STUDIES ON LOW MOLECULAR MASS CYSTEINE PROTEASE INHIBITOR FROM ACTINOMYCETES

A THESIS

SUBMITTED TO THE

UNIVERSITY OF PUNE

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

(IN BIOTECHNOLOGY)

BY

JAY PRAKASH SINGH

UNDER THE GUIDANCE OF

DR. (Mrs.) S. M. GAIKWAD

DIVISION OF BIOCHEMICAL SCIENCES

NATIONAL CHEMICAL LABORATORY

PUNE-411008 (INDIA)

JUNE-2009

CONTENTS

CERTIFICATE

Certified that the work incorporated in the thesis entitled, "**Studies on Low Molecular Mass Cysteine Protease Inhibitor from Actinomycetes**" submitted by Mr. Jay Prakash Singh, was carried out under my supervision. Such material as has been obtained from other sources has been duly acknowledged in the thesis.

Dr. (Mrs.) S. M. Gaikwad Date

(Research Guide)

DECLARATION OF THE CANDIDATE

I declare that the thesis entitled, "**Studies on Low Molecular Mass Cysteine Protease Inhibitor from Actinomycetes**." Submitted by me for the degree of Doctor of Philosophy is the record of work carried out by me during the period from **11th Nov. 2005 to 5th June 2009** under the guidance of **Dr. (Mrs.) S. M. Gaikwad** and has not formed the basis for the award of any degree, diploma, associateship, fellowship, titles in this or any other university or institute of higher learning.

I further declare that the material obtained from other sources has been duly acknowledged in the thesis.

Jay Prakash Singh Date

ACKNOWLEDGEMENT

I take this opportunity to gratefully acknowledge my research mentor Dr. M.I. Khan for his invaluable guidance, constant inspiration, unending support and keen interest during the course of this investigation. He has given me the freedom to think and work; and I shall cherish my learning experience under him. His scientific temperament, innovative approach, dedication towards profession and down to earth nature has inspired me the most. Although this eulogy is insufficient, I preserve on everlasting gratitude for him. I am also grateful to Mrs. Khan for love, care and treating me like part of family. I will miss delicious food they served during numerous dinners at their home especially chicken biryani, kababs, fish curry and many more.

I am equally obliged to my research guide, Dr. Sushama for continuous support and specific scientific contribution throughout my stay in the laboratory and devoting her valuable time in putting this thesis into the shape.

I thank Dr G. C. Mishra for his scientific and fundamental teachings and continuous support. He has always been there for me and supported whenever I needed his contributions. It's a fact that I could not finish my research work if he would not have been involved directly or indirectly in almost each and every step during entire six years of my research life. I still remember his words when he once said, "I will remain your guide for whole of my life." I thank you sir …..

I thank Dr. Mohan Vani and Dr Gopal Kundu and their entire groups for helping and encouraging me to finish some of my very important cell biology experiments in their laboratories.

I am obliged to Dr. Rajamohanan and his group from central NMR facility for helping me to record NMR spectra and interpretation. Actually, Dr Rajamohanan himself, is the person whose intellect and effort made it possible.

I am also grateful Dr. M. V. Krishnasastry, Dr. C. G. Suresh, Dr. S.S. Deshmukh, Dr. Absar Ahmed, Dr. Arvind Sahu, Dr Patole, Dr. Manoj Bhatt Dr. Mala Rao and for their valuable suggestions during the course of investigation.

I express my deep feelings and love for my friends and my lab mates Feroz, Anil, Atul, Siddharth, Shashi, Sajid, JP, Late Rohtas, Sreekanth, Ansari, Avinash, Madhurima, Asad, Shabab, Shadab, Ashutosh, Prabhakar, Pooja, Santosh, Apeksha, Vaibhav and Ravi for their help, cooperation and maintaining pleasant and healthy atmosphere throughout my doctorate course. I would like to express my deep felt gratitude to my friends, Abhishek, Jyoti, Abhilash, Ambrish, Anish, Prasanna, Yadav, kanna, Ashish, Ashwini(prajee), Bhuban, Bhushan, Chetan, Fazal, Kishore (bhaiya), Manish, Nishant, Nitin, Nookaraju, Noor, Ajeet, Pandey, Pawan, Poonam, Rhodu, Samir, Santosh (Bond), Santosh (professor) Sarvesh, Satya, Sachin, Sharath, Sushim, Shivshanker, Atul Anand, Gyan, Yashwant, Rishi, Somesh, Mohsin for help, support and charming company.

Atul (gabber), Sudarshan, Mishra, Pankaj, Pandey, Late Rohtas, Chetan, Shashi, Manje, (Arun)Yadav, Noor, Ashish, Vikkoo, Kishore (bhaiya), Amit (sasur) and Akhilesh for countless things they have done for me and for always being there with me whenever I needed them.

I am not able to find words for my parents, my brothers and sister; they have been constant inspiration for me and this work would not have been possible without their constant support and sacrifices.

I thank my friends from NCCS Manu, kamini, Gautam, Shalini, Ram, Sunit, for their support every time whenever I needed them.

Finally, I thank Head, Division of Biochemical Sciences and Directors of National Chemical Laboratory and National Centre for cell Science for permitting me to work avail their facilities required for my research. I should not forget to thank our funding agencies Council of Scientific and Industrial Research and DBT Gov. of India for financial assistance.

Please forgive me if any the contributions are left unacknowledged; it may have happened but, that is unintentional.

Jay Prakash

ABSTRACT:

Deregulated expressions of proteases are convincingly reported to be involved in various pathobiological processes. Thiol proteases such as Cathepsin B expression is increased in many human cancers. The predominant expression of Cathepsin K in osteoclasts has rendered the enzyme a major target for the development of novel anti-resorptive drugs in osteoporosis and Cathepsin S appears to be an attractive drug target candidate for various inflammatory diseases including rheumatoid arthritis. Parasites such as *Trypnosoma*, *Lieshmania*, and *Plasmodium* also take the advantage of thiol protease activity to evade the host immune response. Therefore, it is reasonable to assume that a good thiol protease inhibitor can be a useful compound to target the aforementioned disorders.

Microbes and plants have been the most efficient and convenient source of natural bioactive compounds. Novel secondary metabolites continue to be isolated from Actinomycetes. Their biological activities and chemical structures show a wide range of diversity. Taking into consideration all these understanding, we decided to further explore the actinomycetes in order to isolate small molecular mass cysteine protease inhibitor. During the course of investigation we have isolated a novel peptide (736-842 Da) which has Cathepsin K and L inhibition activity in nanomolar range. This uniquely modified pentapeptide was isolated with the help of solvent extraction and repeated reverse phase HPLC. The peptide was characterized completely with help of various spectroscopic techniques such as LC-MS-MS, FT-IR, 1D and 2D NMR.

The compound was tested on various tumor cell lines (MDA-MB-231, A_{375} and $B_{16}F_{10}$ cells) in order to find its dose of cytotoxicity; interestingly, it was found that more than 90% of cells are viable even at 50 μM CPI-2081 in MTT cell viability assay. However the peptide was able to inhibit the cell migration significantly at 50 μM concentration in wound healing assay. The effect of CPI-2081 was also investigated on osteoclast differentiation. The results suggest that CPI inhibits osteoclast differentiation without any toxic effect on osteoclast precursors.

To summarize, we have isolated a novel modified peptidic thiol protease

inhibitor (CPI-2081) from actinomycetes *(Streptomyces sp* NCIM2081) which exhibit K_i in nanomolar (<100 nm) range and it can inhibit the tumor cell migration without any cytotoxic activity. Also, the CPI is able to inhibit the RANKEL induced osteocalast differentiation without having considerable cytotoxic effect.

Organization of thesis:

Chapter 1: *General Introduction*: This chapter contains the outline of literature survey with reference to properties and types of proteases with special attention to cysteine proteases. This chapter also contains the description of protenatious and small molecule cysteine protease inhibitors. The last part of the chapter contains an evaluation of benefits of the natural product drug discovery with an attempt to justify the need for further effort in the area.

Chapter 2: *Production, purification and characterization of CPI:* This chapter deals with bioactivity guided screening of various *Streptomyces sp.* (source of compound) from NCIM were carried out to look for cysteine protease inhibitor. Culture supernatants of micro-organisms were used to inhibit the activity of papain (cysteine protease). Out of many isolate tasted, culture supernatant of NCIM2081 was found to inhibit protease activity by ~80%. Further experiments were carried out to enhance the production of inhibitor followed by its extraction and isolation. The inseparable mixture of two compounds was obtained and their chemical structures has been elucidated with help of FT-IR, UV, LC-MS and various kinds of NMR spectroscopic techniques, such as 1H, ^{13}C, DEPT, 2D NMR, NOSY, COSY, TOCSY, ROSY, HSQC and HMBC.

A. ***Extraction and purification and identification of CPI:*** This part of chapter-2 contains the detailed procedure of extraction of CPI using various solvents such as n-hexane, dichloromethane and methanol. The complete two round of purification process using RP-HPLC is also been discussed. Then ESI-Ms was carried out to find out the molecular mass of the isolated compound and it was found that there was two moleculer ion

peaks are present in the sample which indicated the prasance of at least two different compounds of moleculer mass of 736 Da and 842 Da, in the sample. FT-IR spectroscopy was done to investigate about the nature of compound, which was found to be comtaining some peptidic amide bonds and aromatic rings.

B. ***NMR spectroscopic studies of CPI and elucidation of its chemical structure:*** This part of the chapter-2 contains the fully detailed account of elucidation of complete chemical structure of compound (CPI-2081) by the help of various 1-D and 2-D NMR spectroscopic techniques such ^{1}H, ^{13}C, ^{13}C DEPT, DOSY, NOSY, COSY, TOCSY, ROSY, HSQC and HMBC.

Chapter 3: Thermodynamics of binding of papain like cysteine proteases and CPI:

CPI-2081 inhibits both the papain like thiol proteases, Cathepsin K and L via competitive mode of action as illustrated by Dixon plot. Thermodynamic parameters have been calculated with the help of Vant Hoff plot for Cathepsin L and Cathepsin K. In case of both the enzymes studied the binding was entropically driven as evidenced by very high change in entropy as compared to the change found in case of enthalpy. Change of free energy during reaction was negative hence, the reaction is spontaneous.

***Chapter 4: In vitro studies on pharmacological implication of CPI in tumor cells and osteoclasts*:** The invasive and metastatic potential of tumor cells has been correlated with activity of various protease activities including thiol proteases such as cathepsins. This chapter deals with the efforts made in the course of finding the lead compound (scaffold) using functional implications of thiol proteases in cell migration. The purified CPI-2081 inhibited the cell motality, up to sixty percent at the concentrations which is not cytotoxic to the tumor cell lines studied. The predominant expression of papain like thiol proteases such as cathepsin K and L, in osteoclasts and their implication in irreversible bone degradation and pathological condition during osteoporosis has

rendered the enzyme a major target for the development of novel anti-resorptive drugs. This chapter deals with the experiments which demonstrate the role of CPI-2081 under study, to inhibit RANKEL induced mature osteoclast and their bone resorptive potential.

Chapter 5: General discussion and conclusion: All experimental results of chapter 2, 3 and 4 have been discussed in detail and it is concluded that our effort to characterize low molecular weight, pentapeptide, competitive protease inhibitor, which has inhibitory constant (K_i) against recombinant human Cathepsin-K and L, <100 nM, probably due to its unique natural modification, could be useful for drug development programs directed against osteoporosis and cancer.

List of Abbreviations

ES-MS	Electrospray mass spectrometer
FTIR	Fourier Transform Infrared Spectroscopy
IC_{50}	50% inhibitory concentration
K_i	Inhibition constant for inhibitor
NCIM	National center for industrial microorganisms
RP-HPLC	Reverse phase high performance liquid chromatography
CPI	Cysteine protease inhibitor
DEPT	Distortionless Enhancement by Polarization Transfer
DMSO	Dimethyl sulphoxide
HMBC	Heteronuclear Multiple Bond Correlation
HMQC	Heteronuclear Multiple Quantum Coherence
HPLC	High Performance Liquid Chromatography
MS	Mass Spectrum
TFA	Trifluoroacetic acid
BAPNA	N-α-benzoyl-L-arginine p-nitroanilide
BAEE	N-α-benzoyl-L-arginine ethyl ester
RB	Rotational Bonds
Å	Angstrom
BAPNA	Nα-Benzoyl-DL-Arginine p-Nitroanilide
CH_3CN	Acetonitrile
CPI	Cysteine Protease Inhibitor
Da	Dalton
DTT	Dithiothreitol
E-64	L-trans-epoxysuccinyl-leucyl-amido (4-guanidino)-butane)
EC	Enzymes classification
EDTA	Ethylene diamine tetra acetic acid
EMIS	Electromagnetic induction spectroscopy
kDa	Kilo Dalton
M	Molar

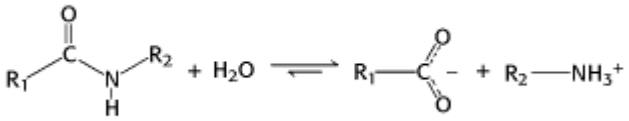
CHAPTER 1

General Introduction

INTRODUCTION

What are proteases?

Proteases (EC: 3.4) are class of hydrolytic enzymes catalyzing hydrolysis, the addition of



water molecule to a peptide bond or amide bonds between amino acid residues in proteins. In other words, a protease is an enzyme that breaks the larger proteins and peptides through the hydrolysis of the peptide bonds between amino acids residues, in a polypeptide chain. The active site of proteases includes features that allow for the activation of water or another nucleophile as well as for the polarization of the peptide carbonyl group and subsequent stabilization of a tetrahedral intermediate. Based on the site of cleavage, the peptidases/ proteases are classified as endopeptidases, which hydrolyze the internal amide bond in peptide or polypeptide chain, and, exopeptidases, cleaving the terminal amide bond in peptide chain. Chemically, the endoproteases have been classified into acidic, basic and neutral proteases. But more convenient and common system of classification is based on the reactive amino acidic residue present in active site. Based on the site of protease action in polypeptide, proteases are classified as exopeptidase and endopeptidases. Mostly, proteins undergo reversible post translational modifications during their life span such as phosphorylation, formation of tertiary structures and transition of the 3-dimentional structures among various energetically feasible forms. However, the hydrolysis of proteins by proteases is an irreversible phenomenon. Therefore, it is not unanticipated that proteolytic enzymes have evolved to facilitate the biological pathways that are themselves often irreversible for example blood coagulation (thrombin), digestion (trypsin, pepsin, chymotrypsin), maturation of growth factors and cytokines (Interleukins and interferon), processing of zymogens and prohormones, apoptosis (caspases), and degradation of unwanted or excess intracellular proteins (lysosomal cathepsins). Proteolysis is a very common process in the cell, essentially used to regulate the function and fortune of proteins[1, 2]. The number of

proteases discovered in entire living system is very high, and many of them are crucial for normal homeostasis of the organism.

Classification protease:

Commonly, proteases follow the general mechanism of a nucleophilic attack on the carbonyl-carbon of a peptide bond[3]. As a result of which, a general acid-base hydrolytic course of action that breaks the covalent bond. Amino acid involved, in the activity of particular proteolytic process, which depends on type of the protease, decides strategies to generate the nucleophile and to juxtapose the nucleophile with the targeted bond. These distinctions provide a basis for classification scheme, and on this basis proteases can be grouped into four major classes: serine, cysteine, aspartate, and metallo. Aspartic and metalloproteases make use of aspartate residues and heavy metals, respectively, to immobilize and polarize a water molecule so that the oxygen atom in water is converted into the nucleophile[4]. Serine and cysteine proteases utilize their HO-and HS-side chains, respectively, directly as nucleophiles. The catalytic mechanisms of serine and cysteine proteases are extremely similar. In general, 3-D structures of these enzymes made up of two large globular domains surrounding a cleft consisting of the active site residues. Substrate entry into the cleft is a prerequisite for cleavage, and efficient entry is controlled sterically by the structural compatibility between the potential substrate and the cleft, a major deciding factor for enzyme specificity. The formation of a spatial compatibility between a targeted bond of the substrate and the active site nucleophile is also an important determinant of substrate specificity. Crystallographic studies of many members of the cysteine and serine class enzymes reveals comprehensive structure of the active site regions and the significance of additional amino acid residues to the catalytic process[5, 6].

Aspartyl proteases The Aspartic proteases (EC 3.4.23) are a family of hydrolytic/ proteolytic enzymes that share the similar catalytic machinery. Members of the aspartic protease family can be found in diverse organisms, ranging from microorganisms like bacteria and fungi, humans, and other higher as well as lower organism in systematics to plants and viruses. The best known sources of aspartic proteases are the mammalian stomach, yeast and fungi. Porcine pepsin A (EC 3.4.23.1) is the typical aspartic protease involving aspartic acid residue in the catalytic reaction. Pepsin is secreted as pepsinogen

containing 44 additional residues which termed as propetide, covalently attached to the N-terminus of pepsin. The propeptide is held in a pepsinogen binding cleft by ionic interactions between acidic and basic amino acid residues of the pepsin part of the protein. The active pepsin is produced by neutralization of the negatively charged residues in pepsin part of the molecule and subsequent release of propeptide by proteolytic cleavage and conformational rearrangement of the pepsin[7- 10]

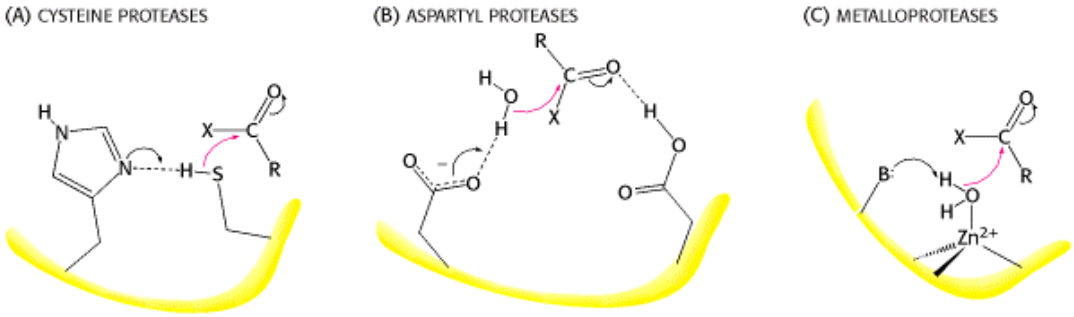


Figure 1: The activation strategies for three major classes of proteases: *The peptide carbonyl group is attacked by (A) a histidine-activated cysteine, in the cysteine proteases; (B) an aspartate-activated water molecule, in the aspartyl proteases; and (C) a metalactivated water molecule, in the metalloproteases. For the metalloproteases, the letter B represents a base (often a glutamate) that helps deprotonate the metal-bound water. Serine protease has similar mechanism of action to cysteine with the only exception that the OH-group of serine residue takes the responsibility of SH-group of cysteine residue in active site.*

Serine proteases: The serine proteases are a family of enzymes that cut certain peptide bonds in other proteins. This activity depends on a set of amino acid residues in the active site of the enzyme; one of which is always a serine (thus accounting for their name). In mammals, serine proteases perform many important functions, especially in digestion, blood clotting and complement system. There are three digestive enzymes secreted in intestine are serine proteases chymotrypsin, trypsin and elastase. These three share closely-similar structures (primary as well as tertiary structures). In fact, their active serine residue is at the same position (Ser-195) in all three. Despite their similarities, they have different substrate specificities; that is, they cleave different peptide bonds during protein digestion. The activating clotting factors are serine proteases such as factor X, factor IX, thrombin and plasmin. Many envolved complement factors are also serine proteases namely C1r, C1s, C3 convertases (C4B, C2a, and C3b) [11, 12].

Threonine proteases: Threonine proteases are proteases that have threonine, an amino acid, bonded at the active site[13]. It is responsible for functioning proteasome, the large protein-degrading apparatus[14]. Threonine proteases have a conserved N-terminal threonine at active site. Pre-proteins, which are catalytic beta subunits, are activated when the N-terminus is cleaved off. This makes threonine the N-terminal residue. Threonine proteases are activated by primary amines. The mechanism for the threonine protease was described first in 1995[14]. The mechanism showed the cleaving of a peptide bond which made an amino acid residue (usually serine, threonine, or cysteine) or a water molecule become a good nucleophile which could perform a nucleophilic attack on the carboxyl group of the peptide. The amino acid residue (in this case threonine) is usually activated by a histidine residue.

Metallo proteases: These are proteolytic enzymes whose catalytic activity depends on presence of metal ion involves a metal. Most of the metalloproteases are zinc-dependent, but, some of them require cobalt for their hydrolytic activity[15]. The metal ion is coordinated to the protein via three Histadine imidazole ligands. The fourth coordination position is taken up by a labile water molecule. The metalloprotease are subdivided into endo-metallopeptidases (E.C: 3.4.24) and exo-metallopeptidases (E.C: 3.4.17) based on the position of cleavage in the protein/ peptide chain. The bacterial enzyme thermolysin and the digestive enzyme carboxypeptidase A are classic examples of the zinc proteases. Thermolysin, but not carboxypeptidase A, is a member of a large and diverse family of homologous zinc proteases that includes the matrix metalloproteases, enzymes that catalyze the reactions in tissue remodeling and degradation. Treatment with chelating agents such as EDTA leads to complete inactivation. EDTA is a metal chelator which removes zinc and cobalt metal ion, essential for the activity, from the reaction medium.

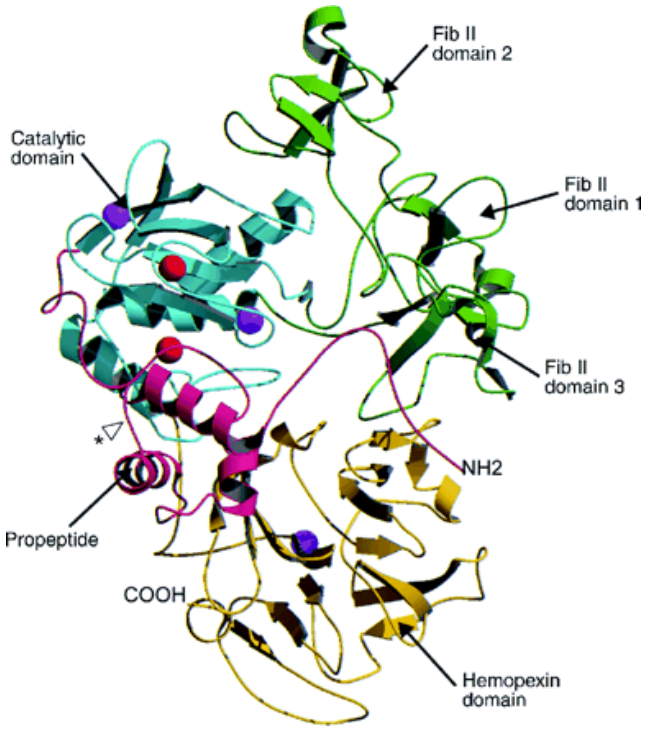


Figure 2: Structure of proMMP-2: The prodomain, catalytic domain, fibronectin domains, and hemopexin domain are shown in red, blue, green, and yellow, respectively. Zn^{2+} ions are indicated in red and Ca^{2+} ions are magenta (24). Asterisk indicates the cleavage site for MT1-MMP.[15]

Cysteine proteases:

In this class of enzymes, a cysteine residue, activated by a histidine residue, plays the role of the nucleophile that attacks the amide bond, in a fashion quite similar to that of the serine residue in serine proteases. Although, the generation of oxyanion or thiolate anion (nucleophile) are very important for the catalytic process in both the serine and cysteine proteases respectively, but the formation of these anions is dependent on ion pair formation between the active site amino acids and neighboring amino acid such as histidine[2, 16]. A model case of these proteins is Papain (EC 3.4.22.2), an enzyme purified from the fruit of the papaya. Mammalian proteases homologous to papain have been discovered, particularly the cathepsins, proteases having a role in antigen presentation[17, 18], bone remodeling[19], and other biological function. The cysteine-based active site arose independently at least twice in the course of evolution; the caspases, enzymes that play a major role in apoptosis[20], have active sites similar to that of papain, but their overall structures are not related.

Table 1:*[21] Clans and Families of Cysteine Peptidases

Clan	Family	Representative enzyme	Identified catalytic Residues
CA	C1	Papain	Cys/His
CA	C2	Calpain	Cys
CA	C10	Streptopain	Cys/His
CB	C3	Polio Virus Picornain 3C	His/Cys
CB	C4	Tobacco etch virus N1a endopeptidase	His/Cys
-	C18	Hepatitis C Virus endopeptidase	His/Cys
-	C5	Adenovirus endopeptidase	His/Cys
CC	C6	Tobacco etch virus HC-proteinase	Cys/His
CC	C7	Chestnut blight virus p29 endopeptidase	Cys/His
-	C8	Chestnut blight virus p48 endopeptidase	Cys/His
-	C9	Sindbis virus nsP2 endopeptidase	Cys/His
-	C16	Mouse hepatitis virus endopeptidase	Cys/His
-	C21	Turnip yellow mosaic virus endopeptidase	Cys/His
-	C11	Clostripain	Cys
-	C12	Deubiquitinating peptidase Yuh 1	Cys/His
-	C19	Deubiquitinating peptidase Ubp1	Cys/His
-	C13	Hemoglobinase	-
-	C14	Interleukin 1 β converting enzyme	Cys
-	C15	Pyroglutamyl-peptidase 1	Cys
-	C17	Microsomal ER60 endopeptidase	-
-	C20	Type IV prepilin leader peptidase	-

**Methods in Enzymology: Serine and Cysteine Peptidase (1994) 244: 1-765*

Papain superfamily (C1)

The papain family contains peptidases with a wide variety of activities, including endopeptidases with broad specificity (such as papain), endopeptidases with very narrow specificity (such as glycyl endopeptidases), aminopeptidases, a dipeptidyl-peptidase and peptidase with both endopeptidases and exopeptidases activities (such as cathepsins B and H) There are also family members that show no catalytic activity. Enzymes of this family have been identified in very diverse biological systems: like baculovirus, eubacteria (*Porphyromonas* and *Lactococuus*), yeast and probably all protozoa, plants and animals. The catalytic domains of papain-like cysteine proteases are between 220 and

260 amino acids in length. Exceptions are several parasite-derived cysteine proteases which contain C-terminal extension of unknown function.

In addition to the active site residues, several other amino acid sequence regions show a high degree of conservation. Among those are several cysteine residues which are involved in the formation of disulfide bridges, Pro2 which might prevent the N-terminal truncation and thus inactivation of the mature protease by aminopeptidases, the Gly-Cys-X-Gly-Gly motif forming a wall along the nonprime subsite binding area and the Gly-Pro motif separating a β and α domain at the interface between the major L and R domains.

The papain superfamily of cysteine proteases is composed of papain group, which included such stereotypical cysteine proteases as cathepsins B, C, H, L and S; and the bleomycin hydrolase and calpain groups. The papain group is the most abundant and complex that includes the mammalian lysosomal enzymes. The phylogenetic account of Papain group was prepared by Berti & Storer and is particularly useful for its alignment of 48 representative sequences of papain group members[22]. Papain is the model for the eponymous group and is purified for the first time from a plant product (the latex of *Carica papaya*). More interesting is that the members of the group are found in a broad range of life forms including fungi[23], plants, protozoans[24, 25, 26], invertebrates[27, 28], fish[29], and mammalian systems. Although, their main role is in protein degradation, the redundant nature of their function in many higher organisms makes their essentiality very ambiguous. Whereas, in simpler organisms, a cysteine protease may take on more key roles, such as the cathepsin L-like enzyme responsible for digestion in sponges[28] or cruzain from the protozoan *Trypanosoma cruzi*, which is necessary for parasite replication[25, 30]. Development of knockout and transgenic technologies have contributed to the isolation of new cysteine proteases with limited tissue distribution and distinct roles in most of the sophisticated pathways in cells tissue distributions have contributed to the discovery of more sophisticated roles for these enzymes such as peptide presentation in context of MHC II[31, 32]. As detection and identification techniques become progressively more refined and sensitive, more and more cysteine proteases that exhibit narrow tissue distributions are being discovered. Some of them will almost definitely function in critical physiological processes. Thus, structural information is important for understanding the specificities of how these proteins function and for the accurate blueprint of inhibitors to serve as therapeutics when activity leads to pathology.

Table 2: * Peptidases of Papain (Family C1)

Peptidase	EC	Database Code
Family C1:Papain		
Actinidin	3.4.22.14	ACTN_ ACTCH
Aleurain(barley)	-	ALEU_HORVU
Allergen (*Dermatophagoides)*	-	MMAL_DERPT
Allergen (*Eutoglyphus*)	-	EUM1_EURMA
Bleoomycin hydrolase	-	BLMH_YEAST
Calotropin (*Calotropis*)	-	CAL1_CALG1
Caricain	3.4.22.30	PAP3_CARPA
Cathepsin B	3.4.22.1	CATB_*,CYSP_SCHMA
Cathepsin H	3.4.22.16	CATH_*
Cathepsin L	3.4.22.15	CATL_*
Cathepsin S	3.4.22.26	CATS_*
Chymopapain	3.4.22.6	PAP2_CARPA
Cysteine aminopeptidase (*Lactococcus*)	-	(M86245)
Cysteine endopeptidase 2 and 3 (barley)	-	CYS1_HORVU,CYS2 CYS2_HORVU
Cysteine endopeptidase (*Brassica napus*)	-	CYS4_BRANA
Cysteine endopeptidase(*Caenorhabditis*)	-	CYS1_CAEEL
Cysteine endopeptidase 1 and 2 (*Dictyostelium*)	-	CYS_DICDI,CYS2_DICDI
Cysteine endopeptidase 1 and 2 (Entamoeba)	-	(M27307), (M64712) (M64721), M(94163
Cysteine endopeptidase 1 and 2 (*Haemonchus*)	-	(CYS_HAECO, CYS2_DICDI)
Cysteine endopeptidase (*Hemerocallis*)	-	(X74406)
Cysteine endopeptidase 1,2 and 3 (*Homarus*)	-	CYS1_HOMAM, CYS2_HOMAM CYS3_HOMAM
Cysteine endopeptidase (*Leishmania*)	-	LCPA_LEIME, (M97695), (Z14061)
Cysteine endopeptidase (mung bean)	-	CYSP_VIGMU
Cysteine endopeptidase (*Ostertagia*)	-	(M88505)
Cysteine endopeptidase (Pea)	-	CYSP_PEA, (X66061)
Cysteine endopeptidase (*Plasmodium*)	-	CSP_PLACM,CYSP_PLAFA (L08500), (L26362)
Cysteine endopeptidase (*Porphyromonas*)	-	TPR_PORGI
Cysteine endopeptidase (*Tertrahymena*)	-	(L03212
Cysteine endopeptidase (*Theileria*)	-	CYSP_THEPA, CYSP_THEAN
Cysteine endopeptidase (Tobacco)	-	(Z13959), (Z13964)
Cysteine endopeptidase (Tomato)	-	CYSL_LYCES,(Z14028)
Cysteine endopeptidase (*Trypanosoma*)	-	CYSP_TRYBR, (L25130), (M90067)
Dipeptidyl peptidase 1	3.4.14.1	CATC_RAT
Endopeptidase (baculovirus of Autographa	-	CYSP_NPVAC
Endopeptidase Ep-C1(*Phaseolus vulgaris*)	-	CYSP_PHAVU
Glycycl endopeptisase	3.4.22.25	PAP4_CARPA
Oryzin (includes forms α, β and ϒ)	-	ORYA_ORYSA, ORYB_ORYSA
Papain	3.4.22.2	PAPA_CARPA
Stem bromelain	3.4.22.32	BROM_ANACO
Thaumatopain (*Thaumatococcus*)	-	THPA_THADA

* Methods in Enzymology: Serine and Cysteine Peptidase (1994) 244:1-765

Regulation of cysteine proteases:

Proteases represent the group of enzymes having potentially dangerous activities that must be regulated and confined within suitable compartments, failure of which may lead to pathological consequences. Several other mechanisms are used to regulate protease activity endogenously, in addition to transcriptional and post-transcriptional controls[33, 34]. Most enzymes are synthesized as proenzymes, called as zymogens that must be activated by enzymatic hydrolysis. This may be achieved in autocatalytic mode under particular conditions, such as acidic pH. Release of proteases from cell also a controlled and regulated mechanism. The activity of protease with only limited proteolytic potential is regulated by balancing the amount of active enzyme and the amount of active endogenous inhibitors present in the physiological microenvironment. The disturbance in the regulation is caused by surplus or a deficit of enzyme relative to its inhibitors. However, the regulation of cysteine proteases seems to be more complex. Apart from the transcriptional and translational control of gene expression, a number of other factors also regulate the proteolytic activity of cysteine proteases:

1. **Compartmentalized confinement:** Most of the cysteine proteases are optimized to function in acidic compartments of cell.
2. **Redox potential**: The active site cysteine is quickly oxidized; therefore, these enzymes are optimally active in reducing conditions. Cell organelles accumulate cysteine to sustain such an environment[35].
3. **pH**: To prevent the indiscriminate action due to accidental secretion of cysteine proteases, all enzymes require proteolytic activation, cysteine proteases however, is regulated on an extra level by the requirement of low pH condition.
4. **Specific targeting:** Cysteine proteases are targeted to endosomes and lysosomes by N-mannosylation on the basis of phosphomannosyl residues, which promote binding to specific lysosomes targeting receptors.
5. **The presence of endogenous inhibitors:** These protease inhibitors appear to function to inhibit active enzyme that escapes compartmentalization. The cytoplasm and extracellular spaces are endowed with cysteine protease inhibitors in higher molar concentration than enzyme. Some cells, such as macrophages, are capable of recruiting the active enzymes within lysosomals to the cell surface under certain

circumstances[36]. An important phenomenon that the outer surface of the cell becomes an extra compartment during matrix remodeling, separating from endogenous protease inhibitors, by mobilizing acid, enzymes, and possibly other unknown factors in osteoclast[37, 38, 39], a bone macrophage.

Enzyme Inhibitors: As mentioned earlier the enzymes help to execute very significant biological reactions, which can be dangerous if left to act continuously. Therefore, there are mechanisms to regulate the enzyme activities in normal homeostatic conditions most importantly by specific endogenous inhibitors, for example, cystatins for papain like proteases, and 2-microglobulin used to control most of the serum aspartyl proteases. This type of inhibiting enzyme activity serves as a major control apparatus in biological worlds. The activity of enzymes can be inhibited by the binding of specific natural or synthetic organic small molecules. Inhibition by particular chemicals can be exploited to understand the mechanism of enzyme action. Specific inhibitors can be used to identify residues involved in catalysis process. In addition, many drugs and toxins act by inhibiting enzymes activities. Also, these natural mechanisms can be exploited to control the deregulated expression of the enzymes during pathological conditions.

Classification of enzyme inhibitors

The initial classification of inhibitors based on the binding of inhibitors with enzymes is given by

- **Active site directed** –When inhibitor binds with active site of enzyme
- **Allosteric effector**- When inhibitor binds other than active site of enzyme

The inhibitors which attack the active site can be divided according to the type of interaction with **covalent/ noncovalent** and **irreversible/reversible** inhibitors.

A. Reversible inhibitors: These kinds of inhibitor bind to enzymes with non-covalent interactions such as hydrogen bonds, hydrophobic interactions and ionic bonds. Multiple weak bonds between the inhibitor and the active site combine to produce strong and specific binding. In contrast to substrates and irreversible inhibitors, reversible inhibitors generally do not undergo chemical reactions when bound to the enzyme and can be easily removed by dilution or dialysis. There are three kinds of

reversible enzyme inhibitors. They are classified according to the effect of varying the concentration of the enzyme's substrate on the inhibitor.

- **Competitive inhibition**: The substrate and inhibitor compete for access to the enzyme's active site. The substrate and inhibitor cannot bind to the enzyme at the same time. This type of inhibition can be overcome by sufficiently high concentrations of substrate, i.e., by out-competing the inhibitor. Competitive inhibitors are often similar in structure to the real substrate.
- **Mixed inhibition**: The inhibitor binds to a different site on an enzyme. Inhibitor binding to this allosteric site changes the conformation (i.e., tertiary structure or three-dimensional shape) of the enzyme so that the affinity of the substrate for the active site is reduced. However, the binding of the inhibitor affects the binding of the substrate, and vice versa. This type of inhibition can be reduced, but not overcome by increasing concentrations of substrate. Although it is possible for mixed-type inhibitors to bind in the active site, this type of inhibition generally results from an allosteric effect where the inhibitor binds to a different site on an enzyme.
- **Non-competitive inhibition:** Non-competitive inhibitor is a form of mixed inhibition where the binding of the inhibitor to the enzyme reduces its activity but does not affect the binding of substrate. As a result, the extent of inhibition depends only on the concentration of inhibitor.

B. Irreversible inhibitors: These inactivators always bind to the enzyme covalently. Irreversible inhibitors react with the enzyme via a noncovalent transition state and thus lead to rapid reduction of enzyme activity. In practice it is often difficult to differentiate between reversible and irreversible inhibitors, for example if a reversible inhibitor binds to the enzyme with such high affinity that the enzyme-inhibitor complex only dissociates very slowly and thus appears irreversible. This type of inhibitor is known as “tight binding”. Normally a rapid equilibrium is observed with reversible inhibitors, whereas, reactions which result in modification of the enzyme take place relatively slowly.

Protease inhibitor:

The action of proteases is tightly controlled to prevent improper cleavage of signaling molecules. Protease activities are regulated at the transcriptional level by differential expression and at the protein level by activation of inactive zymogens and by the binding of inhibitors and cofactors. During cell and tissue development and organism homeostasis, the protease signaling pathways work normally and are tightly controlled. But what happens when the regulation of protease signaling fails? At the substrate-cleavage level, there is either too little or too much proteolysis. Diminished proteolysis as a result of insufficient protease activity mostly originates from genetic irregularities (endogenous proteases) excessive inhibitory activity or insufficient activation is carried out often by pathogens. By contrast, excessive or inappropriate proteolysis is seldom a result of genetic aberrations. But most often results from numerous endogenous and/or exogenous factors, which result in unwanted activation of protease signaling pathways, such as the effect of atherosclerotic plaque formation or blood vessel injury on the blood coagulation cascade, which leads to the appearance of intravascular thrombin. So far, inappropriate proteolysis has been found to have a major role in cancer as well as cardiovascular, inflammatory, neurodegenerative, bacterial, viral and parasitic diseases[40].

