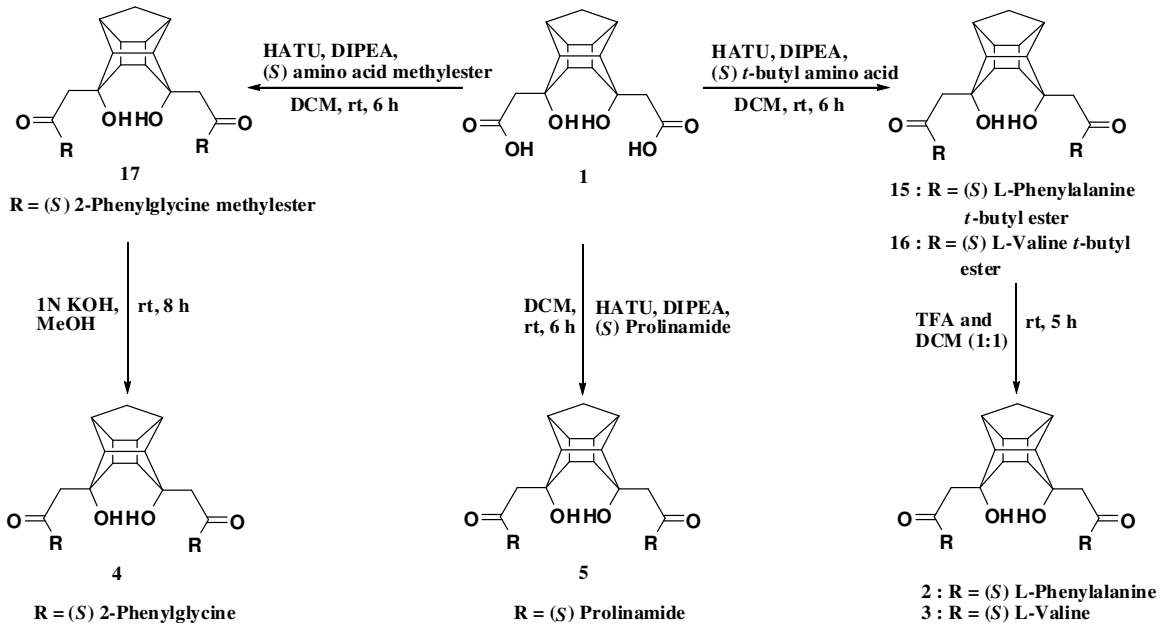
Scheme 1: Synthesis of PCU cage diol dicarboxylic acid (**1**)

Enantiomerically pure *tert*-butyl amino acid esters (**12** and **13**), were synthesized by reacting amino acids (**8** and **9**) with *tert*-butyl acetate (**11**) and $HClO_4$.[71] 2-Phenylglycine (**10**) was reacted with thionyl chloride and methanol to yield the corresponding enantiomerically pure methyl ester **14** (Scheme 2).[72]



Scheme 2. Synthesis of enantiomerically pure *tert*-butyl amino acid esters **12** and **13** and amino acid methyl ester **14**

Compound **1** in dry DCM was reacted with *tert*-butyl amino acid esters **12** and **13** or the methyl ester **14** in the presence of the coupling reagent HATU (2-(1*H*-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyl uronium hexafluorophosphate methanaminium) and DIPEA (*N,N*-diisopropylethylamine) as a base to give the corresponding compounds **5, 15**, **16**, and **17** (Scheme 3). Deprotection of PCU diol *tert*-butyl amino acid esters **15** and **16** was carried out using TFA (trifluoroacetic acid) in DCM (1:1) to yield the corresponding enantiopure cage amino acids **2** and **3** while the deprotection of compound **17** was achieved using aqueous KOH in methanol to yield enantiopure compound **4** (Scheme 3).



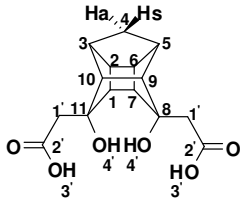
Scheme 3. Synthesis of PCU derivatives **2-5**

3. RESULTS AND DISCUSSION

Compound **1** is mesomeric while compounds **2-5** are optically pure (carbons C-8/C-11 have opposite chiralities). It is known from literature, that the H-4 cage proton signals appear as a pair of doublets (geminal protons) one each for H-4a and H-4s with an AB spin resonances at approximately 1.5 and 1.8 ppm (J ~ 10 Hz) respectively.

For compound **1**, the ^{1}H NMR spectrum exhibits the geminal PCU bridge methylene protons resonances of H-4a and H-4s at 1.02 and 1.49 ppm with a coupling constant of 10.4 Hz. A correlation was observed in the COSY and NOESY spectra for H-4a and H-4s with a signal at

2.36 ppm, which was assigned to H-3/H-5. These protons also show COSY correlations with two other methine protons (H-2/H-6 and H-9/H-10). The H-2/H-6 (2.47 ppm) signal correlates with that of H-1/H-7 (2.58 ppm) while the other signal (H-9/H-10) exhibits no further correlations. The NOESY spectrum also shows correlation of H-4a with H-2/H-6, while H-4s correlates with H-9/H-10. The HMBC and NOESY techniques are vital tools to determine how the side chains interact with the cage. The HMBC spectrum shows correlation of the C-9/C-10 peak with H-1' at 2.23 ppm and H-4a at 1.02 ppm. The quaternary carbon signal of C-8/C-11 (76.4 ppm) and the carbonyl carbon C-2' (172.3 ppm) also show HMBC correlations to H-1' (2.23 ppm), further confirming these assignments. Further verifications were carried out using the HSQC spectrum. The assignments for compound **1** are presented in Table 1.



1

Table 1. NMR data for Compound 1 in $(CD_3)_2SO$

Compound **1**			
Atom	$\delta^{1}H^{a,b}$	***J*(Hz)**	$\delta^{13}C^{a,b}$
1/7	2.58	-	42.2
2/6	2.47	-	38.8
3/5	2.36	-	43.8
4a	1.02	10.4	33.4
4s	1.49	10.4	33.4
8/11	-	-	76.4
9/10	2.30	-	49.9
1'	2.23	-	44.1
2'	-	-	172.3
3' (OH)	11.92	-	-
4' (OH)	7.16	-	-

[a] Solvent $(CD_3)_2SO$.
[b] 400 MHz for ^{1}H and 100 MHz for ^{13}C.

A similar methodology was used to assign the NMR signals of the cage protons and carbons for the remaining compounds. Details of the subsequent NMR assignments of the PCU skeleton will be omitted in further discussions for the sake of brevity.

All carbon atom signals for compound **2** appear as pairs. It is impossible to distinguish between split carbon and proton signals. The two quaternary carbons registering at 77.2 and 77.8 ppm were assigned to C-8/C-11 or C-11/C-8. The C-8 and C-11 signals exhibits a HMBC correlation with a pair of doublets registering at 2.24 and 2.34 ppm (J = 15 Hz) which were assigned to H-1'/H-1". The H-1'/H-1" protons are diastereotopic due to the neighbouring chiral carbon (C-8 and C-11). H-1'/H-1" displays HMBC correlation to two overlapping C-2'/C2" carbonyl carbons at 173.1 – 173.9 ppm. The two multiplets at 4.73 and 4.81 ppm, both integrating to one proton, were assigned to H-3'/H-3". H-3'/H-3" display both COSY and NOESY correlations with two methylene signals (multiplets) registering at 2.96 and 3.27 ppm (H-4'/H-4"). These two methylene protons H-4'/H-4" are also diastereotopic. H-3'/H-3" and H-4'/H-4" display HMBC correlations with a carbonyl carbon at 173.1 – 173.9 ppm, which were assigned to C-9'/C-9" (C-2'/C-2" and C-9'/C-9" are overlapping). H-3'/H-3" show a HMBC correlation to two quaternary carbons registering at 137.2 and 137.5 ppm which were assigned to C-5'/C-5". H-4'/H-4" show HMBC correlations to aromatic carbons registered at 128.8 and 128.9 ppm (C-6'/C-6"). Carbon signals for C7' (129.5 ppm) and C8'/C8" (127.1/127.3) were assigned[73,74] based on the intensity of the peaks. Peculiarly, C7' (129.5 ppm) did not display any splitting of the aromatic carbon signal.

We previously reported the X-ray crystal structure of compound **2** (Figure 2).[75] It is evident from the crystal structure that a conformational difference of the "arms" of the cage exists, with one arm pointing towards the back of the cage while the other is positioned in front of the cage. We have previously reported several cases where the same observation was made from NMR experiments in solution.[15,16,53-55,57,59] From the NOESY spectrum H-4s, H-9 and H-10 display weak correlations with the aromatic protons. These assignments are presented in Table 2.

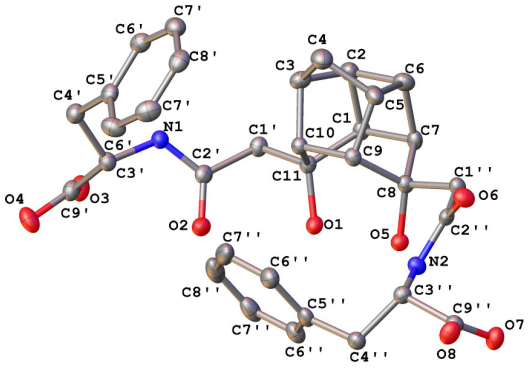
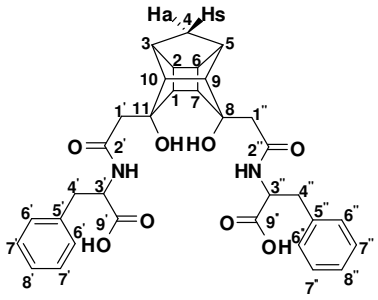
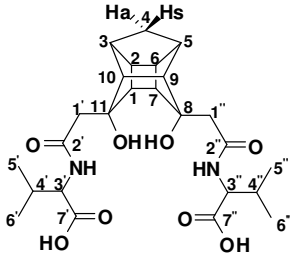


Figure 3. Crystal structure of compound **2**.[75] Hydrogen atoms are ommitted for clarity.

The ^{1}H NMR spectrum of compound **3** shows a muliplet (it should be a pair of doublets) for the chiral protons H-3'/H-3" registering at 4.36 ppm (J = 4.8 Hz) which integrate for two protons. These chiral methine protons exhibit a COSY correlation to a multiplet registering at 2.19 ppm, which was assigned to the adjacent proton H4'. This methylene proton (H4' at 2.19 ppm) displays both NOESY and COSY interactions with two methyl protons at 0.94, 1.00 and a chiral proton registered at 4.36 ppm (C-3'/C-3"). Based on these correlations, the two methyl protons were assigned to H-5'/H-5" and H-6'/H-6". The C-3'/C-3" peaks (58.6/58.7 ppm) also show HMBC correlations to H-4', H-5' and H-6'. Overlapping carbonyl signals in the region of 174.3 – 174.6 ppm (C-2' and/or C-7') show HMBC correlations to H-3'/H-3", H-4' and H-1'/H-1", hence the carbonyl carbons could not be distinguished. Splitting of some carbon signals are observed. This has been previously reported for similar compounds and it was attributed to conformational effects by the side arms.[55-57,59] In the case of compound **3** splitting of carbons for atom H-4' was not observed. These assignments are presented in Table 2.



2



3

Table 2. NMR data for PCU derivatives **2** and **3** in $(CD_3)OD$

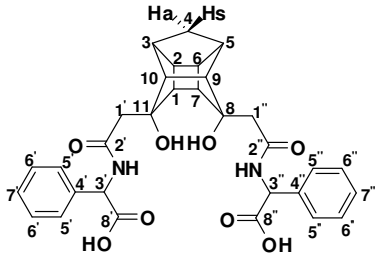
Compound **2**				Compound **3**			
Atom[c]	**δ^1H[a,b]**	***J* (Hz)**	**$\delta^{13}C$[a,b]**	**Atom[c]**	**δ^1H[a,b]**	***J* (Hz)**	**$\delta^{13}C$[a,b]**
1/7	2.09/2.11		44.5/44.8	**1/7**	2.55		44.3
2/6	2.48		39.4/39.7	**2/6**	2.62		40.6/40.7
3/5	2.26/2.30		43.2-43.8	**3/5**	2.49/2.55		45.7
4a	1.04	10.6	33.9	**4a**	1.18	10.8	34.8
4s	1.39	10.5	33.9	**4s**	1.60	10.7	34.8
8/11	-		77.2/77.8	**8/11**	-		78.5/78.6
9/10	1.67/1.89	8.8	49.1/49.4	**9/10**	2.33/2.40		50.7/50.9
1'/1''	2.24/2.34	15	43.2/44.8	**1'/1''**	2.33/2.47		45.2/45.3
2'/2''/9'/9''	-		173.1-173.9	**2'/7'**	-		174.3/174.6
3'/3''	4.73/4.81	4.6	53.8/53.9	**3'/3''**	4.36	4.8	58.6/58.7
4'/4''	2.96/3.27	6.5	37.5/37.6	**4'**	2.19	4.6	31.7
5'/5''	-		137.2/137.5	**5'/5''/6'/6'**	0.94-1.00		18.0/18.2/19.6/1
				'			9.7
6'/6''	7.14-7.32		128.8/128.9				
7'	7.14-7.32		129.5		-		
8'/8''	7.14-7.32		127.1/127.3				

[a] Solvent CD_3OD.
[b] 400 MHz for ^1H and 100 MHz for ^{13}C. Compound **2**: 600 MHz for ^1H and 125 MHz for ^{13}C
[c] It was not possible to distinguish between the atoms on the left and right side of the cage and side arms.

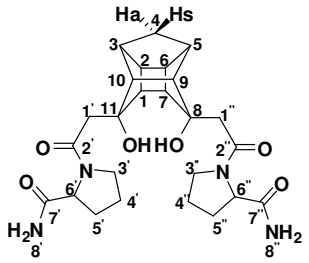
The ^{13}C NMR spectrum of compound **4** shows two overlapping carbonyl peaks registering in the region of 173.8 ppm. These overlapping signals were assigned to C-2' and C-8'. Interestingly the majority of the side chain atoms were not split as for compounds **2** and **3**.

C-8' displays a HMBC correlation to a methine proton registering at 5.43 ppm; this methine proton further exhibits a HMBC correlation to a quaternary carbon at 138.8 ppm and a phenyl carbon registering at 128.7 ppm. The methine proton was assigned to H-3' (5.43 ppm) while the quaternary carbon and the phenyl carbon were assigned to C-4' and C-5' respectively. H-3' also displayed NOESY interaction with the phenyl protons at regions 7.30 – 7.43 ppm. The remaining phenyl carbons registering at 129.4 and 129.8 ppm were assigned to C-6' and C-7' based on the integration of the peaks.[73] C-2' (173.8 ppm) and C-8/11 (a chiral carbon, 78.7 ppm) both display HMBC correlations to overlapping proton signals in the region of 2.17 and 2.24 ppm. The HSQC spectrum shows that these two signals correlate to one carbon (45.3 ppm) suggesting that they are doublets (this observation could not be made exclusively from the ^{1}H spectrum due to the overlapping in this region). These two signals were assigned to H-1' and H1". All remaining carbons and protons were assigned using the HSQC spectrum. These assignments are presented in Table 3.

The ^{13}C NMR spectrum of compound **5** shows two carbonyl signals at 170.6 and 173.8 ppm. The side chains did not reveal split carbon signals. These signals are due to the C-2'/C-2" or C-7'/C-7" carbonyl carbons. Only the signal at 173.8 ppm shows HMBC correlation to the H-8'/H-8" protons at 6.92 and 7.08 ppm respectively; the carbonyl signal at 173.8 ppm was therefore assigned to C-7' and C-7". The signal at 170.6 ppm was assigned to C-2' and C-2". The carbonyl carbon (C-7'/ C-7") signal also shows correlation to the chiral methine proton at 4.16 ppm, which was assigned to H-6' and H-6". The chiral methine proton (H-6'/H-6") shows COSY correlation to H-5'/H-5" at 1.99 pm. This was further confirmed from the HMBC spectrum since the signal of the carbonyl carbon C-7'/C-7" shows correlation to H5'/H-5". H5'/H-5" displays both COSY and NOESY correlations to a signal at 1.81 ppm, which was assigned to H-4' and H-4". The methylene protons registered at 3.50 ppm showed COSY correlation to H-4'/H-4" which was then assigned to H-3' and H-3". In addition, the C-2'/C-2" signal shows a HMBC correlation to H-3'/ H-3". The carbonyl carbon C-2'/C-2" and the quaternary carbons C-8/C-11 display HMBC correlations to methylene protons at 2.04 and 2.37 ppm which were assigned to H-1' and H-1". HMBC and HSQC spectra were further used to confirm the assigned proton and carbon signals. The signal assignments for compounds **4** and **5** are presented in Table 3.



4



5

Table 3. NMR data for PCU derivatives **4** in $(CD_3)OD$ and **5** in $(CD_3)_2SO$

Compound **4**				Compound **5**			
Atom	**$\delta^1H^{b,c}$**	***J* (Hz)**	**$\delta^{13}C^{b,c}$**	**Atom**	**$\delta^1H^{a,c}$**	***J* (Hz)**	**$\delta^{13}C^{a,c}$**
1/7	2.37		45.7	**1/7**	2.67		42.0
2/6	2.51		40.6	**2/6**	2.52		38.5
3/5	2.43		44.4	**3/5**	2.42		43.9
4a	1.10	10.8	34.8	**4a**	1.04	10.6	33.5
4s	1.50	10.7	34.8	**4s**	1.48	10.5	33.5
8/11			78.7	**8/11**	-		77.8
9/10	2.15		50.8	**9/10**	2.10		50.3
1'/1''	2.17/2.24		45.3	**1'/1''**	2.04/2.37		41.3/42.2
2'/8'			173.8	**2'**	-		170.6
3'	5.43		58.2	**3'**	3.50		47.3
4'			138.8	**4'**	1.81		24.0
5'	7.30-7.43		128.7	**5'**	1.99		29.3
6'	7.30-7.43		129.4	**6'**	4.16	5.59	59.3
7'	7.30-7.43		129.8	**7'**	-		173.8
				8'/8''	6.92/7.08		-

[a]Solvent $(CD_3)_2SO$.
[b]Solvent $(CD_3)OD$.
[c]400 MHz for ^1H and 100 MHz for ^{13}C.

4.EXPERIMENTAL SECTION

IR spectra were obtained from a Perkin Elmer Spectrum 100 instrument with an attenuated total reflectance attachment. High Resolution Electron Spray Ionization Mass Spectroscopic analysis was performed for all compounds on a Bruker MicroTOF QII mass spectrometer in positive mode with the exception of compound **1** in which data was collected in the negative mode. All

MS samples were analysed with an internal calibration. Melting point analysis (uncorrected) was performed on a Stuart Scientific digital melting point apparatus SMP3. Tetrahydrofuran (THF) was freshly distilled using sodium wire/benzophenone under nitrogen (N_2) atmosphere. Dichloromethane (DCM) was dried using phosphorus pentoxide as drying agent. The NMR data were recorded on Bruker AVANCE III 400 and 600 MHz instruments; the chemical shifts were referenced to the solvent peak, namely δ = 7.24 ppm for $CDCl_3$, δ = 2.50 ppm for $(CD_3)_2SO$, and δ = 3.34 ppm for CD_3OD at ambient temperature. Analytical analysis was performed on an Agilent 1100 HPLC (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 8A instrument (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. Mobile phase A consisted of 0.1% formic acid/H_2O (v/v) and mobile phase B consisted of 0.1% formic acid/acetonitrile (v/v). All the enantiopure (*S*) amino acids and coupling reagents were purchased from GL Biochem (Shanghai) Ltd. Analytical grade solvents and reagents such as allyl magnesium bromide, *tert*-butyl acetate, DIPEA, prolinamide for synthesis were procured from Sigma-Aldrich (South Africa).

Synthesis of *exo*-8-*exo*-11-diallylcarboxilic acid pentacyclo[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]undecane-*endo*-8-*endo*-11-diol (1)

A solution of diene **7**^{46,70} (15.0 g, 58.0 mmol) in dry methanol (100 mL) was cooled in an external dry ice-isopropanol bath (-78 ºC). Ozone was bubbled into the reaction mixture until a blue-purple colour persisted indicating the presence of excess ozone in the system and hence the completion of the reaction. The excess ozone gas was flushed from the reaction vessel with a stream of nitrogen and the solvent (CH_3OH) removed in *vacuo*. Formic acid (100 mL) was added to the ozonide and the mixture was cooled in a ice bath with stirring. Hydrogen peroxide (150 mL, 30%) was then added drop-wise to the stirred cooled reaction mixture. The reaction was left to attain ambient temperature for 1 h and then gently refluxed for 12 h, the resulting mixture was concentrated in *vacuo* to yield crude product. The crude product was purified *via* column chromatography on silica gel using CH_3COOH: MeOH: EtOAc: Hexane (1:2:20:77, R_f = 0.3, 92% yield) to obtain compound **1** as a viscous pale yellow oil, which solidified on

standing at room temperature to give a white solid. Melting point: 155 – 157 ºC, IR ν_{max}: 3102, 2985, 2864, 1698, 1639, 1451, 1292, 1280, 1181, 1165, 1120, 1068, 1045, 923, 862, 799, 736, 652, 575 and 477 cm^{-1}. ^{1}H NMR [$(CD_3)_2SO$, 400 MHz]: δ_H 1.02 (AB, J_{AB} = 10.4 H_Z, 1H), 1.49 (AB, J_{AB} = 10.4 H_Z, 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), 7.16 (s, 2H), 11.92 (s, 2H). ^{13}C NMR [$(CD_3)_2SO$, 100 MHz]: δ_C 33.4 (CH_2), 38.8 (CH), 42.2 (CH), 43.8 (CH), 44.0 (CH_2), 49.9 (CH), 76.4 (C), 172.3 (C). HR ESI *m/z*: calculated for $C_{15}H_{18}O_6$ $[M-H]^-$: 293.1031 found 293.1030.

General procedure for the synthesis of *exo*-8-*exo*-11-di-(*S*)-amino acid pentacyclo[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]undecane-*endo*-8-*endo*-11-diol (2, 3 and **4)**

A solution of compound **1** (1.0 eq., 0.5 g) in dry DCM (15 mL) was stirred at room temperature for 5 min. To this mixture was added enantiomerically pure *tert*-butyl amino acid esters/(*S*)-amino acid methyl esters (4.0 eq). The reaction mixture was cooled in an ice water bath and stirred for 5 min. This was followed by the addition of HATU (5.0 eq., 3.24 g) and DIPEA (8.0 eq., 1.76 g, 2.4 mL, a base). The solution was then slowly brought to room temperature and stirred for 6 h. The crude reaction mixture was washed with water (100 mL) and then with of 10% HCl (100 mL). The organic layer was dried over anhydrous Na_2SO_4 and filtered. The crude product was evaporated at 40 °C to dryness under vacuum to obtain thick yellow oil. The crude *tert*-butyl esters of the PCU cage diol diamino acids (**15**and **16,** ≈ 0.8 – 1.0 g) were further dissolved in 8 mL of 1:1 (v/v) DCM and TFA solvent mixture and stirred overnight. TFA was removed from the mixture by bubbling air through the peptide and the remaining DCM was removed under vacuum at 30 °C, to obtaincrude PCU cage diol diamino acids (**2** and **3**) as pale yellow oils.

The methyl esters of PCU cage diol di-(*S*)-amino acids (**17,** ≈ 0.9 g) were dissolved in methanol (3 mL) and excess of 1 N KOH (5 mL) and stirred overnight. Glacial acetic acid was added until the mixture was neutral and then diluted with chloroform. The organic solution was washed once with water and brine. The organic layer was dried over anhydrous $MgSO_4$ and concentrated *in vacuo* to give a crude oily compound **4**. These crude compounds were then purified by using preparative HPLC to give enantiomerically pure PCU cage diol diamino acids (**2**, **3** and **4**).

Data for *exo*-8-*exo*-11-diphenyl-(*S*)-alaninepentacyclo[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]undecane-*endo*-8-*endo*-11-diol (2)

A white solid (67%). Melting point: 231 – 232 ºC, $[\alpha]^{20}_{D}$:+14.29 (*c* 0.14 in MeOH). IR ν_{max}: 3261, 2957, 1758, 1638, 1522, 1191, 758, 701 and 489 cm^{-1}. ^{1}H NMR [CD_3OD, 400 MHz]: δ_H1.04 (1H, d, *J* = 10.6 Hz), 1.39 (1H, d, *J* = 10.5 Hz), 1.67 (1H, d, *J* = 8.8 Hz), 1.90 (1H, d, *J* = 8.8 Hz), 2.09-2.34 (8H, m), 2.45-2.47 (2H, m), 2.91-3.01 (2H, m), 3.24-3.31 (2H, m), 4.69-4.73 (1H, q, *J* = 4.7 Hz), 4.77-4.81 (1H, q, , *J* = 4.6 Hz), 7.13-7.17 (1H, m), 7.22-7.32 (9H, m). ^{13}C NMR [CD_3OD, 100 MHz]: δ_C 33.9 (CH_2), 37.5 (CH_2), 37.6 (CH_2), 39.4 (CH), 39.7 (CH), 43.2 (CH, CH_2), 43.8 (CH), 43.8 (CH), 44.5 (CH), 44.8 (CH, CH_2) 49.1 (CH), 49.4 (CH), 53.8 (CH), 53.9 (CH), 77.2 (C), 77.8 (C), 127.1 (CH), 127.3 (CH), 128.8 (CH), 128.9 (CH), 129.5 (CH), 137.2 (C), 137.5 (C), 173.1 (C), 173.9 (C). HR ESI *m/z*: calculated for $C_{33}H_{36}N_2O_8$ $[M+H]^+$: 589.2544 found 589.2541.

Data for *exo*-8-*exo*-11-di-(*S*)-valinepentacyclo[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]undecane-*endo*-8-*endo*-11-diol (3)

A colourless oil (72%). $[\alpha]^{20}_{D}$:+8.69 (*c* 0.23 in MeOH). IR ν_{max}: 2960, 2929, 1721, 1621, 1537, 1464, 1392, 1259, 1209, 1064, 986 and 544 cm^{-1}. ^{1}H NMR [CD_3OD, 400 MHz]: δ_H1.18 (1H, d, *J* = 10.8 Hz), 1.60 (1H, d, *J* = 10.7 Hz), 0.94-1.00 (12H, m), 2.19 (2H, q, *J* = 4.6 Hz), 2.20-2.62 (12H, m), 4.36 (2H, t, 4.8). ^{13}C NMR [CD_3OD, 100 MHz]: δ_C 34.8 (CH), 40.6 (CH), 40.7 (CH), 44.3 (CH), 45.2 (CH_2), 45.3 (CH_2), 45.7 (CH), 50.7 (CH), 50.9 (CH), 78.5 (C), 78.6 (C), 174.3 (C), 174.6 (C). HR ESI *m/z*: calculated for $C_{25}H_{36}N_2O_8$ $[M+H]^+$: 493.2531 found 493.2544.

Data for *exo*-8-*exo*-11-di-(*S*)-2-phenylglycinepentacyclo[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]undecane-*endo*-8-*endo*-11-diol (4)

A colourless oil (58%). $[\alpha]^{20}_{D}$:-30.00 (*c* 0.1 in MeOH). IR ν_{max}: 3060, 2994, 2863, 1723, 1629, 1524, 1452, 1256, 1179, 1133, 1063, 722, 696, 596 and 513 cm^{-1}. ^{1}H NMR [CD_3OD, 400 MHz]: δ_H1.10 (1H, d, *J* = 10.8 Hz), 1.50 (1H, d, *J* = 10.7 Hz), 2.15-2.58 (12H, m) 5.43 (2H, s), 7.30-7.43 (10H, m). ^{13}C NMR [CD_3OD, 100 MHz]: δ_C 34.8 (CH_2), 40.6 (CH), 44.4 (CH), 45.3 (CH_2), 45.7 (CH), 50.8 (CH), 58.2 (CH), 78.7 (C), 128.7 (CH), 129.4 (CH), 129.8 (CH), 173.8 (C). HR ESI *m/z*: calculated for $C_{31}H_{32}N_2O_8$ $[M+H]^+$: 561.2213 found 561.2231.

Synthesis of *exo*-8-*exo*-11-di-(*S*)-prolinamidepentacyclo[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]undecane-*endo*-8-*endo*-11-diol (5)

A solution of compound **1** (1.0 eq., 0.5 g) in dry DCM (15 mL) was stirred at room temperature for 5 min. To this mixture was added enantiopure (*S*)-prolinamide (4.0 eq., 0.8 g) whilst it was cooled in an ice water bath with stirring for 5 min. To the above cooled mixture was added HATU (5.0 eq., 3.24 g) followed by DIPEA (8.0 eq., 1.76 g, 2.4 mL). This reaction mixture was then slowly brought to room temperature and stirred for 6 h. The crude reaction mixtures was diluted with DCM and filtered. The crude product was evaporated to dryness at 40 °C under vacuum using a teflon pump to obtain a thick yellow oil. The product was purified with preparative HPLC to give enantiopure PCU cage diol diprolinamide as a pale yellow solid (82%). Melting point: 83 – 85 ºC. $[\alpha]^{20}_{D}$:-37.50 (*c* 0.16 in MeOH). IR ν_{max}: 3179, 2957, 2872, 1669, 1603, 1448, 1417, 1288, 1191, 1168, 1071, 987, 908, 653, 539 and 421 cm^{-1}. ^{1}H NMR [$(CD_3)_2SO$, 400 MHz]: δ_H1.04 (1H, t, *J* = 10.6 Hz), 1.48 (1H, t, *J* = 10.5 Hz), 1.74-2.72 (18H, m), 3.50-3.68 (2H, m), 4.16-4.22 (2H, m), 6.92-6.95 (2H, m), 7.08-7.22 (2H, m). ^{13}C NMR [$(CD_3)_2SO$, 100 MHz]: δ_C 24.03 (CH_2), 29.3 (CH_2), 33.5 (CH_2), 38.5 (CH), 41.3 (CH_2), 42.0 (CH), 42.2 (CH_2) 43.9 (CH), 47.3 (CH_2), 50.3 (CH), 59.3 (CH), 77.8 (C), 170.6 (C), 173.8 (C). HR ESI *m/z*: calculated for $C_{25}H_{34}N_4O_6$ $[M+H]^+$: 487.2542 found 487.2551.

5. CONCLUSION

The full NMR elucidation of five novel PCU cage derivatives was successfully achieved. Considerable overlapping of proton and carbon signals was observed, however this was overcome by using 2D NMR techniques. This technique proved to be a vital and convenient tool in elucidating PCU cage short peptides derivatives. These compounds are currently being evaluated as potential HIV protease inhibitors. It is unclear why some of the compounds give split side chain signals and others not.

6. SUPPORTING INFORMATION

All the spectra mentioned in the text are available as supporting information (Appendix.1).

7. ACKNOWLEDGMENT

This research was supported by NRF (SA) TG (GUN: 66319), HGK (GUN: 69728), HGK and PIA (SA-Sweden bilateral grant), Aspen Pharmacare and University of KwaZulu-Natal for financial support. The authors thank Dr Katja Petzold for assistance with the EASY-ROESY experiments. We thank Mr. Dilip Jagjivan (UKZN, South Africa) for their assistance with the NMR experiments.

8. REFERENCES

1. Gerzon, K.; Kau, D. *J. Med. Chem.* **1967**, *10*, 189-99.
2. Gerzon, K.; Tobias, D. J., Sr.; Holmes, R. E.; Rathbun, R. E.; Kattau, R. W. *J. Med. Chem.* **1967**, *10*, 603-606.
3. McKervey, M. A. *Chem. Soc. Rev.* **1974**,*3*, 479-512.
4. Brookes, K. B.; Hickmott, P. W.; Jutle, K. K.; Schreyer, C. A. *S. Afr. J. Chem.* **1992**, *45*, 8-11.
5. Ranganathan, D.; Haridas, V.; Madhusudanan, K. P.; Roy, R.; Nagaraj, R.; John, G. B.; Sukhaswami, M. B. *Angew. Chem.,Int. Ed. Engl.* **1996**, *35*, 1105-1107.
6. Spasov, A.; Khamidova, T.; Bugaeva, L.; Morozov, I. *Pharm. Chem. J.* **2000**, *34*, 3-9.
7. Lamoureux, G.; Artavia, G. *Curr. Med. Chem.* **2010**, *17*, 2967-2978.
8. Onajole, O. K.; Govender, P.; van Helden, P. D.; Kruger, H. G.; Maguire, G. E. M.; Wiid, I.; Govender, T. *Eur. J. Med. Chem.* **2010**, *45*, 2075-2079.
9. Marchand, A. P. *Advances in Theoretically Interesting* Molecules **1989**, *1*, 357-399.
10. Oliver, D. W.; Dekker, T. G.; Snyckers, F. O. *Eur. J. Med. Chem.* **1991**, *26*, 375-379.
11. Geldenhuys, W. J.; Malan, S. F.; Bloomquist, J. R.; Marchand, A. P.; Van der Schyf, C. J. *Med. Res. Rev.* **2005**, *25*, 21-48.
12. Grobler, E.; Grobler, A.; Van der Schyf, C. J.; Malan, S. F. *Bioorg. Med. Chem.* **2006**, *14*, 1176-1181.
13. Ito, F. M.; Petroni, J. M.; de Lima, D. P.; Beatriz, A.; Marques, M. R.; de Moraes, M. O.; Costa-Lotufo, L. V.; Montnegro, R. C.; Magalhaes, H. I. F.; Pessoa, C. D. O. Molecules **2007**, *12*, 271-282.
14. Oliver, D. W.; Malan, S. F. *Med. Chem. Res.* **2008**, *17*, 137-151.
15. Makatini, M. M.; Petzold, K.; Sriharsha, S. N.; Soliman, M. E. S.; Honarparvar, B.; Arvidsson, P. I.; Sayed, Y.; Govender, P.; Maguire, G. E. M.; Kruger, H. G.; Govender, T. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 2274-2277.
16. Makatini, M. M.; Petzold, K.; Sriharsha, S. N.; Ndlovu, N.; Soliman, M. E. S.; Honarparvar, B.; Parboosing, R.; Naidoo, A.; Arvidsson, P. I.; Sayed, Y.; Govender, P.; Maguire, G. E. M.; Kruger, H. G.; Govender, T. *Eur. J. Med. Chem.* **2011**, *46*, 3976-3985.
17. Dekker, T. G.; Oliver, D. W. *S. Afr. J. Chem.* **1979**, *32*, 45-48.
18. Dekker, T. G.; Oliver, D. W.; Venter, A. *Tetrahedron Lett.* **1980**, *21*, 3101-3104.
19. Mehta, G.; Srikrishna, A. *J. Chem. Soc, Chem. Commun.* **1982**, 218-219.
20. Mehta, G.; Reddy, A. V.; Srikrishna, A. *J. Chem. Soc, Perkin. Trans.* 1 **1986**, 291-297.
21. Oliver, D. W.; Dekker, T. G. *S. Afr. J. Sci.* **1988**, *84*, 407-409.

22. Marchand, A. P. *Chem. Rev.* **1989**, *89*, 1011-1033.
23. Marchand, A. P.; Reddy, G. M.; Deshpande, M. N.; Watson, W. H.; Nagl, A.; Lee, O. S.; Osawa, E. *J. Am. Chem. Soc.* **1990**, *112*, 3521-3529.
24. Oliver, D. W.; Dekker, T. G.; Snyckers, F. O.; Fourie, T. G. *J. Med. Chem.* **1991**, *34*, 851-854.
25. Oliver, D. W.; Dekker, T. G.; Wessels, P. L. *Magn. Reson. Chem.* **1994**, *32*, 330-334.
26. Kassiou, M.; Nguyen, V. H.; Knott, R.; Christie, M. J.; Hambley, T. W. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 595-600.
27. Fourie, L.; Govender, T.; Hariprakasha, H. K.; Kruger, H. G.; Raasch, T. *Magn. Reson. Chem.* **2004**, *42*, 617-623.
28. Govender, T.; Hariprakasha, H. K.; Kruger, H. G.; Raasch, T. *S. Afr. J. Chem.* **2005**, *58*, 37-40.
29. Bisetty, K.; Govender, P.; Kruger, H. G. Biopolymers **2006**, *81*, 339-349.
30. Albericio, F.; Arvidson, P. I.; Bisetty, K.; Giral, E.; Govender, T.; Jali, S.; Kongsaeree, P.; Kruger, H. G.; Prabpai, S. *Chem. Biol. Drug. Des.* **2008**, *71*, 125-130.
31. Banister, S. D.; Moussa, I. A.; Jordan, M. J. T.; Coster, M. J.; Kassiou, M. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 145-148.
32. Banister, S. D.; Moussa, I. A.; Beinat, C.; Reynolds, A. J.; Schiavini, P.; Jorgensen, W. T.; Kassiou, M. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 38-41.
33. Banister, S. D.; Moussa, I. A.; Jorgensen, W. T.; Chua, S. W.; Kassiou, M. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 3622-3626.
34. Singh, P.; Bisetty, K.; Govender, P.; Kruger, G. *S. Afr. J. Chem.* **2011**, *64*, 49-56.
35. Dekker, J.; Dekker, J. J.; Fourie, L.; Dekker, T. G.; Pachler, K. G. R.; Wessels, P. L. *Tetrahedron Lett.* **1976**, *17*, 1613-1614.
36. Pachler, K. G. R.; Wessels, P. L.; Dekker, J.; Dekker, J. J.; Dekker, T. G. *Tetrahedron Lett.* **1976**, *17*, 3059-3062.
37. Dekker, T. G.; Pachler, K. G. R.; Wessels, P. L.; Dekker, J.; Dekker, J. J.; Fourie, L. *Org. Magn. Reson.***1978**, *11*, 520-523.
38. Dekker, T. G.; Oliver, D. W.; Pachler, K. G. R.; Wessels, P. L.; Woudenberg, M. *Org. Magn. Reson.* **1981**, *15*, 188-192.
39. Dekker, T. G.; Van der Schyf, C. J.; Kruger, G. J.; Van Rooyen, P. H. *S. Afr. J. Chem.* **1983**, *36*, 59-64.
40. Oliver, D. W.; Dekker, T. G. *S. Afr. J. Chem.* **1989**, *42*, 43-44.
41. Martins, F. J. C.; Coetzee, G. H.; Fourie, L.; Venter, H. J.; Viljoen, A. M.; Wessels, P. L. *Magn. Reson. Chem.* **1993**, *31*, 578-584.
42. Martins, F. J. C.; Viljoen, A. M.; Kruger, H. G.; Joubert, J. A. Tetrahedron **1993**, *49*, 9573-9580.
43. Martins, F. J. C.; Viljoen, A. M.; Kruger, H. G.; Wessels, P. L. *Tetrahedron* **1993**, *49*, 6527-6532.
44. Martins, F. J. C.; Viljoen, A. M.; Kruger, H. G.; Joubert, A. J.; Wessels, P. L. Tetrahedron **1994**, *50*, 10783-10790.
45. Martins, F. J. C.; Viljoen, A. M.; Kruger, H. G.; Fourie, L.; Roscher, J.; Joubert, A. J.; Wessels, P. L. Tetrahedron **2001**, *57*, 1601-1607.
46. Govender, T.; Hariprakasha, H. K.; Kruger, H. G.; Marchand, A. P. *Tetrahedron-Asymmetr.* **2003**, *14*, 1553-1557.

47. Boyle, G. A.; Govender, T.; Kruger, H. G.; Maguire, G. E. M. *Tetrahedron-Asymmetr.* **2004**, *15*, 2661-2666.
48. Boyle, G. A.; Govender, T.; Kruger, H. G.; Maguire, G. E. M. *Tetrahedron-Asymmetr.* **2004**, *15*, 3775-3781.
49. Gokul, V.; Kruger, H. G.; Govender, T.; Fourie, L.; Power, T. D. *J. Mol. Struc-Theochem.* **2004**, *672*, 119-125.
50. Odisitse, S.; Jackson, G. E.; Govender, T.; Kruger, H. G.; Singh, A. *Dalton Trans.* **2007**, 1140-1149.
51. Onajole, O. K.; Govender, K.; Govender, P.; van Helden, P. D.; Kruger, H. G.; Maguire, G. E. M.; Muthusamy, K.; Pillay, M.; Wiid, I.; Govender, T. *Eur. J. Med. Chem.* **2009**, *44*, 4297-4305.
52. Wilkes, D. K.; de Vries, A.; Oliver, D. W.; Malan, S. F. *Archiv. Der.* Pharmazie **2009**, *342*, 73-79.
53. Kruger, H. G.; Ramdhani, R. *S. Afr. J. Chem.* **2006**, *59*, 71-74.
54. Kruger, H. G.; Ramdhani, R. *Magn. Reson. Chem.* **2006**, *44*, 1058-1062.
55. Boyle, G. A.; Kruger, H. G.; Maguire, G. E. M.; Singh, A. *Struct. Chem.* **2007**, *18*, 633-639.
56. Boyle, G. A.; Govender, T.; Kruger, H. G.; Maguire, G. E. M.; Naicker, T. *Struct. Chem.* **2008**, *19*, 429-434.
57. Boyle, G. A.; Govender, T.; Kruger, H. G.; Maguire, G. E. M.; Naicker, T. *Magn. Reson. Chem.* **2008**, *46*, 1089-1095.
58. Govender, T.; Kruger, H. G.; Makatini, M.; Onajole, O. K. *Struct. Chem.* **2008**, *19*, 719-726.
59. Onajole, O. K.; Govender, T.; Makatini, M.; Kruger, H. G. *Magn. Reson. Chem.* **2008**, *46*, 1007-1014.
60. Onajole, O. K.; Govender, P.; Govender, T.; Maguire, G. E. M.; Kruger, H. G. *Struct. Chem.* **2009**, *20*, 1067-1076.
61. Altaib, M. S.; Arvidsson, P. I.; Govender, T.; Maguire, G. E. M.; Makatini, M.; Onajole, O. K.; Kruger, H. G. *Magn. Reson. Chem.* **2010**, *48*, 435-442.
62. Onajole, O. K.; Govender, P.; Govender, T.; Maguire, G. E. M.; Kruger, H. G. *Struct. Chem.* **2010**, *21*, 1203-1209.
63. Onajole, O. K.; Makatini, M. M.; Govender, P.; Govender, T.; Maguire, G. E. M.; Kruger, H. G. *Magn. Reson. Chem.* **2010**, *48*, 249-255.
64. Craze, G. A.; Watt, I. *J. Chem. Soc. Perk .Trans. 2.* **1981**, 175-184.
65. Cadd, D. H.; Feast, W. J.; Kenwright, A. M.; Say, J. M. *Magn. Reson. Chem.* **1993**, *31*, 801-807.
66. Martins, F. J. C.; Viljoen, A. M.; Kruger, H. G.; Wessels, P. L. *Magn. Reson. Chem.* **2004**, *42*, 402-408.
67. Baker, W. R.; Condon, S. L. *J. Org. Chem.* **1993**, *58*, 3277-3284.
68. Brik, A.; Wong, C. H. *Org. Biomol. Chem.* **2003,***1*, 5-14.
69. Cookson, R. C.; Crundwell, E.; Hill, R. R.; Hudec, J. *J. Chem. Soc.* **1964**, 3062-3075.
70. Marchand, A. P.; Huang, Z. L.; Chen, Z. B.; Hariprakasha, H. K.; Namboothiri, I. N. N.; Brodbelt, J. S.; Reyzer, M. L. *J. Heterocycl. Chem.* **2001**, *38*, 1361-1368.
71. Chen, H.; Feng, Y.; Xu, Z.; Ye, T. *Tetrahedron* **2005**, *61*, 11132-11140.
72. Rutjes, F. P. J. T.; Veerman, J. J. N.; Meester, W. J. N.; Hiemstra, H.; Schoemaker, H. E. *Eur. J. Org. Chem.* **1999**, 1127-1135.

73. Osborne, A. G.; Ahmet, M. T.; Miller, J. R.; Warmsley, J. F. *Spectrochim. Acta Part A* **1995**, *51A*, 237-246.
74. Katritzky, A. R.; Nikonov, G. N.; Moyano, E. L.; Akhmedov, N. G.; Steel, P. J. *Arkivoc* **2003**, *5*, 121-138.
75. Karpoormath, R.; Govender, P.; Kruger, H. G.; Govender, T.; Maguire, G. E. M. *Acta. Crystallogr. Sect. E: Struct. Rep. Online (CCDC 797748)* **2010**, *E66*, o2537-o2538.

CHAPTER 4

PENTACYCLOUNDECANE DERIVED HYDROXY ACID PEPTIDES: A NEW CLASS OF IRREVERSIBLE NON-SCISSILE ETHER BRIDGED TYPE ISOSTERE AS POTENTIAL HIV-1 WILD TYPE C-SA PROTEASE INHIBITORS

Rajshekhar Karpoormath,[a] Yasien Sayed,[c] Patrick Govender,[b] Thavendran Govender,[d] Hendrik G. Kruger,[a] Mahmoud E. S. Soliman[e*] and Glenn E. M. Maguire[a*]

[a]School of Chemistry, University of KwaZulu-Natal, Durban 4001, South Africa
[b]School of Biochemistry, University of KwaZulu-Natal, Durban 4001, South Africa
[c]Protein Structure-Function Research Unit, School of Molecular and Cell Biology, University of the Witwatersrand, Wits 2050, South Africa
[d]School of Pharmacy and Pharmacology, University of KwaZulu-Natal, Durban 4001, South Africa
[e]Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Zagazig University, Zagazig 44519, Egypt
**** Corresponding authors: email: mess20@bath.ac.uk and maguireg@ukzn.ac.za.***

ABSTRACT

Novel peptides incorporating the PCU derived hydroxy acid (5-Hydroxy-4-oxahexacyclo[5.4.1.$0^{2,6}$.$0^{3,10}$.$0^{5,9}$.$0^{8,11}$]dodecane) were synthesized and their activity against the resistance-prone wild type C-South African (C-SA) HIV-protease is reported. The attachment of peptides and peptoids to the PCU derived hydroxy acid resulted in a series of structurally diverse promising HIV-1 protease inhibitors. Amongst the nine novel compounds, ***16, 17, 20*** *and* ***23*** *gave IC_{50} values ranging from 0.6-5.0 μM against the wild type C-SA HIV-1 protease enzyme. Docking studies and molecular dynamic (MD) simulations have been carried out in order to understand the binding mode of the PCU moiety at the active site of the HIV protease enzyme. A conserved hydrogen bonding pattern between the PCU derived hydroxy ether and theactive site residues, ASP25/ASP25', was observed in all active compounds.*

Keywords**:** Transition state analogs, 5-Hydroxy-4 oxahexacyclo[5.4.1.$0^{2,6}$.$0^{3,10}$.$0^{5,9}$.$0^{8,11}$]dodecane peptides, HIV-1 wild type C-SA protease, PCU derivedpeptides, HIV protease inhibitor, inhibitory concentration (IC_{50}), Docking, Molecular dynamics.

1. INTRODUCTION

HIV is a global health problem and has become a major threat to the existence of the human race. It has already caused an estimated of more than 25 million deaths worldwide and has generated profound demographic changes in the most heavily affected countries. Globally in 2012 alone there are 34.0 million people living with HIV and 2.5 million new HIV infections worldwide.[1] The human immunodeficiency virus (HIV) encodes an aspartic protease (PR) that is essential for the formation of mature and infectious virions.[2,3] The HIV-PR is an established target in the chemotherapy of acquired immunodeficiency syndrome (AIDS) and intensive efforts have been directed to develop potent, orally available, peptidomimetic inhibitors for this enzyme.[3-6] At present, the FDA has approved nine PR inhibitors, and several others are now in clinical trials.[6-8] However, long-term anti-retroviral therapy for HIV infected patients promotes the emergence of resistance mutations of the HIV-PR, and consequently reduces the clinical efficacy of these agents.[9,10] Thus, there is a continuing demand for newer mutant-resistant HIV-PR inhibitors. Fortunately, HIV-PR variants expressing resistance to inhibitors have also been derived in cell culture and therefore new classes of potential inhibitors can be evaluated against drug-resistant HIV strains.[11]

Difficulties associated with drug-delivery have been observed mostly in peptide-based drugs due to their highly hydrophilic nature, which results in poor membrane permeability.[3,4,12] A second limitation of peptide drugs is their susceptibility to *in-vivo* degradation.

Over the past few years it was reported that the incorporation of polycyclic cage frameworks into biologically active molecules often improves the activity due to enhanced transport across cellular membranes.[13-17] The lipophilic nature of the cage skeleton also enables the molecules to cross the blood-brain barrier as well as the central nervous system (CNS).[14,18-20] A second advantage of incorporating the cage skeleton into such compounds is an increased resistance to biodegradability.[13] The computational design,[21-34] NMR elucidation[35-41] and biological studies[42-45] of several types of polycyclic cage molecules were previously reported in our group.

We have recently reported a family of PCU lactam peptides as potential HIV-PR inhibitors.[33,34] EASY-ROESY NMR experiments were used to determine crucial information about the 3D structure of these peptides in solution. The active inhibitors exhibited a turn of the peptide side chain that interacts with the cage protons. Docking of the cage peptides with the C-SA PR was also performed and a correlation of the binding energies with the IC_{50} inhibition results was evident. It was also demonstrated for the first time that the chirality of the PCU moiety was important for inhibition activity. Although the inhibitors approached activities in the nano molar range, it was clear that the need for better cage derived inhibitors should be explored. We decided to investigate a different PCU derivative as a potential transition state analogue as well as to assess whether the introduction of peptoids may have a positive effect on the inhibition activity.

Considerable effort has been invested to improve the pharmacological properties of peptides, mainly by increasing their enzymatic stability while preserving the chemical moieties required for functional activity.[46,47] 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 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.[48-51]

Herein we report the synthesis and efficacy of peptide and peptoid based PCU hydroxy ether derived compounds as potential HIV-1 protease inhibitors. The strategy used for designing HIV protease inhibitors is to mimic the transition state analogue of the Asp-induced enzyme cleavage thereby increasing the enzyme's affinity for the inhibitor over the peptide substrate.

Literature reports indicate that phenylalanine is the most common substituent at the S^1 subsite and it is found in 40% of *Gag-Pol* gene sequences of the virus.[52] According to literature, the $S^1/S^{1'}$ subsites accommodate hydrophobic substituents.[53] $S^2/S^{2'}$ and $S^3/S^{3'}$ subsites accommodate less hydrophobic substituents such as valine and alanine. It is postulated that inhibitors must have peptidic character with the scissile bond of the substrate being replaced by a non-cleavable bond.[54,55] The substrate specificity is determined by subsites located closely to the scissile bond.

Primarily the $S^1/S^{1'}$ and $S^2/S^{2'}$ subsites being responsible for substrate recognition.[56] Based on this concept, we have designed and synthesized nine novel PCU derived hydroxy acid peptides and peptoid analogs. Variations of the amino acid sequence in the peptide chain were introduced. It was shown that the cage ether type bond of **2** is virtually non-hydrolysable under conditions that were harsher than physiological conditions.[57]

Amprenavir (Figure 1) is protease inhibitor and was first developed by GlaxoSmithKline. This drug was approved by the Food and Drug Administration in 1995. Darunavir (Figure 1a) a second generation protease inhibitor named after A. K. Ghosh. He was mainly responsible for developing nonpeptidyl ligands (Figure 1) that displayed enhanced activity when cyclic ethers were incorporated into the structures.[58-60] The rationale behind using the cage ether is that when it is incorporated into a short peptide, this cyclic hydroxy ether moiety could perhaps serve as a cyclic ethertransition state analogue under protease conditions. The structural resemblance between the non-scissile cyclic ether and the cage ether is presented in Figure 1, (structures **1a**, **1b** and **2**). If this proposed interaction between the cage-peptide and the PR is possible, then it should result in effective inactivation of the PR.[53]

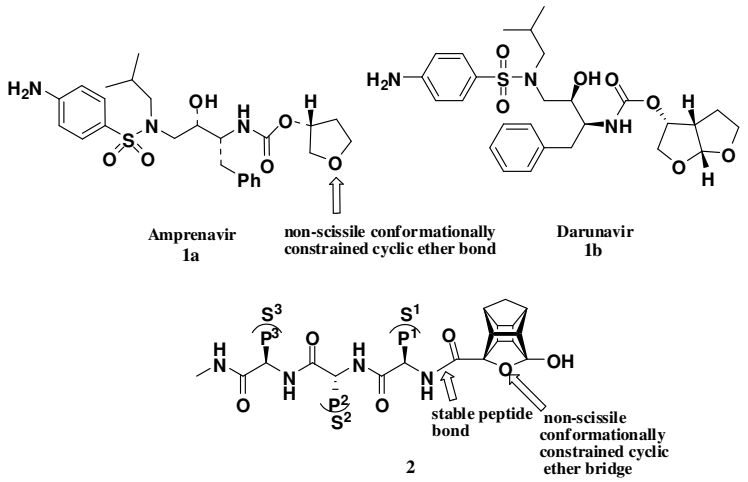


Figure 1. Structure of the proposed 5-Hydroxy-4-oxahexacyclo[5.4.1.$0^{2,6}$.$0^{3,10}$.$0^{5,9}$.$0^{8,11}$]dodecane peptide showing structural resemblance to the constrained cyclic ether bond of Amprenavir and Darunavir

This approach is a supplement to our recent report about the first family of PCU-lactam peptide HIV-PR inhibitors.[33,34] In that study, information about the 3D solution structure obtained from

EASY-ROESY NMR techniques of the cage peptide could be correlated to the inhibitory activity of the peptide. A docking study followed by a QM/MM/MD study of the inhibitor/HIV-PR complex suggestedthat, the hydroxy lactam function of the PCU interacts with the Asp25/Asp25' residues of the protease enzyme.

2. MATERIAL AND METHODS

2.1. EXPERIMENTAL

Analytical analysis was performed on an Agilent 1100 HPLC (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 8A instrument (Ace C18 150 mm x 21.2 mm x 5 microns) with a UV/VIS detector (215 nm) and automated fraction collector. A two-mobile phase system was employed, utilizing formic acid as the ion-pairing agent. Mobile phase A consisted of 0.1% formic acid/H_2O (v/v) and mobile phase B consisted of 0.1% formic acid/acetonitrile (v/v). High Resolution Electron Spray Ionization Mass Spectroscopic (HRESIMS) analysis was performed on a Bruker MicroTOF QII mass spectrometer in positive mode with an internal calibration. Microwave couplings were conducted on a Discovery CEM Liberty microwave peptide synthesizer. The coupling conditions of the peptide synthesizer were adapted from literature.[33,34] The ^{13}C NMR and ^{1}H NMR data were recorded on a Bruker AVANCE III 400 MHz spectrometer. Some samples were analyzed using a Bruker AVANCE III 600MHz when higher sensitivity was required. Protease hydrolytic activity was measured by monitoring the relative decrease in absorbance at 300 nm using an Analytik Jena Specord 210 spectrophotometer. Optical rotations were measured at room temperature in dry methanol using a Perkin Elmer Polarimeter-Model 341.All IR spectra were recorded on a Perkin Elmer Spectrum 100 instrument with a universal ATR attachment. All the amino acids, resins and coupling reagents are commercially available and were purchased from GL Biochem (Shanghai) Ltd. Analytical grade solvents for synthesis were procured from Sigma-Aldrich (South Africa).

In vitro HIV-1 protease activity

The catalytic activity of the HIV-1 protease[34,61] was monitored following the hydrolysis of the chromogenic peptide substrate Lys-Ala-Arg-Val-Nle-*p*-nitro-Phe-Glu-Ala-Nle-NH_2. 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 a 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 1 in supplementary information).

To determine the concentration of the compounds that resulted in 50% inhibition (IC_{50}) 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 an Analytik Jena Specord 210 spectrophotometer. Activity was standardized using commercially available drugs Atazanavir and Lopinavir. All cage peptides were soluble in the aqueous buffer solution. (pH = 5, 50 mM sodium acetate and 0.1 M NaCl)

General Procedure for the synthesis of 5-Hydroxy-4-oxahexacyclododecane -3-carboxylic acid, compound 15

Compound **15** was synthesized according to the previously reported method.[57]

General procedure for manual SPPS loading of first amino acid to 2-chlorotrityl chloride resin[33,34]

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) in DCM (10 mL) and *N*,*N*-diisopropylethylamine (DIPEA) (6.65 mmol), was added to the resin and was mixed for

2 hrs with a stream of nitrogen bubbles. 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 dried under vacuum 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 of SPPS 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 amino acids (0.2 mM), DIPEA (1 mM) and *O*-benzotriazole-*N,N,N',N'*-tetramethyl-uronium-hexafluoro-phosphate(HBTU) (2 mM) were prepared and inserted to their appropriate reaction vessels on the peptide synthesizer. All peptides were synthesized on a 0.5 mmol scale using microwave power. The washing procedure was similar to the one used for manual SPPS. The standard coupling method for Arginine (Arg) was used for attachment of the second amino acid to the resin (1 g) preloaded with the first amino acid. The standard coupling method was used for subsequent reactions and for Fmoc deprotection.[33,34,62] The washing procedures were similar to those used for manual SPPS.

Final cleavage from the resin was performed manually. The resin bound peptide was washed with DCM (3 x 10 mL) in a reaction vessel while nitrogen was bubbled through the solution. A cleavage mixture of 0.5% (v/v) TFA and 95% (v/v) DCM was added to the dried 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 pressure using a Teflon pump at 40 °C and affording the peptide as a white powder.

General procedure for the synthesis of peptoids[48-51]

Activated 2-chlorotrityl chloride resin (1.0 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, 11.51 mmol) and DIPEA (2 mL) were added to the resin and the mixture was bubbled for 2 hrs *via* a stream of nitrogen. The solvent was removed by filtration and the resin was washed with DCM (3 x 10 mL) and

DMF (3 x 10 mL). A solution of the amine in DMF (1 M) was added to the reaction vessel and bubbled for 1 hr with 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, 11.51 mmol) and *N,N'*-diisopropylcarbodiimide (DIC) (2 mL) was then added and the mixture was added to the reaction vessel and bubbled for 2 hrs using nitrogen. The resin was washed as described previously followed by the addition of a solution of the amine (1 M), bromoacetic acid (1.6 g, 11.51 mmol) in DMF and DIC (2 mL). Final cleavage of the peptoid from the resin was carried out similarly as that of thepeptides described previously.

General procedure for coupling of peptides and peptoids to 5-Hydroxy-4-oxahexacyclododecane-3-carboxylic acid, compounds 16-24

The cleaved peptide (1.2 eq.) was dissolved in DCM (3 mL) followed by the addition of 5-Hydroxy-4-oxahexacyclododecane-3-carboxylic acid[57] (1 equiv), HATU (2.5 equiv) in DMF (7 mL) and DIPEA (3 equiv) as a base. The mixture was left to stir at room temperature for 24 hrs. The product was evaporated to dryness under vacuum using a teflon pump at 40 °C. A cleavage mixture (10 mL) of 95% (v/v) TFA and 5% (v/v) DCM was added to the coupled peptide and stirred for 24 hrs at room temperature to remove the *N*-Boc protecting group. The TFA was removed from the mixture by bubbling air through the peptide and the remaining DCM was removed under *vacuum* at 30°C. The product was obtained as yellow oil which was purified by preparative HPLC.

Compound 16

$[\alpha]^{20}_{D}$ = -23.81 (*c* 0.21 in MeOD); ^{1}H NMR (400 MHz, DMSO, 100 °C): δ8.36 (t, *J* = 6.40 Hz, 1H), 7.99 (d, *J* = 9.12 Hz, 1H), 7.44 (m, 1H), 7.23-7.09 (m, 5H), 4.59 (m, 1H), 4.24-4.15 (m, 2H), 3.05-2.89 (m, 2H), 2.73-2.56 (m, 3H), 2.45-2.43 (m, 6H) 1.97-1.92 (m, 1H), 1.79-1.76 (m, 1H),1.43 (d, *J* = 10.16 Hz, 1H), 1.27 (d, *J* = 7.32 Hz, 3H), 0.89-0.83 (m, 6H). ^{13}C NMR (100 MHz, DMSO, 100 °C): δ170.8, 170.5, 170.4, 137.3, 129.3, 127.8, 126.2, 118.1, 89.5, 58.2, 57.9, 57.0, 56.1, 53.0, 48.4, 48.1, 47.4, 46.2, 46.0, 45.9, 42.7, 42.1, 40.9, 37.3, 31.1, 19.0, 18.1, 16.9, IR (neat): 3277.7, 2966.7, 1721.5, 1637.6, 1528.4, 1454.6, 1350.3, 1226.6, 1141.49, 911.2, 743.8, 698.9 cm^{-1}; HRESIMS (*m/z*): calculated for $C_{29}H_{35}N_3NaO_7$ ([M + Na]$^{23+}$) 560.2367 found 560.2351.

Compound 17

$[\alpha]^{20}_{D}$ = -27.27 (*c* 0.22 in MeOD); ^{1}H NMR (400 MHz, DMSO, 100 °C): δ8.38 (m,1H), 8.04 (d, *J* = 9.06, 1H), 7.41 (m, 1H), 7.25-7.10 (m, 5H), 4.62 (m, 1H), 4.24 (m, 1H), 3.84-3.71 (m, 2H), 3.36 (m, 1H), 3.06-2.92 (m, 2H), 2.73-2.57 (m, 3H), 2.45 (m, 5H), 1.97-1.95 (m, 1H), 1.80-1.77 (m, 1H),1.44 (d, *J* = 10.12 Hz, 1H), 0.89-0.84 (m, 6H). ^{13}C NMR (100 MHz, DMSO, 100 °C): δ 171.0, 170.5, 170.3, 137.3, 129.2, 127.9, 126.1, 118.1, 89.5, 79.1, 64.8, 58.2, 57.9, 57.2, 56.1, 53.0, 48.5, 48.1, 46.2, 46.0, 42.7, 42.1, 40.5, 37.3, 30.9, 19.0, 18.0, 15.1, IR (neat): 3288.5, 2965.7, 1728.2, 1638.4, 1524.0, 1327.6, 1206.3, 1138.62, 949.5, 905.3, 699.0 cm^{-1}; HRESIMS (*m/z*): calculated for $C_{28}H_{33}N_3NaO_7$ ([M + Na]$^{23+}$) 546.2110 found 546.2190.

Compound 19

$[\alpha]^{20}_{D}$ = -33.33 (*c* 0.21 in MeOD); ^{1}H NMR (400 MHz, DMSO, 100 °C): δ8.54 (d, *J* = 7.26 Hz, 1H), 8.02 (d, *J* = 9.08, Hz, 1H), 7.44 (m, 1H), 7.39-7.28 (m, 5H), 7.26-7.06 (m, 5H), 5.11 (s, 2H), 4.62 (m, 1H), 4.36 (m, 1H), 4.24 (m, 1H); 3.082-2.88 (m, 2H), 2.76-2.53 (m, 3H), 2.53-2.39 (m, 5H) 1.92 (m, 1H), 1.79 (d, *J* = 10.28 Hz, 1H), 1.43 (d, *J* = 10.16 Hz, 1H), 1.32 (d, *J* = 7.32 Hz, 3H), 0.89-0.78 (m, 6H) ^{13}C NMR (100 MHz, DMSO, 100 °C): δ172.2, 170.7, 170.6, 170.5, 170.3, 137.4, 135.9, 129.3, 128.4, 128.0, 127.9, 127.8, 126.2, 118.2, 89.6, 65.9, 58.3, 57.9, 57.0, 56.2, 53.1, 48.5, 48.2, 47.6, 46.3, 46.1, 45.9, 42.1, 41.0, 40.1, 37.4, 31.1, 19.0, 18.1, 16.7, IR (neat): 3287.8, 2964.4, 1714.9, 1640.5, 1524.1, 1327.2, 1205.2, 1138.4, 737.7, 697.0 cm^{-1}; HRESIMS (*m/z*): calculated for $C_{36}H_{41}N_3NaO_7$ ([M + Na]$^{23+}$) 650.2837 found 650.2828.

Compound 20

$[\alpha]^{20}_{D}$ = -20.00 (*c* 0.20 in MeOD); ^{1}H NMR (400 MHz, DMSO, 100 °C): δ8.43 (m, 1H), 6.91 (d, *J* = 3.1 Hz, 1H), 4.26-4.15 (m, 2H), 2.75-2.66 (m, 4H), 2.51 (m, 2H), 2.47 (m, 2H), 1.97-1.94 (m, 1H), 1.82 (q, *J* = 9.98, 3.14 Hz, 1H), 1.46 (d, *J* = 10.16 Hz, 1H), 1.27-1.24 (m, 3H), 0.87-0.77 (m, 6H). ^{13}C NMR (100 MHz, DMSO, 100 °C): δ 173.8, 170.4, 118.1, 89.6, 58.6, 57.9, 56.2, 55.9, 48.8, 48.1, 47.4, 46.3, 42.7, 42.1, 41.0, 40.0, 31.5, 19.0, 17.7, 16.9, 15.1, IR (neat): 3288.6, 2970.5, 1726.1, 1638.0, 1529.2, 1454.7, 1326.9, 1292.2, 1204.7, 1137.1, 992.5, 950.2, 906.4, 868.8, 641.7, 521.5 cm^{-1}; HRESIMS (*m/z*): calculated for $C_{20}H_{26}N_2NaO_6$ ([M + Na]$^{23+}$) 413.1683 found 413.1687.

Compound 22

$[\alpha]^{20}_{D}$ = -30.00 (*c* 0.21 in MeOD); ^{1}H NMR (400 MHz, DMSO, 100 °C): δ 8.14 (t, *J* = 7.98 Hz, 1H), 8.01 (d, *J* = 6.72 Hz, 1H), 7.84 (d, *J* = 3.82, 1H), 7.36 (d, *J* = 2.32, 1H), 4.39-4.21 (m, 4H), 3.70 (m, 4H), 2.78-2.60 (m, 4H), 2.49-2.47 (m, 2H), 2.18 (q, *J* = 17.02, 8.25, 2H), 1.92-1.69 (m, 4H), 1.46 (d, *J* = 10.16 Hz, 1H), 1.43-1.48 (m, 1H), 1.19-1.16 (m, 3H), 1.10-1.02(m, 1H), 0.84-0.77 (m, 6H). ^{13}C NMR (100 MHz, DMSO, 100 °C): δ 173.9, 171.8, 170.9, 170.8, 170.4, 118.2, 89.6, 61.3, 58.4, 58.2, 56.6, 56.3, 54.6, 51.1, 48.6, 48.4, 48.0, 46.4, 46.1, 42.8, 42.2, 41.1, 37.0, 29.9, 27.7, 24.1, 18.1, 15.2, 11.1. IR (neat): 3281.1, 2967.6, 1717.7, 1633.9, 1525.4, 1452.1, 1330.4, 1205.7, 1138.8, 1056.7, 993.7, 950.4, 904.2, 868.6, 789.4, 640.8, 616.9 cm^{-1}. HRESIMS (*m/z*): calculated for $C_{29}H_{40}N_4NaO_{11}$ ($[M + Na]^{23+}$) 643.2586 found 643.2579.

Compound 23

^{1}H NMR (400 MHz, DMSO, 100 °C): δ 4.75-4.69 (m, 1H), 4.45-4.35 (m, 1H), 4.15 (s, 1H), 3.97-3.95 (m, 2H), 3.84 (d, *J* = 4.32 Hz, 1H), 3.01 (d, *J* = 16.76 Hz, 2H), 2.96-2.89 (m, 2H), 2.79 (d, *J* = 9.56 Hz, 1H), 2.68-2.56 (m, 2H), 2.47-2.34 (m, 4H), 1.83(d, *J* = 10.08 Hz, 1H), 1.45(d, *J* = 10.52 Hz 1H), 1.05-0.99 (m, 6H). ^{13}C NMR (100 MHz, DMSO, 100 °C): δ 170.7, 168.2, 118.6, 90.5, 58.3, 55.2, 49.1, 48.3, 46.7, 45.6, 42.9, 41.0, 41.6, 41.0, 40.8, 37.8, 35.2, 34.5, 20.7, 19.3, 18.7. IR (neat): 3358.2, 2968.3, 2866.4, 1728.1, 1603.1, 1460.6, 1405.6, 1335.5, 1290.9, 1195.7, 1134.4, 1068.2, 1002.7, 904.5, 868.5, 827.9, 741.8, 645.7, 538.9 cm^{-1}; HRESIMS (*m/z*): calculated for $C_{20}H_{26}N_2NaO_6$ ($[M + Na]^{23+}$) 413.1683 found 413.1682.

Compound 24

^{1}H NMR (400 MHz, DMSO, 100 °C): δ 7.35-7.14 (m, 5H), 4.69-4.35 (m, 3H), 4.18-4.05 (m, 3H), 3.99-3.89 (m, 4H), 3.05-2.97 (m, 5H), 2.81-2.72 (m, 1H), 2.66-2.65 (m, 2H), 2.47 (s, 2H), 1.87-1.82 (m, 1H), 1.49-1.45 (m, 1H), 1.05-0.91 (m, 6H). ^{13}C NMR (100 MHz, DMSO, 100 °C): δ 170.6, 170.2, 169.0, 137.3, 128.4, 127.4, 118.7, 90.5, 58.7, 55.3, 48.6, 47.9, 45.6, 44.5, 42.9, 42.6, 41.8, 40.9, 35.2, 35.1, 20.33, 19.4, 19.2. IR (neat): 3332.8, 2971.8, 1731.5, 1629.7, 1453.7, 1405.5, 1343.5, 1291.8, 1197.6, 1135.9, 1075.7, 1007.0, 904.9, 868.7, 741.1, 630.7 cm^{-1}; HRESIMS (*m/z*): calculated for $C_{29}H_{35}N_3NaO_7$ ($[M + Na]^{23+}$) 560.2367 found 560.2358.

General procedure[63] for the synthesis of Benzyl protected 5-Hydroxy-4-oxahexacyclododecane peptides, 18 and 21

5-Hydroxy-4-oxahexacyclododecane peptide (1 eq) and benzybromide (2.5 eq) were dissolved in DMF (8 mL). Cesium carbonate (2 eq) was added to the solution at ambient temperature. After stirring for 2 h, saturated aqueous sodium bicarbonate (100 mL) was added to the solution and extracted with ethyl acetate (3 x 100 mL). The combined organic phases were washed with 5% aqueous citrate and saturated aqueous sodium chloride (100 mL) and dried over sodium sulphate. The solvent was evaporated, and the crude peptide was precipitated in cold ether and dried under reduced pressure. This crude product was further purified by preparative HPLC.

Compound 18

$[\alpha]^{20}_{D}$ = -19.05 (*c* 0.21 in MeOD); ^{1}H NMR (400 MHz, DMSO, 100 °C): δ 8.32 (d, *J* = 7.36 Hz, 1H), 7.95 (d, *J* = 9.06 Hz, 1H), 7.39 (m, 1H), 7.27-7.09 (m, 10H), 4.63 (m, 1H), 4.58 (m, 1H), 4.48-4.43 (m, 1H), 4.28-4.23(m, 1H), 3.55 (s, 1H), 3.36 (m, 1H), 3.09-2.89 (m, 4H), 2.71-2.55 (m, 3H), 2.45 (m, 4H), 1.97-1.93 (m, 1H), 1.78 (d, *J* = 10.32 Hz, 1H),1.43 (d, *J* = 10.12 Hz, 1H), 0.85-0.80 (m, 6H). ^{13}C NMR (100 MHz, DMSO, 100 °C): δ 172.7, 170.7, 170.4, 170.2, 137.4, 137.3, 129.2, 128.9, 128.1, 127.8, 126.3, 118.1, 89.5, 64.8, 58.2, 57.9, 56.9, 56.1, 53.3, 52.9, 48.5, 48.1, 46.2, 45.9, 42.7, 42.1, 40.9, 37.4, 37.3, 36.5, 31.1, 19.0, 17.8, 15.1, IR (neat): 3289.0, 2965.3, 1720.5, 1641.6, 1523.6, 1326.3, 1206.3, 1138.5, 906.4, 698.1 cm^{-1}; HRESIMS (*m/z*): calculated for $C_{35}H_{39}N_3NaO_7$ ($[M + Na]^{23+}$) 636.2719 found 636.2680.

Compound 21

$[\alpha]^{20}_{D}$ = -33.33 (*c* 0.21 in MeOD); ^{1}H NMR (400 MHz, DMSO, 100 °C): δ 8.59 (t, *J* = 6.82 Hz, 1H), 7.38-7.30 (m, 5H), 7.13 (d, *J* = 1.40 Hz, 1H), 6.88 (q, *J* = 9.26, 3.78 Hz, 1H), 5.10 (s, 1H), 4.36-4.31 (m, 1H), 4.27-4.22 (m, 1H), 2.77-2.63 (m, 4H), 2.56-2.51 (m, 2H), 2.47 (m, 1H), 1.96-1.90 (m, 1H), 1.82 (q, *J* = 10.20, 5.60 Hz, 1H), 1.47 (d, *J* = 9.96 Hz, 1H), 1.31-1.29 (m, 3H), 0.85-0.81 (m, 3H), 0.77-0.73 (m, 3H). ^{13}C NMR (100 MHz, DMSO, 100 °C): δ 172.1, 170.4, 170.3, 135.8, 128.3, 128.0, 127.7, 118.1, 89.6, 65.9, 58.6, 57.9, 56.2,55.9, 55.8, 48.8, 48.1, 47.5, 46.3, 46.0, 42.7, 42.1, 31.5, 19.0, 17.6, 16.7. IR (neat): 3286.7, 2966.8, 1742.8, 1643.5, 1524.4,

1455.5, 1347.2, 1203.2, 1160.4, 1138.8, 1052.4, 991.3, 949.3, 904.7, 738.0, 697.3, 640.9 cm^{-1}; HRESIMS (*m/z*): calculated for $C_{27}H_{32}N_2NaO_6$ ($[M + Na]^{23+}$) 503.2153 found 503.2150.

2.2. COMPUTATIONAL SIMULATIONS

Preparation of PCU derivedpeptide inhibitors

The structures of the cage-peptide/peptoid, **16-24**, were constructed and the geometry was optimized using the MMFF94 force field implemented in the Avogadro software.[64] Both diastereomers of the cage peptides were investigated. The minimized structures were then subjected to docking studies.

C-SA PR enzyme model system

Since the X-ray structure of the South African HIV-1 protease subtype C (C-SA) has not yet been reported, the initial 3D structure of the enzyme was taken from the reported X-ray data of subtype B HIV-PR (PDB accession code 1HXW).[10] C-SA HIV-PR differs from subtype B at eight positions; T12S, I15V, L19I, M36I, R41K, H69K, L89M and I93L.[61] The X-ray crystallographic coordinates of 1HXW were modified as follows:

The original substrate included in the crystal structure, Ritonavir, and the crystallographic water molecules were removed and hydrogens were added to the system. Ionization states for ionizable amino acid residues were determined according to their standard p*Ka* values. Mutations at eight positions, T12S, I15V, L19I, M36I, R41K, H69K, L89M and I93L were manually induced. Throughout this work, the modeled C-SA HIV-PR was used for docking and MD simulations. This constructed enzyme was successfully used in our previous study on cage lactam peptides.[33,34]

Docking of the inhibitors into the C-SA PR model

The energy minimized inhibitors, **16-24** were used in docking simulations. The effect of ionization states of docked compounds to the binding scores has been discussed in the literature.[65-67] The ionization states of the cage-peptide inhibitors have not yet been determined experimentally. During docking, the neutral and ionized states (aliphatic amine and carboxylic

acid groups of compounds to be docked were protonated and deprotonated, respectively) were considered and compared, separately. The simulations were performed under physiological pH conditions, which require the correct protonation state of ionizable groups. In particular, one of the aspartates (Asp25) of the catalytic site exhibits an increased p*K*a value of 5.2 in the inhibitor-bound protease[53] while no increased p*K*a was reported for the free form of the protease (p*K*a = 4.5).[68] Therefore, a monoprotonated active site should be prevalent at physiological pH, and thus this protonation state was used throughout these docking studies. However, to ensure that the results of the simulations do not critically depend on the protonation state, a control simulation of a cage-bound CSA-HIV-1 protease with a deprotonated carboxylate oxygen of the Asp25/Asp25' residues was also performed. We have previously reported that no significant difference was observed for docking of the cage lactam peptides with the different protonation states of the enzyme.[33,34]

Docking studies were performed using the Autodock software.[69] Geisteger charges were computed and the Autodock atom types were defined using theAutodock Tools graphical user interface supplied by MGL Tools.[70] The Lamarckian Genetic Algorithm (LGA), which is considered one of the best docking methods available in Autodock[69,71] was employed. This algorithm yields superior docking performance compared to simulated annealing or the simple genetic algorithm and the other search algorithms available in Autodock. The docked conformations of each of diastereomeric ligands were ranked into clusters based on the binding energy. The top ranked conformations was visually analyzed. The suitability of the constructed C-SA PR for docking experiments was tested with a control experiment in which Ritonavir, was docked into active site of the modeled enzyme and compared to the X-ray crystal structure[54] of the Ritonavir-protease complex.

MD simulations of inhibitor-enzyme complex:

MD simulations were performed for the inhibitor- enzyme complex of the most active compound **20**. The diastereomer with the best docked binding energy was used. Partial charges and the force field parameters for the inhibitor weregenerated using the Antechamber program[72] in the Amber10 package.[73] These were described by the general Amber force field (GAFF).[74] All

hydrogen atoms of the proteins were added using the Leap module in Amber10. The standard AMBER force field for bioorganic systems (ff03)[75] was used to describe the HIV PR enzyme parameters. Counter ions were added to neutralize the complex. Then, the system was solvated using atomistic TIP3P water[76] in a cubic box with 8.0 Å distance around the complex.

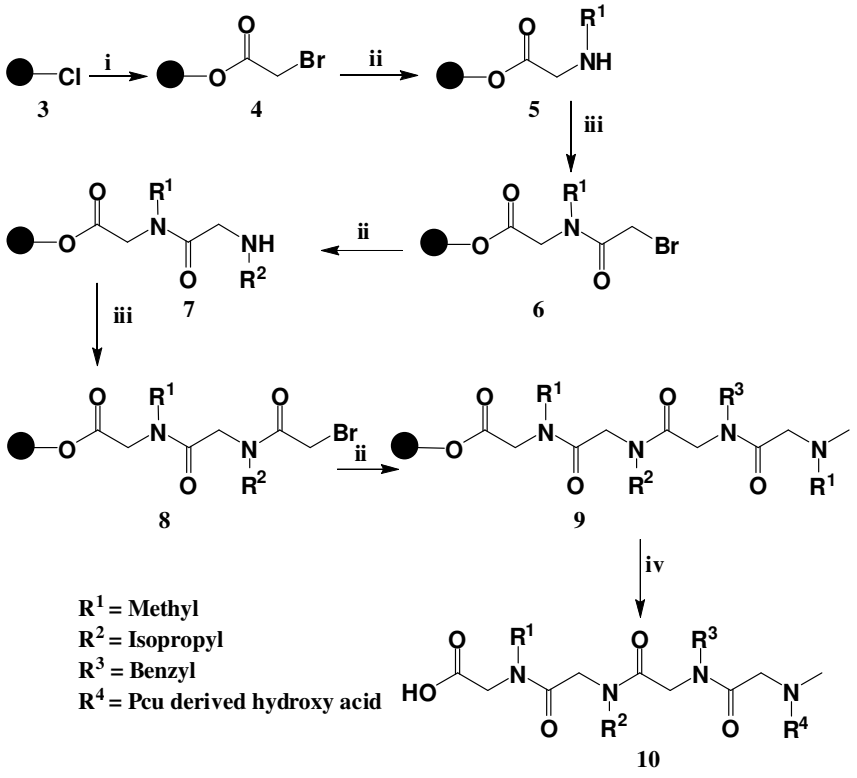
The molecular dynamics package Amber10[73] was used for the minimization and equilibration protocols. Cubic periodic boundary conditions were imposed and the long-range electrostatic interactions were treated with the particle-mesh Ewald method[77] implemented in Amber10 with a non-bonding cutoff distance of 10.0 Å. The energy minimization was first conducted using the steepest descent method in Amber10 for 1000 iterations switched to conjugate gradient for 2000 steps with a restraint potential of 2 kcal/mol $Å^2$ applied to the solute. Then the total system was freely minimized for 1000 iterations. For the equilibration and subsequent production run, the SHAKE algorithm[78] was employed on all atoms covalently bonded to a hydrogen atom, allowing for an integration time step of 2 fs. Harmonic restraints with force constants 2.0 kcal/mol $Å^2$ were applied to all solute atoms. A canonical ensemble (NVT) MD was carried out for 70 ps, during which the system was gradually annealed from 0 to 300 K using a Langevin thermostat with a coupling coefficient of 1.0/ps. Subsequently, the system was equilibrated at constant volume and temperature (300 K) with a 2 fs time step for 100 ps while maintaining the force constants on the restrained solute. With no restraints imposed, a production run was performed for 2 ns in isothermal isobaric (NPT) ensemble using a Berendsen barostat[79] with a target pressure of 1 bar and a pressure coupling constant of 2 ps. The coordinate file was saved every 1 ps and the trajectory was analyzed at every 1 ps using the Ptraj module implemented in Amber10.

3. RESULTS AND DISCUSSION

3.1. SYNTHESIS, SPECTRAL CHARACTERIZATION AND BIOLOGICAL ACTIVITY.

The cage peptides and peptoids were synthesized as described in the experimental section. In most cases, the peptides/peptoids were of sufficient purity to be coupled to the PCU derived hydroxy acid after lyophilisation. The peptoid oligomers were synthesized *via* solid phase synthesis on activated 2-chlorotrityl chloride resin **3** by attaching bromoacetic acid in the

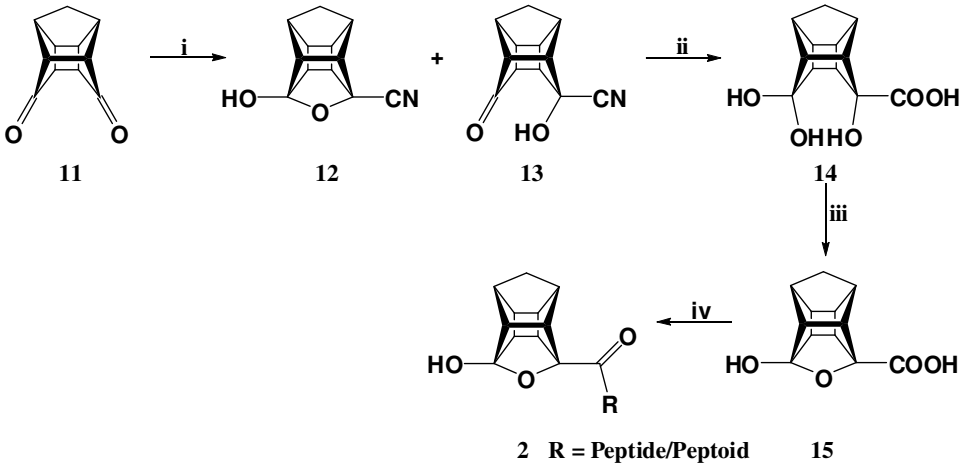
presence of DIPEA in dry dichloromethane (DCM) affording compound **4**. Upon filtration, the amine solution in DMF was added and mixed under a stream of nitrogen bubbles to give compound **5**. The chain was elongated by coupling bromoacetic acid in the presence of *N*,*N*'-diisopropylcarbodiimide (DIC). Similarly the corresponding amines were coupled to the extended chain followed by bromoacetic acid to give compounds **6-10**. The completed peptoids were cleaved from the resin using 5% trifluoroacetic acid (TFA) in DCM as illustrated in Scheme 1.



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

The cage hydroxy acid **15**[57]was synthesized from pentacycloundecane-8, 11-dione **11**[80-82] which can be easily obtained from the photocyclisation of the Diels-Alder adduct of cyclopentadiene and *p*-benzoquinone. A solution of the dione **11** in a mixture of water and acetic acid stirred in an ice bath is treated with an aqueous solution of sodium cyanide to form a mixture of the cyanohydrin (**12** and **13**).[57] This cyanohydrin mixture when treated with concentrated hydrochloric acid gives the tri-hydroxy carboxylic acid **14**.[57] It should be noted that this cage

acid is synthesised as a racemate. The resulting cage peptides are therefore a mixture of two diastereomers.tri-hydroxy carboxylic acid **14** was heated to reflux in 1,4-dioxane to give the desired hydroxy acid **15**. The peptides and peptoids obtained were coupled to **15** with HATU, DIPEA in DMF at ambient temperature as shown in Scheme 2.



Scheme 2. Synthesis of 5-Hydroxy-4-oxahexacyclo[5.4.1.02,6.03,10.05,9.08,11]dodecane peptide and peptoid analogs. **2**; **i** = NaCN, CH_3COOH, H_2O, 0 °C, **ii** = conc. HCL, reflux, **iii** = 1,4-dioxane, reflux, **iv** = HATU, DIPEA, DMF, RT.

The products **16** to **24** (Figure 2) were isolated and purified as described in the experimental section. It was not possible to separate the diastereomeric mixtures by preparative HPLC. The correctness of the cage peptides and peptoids were confirmed with standard NMR and MS techniques. The purity of the peptides were determined with HPLC and were > 98%.

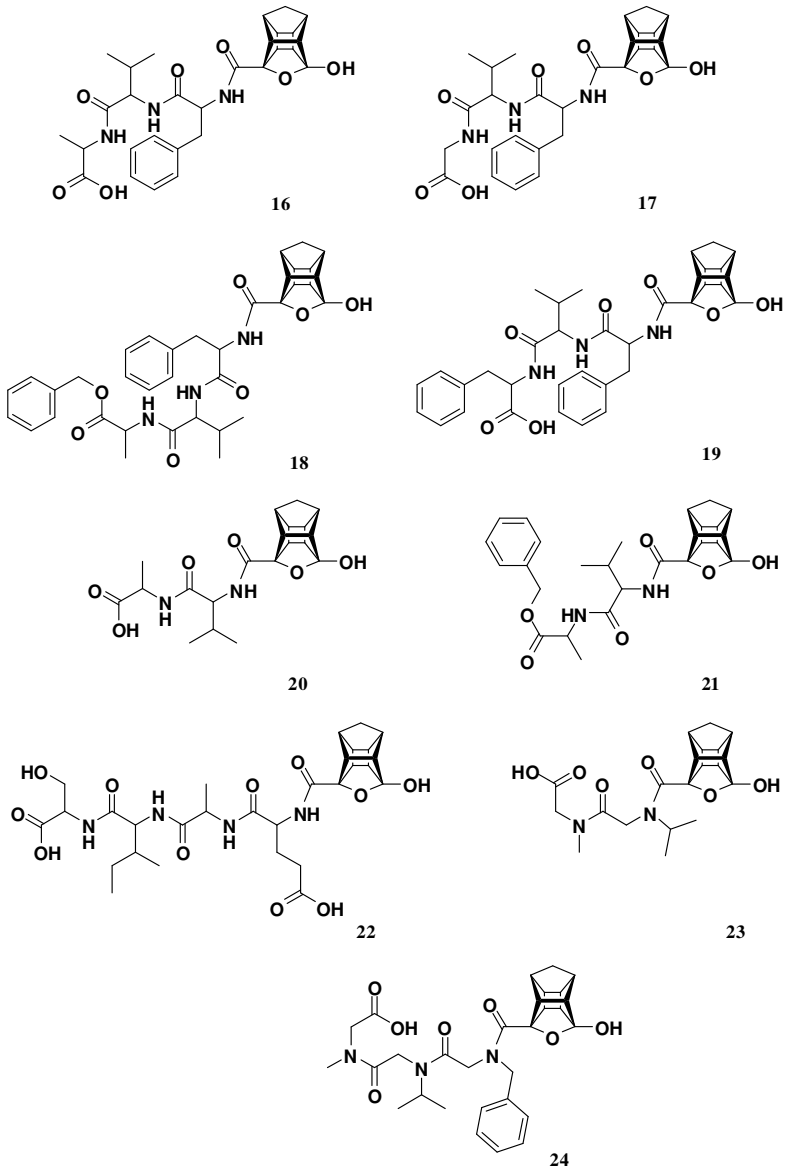


Figure 2. PCU-peptide and PCU-peptoid based inhibitors reported in this work.

The catalytic activity of the HIV-1 protease was monitored by following the hydrolysis of the chromogenic peptide substrate His-Lys-Ala-Arg-Val-Leu-Phe(*p*-NO_2)-Glu-Ala-Nle-Ser as described before.[61] The use of chromogenic substrates has been established for the determination of the HIV-1 protease inhibitors efficacy. The method is used to measure the protease activity by recording an decrease in absorbance as a result of the hydrolysis of a chromogenic substrate by the protease.[83] The inhibitory activities of the synthesized compounds

were tested using an UV spectrophotometric assay in order to determine the IC_{50} values for each compound. Biological results of the synthesized substrates are presented in Table 1.

Table 1. Inhibition of wild type C-SA HIV-1 protease by PCU derived peptides (**16-22**) and peptoids (**23-24**).

Entry	Sr.No.	Compound (R)	MF	% Yield	CLogP	Binding Energy[a]	IC_{50} (μM)[b]
1	**16**	HO-Ala-Val-Phe-	$C_{29}H_{35}N_3O_7$	42	0.792	-7.01	2
2	**17**	HO-Gly-Val-Phe-	$C_{28}H_{33}N_3O_7$	54	0.483	-7.74	1
3	**18**	Bn-O-Ala-Val-Phe-	$C_{36}H_{41}N_3O_7$	46	2.623	-5.53	>60
4	**19**	HO-Phe-Val-Phe-	$C_{35}H_{39}N_3O_7$	38	2.210	-5.02	>60
5	**20**	HO-Ala-Val-	$C_{20}H_{26}N_2O_6$	68	-0.367	-8.98	0.6
6	**21**	Bn-O-Ala-Val-	$C_{27}H_{32}N_2O_6$	51	1.462	-6.54	60
7	**22**	HO-Ser-Ilu-Ala-Glu-	$C_{29}H_{40}N_4O_{11}$	62	-2.310	-4.62	>60
8	**23 (Peptoid)**	R^1=Isopropyl R^2=Methyl	$C_{20}H_{26}N_2O_6$	57	0.138	-7.21	5
9	**24 (Peptoid)**	R^1=Benzyl R^2=Isopropyl R^3=Methyl	$C_{29}H_{35}N_3O_7$	43	1.795	-6.75	>60
10		**Lopinavir**					**0.025**
11		**Atazanavir**					**0.004**

[a] Calculated with Autodock. The diastereomer with the lowest binding energy is reported here. The 3D structures of these docked complexes are available as supplementary material.
[b] IC_{50} = 50% inhibition constant. Standard deviations (SD) for all the compounds were ≤ 2% of the reported IC_{50} values.

As we have previously reported,[34] it appears in general that the cage peptides exhibit very little cytotoxicity towards human cells.

According to previous literature reports, phenylalanine is the preferred substituent (P^1) at the S^1 subsite and it is found in 40% of *Gag-Pol* gene sequences of the virus.[52] Thus, we decided to synthesize sequences, with phenylalanine as the first amino acid (P^1 substituent), attached to the PCU cage. Even though the $S^2/S^{2'}$ and $S^3/S^{3'}$ pockets are also of a hydrophobic nature, both hydrophilic and hydrophobic side chains from the *Gag* and *Gag-Pol* polyproteins can occupy these sites. Based on literature and the HIV substrate sequence, we envisaged that either valine or alanine would be the best starting point in the peptide chain for binding at the $S^2/S^{2'}$ and $S^3/S^{3'}$ subsites.[33,34]

It was observed that compound **16** (entry 1) showed a moderate inhibition of 2 μM (IC_{50}). Compound **17** (entry 2), the activity increased by one-fold, where alanine was replaced by glycine at the proposed $P^3/P^{3'}$ position. When more hydrophobic bulky groups were incorporated as the $P^3/P^{3'}$ substituent, little or no inhibition activity was observed (compound **18**, entry 3 and **19**, entry 4). The reduced potency of these compounds could be as a result of the bulky hydrophobic residues occupying the $S^3/S^{3'}$ subsite. Since the $S^1/S^{1'}$ and $S^3/S^{3'}$ positions are located adjacent to one another in the *cis* configuration, incorporation of bulky substituents in both the subsites could alter the conformational arrangements, which consequently reduces their binding affinity.[6,53,56] This possibility will be discussed with the presentation of the docking results.

To investigate if the PCU-cage is competing with phenylalanine for the S^1 subsite, phenylalanine was removed from peptide **16** (entry 1), to give compound **20** (entry 5). This resulted in a higher binding affinity and preference for the active site with an inhibitory activity of 0.6 μM (IC_{50}). This is a 2 to 4 fold greater activity than compounds **16** (entry 1) and **17** (entry 2), which gave evidence that PCU cage is indeed occupying the S^1 subsite. To examine the role of a hydrophobic group at S^3 subsite we synthesized compound **21** (entry 6) containing a benzyl protecting group on the C terminal of the peptide sequence. This reduced binding affinity to the active site and had no inhibitory affect. This result is similar to that of compound **18** (entry 3).

Compounds **22** (entry 7) containing PCU cage with the natural HIV protease substrate (PCU-EAIS) showed inhibition activity greater than 60 μM (IC_{50}). Oligomers of *N*-substituted glycines are known to be resistant to proteolysis.[48-51] Keeping this in mind compounds **23** (entry 8) and **24** (entry 9) were synthesized to mimic the peptide sequence of compound **20** (entry 5) and **16** (entry 1) respectively. Compound **23** (entry 8) showed inhibition activity of 5 μM (IC_{50}), which is approximately 10-fold less active, when compared to **20** (IC_{50}, 0.6 μM). No inhibition activity was observed for compound **24** (entry 9) as compared to compound **16** (entry 1).

From the compounds in Table 1 the best inhibition activity was obtained for compound **20** (entry 5), which has potential activity when compared with Lopinavir (entry 10) but almost two orders of magnitude less active than Atazanavir (entry 11). To further validate these experimental results, docking studies and MD simulations were performed on the most active compound.

3.2. COMPUTATIONAL SIMULATIONS

Docking and MD simulations

To obtain a better understanding of the molecular and biological behavior of the synthesized PCU-peptide inhibitors, we embarked on a computational investigation involving docking and MD simulations. First, the PCU-peptides were subjected to docking studies to explore their binding pattern; second MD simulations of the docked inhibitor enzyme complex were performed to provide a molecular view of the dynamic behavior of such inhibitors inside the enzyme active site over a suitable time scale. Docking studies are used at different stages in drug discovery such as in the prediction of ligand-receptor complex structures and also to rank the ligand molecules based upon the binding energies of the corresponding ligand-enzyme complexes. Docking protocols aid in elucidation of the most energetically favourable binding mode of the ligand to the receptor. The objective of our docking study was to elucidate the potential interaction mode of the PCU-peptide derivatives with C-SA HIV PR and to see if a correlation between the binding energies and the observed IC_{50} results exists.

Docking of Ritonavir (co-crystallized inhibitor) with subtype B HIV-PR (PDB accession code 1HXW)[10] was previously performed[33,34] to evaluate the efficacy of Autodock for its use in docking experiments with this subtype as a target. Acceptable results were obtained. The interacting amino acids in the C-SA starting structure were conserved in the conformation of the enzyme after docking and the variation in orientation of these amino acids in the binding cavity was unnoticeable. Diastereomers of PCU peptides and peptoids were obtained from the coupling of the enantiopure short peptide or peptoids to the racemic PCU hydroxyl acid,[33,34,57,84] Figure 3 (**a**) and (**b**).

Docking routines of the new PCU-peptide inhibitors with the C-SA enzyme were performed. The binding energies of the docked diastereomeric PCU-peptide inhibitors have been tabulated (Table 2). The docking results are in reasonable agreement with the measured IC_{50} values.

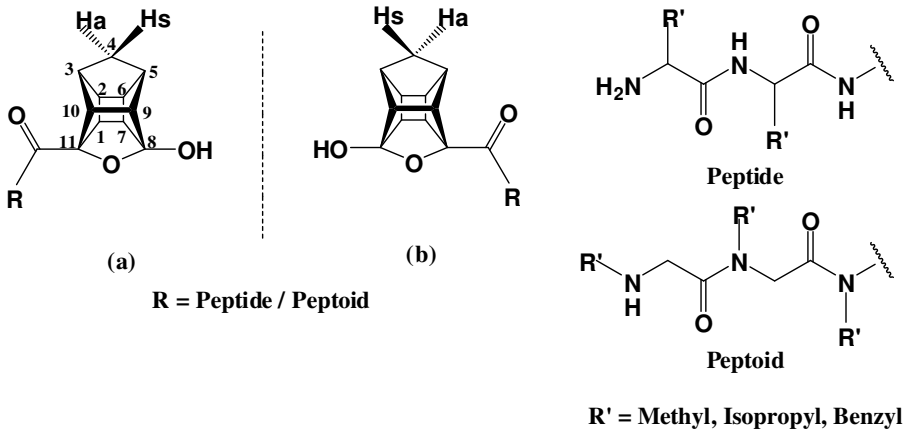


Figure 3. Stereochemistry of the generic cage diastereomers **(a)** and **(b)**.

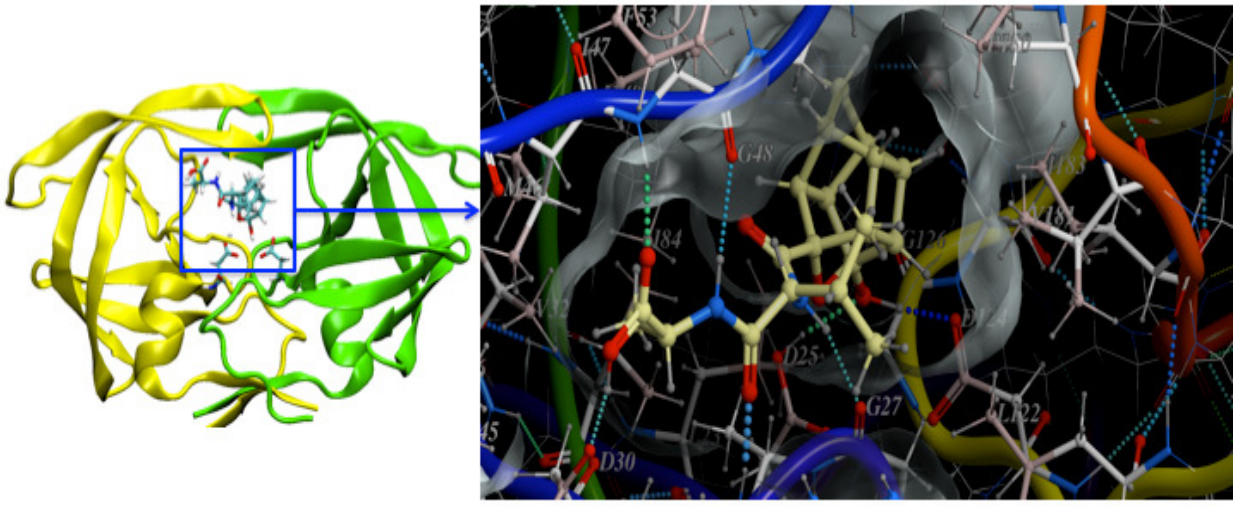
Table 2. Selected automated docking results for the synthesized PCU derived hydroxy ether peptide and peptoid inhibitors and comparison to the experimental HIV-PR inhibition results. (All of the docked inhibitor-enzyme complex results are available in PDB format and are provided with the supplementary material).

Compound[a]	**Binding energy**[b] **(kcal/mol)**	**IC_{50} / μM**
16-a[c]	-6.95	2
16-b[c]	-7.01	
17-a[c]	-7.10	1
17-b[c]	-7.74	
20-a[c]	-8.98	0.6
20-b[c]	-8.01	
21-a[c]	-6.54	60
21-b[c]	-5.92	
23-a[d]	-7.21	5
23-b[d]	-6.61	

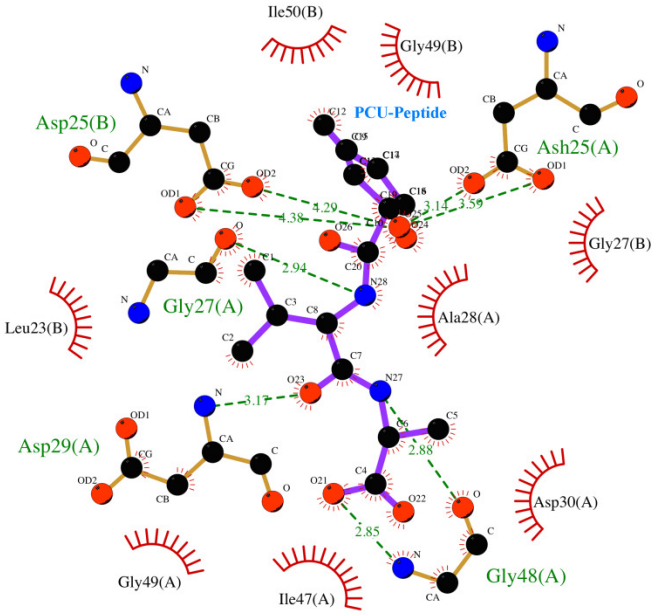
[a]The stereochemistry for the cage peptide is 8-(*R*)-11-(*R*)-PCU-peptide for the "a" diastereomer and 8-(*S*)-11-(*S*)-PCU-peptide for the "b" diastereomer.
[b]Energy calculated from Autodock. The 3D structures of the docked complexes are available as supplementary material.
[c]These two diastereomers were not separated and the IC_{50} values are for the mixture.
[d]These two enantiomers were not separated and the IC_{50} values are for the mixture.

In our previous reports[33,34] it was demonstrated that the chirality of the PCU cage is important for enhanced inhibition. The docking results in this study again confirm this observation.[33,34] This suggests that the diastereomers should ideally be separated when more promising cage peptide inhibitors are discovered.

Docking studies have clearly showed that conserved hydrogen bonds formed between the cage hydroxy ether moiety and Asp25/Asp25' of the dimeric catalytic triad residues, Asp25-Thr26-Gly27 (A/B chains) (Figure 3 and the supplementary information). Such hydrogen bonds anchor the cage skeleton to the $S^1/S^{1'}$ subsite causing the adjacent amino acids to shift to the $S^2/S^{2'}$and $S^3/S^{3'}$. This could explain the finding that less hydrophobic amino acids, such as valine, directly attached to the cage exhibited higher activity.



a) **b)**



c)

Figure 3. (a) Lowest energy docked structure for compound **20** with C-SA HIV-PR. **(b)** A closer view showing the binding mode of compound **20** inside the C-SA HIV-PR active site. Some atoms are deleted for clarity. The two PR monomers are coloured as the yellow and green ribbon. The inhibitor and the two Asp25 residues presented as coloured sticks. **(c)** Electrostatic and hydrogen bonding interaction between **20** and the nearby residues of C-SA HIV-PR active site. This plot was created with the Ligplot software.[85] The 3D presentations for computational results are available as PDB files with the supporting information.

Even though docking calculations cannot provide insight on the dynamic inhibitor-enzyme interactions, they provided us with a general picture of the most energetically favourable binding orientation of inhibitors to the enzyme. To obtain further insight into the dynamic changes of the docked inhibitors within the enzyme active site pocket over time, the lowest energy docked complex of the inhibitor **20** was subjected to unconstrained MD simulations (2 ns).

The MD simulations clearly showed that the inhibitor easily fits into the active enzyme pocket. To assess the quality of our MD simulations, energetic and structural properties are monitored along the entire 2 ns MD trajectory of the complex. Figure 4A presents the plot of the potential energy of the system as a function of time. The fluctuation of potential energy is less than 1000 kcal mol^{-1} over the course of the reaction.

The 2 ns averaged backbone RMSDs for the **20** enzyme complex is 1.16 Å and this is an indication that the generated MD trajectory of the complex is quite stable. The overall orientation of the complex is preserved and very little movement of the PCU cage hydroxy ether inside the enzyme pocket is observed during the course of the MD run. However, slight changes in the orientation in the cage inhibitor peptide side chain were observed. As evident from the analysis of the hydrogen bond interactions along the MD trajectories, the inhibitor is forming hydrogen bonds between the cage hydroxy ether moiety and at least one of the two Asp25 carboxyl groups (Figure 4C).

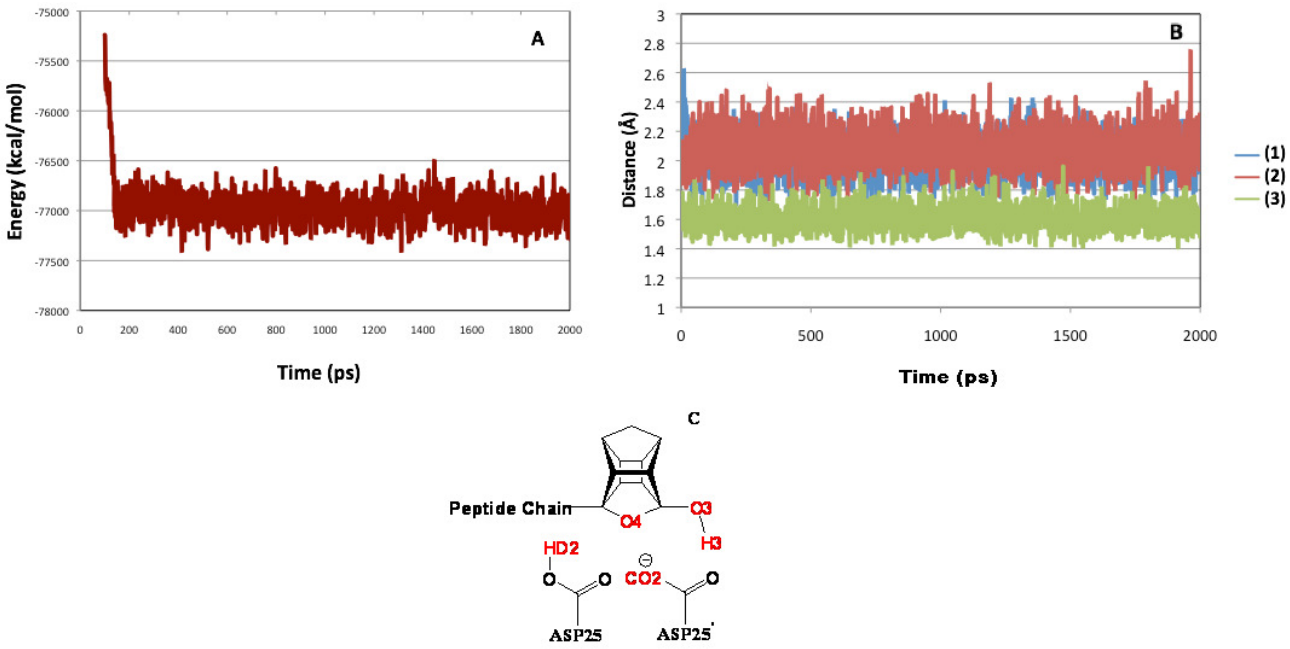


Figure 4. **(A)** The potential energy of **20** C-SA HIV-PR complex observed in MD simulation as a function of time. **(B)** Selected distances obtained from MD simulations: **(1)** cage(O4)-Asp25(HD2), **(2)** cage(O3)-Asp25(HD2) and **(3)** cage(H3)-Asp25'(OD2). **(C)** Labeling of some atoms involved in hydrogen bond formation.