The enzyme inhibitors serve as probes for kinetic and chemical processes during catalysis that led to a detailed understanding of enzyme catalytic mechanisms and has provided effective therapeutic agents for the treatment of diseases. They help in elucidating the mode of ligand binding, where the ligand may be an inhibitor, substrate, or substrate analogue. Inhibitors are also used for *in vivo* studies to localize and quantify enzymes in organs or to mimic certain genetic diseases that involve the absence of an enzyme in given biosynthetic pathway.

Cysteine protease inhibitor

Due to its significant contribution in variety of processes, there is always a demand of low molecular weight as well as high molecular weight cysteine protease inhibitors.

Low-molecular weight inhibitors

The design and synthesis of cysteine protease inhibitors has a long history and has been extensively reviewed in recent years[40-46]. Compounds synthesized included a wide range of peptide aldehydes, methyl ketones, nitriles as reversibly acting inhibitors and diazomethanes, halomethyl ketones, acyloxymethyl ketones, *O*-acylhydroxamates, and epoxysuccinyl derivatives as irreversible inhibitors. Whereas, early developments of cysteine protease inhibitors provided useful tools to study the protease activity. Only recently significant progress has been accomplished to develop cysteine protease inhibitors into drugs. This development was mostly driven by the dramatic increase in understanding of papain like cysteine proteases as pharmaceutically valid targets. Cysteine protease inhibitors have been used to develop new drugs.

Aldehydes

Aldehydes and its analogues continue to be attractive moieties despite their well-established chemical reactivity. "Slow binding" inhibition is observed by aldehydes, i.e. the slow attainment of steady state during inhibition with a lag phase with a half life of several minutes, is not due to induction of conformational changes in the enzyme ("slow binding" in the usual sense) but is rather a result of the low concentration of free aldehyde in solution[47]. Development of peptidyl aldehydes as inhibitors of cysteine and serine proteases is based on the assumption that a tetrahedral intermediate is involved in enzymatic hydrolysis has led to investigation of the effect of carbonyl compounds on these proteases, with the intention of developing analogs of this transition state (e.g. Ac-Phe-Glyal) [48, 49]. These are leupeptines, chymostatins, antipain, elastinal, and α- MAPI[50] .

Acyclic and cyclic Ketones

Investigations of inhibition of the serine protease elastase by peptidyl aldehydes *in vivo* have shown that rapid loss of activity occurs if the aldehyde is oxidized to a carboxylic acid. For this reason, the aldehyde group was replaced by the metabolically more stable trifluoromethyl ketone function (TFMK). This produced significantly more potent inhibitors of serine proteases, also *in vitro*. It has been shown that a tetrahedral hemiketal

is formed as a covalent enzyme-inhibitor adduct. Many derivatives have been developed which inhibit different serine proteases[51-54].

Epoxysuccinyl analogues

In 1978, Hanada *et al.* succeeded in isolating a highly active, irreversible inhibitor of papain from culture extract of *Aspergillus japonicus*[55].The substance was identified as 1-[[*N*-(L-3-*trans*-carboxyoxiran- 2-carbonyl)-L-leucyl]amino]-4-guanidinobutane, E-64[56]. Systematic studies were carried out to investigate the role of the different structural components of the inhibitor in enzyme inhibition and the *trans*-L-(*S*,*S*)-epoxysuccinic acid was discovered to be the reactive group essential for inhibition[57]. A change of configuration of the epoxide residue or the neighboring amino acids reduces the activity by a factor of 10-100[58]. Structural information about the reverse mode of binding of the propeptide to active cathepsin B was utilized in designing novel, extended active site spanning E-64 peptidyl analogues. Selective cathepsin B or cathepsin L inhibitors were synthesized with the most active compounds spanning over six subsites (P4 to P2' residues)[59]. Peptidyl epoxides with a tyrosine and biotin moieties allowing iodination and streptavidin-based detection proved very valuable as functional proteomics tools[60].

Vinyl sulfones

Peptidyl vinyl sulfone inhibitors are remarkably potent irreversible inhibitors of cathepsins[61]. They have been shown to be effective in mice arthritis models by significantly reducing inflammation as well as bone and cartilage erosion[62]. However, due to the irreversible mode of action, vinyl sulfone inhibitors are unlikely to be developed as therapeutic drugs for chronic diseases such as osteoporosis. Structural analysis of the cruzipain inhibitor complex revealed a covalent Michael adduct with the active site cysteine residue and strong hydrogen bonding interactions in the S1' subsite[63]. The same compound was also orally effective in a mouse model of malaria[64, 65].

Nitriles

Recently, nonpeptidyl derivatives of nitriles employing pyrrolidine or azetidine rings have been demonstrated to be potent cysteine protease inhibitors. Interestingly, a four-membered ring derivative was approximately 10-fold more potent than the five-membered ring analogue which was possibly due to the increased chemical reactivity of

the azetidine ring. Nonpeptidyl nitrile acts as reversible, but time dependent inhibitors by forming a cleavable isothiourea ester link with the enzyme [66].

β-Lactams

The development of β-lactams as cysteine protease inhibitors is, however, very recent. Single rings as well as bicyclic ring β-lactam moieties have been evaluated as inhibitors of cysteine proteases. The β-lactam moiety serves as thiol reactive species and is linked with nonpeptidyl or amino acid or peptide portions targeting binding site pockets of relevant cathepsins. 2-Substituted oxapenams employing a nonpeptidyl aromatic or alkyl moiety as subsite motif displayed surprisingly potent inhibitory efficacy in the mid-nanomolar range. As expected, lacking a specific targeting moiety, these compounds were not selective. The mode of inhibition was time dependent and no recovery of enzymatic activity was observed. On the other hand, the incorporation of amino acid or peptidyl moieties as targeting sequence for individual proteases at the other side of the ring resulted in a significant enhancement of the potency and specificity of the compounds [67-69].

Diacyl Bis Hydrazides

Diacyl bis hydrazides evolved from previously reported diaminopropanones which have been developed as potent and selective cathepsin K inhibitors spanning both S and S' subsites of the substrate binding cleft [70]. The incorporation of bis aza analogues increased the potency while maintaining the selective profile for cathepsin K. The most potent compounds possess a leucine residue either in the P2 or P2' position or in both. Cyclic reversible cysteine protease inhibitor contains a peptidomimetic thiazole ring in place of an amide bond and may form an acyl adduct with the enzyme.

Proteinaceous cysteine protease inhibitors

The recent decade has witnessed tremendous development in the field of proteinaceous cysteine protease inhibitors. Though the prototype cystatins discovered in the 1960s remain the best-characterized group, several new large and a few smaller families are now recognized. Currently 10 are described, and the accompanying growth in a number of known inhibitors and processes involving these molecules is even higher. Most, if not

all, aspects of such an important activity as proteolysis are accompanied by inhibitors exerting regulatory or protective functions. Advances in genomic studies have already enabled rough estimates of the number of inhibitors and proteases produced by many organisms, however much remains to be done to understand the detailed function of these proteins. Moreover, some inhibitors constituting yet unknown new families might have been easily overlooked in such comparative studies. It is probable that such families still await discovery, especially since known proteinaceous inhibitors of cysteine proteases outside the papain family are scarce.

Two major inhibitor classes are known: (i) cystatins and (ii) serpins, which are active toward cysteine proteases. Cystatins have been implicated in a wide range of regulatory and disease-related processes such as in immune responses, cancer and immune evasion by parasites[71]. In plants, cystatin inhibits papain activity and is one of six identified *Arabidopsis* cystatin genes. Over expression of this cystatin in *Arabidopsis* cell cultures blocked cell death in response to avirulent bacteria and NO[72].

Cystatins

The cystatins are tight and reversibly binding inhibitors of the papain-like cysteine protease. They are present from protozoa to mammals. They are divided into three groups (types) based on distinct structural details, but also reflecting their distribution in the body and physiological roles in humans[73]. Type 1 cystatins are polypeptides with ~100 amino acid residues, which possess neither disulphide bonds nor carbohydrate side chains (cystatins A and B). They are found mainly intracellularly, but can also appear in body fluids at significant concentrations. They are also known as stefins. Type 2 encompasses cystatins C, D, E/M, F, G (CRES), S, SN and SA characterized by two conserved disulphide bridges, larger size (~120 residues), and a presence of a signal peptide for extracellular targeting.

Cystatins belonging to this group are found in most body fluids. Finally, type 3 contains Kininogens, high molecular mass proteins with three tandemly repeated type 2`-like cystatin domains, only two of which are able to inhibit cysteine proteases. Kininogens are intravascular proteins of blood plasma.

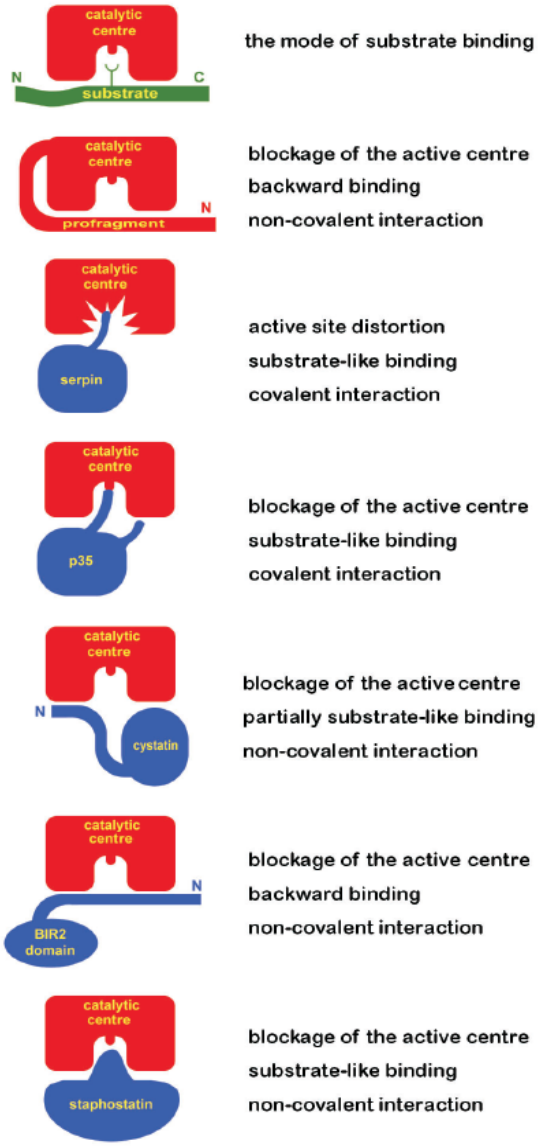


Figure 3. *Schematic representation of inhibitory mechanisms directed against cysteine proteinases. Inhibitors are shown in blue, enzymes in red and substrate in green.*

Phytocystatins

Studies on phyto-cystatins started in 1987 with the cloning of the first phyto-cystatin from rice seed[74]. Subsequent studies in other species were based on this oryzacystatin probe and focused mainly on seeds. As a consequence, most of the newly discovered phyto-cystatins were highly homologous to oryzacystatin and very few results concerned the occurrence of phyto-cystatins in other parts of the plant. Generally, their amino acid sequences show higher similarity with the family II cystatins. A phyto-cystatin-specific consensus sequence has also been defined[75] 17 amino acids after the signature of the cystatin super-family: [LVI]-[AGT]-[RKE]-[FY]-[AS]-[VI]-X-[EDQV]-[HYFQ]-N.

Phyto-cystatins present an interest not only for fundamental studies aiming at the understanding of their role in plant development and defence against pathogens, but also for applied research in a variety of fields. The use of cystatins as bio-pesticides has been successfully tested against a large collection of plant pathogens[76]. Oryzacystatin I has been tested for its antiviral action against the HSV-1 virus and its antiherpetic effect was similar to that of acyclovir[77]. Therefore, phyto-cystatins can be considered as promising candidates in crop protection as well as in human and veterinary medicine.

Other in-depth investigations concerned the role of phyto-cystatins in plant defence mechanisms. They revealed an induction of phyto-cystatins by cell damage, fungal infection or jasmonic acid treatment[78]. A role of the phytocystatins in the defence against pathogens has been described for virus[79], fungus[80], nematodes, mites [81] and insects[82, 83]. *Pernas et al.* also found that the transcript level of a chestnut cystatin was induced by abiotic stress, namely by cold stress, saline shock or heat stress. Consequently, a key function of phyto-cystatins lies in the defence against biotic and abiotic stress. Another documented role of phyto-cystatin concerns programmed cell death. *Solomon et al*. [84] demonstrated the regulation of programmed cell death (PCD) by phytocystatins. The results suggest their potential implication in xylogenesis, senescence and hypersensitive response.

Inhibition of legumain and calpain: In general, cystatins are characterized as inhibitors of C1 family proteases; however, separate cases of cystatin fold adaptation to the inhibition of cysteine proteases belonging to other families have been reported. Legumain, a C13 family cysteine protease, is potently inhibited by cystatins C, E/M and, with a higher K_i by cystatin F. Interestingly, it has been demonstrated that the binding site for legumain is distinct from that of papain. A ternary complex may be formed allowing the cystatin to inhibit both proteases simultaneously[85]. Moreover, the main physiological function of cystatin E/M seems to be legumain and not C1 family protease inhibition (see above). The calcium-dependent cysteine proteases – calpains (family C2) are inhibited by the second cystatin domain of Kininogens. Due to exclusively intracellular localization of calpains, the ability of intravascular Kininogens to inhibit those enzymes is thought to serve as a protection against their activity at accidental release in pathological states.

Clitocypin: Clitocypin has been isolated in large quantities from fruit bodies of an edible mushroom *Clitocybe nebularis* by affinity chromatography on carboxymethylpapain Sepharose. It is the first proteinaceous inhibitor characterized from higher fungi.

Staphostatins: Specific inhibitors of staphylococcal cysteine proteases (staphostatins) constitute yet another, recently described novel inhibitor family. Unlike chagasins and clitocypin, however, detailed structural characterization has been accomplished, and clues for the function are available. Three members of the family have been described

staphostatins A and B from *Staphylococcus aureus* and staphostatin A from *Staphylococcus epidermidis* [86].

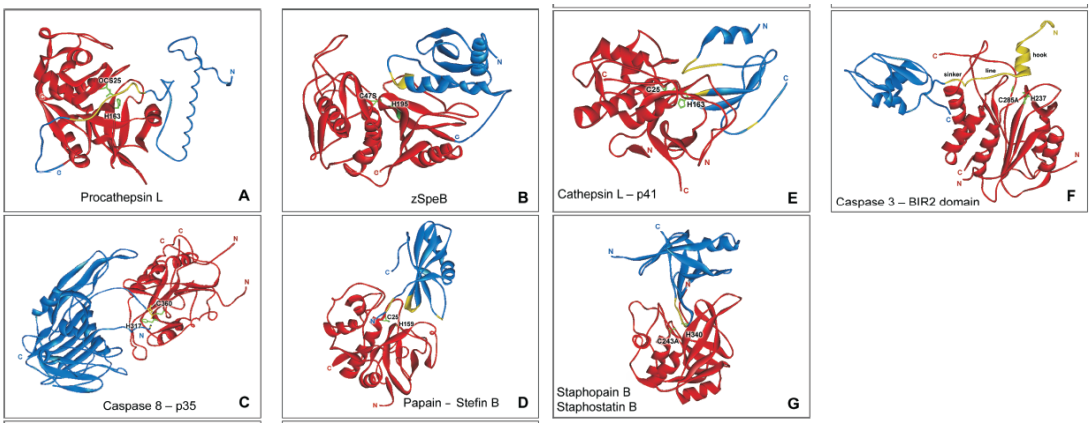


Figure 4. Ribbon drawing of inhibitory complexes formed by A, procathepsin L (PDB code: 1cs8); B, SpeB zymogen (1dki); C, p35 and caspase 8 (1i4e); D, stefin B and papain (1stf); E, p41 and cathepsin L (1icf); F, XIAP and caspase 3 (1i3o); G, staphostatin B and staphopain B (1pxv). Inhibitors are shown in blue, and enzymes in red. The location of the catalytic Cys and His residues is shown in green; regions of inhibitor that are most important for interaction with enzyme are labeled yellow.*
**Adapted from, Rzychon, M., Chmiel D. and Stec-Niemczyk J. Acta. Biochemical Polonica (2004) 51:861–873.*

Chagasin family: Protozoan parasites transmitted by insect bites, such as *Trypanosoma cruzi* (a causative agent of Chagas' heart disease in Latin America), *Trypanosoma brucei* (responsible for sleeping sickness in sub-Saharan Africa), *Leishmania* sp. (*leishmaniasis*) and *Plasmodium* sp. (malaria) produce large quantities of papain-like (family C1) lysosomal proteases. In plants and mammals, proteases like – cathepsins are regulated by endogenous inhibitors of the cystatin family. A search for similar regulators of parasite cysteine proteases resulted in the discovery of a novel inhibitor family distinct from cystatins and other known groups. Chagasin is a *T. cruzi* inhibitor of the endogenous cysteine protease cruzipain[87]. Subsequently, homologues were identified in the genomes of all mentioned parasites as well as in bacteria [88]. While genomes of all but one of those organisms encode cruzipain-like proteins, the function of *Pseudomonas aeruginosa* chagasin homologue remains puzzling since the bacterium does not seem to produce any family C1 proteases. Some of the identified proteins were recombinantly expressed and demonstrated to inhibit papain like cysteine proteases [89].

Calpastatin: Calpastatin is a specific inhibitor of calcium-dependent cysteine proteases – calpains. It does not inhibit any other cysteine proteases tested to date. Likewise, activity of calpains, is regulated only by calpastatin. (Though Kininogens are also able to inhibit calpain proteases, their extracellular localization excludes them as regulators of intracellular calpains, and they are thought to serve only as protectants against accidental calpain release [90].

Thyropins (thyroglobulin type-1 domain proteinase inhibitors): Thyroglobulin type-1 domain (Thyr-1) is a cysteine-rich structural element found either as a single or repetitive module in a variety of functionally unrelated proteins such as a precursor of thyroid hormones, thyroglobulin, basement membrane protein, nidogen (entactin), testicular proteoglycan, testican, insulin-like growth factor binding proteins (IGFBPs), pancreatic carcinoma marker proteins (GA733), major histocompatibility complex class II (MHC II)-associated p41 invariant chain, equistatin, chum salmon egg cysteine protease inhibitor (ECI) and saxiphilin [91-94].

MHC class II-associated p41-invariant chain fragment: A crucial function of MHC II is to present bound peptides, derived mainly from extracellular antigens, to CD4+ T lymphocytes. The invariant chain (Ii) associates with MHC class II molecules in endoplasmic reticulum, its main function being to block the peptide-binding site during early stages of intracellular transport. Upon arrival at endosomes/lysosomes, Ii, invariant chain is proteolytically degraded, thus allowing the MHC II to bind peptides derived from endocytosed extracellular antigens. Both antigen processing by means of limited proteolysis and Ii degradation are mediated by endosomal proteases. A protein fragment corresponding to p41 Thyr-1 domain was demonstrated to efficiently inhibit cathepsin L, cruzipain and to a lesser extent cathepsin H as well as papain[95]. The inhibitory properties towards antigen processing proteases implicate possible modulation of antigen presentation by alternative splicing of Ii. Different lines of evidence in favor of the invariant chain exerting inhibitory and modulatory activity in vivo are delineated below. Proportions of p41 in relation to p31 vary in different cells, suggesting a probable way to regulate antigen presentation. The former can modify the degradation pattern of the latter, demonstrating direct interference with proteolytic pathways.

IAP family: Inhibitors of apoptosis (IAPs) are a family of proteins distinguished by encompassing one or more characteristic, ~70-residue zinc binding BIR (baculovirus IAP repeat) domains. They were primarily characterized as inhibitors of apoptosis[96], though currently some BIR-containing proteins are known which do not seem to confer such a function. Apoptosis (programmed cell death) is one of the crucial events controlling the number of cells in multicellular eucaryotes.

Propeptide-like inhibitors: Papain-like cysteine proteases are produced as inactive precursors. The lack of activity is due to potent inhibition by the N-terminal propeptides. Only at the target location does limited proteolysis of the propeptide occur, allowing the enzyme to exert its action. Fox and colleagues[97] were the first to show that synthetic proregions inhibit their cognate proteases. Later studies demonstrated that propeptides are competitive, slow-binding inhibitors and unlike the broad-spectrum cystatins, possess high selectivity for the enzymes from which they originate. The inhibition is pH dependent (acidification decreases the affinity) being in good agreement with the process of enzyme activation, since pH is one of the most common environmental parameters in triggering the process. Based on structural data the mode of inhibition was elucidated. The propeptide spans the protease active site in an orientation opposite to that of a substrate, thus escaping cleavage. Besides being inhibitors, the propeptides assist in proper folding and targeting, and stabilize the cognate enzymes [98].

Natural products: Natural products are mostly secondary metabolites that have various roles in host organism. It is considered that these compounds are not essential for normal growth and development of the organism. But, the generation of natural products requires large amount of energy, therefore, it must have some role in the survival of the host organism. In agreement with this opinion, it has been anticipated that they protect the food source from competitors[99]. Microbial secondary metabolites are compounds of low molecular weight (<3000) and are accumulated after vegetative growth has ceased, as families of structurally related components. They are primarily produced by a comparatively limited group of bacteria and fungi, but their intergeneric, interspecific, and intraspecific variation is extremely high[100]. Many of them are armed with special biological activities, such as antibiotics, toxins, ionophores, bioregulators, and intra- and

interspecific signaling[101, 102]. The secondary metabolites have also been suggested to be possible waste products or shunt metabolites, but this is hard to reconcile with the complexity of their biosynthetic pathways and of the underlying genetic information[103]. They may act at the low concentrations present in the environment as chemical signals to regulate metabolic processes in bacteria by stimulating or suppressing gene expression at the transcriptional level and thereby influencing population composition and dynamics[104, 101]. For thousands of years the natural products are being used as traditional medicine and poisons. Further studies such as chemical, pharmacological and clinical, led to the discovery of most early medicines, aspirin, digitoxin, morphine, quinine and pilocarpine[105-109]. The discovery of antibacterial culture supernatant, penicillin, by Fleming in 1928 and commercialization of synthetic penicillin revolutionized drug discovery programs, leading to the discovery of streptomycin, chloramphenicol, chlortetracycline, cephalosporin C, erythromycin and vancomycine[110-115]. Use of mechanism-based bioassay guided fractionation and other improvement in screening format lead to the discovery of alpha lactamase inhibitor, clavulanic acid from *Streptomyces clavuligerus*[114]and HMG-CoA reductase inbhibitor, mevastain from *Penicilium citrinum*[115]. Natural products have been isolated from various plant sources i.e. phloroglucinol-monoterpene adducts from *Eucalyptus grandis*[116], Immunostimulatory compounds from *Vitex negundo*[117], Shogaols from *Zingiber officinale* as promising anti-fouling agents[118]. There are few good reviews have been published which deals with various natural products isolation from plant sources[119, 120]. Bhutani *et.al.* have presented a detailed review of Anti HIV natural products[121].

Table 3. Top 35 Worldwide Ethical Drug Sales for 2000, 2001, and 2002[122], with Natural Product-Derived Drugs in Blue, Biologically Derived Drugs in Red, and Synthetically Derived Drugs in Black.

Rank	2000	2001	2002
1	Omeparazole	Atorvastatin	Atorvastatin
2	Atorvastatin	Omeperazole	Simvastatin
3	Simvastatin	Simvastatin	Omeparazole
4	Amlodipine	Lansoprazole	Erythropoietin (J&J)
5	Lansoprazole	Amlodipine	Amlodipine
6	Loratadine	Erythropoietin (J&J)	Lansoprazole
7	Erythropoietin (J&J)	Loratadine	Olanzapine
8	Celecoxib	Celecoxib	Paroxetine
9	Fluoxetine	Olanzapine	Celecoxib
10	Olanzapine	Paroxetine	Srttraline
11	Paroxetine	Sertraline	Interferon α 2 b + Ribarvarin
12	Sertraline	Metformin/Metformin+Glyburide	Rofecoxib
13	Rofecoxib	Rofecoxib	Salmeterol+Fluticasone propionate
14	Erythropoietin (Amgen)	Erythropoietin (Amgen)	Gabapentin
15	Metformin/Metformin+Glyburide	Pravastatine (BMS)	Pravastatin (BMS)
16	Estrone	Estrone	Erythropoietin (Amgen)
17	Amoxicillin +Clavulanic acid	Amoxicillin +Clavulanic acid	Alendronate Sodium
18	Enalapril	Fluoxetine	Losartan/Losartan+Hydrothiazide
19	Parvastatin (BMS)	Risperidone	Risperidone
20	Insulin	Losartan/Losarton+Hydrothazide	Venlafaxine
21	Ciprofloxacin	Insulin	Esomeprazole magnesium
22	Losartan/Losarton+Hydrothazide	Ciprofloxacin	Fexofenadine
23	Pravastatin (Sankyo)	Gabapentin	Clopidogrel bisulfate
24	Risperidone	Alendronate sodium	Insulin
25	Paclitaxel	Leuprolide acetate	Estrone
26	Leuprolide acetate	Fexofenadine	Loratadine
27	Azithromycin	Venlafaxine	Amoxicillin +Clavulin acid
28	Interferon α 2 b+ Ribarvarin	Sidenafil	Valsartan
29	Sidenafil	Azithromycin	Valsartan
30	Gabapentin	Interferon α 2 b + Ribarvarin	Citalopran hydrobromide
31	Fluticasone propionate	Pravastatin (Sankyo)	Leuprolide acetate
32	Clarithromycin	Filgrastim	Oxycodone HCl
33	Filgrastim	Fluticasone propionate	Azithromycin
34	Cyclosporin	Enoxaparin	Montelukast sodium
35	Lisinopril	Vaccines (Aventis)	Rituximab

Status of natural products in drug discovery today: The status of natural product drug discovery has been extensively reviewed by Mark S. Butler[122]. In which the author has further reviewed the work of Newman *et al.* who analyzed a number of natural product derived drugs in the total drugs launched during the year 1981-2002[123, 124]. They concluded that NPs were still a significant source of new drugs, especially in the anticancer and antihypertensive therapeutic areas.[125] In another study, Proudfoot reported that 8 out of 29 small molecule drugs launched in 2000 were derived from NPs (Table: 4) or hormones and concluded that latest high throughput technologies such as HTS did not have a significant impact on the derivation of these drugs.[126] NP-derived drugs are well represented in the top 35 worldwide selling ethical drug sales of 2000, 2001, and 2002 (table 3).

Table 4: Drugs derived from natural products launched in Europe, Japan and the United States 2001–2005.[127]

Year	Generic name(trade name)	Natural product	Indications
2001	Caspofungin (Cancidas1)	Pneumocandin B	Antifungal
2001	Pimecrolimus (Elidel1)	Ascomycin	Atopic dermatitis
2001	Telithromycin (Ketek1)	Erythromycin	Antibacterial
2002	Amrubicin hydrochloride (Calsed1)	Doxorubicin	Anticancer
2002	Biapenem (Omegacin1)	Thienamycin	Antibacterial
2002	Ertapenem (Invanz™)	Thienamycin	Antibacterial
2002	Fulvestrant (Faslodex1)	Estradiol	Anticancer
2002	Galantamine (Reminyl1)	Galantamine	Alzheimer's disease
2002	Micafungin (Funguard1)	FR901379	Antifungal
2002	Nitisinone (Orfadin1)	Leptospermone	Antityrosinaemia
2003	Daptomycin (Cubicin™)	Daptomycin	Antibacterial
2003	Miglustat (Zavesca1)	1-deoxynojirimycin Type 1	Gaucher disease
2003	Mycophenolate sodium (Myfortic1)	Mycophenolic acid	Immunosuppression
2003	Pitavastatin (Livalo1)	Mevastatin	Dyslipidemia
2003	Rosuvastatin (Crestor1)	Mevastatin	Dyslipidemia
2004	Everolimus (Certican™)	Sirolimus	Immunosuppression
2004	Talaporfin sodium (Laserphyrin1)	Chlorophyll and L-aspartic acid	Anticancer
2005	Doripenem (Finibax1)	Carbapenem	Antibacterial
2005	Extenatide (Byetta1)	Incretin	Anti-diabetic
2005	Paclitaxel nanoparticles (Abraxane1)	Taxol	Anticancer
2005	Pramlintide acetate (Symlin1)	Amylin	Anti-diabetic
2005	Tigecycline (Tigacil1)	Tetracycline	Antibacterial
2005	Zicontide (Prialt™)	MVIIA	Pain management

Adapted from Butler, M. S. (2004) *J. Nat. Prod.* 67, 2141-2153

Further efforts are needed to in isolation of new drug leads from natural product: In order for natural product drug discovery to continue to be successful, new approaches are required. By applying the new approaches such as latest development in structure elucidation by various spectroscopic techniques, in an efficient fashion to natural product drug discovery, it might be possible to increase the current efficiency in identifying and developing new drugs leads from natural products. Only a small fraction of the microbial world has been explored and the exponential growth in the microbial genomic database, the scope for discovering new biologically active natural products is promising. After identifying valid natural product leads, applying new organic synthetic methodologies, biotransformation, combinatorial biosynthesis and combinations of these techniques for the modification of natural product leads would generate a large number of novel, structurally diverse analogs that can be screened for improved properties or new activities. We currently have the proper tools to prepare analogs and to discover the structure–activity relationships that would enable us to truly harness the potential of natural products and enhance the success of natural products in drug discovery. Because attention has now shifted to smaller high-quality libraries, natural product leads are a legitimate starting template for combinatorial chemistry. Novel biologically active analogs with improved properties or new activities can be discovered by this process. The same combinatorial chemistry that initially caused the decline in natural product screening is now an essential tool for generating analogs of natural products. By coupling these approaches with the advances in purification, structure elucidation, and reorganization of the screening process, the timeline for natural products drug discovery is shortened similar to that expected for synthetic compounds. Despite the lack of effort by most of the large pharmaceutical companies, natural product research has still been active during the past decade. Therefore, the question that should be asked is not whether we should screen natural products for drug discovery, but whether we can afford not to screen natural products.

Microorganisms as source of small molecular mass CPI: The majority of known microbial antibiotics (65%) in addition to a number of other commercially important pharmaceuticals and agrochemicals are produced by members of the *actinomycetes*, especially those from the *Streptomyces* genus[128, 129]. These compounds encompass most natural product classes, including L-lactams, peptides, polyketides, oligosaccharides,

terpenes and alkaloids. Several low molecular mass protease inhibitors such as leupetin, antipain, pepstain have been isolated from streptomyces[130]. One more irreversible cysteine protease inhibitor named as E-64 which is used in academic and commercial research laboratories have been isolated from a fungus. One novel antibiotic, bovicine-HJ 50 has been purified and charrecterized from *Streptococcus sp*[131]and antimicrobial peptides Subpeptin JM4-A and Subpeptin JM4-B, have been isolated from *Bacillus subtilis*[131]. Many phytochemicals have been isolated and characterized from natural products that have been reported to be harboring anti microbial properties.[132, 133, 134]. In spite of having these inhibitors from microorganisms, there is much which still needs to be done in order to get new drug leads.

Present Investigation: The present investigation is an effort to find out new small molecular mass cysteine protease inhibitor from natural source. We have selected the *Streptomyces sp* NCIM2081 for isolation and characterization cysteine protease inhibitor, using activity based assay. We have isolated and elucidated the complete chemical structure, which is a modified peptide, CPI-2081. We have also studied the energetics of binding of CPI-2081 to the pharmacologically important thiol proteases, cathepsin-K and cathepsin-L. Furthermore, the efficacy of the CPI-2081 has also been evaluated on tumor cells and osteoclast.

REFERENCES

1. Neurath, H.and Walsh, K. A. (1976) *Proc. Natl. Acad. Sci. USA* 73:3825–32
2. Polgar, L. (1989) *Mechanisms of Protease Action*, ed. L Polar, 43–76. Boca Ratan, FL: CRC Press
3. Polgar L, ed. (1989) In *Mechanisms of Protease Action,* 208–210. Boca Ratan, FL: CRC Press
4. Menard, R.and Storer, A. (1992) *Biol. Chem.*373,393–400.
5. Matthews, B. W.; Sigler, P. B.; Henderson, R.and Blow, D. M. (1967) *Nature* 214:652–56.
6. Varughese, K. L.; Ahmed, F. R.; Careys, P. R.; Hasnain, S.; Huber, C. P.and Storer, A. C. (1989) *Biochemistry* 28:1330–32
7. Sielecki, A.R.; Fujinaga, M.; Read, R. J.and James, M.N.G. (1991) *J. Mol. Biol.*, 219, 671–692.
8. Hartsuck, J. A.; Koelsch, G.and Remington, J. S. (1992) *Proteins*, 13, 1–25.
9. Glick, M. D.; Auer,H.E.; Rich, D.H.; Kawai, M.and Kamath, A. (1986) *Biochemistry*, 25, 1858–1864.
10. Glick, M. D.; Hilt, C. R.and Mende-Mueller, L. (1991) *Int. J. Pept. Protein Res.*, 37, 230–235.
11. Borensztajn, K.; Stiekema, J.; Nijmeijer, S; Reitsma, P. H.; Peppelenbosch, M. P.and Spek, C. A. (2008) *Am J Pathol.* 172, 309-20.
12. Cohen, G. H.; Silverton, E. W.and Davies, D. R.(1981) *J Mol Biol.*148, 449-79
13. Ekici, O. D.; Paetzel, M.and Dalbey, R. E.(2008)*Protein Sci.* 17, 2023-2037.
14. Seemuller, E.; Lupas, A.; Stock, D.; Lowe, J.; Huber, R.and Baumeister, W. (1995)*Science***,** 268, 579 – 582.
15. Morgunova, E.; Tuuttila, A.; Bergmann, U.; Isupov, M.; Lindqvist, Y.; Schneider, G.and Tryggvason, K. (1999) *Science,* 284, 1667 – 1670.
16. Mason, R. W.and Wilcox, D. (1993) *Adv. Cell Mol. Biol. Membr.* 1, 81–116.
17. Wolf, P. R.and Ploegh, H.L. (1995) *Annu Rev Cell Biol,* 11, 267–306.
18. Watts, C. (1997) *Annu Rev Immunol*,15, 821–850.
19. Troen, B. R. (2004)*Drug News Perspect.*17, 19-28.

20. Harvey, N. L.and Kumar, S (1998) *Advances in Biochemical Engineering/ Biotechnology, 62, 107-128.*
21. Barrett A. J. (1994) *Methods Enzymol.* 244, 1-15.
22. Berti, P. J.and Storer, A. C.(1995) *J. Mol. Biol.* 246, 273–83.
23. Sharma, A.; Padwal-Desai, S.and Ninjoor, V. (1989)*Biochem. Biophys. Res. Commun.*159, 464–71.
24. Bart, G.; Frame, M. J.; Carter, R.; Coombs, G. H. and Mottram, J. C. (1997) *Mol. Biochem. Parasitol.* 88, 53–61.
25. Hart, G.; Andrews, N.; Mills, A. A.; Engel, J.; Smith, R.(1993) *Mol. Biochem. Parasitol.* 58,17–24.
26. Rosenthal, P. J.(1995) *Exp. Parasitol.* 80, 272–81.
27. Caffrey, C. R. and Ruppel, A. (1997) *J. Parasitol.* 83,1112–18.
28. Krasko, A.; Gamulin, V.; Seack, J.; Steffen, R.and Schroder, H. C. (1997) *Mol. Mar. Biol. Biotechnol.* 6, 296–307.
29. Heu, M. S.; Kim, H. R.; Cho, D. M.; Godber, J. S.and Pyeun, J. H. (1997) *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* 118, 523–29.
30. Meirelles, M. N. L.; Juliano, L.; Carmona, E.; Silva, S. G.and Costa, E. M.(1992) *Mol. Biochem. Parasitol.* 52, 175–84.
31. Nakagawa, T.; Roth, W.; Wong, P.; Nelson, A.and Farr, A. (1998) *Science,* 280, 450–53.
32. Riese, R. J.; Mitchell, R. N.; Villadangos, J. A. ; Karp, E. R. and DeSanctis, G. T.; (1998) *J. Clin. Invest.*101, 2351–2363.
33. Twining, S. S.; Fukuchi, T.; Yue, B. Y.; Wilson, P. M.; Zhou, X.and Loushin, G. (1994) *Invest Ophthalmol Vis Sci.*35, 3226-33.
34. Chapman, H. A.; Riese, R. J.and Sh*i* G. *(*1997)*Annu. Rev. Physiol. 1997. 59:63–88.*
35. Pisoni, R. L.; Acker, T. L.; Lisowski, K. M.; Lemons, R. M.and Theone, J. G.(1990) *J. Cell Biol.* 110, 327–35.
36. Reddy, V. Y.; Zhang, Q. Y.and Weiss, S. J.(1995) *Proc. Natl. Acad. Sci. USA* 92:3849–53.
37. Baron, R.; Neff, L.; Louvard, D.and Courtoy, P. J. (1985) *J.Cell Biol.* 101, 2210–28.