4. CONCLUSION

Pentacycloundecane derived hydroxy acid peptides have emerged as a new class of irreversible non-scissile ether bridge type isosters. The results demonstrate the possibility of employing the cage peptides as potential protease inhibitors. Docking and MD simulations suggest that the cage or part thereof occupies the $S^1/S^{1'}$ subsite. The incorporation of bulky substituents such as

phenylalanine at the $S^2/S^{2'}$ and $S^3/S^{3'}$ subsites can alter the conformational arrangement of the inhibitor, which consequently reduces their binding affinity.

Further optimizations as well as in-depth structural and biological studies of the selected PCU derived-peptide/peptoid based protease inhibitors are the subject of ongoing investigations.

5. SUPPORTING INFORMATION:

All the spectra mentioned in the text are available as supporting information (Appendix 4).

6. ACKNOWLEDGMENT

This research was supported by NRF (SA) TG (GUN: 66319), KP (GUN: 69728), HGK and PIA (SA-Sweden bilateral grant), Aspen Pharmacare and University of KwaZulu-Natal for financial support. Y.S would like to thank the Carnegie Corporation of New York and the NRF for financial support. The authors thank Prof. Jürgen Schleucher (Umeå University, Sweden) and Mr. Dilip Jagjivan (UKZN, South Africa) for their assistance with the NMR experiments.

7. REFERENCES

1. In *UNAIDS (2012)*; Joint United Nations Programme on HIV/AIDS (UNAIDS), (**2012**); pp. 1-96.
2. Fitzgerald, P. M. D.; Springer, J. P. *Annu. Rev. Bioph. Biom.* **1991**, *20*, 299-320.
3. West, M. L.; Fairlie, D. P. *Trends Pharmacol. Sci.* **1995**, *16*, 67-75.
4. Kick, E. K.; Ellman, J. A. *J. Med. Chem.* **1995**, *38*, 1427-1430.
5. Martinez-Cajas, J. L.; Wainberg, M. A. *Antivir. Res.* **2007**, *76*, 203-221.
6. Babine, R. E.; Bender, S. L. *Chem. Rev.* **1997**, *97*, 1359-1472.
7. Wlodawer, A.; Vondrasek, J. *Annu. Rev. Bioph. Biom.* **1998**, *27*, 249-284.
8. Wu, X. Y.; Ohrngren, P.; Ekegren, J. K.; Unge, J.; Unge, T.; Wallberg, H.; Samuelsson, B.; Hallberg, A.; Larhed, M. *J. Med. Chem.* **2008**, *51*, 2586-2586.
9. Lee, T.; Laco, G. S.; Torbett, B. E.; Fox, H. S.; Lerner, D. L.; Elder, J. H.; Wong, C. H. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 939-944.
10. Velaquez-Campoy, A.; Vega, S.; Fleming, E.; Bacha, U.; Sayed, Y.; Dirr, H. W.; Freire, E. *AIDS Rev.* **2003**, *5*, 165-171.
11. Partaledis, J. A.; Yamaguchi, K.; Tisdale, M.; Blair, E. E.; Falcione, C.; Maschera, B.; Myers, R. E.; Pazhanisamy, S.; Futer, O.; Cullinan, A. B.; Stuver, C. M.; Byrn, R. A.; Livingston, D. J. *J. Virol.* **1995**, *69*, 5228-5235.
12. Staszewska, A.; Stefanowicz, P.; Szewczuk, Z. *Tetrahedron Lett.* **2005**, *46*, 5525-5528.
13. Brookes, K. B.; Hickmott, P. W.; Jutle, K. K.; Schreyer, C. A. *S. Afr. J. Chem.* **1992**, *45*, 8-11.
14. Geldenhuys, W. J.; Malan, S. F.; Bloomquist, J. R.; Marchand, A. P.; Van der Schyf, C. J. *Med. Res. Rev.* **2005**, *25*, 21-48.

15. Gerzon, K.; Kau, D. *J. Med. Chem.* **1967**, *10*, 189-199.
16. Ito, F. M.; Petroni, J. M.; de Lima, D. P.; Beatriz, A.; Marques, M. R.; de Moraes, M. O.; Costa-Lotufo, L. V.; Montnegro, R. C.; Magalhaes, H. I. F.; Pessoa, C. D. O. *Molecules* **2007**, *12*, 271-282.
17. Ranganathan, D.; Haridas, V.; Madhusudanan, K. P.; Roy, R.; Nagaraj, R.; John, G. B.; Sukhaswami, M. B. *Angew. Chem. Int. Ed.* **1996**, *35*, 1105-1107.
18. Grobler, E.; Grobler, A.; Van der Schyf, C. J.; Malan, S. F. *Bioorg. Med. Chem.* **2006**, *14*, 1176-1181.
19. Oliver, D. W.; Dekker, T. G.; Snyckers, F. O.; Fourie, T. G. *J. Med. Chem.* **1991**, *34*, 851-854.
20. Oliver, D. W.; Dekker, T. G.; Snyckers, F. O. *Eur. J. Med. Chem.* **1991**, *26*, 375-379.
21. Gokul, V.; Kruger, H. G.; Govender, T.; Fourie, L.; Power, T. D. *J. Mol. Struct. THEOCHEM* **2004**, *672*, 119-125.
22. Bisetty, K.; Gomez-Catalan, J.; Aleman, C.; Giralt, E.; Kruger, H. G.; Perez, J. J. *J. Pept. Sci.* **2004**, *10*, 274-284.
23. Bisetty, K.; Catalan, J. G.; Kruger, H. G.; Perez, J. J. *J. Mol. Struct. THEOCHEM* **2005**, *731*, 127- 137.
24. Bisetty, K.; Corcho, F. J.; Canto, J.; Kruger, H. G.; Perez, J. J. *J. Pept. Sci.* **2006**, *12*, 92-105.
25. Bisetty, K.; Govender, P.; Kruger, H. G. *Biopolymers* **2006**, *81*, 339-349.
26. Bisetty, K.; Corcho, F. J.; Canto, J.; Kruger, H. G.; Perez, J. J. *J. Mol. Struct. THEOCHEM* **2006**, *759*, 145-157.
27. Kruger, H. G.; Mdluli, P.; Power, T. D.; Raasch, T.; Singh, A. *J. Mol. Struct. THEOCHEM* **2006**, *771*, 165-170.
28. Bisetty, K.; Corcho, F. J.; Canto, J.; Kruger, H. G.; Perez, J. J. *J. Mol. Struct. THEOCHEM* **2006**, *770*, 221-228.
29. Singh, T.; Bisetty, K.; Kruger, H. G. *Trends Perspect. Mod. Comput. Sci.* **2006**, *7*, 511-514.
30. Bisetty, K.; Kruger, H. G.; Perez, J. J. *Mol. Simul.* **2007**, *33*, 1105-1108.
31. Albericio, F.; Arvidson, P. I.; Bisetty, K.; Giralt, E.; Govender, T.; Jali, S.; Kongsaeree, P.; Kruger, H. G.; Prabpai, S. *Chem. Biol. Drug Des.* **2008**, *71*, 125-130.
32. Singh, P.; Bisetty, K.; Govender, P.; Kruger, G. *S. Afr. J. Chem.* **2011**, *64*, 49-56.
33. Makatini, M. M.; Petzold, K.; Sriharsha, S. N.; Soliman, M. E. S.; Honarparvar, B.; Arvidsson, P. I.; Sayed, Y.; Govender, P.; Maguire, G. E. M.; Kruger, H. G.; Govender, T. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 2274-2277.
34. Makatini, M. M.; Petzold, K.; Sriharsha, S. N.; Ndlovu, N.; Soliman, M. E. S.; Honarparvar, B.; Parboosing, R.; Naidoo, A.; Arvidsson, P. I.; Sayed, Y.; Govender, P.; Maguire, G. E. M.; Kruger, H. G.; Govender, T. *Eur. J. Med. Chem.* **2011**, *46*, 3976-3985.
35. Kruger, H. G.; Mdluli, P. S. *Struct. Chem.* **2006**, *17*, 121-125.
36. Onajole, O. K.; Govender, T.; Makatini, M.; Kruger, H. G. *Struct. Chem.* **2008**, *19*, 719-726.
37. Onajole, O. K.; Govender, T.; Makatini, M.; Kruger, H. G. *Magn. Reson. Chem.* **2008**, *46*, 1007-1014.
38. Onajole, O. K.; Govender, P.; Govender, T.; Maguire, G. E. M.; Kruger, H. G. *Struct. Chem.* **2009**, *20*, 1067-1076.

39. Altaib, M. S.; Arvidsson, P. I.; Govender, T.; Maguire, G. E. M.; Makatini, M.; Onajole, O. K.; Kruger, H. G. *Magn. Reson. Chem.* **2010**, *48*, 435-442.
40. Onajole, O. K.; Makatini, M. M.; Govender, P.; Govender, T.; Maguire, G. E. M.; Kruger, H. G. *Magn. Reson. Chem.* **2010**, *48*, 249-255.
41. Onajole, O. K.; Govender, P.; Govender, T.; Maguire, G. E. M.; Kruger, H. G. *Struct. Chem.* **2010**, *21*, 1203-1209.
42. Onajole, O. K.; Belewa, X. V.; Coovadia, Y.; Govender, T.; Kruger, H. G.; Maguire, G. E. M.; Naidu, D.; Somai, B.; Singh, N.; Govender, P. *Med. Chem. Res.* **2011**, *20*, 1394-1401.
43. Onajole, O. K.; Govender, K.; Govender, P.; van Helden, P. D.; Kruger, H. G.; Maguire, G. E. M.; Muthusamy, K.; Pillay, M.; Wiid, I.; Govender, T. *Eur. J. Med. Chem.* **2009**, *44*, 4297-4305.
44. Onajole, O. K.; Govender, P.; van Helden, P. D.; Kruger, H. G.; Maguire, G. E. M.; Wiid, I.; Govender, T. *Eur. J. Med. Chem.* **2010**, *45*, 2075-2079.
45. Onajole, O. K.; Coovadia, Y.; Govender, T.; Kruger, H. G.; Maguire, G. E. M.; Naidu, D.; Singh, N.; Govender, P. *Chem. Biol. Drug Des.* **2011**, *77*, 295-299.
46. Adessi, C.; Soto, C. *Curr. Med. Chem.* **2002**, *9*, 963-978.
47. Hruby, V. J.; Balse, P. M. *Curr. Med. Chem.* **2000**, *7*, 945-970.
48. Zuckermann, R. N.; Kerr, J. M.; Kent, S. B. H.; Moos, W. H. *J. Am. Chem. Soc.* **1992**, *114*, 10646-10647.
49. Miller, S. M.; Simon, R. J.; Ng, S.; Zuckermann, R. N.; Kerr, J. M.; Moos, W. H. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2657-2662.
50. Miller, S. M.; Simon, R. J.; Ng, S.; Zuckermann, R. N.; Kerr, J. M.; Moos, W. H. *Drug Dev. Res.* **1995**, *35*, 20-32.
51. Simon, R. J.; Kania, R. S.; Zuckermann, R. N.; Huebner, V. D.; Jewell, D. A.; Banville, S.; Ng, S.; Wang, L.; Rosenberg, S.; Marlowe, C. K.; Spellmeyer, D. C.; Tan, R. Y.; Frankel, A. D.; Santi, D. V.; Cohen, F. E.; Bartlett, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9367-9371.
52. Clemente, J. C.; Coman, R. M.; Thiaville, M. M.; Janka, L. K.; Jeung, J. A.; Nukoolkarn, S.; Govindasamy, L.; Agbandje-McKenna, M.; McKenna, R.; Leelamanit, W.; Goodenow, M. M.; Dunn, B. M. *Biochemistry* **2006**, *45*, 5468-5477.
53. Brik, A.; Wong, C. H. *Org. Biomol. Chem.* **2003**, *1*, 5-14.
54. 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. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2484-2488.
55. 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. *J. Med. Chem.* **2000**, *43*, 3495-3504.
56. Ridky, T. W.; Cameron, C. E.; Cameron, J.; Leis, J.; Copeland, T.; Wlodawer, A.; Weber, T. I.; Harrison, R. W. *J. Biol. Chem.* **1996**, *271*, 4709-4717.
57. Martins, F. J. C.; Viljoen, A. M.; Kruger, H. G.; Wessels, P. L. *Tetrahedron* **1993**, *49*, 6527-6532.
58. Ghosh, A. K.; Thompson, W. J.; McKee, S. P.; Duong, T. T.; Lyle, T. A.; Chen, J. C.; Darke, P. L.; Zugay, J. A.; Emini, E. A. *J. Med. Chem.* **1993**, *36*, 292-294.

59. Ghosh, A. K.; Kincaid, J. F.; Cho, W. H.; Walters, D. E.; Krishnan, K.; Hussain, K. A.; Koo, Y.; Cho, H.; Rudall, C.; Holland, L.; Buthod, J. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 687-690.
60. Ghosh, A. K.; Leshchenko, S.; Noetzel, M. *J. Org. Chem.* **2004**, *69*, 7822-7829.
61. Mosebi, S.; Morris, L.; Dirr, H. W.; Sayed, Y. *J. Virol.* **2008**, *82*, 11476-11479.
62. Yu, H. M.; Chen, S. T.; Wang, K. T. *J. Org. Chem.* **1992**, *57*, 4781-4784.
63. Yamamoto, T.; Nair, P.; Jacobsen, N. E.; Davis, P.; Ma, S.-w.; Navratilova, E.; Moye, S.; Lai, J.; Yamamura, H. I.; Vanderah, T. W.; Porreca, F.; Hruby, V. J. *J. Med. Chem.* **2008**, *51*, 6334-6347.
64. Avogadro: An open-source molecular builder and visualization tool. Version 1.0.0. Http://avogadro.Openmolecules.Net/
65. Soares, T. A.; Goodsell, D. S.; Ferreira, R.; Olson, A. J.; Briggs, J. M. *J. Mol. Recognit.* **2000**, *13*, 146-156.
66. Frimurer, T. M.; Peters, G. H.; Iversen, L. F.; Andersen, H. S.; Moller, N. P. H.; Olsen, O. H. *Biophys. J.* **2003**, *84*, 2273-2281.
67. Locuson, C. W.; Gannett, P. M.; Ayscue, R.; Tracy, T. S. *J. Med. Chem.* **2007**, *50*, 1158-1165.
68. Smith, R.; Brereton, I. M.; Chai, R. Y.; Kent, S. B. H. *Nat. Struct. Biol.* **1996**, *3*, 946-950.
69. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639-1662.
70. Sanner, M. F. *J. Mol. Graphics Modell.* **1999**, *17*, 57-61.
71. Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. *J. Comput. Chem.* **2007**, *28*, 1145-1152.
72. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Modell.* **2006**, *25*, 247-260.
73. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668-1688.
74. Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157-1174.
75. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M.; Kollman, P. A. *J. Comput. Chem.* **2003**, *24*, 1999–2012.
76. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
77. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577-8593.
78. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
79. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
80. Cookson, R. C.; Crundwell, E.; Hill, R. R.; Hudec, J. *J. Chem. Soc.* **1964**, 3062-3075.
81. Marchand, A. P.; Allen, R. W. *J. Org. Chem.* **1974**, *39*, 1596-1596.
82. Marchand, A. P. *Adv. Theo. Interesting Mol.* **1989**, *1*, 357-399.
83. Gutierrez, O. A.; Salas, E.; Hernandez, Y.; Lissi, E. A.; Castrillo, G.; Reyes, O.; Garay, H.; Aguilar, A.; Garcia, B.; Otero, A.; Chavez, M. A.; Duarte, C. A. *Anal. Biochem.* **2002**, *307*, 18-24.

84. Martins, F. J. C.; Viljoen, A. M.; Kruger, H. G.; Joubert, J. A. *Tetrahedron* **1993**, *49*, 9573-80.
85. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. *Protein Eng.* **1995**, *8*, 127-134.

CHAPTER 5

SYNTHESIS AND NMR ELUCIDATION OF PENTACYCLOUNDECANE DERIVED HYDROXY ACID PEPTIDES AS POTENTIAL ANTI-HIV-1 AGENTS

Rajshekhar Karpoormath,[a,b] Fernando Albericio,[a,c,d,e] Thavendran Govender,[b] Glenn E. M. Maguire,[a] and Hendrik G. Kruger[a*]

[a]School of Chemistry, University of KwaZulu-Natal, Durban 4001, South Africa
[b]School of Pharmaceutical Sciences, University of KwaZulu-Natal, Durban 4001, South Africa
[c]Institute for research in Biomedicine, Barcelona Science Park, Baldiri Reixac 10, 08028-Barcelona, Spain
[d]CIBER-BBN, Networking Centre on Bioengineering, Biomaterials and Nanomedicine, Barcelona Science Park, Baldiri Reixac 10, 08028-Barcelona, Spain
[e]Department of Organic Chemistry, University of Barcelona, Martí i Franque´s 1-11, 08028-Barcelona, Spain
Corresponding author: *Email: kruger@ukzn.ac.za, Fax: +27-31-2603091

ABTRACT

The synthesis and NMR elucidation of eight novel peptides incorporating the pentacycloundecane (PCU) derived hydroxy acid are reported. The PCU cage amino acids were synthesized as racemates and the incorporation of the PCU derived hydroxy acid with natural (S)-amino acids produced inseparable diastereomeric peptides. A series of overlapping signals from the cage and that of the peptide side chain was observed in the ^{1}H and ^{13}C NMR spectra, complicating the elucidation thereof. Two-dimensional NMR techniques proved to be a very useful tool in overcoming these difficulties. These compounds are potential HIV protease inhibitors.

Keywords: 2D NMR, Pentacyclo undecane, PCU cage peptides.

1. INTRODUCTION

Polycyclic cage compounds have been studied extensively over the past few decades.[1-24] South African authors have contributed considerably towards utilizing NMR spectroscopy as a tool to

determine the structures of these cage compounds, several of which have found application in medicinal chemistry.[10,15,19,24-44] The assignments of ^{1}H spectra of the cage compounds becomes more challenging due to broad unresolved overlapping resonances as a result of through-space effects, geminal and vicinal proton-proton coupling and long range proton-proton interactions. Our research group has extensively used two-dimensional (2D) NMR spectroscopic techniques to study the structure of the cage skeleton in relation to its side 'arms'.[19,24,36-44] As part of our ongoing research program the synthesis of eight PCU derived hydroxy acid peptides (**1-8**) was attempted (Figure 1).

The use of cage compounds in drug design has found several applications in the pharmaceutical drug discovery research.[11,15,16,19-24,43-47] Cage based compounds have been reported to improve drug lipophilicity; it also serves as a transport medium in facilitating drug delivery through cell membranes such as the blood brain barrier (BBB) and the central nervous system (CNS).[2,4,23,48] The cage compounds increase the duration of drug action by retarding the action of drug metabolism.[13,49]

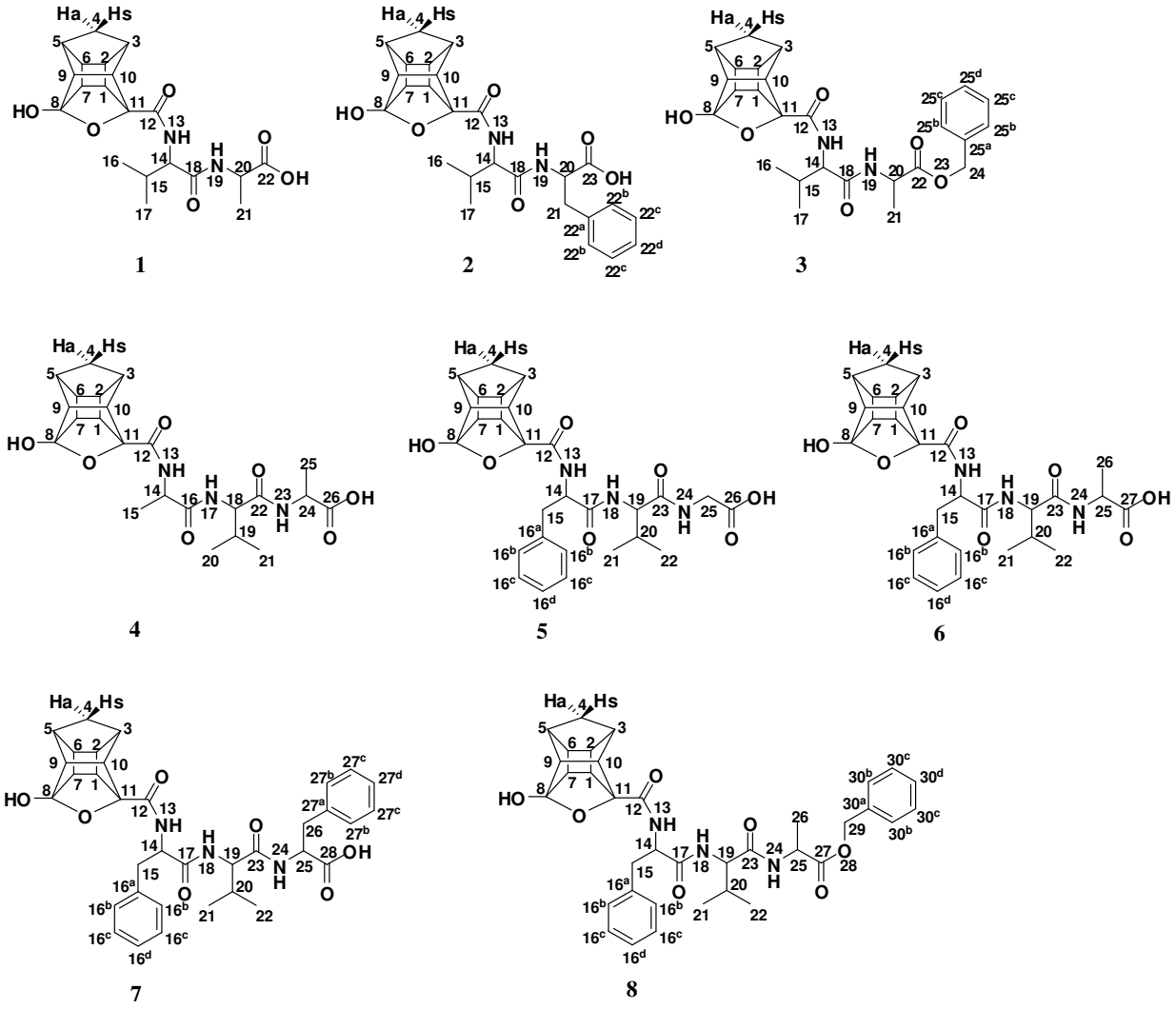


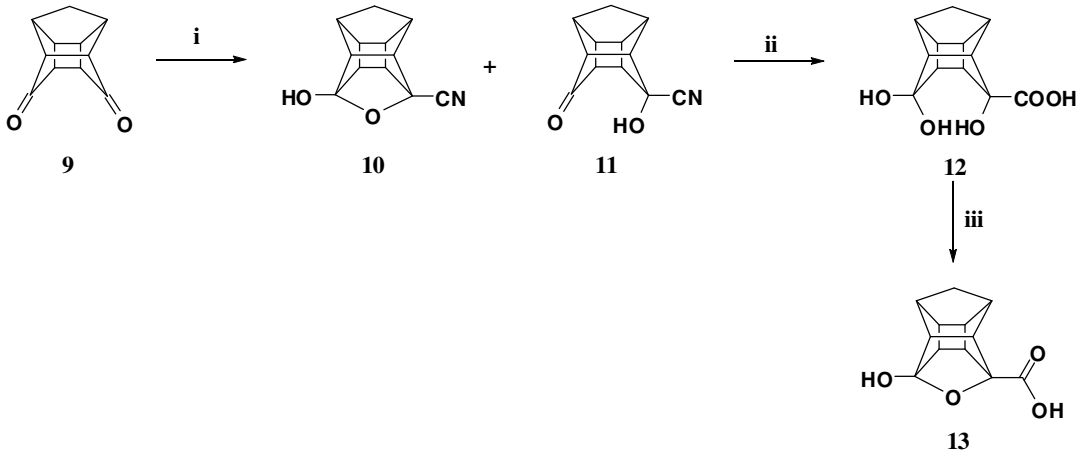
Figure 1. Structures of compounds **1-8**

In this study the PCU hydroxy acid[32] incorporated into short peptides was used to mimic an irreversible non-scissile ether bridged type isostere as potential HIV-1 protease inhibitors analogues to a non-cage ether reported by Ghosh et. al.[50] Compounds **1**, **5** and **6** (Fig.1) gave promising *in vitro* activity against HIV-1 wild type C-SA protease enzyme with IC_{50} values ranging from 0.6 – 2 μM.[22]

2. SYNTHESIS

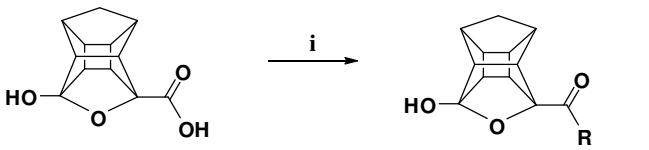
The pentacycloundecane peptide derivatives **1-8** were synthesized from the cage ether **13**, with Cookson's dione **9**[3,7,51] as the starting material. The latter can be easily obtained from the

photocyclisation of the Diels-Alder adduct of cyclopentadiene and *p*-benzoquinone. Treatment of the dione **9** with aqueous sodium cyanide stirred in a mixture of water and acetic acid resulted in a mixture of cyanohydrins (**10** and **11**).[32] This cyanohydrin mixture, when treated with concentrated hydrochloric acid, gave the tri-hydroxy carboxylic acid **12**.[32] The resulting tri-hydroxy carboxylic acid **12** was heated to reflux in 1,4-dioxane to give the desired hydroxy acid **13** as a racemate. These reactions are outlined in Scheme 1.



Scheme 1. Synthesis of 5-Hydroxy-4-oxahexacyclo[5.4.1.$0^{2,6}$.$0^{3,10}$.$0^{5,9}$.$0^{8,11}$]dodecane -3-carboxylic acid **13**; **i =** NaCN, CH_3COOH, H_2O, 0 °C, **ii =** conc. HCL, reflux, **iii =** 1,4-dioxane, reflux.

Compounds **1-8** were obtained by coupling the corresponding short peptides to **13** in the presence of coupling reagents HATU, DIPEA in DMF at ambient temperatures as shown in Scheme 2. It should be noted that the resulting cage peptides are mixtures of two diastereomers since the PCU hydroxy cage acid was synthesized as a racemate.



Scheme 2. Synthesis of 5-Hydroxy-4- oxahexacyclo[5.4.1.$0^{2,6}$.$0^{3,10}$.$0^{5,9}$.$0^{8,11}$]dodecane peptide analogs. **i =** HATU, DIPEA, DMF, rt.

The products **1** to **8** (Figure 1) were isolated and purified as described in experimental section. It was not possible to separate the diastereomeric mixtures by preparative HPLC. The correctness

of the cage peptides were confirmed with standard NMR and MS techniques. The synthesized cage peptides were obtained in high purity (> 98%), which was determined with HPLC.

3. RESULTS AND DISCUSSION

The PCU cage structures show complex overlapping of the protons as the PCU cage derivatives (**1-8**) are diastereomeric, asymmetric and non-meso compounds which further complicates the structural elucidation of these compounds. Due to the diastereomeric nature of these compounds, all the ^{13}C NMR signals of these compounds (**1-8**) are split. It is known from the literature that the proton NMR spectrum of PCU cage compounds usually show a pair of doublets (geminal protons) with an AB spin resonance at approximately 1.5 and 1.8 ppm, having a coupling constant of about ≈ 10 Hz. The assignments for all PCU cage compounds **1-8** are presented in Tables 1 and 2.

The geminal bridge methylene proton (H-4a) for the PCU hydroxy peptide **1** register as a doublet (AB system) at 1.47 ppm (H-4a) with a coupling constant of 10.4 Hz, but in the case of H-4s a pair of doublets is observed resonating at 1.82 ppm and 1.84 ppm with a *J* value of 6.0 Hz. Similar observations are recorded for compounds **5** and **6**, which can perhaps be attributed to a side chain conformational effect. For compound **3** the H-4s signal at 1.83 ppm (J = 9.4 Hz) appears as a triplet due to coalescence of the pair of doublets. In PCU cage compounds proton H-4a (H-4s) normally exhibits COSY and NOESY/ROESY interactions with H-3/5 and H-2/6 (H-3/5 and H-9/10) respectively, which has proven to be a convenient handle in solving the NMR spectra of these compounds. A similar interaction is also observed between H-2/6 and H-1/7. HMBC, NOESY and ROESY techniques are useful methods to determine the position of the "arms" /substituents in relation to the cage moiety especially when the cage is asymmetric.

In particular both H-4a (1.47 ppm) and H-4s (1.83 ppm) of **1** display COSY correlations to the resonances at 2.52 ppm and 2.46 ppm. These resonances were assigned to H-3/H-5. Since the molecule is asymmetric, the specific assignment is still pending. In addition the H-4a resonance shows a NOESY correlation to 2.66 ppm and 2.68 ppm (H-2/H-6) and the H-4s resonance to 2.52 ppm and 2.71 ppm (H-9/H-10). In order to distinguish between the proton resonances of H-9 and H-10, a HMBC correlation between carbonyl resonance C-12 (170.5 ppm) and the protons resonance at 2.71 ppm indicated that this resonance is associated to H-10. The position of H-10

was further confirmed as it displayed COSY correlation to a proton signal at 2.52 ppm, which was assigned to H-9. The carbonyl resonance of C-12 also displays a HMBC correlation to H-10 and H-1 (2.75 ppm). The latter in turn showed a COSY correlation to H-2 (2.66 ppm). A COSY correlation between the resonance at 2.52 ppm (also H-9) and the signal of H-10 as well as H-2 was observed, indicating that the proton signal of H-3 and H-9 are overlapping. This was further confirmed by the HMBC spectrum, wherein the carbon peaks of C-3 (46.1 ppm) and C-9 (56.2 ppm) displayed correlation to the same proton signal resonating at 2.52 ppm (H-3). Since now the position of H-3 (2.52 ppm) is confirmed, therefore the resonance at 2.46 ppm was attributed to H-5. In addition to the COSY correlation between H-1 and H-2, a further correlation between H-1 and 2.48 ppm allowed this resonance to be assigned to H-7. In the ^{13}C NMR spectrum the quaternary carbon signals appearing at 118.2 ppm and 89.75 ppm were assigned to C-8 and C-11 based on their deshielding effect and chemical environment. All the remaining cage carbon signals were obtained from the HSQC spectrum.

The next section describes the NMR correlations for the rest of the peptide sequence attached to the PCU cage. The methyl (alanine) and isopropyl methyl (valine) carbons register at 17.1 ppm, 17.8 ppm and 19.1 ppm and were identified from the APT ^{13}C NMR spectrum. The carbon signal at 17.1 pm show HSQC correlations to the ^{1}H NMR signals appearing as a pair doublets at 1.26 ppm and 1.27 ppm (J = 5.20 Hz), similarly the isopropyl methyl carbon signals 17.8 ppm and 19.1 ppm display correlations to a two pairs of doublets at 0.79 / 0.81 ppm and 0.86 / 0.88 ppm with a coupling constant approximately ≈ 6.7 Hz. The isopropyl methyl groups are also split due to the asymmetric nature of the environment close by. Hence these side chain methyl protons were assigned to H-21 (1.26 ppm / 1.27 ppm), H-16 (0.79 ppm / 0.81 ppm) and H-17 (0.86 ppm / 0.88 ppm) respectively. The H-15 methine signal at 1.97 ppm was identified by a COSY correlation to the isopropyl methyl protons H-16 and H-17. This H-15 resonance in turn shows a COSY interaction with the chiral H-14 methine proton signal appearing as a multiplet at 4.25 ppm. The chiral H-14 proton further displays COSY correlation to H-13 (NH) a doublet signal resonating at 6.91 ppm (J = 4.1 Hz). For this resonance a pair of doublets is seen because of the diastereomeric nature of the compound, which is further validated by the APT ^{13}C NMR spectrum where the carbonyl carbon (C-12) is registered as a split signal. The H-15 resonance

shows a HMBC correlation to a carbonyl carbon registered at 170.2 ppm. This was then assigned to C-18.

The H-21 methyl signal shows a correlation with the chiral methine proton H-20 appearing as a quartet signal at 4.19 ppm (J = 7.40 Hz). H-20 further shows a COSY correlation to a triplet proton signal at 8.44 ppm (H-19, J = 8.04 Hz). This pair of doublets coalesces into a triplet signal at 8.44 ppm. All the compounds **1**-**8** give perfect pairs of doublets for the corresponding NH proton signals. From literature it is known that NH proton signals in linear peptides usually appear as doublets.[52] Since these PCU cage peptides are a mixture of diastereomers, the amide protons appear as a pair of doublets. The remaining carbonyl peak registers at 173.8 ppm in the ^{13}C NMR spectrum and was assigned to C-22. This was further confirmed by HMBC spectra wherein H-21 displays correlation to the terminal carbonyl carbon C-22. All other assignments were further verified utilizing HSQC and HMBC spectra. ROESY correlations were observed between H-13 of the peptide chain and the PCU cage protons (H-1, H-2 and H-10). Further to this, the isopropyl methyl protons H-16/H-17 displayed ROESY correlations to H-1, H-4a and H-4s. ROESY experiments were carried out for compounds **1**, **4**, **5**, **6** and **7** to obtain better through space correlations between the PCU cage protons and the peptide chain since normal NOESY experiments did not display clear correlations. For the remainder of the compounds **2**-**8**, the same procedure was used for the elucidation of the cage signals. This discussion for the remaining compounds will be omitted for the sake of brevity. For the peptide side chains only differences to previous compounds will be highlighted. The assignments for compound **1** are summarized in Tables 1 and 2.

Table 1. ^{1}H NMR data for PCU hydroxy peptides **1**-**8**.

Atom	**1** $\delta^{1}H^{\ddagger,*}$	**2** $\delta^{1}H^{\dagger,*}$	**3** $\delta^{1}H^{\ddagger,*}$	**4** $\delta^{1}H^{\dagger,*}$	**5** $\delta^{1}H^{\ddagger,*}$	**6** $\delta^{1}H^{\ddagger,*}$	**7** $\delta^{1}H^{\dagger,*}$	**8** $\delta^{1}H^{\dagger,*}$
1	2.75	2.77	2.72	2.76	2.73	2.73	2.71	2.72
2	2.66	2.65	2.58	2.63	2.59	2.59	2.57	2.58
3	2.52	2.51	2.45	2.47	2.44	2.46	2.43	2.45
4a	1.47 d, *J*=10.4 Hz	1.47 d, *J*=10.1 Hz	1.47 d, *J*=9.3 Hz	1.48 d, *J*=10.2 Hz	1.44 d, *J*=10.2 Hz	1.45 d, *J*=10.1 Hz	1.44 d, *J*=10.1 Hz	1.44 d, *J*=10.1 Hz
4s	1.82/1.84 d, *J*=6.0 Hz, $^{(\delta)}$H-4s/4sb	1.83 d, *J*=10.2 Hz	1.82 t, *J*=9.4 Hz, $^{(€)}$H-4s	1.83 d, *J*=10.2 Hz	1.79/1.80 d, *J*=6.0 Hz,	1.79/1.80 d, *J*=6.1 Hz,	1.78 d, *J*=10.3 Hz	1.79 d, *J*=10.2 Hz
5	2.46	2.49	2.44	2.47	2.45	2.45	2.43	2.44
6	2.68	2.66	2.64	2.67	2.65	2.67	2.66	2.64
7	2.48	2.51	2.47	2.50	2.44	2.42	2.46	2.45
8	-	-	-	-	-	-	-	-
9	2.52	2.55	2.49	2.51	2.47	2.48	2.45	2.50
10	2.71	2.69	2.66	2.68	2.65	2.65	2.66	2.64
11	-	-	-	-	-	-	-	-
12	-	-	-	-	-	-	-	-
13	6.91/6.93 d, *J*=3.1 Hz, $^{(\delta)}$NH-13a/13b	6.95 d, *J*=9.3 Hz	6.88/6.90 d, *J*=5.7 Hz, $^{(\delta)}$NH-13a/13b	7.48/7.50 d, *J*=3.2 Hz, $^{(\delta)}$NH-13a/13b	7.40 t, *J*=9.0 Hz, $^{(€)}$NH-13	7.44 t, *J*=9.0 Hz, $^{(€)}$NH-13	7.38/7.41 d, *J*=5.5 Hz, $^{(\delta)}$NH-13a/13b	7.44/7.46 d, *J*=6.3 Hz, $^{(\delta)}$NH-13a/13b
14	4.25	4.16	4.24	4.37	4.62	4.60	4.61	4.60
15	1.97	1.92	1.93	1.22	2.99	2.99	2.90/2.98	2.97

†400 MHz
‡600 MHz
*Solvent $(CD_3)_2SO$
$^{(\delta)}$ Two separate resonances potentially due to a side chain conformational effect.
$^{(€)}$ Due to coalescence of the pair of doublets, the signal appears as a triplet.

For compound **2**, the aromatic protons in the ^{1}H NMR spectra are not well resolved but the carbons for aromatic region in the ^{13}C NMR spectra are very well separated. In the HMBC spectrum, the chiral methine proton H-20 (4.44 ppm) displays a correlation to the quaternary carbon C-22^{a} (137.4 ppm) of the phenyl ring and the diastereomeric methylene protons H-21^{a} / H-21^{b} (2.85 ppm / 3.07 ppm) show correlations to the aromatic carbon C-22^{b} (128.9 ppm). Carbon signals for C-22^{c} (128.1 ppm) and C-22^{d} (126.3) were assigned[53,54] based on the intensity of the peaks.

Table 1. Continued.

Atom	1 $\delta^1H^{‡,*}$	2 $\delta^1H^{†,*}$	3 $\delta^1H^{‡,*}$	4 $\delta^1H^{†,*}$	5 $\delta^1H^{‡,*}$	6 $\delta^1H^{‡,*}$	7 $\delta^1H^{†,*}$	8 $\delta^1H^{†,*}$
16/16[b-d]	0.79/0.81 d, *J*=6.7 Hz,	0.73-0.77	0.74/0.77 d, *J*=6.7 Hz,	-	7.08-7.23	7.11-7.24	7.08-7.16-	7.10-7.22
17	0.86/0.88 d, *J*=6.7 Hz,	0.73-0.77	0.82/0.84 d, *J*=6.7 Hz,	7.81/7.87 d, *J*=9.0 Hz,	-	-	-	-
18	-	-	-	4.21	8.01/8.05 d, *J*=9.0Hz,	7.97 d/8.02, *J*=9.0 Hz,	7.93/7.98 d, *J*=9.1 Hz,	8.01/8.06 d, *J*=9.0 Hz,
19	8.44 t, *J*=8.0 Hz	8.40 t, *J*=7.1 Hz	8.60/8.61 d, *J*=6.8 Hz,	1.94	4.24	4.23	4.25	4.23
20	4.19 q, *J*=7.4 Hz	4.44	4.33	0.82/0.84 d, *J*=3.9 Hz,	1.97	1.96	1.96	1.92
21	1.26/1.27 d, *J*=5.2 Hz,	2.85/3.07	1.30/1.31 d, *J*=5.2 Hz,	0.87/0.88 d, *J*=3.4 Hz,	0.87	0.86 t, *J*=6.0 Hz,	0.83	0.83
22/22[b-d]	-	7.17-7.27	-	-	0.87	0.89/0.90 d, *J*=4.6 Hz,	0.83	0.83
23	-	-	-	8.25/8.27 d, *J*=3.4 Hz,	-	-	-	-
24	-	-	5.10 d, *J*=1.3 Hz	4.21	8.36	8.36 t, *J*=6.3 Hz	8.33 d, *J*=7.2 Hz	8.55/8.57 d, *J*=4.6 Hz,
25/25[b-d]	-	-	7.30-7.37	1.25	3.78	4.20	4.45	4.35
26	-	-	-	-	-	1.29 d, *J*=7.2 Hz	2.92/3.05	1.32 d, *J*=7.2 Hz
27/27[b-d]	-		-	-	-	-	7.17-7.25	-
28	-		-	-	-	-	-	-
29	-		-	-	-	-	-	5.11
30/30[b-d]	-		-	-	-	-	-	7.33-7.37

[†]400 MHz
[‡]600 MHz
[*]Solvent $(CD_3)_2SO$
(δ) Two separate resonances potentially due to a side chain conformational effect.
(€) Due to coalescence of the pair of doublets, the signal appears as a triplet.
(Ω) The methyl protons appear as a pair of doublets of doublets.

The carbonyl signal at 170.4 ppm for compound **2** shows HMBC correlations to various proton signals (H-20, H-19, H-15, H-14, H-13, H-10 and H-1), which confirms the overlapping of the carbonyl carbon signals of C-18 and C-12. The methodology to elucidate the remaining signals of the peptide side chain was similar to that described for compound **1**. NOESY correlations were observed between H-13 and H-1. The aromatic protons (H-22) displayed through space correlations to H-9 and H-6 of the cage protons. The assignments of the peaks were further confirmed by HMBC and HSQC spectra. These assignments are presented in Tables 1 and 2.

In the ^1H NMR spectra of compound **3**, a doublet (diastereomeric) methylene proton signal (H-24) resonating at 5.10 ppm (J = 1.38 Hz) and integrating to two protons shows HMBC correlation to C-22 (172.1 ppm). H-24 further shows HMBC correlations to a quaternary and

aromatic carbon signals registered at 138.8 ppm and 127.8 ppm, which were assigned to C-25[a] and C-25[b] respectively. These assignments were further confirmed from the ^{13}C-APT spectra. Based on the intensity of the peaks[53,54] the carbon signals registered at 128.0 ppm and 128.4 ppm were assigned to C-25[c] and C-25[d]. Through space interactions were observed in the NOESY spectrum between the PCU cage protons (H-4a, H-1, H-2 and H-10) and the methyl protons of the peptide chain (H-16 and H-17). The ^{1}H and ^{13}C NMR assignments are shown in Tables 1 and 2.

Table 2. ^{13}C NMR chemical shifts of compounds **1-8**.

Atom	**1** $\delta^{13}C^{\ddagger,*}$	**2** $\delta^{13}C^{\dagger,*}$	**3** $\delta^{13}C^{\ddagger,*}$	**4** $\delta^{13}C^{\dagger,*}$	**5** $\delta^{13}C^{\ddagger,*}$	**6** $\delta^{13}C^{\ddagger,*}$	**7** $\delta^{13}C^{\dagger,*}$	**8** $\delta^{13}C^{\dagger,*}$
1	48.9 (CH)	48.7 (CH)	48.2 (CH)	48.4 (CH)	48.1 (CH)	48.2 (CH)	48.1 (CH)	48.2 (CH)
2	42.2 (CH)	42.1 (CH)	42.2 (CH)	42.2 (CH)	42.1 (CH)	42.1 (CH)	42.0 (CH)	42.1 (CH)
3	46.1 (CH)	46.0 (CH)	46.2 (CH)	46.1 (CH)	46.2 (CH)	46.1 (CH)	46.0 (CH)	46.1 (CH)
4a	42.8 (CH_2)	42.8 (CH_2)	42.8 (CH_2)	42.8 (CH_2)	42.8 (CH_2)	42.8 (CH_2)	42.8 (CH_2)	42.8 (CH_2)
4s	42.8 (CH_2)	42.8 (CH_2)	42.8 (CH_2)	42.8 (CH_2)	42.8 (CH_2)	42.8 (CH_2)	42.8 (CH_2)	42.8 (CH_2)
5	42.7 (CH)	42.7 (CH)	42.8 (CH)	42.7 (CH)	42.7 (CH)	42.7 (CH)	42.7 (CH)	42.8 (CH)
6	41.1 (CH)	40.9 (CH)	41.0 (CH)	41.0 (CH)	40.9 (CH)	40.9 (CH)	41.0 (CH)	41.0 (CH)
7	46.3 (CH)	46.2 (CH)	46.4 (CH)	46.3 (CH)	46.2 (CH)	46.3 (CH)	46.3 (CH)	46.4 (CH)
8	118.2 (C)	118.1 (C)	118.2 (C)	118.2 (C)	118.1 (C)	118.2 (C)	118.1 (C)	118.2 (C)
9	56.2 (CH)	56.4 (CH)	56.2 (CH)	56.2 (CH)	56.1 (CH)	56.1 (CH)	56.1 (CH)	56.2 (CH)
10	58.6 (CH)	58.5 (CH)	57.9 (CH)	58.2 (CH)	57.9 (CH)	57.9 (CH)	58.2 (CH)	57.9 (CH)
11	89.8 (C)	89.6 (C)	89.7 (C)	89.5 (C)	89.6 (C)	89.5 (C)	89.5 (C)	89.6 (C)
12	170.5 (C)	170.4 (C)	170.5 (C)	170.3 (C)	170.5 (C)	170.4 (C)	170.7 (C)	170.5 (C)
13	-	-	-	-	-	-	-	-
14	56.4 (CH)	56.5 (CH)	55.9 (CH)	47.8 (CH)	52.9 (CH)	53.1 (CH)	52.9 (CH)	53.1 (CH)
15	31.5 (CH)	31.4 (CH)	31.5 (CH)	18.6 (CH_3)	37.3 (CH_2)	37.4 (CH_2)	37.4 (CH_2)	37.3 (CH_2)
16/16[a]	17.8/19.1 (CH_3)	17.8/19.0 (CH_3)	17.7/19.0 (CH_3)	171.8 (C)	137.3 (C)	137.3 (C)	137.4 (C)	137.4 (C)
16[b]	-	-	-	-	129.3 (CH)	129.2 (CH)	129.2 (CH)	129.2 (CH)
16[c]	-	-	-	-	127.9 (CH)	127.9 (CH)	128.1 (CH)	128.0 (CH)
16[d]	-	-	-	-	126.2 (CH)	126.2 (CH)	126.3 (CH)	126.2 (CH)
17	17.8/19.1 (CH_3)	17.8/19.0 (CH_3)	17.7/19.0 (CH_3)	-	170.3 (C)	170.2 (C)	170.3 (C)	170.3 (C)
18	170.1 (C)	170.4 (C)	170.4	56.8 (CH)	-	-	-	-

†100 MHz for ^{13}C
‡150 MHz for ^{13}C
*Solvent $(CD_3)_2SO$

Table 2. Continued

Atom	1 $\delta^{13}C$	2 $\delta^{13}C$	3 $\delta^{13}C$	4 $\delta^{13}C$	5 $\delta^{13}C$	6 $\delta^{13}C$	7 $\delta^{13}C$	8 $\delta^{13}C$
19	-	-	-	31.1 (CH)	57.2 (CH)	57.0 (CH)	57.0 (CH)	56.9 (CH)
20	47.2 (CH)	53.3 (CH)	47.6 (CH)	17.9/19.1 (CH_3)	30.9 (CH)	31.1 (CH)	31.1 (CH)	31.1 (CH)
21	17.1 (CH_3)	36.6 (CH_2)	16.7 (CH_3)	17.9/19.1 (CH_3)	18.0/19.0 (CH_3)	18.1/19.0 (CH_3)	17.9/19.1 (CH_3)	18.1/19.0 (CH_3)
22/22[a]	173.9 (C)	137.4 (C)	172.1 (C)	170.3 (C)	18.0/19.0 (CH_3)	18.1/19.0 (CH_3)	17.9/19.1 (CH_3)	18.1/19.0 (CH_3)
22[b]	-	128.9 (CH)	-	-	-	-	-	-
22[c]	-	128.1 (CH)	-	-	-	-	-	-
22[d]	-	126.3 (CH)	-	-	-	-	-	-
23	-	172.7 (C)	-	-	171.0 (C)	170.5 (C)	170.5 (C)	170.6 (C)
24	-	-	66.0 (CH_2)	47.4 (CH)	-	-	-	-
25/25[a]	-	-	138.8 (C)	17.0 (CH_3)	40.5 (CH_2)	47.4 (CH)	53.4 (CH)	47.6 (CH)
25[b]	-	-	127.8 (CH)	-	-	-	-	-
25[c]	-	-	128.0 (CH)	-	-	-	-	-
25[d]	-	-	128.4 (CH)	-	-	-	-	-
26	-	-	-	173.9 (C)	171.1 (C)	16.9 (CH_3)	36.5 (CH_2)	16.7 (CH_3)
27/27[a]	-	-	-	-	-	173.9 (C)	137.4 (C)	172.2 (C)
27[b]	-	-	-	-	-	-	128.9 (CH)	-
27[c]	-	-	-	-	-	-	127.8 (CH)	-
27[d]	-	-	-	-	-	-	126.2 (CH)	-
28	-	-	-	-	-	-	172.7 (C)	-
29	-	-	-	-	-	-	-	65.9 (CH_2)
30[a]	-	-	-	-	-	-	-	135.9 (C)
30[b]	-	-	-	-	-	-	-	127.8 (CH)
30[c]	-	-	-	-	-	-	-	128.4 (CH)
30[d]	-	-	-	-	-	-	-	127.9 (CH)

[†]100 MHz for ^{13}C
[‡]150 MHz for ^{13}C
[*]Solvent $(CD_3)_2SO$

For compound **4** no ROESY correlations were observed between the PCU cage protons and the peptide chain. In ROESY spectrum of compound **5**, the aromatic protons H-16 displayed through space correlation with the PCU cage protons H-5, H-7, and H-9 where as in compound **6**, H-13 displays ROESY correlations to the proton signals H-3, H-10, H-21/22 and H-26. For compound **7**, the aromatic protons H-16 and H-27 display ROESY correlations to the isopropyl methyl protons H-21/22. Similarly in compound **8**, the isopropyl protons H-21/22 display

NOESY correlations to the aromatic protons H-30 and the cage proton H-1. The ^{1}H and ^{13}C NMR peak assignments of compounds **4**-**8** are given in Table 1 and 2.

4. EXPERIMENTAL SECTION

Analysis was performed on an Agilent 1100 HPLC (Waters Xbridge C18 150 mm x 4.6 mm x 5 μs) 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 8A instrument (Ace C18 150 mm x 21.2 mm x 5 μs) with a UV/VIS detector (215 nm) and automated fraction collector. A two-mobile phase system was employed, utilizing formic acid as the ion-pairing agent. Mobile phase A consisted of 0.1% formic acid/H_2O (v/v) and mobile phase B consisted of 0.1% formic acid/acetonitrile (v/v). High Resolution Electron Spray Ionization Mass Spectroscopic (HR-ESI-MS) analysis was performed on a Bruker MicroTOF QII mass spectrometer in positive mode with an internal calibration. Microwave couplings were conducted on a Discovery CEM Liberty microwave peptide synthesizer. The coupling conditions of the peptide synthesizer were adapted from literature.[19,20] The ^{13}C NMR and ^{1}H NMR data were recorded on a Bruker AVANCE III 400 MHz spectrometer. Some samples were analyzed using a Bruker AVANCE III 600MHz when higher sensitivity was required. Optical rotations were measured at room temperature in dry methanol using a Perkin Elmer Polarimeter-Model 341. All IR spectra were recorded on a Perkin Elmer Spectrum 100 instrument with a universal ATR attachment. All the amino acids, resins and coupling reagents are commercially available and were purchased from GL Biochem (Shanghai) Ltd. Analytical grade solvents for synthesis were procured from Sigma-Aldrich (South Africa)

General Procedure for the synthesis of 5-Hydroxy-4-oxahexacyclododecane-3-carboxylic acid, compound 13

Compound **13** was synthesized according to the previously reported method.[32]

General procedure for manual SPPS loading of first amino acid to the 2-chlorotrityl chloride resin[19,20]

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) in DCM (10 mL)

and *N*,*N*-diisopropylethylamine (DIPEA) (6.65 mmol), was added to the resin and was mixed for 2 hrs with a stream of nitrogen bubbles. 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 dried under vacuum 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 of SPPS 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 amino acids (0.2 mM), DIPEA (1 mM) and *O*-benzotriazole-1-yl-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate (HBTU) (2 mM) were prepared and inserted into their appropriate reaction vessels on the peptide synthesizer. All peptides were synthesized on a 0.5 mmol scale using microwave power (see Table 1 of the supplimentary information). The standard coupling method for arginine (Arg)[55] was used for attachment of the second amino acid to the resin (1 g) preloaded with the first amino acid. The standard coupling method was used for subsequent reactions and for Fmoc deprotection.[19,20,56] The washing procedures were similar to those used for manual SPPS. Final cleavage from the resin was performed manually. The resin bound peptide was washed with DCM (3 x 10 mL) in a reaction vessel while nitrogen was bubbled through the solution. A cleavage mixture of 0.5% (v/v) TFA and 95% (v/v) DCM was added to the dried 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 in solution was seperated from the resin by filtration and collected in a flask containing water (100 mL). The filtrate was extracted several times with DCM to remove the peptide from the water layer. DCM was removed under reduced pressure using a teflon pump at 40 °C and affording the peptide as a white powder.

General procedure for coupling of peptides to 5-Hydroxy-4-oxahexacyclododecane-3-carboxylic acid to yield compounds 1-8

The cleaved peptide (1.2 eq.) was dissolved in DCM (3 mL) followed by the addition of 5-Hydroxy-4-oxahexacyclododecane-3-carboxylic acid[32] (1 equiv), HATU (2.5 equiv) in DMF (7

mL) and DIPEA (3 equiv) as a base. The mixture was left to stir at room temperature for 24 hrs. The product was evaporated to dryness under vacuum using a teflon pump attached to the Buchi rotary evaporator, which was maintained at 40 °C. A cleavage mixture (10 mL) of 95% (v/v) TFA and 5% (v/v) DCM was added to the coupled peptide and stirred for 24 hrs at room temperature to remove the *N*-Boc protecting group. The TFA was removed from the mixture by bubbling air through the peptide and the remaining DCM was removed under *vacuum* at 30°C. The product was obtained as yellow oil which was purified by preparative HPLC [Mobile phase solution A: 0.1% formic acid/H_2O (v/v) and mobile phase solution B: 0.1% formic acid/acetonitrile (v/v)].

Compound 1

A colourless oil (76%). Retention time: 14.13 min (HPLC),$[\alpha]^{20}_D$ = -20.00 (*c* 0.20 in MeOD); IR (neat): 3288.6, 2970.5, 1726.1, 1638.0, 1529.2, 1454.7, 1326.9, 1292.2, 1204.7, 1137.1, 992.5, 950.2, 906.4, 868.8, 641.7, 521.5 cm^{-1}; HR-ESI-MS (*m/z*): calculated for $C_{20}H_{26}N_2NaO_6$ ([M + Na]$^{23+}$) 413.1683 found 413.1687. The NMR data for **1** are presented in Tables 1 and 2.

Compound 2

A white solid (62%). Retention time: 15.23 min (HPLC), Melting point: 129 – 130 ºC, $[\alpha]^{20}_D$ = -11.76 (*c* 0.17 in MeOD); IR (neat): 3300.5, 2967.2, 2868.8, 1725.5, 1640.5, 1527.3, 1455.4, 1326.0, 1292.2, 1204.9, 1138.1, 1105.6, 991.9, 949.5, 905.6, 868.6, 740.6, 698.9, 521.6, 495.4 cm^{-1}; HR-ESI-MS (*m/z*): calculated for $C_{26}H_{30}N_2O_6$ ([M + H]$^+$) 467.2177 found 467.2156. The NMR data for **2** are listed in Tables 1 and 2.

Compound 4

A pale white solid (48%). Retention time: 14.28 min (HPLC), Melting point: 159 – 160 ºC, $[\alpha]^{20}_D$ = -41.67 (*c* 0.12 in MeOD); IR (neat): 3293.0, 2969.8, 1725.6, 1639.4, 1527.3, 1454.4, 1330.9, 1292.1, 1206.3, 1138.7, 993.7, 951.8, 905.6, 868.9, 789.5, 641.5, 521.7 cm^{-1}; HR-ESI-MS (*m/z*): calculated for $C_{23}H_{31}N_3O_7$ ([M + H]$^+$) 462.2235 found 462.2269. The NMR data for **4** are presented in Tables 1 and 2.

Compound 5

A pale yellow oil (54%). Retention time: 15.28 min (HPLC), $[\alpha]^{20}_{D}$ = -27.27 (*c* 0.22 in MeOD); IR (neat): 3288.5, 2965.7, 1728.2, 1638.4, 1524.0, 1327.6, 1206.3, 1138.62, 949.5, 905.3, 699.0 cm^{-1}; HR-ESI-MS (*m/z*): calculated for $C_{28}H_{33}N_3NaO_7$ ($[M + Na]^{23+}$) 546.2110 found 546.2190. The NMR data for **5** are listed in Tables 1 and 2.

Compound 6

A pale yellow oil (42%). Retention time: 15.35 min (HPLC), $[\alpha]^{20}_{D}$ = -23.81 (*c* 0.21 in MeOD); IR (neat): 3277.7, 2966.7, 1721.5, 1637.6, 1528.4, 1454.6, 1350.3, 1226.6, 1141.49, 911.2, 743.8, 698.9 cm^{-1}; HR-ESI-MS (*m/z*): calculated for $C_{29}H_{35}N_3NaO_7$ ($[M + Na]^{23+}$) 560.2367 found 560.2351. The NMR data for **5** are given in Tables 1 and 2.

Compound 7

A yellow oil (42%). Retention time: 15.62 min (HPLC), $[\alpha]^{20}_{D}$ = -33.33 (*c* 0.21 in MeOD); IR (neat): 3287.8, 2964.4, 1714.9, 1640.5, 1524.1, 1327.2, 1205.2, 1138.4, 737.7, 697.0 cm^{-1}. HR-ESI-MS (*m/z*): calculated for $C_{35}H_{39}N_3NaO_7$ ($[M + Na]^{23+}$) 636.2719 found 636.2680. The NMR data are presented in Tables 1 and 2.

General procedure[57] for the synthesis of benzyl protected 5-Hydroxy-4-oxahexacyclododecane peptides, 3 and 8

5-Hydroxy-4-oxahexacyclododecane peptide (1 eq) and benzylbromide (2.5 eq) were dissolved in DMF (8 mL). Cesium carbonate (2 eq) was added to the solution at ambient temperature. After stirring for 2 h, saturated aqueous sodium bicarbonate (100 mL) was added to the solution and extracted with ethyl acetate (3 x 100 mL). The combined organic phases were washed with 5% aqueous citrate and saturated aqueous sodium chloride (100 mL) and dried over sodium sulphate. The solvent was evaporated, and the crude peptide was precipitated in cold ether and dried under reduced pressure. This crude product was further purified by preparative HPLC [Mobile phase solution A: 0.1% formic acid/H_2O (v/v) and mobile phase solution B: 0.1% formic acid/acetonitrile (v/v)].

Compound 3

A pale yellow oil (51%). Retention time: 15.90 min (HPLC), $[\alpha]^{20}_{D}$ = -13.33 (*c* 0.21 in MeOD); IR (neat): 3286.7, 2966.8, 1742.8, 1643.5, 1524.4, 1455.5, 1347.2, 1203.2, 1160.4, 1138.8, 1052.4, 991.3, 949.3, 904.7, 738.0, 697.3, 640.9 cm^{-1}; HR-ESI-MS (*m/z*): calculated for $C_{27}H_{32}N_2NaO_6$ ($[M + Na]^{23+}$) 503.2153 found 503.2150. The NMR data for **3** are listed in Tables 1 and 2.

Compound 8

A yellow oil (46%). Retention time: 15.68 min (HPLC), $[\alpha]^{20}_{D}$ = -19.05 (*c* 0.21 in MeOD); IR (neat): 3289.0, 2965.3, 1720.5, 1641.6, 1523.6, 1326.3, 1206.3, 1138.5, 906.4, 698.1 cm^{-1}; HR-ESI-MS (*m/z*): calculated for $C_{36}H_{41}N_3NaO_7$ ($[M + Na]^{23+}$) 650.2837 found 650.2828. The NMR data for **8** are listed in Tables 1 and 2.

5. CONCLUSION

The synthesis and complete NMR elucidation of eight novel PCU cage peptides were successfully carried out. The synthesis was achieved utilizing known protocols from the literature. The NOESY and ROESY experiments showed interesting through space correlations between the cage side chain and the PCU cage protons. Compounds **1-8** are all non-separable diastereomers and are unsymmetrical in nature, which complicated the assignments of PCU cage protons. However 2D NMR spectroscopy proved to be a crucial tool for effective structural elucidation of these compounds in spite of major overlapping of proton signals. All the ^{13}C carbon signals for these compounds are split. Incorporation of the PCU hydroxy cage acid **13** into the peptides as potential HIV-1 protease inhibitors was successful.

6. SUPPORTING INFORMATION:

All the spectra mentioned in the text are available as supporting information (Appendix 2).

7. ACKNOWLEDGMENT

This research was supported by NRF (SA) TG (GUN: 66319), HGK (GUN: 69728), HGK and PIA (SA-Sweden bilateral grant), Aspen Pharmacare and University of KwaZulu-Natal for

financial support. We thank Mr. Dilip Jagjivan (UKZN, South Africa) for his assistance with the NMR experiments.

8. REFERENCES

1. Gerzon, K.; Kau, D. *J. Med. Chem.* **1967**, *10*, 189-199.
2. Nagasawa, H. T.; Elberlin.Ja; Shirota, F. N. *J. Med. Chem.* **1973**, *16*, 823-826.
3. Marchand, A. P.; Allen, R. W. *J. Org. Chem.* **1974**, *39*, 1596-1596.
4. Nagasawa, H. T.; Elberling, J. A.; Shirota, F. N. *J. Med. Chem.* **1975**, *18*, 826-830.
5. Craze, G. A.; Watt, I. *J. Chem. Soc. Perk. T. 2* **1981**, 175-184.
6. Griffin, G. W.; Marchand, A. P. *Chem. Rev.* **1989**, *89*, 997-1010.
7. Marchand, A. P. *Adv. Theo. Interesting Mol.* **1989**, *1*, 357-99.
8. Marchand, A. P. *Chem. Rev.* **1989**, *89*, 1011-1033.
9. Marchand, A. P. In *A Research Annual;* Thummel, R. P. Ed.; JAI Press: Greenwich,CT, **1989**; p. 467.
10. Fourie, L.; Govender, T.; Hariprakasha, H. K.; Kruger, H. G.; Raasch, T. *Magn. Reson. Chem.* **2004**, *42*, 617-623.
11. Geldenhuys, W. J.; Malan, S. F.; Bloomquist, J. R.; Marchand, A. P.; Van der Schyf, C. J. *Med. Res. Rev.* **2005**, *25*, 21-48.
12. Bisetty, K.; Govender, P.; Kruger, H. G. *Biopolymers* **2006**, *81*, 339-349.
13. Ito, F. M.; Petroni, J. M.; de Lima, D. P.; Beatriz, A.; Marques, M. R.; de Moraes, M. O.; Costa-Lotufo, L. V.; Montnegro, R. C.; Magalhaes, H. I. F.; Pessoa, C. D. O. *Molecules* **2007**, *12*, 271-282.
14. Albericio, F.; Arvidson, P. I.; Bisetty, K.; Giral, E.; Govender, T.; Jali, S.; Kongsaeree, P.; Kruger, H. G.; Prabpai, S. *Chem. Biol. Drug Des.* **2008**, *71*, 125-130.
15. Oliver, D. W.; Malan, S. F. *Med. Chem. Res.* **2008**, *17*, 137-151.
16. Onajole, O. K.; Govender, K.; Govender, P.; van Helden, P. D.; Kruger, H. G.; Maguire, G. E. M.; Muthusamy, K.; Pillay, M.; Wiid, I.; Govender, T. *Eur. J. Med. Chem.* **2009**, *44*, 4297-4305.
17. Lamoureux, G.; Artavia, G. *Curr. Med. Chem.* **2010**, *17*, 2967-2978.
18. Banister, S. D.; Moussa, I. A.; Beinat, C.; Reynolds, A. J.; Schiavini, P.; Jorgensen, W. T.; Kassiou, M. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 38-41.
19. Makatini, M. M.; Petzold, K.; Sriharsha, S. N.; Soliman, M. E. S.; Honarparvar, B.; Arvidsson, P. I.; Sayed, Y.; Govender, P.; Maguire, G. E. M.; Kruger, H. G.; Govender, T. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 2274-2277.
20. Makatini, M. M.; Petzold, K.; Sriharsha, S. N.; Ndlovu, N.; Soliman, M. E. S.; Honarparvar, B.; Parboosing, R.; Naidoo, A.; Arvidsson, P. I.; Sayed, Y.; Govender, P.; Maguire, G. E. M.; Kruger, H. G.; Govender, T. *Eur. J. Med. Chem.* **2011**, *46*, 3976-3985.
21. Onajole, O. K.; Belewa, X. V.; Coovadia, Y.; Govender, T.; Kruger, H. G.; Maguire, G. E. M.; Naidu, D.; Somai, B.; Singh, N.; Govender, P. *Med. Chem. Res.* **2011**, *20*, 1394-1401.
22. Karpoormath, R.; Sayed, Y.; Govender, P.; Govender, T.; Kruger, H. G.; Soliman, M. E. S.; Maguire, G. E. M. *Bioorg. Chem.* **2012**, *40*, 19-29.

23. Joubert, J.; Geldenhuys, W. J.; Van der Schyf, C. J.; Oliver, D. W.; Kruger, H. G.; Govender, T.; Malan, S. F. *Chem. Med. Chem.* **2012**, *7*, 375-384.
24. Honarparvar, B.; Makatini, M. M.; Pawar, S. A.; Petzold, K.; Soliman, M. E. S.; Arvidsson, P. I.; Sayed, Y.; Govender, T.; Maguire, G. E. M.; Kruger, H. G. *Chem. Med. Chem.* **2012**, *7*, 1009-1019.
25. Dekker, J.; Dekker, J. J.; Fourie, L.; Dekker, T. G.; Pachler, K. G. R.; Wessels, P. L. *Tetrahedron Lett.* **1976**, 1613-1614.
26. Pachler, K. G. R.; Wessels, P. L.; Dekker, J.; Dekker, J. J.; Dekker, T. G. *Tetrahedron Lett.* **1976**, *17*, 3059-3062.
27. Dekker, T. G.; Pachler, K. G. R.; Wessels, P. L.; Dekker, J.; Dekker, J. J.; Fourie, L. *Org. Magn. Reson.* **1978**, *11*, 520-523.
28. Boeyens, J. C. A.; Burger, J.; Dekker, J. J.; Fourie, L. *S. Afr. J. Chem.* **1978**, *31*, 101-106.
29. Oliver, D. W.; Dekker, T. G.; Snyckers, F. O.; Fourie, T. G. *J. Med. Chem.* **1991**, *34*, 851-854.
30. Martins, F. J. C.; Coetzee, G. H.; Fourie, L.; Venter, H. J.; Viljoen, A. M.; Wessels, P. L. *Magn. Reson. Chem.* **1993**, *31*, 578-584.
31. Martins, F. J. C.; Viljoen, A. M.; Kruger, H. G.; Joubert, J. A. *Tetrahedron* **1993**, *49*, 9573-80.
32. Martins, F. J. C.; Viljoen, A. M.; Kruger, H. G.; Wessels, P. L. *Tetrahedron* **1993**, *49*, 6527-6532.
33. Martins, F. J. C.; Viljoen, A. M.; Kruger, H. G.; Joubert, A. J.; Wessels, P. L. *Tetrahedron* **1994**, *50* 10783-10790.
34. Govender, T.; Hariprakasha, H. K.; Kruger, H. G.; Marchand, A. P. *Tetrahedron Asymmetr.* **2003**, *14*, 1553-1557.
35. Martins, F. J. C.; Viljoen, A. M.; Kruger, H. G.; Wessels, P. L. *Magn. Reson. Chem.* **2004**, *42*, 402-408.
36. Govender, T.; Hariprakasha, H. K.; Kruger, H. G.; Raasch, T. *S. Afr. J. Chem.* **2005**, *58*, 37-40.
37. Kruger, H. G.; Ramdhani, R. *Magn. Reson. Chem.* **2006**, *44*, 1058-1062.
38. Kruger, H. G.; Ramdhani, R. *S. Afr. J. Chem.* **2006**, *59*, 71-74.
39. Govender, T.; Kruger, H. G.; Makatini, M.; Onajole, O. K. *Struct. Chem.* **2008**, *19*, 719-726.
40. Onajole, O. K.; Govender, T.; Makatini, M.; Kruger, H. G. *Magn. Reson. Chem.* **2008**, *46*, 1007-1014.
41. Onajole, O. K.; Govender, P.; Govender, T.; Maguire, G. E. M.; Kruger, H. G. *Struct. Chem.* **2009**, *20*, 1067-1076.
42. Onajole, O. K.; Makatini, M. M.; Govender, P.; Govender, T.; Maguire, G. E. M.; Kruger, H. G. *Magn. Reson. Chem.* **2010**, *48*, 249-255.
43. Makatini, M.; Chetty, T.; Onajole, O. K.; Govender, T.; Govender, P.; Maguire, G. E. M.; Kruger, H. G. *J. Pept. Sci.* **2012**, *18*, 114-21.
44. Makatini, M. M.; Petzold, K.; Alves, C. u. N.; Arvidsson, P. I.; Honarparvar, B.; Govender, P.; Govender, T.; Kruger, H. G.; Sayed, Y.; Jer√¥nimoLameira; Maguire, G. E. M.; Soliman, M. E. S. *J. Enzyme Inhib. Med. Chem.* **2012**, 1-12.
45. Onajole, O. K.; Govender, P.; van, H. P. D.; Kruger, H. G.; Maguire, G. E. M.; Wiid, I.; Govender, T. *Eur. J. Med. Chem.* **2010**, *45*, 2075-2079.

46. Onajole, O. K.; Coovadia, Y.; Govender, T.; Kruger, H. G.; Maguire, G. E. M.; Naidu, D.; Singh, N.; Govender, P. *Chem. Biol. Drug Des.* **2011**, *77*, 295-9.
47. Onajole, O. K.; Sosibo, S.; Govender, P.; Govender, T.; van, H. P. D.; Maguire, G. E. M.; Mlinaric-Majerski, K.; Wiid, I.; Kruger, H. G. *Chem. Biol. Drug Des.* **2011**, *78*, 1022-30.
48. Zah, J.; Terre'Blanche, G.; Erasmus, E.; Malan, S. F. *Bioorg. Med. Chem.* **2003,** *11*, 3569-3578.
49. Brookes, K. B.; Hickmott, P. W.; Jutle, K. K.; Schreyer, C. A. *S. Afr. J. Chem.* **1992**, *45*, 8-11.
50. Ghosh, A. K.; Kincaid, J. F.; Cho, W. H.; Walters, D. E.; Krishnan, K.; Hussain, K. A.; Koo, Y.; Cho, H.; Rudall, C.; Holland, L.; Buthod, J. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 687-690.
51. Cookson, R. C.; Crundwell, E.; Hill, R. R.; Hudec, J. *J. Chem. Soc.* **1964**, 3062-3075.
52. Williams, D. E.; Dalisay, D. S.; Patrick, B. O.; Matainaho, T.; Andrusiak, K.; Deshpande, R.; Myers, C. L.; Piotrowski, J. S.; Boone, C.; Yoshida, M.; Andersen, R. J. *Org. Lett.* **2011**, *13*, 3936-3939.
53. Osborne, A. G.; Ahmet, M. T.; Miller, J. R.; Warmsley, J. F. *Spectrochim. Acta, Part A* **1995**, *51A*, 237-246.
54. Katritzky, A. R.; Nikonov, G. N.; Moyano, E. L.; Akhmedov, N. G.; Steel, P. J. *Arkivoc* **2003**, *5*, 121-138.
55. Amblard, M.; Fehrentz, J.-A.; Martinez, J.; Subra, G. *Mol. Biotechnol.* **2006**, *33*, 239-254.
56. Yu, H. M.; Chen, S. T.; Wang, K. T. *J. Org. Chem.* **1992**, *57*, 4781-4784.
57. Yamamoto, T.; Nair, P.; Jacobsen, N. E.; Davis, P.; Ma, S.-w.; Navratilova, E.; Moye, S.; Lai, J.; Yamamura, H. I.; Vanderah, T. W.; Porreca, F.; Hruby, V. J. *J. Med. Chem.* **2008**, *51*, 6334-6347.

CHAPTER 6

SUMMARY AND CONCLUSION

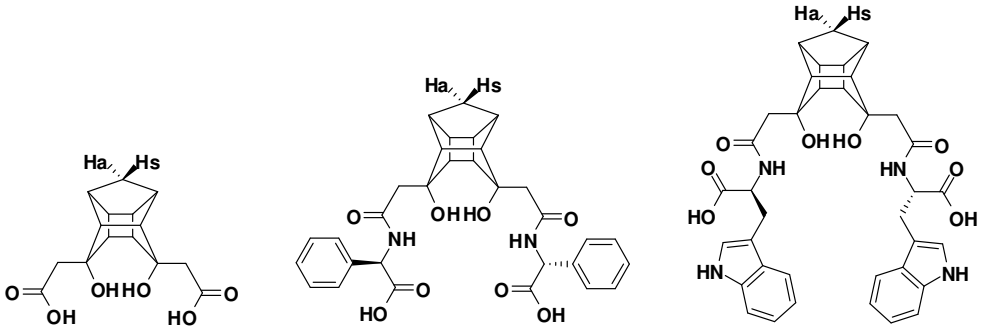
Acquired Immune Deficiency Syndrome (AIDS) continues to be a major epidemic and challenge in the world. Currently, the most effective course of treatment consists of a combination therapy of HIV protease and reverse transcriptase inhibitors. Due to their severe side effects, toxicities, high dosage and high treatment cost, there is an urgent and constant demand for the development of new anti HIV drugs. The proteolytic HIV protease enzyme is still the most sought after target for AIDS chemotherapy by researchers in all continents.

In this study, peptides and peptoids incorporated into the PCU cage moieties were designed and synthesized to potentially improve the pharmacokinetic and pharmacodynamic properties of the molecules. The rigid carbon skeleton of the polycyclic cages have shown inherent chemical and biochemical properties such as retarding metabolic degradation, stability, receptor site specificity and enhancing the ability of the drugs to cross the blood-brain barrier of the central nervous system (CNS). Two different polycyclic cage molecules namely, the PCU diol diacid and PCU hydroxy acid were used as scaffolds for the attachment to different short peptide/peptoid chain sequences. The sequences were chosen to fit the protease subsites and the synthesized compounds resulted in novel HIV protease inhibitors. This research project has resulted in the synthesis of 30 intermediates and 22 novel compounds in total.

1. PCU-DIOL INHIBITORS

The symmetric and mesomeric PCU diol diacids (**1** and **14**) in Chapter 2 and Chapter 1 respectively were successfully synthesized, isolated, characterized and reported for the first time in this project. The PCU diol diacid moiety resembles a dihydroxyethylene type isostere which was previously utilized as a HIV protease inhibitor. Nine novel PCU diol peptides/peptoids were synthesized as HIV-PR inhibitors. It was observed from the crystal structure of PCU diol diphenylalanine (**5**), in Chapter 1 (CCDC 797748), that the distance between the diols (2.3 Å) was quite close to the calculated diol distance (≈ 2.3 Å) of the transition state (R1-NH-C$(OH)_2$-CH-NH-R2)[1,2] resulting from the hydrolysis of the natural HIV-PR protein substrate. Peptides **7**

and **10**, in Chapter 1, contain two phenylglycine and two tryptophan amino acids respectively as side arms to the cage, displayed moderate inhibitory activity with IC_{50} values of 5 μM. From the docking and molecular dynamic (MD) studies it was observed that the cage diol peptides fit comfortably inside the active site of the enzyme.



14 **7** **10**

Figure 1. Structure of PCU diol diacid **14** and the most active derivatives **7** and **10** (Chapter 2).

The hetero-atoms displayed strong hydrogen bonding interactions with the enzymatic aspartic acid (Asp25/Asp25') residues and very little movement of the PCU cage hydroxyl groups inside the enzyme pocket was observed during the course of the MD simulations. The NMR elucidation of these novel pentacycloundecane (PCU) derived short peptides (Chapter 2) was also presented. Major overlapping of the methine resonances of the PCU 'cage' in the ^{1}H and ^{13}C NMR spectra was observed. Conformational differences of the peptide side chains due to diastereomeric interactions between the cage skeleton and the chiral side chains were also observed. The side arms displayed weak NOESY/ROESY interactions with the cage protons, which enabled us to deduce the orientation of the side arms with respect to the PCU cage skeleton.

2. PCU HYDROXY ACID INHIBITORS

A series of novel structurally diverse diastereomeric PCU derived hydroxy acid peptide and peptoid analogs were synthesized and tested against the resistant prone wild type C-South African (C-SA) HIV-protease enzyme. The rationale behind using the cage ether peptide/peptoid was that the cyclic PCU hydroxy ether moiety would perhaps serve as a cyclic

ether transition state[3-5] analogue under protease conditions. Amongst these novel compounds **16**, **17**, **20** and **23** (Figure 2) gave IC_{50} values ranging from 0.6 to 5.0 μM against the wild type C-SA HIV-1 protease enzyme. The binding energies of the docked diastereomeric PCU-peptide inhibitors were in reasonable agreement with the measured IC_{50} values.

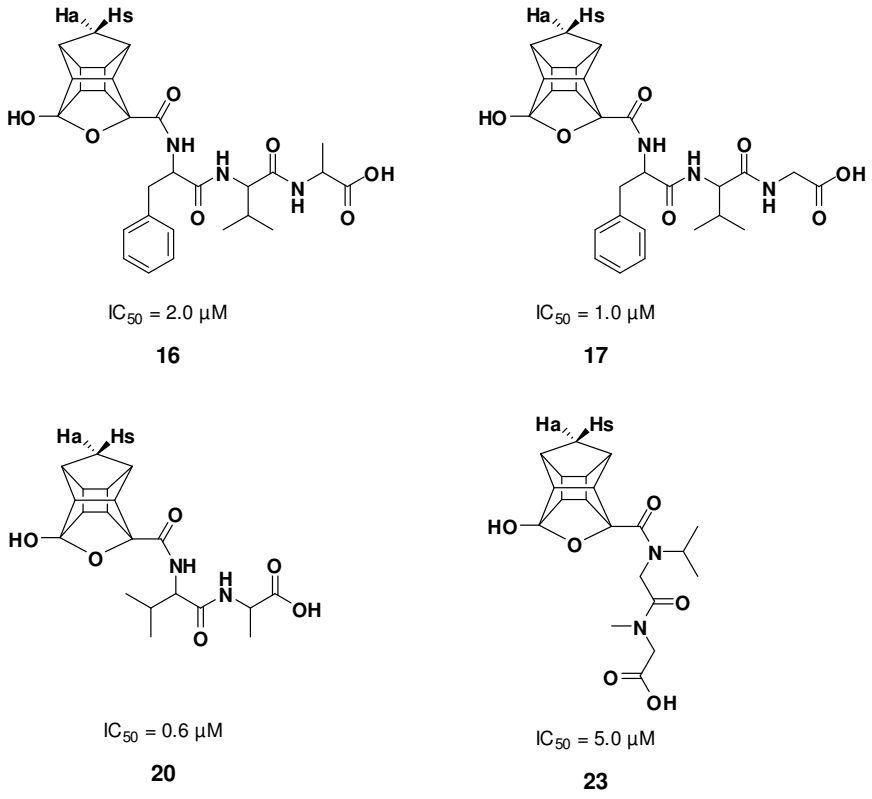


Figure 2. PCU hydroxy acid derived peptides/peptoids **16**, **17**, **20** and **23** (Chapter 4).

The docked structures displayed a conserved hydrogen-bonding pattern between the cage hydroxy ether and the active site residues (Asp25/Asp25') confirming the proposed transition state role played by the cage ether. This was further evident from the analysis of the hydrogen bond interactions along the MD trajectories.

2D NMR techniques proved to be very important tools in elucidating the structures of the diastereomeric PCU hydroxy peptides (Chapter 5). It was observed from the ^{13}C NMR spectra that all the carbon signals of the cage and their side chains were split, confirming that the compounds were diastereomeric mixtures. Through space NOESY/ROESY correlations were

also observed between the cage protons and the side chains, which also enabled us to have a probable 3D view of these peptides.

3. CONCLUSION

A series of two different PCU derived peptide/peptoid analogues were synthesized and evaluated as potential HIV-1 protease inhibitors. In the case of PCU cage diol peptide/peptoid, the best IC_{50} value observed was 5 μM for the cage diol peptides having amino acids 2-phenylglycine (**7**) and tryptophan (**10**) as side arms, as shown in Figure 1. Enhanced IC_{50} values can perhaps be achieved by changing the length of the link between cage moiety and side chain peptide.

The PCU hydroxy acid derived peptides and peptoids have emerged as a new class of irreversible non-scissile ether bridge type isosters. Promising IC_{50} values of 0.6 μM were obtained from this peptide/peptoid series for cage peptides (**20**) and 1 μM (**17**) (Figure 2). Variation of the amino acid sequence and protection of the C-terminal of the cage peptide/peptoid enabled us to determine a structure activity relationship for these inhibitors. Substitution of bulky hydrophobic groups for the $S^{1}/S^{1'}$ and $S^{3}/S^{3'}$ resulted in considerable loss of inhibitory activity. From docking and MD simulations it was observed that the bulky cage or part thereof occupies the $S^{1}/S^{1'}$ subsite.

The 2D NMR techniques especially NOESY/ROESY experiments have proved to be extremely important tools in elucidating these complex PCU cage compounds with overlapping signals. The binding energies calculated from the docking studies were in reasonable agreement with the measured IC_{50} values. Conserved hydrogen bonds were clearly observed between the cage hetero-atoms and one of the triad residues, Asp25-Thr26-Gly27 (A/B chains) which in turn helped in anchoring the cage skeleton to the $S^{1}/S^{1'}$ subsite of the C-SA HIV-PR. The MD trajectories revealed that the enzyme complex was stable and the cage inhibitors easily fit into the enzyme pocket/active site. These computational techniques allowed us to rationalize the observed HIV protease inhibition for different cage peptides and peptoids. Future studies should explore the activity of these cage peptide/peptoids against other protease related diseases.

4. REFERENCES

1. Wlodawer, A.; Vondrasek, J. *Annu. Rev. Bioph. Biom.* **1998**, *27*, 249-284.
2. Brik, A.; Wong, C. H. *Org. Biomol. Chem.* **2003**, *1*, 5-14.
3. Ghosh, A. K.; Thompson, W. J.; McKee, S. P.; Duong, T. T.; Lyle, T. A.; Chen, J. C.; Darke, P. L.; Zugay, J. A.; Emini, E. A. *J. Med. Chem.* **1993**, *36*, 292-294.
4. Ghosh, A. K.; Kincaid, J. F.; Cho, W. H.; Walters, D. E.; Krishnan, K.; Hussain, K. A.; Koo, Y.; Cho, H.; Rudall, C.; Holland, L.; Buthod, J. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 687-690.
5. Ghosh, A. K.; Leshchenko, S.; Noetzel, M. *J. Org. Chem.* **2004**, *69*, 7822-7829.

APPENDIX 1

SUPPLEMENTARY INFORMATION FOR CHAPTER 2 AND 3

Chapter 2

Novel PCU cage diol peptides as potential targets against wild type CSA HIV-1 protease; Synthesis, biological screening and molecular modeling studies.

Rajshekhar Karpoormath,[a] Yasien Sayed,[b]Thavendran Govender,[d] Hendrik G. Kruger,[a] Glenn E. M. Maguire,[a*] and Mahmoud E. S. Soliman[d,e*]

[a]School of Chemistry, University of KwaZulu Natal, Durban 4001, South Africa
[b]School of Biochemistry, University of KwaZulu Natal, Durban 4001, South Africa
cProtein Structure-Function Research Unit, School of Molecular and Cell Biology, University of the Witwatersrand, Wits 2050, South Africa
[d]School of Pharmacy and Pharmacology, University of KwaZulu Natal, Durban 4001, South Africa
[e]Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Zagazig University, Zagazig 44519, Egypt
*** *Corresponding authors: Email:*kruger@ukzn.ac.za, Fax: +27-31-2603091**

Chapter 3

Synthesis and NMR elucidation of Novel Pentacycloundecane-Derived Peptides

Rajshekhar Karpoormath,[a,b] Oluseye K. Onajole,[a] Tricia Naicker,[b] Thavendran Govender,[b] Glenn E. M. Maguire,[a] and Hendrik G. Kruger[a*]

[a]School of Chemistry, University of KwaZulu-Natal, Durban 4001, South Africa
[b]School of Pharmacy and Pharmacology, University of KwaZulu-Natal, Durban 4001, South Africa
*** *Corresponding authors: Email:*kruger@ukzn.ac.za, Fax: +27-31-2603091**

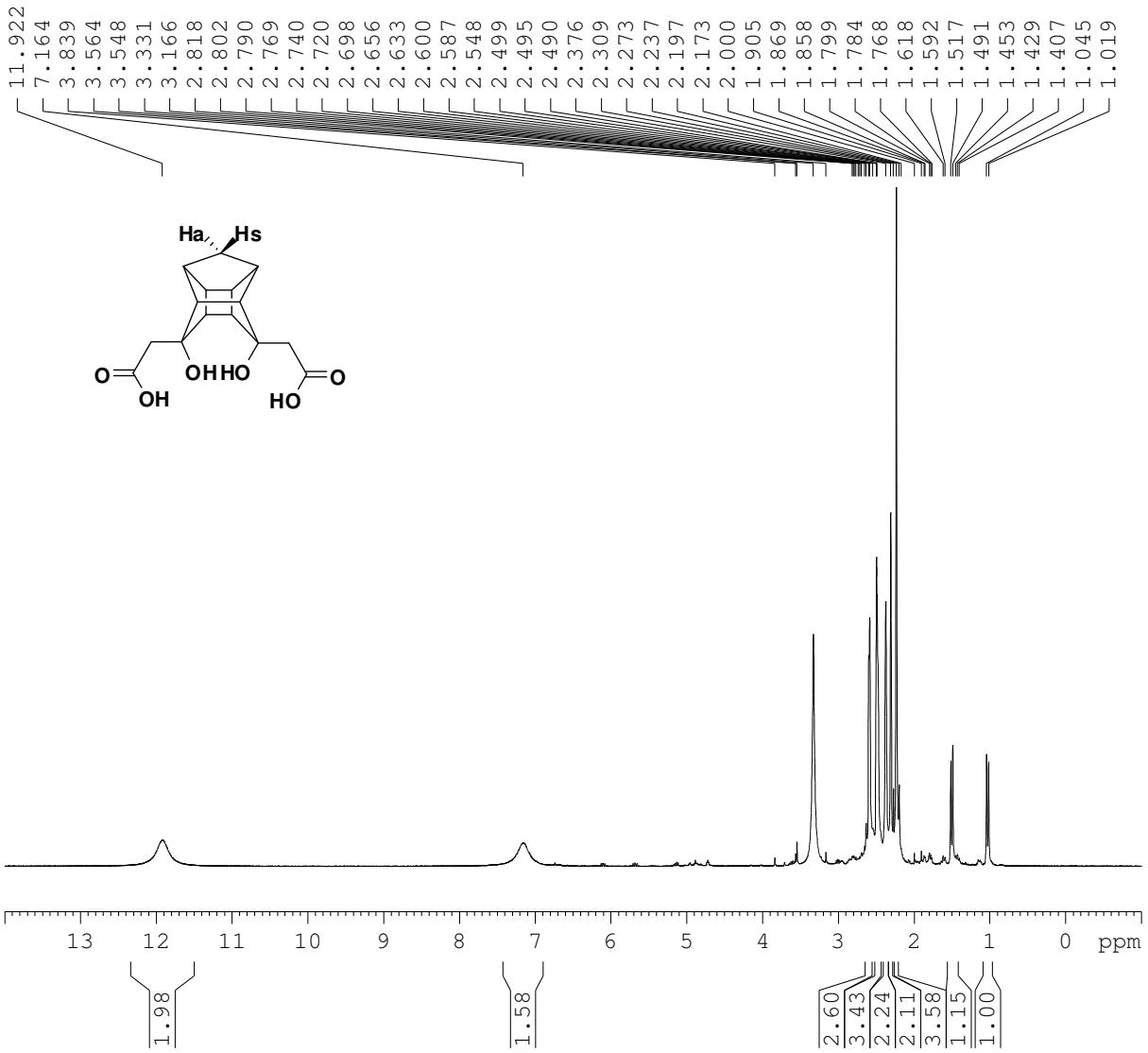
Over-expression, extraction and purification of the C-SA protease (Chapter 1)

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 (OD_{600}) 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 $MgCl_2$ 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.

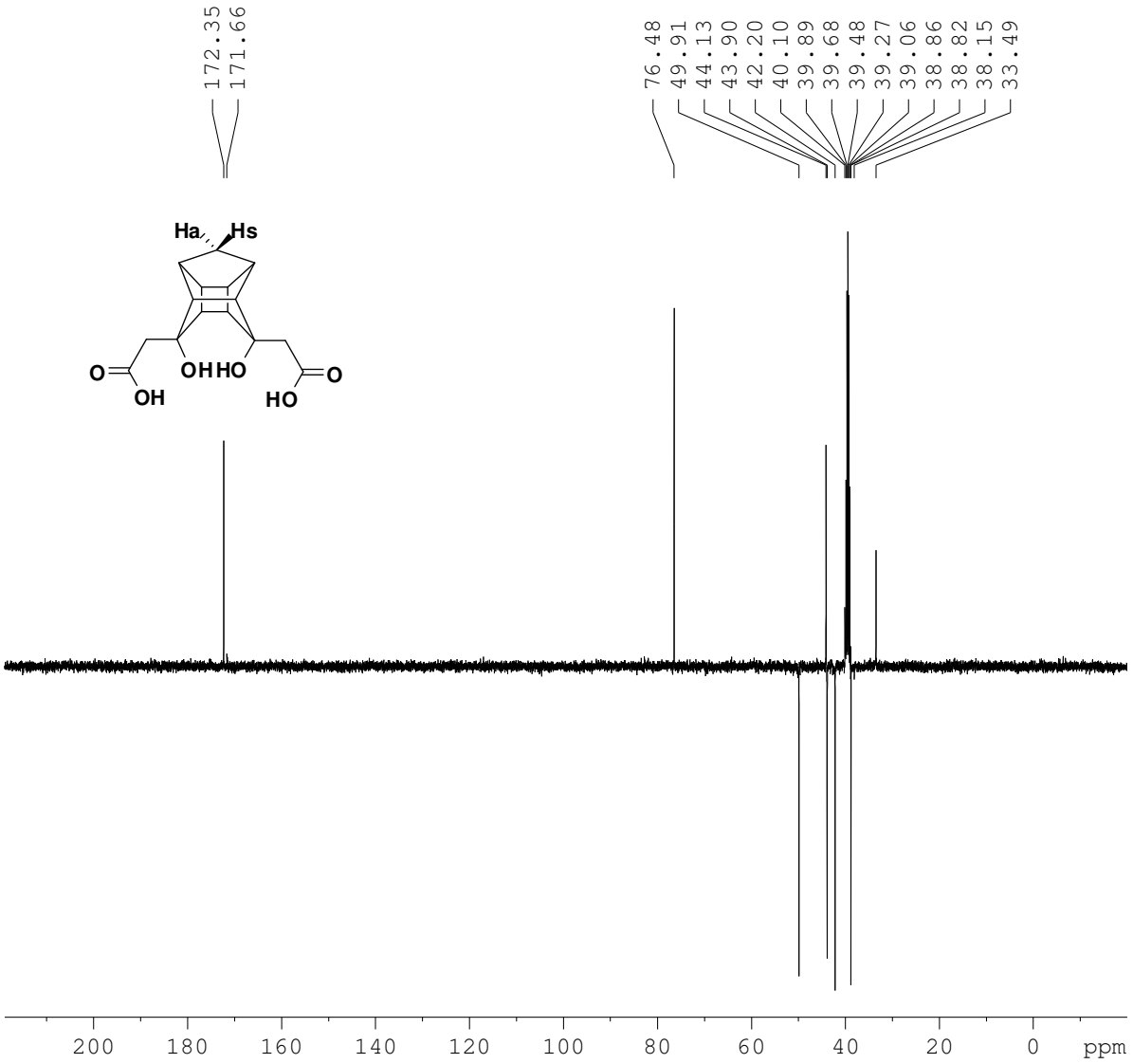
The protease, in the supernatant, was purified by passing through an anion exchange (DEAE) column previously equilibrated with solubilization buffer. Upon elution from the column, the protease was acidified by adding formic acid to a final concentration of 25 mM. Precipitation of significant amount of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



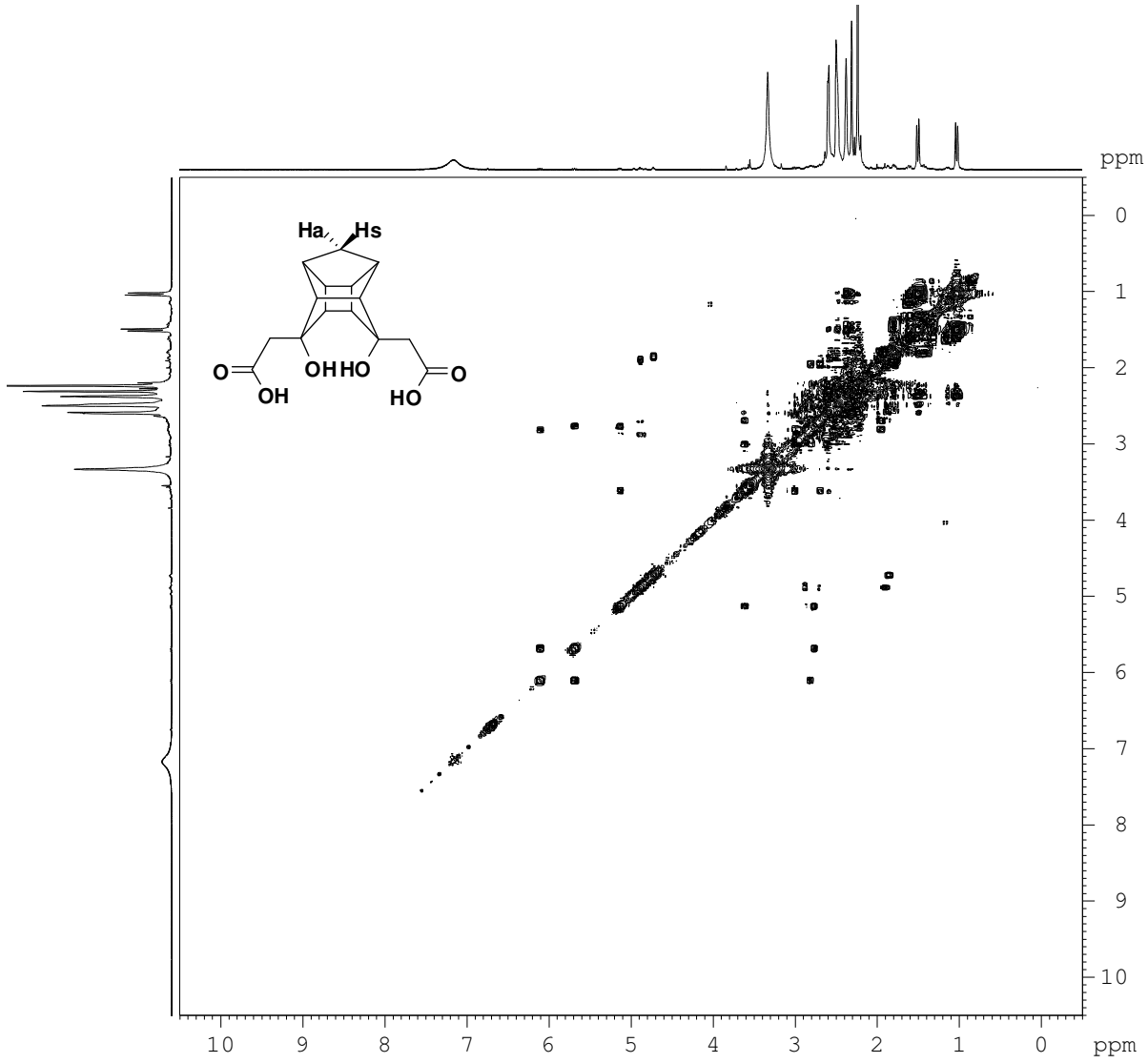

```
Current Data Parameters
NAME      Apr06-2011-GK-Raj
EXPNO                   20
PROCNO                   1

F2 - Acquisition Parameters
Date_           20110407
Time               10.39
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG             zg30
TD                 32768
SOLVENT             DMSO
NS                    16
DS                     2
SWH             8223.685 Hz
FIDRES          0.250967 Hz
AQ            1.9923444 sec
RG                   144
DW                60.800 usec
DE                 6.50 usec
TE                298.2 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                  1H
P1                10.00 usec
PL1               -3.00 dB
PL1W        15.48668575 W
SFO1        400.2224715 MHz

F2 - Processing parameters
SI                 16384
SF          400.2200000 MHz
WDW                   EM
SSB                    0
LB                0.30 Hz
GB                     0
PC                1.00
```

^{1}H NMR Spectrum of Compound 14 (Chapter 1) and Compound 1 (Chapter 2)



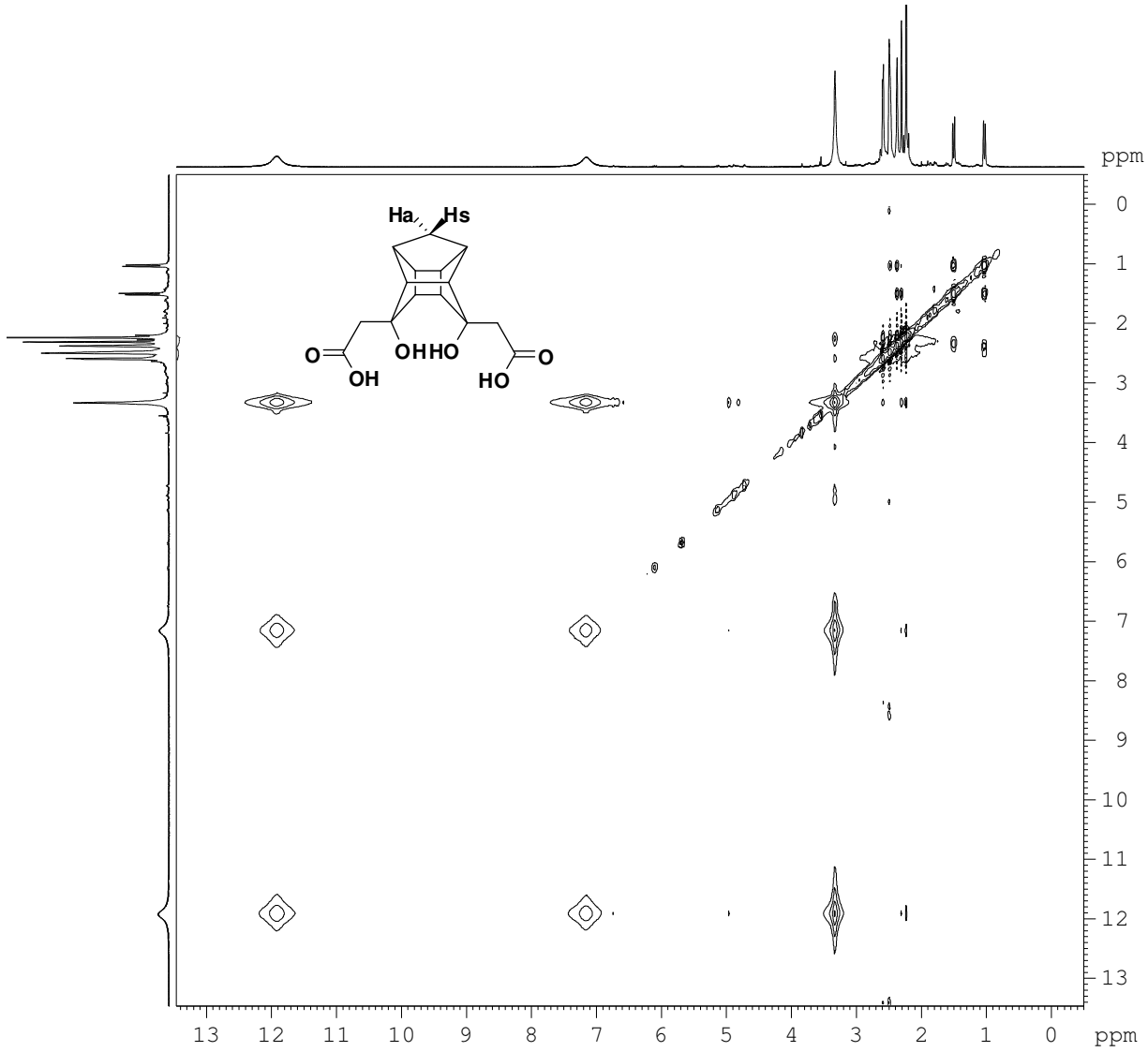

```
Current Data Parameters
NAME      Apr06-2011-GK-Raj
EXPNO                   21
PROCNO                   1

F2 - Acquisition Parameters
Date_           20110407
Time              11.04
INSTRUM           spect
PROBHD   5 mm PABBO BB-
PULPROG            jmod
TD                65536
SOLVENT            DMSO
NS                   512
DS                     4
SWH            24038.461 Hz
FIDRES          0.366798 Hz
AQ              1.3631988 sec
RG                  2050
DW                20.800 usec
DE                 6.50 usec
TE                298.2 K
CNST2         145.0000000
CNST11          1.0000000
D1            2.00000000 sec
D20           0.00689655 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                 13C
P1                  8.40 usec
P2                 16.80 usec
PL1                -2.00 dB
PL1W         54.14257431 W
SFO1         100.6454626 MHz

======== CHANNEL f2 ========
CPDPRG2          waltz16
NUC2                 1H
PCPD2              90.00 usec
PL2                -3.00 dB
PL12               15.60 dB
PL2W         15.48668575 W
PL12W         0.21377575 W
SFO2         400.2216009 MHz

F2 - Processing parameters
SI                 32768
SF           100.6354493 MHz
WDW                   EM
SSB                    0
LB                  1.00 Hz
GB                     0
PC                  1.40
```

^{13}C NMR Spectrum of Compound 14 (Chapter 1) and Compound 1 (Chapter 2)



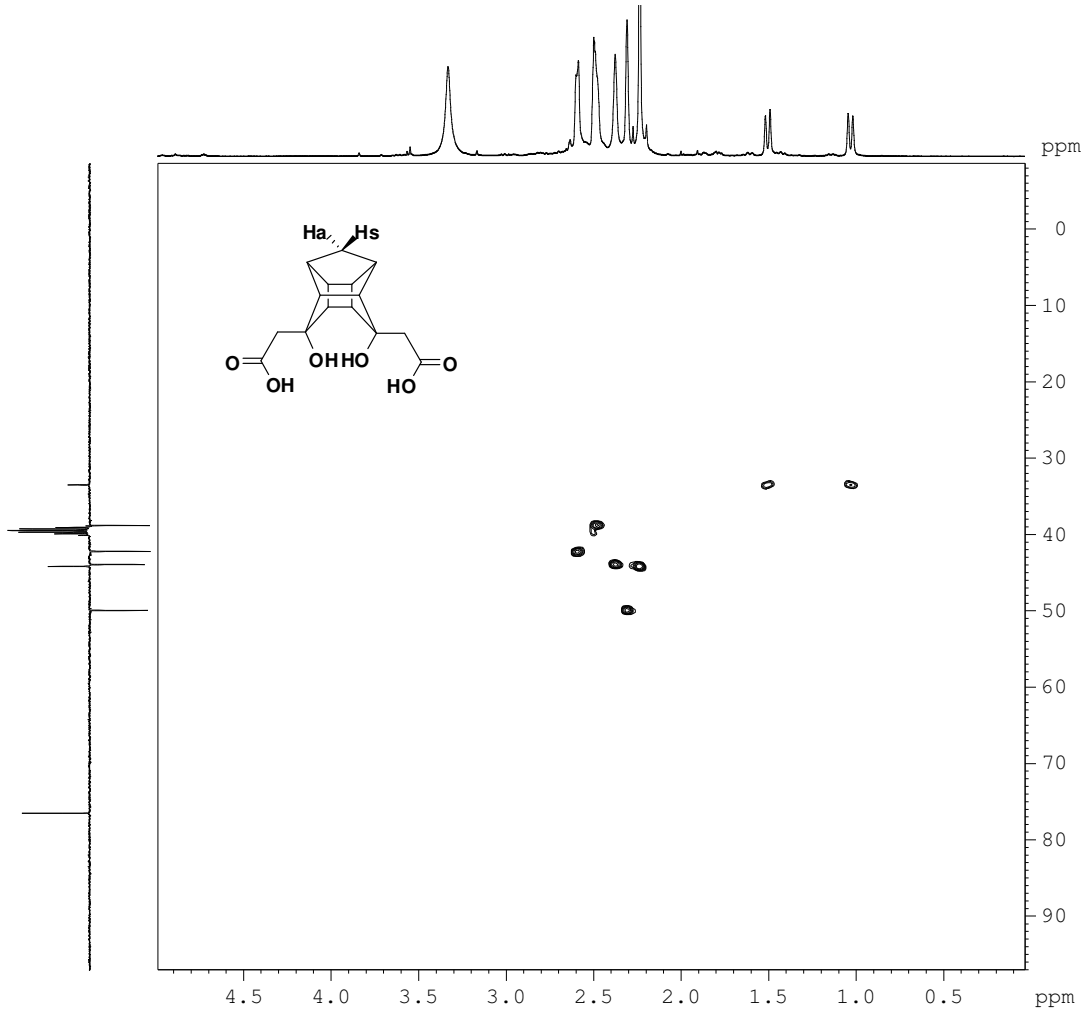

```
Current Data Parameters
NAME      Apr06-2011-GK-Raj
EXPNO                   23
PROCNO                   1

F2 - Acquisition Parameters
Date_              20110408
Time                  0.53
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG          cosygpqf
TD                    2048
SOLVENT               DMSO
NS                       2
DS                       8
SWH              5376.344 Hz
FIDRES           2.625168 Hz
AQ             0.1905140 sec
RG                      64
DW                 93.000 usec
DE                  6.50 usec
TE                 298.2 K
D0            0.00000300 sec
D1            1.48812103 sec
D13           0.00000400 sec
D16           0.00020000 sec
IN0           0.00018600 sec

======== CHANNEL f1 ========
NUC1                    1H
P0                  10.00 usec
P1                  10.00 usec
PL1                 -3.00 dB
PL1W         15.48668575 W
SFO1         400.2227004 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPZ1                10.00 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                      1
TD                     128
SFO1             400.2227 MHz
FIDRES          42.002689 Hz
SW                 13.433 ppm
FnMODE                  QF

F2 - Processing parameters
SI                    1024
SF           400.2200000 MHz
WDW                   SINE
SSB                      0
LB                    0.00 Hz
GB                       0
PC                    1.40

F1 - Processing parameters
SI                    1024
MC2                     QF
SF           400.2200000 MHz
WDW                   SINE
SSB                      0
LB                    0.00 Hz
```

COSY NMR Spectrum of Compound 1 (Chapter 2)




```
Current Data Parameters
NAME      Apr06-2011-GK-Raj
EXPNO                    24
PROCNO                    1

F2 - Acquisition Parameters
Date_              20110408
Time                   1.02
INSTRUM               spect
PROBHD   5 mm PABBO BB-
PULPROG             noesyph
TD                     2048
SOLVENT                DMSO
NS                        4
DS                        4
SWH               5376.344 Hz
FIDRES            2.625168 Hz
AQ                0.1905140 sec
RG                       64
DW                  93.000 usec
DE                   6.50 usec
TE                  298.2 K
D0              0.00008027 sec
D1              2.06021094 sec
D8              0.30000001 sec
IN0             0.00018600 sec

======== CHANNEL f1 ========
NUC1                     1H
P1                   10.00 usec
PL1                  -3.00 dB
PL1W          15.48668575 W
SFO1          400.2227004 MHz

F1 - Acquisition parameters
ND0                       1
TD                      256
SFO1             400.2227 MHz
FIDRES           21.001345 Hz
SW                  13.433 ppm
FnMODE         States-TPPI

F2 - Processing parameters
SI                     1024
SF            400.2200000 MHz
WDW                   QSINE
SSB                       2
LB                     0.00 Hz
GB                        0
PC                     1.00

F1 - Processing parameters
SI                     1024
MC2            States-TPPI
SF            400.2200000 MHz
WDW                   QSINE
SSB                       2
LB                     0.00 Hz
GB                        0
```

NOESY NMR Spectrum of Compound 1 (Chapter 2)



Current Data Parameters
NAME Apr06-2011-GK-Raj
EXPNO 25
PROCNO 1

F2 - Acquisition Parameters
Date_ 20110408
Time 1.47
INSTRUM spect
PROBHD 5 mm PABBO BB-
PULPROG hsqcedetgp
TD 1024
SOLVENT DMSO
NS 2
DS 16
SWH 5376.344 Hz
FIDRES 5.250336 Hz
AQ 0.0952820 sec
RG 2050
DW 93.000 usec
DE 6.50 usec
TE 298.2 K
CNST2 145.0000000
D0 0.00000300 sec
D1 1.50061405 sec
D4 0.00172414 sec
D11 0.03000000 sec
D13 0.00000400 sec
D16 0.00020000 sec
D21 0.00345000 sec
IN0 0.00003000 sec
ZGOPTNS

======== CHANNEL f1 ========
NUC1 1H
P1 10.00 usec
P2 20.00 usec
P28 1000.00 usec
PL1 -3.00 dB
PL1W 15.48668575 W
SFO1 400.2227004 MHz

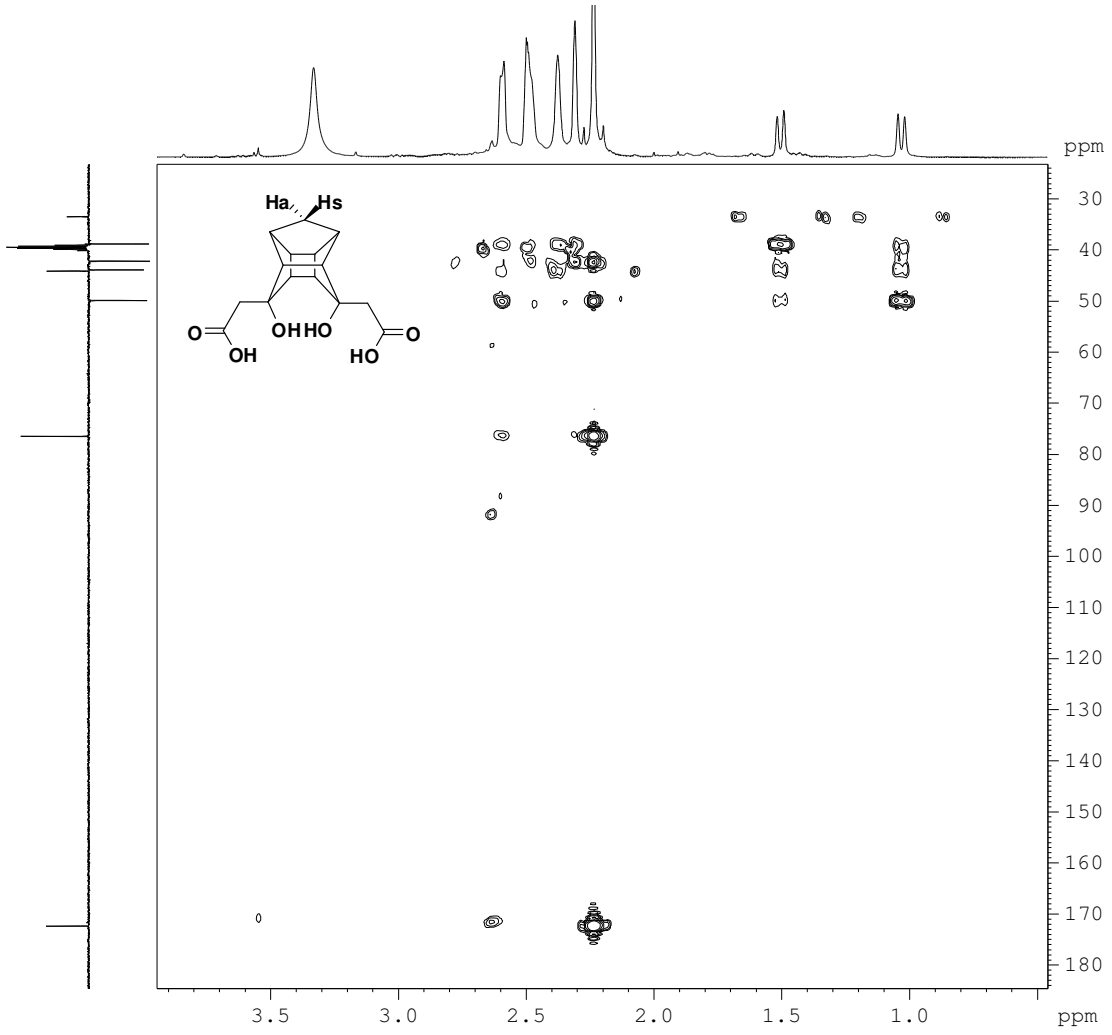
======== CHANNEL f2 ========
CPDPRG2 garp
NUC2 13C
P3 9.00 usec
P4 18.00 usec
PCPD2 70.00 usec
PL2 -2.00 dB
PL12 15.60 dB
PL2W 54.14257431 W
PL12W 0.94089001 W
SFO2 100.6429049 MHz

====== GRADIENT CHANNEL =====
GPNAM1 SINE.100
GPNAM2 SINE.100
GPZ1 80.00 %
GPZ2 20.10 %
P16 1000.00 usec

F1 - Acquisition parameters
ND0 2
TD 256
SFO1 100.6429 MHz
FIDRES 65.118813 Hz
SW 165.639 ppm
FnMODE Echo-Antiecho

F2 - Processing parameters
SI 1024
SF 400.2199980 MHz
WDW QSINE

HSQC NMR Spectrum of Compound 1 (Chapter 2)




```
Current Data Parameters
NAME      Apr06-2011-GK-Raj
EXPNO                   26
PROCNO                   1

F2 - Acquisition Parameters
Date_           20110408
Time                2.03
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG      hmbcgpndqf
TD                  4096
SOLVENT              DMSO
NS                      8
DS                     16
SWH            5376.344 Hz
FIDRES         1.312584 Hz
AQ             0.3809780 sec
RG                   2050
DW                93.000 usec
DE                 6.50 usec
TE                298.2 K
CNST13         8.0000000
D0            0.00000300 sec
D1            1.50327694 sec
D6            0.06250000 sec
D16           0.00020000 sec
IN0           0.00002235 sec

======== CHANNEL f1 ========
NUC1                  1H
P1                10.00 usec
P2                20.00 usec
PL1               -3.00 dB
PL1W        15.48668575 W
SFO1        400.2227004 MHz

======== CHANNEL f2 ========
NUC2                 13C
P3                 9.00 usec
PL2               -2.00 dB
PL2W        54.14257431 W
SFO2        100.6454466 MHz

====== GRADIENT CHANNEL =====
GPNAM1          SINE.100
GPNAM2          SINE.100
GPNAM3          SINE.100
GPZ1               50.00 %
GPZ2               30.00 %
GPZ3               40.10 %
P16              1000.00 usec

F1 - Acquisition parameters
ND0                     2
TD                    128
SFO1           100.6454 MHz
FIDRES       174.631683 Hz
SW             222.095 ppm
FnMODE                 QF

F2 - Processing parameters
SI                    1024
SF           400.2199980 MHz
WDW                 SINE
SSB                    0
LB                  0.00 Hz
GB                     0
PC                  1.40

F1 - Processing parameters
SI                    1024
MC2                    QF
SF           100.6354390 MHz
WDW                 SINE
SSB                    0
LB                  0.00 Hz
```

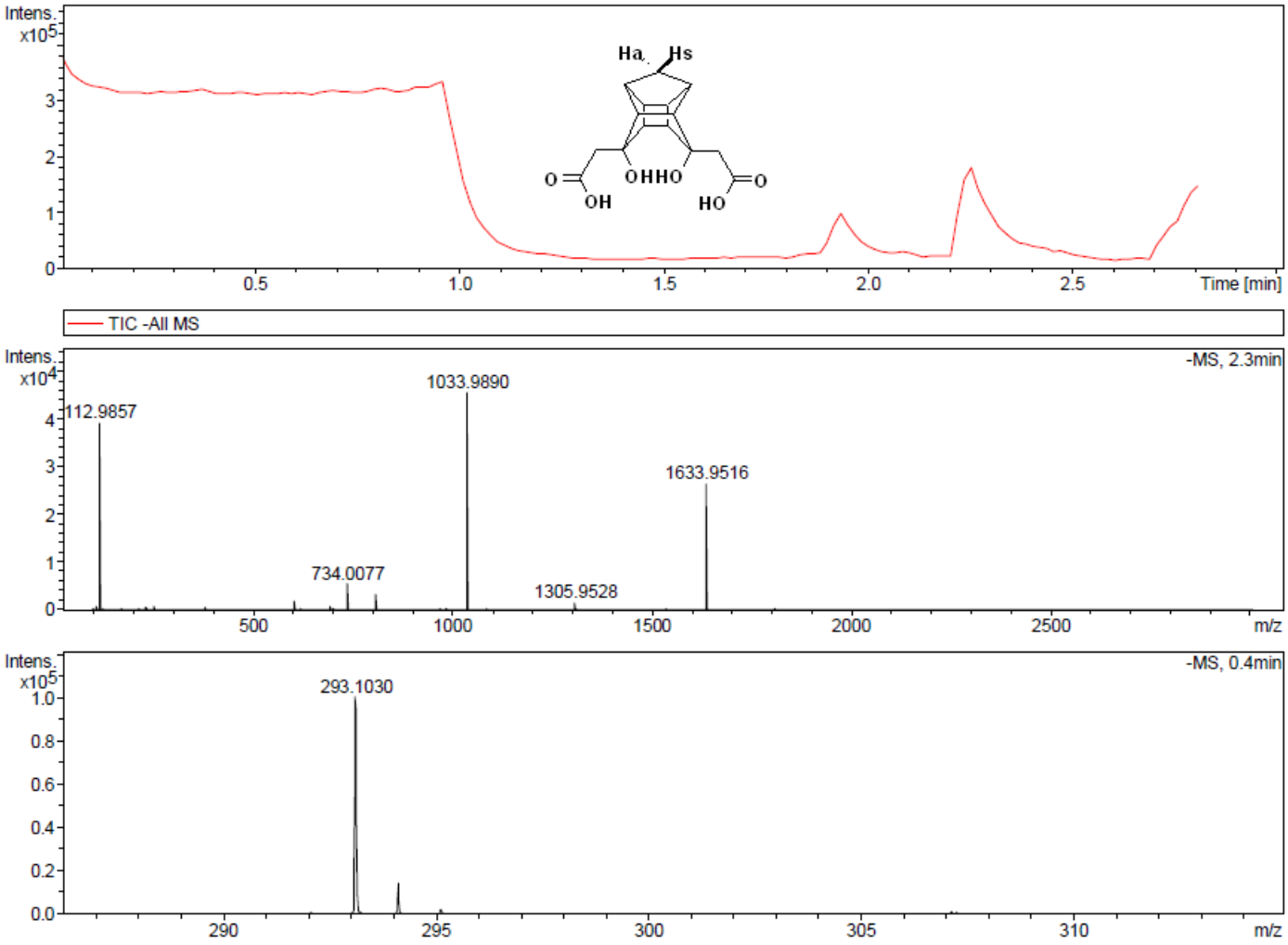
HMBC NMR Spectrum of Compound 1 (Chapter 2)

Analysis Info

Analysis Name	D:\Data\Raj\Raj C15H18O6.d	Acquisition Date	5/19/2011 1:12:59 PM
Method	tune_low_expert.m	Operator	BDAL@DE
Sample Name	Raj C15H18O6	Instrument / Ser#	micrOTOF-Q 10139
Comment	M/Z= 294.11		

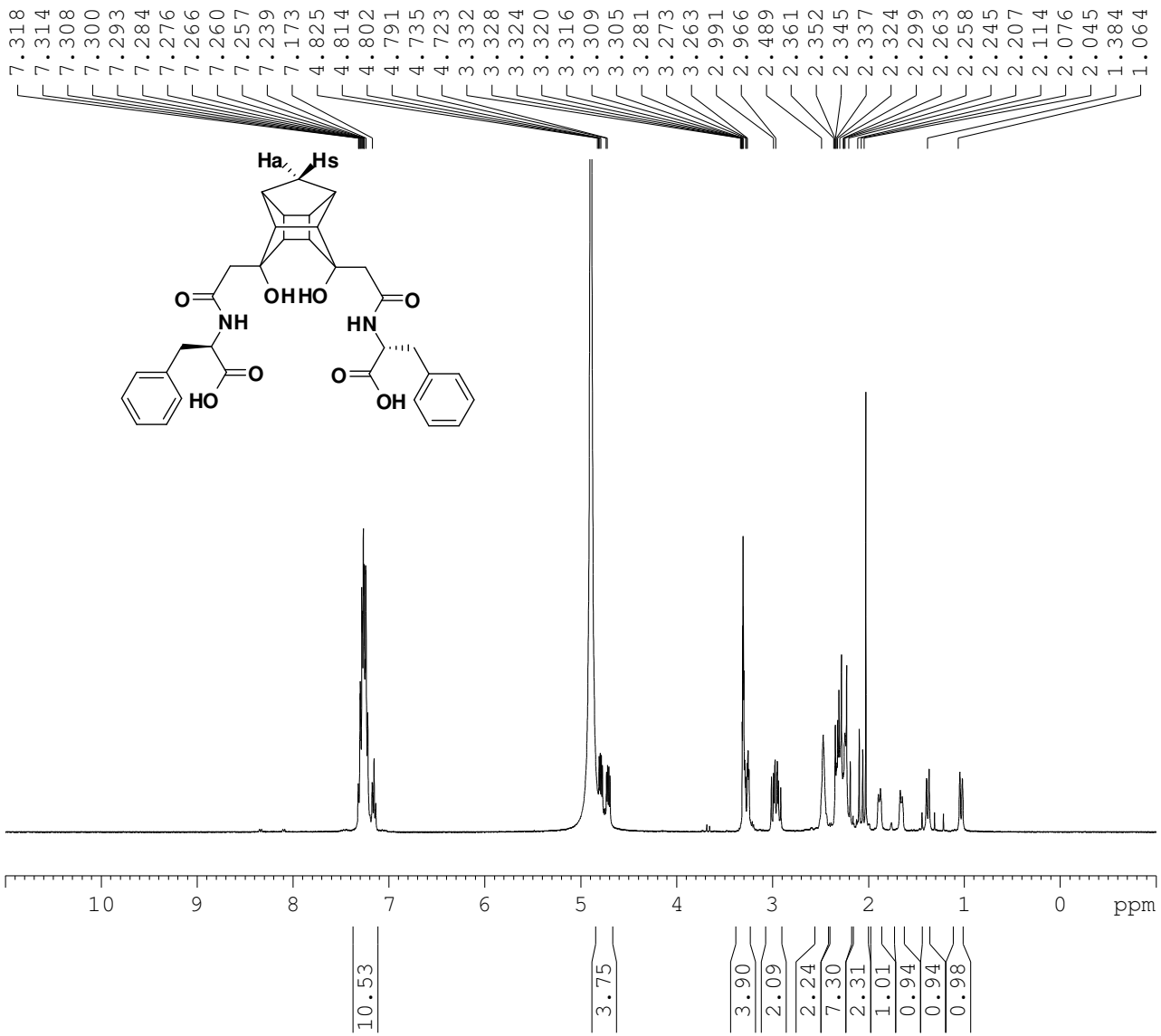
Acquisition Parameter

Source Type	ESI	Ion Polarity	Negative	Set Nebulizer	14.5 psi
Focus	Not active	Set Capillary	6000 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	120.0 Vpp	Set Divert Valve	Source



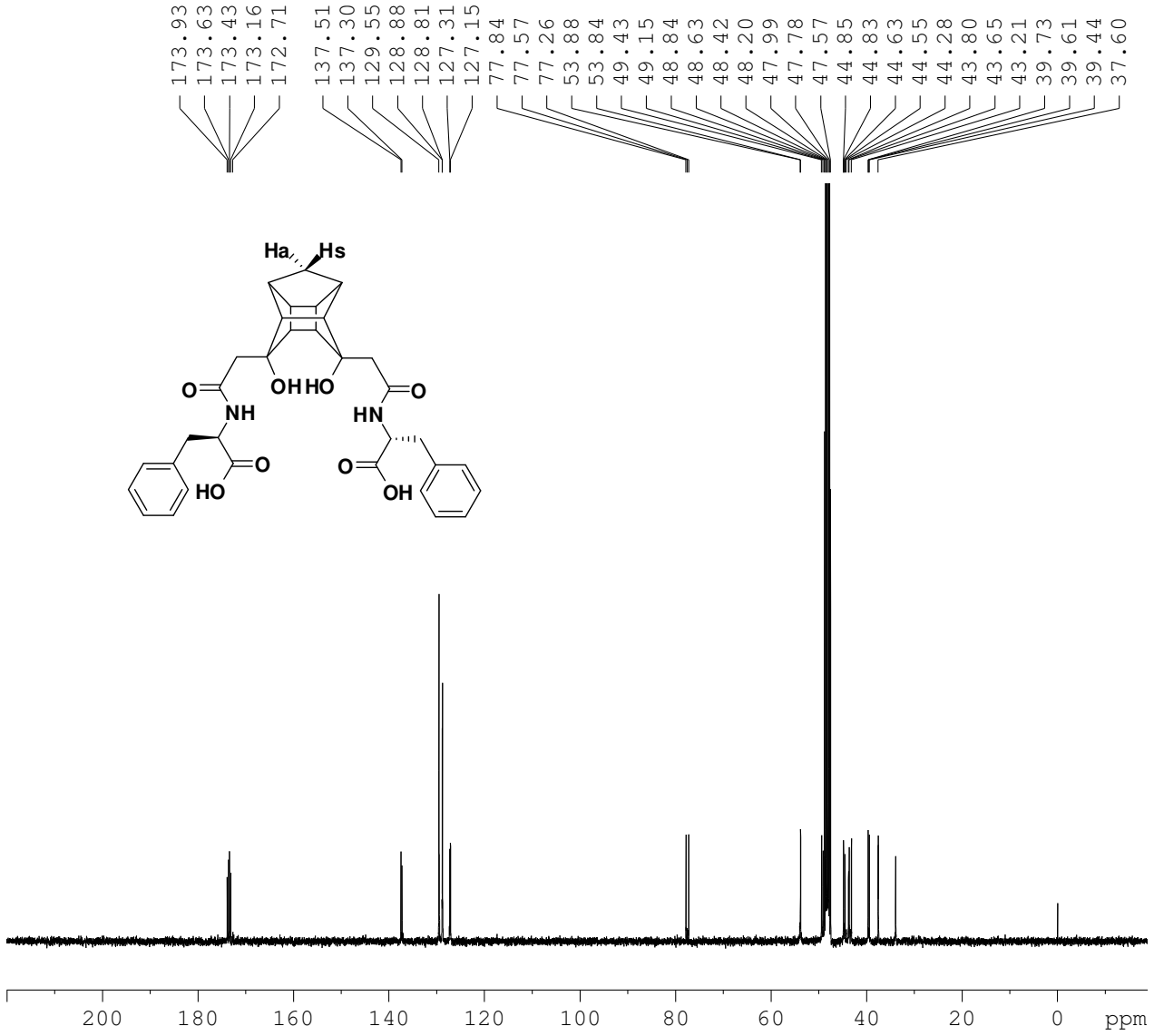
Meas. m/z	#	Formula	m/z	err [ppm]	Mean err [ppm]	rdb	N-Rule	e⁻ Conf	mSigma	Std I	Std Mean m/z	Std I VarNorm	Std m/z Diff	Std Comb Dev
293.1030	1	C 15 H 17 O 6	293.1031	0.2	0.4	7.5	ok	even	11.60	0.0228	0.0002	0.0095	0.0005	0.8427

HRESIMS (*m/z*) of Compound 14(Chapter 1) and Compound 1 (Chapter 2)



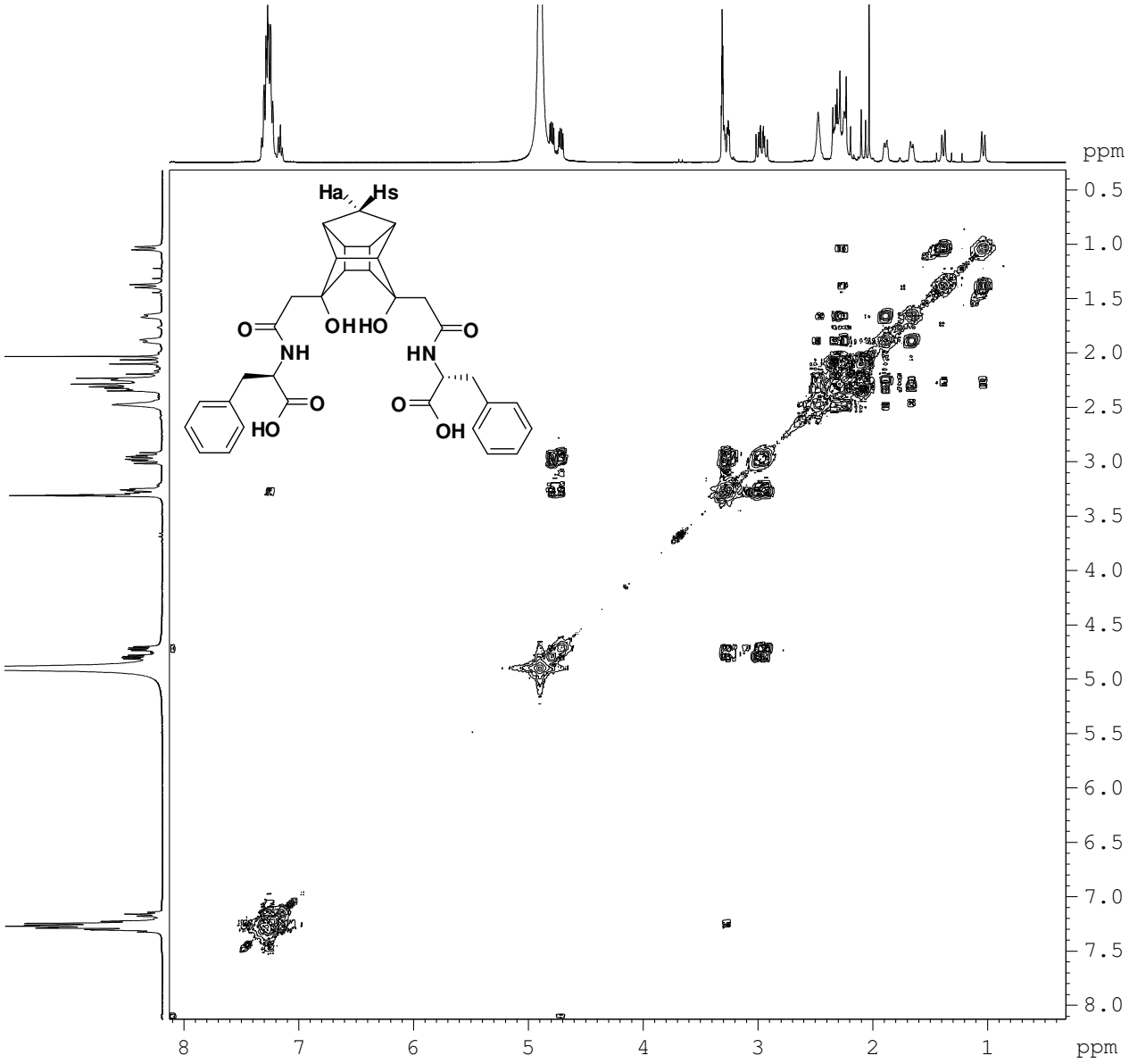

```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                      40
PROCNO                      1

F2 - Acquisition Parameters
Date_              20081106
Time                  17.21
INSTRUM                spect
PROBHD   5 mm PABBO BB-
PULPROG                zg30
TD                    65536
SOLVENT                MeOD
NS                       16
DS                        2
SWH                8223.685 Hz
FIDRES            0.125483 Hz
AQ                3.9846387 sec
RG                      161
DW                   60.800 usec
DE                    6.50 usec
TE                   296.6 K
D1              1.00000000 sec
TD0                       1

======== CHANNEL f1 ========
NUC1                      1H
P1                    10.00 usec
PL1                   -3.00 dB
PL1W           15.48668575 W
SFO1           400.2224715 MHz

F2 - Processing parameters
SI                    32768
SF             400.2200057 MHz
WDW                      EM
SSB                       0
LB                     0.30 Hz
GB                        0
PC                     1.00
```

^{1}H NMR Spectrum of Compound 5 (Chapter 1) and Compound 2 (Chapter 2)



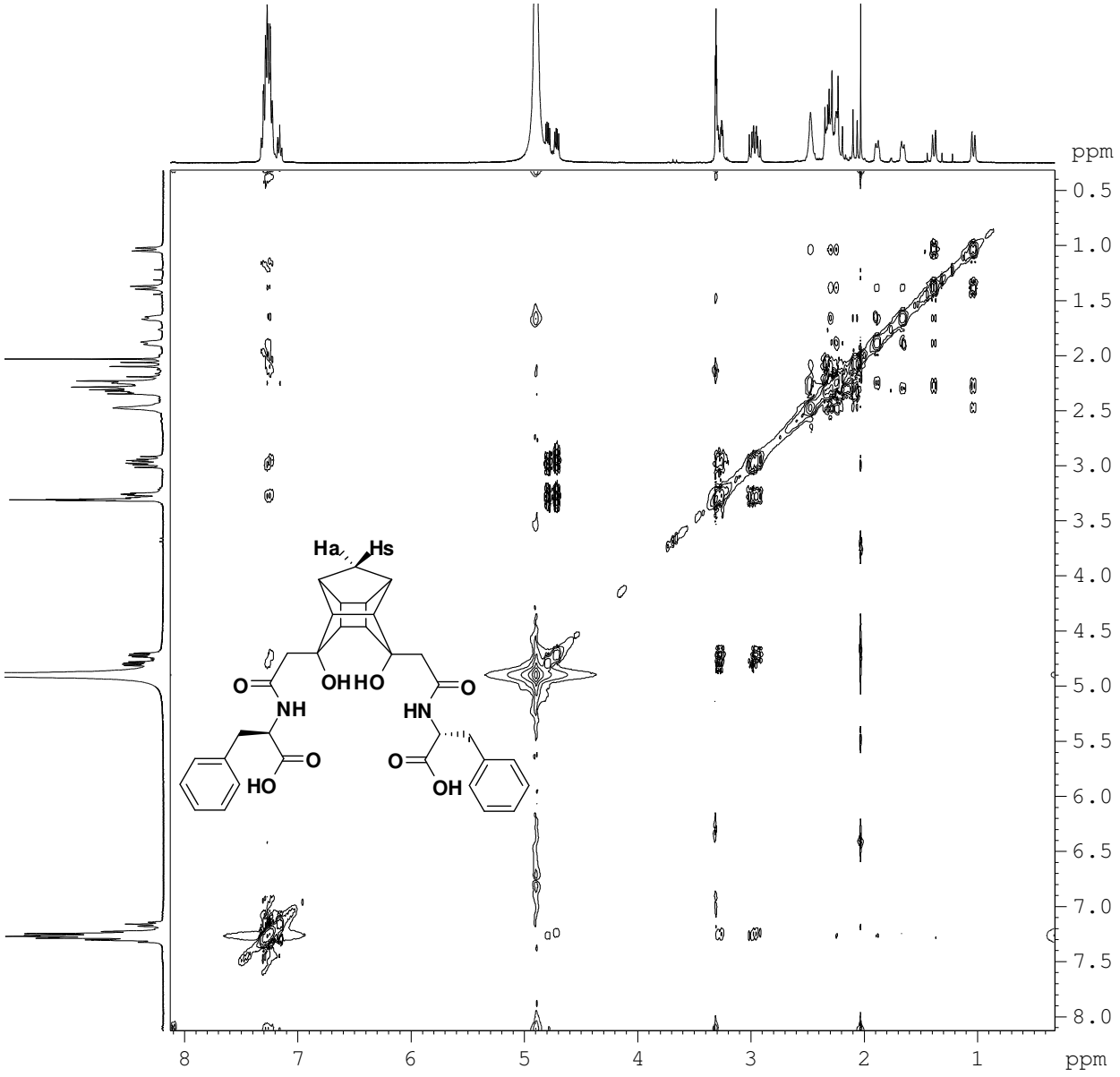

```
Current Data Parameters
NAME      Nov07-2008-GK-Raj
EXPNO                    40
PROCNO                    1

F2 - Acquisition Parameters
Date_              20081108
Time                   5.47
INSTRUM               spect
PROBHD   5 mm PABBO BB-
PULPROG              zgpg30
TD                    65536
SOLVENT                MeOD
NS                     3000
DS                        4
SWH               24038.461 Hz
FIDRES             0.366798 Hz
AQ                 1.3631988 sec
RG                      1150
DW                    20.800 usec
DE                     6.50 usec
TE                    295.3 K
D1                2.00000000 sec
D11               0.03000000 sec
TD0                       1

======== CHANNEL f1 ========
NUC1                    13C
P1                     8.40 usec
PL1                   -2.00 dB
PL1W           54.14257431 W
SFO1           100.6454626 MHz

======== CHANNEL f2 ========
CPDPRG2             waltz16
NUC2                    1H
PCPD2                90.00 usec
PL2                  -3.00 dB
PL12                 15.60 dB
PL13                 18.00 dB
PL2W           15.48668575 W
PL12W           0.21377575 W
PL13W           0.12301511 W
SFO2           400.2216009 MHz

F2 - Processing parameters
SI                    32768
SF           100.6353389 MHz
WDW                      EM
SSB                       0
LB                    1.00 Hz
GB                        0
PC                    1.40
```

^{13}C NMR Spectrum of Compound 5 (Chapter 1) and Compound 2 (Chapter 2)



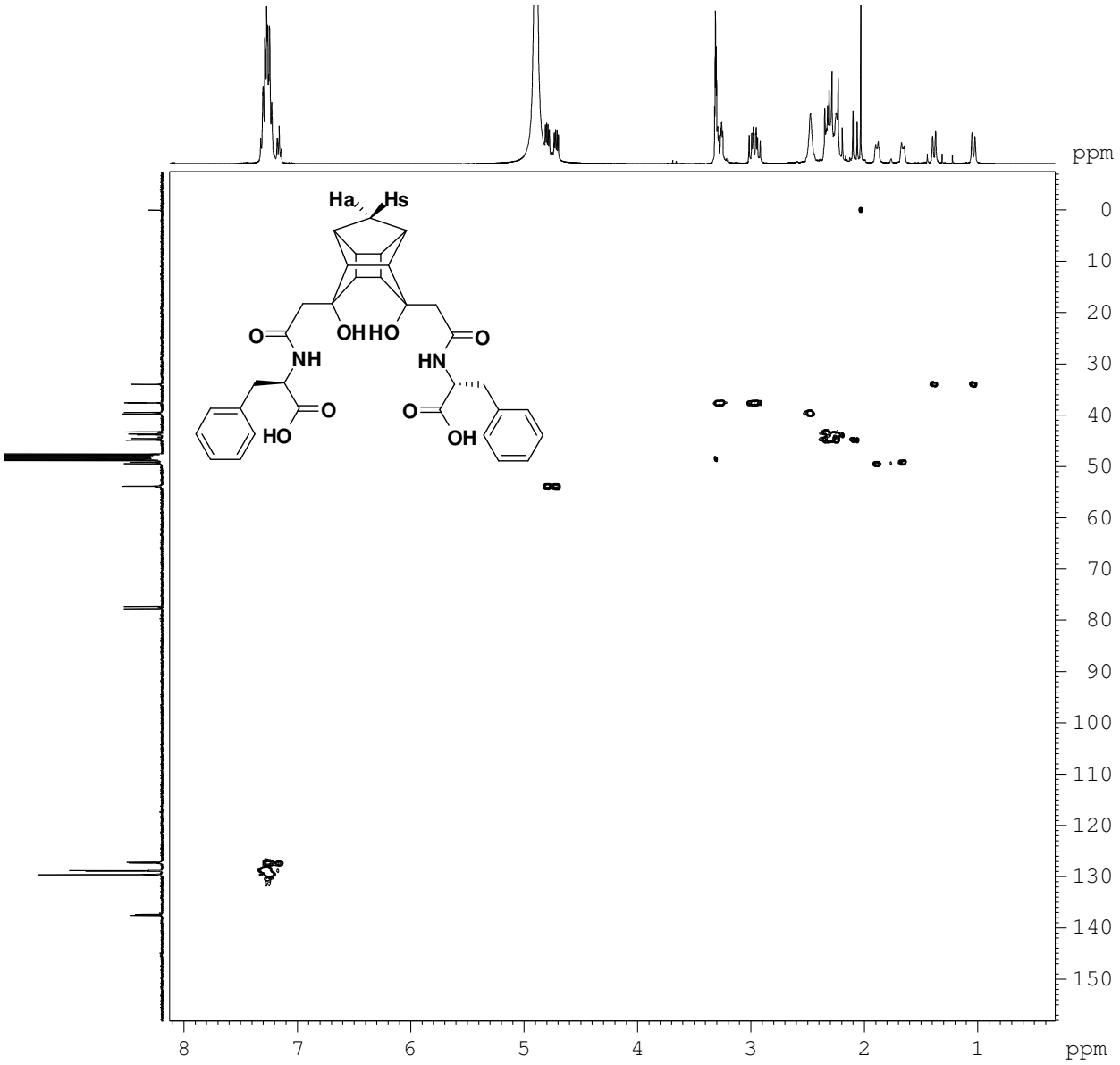

```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                  41
PROCNO                  1

F2 - Acquisition Parameters
Date_            20081106
Time                17.22
INSTRUM             spect
PROBHD    5 mm PABBO BB-
PULPROG          cosygpqf
TD                   2048
SOLVENT               MeOD
NS                       2
DS                       8
SWH              3125.000 Hz
FIDRES           1.525879 Hz
AQ               0.3277300 sec
RG                      64
DW                160.000 usec
DE                  6.50 usec
TE                 296.6 K
D0               0.00000300 sec
D1               1.35090494 sec
D13              0.00000400 sec
D16              0.00020000 sec
IN0              0.00032000 sec

======== CHANNEL f1 ========
NUC1                    1H
P0                   10.00 usec
P1                   10.00 usec
PL1                  -3.00 dB
PL1W          15.48668575 W
SFO1          400.2216946 MHz

====== GRADIENT CHANNEL =====
GPNAM1            SINE.100
GPZ1                10.00 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                      1
TD                     128
SFO1             400.2217 MHz
FIDRES          24.414063 Hz
SW                 7.808 ppm
FnMODE                  QF

F2 - Processing parameters
SI                    1024
SF            400.2200050 MHz
WDW                   SINE
SSB                      0
LB                    0.00 Hz
GB                       0
PC                    1.40

F1 - Processing parameters
SI                    1024
MC2                     QF
SF            400.2200050 MHz
WDW                   SINE
SSB                      0
LB                    0.00 Hz
GB                       0
```

COSY NMR Spectrum of Compound 2 (Chapter 2)



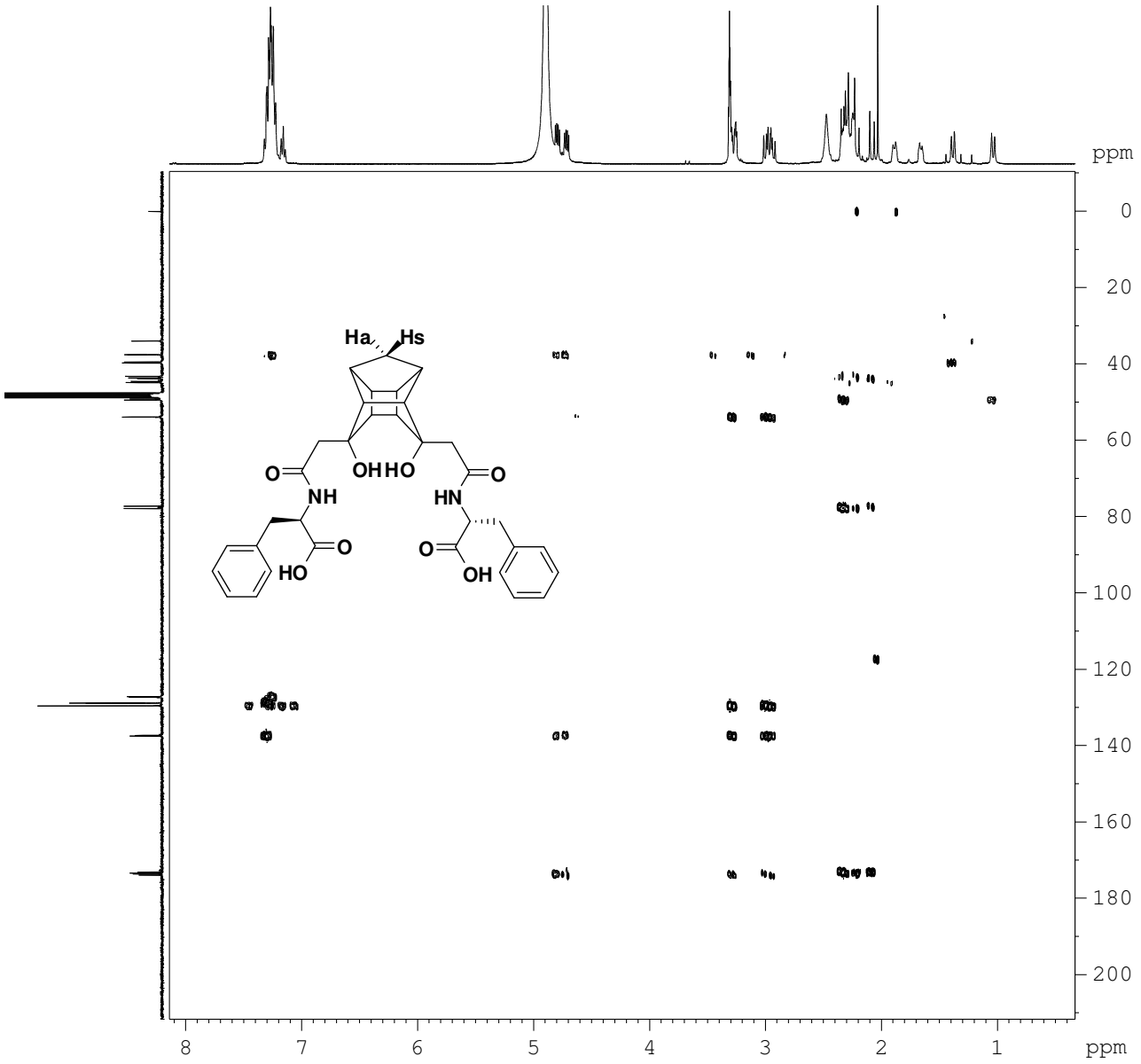

```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                  42
PROCNO                  1

F2 - Acquisition Parameters
Date_            20081106
Time                17.31
INSTRUM             spect
PROBHD    5 mm PABBO BB-
PULPROG           noesyph
TD                   2048
SOLVENT              MeOD
NS                      4
DS                      4
SWH              3125.000 Hz
FIDRES           1.525879 Hz
AQ               0.3277300 sec
RG                     64
DW                160.000 usec
DE                  6.50 usec
TE                 296.5 K
D0             0.00014727 sec
D1             1.92299497 sec
D8             0.30000001 sec
IN0            0.00032000 sec

======== CHANNEL f1 ========
NUC1                   1H
P1                  10.00 usec
PL1                 -3.00 dB
PL1W          15.48668575 W
SFO1          400.2216946 MHz

F1 - Acquisition parameters
ND0                     1
TD                    256
SFO1            400.2217 MHz
FIDRES          12.207031 Hz
SW                 7.808 ppm
FnMODE       States-TPPI

F2 - Processing parameters
SI                   1024
SF            400.2200050 MHz
WDW                 QSINE
SSB                     2
LB                   0.00 Hz
GB                      0
PC                   1.00

F1 - Processing parameters
SI                   1024
MC2          States-TPPI
SF            400.2200050 MHz
WDW                 QSINE
SSB                     2
LB                   0.00 Hz
GB                      0
```

NOESY NMR Spectrum of Compound 2 (Chapter 2)



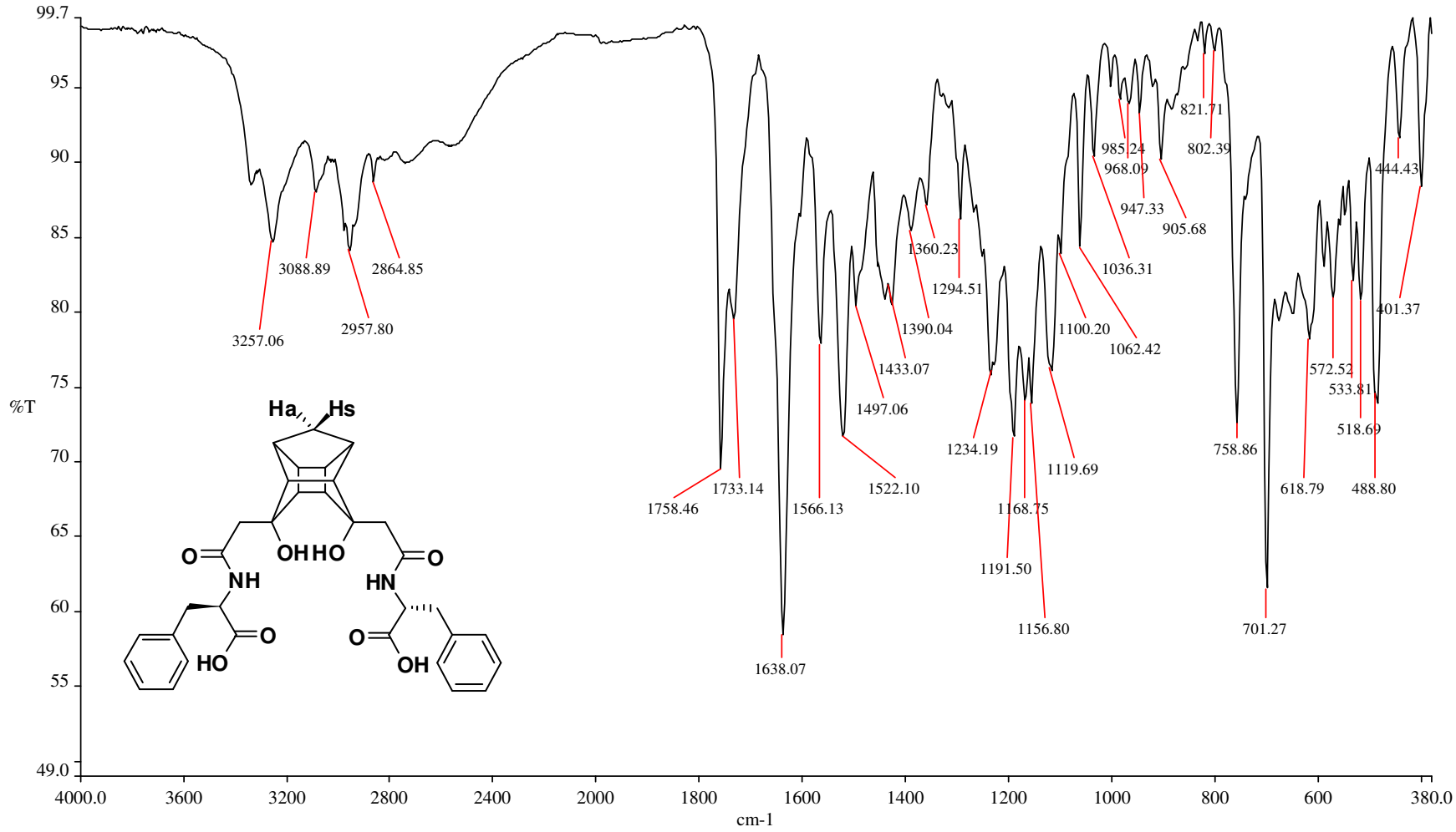

```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                  44
PROCNO                  1

F2 - Acquisition Parameters
Date_           20081106
Time               18.52
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG          hsqcetgp
TD                 1024
SOLVENT             MeOD
NS                     2
DS                    16
SWH             3125.000 Hz
FIDRES          3.051758 Hz
AQ             0.1638900 sec
RG                  2050
DW               160.000 usec
DE                 6.50 usec
TE                296.4 K
CNST2        145.0000000
D0            0.00000300 sec
D1            1.43200600 sec
D4            0.00172414 sec
D11           0.03000000 sec
D13           0.00000400 sec
D16           0.00020000 sec
IN0           0.00003000 sec
ZGOPTNS

======== CHANNEL f1 ========
NUC1                  1H
P1                10.00 usec
P2                20.00 usec
P28             1000.00 usec
PL1                -3.00 dB
PL1W         15.48668575 W
SFO1        400.2216946 MHz

======== CHANNEL f2 ========
CPDPRG2             garp
NUC2                 13C
P3                 9.00 usec
P4                18.00 usec
PCPD2             70.00 usec
PL2                -2.00 dB
PL12               15.60 dB
PL2W         54.14257431 W
PL12W         0.94089001 W
SFO2        100.6429049 MHz

====== GRADIENT CHANNEL =====
GPNAM1          SINE.100
GPNAM2          SINE.100
GPZ1               80.00 %
GPZ2               20.10 %
P16              1000.00 usec

F1 - Acquisition parameters
ND0                    2
TD                   256
SFO1           100.6429 MHz
FIDRES         65.118813 Hz
SW               165.639 ppm
FnMODE    Echo-Antiecho

F2 - Processing parameters
SI                  1024
SF           400.2200050 MHz
WDW                QSINE
SSB                    2
LB                  0.00 Hz
GB                     0
PC                  1.40

F1 - Processing parameters
```

HSQC NMR Spectrum of Compound 2 (Chapter 2)




```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                     43
PROCNO                     1

F2 - Acquisition Parameters
Date_           20081106
Time              18.17
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG       hmbcgpndqf
TD                  4096
SOLVENT             MeOD
NS                     8
DS                    16
SWH             3125.000 Hz
FIDRES          0.762939 Hz
AQ              0.6554100 sec
RG                  2050
DW               160.000 usec
DE                 6.50 usec
TE                296.3 K
CNST13          8.0000000
D0              0.00000300 sec
D1              1.22884500 sec
D6              0.06250000 sec
D16             0.00020000 sec
IN0             0.00002235 sec

======== CHANNEL f1 ========
NUC1                   1H
P1                 10.00 usec
P2                 20.00 usec
PL1                -3.00 dB
PL1W         15.48668575 W
SFO1         400.2216946 MHz

======== CHANNEL f2 ========
NUC2                  13C
P3                  9.00 usec
PL2                -2.00 dB
PL2W         54.14257431 W
SFO2         100.6454466 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPNAM2           SINE.100
GPNAM3           SINE.100
GPZ1                50.00 %
GPZ2                30.00 %
GPZ3                40.10 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                    2
TD                   128
SFO1           100.6454 MHz
FIDRES        174.631683 Hz
SW              222.095 ppm
FnMODE                 QF

F2 - Processing parameters
SI                  1024
SF          400.2200000 MHz
WDW                 SINE
SSB                    0
LB                  0.00 Hz
GB                     0
PC                  1.40

F1 - Processing parameters
SI                  1024
MC2                   QF
SF          100.6353190 MHz
WDW                 SINE
SSB                    0
LB                  0.00 Hz
```

HMBC NMR Spectrum of Compound 2 (Chapter 2)



IR Spectrum of Compound 5 (Chapter 1) and Compound 2 (Chapter 2)

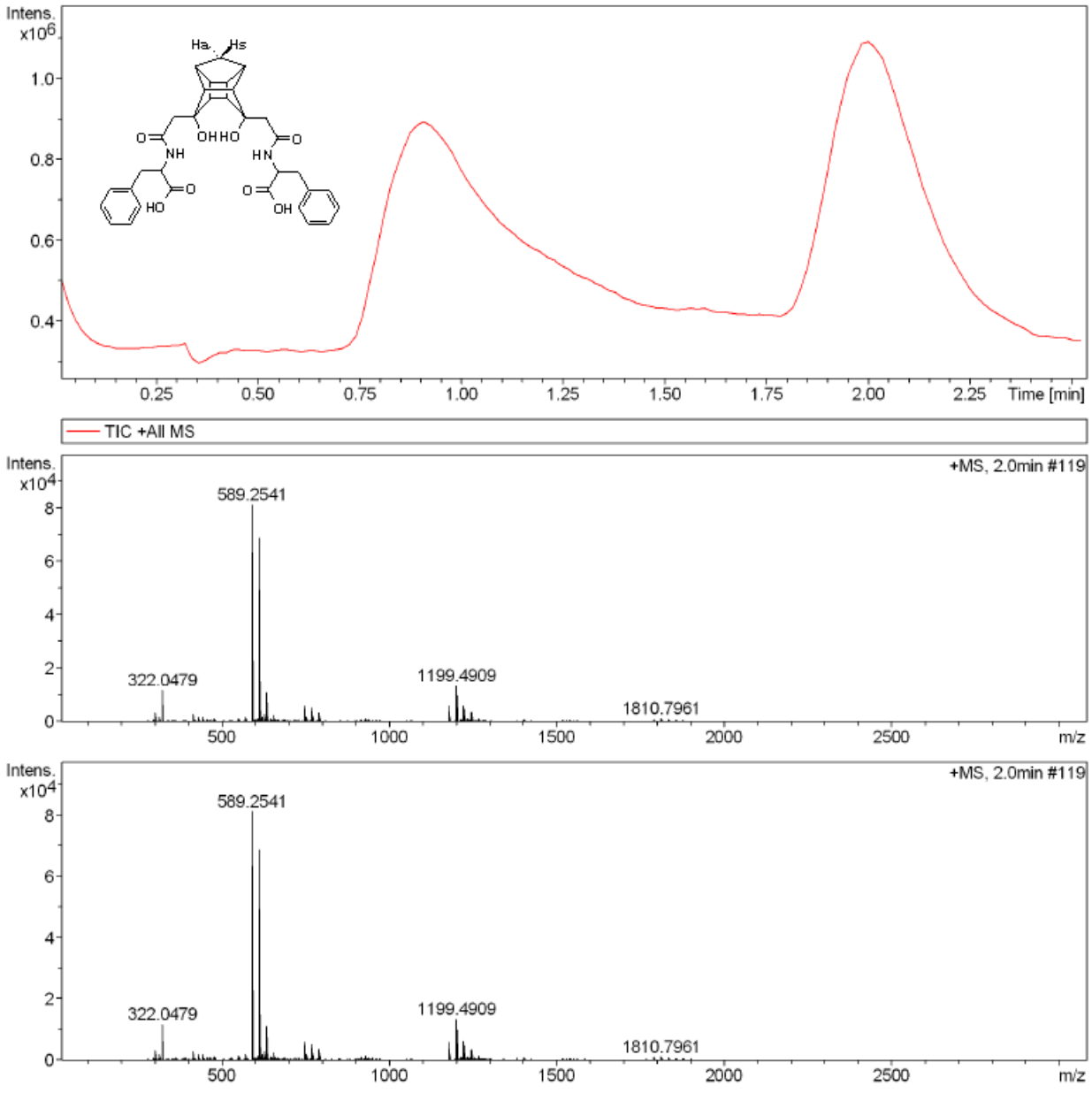
Display Report

Analysis Info

Analysis Name	D:\Data\Raj\ACA-Phe-OH2.d	Acquisition Date	1/1/2010 5:40:47 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	ACA-Phe-OH2	Instrument	micrOTOF-Q 10139
Comment	588.25		

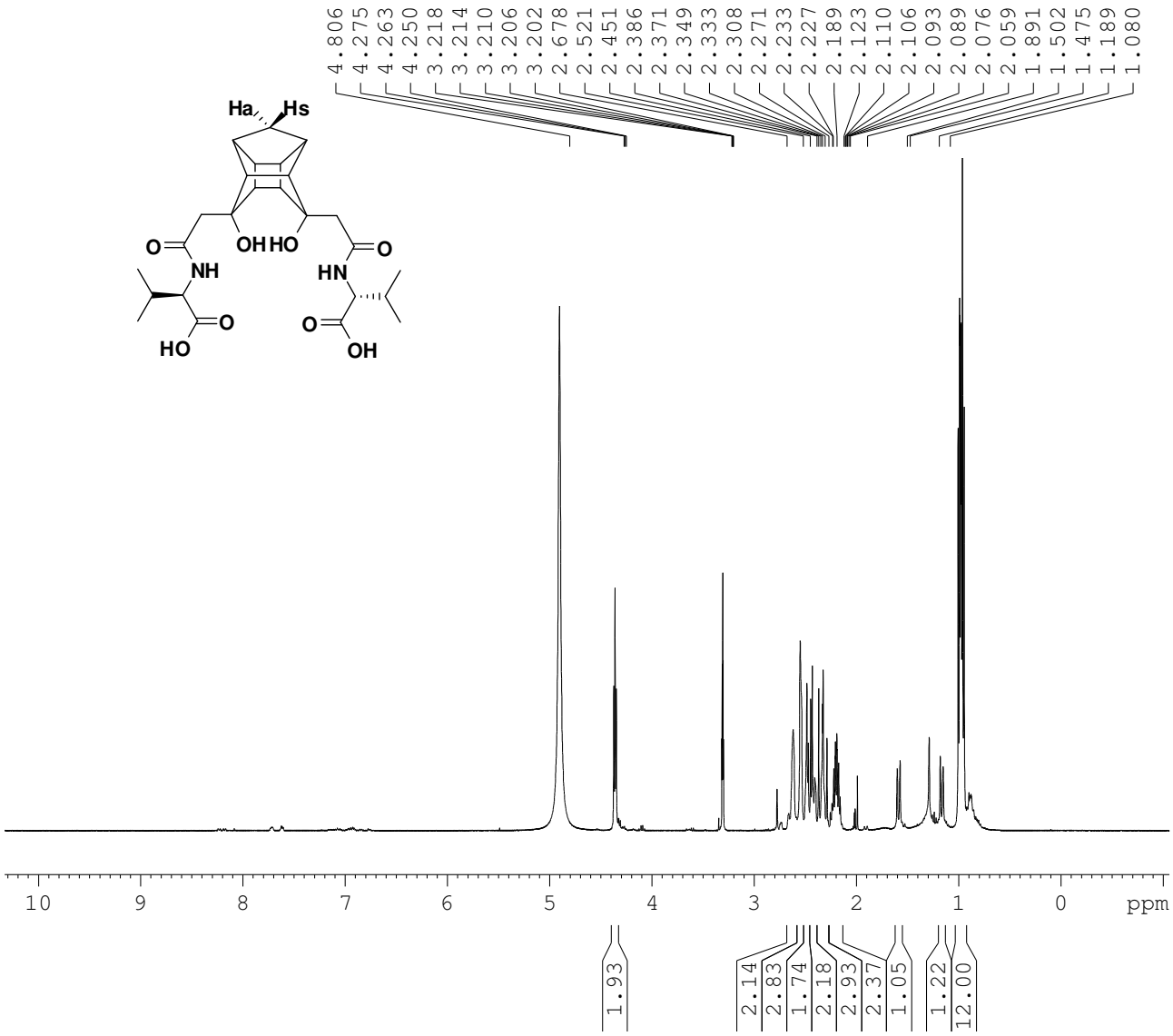
Acquisition Parameter

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



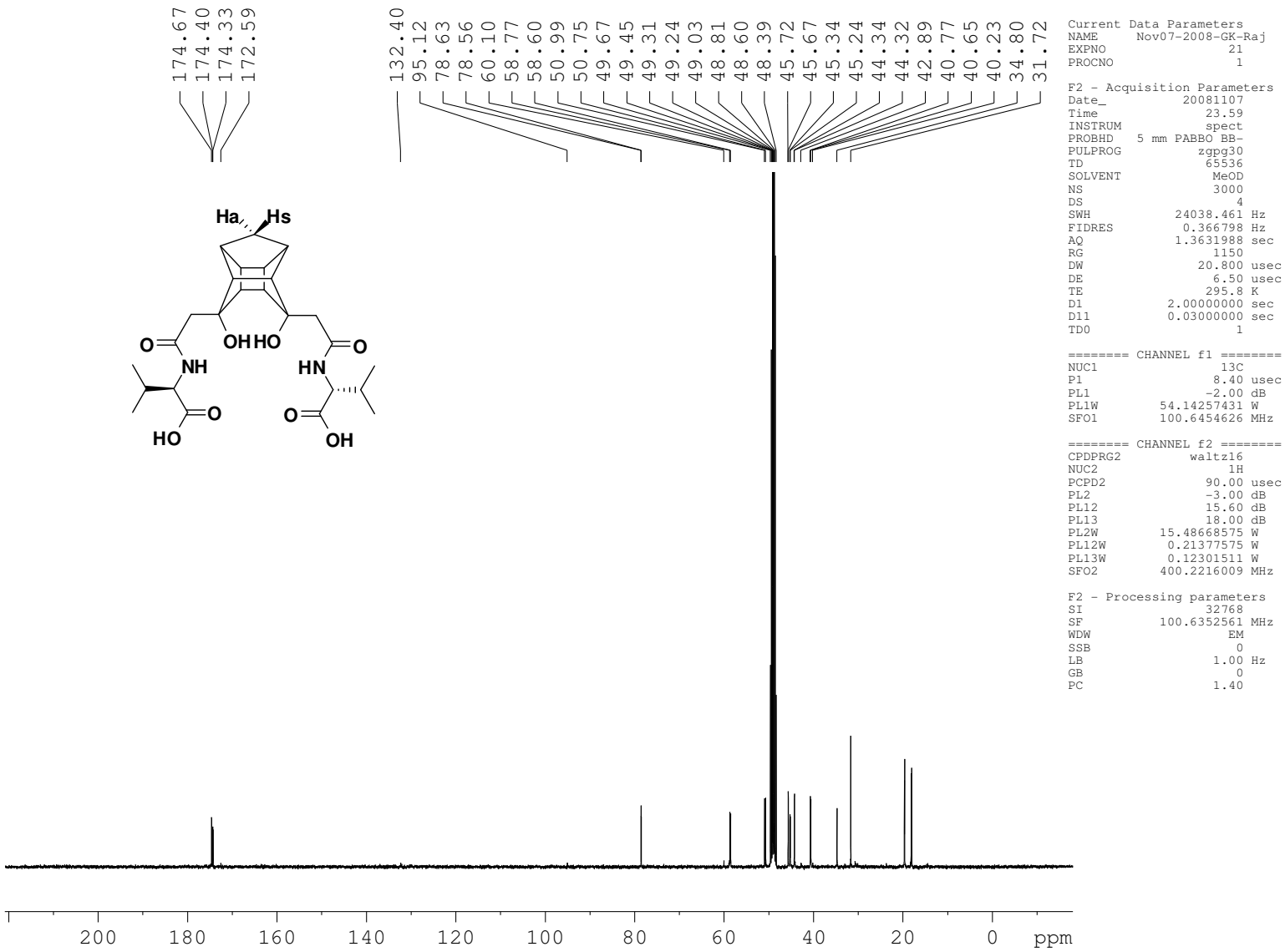
Bruker Compass DataAnalysis 4.0 printed: 1/1/2010 5:47:57 PM Page 1 of 1

HRESIMS (*m/z*) of Compound 5 (Chapter 1) and Compound 2 (Chapter 2)



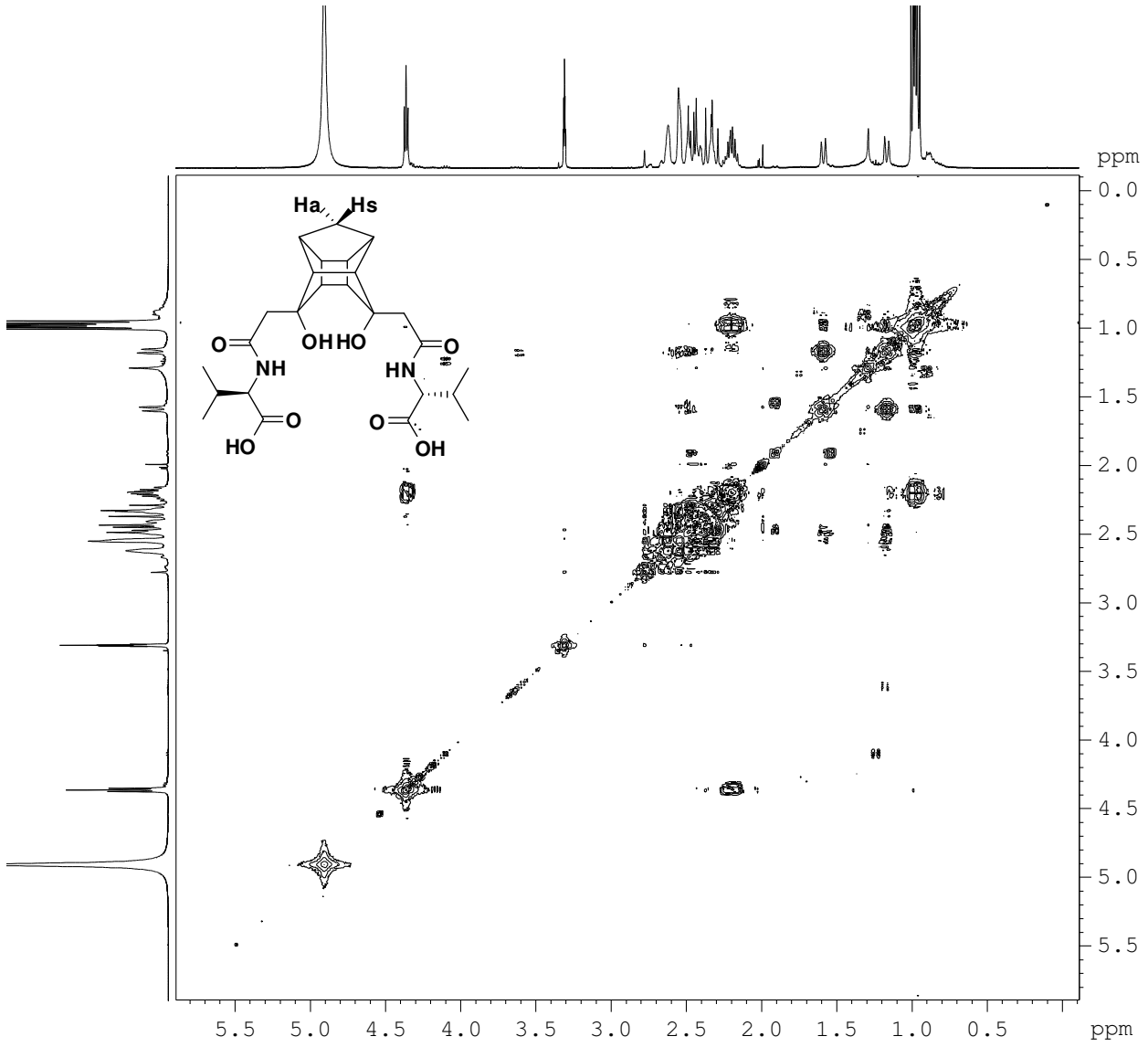

```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                      20
PROCNO                      1

F2 - Acquisition Parameters
Date_              20081106
Time                   9.29
INSTRUM               spect
PROBHD   5 mm PABBO BB-
PULPROG                zg30
TD                    65536
SOLVENT                MeOD
NS                       16
DS                        2
SWH                8223.685 Hz
FIDRES             0.125483 Hz
AQ                 3.9846387 sec
RG                       128
DW                  60.800 usec
DE                   6.50 usec
TE                   295.7 K
D1              1.00000000 sec
TD0                       1

======== CHANNEL f1 ========
NUC1                      1H
P1                   10.00 usec
PL1                  -3.00 dB
PL1W           15.48668575 W
SFO1           400.2224715 MHz

F2 - Processing parameters
SI                    32768
SF             400.2200057 MHz
WDW                      EM
SSB                       0
LB                   0.30 Hz
GB                        0
PC                    1.00
```

^{1}H NMR Spectrum of Compound 6 (Chapter 1) and Compound 3 (Chapter 2)



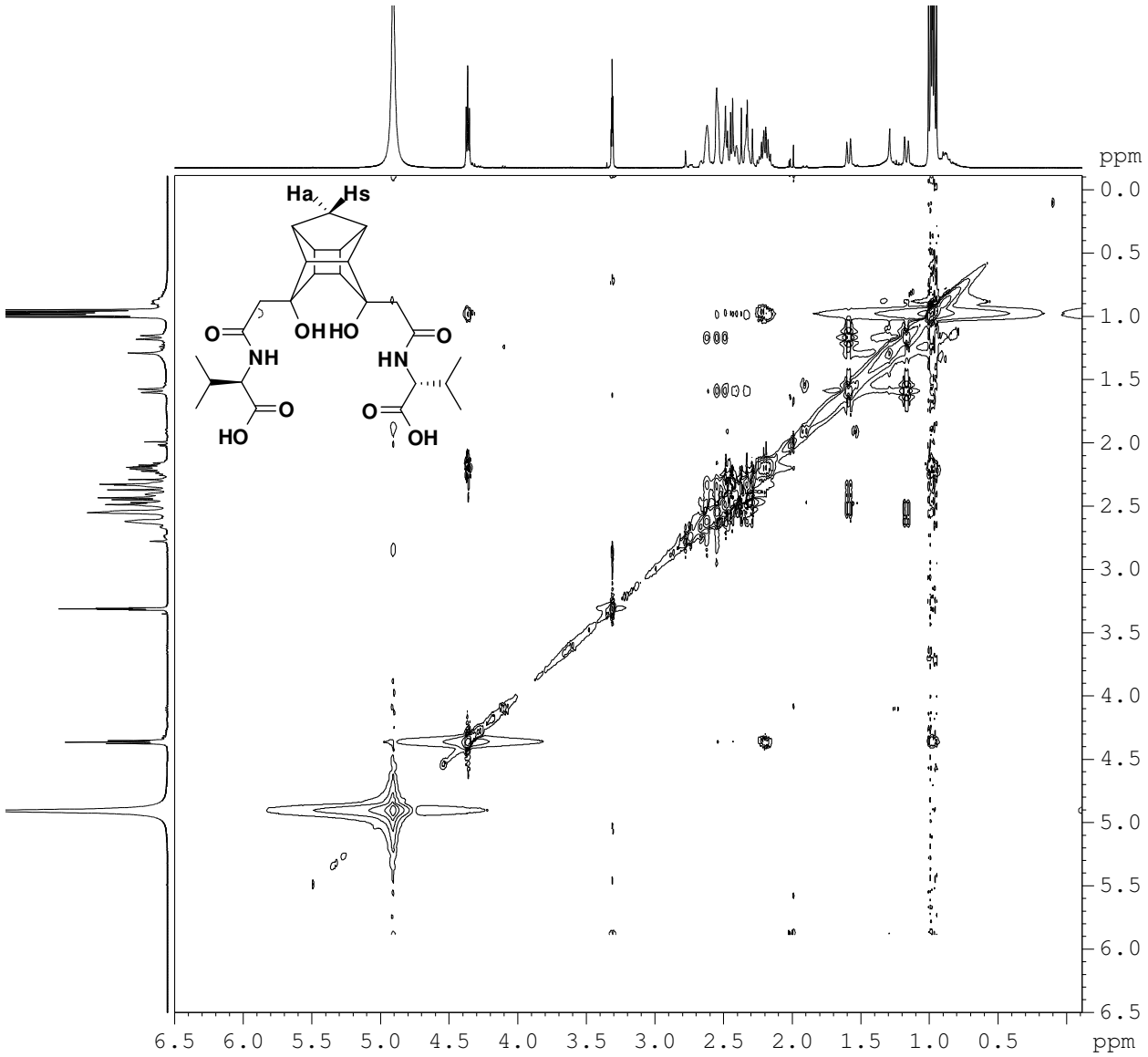

```
Current Data Parameters
NAME      Nov07-2008-GK-Raj
EXPNO                   21
PROCNO                   1

F2 - Acquisition Parameters
Date_             20081107
Time                 23.59
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG             zgpg30
TD                   65536
SOLVENT                MeOD
NS                    3000
DS                       4
SWH              24038.461 Hz
FIDRES               0.366798 Hz
AQ                1.3631988 sec
RG                     1150
DW                   20.800 usec
DE                    6.50 usec
TE                   295.8 K
D1             2.00000000 sec
D11            0.03000000 sec
TD0                      1

======== CHANNEL f1 ========
NUC1                   13C
P1                    8.40 usec
PL1                  -2.00 dB
PL1W          54.14257431 W
SFO1         100.6454626 MHz

======== CHANNEL f2 ========
CPDPRG2             waltz16
NUC2                   1H
PCPD2                90.00 usec
PL2                  -3.00 dB
PL12                 15.60 dB
PL13                 18.00 dB
PL2W          15.48668575 W
PL12W          0.21377575 W
PL13W          0.12301511 W
SFO2         400.2216009 MHz

F2 - Processing parameters
SI                   32768
SF           100.6352561 MHz
WDW                     EM
SSB                      0
LB                    1.00 Hz
GB                       0
PC                    1.40
```

^{13}C NMR Spectrum of Compound 6 (Chapter 1) and Compound 3 (Chapter 2)



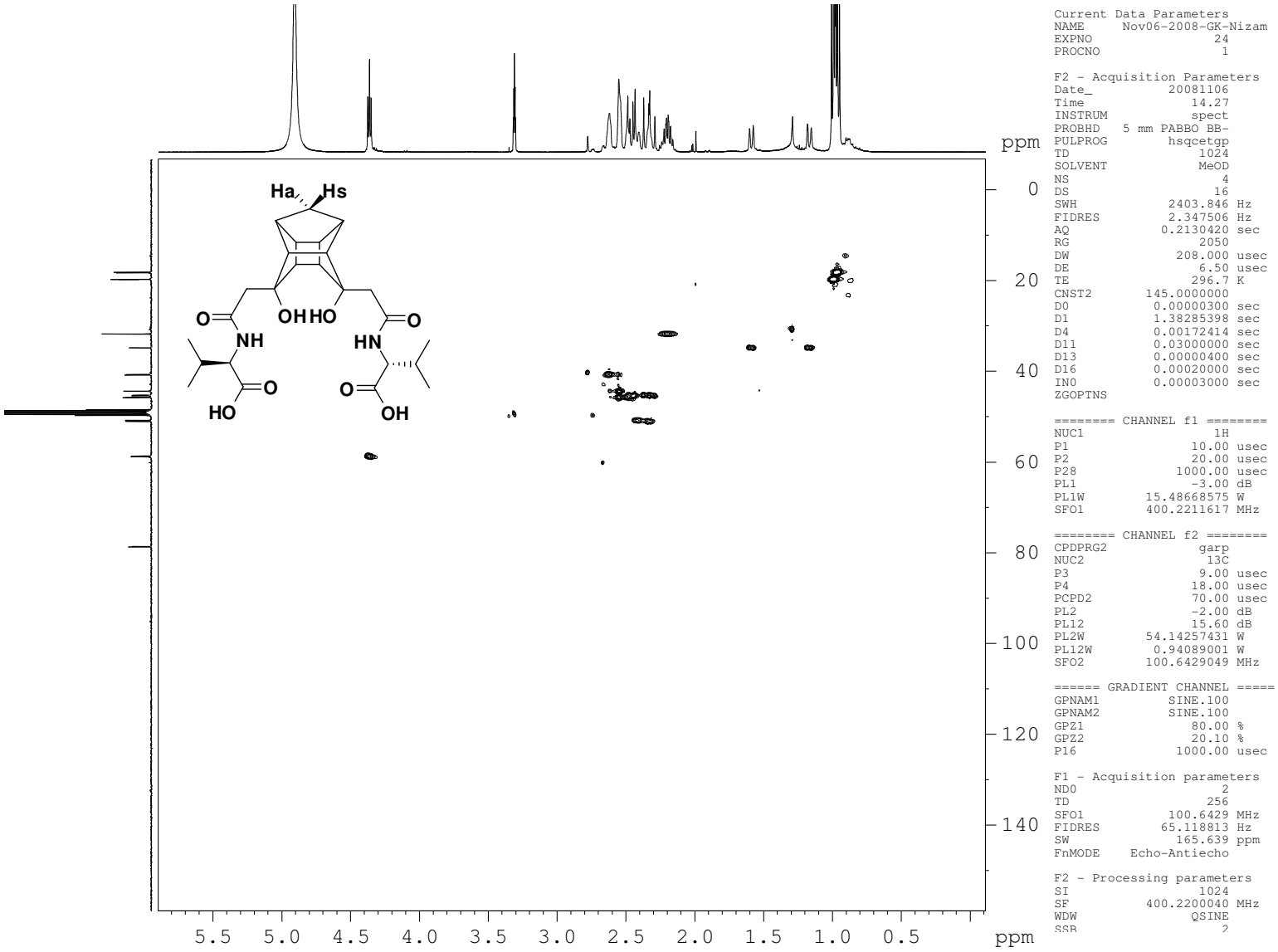

```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                  21
PROCNO                  1

F2 - Acquisition Parameters
Date_            20081106
Time                9.30
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG         cosygpqf
TD                 2048
SOLVENT             MeOD
NS                    2
DS                    8
SWH             2403.846 Hz
FIDRES          1.173753 Hz
AQ              0.4260340 sec
RG                   64
DW               208.000 usec
DE                 6.50 usec
TE                295.7 K
D0             0.00000300 sec
D1             1.25260103 sec
D13            0.00000400 sec
D16            0.00020000 sec
IN0            0.00041600 sec

======== CHANNEL f1 ========
NUC1                 1H
P0                10.00 usec
P1                10.00 usec
PL1                -3.00 dB
PL1W        15.48668575 W
SFO1       400.2211617 MHz

====== GRADIENT CHANNEL =====
GPNAM1          SINE.100
GPZ1               10.00 %
P16              1000.00 usec

F1 - Acquisition parameters
ND0                   1
TD                  128
SFO1           400.2212 MHz
FIDRES         18.780048 Hz
SW                 6.006 ppm
FnMODE               QF

F2 - Processing parameters
SI                  1024
SF         400.2200050 MHz
WDW                 SINE
SSB                    0
LB                  0.00 Hz
GB                     0
PC                  1.40

F1 - Processing parameters
SI                  1024
MC2                   QF
SF         400.2200050 MHz
WDW                 SINE
SSB                    0
```

COSY NMR Spectrum of Compound 3 (Chapter 2)



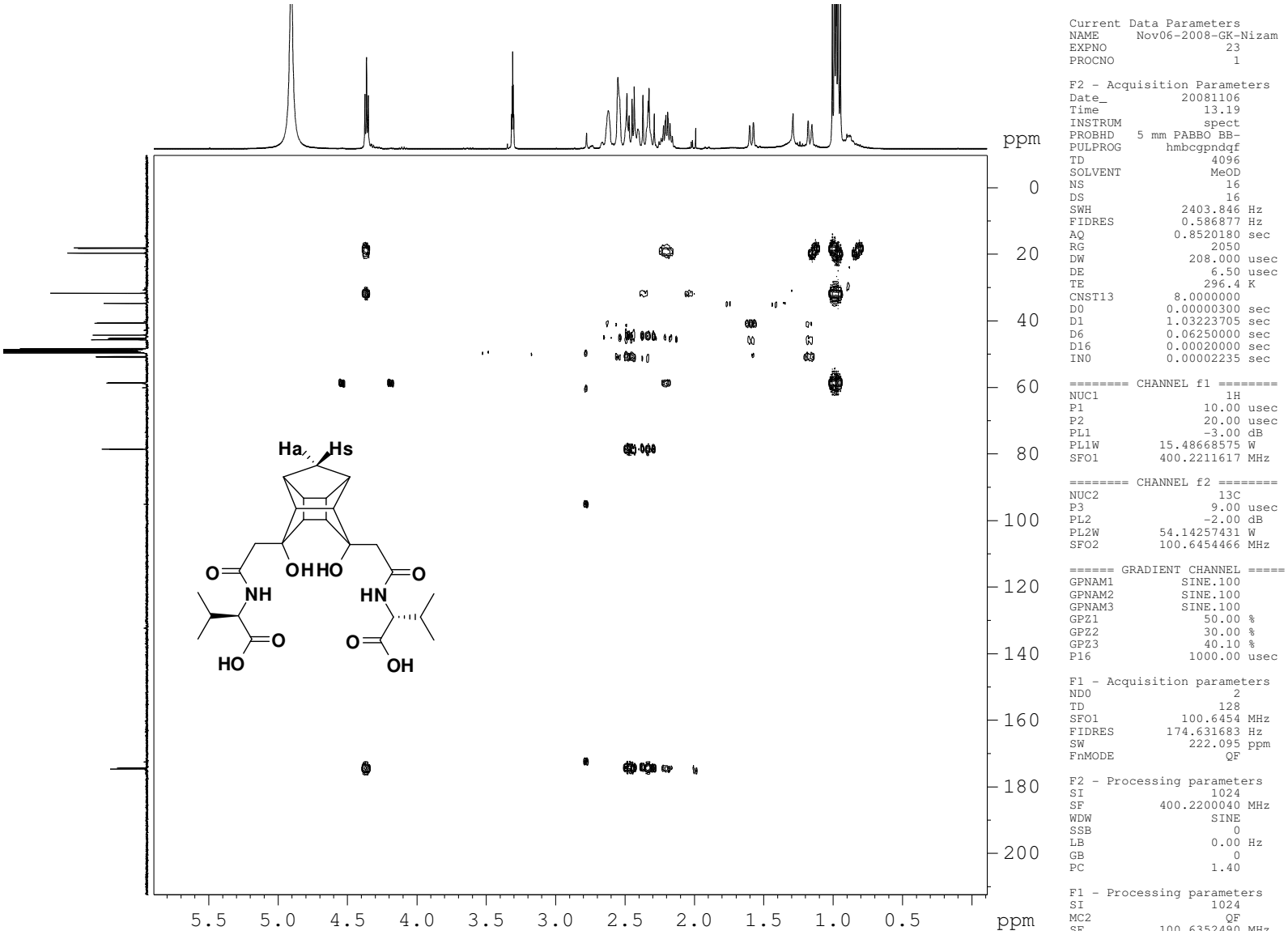

```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                  22
PROCNO                  1

F2 - Acquisition Parameters
Date_              20081106
Time                  9.39
INSTRUM              spect
PROBHD    5 mm PABBO BB-
PULPROG            noesyph
TD                    2048
SOLVENT               MeOD
NS                      16
DS                       4
SWH               2403.846 Hz
FIDRES               1.173753 Hz
AQ                0.4260340 sec
RG                      64
DW                 208.000 usec
DE                   6.50 usec
TE                  295.7 K
D0              0.00019527 sec
D1              1.82469106 sec
D8              0.30000001 sec
IN0             0.00041600 sec

======== CHANNEL f1 ========
NUC1                    1H
P1                  10.00 usec
PL1                 -3.00 dB
PL1W           15.48668575 W
SFO1           400.2211617 MHz

F1 - Acquisition parameters
ND0                      1
TD                     256
SFO1              400.2212 MHz
FIDRES               9.390024 Hz
SW                    6.006 ppm
FnMODE         States-TPPI

F2 - Processing parameters
SI                    1024
SF              400.2200050 MHz
WDW                  QSINE
SSB                      2
LB                    0.00 Hz
GB                       0
PC                    1.00

F1 - Processing parameters
SI                    1024
MC2            States-TPPI
SF              400.2200050 MHz
WDW                  QSINE
SSB                      2
LB                    0.00 Hz
GB                       0
```

NOESY NMR Spectrum of Compound 3 (Chapter 2)



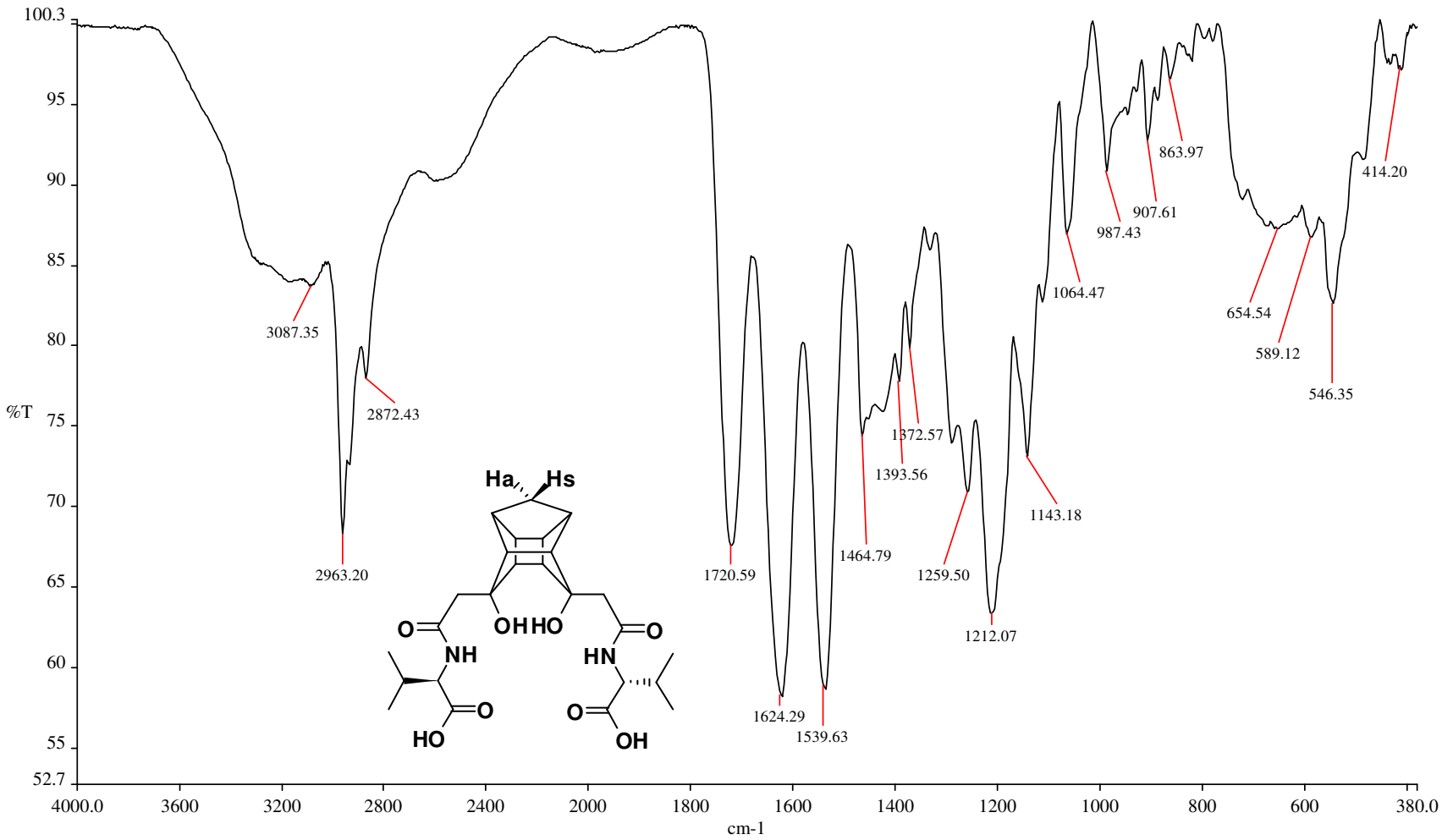

```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                      24
PROCNO                      1

F2 - Acquisition Parameters
Date_              20081106
Time                  14.27
INSTRUM               spect
PROBHD   5 mm PABBO BB-
PULPROG           hsqcetgp
TD                    1024
SOLVENT                MeOD
NS                        4
DS                       16
SWH                2403.846 Hz
FIDRES             2.347506 Hz
AQ                 0.2130420 sec
RG                      2050
DW                   208.000 usec
DE                     6.50 usec
TE                    296.7 K
CNST2             145.0000000
D0                 0.00000300 sec
D1                 1.38285398 sec
D4                 0.00172414 sec
D11                0.03000000 sec
D13                0.00000400 sec
D16                0.00020000 sec
IN0                0.00003000 sec
ZGOPTNS

======== CHANNEL f1 ========
NUC1                       1H
P1                      10.00 usec
P2                      20.00 usec
P28                   1000.00 usec
PL1                     -3.00 dB
PL1W              15.48668575 W
SFO1             400.2211617 MHz

======== CHANNEL f2 ========
CPDPRG2                 garp
NUC2                     13C
P3                       9.00 usec
P4                      18.00 usec
PCPD2                   70.00 usec
PL2                     -2.00 dB
PL12                    15.60 dB
PL2W              54.14257431 W
PL12W              0.94089001 W
SFO2             100.6429049 MHz

====== GRADIENT CHANNEL =====
GPNAM1               SINE.100
GPNAM2               SINE.100
GPZ1                    80.00 %
GPZ2                    20.10 %
P16                   1000.00 usec

F1 - Acquisition parameters
ND0                         2
TD                        256
SFO1             100.6429 MHz
FIDRES            65.118813 Hz
SW                  165.639 ppm
FnMODE    Echo-Antiecho

F2 - Processing parameters
SI                       1024
SF             400.2200040 MHz
WDW                     QSINE
SSB                         2
```

HSQC NMR Spectrum of Compound 3 (Chapter 2)




```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                  23
PROCNO                  1

F2 - Acquisition Parameters
Date_           20081106
Time               13.19
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG       hmbcgpndqf
TD                 4096
SOLVENT             MeOD
NS                   16
DS                   16
SWH             2403.846 Hz
FIDRES          0.586877 Hz
AQ              0.8520180 sec
RG                 2050
DW               208.000 usec
DE                 6.50 usec
TE                296.4 K
CNST13          8.0000000
D0             0.00000300 sec
D1             1.03223705 sec
D6             0.06250000 sec
D16            0.00020000 sec
IN0            0.00002235 sec

======== CHANNEL f1 ========
NUC1                1H
P1                10.00 usec
P2                20.00 usec
PL1               -3.00 dB
PL1W         15.48668575 W
SFO1         400.2211617 MHz

======== CHANNEL f2 ========
NUC2                13C
P3                 9.00 usec
PL2               -2.00 dB
PL2W         54.14257431 W
SFO2         100.6454466 MHz

====== GRADIENT CHANNEL =====
GPNAM1          SINE.100
GPNAM2          SINE.100
GPNAM3          SINE.100
GPZ1               50.00 %
GPZ2               30.00 %
GPZ3               40.10 %
P16              1000.00 usec

F1 - Acquisition parameters
ND0                  2
TD                 128
SFO1          100.6454 MHz
FIDRES        174.631683 Hz
SW               222.095 ppm
FnMODE              QF

F2 - Processing parameters
SI                 1024
SF          400.2200040 MHz
WDW                SINE
SSB                   0
LB                 0.00 Hz
GB                    0
PC                 1.40

F1 - Processing parameters
SI                 1024
MC2                  QF
SF          100.6352490 MHz
```

HMBC NMR Spectrum of Compound 3 (Chapter 2)



IR Spectrum of Compound 6 (Chapter 1) and Compound 3 (Chapter 2)

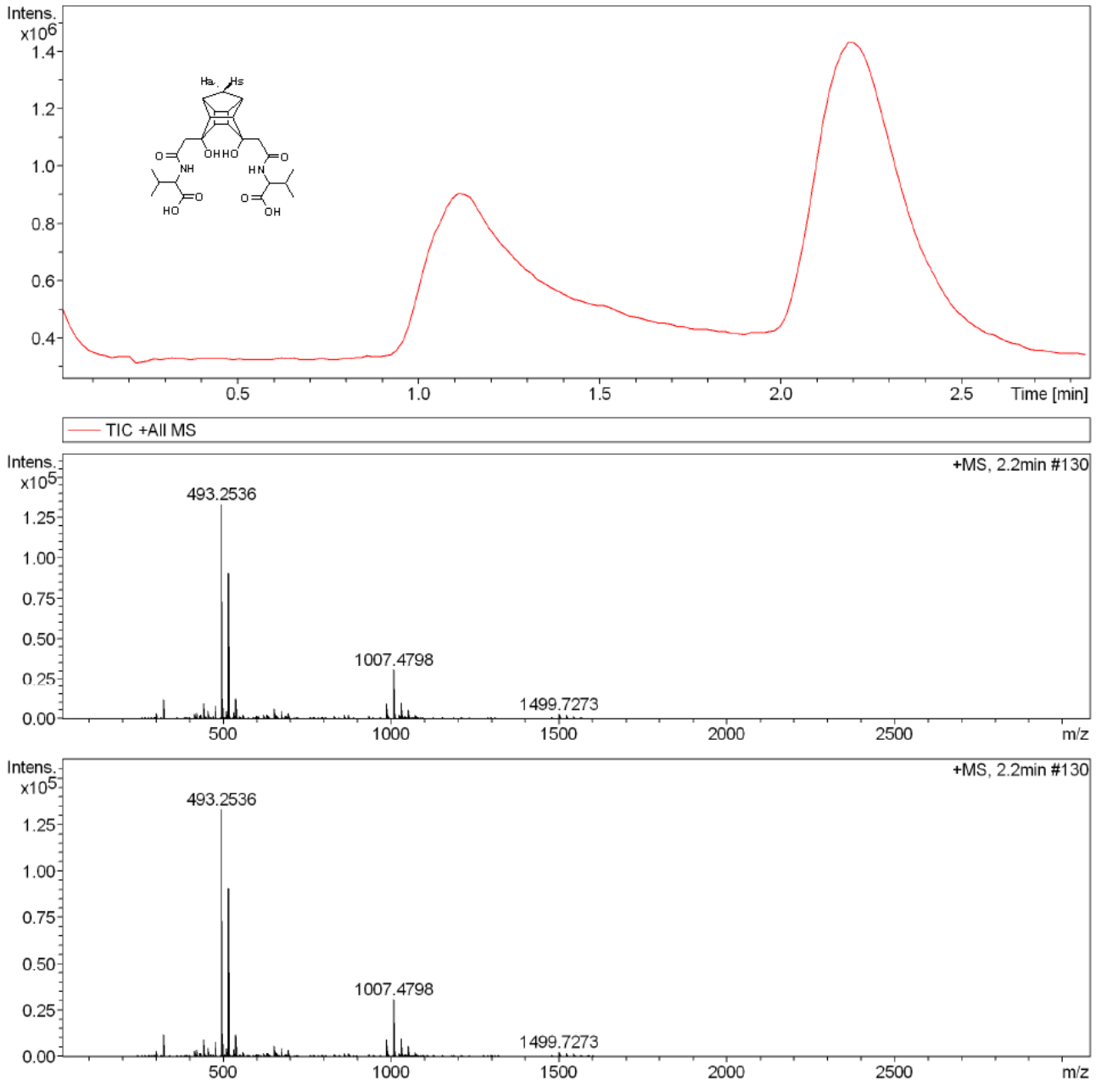
Display Report

Analysis Info

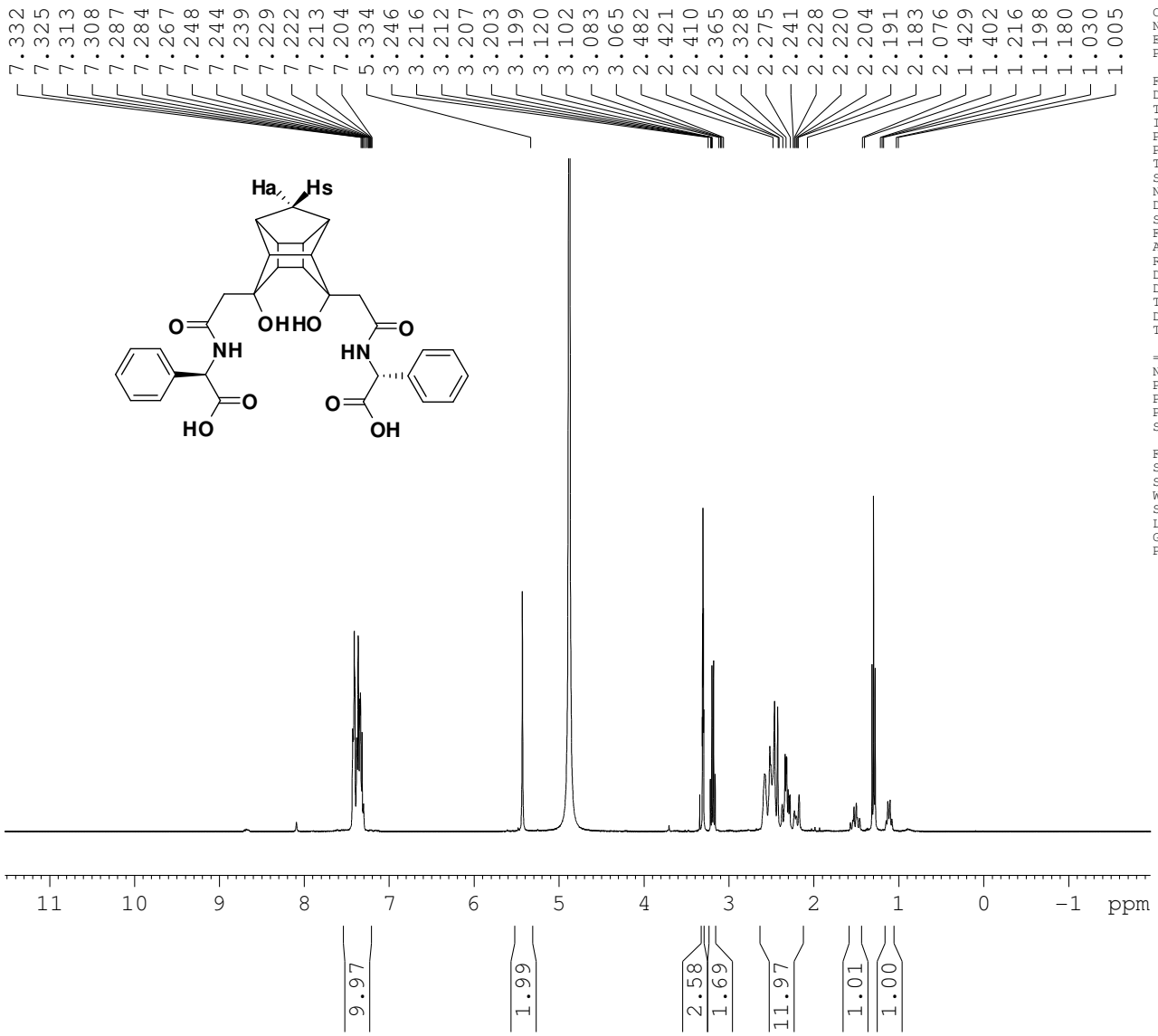
Analysis Name	D:\Data\Raj\ACA-Val-OH1.d	Acquisition Date	1/1/2010 5:49:25 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	ACA-Val-OH1	Instrument	micrOTOF-Q 10139
Comment	492.25		

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

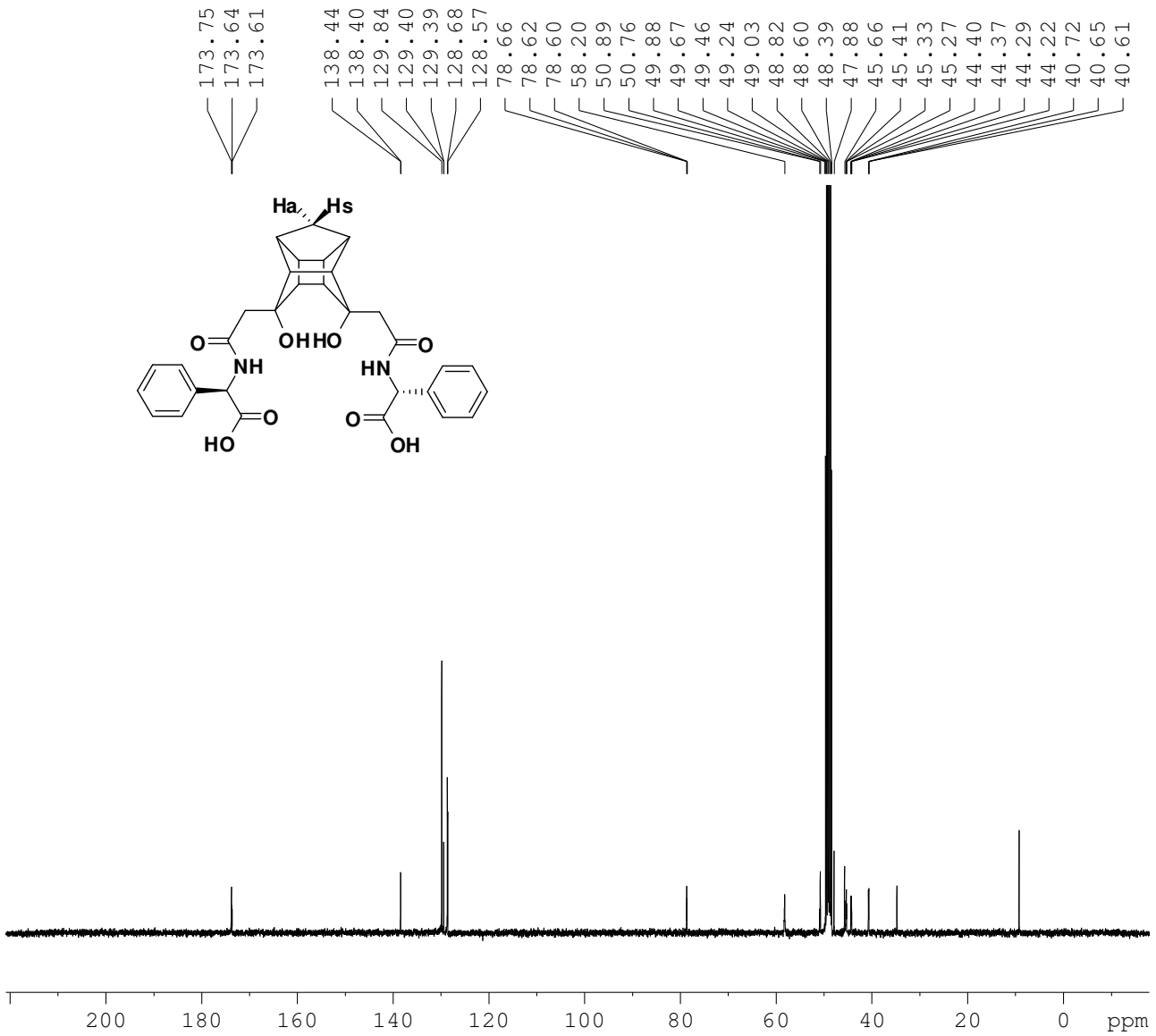


HRESIMS (*m/z*) of Compound 6 (Chapter 1) and Compound 3 (Chapter 2)



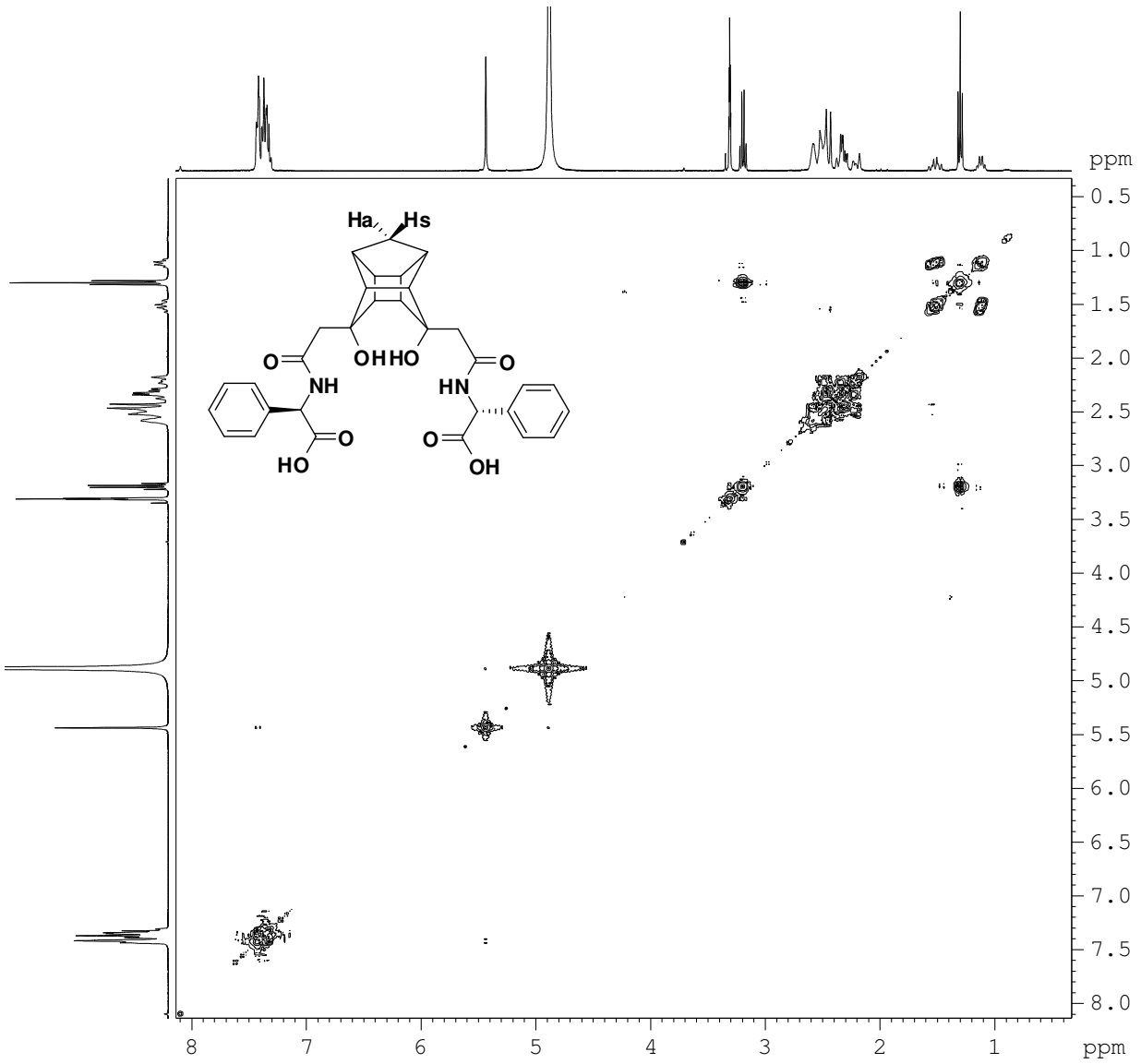

```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                      30
PROCNO                      1

F2 - Acquisition Parameters
Date_                20081106
Time                    14.59
INSTRUM                 spect
PROBHD    5 mm PABBO BB-
PULPROG                 zg30
TD                     65536
SOLVENT                 MeOD
NS                        16
DS                         2
SWH                 8223.685 Hz
FIDRES              0.125483 Hz
AQ                3.9846387 sec
RG                       181
DW                    60.800 usec
DE                     6.50 usec
TE                    296.7 K
D1                1.00000000 sec
TD0                        1

======== CHANNEL f1 ========
NUC1                      1H
P1                     10.00 usec
PL1                    -3.00 dB
PL1W             15.48668575 W
SFO1            400.2224715 MHz

F2 - Processing parameters
SI                     32768
SF              400.2200058 MHz
WDW                       EM
SSB                        0
LB                    0.30 Hz
GB                         0
PC                     1.00
```

^{1}H NMR Spectrum of Compound 7 (Chapter 1) and Compound 4 (Chapter 2)