38. Baron, R. (1989) *Anat.Rec.* 224:2317–429.
39. Dalaisse, J. M.; Eeckhout, Y. and Vaes, G. (1980) *Biochem. J.* 192:365–68.
40. In Handbook of Proteolytic Enzymes, 2 Edn (Barrett, A.J., Rawlings, N.D. & Woessner, J.F. eds), 1886-1887, Elsevier, London (2004) V
41. Lecaille, F.; Kaleta, J. and Bro¨mme, D. (2002) *Chem. Rev.* 102: 4459-4488
42. Jasmer, D. P.; Roth, J. and Myler, P. (2001) *J. Mol. Biochem. Parasitol.* 116: 159-162.
43. Falcone, F. H.; Tetteh, K. K.,;Hunt, P.; Blaxter, M. L.; Loukas, A. and Maizels, R. M. (2000) *Exp. Parasitol.* 94, 201-205.
44. Pierre, P. and Mellman, I. (1998) *Cell* 93, 1135-1137.
45. Kos, J. and Lah, T. T. (1998) *Oncol. Rep.* 5, 1349.
46. Maizels, R. M.; Gomez-Escobar, N.; Gregory, W. F.; Murray, J. and Zang, X. (2001) *Int. J. Parasitol.* 31, 889-894.
47. Schultz, R.; Varma-Nelson, P.; Oritz, R.; Kozlowski, K.; Orawski, A.; Pagast, P. and Frankfater, A. (1989) *J. Biol. Chem* 264, 1497- 1507.
48. Westerik, J. and Wolfenden, R. (1972) *J. Biol. Chem.* 247, 8195- 8197.
49. Mattis, J.; Henes, J. and Fruton, J. (1977) *J. Biol. Chem.* 252, 6776- 6782.
50. Frommer, W.; Junge, B.; Muöller, L.; Schmidt, D. and Truscheit, E. (1979) *Planta Med.* 35, 195-215.
51. Imperiali, B. and Abeles, R. (1986) *Biochemistry* 25, 3760-3767.
52. Dunlap, R.; Stone, P. and Abeles, R. (1987) *Biochem. Biophys. Res. Commun.* 145, 509-513.
53. Stein, R.; Strimpler, A.; Edwards, P.; Lewis, J.; Mauger, R.; Schwartz, J.; Stein, M.; Trainor, D.; Wildonger, R. and Zottola, M. (1987) *Biochemistry* 26, 2682-2689.
54. Peet, N.; Burkhardt, J.; Angelastro, M.; Giroux, E.; Mehdi, S.; Bey, P.; Kolb, M.; Neises, B. and Schirlin, D. (1990) *J. Med. Chem.* 33, 394-407.
55. Hanada, K.; Tamai, M.; Yamagishi, M.; Ohmura, S.; Sawada, J. and Tanaka, I. (1978) *Agric. Biol. Chem.* 42, 523-528.
56. Hanada, K.; Tamai, M.; Ohmura, S.; Sawada, J.; Seki, T. and Tanaka, I. (1978) *Agric. Biol. Chem.* 42, 529-536.

57. Hanada, K.; Tamai, M.; Morimoto, S.; Adachi, T.; Ohmura, S.; Sawada, J. and Tanaka, I. (1987) *Agric. Biol. Chem*. 42, 537-541.
58. Barrett, A.; Kembhavi, A.; Brown, M.; Kirschke, H.; Knight, C.; Tamai, M. and Hanada, K. (1982) *Biochem. J.* 201, 189-198.
59. Schaschke, N.; Assfalg-Machleidt, I.; Machleidt, W.; Moroder, L. (1998) *FEBS Lett*. 421, 80-84.
60. Greenbaum, D.; Baruch, A.; Hayrapetian, L.; Darula, Z.; Burlingame, A.; Medzihradszky, K.F. and Bogyo, M. (2002) *Mol. Cell. Proteomics* 1: 60–68.
61. Brinen, L. S.; Hansell, E.; Cheng, J.; Roush, W. R.; McKerrow, J. H. and Fletterick, R. (2000) *J. Struct. Fold Des.* 8, 831-834.
62. Rosenthal, P. J.; Olson, J. E.; Lee, G. K.; Palmer, J. T.; Klaus, J. L. and Rasnick, D. (1996) *Antimicrob. Agents Chemother*. 40, 1600-1607.
63. Olson, J. E.; Lee, G. K.; Semenov, A. and Rosenthal, P. J. (1999) *Bioorg. Med. Chem.* 7, 633-637.
64. Falgueyret, J. P.; Oballa, R. M.; Okamoto, O.; Wesolowski, G.; Aubin, Y.; Rydzewski, R. M.; Prasit, P.; Riendeau, D.; Rodan, S. B. and Percival, M. D.(2001) *J. Med. Chem.* 44, 94-96.
65. Greenspan, P. D.; Clark, K. L.; Tommasi, R. A.; Cowen, S. D.; McQuire, L. W.; Farley, D. L.; van Duzer, J. H.; Goldberg, R. L.; Zhou, H.; Du, Z.; Fitt, J. J.; Coppa, D. E.; Fang, Z.; Macchia, W.; Zhu, L.; Capparelli, M. P.; Goldstein, R.; Wigg, A. M.; Doughty, J. R.; Bohacek, R. S. and Knap, A. K. (2001) *J. Med. Chem.* 44, 4524-4529.
66. Moon, J.; Coleman, R. and Hanzlik, R. (1986) *J. Am. Chem. Soc.* 108, 1350-1351.
67. Singh, R.; Zhou, N. E.; Guo, D.; Kaleta, J.; Cameron, A.; Purisima, E.; Menard, R. and Micetich, *R. G. WO 97/38008,* 1997.
68. Singh, R.; Zhou, N. E.; Guo, D.; Kaleta, J. and Micetich, R. G. US 5,925,633 A 1999.
69. Singh, R.; Zhou, N. E.; Guo, D. and Micetich, *R. G. US 6,232,305 B1* 2001.
70. Yamashita, D. S.; Smith, W. W.; Zhao, B.; Janson, C. A.; Tomaszek, T. A.; Bossard, M. J. Levy, M. A.; Oh, H.-J.; Carr, T. J.; Thomson, S. K.; Ijames, C. F.; McQueney, M.; D'Alessio, K. J.; Amegadzie, B. Y.; Hanning, C. R.; Abdel-

Meguid, S.; Desjarlais, R. L.; Gleason, J. G. and Veber, D. F. (1997) *J. Am. Chem. Soc.* 119, 11351-11354.

71. Abrahamson, M.; Alvarez-Fernandez, M. and Nathanson, C.M. (2003) *Biochem. Soc. Symp*. 70, 179-199.
72. Belenghi, B.; Acconcia, F.; Trovato, M.; Perazzolli, M.; Bocedi, A.; Poticelli, F.; Ascenzi, P. and Delledonne M. (2003) *Eur. J. Biochem.* 270:2593-2604.
73. Dublin, G. (2005) *Cell.Mol.Life Sci*.62, 653-669.
74. Kondo, H.; Abe, K.; Emori, Y. and Arai, S. (1991) *FEBS Lett.* 278, 87-90.
75. Margis, R.; Reis, E.M. and Villeret, V. (1998) *Arch. Biochem. Biophys.* 359, 24-30.
76. Solomon, M.; Urwin, P.; Atkinson, H.J.; Waller, D.A. and McPherson, M.J. (1995) *Plant J*. 8, 121-131
77. Arai, S.; Matsumoto, I.; Emori, Y. and Abe, K. J. (2002) Agric. *Food Chem.* 50, 6612-6617
78. Pernas, M.; Sanchez-Mong, R. and Salcedo, G. (2000) *FEBS Lett.* 467, 206– 210.
79. Aoki, H.; Akaike, T.; Abe, K.; Kuroda, M.; Arai, S.; Okamura, R.; Negi, A. and Maeda, H. (1995) *Antimicrob. Agents Chemother.* 39, 846–849.
80. Soares-Costa, A.; Beltramini, L. M.; Thiemann, O.H. and Henrique-Silva, F. (2002) *Biochem. Biophys. Res. Commun*. 296, 1194– 1199.
81. Pernas, M. R.; Sanchez-Monge, L.; Gomez, and G. Salcedo, (1998) *Plant Mol. Biol.* 38, 1235–1242.
82. Koiwa, H. R.; Shade, E.; Zhu-Salzman, K.; D'Urzo, M.P.; Murdock, L.L.; Bressan, R.A. and Hasegawa, P.M. (2000) *FEBS Lett*. 471, 67–70.
83. Oppert, B. T.D.; Morgan, K.; Hartzer, B.; Lenarcic, K.; Galesa, J.; Brzin, V.; Turk, K.; Yoza, K.; Ohtsubo, Kramer, K. J. (2003) *Comp. Biochem. Physiol.* 134, 481–490.
84. Solomon, M.; Belenghi, B.; Delledonne, M.; Menachem, E. and Levine A. (1999) *Plant Cell* 11, 431–443.
85. Strobl, S.; Fernandez-Catalan, C.; Braun, M.; Huber, R.; Masumoto, H. and Nakagawa, K. (2000) *Proc. Natl. Acad. Sci. USA* 97**,** 588–592.
86. Rzychon M., Sabat A., Kosowska K., Potempa J. and Dubin A. (2003) *Mol. Microbiol*. 49**,** 1051–1066.

87. Monteiro, A. C. S.; Abrahamson, M.; Lima, A. P. C. A.; Vannier- Santo, M. A. and Scharfstein, J. (2001) *J. Cell Sci.* 114, 3933–3942.
88. Rigden, D. J.; Mosolov, V. V. and Galperin, M. Y. (2002) *Prot. Sci.* 11: 1971–1977.
89. Sanderson, S. J.; Westrop, G. D.; Scharfstein, J.; Mottram, J. C. and Coombs, G. H. (2003) *FEBS Lett.* 542**:** 12–16.
90. Hao, L. Y.; Kameyama, A.; Kuroki, S.; Takano, J.; Takano, E. and Maki, M. (2000) *Biochem. Biophys. Res. Comm.* 279, 756–761.
91. Ogrinc, T.; Dolenc, I.; Ritonja, A. and Turk, V. (1993) *FEBS Lett.* 336**,** 555–559.
92. Yamashita, M. and Konagaya, S. (1996) *J. Biol. Chem.* 271**,** 1282–1284.
93. Lenarˇciˇc, B.; Ritonja, A.; Strukelj, B.; Turk, B. and Turk, V. (1997) *J. Biol. Chem.* 272, 13899–13903.
94. Lenarˇciˇc, B.; Krishan, G.; Borukhovic, R.; Ruck, B.; Turk, V. and Moczydlowski, E. (2000) *J. Biol. Chem.* 275, 15572–15577.
95. Bevec, T.; Stoka, V.; Pungerˇciˇc, G.; Dolenc, I. and Turk V. (1996) *J. Exp. Med.* 183, 1331–1338.
96. Crook, N. E., Clem, R. J. and Miller, L. K. (1993) *J. Virol.* 67: 2168–2174.
97. Fox, T.; de Miguel, E.; Mort, J. S. and Storer, A. C. (1992) *Biochemistry* 31, 12571–12576.
98. Wiederanders, B. (2003) *Acta. Biochim. Pol.* 50, 691–713.
99. Challis, G.L. and Hopwood, G. A. (2003) *Proc. Natl. Acad. Sci. USA* 100, 14555–14561.
100. Marinelli, F. (2009) *Methods in Enzymology*, 458, 29-58.
101. Demain, A. L. and Adrio, J. L. (2008) *Mol. Biotechnol.* 38, 41–55.
102. Fajardo, A. and Martinez, j. L.(2008) *Curr. Opin. Microbiol.* **11**, 161–167.
103. Vining, L. C. (1990) *Annu. Rev. Microbiol.* 44, 395–427.
104. Davies, J.; Spiegelman, G. B. and Yim, G.(2006) *Curr. Opin. Microbiol.* 9, 445–453.
105. Newman, D. J.; Cragg, G. M. and Snader, K. M. (2000) *Nat. Prod. Rep.* **2000**, *17*, 215-234.

106. Buss, A. D.; Cox, B. and Waigh, R. D. In *Burger's Medicinal Chemistry and Drug Discovery,* 6th ed.; *Volume 1: Drug Discovery*; Abraham, D. J., Ed.; Wiley: Hoboken, NJ, 2003; Chapter 20, pp 847-900.
107. Grabley, S. and Thiericke, R. In *Drug Discovery from Nature*; Grabley, S., Thiericke, R., Eds.; Springer: Berlin, 2000; Chapter 1, pp 3-37.
108. Sneader, W. *Drug Prototypes and their Exploitation*; Wiley: Chichester, UK, 1996.
109. Mann, J. *Murder, Magic and Medicine,* 2nd ed.; Oxford University Press: Oxford, UK, 2000.
110. Alder, A. L., Ed. *The History of Penicillin Production*; American Institute of Chemical Engineers: New York, 1970.
111. Lax, E. *The Mold in Dr. Florey's Coat*; Henry Holt Publishers: New York, 2004.
112. Wainwright, M. *Miracle Cure: The Story of Penicillin and the Golden Age of Antibiotics*; Blackwell: Oxford, UK, 1990.
113. Mann, J. *The Elusive Magic Bullet: The Search for the Perfect Drug*; Oxford University Press: Oxford, UK, 1999; pp 39-78.
114. (a)Brown, A. G.; Butterworth, D.; Cole, M.; Hanscomb, G.; Hood, J. D.; Reading, C. and Rolinson, G. N.(1976) *J. Antibiot. 29*, 668-669. (b)Howarth, T. T.; Brown, A. G. and King, T. J. (1976) *J. Chem. Soc., Chem. Commun*. 266-267. (c) Brown, A. G. (1986) *Drug Des. Delivery*, *1*, 1-21.
115. (a) Endo, A.; Kuroda, M. and Tsujita, Y.(1976) *J. Antibiot. 29*, 1346-1348. (b) Endo, A. (1992) *J. Lipid Res. 33*, 1569-1582
116. Singh, I. P.; Umehara, K.; Etoh, H.; Takasaki, M. and Konoshima. T. (1998) *Phytochemistry* 47, 1157-1159.
117. Singh, D. D.; Chitra, G.; Singh, I. P. and Bhutani. K. K. (2005) *Ind. J. Chem.*, 44B, 1288-1290.
118. Etoh, H.; Kondoh, T.; Noda, R.; Singh, I. P.; Sekiwa, Y.; Morimitsu, K. and Kubota. K. (2002) *Biosci. Biotechnol. Biochem.* 66, 1748-1750.
119. Singh, I. P. and Bharate, S. B. (2006) *Natural Product Reports* 23, 558 - 591.
120. Singh, I. P.; Lal, U. R.; Bodiwala, H. S.; Mahajan, R. P. and Bhutani. K. K.; (2006) STUDI UM PRESS LLC, P.O. Box-722200, Houston, Texas 77072, U.S.A., 13, 116-149.

121. Singh, I. P.; Bharate, S. B. and Bhutani. K. K. (2005) *Current Science*. 89, 269-290.
122. Butler, M. S. (2004) *J. Nat. Prod.* 67, 2141-2153.
123. Newman, D. J.; Cragg, G. M. and Snader, K. M. (2003) *J. Nat. Prod.*, *66*, 1022-1037.
124. Cragg, G. M.; Newman, D. J. and Snader, K. M. (1997) *J. Nat. Prod.*, *60*, 52-60.
125. Proudfoot, J. R.(2002) *Bioorg. Med. Chem. Lett. 12*, 1647-1650.
126. Lam, K. S. (2007) TRENDS in Microbiology, 15, 279-289.
127. Challis, G. L. and Ravel, J. (2000) *FEMS Microbiology Letters* 187, 111-114.
128. Demain, A.L. (1999) *Appl. Microbiol. Biotechnol.* 52, 455-463.
129. Umezawa, H. (1982) *Ann. Rev. Microbiol.* 36:75-99.
130. Xiao, H.; Chen, X.; Chen, M.; Tang, S.; Zhao, X. and Huan. L. (2004) *Microbiology* 150, 103-108.
131. Wu, S.; Jia, S.; Sun, D.; Chen, M.; Chen, X.; Zhong, J. and Huan, L.(2005) *Current Microbiology* 51, 1-6.
132. Chacha, M.; Mapitse, R.; Afolayan, A. J. and Majinda, R. R. T. (2008) *Natural Products Communications* 3, 17-20.
133. Mbukwa, E.; Chacha, M. and Majinda, R. R. T. (2007) ARKIVOC ix, 104-112.
134. Chacha, M., Bojase-Moleta, G. and Majinda, R. R. T. (2005) *Phytochemistry* 66, 99-104.

CHAPTER 2

Production, purification and characterization of CPI

Summary:

Deregulated expressions of proteases are convincingly reported to be involved in various pathobiological processes. Thiol proteases such as Cathepsin B expression is increased in many human cancers. The predominant expression of Cathepsin K in osteoclasts has rendered the enzyme a major target for the development of novel anti-resorptive drugs in osteoporosis and Cathepsin S appears to be an attractive drug target candidate for various inflammatory diseases including rheumatoid arthritis. Parasites such as Trypanosoma, Lieshmania, and Plasmodium also take the advantage of thiol protease activity to evade the host immune response. Therefore it is reasonable to assume that a good thiol protease inhibitor can be a useful compound to target the aforementioned disorders. Microbes and plants have been the most efficient and convenient source of natural bioactive compounds. Novel secondary metabolites continue to be isolated from Actinomycetes. Their biological activities and chemical structures show a wide range of diversity. Taking into account all these understandings, we decided to further explore the actinomycetes in order to isolate thiol protease inhibitor. During the course of investigation we have isolated a mixture of two novel peptides (736 and 842 Da) which inhibit papain (model thiol protease) in competitive mode of inhibition. The IC_{50} value for Papain inhibition by CPI-2081 was 36.9 ±1.8 nM and the K_i (Inhibitory constant) was found to be 49.14± 2.45 nM. This uniquely modified penta peptide was isolated with the help of solvent extraction and repeated reverse phase HPLC. The peptides were characterized completely with help of various techniques such as LC-MS-MS, FT-IR spectroscopy, 1D and 2D NMR Spectroscopy.

Introduction:

- **Screening of natural product in drug discovery programs have lost their momentum:** Nearly 10,000 different bio-chemicals of pharmacological importance have been isolated from microbial sources, out of which more than fifty percent have been isolated from actinomycetes. In early seventies a series of efforts has been made to successfully isolate many low molecular weight protease inhibitors from actimomycetes such as leupeptin, antipain, etc[1]. These are being used by cell and molecular biologists as the essential components of protease

inhibitor cocktail to protect the proteins from attack of proteases. All attempts to utilize these inhibitors to combat pathological consequences which involve proteases to degrade the vital organs, were unsuccessful in later stages of drug development (in vivo activities and clinical trails) due to metabolic side effects and poor permeability across the plasma membrane. These days, rational or structure-based drug design is the popular method of finding new protease inhibitors for pharmaceutical purposes whereas activity based search for new protease inhibitors from natural sources has lost its momentum[2]. It is essentially worthwhile to put serious effort in finding out protease inhibitors from natural sources which will help to gain the momentum to find out new natural scaffolds for synthetic chemists who are involved in synthesizing new drugs.

- **Natural products are more important than combinatorial compound libraries as a potential source of new drug:** Although, formation of natural products have been the primary source of lead scaffold to the organic chemists involved in development medicinal drugs, it serves as a selectable evolutionary gain to the producing organism. Even if, the technological advancements such as high throughput screening systems have solved the limitations up to some extent while screening of natural products confronted due to its time consuming procedures; there are several advantages to the screening of natural products for drug discovery that outweigh their limitations. Some of them are needed to be considered seriously to justify the need to carry out these kinds of work as tool for development of more medicinally useful lead compounds.

  1. Natural products offer unparalleled chemical variety with structural complexity and biological effectiveness[3]. Natural product resources are mostly unexplored.
  2. Applying the recent technological improvement in high performance liquid chromatography, mass spectrometry and high quality NMR techniques will certainly lead to the discovery of novel bioactive compounds. It is difficult to estimate the magnitude of this advantage, but, J. Be´rdy, in his review estimates a 100-fold higher hit rate for natural products as compared to synthetic compounds[4].

3. Man-Ling Lee and Gisbert Schneider have done very extensive study to find out the similarities of scaffold (ring system) between current trade drugs on one hand and natural product or drugs obtained by combinatorial synthetic chemistry on the other. They found that the natural products occupy more complementary region of chemical space than compounds obtained by combinatorial synthetic chemistry. Some other investigators have also reached to the similar conclusion despite using different approaches and chemical databases [5, 6, 7] .
4. Sometimes, natural products occupy completely unusual chemical structure that is difficult to explain such as occurrence of uncommon amino acid or structural groups. Natural product database contain numerous scaffolds, but some of very important ring system is not available in databases at all. Such unexploited scaffolds represent potential new starting points in drug discovery.
5. Furthermore, the use of natural products as templates for combinatorial chemistry enables the generation of libraries of biologically important natural product analogs, which might have improved pharmacokinetic properties. This application is an important turning point because combinatorial chemistry was previously thought to substitute the need to hunt for novel natural products. Now it is a very well established and accepted fact that both the disciplines complement each other.
6. The structural diversity of the natural product libraries can be further enhanced by combining the techniques of biotransformation and combinatorial biosynthesis. Using our increased technological advancements of molecular biology and biosynthesis of natural products, we can manipulate and optimize the regulation of natural product biosynthesis to yield more and better quality[8, 9]. Now a days these can be done more easily due to the availability of complete genome sequences of pharmacologically important organisms, where, identifying new natural products can be done by genome mining techniques.

7. Natural product compounds not only serve as drugs or templates for drugs, but also sometimes, lead to the discovery and validation of drug targets and pathways involved in the disease process, such as the protein–protein complexes B-catenin in the WNT pathway and HIF-1/p300[10], that have validated these anticancer targets and pathways.
8. Natural products also create opportunities for more drug targets to be discovered in these pathways. The explanation of the anti-inflammatory mechanism of action of aspirin led to the discovery of the cyclooxygenase isozymes COX-1 and -2, which were used in the development of novel anti-inflammatory drugs[11].
9. Synthetic drugs are usually the product of various structural modifications over the course of an extensive drug discovery program; a natural product can go directly from a biologically active compound to drug.
10. Natural products are important not only for their prospective therapeutic actions, but also for the fact that they often have the required pharmacokinetic properties for clinical purposes. Antibacterial agents erythromycin A, vancomycin, penicillin G, streptomycin and tetracycline; antifungal agents amphotericin B and griseofulvin; the cholesterol-lowering agent lovastatin; anticancer agents daunorubicin, mitomycin C and bleomycin; and immunosuppressants rapamycin, mycophenolic acid and cyclosporine A are a few of the many microbial natural products that reached the market without requiring any chemical modifications. These examples clearly display the significant capability of microorganisms to generate drug-like small molecules.

❖ **Advantages of small molecular mass natural protease inhibitors in drug development strategy:** Many disease states manifestations are a result of altered protease expression and substrate proteolysis. In particular, excessive protease activity is often observed, as in case of osteoporosis and arthritic patients, offering targets for therapeutic inhibition [12, 13, 14]**.** There are at least 475 known putative proteases and 103 homolog to known proteases in humans, and it is estimated that up to 1,200 human genes encode proteases. Proteases, representative of all five

known classes have been implicated in human disease. Some parasites, Trypanosoma, malarial and Leishmania often use the attributes of proteases to facilitate the pathology and win over the host defense to survive in hostile microenvironment [15, 16]. Many protease inhibitors have been discovered, but due to high molecular weight, lack of specificity and larger size are disqualifed for therapeutic applications. Protease inhibitors are also required for the study of mechanism of enzyme action. There have been many reports of production of antimicrobial molecules in microbial cultures itself[17]. The best known example for this is antibiotics, penicillin and streptomycin etc and proteinaceous protease inhibitors such as cystatins. Chagasin, clitocypins, staphostatins, calpastatins etc[18]**.** After getting the high molecular weight proteinaceous protease inhibitors, it becomes necessary to characterize the mechanism of inhibition of enzyme action and molecular details which take part in course of action. The knowledge acquired regarding the proteinaceous inhibitor is utilized to design low molecular weight synthetic molecules which can be used for therapeutic purposes. Novel protease inhibitors from natural sources, however, continue to dominate either in its native or chemically modified form with enhanced functional implications. Therefore, further attempt to isolate new low molecular weight, stable in hostile homeostatic conditions and pharmacologically selective natural protease inhibitor which work in minimum possible molar concentration, is highly essential.

- **Microorganisms as source of low molecular mass protease inhibitors and scope of the work:** Among all kinds of sources for natural bioactive compounds, actimomycetes appear to have competitive edge over others with respect to their ability to produce bioactive small molecules for drug development. There are many novel bioactive compounds which have been isolated from actinomycetes in recent past. Lanophylins, A1, A2, B1 and B2 were isolated from *Streptomyces sp.*[19]. Lanophylins are the first naturally occurring (2E)-methyllidine-2methyl-1pyrroline backbone which shows lanosterol synthase inhibitory activity and can be an intresting tool for investigating the cholesterol maintenance. Some alkaloids isolated from *Dactylosporangium sp*. inhibited recombinant human kinase and serine proteases, cathepsin G and chymotrypsin[20, 21]. Mannopeptimycins,

glycopeptides exhibiting antibacterial activities were isolated from *Streptomyces hygroscopicus*[22]. Few small molecule non proteinaceous novel cysteine protease inhibitors, which are used in biomedical research as protease inhibitor cocktail, have been isolated from *Streptomyces sp*[23, 24]. FA-70C, an Antipain (cysteine protease inhibitor) analog, was isolated from the culture supernatant of *Streptomyces species* strain FA-70. It inhibits Arg-gingipain (Rgp), a key cysteine protease produced by *Porphyromonas gingivalis*, a major pathogen of advanced periodontal diseases[25]**.** A new compound migrastatin, an inhibitor of cell migration, was isolated from Streptomyces sp. Mk929-43F1[26]. A series of successful attempts to isolate protease inhibitors from S*treptomyces sp.* lead to discovery of few small molecule Thiol and serine protease inhibitors such as leupeptin[27], E64[28], antipain[29]. But due to requirement of high molar concentration for inhibitory activity these compounds have not found their applicability in therapeutic formulations to target proteases in human diseases, and are maximally restricted to use in research laboratories as protease inhibitor cocktails. Therefore, this work is about isolation and characterization of new small molecular mass thiol protease inhibitor which requires nanomolar range of concentration for Thiol protease inhibition.

A. Production purification and identification of CPI-2081:

Experimental:

Materials: Papain, N α-Benzoyl-DL-Arginine p-Nitroanilide (BAPNA, extinction coefficient of 8800 M^{-1} cm^{-1} at 410 nm), Trifluoroacetate (TFA) and other chromogenic substrates were purchased from Sigma Chem. Co. (USA). Acetonitrile was purchased from E-Merck (Germany). NaCl, malt extract, glucose, yeast extract, peptone, Tris and agar were purchased from Himedia and Qualigens (India). All other chemicals were of analytical grade. All reverse phase HPLC (High Performance Liquid Chromatography) columns and separation modules were purchased from Waters, USA .

Organism and strain maintenance and fermentation: *Streptomyces sp* .strains were obtained from NCIM, NCL (National Collection of Industrial Micro-organisms, National Chemical Laboratory, Pune) Time course of inhibitor production: *Streptomyces sp.*

NCIM2081 culture was seeded in 1000ml flasks containing MGYP (0.3% Malt extract, 1% Glucose, 0.3% Yeast extract and 0.5%Peptone) culture media containing 2% soyabean meal and kept in incubator shaker at 28 ^{0}C temprature and speed of shaker 200 RPM. After 72 hrs three flasks were removed and kept at 4 ^{0}C for harvesting culture broth followed by protease inhibition activity assay, in the centrifuged culture broth, using papain as model enzyme.

Assay for inhibitory activity of CPI-2081 towards papain

The inhibitory activity of CPI-2081 against papain was determined by assaying the proteolytic activity of 50 µl of papain (1mg/ml) in Tris-HCL buffer, pH 6.5 in the presence of 10 mM DDT and 2 mM EDTA, using BAPNA (1.5 mM) as the substrate in the presence and absence of CPI-2081 at 37 ^{0}C.

RP-HPLC: Reverse phase HPLC was carried out using waters separation module with Waters C_{18}-symmetry and delta pack columns. Running conditions were, 45 minutes gradient of 5 to 95 percent acetonitrile in glass double distilled miliQ water containing 0.1% TFA at a flow rate of 3ml/min of semipreparative HPLC (7.9mm x 300mm) whereas 1 ml/min for analytical (4.6mm x 250mm) columns. Detector having UV/Vis dual wavelength spectrophotometer operating at 220nm and 280 nm was used to monitor the progress of HPLC (high performance liquid chromatography).

FT-IR spectroscopy: The compound in the form of dry powder was taken in KBr pellet was carried out using a Perkin–Elmer Spectrum One instrument. Spectrometer operated in the diffuse reflectance mode at a resolution of 2 cm^{-1}.To obtain good signal to noise ratio, 128 scans of the film were taken in the range 450–4000 cm^{-1}.

LC-MS analysis: Mass spectrometry was carried out with a QTOF2 (Waters) equipped with a standard ESI source. All mass spectral measurements were taken at capillary voltage 3.75-3.68 KV, cone voltage 73-75.46V, source temperature 120 °C, desolvation temperature 350 °C and desolvation gas 750L/hr. Each 0.4s scan spanned m/z 100-800 with an interscan delay time of 0.1 s. LC separations were performed in a Waters ACQUITY-UPLC^{TM} system (Waters Corp., Milford, USA) using a 50mm×2.1mm ACQUITY-BEH C_{18}, 1.7 m column. A sample volume of 10 µL was injected with a Acquity UPLC^{TM} autosampler. The mobile phase was composed of Solvent A (20%

water) and solvent B (80% methanol) at a constant flow of 0.3 ml/min. Total run time was 10minutes.

Results and Discussion:

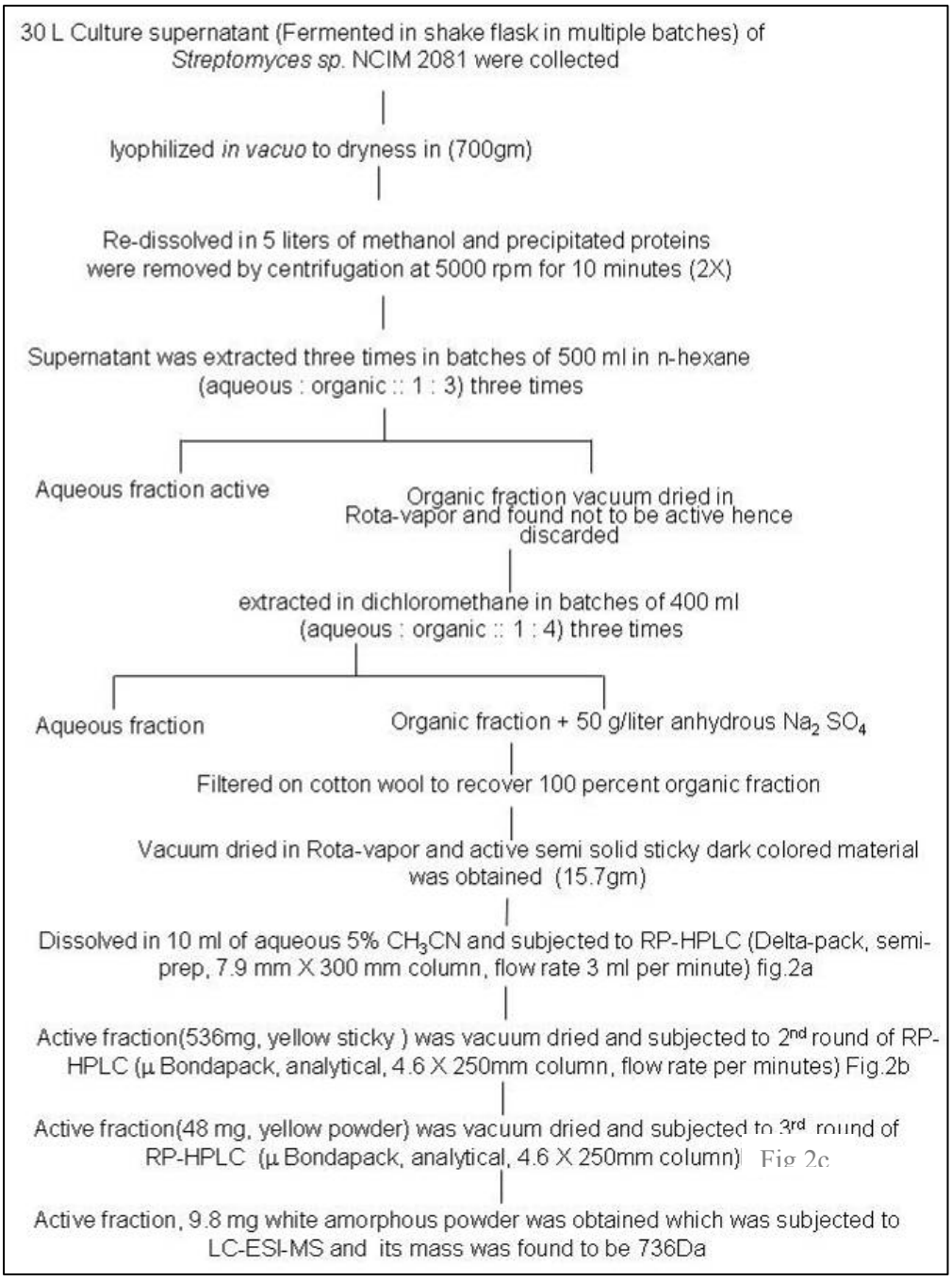
The design of the bioassay consisted of screening of various *Streptomyces sp* fermentation broth on the basis of its ability to inhibit the thiol protease activity in cell free system using papain as model enzyme. The fermentation broth of strains that are active and exert their Thiol protease inhibitory activity, were selected for identification, extraction and isolation of responsible compound. Culture broth from *Streptomyces sp.* NCIM 2081 was found to be able to inhibit the hydrolysis of synthetic substrate, BAPNA (N-benzoyl-DL-arginine-p-nitroanilide hydrochloride) by thiol proteases papain. BAPNA contains the Thiol protease cleaving site and p-nitroanilide is released as a result of protease (Papain) action which can be monitored spectrophotometrically by measuring the absorption at 410nm. The organism was grown in MGYP media containing malt extract, glucose, yeast extract and peptone. The strain NCIM 2081 was cultured in shake flasks in submerged conditions for various time durations, 24 hours, 48 hours, 72 hours and 96 hours, to carry out the time course for production of papain inhihitory activity. It was found that after 72 hours of fermentation time duration, the protease inhibitory activity was found to be the highest (Figure: 1). Thereafter, culture supernatant was harvested by centrifuging at 5000 RPM in sorval SH-32 rotor, lyophilized till complete dehydration. This dehydrated powder was used for extraction of CPI-2081 (cysteine protease inhibitor) as described in the flow chart for extraction procedure. In brief, the anhydrous powder thus obtained was dissolved in aqueous methanol to precipitate all proteins and extracted with n-hexane followed by dichlormethane. The Aliquot of the dichloromethane extracts of the fermentation broth was able to inhibit the substrate hydrolysis by Papain (model Thiol protease). Dichloromethane was removed under v*acuo* after filtration and extracted sample was chromatographed on semi preparative Delta pack C18 column, a reversed-phase HPLC resin, and eluted with an aqueous acetonitrile gradient (5% to 95%) containing 0.1% TFA in 40 minutes. The fractions were analyzed in the cell free protease inhibition assay. The active fractions were pooled and rechromatogarphed by reversed-phase HPLC which showed many closely eluted compounds. Chromatography of this fraction on analytical, C_{18} reversed-phase HPLC

using gradient of aqueous acetonitrile containing of 0.1% TFA led to separation of compounds and active fraction was collected (Figure: 2a), which was further purified by a second reversed-phase HPLC (Figure: 2b) to yield 9mg from 35 liter of the original fermentation broth. The complete profile of purified compound is shown in figure 2C. ESI-MS analysis of purified compound showed two protonated molecular ions at *m*/*z* 736.4058 and *m*/*z* 842.22 (Figure 3), which indicated the presence of mixture of two inseparable compounds (CPI-2081). There are many low molecular mass cysteine proteases inhibitors are reported from microorganisms, particularly *Streptomyces sp.* that are used as protease inhibitor cocktail during cell biology experiments and very extensively reviewed by H. Umezawa 1982[1]. One of that is the leupeptin, isolated from *Streptomyces sp.* is an N-acetylated tetrapeptide of 463.01 Da molecular mass cysteine and serine protease inhibitor showing inhibition constant value of ~15 nM against cathepsin B and papain.[30]. Therefore, we decided to find out the inhibitory constant of CPI-2081 isolated from *Streptomyces sp* 2081. We prepared the Dixon plot by plotting the rate of BAPNA (synthetic substrate for Papain) hydrolysis by Papain (putative cysteine protease). We found that the CPI-2081 is able to inhibit the activity of Papain significantly showing K_i value of 49.14± 2.45 nM.

Various unsuccessful attempts to separate these two compounds were made by changing the mobile phase and their gradient conditions. The UV spectrum of both compounds was identical and displayed absorption maxima at λ_{max} 280 and 220 analogous to the UV spectrum of peptides containing aromatic amino acids. The FT-IR spectra (Figure: 4) in solid KBr revealed that it contain peptide amide core. Some of the important FT-IR signals are explained as, 1530 cm^{-1} (N-H bending), 1633 cm^{-1}(amide C=O stretching), 1453 cm-1(Aromatic C=C stretching), 3275 cm^{-1} (N-H stretching), 2959 cm^{-1}(C-H stretching) and 1227 cm^{-1}(C-O stretching). After getting this initial indication about the purified compound that it is a mixture of two inseparable compounds which may have nearly similar physiochemical properties. There are similar reports in literature by Reddy *et al,* in which they got three related tri peptides, eurypamide-A, B and D, as an inseparable mixture[31] . In another report the bioassay-guided fractionation of the fermentation broths of the endophyte, Streptomyces *sp.* MSU 2110, on silica gel and HPLC chromatography yielded two principal, inseparable, peptides with masses of

1217.9 and 1203.8 Da[32] . Boryana *et al*. also obtained an inseparable mixture of two napthoquinone epoxides from Brazilian Red Propolis [33]. Globulifera et al have analyzed the mixture of anti HIV benzophenones using various MS-MS and NMR techniques[34]. In later part of this chapter we have explained the complete structure elucidation of our inseparable mixture of both the compounds separately by the help of various NMR and other spectroscopic techniques.