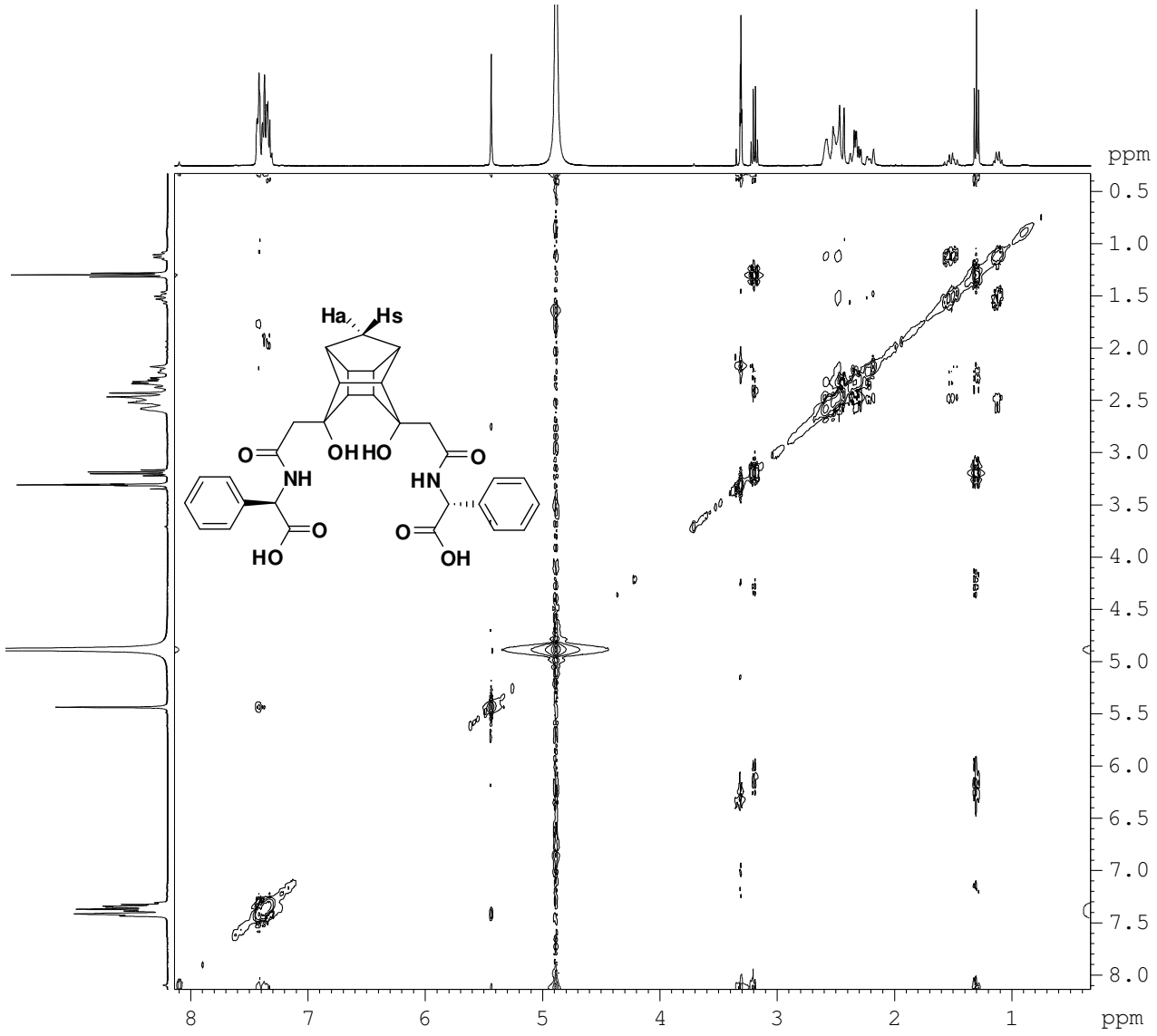

```
Current Data Parameters
NAME      Nov07-2008-GK-Raj
EXPNO                   30
PROCNO                   1

F2 - Acquisition Parameters
Date_            20081108
Time                 2.53
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG             zgpg30
TD                   65536
SOLVENT               MeOD
NS                    3000
DS                       4
SWH             24038.461 Hz
FIDRES           0.366798 Hz
AQ               1.3631988 sec
RG                    1290
DW                 20.800 usec
DE                  6.50 usec
TE                 295.5 K
D1              2.00000000 sec
D11             0.03000000 sec
TD0                      1

======== CHANNEL f1 ========
NUC1                   13C
P1                    8.40 usec
PL1                  -2.00 dB
PL1W          54.14257431 W
SFO1         100.6454626 MHz

======== CHANNEL f2 ========
CPDPRG2          waltz16
NUC2                   1H
PCPD2                90.00 usec
PL2                  -3.00 dB
PL12                 15.60 dB
PL13                 18.00 dB
PL2W          15.48668575 W
PL12W          0.21377575 W
PL13W          0.12301511 W
SFO2         400.2216009 MHz

F2 - Processing parameters
SI                   32768
SF           100.6352487 MHz
WDW                     EM
SSB                      0
LB                    1.00 Hz
GB                       0
PC                    1.40
```

^{13}C NMR Spectrum of Compound 7 (Chapter 1) and Compound 4 (Chapter 2)



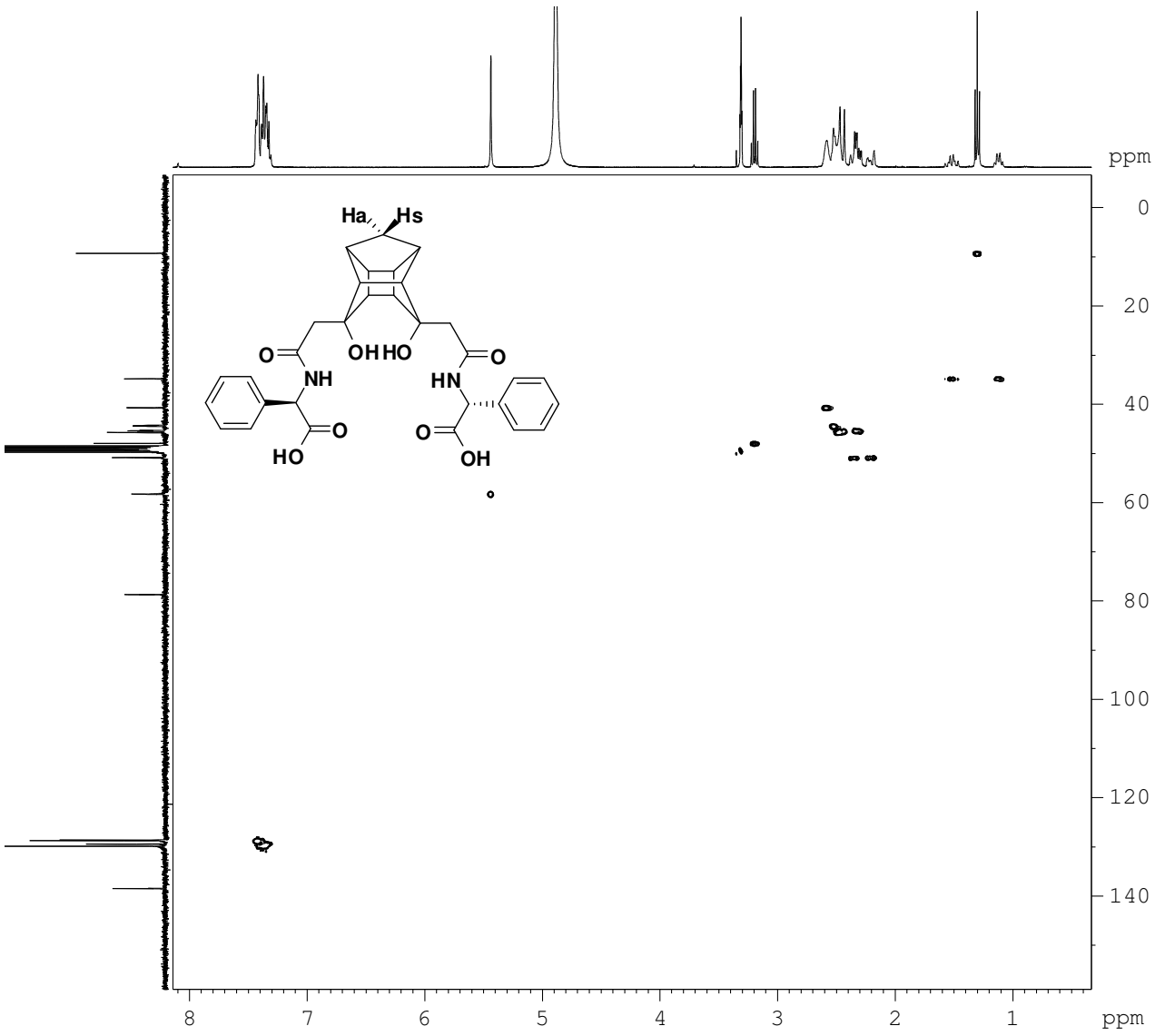

```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                  31
PROCNO                  1

F2 - Acquisition Parameters
Date_           20081106
Time              15.00
INSTRUM           spect
PROBHD   5 mm PABBO BB-
PULPROG         cosygpqf
TD                2048
SOLVENT           MeOD
NS                   2
DS                   8
SWH            3125.000 Hz
FIDRES            1.525879 Hz
AQ             0.3277300 sec
RG                  64
DW               160.000 usec
DE                 6.50 usec
TE               296.6 K
D0             0.00000300 sec
D1             1.35090494 sec
D13            0.00000400 sec
D16            0.00020000 sec
IN0            0.00032000 sec

======== CHANNEL f1 ========
NUC1                1H
P0                10.00 usec
P1                10.00 usec
PL1               -3.00 dB
PL1W         15.48668575 W
SFO1         400.2216991 MHz

====== GRADIENT CHANNEL =====
GPNAM1          SINE.100
GPZ1              10.00 %
P16             1000.00 usec

F1 - Acquisition parameters
ND0                  1
TD                 128
SFO1            400.2217 MHz
FIDRES         24.414063 Hz
SW                7.808 ppm
FnMODE              QF

F2 - Processing parameters
SI                1024
SF           400.2200050 MHz
WDW               SINE
SSB                  0
LB                0.00 Hz
GB                   0
PC                1.40

F1 - Processing parameters
SI                1024
MC2                 QF
SF           400.2200050 MHz
WDW               SINE
SSB                  0
LB                0.00 Hz
```

COSY NMR Spectrum of Compound 4 (Chapter 2)



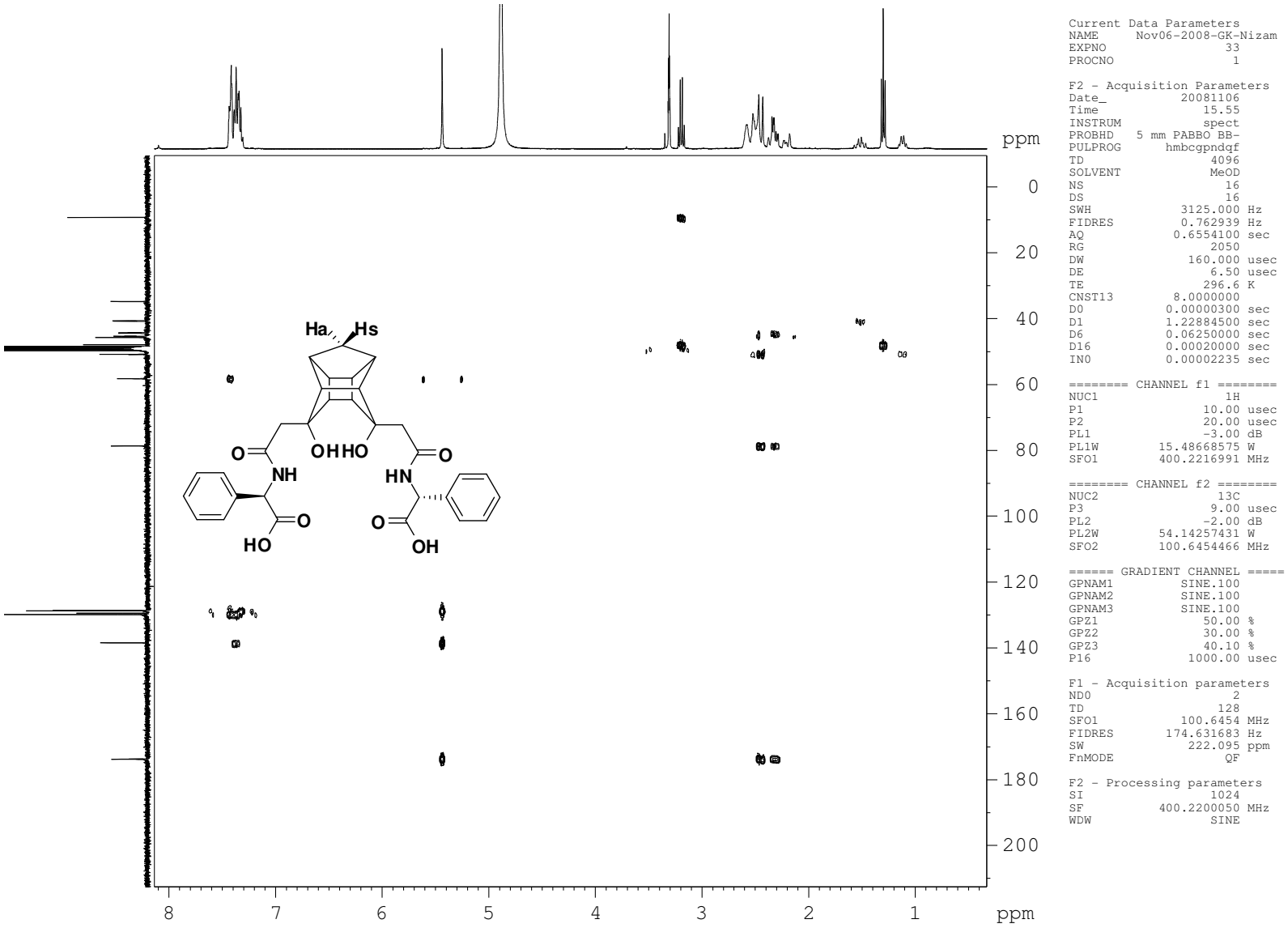

```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                  32
PROCNO                  1

F2 - Acquisition Parameters
Date_            20081106
Time                15.09
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG           noesyph
TD                   2048
SOLVENT               MeOD
NS                       4
DS                       4
SWH              3125.000 Hz
FIDRES              1.525879 Hz
AQ               0.3277300 sec
RG                      64
DW                160.000 usec
DE                  6.50 usec
TE                 296.6 K
D0              0.00014727 sec
D1              1.92299497 sec
D8              0.30000001 sec
IN0             0.00032000 sec

======== CHANNEL f1 ========
NUC1                    1H
P1                  10.00 usec
PL1                 -3.00 dB
PL1W          15.48668575 W
SFO1          400.2216991 MHz

F1 - Acquisition parameters
ND0                     1
TD                    256
SFO1            400.2217 MHz
FIDRES          12.207031 Hz
SW                 7.808 ppm
FnMODE        States-TPPI

F2 - Processing parameters
SI                   1024
SF            400.2200050 MHz
WDW                 QSINE
SSB                     2
LB                   0.00 Hz
GB                      0
PC                   1.00

F1 - Processing parameters
SI                   1024
MC2           States-TPPI
SF            400.2200050 MHz
WDW                 QSINE
SSB                     2
LB                   0.00 Hz
GB                      0
```

NOESY NMR Spectrum of Compound 4 (Chapter 2)



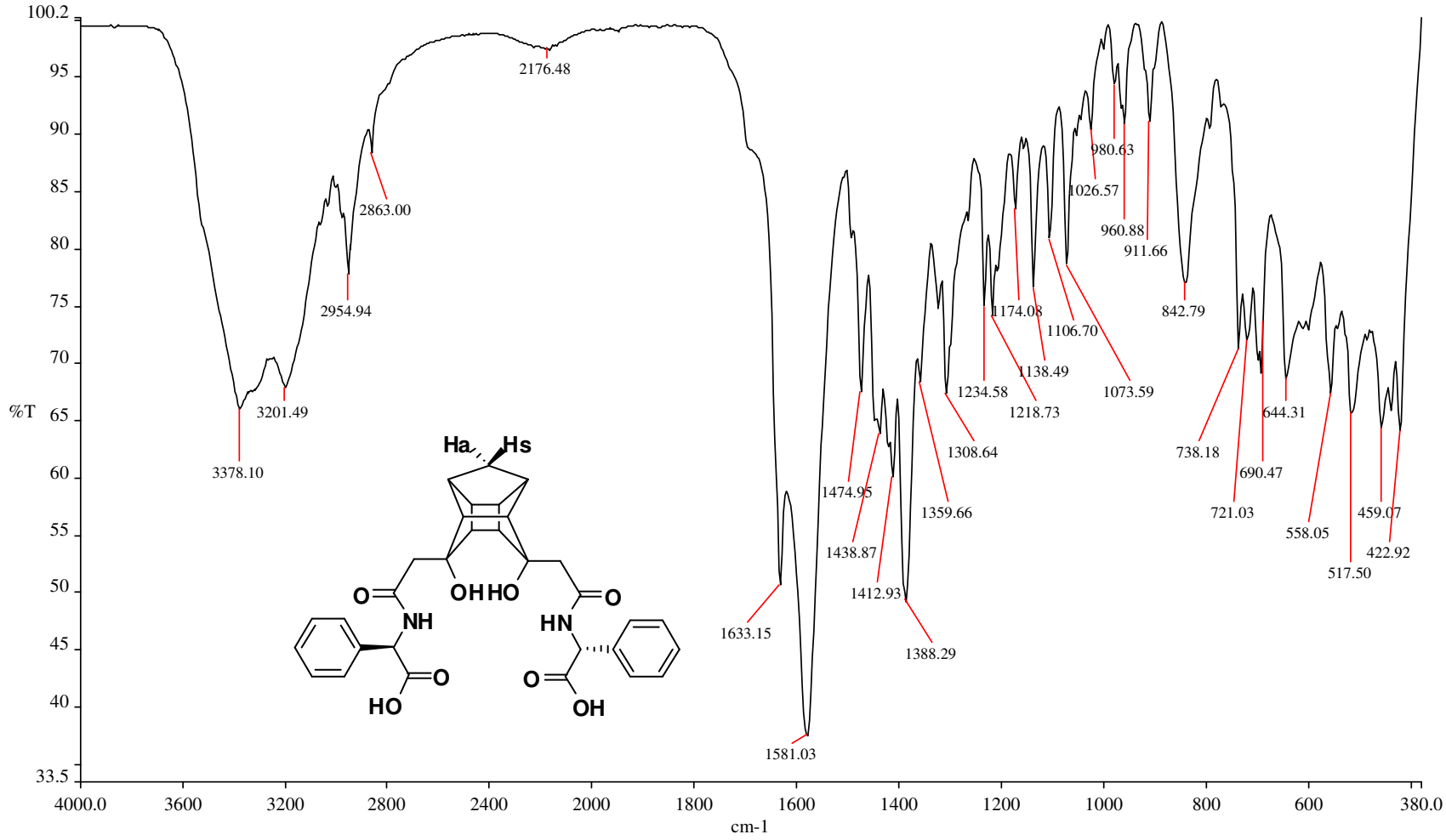

```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                  34
PROCNO                  1

F2 - Acquisition Parameters
Date_            20081106
Time                17.03
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG          hsqcetgp
TD                  1024
SOLVENT              MeOD
NS                      2
DS                     16
SWH              3125.000 Hz
FIDRES           3.051758 Hz
AQ               0.1638900 sec
RG                   2050
DW                160.000 usec
DE                  6.50 usec
TE                 296.7 K
CNST2         145.0000000
D0             0.00000300 sec
D1             1.43200600 sec
D4             0.00172414 sec
D11            0.03000000 sec
D13            0.00000400 sec
D16            0.00020000 sec
IN0            0.00003000 sec
ZGOPTNS

======== CHANNEL f1 ========
NUC1                   1H
P1                  10.00 usec
P2                  20.00 usec
P28               1000.00 usec
PL1                 -3.00 dB
PL1W          15.48668575 W
SFO1         400.2216991 MHz

======== CHANNEL f2 ========
CPDPRG2              garp
NUC2                  13C
P3                   9.00 usec
P4                  18.00 usec
PCPD2               70.00 usec
PL2                 -2.00 dB
PL12                15.60 dB
PL2W          54.14257431 W
PL12W          0.94089001 W
SFO2         100.6429049 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPNAM2           SINE.100
GPZ1                80.00 %
GPZ2                20.10 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                     2
TD                    256
SFO1            100.6429 MHz
FIDRES         65.118813 Hz
```

HSQC NMR Spectrum of Compound 4 (Chapter 2)




```
Current Data Parameters
NAME      Nov06-2008-GK-Nizam
EXPNO                  33
PROCNO                  1

F2 - Acquisition Parameters
Date_           20081106
Time                15.55
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG       hmbcgpndqf
TD                  4096
SOLVENT              MeOD
NS                     16
DS                     16
SWH            3125.000 Hz
FIDRES          0.762939 Hz
AQ             0.6554100 sec
RG                  2050
DW               160.000 usec
DE                 6.50 usec
TE                296.6 K
CNST13         8.0000000
D0             0.00000300 sec
D1             1.22884500 sec
D6             0.06250000 sec
D16            0.00020000 sec
IN0            0.00002235 sec

======== CHANNEL f1 ========
NUC1                   1H
P1                 10.00 usec
P2                 20.00 usec
PL1                -3.00 dB
PL1W         15.48668575 W
SFO1         400.2216991 MHz

======== CHANNEL f2 ========
NUC2                  13C
P3                  9.00 usec
PL2                -2.00 dB
PL2W         54.14257431 W
SFO2         100.6454466 MHz

====== GRADIENT CHANNEL =====
GPNAM1          SINE.100
GPNAM2          SINE.100
GPNAM3          SINE.100
GPZ1               50.00 %
GPZ2               30.00 %
GPZ3               40.10 %
P16              1000.00 usec

F1 - Acquisition parameters
ND0                     2
TD                    128
SFO1           100.6454 MHz
FIDRES        174.631683 Hz
SW              222.095 ppm
FnMODE                 QF

F2 - Processing parameters
SI                   1024
SF           400.2200050 MHz
WDW                  SINE
```

HMBC NMR Spectrum of Compound 4 (Chapter 2)



IR Spectrum of Compound 7 (Chapter 1) and Compound 4 (Chapter 2)

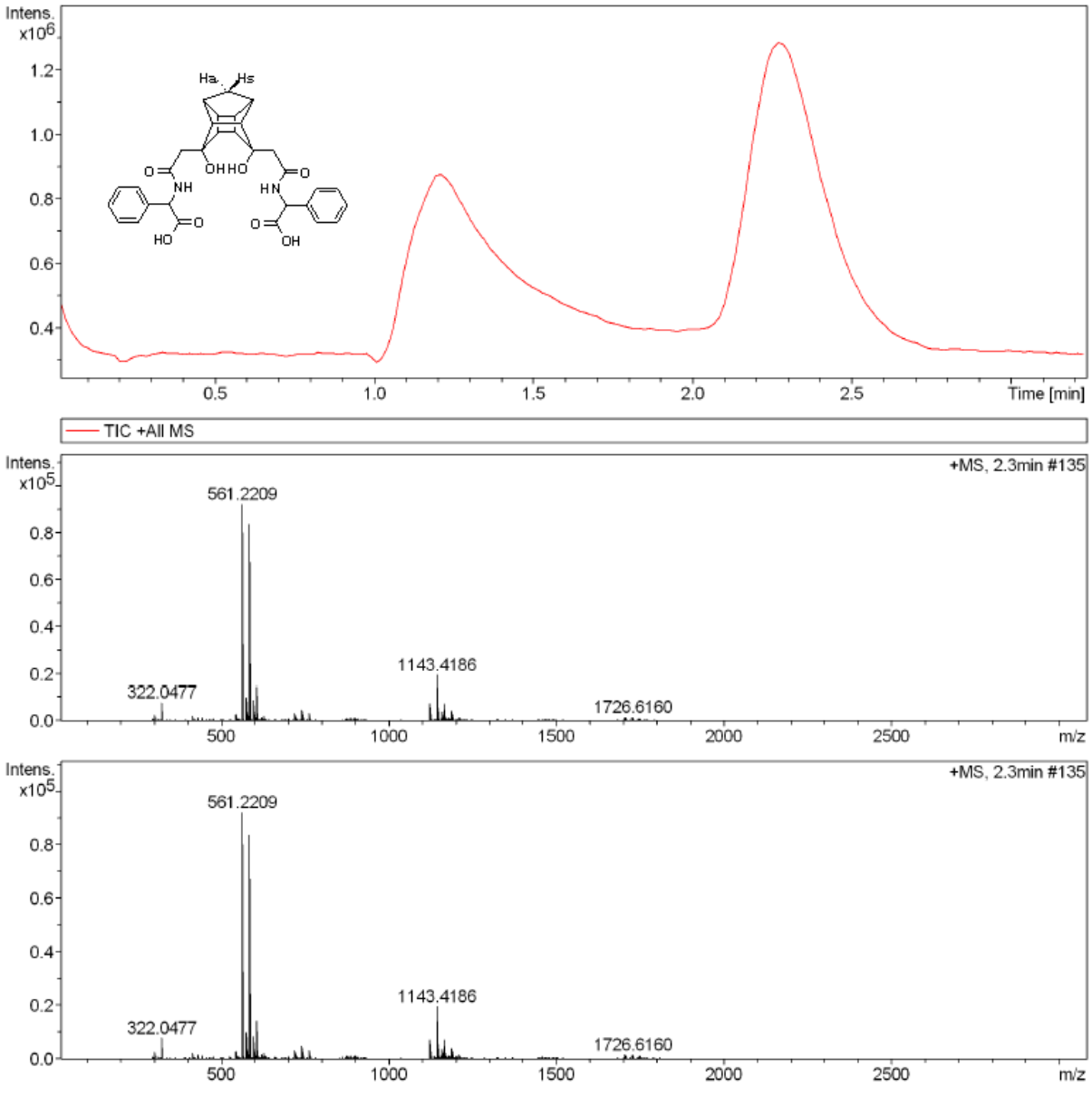
Display Report

Analysis Info

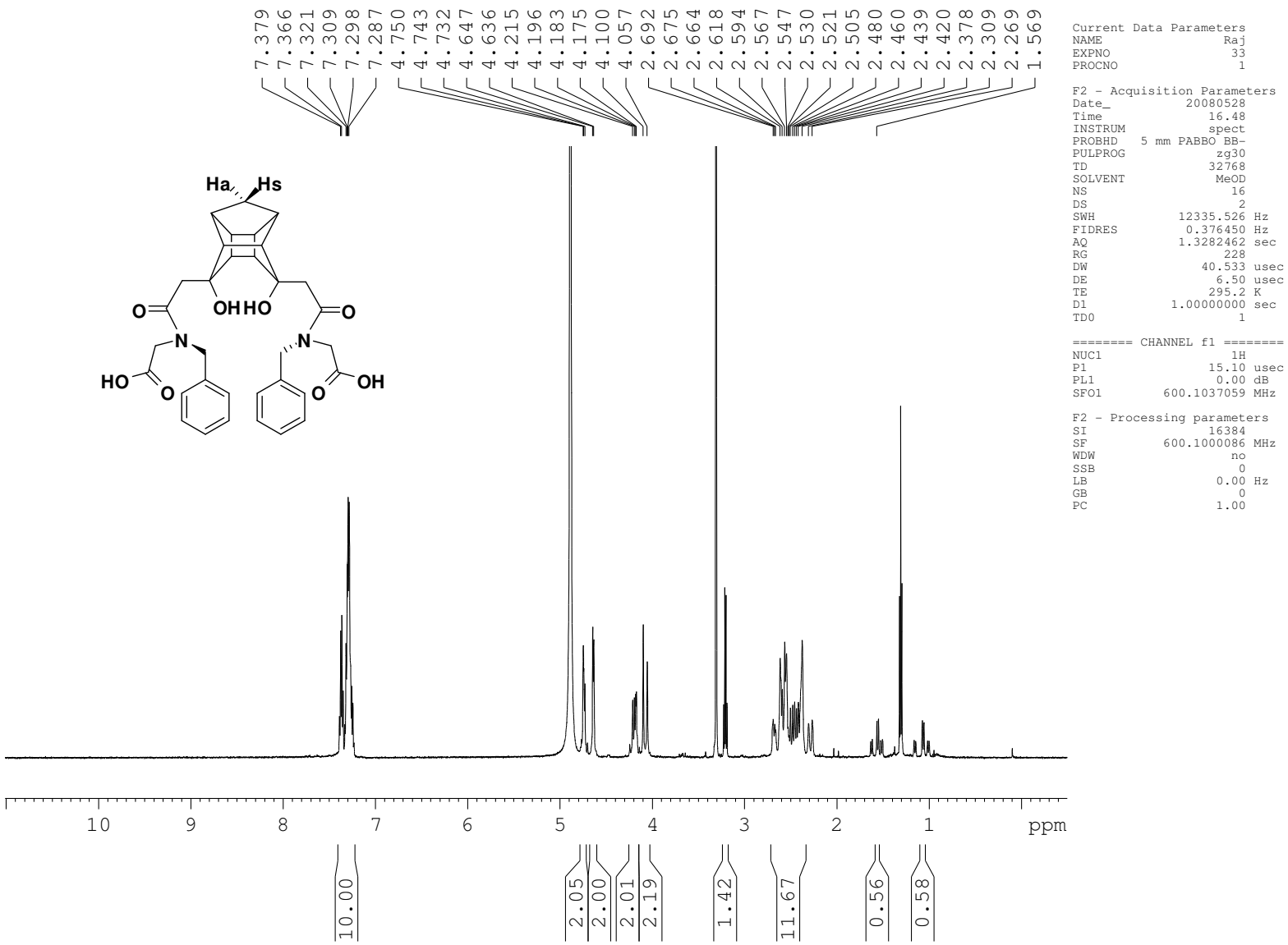
Analysis Name	D:\Data\Raj\ACA-2-Phe-Gly-OH1.d	Acquisition Date	1/1/2010 6:03:32 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	ACA-2-Phe-Gly-OH1	Instrument	micrOTOF-Q 10139
Comment	560.22		

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

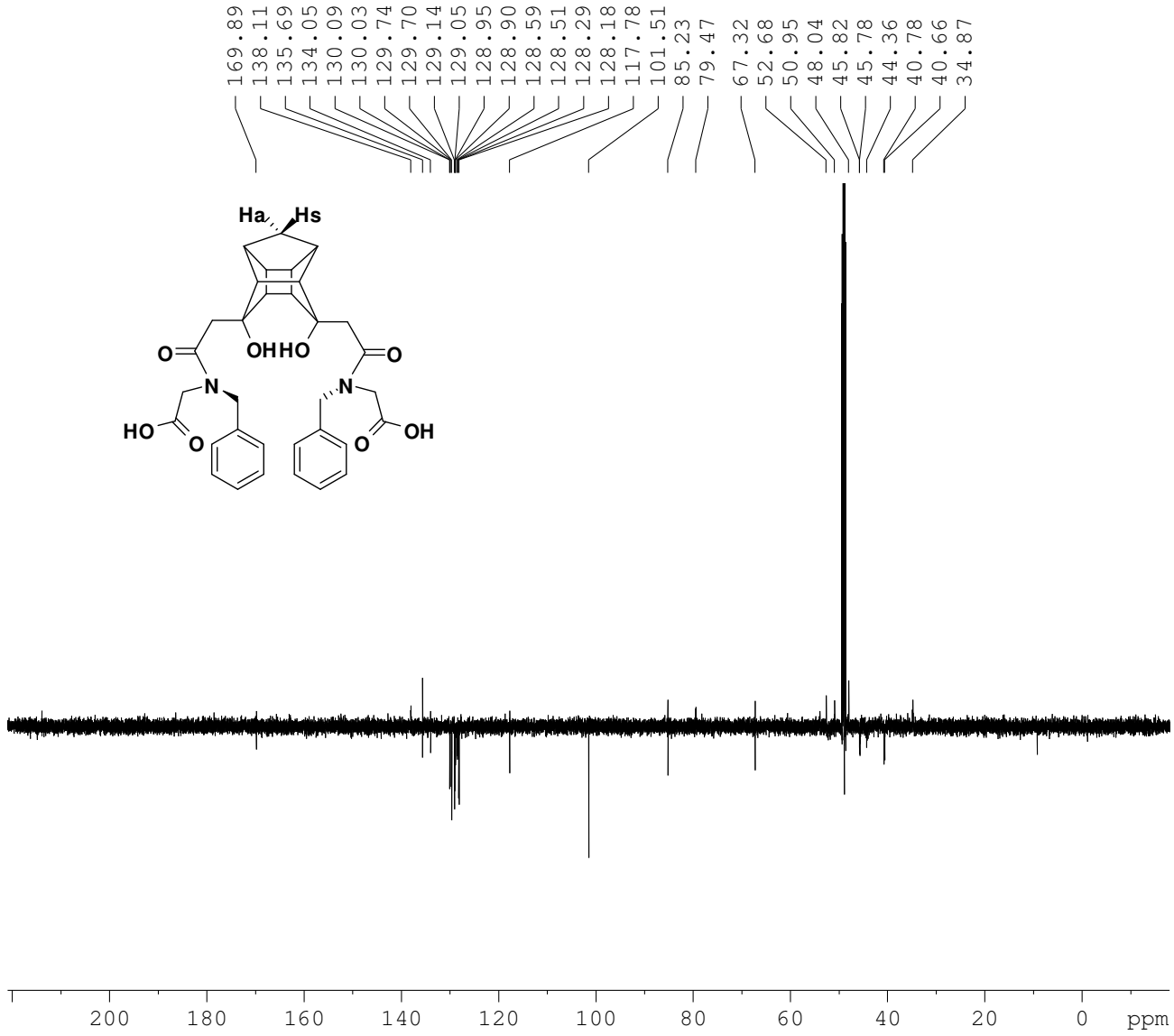


HRESIMS (*m/z*) of Compound 7 (Chapter 1) and Compound 4 (Chapter 2)



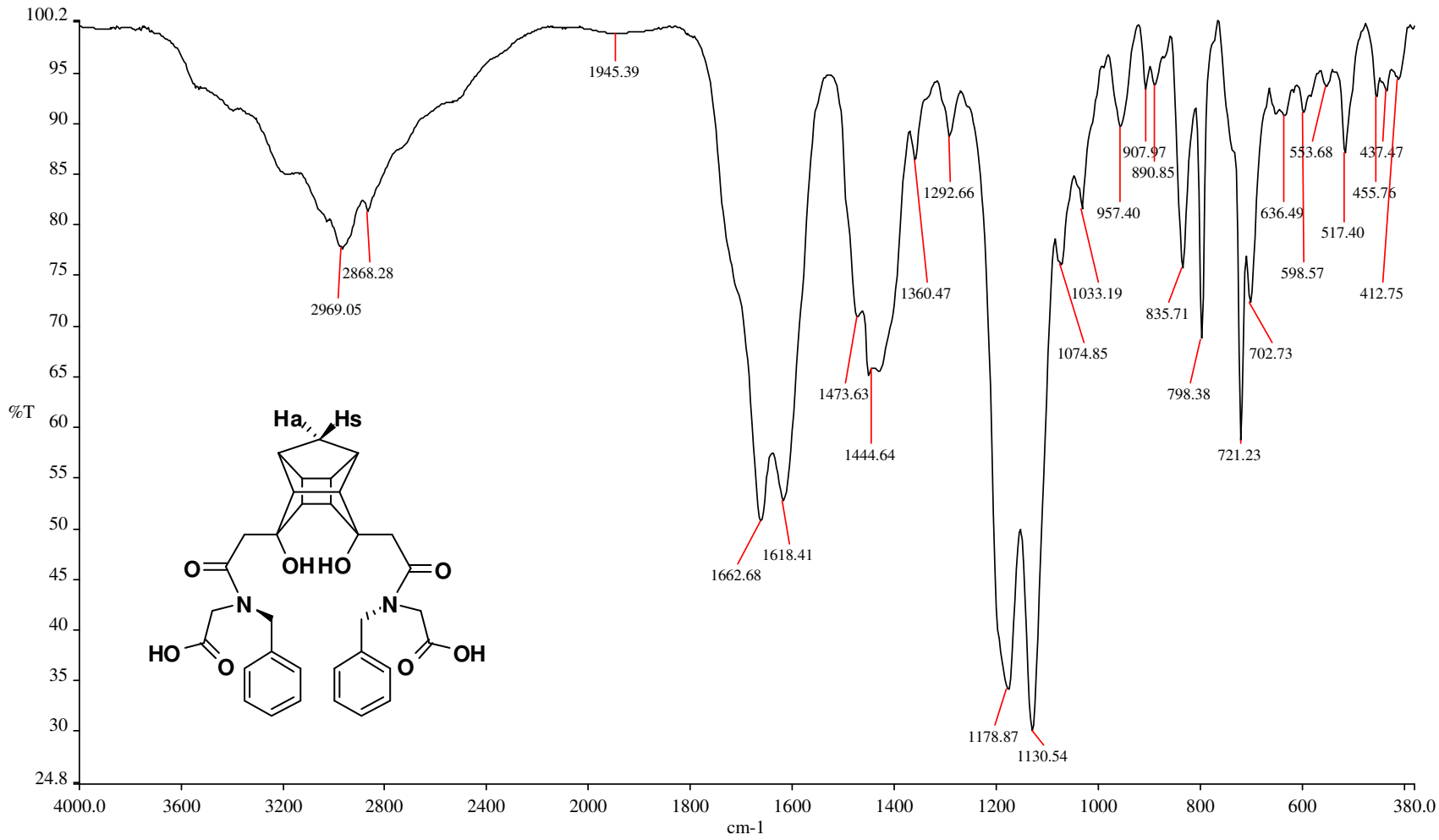

```
Current Data Parameters
NAME                  Raj
EXPNO                  33
PROCNO                  1

F2 - Acquisition Parameters
Date_            20080528
Time                16.48
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG             zg30
TD                 32768
SOLVENT             MeOD
NS                    16
DS                     2
SWH            12335.526 Hz
FIDRES          0.376450 Hz
AQ              1.3282462 sec
RG                   228
DW                40.533 usec
DE                 6.50 usec
TE                295.2 K
D1            1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                  1H
P1                 15.10 usec
PL1                 0.00 dB
SFO1         600.1037059 MHz

F2 - Processing parameters
SI                 16384
SF           600.1000086 MHz
WDW                   no
SSB                    0
LB                  0.00 Hz
GB                     0
PC                  1.00
```

^{1}H NMR Spectrum of Compound 8 (Chapter 1)




```
Current Data Parameters
NAME                  Raj
EXPNO                  36
PROCNO                  1

F2 - Acquisition Parameters
Date_           20080528
Time               22.56
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG             jmod
TD                 65536
SOLVENT             MeOD
NS                  5000
DS                     4
SWH            36057.691 Hz
FIDRES         0.550197 Hz
AQ             0.9088159 sec
RG                 16400
DW               13.867 usec
DE                 6.50 usec
TE                297.8 K
CNST11         1.0000000
CNST2        145.0000000
D1             2.00000000 sec
d20            0.00689655 sec
DELTA          0.00001146 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                  13C
P1                  9.00 usec
p2                 18.00 usec
PL1                1.50 dB
SFO1         150.9103545 MHz

======== CHANNEL f2 ========
CPDPRG2          waltz16
NUC2                  1H
PCPD2              96.00 usec
PL12               15.40 dB
PL2                -0.50 dB
SFO2         600.1024004 MHz

F2 - Processing parameters
SI                 32768
SF           150.8950372 MHz
WDW                   no
SSB                    0
LB                 0.00 Hz
GB                     0
PC                 1.40
```

^{13}C NMR Spectrum of Compound 8 (Chapter 1)



IR Spectrum of Compound 8 (Chapter 1)

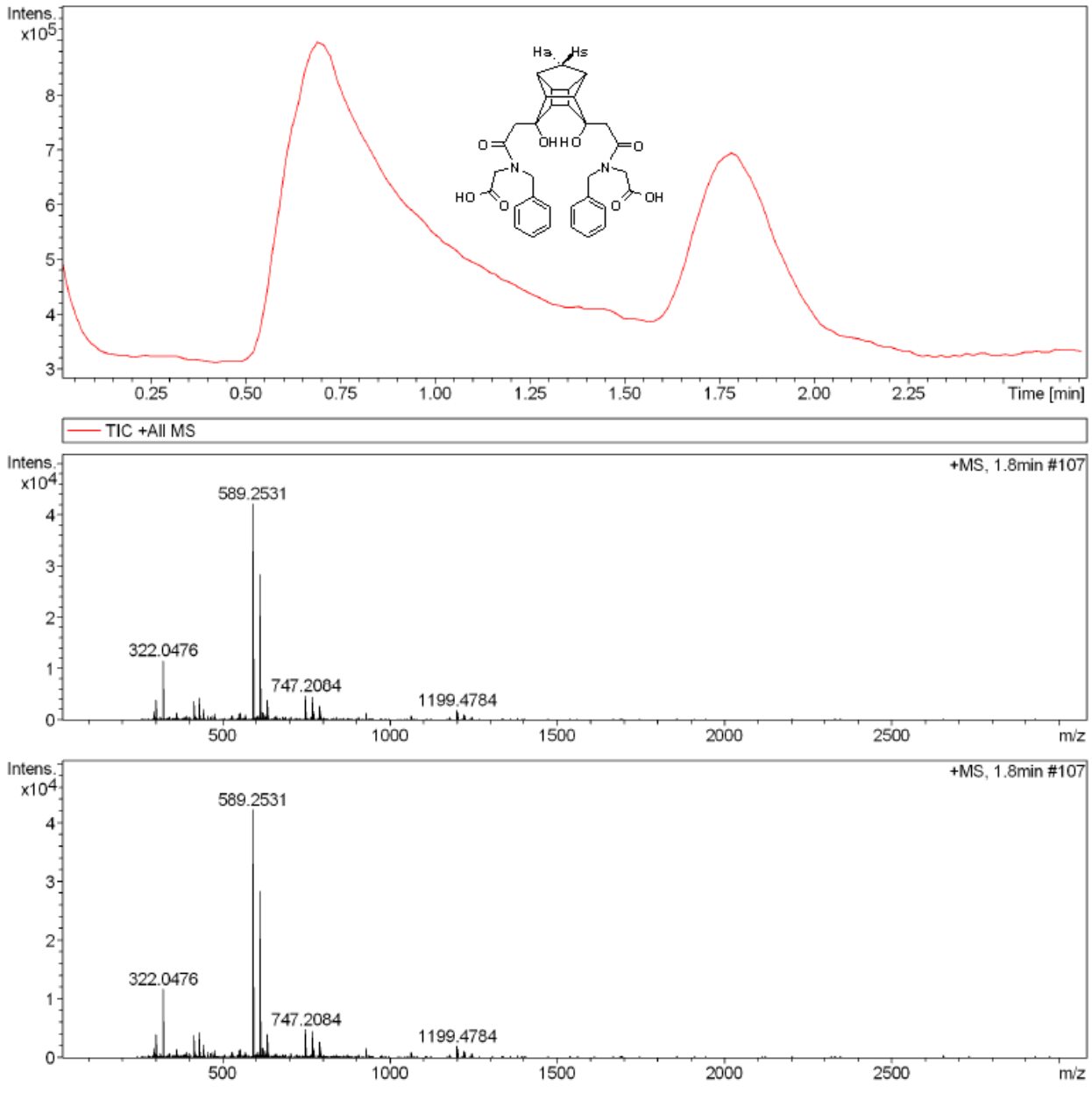
Display Report

Analysis Info

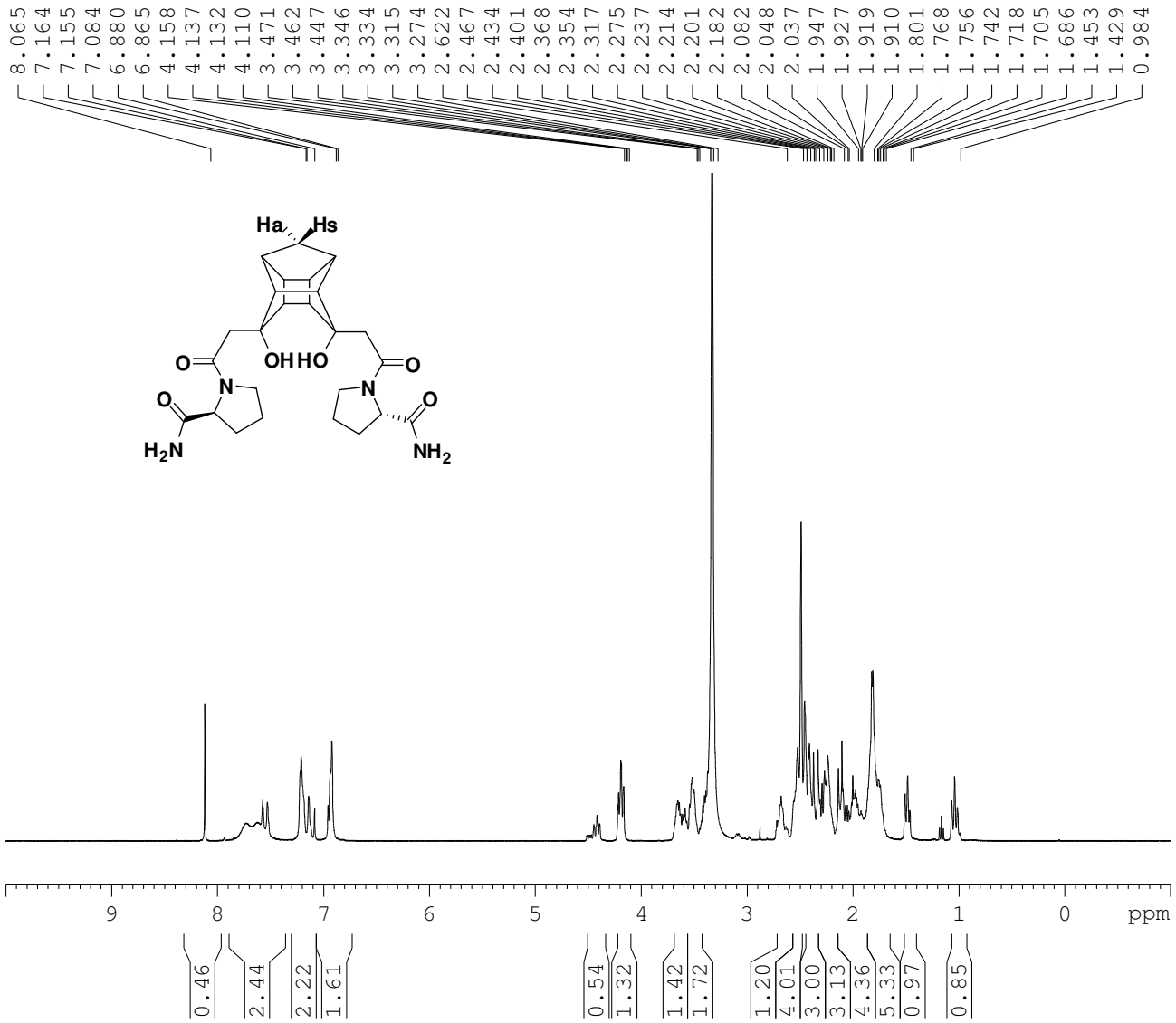
Analysis Name	D:\Data\Raj\ACA-N-Bz-Gly-OH1.d	Acquisition Date	1/1/2010 6:14:58 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	ACA-N-Bz-Gly-OH1	Instrument	micrOTOF-Q 10139
Comment	588.25		

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

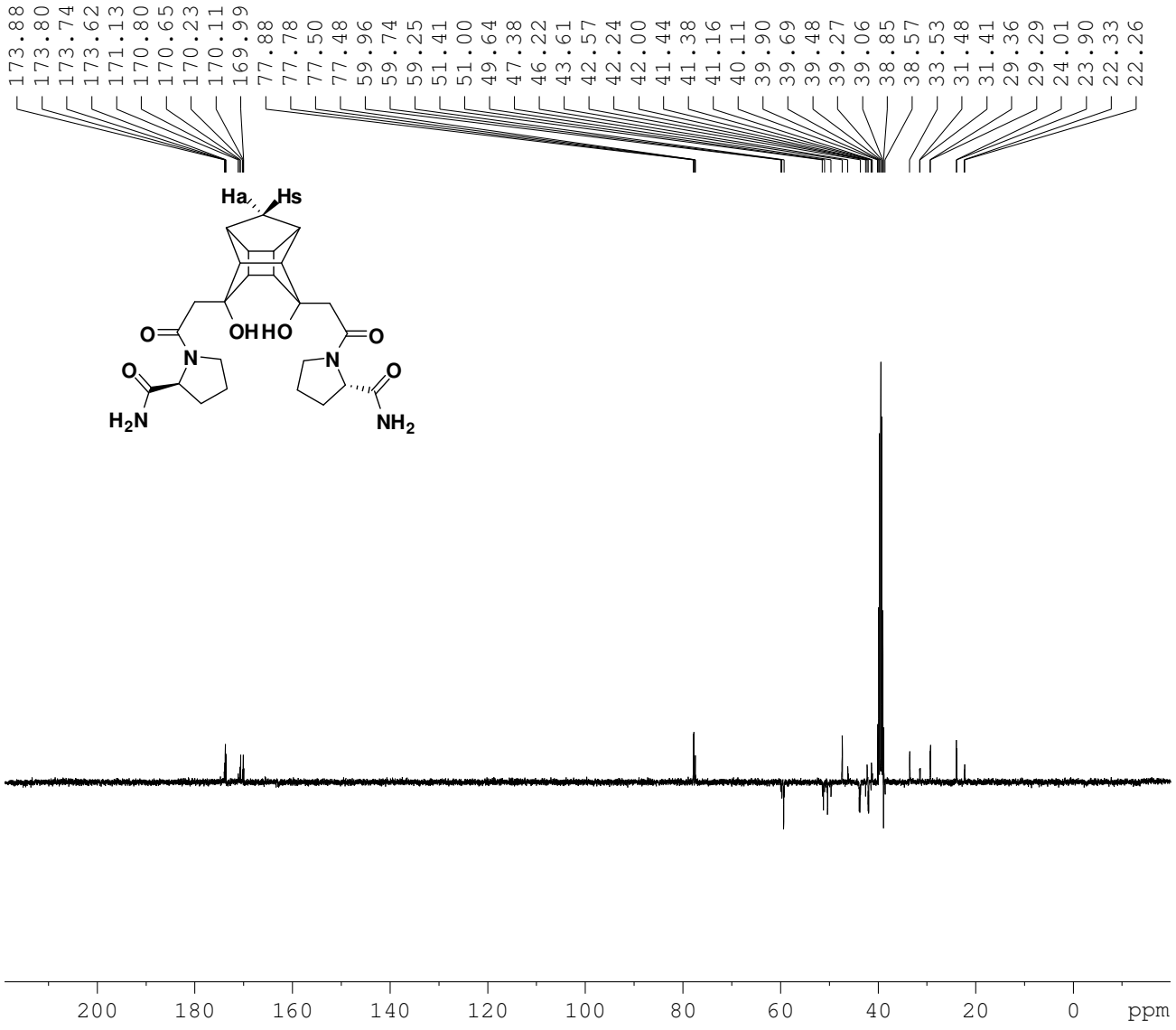


HRESIMS (*m/z*) of Compound 8 (Chapter 1)



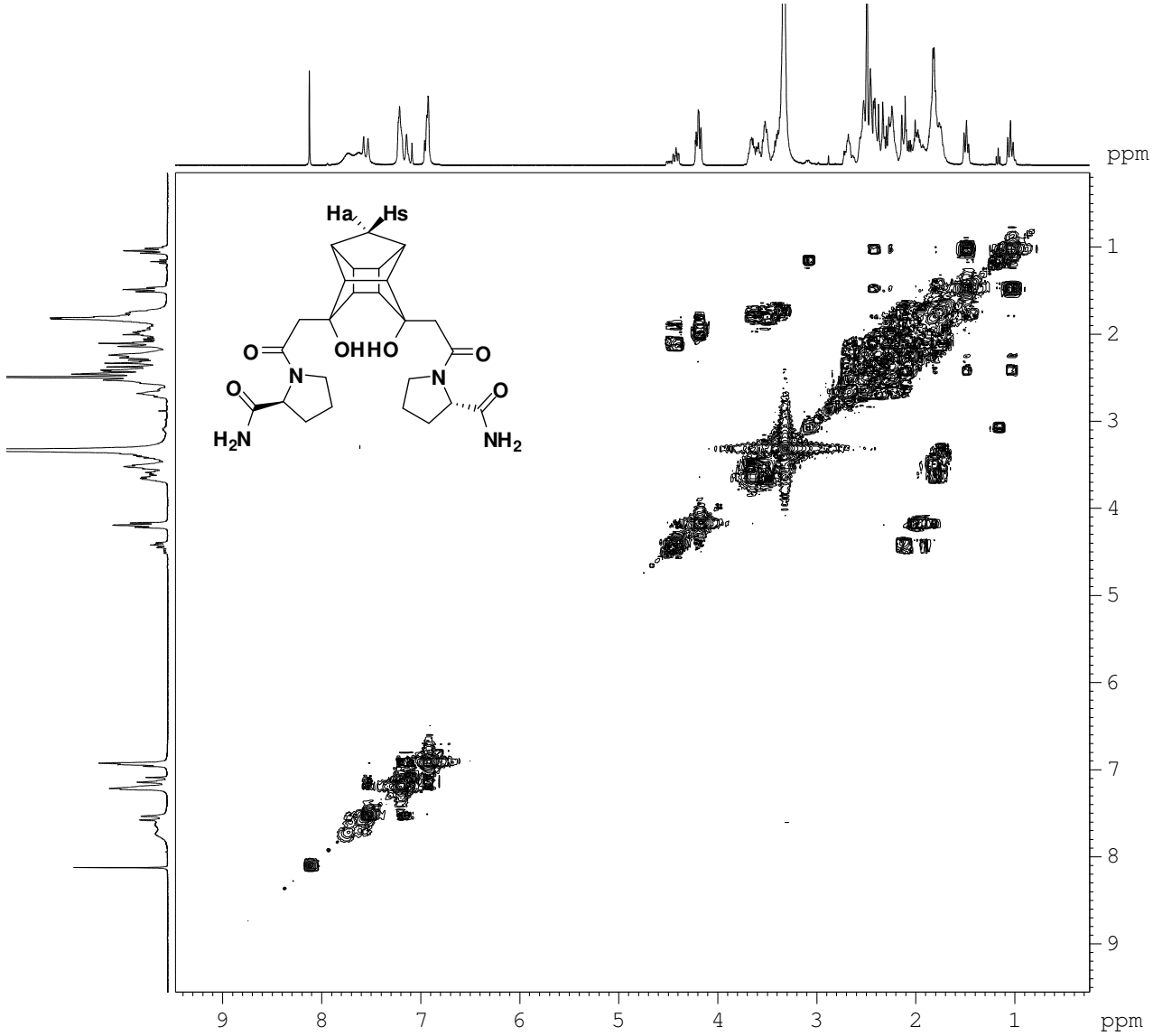

```
Current Data Parameters
NAME      Apr06-2011-GK-Raj
EXPNO                   10
PROCNO                   1

F2 - Acquisition Parameters
Date_             20110407
Time                  7.48
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG               zg30
TD                   32768
SOLVENT               DMSO
NS                      16
DS                       2
SWH               8223.685 Hz
FIDRES            0.250967 Hz
AQ               1.9923444 sec
RG                     128
DW                  60.800 usec
DE                   6.50 usec
TE                  298.2 K
D1              1.00000000 sec
TD0                      1

======== CHANNEL f1 ========
NUC1                    1H
P1                   10.00 usec
PL1                  -3.00 dB
PL1W          15.48668575 W
SFO1          400.2224715 MHz

F2 - Processing parameters
SI                   16384
SF            400.2200034 MHz
WDW                     EM
SSB                      0
LB                    0.30 Hz
GB                       0
PC                    1.00
```

^{1}H NMR Spectrum of Compound 9 (Chapter 1) and Compound 5 (Chapter 2)



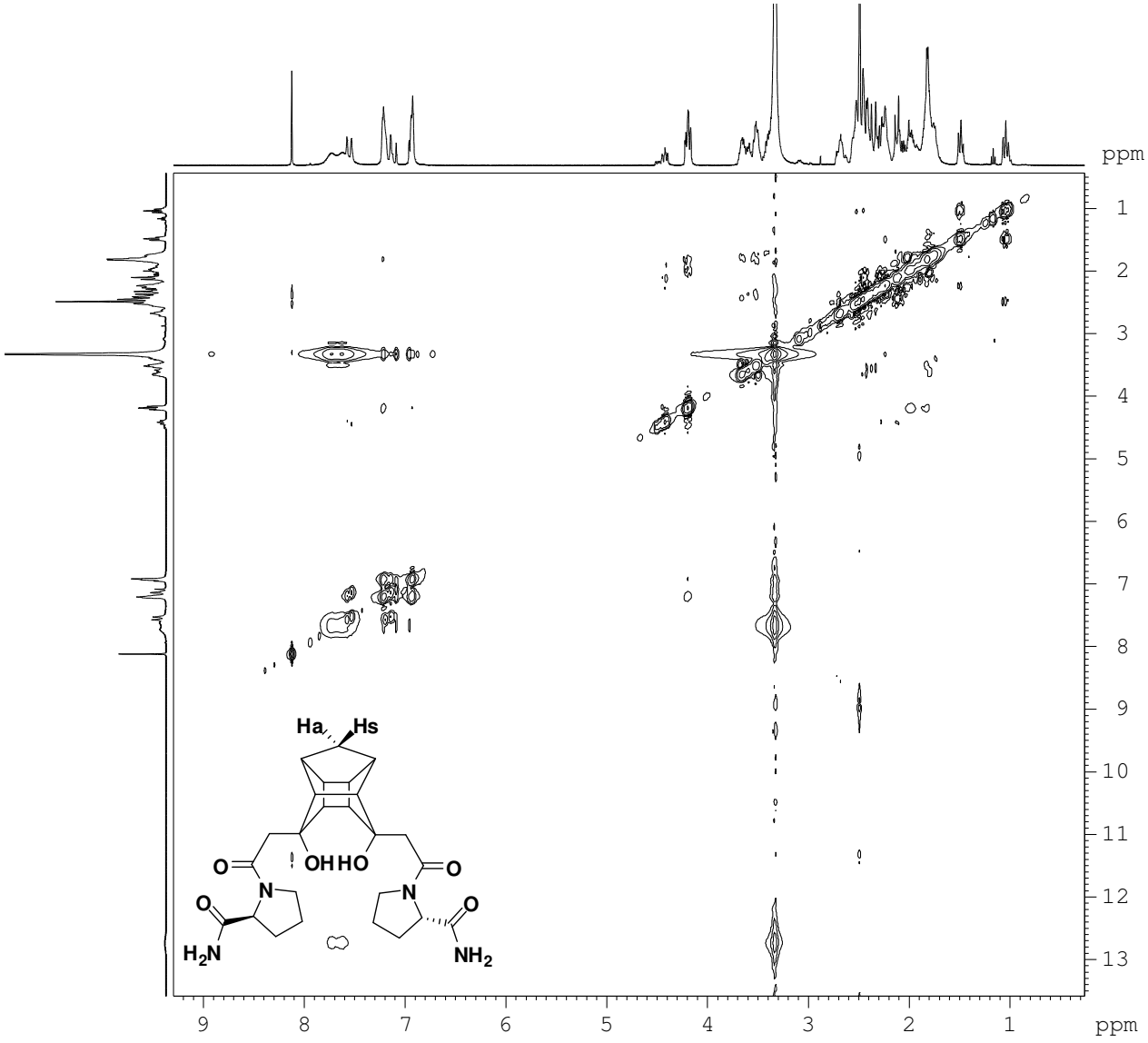

```
Current Data Parameters
NAME      Apr06-2011-GK-Raj
EXPNO                   11
PROCNO                   1

F2 - Acquisition Parameters
Date_             20110407
Time                 8.36
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG              jmod
TD                  65536
SOLVENT              DMSO
NS                   1024
DS                      4
SWH           24038.461 Hz
FIDRES         0.366798 Hz
AQ            1.3631988 sec
RG                   2050
DW                20.800 usec
DE                 6.50 usec
TE                298.2 K
CNST2        145.0000000
CNST11         1.0000000
D1            2.00000000 sec
D20           0.00689655 sec
TD0                     1

======== CHANNEL f1 ========
NUC1                  13C
P1                   8.40 usec
P2                  16.80 usec
PL1                 -2.00 dB
PL1W         54.14257431 W
SFO1         100.6454626 MHz

======== CHANNEL f2 ========
CPDPRG2           waltz16
NUC2                  1H
PCPD2               90.00 usec
PL2                 -3.00 dB
PL12                15.60 dB
PL2W         15.48668575 W
PL12W         0.21377575 W
SFO2         400.2216009 MHz

F2 - Processing parameters
SI                  32768
SF           100.6354474 MHz
WDW                    EM
SSB                     0
LB                   1.00 Hz
GB                      0
PC                   1.40
```

^{13}C NMR Spectrum of Compound 9 (Chapter 1) and Compound 5 (Chapter 2)



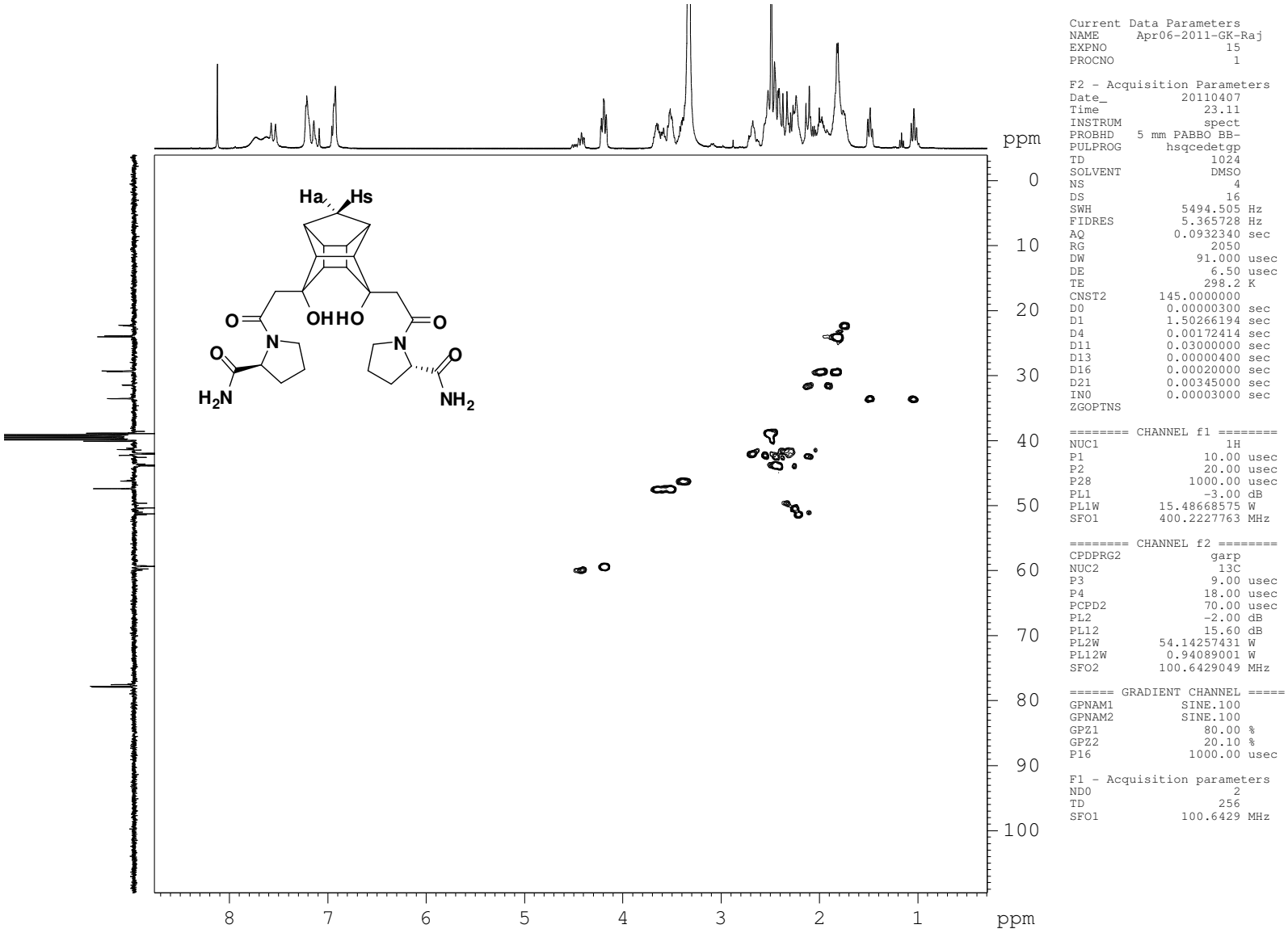

```
Current Data Parameters
NAME      Apr06-2011-GK-Raj
EXPNO                   13
PROCNO                   1

F2 - Acquisition Parameters
Date_             20110407
Time                 22.17
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG          cosygpqf
TD                   2048
SOLVENT               DMSO
NS                       2
DS                       8
SWH              5494.505 Hz
FIDRES           2.682864 Hz
AQ               0.1864180 sec
RG                      64
DW                 91.000 usec
DE                  6.50 usec
TE                 298.2 K
D0              0.00000300 sec
D1              1.49221694 sec
D13             0.00000400 sec
D16             0.00020000 sec
IN0             0.00018200 sec

======== CHANNEL f1 ========
NUC1                    1H
P0                  10.00 usec
P1                  10.00 usec
PL1                 -3.00 dB
PL1W         15.48668575 W
SFO1         400.2227763 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPZ1                10.00 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                      1
TD                     128
SFO1            400.2228 MHz
FIDRES          42.925823 Hz
SW                13.729 ppm
FnMODE                  QF

F2 - Processing parameters
SI                    1024
SF           400.2200050 MHz
WDW                   SINE
SSB                      0
LB                  0.00 Hz
GB                       0
PC                    1.40

F1 - Processing parameters
SI                    1024
MC2                     QF
SF           400.2200080 MHz
WDW                   SINE
SSB                      0
LB                  0.00 Hz
```

COSY NMR Spectrum of Compound 9 (Chapter 1) and Compound 5 (Chapter 2)



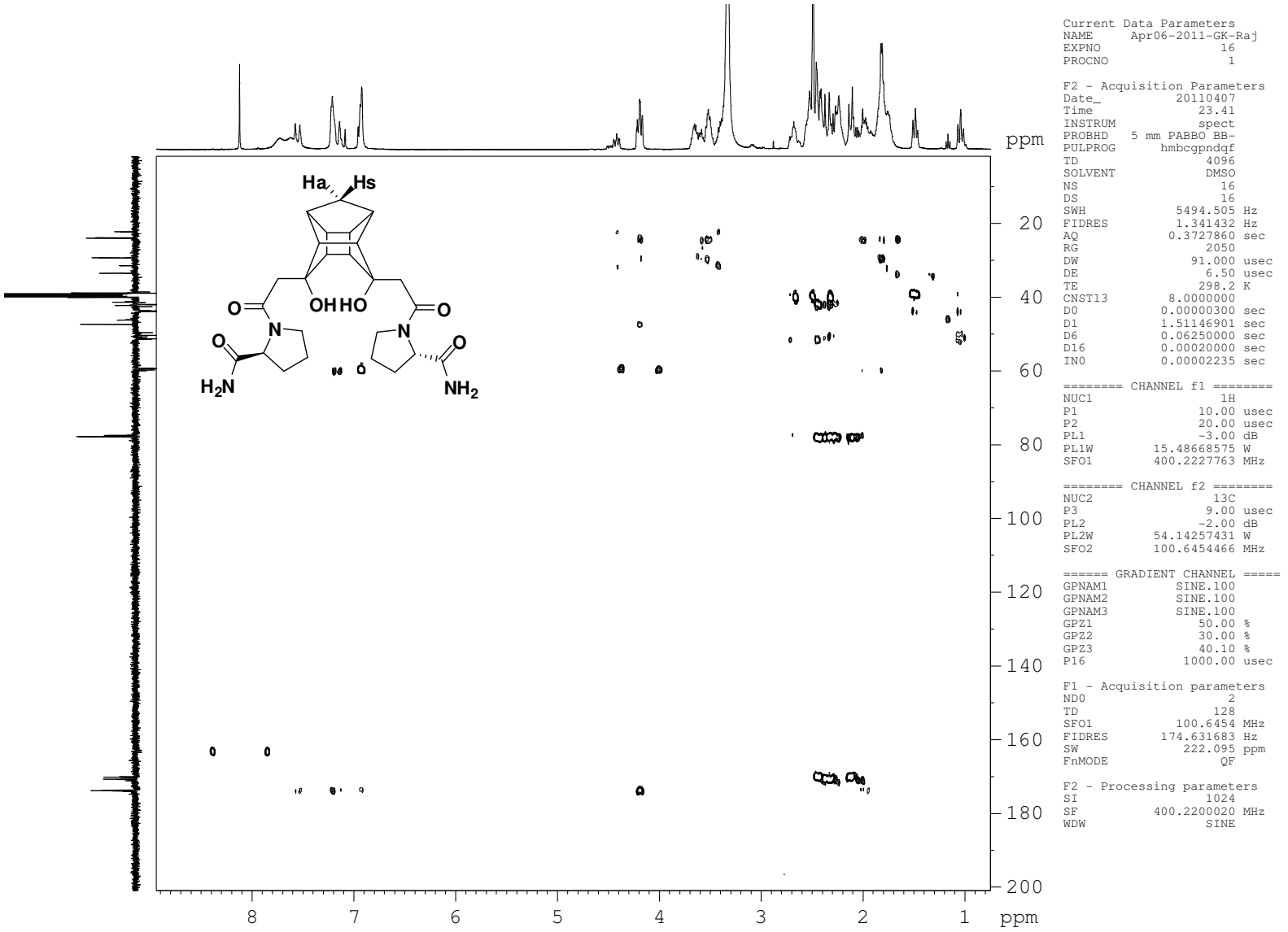

```
Current Data Parameters
NAME      Apr06-2011-GK-Raj
EXPNO                   14
PROCNO                   1

F2 - Acquisition Parameters
Date_            20110407
Time                22.26
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG           noesyph
TD                   2048
SOLVENT              DMSO
NS                      4
DS                      4
SWH              5494.505 Hz
FIDRES           2.682864 Hz
AQ               0.1864180 sec
RG                     64
DW                 91.000 usec
DE                  6.50 usec
TE                 298.2 K
D0            0.00007827 sec
D1            2.06430697 sec
D8            0.30000001 sec
IN0           0.00018200 sec

======== CHANNEL f1 ========
NUC1                   1H
P1                  10.00 usec
PL1                 -3.00 dB
PL1W         15.48668575 W
SFO1         400.2227763 MHz

F1 - Acquisition parameters
ND0                     1
TD                    256
SFO1            400.2228 MHz
FIDRES          21.462912 Hz
SW                13.729 ppm
FnMODE       States-TPPI

F2 - Processing parameters
SI                   1024
SF           400.2200020 MHz
WDW                 QSINE
SSB                     2
LB                   0.00 Hz
GB                      0
PC                   1.00

F1 - Processing parameters
SI                   1024
MC2          States-TPPI
SF           400.2200030 MHz
WDW                 QSINE
SSB                     2
LB                   0.00 Hz
GB                      0
```

NOESY NMR Spectrum of Compound 9 (Chapter 1) and Compound 5 (Chapter 2)




```
Current Data Parameters
NAME      Apr06-2011-GK-Raj
EXPNO                   15
PROCNO                   1

F2 - Acquisition Parameters
Date_           20110407
Time              23.11
INSTRUM           spect
PROBHD   5 mm PABBO BB-
PULPROG       hsqcedetgp
TD                 1024
SOLVENT             DMSO
NS                     4
DS                    16
SWH             5494.505 Hz
FIDRES          5.365728 Hz
AQ             0.0932340 sec
RG                  2050
DW                91.000 usec
DE                 6.50 usec
TE                298.2 K
CNST2        145.0000000
D0            0.00000300 sec
D1            1.50266194 sec
D4            0.00172414 sec
D11           0.03000000 sec
D13           0.00000400 sec
D16           0.00020000 sec
D21           0.00345000 sec
IN0           0.00003000 sec
ZGOPTNS

======== CHANNEL f1 ========
NUC1                 1H
P1                10.00 usec
P2                20.00 usec
P28             1000.00 usec
PL1               -3.00 dB
PL1W         15.48668575 W
SFO1        400.2227763 MHz

======== CHANNEL f2 ========
CPDPRG2             garp
NUC2                13C
P3                 9.00 usec
P4                18.00 usec
PCPD2             70.00 usec
PL2               -2.00 dB
PL12              15.60 dB
PL2W         54.14257431 W
PL12W         0.94089001 W
SFO2        100.6429049 MHz

====== GRADIENT CHANNEL =====
GPNAM1          SINE.100
GPNAM2          SINE.100
GPZ1               80.00 %
GPZ2               20.10 %
P16              1000.00 usec

F1 - Acquisition parameters
ND0                    2
TD                   256
SFO1           100.6429 MHz
```

HSQC NMR Spectrum of Compound 9 (Chapter 1) and Compound 5 (Chapter 2)



Current Data Parameters
NAME Apr06-2011-GK-Raj
EXPNO 16
PROCNO 1

F2 - Acquisition Parameters
Date_ 20110407
Time 23.41
INSTRUM spect
PROBHD 5 mm PABBO BB-
PULPROG hmbcgpndqf
TD 4096
SOLVENT DMSO
NS 16
DS 16
SWH 5494.505 Hz
FIDRES 1.341432 Hz
AQ 0.3727860 sec
RG 2050
DW 91.000 usec
DE 6.50 usec
TE 298.2 K
CNST13 8.0000000
D0 0.00000300 sec
D1 1.51146901 sec
D6 0.06250000 sec
D16 0.00020000 sec
IN0 0.00002235 sec

======== CHANNEL f1 ========
NUC1 1H
P1 10.00 usec
P2 20.00 usec
PL1 -3.00 dB
PL1W 15.48668575 W
SFO1 400.2227763 MHz

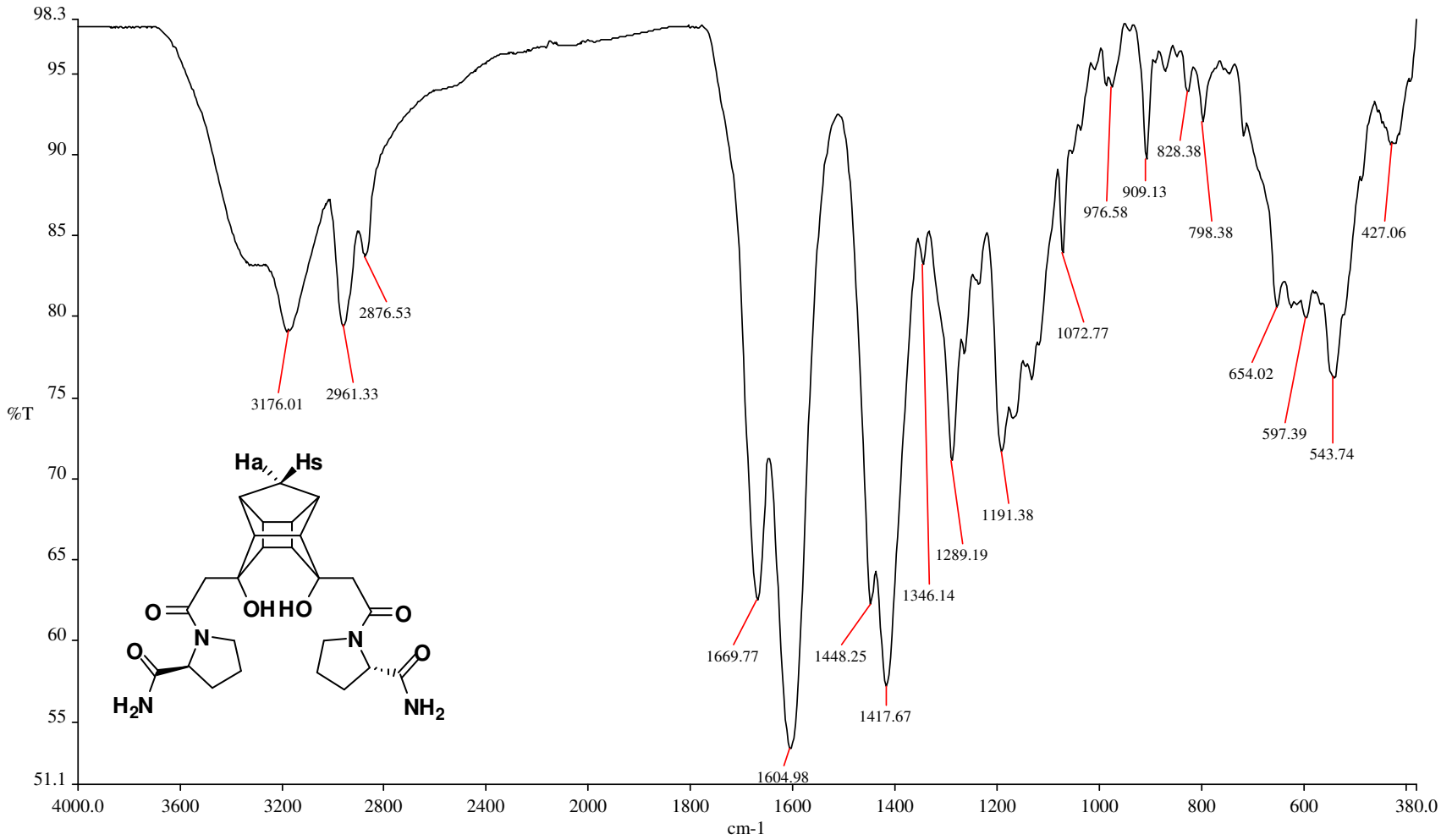
======== CHANNEL f2 ========
NUC2 13C
P3 9.00 usec
PL2 -2.00 dB
PL2W 54.14257431 W
SFO2 100.6454466 MHz

====== GRADIENT CHANNEL =====
GPNAM1 SINE.100
GPNAM2 SINE.100
GPNAM3 SINE.100
GPZ1 50.00 %
GPZ2 30.00 %
GPZ3 40.10 %
P16 1000.00 usec

F1 - Acquisition parameters
ND0 2
TD 128
SFO1 100.6454 MHz
FIDRES 174.631683 Hz
SW 222.095 ppm
FnMODE QF

F2 - Processing parameters
SI 1024
SF 400.2200020 MHz
WDW SINE

HMBC NMR Spectrum of Compound 9 (Chapter 1) and Compound 5 (Chapter 2)



IR Spectrum of Compound 9 (Chapter 1) and Compound 5 (Chapter 2)

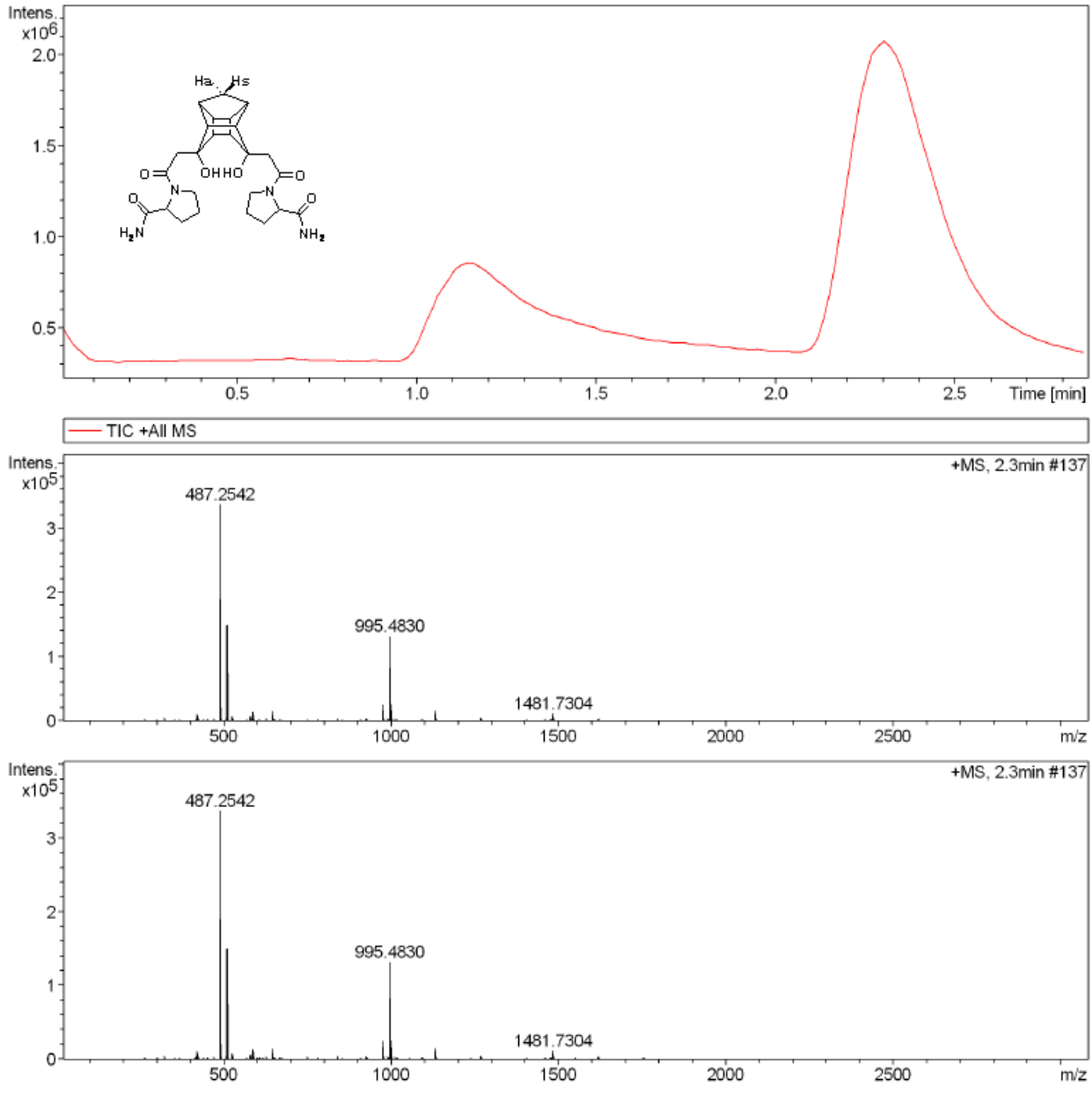
Display Report

Analysis Info

Analysis Name	D:\Data\Raj\ACA-Pro1.d	Acquisition Date	1/1/2010 6:23:03 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	ACA-Pro1	Instrument	micrOTOF-Q 10139
Comment	486.25		

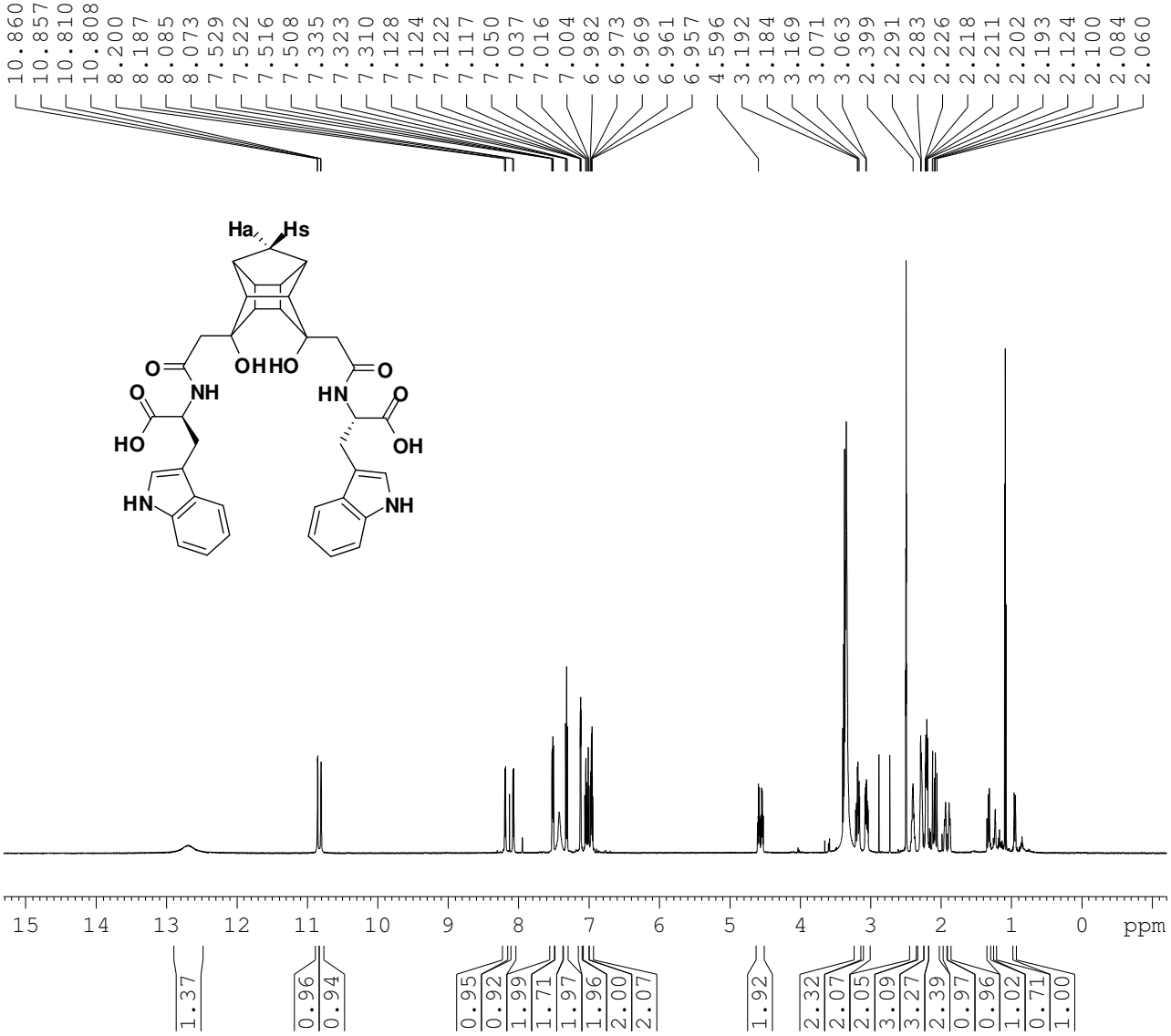
Acquisition Parameter

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



Bruker Compass DataAnalysis 4.0 printed: 1/1/2010 6:27:28 PM Page 1 of 1

HRESIMS (*m/z*) of Compound 9 (Chapter 1) and Compound 5 (Chapter 2)



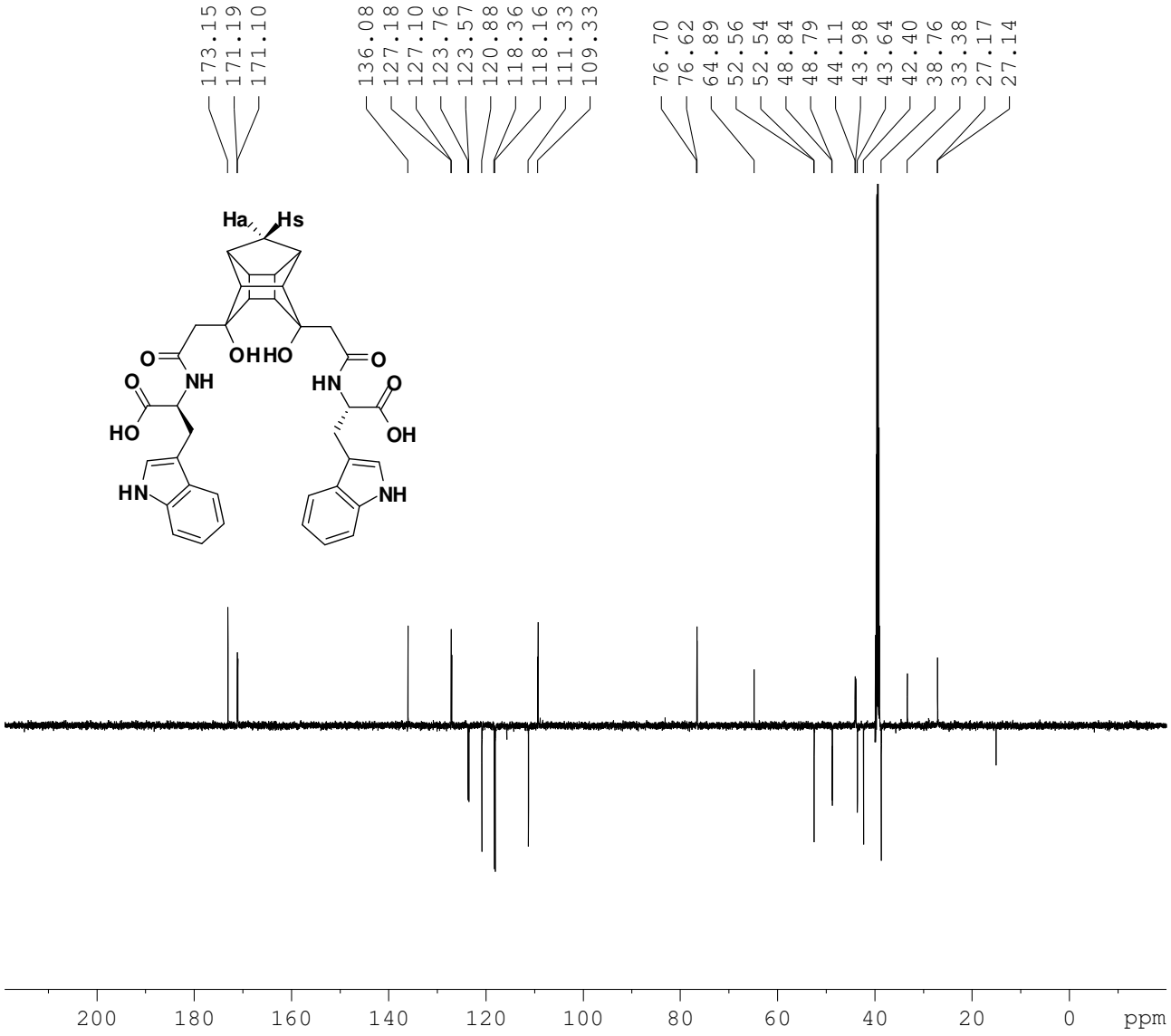
Current Data Parameters
NAME Raj
EXPNO 201
PROCNO 1

F2 - Acquisition Parameters
Date_ 20120212
Time 20.01
INSTRUM spect
PROBHD 5 mm PABBO BB-
PULPROG zg30
TD 32768
SOLVENT DMSO
NS 16
DS 2
SWH 12335.526 Hz
FIDRES 0.376450 Hz
AQ 1.3282462 sec
RG 114
DW 40.533 usec
DE 6.50 usec
TE 298.4 K
D1 1.00000000 sec
TD0 1

======== CHANNEL f1 ========
NUC1 1H
P1 15.10 usec
PL1 0.00 dB
PL1W 16.62868881 W
SFO1 600.1037059 MHz

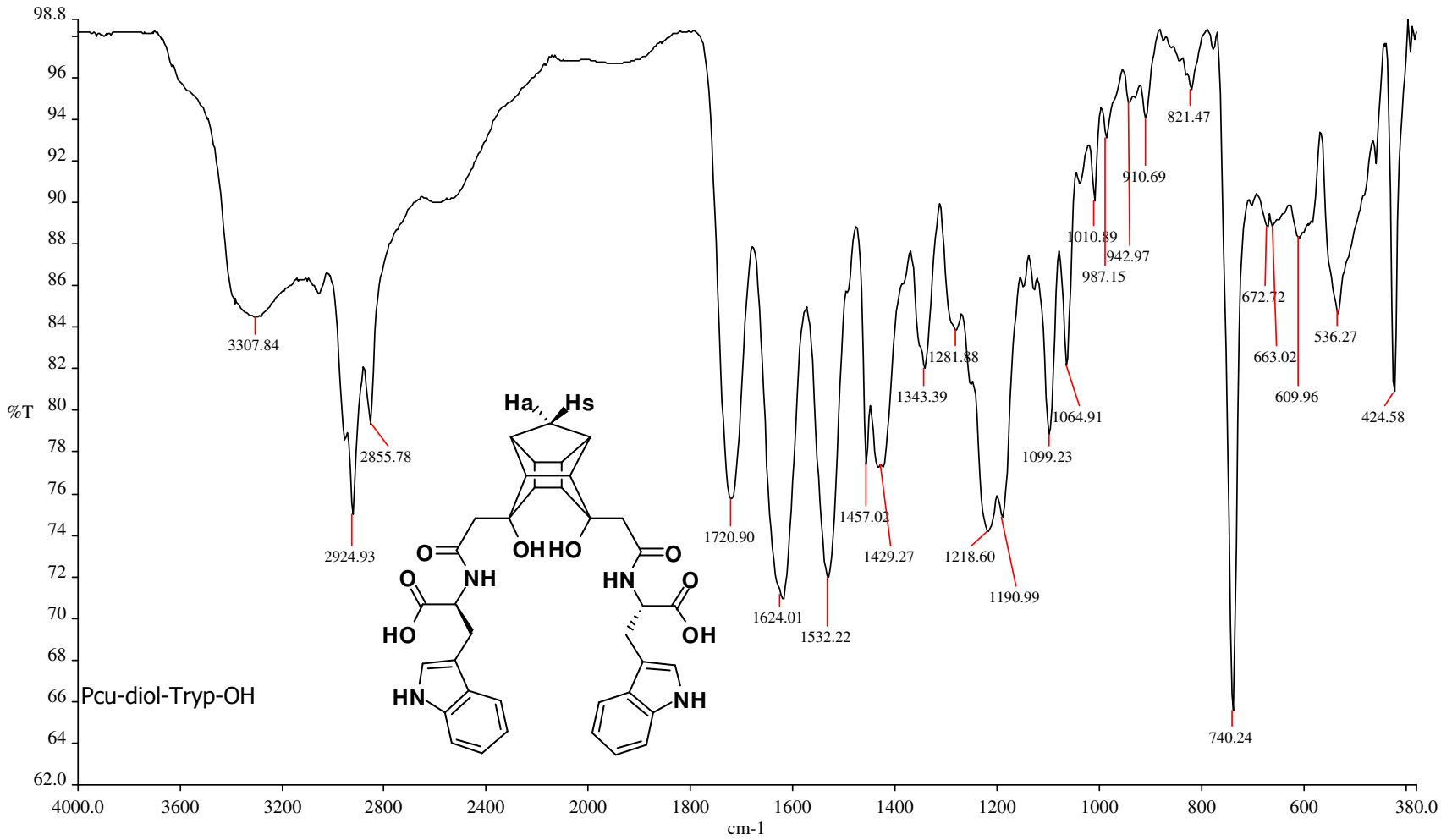
F2 - Processing parameters
SI 16384
SF 600.1000061 MHz
WDW no
SSB 0
LB 0.00 Hz
GB 0
PC 1.00

^{1}H NMR Spectrum of Compound 10 (Chapter 1)




```
Current Data Parameters
NAME                  Raj
EXPNO                 202
PROCNO                  1

F2 - Acquisition Parameters
Date_            20120213
Time                13.46
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG             jmod
TD                 65536
SOLVENT             DMSO
NS                  3072
DS                     4
SWH           36057.691 Hz
FIDRES         0.550197 Hz
AQ             0.9088159 sec
RG                 16400
DW                13.867 usec
DE                 6.50 usec
TE                300.9 K
CNST2        145.0000000
CNST11         1.0000000
D1           2.00000000 sec
D20          0.00689655 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                 13C
P1                  9.00 usec
P2                 18.00 usec
PL1                 1.50 dB
PL1W         59.57236099 W
SFO1         150.9103545 MHz

======== CHANNEL f2 ========
CPDPRG2           waltz16
NUC2                 1H
PCPD2               96.00 usec
PL2                 -0.50 dB
PL12                15.40 dB
PL2W         18.65769577 W
PL12W         0.47957665 W
SFO2         600.1024004 MHz

F2 - Processing parameters
SI                 32768
SF           150.8953395 MHz
WDW                  EM
SSB                   0
LB                 1.00 Hz
GB                    0
PC                 1.40
```

^{13}C NMR Spectrum of Compound 10 (Chapter 1)



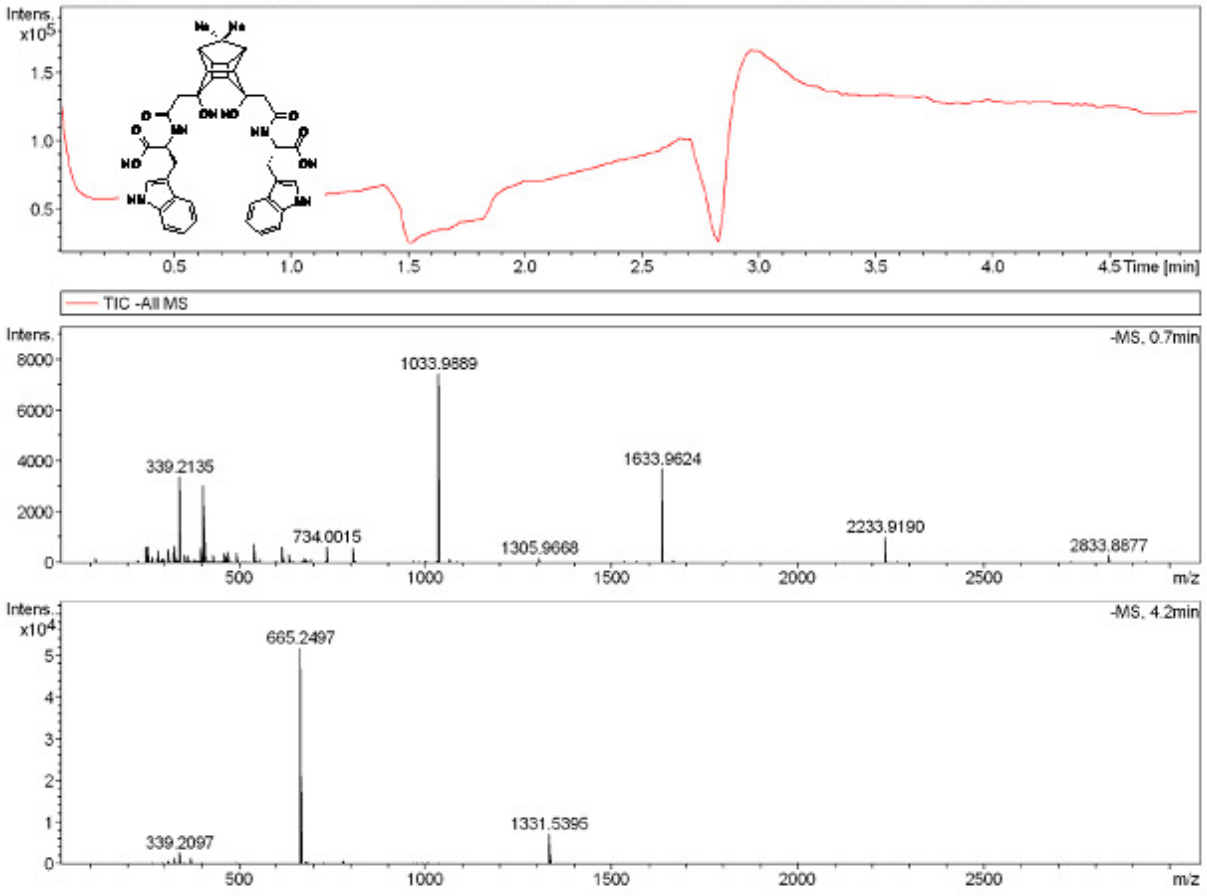
IR Spectrum of Compound 10 (Chapter 1)

Analysis Info

Analysis Name	D:\Data\Raj\4-3.d	Acquisition Date	2/28/2012 8:09:53 PM
Method	tune_wide_expert_lactamEAIS.m	Operator	BDAL@DE
Sample Name	4-3	Instrument / Ser#	micrOTOF-Q 10139
Comment			

Acquisition Parameter

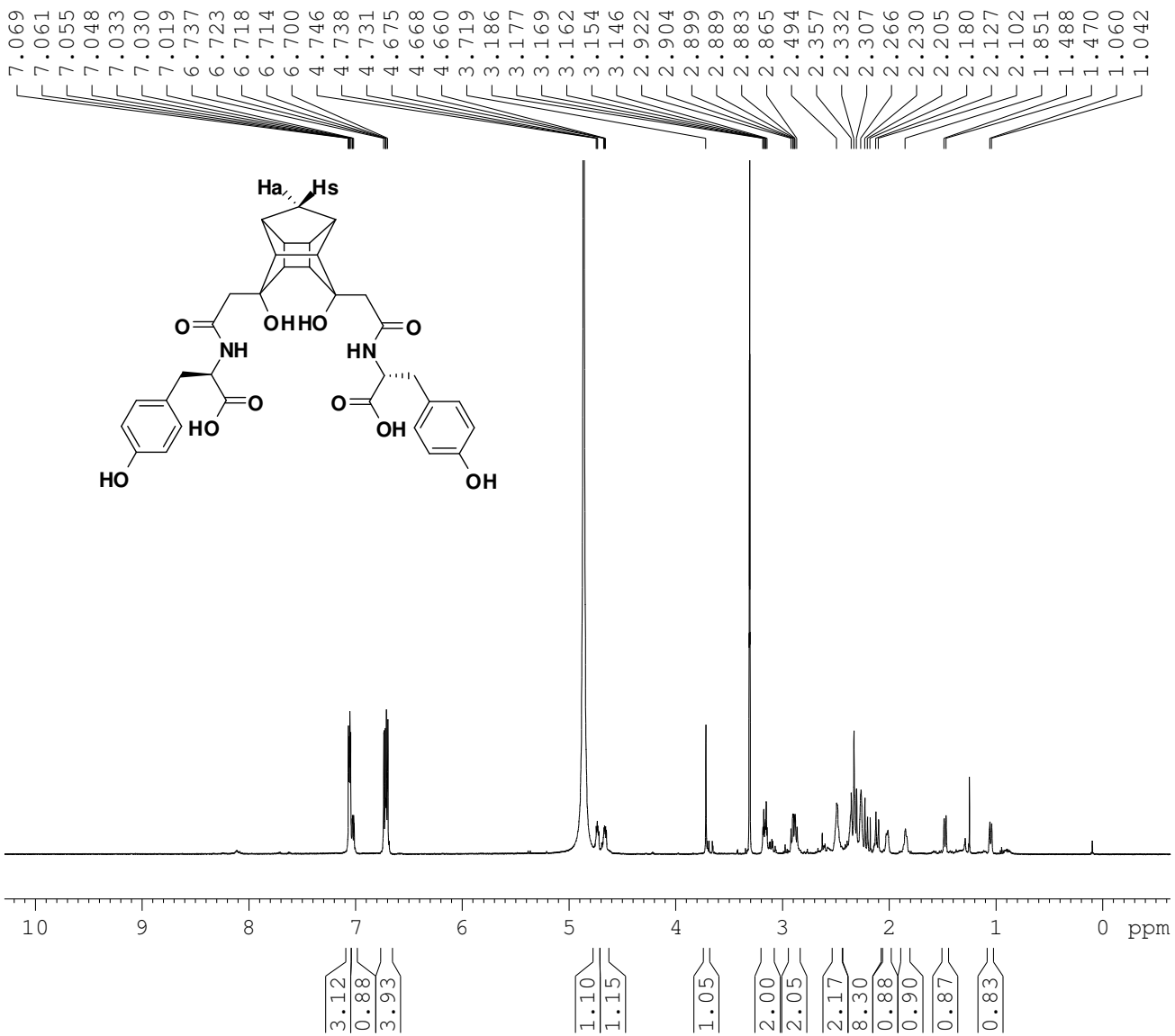
Source Type	ESI	Ion Polarity	Negative	Set Nebulizer	14.5 psi
Focus	Not active	Set Capillary	3500 V	Set Dry Heater	200 °C
Scan Begin	100 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	500.0 Vpp	Set Divert Valve	Source



Meas. m/z	#	Formula	m/z	err [ppm]	Mean err [ppm]	rdb	N-Rule	e⁻ Conf	mSigma	Std I	Std Mean m/z	Std I VarNorm	Std m/z Diff	Std Comb Dev
665.2497														
	1	C 32 H 37 N 6 O 10	665.2577	12.0	12.2	17.5	ok	even	7.10	0.0127	0.0082	0.0051	0.0017	0.8730
	2	C 33 H 37 N 4 O 11	665.2464	-4.9	-4.6	17.5	ok	even	9.53	0.0170	0.0031	0.0066	0.0017	0.7793
	3	C 28 H 37 N 6 O 13	665.2424	-10.9	-10.7	13.5	ok	even	19.12	0.0262	0.0071	0.0089	0.0017	0.8977
	4	C 27 H 37 N 8 O 12	665.2536	6.0	6.2	13.5	ok	even	20.88	0.0295	0.0042	0.0097	0.0017	0.8506