Flow chart for the extraction of CPI-2081 from culture supernatant of Streptomyces sp. NCIM2081



Figures:

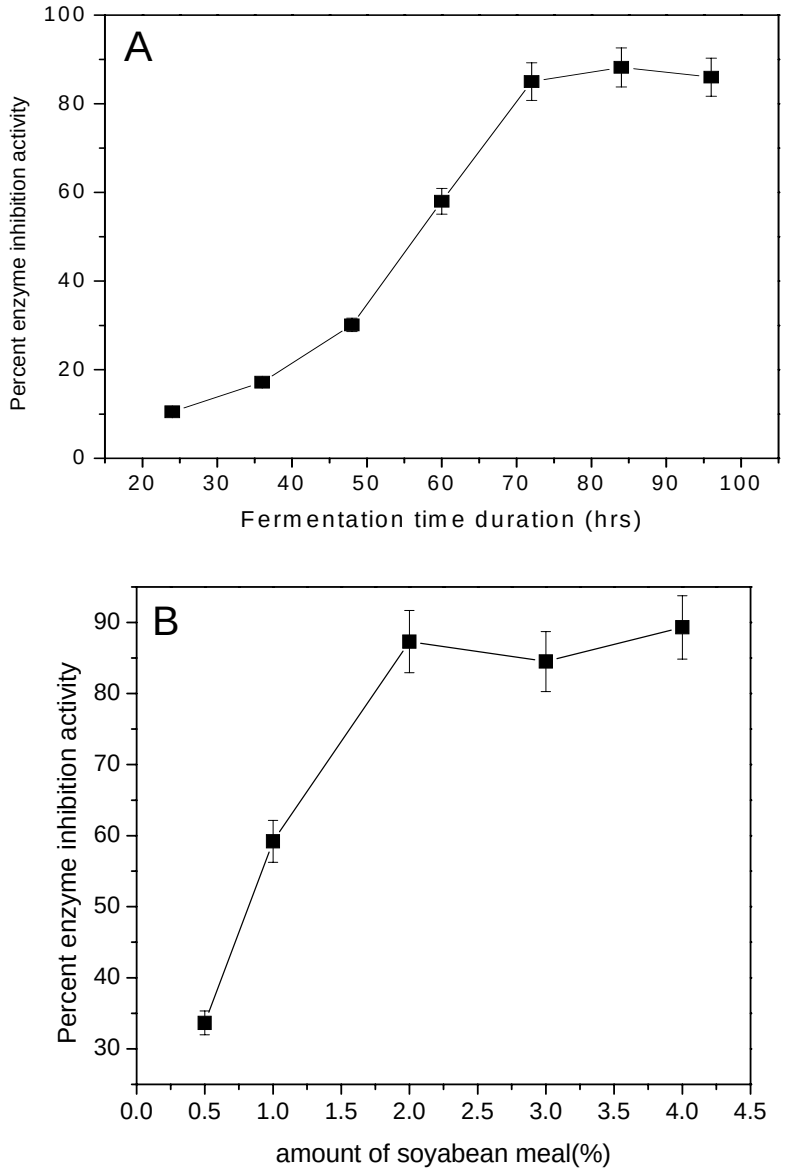


Figure 1. (A) Time course of inhibitor production: Streptomyces sp.(NCIM2081) culture was seeded in various 250ml flask in MGYP culture media containing 2% soyabean meal and kept in incubator shaker at 28 ^{0}C temperature and 200 RPM. followed by protease inhibition activity assay in culture supernatant using papain as model enzyme. The percent enzyme inhibition activity of culture supernatant is plotted against time duration in hours (hrs) (B) Dose response of soyabean meal for inhibition activity in culture supernatant of Streptomyces sp: Culture was seeded in 250 ml flask and kept for fermentation at 28 ^{0}C and 200 RPM for 72 hours (hrs) in MGYP culture media having different concentrations of soyabean meal supplement. The percent enzyme inhibition activity of culture supernatant is plotted against amount of soyabean meal supplement.

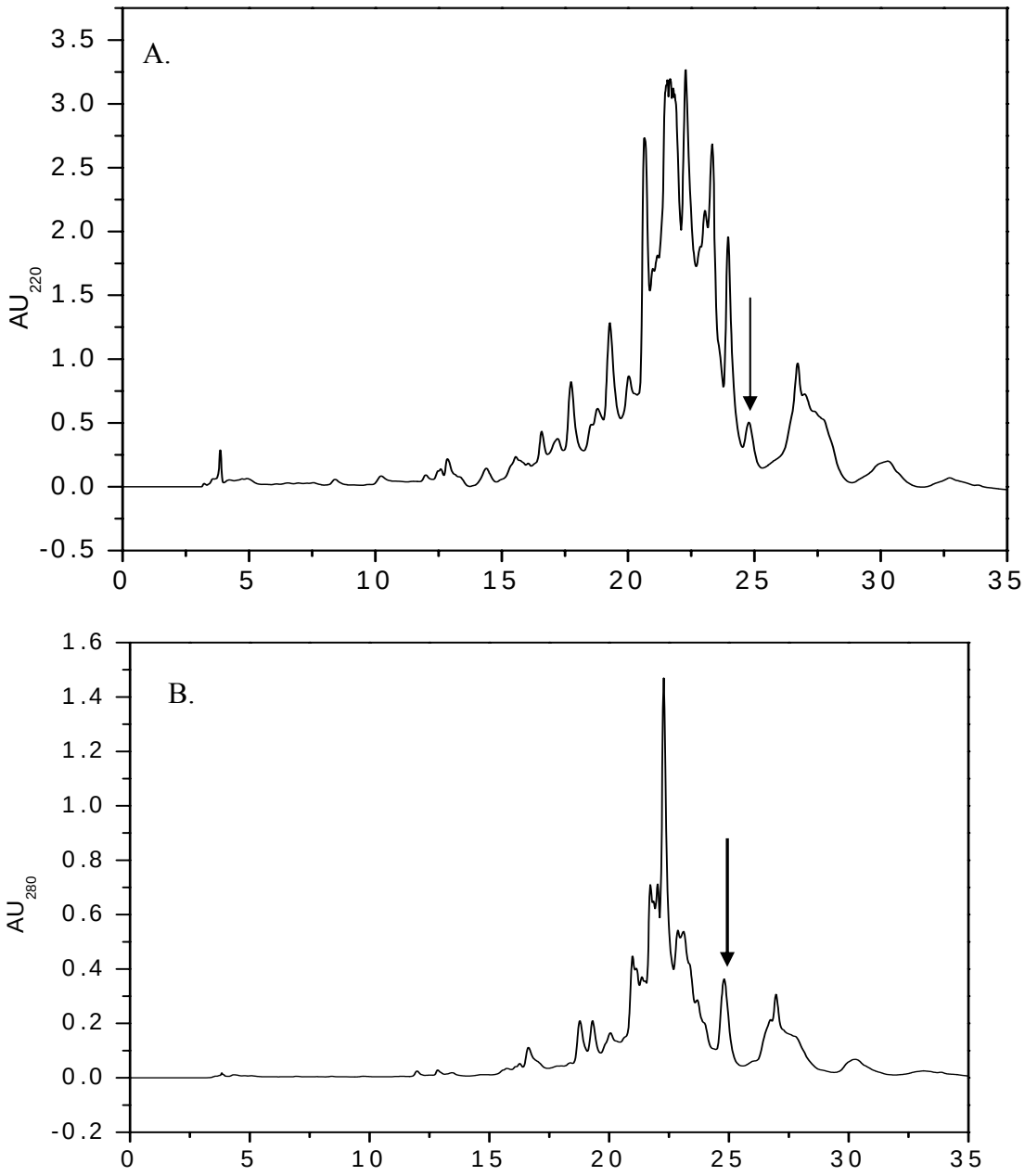


Figure 2a: RP-HPLC profile of dichloromethane extract: HPLC was carried out using waters separation module. DCM extract was loaded on semi preparative RP-Delta pack C18 column (300mm x 7.9mm). Running conditions were, 45 minutes gradient of 5% to 95% acetonitrile in glass double distilled miliQ water containing 0.1% TFA at a flow rate of 3ml/min the progress of HPLC was monitored using duel wavelength uv spectrophotometer (A) RP-HPLC chromatogram at 220nm. (B) RP-HPLC chromatogram at 280nm. Active fraction is been pointed using arrow.

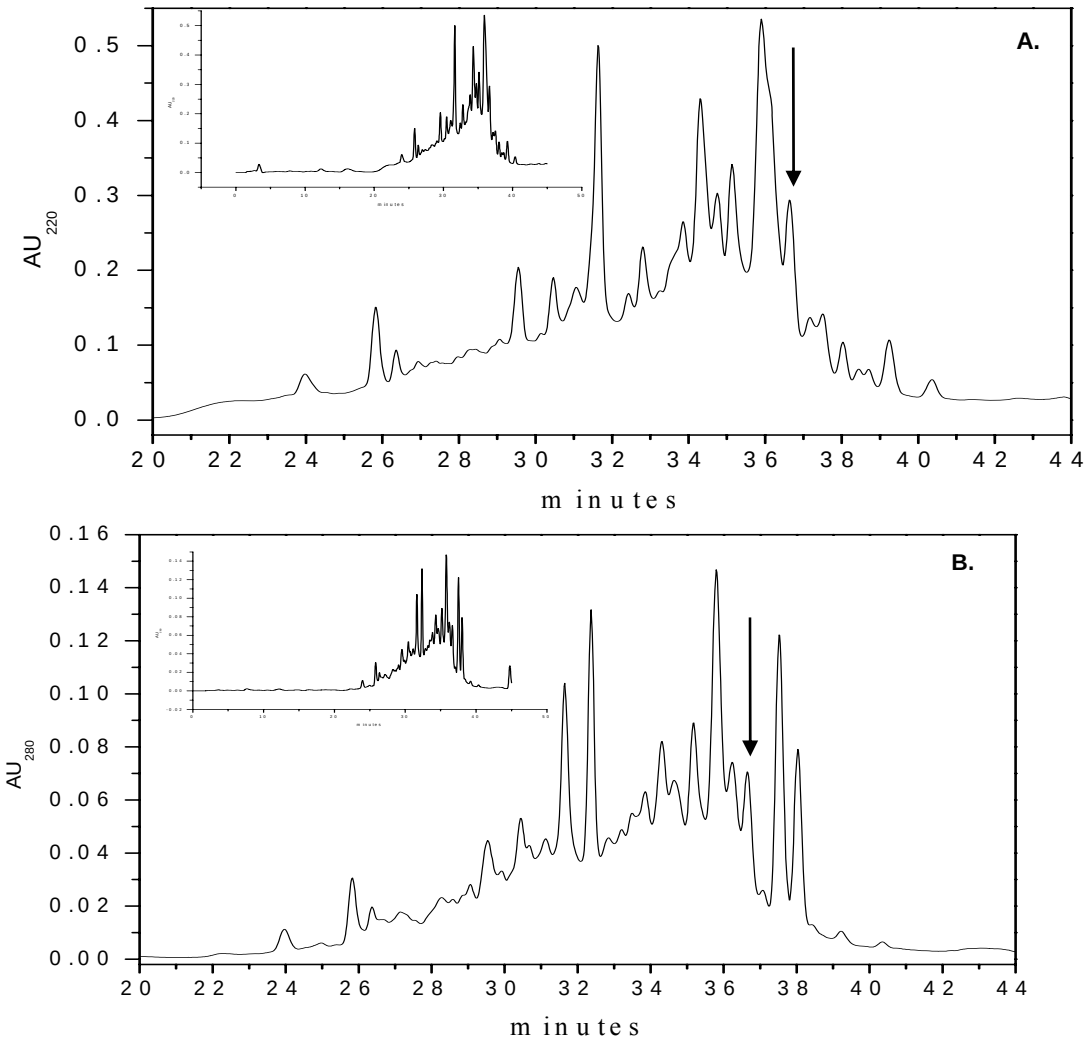


Figure 2b: HPLC profile of active fraction eluted from RP-Semi preparative HPLC, using waters separation module. Active fraction elute from semi prep column is loaded on analytical RP–µBondapack C18 column (250mm x 4.6mm). Running conditions were, 45 minutes gradient of 5% to 95% acetonitrile in glass double distilled miliQ water containing 0.1% TFA at a flow rate of 1ml/min. Full spectrum is shown smaller and relevant part is expanded for clearer view progress of HPLC was monitored using duel wavelength uv spectrophotometer (A) RP-HPLC chromatogram at 220nm. (B) RP-HPLC chromatogram at 280nm. Active fraction has been pointed using arrow.

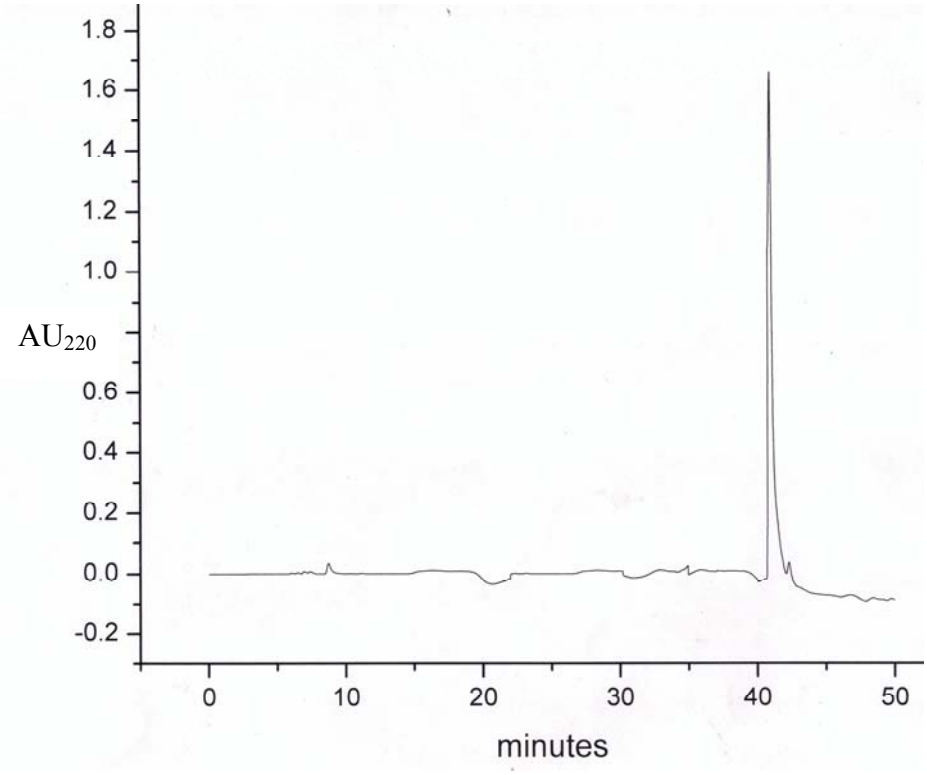


Figure 2c: HPLC profile of purified active fraction eluted from RP-Semi preparative HPLC chromatography, using waters separation module: *Active fraction elute, pure CPI-2081 from semi prep column is loaded on analytical RP-mBondapack C_{18} column (250mm x 4.6mm). Running conditions were, 45 minutes gradient of 5% to 95% acetonitrile in glass double distilled miliQ water containing 0.1% TFA at a flow rate of 1ml/min.*

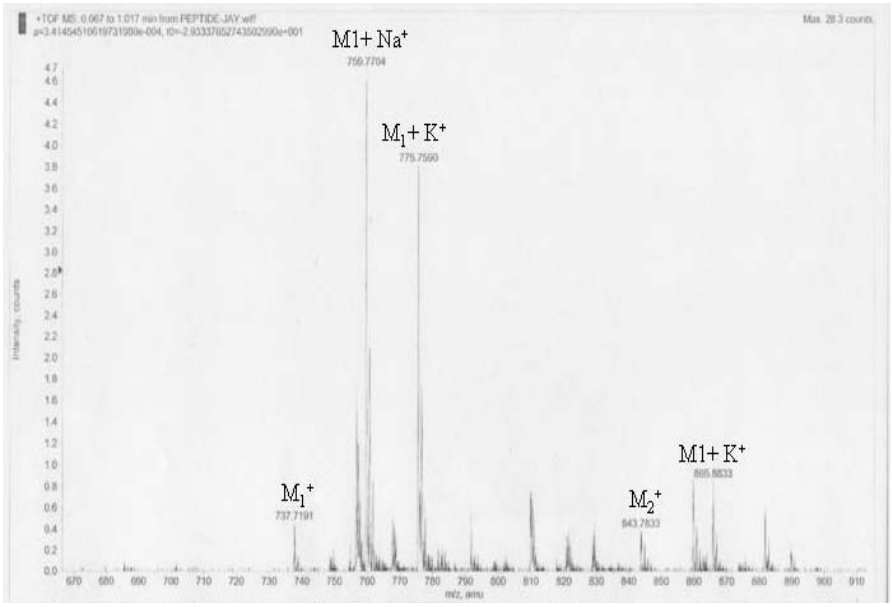


Figure 3: ESI-MS spectrum of purified CPI-2081 showing mixture of two compounds as there are two molecular ion peaks are visible. The major (736 Da) and minor (842 Da) peptides are depicted as M1 and M2 respectively

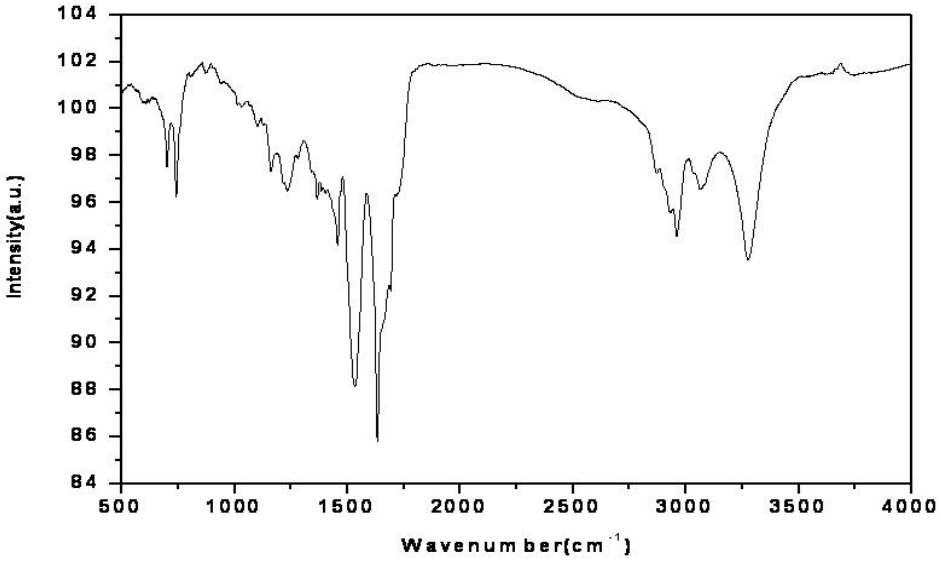


Figure 4: The FTIR spectroscopy measurement of CPI-2081: *The compound in the form of dry powder taken in KBr pellet were carried out using a Perkin–Elmer Spectrum One instrument. Spectrometer operated in the diffuse reflectance mode at a resolution of 2 cm^{-1}.To obtain good signal to noise ratio, 128 scans of the film were taken in the range 450–4000 cm^{-1}, 1530 cm^{-1} (N-H bending), 1633 cm^{-1}(amide C=O stretching), 1453 cm^{-1}(Aromatic C=C stretching), 3275 cm^{-1} (N-H stretching), 2959 cm^{-1}(C-H stretching) and 1227 cm^{-1}(C-O stretching).*

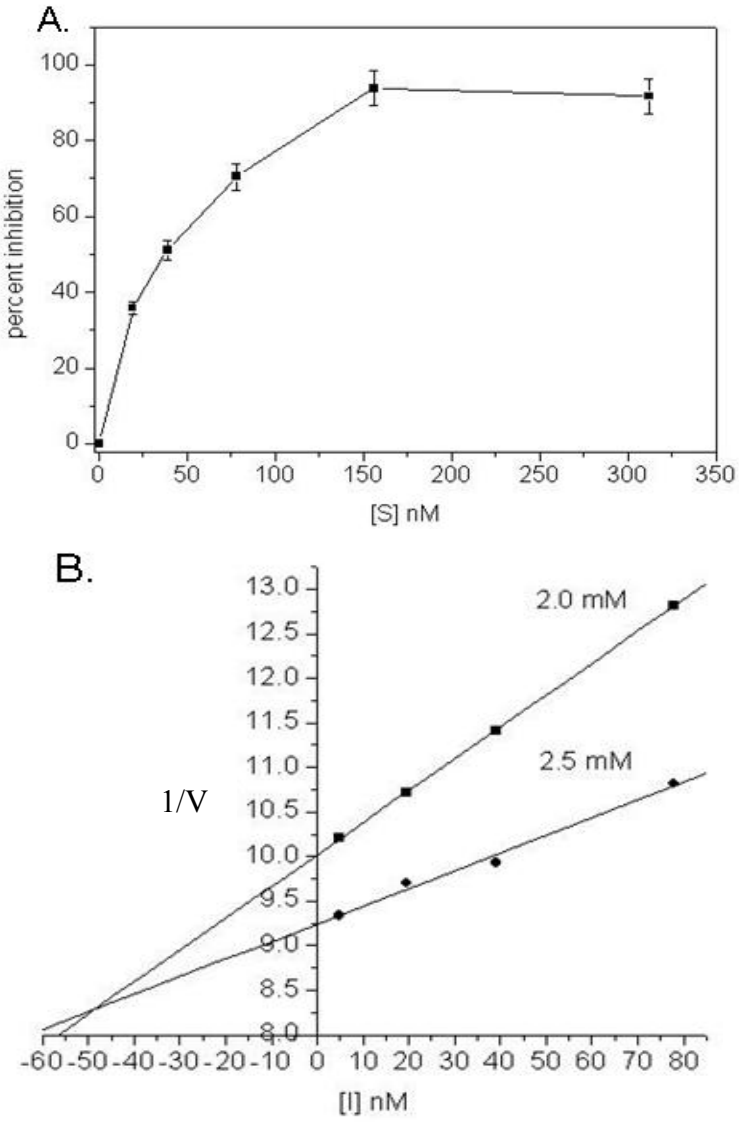
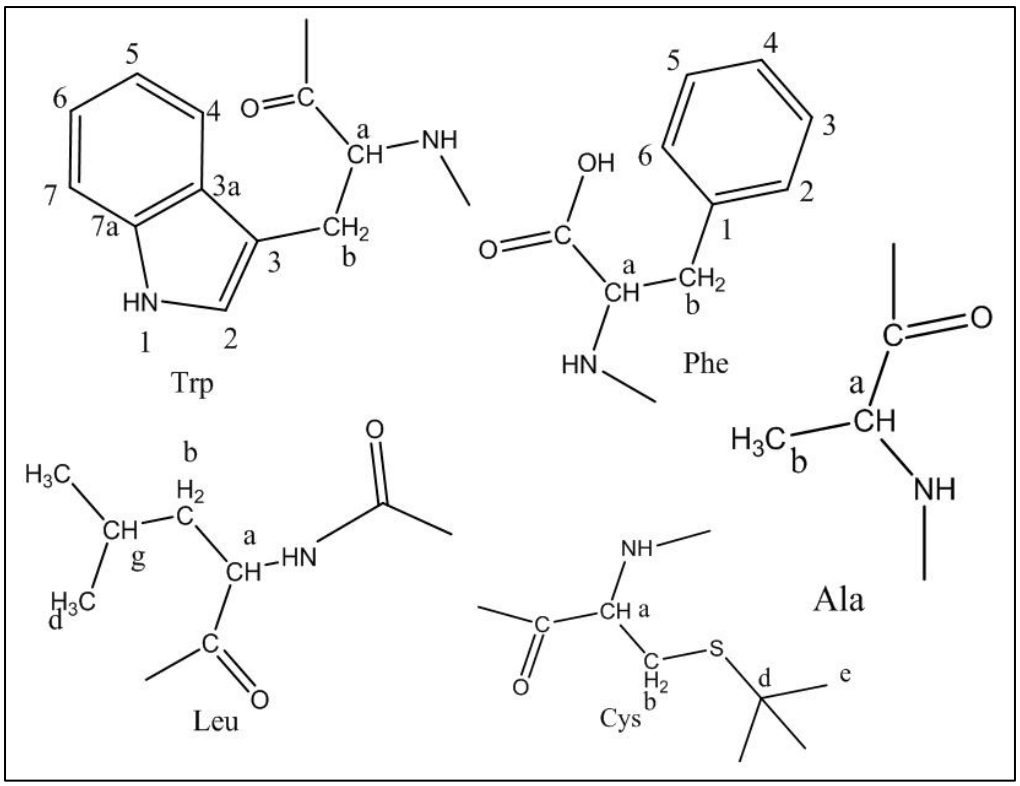


Figure 5: Determination of IC_{50} and K_i of CPI-2081 for papain inhibition *(A) Determination of IC_{50} (36.9* ±1.8 nM*) for inhibition of hydrolysis of BAPNA by Papain. Percent inhibition ± SEM is plotted against concentration of CPI-2081 in nanomolar (nm) (B) Dixon plot for determination of inhibitory constant (K_i=* 49.14± 2.45 nM *) for inhibition of proteolysis by CPI-2081. Reciprocal rate of reaction is plotted against concentration of CPI-2081 in nanomolar (nm).*

B. NMR spectroscopic studies of CPI-2081 and elucidation of its chemical structure:

Materials and Method: All the NMR spectroscopic (**N**uclear **M**agnetic **R**esonance) measurements were carried out on a Bruker AV 500 spectrometer operating at 500.13 MHz, 125.75 MHz, respectively for ^{1}H, ^{13}C. ~ 10 mg of the sample isolated was dissolved in DMSO-d_6 in a standard 5mm NMR tube and the ^{1}H, COSY (**Co**rrelation **S**pectroscop**y**), NOESY (**N**uclear **O**verhauser **E**nhancement **S**pectroscop**y**), ROESY(**R**otational **N**uclear **O**verhauser **E**ffect **S**pectroscop**y)**, ^{13}C CPD (**C**omposite-**P**ulse **D**ecoupling), ^{13}C DEPT (a spin-echo sequence of pulses to select for spatial location within the static field), ^{13}C-^{1}H HSQC (**H**eteronuclear **S**ingle **Q**uantum **C**orrelation), ^{13}C-^{1}H HMBC **(H**eteronuclear **M**ultiple **B**ond **C**orrelation), ^{15}N DEPT, and ^{15}N-^{1}H HSQC experiments were performed. The one dimensional ^{13}C experiments were performed on 5mm QNP probe at ambient temperature (~ 28 ^{o}C). 20000 and 11000 transients were collected for ^{13}C CPD and DEPT135 spectra. All the 2D experiments except the ^{15}N-^{1}H HSQC were conducted on a 5mm broad band inverse gradient probe. The ^{15}N - ^{-1}H HSQC experiment was carried out on a Bruker AV 400 NMR spectrometer operating at 40. 56 MHz for ^{15}N using a Broad band observe (BBO) gradient probe. Gradient spectroscopic techniques were employed for all the 2D experiments. 400 experiments (t_1 increments) of 24 scans were used for COSY, NOESY and ROESY measurements. The COSY and the HMBC spectra were collected in a magnitude mode while a phase sensitive (States-TPPI) mode was used for HSQC, NOESY and ROESY measurements. A mixing time of 1 sec and 300 m sec was employed for NOESY and ROESY experiments, respectively. The numbers of scans used for each t_1 increment for other 2D experiment were as follows: 24 (^{13}C HSQC,), 80 (^{13}C HMBC), 64 (^{15}N HSQC). The ^{13}C HMBC data were optimized for a long range coupling constant of 6 Hz. A pulse sequence employing a double low pass filter was found to give better results for ^{13}C HMBC due to spread in $^{1}J_{C-H}$ values (160 -135 Hz). The HMBC spectra were acquired without proton decoupling during detection. The 90º pulse lengths for ^{1}H, ^{13}C and ^{15}N were 13.5, 10 and 14 µ sec, respectively. Appropriate window functions *viz*. sine squared bell with no phase shift for all magnitude mode and phase shifted (ssb = 2) sine squared bell for phase sensitive mode were used for data processing In general a 2Kx2K data matrix size

was used for the 2D experiments. The ^{1}H and ^{13}C chemical shifts were referred to the residual solvent peak (2.50 and 39.95 the central signal of the solvent, respectively for ^{1}H and ^{13}C). The ^{15}N chemical shifts were referred to an external sample of nitromethane (0 ppm). DOSY experiments were carried out on a Burker Av 500 NMR spectrometer equipped with a 5mm broad band inverse gradient probe by systematic variation of the strength of the gradient amplitude and the data was processed with the standard Bruker software. *The numbering of atoms followed for different amino acids throughout the study is given below in the box.*



Results and discussion:

NMR Characterization: The isolated sample being hydrophobic did not show any solubility in D_2O hence the initial NMR measurements were carried out in Methanol d_4. Though the 1H, ^{13}C and 2D COSY, NOESY ^{13}C-^{1}H, HSQC and ^{13}C-^{1}H HMBC experiments performed on this sample gave some idea about the residues the sequence determination was difficult due to the absence of the NH protons due to their exchange with deuterium of the alcohol. Hence, this sample, after removal of the solvent and repeated exchange of the deuterated amide protons with ordinary methanol, was dissolved in DMSO-d6 and the NMR characterization was carried out using a combination of ^{1}H, ^{13}C, ^{15}N, 1D and 2D experiments. The result of this study is presented below. Figs-1 to 3 shows the full and expanded 500 MHz ^{1}H NMR spectrum of the oligo peptide obtained in DMSO-d6 from which the following features are noticed.

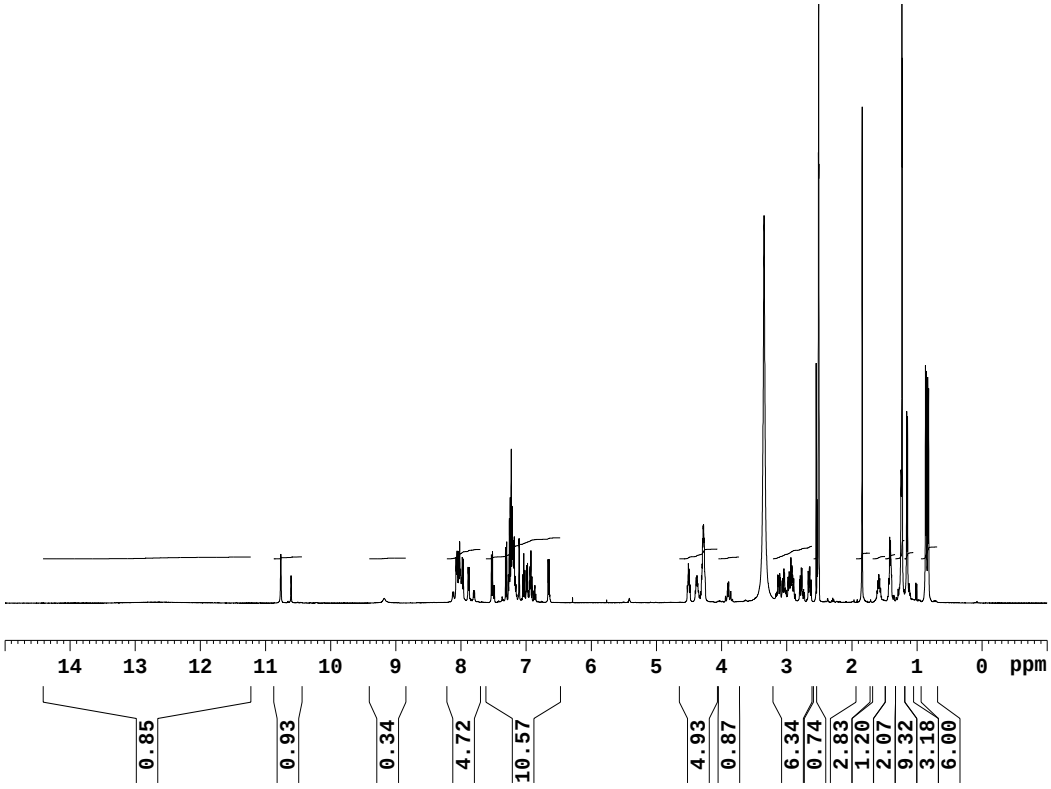


Fig-1: 500 MHz ^{1}H NMR spectrum of the oligo peptide in DMSO-d_6

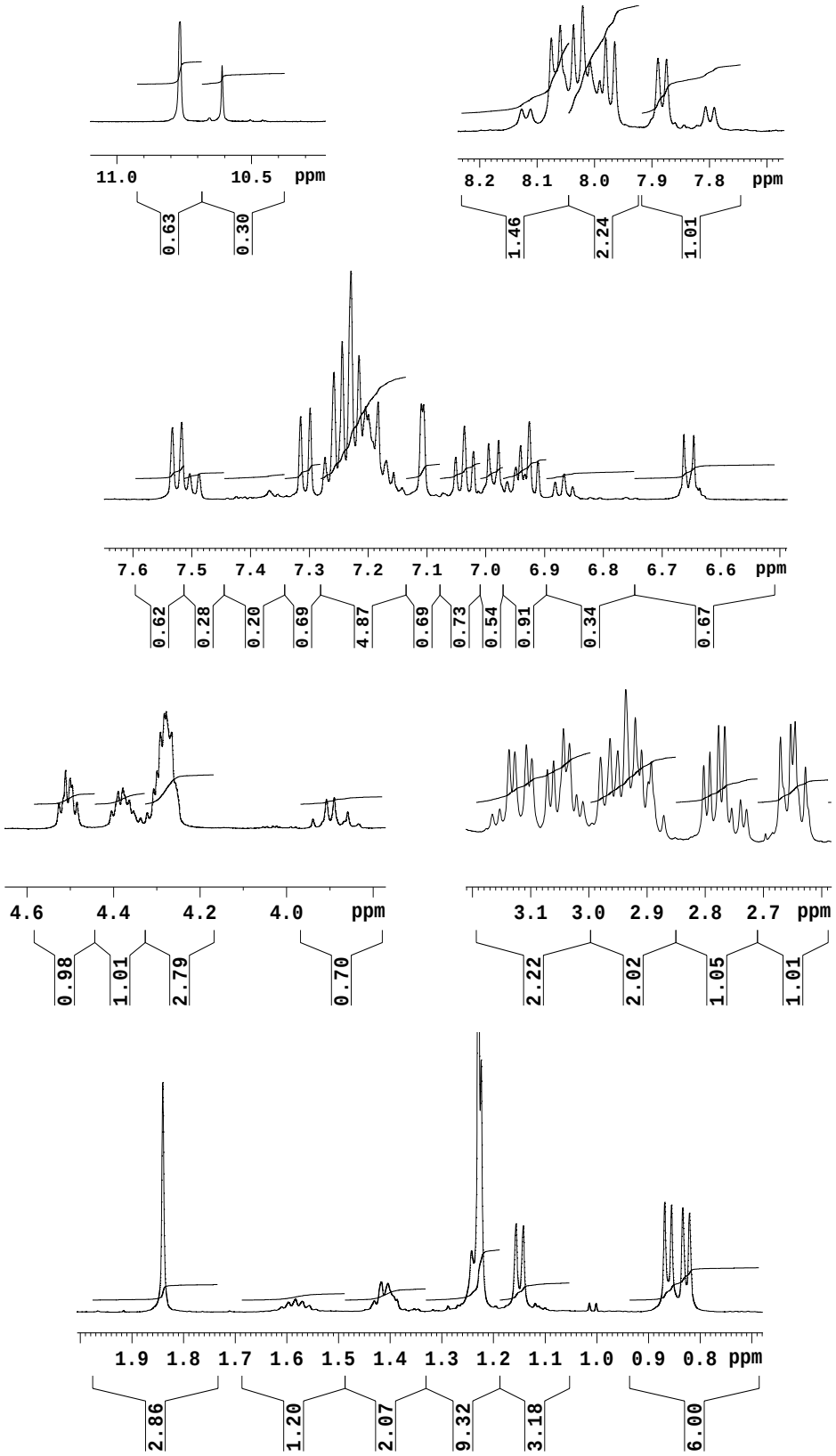


Fig-2:500 MHz ^{1}H NMR spectrum of the oligo peptide in DMSO-d_6 expansion of the, β, γ and δ, NH and aromatic region proton regions.

1) Broad peak at ~ 13 ppm is likely to be from a carboxylic acid.
2) 11-10 ppm two peaks (10.8 and 10.6) total area of which nearly corresponds to nearly one proton. These two protons are nearly in the ratio 65:30 and the low field peak of this shows a weaker J coupling.
3) Weak signal T ~ 9.2 ppm corresponding to ~ 30% of a single proton intensity.
4) The amide NH region (8.3-7.8 ppm) corresponds to ~ five protons which indicate that the compound is likely to be a penta peptide. Some minor signals can also be noticed in this region.
5) Aromatic region (7.6 – 6.6 ppm): correspond to ~ 11 protons and shows the presence of aromatic amino acids. The doublet signals at ~ 7.5 and 6.5 are typical of Trp and Tyr moieties while the multiplet pattern centered around 7.2 ppm can be from five protons of Phe moiety. Some minor signals corresponding to ~ 30% of intensity can be noticed for some of the peaks (eg. Doublet at 7.54 ppm, triplet at ~ 6.85 ppm). The doublet signal 6.65 ppm which is characteristic of a Tyr moiety shows an integrated intensity of ~0.67 which corresponds to ~ 0.34 proton.
6) The alpha proton region (4.5 to 4.0 ppm) shows the presence of five protons and confirms the presence of a pentapeptide.
7) An additional signal with a quartet pattern at ~ 3.9 ppm
8) The beta proton region (3.2 -2.6 ppm) shows the presence of six protons, probably due to three methylenes.
9) The remaining region (2-.5 ppm) shows singlets at ~ 1.84 ppm (~ 3protons) and 1.2 ppm (9 protons), multiplets at 1.6 (1 proton) and 1.4 ppm (2 protons) and doublets at 1.15 (three protons), 0.86 (three protons) and 0.82 (three protons). The latter two methyl groups may be due to a Leu moiety, while the doublet at 1.15 likely to be from an Ala moiety. The singlets at 1.84 and 1.2 ppm can be form an acetyl and tertiary butyl groups, respectively.

^{13}C NMR

The ^{1}H decoupled ^{13}C spectrum and the DEPT-135 spectrum of the sample is compared in Figs 4-6. From these spectra, the number of carbonyl carbons, the other quaternary carbons, the CH, CH_2 and CH_3 carbons can be identified. The fsummary of the observations are given below:

1) Carbonyl region (175 to 168 ppm): Shows the presence of six major peaks along with nearly six minor peaks. The presence of five aminoacids with an extra carbonyl at the N terminal end is implied from this.
2) Aromatic region (160-100ppm): As many as 11 aromatic protonated carbons could be identified from the DEPT spectrum a part form the minor signals. Out of these two resonances at 129.63 and 128.28 shows stronger intensity and the chemical shifts falls in the region expected for the *ortho* and *meta* carbons of a Phe moiety. The presence of this many aromatic protonated carbons suggest the presence of more than two amino acids with aromatic side chain. The weak signal at ~ 156 ppm is typical of a Tyr moiety. Considering the presence of Phe and Tyr moieties, which can give rise to as many as 5 protonated carbon signals, five of the remaining signals are likely to be due to a Trp moiety as it is the only aromatic amino acid that can give five distinct CH carbon signals. The aromatic region also has six quaternary carbons in addition to the carbon at 156 ppm. One of the quaternary carbons that resonate at ~ 110 ppm is typical of the indole ring of Trp moiety. Besides these signals weaker signals are also seen for almost all carbons.
3) The C alpha region (60 - 43 ppm) shows the presence of five resonances, (54.09, 53.86, 53.77, 51.68, and 48.56) confirming the presence of five amino acids.
4) Four CH_2 carbons could easily be identified for the DEPT spectrum at 41.10, 37.19, 30.35 and 27.85 ppm. This clearly shows that out of five amino acids, most likely, four of them have a beta CH_2 group.
5) There are six remaining resonances at 31.13, 24.65, 23.41, 22.97, 22.10 and 18.62 which could be coming from methyl/CH groups present in the molecule. The signal at 31.3 is very intense and based on this and its chemical shift, it can be assigned to a tertiary butyl group as it also explains the 9 proton intensity signal in the proton spectrum at ~ 1.2 ppm. Four of the remaining signals can be assigned

to the methyl groups corresponding to doublets in the high filed region of the proton spectrum and to the carbon due to an acetyl methyl (3H, 1.84 ppm in 1H spectrum).

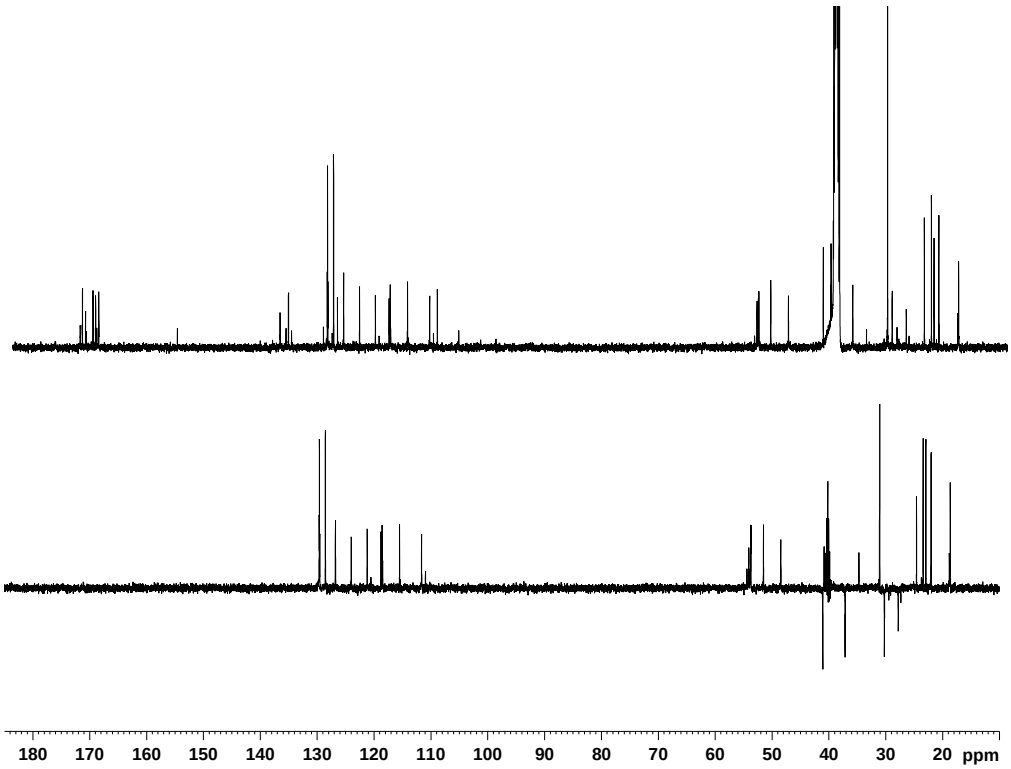


Fig-3A: 125.75 MHz ^{13}C CPD and DEPT NMR spectrum of the oligo peptide in DMSO-d_6.

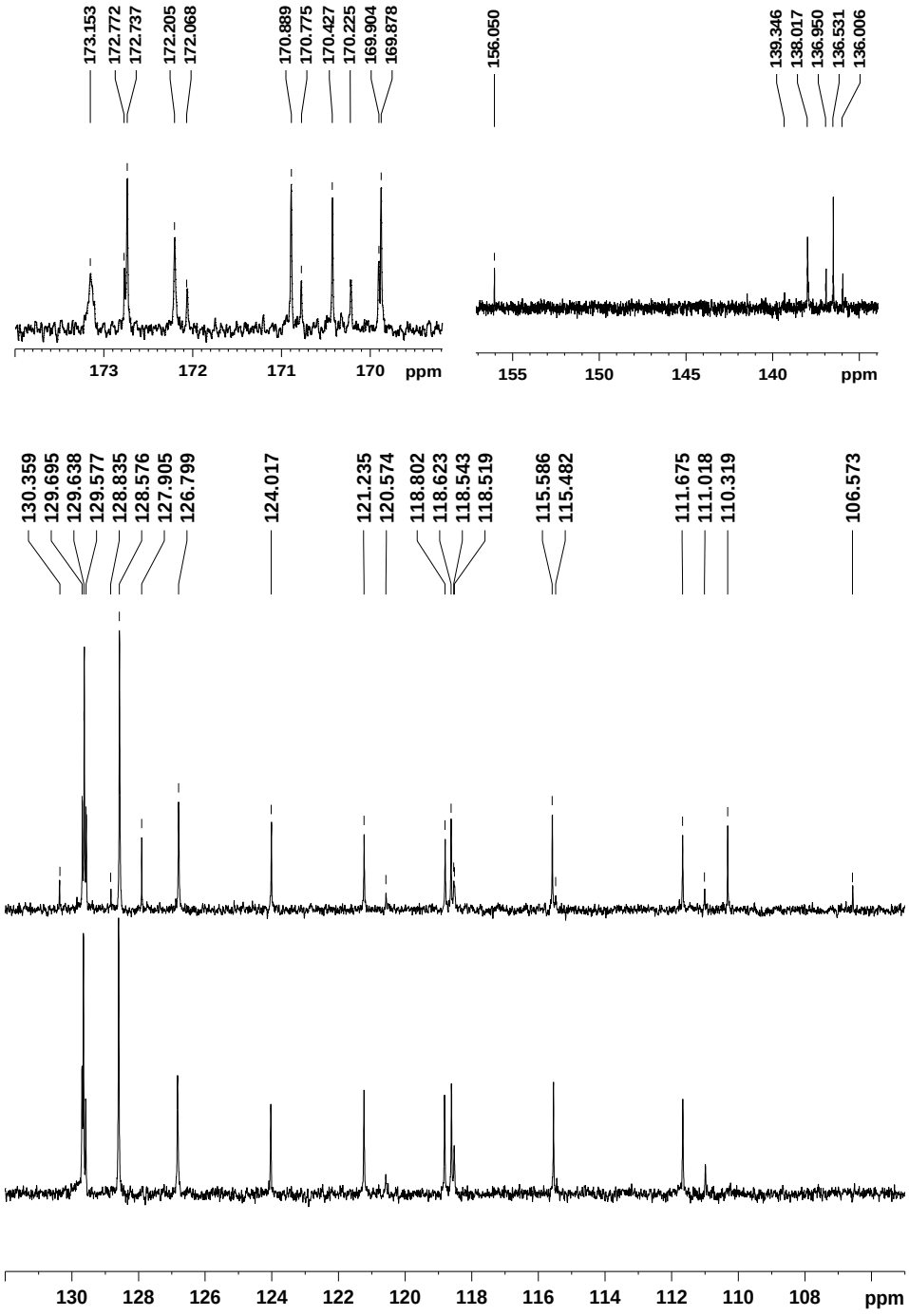


Fig-3B: 125.75 MHz ^{13}C CPD and DEPT NMR spectrum of the peptide in DMSO-d_6.

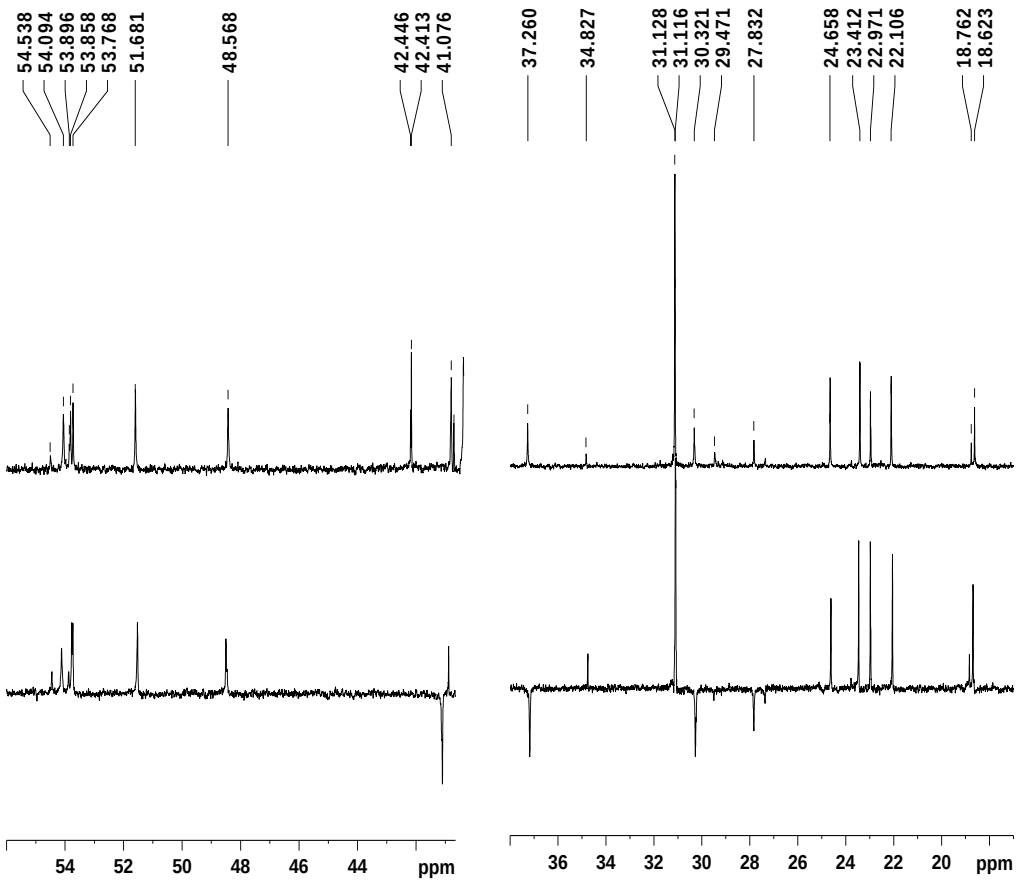


Fig-3C: 125.75 MHz ^{13}C CPD and DEPT NMR spectrum of the aliphatic region of the oligo peptide in DMSO-d_6.

COSY spectrum:

2D COSY spectrum correlates the protons that are involved in scalar coupling and hence can be used to identify immediate neighbors in spin systems of amino acids. A fig 4A and 4B shows the full and expanded COSY spectra of the compound.

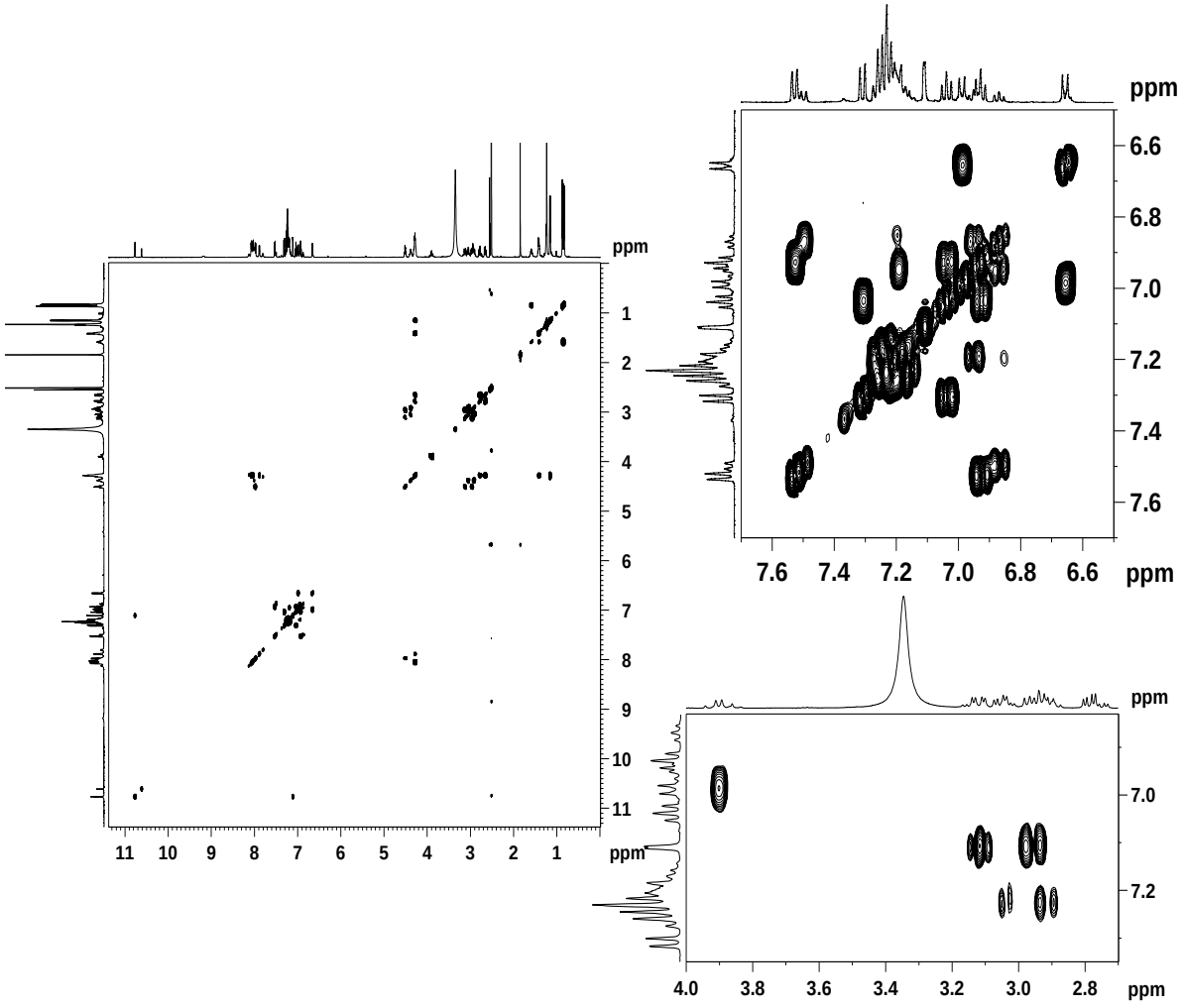


Fig-4A: 500 MHz COSY spectrum of the oligo peptide in DMSO-d_6.: Full spectrum, and the expanded spectra of the aromatic regions showing the correlation to the aromatic and β protons

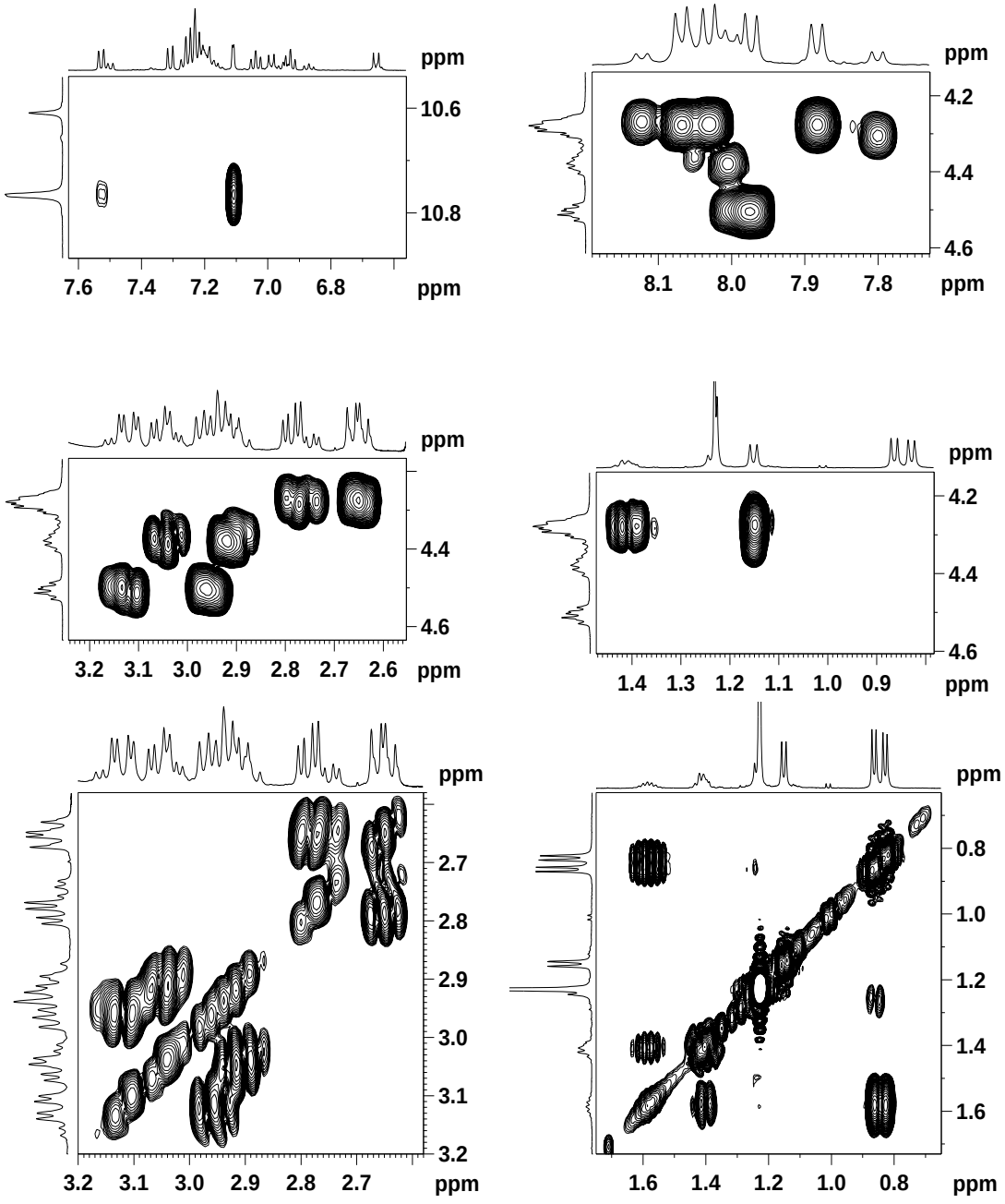


Fig-4B: 500 MHz expanded COSY spectrum of the oligo peptide in DMSO-d_6.: Correlations of low field NH to aromatic protons, α protons to NH protons, α protons to β protons, between β protons and γ and δ protons are shown.

The following details about the coupled spin system present in the compound being investigated could be gathered form the COSY spectrum:

The proton at ~10.7 ppm shows COSY connectivities to two different aromatic protons and following the COSY cross peaks the following spin net work could be identified.

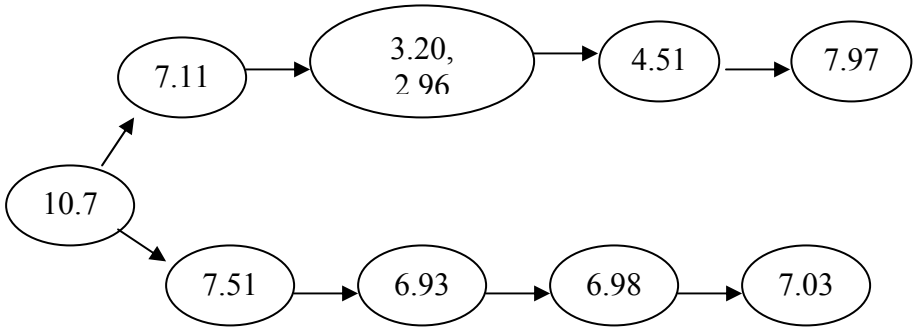


Fig (5A) Observed COSY connectivities of low field NH proton.

The presence of weak cross peak between the proton at 7.11 to two sets of protons (3.20 and 2.96 ppm) in the beta proton range (3.5 to 2.5 ppm) is good enough for the assignment of them. The chemical shifts of the protons in this spin system observed very well correspond to a Trp moiety. It is interesting to note that the minor signal at 10.6 ppm did not show any strong connectivity to any of the aromatic protons. Nevertheless, the following connectivities could be established from the COSY data for the minor peaks in the aromatic region. The chemical shift pattern observed also implies that most likely, Trp is also a part of the minor component.

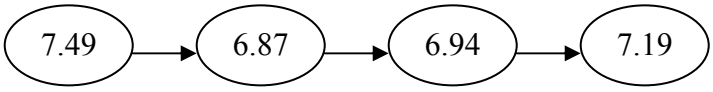


Fig. (5B)The observed COSY connectivities of the weak NH proton at 10.6 ppm.

A correlation between the doublet at 6.65 and 6.98 could also be established which most likely to be due to a Typ type of moiety. Most important aspect that has to be noticed is that the integral of these peaks corresponds only to ~30%, similar to that of the minor component. Hence, the likelihood of the Tyr being a part of the minor component can not

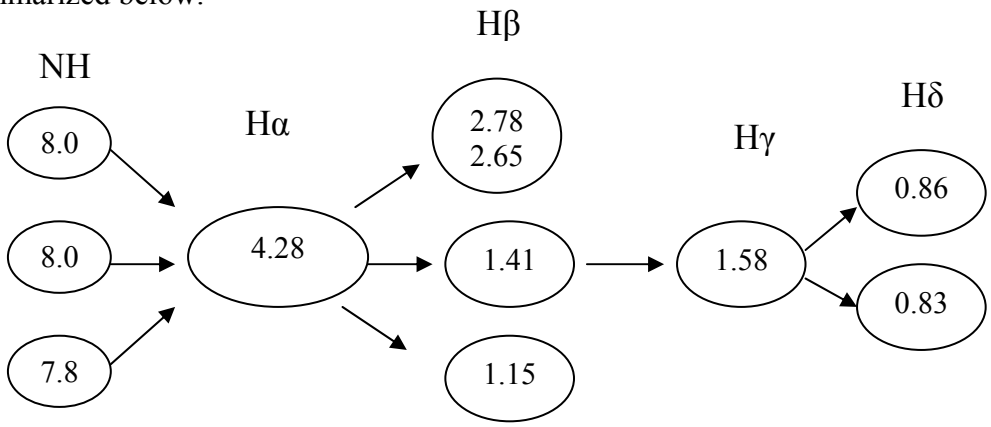
be ruled out. The low field signal also found to have a weak scalar coupling with the protons at ~ 3.98 ppm which did not show any further connetivities.

Fig-5C: The observed COSY connectivities of the aromatic proton at 6.65 ppm

Protons in the region 7.28 to 7.16 form a mutually coupled network which is likely to be of a Phe residue. Protons in this region also exhibited weak couplings to protons at 2.04 and 2.92, which are likely to its beta CH2 protons.

Correlations of the amide NH signals:

From the COSY spectrum of this region, the direct correlation of NH proton to corresponding α protons can be identified and walking through the cross peaks, the β,γ and δ protons can be sequentially assigned. Three of the NH protons are found to correlate to the α proton resonances appearing at 4.28 ppm which further shows three distinct correlations to three different types of amino acid residues. Two of these amino acids can be identified as Ala (methyl group in the β position) and Leu (methyl groups in the δ position) while the third amino acid does not have any γ protons and hence can be one of the following aminoacids: Cys, Asp, Asn, Ser. From the observed chemical shift of the β proton alone, the presence of Ser can be ruled out. The correlations observed are summarized below.



The NH proton at 7.97 correlates to the most deshielded α proton (4.51 ppm) which can be further connected to the NH of the indole ring of the Trp residue. The remaining NH at 8.00 ppm also can be correlated to the aromatic amino acid. These correlations are summarized below (Figure.5D). Interestingly, no NH peak corresponding to the Tyr residue could be seen in the spectrum.

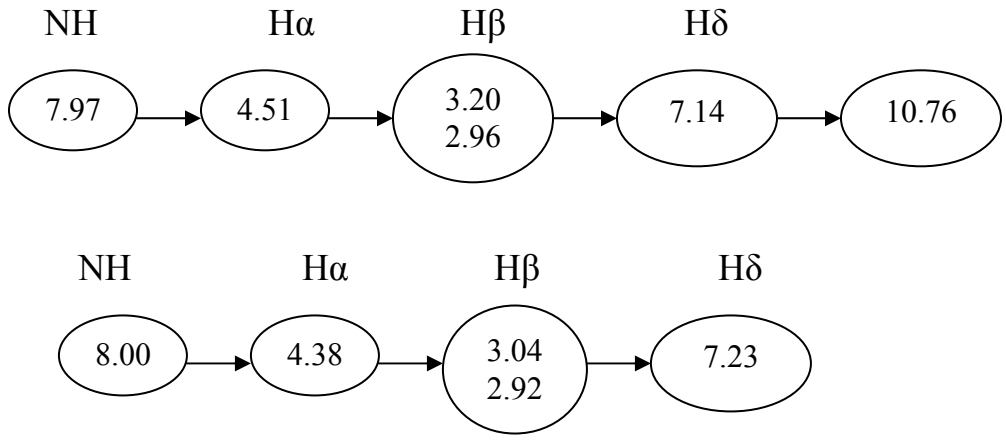


Fig-5D: The observed COSY correlations of two of the peptide NH protons

Three sets of chemically inequivalent β-CH2 protons can also be identified form the COSY spectrum. All these observations can be further supported by the TOCSY spectrum (Figure: 6A and 6B) in which correlation among-

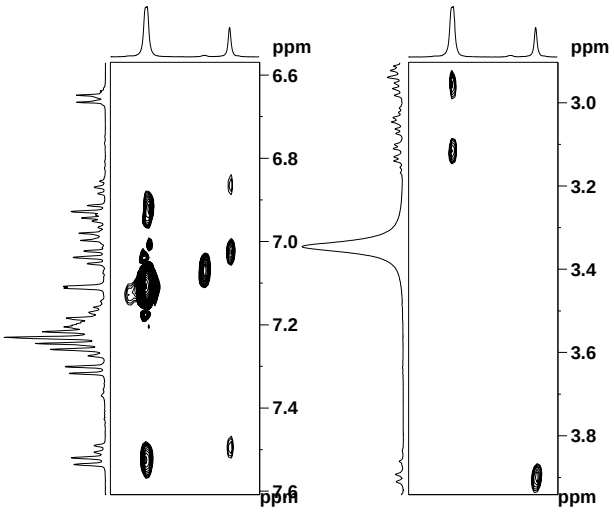


Fig-6A: 500 Mhz TOCSY correlations of the low field NH protons to aromatic and β protons

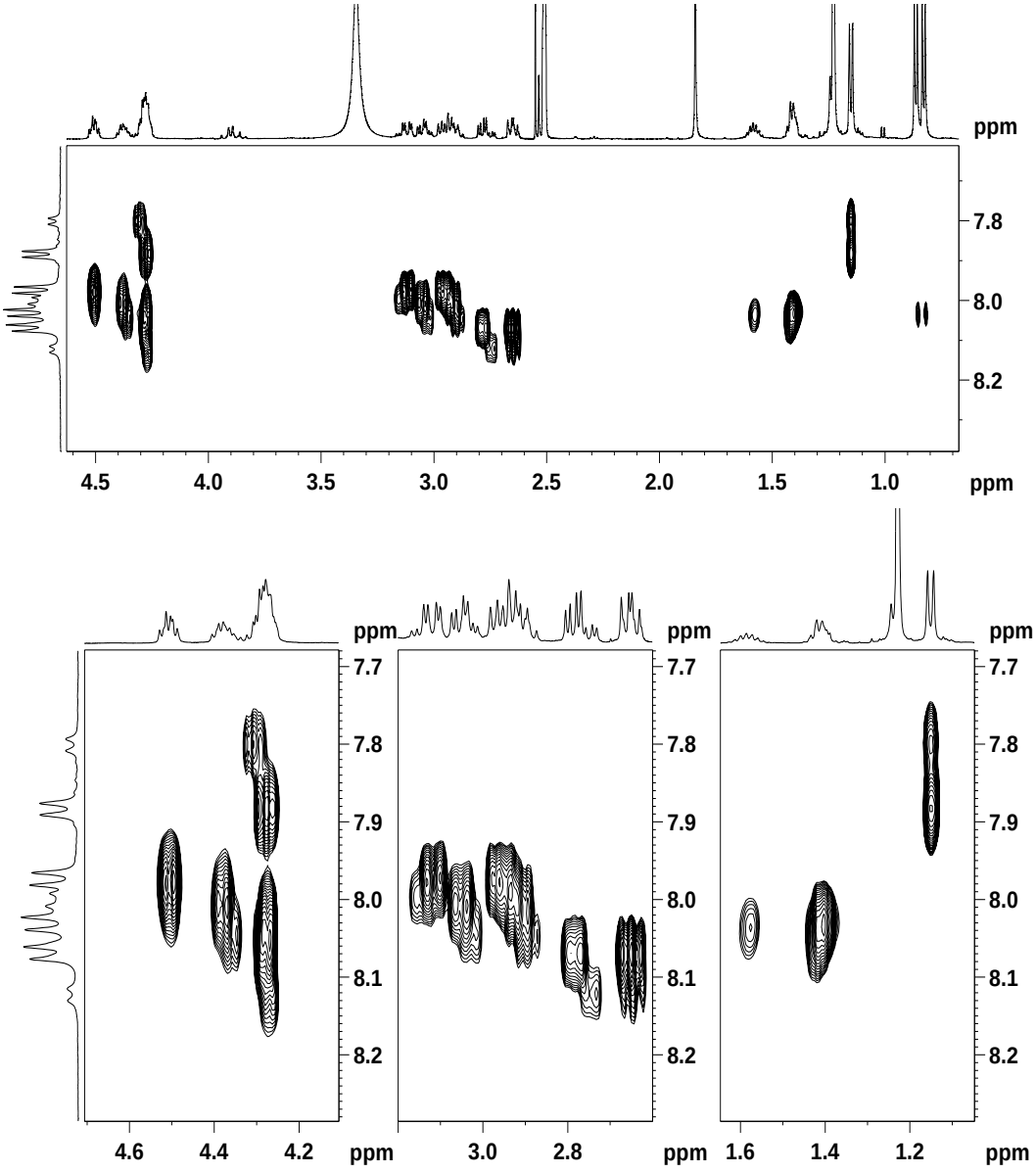


Fig-6B: 500 Mhz TOCSY correlations of the peptide NH protons to α , β, γ and δ protons

The TOCSY data clearly shows that the NH at 7.88, 7.97, 8.00, and 8.03 belongs to the Ala, Trp, Phe and Leu residue, respectively. The remaining amino acid residue has the most deshielded amide NH proton resonance.

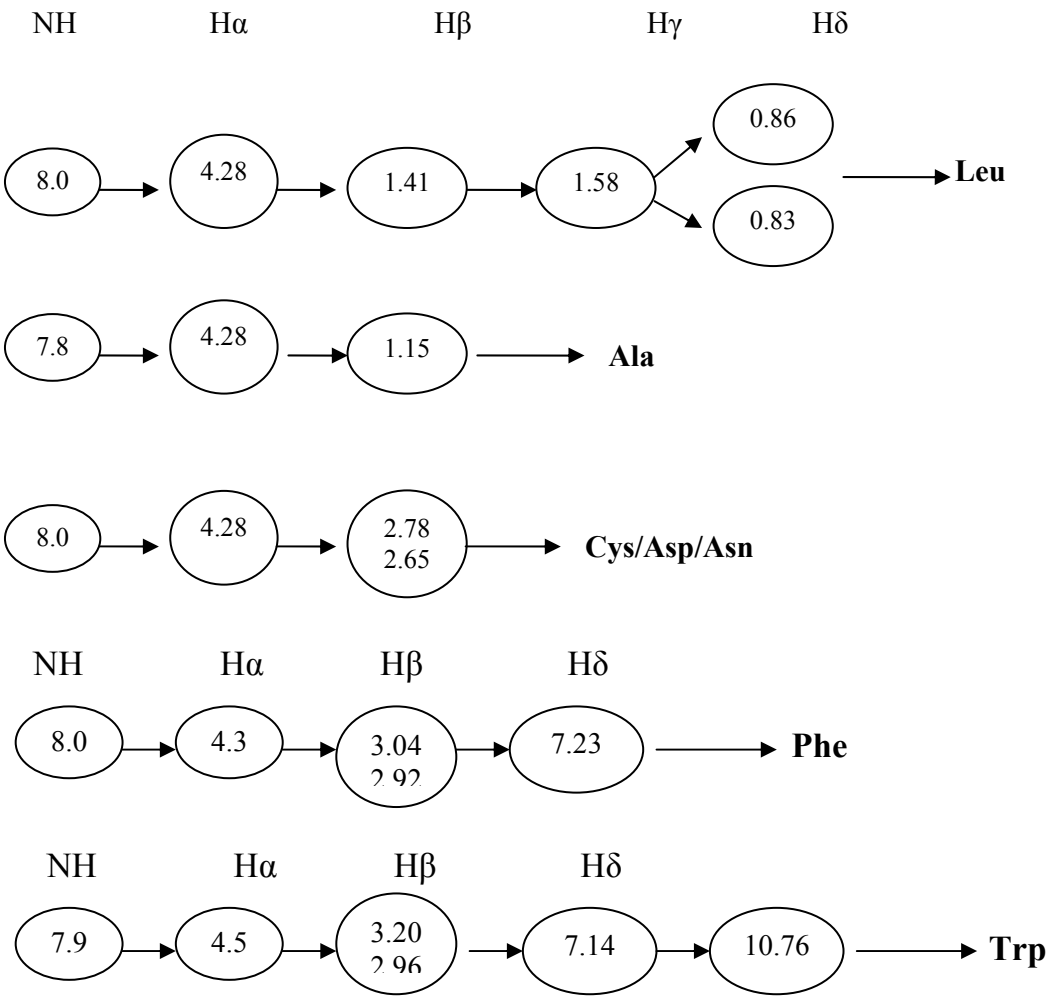


Fig-6C: The observed connectivities of various peptide NH protons

Another interesting feature that has borne out from the TOCSY experiment is the connectivity of the minor signals to the Tyr like moiety and the quartet at ~ 3.89 ppm. In addition, the correlations of all the aromatic protons can also be traced out from the spectrum. Identification of the fifth amino acid residue and the sequencing of the penta peptide require a detailed investigation by hetero nuclear correlation (^{1}H-^{13}C HSQC, HMBC) and Overhausser enhancement spectroscopy (NOESY and ROESY). The hetero nuclear experiments will not only help in the unambiguous assignment of the carbon

resonances but also to identify the adjacent amino acid residues. The observed chemical shift of the β methylene group of the unidentified amino acid (2.784 and 2.65) and the presence of only six carbonyl groups, five for the amino acid residues and one for the acetyl group, seem to suggest the presence of a Cys residue.

Hetero-nuclear correlation studies:

a) 1H-13C HSQC

From the 1H-13C HSQC spectrum information about the carbons with attached protons can be obtained. The various correlations obtained from figure 7A are given in the following table:

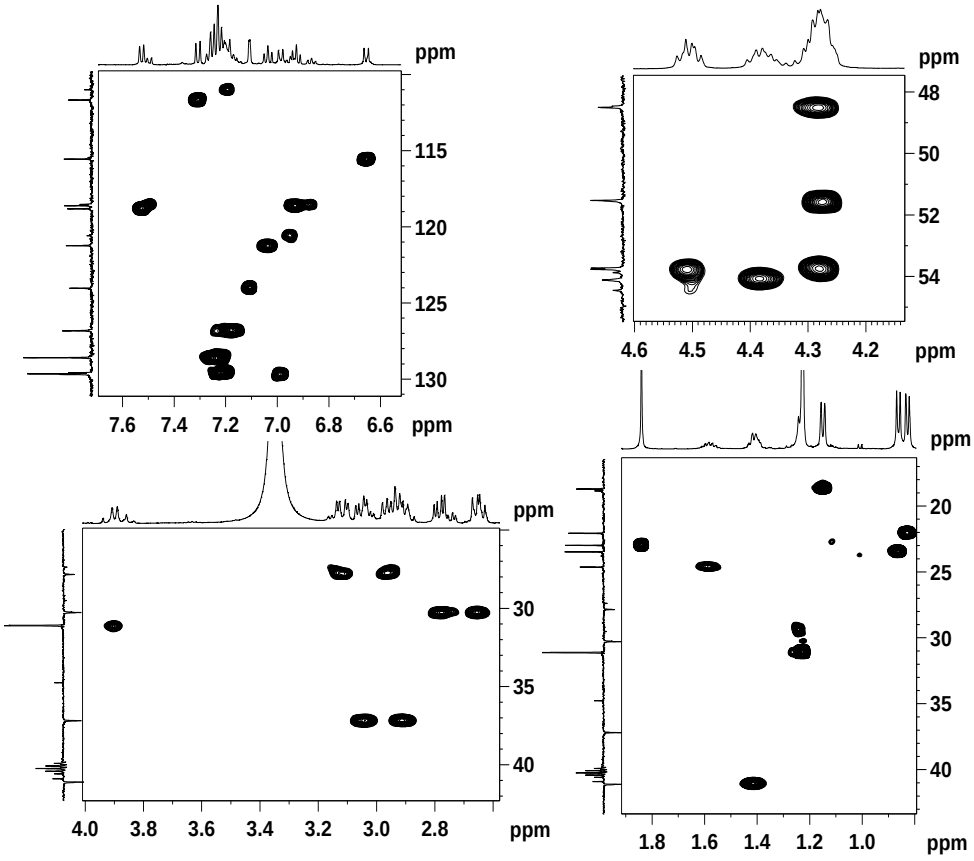


Fig-7A: Expanded 1H-13C HSQC spectrum of the penta peptide: aromatic protons, α protons, β protons and other aliphatic protons

The signals at 10.76 ans 10.6 did not show correlation to any of the carbons and suggest that they are indeed the NH proton of the indole ring in the Trp residue. Similarly the weak signal at 9.18 also did not show any carbon connectivity.

b) 1H-13C HMBC:

In the HMBC spectrum, correlations between the hetero atom and the proton which are two or three bonds away are obtained. Hence, a complete assignments of all the cross peaks obtained should not only provide information about the assignment of the quaternary carbons but also should throw light on the inter residue connectivities. The first step in determining inter residue connectivities is a proper assignment of all the carbon signals of each residue, especially that of the carbonyls.