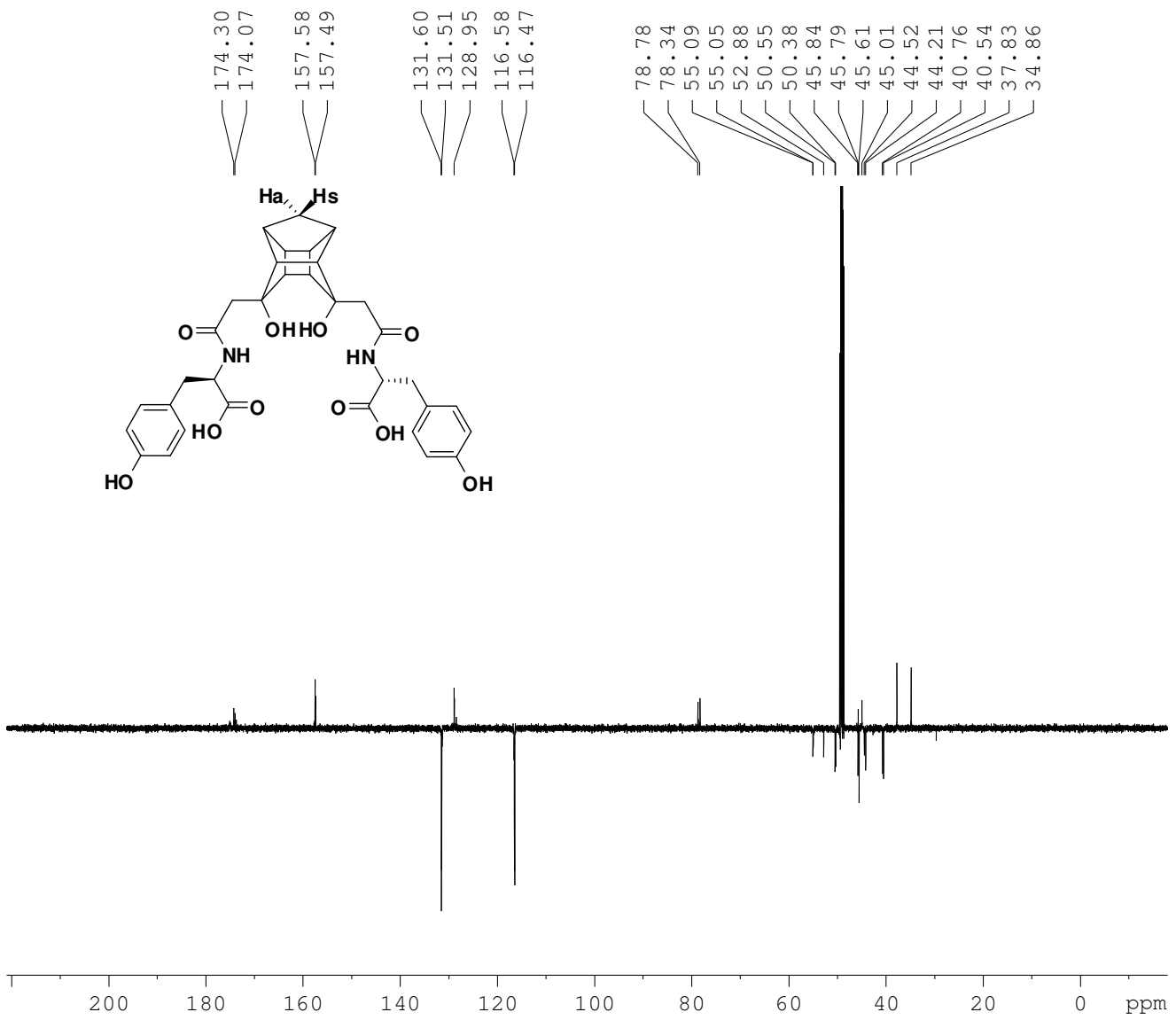
Meas. m/z	#	Formula	m/z	err [ppm]	Mean err [ppm]	rdb	N-Rule	e⁻ Conf	mSigma	Std I	Std Mean m/z	Std I VarNorm	Std m/z Diff	Std Comb Dev
	5	C 37 H 37 N 4 O 8	665.2617	18.1	18.3	21.5	ok	even	27.04	0.0418	0.0122	0.0137	0.0017	0.9632
	6	C 26 H 41 N 4 O 16	665.2523	4.0	4.2	8.5	ok	even	32.70	0.0474	0.0029	0.0156	0.0017	0.8716

HRESIMS (*m/z*) of Compound 10 (Chapter 1)



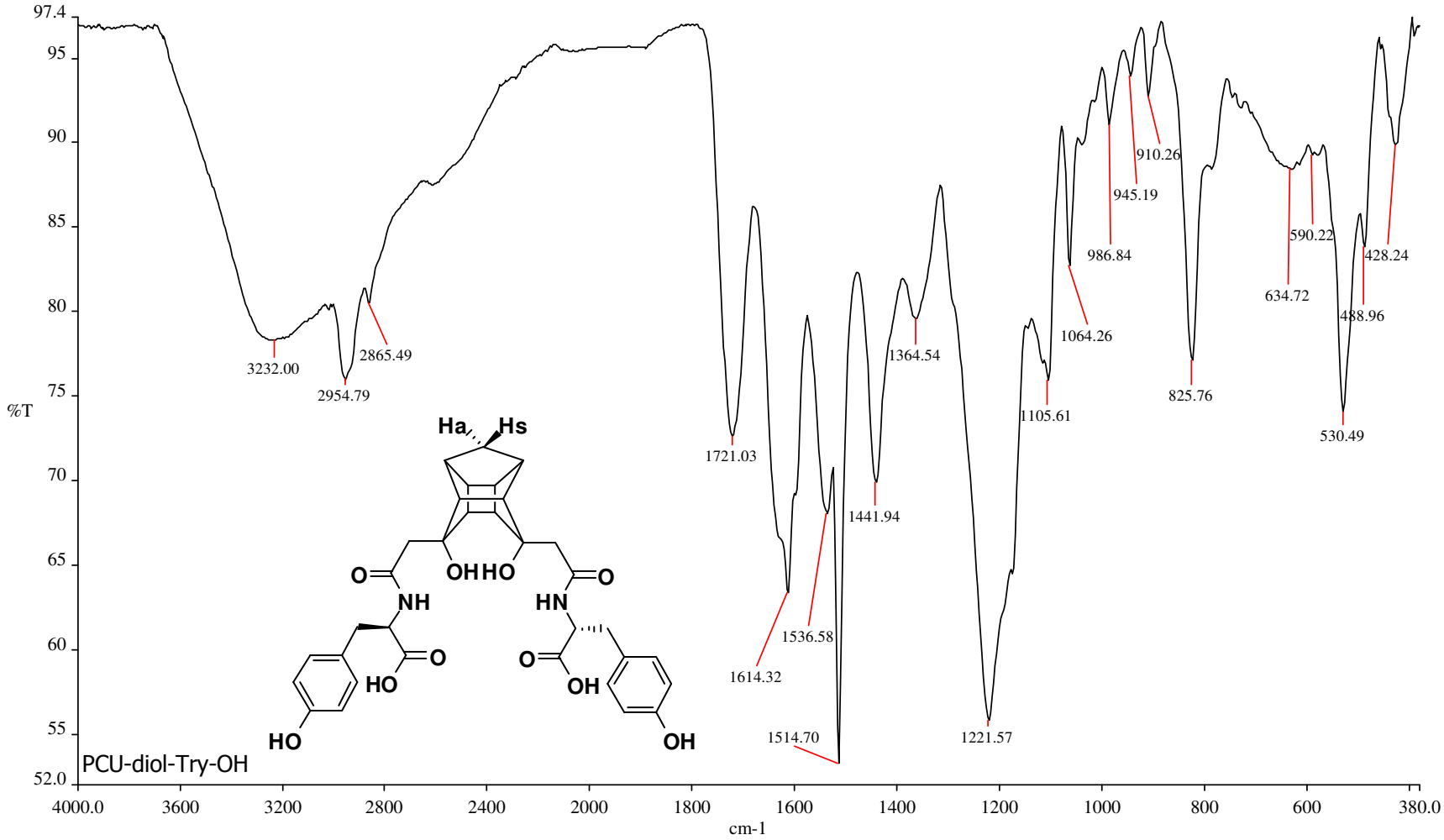

```
Current Data Parameters
NAME                  Raj
EXPNO                 195
PROCNO                  1

F2 - Acquisition Parameters
Date_              20120211
Time                 17.58
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG               zg30
TD                   32768
SOLVENT               MeOD
NS                      16
DS                       2
SWH              12335.526 Hz
FIDRES           0.376450 Hz
AQ               1.3282462 sec
RG                     144
DW                  40.533 usec
DE                   6.50 usec
TE                  298.4 K
D1              1.00000000 sec
TD0                     1

======== CHANNEL f1 ========
NUC1                   1H
P1                  14.19 usec
PL1                  0.00 dB
PL1W          16.62868881 W
SFO1          600.1037059 MHz

F2 - Processing parameters
SI                   16384
SF            600.1000086 MHz
WDW                    no
SSB                     0
LB                   0.00 Hz
GB                      0
PC                   1.00
```

^{1}H NMR Spectrum of Compound 11 (Chapter 1)




```
Current Data Parameters
NAME                  Raj
EXPNO                 200
PROCNO                  1

F2 - Acquisition Parameters
Date_            20120212
Time                6.18
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG             jmod
TD                 65536
SOLVENT             MeOD
NS                  6000
DS                     4
SWH           36057.691 Hz
FIDRES         0.550197 Hz
AQ             0.9088159 sec
RG                  2050
DW                13.867 usec
DE                 6.50 usec
TE                301.1 K
CNST2        145.0000000
CNST11         1.0000000
D1           2.00000000 sec
D20          0.00689655 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                 13C
P1                  9.00 usec
P2                 18.00 usec
PL1                 1.50 dB
PL1W         59.57236099 W
SFO1        150.9103545 MHz

======== CHANNEL f2 ========
CPDPRG2          waltz16
NUC2                 1H
PCPD2              96.00 usec
PL2                -0.50 dB
PL12               15.40 dB
PL2W         18.65769577 W
PL12W         0.47957665 W
SFO2        600.1024004 MHz

F2 - Processing parameters
SI                 32768
SF          150.8950302 MHz
WDW                   no
SSB                    0
LB                  0.00 Hz
GB                     0
PC                  1.40
```

^{13}C NMR Spectrum of Compound 11 (Chapter 1)



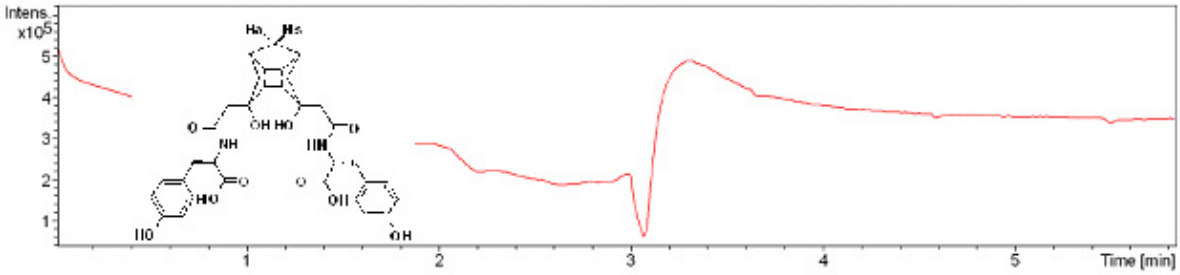
IR Spectrum of Compound 11 (Chapter 1)

Analysis Info

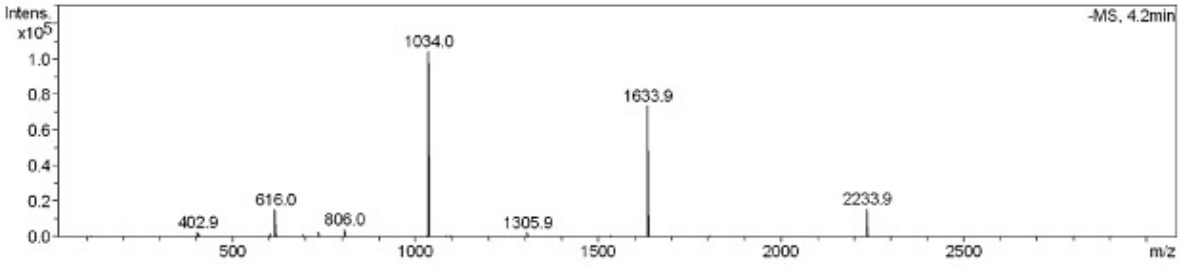
Analysis Name	D:\Data\Raj\1-3.d	Acquisition Date	2/28/2012 8:00:29 PM
Method	tune_wide_expert_lactamEAIS.m	Operator	BDAL@DE
Sample Name	1-3	Instrument / Ser#	micrOTOF-Q 10139
Comment			

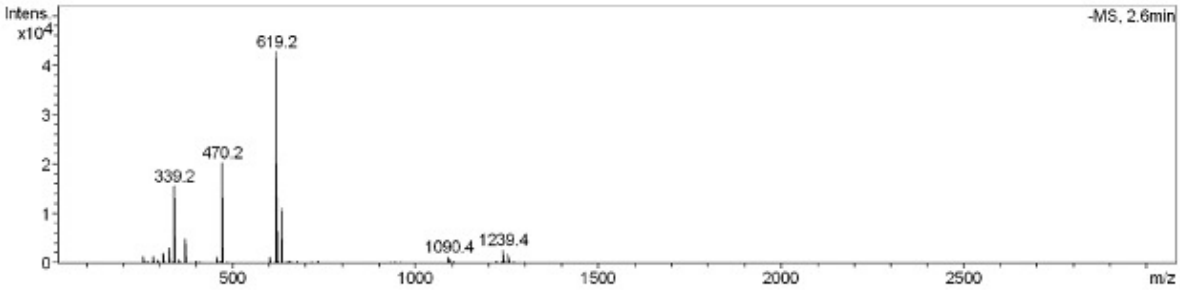
Acquisition Parameter

Source Type	ESI	Ion Polarity	Negative	Set Nebulizer	14.5 psi
Focus	Not active	Set Capillary	3500 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

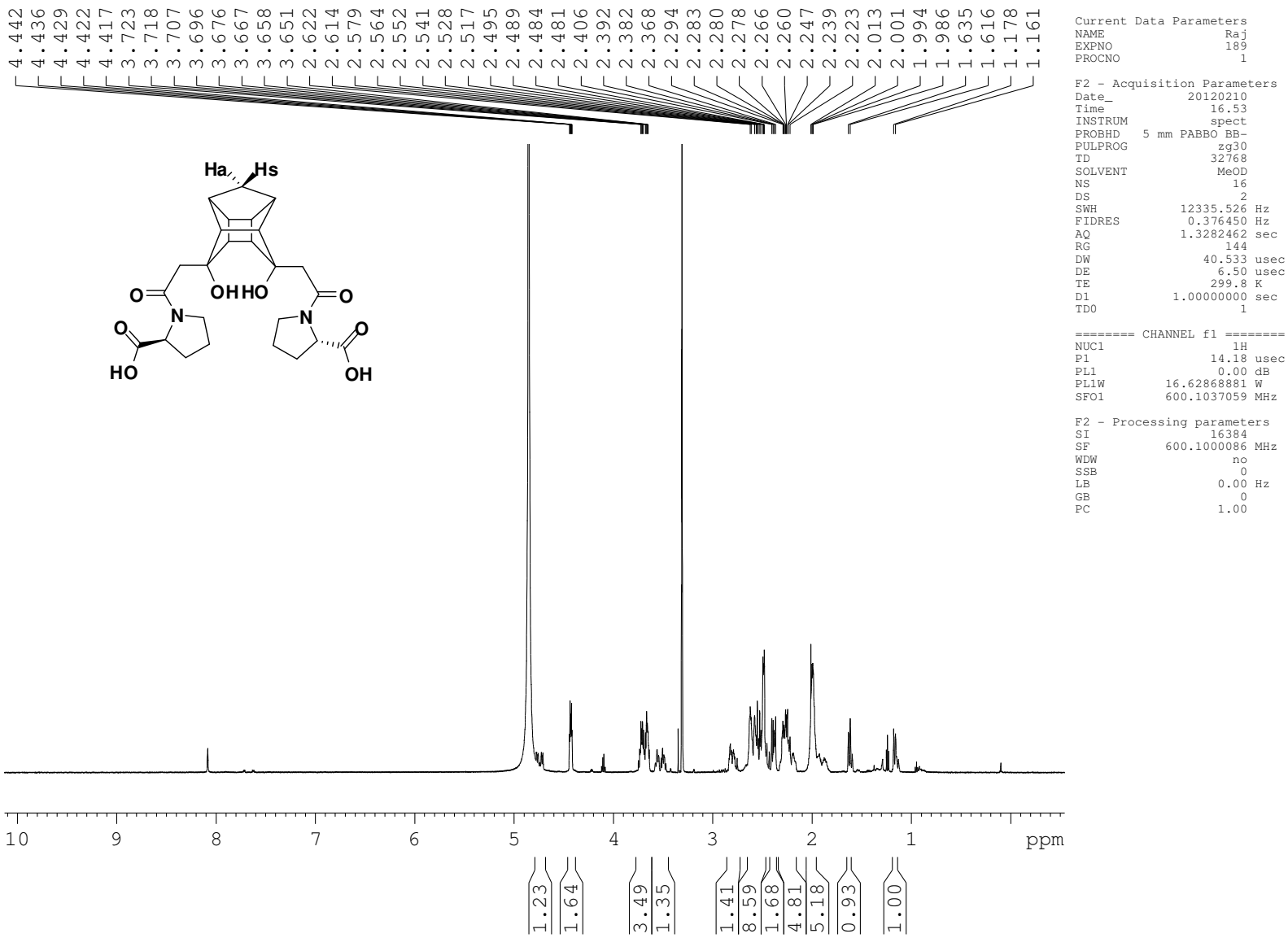




Meas. m/z	#	Formula	m/z	err [ppm]	Mean err [ppm]	rdb	N-Rule	e⁻ Conf	mSigma	Std I	Std Mean m/z	Std I VarNorm	Std m/z Diff	Std Comb Dev
619.2														
	1	C 28 H 35 N 4 O 12	619.2	37.1	37.8	13.5	ok	even	5.90	0.0109	0.0234	0.0046	0.0015	0.8565
	2	C 29 H 35 N 2 O 13	619.2	19.0	19.6	13.5	ok	even	7.99	0.0152	0.0122	0.0061	0.0015	0.7928
	3	C 25 H 35 N 2 O 16	619.2	-5.6	-5.0	9.5	ok	even	19.05	0.0279	0.0032	0.0099	0.0015	0.7140
	4	C 24 H 35 N 4 O 15	619.2	12.5	13.1	9.5	ok	even	20.67	0.0312	0.0082	0.0106	0.0015	0.8098

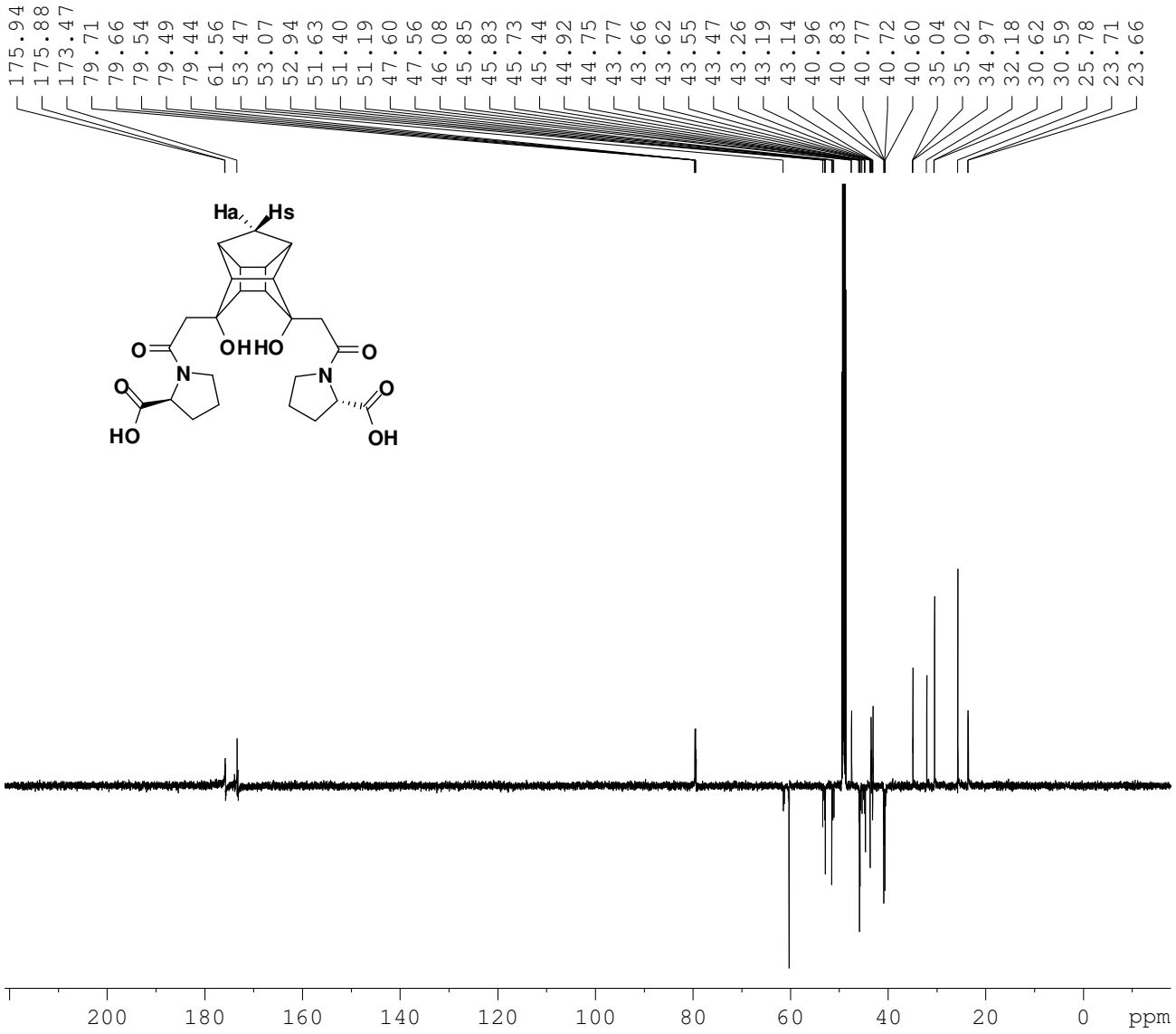
Meas. m/z	#	Formula	m/z	err [ppm]	Mean err [ppm]	rdb	N-Rule	e⁻ Conf	mSigma	Std I	Std Mean m/z	Std I VarNorm	Std m/z Diff	Std Comb Dev
	5	C 23 H 35 N 6 O 14	619.2	30.6	31.3	9.5	ok	even	22.44	0.0349	0.0194	0.0115	0.0014	0.9016
	6	C 33 H 35 N 2 O 10	619.2	43.6	44.3	17.5	ok	even	26.34	0.0436	0.0274	0.0142	0.0015	0.9434

HRESIMS (*m/z*) of Compound 11 (Chapter 1)



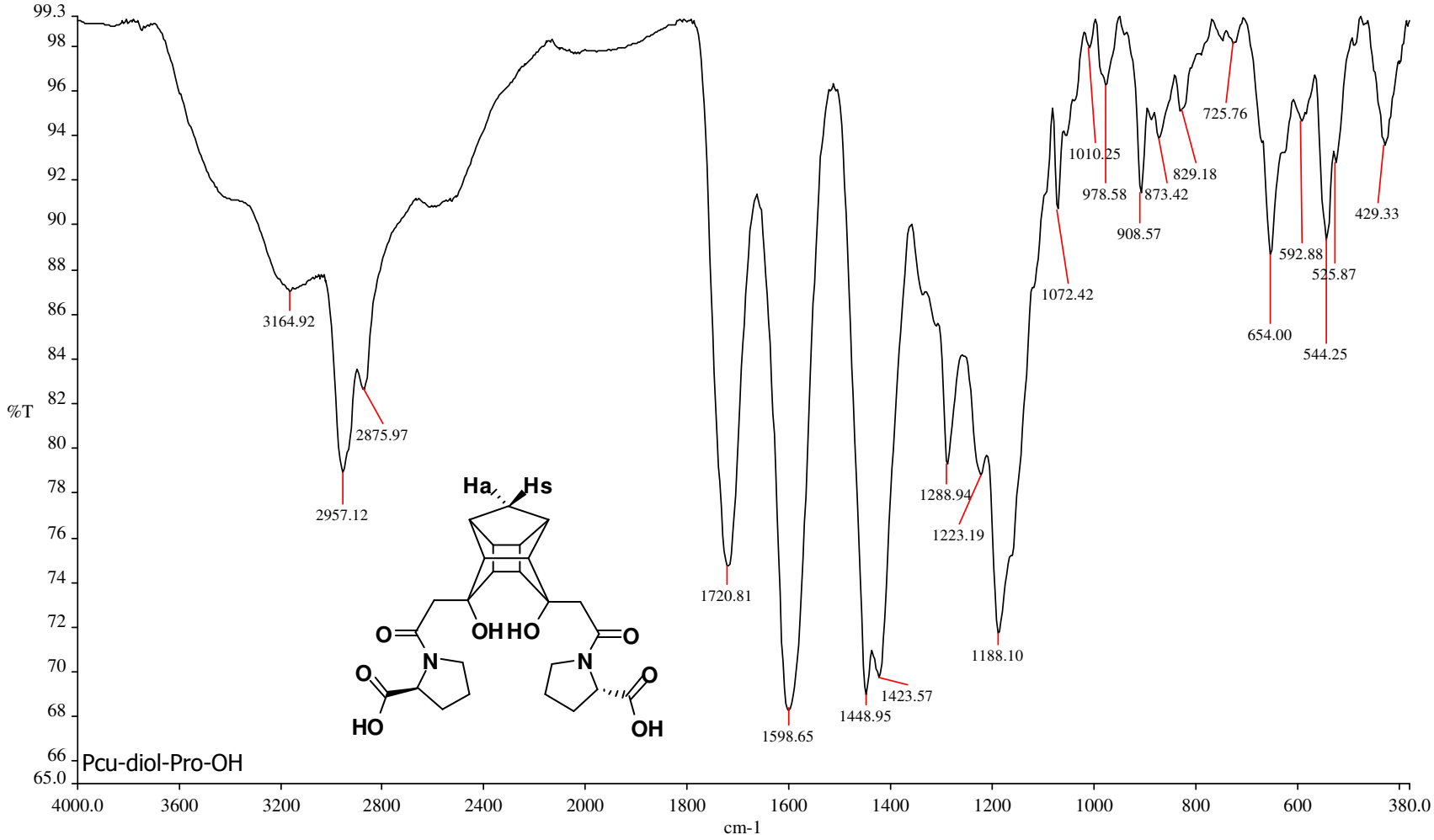

```
Current Data Parameters
NAME                  Raj
EXPNO                 189
PROCNO                  1

F2 - Acquisition Parameters
Date_              20120210
Time                 16.53
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG              zg30
TD                  32768
SOLVENT              MeOD
NS                     16
DS                      2
SWH              12335.526 Hz
FIDRES           0.376450 Hz
AQ               1.3282462 sec
RG                    144
DW                 40.533 usec
DE                  6.50 usec
TE                 299.8 K
D1              1.00000000 sec
TD0                     1

======== CHANNEL f1 ========
NUC1                   1H
P1                  14.18 usec
PL1                  0.00 dB
PL1W         16.62868881 W
SFO1         600.1037059 MHz

F2 - Processing parameters
SI                  16384
SF           600.1000086 MHz
WDW                    no
SSB                     0
LB                   0.00 Hz
GB                      0
PC                   1.00
```

^{1}H NMR Spectrum of Compound 12 (Chapter 1)




```
Current Data Parameters
NAME                  Raj
EXPNO                 190
PROCNO                  1

F2 - Acquisition Parameters
Date_           20120211
Time               17.12
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG             jmod
TD                 65536
SOLVENT             MeOD
NS                  4096
DS                     4
SWH           36057.691 Hz
FIDRES         0.550197 Hz
AQ            0.9088159 sec
RG                  2050
DW               13.867 usec
DE                6.50 usec
TE               300.6 K
CNST2        145.0000000
CNST11         1.0000000
D1            2.00000000 sec
D20           0.00689655 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                 13C
P1                  9.00 usec
P2                 18.00 usec
PL1                 1.50 dB
PL1W         59.57236099 W
SFO1        150.9103545 MHz

======== CHANNEL f2 ========
CPDPRG2          waltz16
NUC2                 1H
PCPD2              96.00 usec
PL2                -0.50 dB
PL12               15.40 dB
PL2W         18.65769577 W
PL12W         0.47957665 W
SFO2        600.1024004 MHz

F2 - Processing parameters
SI                 32768
SF          150.8950300 MHz
WDW                   EM
SSB                    0
LB                  1.00 Hz
GB                     0
PC                  1.40
```

^{13}C NMR Spectrum of Compound 12 (Chapter 1)



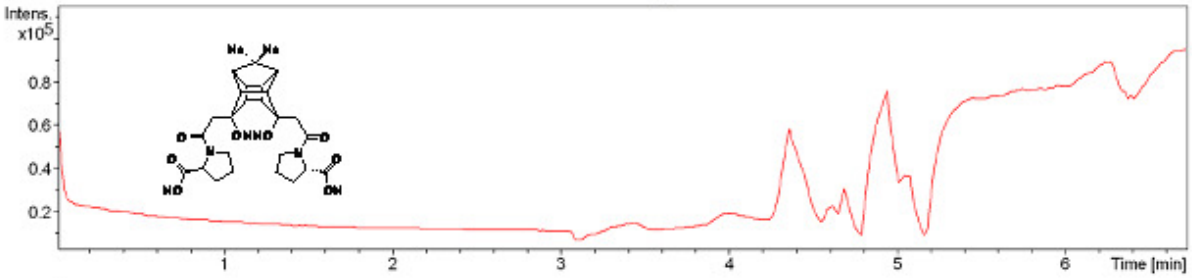
IR Spectrum of Compound 12 (Chapter 1)

Analysis Info

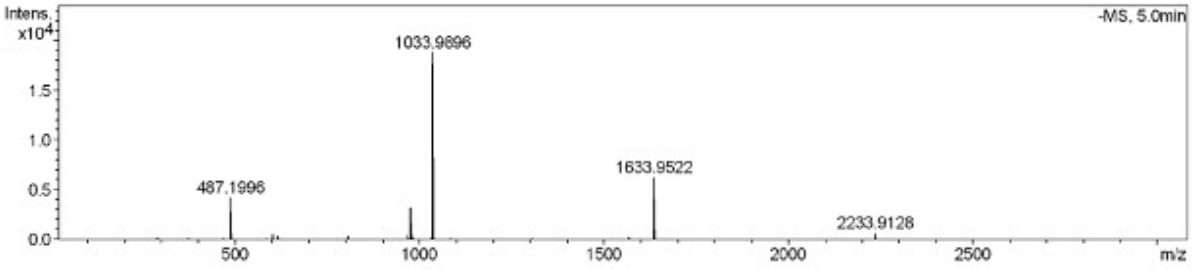
Analysis Name	D:\Data\Raj\P-Proline 2.d	Acquisition Date	2/24/2012 6:23:15 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	P-Proline 2	Instrument / Ser#	micrOTOF-Q 10139
Comment			

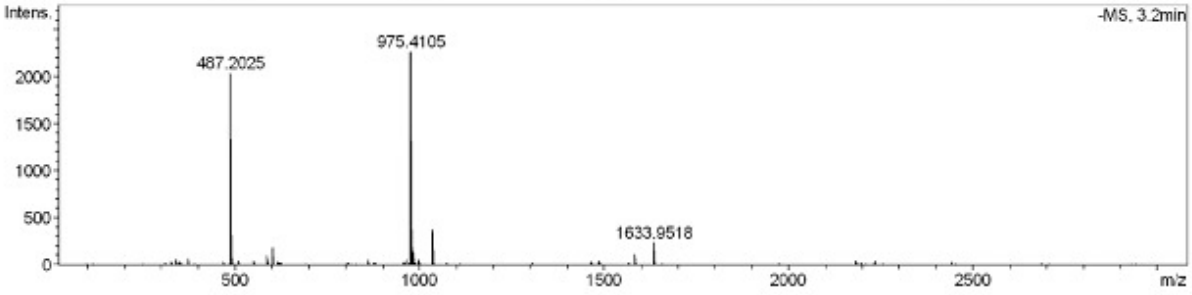
Acquisition Parameter

Source Type	ESI	Ion Polarity	Negative	Set Nebulizer	14.5 psi
Focus	Not active	Set Capillary	3500 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

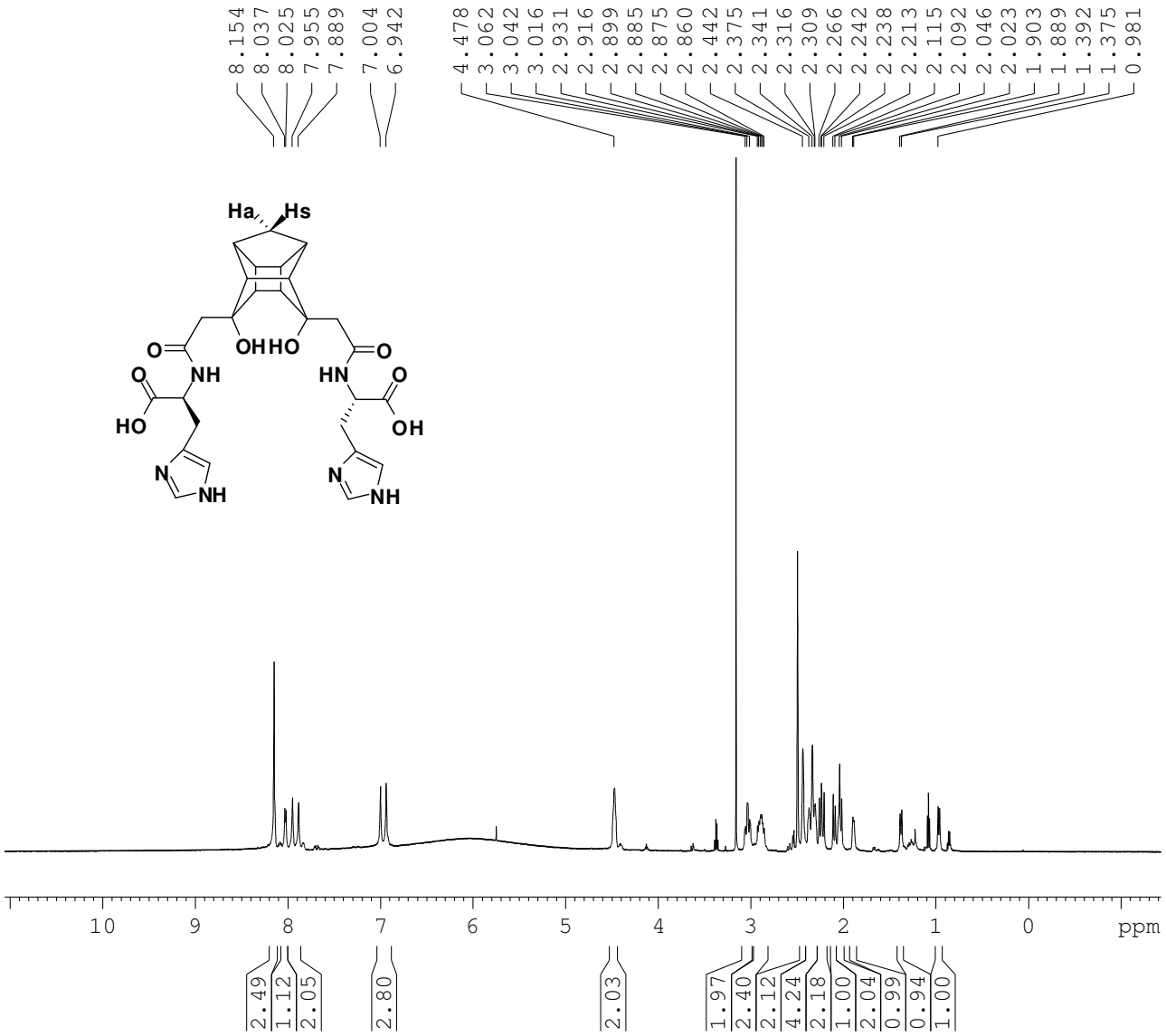




Meas. m/z	#	Formula	m/z	err [ppm]	Mean err [ppm]	rdb	N-Rule	e⁻ Conf	mSigma	Std I	Std Mean m/z	Std I VarNorm	Std m/z Diff	Std Comb Dev
487.2025														
	1	C 26 H 31 O 9	487.1974	-10.6	-9.8	11.5	ok	even	4.14	0.0059	0.0048	0.0033	0.0016	0.7995
	2	C 25 H 31 N 2 O 8	487.2086	12.4	13.2	11.5	ok	even	4.31	0.0070	0.0065	0.0038	0.0015	0.8504
	3	C 27 H 27 N 4 O 5	487.1987	-7.9	-7.1	16.5	ok	even	14.28	0.0214	0.0035	0.0081	0.0015	0.8427
	4	C 28 H 23 N 8 O	487.2000	-5.1	-4.4	21.5	ok	even	26.74	0.0402	0.0022	0.0135	0.0015	0.8610

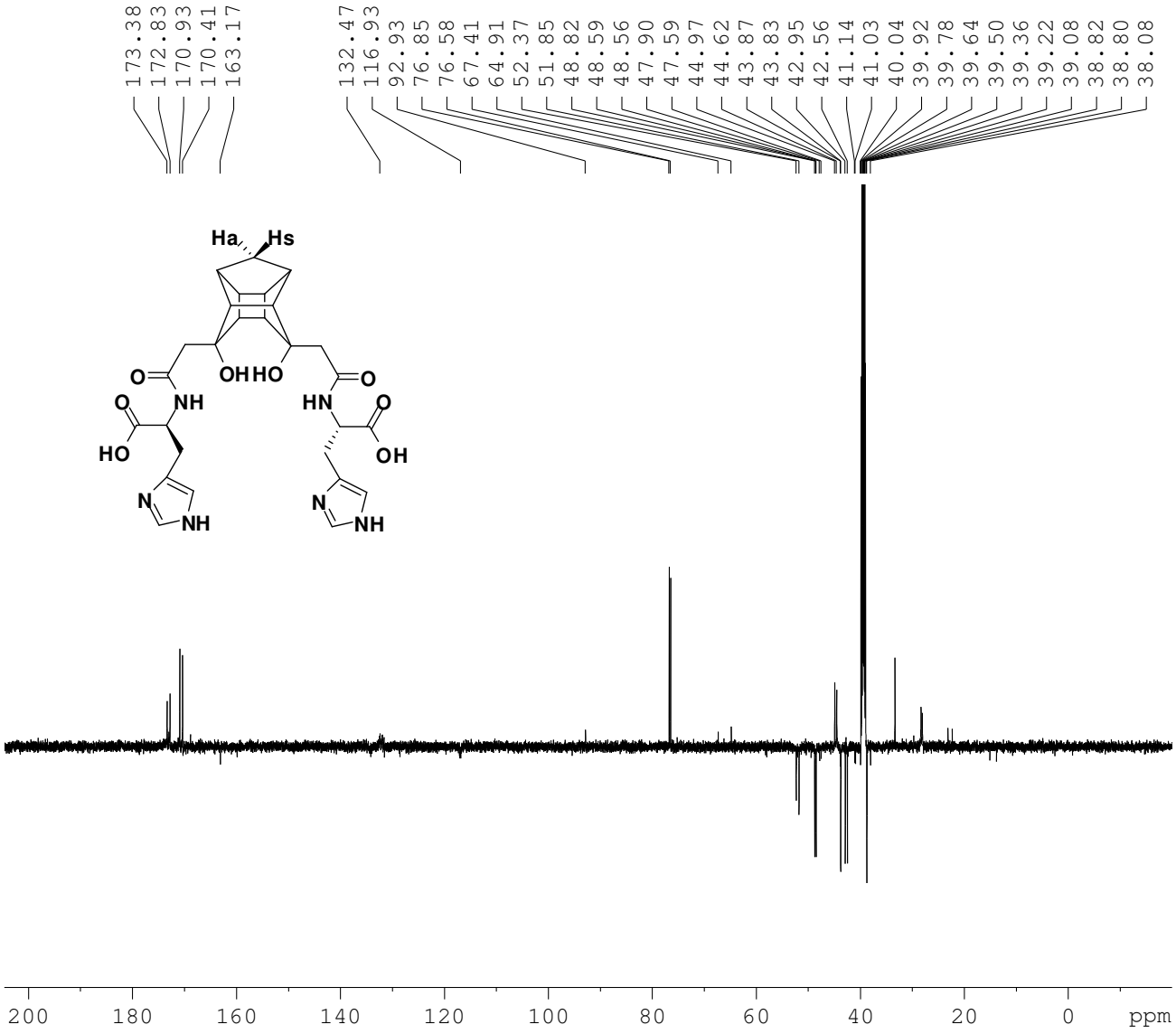
Meas. m/z	#	Formula	m/z	err [ppm]	Mean err [ppm]	rdb	N-Rule	e⁻ Conf	mSigma	Std I	Std Mean m/z	Std I VarNorm	Std m/z Diff	Std Comb Dev
	5	C 32 H 27 N 2 O 3	487.2027	0.4	1.2	20.5	ok	even	37.20	0.0581	0.0009	0.0186	0.0016	0.8460
	6	C 37 H 27 O	487.2067	8.6	9.4	24.5	ok	even	60.94	0.0933	0.0046	0.0292	0.0016	0.9691

HRESIMS (*m/z*) of Compound 12 (Chapter 1)



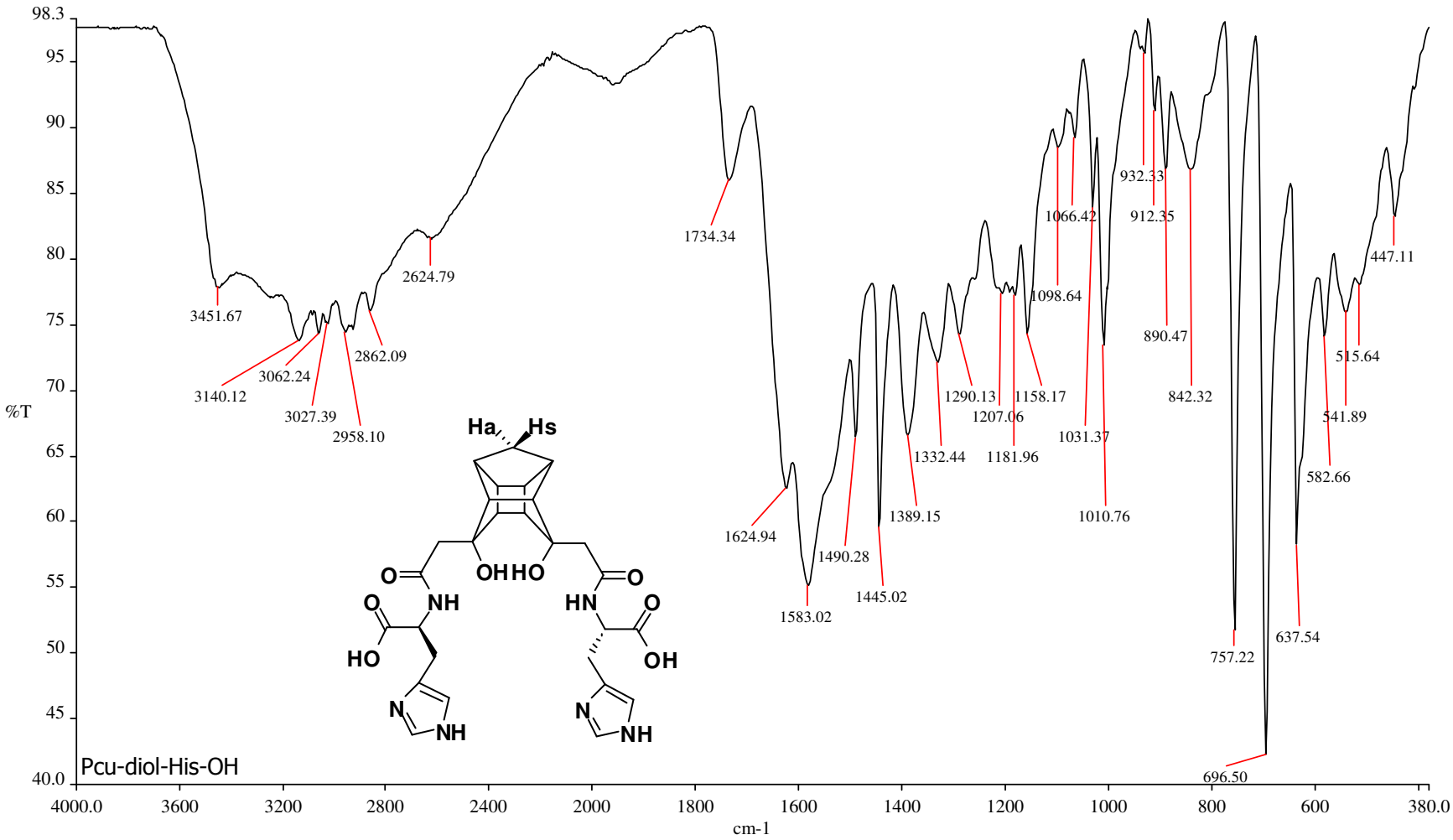

```
Current Data Parameters
NAME                  Raj
EXPNO                 219
PROCNO                  1

F2 - Acquisition Parameters
Date_           20120402
Time               22.17
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG             zg30
TD                 32768
SOLVENT             DMSO
NS                    16
DS                     2
SWH            12335.526 Hz
FIDRES         0.376450 Hz
AQ             1.3282462 sec
RG                   144
DW                40.533 usec
DE                 6.50 usec
TE                295.0 K
D1             1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                  1H
P1                 15.10 usec
PL1                0.00 dB
PL1W        16.62868881 W
SFO1        600.1037059 MHz

F2 - Processing parameters
SI                 16384
SF          600.1000060 MHz
WDW                   EM
SSB                    0
LB                 0.30 Hz
GB                     0
PC                 1.00
```

^{1}H NMR Spectrum of Compound 13 (Chapter 1)




```
Current Data Parameters
NAME                  Raj
EXPNO                 220
PROCNO                  1

F2 - Acquisition Parameters
Date_           20120403
Time               23.22
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG             jmod
TD                 65536
SOLVENT              DMSO
NS                   4800
DS                      4
SWH            36057.691 Hz
FIDRES          0.550197 Hz
AQ              0.9088159 sec
RG                   2050
DW                13.867 usec
DE                 6.50 usec
TE                298.0 K
CNST2         145.0000000
CNST11          1.0000000
D1             2.00000000 sec
D20            0.00689655 sec
TD0                     1

======== CHANNEL f1 ========
NUC1                  13C
P1                   9.00 usec
P2                  18.00 usec
PL1                  1.50 dB
PL1W         59.57236099 W
SFO1        150.9103545 MHz

======== CHANNEL f2 ========
CPDPRG2           waltz16
NUC2                  1H
PCPD2               96.00 usec
PL2                 -0.50 dB
PL12                15.40 dB
PL2W         18.65769577 W
PL12W         0.47957665 W
SFO2        600.1024004 MHz

F2 - Processing parameters
SI                   32768
SF          150.8953361 MHz
WDW                    EM
SSB                     0
LB                   1.00 Hz
GB                      0
PC                   1.40
```

^{13}C NMR Spectrum of Compound 13 (Chapter 1)



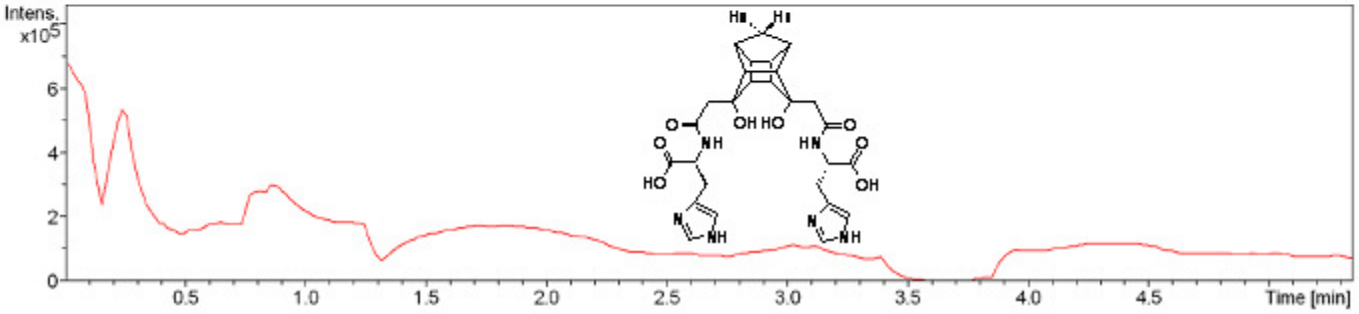
IR Spectrum of Compound 13 (Chapter 1)

Analysis Info

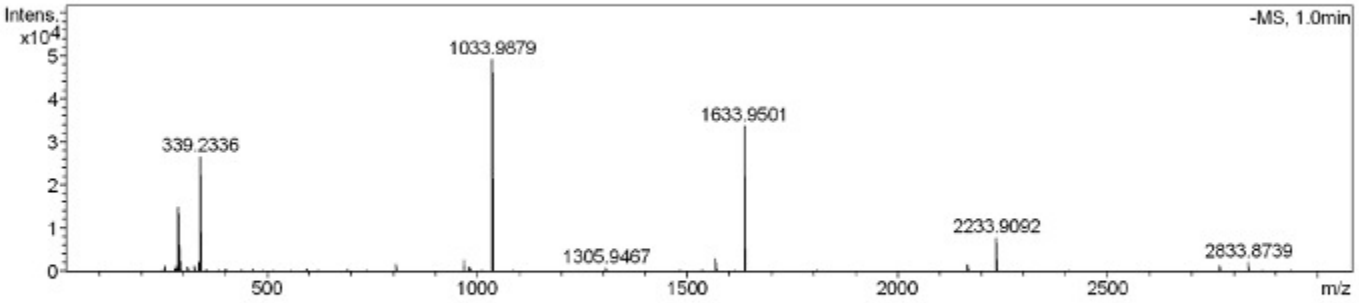
Analysis Name	D:\Data\Raj\P-Histidine 3.d	Acquisition Date	2/24/2012 6:51:20 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	P-Histidine 3	Instrument / Ser#	micrOTOF-Q 10139
Comment			

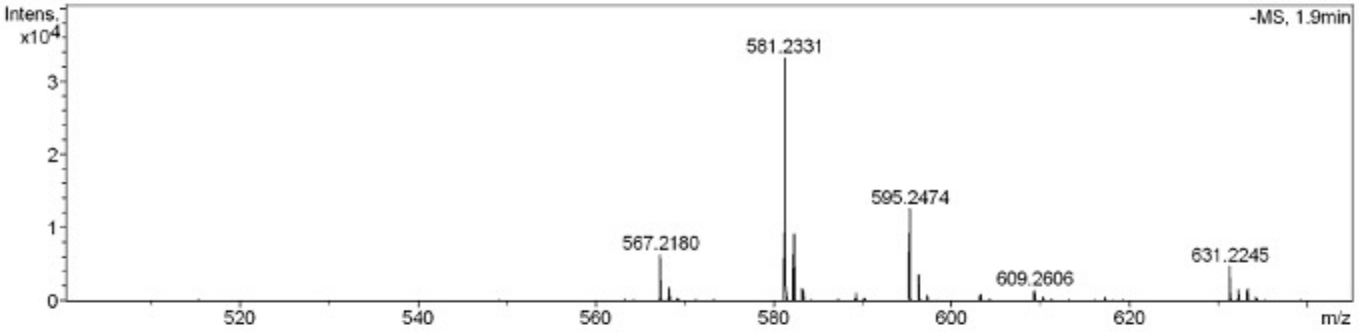
Acquisition Parameter

Source Type	ESI	Ion Polarity	Negative	Set Nebulizer	14.5 psi
Focus	Not active	Set Capillary	3500 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	mSigma	Std I	Std Mean m/z	Std I VarNorm	Std m/z Diff	Std Comb Dev
567.2180														
	1	C 26 H 35 N 2 O 12	567.2195	2.7	3.4	10.5	ok	even	4.83	0.0071	0.0020	0.0045	0.0015	0.7397
	2	C 23 H 27 N 12 O 6	567.2182	0.3	0.9	16.5	ok	even	6.94	0.0116	0.0008	0.0063	0.0011	0.6481
	3	C 27 H 31 N 6 O 8	567.2209	5.1	5.7	15.5	ok	even	10.14	0.0159	0.0033	0.0068	0.0015	0.8431

HRESIMS (*m/z*) of Compound 13 (Chapter 1)

APPENDIX 2

SUPPLEMENTARY INFORMATION FOR CHAPTER 4 AND 5

Chapter 4

Pentacycloundecane derived hydroxyl acid peptides: A new class of irreversible non-scissile bridgedethertype isoster as potential HIV-1 wild type C-SA protease inhibitors

Rajshekhar Karpoormath,[a] Yasien Sayed,[c] Patrick Govender,[b] Glenn E. M. Maguire,[a] Thavendran Govender,[d] Hendrik G. Kruger[a*] and Mahmoud E. S. Soliman[e*]

[a]School of Chemistry, University of KwaZulu Natal, Durban 4001, South Africa
[b]School of Biochemistry, University of KwaZulu Natal, Durban 4001, South Africa
[c]Protein Structure-Function Research Unit, School of Molecular and Cell Biology, University of the Witwatersrand, Wits 2050, South Africa

[d]School of Pharmacy and Pharmacology, University of KwaZulu Natal, Durban 4001, South Africa
[e]Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Zagazig University, Zagazig 44519, Egypt
**** Corresponding authors: email: mess20@bath.ac.uk and kruger@ukzn.ac.za***

Chapter 5

Synthesis and NMR elucidation of Pentacycloundecane derived hydroxy acid peptides as potential anti-HIV-1 agents

Rajshekhar Karpoormath,[a,b] Fernando Albericio,[c,d,e] Thavendran Govender,[b] Glenn E. M. Maguire,[a] and Hendrik G. Kruger[a*]

[a]School of Chemistry, University of KwaZulu-Natal, Durban 4001, South Africa
[b]School of Pharmaceutical Sciences, University of KwaZulu-Natal, Durban 4001, South Africa
[c]Institute for research in Biomedicine, Barcelona Science Park, Baldiri Reixac 10, 08028-Barcelona, Spain

[d]CIBER-BBN, Networking Centre on Bioengineering, Biomaterials and Nanomedicine, Barcelona Science Park, Baldiri Reixac 10, 08028-Barcelona, Spain
[e]Department of Organic Chemistry, University of Barcelona, Martí i Franque´s 1-11, 08028-Barcelona, Spain
Corresponding author: *Email: kruger@ukzn.ac.za, Fax: +27-31-2603091

Supporting Information

Materials and methods

Analytical analysis was performed on an Agilent 1100 HPLC (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 Instrument (Ace C18 150 mm x 21.2 mm x 5 microns) with a UV/VIS detector (215 nm) and automated fraction collector. A two- mobile phase system was employed, utilizing formic acid as the ion-pairing agent. Mobile phase A consisted of 0.1% formic acid/H_2O (v/v) and mobile phase 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. Microwave couplings were conducted on a Discovery CEM Liberty microwave peptide synthesizer (see table 1 for conditions, not same as supplier). The ^{13}C NMR and ^{1}H NMR data were recorded on a Bruker AVANCE III 400 MHz spectrometer. Some samples were analyzed using a Bruker AVANCE III 600MHz when higher sensitivity was required.

Supplementary table 1: Microwave conditions for coupling and deprotection

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

Over-expression, extraction and purification of the C-SA protease (Chapter 4)

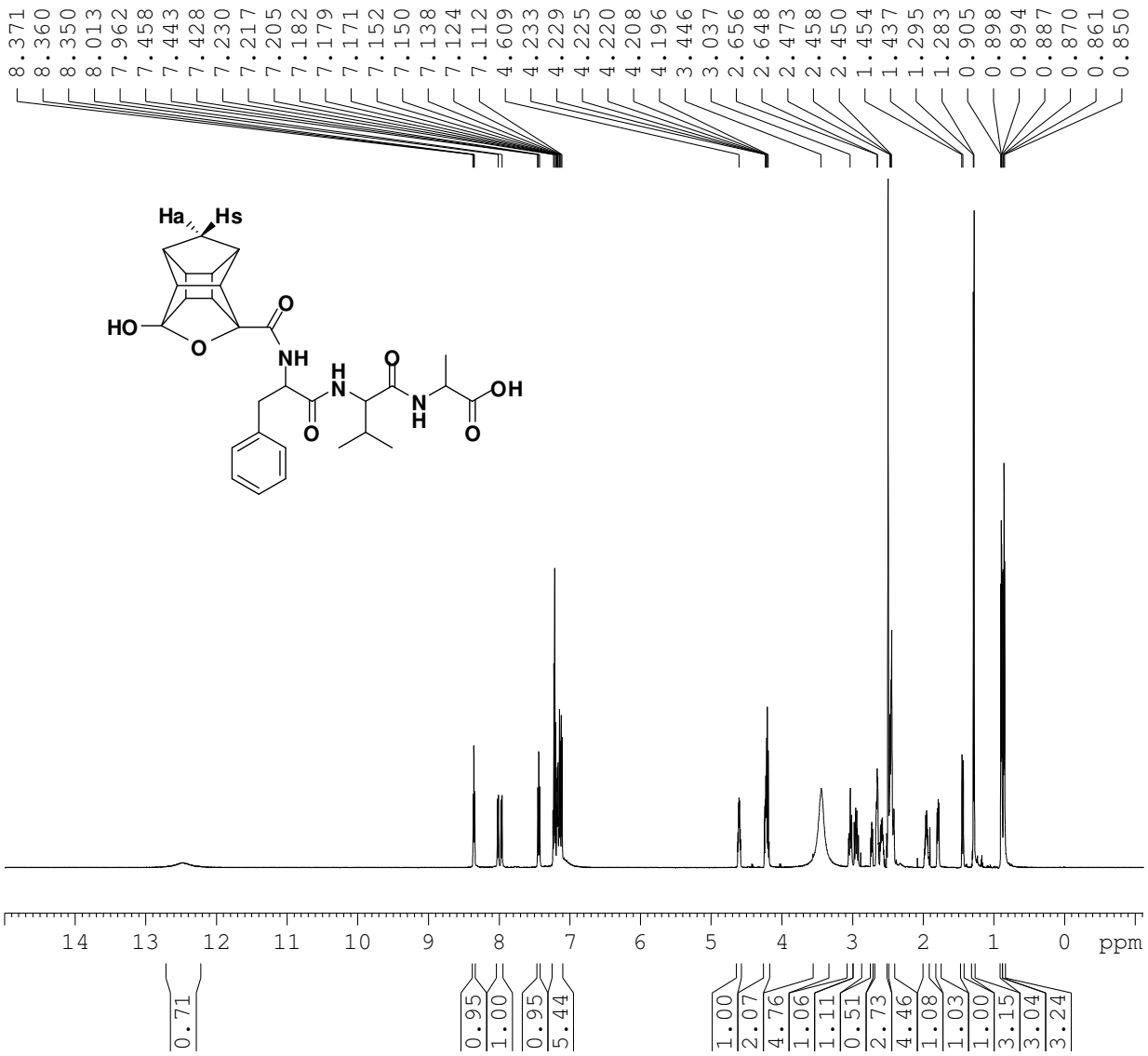
Plasmid encoding HIV-1 subtype C protease (containing the mutation Q7K designed to reduce the hypersensitive autolytic site) is expressed as inclusion bodies (Ido *et al*., 1991) in *Escherichia coli* BL21 (DE3) pLysS cells. Briefly, *Escherichia coli* cells harboring the plasmid DNA were grown at 37 °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 (OD_{600}) 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 $MgCl_2$ 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.

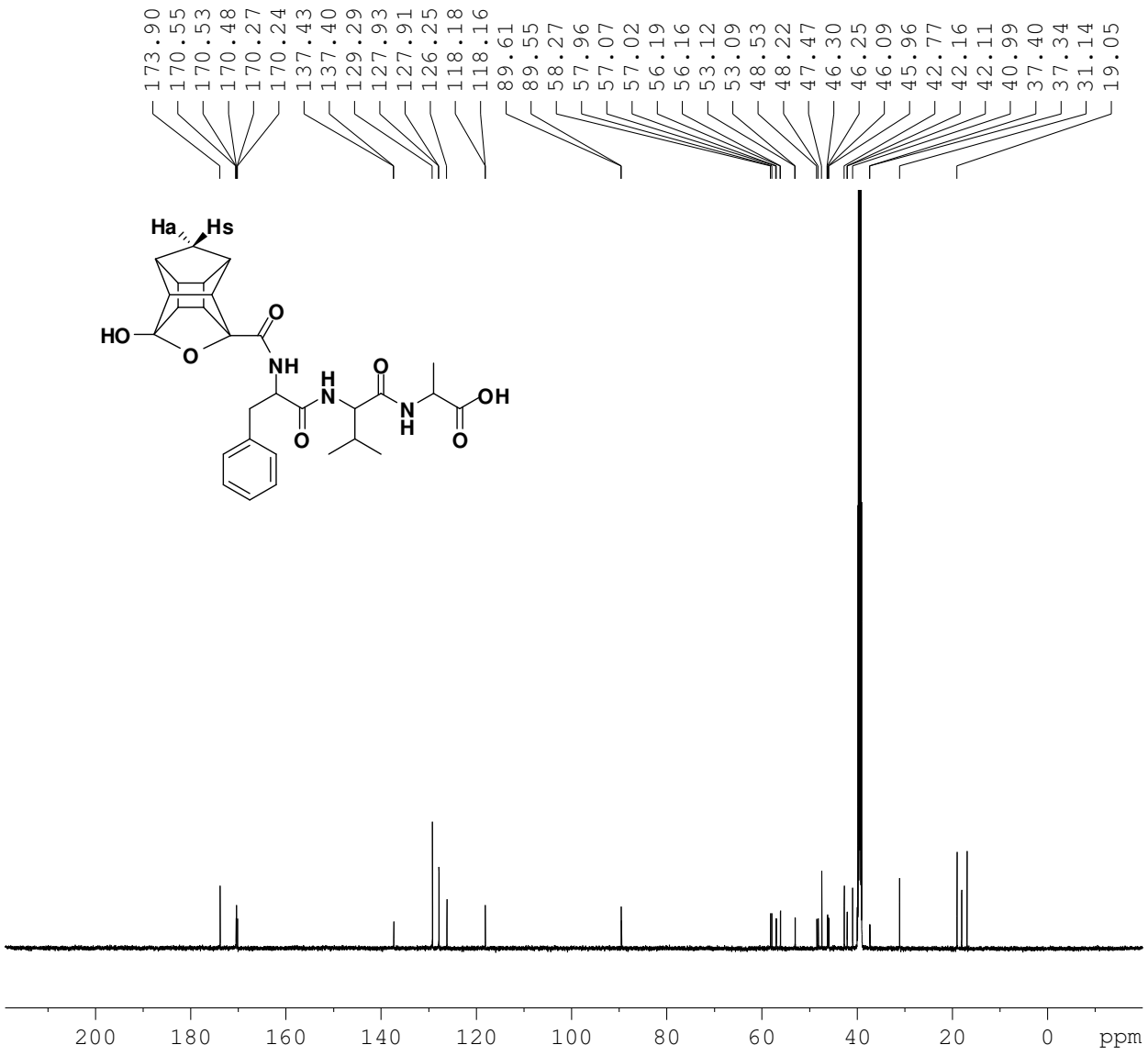
The protease, in the supernatant, was purified by passing through an anion exchange (DEAE) column previously equilibrated with solubilization buffer. Upon elution from the column, the protease was acidified by adding formic acid to a final concentration of 25 mM. Precipitation of significant amount of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



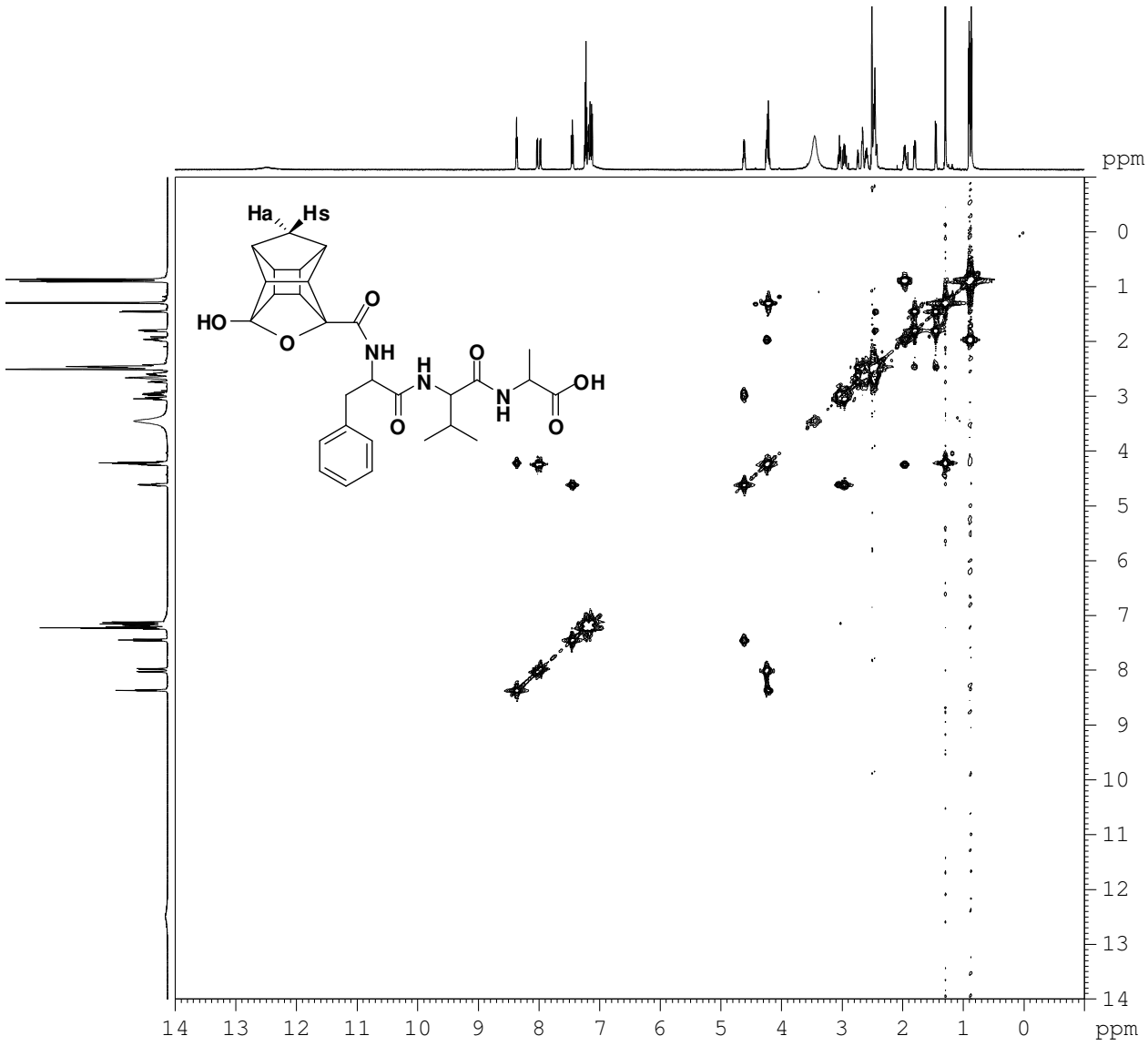

```
Current Data Parameters
NAME                 Raj 2
EXPNO                  113
PROCNO                   1

F2 - Acquisition Parameters
Date_             20101109
Time                 16.53
INSTRUM              spect
PROBHD    5 mm PABBO BB-
PULPROG              zg30
TD                  32768
SOLVENT              DMSO
NS                     16
DS                      2
SWH              12335.526 Hz
FIDRES           0.376450 Hz
AQ               1.3282462 sec
RG                    161
DW                 40.533 usec
DE                  6.50 usec
TE                 298.1 K
D1              1.00000000 sec
TD0                     1

======== CHANNEL f1 ========
NUC1                   1H
P1                 13.57 usec
PL1                 0.00 dB
PL1W         16.62868881 W
SFO1         600.1037059 MHz

F2 - Processing parameters
SI                  16384
SF           600.1000053 MHz
WDW                    no
SSB                     0
LB                   0.00 Hz
GB                      0
PC                   1.00
```

^{1}H NMR Spectrum of Compound 16 (Chapter 4) and Compound 6 (Chapter 5)



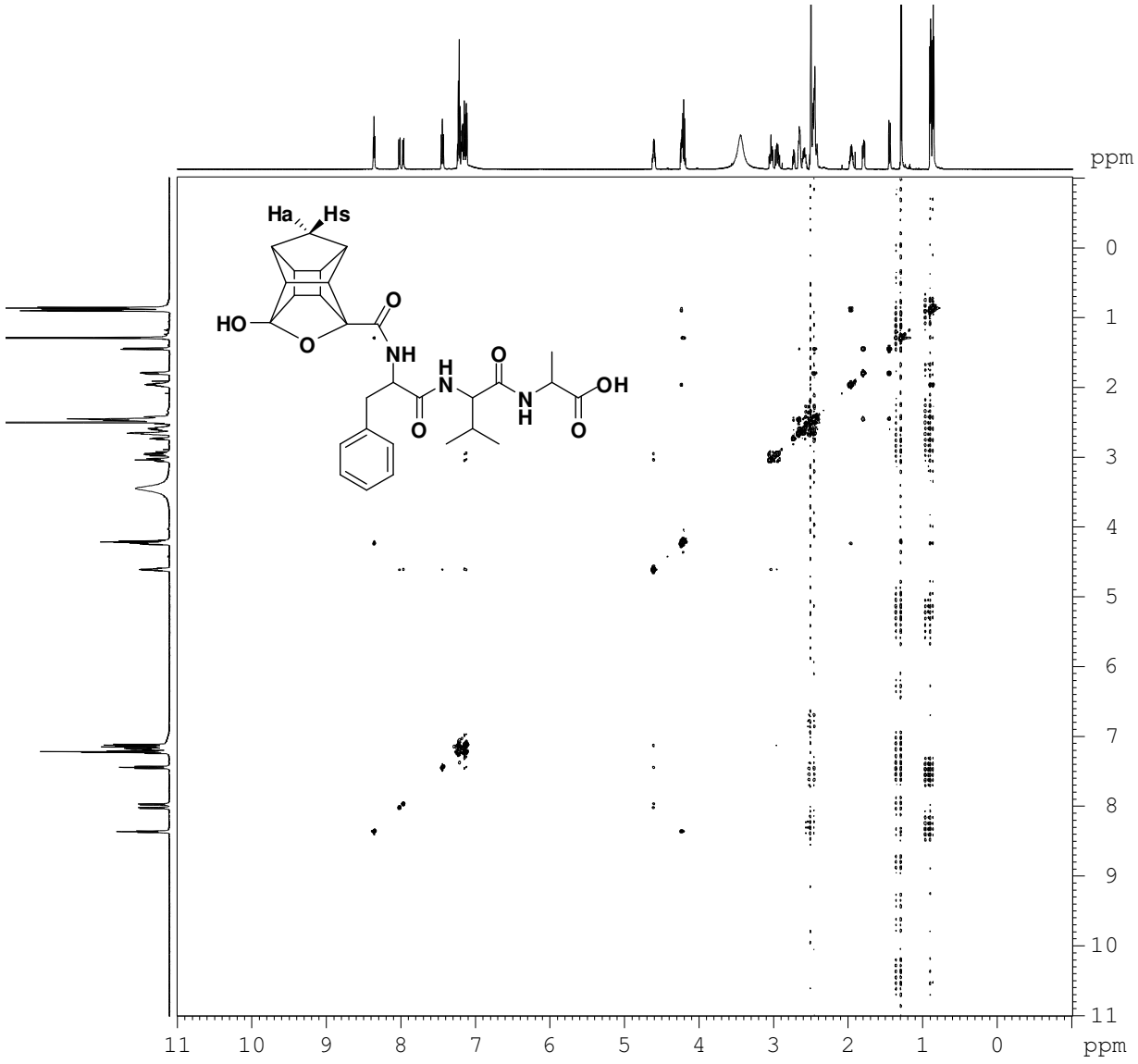

```
Current Data Parameters
NAME                 Raj 2
EXPNO                  118
PROCNO                   1

F2 - Acquisition Parameters
Date_             20101110
Time                 6.00
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG            zgpg30
TD                  65536
SOLVENT               DMSO
NS                     4800
DS                        4
SWH              36057.691 Hz
FIDRES            0.550197 Hz
AQ                0.9088159 sec
RG                     2050
DW                  13.867 usec
DE                    6.50 usec
TE                   298.2 K
D1              2.00000000 sec
D11             0.03000000 sec
TD0                      1

======== CHANNEL f1 ========
NUC1                   13C
P1                    9.00 usec
PL1                   1.50 dB
PL1W          59.57236099 W
SFO1         150.9103545 MHz

======== CHANNEL f2 ========
CPDPRG2            waltz16
NUC2                   1H
PCPD2                96.00 usec
PL2                  -0.50 dB
PL12                 15.40 dB
PL13                 18.00 dB
PL2W          18.65769577 W
PL12W          0.47957665 W
PL13W          0.26354694 W
SFO2         600.1024004 MHz

F2 - Processing parameters
SI                    32768
SF           150.8953370 MHz
WDW                     EM
SSB                      0
LB                    1.00 Hz
GB                       0
PC                    1.40
```

^{13}C NMR Spectrum of Compound 16 (Chapter 4) and Compound 6 (Chapter 5)



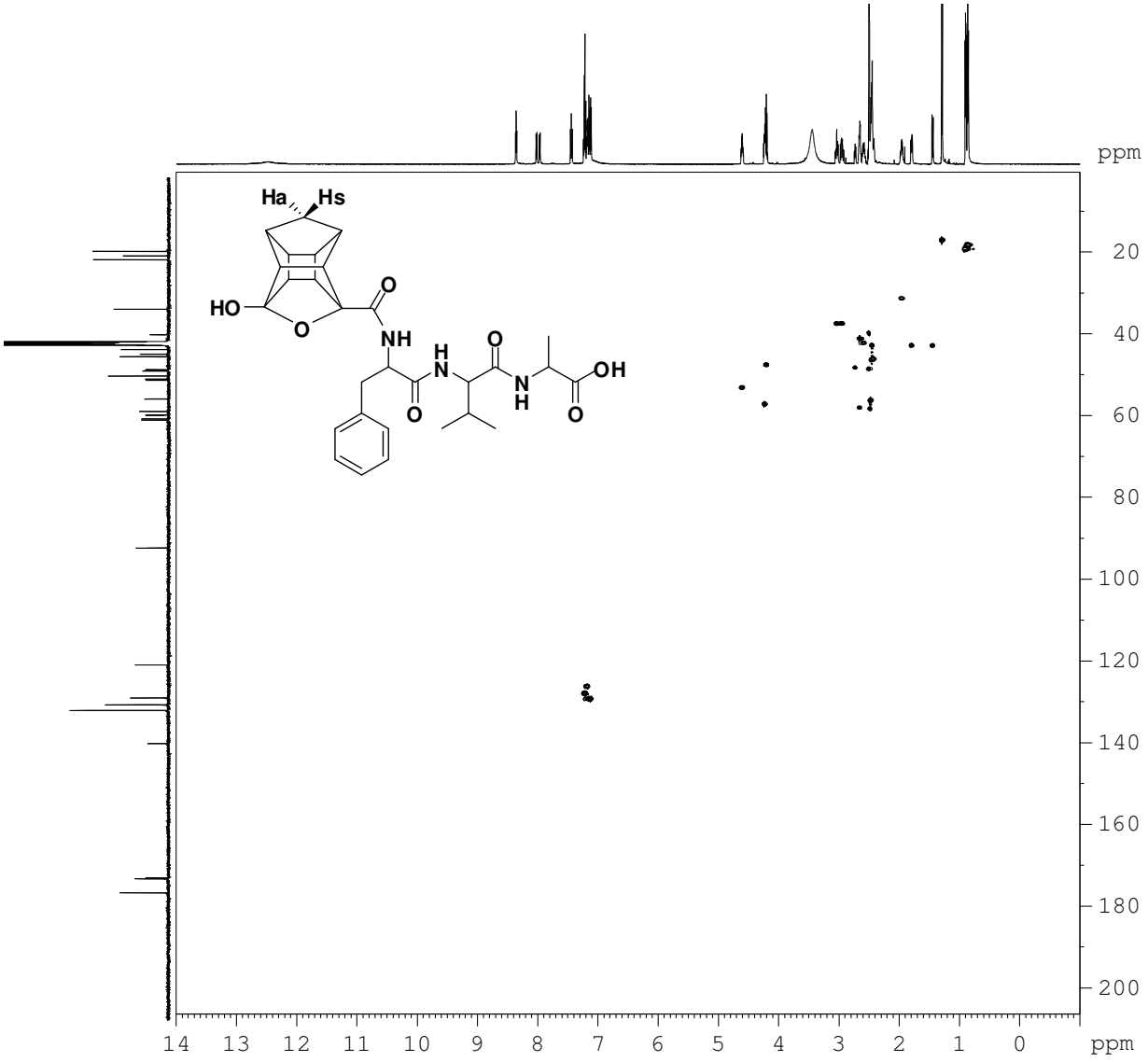

```
Current Data Parameters
NAME                 Raj 2
EXPNO                  115
PROCNO                   1

F2 - Acquisition Parameters
Date_             20101109
Time                19.28
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG          cosygpqf
TD                   2048
SOLVENT               DMSO
NS                       2
DS                       8
SWH              12335.526 Hz
FIDRES               6.023206 Hz
AQ                0.0830623 sec
RG                      64
DW                   40.533 usec
DE                    6.50 usec
TE                   298.0 K
D0              0.00000300 sec
D1              3.00000000 sec
D13             0.00000400 sec
D16             0.00020000 sec
IN0             0.00008120 sec

======== CHANNEL f1 ========
NUC1                    1H
P0                   13.56 usec
P1                   13.57 usec
PL1                   0.00 dB
PL1W          16.62868881 W
SFO1          600.1037059 MHz

====== GRADIENT CHANNEL =====
GPNAM1            SINE.100
GPZ1                10.00 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                      1
TD                     512
SFO1             600.1037 MHz
FIDRES           24.053263 Hz
SW                  20.522 ppm
FnMODE                  QF

F2 - Processing parameters
SI                    1024
SF            600.0999970 MHz
WDW                   SINE
SSB                      0
LB                    0.00 Hz
GB                       0
PC                    1.00

F1 - Processing parameters
SI                    1024
MC2                     QF
SF            600.0999950 MHz
WDW                   SINE
SSB                      0
LB                    0.00 Hz
```

COSY NMR Spectrum of Compound 6 (Chapter 5)



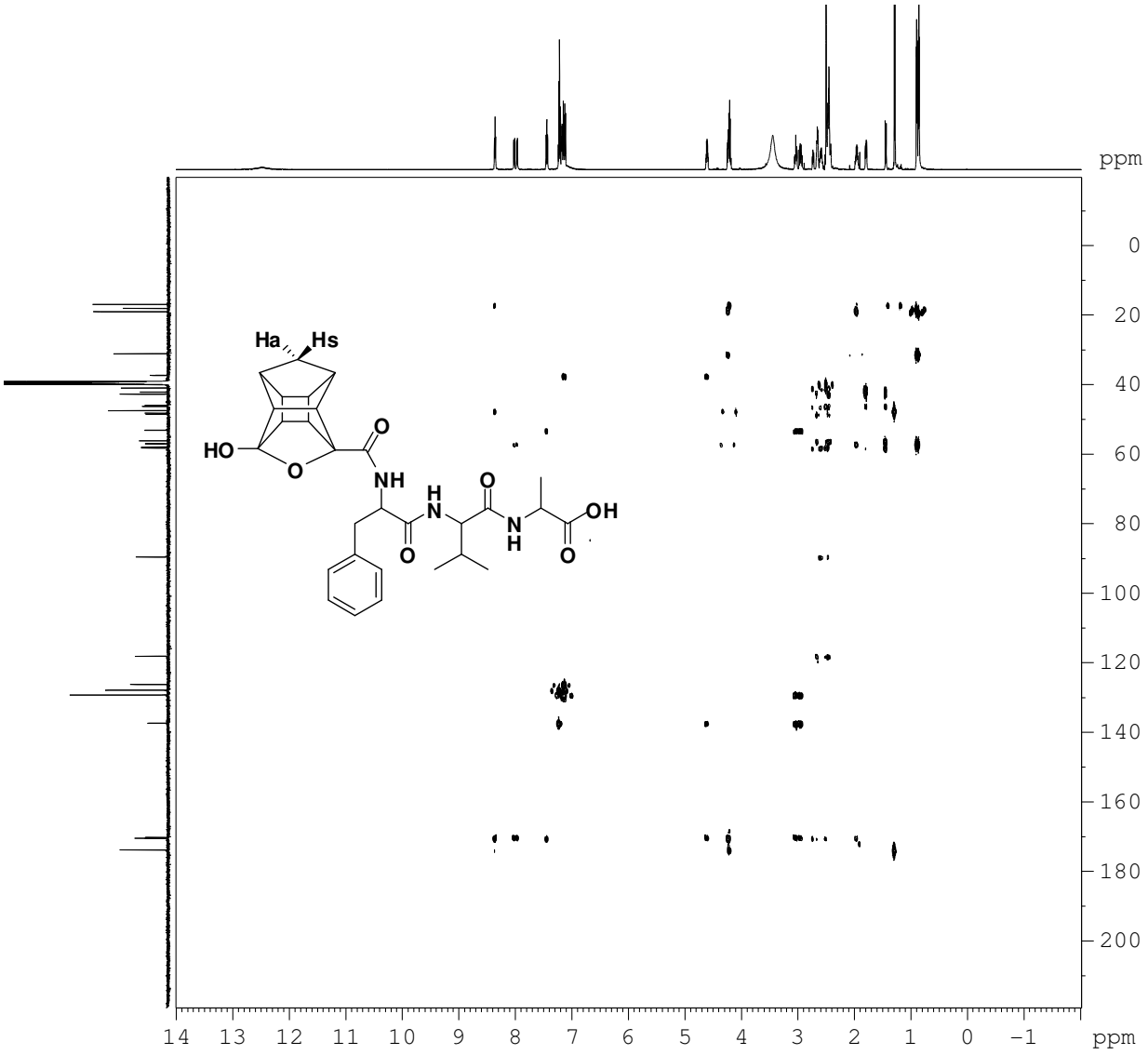

```
Current Data Parameters
NAME                  Raj 2
EXPNO                   114
PROCNO                    1

F2 - Acquisition Parameters
Date_              20101109
Time                  16.58
INSTRUM               spect
PROBHD   5 mm PABBO BB-
PULPROG             kproesy
TD                     2048
SOLVENT                DMSO
NS                        8
DS                       64
SWH                7211.539 Hz
FIDRES             3.521259 Hz
AQ                 0.1420447 sec
RG                      22.6
DW                    69.333 usec
DE                     6.00 usec
TE                    298.3 K
CNST2           6849.0000000
D0                 0.00005206 sec
D1                 1.50000000 sec
D11                0.03000000 sec
D12                0.00002000 sec
D13                0.00000400 sec
D16                0.00030000 sec
IN0                0.00013865 sec

======== CHANNEL f1 ========
CNST1               0.0000000
CNST3            3954.3879395
CNST4           -3954.3879395
NUC1                       1H
P1                      13.57 usec
P7                   250000.00 usec
P11                   1000.00 usec
PL1                      0.00 dB
PL9                     70.00 dB
PL1W               16.62868881 W
PL9W                0.00000166 W
SFO1               600.1030005 MHz
SP1                     13.00 dB
SP2                     13.00 dB
SP3                     13.00 dB
SP4                     13.00 dB
SP5                     13.00 dB
SP6                     13.00 dB
SPNAM1    gaussramp-down
SPNAM2         square.100
SPNAM3       gaussramp-up
SPNAM4    gaussramp+down
SPNAM5         square.100
SPNAM6       gaussramp+up
SPOAL1                1.000
SPOAL2                0.500
SPOAL3                0.000
SPOAL4                1.000
SPOAL5                0.500
SPOAL6                0.000
SPOFFS1                0.00 Hz
SPOFFS2                0.00 Hz
SPOFFS3                0.00 Hz
```

ROESY NMR Spectrum of Compound 6 (Chapter 5)



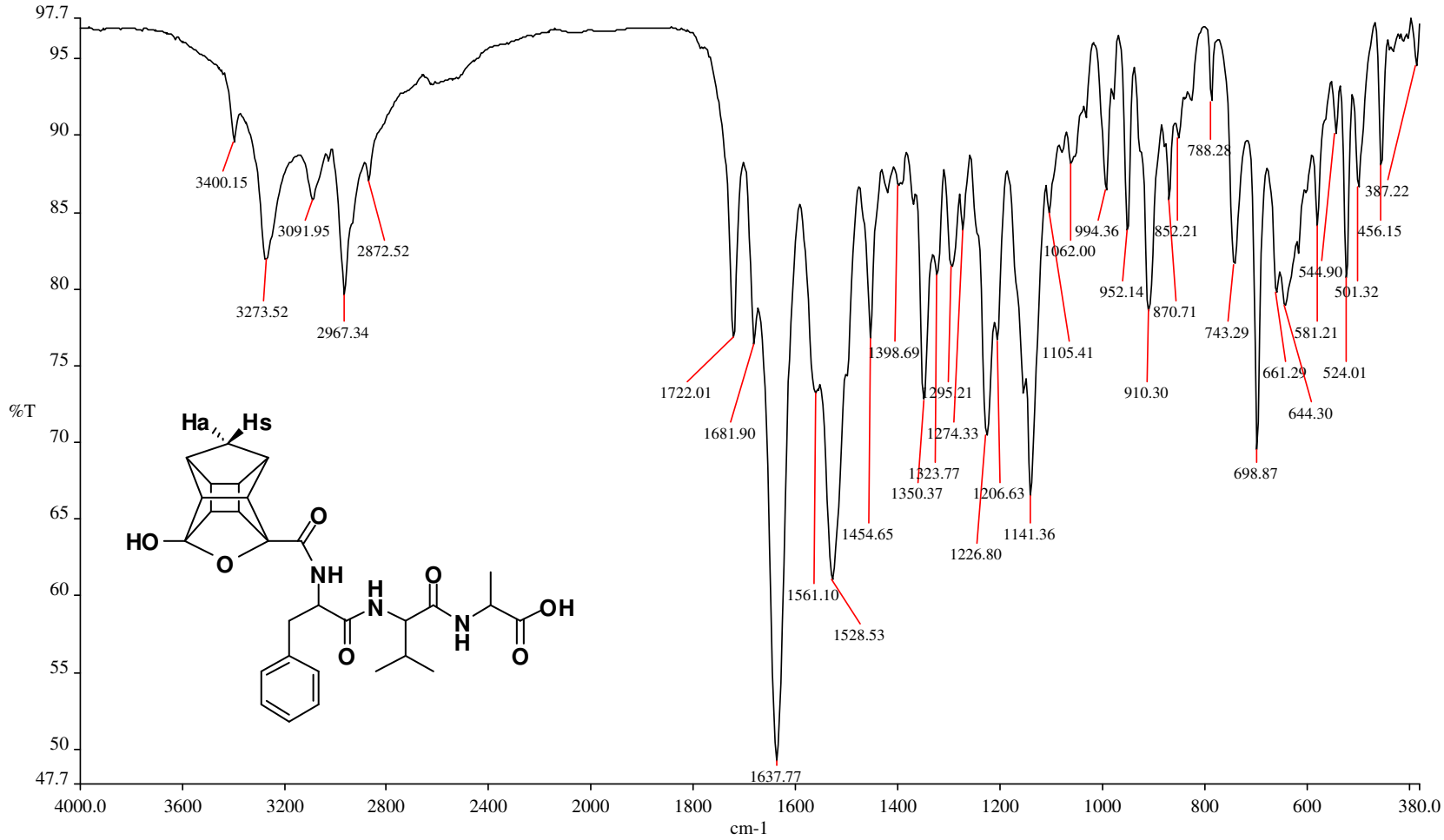

```
Current Data Parameters
NAME                 Raj 2
EXPNO                  116
PROCNO                   1

F2 - Acquisition Parameters
Date_             20101109
Time                 20.23
INSTRUM              spect
PROBHD    5 mm PABBO BB-
PULPROG          hsqcetgp
TD                   2048
SOLVENT               DMSO
NS                       8
DS                      16
SWH              12335.526 Hz
FIDRES           6.023206 Hz
AQ               0.0830623 sec
RG                    2050
DW                 40.533 usec
DE                   6.50 usec
TE                  298.3 K
CNST2          145.0000000
D0              0.00000300 sec
D1              1.00000000 sec
D4              0.00172414 sec
D11             0.03000000 sec
D13             0.00000400 sec
D16             0.00020000 sec
IN0             0.00001385 sec
ZGOPTNS

======== CHANNEL f1 ========
NUC1                    1H
P1                  13.57 usec
P2                  27.14 usec
P28               1000.00 usec
PL1                  0.00 dB
PL1W           16.62868881 W
SFO1          600.1037059 MHz

======== CHANNEL f2 ========
CPDPRG2               garp
NUC2                   13C
P3                   9.00 usec
P4                  18.00 usec
PCPD2               60.00 usec
PL2                  1.50 dB
PL12                17.98 dB
PL2W           59.57236099 W
PL12W           1.33981502 W
SFO2          150.9103545 MHz

====== GRADIENT CHANNEL =====
GPNAM1            SINE.100
GPNAM2            SINE.100
GPZ1                80.00 %
GPZ2                20.12 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                      2
TD                     512
SFO1            150.9104 MHz
FIDRES          70.426971 Hz
```

HSQC NMR Spectrum of Compound 6 (Chapter 5)




```
Current Data Parameters
NAME                  Raj 2
EXPNO                   121
PROCNO                    1

F2 - Acquisition Parameters
Date_              20101109
Time                 21.39
INSTRUM              spect
PROBHD    5 mm PABBO BB-
PULPROG     hmbcgplpndqf
TD                   4096
SOLVENT               DMSO
NS                      16
DS                      16
SWH              9615.385 Hz
FIDRES           2.347506 Hz
AQ               0.2130420 sec
RG                    2050
DW                 52.000 usec
DE                  6.50 usec
TE                 298.1 K
CNST2          145.0000000
CNST13          10.0000000
D0              0.00000300 sec
D1              3.00000000 sec
D2              0.00344828 sec
D6              0.05000000 sec
D16             0.00020000 sec
IN0             0.00001385 sec

======== CHANNEL f1 ========
NUC1                    1H
P1                  13.57 usec
P2                  27.14 usec
PL1                  0.00 dB
PL1W          16.62868881 W
SFO1          600.1036006 MHz

======== CHANNEL f2 ========
NUC2                   13C
P3                   9.00 usec
PL2                  1.50 dB
PL2W          59.57236099 W
SFO2          150.9103545 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPNAM2           SINE.100
GPNAM3           SINE.100
GPZ1                60.00 %
GPZ2                20.00 %
GPZ3                60.12 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                     2
TD                    256
SFO1            150.9104 MHz
FIDRES        140.850357 Hz
SW              238.935 ppm
FnMODE                 QF

F2 - Processing parameters
SI                    2048
```

HMBC NMR Spectrum of Compound 6 (Chapter 5)



IR Spectrum of Compound 16 (Chapter 4) and Compound 6 (Chapter 5)

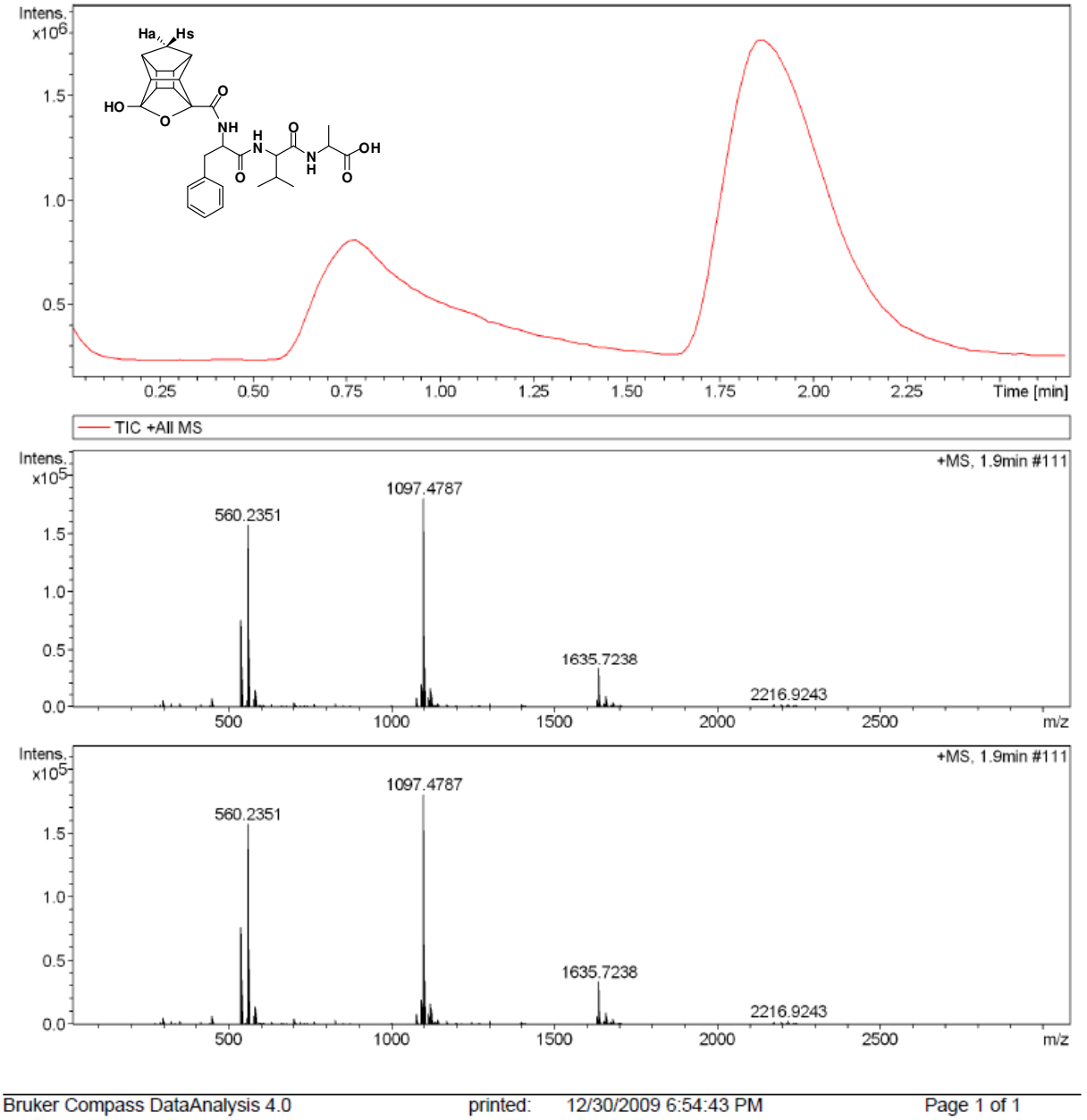
Display Report

Analysis Info

Analysis Name	D:\Data\Raj\OH-PCU-PHE-VAL-ALA-OH2.d	Acquisition Date	12/30/2009 6:49:25 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	OH-PCU-PHE-VAL-ALA-OH2	Instrument	micrOTOF-Q 10139
Comment	537.25		

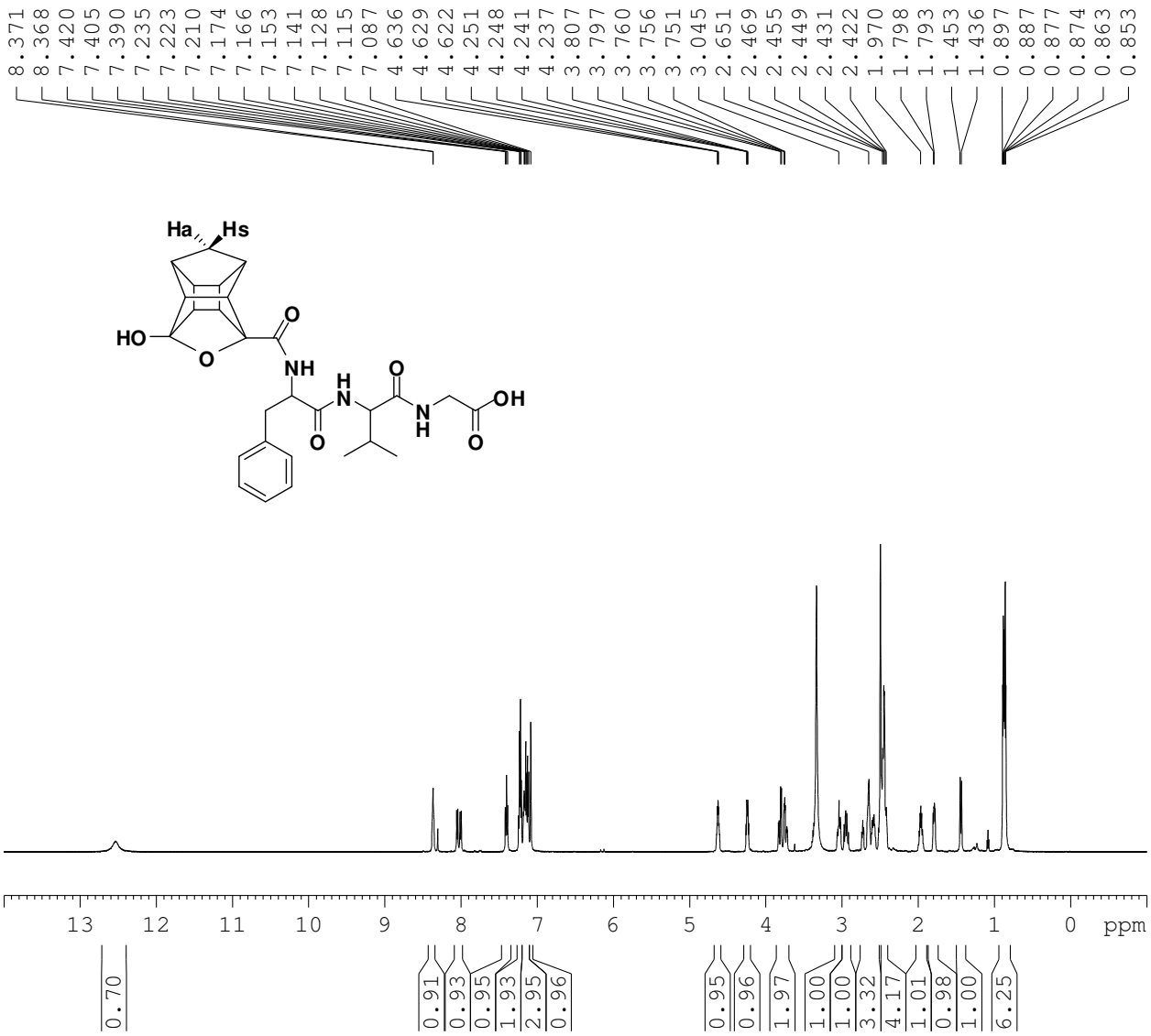
Acquisition Parameter

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



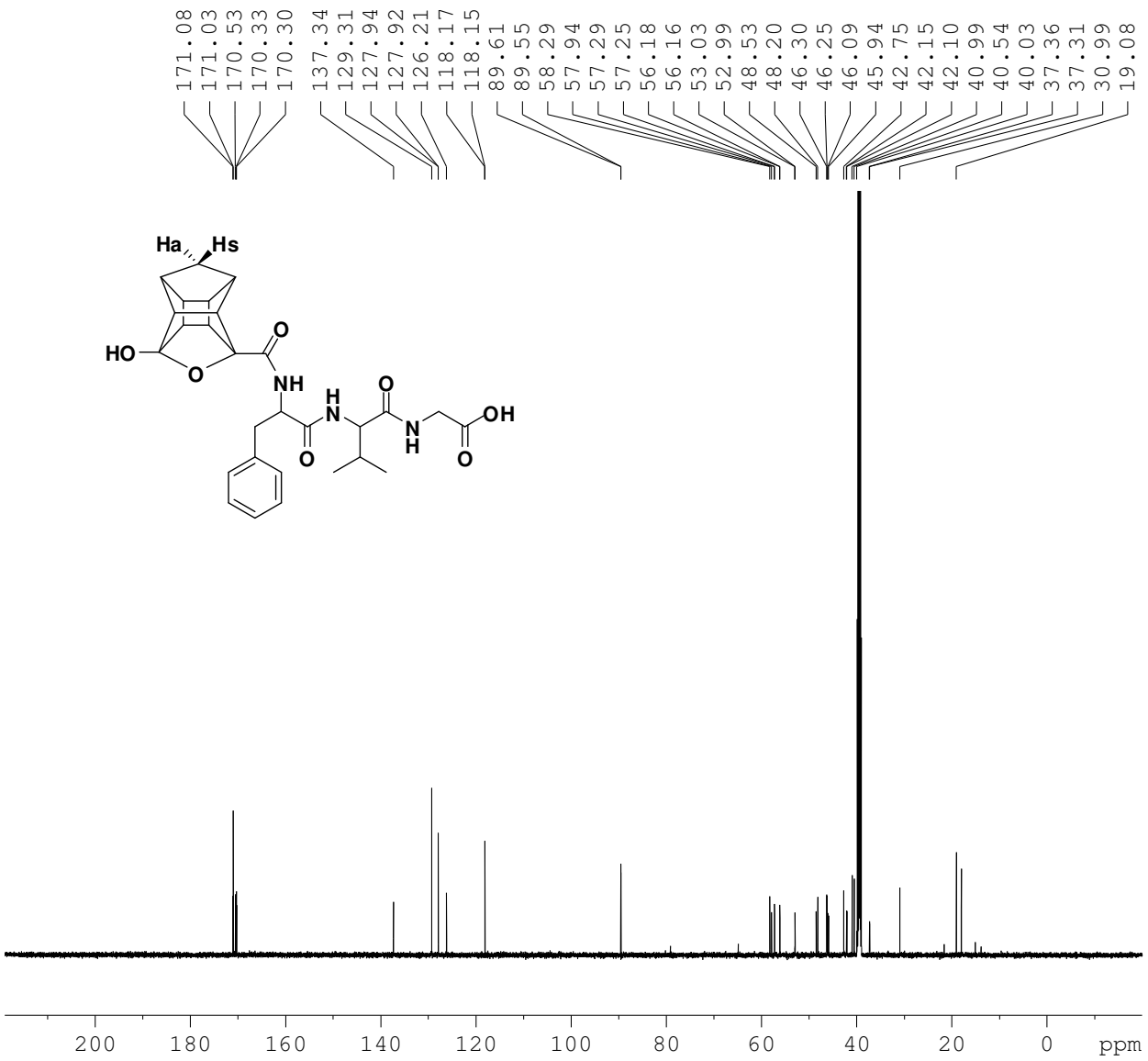
Bruker Compass DataAnalysis 4.0 printed: 12/30/2009 6:54:43 PM Page 1 of 1

HRESIMS (*m/z*) of Compound 16 (Chapter 4) and Compound 6 (Chapter 5)



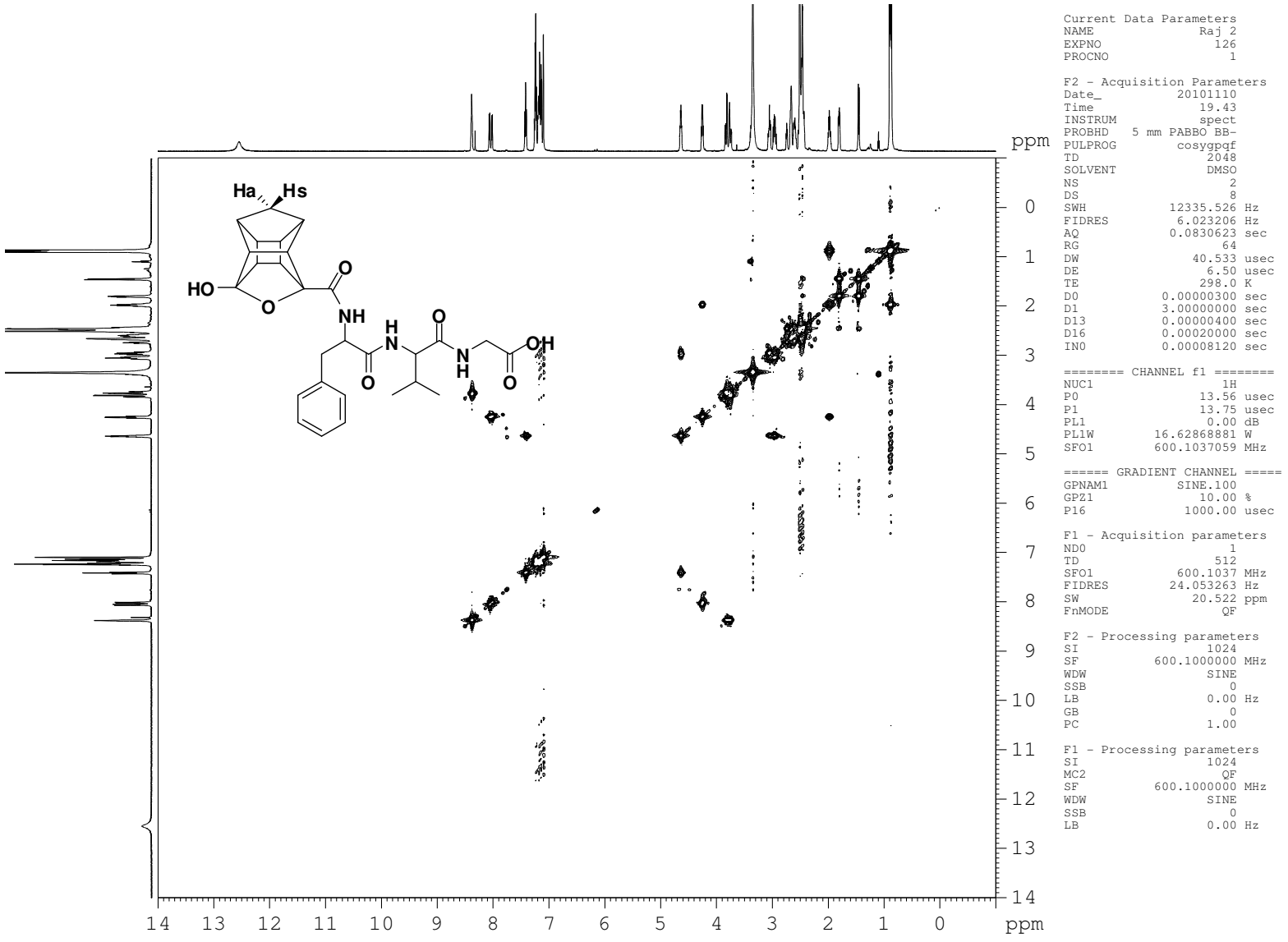

```
Current Data Parameters
NAME                 Raj 2
EXPNO                  123
PROCNO                   1

F2 - Acquisition Parameters
Date_             20101110
Time                 14.38
INSTRUM              spect
PROBHD    5 mm PABBO BB-
PULPROG              zg30
TD                  32768
SOLVENT              DMSO
NS                     16
DS                      2
SWH              12335.526 Hz
FIDRES           0.376450 Hz
AQ               1.3282462 sec
RG                    80.6
DW                40.533 usec
DE                 6.50 usec
TE                 298.3 K
D1              1.00000000 sec
TD0                      1

======== CHANNEL f1 ========
NUC1                   1H
P1                  13.75 usec
PL1                  0.00 dB
PL1W          16.62868881 W
SFO1          600.1037059 MHz

F2 - Processing parameters
SI                   16384
SF            600.1000053 MHz
WDW                    no
SSB                     0
LB                   0.00 Hz
GB                      0
PC                   1.00
```