Assignments of carbonyl carbons:

The amide carbonyl carbon of each amino acid residue should show a three bond connectivity to its β protons which can be used for the identification of the carbonyl group of each moiety. The correlations of the carbonyl carbon derived from the ^{1}H-^{13}C HMBC spectrum is given below.

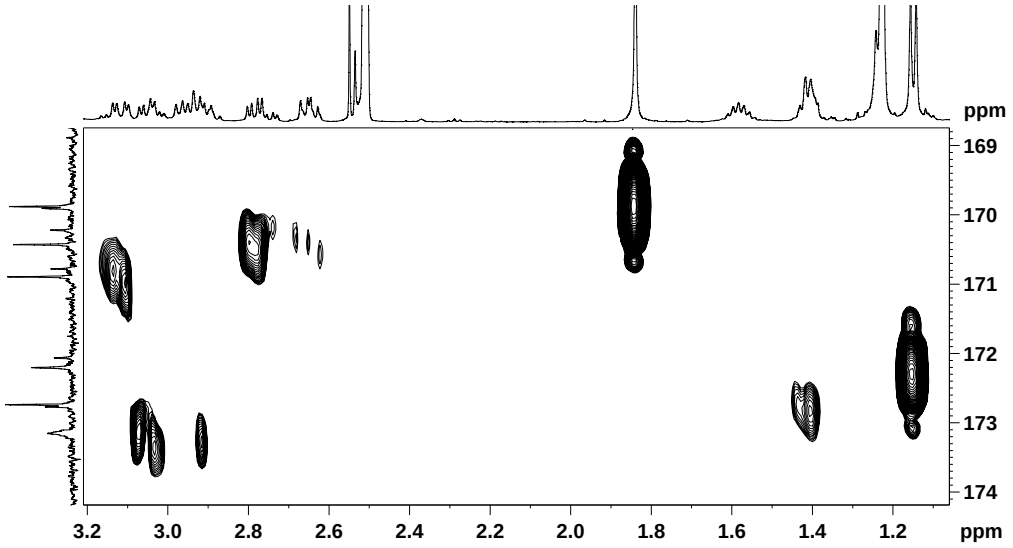


Fig-7B: Expanded ^{1}H-^{13}C HMBC correlations of the carbonyl carbons to β protons.

Correlations of carbonyl group and α proton

Carbonyl carbon of each peptide linkage in the sample studied found to show correlation to its own α protons in the HMBC experiments optimized for a long range coupling of 6 Hz. The carbonyl group of Ala residue showed a weak correlation to α proton of Phe residue which is suggestive of the proximity of them.

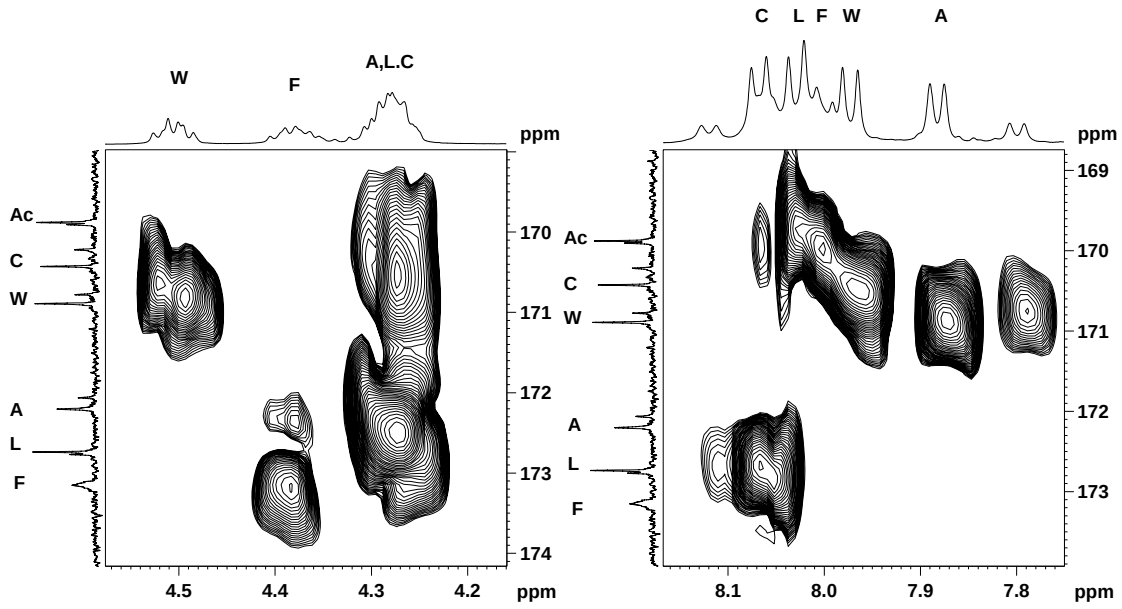


Fig-7C: Expanded ^{1}H-^{13}C HMBC correlations of the carbonyl carbons to α , Protons. Correlations of the NH protons to carbonyl carbons.

The conclusions that can be arrived at from these 1H-13C HMBC correlations are given below:

1) The C terminal amino acid is likely to be Phe since its carbonyl does not show any correlation.
2) Leu is the terminal amino acid and it is acetylated (the acetyl carbonyl at 169.88 shows correlation to Leu NH)
3) The Leu carbonyl is linked to the Cys NH which implies the presence of NHOAc-Leu-Cys- moiety.
4) The Cys carbonyl shows correlation to the Trp NH. Hence the presence of NHOAc-Leu-Cys- Trp moiety could be inferred.
5) The Trp carbonyl (170.49) is linked to Ala NH at 7.883 and the presence of NHOAc-Leu-Cys-Trp-Ala- moiety could be envisaged.

6) Considering Phe as the C-terminal amino acid, the following sequence satisfies the observed correlations.

NHOAc-Leu-Cys-Trp-Ala-PheCOOH

Location of the t-Bu group:

The strong ^{1}H signal at 1.23corresponding to 9 protons and the ^{13}C signal at ~ 31.23 to a t-butyl group (figure-7D). This protons signal shows strong HMBC correlation to a quaternary carbon at 42.52 and the methyl carbon at 31.23. The quaternary carbon at 42.52 showed a weak

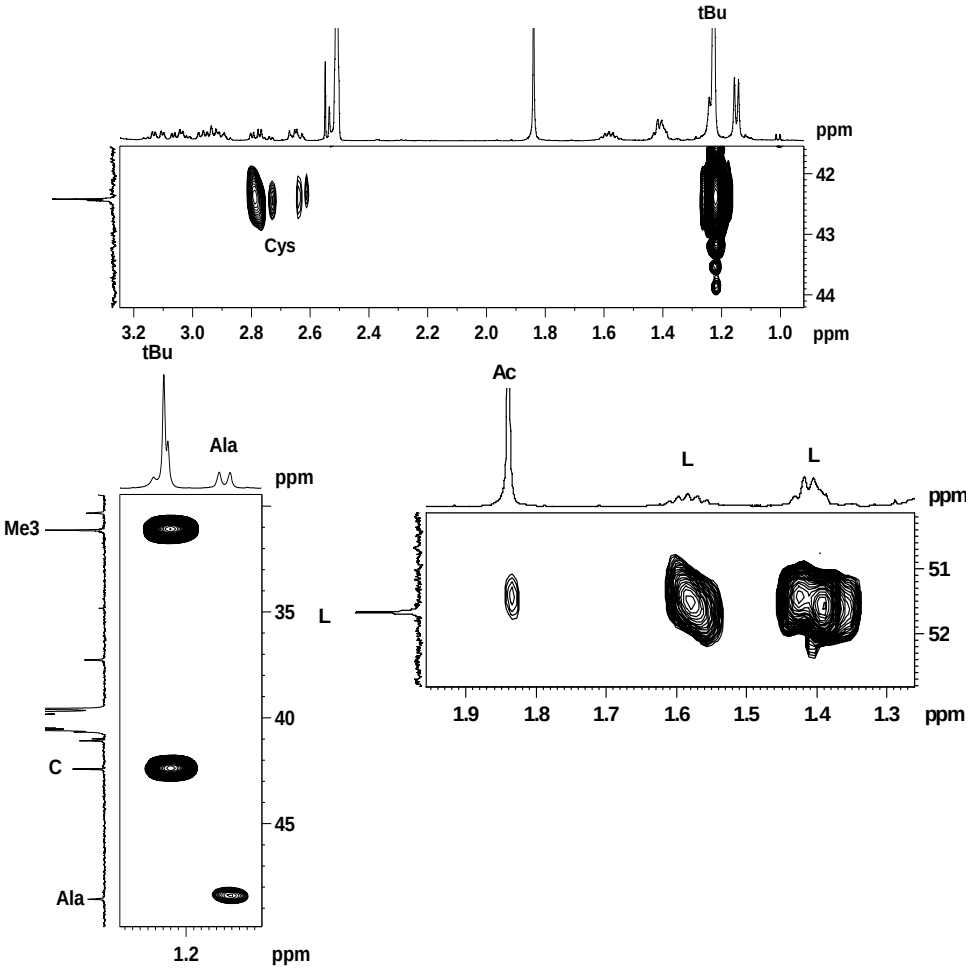


Fig-7D: Expanded ^{1}H-^{13}C HMBC correlations of the t-butyl and acetyl carbons.

Correlation to the methylene protons at 2.786 and 2.659, the β protons of the Cys residue. This gives an unambiguous evidence for the modification of the thiol proton of Cys by a tBu group or in other words the presence of CysS-tBu functionality. Evidence for the modification of the NH_2 group of N terminal Leu by the aceyl group can also be obtained from a weak cross peak between the α carbon of Leu at 51.68 and the methyl proton signal at 1.84 ppm. Thus, the structure of the penta peptide can be arrived at as:

NHOAc-Leu-Cys (S-tBu) -Trp-Ala-PheCOOH

NH to Cα, Cβ Correlations:

Correlation of the NH protons to its own α and β carbons can be used for the confirmation of their assignment. Fig – and Table summarizes the observed correlations.

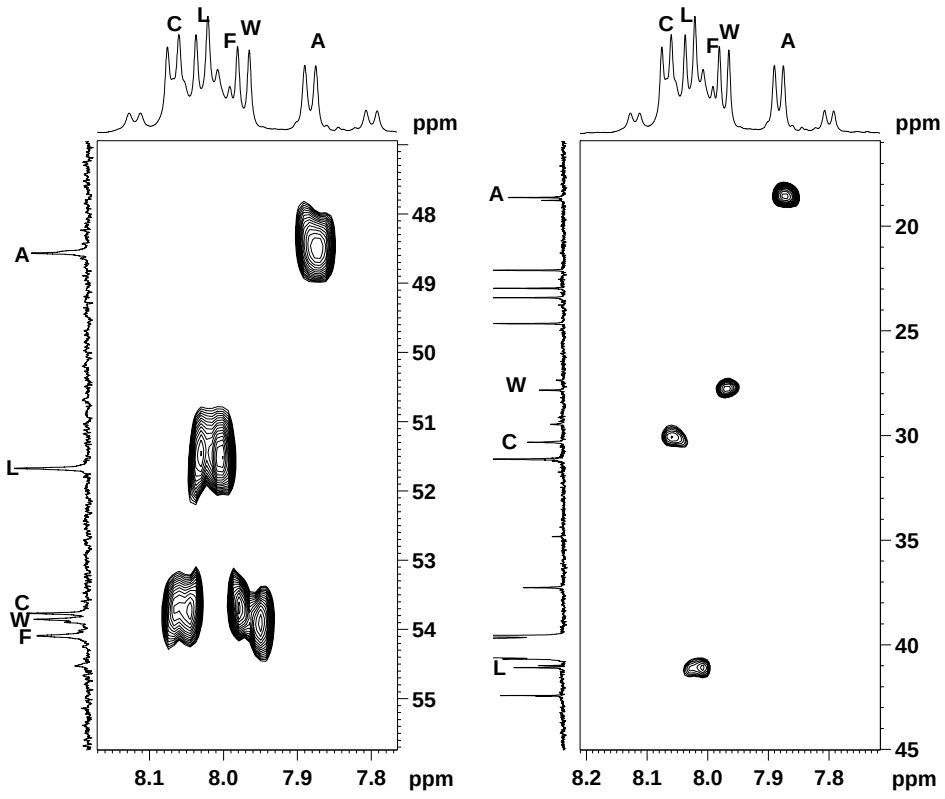


Fig-7E: Expanded ^{1}H-^{13}C HMBC correlations of the NH protons to α and β carbons.

Correlations of aromatic moieties: (Major resonances)

The HMBC correlations observed for various aromatic carbons (Figure 8A) to the protons can be used for the unambiguous assignments of aromatic amino acid residues.

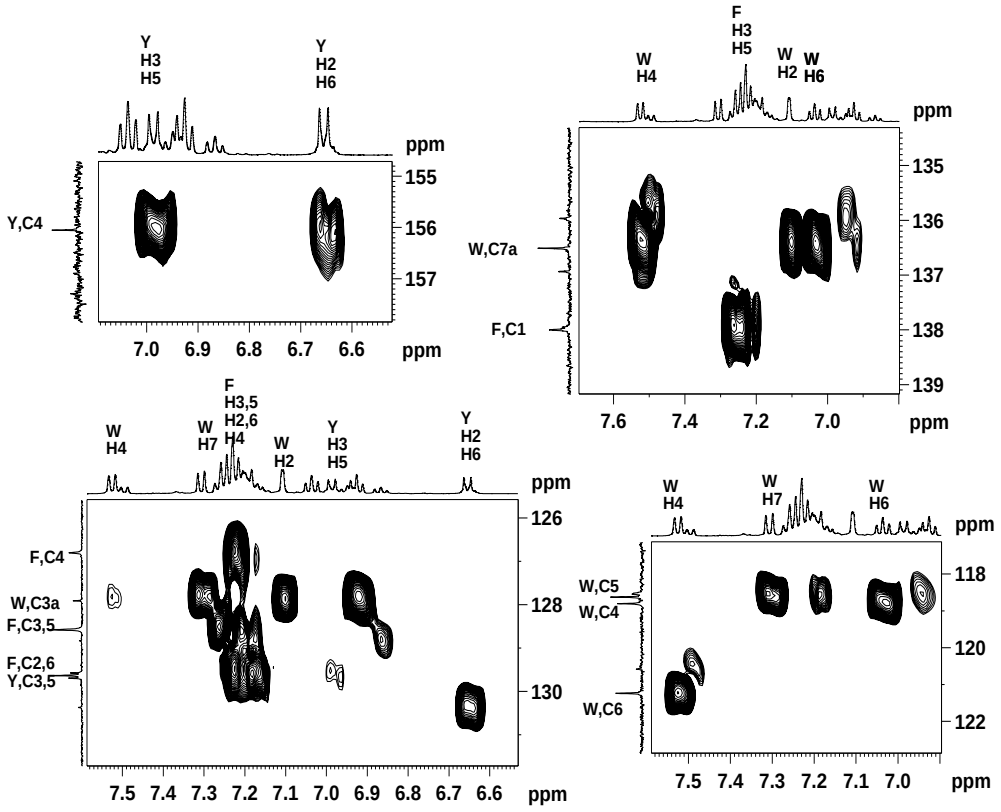


Fig-8A: Expanded ^{1}H-^{13}C HMBC correlations of aromatic protons to aromatic carbons.

We have already mentioned about the presence of weak signals which has been attributed to the co existence of another oligo pepetide, mostly a penta peptide with some modified aromatic amino acid. The details of the correlation of these weak signals could also be picked up in the 1H-13C HMBC experiments which provide very useful structural information. These are presented in Table-1and can also be seen in Figs1 and 2.

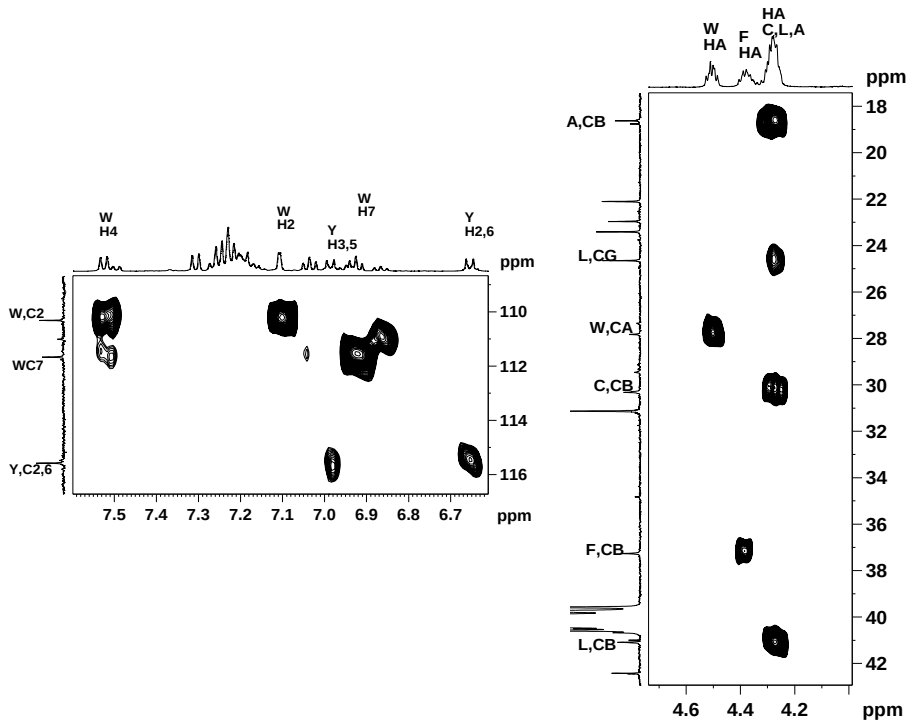


Fig-8B: Expanded ^{1}H-^{13}C HMBC correlations of aromatic protons to aromatic carbons and α protons to aliphatic carbons.

Correlations of α proton:

The α proton of each amino acid residue will show correlation to carbons on the side chain (Figure 8B) and also to carbonyl groups. The connectivities of α protons to the side chain carbons are presented in the table: 1.

Correlations of β protons:

The β protons of each amino residue show HMBC correlation to its own carbonyl, carbons which are two (α, γ carbons) and three bonds away from (δ carbon) it. These are useful in identifying the proton and carbons resonances of a particular residue. The carbonyl to β proton correlation has already been discussed before. The other correlations observed are presented in the table.

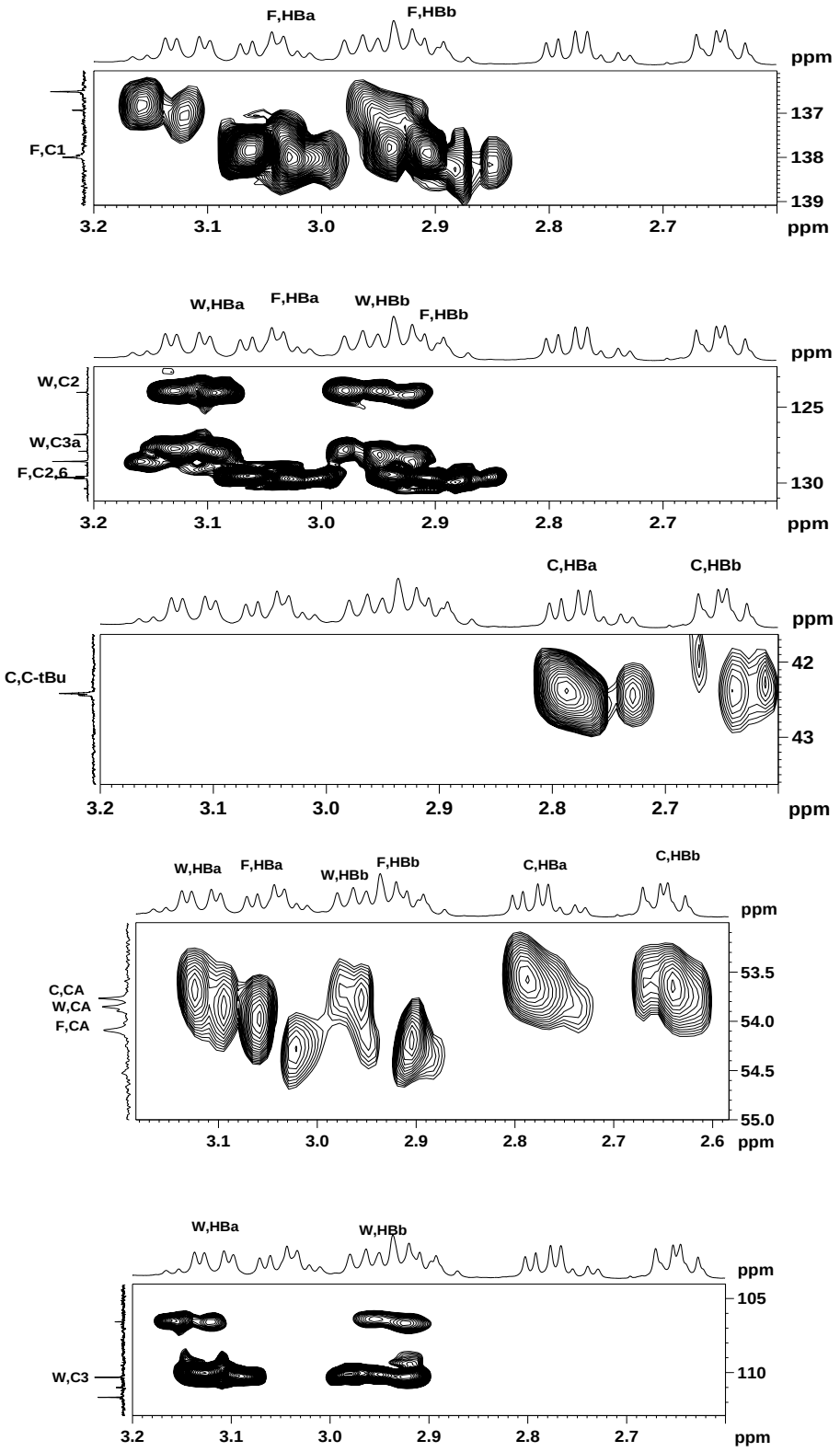


Fig-9A: Expanded ^{1}H-^{13}C HMBC correlations of aromatic protons to aromatic carbons and α protons to aliphatic carbons.

Apart from the different types of protons discussed so far the system also contains protons arising form medication of the amino acid as well as the γ and δ protons. These correlations are depicted in following table and figure 9B.

Table-15. HMBC connectivities of other aliphatic protons

Sr No	^{1}H chemical shift	^{13}C connectivity
1	1.582(Leuγ)	51.65 (Leu α),41.1(Leuβ),23.4 (Leuδ),22.1 (Leuδ),
2	1.23 (t-Bu Me)	31.1 (Me, tBu), 42.42(CtBu)
3	0.862 (Leuδ)	41.1(Leuβ), 24.66 (Leu γ),22.1 (Leuδ),
4	0.826 (Leuδ)	41.1(Leuβ), 24.66 (Leu γ),23.43(Leuδ)
5		

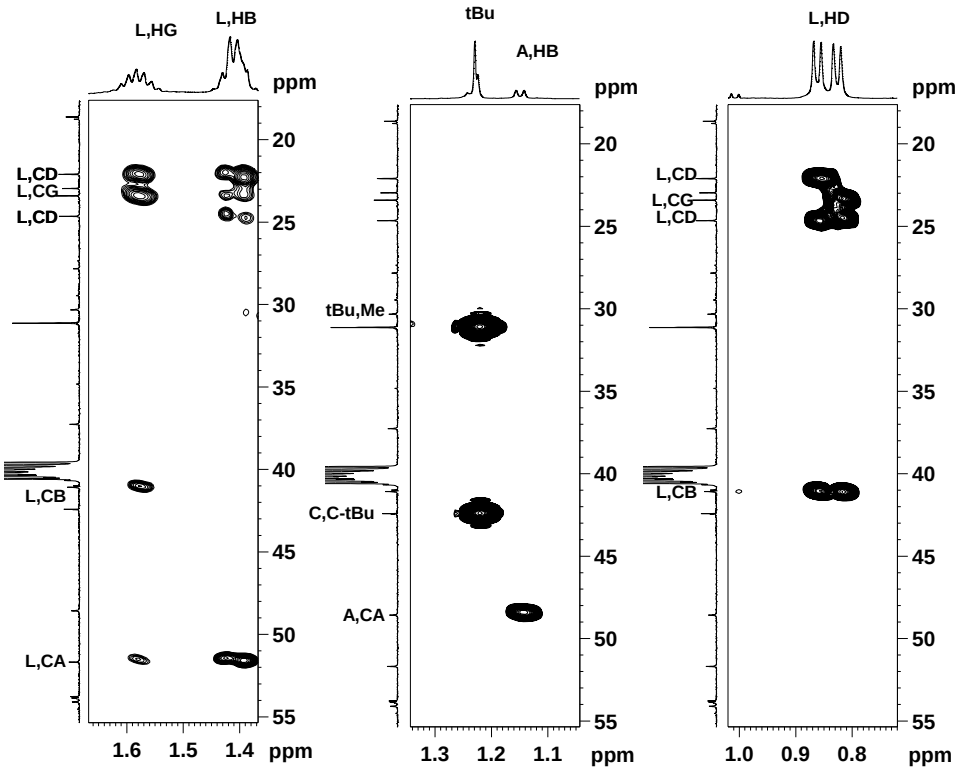


Fig-9B: ^{1}H-^{13}C HMBC correlations of β, γ and δ protons.

Structure of the minor CPI-2081

The presence of a minor product is very clear from the ^{1}H and ^{13}C spectra itself as weaker signals are seen along with the signals of the penta peptide. ^{1}H NMR shows that the minor product is present in ~ 30 Mole%. The detailed investigation carried out also shows that the minor product is also likely to be a penta peptide containing more or less the same amino acid residues. The intensities observed for the α and β protons compared to the NH and aromatic protons in the ^{1}H spectrum, and overall patterns obtained for the ^{1}H and ^{13}C spectra can be rationalized only by considering a penta peptide having identical residues with modification of one of the amino acid residues. The multiplicities of the Trp residue signals observed in the ^{1}H spectrum seem to suggest the presence of a modified Trp residue in the minor product. The most interesting observation is the presence of a Tyr type of moiety which also corresponds to ~ 30 mole% in the ^{1}H spectrum. ^{13}C NMR also confirms the presence of Tyr type moiety in the sample (quaternary carbons at 156.6 and 130.6ppm and aromatic CH carbons at129.7 and 115.6 ppm). Apart from this an additional quartet (dd) is also seen at 3.90 ppm in the ^{1}H spectrum which also corresponds to 30 mole %. The ^{13}C-^{1}H HSQC spectrum correlates this quartet to a carbon at ~ 31.1 ppm where the ^{13}C signal from the t-Bu group also incidentally appears. These protons were found to show weak J couplings in the COSY and TOCSY (**To**tal **C**orrelated **S**pectroscop**y)** spectra to the protons on Tyr like moieties. The COSY and TOCSY do not show any further weak or strong couplings. The AB quartet pattern observed with a large coupling constant (15.7 Hz) is indicative of a methylene group. The absence of any ^{1}H -^{1}H J couplings to other protons rules out the possibility of a Tyr residue as it should have shown further J couplings to its α proton. All these observations can be rationalized by considering the presence of only a Tyr side chain (-CH2 –C6H5-OH) group. The presence of such a group has been confirmed from the ^{13}C-^{1}H HMBC spectrum where a connectivity of this CH_2 group to the phenolic residue from the cross peaks seen to the aromatic CH carbon at 129.60 ppm (corresponding to protons at 6.98 ppm) and a quaternary carbon at 130.36 ppm (C4 carbon).

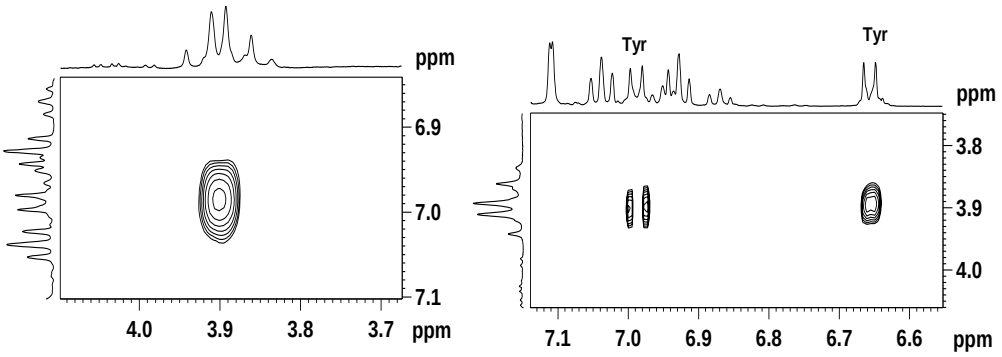


Fig-10A: COSY and TOCSY spectral correlations of the CH_2 protons at 3.98

Information regarding the connectivity of this CH_2 group is also obtained from the HMBC spectrum. The CH_2 protons at 3.98 ppm shows a correlation to a weak signal at 106.6 ppm of a quaternary carbon which is attributed to the C3 carbon as it shows correlation to the α proton of Trp residue. This correlation also suggest that the α proton of the both the major and minor signal has nearly same chemical shift. The major α proton correlates to the Trp C3 carbon at 110.4.The correlation of the weak signal at

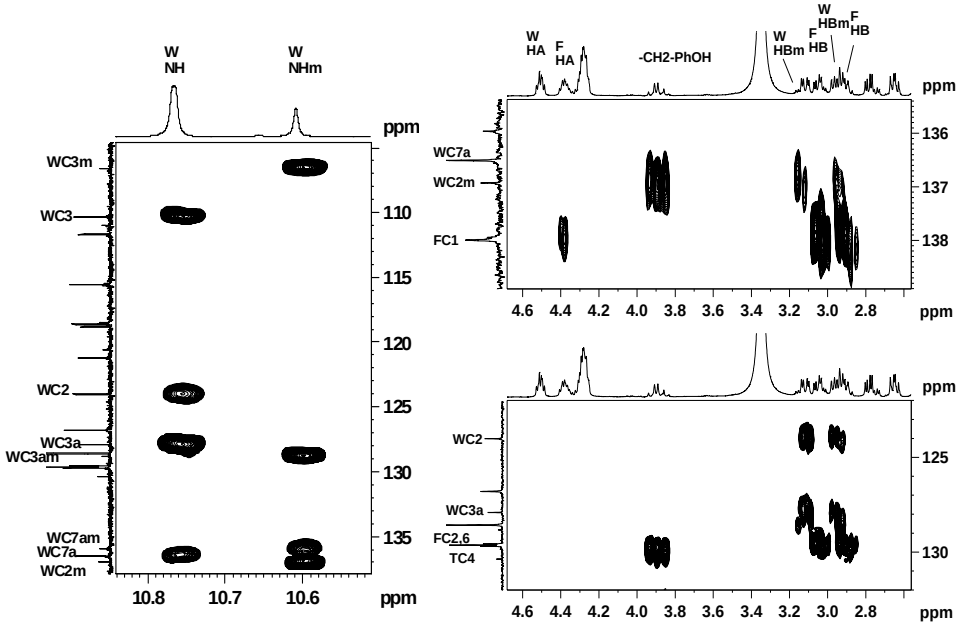


Fig-10B: ^{1}H-^{13}C HMBC correlations of the Trp NH protons and carbons of the minor signals.

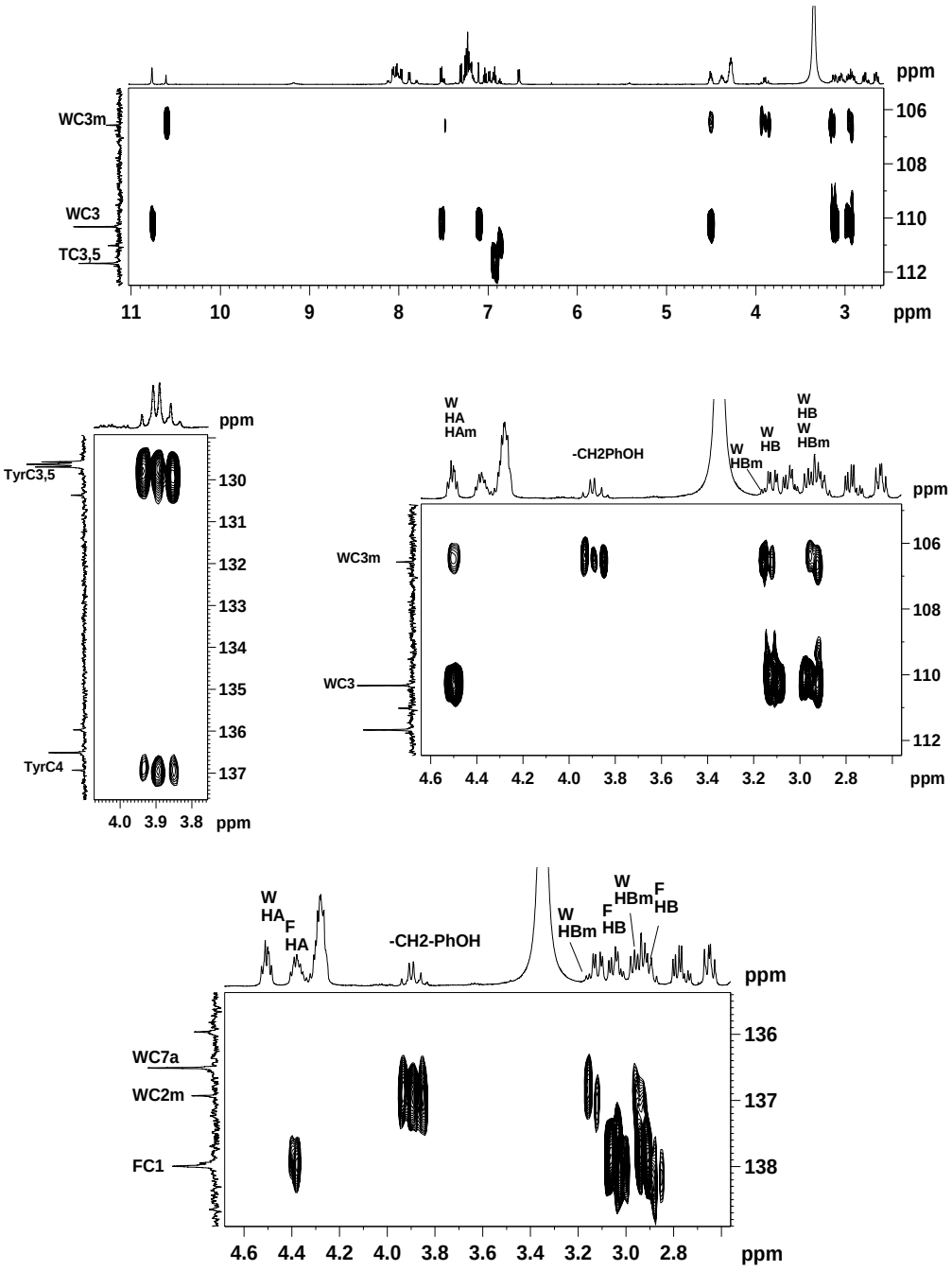


Fig-10C: Comparison of the ^{1}H-^{13}C HMBC correlations of C3 carbon of the major and minor signal of Trp residue. And ^{1}H-^{13}C HMBC correlations of the CH2 proton (a) and carbons (b, c) of the minor signals.

Hence, the proposed structure of minor CPI-2081 could be as follows: NHOAc-Leu-Cys (S-tBu) –Trp(2-CH_2 C_6H_4-OH)-Ala-PheCOOH**.** Since the modification is on only one of the aminoacids of the pentapetide, the ^{1}H and ^{13}C chemical shifts observed for the minor component is very close to that of the major signals except for the signals of Trp moiety. Almost all the carbonyl carbons show the presence of weak signals due to the modified penta peptide. The 2D DOSY experiment performed on a sample dissolved in Methanol d4 clearly indicated the presence of compounds having very close but distinct diffusion coefficients.