^{1}H NMR Spectrum of Compound 17 (Chapter 4) and Compound 5 (Chapter 5)



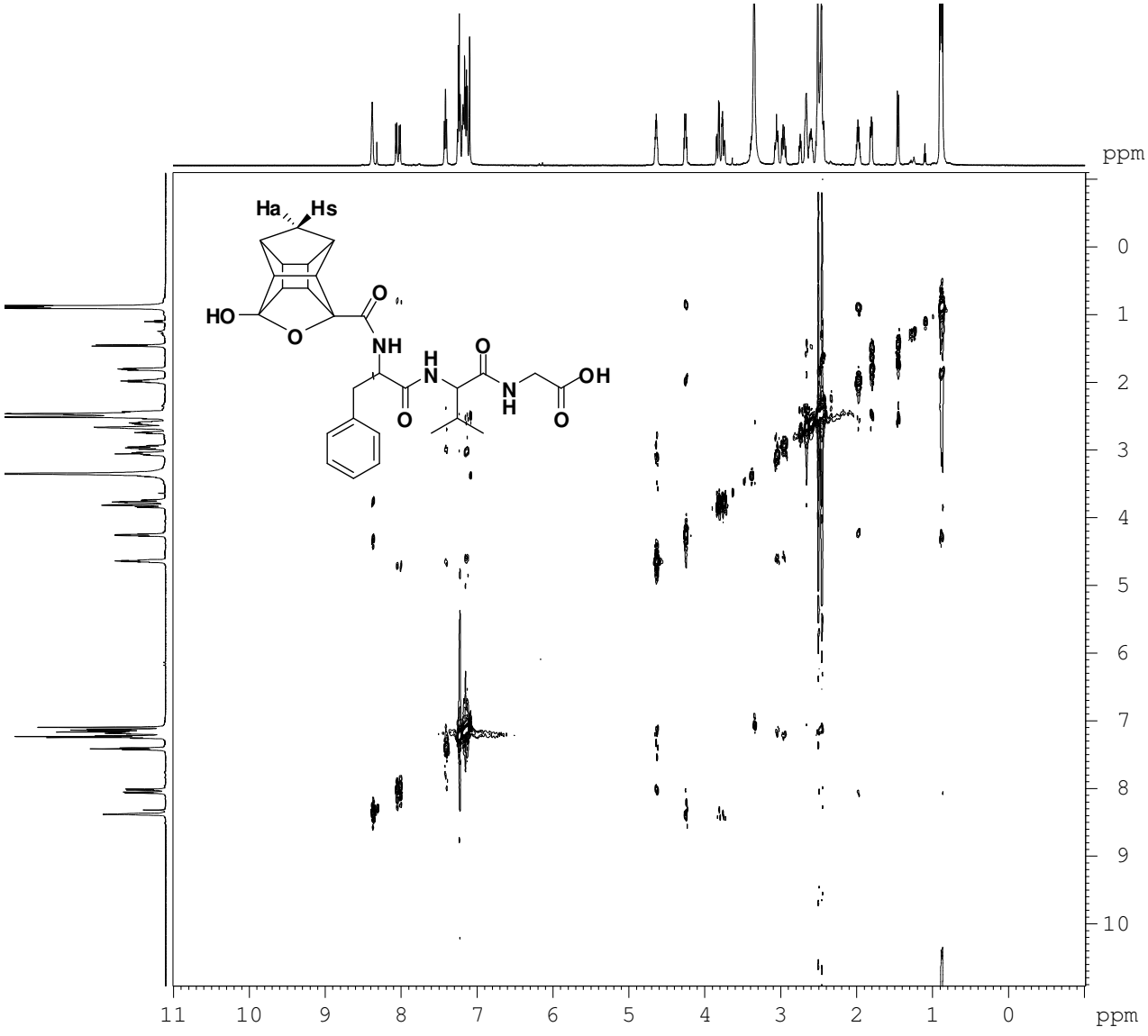

```
Current Data Parameters
NAME                 Raj 2
EXPNO                  129
PROCNO                   1

F2 - Acquisition Parameters
Date_             20101111
Time                 5.35
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG           zgpg30
TD                 65536
SOLVENT              DMSO
NS                    4800
DS                       4
SWH              36057.691 Hz
FIDRES           0.550197 Hz
AQ               0.9088159 sec
RG                    2050
DW                 13.867 usec
DE                  6.50 usec
TE                  298.2 K
D1             2.00000000 sec
D11            0.03000000 sec
TD0                      1

======== CHANNEL f1 ========
NUC1                  13C
P1                   9.00 usec
PL1                  1.50 dB
PL1W          59.57236099 W
SFO1         150.9103545 MHz

======== CHANNEL f2 ========
CPDPRG2           waltz16
NUC2                  1H
PCPD2               96.00 usec
PL2                 -0.50 dB
PL12                15.40 dB
PL13                18.00 dB
PL2W          18.65769577 W
PL12W          0.47957665 W
PL13W          0.26354694 W
SFO2         600.1024004 MHz

F2 - Processing parameters
SI                   32768
SF           150.8953385 MHz
WDW                     no
SSB                      0
LB                    0.00 Hz
GB                       0
PC                    1.40
```

^{13}C NMR Spectrum of Compound 17 (Chapter 4) and Compound 5 (Chapter 5)



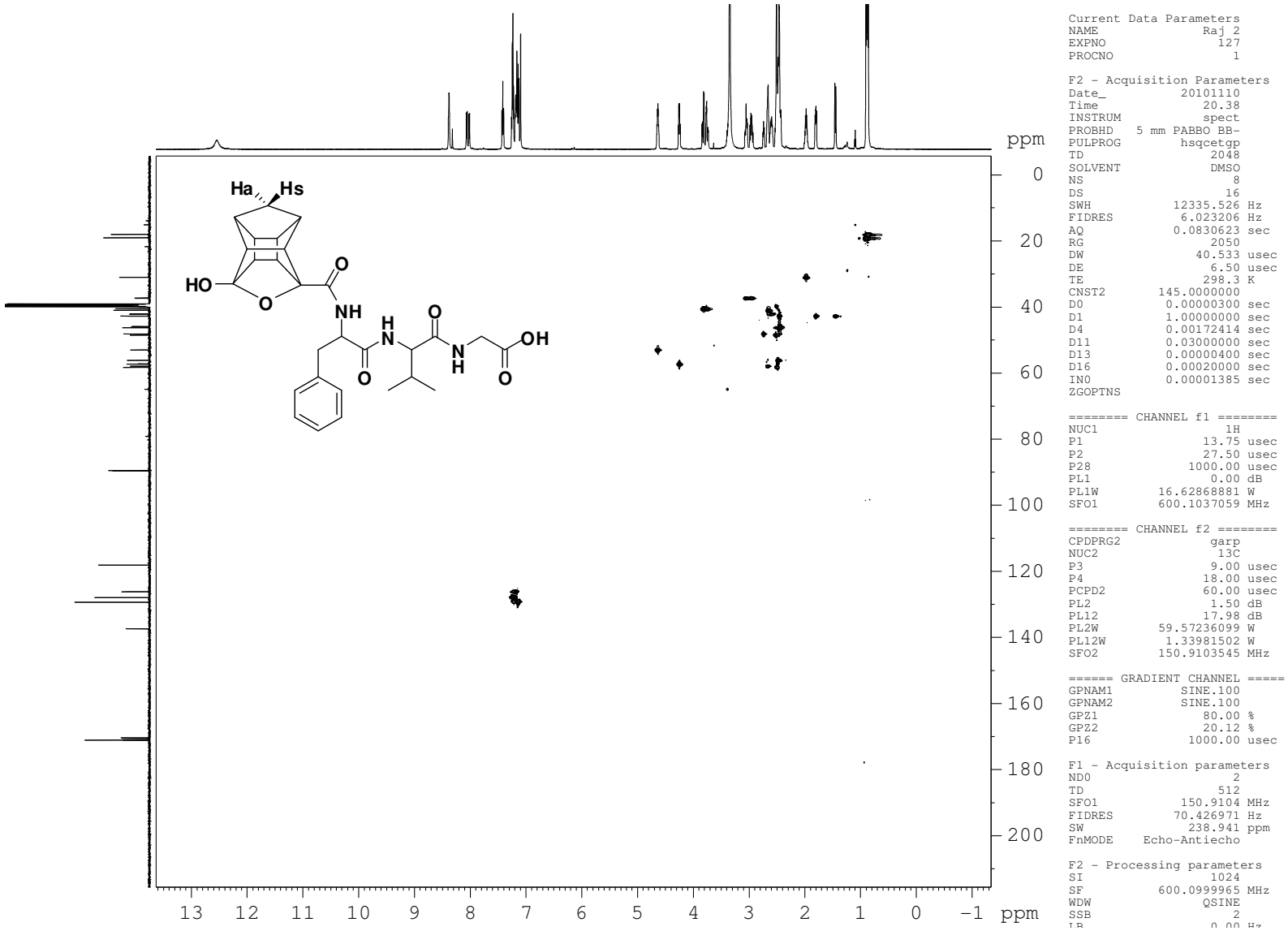

```
Current Data Parameters
NAME                 Raj 2
EXPNO                  126
PROCNO                   1

F2 - Acquisition Parameters
Date_             20101110
Time                 19.43
INSTRUM              spect
PROBHD    5 mm PABBO BB-
PULPROG          cosygpqf
TD                   2048
SOLVENT               DMSO
NS                      2
DS                      8
SWH              12335.526 Hz
FIDRES               6.023206 Hz
AQ                0.0830623 sec
RG                      64
DW                   40.533 usec
DE                    6.50 usec
TE                   298.0 K
D0              0.00000300 sec
D1              3.00000000 sec
D13             0.00000400 sec
D16             0.00020000 sec
IN0             0.00008120 sec

======== CHANNEL f1 ========
NUC1                    1H
P0                   13.56 usec
P1                   13.75 usec
PL1                   0.00 dB
PL1W          16.62868881 W
SFO1          600.1037059 MHz

====== GRADIENT CHANNEL =====
GPNAM1            SINE.100
GPZ1                 10.00 %
P16                1000.00 usec

F1 - Acquisition parameters
ND0                     1
TD                    512
SFO1             600.1037 MHz
FIDRES           24.053263 Hz
SW                  20.522 ppm
FnMODE                  QF

F2 - Processing parameters
SI                    1024
SF           600.1000000 MHz
WDW                   SINE
SSB                      0
LB                    0.00 Hz
GB                       0
PC                    1.00

F1 - Processing parameters
SI                    1024
MC2                     QF
SF           600.1000000 MHz
WDW                   SINE
SSB                      0
LB                    0.00 Hz
```

COSY NMR Spectrum of Compound 5 (Chapter 5)



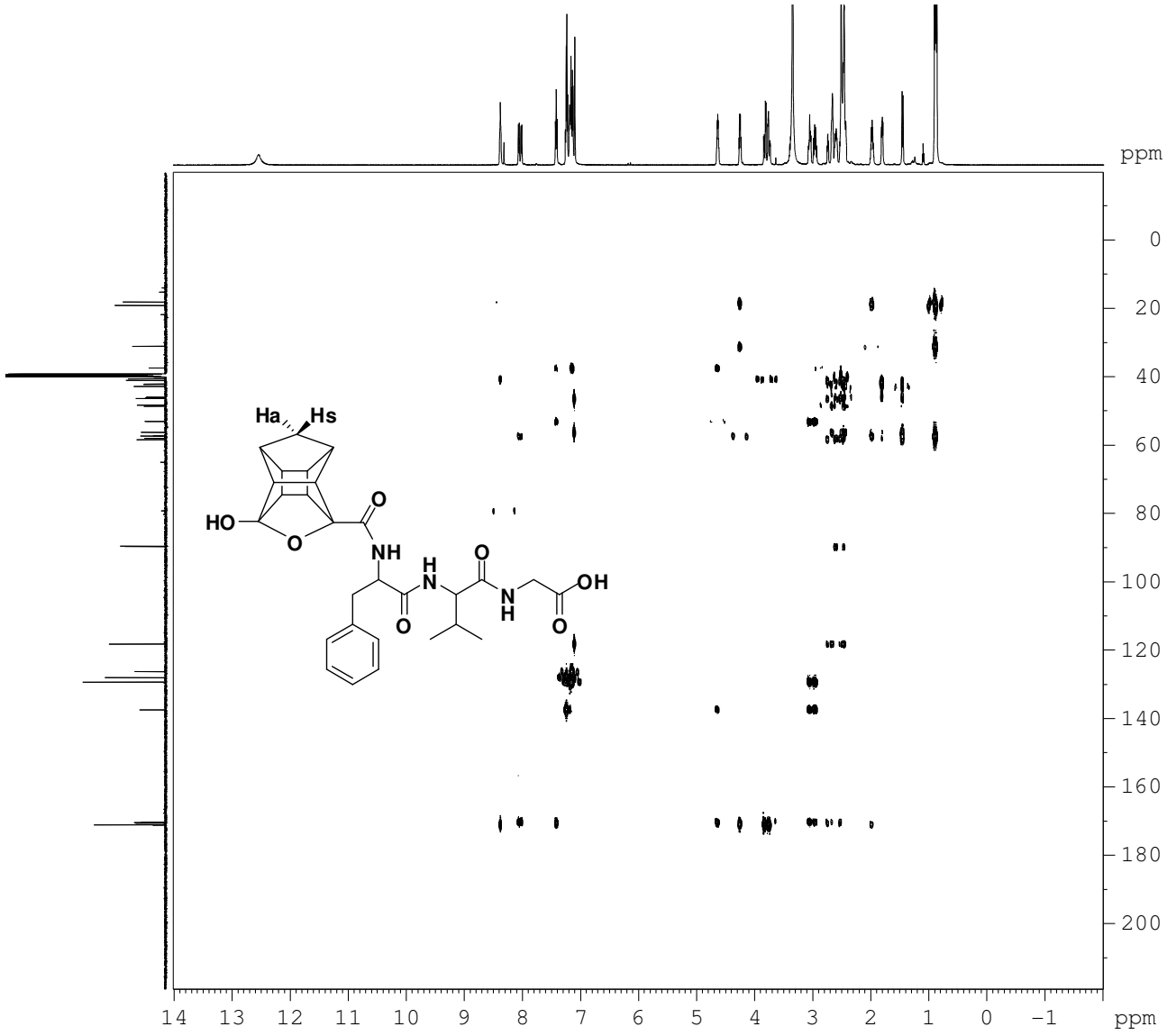

```
Current Data Parameters
NAME                 Raj 2
EXPNO                  125
PROCNO                   1

F2 - Acquisition Parameters
Date_             20101110
Time                 17.13
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG            kproesy
TD                    2048
SOLVENT               DMSO
NS                       8
DS                      64
SWH                7211.539 Hz
FIDRES             3.521259 Hz
AQ                 0.1420447 sec
RG                      22.6
DW                   69.333 usec
DE                     6.00 usec
TE                    298.2 K
CNST2          6849.0000000
D0                0.00005183 sec
D1                1.50000000 sec
D11               0.03000000 sec
D12               0.00002000 sec
D13               0.00000400 sec
D16               0.00030000 sec
IN0               0.00013865 sec

======== CHANNEL f1 ========
CNST1             0.0000000
CNST3          3954.3879395
CNST4         -3954.3879395
NUC1                    1H
P1                     13.75 usec
P7                250000.00 usec
P11                 1000.00 usec
PL1                    0.00 dB
PL9                   70.00 dB
PL1W           16.62868881 W
PL9W            0.00000166 W
SFO1           600.1030005 MHz
SP1                    13.00 dB
SP2                    13.00 dB
SP3                    13.00 dB
SP4                    13.00 dB
SP5                    13.00 dB
SP6                    13.00 dB
SPNAM1    gaussramp-down
SPNAM2         square.100
SPNAM3      gaussramp-up
SPNAM4    gaussramp+down
SPNAM5         square.100
SPNAM6      gaussramp+up
SPOAL1                1.000
SPOAL2                0.500
SPOAL3                0.000
SPOAL4                1.000
SPOAL5                0.500
SPOAL6                0.000
SPOFFS1                0.00 Hz
SPOFFS2                0.00 Hz
SPOFFS3                0.00 Hz
SPOFFS4                0.00 Hz
SPOFFS5                0.00 Hz
SPOFFS6                0.00 Hz

====== GRADIENT CHANNEL =====
GPNAM1              RECT.1
GPNAM3              RECT.1
GPZ1                27.00 %
GPZ3                46.00 %
P16                500.00 usec

F1 - Acquisition parameters
ND0                      1
TD                     512
```

ROESY NMR Spectrum of Compound 5 (Chapter 5)



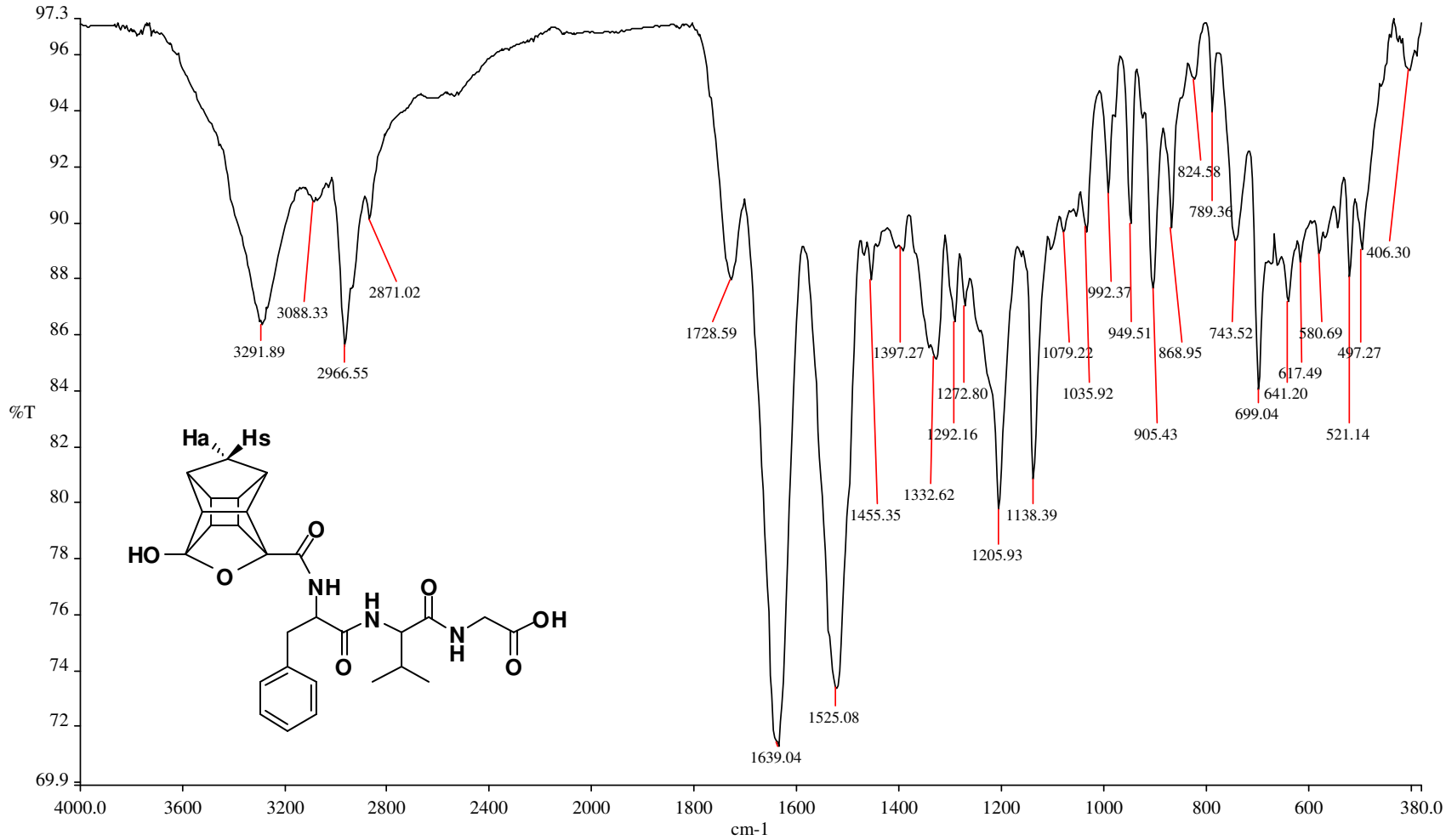

```
Current Data Parameters
NAME                  Raj 2
EXPNO                   127
PROCNO                    1

F2 - Acquisition Parameters
Date_              20101110
Time                  20.38
INSTRUM               spect
PROBHD    5 mm PABBO BB-
PULPROG          hsqcetgp
TD                    2048
SOLVENT               DMSO
NS                       8
DS                      16
SWH              12335.526 Hz
FIDRES            6.023206 Hz
AQ               0.0830623 sec
RG                    2050
DW                 40.533 usec
DE                  6.50 usec
TE                 298.3 K
CNST2           145.0000000
D0               0.00000300 sec
D1               1.00000000 sec
D4               0.00172414 sec
D11              0.03000000 sec
D13              0.00000400 sec
D16              0.00020000 sec
IN0              0.00001385 sec
ZGOPTNS

======== CHANNEL f1 ========
NUC1                    1H
P1                   13.75 usec
P2                   27.50 usec
P28                1000.00 usec
PL1                   0.00 dB
PL1W           16.62868881 W
SFO1           600.1037059 MHz

======== CHANNEL f2 ========
CPDPRG2               garp
NUC2                   13C
P3                    9.00 usec
P4                   18.00 usec
PCPD2                60.00 usec
PL2                   1.50 dB
PL12                 17.98 dB
PL2W           59.57236099 W
PL12W           1.33981502 W
SFO2           150.9103545 MHz

====== GRADIENT CHANNEL =====
GPNAM1             SINE.100
GPNAM2             SINE.100
GPZ1                 80.00 %
GPZ2                 20.12 %
P16                1000.00 usec

F1 - Acquisition parameters
ND0                      2
TD                     512
SFO1           150.9104 MHz
FIDRES          70.426971 Hz
SW                238.941 ppm
FnMODE     Echo-Antiecho

F2 - Processing parameters
SI                    1024
SF             600.0999965 MHz
WDW                  QSINE
SSB                      2
LB                    0.00 Hz
```

HSQC NMR Spectrum of Compound 5 (Chapter 5)




```
Current Data Parameters
NAME                  Raj 2
EXPNO                   182
PROCNO                    1

F2 - Acquisition Parameters
Date_              20110121
Time                  12.35
INSTRUM               spect
PROBHD   5 mm PABBO BB-
PULPROG     hmbcgplpndqf
TD                      4096
SOLVENT                 DMSO
NS                        16
DS                        16
SWH                9615.385 Hz
FIDRES             2.347506 Hz
AQ                0.2130420 sec
RG                     2050
DW                   52.000 usec
DE                    6.50 usec
TE                   298.2 K
CNST2           145.0000000
CNST13           10.0000000
D0                0.00000300 sec
D1                3.00000000 sec
D2                0.00344828 sec
D6                0.05000000 sec
D16               0.00020000 sec
IN0               0.00001385 sec

======== CHANNEL f1 ========
NUC1                      1H
P1                     13.74 usec
P2                     27.48 usec
PL1                     0.00 dB
PL1W           16.62868881 W
SFO1           600.1036006 MHz

======== CHANNEL f2 ========
NUC2                     13C
P3                      9.00 usec
PL2                     1.50 dB
PL2W           59.57236099 W
SFO2           150.9103545 MHz

====== GRADIENT CHANNEL =====
GPNAM1               SINE.100
GPNAM2               SINE.100
GPNAM3               SINE.100
GPZ1                    60.00 %
GPZ2                    20.00 %
GPZ3                    60.12 %
P16                   1000.00 usec

F1 - Acquisition parameters
ND0                       2
TD                      181
SFO1              150.9104 MHz
FIDRES          199.213760 Hz
SW                 238.935 ppm
FnMODE                  QF

F2 - Processing parameters
SI                     2048
SF             600.0999980 MHz
WDW                    SINE
SSB                       0
LB                     0.00 Hz
GB                        0
PC                     1.00

F1 - Processing parameters
```

HMBC NMR Spectrum of Compound 5 (Chapter 5)



IR Spectrum of Compound 17 (Chapter 4) and Compound 5 (Chapter 5)

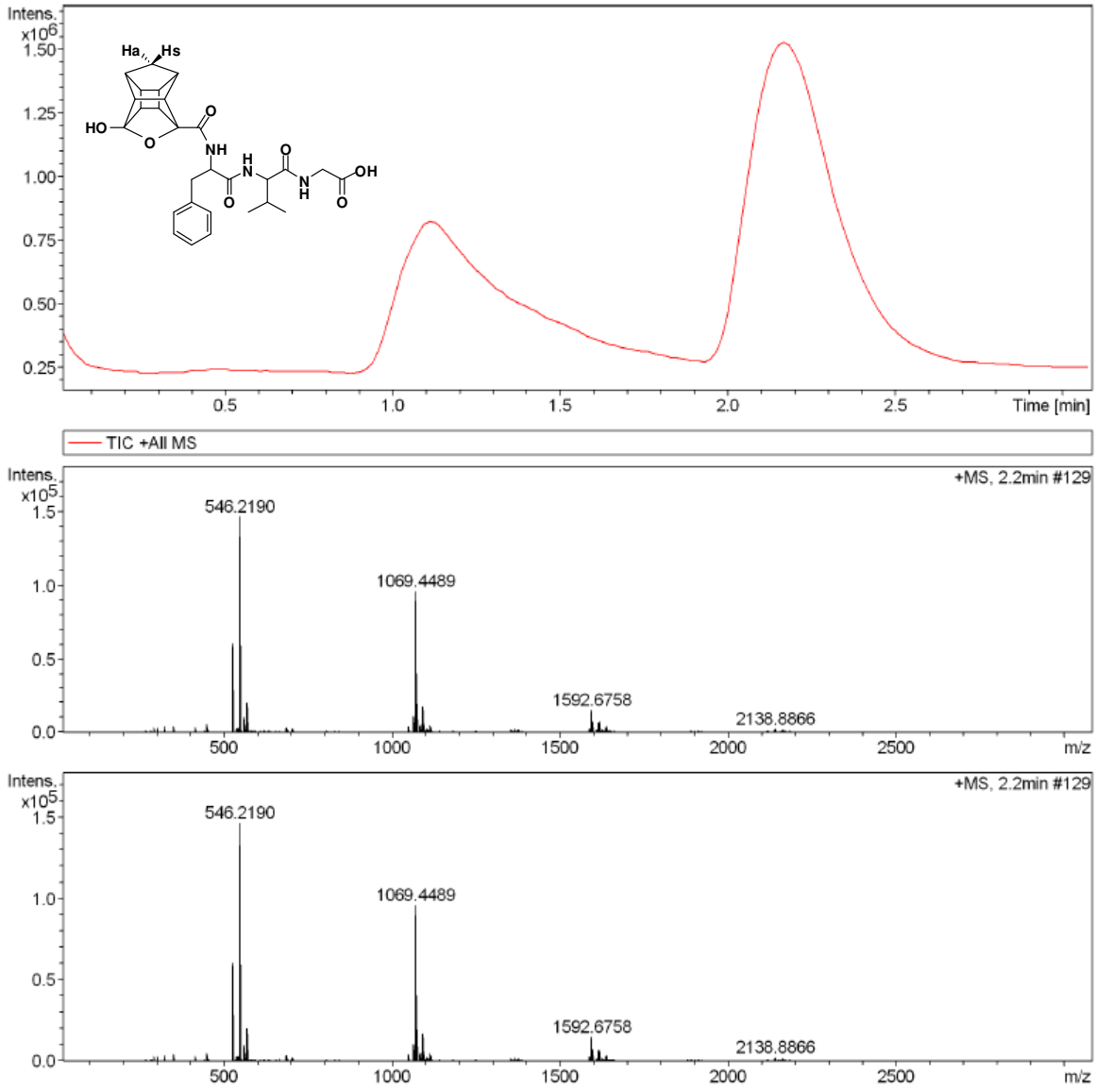
Display Report

Analysis Info

Analysis Name	D:\Data\Raj\OH-PCU-PHE-VAL-GLY-OH1.d	Acquisition Date	12/30/2009 6:23:13 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	OH-PCU-PHE-VAL-GLY-OH1	Instrument	micrOTOF-Q 10139
Comment	523.23		

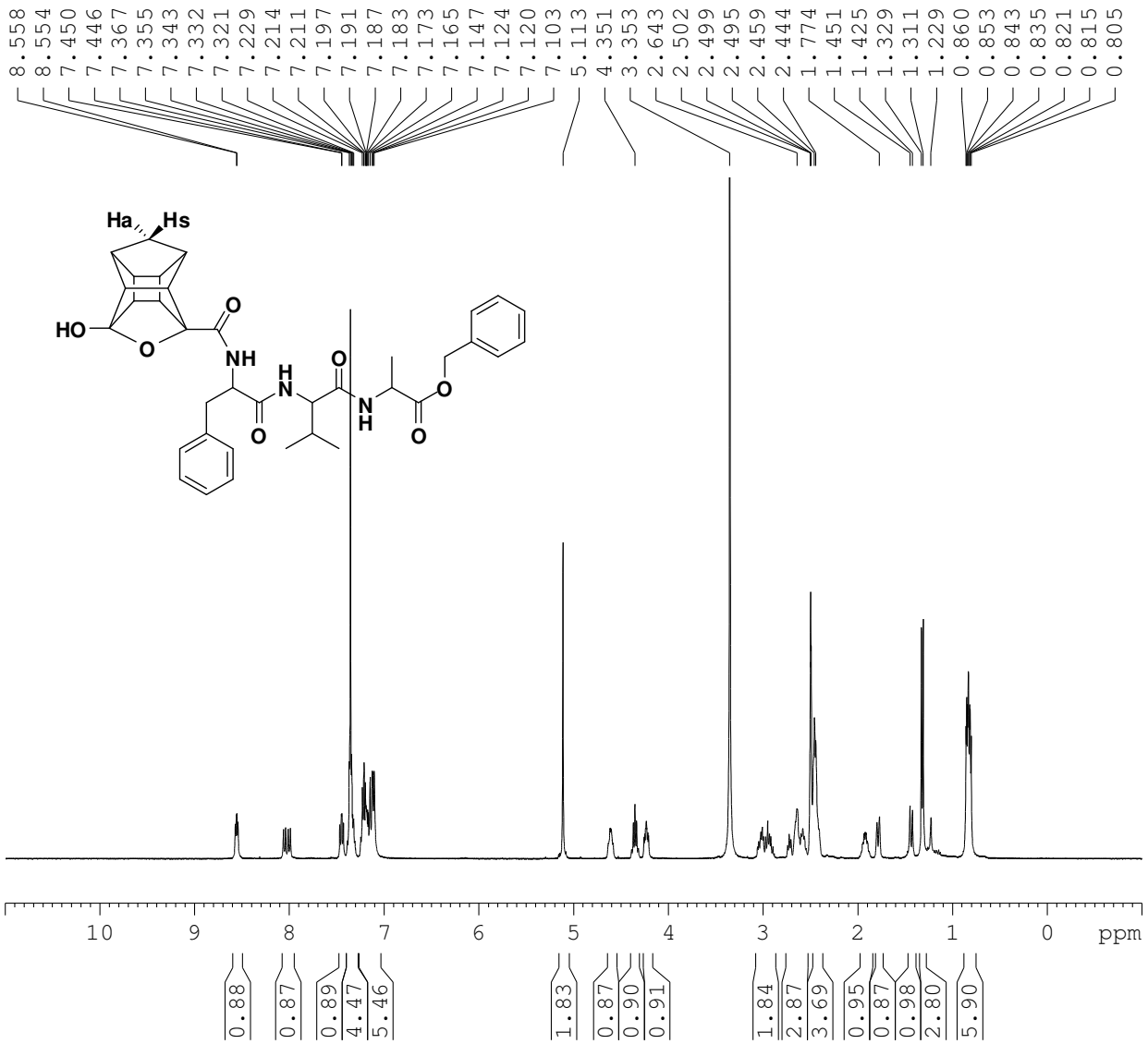
Acquisition Parameter

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



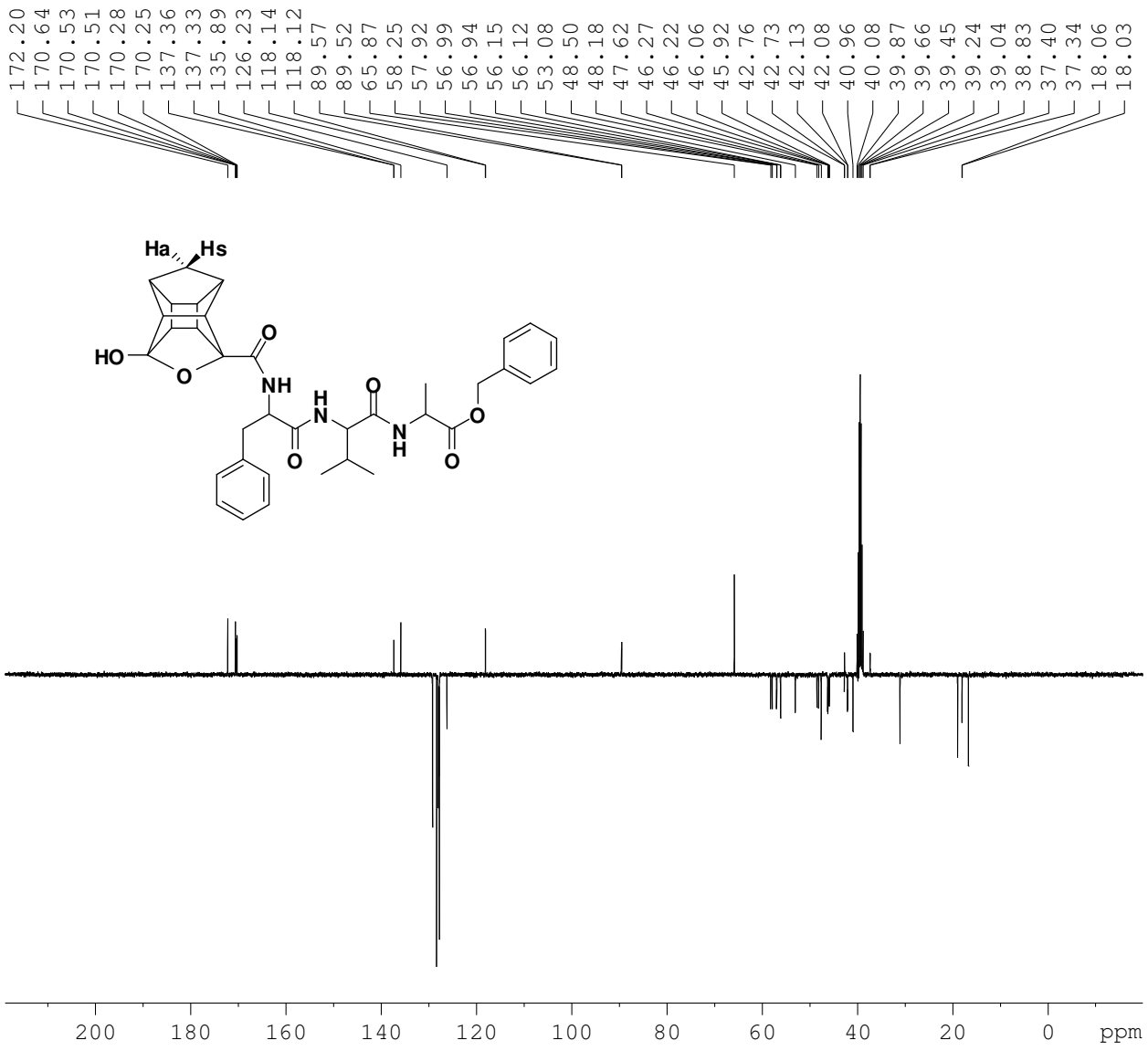
Bruker Compass DataAnalysis 4.0 printed: 12/30/2009 6:29:42 PM Page 1 of 1

HRESIMS (*m/z*) of Compound 17 (Chapter 4) and Compound 5 (Chapter 5)



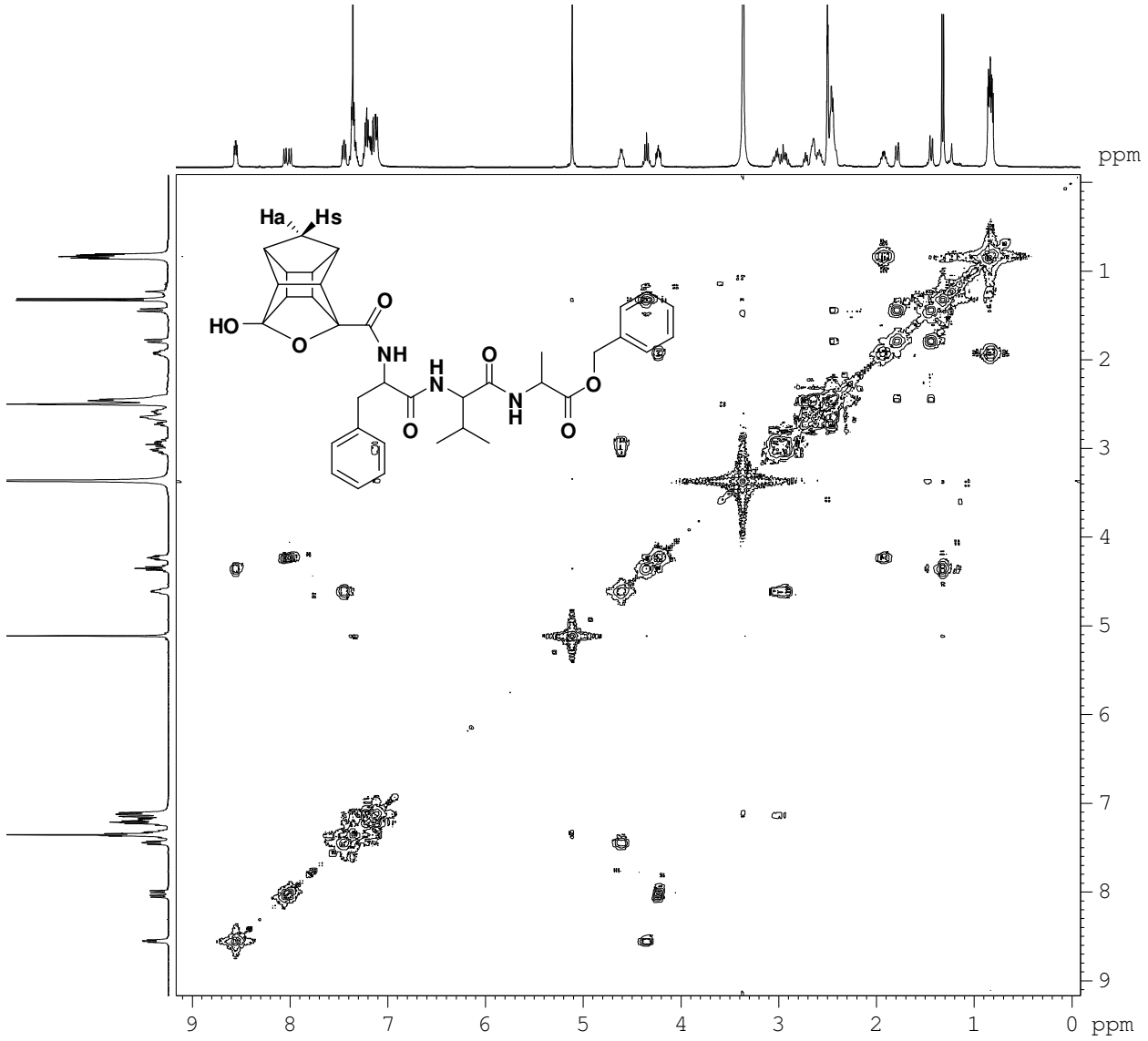

```
Current Data Parameters
NAME      Oct09-2009-GK-RAJ
EXPNO                   20
PROCNO                   1

F2 - Acquisition Parameters
Date_           20091009
Time               19.54
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG             zg30
TD                 32768
SOLVENT             DMSO
NS                    16
DS                     2
SWH             8223.685 Hz
FIDRES          0.250967 Hz
AQ             1.9923444 sec
RG                   101
DW                60.800 usec
DE                 6.50 usec
TE                295.8 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                  1H
P1                10.00 usec
PL1               -3.00 dB
PL1W        15.48668575 W
SFO1        400.2224715 MHz

F2 - Processing parameters
SI                 16384
SF          400.2199996 MHz
WDW                   EM
SSB                    0
LB                  0.30 Hz
GB                     0
PC                  1.00
```

^{1}H NMR Spectrum of Compound 18 (Chapter 4) and Compound 8 (Chapter 5)



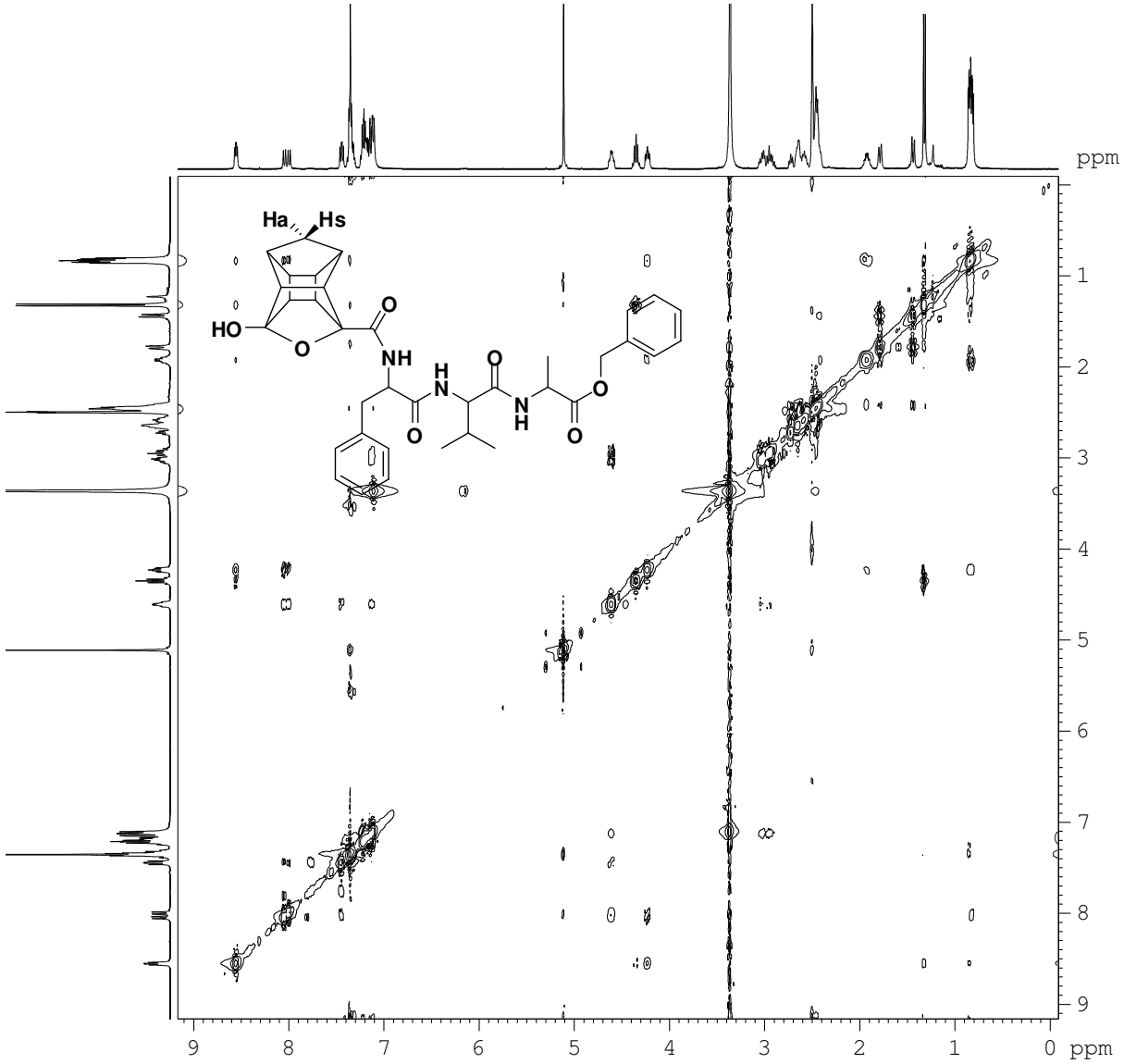

```
Current Data Parameters
NAME      Oct09-2009-GK-RAJ
EXPNO                   22
PROCNO                   1

F2 - Acquisition Parameters
Date_             20091011
Time                 18.38
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG              jmod
TD                  65536
SOLVENT              DMSO
NS                   1024
DS                      4
SWH            24038.461 Hz
FIDRES          0.366798 Hz
AQ             1.3631988 sec
RG                   2050
DW                20.800 usec
DE                 6.50 usec
TE                297.4 K
CNST2         145.0000000
CNST11          1.0000000
D1            2.00000000 sec
D20           0.00689655 sec
TD0                     1

======== CHANNEL f1 ========
NUC1                  13C
P1                   8.40 usec
P2                  16.80 usec
PL1                 -2.00 dB
PL1W         54.14257431 W
SFO1        100.6454626 MHz

======== CHANNEL f2 ========
CPDPRG2           waltz16
NUC2                  1H
PCPD2               90.00 usec
PL2                 -3.00 dB
PL12                15.60 dB
PL2W         15.48668575 W
PL12W         0.21377575 W
SFO2        400.2216009 MHz

F2 - Processing parameters
SI                   32768
SF          100.6354435 MHz
WDW                     EM
SSB                      0
LB                    1.00 Hz
GB                       0
PC                    1.40
```

^{13}C NMR Spectrum of Compound 18 (Chapter 4) and Compound 8 (Chapter 5)



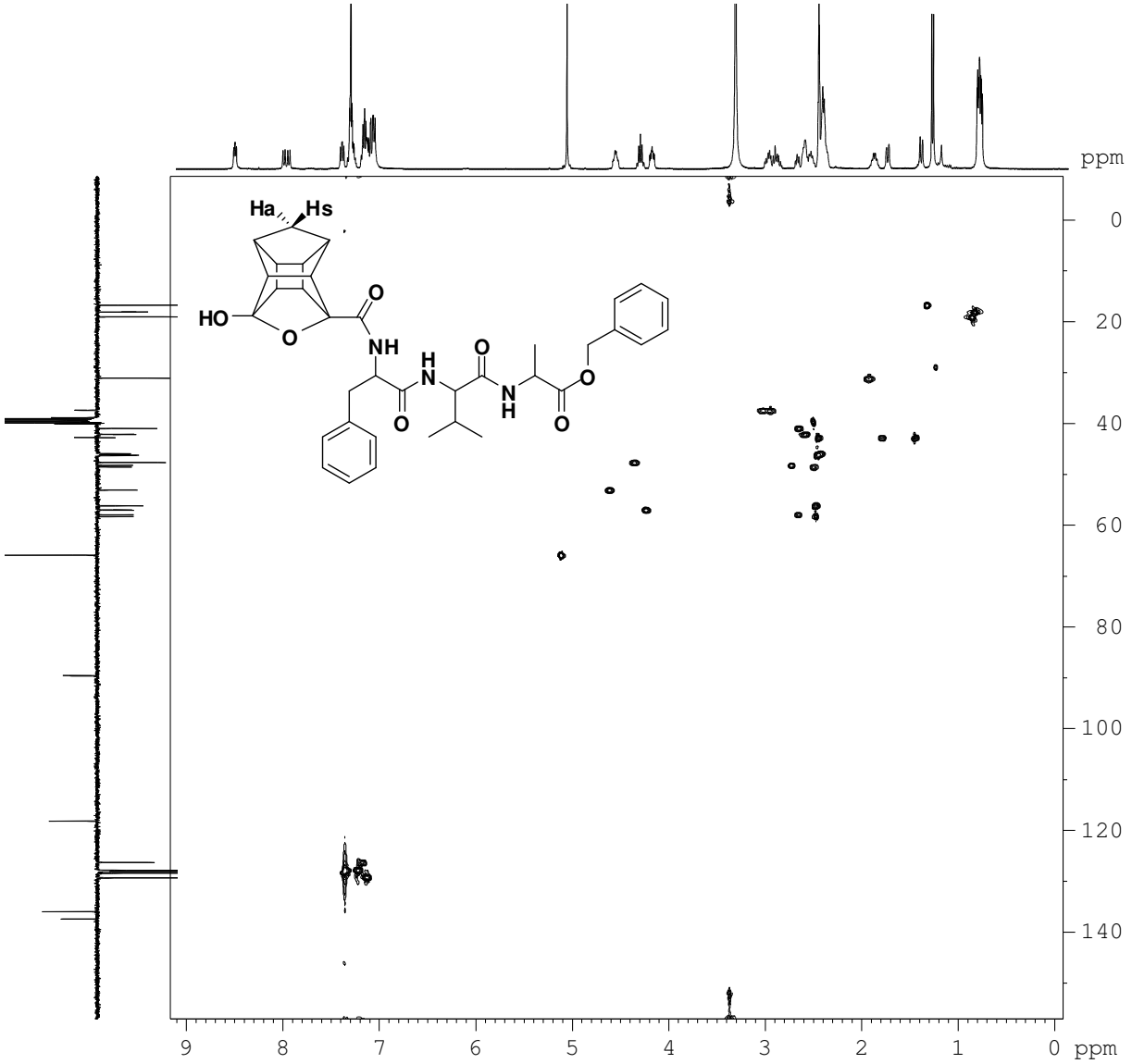

```
Current Data Parameters
NAME      Oct09-2009-GK-RAJ
EXPNO                   26
PROCNO                   1

F2 - Acquisition Parameters
Date_           20091011
Time              20.42
INSTRUM           spect
PROBHD   5 mm PABBO BB-
PULPROG        cosygpqf
TD                 2048
SOLVENT             DMSO
NS                    2
DS                    8
SWH            3703.704 Hz
FIDRES         1.808449 Hz
AQ           0.2765300 sec
RG                   64
DW             135.000 usec
DE               6.50 usec
TE              296.4 K
D0           0.00000300 sec
D1           1.40210497 sec
D13          0.00000400 sec
D16          0.00020000 sec
IN0          0.00027000 sec

======== CHANNEL f1 ========
NUC1                 1H
P0               10.00 usec
P1               10.00 usec
PL1              -3.00 dB
PL1W        15.48668575 W
SFO1       400.2218145 MHz

====== GRADIENT CHANNEL =====
GPNAM1         SINE.100
GPZ1             10.00 %
P16            1000.00 usec

F1 - Acquisition parameters
ND0                  1
TD                 128
SFO1          400.2218 MHz
FIDRES       28.935184 Hz
SW            9.254 ppm
FnMODE              QF

F2 - Processing parameters
SI                1024
SF          400.2199980 MHz
WDW               SINE
SSB                  0
LB                0.00 Hz
GB                   0
PC                1.40

F1 - Processing parameters
SI                1024
MC2                 QF
SF          400.2199980 MHz
WDW               SINE
SSB                  0
LB                0.00 Hz
GB                   0
```

COSY NMR Spectrum of Compound 8 (Chapter 5)



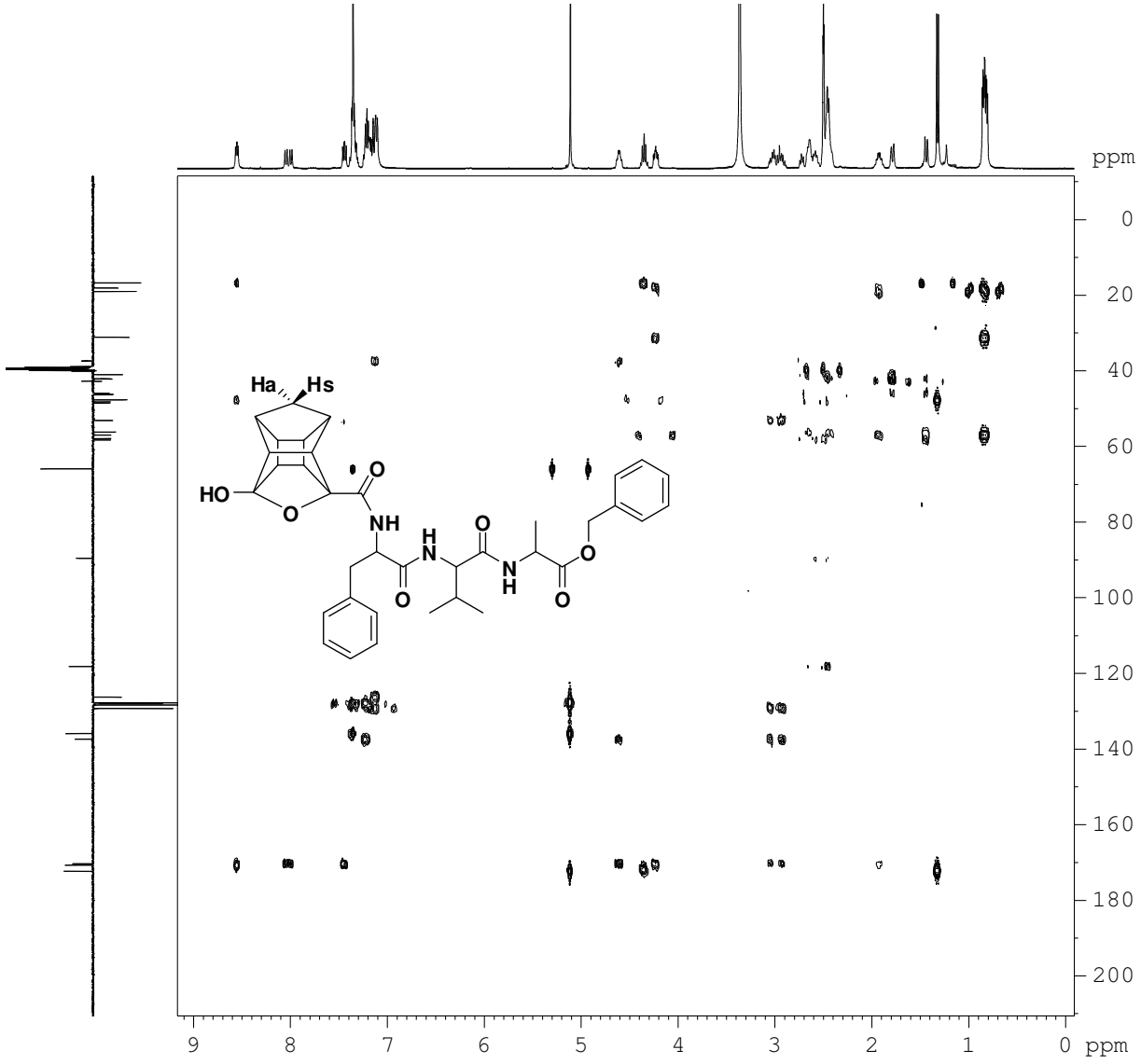

```
Current Data Parameters
NAME      Oct09-2009-GK-RAJ
EXPNO                   27
PROCNO                   1

F2 - Acquisition Parameters
Date_              20091011
Time                 20.50
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG           noesyph
TD                    2048
SOLVENT               DMSO
NS                       8
DS                       4
SWH               3703.704 Hz
FIDRES            1.808449 Hz
AQ               0.2765300 sec
RG                      64
DW                 135.000 usec
DE                   6.50 usec
TE                  296.4 K
D0              0.00012227 sec
D1              1.97419500 sec
D8              0.30000001 sec
IN0             0.00027000 sec

======== CHANNEL f1 ========
NUC1                    1H
P1                   10.00 usec
PL1                  -3.00 dB
PL1W          15.48668575 W
SFO1          400.2218145 MHz

F1 - Acquisition parameters
ND0                      1
TD                     256
SFO1             400.2218 MHz
FIDRES          14.467592 Hz
SW                   9.254 ppm
FnMODE         States-TPPI

F2 - Processing parameters
SI                    1024
SF            400.2199980 MHz
WDW                  QSINE
SSB                      2
LB                    0.00 Hz
GB                       0
PC                    1.00

F1 - Processing parameters
SI                    1024
MC2           States-TPPI
SF            400.2200000 MHz
WDW                  QSINE
SSB                      2
LB                    0.00 Hz
GB                       0
```

NOESY NMR Spectrum of Compound 8 (Chapter 5)



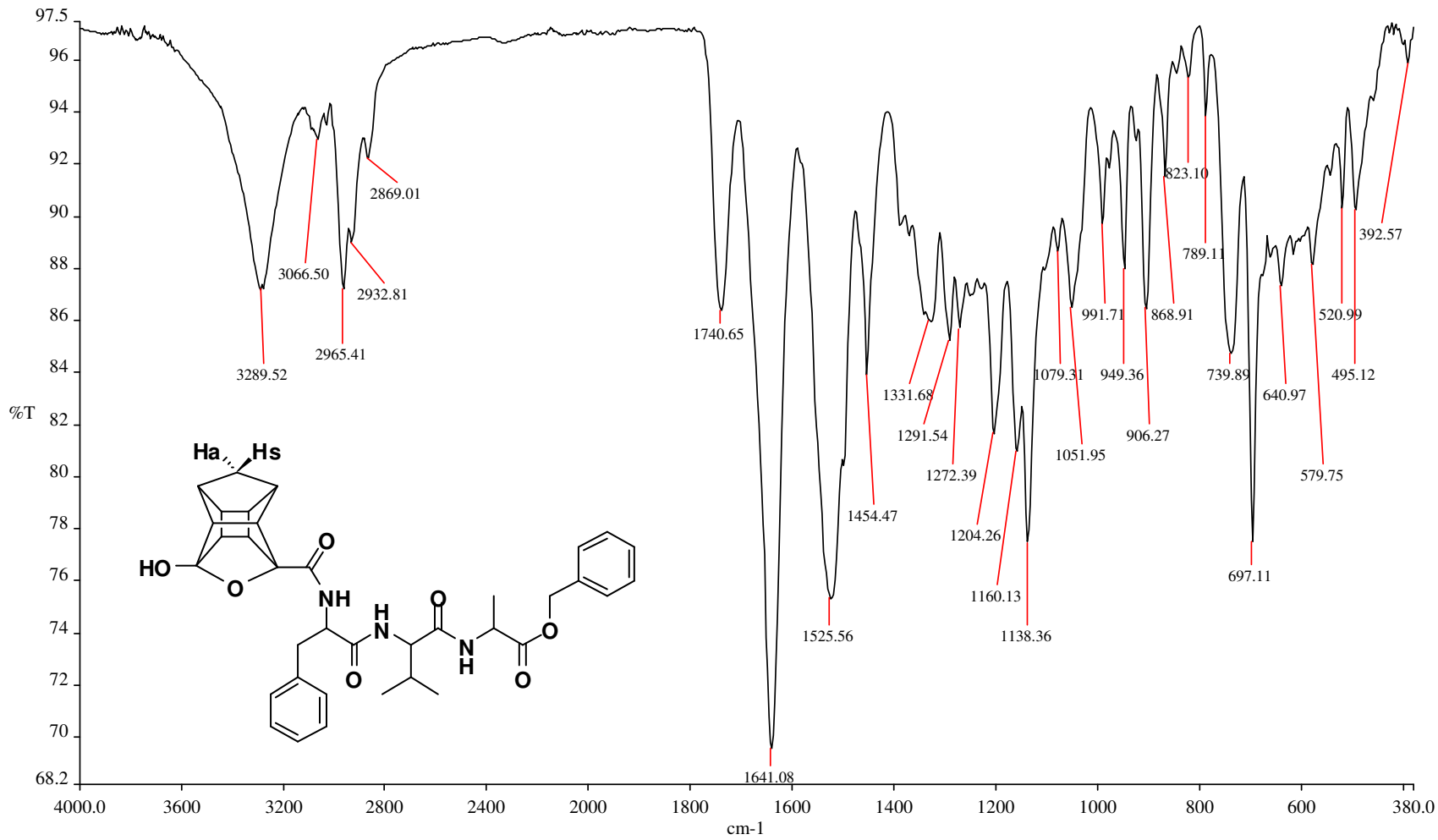

```
Current Data Parameters
NAME      Oct09-2009-GK-RAJ
EXPNO                   28
PROCNO                   1

F2 - Acquisition Parameters
Date_             20091011
Time                 22.20
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG       hsqcedetgp
TD                    1024
SOLVENT               DMSO
NS                       4
DS                      16
SWH              3703.704 Hz
FIDRES           3.616898 Hz
AQ               0.1382900 sec
RG                    2050
DW                 135.000 usec
DE                   6.50 usec
TE                  296.4 K
CNST2          145.0000000
D0               0.00000300 sec
D1               1.45760596 sec
D4               0.00172414 sec
D11              0.03000000 sec
D13              0.00000400 sec
D16              0.00020000 sec
D21              0.00345000 sec
IN0              0.00003000 sec
ZGOPTNS

======== CHANNEL f1 ========
NUC1                    1H
P1                   10.00 usec
P2                   20.00 usec
P28                1000.00 usec
PL1                  -3.00 dB
PL1W           15.48668575 W
SFO1          400.2218145 MHz

======== CHANNEL f2 ========
CPDPRG2               garp
NUC2                   13C
P3                    9.00 usec
P4                   18.00 usec
PCPD2                70.00 usec
PL2                  -2.00 dB
PL12                 15.60 dB
PL2W           54.14257431 W
PL12W           0.94089001 W
SFO2          100.6429049 MHz

====== GRADIENT CHANNEL =====
GPNAM1            SINE.100
GPNAM2            SINE.100
GPZ1                 80.00 %
GPZ2                 20.10 %
P16                1000.00 usec

F1 - Acquisition parameters
ND0                      2
TD                     256
SFO1            100.6429 MHz
```

HSQC NMR Spectrum of Compound 8 (Chapter 5)




```
Current Data Parameters
NAME      Oct09-2009-GK-RAJ
EXPNO                   29
PROCNO                   1

F2 - Acquisition Parameters
Date_              20091011
Time                 22.50
INSTRUM              spect
PROBHD    5 mm PABBO BB-
PULPROG       hmbcgpndqf
TD                     4096
SOLVENT                DMSO
NS                        8
DS                       16
SWH               3703.704 Hz
FIDRES            0.904225 Hz
AQ               0.5530100 sec
RG                     2050
DW                135.000 usec
DE                  6.50 usec
TE                  296.4 K
CNST13           8.0000000
D0               0.00000300 sec
D1               1.33124495 sec
D6               0.06250000 sec
D16              0.00020000 sec
IN0              0.00002235 sec

======== CHANNEL f1 ========
NUC1                    1H
P1                  10.00 usec
P2                  20.00 usec
PL1                 -3.00 dB
PL1W           15.48668575 W
SFO1          400.2218145 MHz

======== CHANNEL f2 ========
NUC2                   13C
P3                   9.00 usec
PL2                 -2.00 dB
PL2W           54.14257431 W
SFO2          100.6454466 MHz

====== GRADIENT CHANNEL =====
GPNAM1            SINE.100
GPNAM2            SINE.100
GPNAM3            SINE.100
GPZ1                50.00 %
GPZ2                30.00 %
GPZ3                40.10 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                      2
TD                     128
SFO1            100.6454 MHz
FIDRES         174.631683 Hz
SW               222.095 ppm
FnMODE                  QF

F2 - Processing parameters
SI                    1024
SF            400.2199980 MHz
WDW                   SINE
```

HMBC NMR Spectrum of Compound 8 (Chapter 5)



IR Spectrum of Compound 18 (Chapter 4) and Compound 8 (Chapter 5)

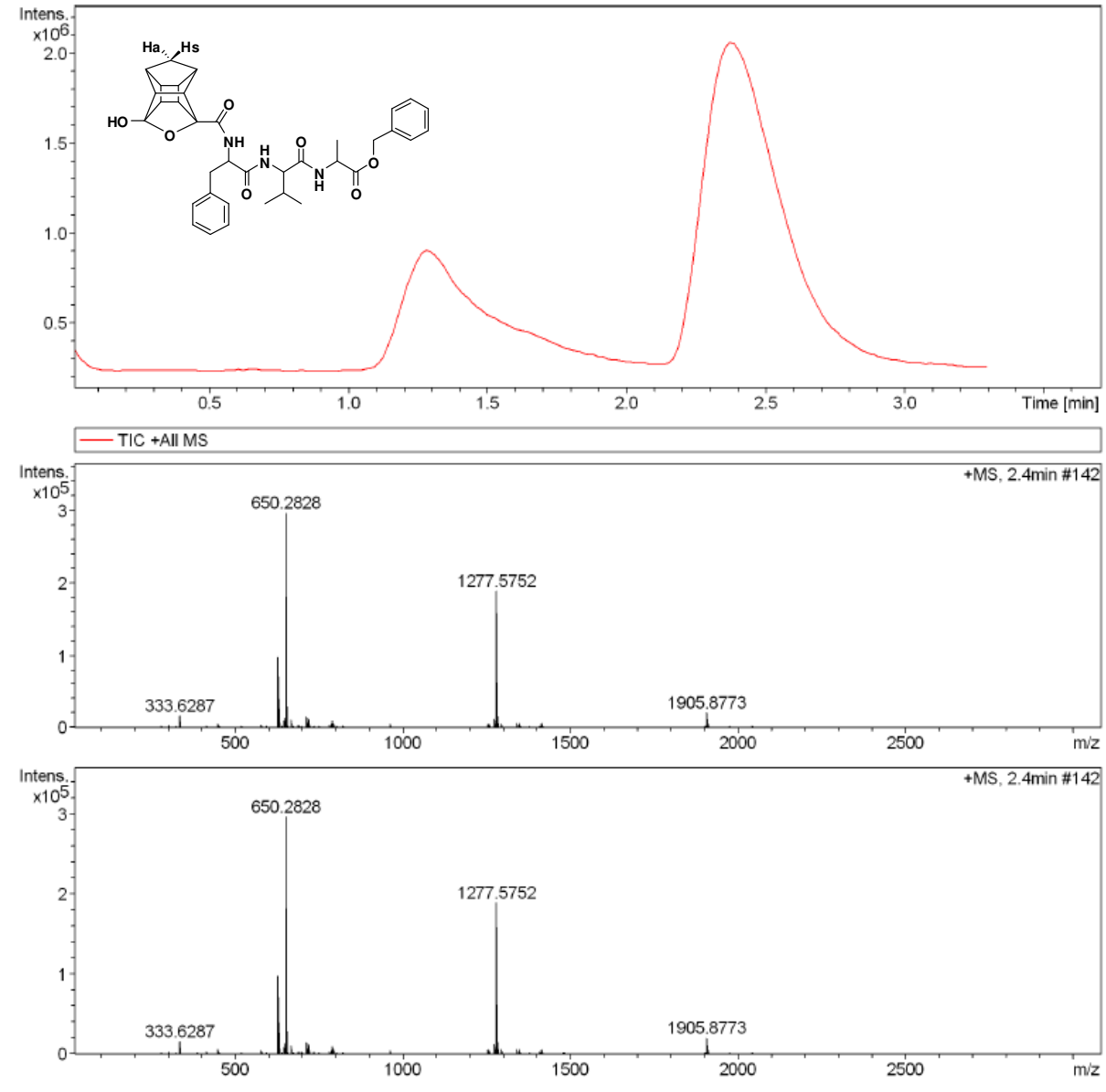
Display Report

Analysis Info

Analysis Name	D:\Data\Raj\OH-PCU-PHE-VAL-ALABZ.d	Acquisition Date	12/30/2009 5:57:33 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	OH-PCU-PHE-VAL-ALABZ	Instrument	micrOTOF-Q 10139
Comment	627.29		

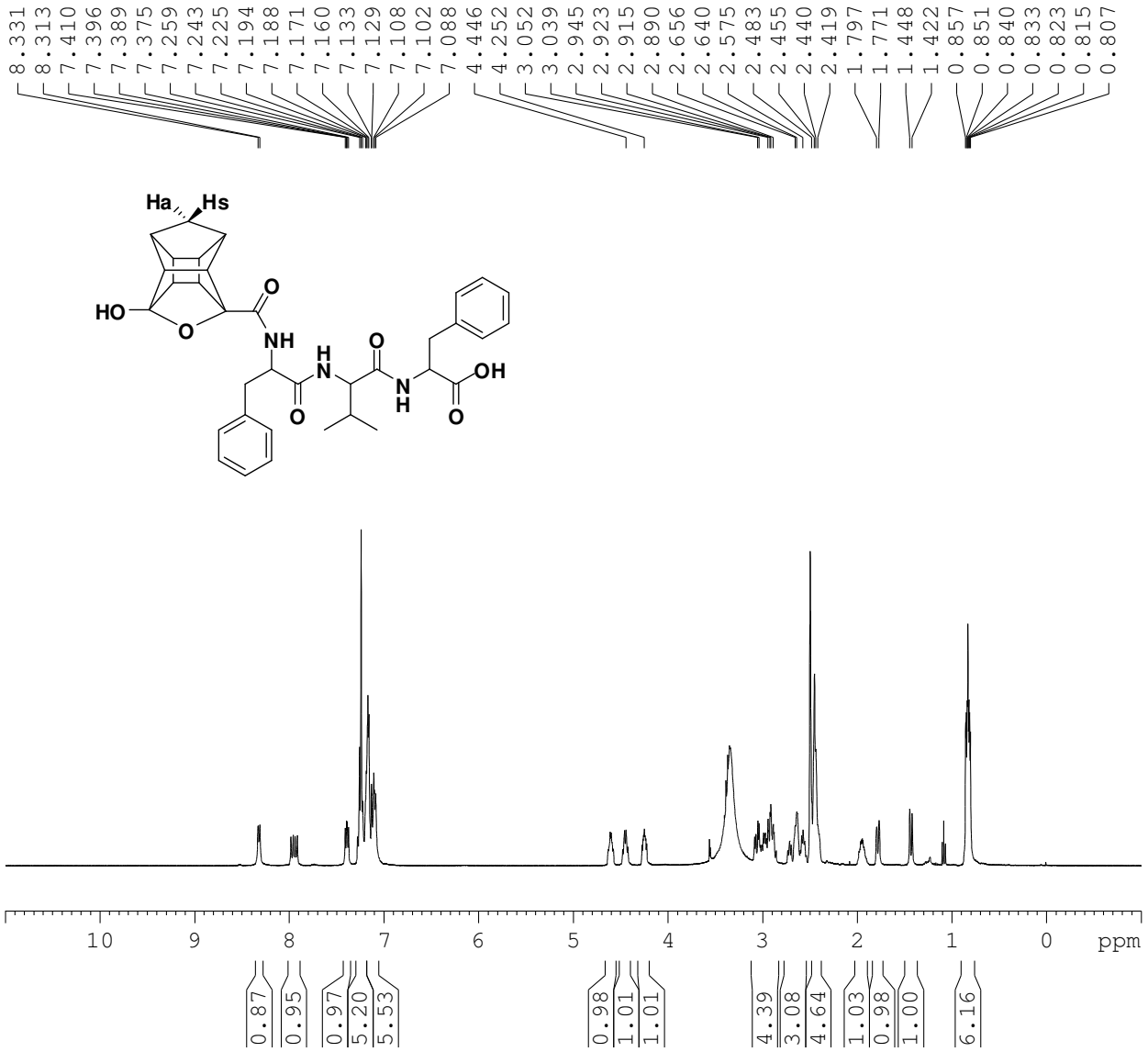
Acquisition Parameter

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



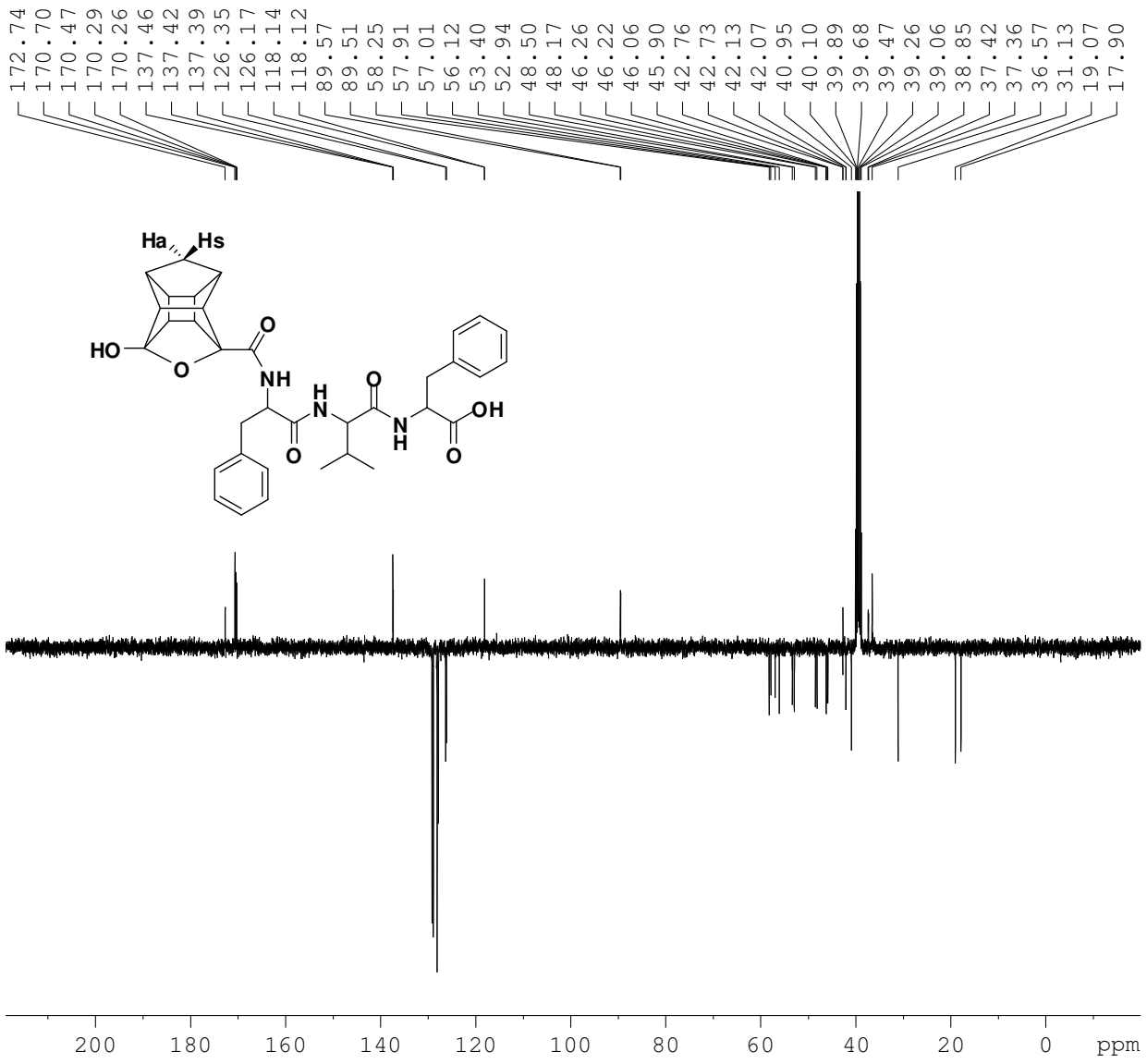
Bruker Compass DataAnalysis 4.0 printed: 12/30/2009 6:07:37 PM Page 1 of 1

HRESIMS (*m/z*) of Compound 18 (Chapter 4) and Compound 8 (Chapter 5)



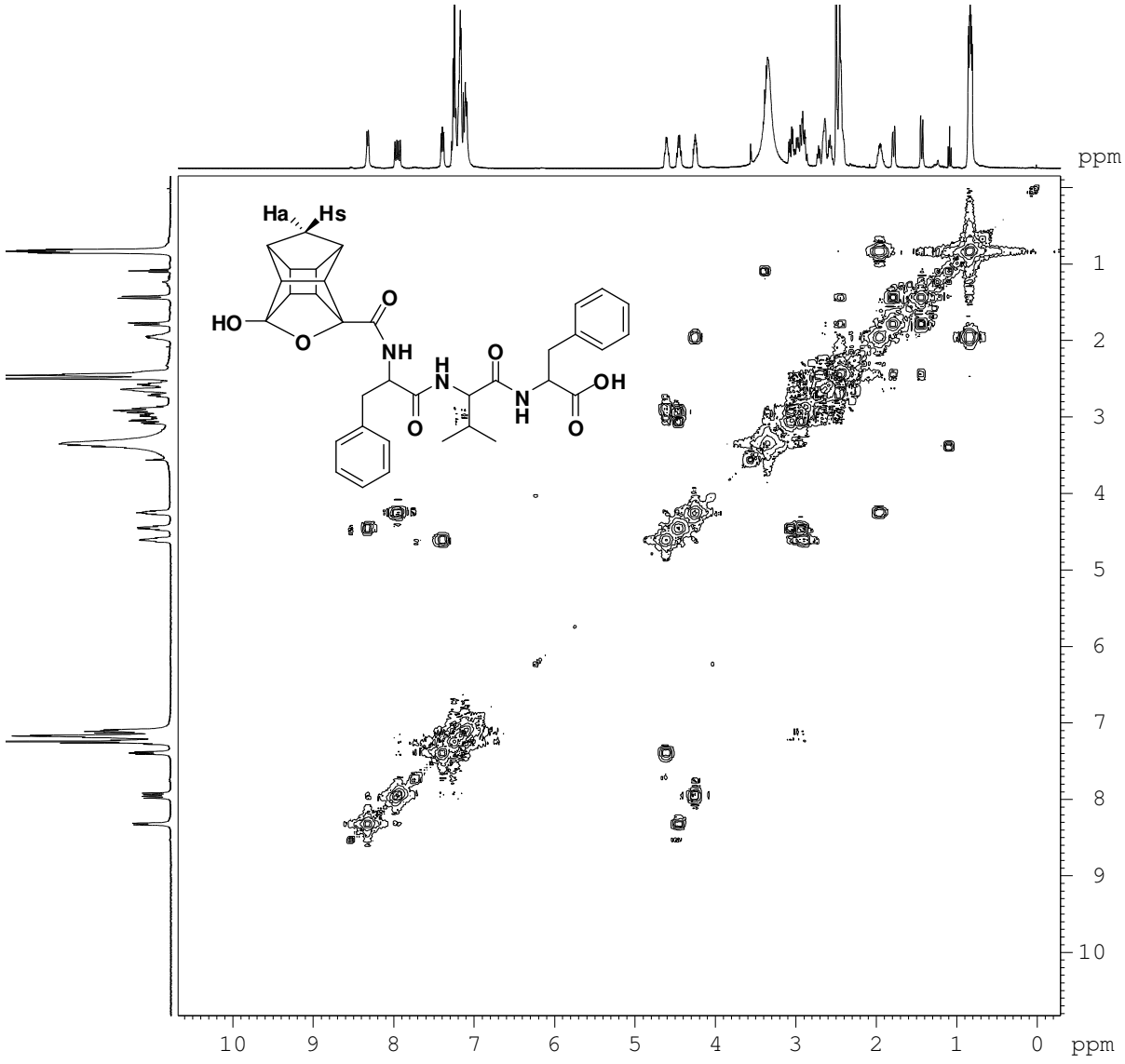

```
Current Data Parameters
NAME      Jul08-2011-GK-Raj
EXPNO                   10
PROCNO                   1

F2 - Acquisition Parameters
Date_           20110708
Time               14.49
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG             zg30
TD                 32768
SOLVENT             DMSO
NS                    16
DS                     2
SWH             8223.685 Hz
FIDRES          0.250967 Hz
AQ              1.9923444 sec
RG                   128
DW                60.800 usec
DE                 6.50 usec
TE                298.2 K
D1            1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                   1H
P1                 10.00 usec
PL1                -3.00 dB
PL1W         15.48668575 W
SFO1         400.2224715 MHz

F2 - Processing parameters
SI                 16384
SF          400.2200001 MHz
WDW                   EM
SSB                    0
LB                  0.30 Hz
GB                     0
PC                  1.00
```

^{1}H NMR Spectrum of Compound 19 (Chapter 4) and Compound 7 (Chapter 5)



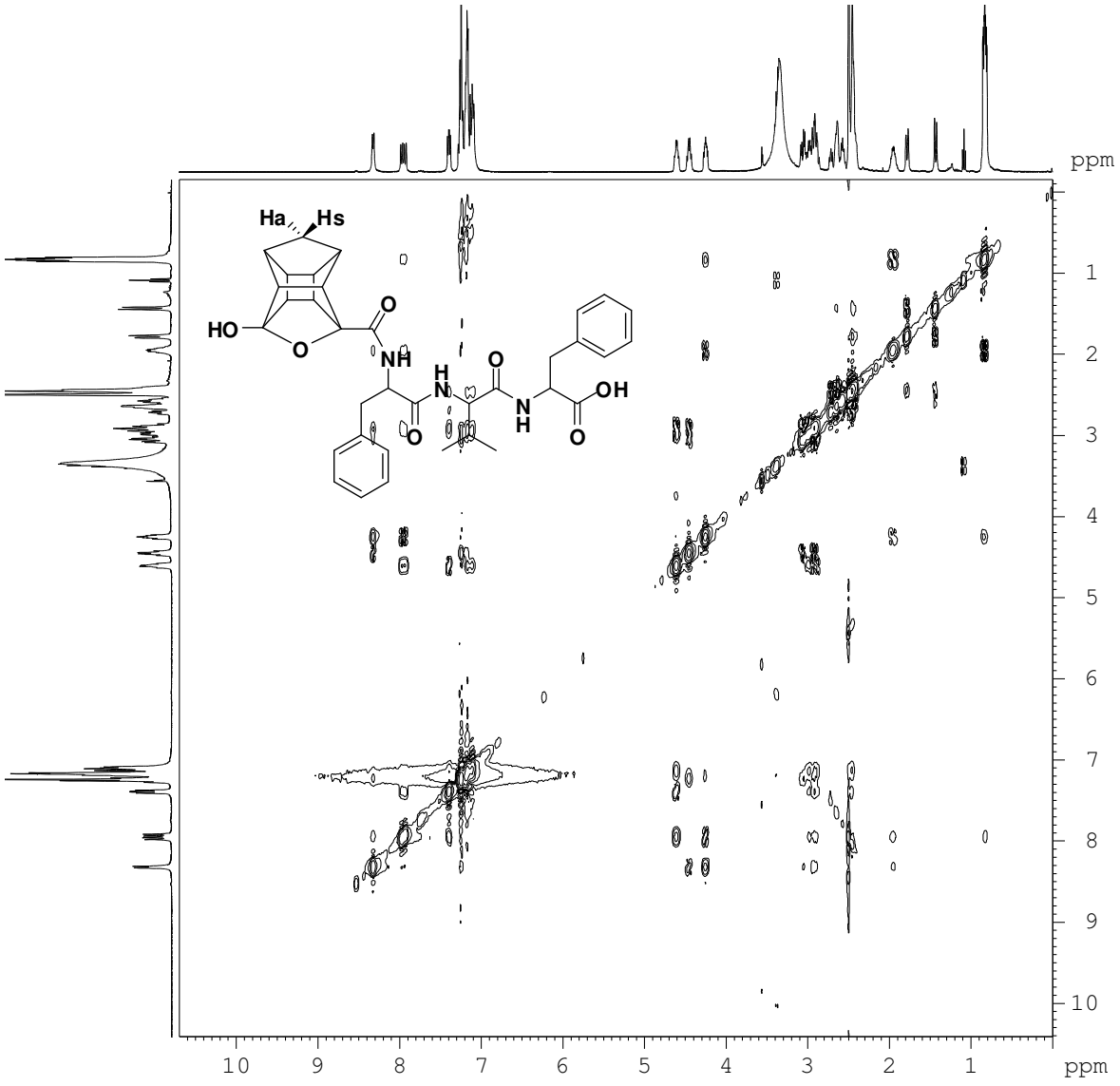

```
Current Data Parameters
NAME      Jul08-2011-GK-Raj
EXPNO                   11
PROCNO                   1

F2 - Acquisition Parameters
Date_           20110708
Time              15.48
INSTRUM           spect
PROBHD   5 mm PABBO BB-
PULPROG             jmod
TD                65536
SOLVENT             DMSO
NS                  512
DS                    4
SWH           24038.461 Hz
FIDRES         0.366798 Hz
AQ             1.3631988 sec
RG                 2050
DW                20.800 usec
DE                 6.50 usec
TE                298.2 K
CNST2        145.0000000
CNST11         1.0000000
D1           2.00000000 sec
D20          0.00689655 sec
TD0                   1

======== CHANNEL f1 ========
NUC1                 13C
P1                  8.40 usec
P2                 16.80 usec
PL1                -2.00 dB
PL1W         54.14257431 W
SFO1         100.6454626 MHz

======== CHANNEL f2 ========
CPDPRG2          waltz16
NUC2                 1H
PCPD2               90.00 usec
PL2                 -3.00 dB
PL12                15.60 dB
PL2W         15.48668575 W
PL12W         0.21377575 W
SFO2         400.2216009 MHz

F2 - Processing parameters
SI                  32768
SF           100.6354466 MHz
WDW                  EM
SSB                   0
LB                 1.00 Hz
GB                    0
PC                 1.40
```

^{13}C NMR Spectrum of Compound 19 (Chapter 4) and Compound 7 (Chapter 5)



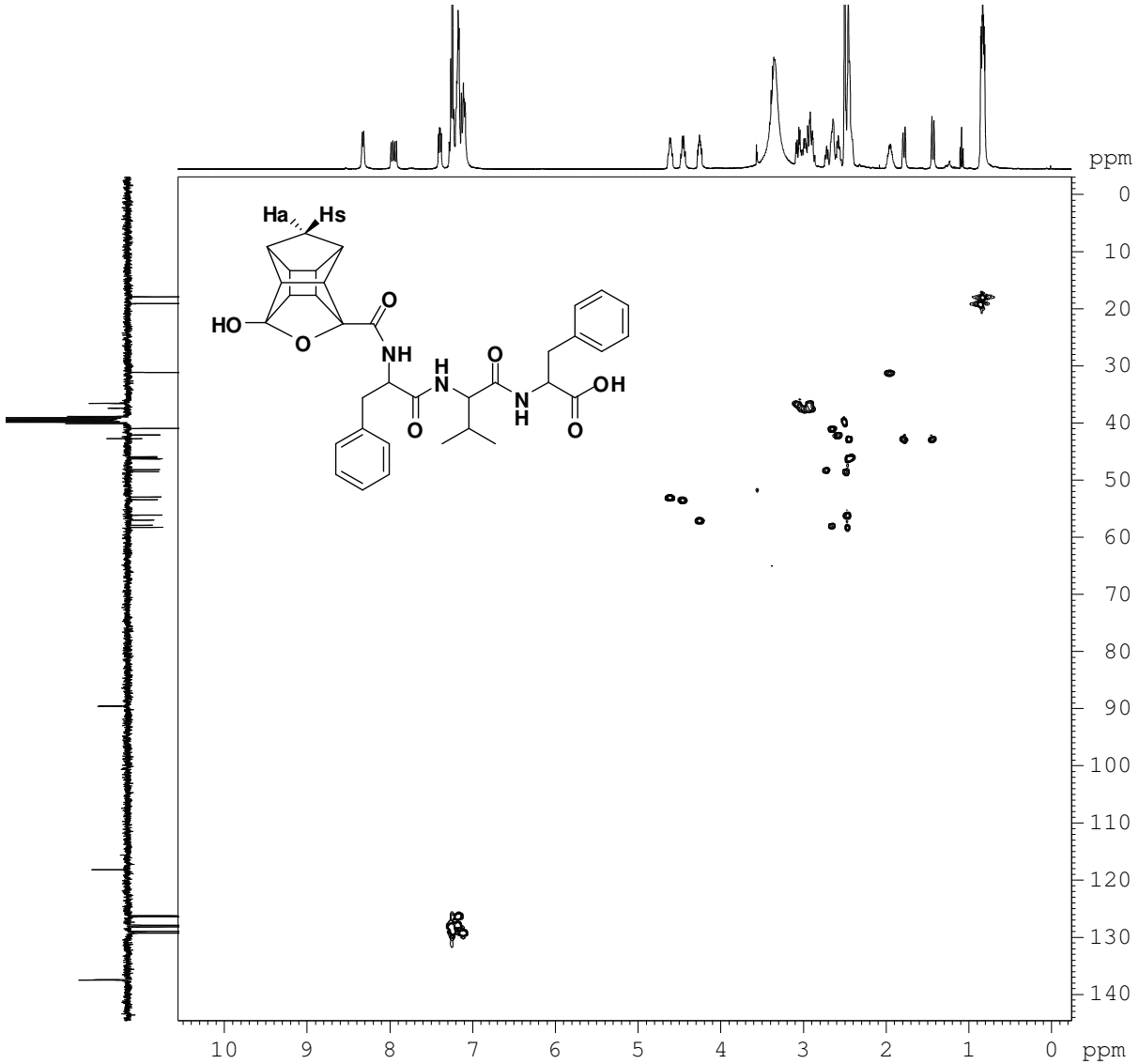

```
Current Data Parameters
NAME      Jul08-2011-GK-Raj
EXPNO                   13
PROCNO                   1

F2 - Acquisition Parameters
Date_              20110708
Time                 15.49
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG          cosygpqf
TD                   2048
SOLVENT               DMSO
NS                       4
DS                       8
SWH               5747.126 Hz
FIDRES            2.806214 Hz
AQ               0.1782260 sec
RG                      64
DW                 87.000 usec
DE                  6.50 usec
TE                 298.2 K
D0              0.00000300 sec
D1              1.50040901 sec
D13             0.00000400 sec
D16             0.00020000 sec
IN0             0.00017400 sec

======== CHANNEL f1 ========
NUC1                    1H
P0                  10.00 usec
P1                  10.00 usec
PL1                 -3.00 dB
PL1W          15.48668575 W
SFO1          400.2226883 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPZ1                10.00 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                      1
TD                     128
SFO1            400.2227 MHz
FIDRES          44.899426 Hz
SW                14.360 ppm
FnMODE                  QF

F2 - Processing parameters
SI                    1024
SF            400.2199980 MHz
WDW                   SINE
SSB                      0
LB                    0.00 Hz
GB                       0
PC                    1.40

F1 - Processing parameters
SI                    1024
MC2                     QF
SF            400.2199990 MHz
WDW                   SINE
SSB                      0
LB                    0.00 Hz
```

COSY NMR Spectrum of Compound 7 (Chapter 5)



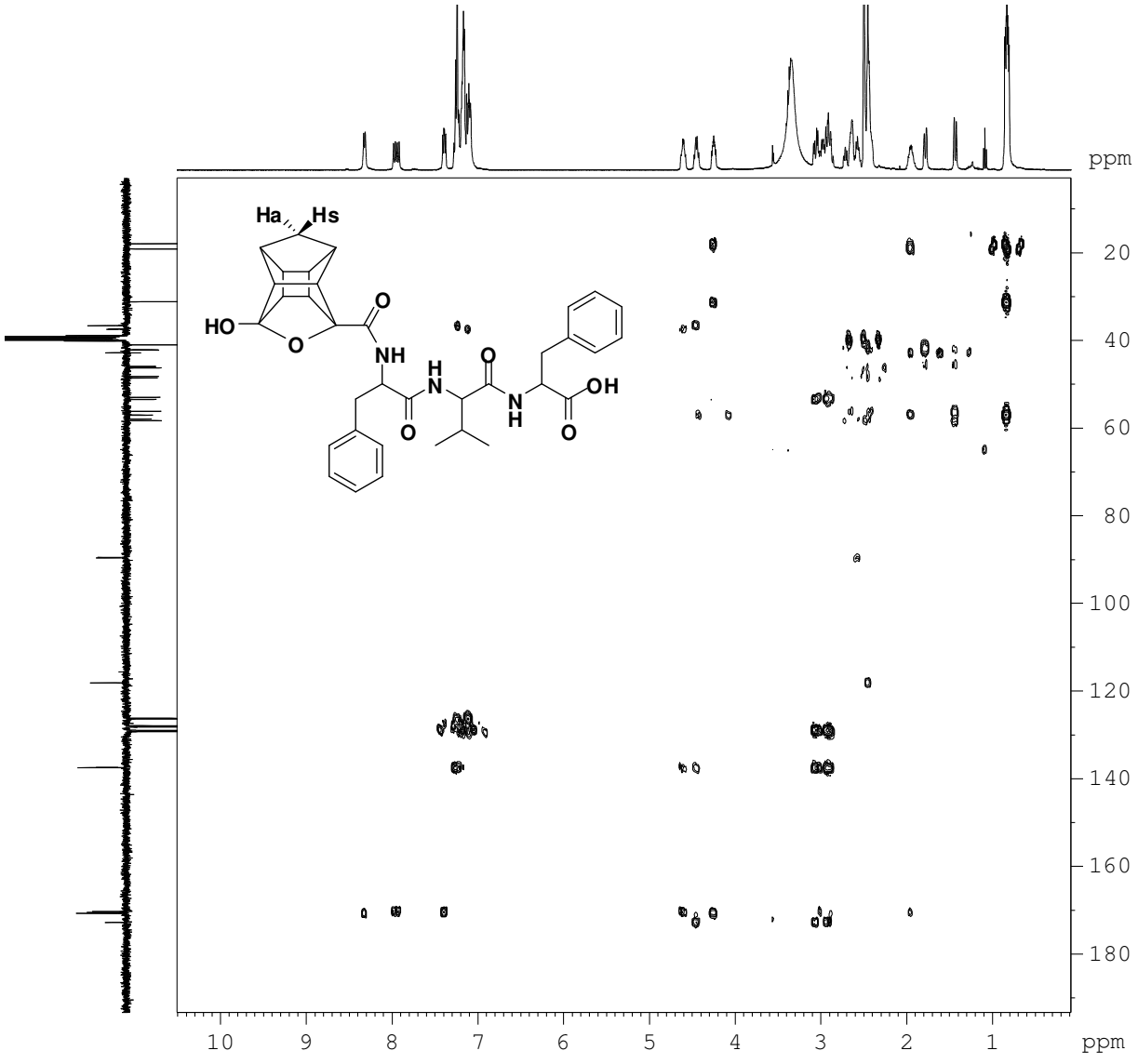

```
Current Data Parameters
NAME      Jul08-2011-GK-Raj
EXPNO                   15
PROCNO                   1

F2 - Acquisition Parameters
Date_              20110708
Time                 17.33
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG           roesyph
TD                    2048
SOLVENT               DMSO
NS                       4
DS                       4
SWH               5747.126 Hz
FIDRES            2.806214 Hz
AQ                0.1782260 sec
RG                      64
DW                 87.000 usec
DE                  6.50 usec
TE                 298.2 K
D0              0.00007663 sec
D1              2.07208896 sec
D12             0.00002000 sec
IN0             0.00017400 sec

======== CHANNEL f1 ========
NUC1                    1H
P1                  10.00 usec
P15             200000.00 usec
PL1                 -3.00 dB
PL11                24.00 dB
PL1W          15.48668575 W
PL11W          0.03090000 W
SFO1          400.2226883 MHz

F1 - Acquisition parameters
ND0                      1
TD                     256
SFO1            400.2227 MHz
FIDRES          22.449713 Hz
SW                14.360 ppm
FnMODE       States-TPPI

F2 - Processing parameters
SI                    1024
SF            400.2199980 MHz
WDW                  QSINE
SSB                      2
LB                    0.00 Hz
GB                       0
PC                    1.00

F1 - Processing parameters
SI                    1024
MC2          States-TPPI
SF            400.2199980 MHz
WDW                  QSINE
SSB                      2
LB                    0.00 Hz
GB                       0
```

ROESY NMR Spectrum of Compound 7 (Chapter 5)



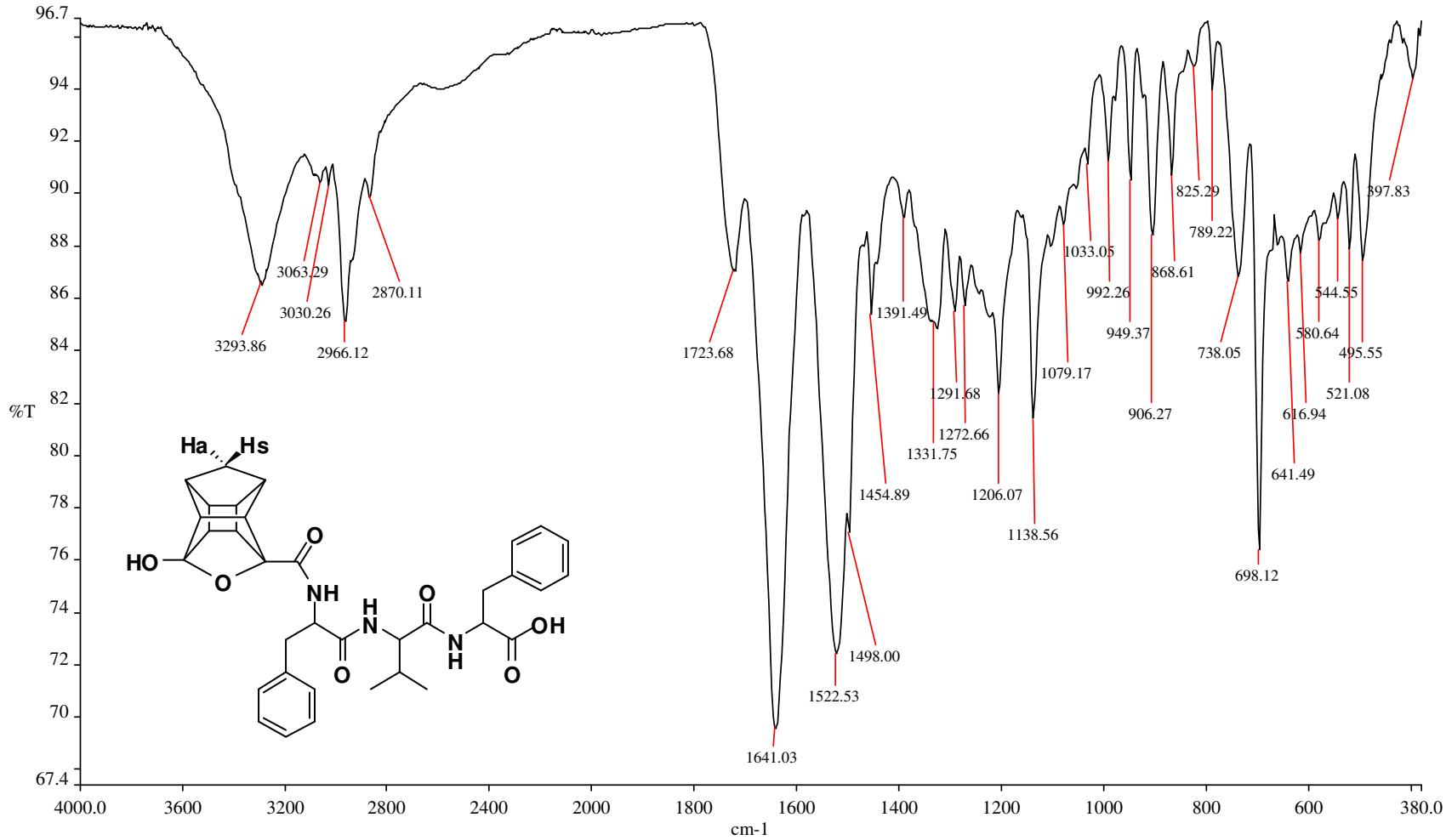

```
Current Data Parameters
NAME      Jul08-2011-GK-Raj
EXPNO                   16
PROCNO                   1

F2 - Acquisition Parameters
Date_              20110708
Time                 18.17
INSTRUM              spect
PROBHD    5 mm PABBO BB-
PULPROG       hsqcedetgp
TD                   1024
SOLVENT               DMSO
NS                       6
DS                      16
SWH              5747.126 Hz
FIDRES           5.612428 Hz
AQ               0.0891380 sec
RG                    2050
DW                 87.000 usec
DE                  6.50 usec
TE                 298.2 K
CNST2         145.0000000
D0             0.00000300 sec
D1             1.50675797 sec
D4             0.00172414 sec
D11            0.03000000 sec
D13            0.00000400 sec
D16            0.00020000 sec
D21            0.00345000 sec
IN0            0.00003000 sec
ZGOPTNS

======== CHANNEL f1 ========
NUC1                    1H
P1                   10.00 usec
P2                   20.00 usec
P28                1000.00 usec
PL1                  -3.00 dB
PL1W          15.48668575 W
SFO1          400.2226883 MHz

======== CHANNEL f2 ========
CPDPRG2               garp
NUC2                   13C
P3                    9.00 usec
P4                   18.00 usec
PCPD2                70.00 usec
PL2                  -2.00 dB
PL12                 15.60 dB
PL2W          54.14257431 W
PL12W          0.94089001 W
SFO2          100.6429049 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPNAM2           SINE.100
GPZ1                80.00 %
GPZ2                20.10 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                      2
TD                     256
SFO1            100.6429 MHz
FIDRES          65.118813 Hz
SW                165.639 ppm
FnMODE     Echo-Antiecho

F2 - Processing parameters
SI                    1024
SF            400.2199980 MHz
WDW                  QSINE
```

HSQC NMR Spectrum of Compound 7 (Chapter 5)




```
Current Data Parameters
NAME      Jul08-2011-GK-Raj
EXPNO                    17
PROCNO                    1

F2 - Acquisition Parameters
Date_              20110708
Time                  19.01
INSTRUM               spect
PROBHD   5 mm PABBO BB-
PULPROG       hmbcgpndqf
TD                    4096
SOLVENT               DMSO
NS                      16
DS                      16
SWH              5747.126 Hz
FIDRES           1.403107 Hz
AQ               0.3564020 sec
RG                    2050
DW                 87.000 usec
DE                  6.50 usec
TE                 298.2 K
CNST13          8.0000000
D0              0.00000300 sec
D1              1.52785301 sec
D6              0.06250000 sec
D16             0.00020000 sec
IN0             0.00002235 sec

======== CHANNEL f1 ========
NUC1                    1H
P1                   10.00 usec
P2                   20.00 usec
PL1                  -3.00 dB
PL1W          15.48668575 W
SFO1          400.2226883 MHz

======== CHANNEL f2 ========
NUC2                   13C
P3                    9.00 usec
PL2                  -2.00 dB
PL2W          54.14257431 W
SFO2          100.6454466 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPNAM2           SINE.100
GPNAM3           SINE.100
GPZ1                50.00 %
GPZ2                30.00 %
GPZ3                40.10 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                      2
TD                     128
SFO1            100.6454 MHz
FIDRES         174.631683 Hz
SW               222.095 ppm
FnMODE                  QF

F2 - Processing parameters
SI                    1024
SF            400.2199970 MHz
WDW                   SINE
SSB                      0
LB                    0.00 Hz
GB                       0
PC                    1.40

F1 - Processing parameters
SI                    1024
MC2                     QF
```

HMBC NMR Spectrum of Compound 7 (Chapter 5)



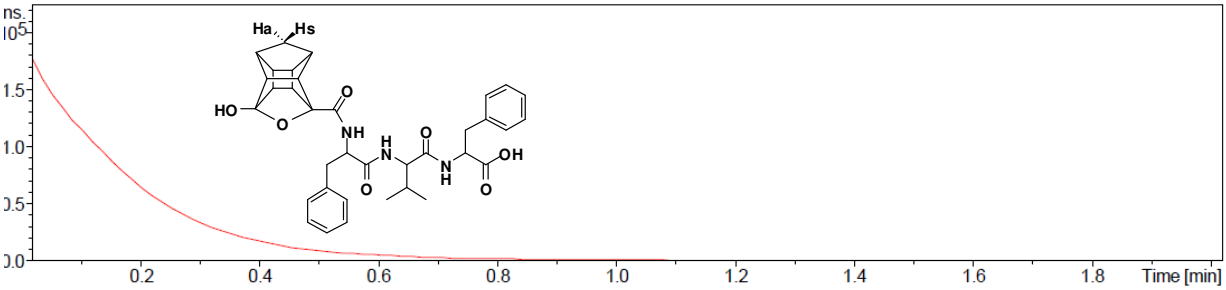
IR Spectrum of Compound 19 (Chapter 4) and Compound 7 (Chapter 5)

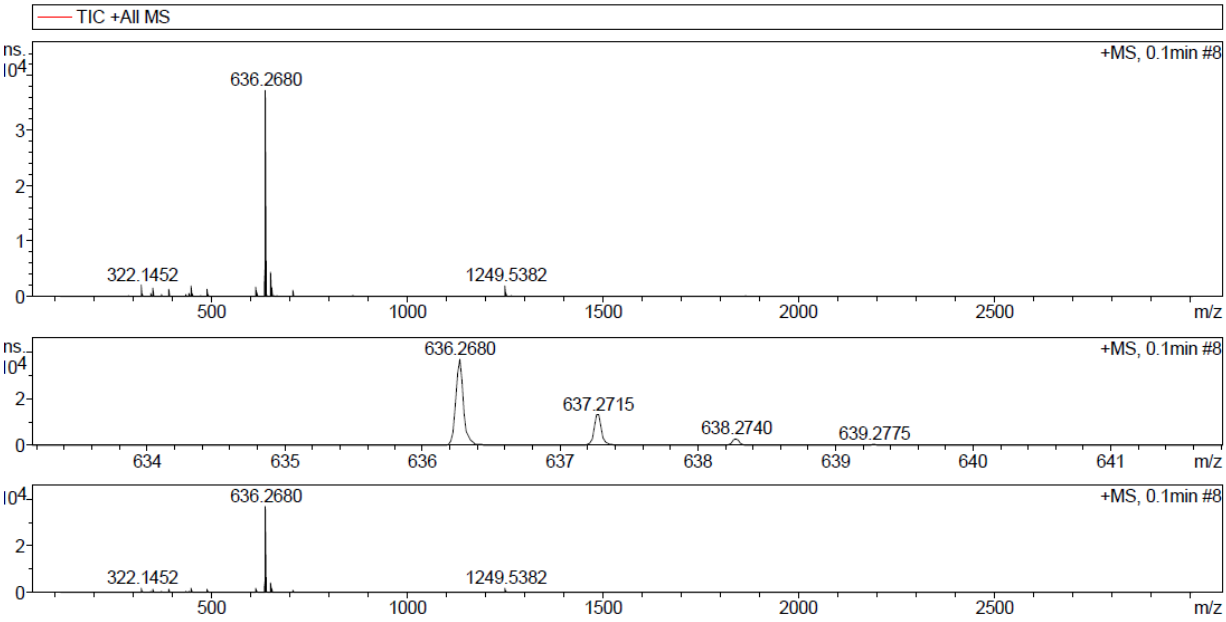
Analysis Info

Analysis Name	D:\Data\Raj\OHpcupvp000001.d	Acquisition Date	8/31/2010 11:55:48 AM
Method	tune_high_expert.m	Operator	BDAL@DE
Sample Name	OHpcupvp	Instrument / Ser#	micrOTOF-Q 10139
Comment	mass = 630		