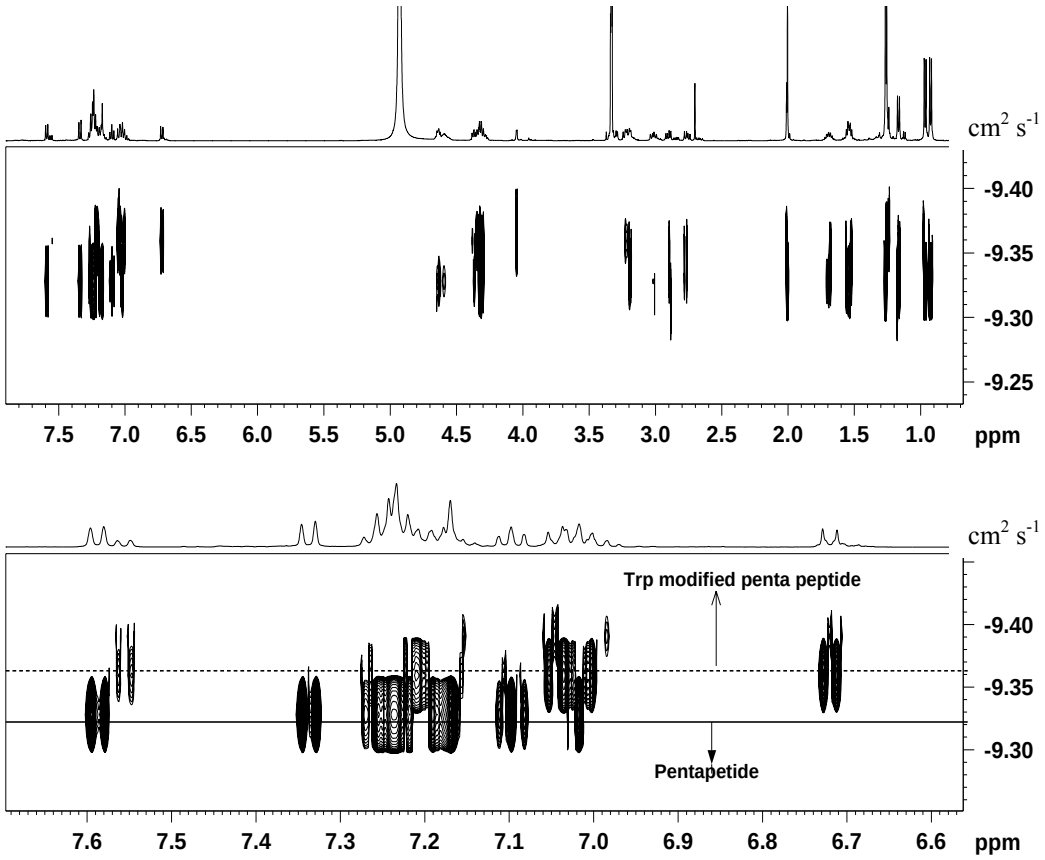


Fig-11: 500 MHz ^{1}H DOSY spectrum of the isolated oligo peptide in Methanol-d_4 (a) The aromatic region of the DOSY spectrum is expanded in (b) The dotted line and solid lines corresponds to the self diffusion coefficient of the minor and major penta peptide.

Overhauser Enhancement spectroscopic studies and conformation of the penta-peptide: In order to get information about the conformation and verifying the sequence proposed, 2D NOESY and ROESY experiments were performed on the sample in DMSO-d6.The NOESY experiment showed strong and positive cross peaks indicating

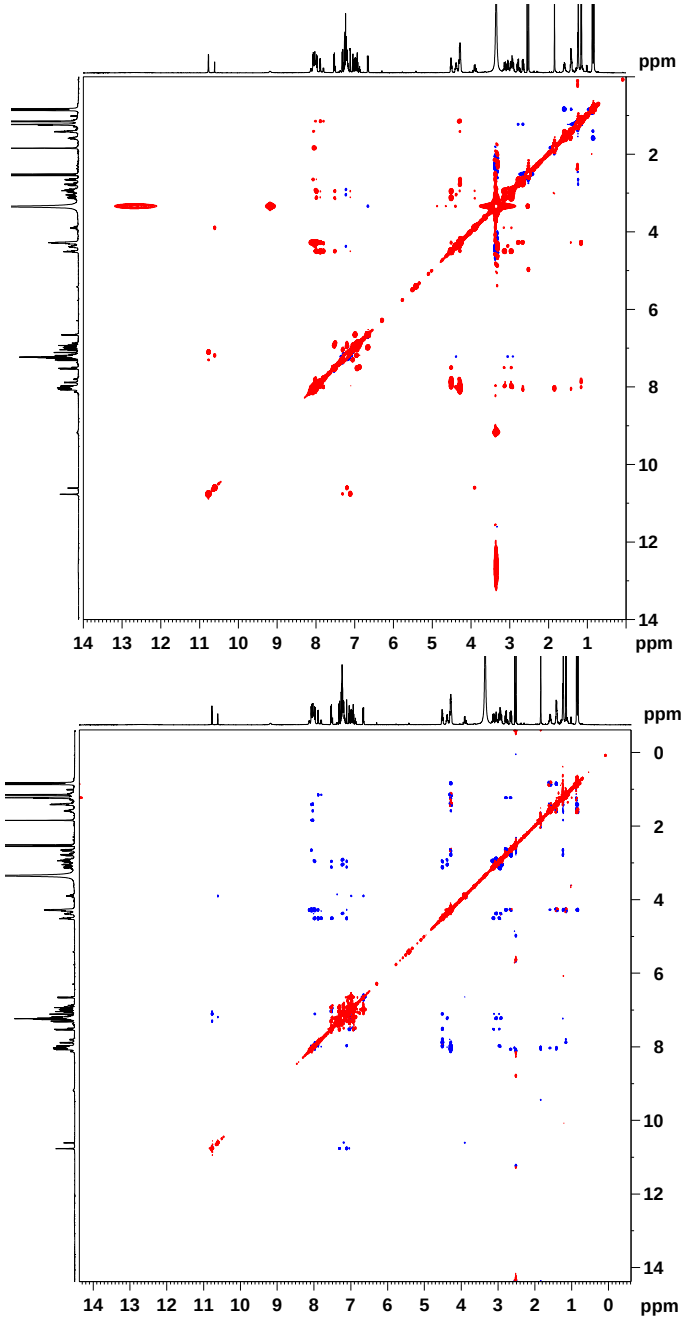


Fig-12A: 500 MHz NOESY (top) and ROESY (bottom) spectra of the penta peptide in DMSO d_6 .

the molecule tumbling rate tends towards the long correlation limit. In this regime and at the mixing time used (750 ms), the cross peak intensities can not be directly correlated to inter nuclear distance due to the presence of a phenomenon called spin diffusion. Hence, a ROESY experiment, where the real NOE's are manifested as negative cross peaks, was performed and the results are shown in the following figures. Fig- 500 MHz NOESY (top) and ROESY (bottom) spectra of the penta peptide in DMSO d_6 The ROE cross peaks seen between the indole NH at 10.8 and 10.6 ppm to the other aromatic protons confirms the assignments of the both major and minor proton signals of the Trp residues. The weak NH signal shows additional cross peaks to the protons attached to the phenolic ring and the CH2 group at 3.98 ppm.

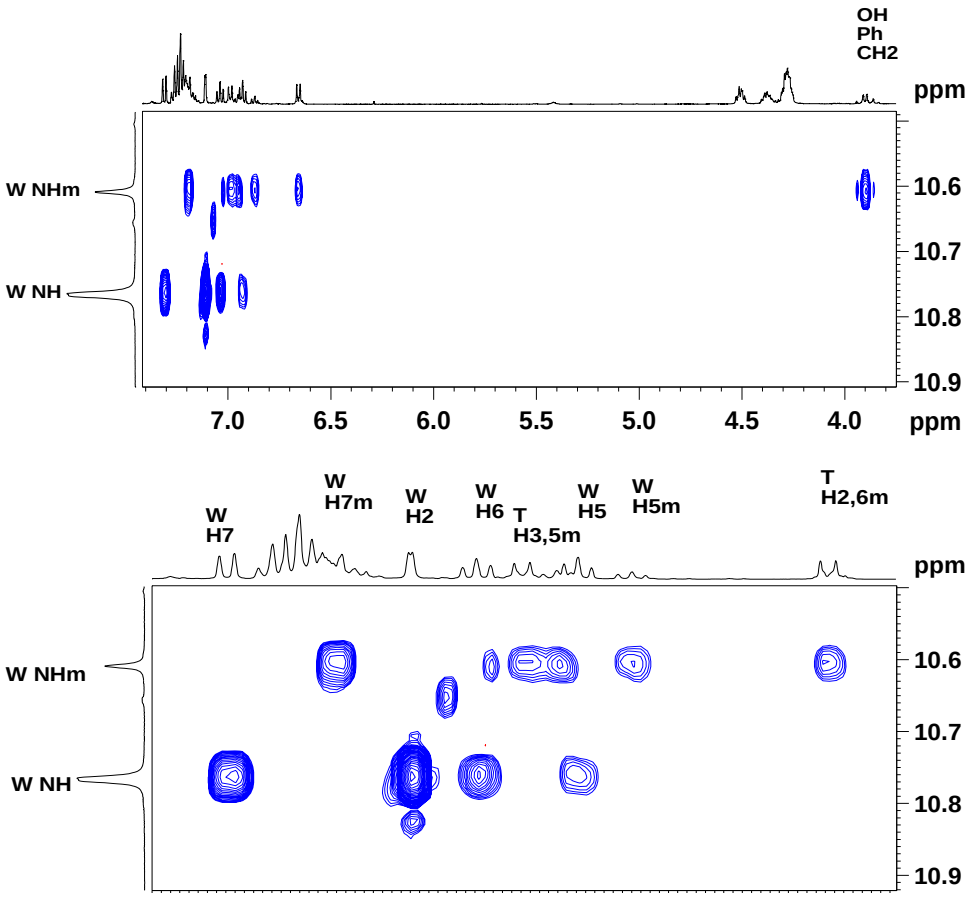


Fig- 12B: 500 MHz ROESY spectrum of the penta peptide in DMSO d_6 : ROE correlations observed for the Trp indole NH. The spectrum below is the expanded spectrum of relevant part.

Cross peaks obtained between the Cys NH at 8.06 to Trp NH at 7.98 and the correlation of the latter to Ala NH at 7.88 confirms that these three residues are connected by the peptide linkage and the Trp residue is flanked by the Cys and Ala residues. The Ala NH also shows cross peaks to the α (4.504) and β (3.052, 2.943) protons of Trp residue. Similarly, strong cross peaks are also observed between the Trp NH at 7.98 and the β protons of the Cys residue at (2.786, 2.659). Strong correlation of Ala methyl group to aromatic protons of Phe residues are also observed and thus provide information about their proximity. The Cys NH at 8.06 also found to have strong ROE to the Leu β (1.41) and γ (1.58) protons as well as to the acetyl group at 1.84. Strong cross peak obtained for the acetyl methyl and Leu NH at 8.03 proves the presence of an N terminal acetyl group attached to the Leu residue. The evidence for the substitution of thiol proton of Cys by t- butyl group is confirmed form the strong NOE cross peak between th et Bu Me and the β protons of Cys.

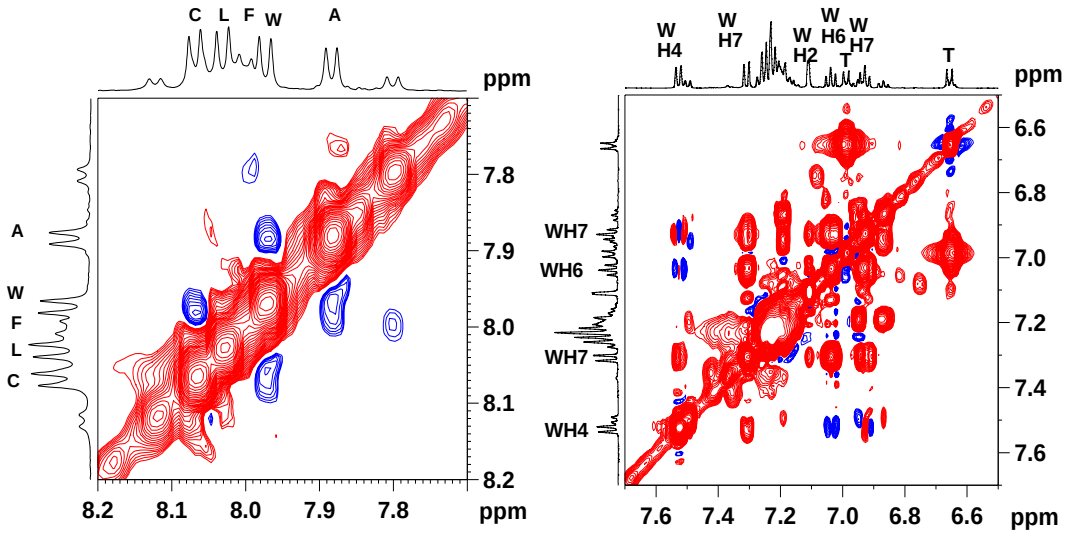


Fig- 13C: 500 MHz ROESY spectrum of the penta peptide in DMSO d_6 : ROE correlations observed for the peptide NH and aromatic protons.

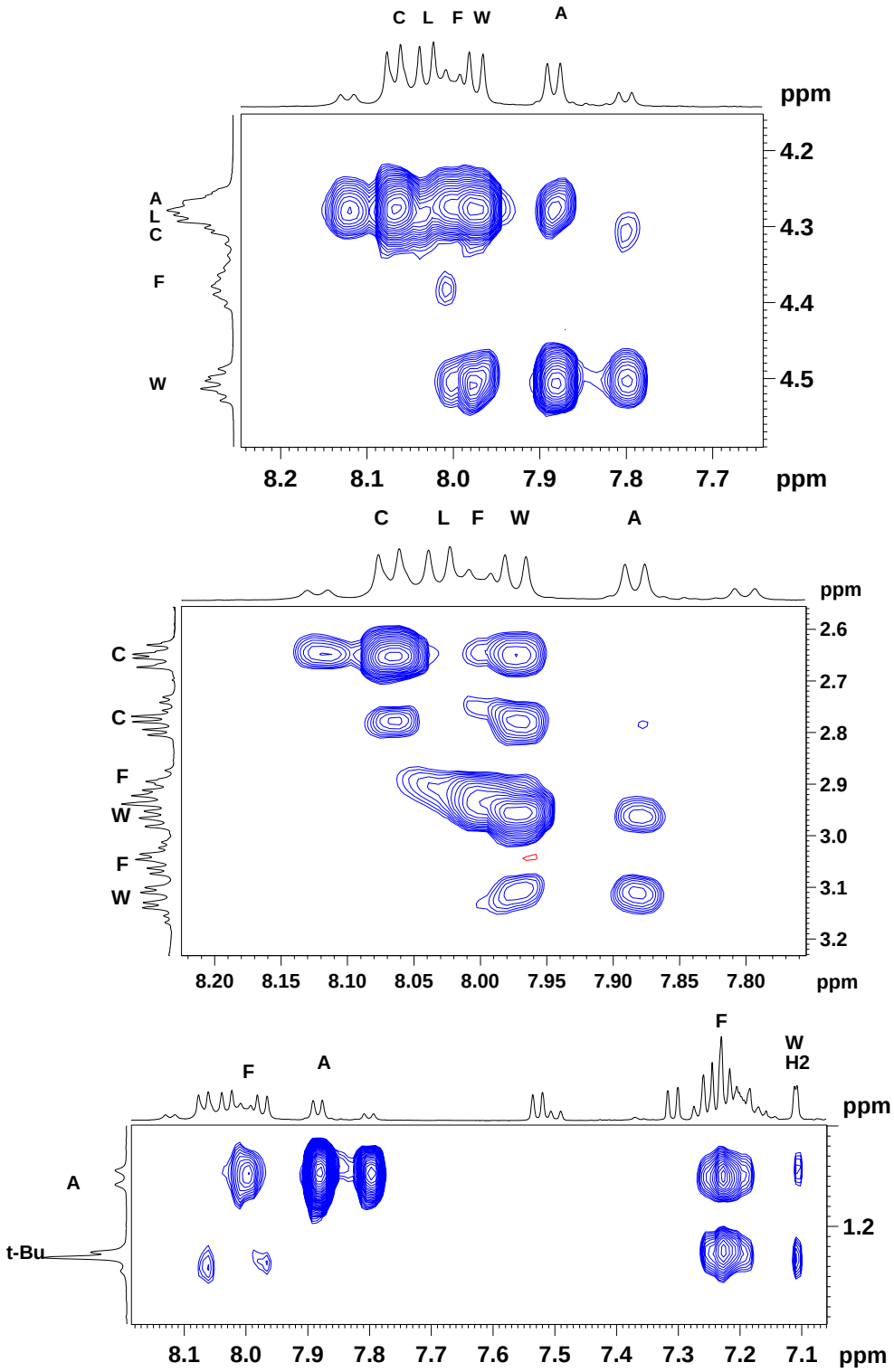


Fig- 13D: 500 MHz ROESY spectrum of the penta peptide in DMSO d_6 : NH to α, β, γ and δ protons.

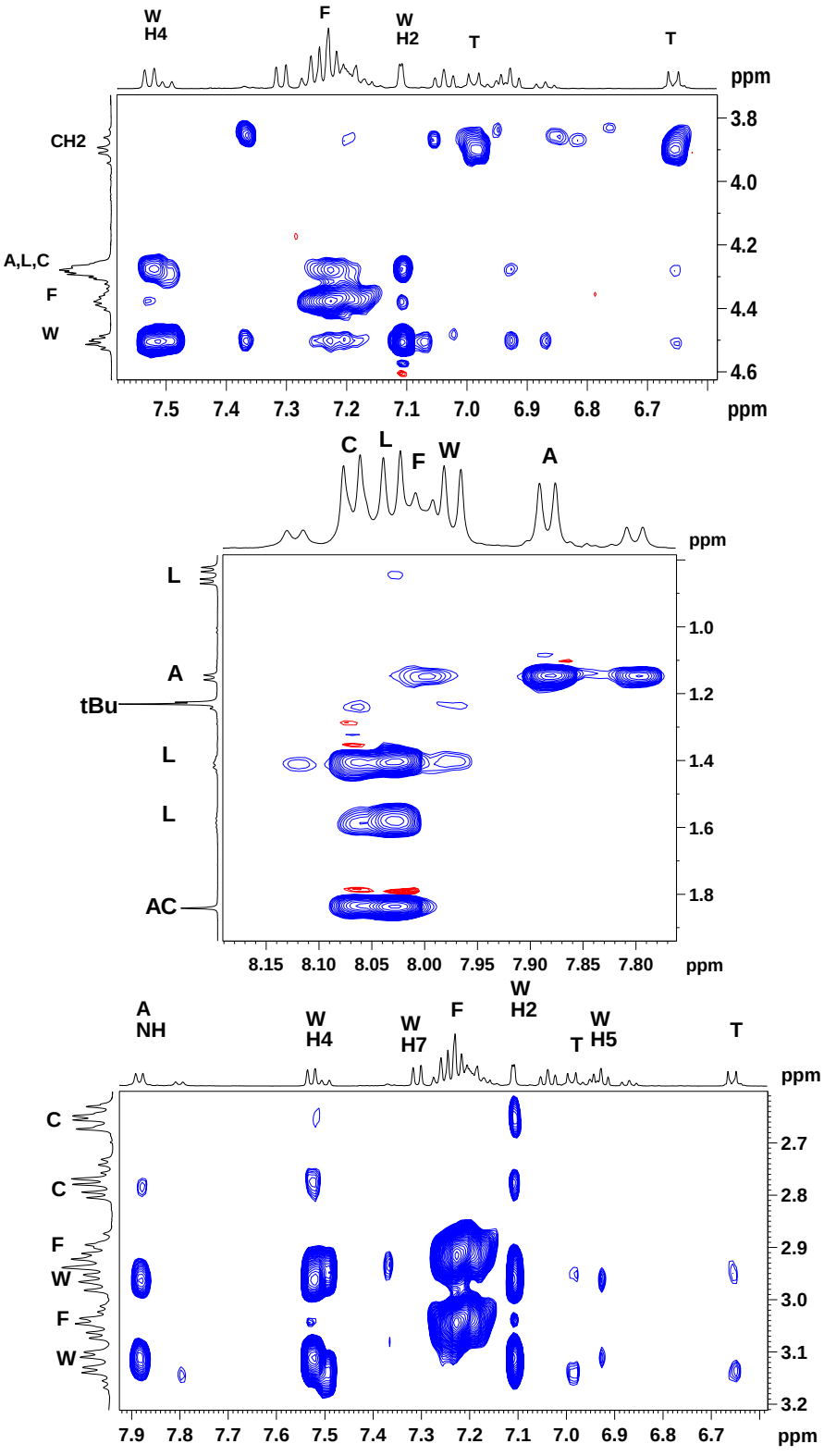


Fig- 13E: 500 MHz ROESY spectrum of the penta peptide in DMSO d_6: α protons to aromatic protons, NH to aliphatic protons and β protons to aromatic protons.

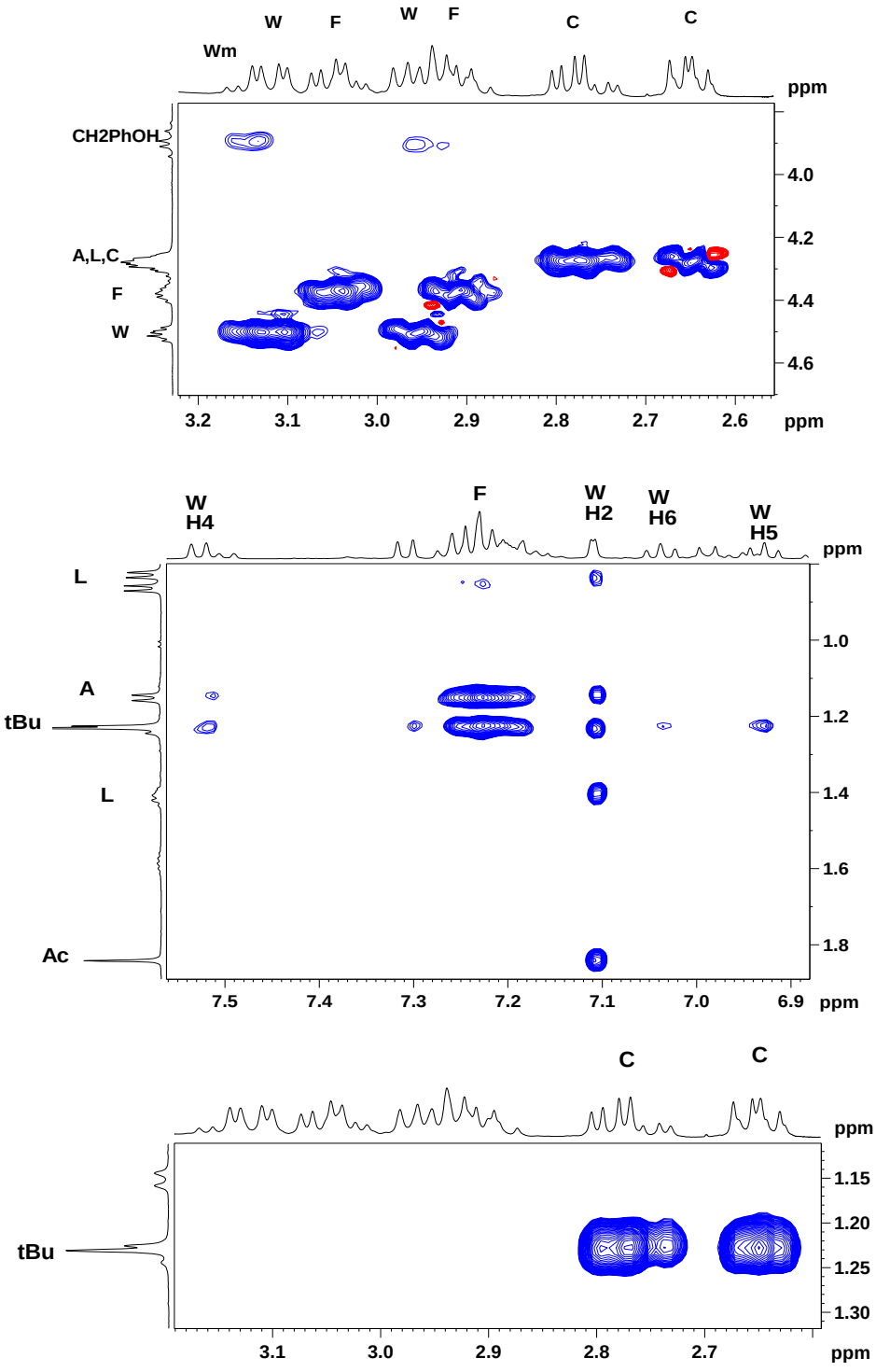


Fig-13F: 500 MHz ROESY spectrum of the penta peptide in DMSO d_6 : β protons to α protons, aliphatic protons to aromatic protons and tBu protons β protons of Cys.

^{1}H-^{15}N HSQC measurements:

The ^{15}N HSQC experiment showed six distinct NH protons signals. One of them, the most deshielded signal (-248.74 ppm) is assigned to the indole NH of the Trp residue and the other five are originating from the five peptide NH of the penta peptide. The chemical shift extracted from the HSQC spectrum is given in Table-1.

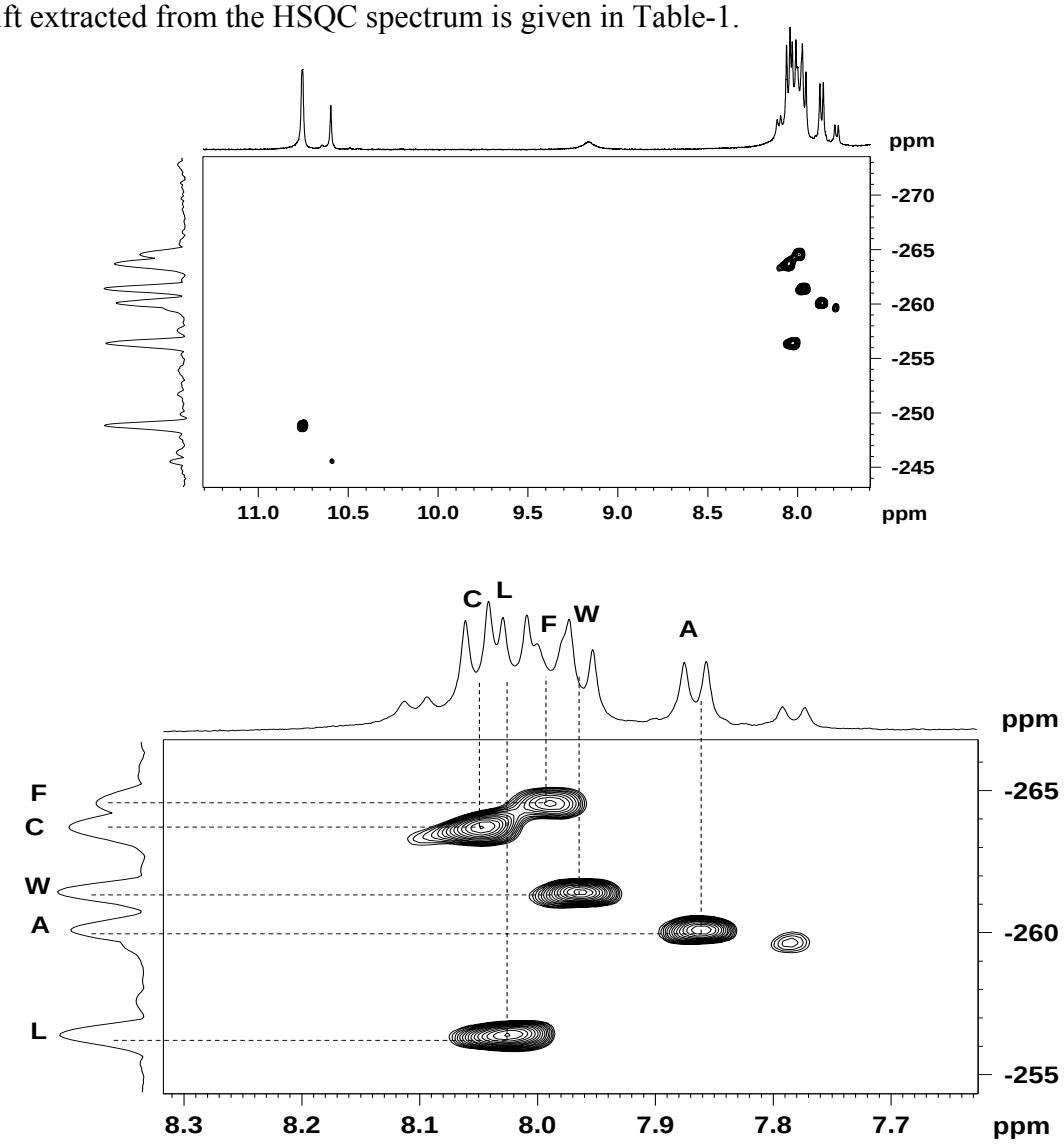


Fig- 14: 56 MHz 1H-15N HSQC spectrum of the penta peptide in DMSO d_6.

Table 1: NMR data for CPI-2081 500.13 MHz (^{1}H), 125.75MHz (^{13}C), 40.56 MHz (^{15}N) in DMSO-d_6:

	CPI-2081 (major)				CPI-2081 (Minor)		
Amino acid	Position	δH(int; mult; J(Hz)	δC	15 N	δH(int; mult; J(Hz)	δC	^{15} N
Leu	α	4.28 (1H, m)	51.65		4.27 (1H, m)	51.65	
	β	1.41(2H, m)	41.27		1.41 (2H, m)	41.27	
	γ	1.58(1H, sep, 6.72)	24.61		1.58 (1H,sept, 6.72)	24.61	
	δ	0.86(d), (3H, d, 6.72)	23.48		0.86 (3H,d,6.62)	23.48	
		0.83(d) (3H, d, 6.72)	22.08		0.83 (3H,d,6.72)		
	C=O	-	172.74		-	22.08	
	NH	8.03(1H, d, 8.05)	-	-256.25	8.03 (1H,d, 8.05)	172.77	-256.25
Acetyl	OAc	1.84 (3H, s)	22.96		1.84 (3H, s)		
	C=O	-	169.88		-	22.96	
						169.91	
Cys	α	4.28 (1H, m)	53.77		4.28 (1H,m)	53.87	
	β	2.78(1H, dd,5.26, 12.72)	30.30		2.75 (1H,dd,5.08,12.72)	30.24	
		2.65(1H, dd 8.84, 12.72)			2.64 (H, dd,8.64,12.72		
	C=O	-	170.43		-	170.21	
	NH	8.06 (1H, d, 7.85)	-	-263.68	8.11 (1H, d, 7.79)	-	-263.29
t-Bu	Me	1.24 (9H, s)	31.1		1.22 (9H, s)	31.1	
	CtBu	-	42.42		-	42.42	
Trp	α	4.51 (1H, m)	53.78		4.51 (1H, m)	54.44	
	β	3.12(1H, dd, 4.92,14.86)	27.86		3.15 (1H,dd,6.37,14.82)	27.36	
		2.96 (1H, dd,8.10,14.86)			2.943 (1H,dd)		
	NH(Indole)	10.76 (1H, Br, 1.8)	-	-248.91	10.61 (1H,s)	-	-245.6
	2	7.108 (1H, d, 2.12)	124.03		7.108 (1H,m)	136.99	
	3	-	110.31		-	106.55	
	3a	-	127.91		-	128.33	
	4	7.527 (1H, d, 7.88)	118.80		7.495 (1H,d,7.88)	118.80	
	5	6.926 (1H, t, 7.44)	118.62		6.869 (1H,t, 7.6)	118.52	
	6	7.033 (1H, d,7.45)	121.23		6.95(1H,d, 7.5)	120.50	
	7	7.305 (1H, d, 8.1)	111.67		7.20(1H,d)	111.01	
	7a	-	136.53		-	135.99	
	8	-	-		3.90(2H,dd)	31.1	
	9	-	-		-	130.38	
	10	-	-		6.659 (1H,d,8.16)	115.60	
	11	-	-		6.98 (1H,d,8.30)	129.69	
	12	-	-		-	156.06	
	13	-	-		6.98 (1H,d,8.30)		
	14	-	170.89		6.659 (1H,d,8.30)	129.69	
	C=O				-		
						115.60	
	NH	7.97 (1H, d, 8.00)		-261.38	7.97		-261.38
						170.47	
Ala	α	4.28 (1H, m)	48.63		4.30 (1H, m)	48.63	
	β	1.15 (3H, d, 7.26)	18.69		1.15 (3H, d, 7.26)	18.69	
	C=O	-	172.20		-	172.07	
	NH	7.88 (1H, d, 7.44)		-260.05	7.80 (1H, d,7.44)		-259.52
Phe	α	4.378 (1H, m)	54.13		4.36 (1H, m)	54.13	
	β	3.052(1H, dd, 5.38,3.82)	37.17		3.03 (1H,dd,5.48,13.82)	37.17	
		2.93(1H,dd,8.28,13.82))	138.05		2.895(1H,dd,8.64,13.82)		
	1	-	129.63		-	138.05	
	2	7.23 (1H, d)	128.57		7.23 (2H, d)	129.63	
	3	7.25 (1H, t)	126.80		7.25 (2H,t)	128.57	
	4	7.18(1H, t)	128.57		7.23(1H)	126.80	
	5	7.25 (1H, m)	129.63		7.25 (2H,t)	128.57	
	6	7.22 (1H, m)			7.22 (2H,d)	129.63	
	NH	8.00 (1H, d, 8.17)	-	-264.38	8.05 (1H, d, 7.72)	-	-264.38
	COOH	12.72(1H, Br)	173.14		12.72(1H,Br)	173.14	

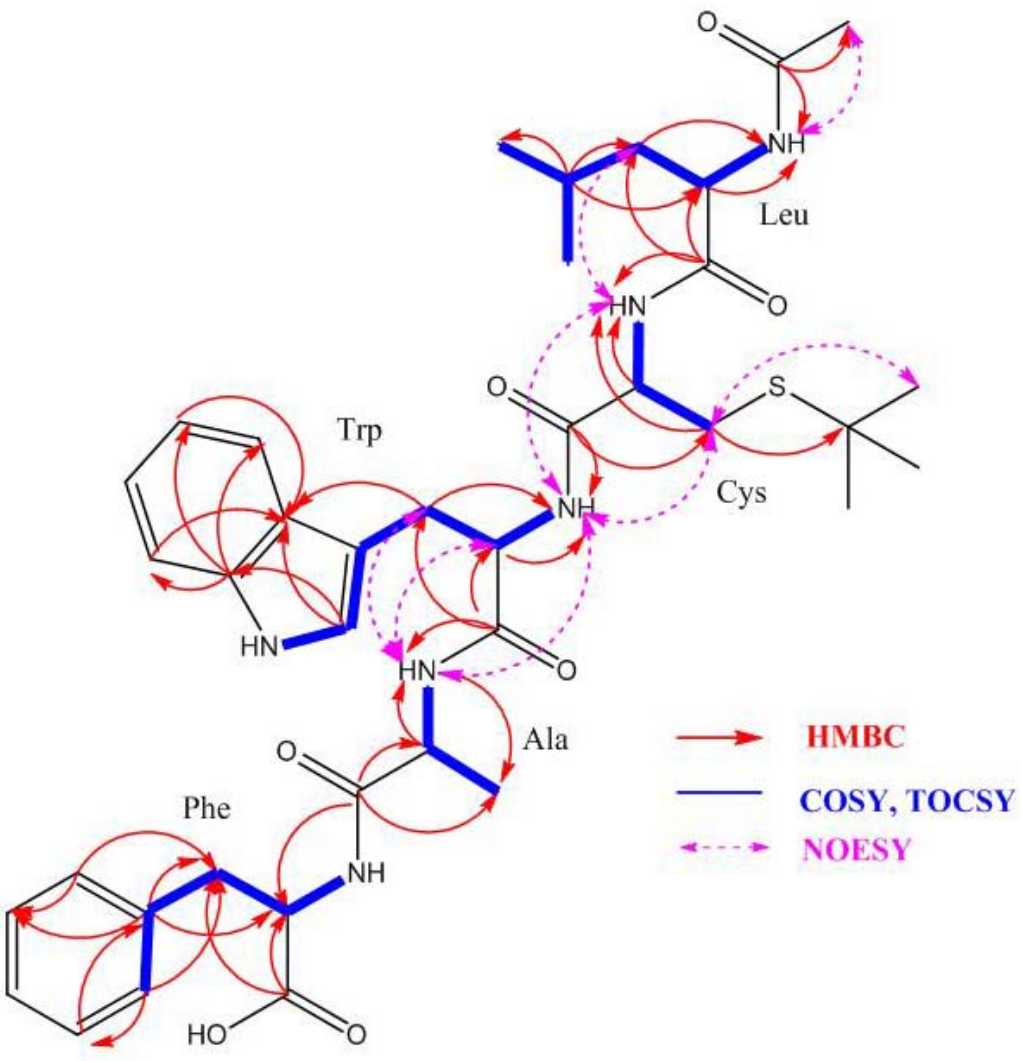
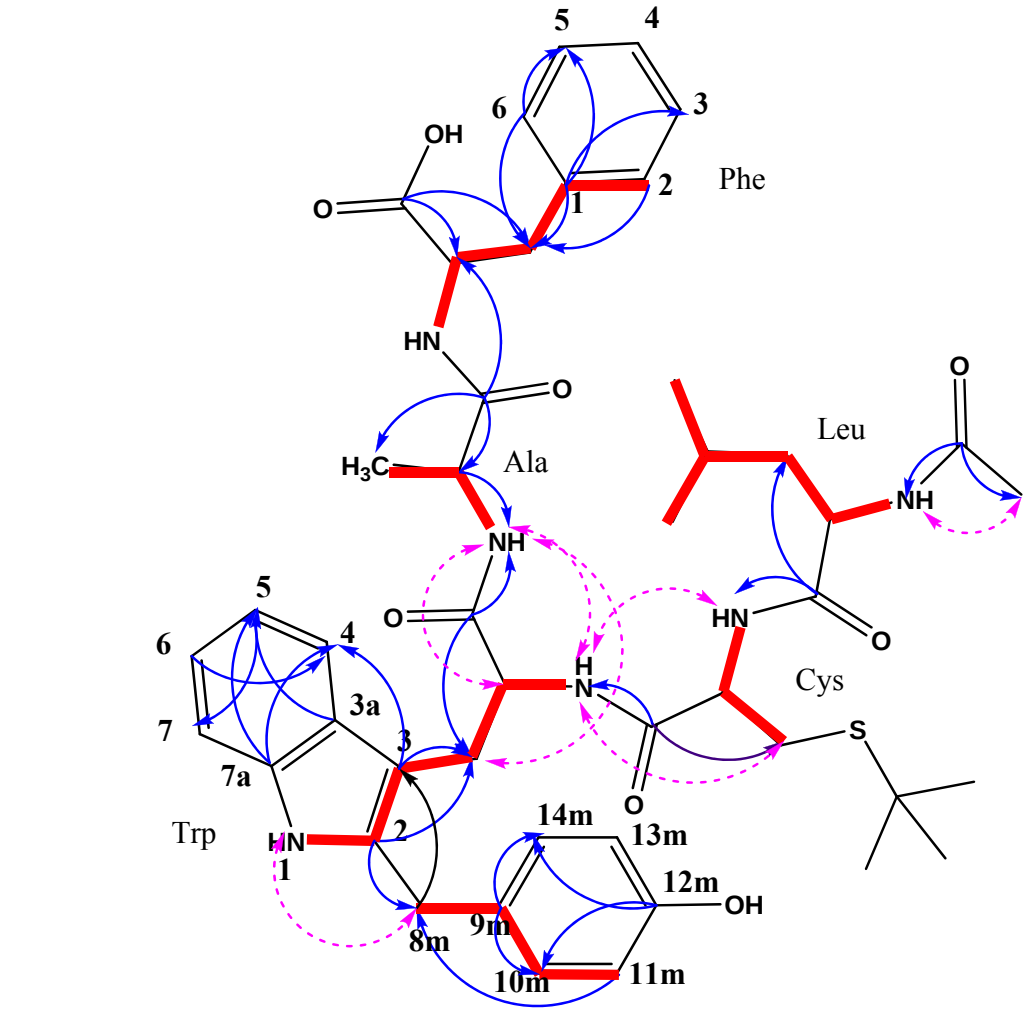


Fig-15A: Selected HMBC, COSY, TCOSY and NOSY correlations of CPI-2081 (Major)



^{1}H-^{1}H COSY, TOCSY

^{1}H-^{1}H NOESY

^{1}H-^{13}C HMBC

Fig- 15B: Selected HMBC, COSY, TCOSY and NOSY correlations of CPI-2081 (Minor)

DISCUSSION:

We have described the isolation and structure elucidation of a novel competitive inhibitor of Thiol protease (CPI-2081), a mixture of two very similar but distinct modified peptides in approximate molar proportion of 35: 65. Four out of five amino acids both peptides (major and minor) are common, therefore, they exhibit very similar physiochemical properties hence, it was not possible to isolate separately even with the help of analytical reverse phase HPLC using various solvent conditions. But, It is possible to validate the structure if one can synthesize all stereoisomer of the both the compounds and selecting the peptide having Thiol proteolytic inhibition activity. The structure elucidated with the help of UV-Vis, FT-IR, MS-MS and extensive NMR spectroscopic techniques contains unique groups and modifications such as acetylation of N-terminal leucine and tertiary butylation of cysteine residue. N-terminal acetylating and complete cyclisation of bioactive peptide confers the extra protection from endogenous peptidases. Compounds, bearing *t*-butyl group is very rare in nature which are mostly isolated from marine sponges, include peptides, terpenes, carbinols, esters and ketone [35, 36, 37, 38]. A very unusual *tert*-butyl ketone coumarin swietenone was isolated from *Chloroxylon swietenia,* a tree found in India[39].