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	120 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	3000 m/z	Set Collision Cell RF	600.0 Vpp	Set Divert Valve	Source

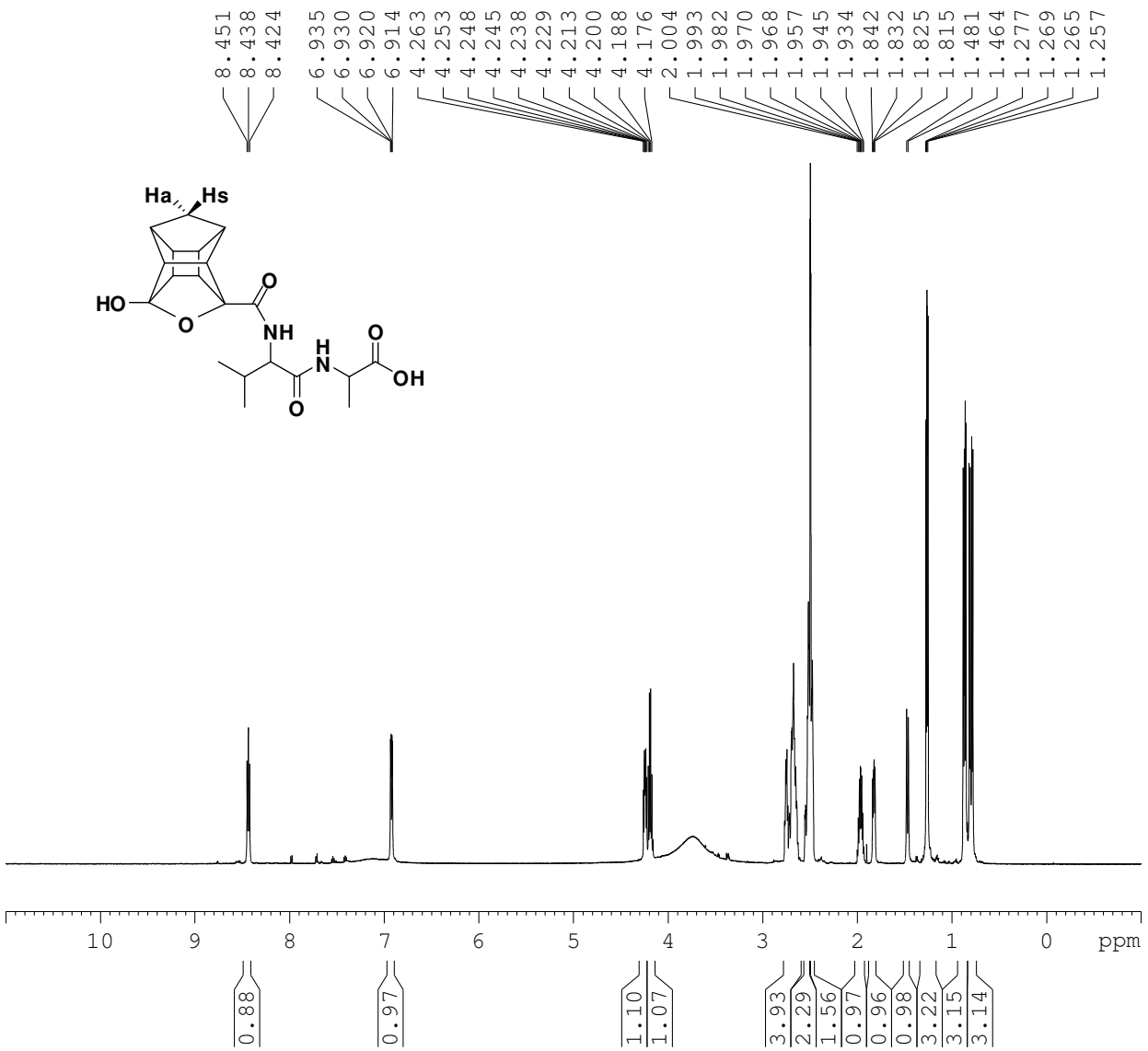




Meas. m/z	#	Formula	m/z	err [ppm]	Mean err [ppm]	rdb	N-Rule	e⁻ Conf	mSigma	Std I	Std Mean m/z	Std I VarNorm	Std m/z Diff	Std Comb Dev
636.2680														
	1	C 33 H 43 Na 3 O 8	636.2646	-5.5	-5.5	11.0	ok	odd	2.39	0.0032	0.0035	0.0012	0.0005	0.7330
	2	C 32 H 41 N 2 Na O 10	636.2653	-4.3	-4.3	13.0	ok	odd	2.82	0.0050	0.0027	0.0018	0.0005	0.7073
	3	C 31 H 38 N 6 Na 2 O 6	636.2643	-5.9	-5.9	15.0	ok	odd	3.71	0.0054	0.0038	0.0025	0.0005	0.7859

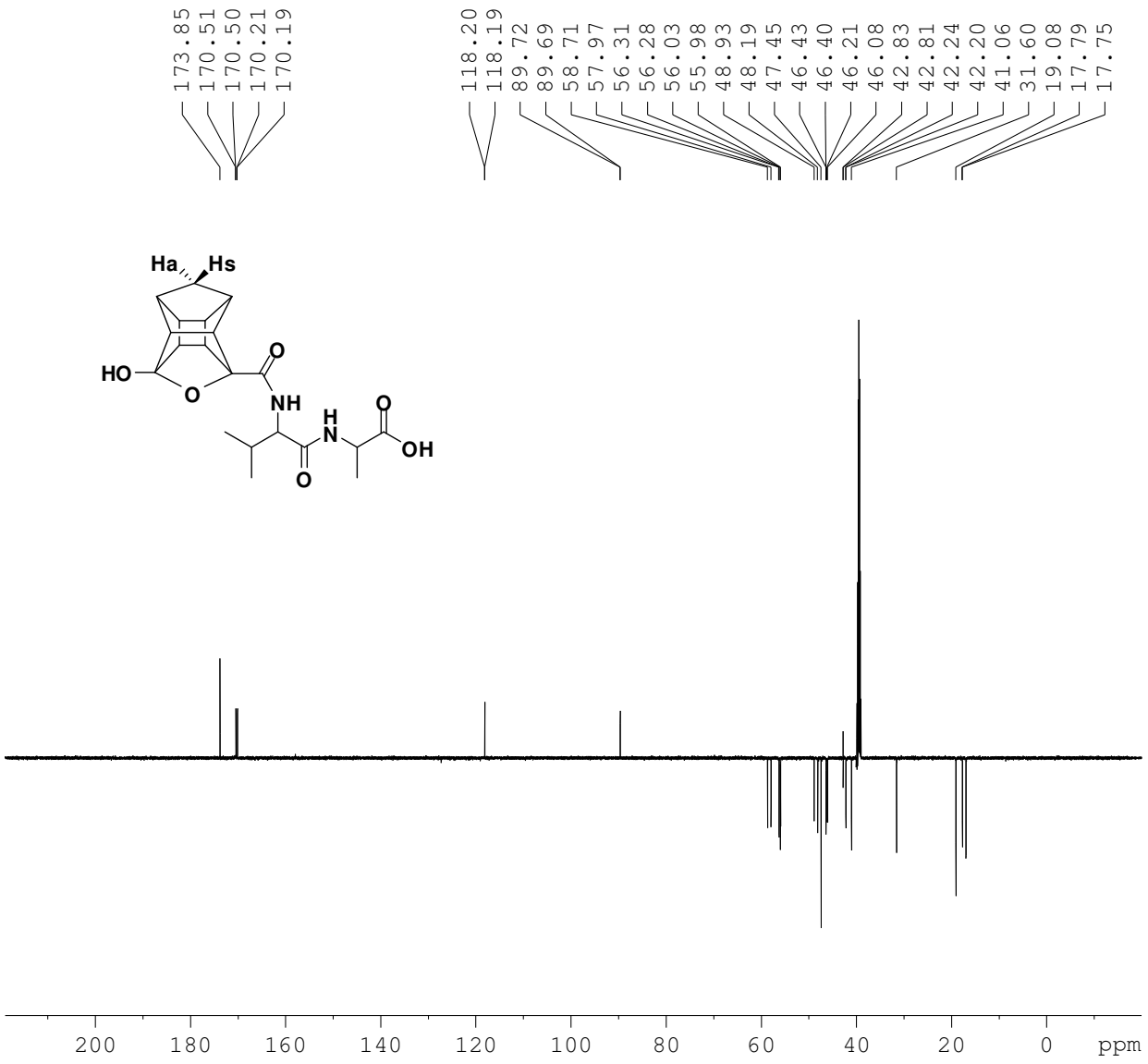
Bruker Compass DataAnalysis 4.0 printed: 10/8/2010 5:41:49 PM Page 1 of 2

HRESIMS (*m/z*) of Compound 19 (Chapter 4) and Compound 7 (Chapter 5)



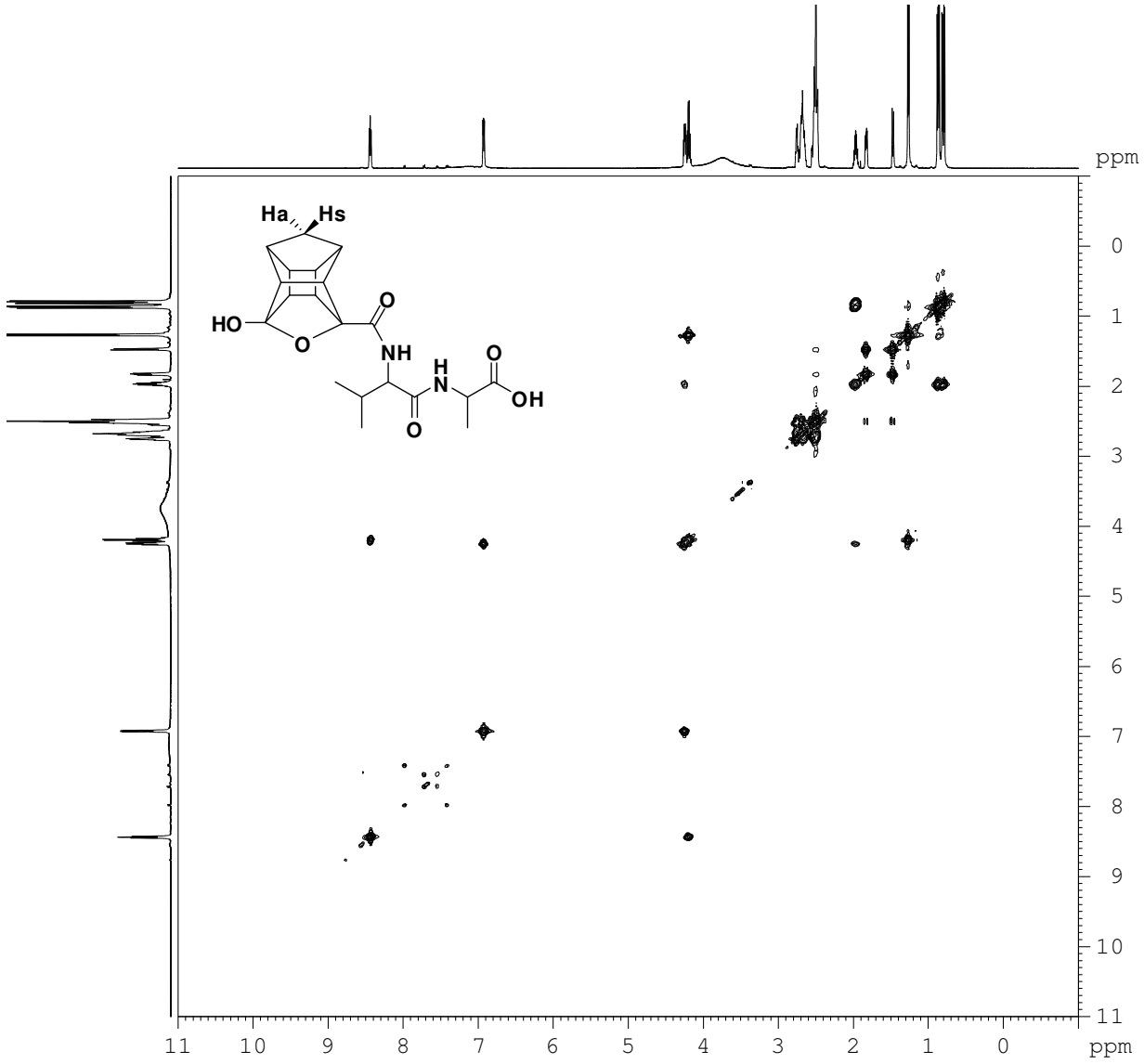

```
Current Data Parameters
NAME                  Raj
EXPNO                 105
PROCNO                  1

F2 - Acquisition Parameters
Date_              20101104
Time                 16.39
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG               zg30
TD                   32768
SOLVENT               DMSO
NS                      16
DS                       2
SWH              12335.526 Hz
FIDRES           0.376450 Hz
AQ               1.3282462 sec
RG                     161
DW                40.533 usec
DE                 6.50 usec
TE                298.1 K
D1             1.00000000 sec
TD0                     1

======== CHANNEL f1 ========
NUC1                   1H
P1                 13.57 usec
PL1                 0.00 dB
PL1W         16.62868881 W
SFO1         600.1037059 MHz

F2 - Processing parameters
SI                  16384
SF           600.1000060 MHz
WDW                    no
SSB                     0
LB                   0.00 Hz
GB                      0
PC                   1.00
```

^{1}H NMR Spectrum of Compound 20 (Chapter 4) and Compound 1 (Chapter 5)



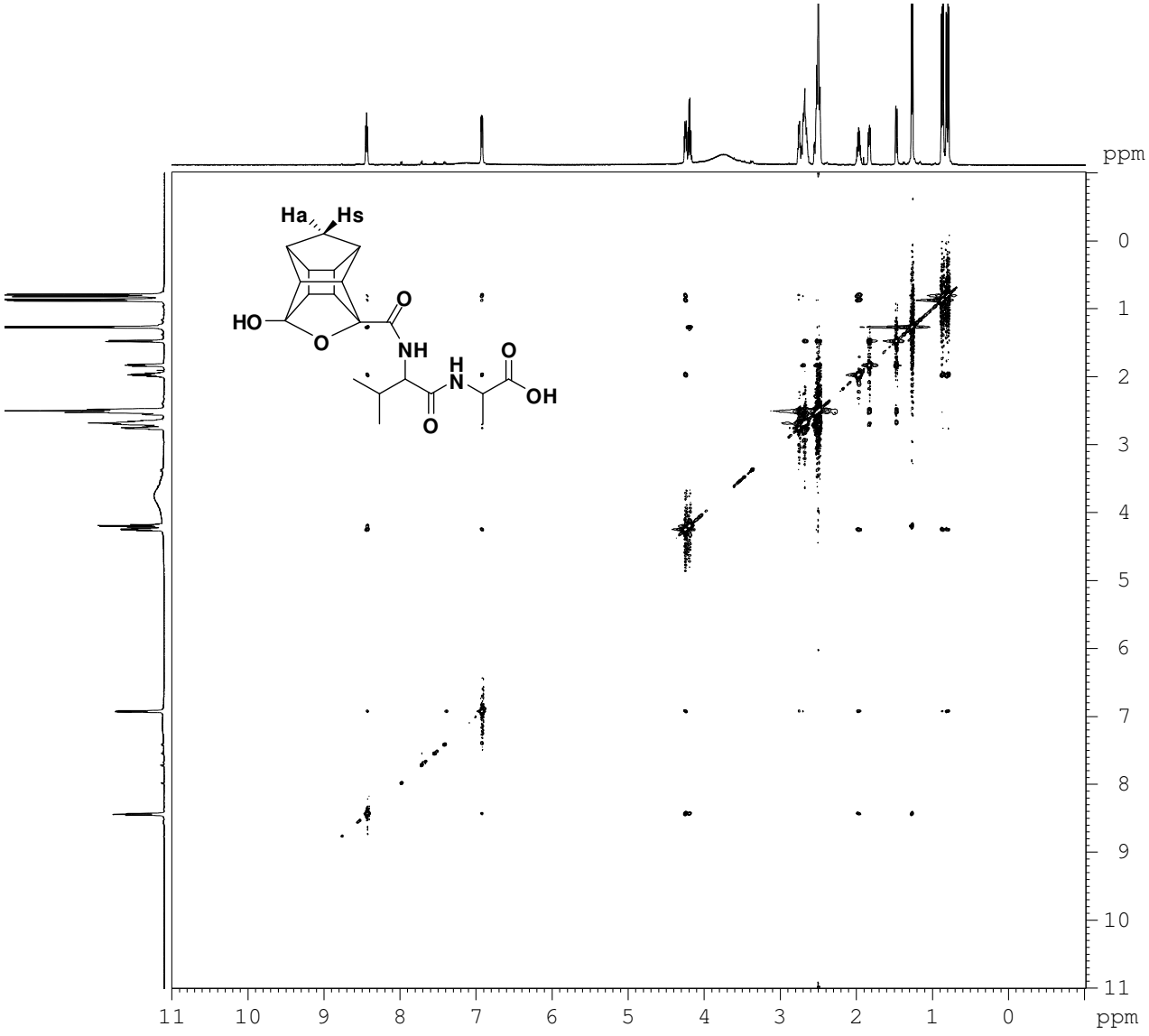

```
Current Data Parameters
NAME                  Raj
EXPNO                 112
PROCNO                  1

F2 - Acquisition Parameters
Date_           20101105
Time                8.42
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG             jmod
TD                 65536
SOLVENT             DMSO
NS                  8160
DS                     4
SWH            36057.691 Hz
FIDRES             0.550197 Hz
AQ                 0.9088159 sec
RG                    2050
DW                   13.867 usec
DE                    6.50 usec
TE                   298.2 K
CNST2          145.0000000
CNST11           1.0000000
D1               2.00000000 sec
D20              0.00689655 sec
TD0                     1

======== CHANNEL f1 ========
NUC1                  13C
P1                   9.00 usec
P2                  18.00 usec
PL1                  1.50 dB
PL1W          59.57236099 W
SFO1          150.9103545 MHz

======== CHANNEL f2 ========
CPDPRG2           waltz16
NUC2                  1H
PCPD2               96.00 usec
PL2                 -0.50 dB
PL12                15.40 dB
PL2W          18.65769577 W
PL12W          0.47957665 W
SFO2          600.1024004 MHz

F2 - Processing parameters
SI                  32768
SF            150.8953375 MHz
WDW                    no
SSB                     0
LB                   0.00 Hz
GB                      0
PC                   1.40
```

^{13}C NMR Spectrum of Compound 20 (Chapter 4) and Compound 1 (Chapter 5)



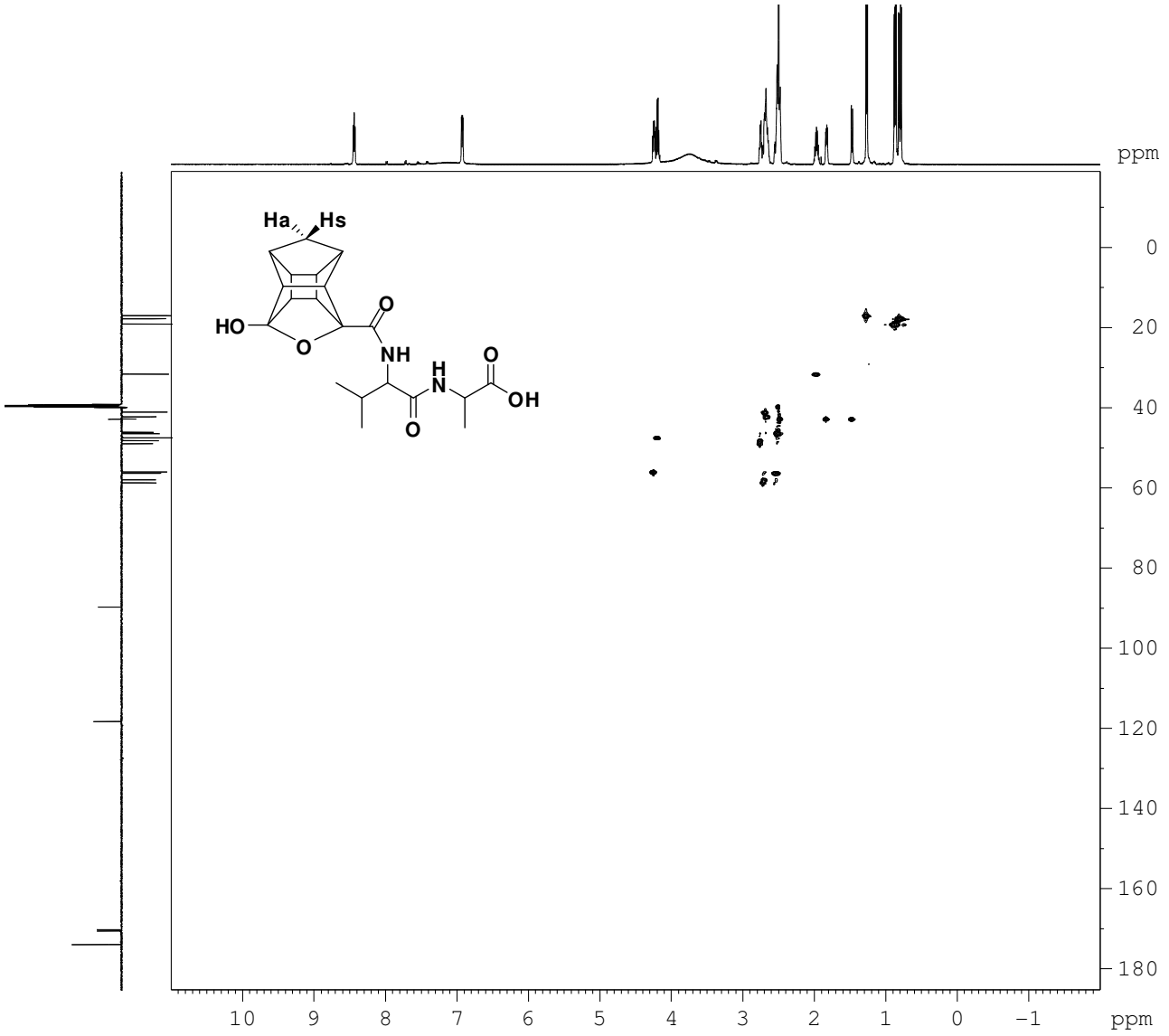

```
Current Data Parameters
NAME                  Raj
EXPNO                 108
PROCNO                  1

F2 - Acquisition Parameters
Date_              20101104
Time                 19.25
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG           cosygpqf
TD                    2048
SOLVENT               DMSO
NS                       2
DS                       8
SWH              12335.526 Hz
FIDRES            6.023206 Hz
AQ               0.0830623 sec
RG                      64
DW                 40.533 usec
DE                  6.50 usec
TE                 298.1 K
D0              0.00000300 sec
D1              3.00000000 sec
D13             0.00000400 sec
D16             0.00020000 sec
IN0             0.00008120 sec

======== CHANNEL f1 ========
NUC1                   1H
P0                  13.56 usec
P1                  13.57 usec
PL1                  0.00 dB
PL1W          16.62868881 W
SFO1         600.1037059 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPZ1                10.00 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                     1
TD                    512
SFO1            600.1037 MHz
FIDRES          24.053263 Hz
SW                20.522 ppm
FnMODE                 QF

F2 - Processing parameters
SI                    1024
SF           600.1000000 MHz
WDW                   SINE
SSB                      0
LB                  0.00 Hz
GB                       0
PC                    1.00

F1 - Processing parameters
SI                    1024
MC2                    QF
SF           600.1000000 MHz
WDW                   SINE
SSB                      0
LB                  0.00 Hz
```

COSY NMR Spectrum of Compound 1 (Chapter 5)



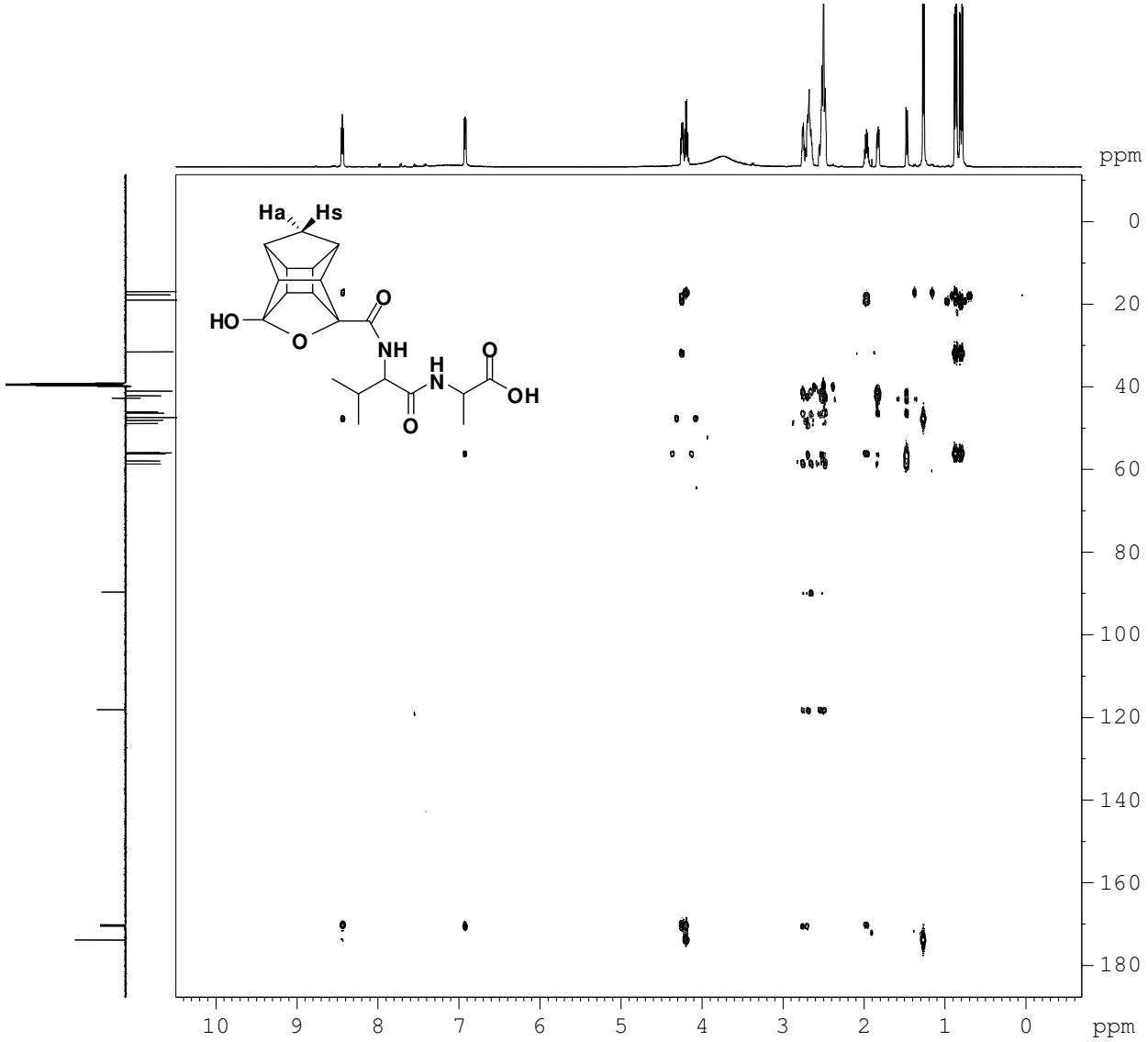

```
Current Data Parameters
NAME                  Raj
EXPNO                 107
PROCNO                  1

F2 - Acquisition Parameters
Date_           20101110
Time                9.58
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG          kproesy
TD                  2048
SOLVENT              DMSO
NS                      8
DS                     64
SWH             7211.539 Hz
FIDRES          3.521259 Hz
AQ             0.1420447 sec
RG                  22.6
DW                69.333 usec
DE                  6.00 usec
TE                298.3 K
CNST2       6849.0000000
D0             0.00005206 sec
D1             1.50000000 sec
D11            0.03000000 sec
D12            0.00002000 sec
D13            0.00000400 sec
D16            0.00030000 sec
IN0            0.00013865 sec

======== CHANNEL f1 ========
CNST1          0.0000000
CNST3       3954.3879395
CNST4      -3954.3879395
NUC1                1H
P1                13.57 usec
P7             250000.00 usec
P11              1000.00 usec
PL1                0.00 dB
PL9               70.00 dB
PL1W         16.62868881 W
PL9W          0.00000166 W
SFO1         600.1030005 MHz
SP1               13.00 dB
SP2               13.00 dB
SP3               13.00 dB
SP4               13.00 dB
SP5               13.00 dB
SP6               13.00 dB
SPNAM1   gaussramp-down
SPNAM2       square.100
SPNAM3     gaussramp-up
SPNAM4   gaussramp+down
SPNAM5       square.100
SPNAM6     gaussramp+up
SPOAL1            1.000
SPOAL2            0.500
SPOAL3            0.000
SPOAL4            1.000
SPOAL5            0.500
SPOAL6            0.000
SPOFFS1            0.00 Hz
SPOFFS2            0.00 Hz
SPOFFS3            0.00 Hz
```

ROESY NMR Spectrum of Compound 1 (Chapter 5)



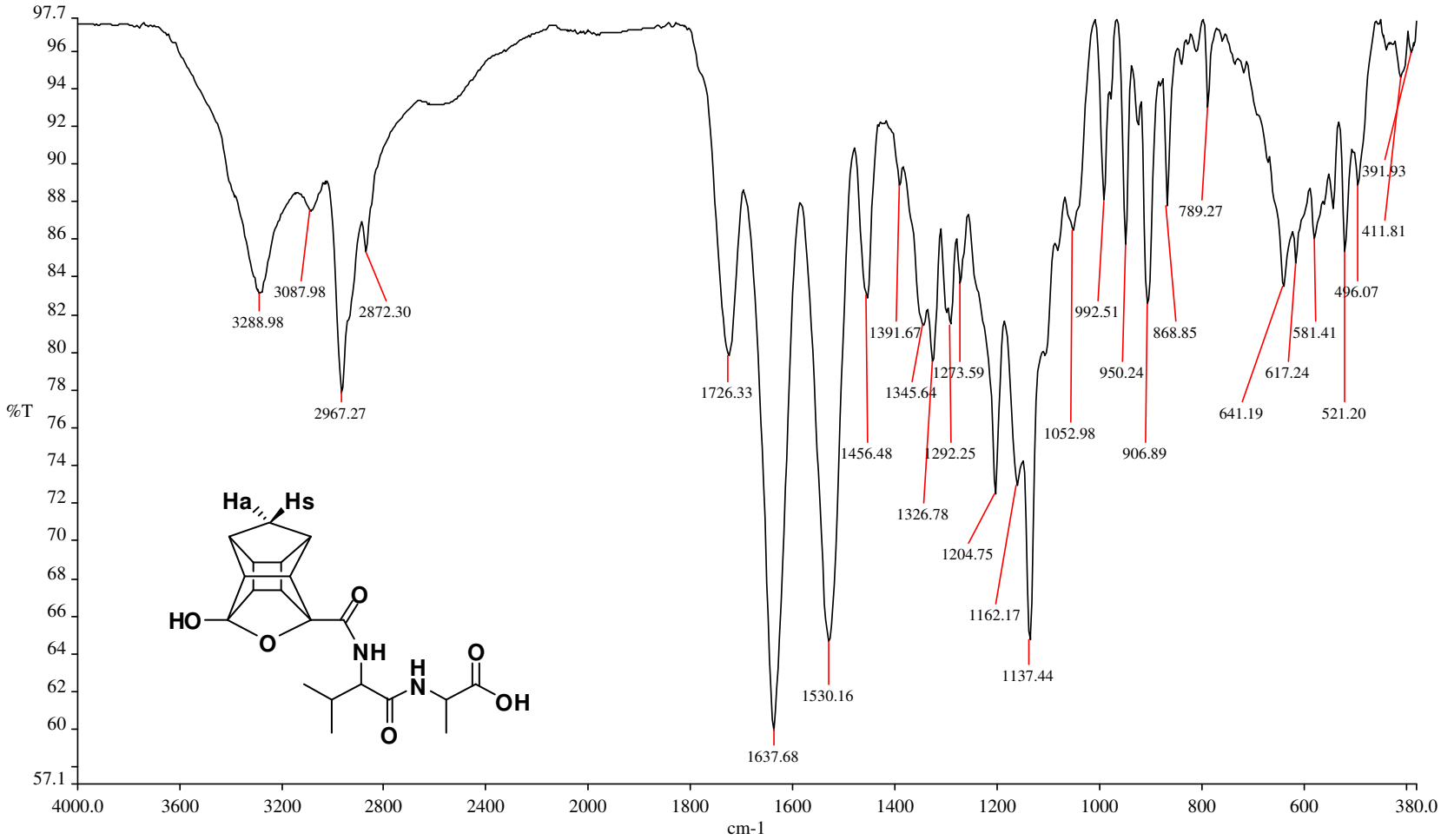

```
Current Data Parameters
NAME                  Raj
EXPNO                 109
PROCNO                  1

F2 - Acquisition Parameters
Date_           20101104
Time               20.20
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG         hsqcetgp
TD                 2048
SOLVENT             DMSO
NS                     8
DS                    16
SWH           12335.526 Hz
FIDRES         6.023206 Hz
AQ            0.0830623 sec
RG                  2050
DW                40.533 usec
DE                 6.50 usec
TE                298.3 K
CNST2        145.0000000
D0            0.00000300 sec
D1            1.00000000 sec
D4            0.00172414 sec
D11           0.03000000 sec
D13           0.00000400 sec
D16           0.00020000 sec
IN0           0.00001385 sec
ZGOPTNS

======== CHANNEL f1 ========
NUC1                  1H
P1                 13.57 usec
P2                 27.14 usec
P28              1000.00 usec
PL1                0.00 dB
PL1W         16.62868881 W
SFO1        600.1037059 MHz

======== CHANNEL f2 ========
CPDPRG2             garp
NUC2                 13C
P3                 9.00 usec
P4                18.00 usec
PCPD2             60.00 usec
PL2                1.50 dB
PL12              17.98 dB
PL2W         59.57236099 W
PL12W         1.33981502 W
SFO2        150.9103545 MHz

====== GRADIENT CHANNEL =====
GPNAM1          SINE.100
GPNAM2          SINE.100
GPZ1               80.00 %
GPZ2               20.12 %
P16              1000.00 usec

F1 - Acquisition parameters
ND0                    2
TD                   512
SFO1           150.9104 MHz
FIDRES        70.426971 Hz
```

HSQC NMR Spectrum of Compound 1 (Chapter 5)




```
Current Data Parameters
NAME                  Raj
EXPNO                 110
PROCNO                  1

F2 - Acquisition Parameters
Date_              20101104
Time                 21.37
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG    hmbcgplpndqf
TD                   4096
SOLVENT               DMSO
NS                      16
DS                      16
SWH              9615.385 Hz
FIDRES           2.347506 Hz
AQ               0.2130420 sec
RG                    2050
DW                  52.000 usec
DE                   6.50 usec
TE                  298.1 K
CNST2          145.0000000
CNST13          10.0000000
D0               0.00000300 sec
D1               3.00000000 sec
D2               0.00344828 sec
D6               0.05000000 sec
D16              0.00020000 sec
IN0              0.00001385 sec

======== CHANNEL f1 ========
NUC1                   1H
P1                  13.57 usec
P2                  27.14 usec
PL1                  0.00 dB
PL1W          16.62868881 W
SFO1          600.1036006 MHz

======== CHANNEL f2 ========
NUC2                  13C
P3                   9.00 usec
PL2                  1.50 dB
PL2W          59.57236099 W
SFO2          150.9103545 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPNAM2           SINE.100
GPNAM3           SINE.100
GPZ1                60.00 %
GPZ2                20.00 %
GPZ3                60.12 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                     2
TD                    256
SFO1            150.9104 MHz
FIDRES        140.850357 Hz
SW               238.935 ppm
FnMODE                QF

F2 - Processing parameters
SI                    2048
```

HMBC NMR Spectrum of Compound 1 (Chapter 5)



IR Spectrum of Compound 20 (Chapter 4) and Compound 1 (Chapter 5)

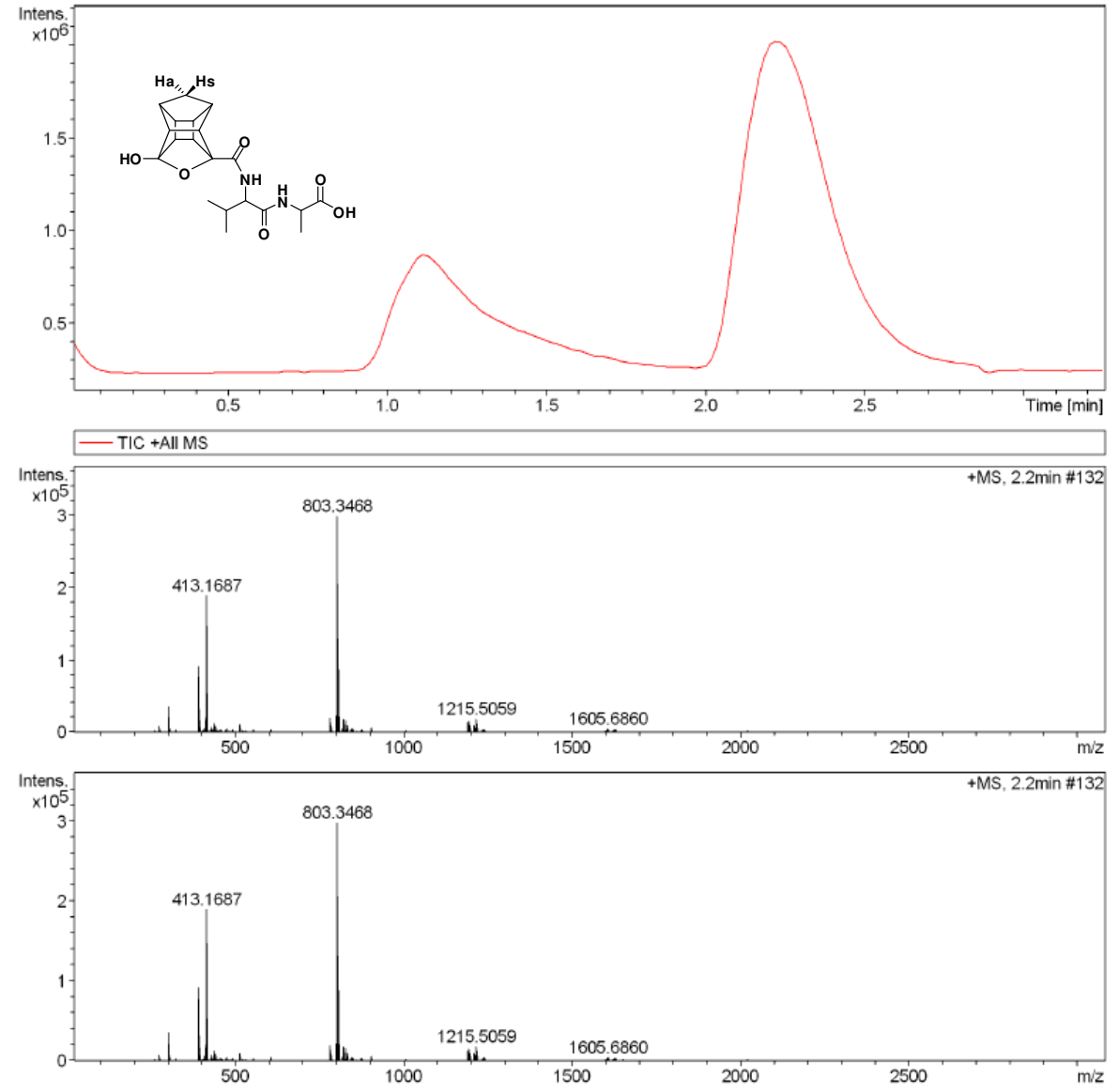
Display Report

Analysis Info

Analysis Name	D:\Data\Raj\OH-PCU-VAL-ALA-OH.d	Acquisition Date	12/30/2009 6:10:56 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	OH-PCU-VAL-ALA-OH	Instrument	micrOTOF-Q 10139
Comment	390.18		

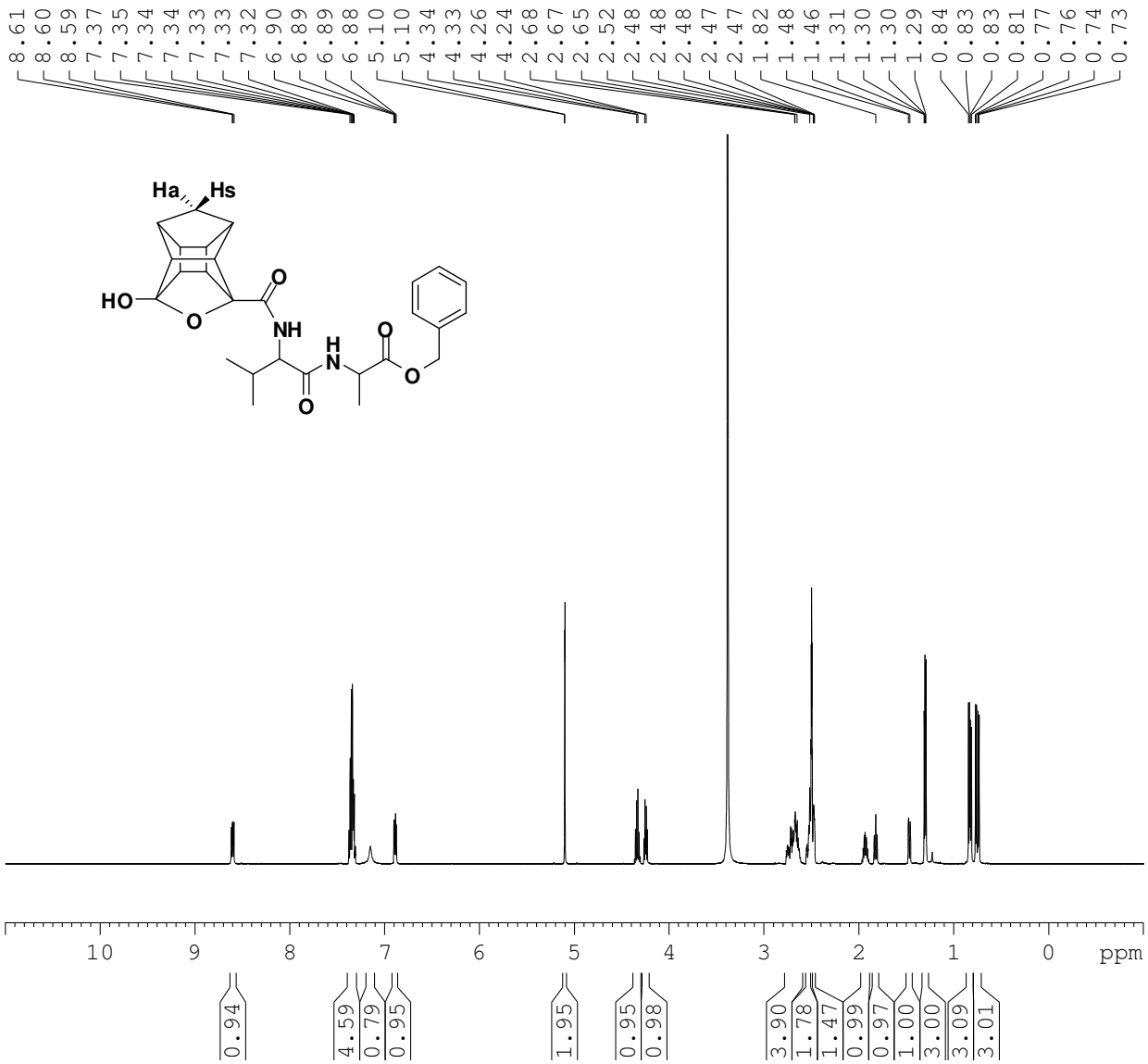
Acquisition Parameter

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



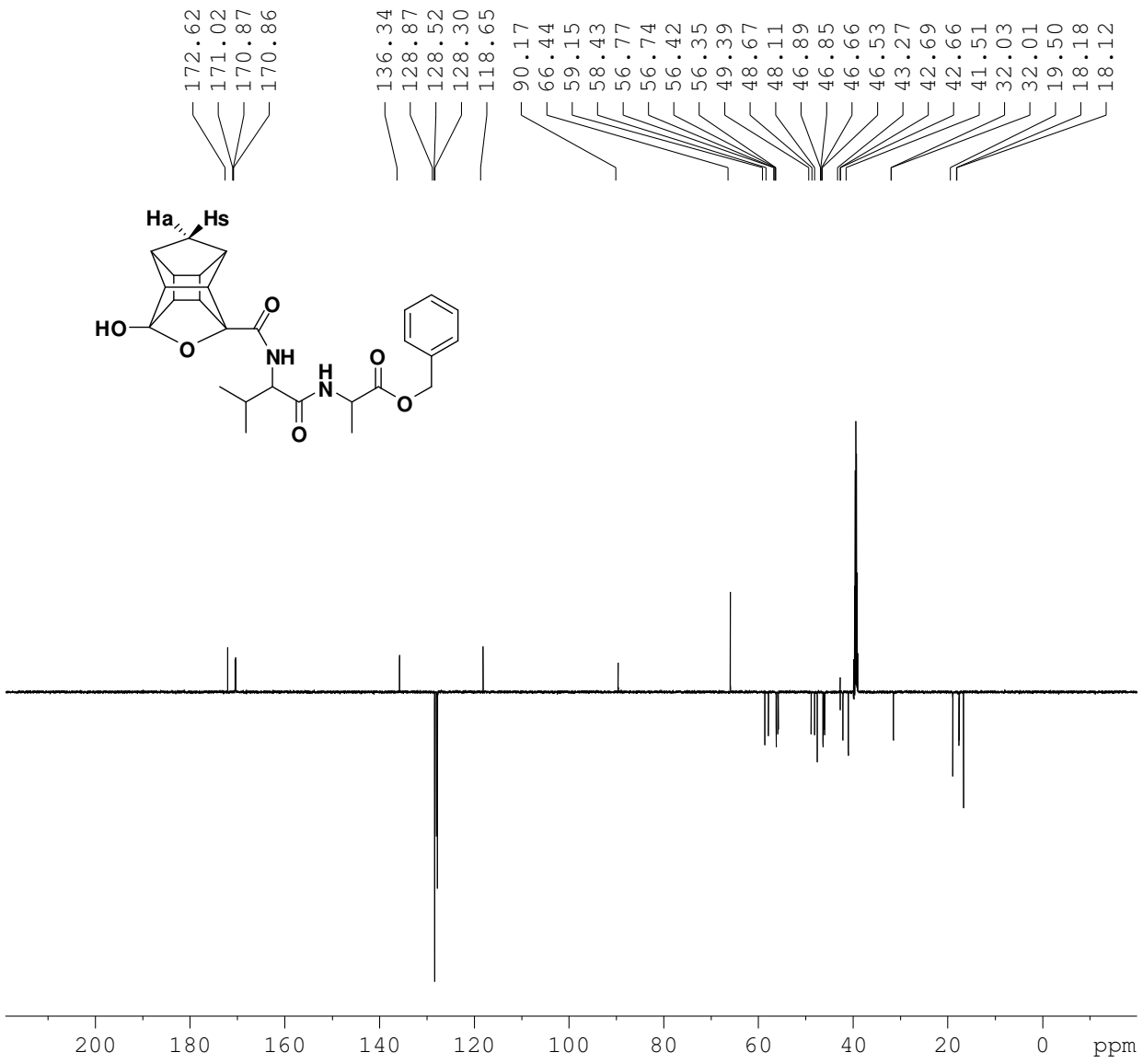
Bruker Compass DataAnalysis 4.0 printed: 12/30/2009 6:17:24 PM Page 1 of 1

HRESIMS (*m/z*) of Compound 20 (Chapter 4) and Compound 1 (Chapter 5)



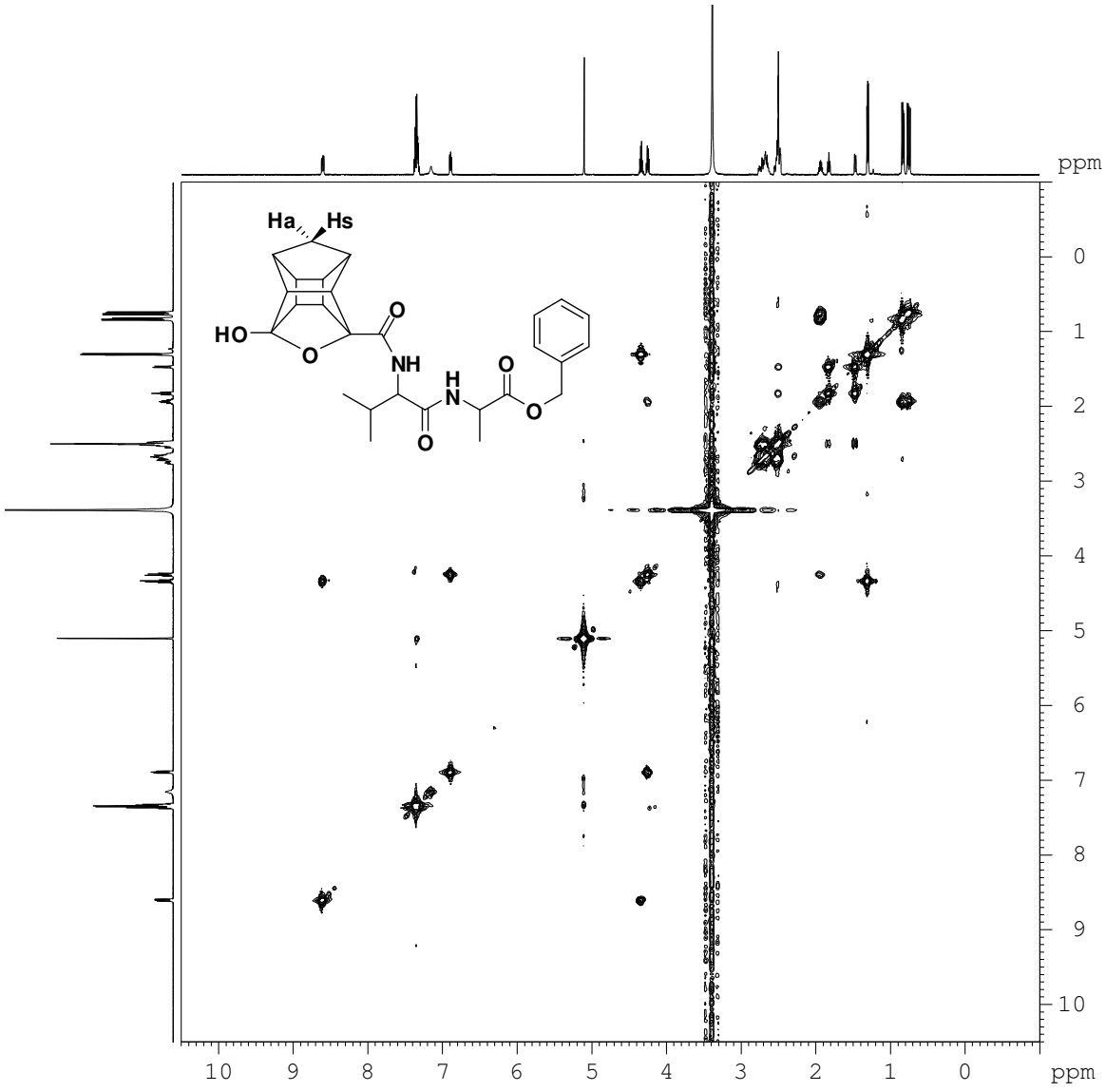

```
Current Data Parameters
NAME                  Raj
EXPNO                 170
PROCNO                  1

F2 - Acquisition Parameters
Date_           20101217
Time                9.40
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG             zg30
TD                 32768
SOLVENT             DMSO
NS                    16
DS                     4
SWH            12335.526 Hz
FIDRES          0.376450 Hz
AQ              1.3282462 sec
RG                    57
DW                40.533 usec
DE                 6.50 usec
TE                298.3 K
D1            1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                  1H
P1                13.75 usec
PL1                0.00 dB
PL1W         16.62868881 W
SFO1         600.1037059 MHz

F2 - Processing parameters
SI                 16384
SF           600.1000060 MHz
WDW                   no
SSB                    0
LB                 0.00 Hz
GB                     0
PC                  1.00
```

^{1}H NMR Spectrum of Compound 21 (Chapter 4) and Compound 3 (Chapter 5)



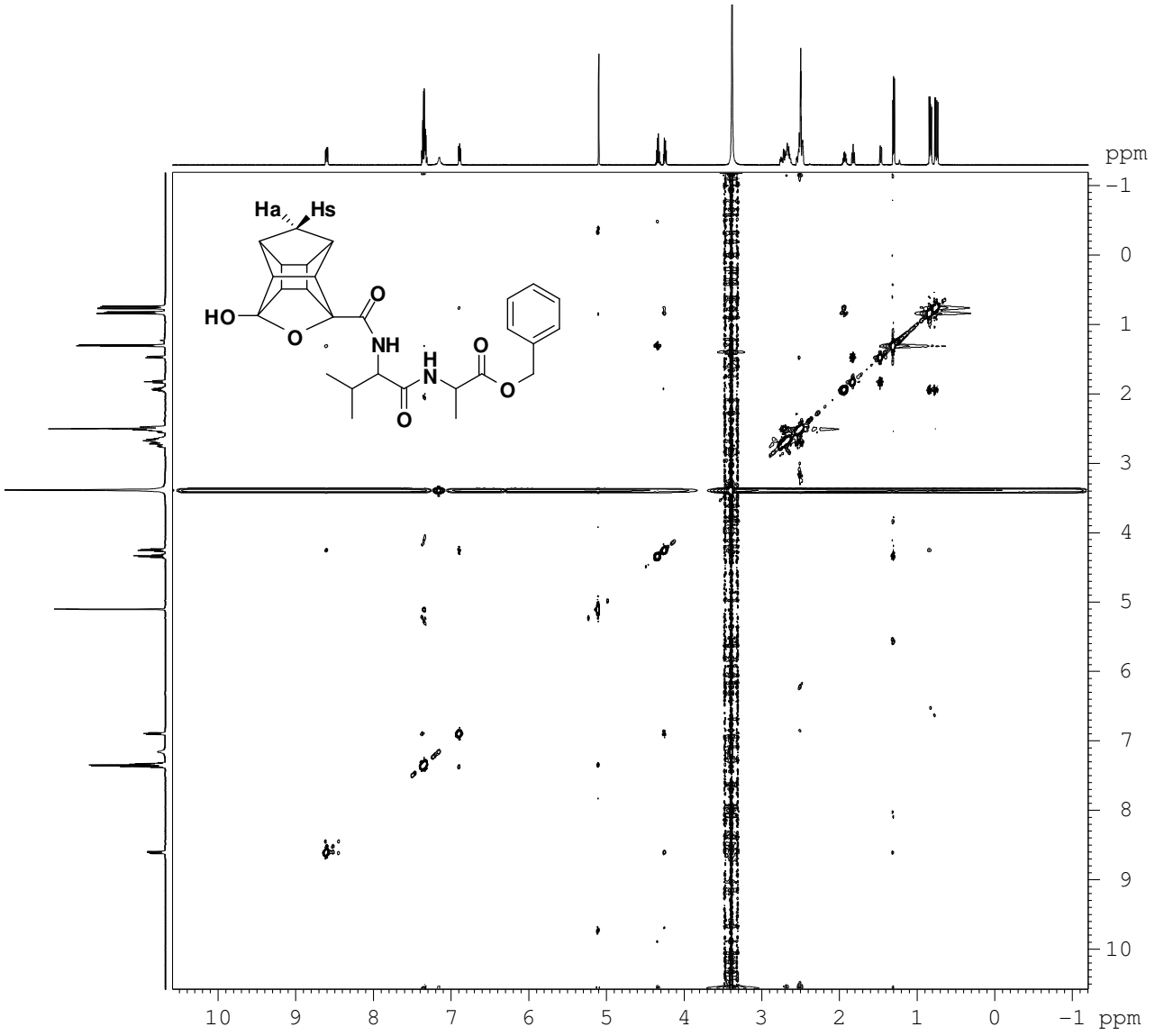

```
Current Data Parameters
NAME                  Raj
EXPNO                 171
PROCNO                  1

F2 - Acquisition Parameters
Date_            20101217
Time                11.07
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG             jmod
TD                 65536
SOLVENT             DMSO
NS                  1600
DS                     4
SWH            36057.691 Hz
FIDRES          0.550197 Hz
AQ             0.9088159 sec
RG                  2050
DW                13.867 usec
DE                 6.50 usec
TE                298.2 K
CNST2         145.0000000
CNST11          1.0000000
D1             2.00000000 sec
D20            0.00689655 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                 13C
P1                  9.00 usec
P2                 18.00 usec
PL1                 1.50 dB
PL1W         59.57236099 W
SFO1         150.9103545 MHz

======== CHANNEL f2 ========
CPDPRG2          waltz16
NUC2                 1H
PCPD2              96.00 usec
PL2                -0.50 dB
PL12               15.40 dB
PL2W         18.65769577 W
PL12W         0.47957665 W
SFO2         600.1024004 MHz

F2 - Processing parameters
SI                 32768
SF           150.8953320 MHz
WDW                  EM
SSB                   0
LB                 1.00 Hz
GB                    0
PC                 1.40
```

^{13}C NMR Spectrum of Compound 21 (Chapter 4) and Compound 3 (Chapter 5)



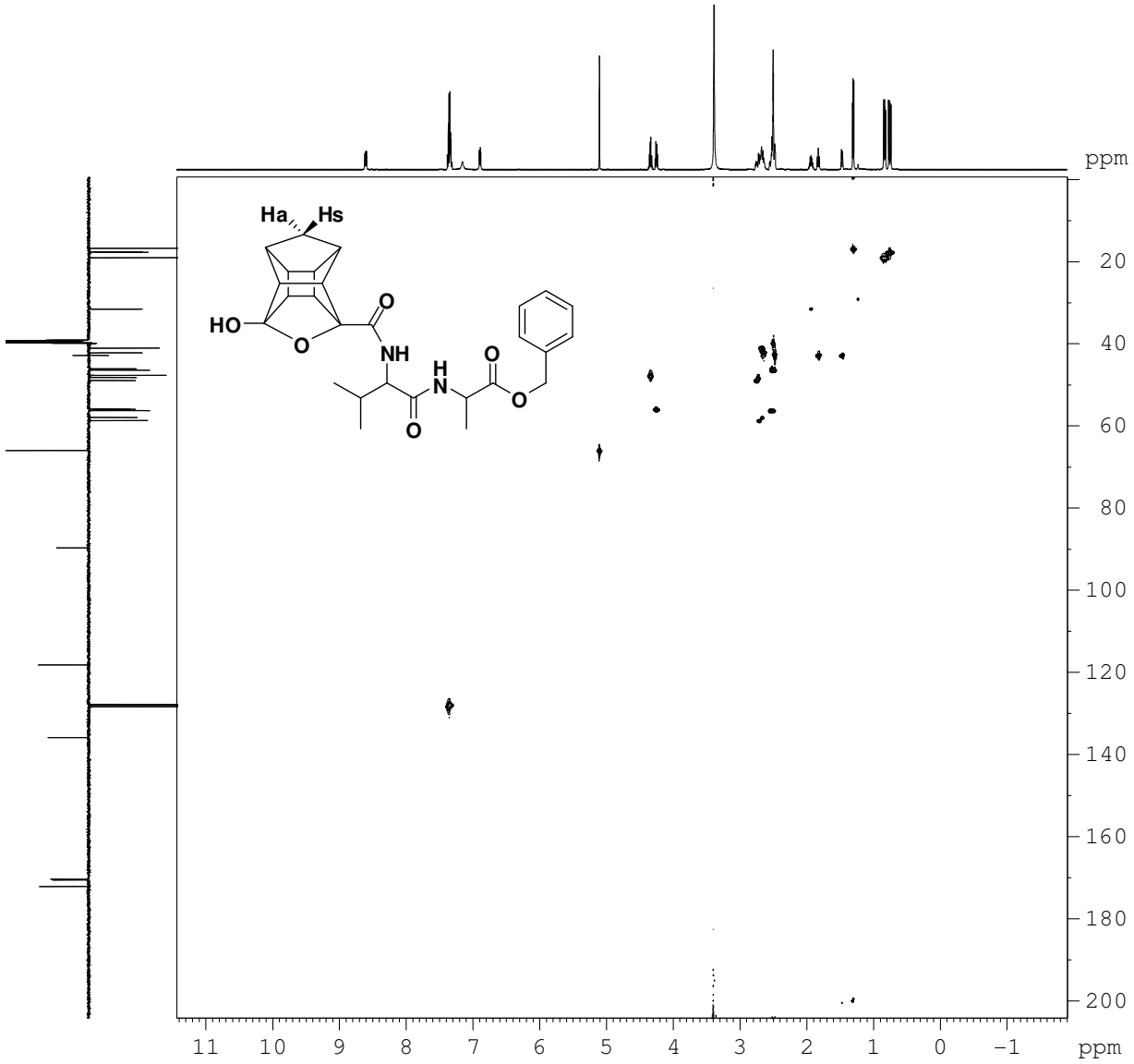
COSY NMR Spectrum of Compound 3 (Chapter 5)

```
Current Data Parameters
NAME                  Raj
EXPNO                 172
PROCNO                  1

F2 - Acquisition Parameters
Date_            20101217
Time                11.07
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG          cosygpqf
TD                  2048
SOLVENT              DMSO
NS                      2
DS                      8
SWH             12335.526 Hz
FIDRES           6.023206 Hz
AQ              0.0830623 sec
RG                     64
DW                 40.533 usec
DE                  6.50 usec
TE                 297.6 K
D0              0.00000300 sec
D1              3.00000000 sec
D13             0.00000400 sec
D16             0.00020000 sec
IN0             0.00008120 sec

======== CHANNEL f1 ========
NUC1                   1H
P0                  13.56 usec
P1                  13.75 usec
PL1                  0.00 dB
PL1W         16.62868881 W
SFO1        600.1037059 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPZ1                10.00 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                     1
TD                    512
SFO1             600.1037 MHz
FIDRES          24.053263 Hz
SW                20.522 ppm
FnMODE                 QF

F2 - Processing parameters
SI                    1024
SF          600.1000000 MHz
WDW                 SINE
SSB                    0
LB                  0.00 Hz
GB                     0
PC                  1.00

F1 - Processing parameters
SI                    1024
MC2                    QF
SF          600.1000000 MHz
WDW                 SINE
SSB                    0
LB                  0.00 Hz
```



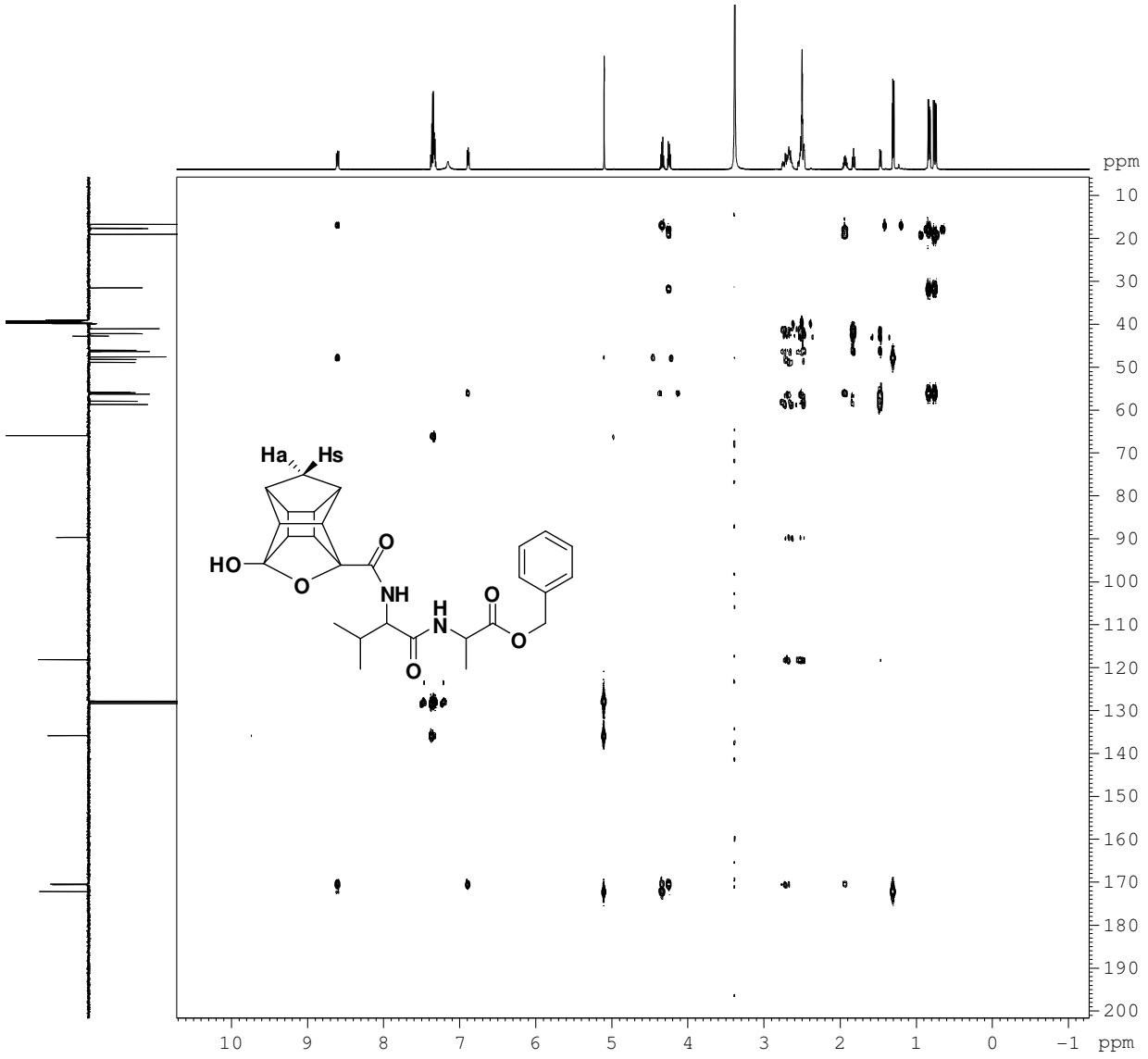

```
Current Data Parameters
NAME                  Raj
EXPNO                 175
PROCNO                  1

F2 - Acquisition Parameters
Date_           20101217
Time               16.14
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG          noesyph
TD                 2048
SOLVENT             DMSO
NS                    4
DS                    4
SWH             7075.472 Hz
FIDRES          3.454820 Hz
AQ             0.1447753 sec
RG                   64
DW                70.667 usec
DE                 6.50 usec
TE                298.2 K
D0             0.00005329 sec
D1             1.00000000 sec
D8             0.30000001 sec
IN0            0.00014160 sec

======== CHANNEL f1 ========
NUC1                  1H
P1                 13.75 usec
PL1                 0.00 dB
PL1W         16.62868881 W
SFO1        600.1028162 MHz

F1 - Acquisition parameters
ND0                   1
TD                  512
SFO1          600.1028 MHz
FIDRES         13.793256 Hz
SW                11.768 ppm
FnMODE       States-TPPI

F2 - Processing parameters
SI                 1024
SF          600.1000000 MHz
WDW                QSINE
SSB                   2
LB                 0.00 Hz
GB                    0
PC                 1.00

F1 - Processing parameters
SI                 1024
MC2          States-TPPI
SF          600.1000000 MHz
WDW                QSINE
SSB                   2
LB                 0.00 Hz
GB                    0
```

NOESY NMR Spectrum of Compound 3 (Chapter 5)



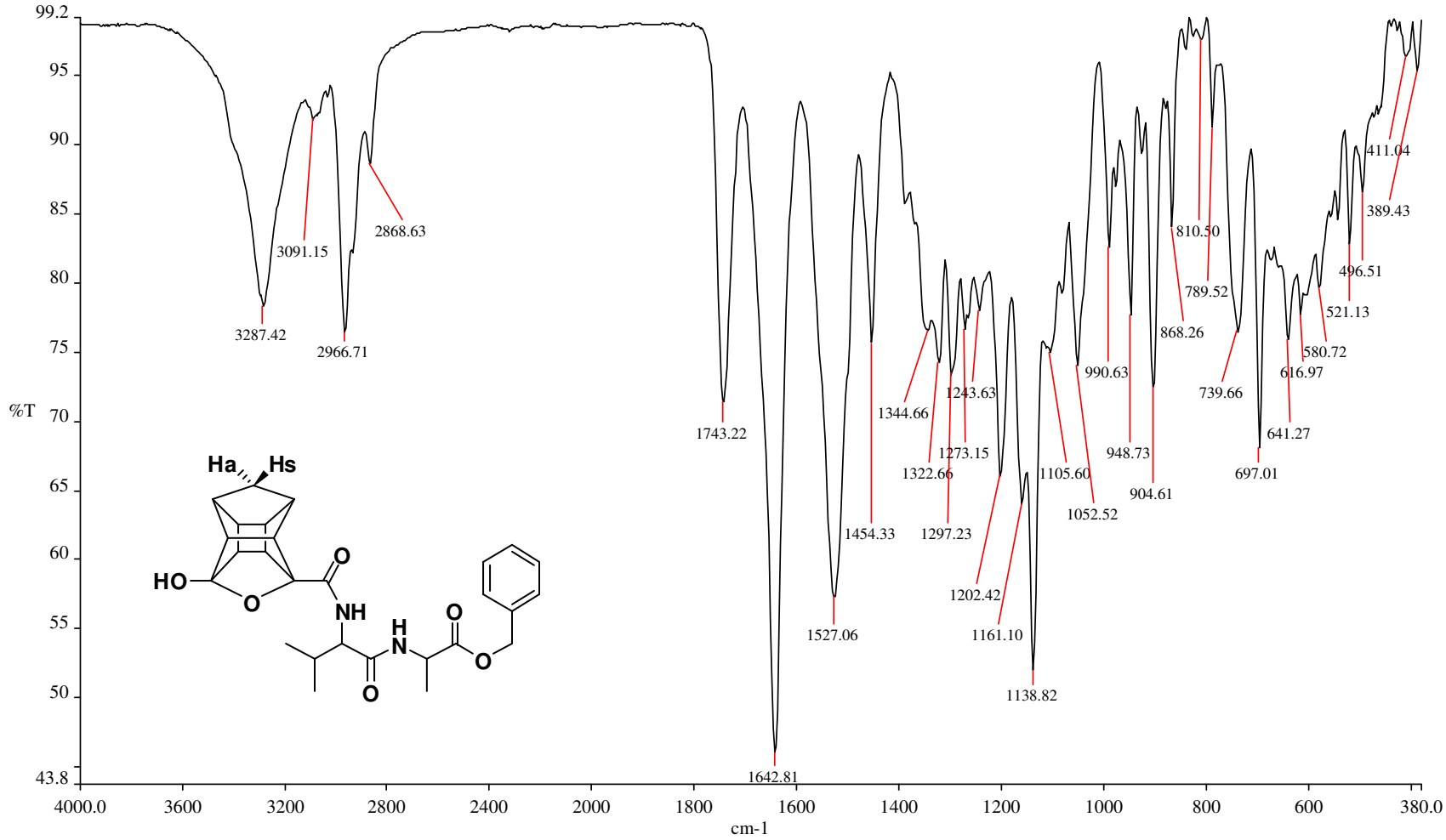

```
Current Data Parameters
NAME                  Raj
EXPNO                 173
PROCNO                  1

F2 - Acquisition Parameters
Date_            20101217
Time                12.02
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG          hsqcetgp
TD                  2048
SOLVENT              DMSO
NS                      4
DS                     16
SWH             12335.526 Hz
FIDRES              6.023206 Hz
AQ               0.0830623 sec
RG                   2050
DW                  40.533 usec
DE                   6.50 usec
TE                  298.3 K
CNST2          145.0000000
D0              0.00000300 sec
D1              1.00000000 sec
D4              0.00172414 sec
D11             0.03000000 sec
D13             0.00000400 sec
D16             0.00020000 sec
IN0             0.00001385 sec
ZGOPTNS

======== CHANNEL f1 ========
NUC1                   1H
P1                  13.51 usec
P2                  27.02 usec
P28               1000.00 usec
PL1                  0.00 dB
PL1W          16.62868881 W
SFO1         600.1037059 MHz

======== CHANNEL f2 ========
CPDPRG2              garp
NUC2                  13C
P3                   9.00 usec
P4                  18.00 usec
PCPD2               60.00 usec
PL2                  1.50 dB
PL12                17.98 dB
PL2W          59.57236099 W
PL12W          1.33981502 W
SFO2         150.9103545 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPNAM2           SINE.100
GPZ1                80.00 %
GPZ2                20.12 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                     2
TD                    512
SFO1            150.9104 MHz
FIDRES         70.426971 Hz
```

HSQC NMR Spectrum of Compound 3 (Chapter 5)




```
Current Data Parameters
NAME                  Raj
EXPNO                 174
PROCNO                  1

F2 - Acquisition Parameters
Date_           20101217
Time                12.40
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG    hmbcgplpndqf
TD                  4096
SOLVENT             DMSO
NS                    16
DS                    16
SWH             9615.385 Hz
FIDRES          2.347506 Hz
AQ              0.2130420 sec
RG                  2050
DW                52.000 usec
DE                 6.50 usec
TE                298.0 K
CNST2         145.0000000
CNST13         10.0000000
D0              0.00000300 sec
D1              2.00000000 sec
D2              0.00344828 sec
D6              0.05000000 sec
D16             0.00020000 sec
IN0             0.00001385 sec

======== CHANNEL f1 ========
NUC1                  1H
P1                 13.51 usec
P2                 27.02 usec
PL1                 0.00 dB
PL1W         16.62868881 W
SFO1         600.1036006 MHz

======== CHANNEL f2 ========
NUC2                 13C
P3                  9.00 usec
PL2                 1.50 dB
PL2W         59.57236099 W
SFO2         150.9103545 MHz

====== GRADIENT CHANNEL =====
GPNAM1          SINE.100
GPNAM2          SINE.100
GPNAM3          SINE.100
GPZ1               60.00 %
GPZ2               20.00 %
GPZ3               60.12 %
P16              1000.00 usec

F1 - Acquisition parameters
ND0                   2
TD                  256
SFO1           150.9104 MHz
FIDRES       140.850357 Hz
SW             238.935 ppm
FnMODE               QF

F2 - Processing parameters
SI                  2048
```

HMBC NMR Spectrum of Compound 3 (Chapter 5)



IR Spectrum of Compound 21 (Chapter 4) and Compound 3 (Chapter 5)

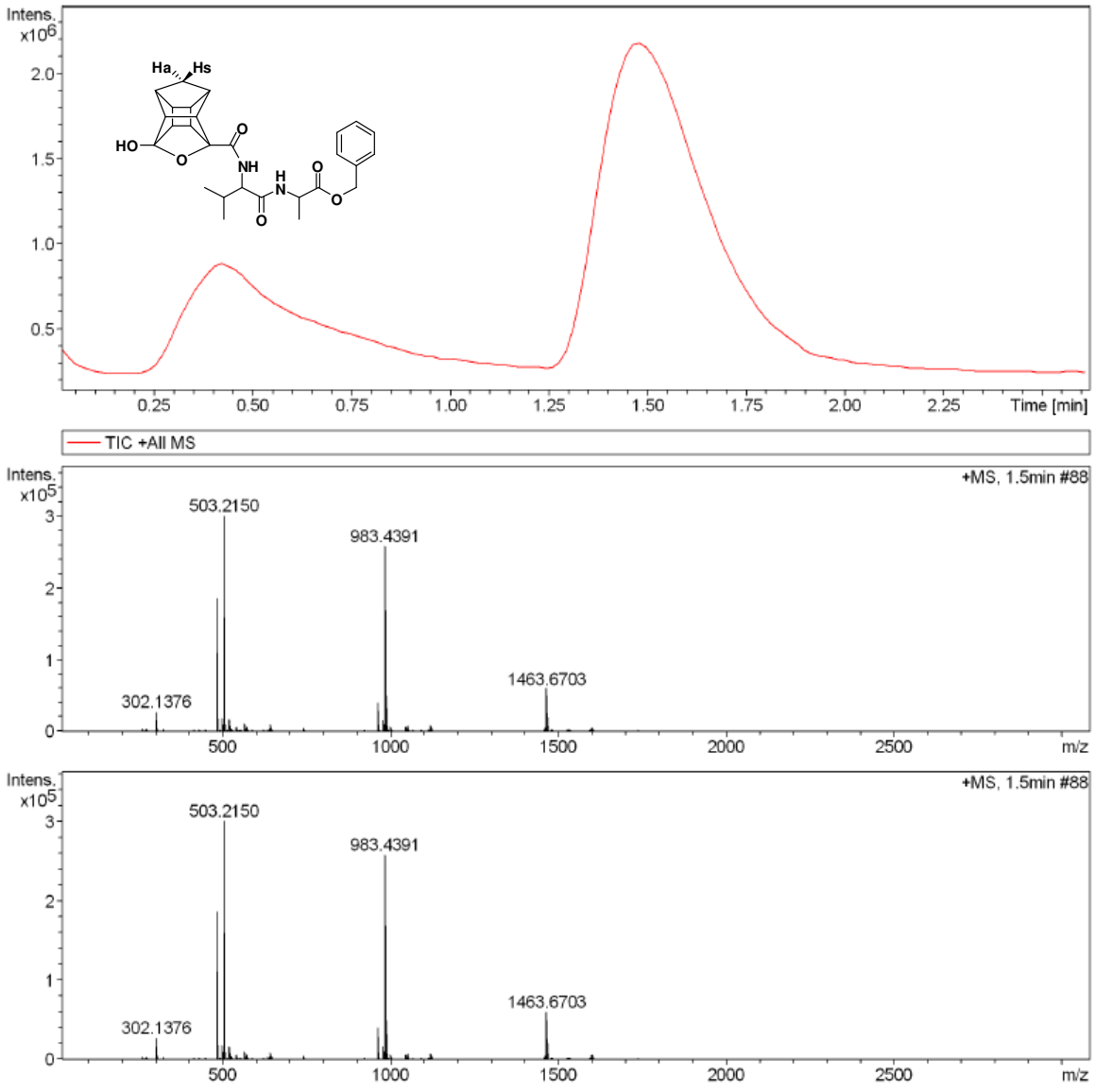
Display Report

Analysis Info

Analysis Name	D:\Data\Raj\OH-PCU-VAL-ALABZ.d	Acquisition Date	12/30/2009 6:05:17 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	OH-PCU-VAL-ALABZ	Instrument	micrOTOF-Q 10139
Comment	480.23		

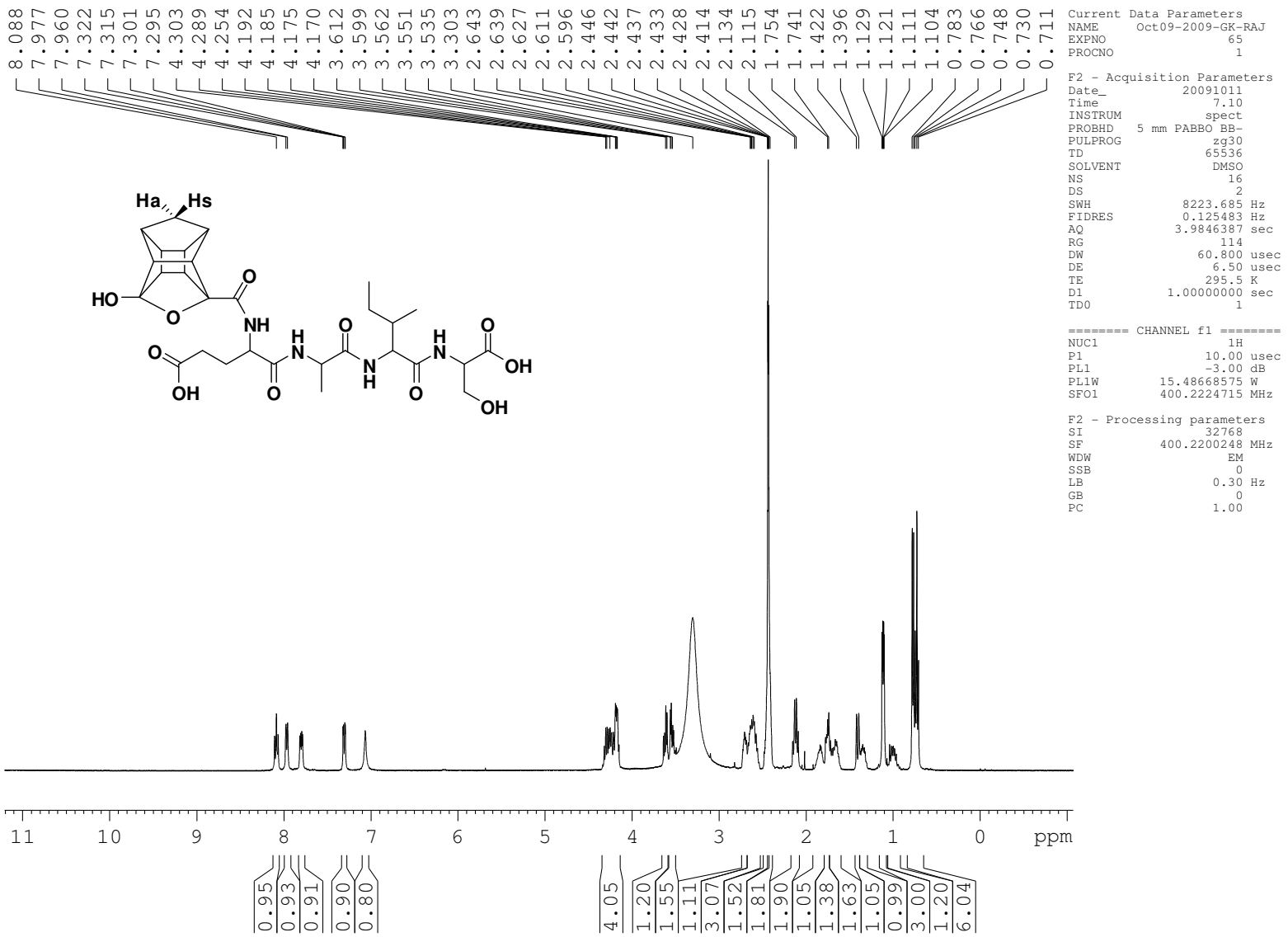
Acquisition Parameter

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



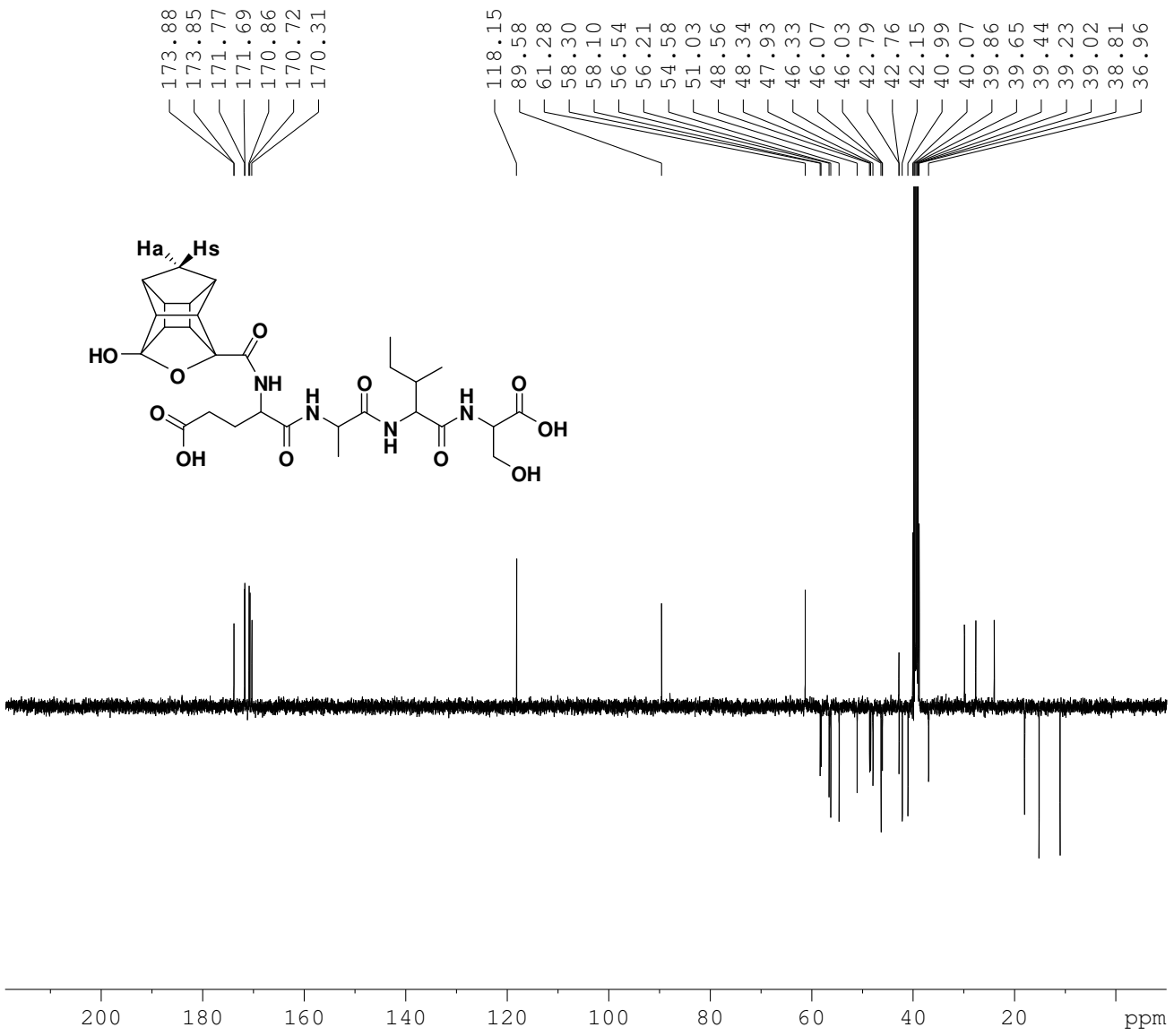
Bruker Compass DataAnalysis 4.0 printed: 12/30/2009 6:10:47 PM Page 1 of 1

HRESIMS (*m/z*) of Compound 20 (Chapter 4) and Compound 1 (Chapter 5)




```
Current Data Parameters
NAME      Oct09-2009-GK-RAJ
EXPNO                   65
PROCNO                   1

F2 - Acquisition Parameters
Date_           20091011
Time                7.10
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG             zg30
TD                 65536
SOLVENT             DMSO
NS                    16
DS                     2
SWH             8223.685 Hz
FIDRES          0.125483 Hz
AQ             3.9846387 sec
RG                   114
DW                60.800 usec
DE                 6.50 usec
TE                295.5 K
D1             1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                  1H
P1                10.00 usec
PL1                -3.00 dB
PL1W         15.48668575 W
SFO1         400.2224715 MHz

F2 - Processing parameters
SI                 32768
SF           400.2200248 MHz
WDW                   EM
SSB                    0
LB                0.30 Hz
GB                     0
PC                1.00
```

^{1}H NMR Spectrum of Compound 22 (Chapter 4)



Current Data Parameters
NAME Oct09-2009-GK-RAJ
EXPNO 62
PROCNO 1

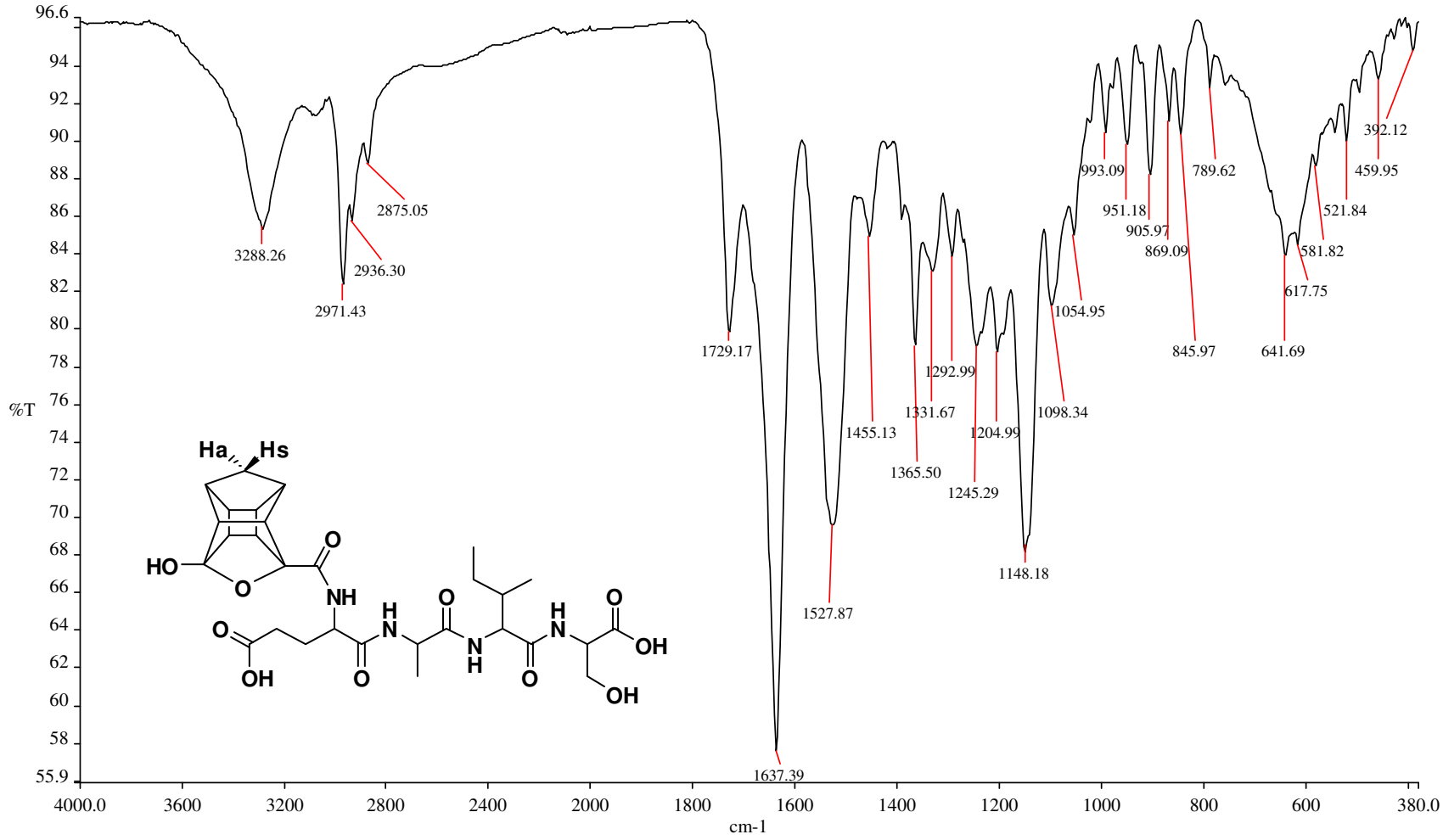
F2 - Acquisition Parameters
Date_ 20091011
Time 5.08
INSTRUM spect
PROBHD 5 mm PABBO BB-
PULPROG jmod
TD 65536
SOLVENT DMSO
NS 1024
DS 4
SWH 24038.461 Hz
FIDRES 0.366798 Hz
AQ 1.3631988 sec
RG 2050
DW 20.800 usec
DE 6.50 usec
TE 296.5 K
CNST2 145.0000000
CNST11 1.0000000
D1 2.00000000 sec
D20 0.00689655 sec
TD0 1

======== CHANNEL f1 ========
NUC1 13C
P1 8.40 usec
P2 16.80 usec
PL1 -2.00 dB
PL1W 54.14257431 W
SFO1 100.6454626 MHz

======== CHANNEL f2 ========
CPDPRG2 waltz16
NUC2 1H
PCPD2 90.00 usec
PL2 -3.00 dB
PL12 15.60 dB
PL2W 15.48668575 W
PL12W 0.21377575 W
SFO2 400.2216009 MHz

F2 - Processing parameters
SI 32768
SF 100.6354493 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40

^{13}C NMR Spectrum of Compound 22 (Chapter 4)



IR Spectrum of Compound 22 (Chapter 4)

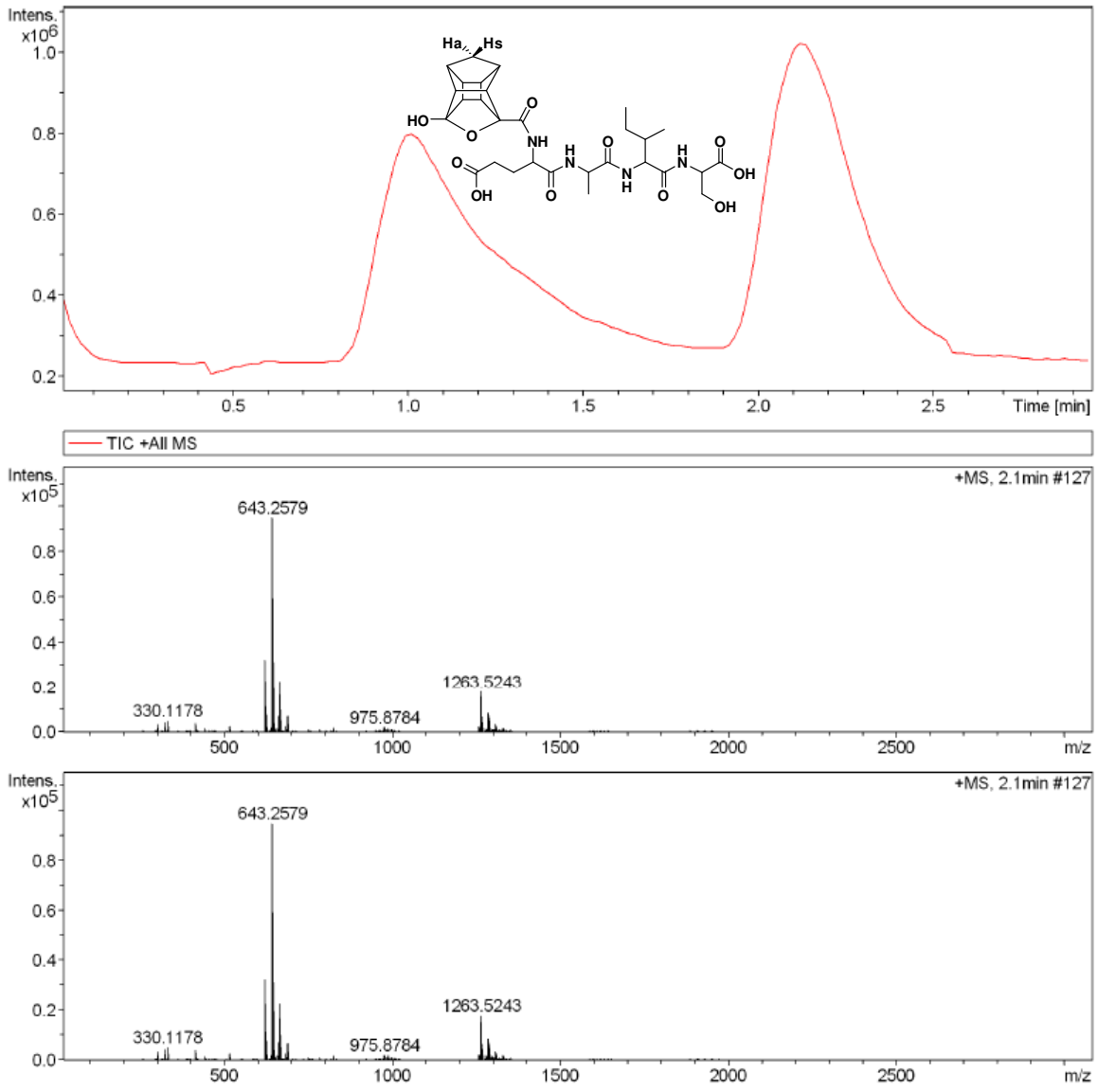
Display Report

Analysis Info

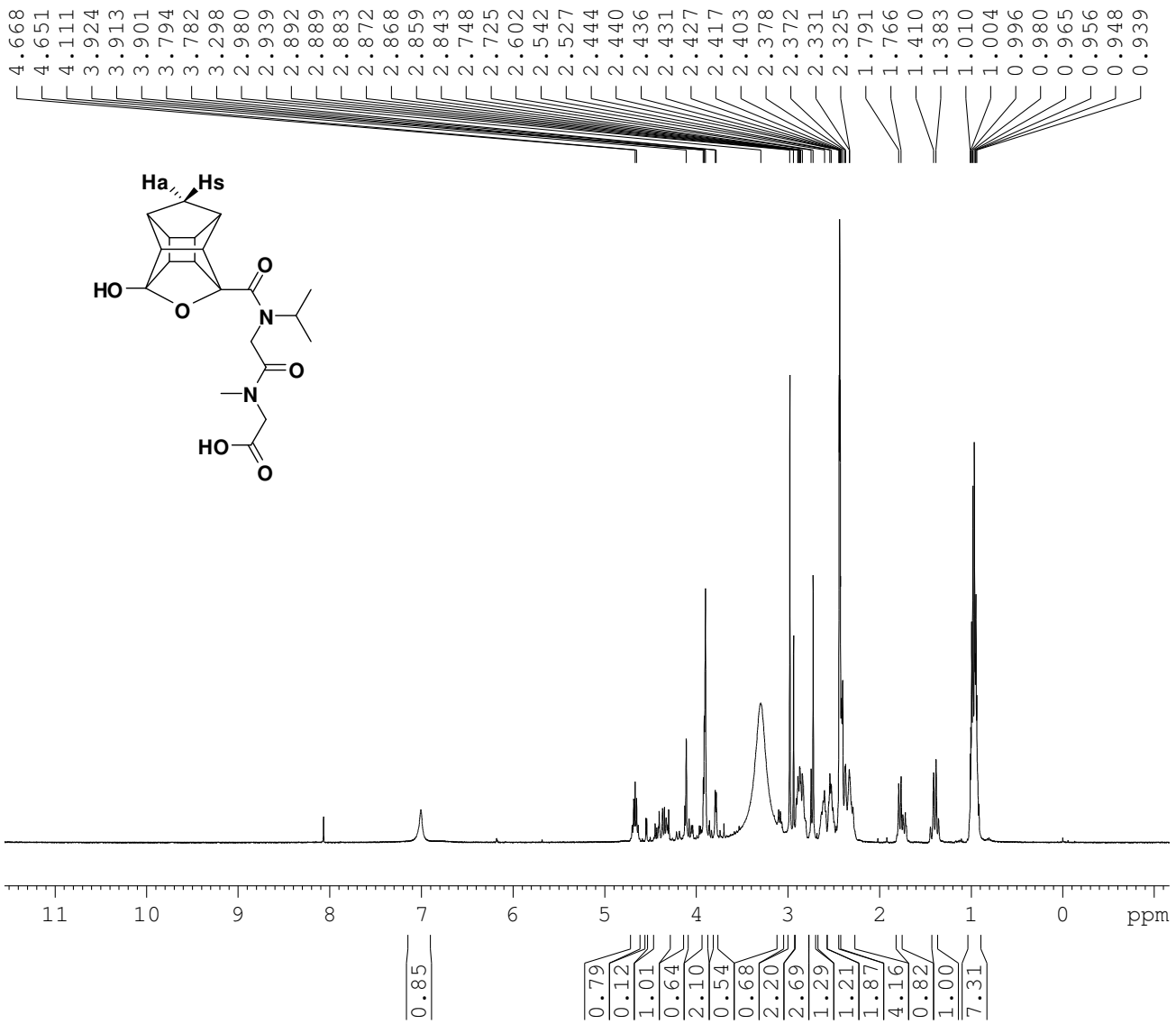
Analysis Name	D:\Data\Raj\OH-PCU-GLU-ALA-ILU-SER.d	Acquisition Date	12/30/2009 6:42:52 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	OH-PCU-GLU-ALA-ILU-SER	Instrument	micrOTOF-Q 10139
Comment	619		

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

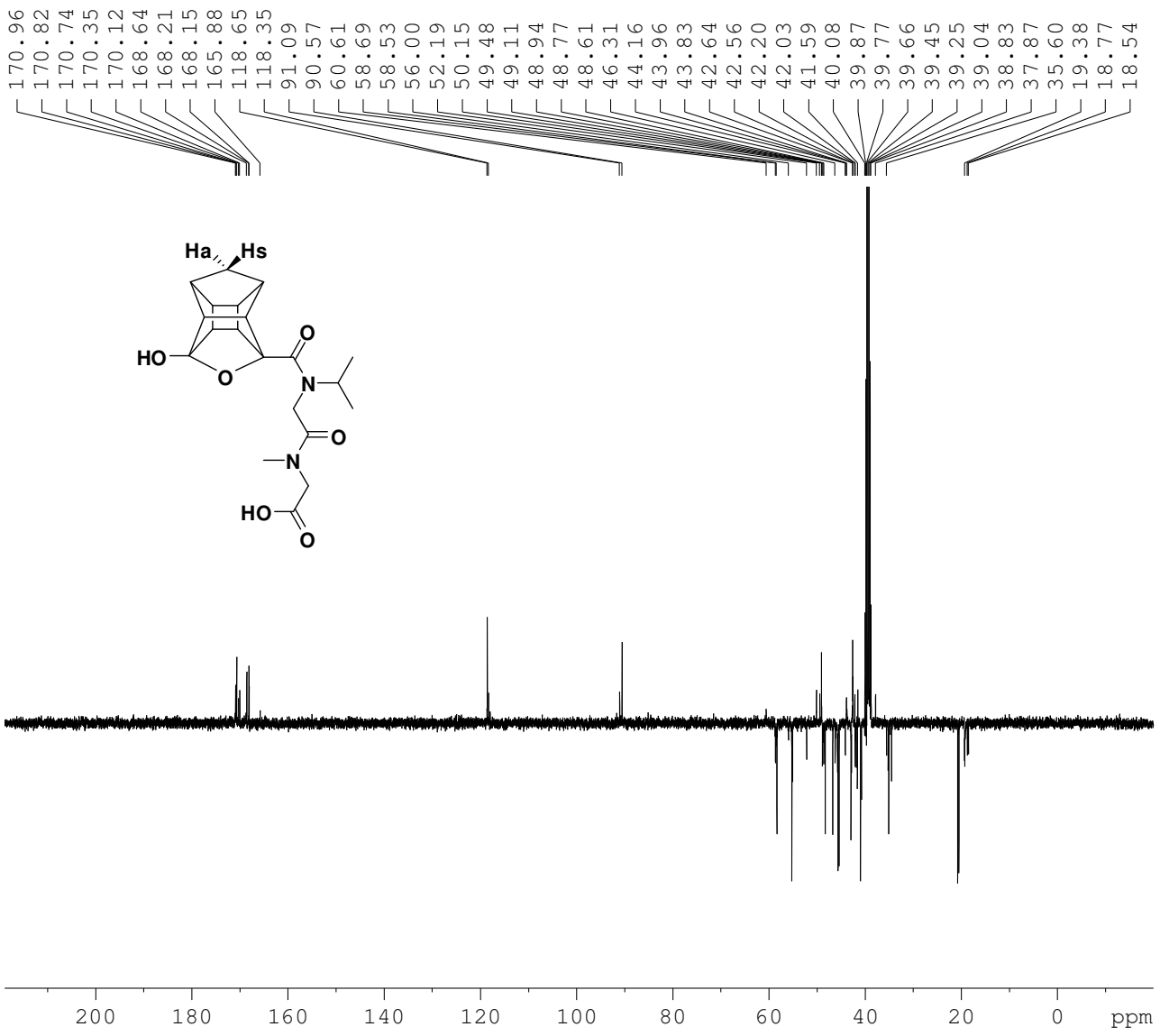


HRESIMS (*m/z*) of Compound 22 (Chapter 4)



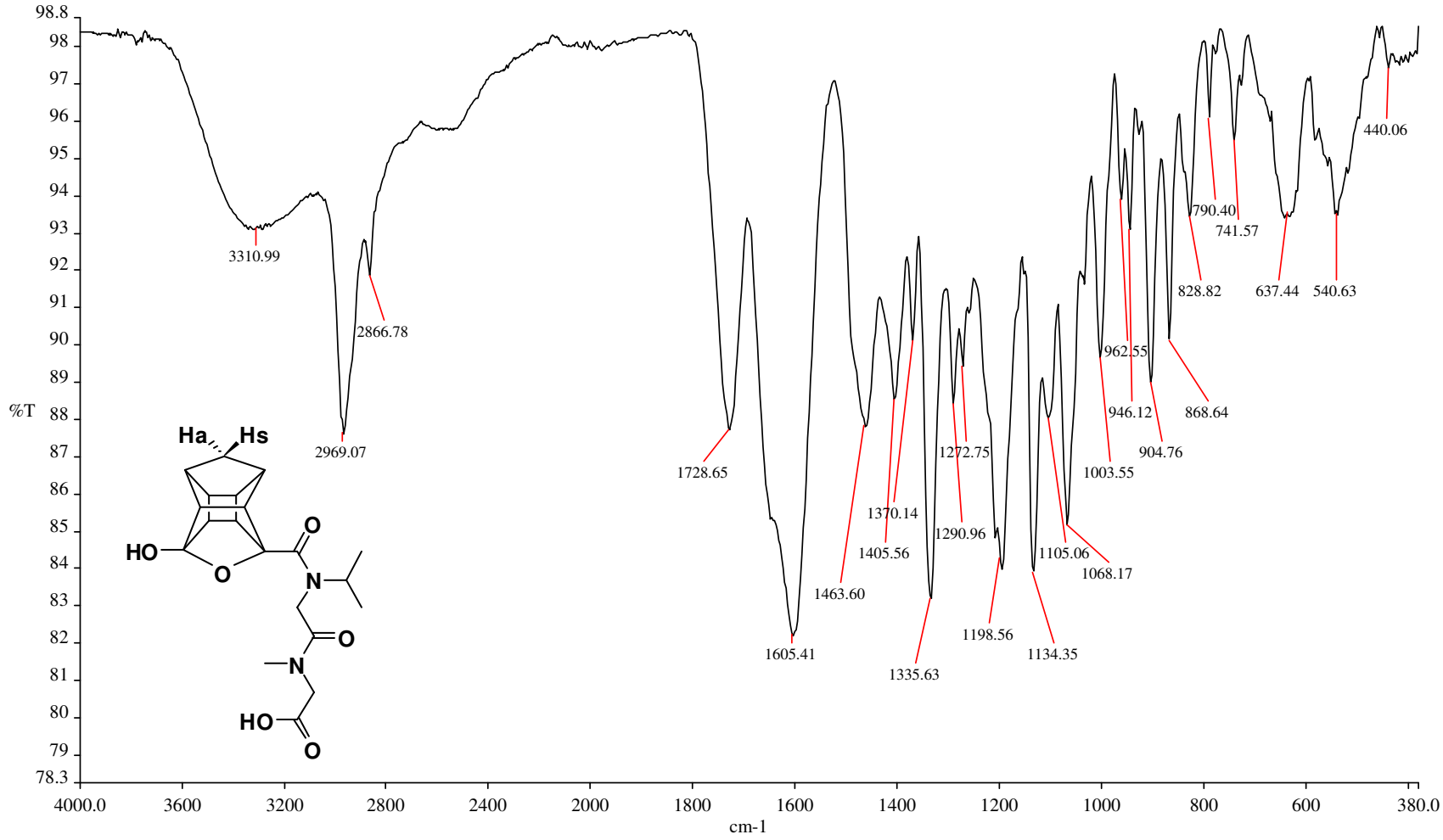

```
Current Data Parameters
NAME      Oct12-2009-GK-Raj
EXPNO                   13
PROCNO                   1

F2 - Acquisition Parameters
Date_           20091013
Time               12.25
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG             zg30
TD                 65536
SOLVENT             DMSO
NS                    16
DS                     2
SWH             8223.685 Hz
FIDRES          0.125483 Hz
AQ             3.9846387 sec
RG                   114
DW                60.800 usec
DE                 6.50 usec
TE                296.8 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                  1H
P1                10.00 usec
PL1               -3.00 dB
PL1W        15.48668575 W
SFO1        400.2224715 MHz

F2 - Processing parameters
SI                 32768
SF          400.2200255 MHz
WDW                   EM
SSB                    0
LB                  0.30 Hz
GB                     0
PC                  1.00
```

^{1}H NMR Spectrum of Compound 23 (Chapter 4)




```
Current Data Parameters
NAME      Oct12-2009-GK-Raj-1
EXPNO                      10
PROCNO                      1

F2 - Acquisition Parameters
Date_              20091013
Time                  10.23
INSTRUM               spect
PROBHD   5 mm PABBO BB-
PULPROG                jmod
TD                    65536
SOLVENT                DMSO
NS                     1024
DS                        4
SWH               24038.461 Hz
FIDRES            0.366798 Hz
AQ                1.3631988 sec
RG                     2050
DW                   20.800 usec
DE                    6.50 usec
TE                   297.6 K
CNST2           145.0000000
CNST11            1.0000000
D1                2.00000000 sec
D20               0.00689655 sec
TD0                        1

======== CHANNEL f1 ========
NUC1                    13C
P1                    8.40 usec
P2                   16.80 usec
PL1                  -2.00 dB
PL1W           54.14257431 W
SFO1           100.6454626 MHz

======== CHANNEL f2 ========
CPDPRG2             waltz16
NUC2                    1H
PCPD2                90.00 usec
PL2                  -3.00 dB
PL12                 15.60 dB
PL2W           15.48668575 W
PL12W           0.21377575 W
SFO2           400.2216009 MHz

F2 - Processing parameters
SI                     32768
SF             100.6354493 MHz
WDW                       EM
SSB                        0
LB                    1.00 Hz
GB                         0
PC                    1.40
```

^{13}C NMR Spectrum of Compound 23 (Chapter 4)



IR Spectrum of Compound 23 (Chapter 4)

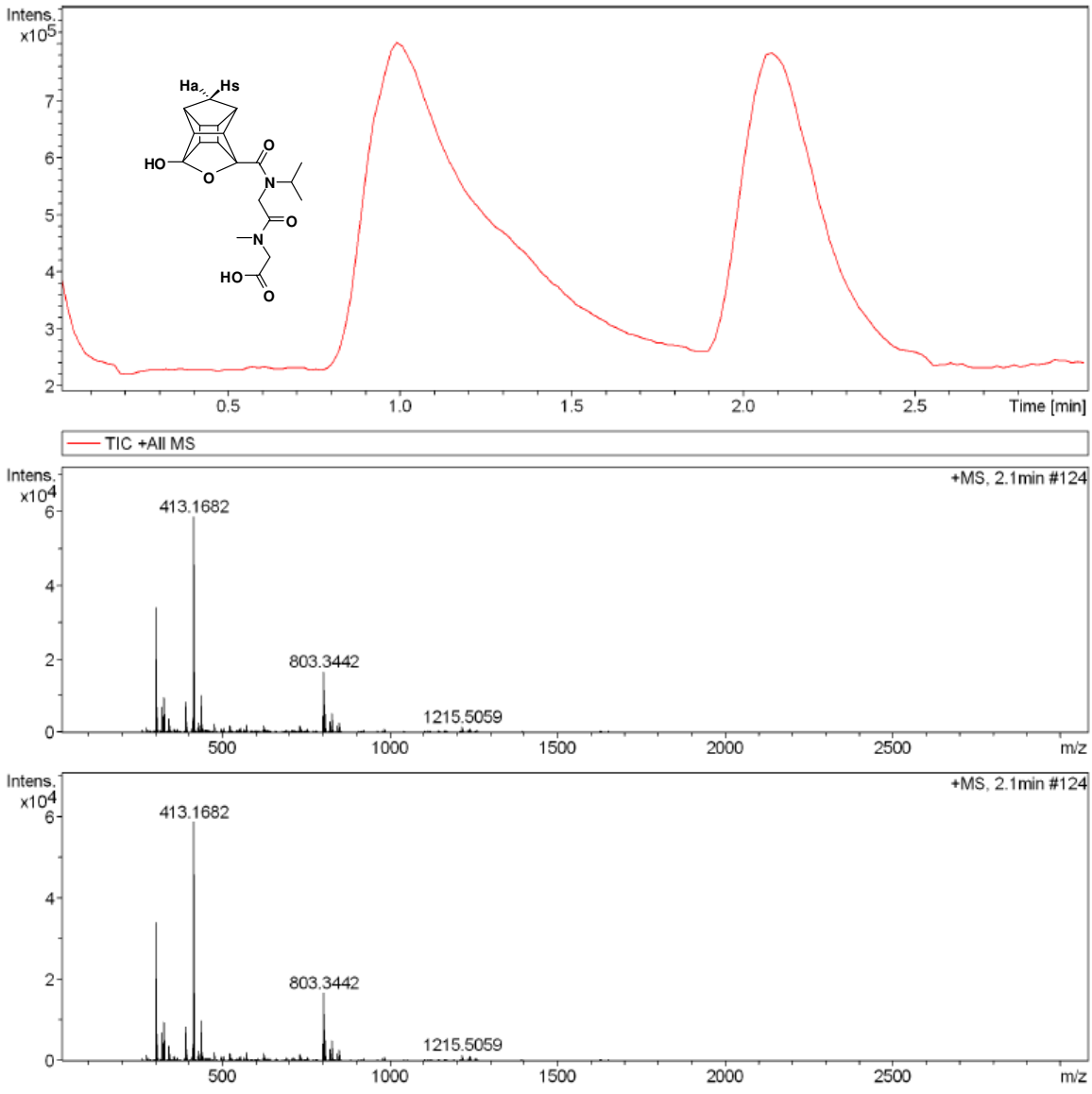
Display Report

Analysis Info

Analysis Name	D:\Data\Raj\OH-PCU-VAL-ALA-PEPTOID1.d	Acquisition Date	12/30/2009 7:02:01 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	OH-PCU-VAL-ALA-PEPTOID1	Instrument	micrOTOF-Q 10139
Comment	390		

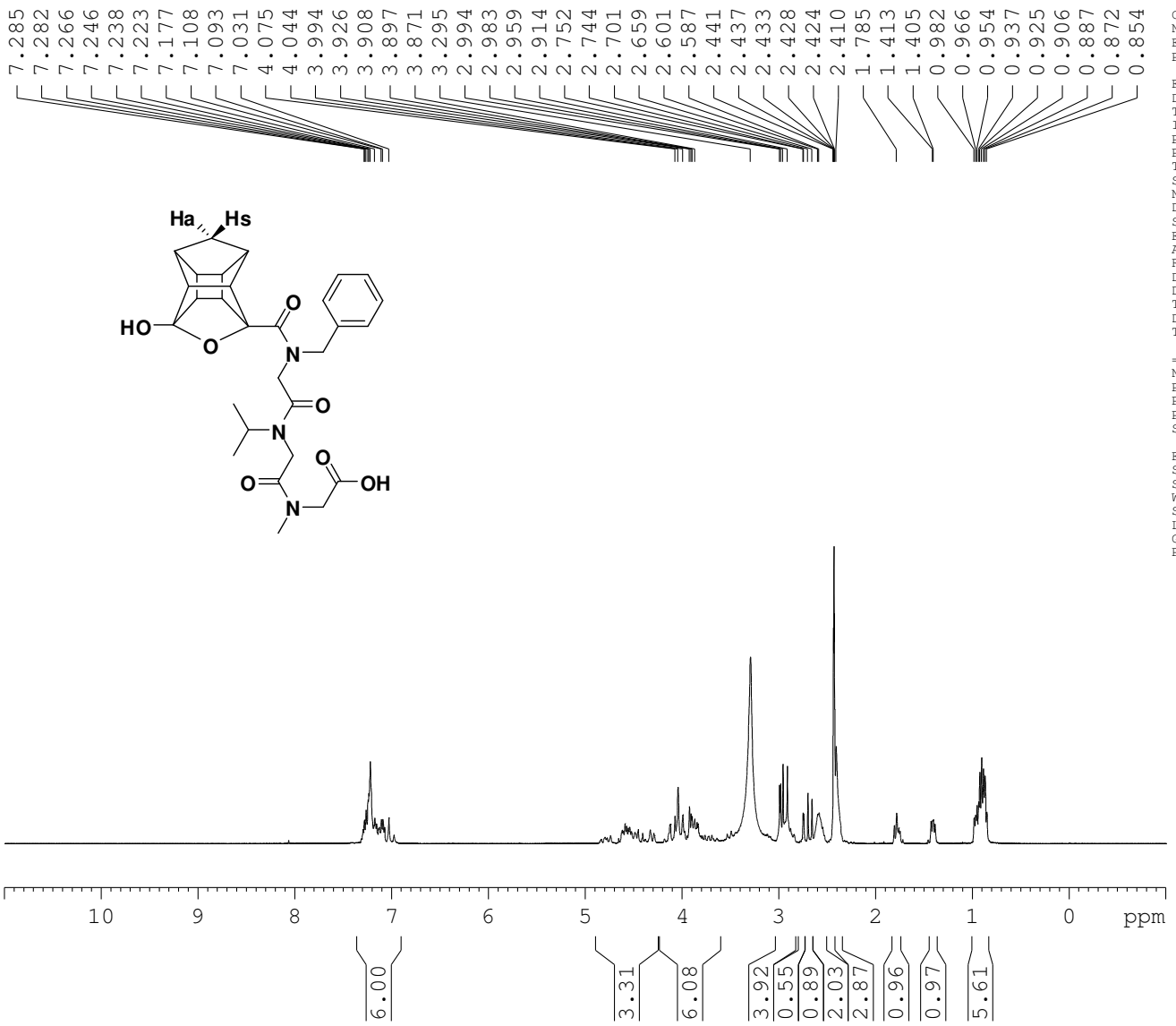
Acquisition Parameter

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



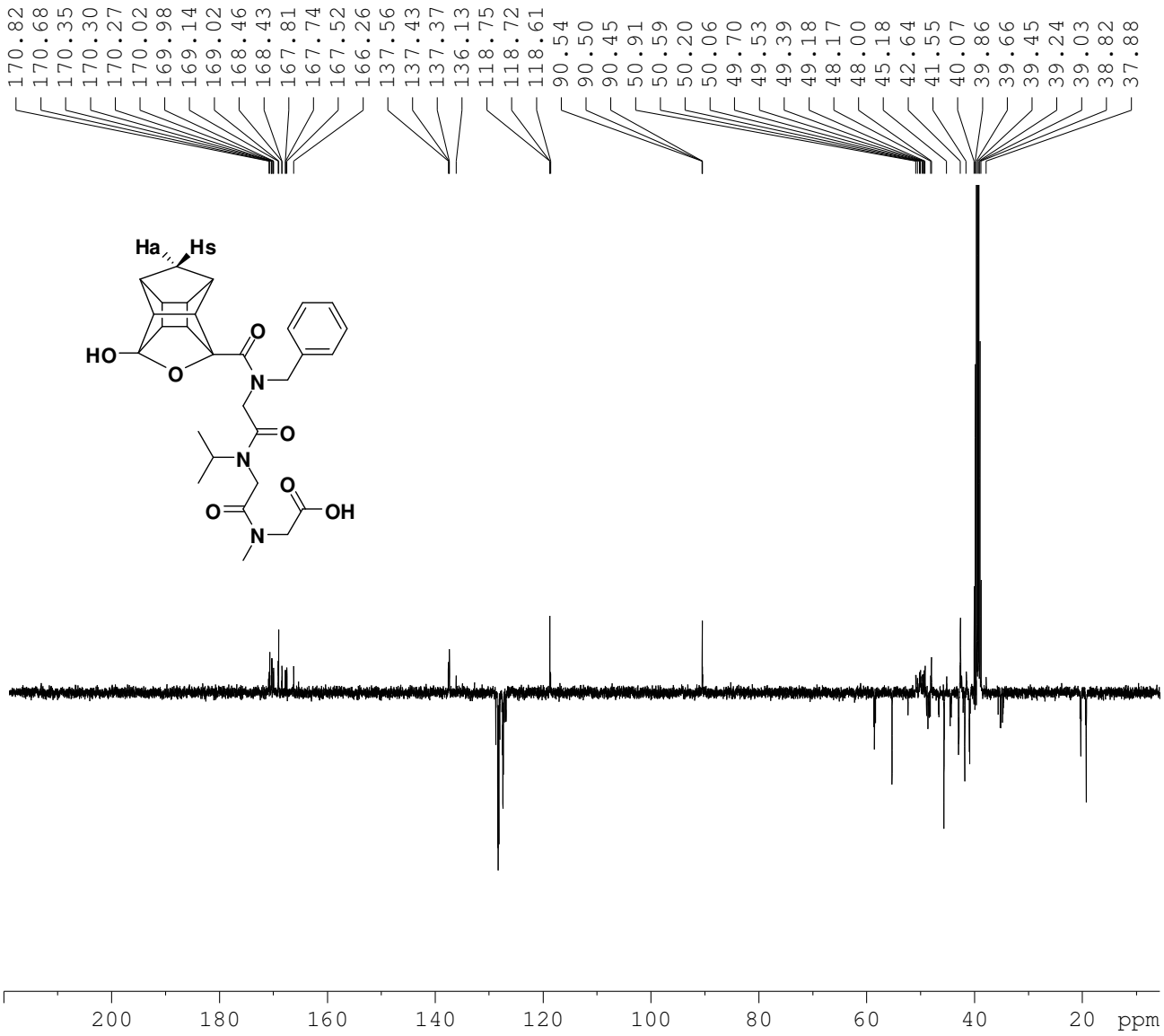
Bruker Compass DataAnalysis 4.0 printed: 12/30/2009 7:06:39 PM Page 1 of 1

HRESIMS (*m/z*) of Compound 23 (Chapter 4)



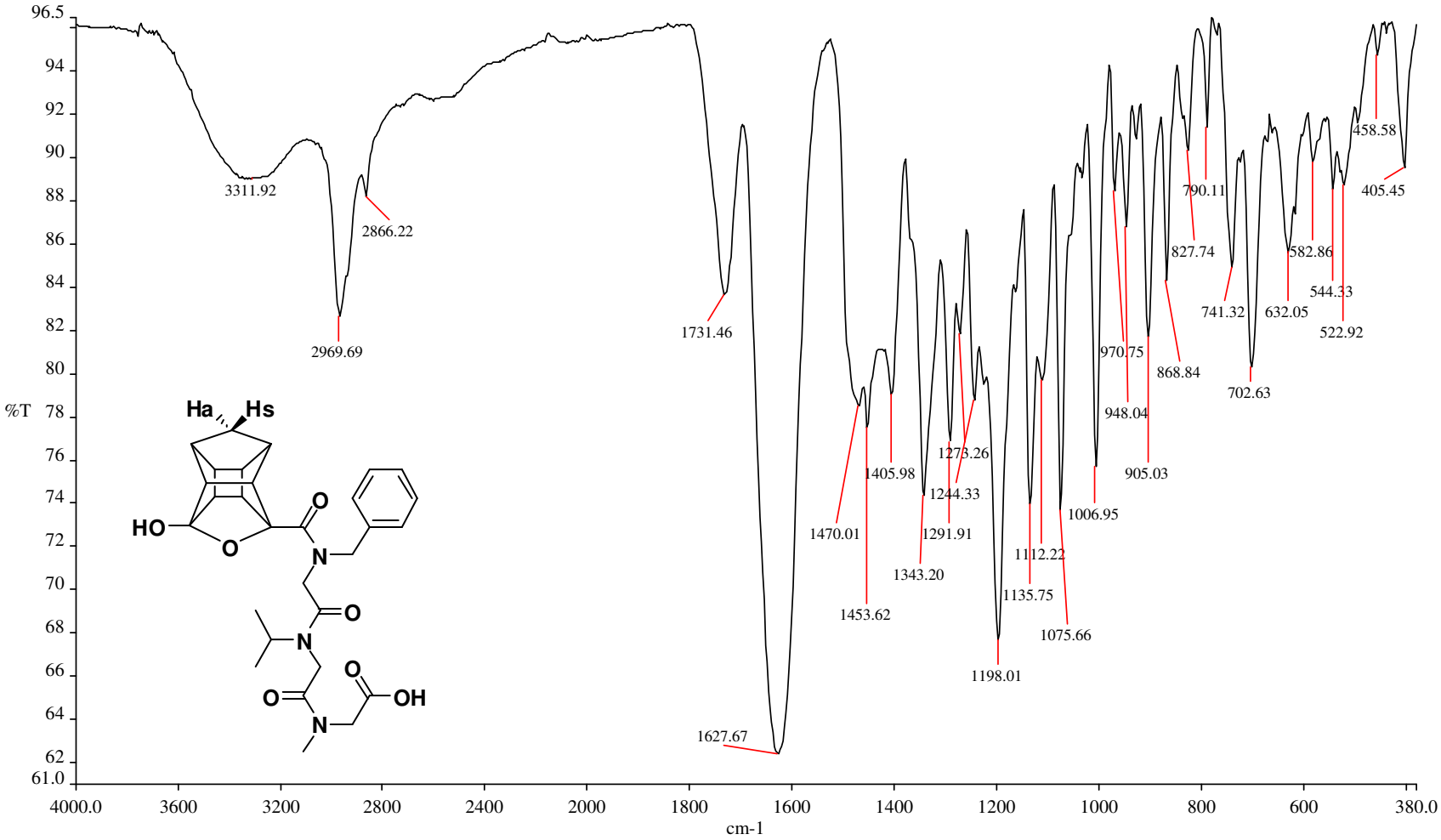

```
Current Data Parameters
NAME      Oct12-2009-GK-Raj-1
EXPNO                      23
PROCNO                      1

F2 - Acquisition Parameters
Date_               20091014
Time                    4.29
INSTRUM                spect
PROBHD   5 mm PABBO BB-
PULPROG                 zg30
TD                     65536
SOLVENT                 DMSO
NS                        16
DS                         2
SWH                 8223.685 Hz
FIDRES              0.125483 Hz
AQ                  3.9846387 sec
RG                        101
DW                     60.800 usec
DE                      6.50 usec
TE                     295.4 K
D1                 1.00000000 sec
TD0                        1

======== CHANNEL f1 ========
NUC1                       1H
P1                     10.00 usec
PL1                    -3.00 dB
PL1W            15.48668575 W
SFO1            400.2224715 MHz

F2 - Processing parameters
SI                         32768
SF              400.2200266 MHz
WDW                        EM
SSB                         0
LB                      0.30 Hz
GB                          0
PC                      1.00
```

^{1}H NMR Spectrum of Compound 24 (Chapter 4)




```
Current Data Parameters
NAME      Oct12-2009-GK-Raj-1
EXPNO                      20
PROCNO                      1

F2 - Acquisition Parameters
Date_            20091014
Time                 2.27
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG              jmod
TD                  65536
SOLVENT              DMSO
NS                   1024
DS                      4
SWH            24038.461 Hz
FIDRES          0.366798 Hz
AQ              1.3631988 sec
RG                   2050
DW                 20.800 usec
DE                  6.50 usec
TE                 296.6 K
CNST2         145.0000000
CNST11          1.0000000
D1              2.00000000 sec
D20             0.00689655 sec
TD0                     1

======== CHANNEL f1 ========
NUC1                   13C
P1                    8.40 usec
P2                   16.80 usec
PL1                  -2.00 dB
PL1W         54.14257431 W
SFO1         100.6454626 MHz

======== CHANNEL f2 ========
CPDPRG2          waltz16
NUC2                   1H
PCPD2                90.00 usec
PL2                  -3.00 dB
PL12                 15.60 dB
PL2W         15.48668575 W
PL12W         0.21377575 W
SFO2         400.2216009 MHz

F2 - Processing parameters
SI                   32768
SF           100.6354493 MHz
WDW                     EM
SSB                      0
LB                    1.00 Hz
GB                       0
PC                    1.40
```

^{13}C NMR Spectrum of Compound 24 (Chapter 4)



IR Spectrum of Compound 24 (Chapter 4)

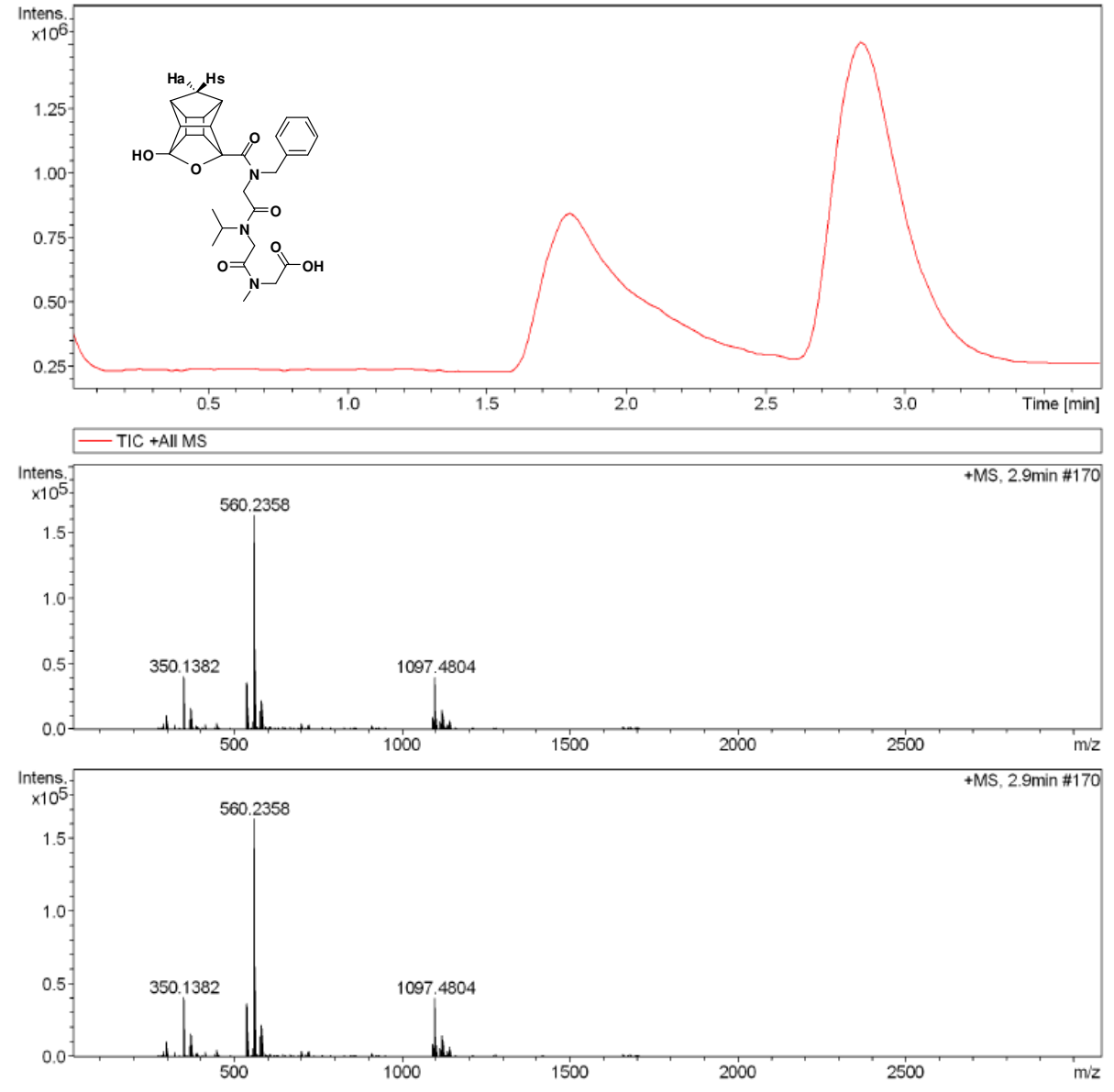
Display Report

Analysis Info

Analysis Name	D:\Data\Raj\OH-PCU-PHE-VAL-ALA,537,PEPTOID.d	Acquisition Date	12/30/2009 5:46:41 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	OH-PCU-PHE-VAL-ALA,537,PEPTOID	Instrument	micrOTOF-Q 10139
Comment	537.25		

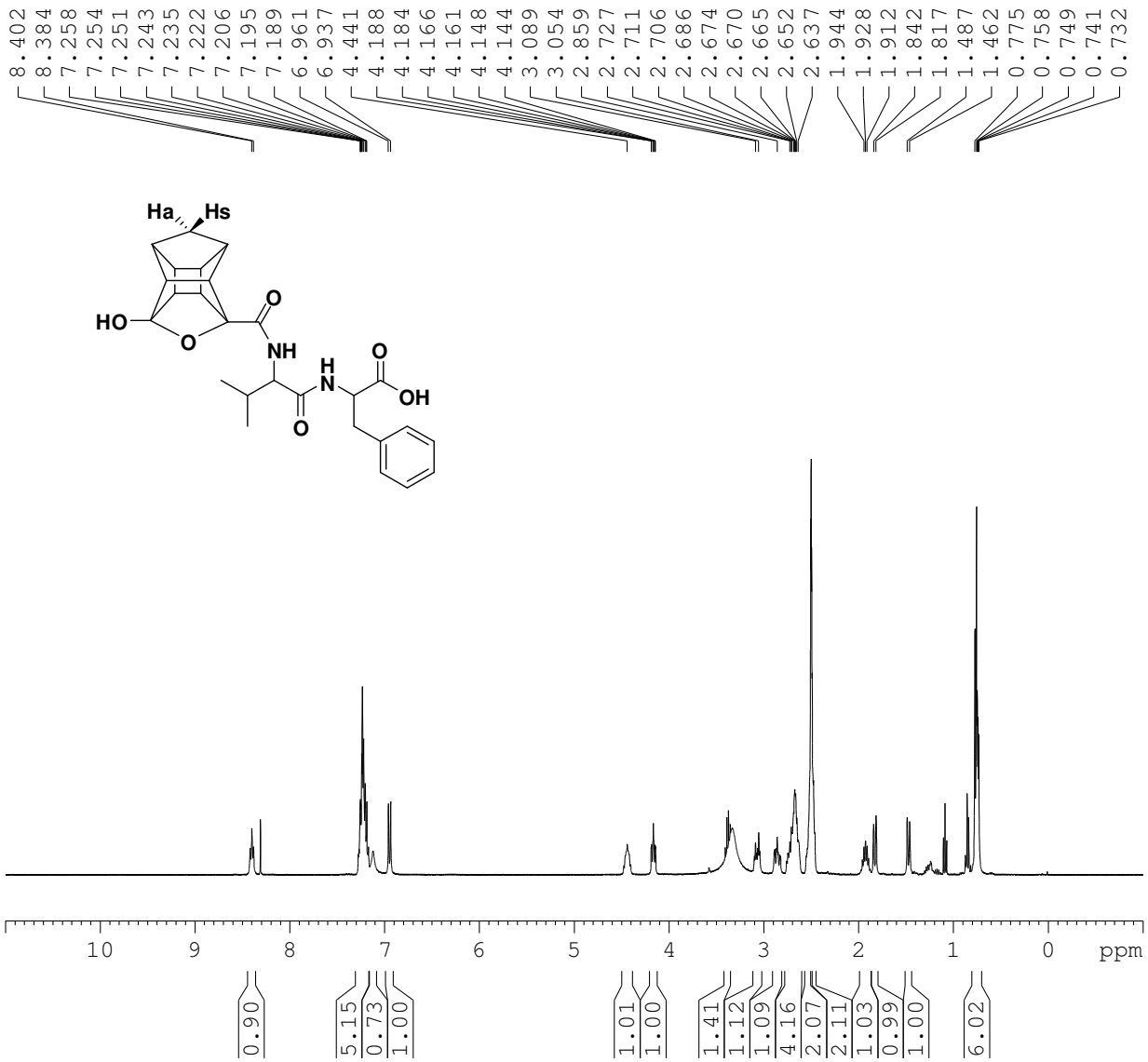
Acquisition Parameter

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



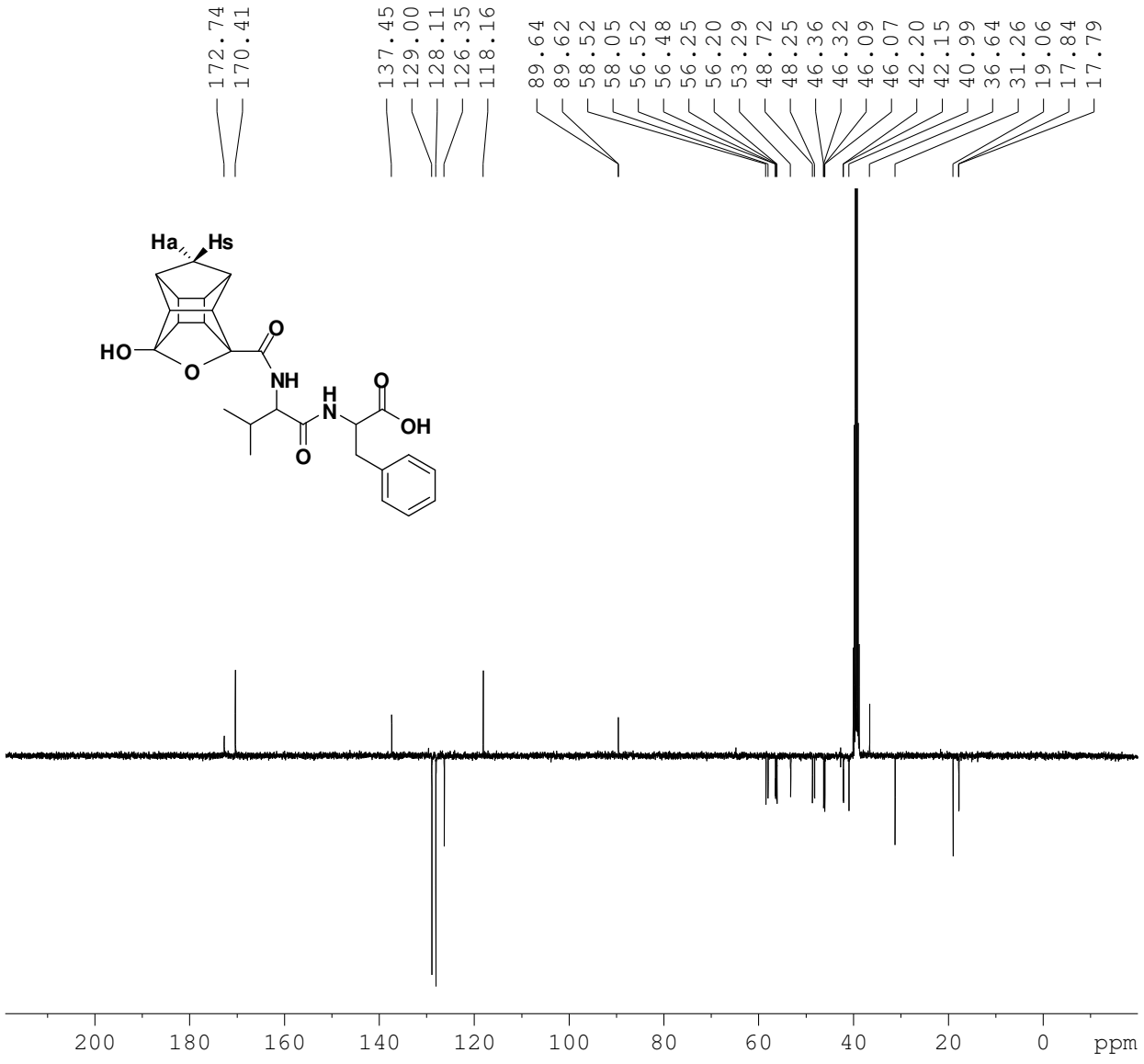
Bruker Compass DataAnalysis 4.0 printed: 12/30/2009 5:54:51 PM Page 1 of 1

HRESIMS (*m/z*) of Compound 24 (Chapter 4)



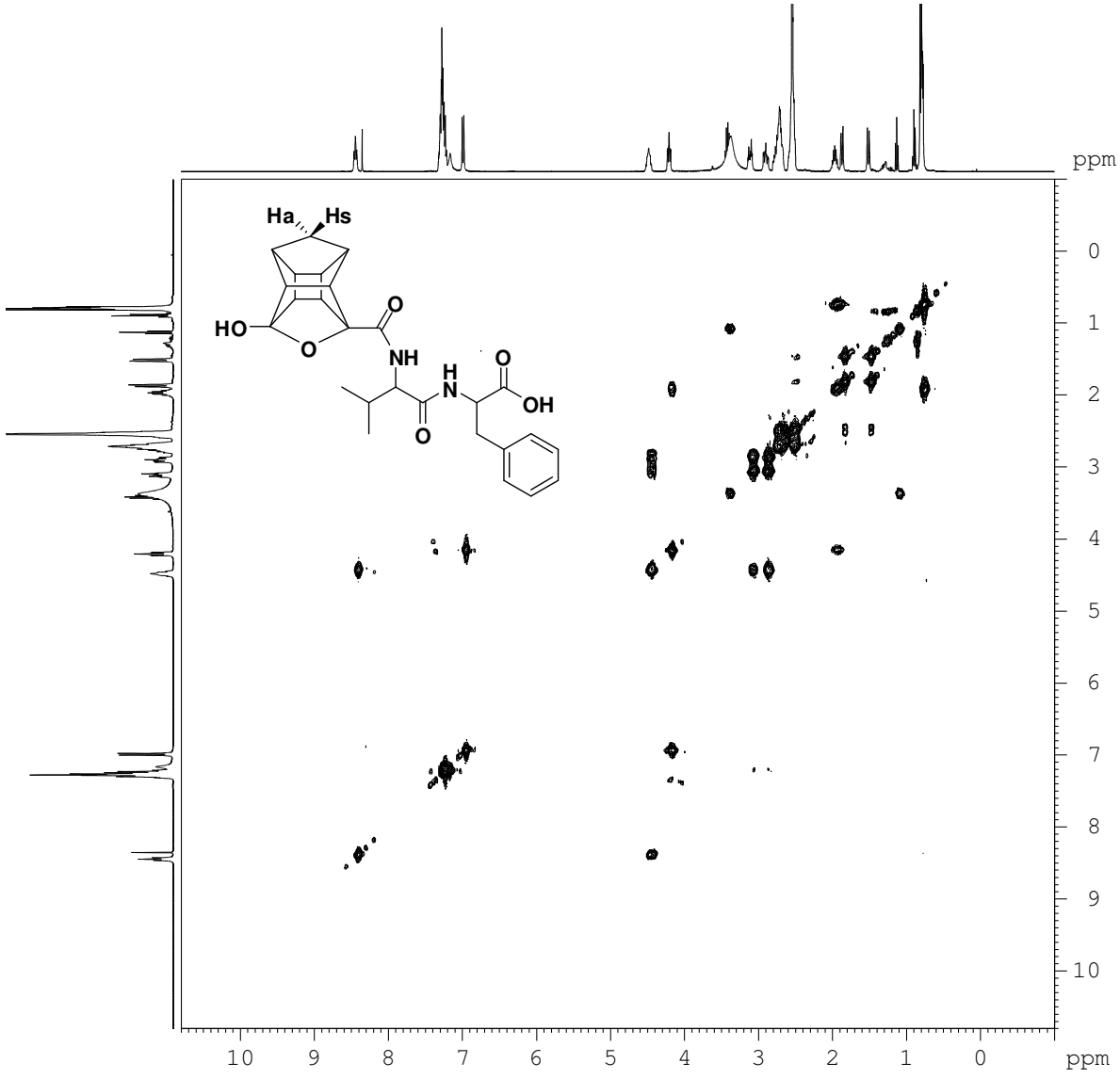

```
Current Data Parameters
NAME      Sep04-2010-GK-Raj-1
EXPNO                  10
PROCNO                  1

F2 - Acquisition Parameters
Date_            20100904
Time                18.17
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG              zg30
TD                  32768
SOLVENT              DMSO
NS                     16
DS                      2
SWH              8223.685 Hz
FIDRES           0.250967 Hz
AQ              1.9923444 sec
RG                    128
DW                 60.800 usec
DE                  6.50 usec
TE                 298.2 K
D1            1.00000000 sec
TD0                     1

======== CHANNEL f1 ========
NUC1                   1H
P1                  10.00 usec
PL1                 -3.00 dB
PL1W         15.48668575 W
SFO1         400.2224715 MHz

F2 - Processing parameters
SI                  16384
SF           400.2199996 MHz
WDW                    EM
SSB                     0
LB                   0.30 Hz
GB                      0
PC                   1.00
```

^{1}H NMR Spectrum of Compound 2 (Chapter 5)



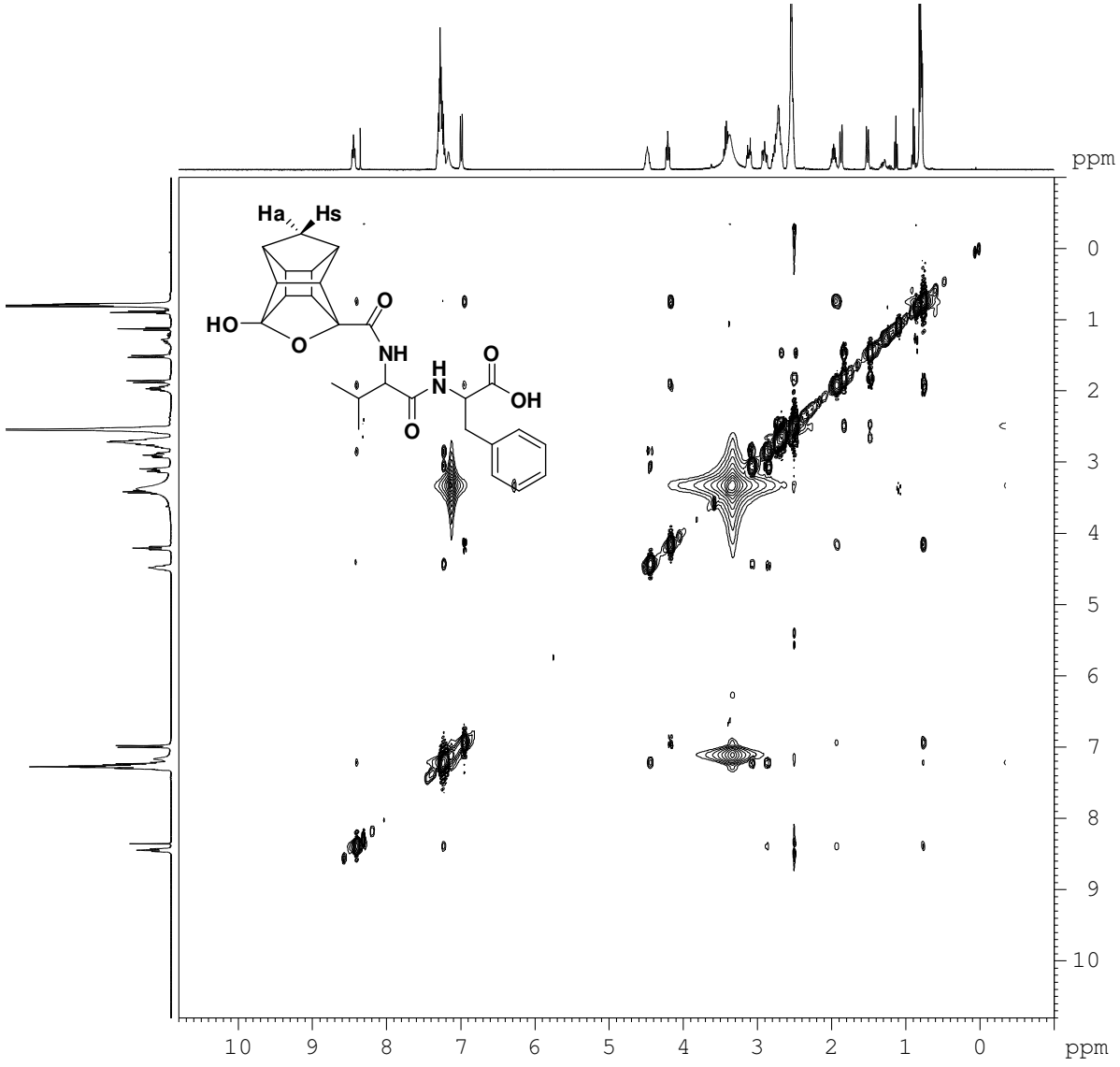

```
Current Data Parameters
NAME      Sep04-2010-GK-Raj-1
EXPNO                      14
PROCNO                      1

F2 - Acquisition Parameters
Date_              20100904
Time                  22.16
INSTRUM               spect
PROBHD   5 mm PABBO BB-
PULPROG                jmod
TD                    65536
SOLVENT                DMSO
NS                     1024
DS                        4
SWH               24038.461 Hz
FIDRES             0.366798 Hz
AQ                1.3631988 sec
RG                     2050
DW                   20.800 usec
DE                    6.50 usec
TE                   298.2 K
CNST2           145.0000000
CNST11            1.0000000
D1             2.00000000 sec
D20            0.00689655 sec
TD0                       1

======== CHANNEL f1 ========
NUC1                    13C
P1                     8.40 usec
P2                    16.80 usec
PL1                   -2.00 dB
PL1W           54.14257431 W
SFO1           100.6454626 MHz

======== CHANNEL f2 ========
CPDPRG2            waltz16
NUC2                    1H
PCPD2                 90.00 usec
PL2                   -3.00 dB
PL12                  15.60 dB
PL2W           15.48668575 W
PL12W           0.21377575 W
SFO2           400.2216009 MHz

F2 - Processing parameters
SI                     32768
SF             100.6354493 MHz
WDW                       EM
SSB                        0
LB                      1.00 Hz
GB                         0
PC                      1.40
```

^{13}C NMR Spectrum of Compound 2 (Chapter 5)



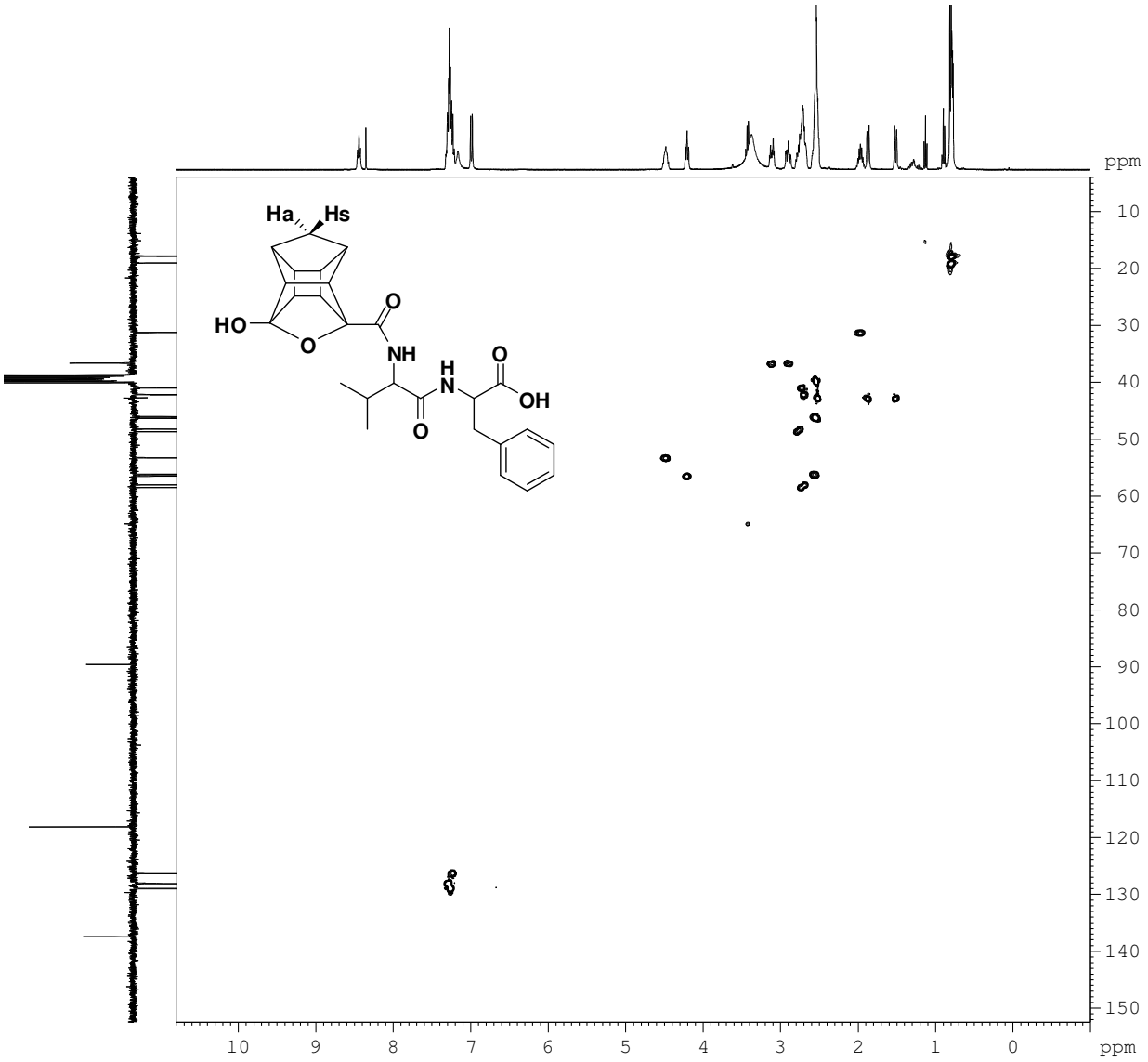

```
Current Data Parameters
NAME      Sep04-2010-GK-Raj-1
EXPNO                  16
PROCNO                  1

F2 - Acquisition Parameters
Date_           20100904
Time               22.20
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG       cosygpmfqf
TD                 2048
SOLVENT             DMSO
NS                     4
DS                     8
SWH             5813.954 Hz
FIDRES          2.838845 Hz
AQ             0.1761780 sec
RG                  2050
DW                86.000 usec
DE                 6.50 usec
TE                298.2 K
D0           0.00000300 sec
D1           2.01556492 sec
D13          0.00000400 sec
D16          0.00020000 sec
IN0          0.00017200 sec

======== CHANNEL f1 ========
NUC1                1H
P1                 10.00 usec
PL1                -3.00 dB
PL1W         15.48668575 W
SFO1         400.2227616 MHz

====== GRADIENT CHANNEL =====
GPNAM1          SINE.100
GPNAM2          SINE.100
GPNAM3          SINE.100
GPZ1               16.00 %
GPZ2               12.00 %
GPZ3               40.00 %
P16              1000.00 usec

F1 - Acquisition parameters
ND0                    1
TD                   256
SFO1           400.2228 MHz
FIDRES         22.710756 Hz
SW               14.527 ppm
FnMODE                QF

F2 - Processing parameters
SI                  1024
SF           400.2199980 MHz
WDW                 SINE
SSB                    0
LB                  0.00 Hz
GB                     0
PC                  1.40

F1 - Processing parameters
SI                  1024
MC2                   QF
SF           400.2200020 MHz
```

COSY NMR Spectrum of Compound 2 (Chapter 5)



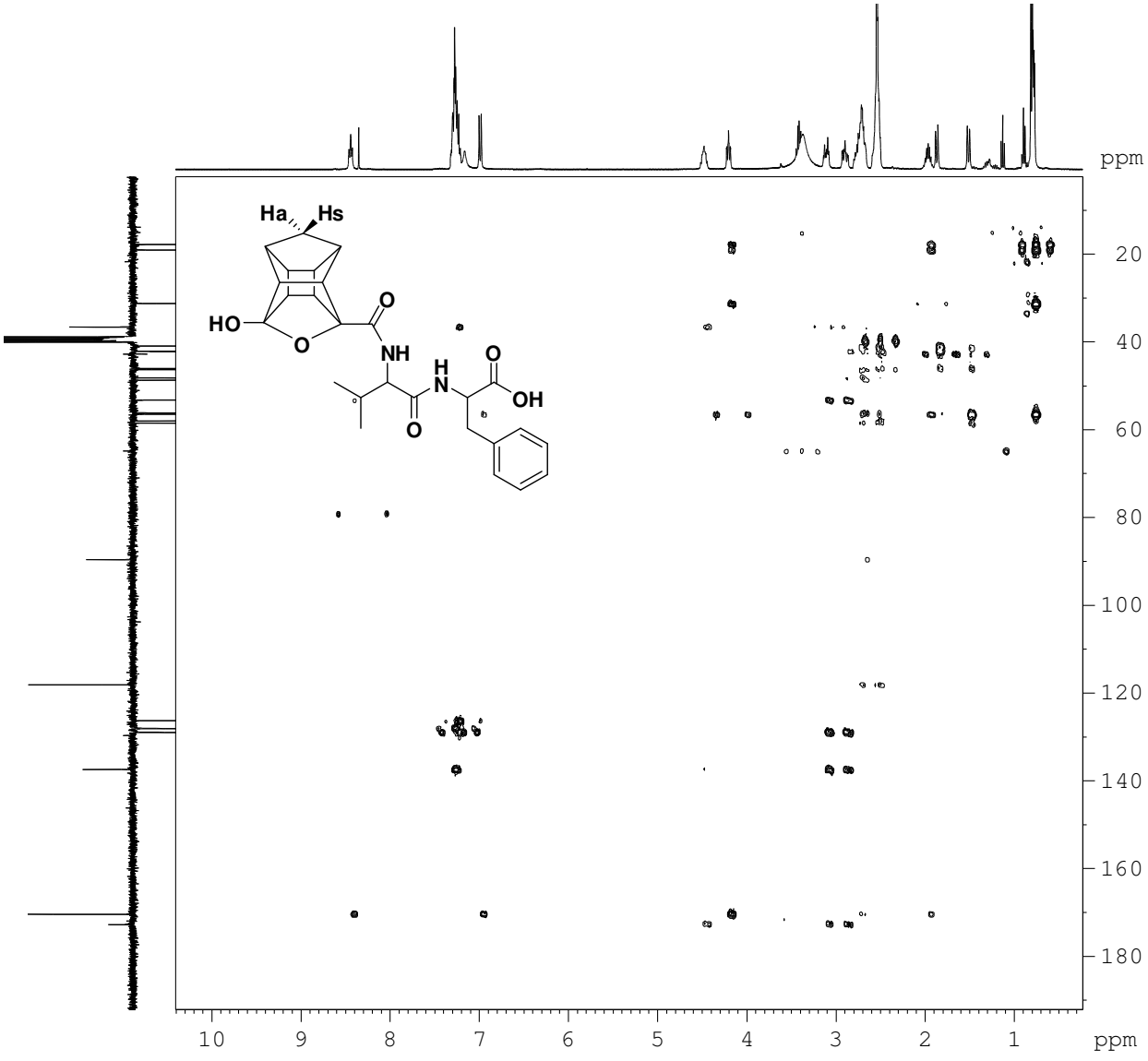

```
Current Data Parameters
NAME      Sep04-2010-GK-Raj-1
EXPNO                      17
PROCNO                      1

F2 - Acquisition Parameters
Date_               20100904
Time                   23.00
INSTRUM                spect
PROBHD    5 mm PABBO BB-
PULPROG             noesyph
TD                     2048
SOLVENT                DMSO
NS                       16
DS                        4
SWH                5813.954 Hz
FIDRES             2.838845 Hz
AQ                 0.1761780 sec
RG                        64
DW                   86.000 usec
DE                    6.50 usec
TE                   298.2 K
D0                 0.00007327 sec
D1                 2.07454705 sec
D8                 0.30000001 sec
IN0                0.00017200 sec

======== CHANNEL f1 ========
NUC1                       1H
P1                      10.00 usec
PL1                     -3.00 dB
PL1W              15.48668575 W
SFO1             400.2227616 MHz

F1 - Acquisition parameters
ND0                        1
TD                       256
SFO1              400.2228 MHz
FIDRES           22.710756 Hz
SW                  14.527 ppm
FnMODE         States-TPPI

F2 - Processing parameters
SI                      1024
SF               400.2199980 MHz
WDW                    QSINE
SSB                        2
LB                      0.00 Hz
GB                         0
PC                      1.00

F1 - Processing parameters
SI                      1024
MC2              States-TPPI
SF               400.2200010 MHz
WDW                    QSINE
SSB                        2
LB                      0.00 Hz
GB                         0
```

NOESY NMR Spectrum of Compound 2 (Chapter 5)



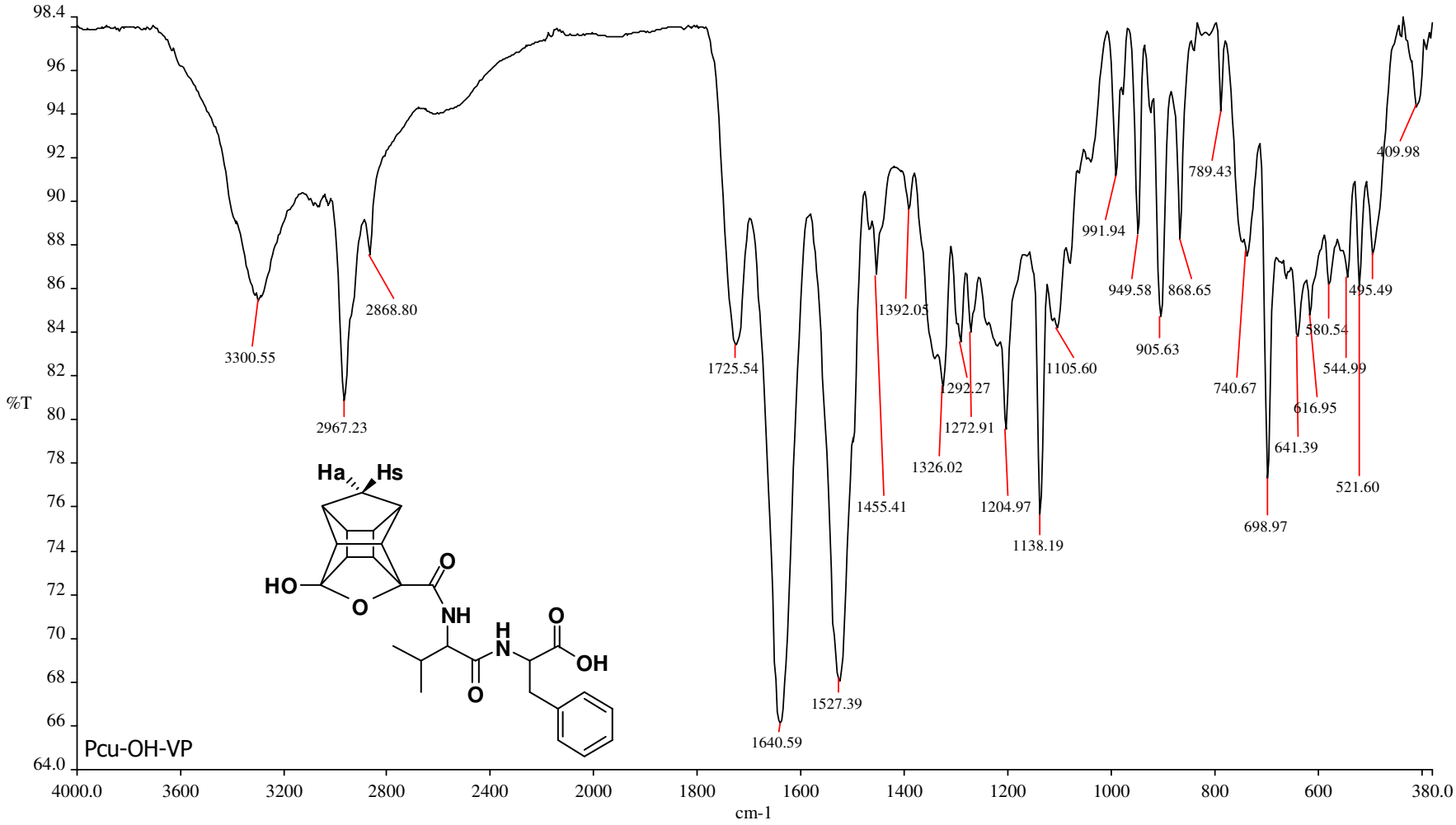

```
Current Data Parameters
NAME      Sep04-2010-GK-Raj-1
EXPNO                  18
PROCNO                  1

F2 - Acquisition Parameters
Date_            20100905
Time                1.56
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG       hsqcedetgp
TD                 1024
SOLVENT             DMSO
NS                     8
DS                    16
SWH             5813.954 Hz
FIDRES          5.677689 Hz
AQ             0.0881140 sec
RG                  2050
DW                86.000 usec
DE                 6.50 usec
TE                298.2 K
CNST2        145.0000000
D0            0.00000300 sec
D1            1.50778198 sec
D4            0.00172414 sec
D11           0.03000000 sec
D13           0.00000400 sec
D16           0.00020000 sec
D21           0.00345000 sec
IN0           0.00003000 sec
ZGOPTNS

======== CHANNEL f1 ========
NUC1                   1H
P1                 10.00 usec
P2                 20.00 usec
P28              1000.00 usec
PL1                -3.00 dB
PL1W         15.48668575 W
SFO1         400.2227616 MHz

======== CHANNEL f2 ========
CPDPRG2              garp
NUC2                  13C
P3                  9.00 usec
P4                 18.00 usec
PCPD2              70.00 usec
PL2                -2.00 dB
PL12               15.60 dB
PL2W         54.14257431 W
PL12W         0.94089001 W
SFO2         100.6429049 MHz

====== GRADIENT CHANNEL =====
GPNAM1          SINE.100
GPNAM2          SINE.100
GPZ1               80.00 %
GPZ2               20.10 %
P16              1000.00 usec

F1 - Acquisition parameters
ND0                    2
TD                   256
SFO1           100.6429 MHz
```

HSQC NMR Spectrum of Compound 2 (Chapter 5)




```
Current Data Parameters
NAME      Sep04-2010-GK-Raj-1
EXPNO                      19
PROCNO                      1

F2 - Acquisition Parameters
Date_            20100905
Time                 2.54
INSTRUM             spect
PROBHD   5 mm PABBO BB-
PULPROG       hmbcgpndqf
TD                   4096
SOLVENT              DMSO
NS                     16
DS                     16
SWH              5813.954 Hz
FIDRES           1.419422 Hz
AQ               0.3523060 sec
RG                    2050
DW                 86.000 usec
DE                  6.50 usec
TE                 298.2 K
CNST13          8.0000000
D0              0.00000300 sec
D1              1.53194904 sec
D6              0.06250000 sec
D16             0.00020000 sec
IN0             0.00002235 sec

======== CHANNEL f1 ========
NUC1                    1H
P1                  10.00 usec
P2                  20.00 usec
PL1                 -3.00 dB
PL1W         15.48668575 W
SFO1         400.2227616 MHz

======== CHANNEL f2 ========
NUC2                   13C
P3                   9.00 usec
PL2                 -2.00 dB
PL2W         54.14257431 W
SFO2         100.6454466 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPNAM2           SINE.100
GPNAM3           SINE.100
GPZ1                50.00 %
GPZ2                30.00 %
GPZ3                40.10 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                     2
TD                    256
SFO1           100.6454 MHz
FIDRES         87.315842 Hz
SW               222.095 ppm
FnMODE                 QF

F2 - Processing parameters
SI                    1024
SF           400.2199980 MHz
WDW                 SINE
```

HMBC NMR Spectrum of Compound 2 (Chapter 5)



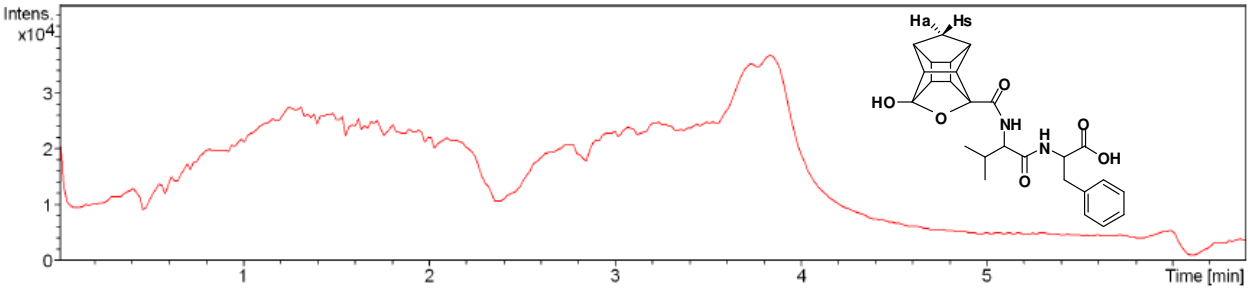
IR Spectrum of Compound 2 (Chapter 5)

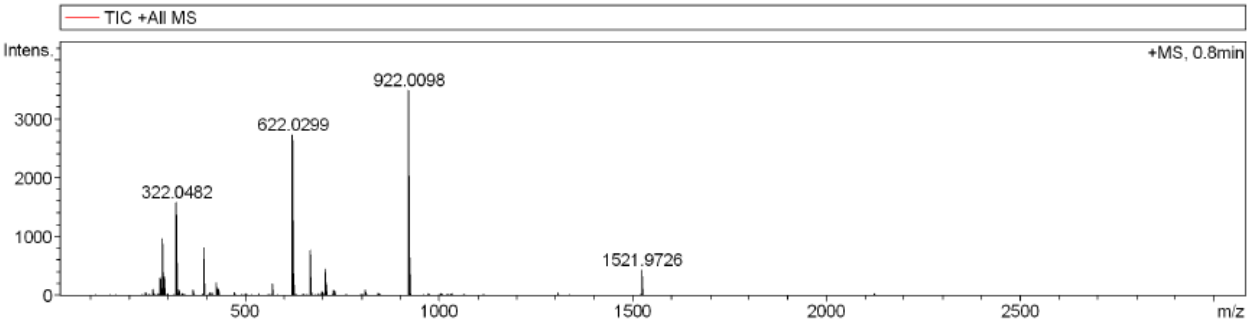
Analysis Info

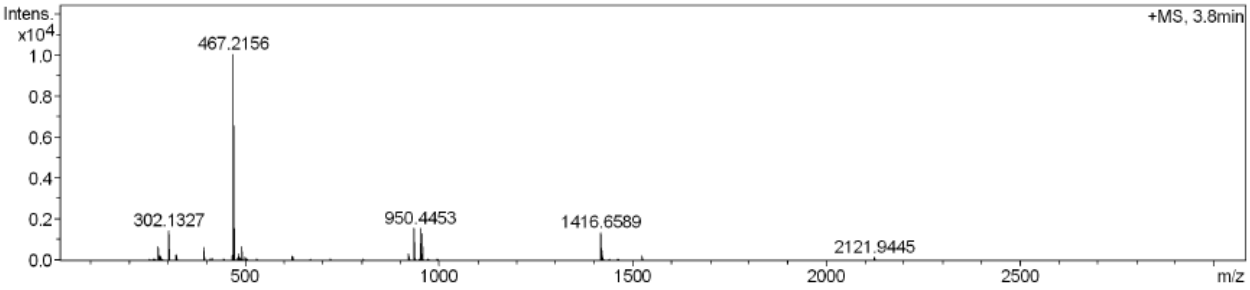
Analysis Name	D:\Data\Raj\5.d	Acquisition Date	2/24/2012 7:21:39 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	5	Instrument / Ser#	micrOTOF-Q 10139
Comment			

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

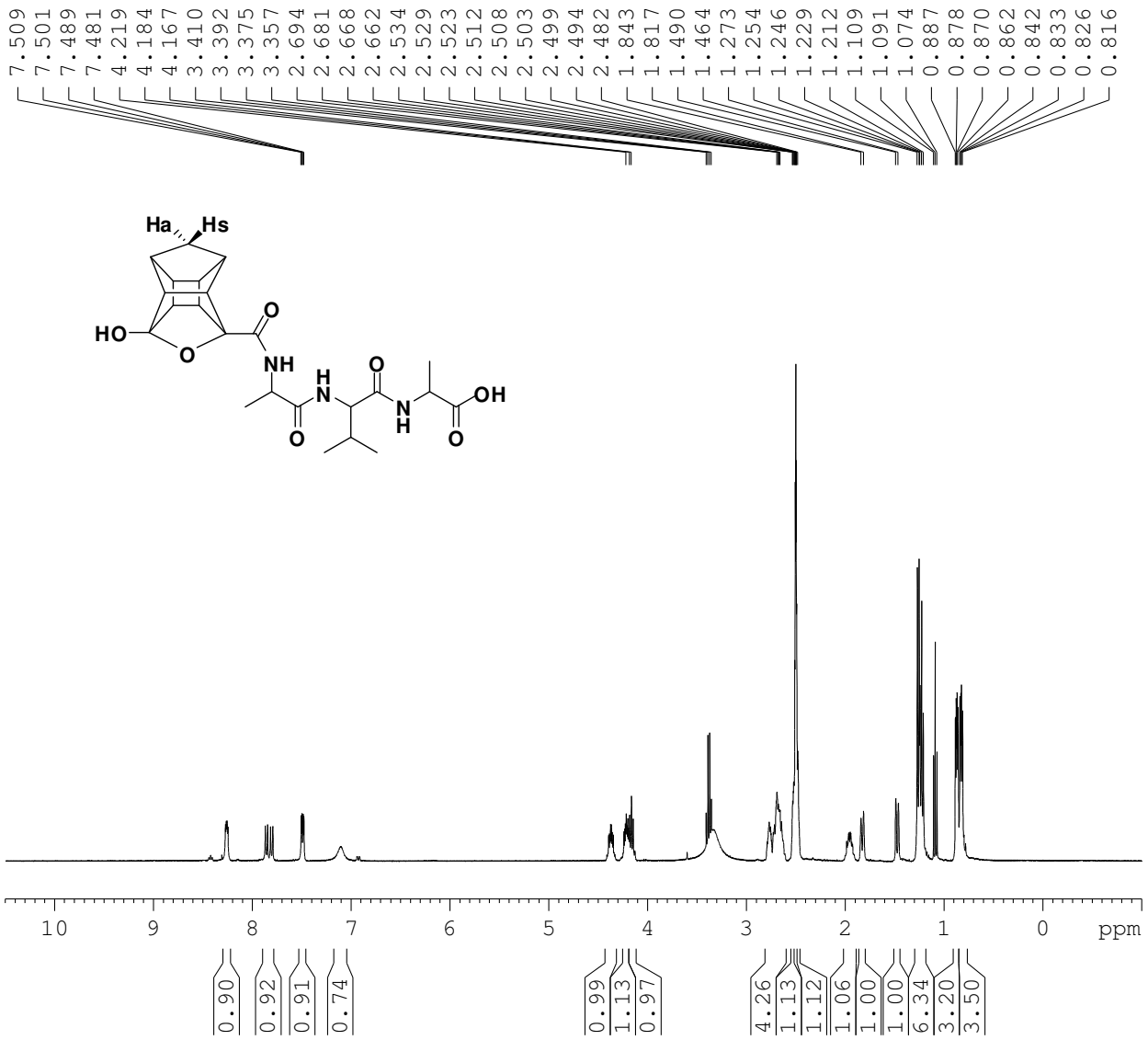






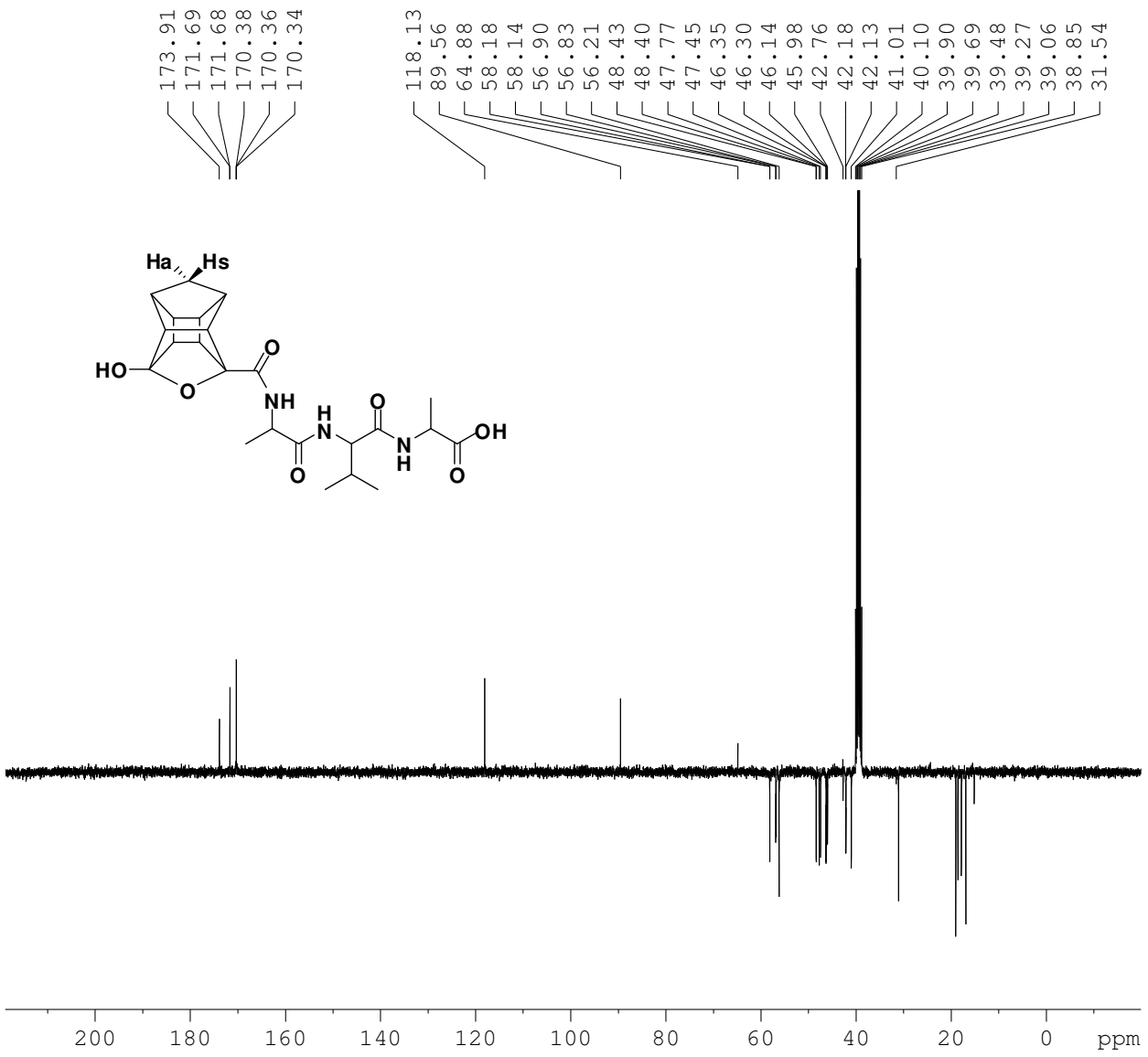
Meas. m/z	#	Formula	m/z	err [ppm]	Mean err [ppm]	rdb	N-Rule	e⁻ Conf	mSigma	Std I	Std Mean m/z	Std I VarNorm	Std m/z Diff	Std Comb Dev
467.2156														
	1	C 22 H 27 N 8 O 4	467.2150	-1.3	-1.2	13.5	ok	even	3.98	0.0064	0.0006	0.0043	0.0002	0.5891
	2	C 21 H 31 N 4 O 8	467.2136	-4.2	-4.1	8.5	ok	even	11.15	0.0200	0.0019	0.0066	0.0003	0.7818
	3	C 26 H 31 N 2 O 6	467.2177	4.4	4.5	12.5	ok	even	14.77	0.0221	0.0021	0.0097	0.0003	0.8347

HRESIMS (*m/z*) of Compound 2 (Chapter 5)



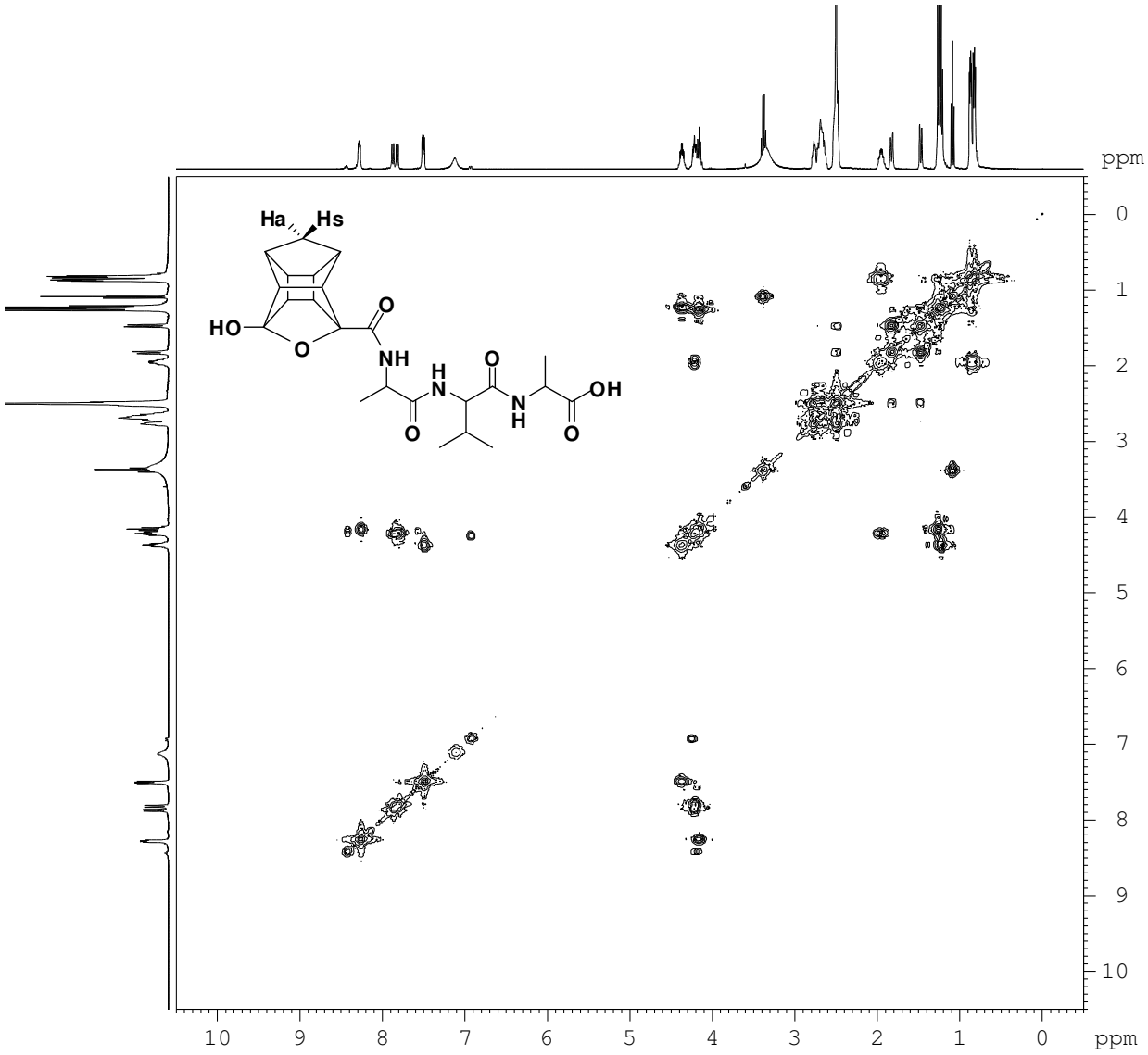

```
Current Data Parameters
NAME      Aug30-2010-GK-Raj
EXPNO                   22
PROCNO                   1

F2 - Acquisition Parameters
Date_           20100902
Time               13.03
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG             zg30
TD                 65536
SOLVENT             DMSO
NS                    16
DS                     2
SWH             8223.685 Hz
FIDRES          0.125483 Hz
AQ             3.9846387 sec
RG                   128
DW                60.800 usec
DE                 6.50 usec
TE                298.2 K
D1           1.00000000 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                  1H
P1                10.00 usec
PL1               -3.00 dB
PL1W        15.48668575 W
SFO1        400.2224715 MHz

F2 - Processing parameters
SI                 32768
SF          400.2199985 MHz
WDW                   EM
SSB                    0
LB                0.30 Hz
GB                     0
PC                1.00
```

^{1}H NMR Spectrum of Compound 4 (Chapter 5)



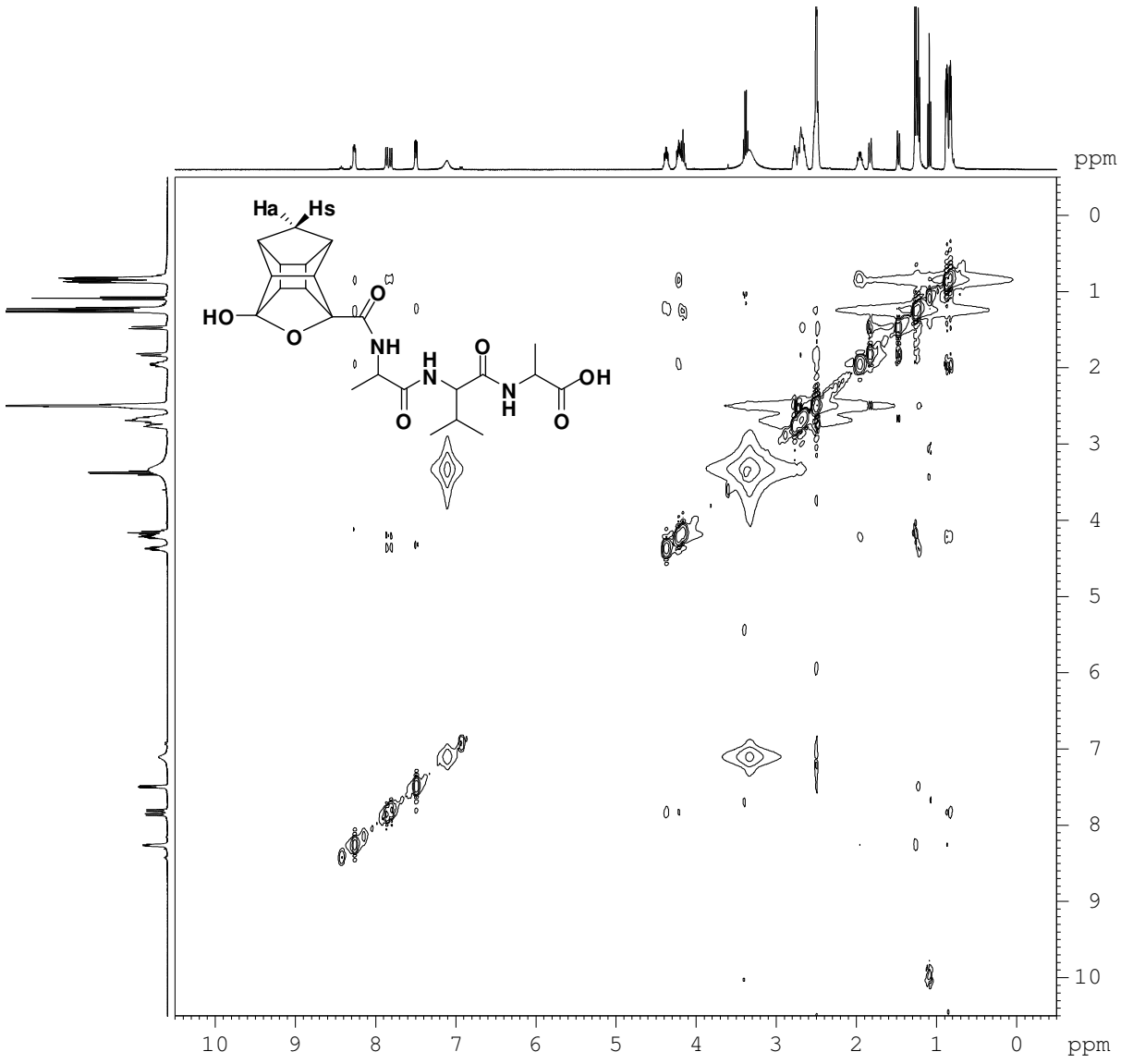

```
Current Data Parameters
NAME      Aug30-2010-GK-Raj
EXPNO                   27
PROCNO                   1

F2 - Acquisition Parameters
Date_           20100902
Time               22.04
INSTRUM            spect
PROBHD   5 mm PABBO BB-
PULPROG             jmod
TD                 65536
SOLVENT             DMSO
NS                  1024
DS                     4
SWH            24038.461 Hz
FIDRES             0.366798 Hz
AQ                 1.3631988 sec
RG                    2050
DW                 20.800 usec
DE                  6.50 usec
TE                 298.2 K
CNST2         145.0000000
CNST11          1.0000000
D1              2.00000000 sec
D20             0.00689655 sec
TD0                    1

======== CHANNEL f1 ========
NUC1                  13C
P1                  8.40 usec
P2                 16.80 usec
PL1                -2.00 dB
PL1W          54.14257431 W
SFO1          100.6454626 MHz

======== CHANNEL f2 ========
CPDPRG2           waltz16
NUC2                  1H
PCPD2              90.00 usec
PL2                -3.00 dB
PL12               15.60 dB
PL2W          15.48668575 W
PL12W          0.21377575 W
SFO2          400.2216009 MHz

F2 - Processing parameters
SI                 32768
SF            100.6354471 MHz
WDW                   EM
SSB                    0
LB                  1.00 Hz
GB                     0
PC                  1.40
```

^{13}C NMR Spectrum of Compound 4 (Chapter 5)



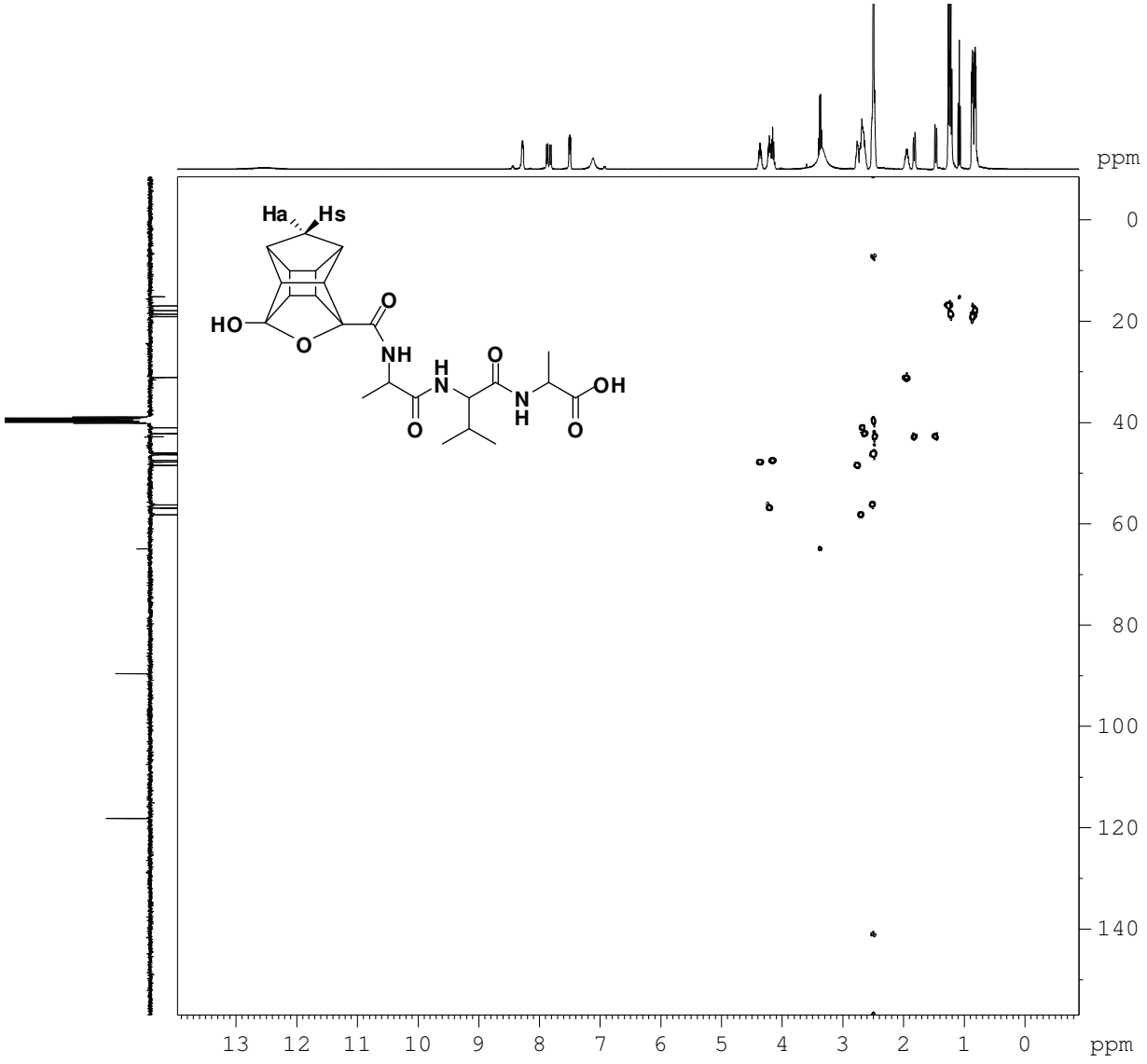

```
Current Data Parameters
NAME      Aug30-2010-GK-Raj
EXPNO                   23
PROCNO                   1

F2 - Acquisition Parameters
Date_              20100902
Time                 13.05
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG          cosygpqf
TD                   2048
SOLVENT               DMSO
NS                       2
DS                       8
SWH               5952.381 Hz
FIDRES            2.906436 Hz
AQ               0.1720820 sec
RG                      64
DW                 84.000 usec
DE                  6.50 usec
TE                 298.2 K
D0             0.00000300 sec
D1             1.50655305 sec
D13            0.00000400 sec
D16            0.00020000 sec
IN0            0.00016800 sec

======== CHANNEL f1 ========
NUC1                    1H
P0                  10.00 usec
P1                  10.00 usec
PL1                 -3.00 dB
PL1W         15.48668575 W
SFO1         400.2226137 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPZ1                10.00 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                      1
TD                     256
SFO1             400.2226 MHz
FIDRES          23.251488 Hz
SW                 14.873 ppm
FnMODE                  QF

F2 - Processing parameters
SI                    1024
SF            400.2199985 MHz
WDW                   SINE
SSB                      0
LB                    0.00 Hz
GB                       0
PC                    1.40

F1 - Processing parameters
SI                    1024
MC2                     QF
SF            400.2199985 MHz
WDW                   SINE
SSB                      0
LB                    0.00 Hz
```

COSY NMR Spectrum of Compound 4 (Chapter 5)



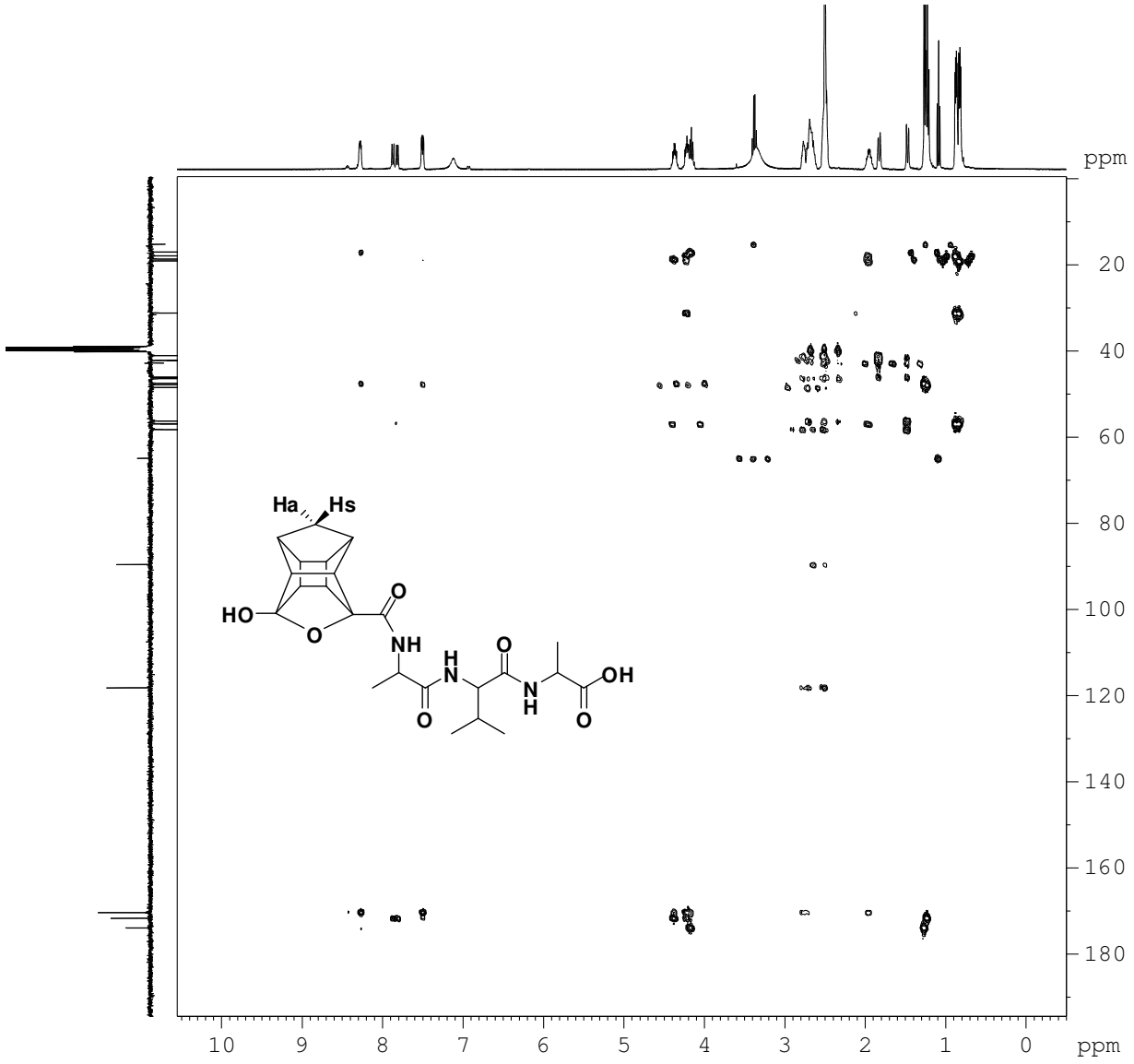

```
Current Data Parameters
NAME      Aug30-2010-GK-Raj
EXPNO                   24
PROCNO                   1

F2 - Acquisition Parameters
Date_             20100902
Time                 13.21
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG           noesyph
TD                    2048
SOLVENT               DMSO
NS                       8
DS                       4
SWH               5952.381 Hz
FIDRES            2.906436 Hz
AQ               0.1720820 sec
RG                      64
DW                 84.000 usec
DE                  6.50 usec
TE                 298.2 K
D0              0.00007127 sec
D1              2.07864308 sec
D8              0.30000001 sec
IN0             0.00016800 sec

======== CHANNEL f1 ========
NUC1                    1H
P1                  10.00 usec
PL1                 -3.00 dB
PL1W          15.48668575 W
SFO1          400.2226137 MHz

F1 - Acquisition parameters
ND0                      1
TD                     256
SFO1             400.2226 MHz
FIDRES           23.251488 Hz
SW                 14.873 ppm
FnMODE       States-TPPI

F2 - Processing parameters
SI                    1024
SF            400.2199985 MHz
WDW                  QSINE
SSB                      2
LB                    0.00 Hz
GB                       0
PC                    1.00

F1 - Processing parameters
SI                    1024
MC2          States-TPPI
SF            400.2199985 MHz
WDW                  QSINE
SSB                      2
LB                    0.00 Hz
GB                       0
```

NOESY NMR Spectrum of Compound 4 (Chapter 5)



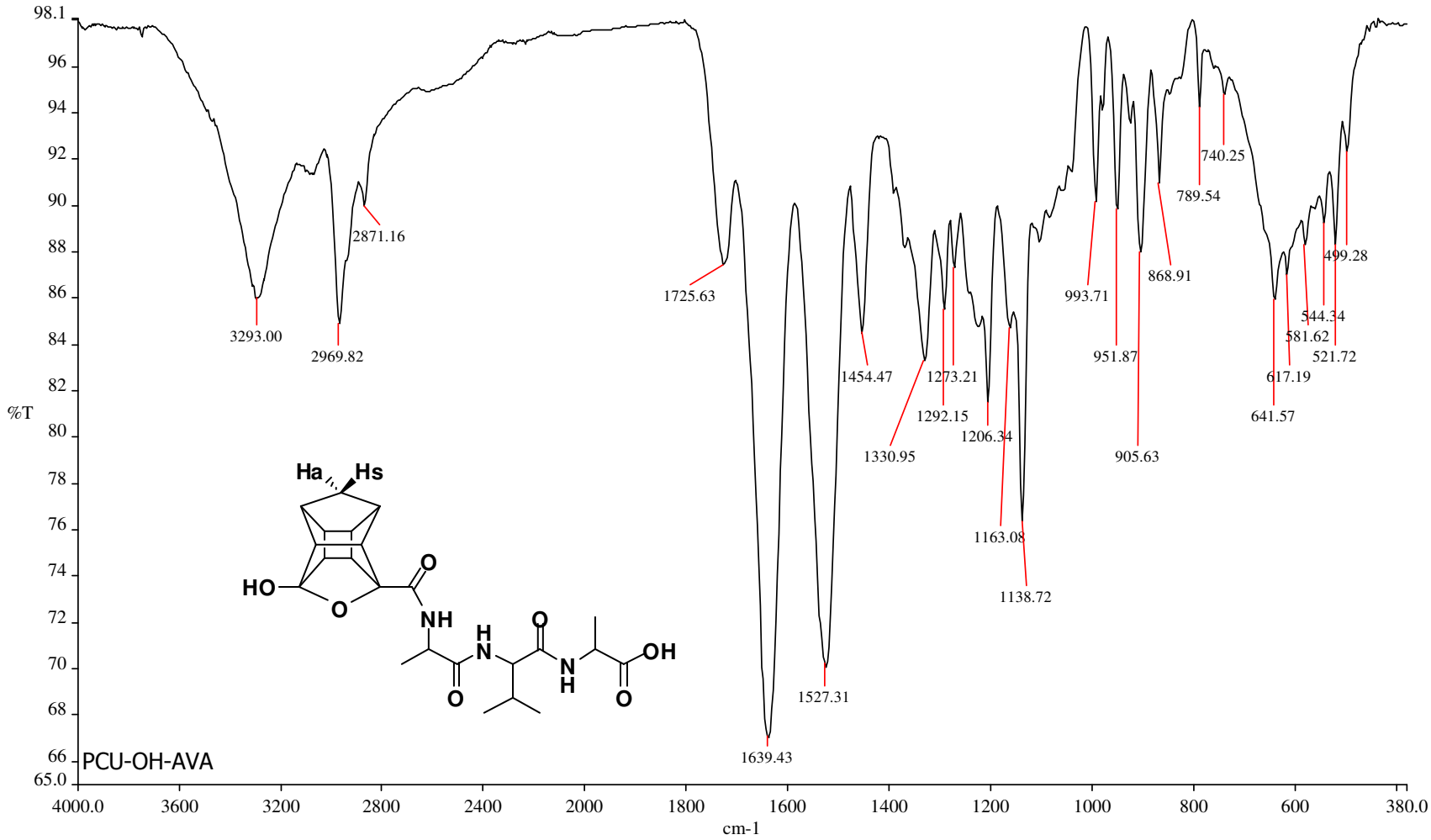

```
Current Data Parameters
NAME      Aug30-2010-GK-Raj
EXPNO                   25
PROCNO                   1

F2 - Acquisition Parameters
Date_              20100902
Time                 14.50
INSTRUM              spect
PROBHD   5 mm PABBO BB-
PULPROG       hsqcedetgp
TD                    1024
SOLVENT               DMSO
NS                       4
DS                      16
SWH              5952.381 Hz
FIDRES           5.812872 Hz
AQ             0.0860660 sec
RG                    2050
DW                 84.000 usec
DE                  6.50 usec
TE                 298.2 K
CNST2         145.0000000
D0             0.00000300 sec
D1             1.50983000 sec
D4             0.00172414 sec
D11            0.03000000 sec
D13            0.00000400 sec
D16            0.00020000 sec
D21            0.00345000 sec
IN0            0.00003000 sec
ZGOPTNS

======== CHANNEL f1 ========
NUC1                    1H
P1                  10.00 usec
P2                  20.00 usec
P28               1000.00 usec
PL1                 -3.00 dB
PL1W          15.48668575 W
SFO1         400.2226137 MHz

======== CHANNEL f2 ========
CPDPRG2               garp
NUC2                   13C
P3                   9.00 usec
P4                  18.00 usec
PCPD2               70.00 usec
PL2                 -2.00 dB
PL12                15.60 dB
PL2W          54.14257431 W
PL12W          0.94089001 W
SFO2         100.6429049 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPNAM2           SINE.100
GPZ1                80.00 %
GPZ2                20.10 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                      2
TD                     256
SFO1             100.6429 MHz
```

HSQC NMR Spectrum of Compound 4 (Chapter 5)




```
Current Data Parameters
NAME      Aug30-2010-GK-Raj
EXPNO                   26
PROCNO                   1

F2 - Acquisition Parameters
Date_              20100902
Time                 17.24
INSTRUM              spect
PROBHD    5 mm PABBO BB-
PULPROG       hmbcgpndqf
TD                     4096
SOLVENT                DMSO
NS                       16
DS                       16
SWH               5952.381 Hz
FIDRES            1.453218 Hz
AQ                0.3441140 sec
RG                     2050
DW                  84.000 usec
DE                   6.50 usec
TE                  298.2 K
CNST13           8.0000000
D0               0.00000300 sec
D1               1.54014099 sec
D6               0.06250000 sec
D16              0.00020000 sec
IN0              0.00002235 sec

======== CHANNEL f1 ========
NUC1                    1H
P1                   10.00 usec
P2                   20.00 usec
PL1                  -3.00 dB
PL1W          15.48668575 W
SFO1          400.2226137 MHz

======== CHANNEL f2 ========
NUC2                   13C
P3                    9.00 usec
PL2                  -2.00 dB
PL2W          54.14257431 W
SFO2          100.6454466 MHz

====== GRADIENT CHANNEL =====
GPNAM1           SINE.100
GPNAM2           SINE.100
GPNAM3           SINE.100
GPZ1                50.00 %
GPZ2                30.00 %
GPZ3                40.10 %
P16               1000.00 usec

F1 - Acquisition parameters
ND0                      2
TD                     256
SFO1            100.6454 MHz
FIDRES          87.315842 Hz
SW                222.095 ppm
FnMODE                  QF

F2 - Processing parameters
SI                    1024
SF            400.2199950 MHz
WDW                   SINE
```

HMBC NMR Spectrum of Compound 4 (Chapter 5)



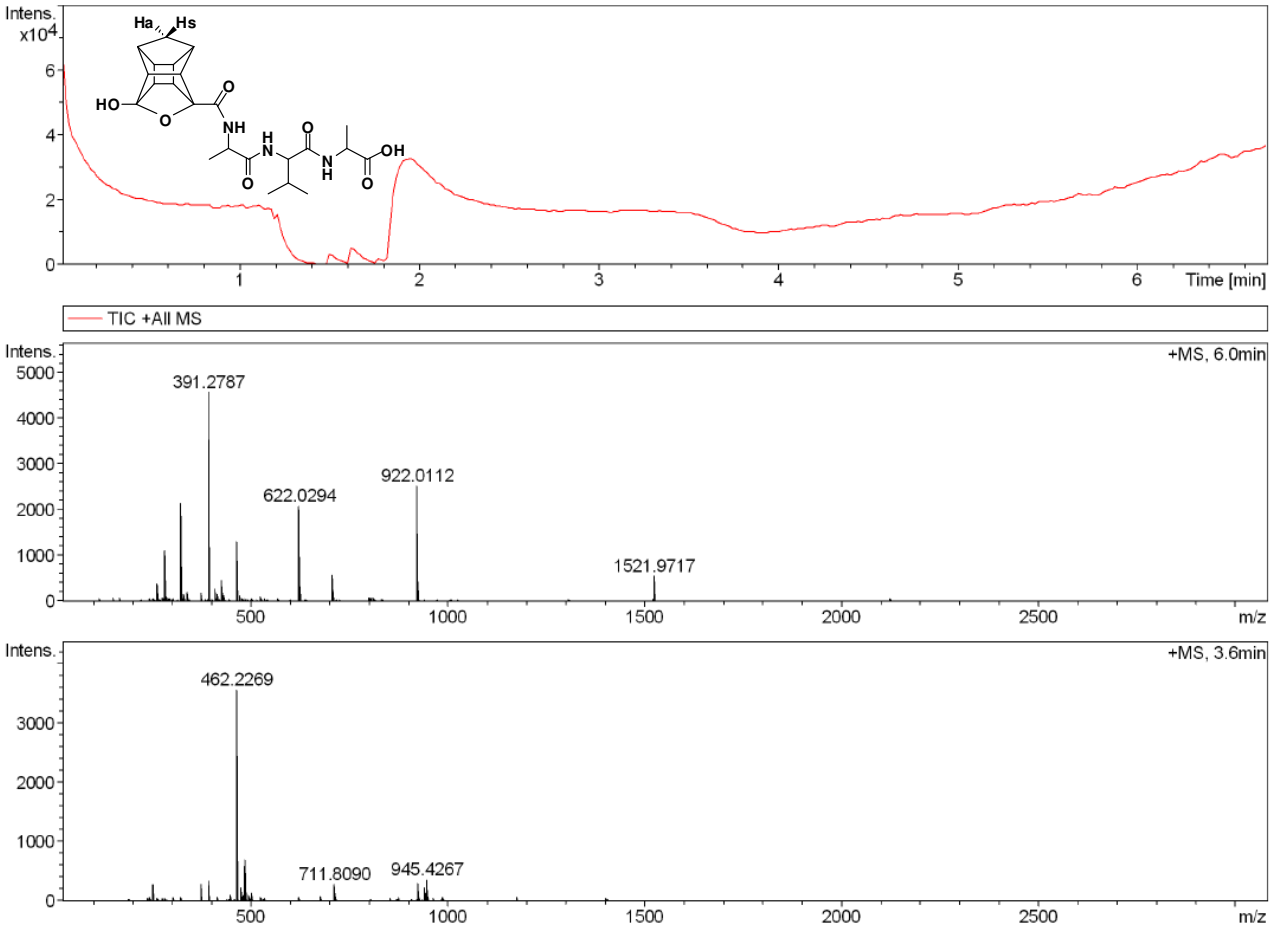
IR Spectrum of Compound 4 (Chapter 5)

Analysis Info

Analysis Name	D:\Data\Raj\6.d	Acquisition Date	2/24/2012 7:32:09 PM
Method	tune_wide_expert.m	Operator	BDAL@DE
Sample Name	6	Instrument / Ser#	micrOTOF-Q 10139
Comment			

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]	Mean err [ppm]	rdb	N-Rule	e⁻ Conf	mSigma	Std I	Std Mean m/z	Std I VarNorm	Std m/z Diff	Std Comb Dev
462.2269														
	1	C 23 H 32 N 3 O 7	462.2235	-7.5	-7.0	9.5	ok	even	6.29	0.0109	0.0033	0.0039	0.0016	0.7779
	2	C 24 H 28 N 7 O 3	462.2248	-4.6	-4.2	14.5	ok	even	19.46	0.0279	0.0020	0.0124	0.0005	0.8060
	3	C 17 H 32 N 7 O 8	462.2307	8.1	8.3	5.5	ok	even	22.68	0.0375	0.0039	0.0162	0.0005	0.9016

HRESIMS (*m/z*) of Compound 4 (Chapter 5)