K. Nakanishi and K. Habagushi have described the mevalonate pathway for the biosynthetic origin of *t*-butyl group in nature[40]. In their scheme, the *t*-butyl group arises by cleavage of the C–C bond adjacent to a *gem*-dimethyl unit followed by methylation elicited by *S*-adenosylmethionine. However, the recent investigations have demonstrated that mevalonate pathway is only a minor route to the *t*-butyl biosynthesis and a non mevalonate pathways are also responsible for its biosynthesis[41, 42]. The compounds bearing *t*-butyl are very stearically protected which can either hinder the chemienzymatic process or nature can use these groups to attend sterioselectivity during the reaction. Therefore, the CPI-2081 bearing this group may prove to be very sterioselective in addition to increasing the extent to hydrophobic interaction during chemienzymatic reaction and due to *N*-terminus modification, it will also be resistant to enzymatic degradation *in-vivo*.

CONCLUSION:

In conclusion, we have been able to purify a pentapeptide which acts as an inhibitor of cysteine protease, Papain, with a very low k_i (<100 nM). This pentapeptide is uniquely modified such as N-terminal acetylation of leucine residue and t-butylation of cysteine residue. Therefore, this peptide cysteine protease inhibitor should have distinct biological functions.

REFERENCES:

1. Umezawa, H. (1982) *Ann. Rev. Microbiol.* 36:75-99.
2. Lam, K. S. (2007) *TRENDS in Microbiol.* 15, 279-289.
3. Clardy, J. and Walsh, C. (2004) *Nature,* 432, 829–837.
4. Be´rdy, J. (2005) *J. Antibiot. (Tokyo),* 58, 1–26.
5. Lee, M.L. and Schneider, G. (2001) *J. Comb. Chem.* 3, 284–289.
6. Bajorath, J. (2002) *J. Comput. Aided Mol. Des.* 16, 431–439.
7. Feher, M. and Schmidt, J.M. (2003) J. Chem. Inf. Comput. Sci. 43, 218-227.
8. Bibb, M.J. (2005) *Curr. Opin. Microbiol.* 8, 208–215.
9. Cundliffe, E. (2006) *J. Ind. Microbiol. Biotechnol.* 33, 500–506.
10. Gullo, V.P. (2006) *J. Ind. Microbiol. Biotechnol.* 33, 523–531.
11. Kiefer, W. and Dannhardt, G. (2004) *Curr. Med. Chem.* 11, 3147–3161.
12. Chapman, H. A.; Riese, R. J.; Shi, G. P. (1997) *Annu Rev Physiol*.; 59, 63-88.
13. Otto, H. H. and Schirmeister, T. (1997) *Chem Rev*. 97, 133-17
14. Mc Grath, ME. (1999) *Annu. Rev. biophys. Biomol. Struct*. 28, 181-204.
15. Greenbaum, D. C.; Baruch, A.; Grainger, M.; Bozdech, Z.; Medzihradszky, K. F.; Engel, J. DeRisi, J. Holder, A. A. and Bogyo, M. (2002) *Science.* 298, 2002-6.
16. Selzer, P. M.; Pingel, S.; Hsieh, I.; Ugele, B.; Chan, V. J.; Engel, J. C.; Bogyo, M.; Russell, D. G.; Sakanari, J. A. and McKerrow, J. H. (1999) *Proc Natl Acad Sci U S A*. 96, 11015-22.
17. Umezawa H. (1982) *Annu Rev Microbiol.* 36, 75-99.
18. Dubin, G. (2005) *Cellular and Molecular Life Sciences* 62, 653–669.
19. Sakano, Y.; Shibuya, M.; Matsumoto, A.; Takahashi, Y.; Tomoda, H.; Omura, S. and Ebizuka, Y. (2003) *J Antibiot (Tokyo).* 56, 817-826.
20. Tani, M.; Harimaya, K.; Gyobu, Y.; Sasaki, T.; Takenouchi, O.; Kawamura, T.; Kamimura, T. and Harada, T. *(*2004) *J Antibiot (Tokyo).* 57, 89-96.
21. Tani, M.; Gyobu, Y.; Sasaki, T.; Takenouchi, O.; Kawamura, T.; Kamimura, T. and Harada, T. (2004) *J Antibiot (Tokyo).* 57, 83-88.

22. Singh, M. P.; Petersen, P. J.; Weiss, W. J.; Janso, J. E.; Luckman, S. W.; Lenoy, E. B.; Bradford, P. A.; Testa, R. T. and Greenstein, M. (2003) *Antimicrob Agents Chemother.* 47, 62-69.
23. Suda, H.; Aoyagi, T.; Hamada, M.; Takeuchi, T. and Umezawa, H. (1972) *J Antibiot (Tokyo)*. 25, 263-6.
24. Aoyagi, T.; Takeuchi, T.; Matsuzaki, A.; Kawamura, K. and Kondo, S. (1969) *J. Antibiot (Tokyo)* .22, 283-6
25. Kadowaki, T.; Kitano, S.; Baba, A.; Takii, R.; Hashimoto, M.; Katunuma, N. and Yamamoto, K. (2003) *Biol Chem.*, 384, 911-20.
26. Nakae, K.; Yoshimoto, Y.; Sawa, T.; Homma, Y.; Hamada, M.; Takeuchi, T.; Imoto, M.. (2000) *J. Antibiot (Tokyo).* 55, 442-444.
27. Aoyagi, T.; Takeuchi,T.; Matsuzaki, A.; Kawamura, K., Kondo, S.; Hamada, M.; Maeda, K.; and Umezawa, H. (1969) *J. Antibiotics Tokyo. 22, 283-286.*
28. Hanada, K.; Tamai, M. and Yamagishi, M. (1978)*Agricultural and Biological Chemistry*., 42, 523-528.
29. Suda, H.; Aoyagi, T.; Hamada, M.; Takeuchi, T. and Umezawa, H. *(*1972)*J Antibiot (Tokyo).* 25, 263-266.
30. Takahashi, S.; Murakami, K. and Miyake, Y. 1981. *J. Biochem.* 90, 1677-1684
31. Reddy, M. V. R.; Harper, M. K. and Faulkner, D. J. 1998.*Tetrahedron,* 54, 10649-10656.
32. Ezra, D.; Castillo, U. F.; Strobel, G. A.; Hess, W. M.; Porter, H.; Jensen, J. B. Condron, M. A. M.; Teplow, D. B.; Sears, J.; Maranta, M.; Hunter, M.; Weber, B. and Debbie Yaver, D. (2004) *Microbiology*, 150, 785–793.
33. Trusheva, B.; Popova, M.; Bankova, V.; Simova, S.; Marcucci, C. M.; Miorin, P.L.; Pasin, F. R. and Tsvetkova, I. (2006). *eCAM*, 3, 249-254.
34. Alobulifera, S.; livinastonei, G.; Ovalifolia, G. and Rosea, C. 1992. *Tetrahedron*, 48, 10093-10102.
35. Chevallier, C.; Richardson, A. D.; Edler, M. C.; Hamel, E.; Harper, M. K. and Ireland, C. M. (2003)*Org. Lett.*, 5, 3737.
36. Hamel, E. *Biopolymers* (2002) 56, 142.

37. Matsunaga, S.; Fusetani, N.; Konosu, S. (1985) *Tetrahedron Lett.* 26, 855.

38. Ryu, G.; Matsunaga, S. and Fusetani, N. (1994) *Tetrahedron Lett.*, 35, 8251.

39. Mujumdar, R. B.; Rao, A. V. R.; Rathi, S. S. and Venkataraman, K. (1975) *Tetrahedron Lett.*, 16, 867.

40. Nakanishi, K.; Habaguchi, K.; Nakadaira, Y.; Woods, M. C.; Maruyama, M.; Major, R. T.; Alauddin, M.; Patel, A. R.; Weinges K. and ahr, W. B¨ (1971) *J. Am. Chem. Soc.*, , 93, 3544.

41. Rohdich, F.; Kis, K.; Bacher, A. and Eisenreich, W. *Curr. Opin. Chem. Biol.*, 2001, 5, 535.

42. Stromgaard, K. and K. Nakanishi, K. *Angew. Chem., Int. Ed.*, 2004, 43, 1640.

CHAPTER 3

Thermodynamics of binding of papain like cysteine proteases and CPI

SUMMARY:

The attempt made in activity based screening of *Streptomyces sp* culture broth lead us to the isolation of a cysteine protease inhibitor (CPI-2081). Its complete structure was elucidated as described in previous chapter which was found to be a uniquely modified penta peptide. The CPI-2081 from *Streptomyces sp.* NCIM2081 displays both the inhibition constant (K_i) and IC_{50} for inhibition of Papain (putative Thiol protease) of 49.14± 2.45 nM and 36.9 ±1.8 nM, respectively. In this chapter, the effect of CPI-2081 on substrate hydrolytic activity of other prominent cysteine (thiol) proteases, Cathepsin-L and Cathepsin-K have been described. The initial kinetic studies revealed that the CPI-2081 inhibits both the cathepsins via competitive mode of inhibition. The equilibrium for association constant (K_i) was found to be 61.2 nM for Cathepsin-L and 55.8 nM for Cathepsin-K at 298 K. The thermodynamics of association of CPI-2081 with both of these Cathepsins was carried out independently, which indicates that, predominantly, the entropy contributes to the free energy of reaction spontaneity for association of CPI-2081 with Cathepsin-L and Cathepsin-K.

INTRODUCTION:

Structure of Active site of cysteine proteases: Cysteine proteases signify a large family of enzymes that utilize an active-site cysteine to hydrolyze amide bonds in proteins or peptides[1, 2]. Although, a key cysteine thiol group is utilized as the nucleophile during proteolysis, the active site has many features in common with that of serine proteases. Likewise, the transition state for cysteine proteases involves formation of a tetrahedral intermediate. However, in this case the sulfur atom of Cys adds to the amide carbonyl carbon to give an oxido-orthothioamide, which interacts with an oxyanion hole[1, 2]**.** Again, there are multiple hydrogen- bonding interactions between the amide backbone of the substrate and the amide backbone of the enzyme. The catalytic site of papain-like cysteine proteases is highly conserved and defined by a Cys-His-Asn catalytic triad (Cys-25, His-159, Asn-175)[3]**.** The nucleophilic Cys-c-S is apparently ionized prior to substrate binding, unlike the Ser- c-O in serine proteases[3]**.** Since papain-like cysteine proteases have diverse roles in physiology and pathology, there is opportunity for inhibitors to serve as drugs for diseases such as osteoporosis, arthritis, asthma, atherosclerosis, cancer, and parasitic infection[3]**.** From a drug discovery point of view, efforts have been aimed at inhibitors of cathepsins B, K, L, and S.[3,4] Compounds that target cathepsins S and K have entered clinical development, and the cathepsin K inhibitor balicatib (AAE581) has advanced into phase two studies [4, 5]**.**

Role of solvent in enzyme inhibitor reaction kinetics: It is considered that standard enthalpy is a quantitative indicator of the changes in intermolecular bond energies, such as hydrogen bonding and Van der Waals interactions, occurring during the binding. In addition standard entropy can be considered as an indicator of the rearrangements undergone by the solvent, usually water molecules, during the ligand binding[6]. The transfer of solvent molecules between the bulk of the solvent and the solvation shells of solutes is termed as 'solvent reorganization'. It is established that solvent reorganization and the change in solvent reorganization that may accompany reactions and spectral transitions, have thermodynamic consequences, especially in hydrogen-bonding solvents such as aqueous buffers, prevalent in biochemical reactions. As a result of this, there is enthalpy-entropy compensation to make the overall standard change in free energy (ΔG) negative and keep the reaction spontaneous. Although the change in enthalpy is

measured directly by calorimetric techniques, it is frequently convenient to compare with indirect estimates using classical Vant Hoff equations[7], arising from temperature dependence of equilibrium constant as we might get using circular dichroism, fluorescence and other indirect techniques. The affinity of inhibitor is measured as the free energy difference between the free, solvated ligand and the free, solvated active site on the one hand and the complex on the other hand. The changes in affinity will reflect not only changes in ligand-active site interactions but also changes in the solution conformation of the ligand[8]**.** The high-quality thiol protease inhibitor should have lower inhibition constant, hence higher affinity.

As stated above the effectiveness of the inhibitor depends on structural features of inhibitor as well as that of enzyme. But it should not be forgotten that the nature of solvent too affects the extent of association and dissociation between inhibitor and enzyme. In this chapter, the general features of CPI-2081 have been described such as inhibition constants of CPI-2081 for thiol proteases such as cathepsin-L and cathepsin K. The thermodynamics of interaction of CPI-2081 with Thiol proteases has also been described for reactions happening in aqueous buffers, which is the sole solvent for general biochemical reactions.

MATERIALS AND METHODS:

Reagents: MES buffer components, 2–(N–morpholino) ethanesulphonic acid (MES), EDTA, DMSO and DTT were purchased from Sigma Aldrich, USA. Enzymes, Recombinant human cathepsin-K and cathepsin-L were purchased from EMD Calbiochem. The fluogenic substrate Cbz-Phe-Arg-AMC (extinction co-efficient= 22000 M^{-1} cm^{-1}) was also obtained from EMD Calbiochem.

Cathepsin K activity inhibition assay and determination of K_i: Method of McGrath et *al*. was followed for measurements of Cathepsin K activities using Z-Phe-Arg-AMC as substrate [9, 10]. Human Cathepsin K and human Cathepsin L inhibition studies were performed in 50 mM MES (pH 5.5), 2.5 mM EDTA, 2.5 mM DTT, and 10% DMSO. The substrate used to monitor Cathepsin K and L activity was Cbz-Phe-Arg-AMC (extinction co-efficient= 22000 M^{-1} cm^{-1}). In both cases, the substrate concentration was near K_m

value fixed at 40 μM for Cathepsin K and 10 μM for Cathepsin L. To calculate the $K_{i,}$ enzyme was incubated with inhibitor, present at varying concentrations, for 30 min at three different temperatures (298, 310 and 315 K) in 96-well microtiter plates to allow for equilibrium to be established for association of enzymes and inhibitors. After preincubation, reactions were initiated with the addition of the fluorogenic substrate specified previously. The hydrolysis of this substrate yields AMC, which was monitored fluorometrically. The amount of 7-amino-4-methylcoumarin (AMC) liberated from the substrate was monitored fluorometrically with excitation at 355 nm and emission at 460 nm in a spectrofluorimeter. Type of inhibition and K_i was determined using graphical method of Dixon Plot[11, 12] in which the effect on the enzymatic rate of reaction was determined for four different substrate concentrations, and over a range of inhibitor concentrations (I). In a plot of reciprocal relative fluorescent unit, (1/RFU) against [CPI-2081], data for each substrate concentration fall on straight lines that intersect at [CPI-2081] = $-K_i$.

Thermodynamic data determination: Free energy changes of inhibition of cathepsin K against CPI-2081 (ΔG) were determined by the equation,

$$\Delta G = -RT \ln K_a \qquad (1)$$

Temperature dependence of the inhibition constants was used to determine the thermodynamic parameters. Changes in enthalpy (ΔH) were determined from the Van't Hoff plots by using the equation,

$$\ln K_a = -(\Delta H/RT) - \Delta S/R \qquad (2)$$

Where ΔH, R (8.314 J/K mol), ΔS and T are change in enthalpy, gas constant, change in entropy and the absolute temperature respectively. The standard free energy was calculated as $\Delta G° = -RT \ln K_a$ at 298.15 K, K_a is obtained as, $1/K_{i\,,}$ the standard enthalpy, $\Delta H°$, from the Van't Hoff plot, ln K_a versus $(1/T)$ (the slope of which is $-\Delta H°/R$) and the entropy as $\Delta S° = (\Delta H° - \Delta G°)/ T$ with T and R (8.314 JK^{-1} mol) [6, 13, 14].

RESULTS AND DISCUSSION:

Initial kinetics and determination of K_i for cathepsin-K and cathepsin-L:

The inhibition constant of inhibitor for an enzyme (K_i) signifies the high affinity for enzyme. Since, Cathepsin-K and Cathepsin-L are crucial thiol proteases involved in bone remodeling, any disturbance in their regulation results in pathological consequences such as osteoporosis and arthritis. Therefore we decided to analyze the initial kinetics of substrate hydrolysis inhibition by CPI-2081. The percent of substrate hydrolyed decreased as a function of concentration of CPI-2081 in dose dependent manner displaying the K_i value of 68.2 ± 3.4nM against cathepsin-L (figure 1A)and 55.83 ± 2.79 nM against cathepsin-K (figure 1B). The inhibitory constants for CPI-2081 against Cathepsin (K_i) was calculated with the help of Dixon plot using ORIGIN 6.1 software, where inverse of rate of release of fluorescent AMC was plotted against a range of concentration of CPI-2081 (0- 78 nM) at 20μM, 40 μM , 60 μM and 80 μM CBZ-Phe-Arg-AMC at 298 K temperature and 0.08nM enzyme concentration. Furthermore, the double reciprocal L. B. Plot was also constructed in which the inverse rate of release of fluorescent AMC as a result of protease action was plotted against inverse concentration of inhibitor (CPI-2081). The double reciprocal L. B. demonstrated the competitive mode of inhibition of both the cathepsin-L and cathepsin-K by CPI-2081.

Thermodynamics of association of CPI-2081 and Cathepsin K and Cathepsin-L: The van't Hoff plots show that the effect of temperature on the association constants, K_A, which is inverse of dissociation constant K_D (also called inhibition constant, K_i) appears to be essentially linear in the range 298 K to 310 K for CPI-2081 examined with human cathepsin-K and cathepsin-L (Figure 2A). Slopes of van't Hoff plots are negative for CPI-2081 in both the cases cathepsin-L and cathepsin-K, showing that the affinities improve with the increase of the temperature (Table 1 & 2). Since, in many cases the enzyme activity decreases with increase in temperature, we also carried out temperature dependent cathepsin-L and cathepsin-K activity experiments (Figure 2B) which show that there is no effect of temperature on enzyme activity for 15 minutes which is the time duration of enzyme activities in all thermodynamic investigations in this study. Final thermodynamic parameters, ΔG, ΔH and ΔS (expressed as mean values ± standard error of three independent experiments), were calculated for the binding equilibria of the

different CPI-2081 and cysteine proteases Cathepsin-L and Cathepsin-K as described in materials and methods. There is a negative change in Gibbs free energy in binding of CPI-2081 to both recombinants human Cathepsin-L and Cathepsin-k; therefore, the reaction is thermodynamically spontaneous. Binding of CPI-2081 to both the enzymes Cathepsin-K & L is endothermic and change in enthalpy is positive (tables-1 &2). Whereas, the change in entropy in both cases are positive, which make TΔS larger as compared to ΔH (tables-1 &2). Therefore, the energy compensation for the spontaneity of reaction is provided by large positive change in entropy, hence the binding of CPI-2081 to Cathepsin-K is entropicaly driven (table 1). Similarly, in case of Cathepsin-L the large positive change in entropy drives the binding reaction of CPI-2081 to enzyme (table 1).

Discussion:

Thermodynamic parameters have been collected for a newly characterized CPI-2081 against Cathepsin-K and Cathepsin-L. The information provided by these data could be useful from a pharmacological point of view to discover new thermodynamic relationships related to drug–receptor interactions and their molecular mechanisms[15, 16, 17]. In the last few years, it has been reported that change in enthalpy (Δ*H*) and entropy (Δ*S*) values of drug interaction with a defined receptor can often give a simple "*in vitro*" way to differentiate the ability of the drug to interfere with the signal transduction pathways[6, 18, 19, 20, 21]. The thermodynamic study of favorability has been evaluated for various protease inhibitors[22, 23]. All these data suggest an intercorrelation between specific binding and the variation of water molecules present to receptor surfaces. In addition, based on the thermodynamic compensation a general model of drug–receptor interaction has been proposed. In this model the solvent molecules do not modify the values of the affinity constant of the drug–receptor interaction because the standard free energy for solvent reorganization can be near to zero and the values of binding parameters are due to specific features of the ligand and receptor in the binding process and not by the solvent[24]. On the other hand, Δ*H*° and −*T*Δ*S*° values are related to the rearrangements occurring during the binding, in the solvent–drug and solvent–receptor interfaces[24]. It seems reasonable to assume that solvent effects might be responsible for the *in vitro* thermodynamic differentiation between various kinds of drugs against a

target. From this background, it can be said that the presence of the linearity of Van't Hoff plots for CPI-2081 and Cathepsin under this study is a significant observation. Similar to what is observed for other protease inhibitors showing that $\Delta H°$ (standard enthalpy) and $\Delta S°$ (standard entropy) values are independent of temperature and obtained by linear van't Hoff plots[6, 18, 19, 20]. Van't Hoff plot turns out to be linear for CPI-2081 considered in the present study, which means that the value of $\Delta H°$ is not significantly affected by temperature variation in the range investigated. The binding of CPI-2081 with both the cathepsin-L and cathepsin-K is entropicaly driven, which indicates that the displacement of water molecules from aromatic rings of CPI-2081 as well as the enzymes generate this heat of entropy. This idea is strengthened by the study of Ravishanker *et al.* that the potential energy change on pair formation between aromatic rings and other hydrophobic interaction in the presence of solvent makes entropic change resulting from displacement of water molecule positive[25, 26]**.** Therefore it is evident that the binding of CPI-2081 with target enzymes, cathepsin-L& K, is made favorable by positive change in entropy, or by the energy released while rearrangement of the solvent (H_2O) molecules between CPI-2081-target enzyme and conformational change due to rearrangements of various hydrophobic interactions during binding process.

Conclusion:

In conclusion, the thermodynamic results described the mode of inhibition of CPI-2081 and Cathepsin-L and K are entropy driven and this could be due to the unique modification of the penta peptide such as t-butylation of cystein and N-terminus acetylation of leucine moity.

FIGURES AND TABLES:

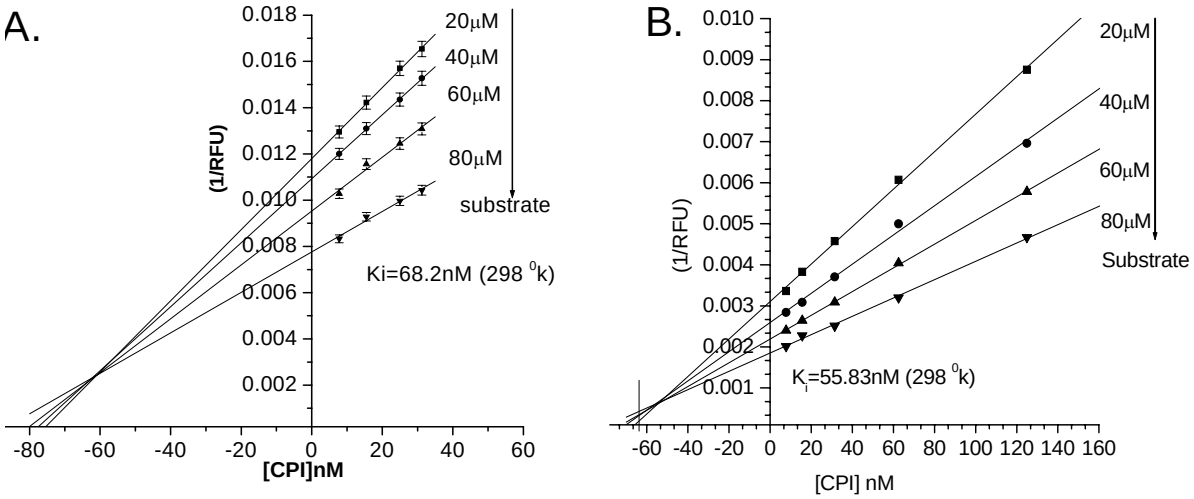


Figure1: Dixon plot for determination of K_i for CPI-2081 as a cathepsin-K inhibitor: *Dixon plot of reciprocal rates of formation of fluorescent AMC, measured as relative fluorescent unit (1/RFU), as a result of substrate hydrolysis. Each line represents linear regression analysis of reciprocal of average rates of substrate hydrolysis for different substrate concentrations as a function of inhibitor concentration [CPI-2081] in nanomolar (nm). (A) Cathepsin –L, 0.08nM, pH=5.5, T=298 ^{K} and (B) Cathepsin-K, 0.08nM, pH=5.5, T=298 K.*

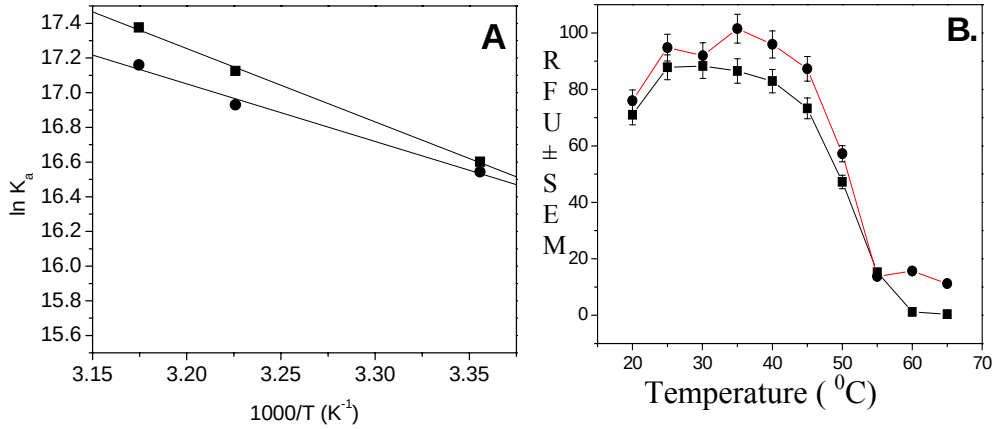


Figure 2: Temperature dependence of K_i as shown by liner fit of Van't Hoff plot: A. *Ligand (CPI-2081) association to thiol proteases, Cathepsin L (■), n=3, slope (m) = - (4.222± 0.189) and r=0.9995; Cathepsin K (●), n=3, slope (m) = - (3.319± 0.33), r=0.9949. Natural logarithm of association constant, ln/K_a), is plotted against 1000/T ($^{K-1}$).* ***B.*** *Effect of temperature on enzyme activity of 0.8 nM, Cathepsin-L (●) and Cathepsin-K (■). The relative fluorescent unit ± Standard Error of Mean (SEM) of three independent experiments is plotted against temperature in degree Celsius.*

Table 1: Inhibition constant and thermodynamic parameter$^{\Psi}$ for inhibition of Cathepsin L by CPI-2081 at different temperatures:

Temperature (^{0} K)	*K_i (nM)*	*ΔG (KJ/mole)*	*Δ H* (KJ/mole)*	*Δ S (J/mole/K)*
298	61.2 ± 2.04	-(40.88± 2.04)	35.10 ±1.089	255.81 ± 4.23
310	42.17± 2.18	-(43.76 ± 2.08)	35.10 ±1.089	254.39 ± 4.01
315	25.90± 2.28	-(45.74 ± 2.28)	35.10 ±1.089	256.70 ± 4.31

Table 2: Inhibition constant and thermodynamic parameter$^{\Psi}$ for inhibition of Cathepsin K by CPI-2081 at different temperatures:

Temperature (0 K)	*Ki (nM)*	*Δ G (KJ/mole)*	*Δ H *(KJ/mole)*	*Δ S (J/mole/K)*
298	55.83±2.79	-(41.37 ± 1.21)	27.59 ± 0.97	174.00 ± 3.89
310	49.57±2.47	-(43.35 ±1.27)	27.59 ± 0.97	171.40 ± 3.98
315	34.35±1.71	-(45.01 ± 1.34)	27.59 ± 0.97	173.02 ± 3.18

$^{\Psi}$All Values expressed in table 1 and table 2, as ± Standard Error

*Values expressed at 310 ^{o}K.

REFERENCES:

1. Powers, J. C.; Asgian, J. L.; Ekici, O. D. and James, K. E. (2002) *Chem. Rev.* 102, 4639.
2. Otto, H.-H. and Schirmeister, T. (1997)*Chem. Rev.* 97, 133.
3. Lecaille, F.; Kaleta, J. and Bro¨mme, D. (2002) *Chem. Rev.* 102, 4459.
4. Chung, J.-Y.; Bae, Y.-A.; Na, B.-K. and Kong, Y. (2005) *Expert Opin. Ther. Pat.*15, 995.
5. Vasiljeva, O.; Reinheckel, T.; Peters, C.; Turk, D.; Turk, V. and Turk, B. (2007)*Curr. Pharm. Des.*13, 387.
6. Borea, P. A.; Dalpiaz, A.; Varani, K.; Gilli, P.; Gilli, G. (2000) *Biochem Pharmacol* 60**,** 1549–1556.
7. Varani, K.; Surprenant, A.; Vincenzi, F.; Tosi, A.; Gessi, S.; Merighi, S. and Borea, P. A. (2008) B*iochemical Pharmacol.* 7 5, 1198 – 1208.
8. Epps, D. E.; Cheney, J.; . Schostarez, H.; Sawyer, T. K.; Prairie, M.; Krueger, W. C. and Mandell, F. (1990) *Journal of Medicinal Chemistry* 33**,** 2080-2086.
9. McQueney, M. S.; Feild, J.; Hanning, C. R.; Brun, K.; Ramachandran, C.; Connor, J.; Drake, F.; Jones, C. S. and Bernard, Y. (1998) *Protein Expression and Purifi.* 14, 387-394.
10. McGrath, M. E.; Sprengeler, P.A.; Hill, C. M.; Martichonok, V.; Cheung, H. Somoza, J. R.; Palmer, J. T. and Janc, J. W. (2003) *Biochemistry* 42, 15018-15028.
11. Dixon. M.(1972) *Biochem J*. 9,197-202.
12. Sapre, M. P.; Jha, H.; Patil, M. B. (2005) *J Gen Appl Microbiol.* 51, 327-34.
13. Borea, P.A.; Varani, K.; Gessi, S.; Merighi, S.; Dalpiaz, A.and Gilli, P.(2004) *Curr Top in Med Chem.* 4, 361–368.
14. Borea, P.A.; Dalpiaz, A.; Varani, K.; Guerra, L. and Gilli, G.(1995) *Biochem Pharmacol* 49, 461–469.
15. Gilli, P.; Ferretti, V.; Gilli, G. and Borea, P. A. (1994) *J Phys Chem* 98, 1515–1518.
16. Grunwald, E. and Steel, C. (1995) *J Am Chem Soc* 117**,** 5687–5692.
17. Raffa, R. B. In: R.B. Raffa, Editor, Wiley & Sons, New York (2001), 221–416.

18. Borea, P. A.; Dalpiaz, A.; Varani, K.; Gessi, S. and Gilli, G. (1996) *Life Sci* 59**,** 1373–1388.
19. Borea, P. A.; Varani, K.; Gessi, S. and Gilli, G. (1998)*Biochem Pharmacol* 55**,** 1189–1197.
20. S. Merighi, S.; Varani, K.; Gessi, S.; Klotz, K. N.; Leung, E. and Baraldi, P.G. (2002) *Biochem Pharmacol* 63, 57–161.
21. Gilli, P.; Gilli, G.; Borea, P. A.; Varani, K.; Scatturin, A. and Dalpiaz, A. (2005) *J Med Chem.* 48, 2026–2035.
22. Ohtaka, H.; Velázquez-Campoy, A.; Xie, D.and Freire, E. (2002) *Protein Sci.* 11, 1908-16.
23. Velazquez-Campoy, A.; Kiso, Y. and Freire, E. (2001) *Arch Biochem Biophys.* 390, 169-75.
24. Grunwald, E. and Steel, C. (1995) *J Am Chem Soc* 117, 5687–5692.
25. Burley, S. K. and Petsko, G.A. (1984) *Science*, 229, 23-29.
26. G. Ravishanker, G.; Mehrotra, P. K.; Mezei, M. and Beveridge, D. L.(1984) *J. Am. Chem. Soc.* 106, 4102.

CHAPTER 4

In vitro studies on pharmacological implication of CPI in tumor cells and osteoclasts

Summary:

Osteoporosis, one of the principal causes of ill health in the elderly (>60years of age), is characterized by a continual surplus of osteoclastic bone degradation. This is facilitated by the degradation of organic matter by secreted cathepsin K or cathepsin L and Matrix Metallo Proteinases (MMPs), in absence of cathepsin K, from osteoclasts after dissolution of mineral crystallites due to lowered pH at the attachment site between osteoclasts and the bone surface. Therefore, cathepsin K and L seems to be an attractive target for therapeutic intervention to ameliorate the significant deleterious impact of bone resorption. In this chapter we report on the basis of study carried out on few human and murine tumor cell lines that, the newly isolated thiol protease inhibitor (CPI-2081) from *Streptomyces sp.* NCIM 2081 is able to inhibit directional in vitro tumor cell motility at lower molar concentration of CPI-2081 (25-50 μM) than the molarity required for cytotoxicity (≥100μM). In the second significant study it was found that the CPI-2081 inhibits the RANKL (**R**eceptor **A**ctivator of **N**F-kappaB **L**igand) induced bone resorption. It was also found that CPI-2081 significantly inhibited the formation of osteoclasts at 100 μM. Furthermore, CPI-2081 showed decrease in bone resorption at 100 μM as compared to control.

Introduction:

Osteoclastic bone resorption is one of the major health concerns in the world because it causes osteoporosis. Osteoporosis and low bone mass are currently estimated to be a major public health threat for elderly U.S. women and men aged 50 yrs and older. In 2002, it is estimated that over 10 million people already have osteoporosis in United States. This figure will rise to almost 12 million by 2010 and to approximately 14 million by 2020, if additional efforts are not made to curtail this disease. The structure and density of the skeleton is well maintained in healthy individual primarily by two types of cells: the bone-forming ones called osteoblasts and the bone-resorbing ones called osteoclasts. The osteoclasts and osteoblast differentiate from the hematopoietic myeloid precursors of monocyte lineage that play pivotal roles in bone morphogenesis,

remodeling and resorption[1, 2, 3, 4, 5]. However, due to hormonal imbalances and other reasons, the population and activity of osteoclasts can outnumber osteoblasts, which eventually lead to osteoporosis[6]. It has been concluded in various investigations that deregulated expression of thiol proteases is one of the critical moment which facilitates the osteoclastic bone resorption[7, 8]. The osteoclast resorbs bone by the action of proteolytic enzymes at low pH in and around resorption lacuna. The enzymes that are found to be crucial in resorption are thiol proteinases, cathepsin K and L[9, 10]. Cathepsin K is reported to be independently involved in bone resorption activity in lower pH, whereas, cathepsin L is involved in MMP mediated bone resorption in absence of cathepsin K[11]. Cathepsin K is a 24 kDa cysteine protease of the papain superfamily. It was first discovered as OC-2 in cDNA library of a rabbit osteoclast[11]. Subsequently, the human equivalent of the protein was also cloned by several investigators and named cathepsin O[12], K[13], X[14], O_2[15]. Cathepsin L was purified, as 23-24 kDa protein. It exists in various forms and can be derived from sub cellular fraction of rat liver by cell fractionation and osmotic disruption of the lysosomes in the lysosomal mitochondrial pellet[16]. The predominant expression of Cathepsin K in osteoclasts has rendered the enzyme into a major target for the development of novel anti-resorptive drugs[17] and Cathepsin S appears to be a considerable drug target for various inflammatory diseases including rheumatoid arthritis[18, 19]. Parasites such as *Trypnosoma*, *Lieshmania*, and *Plasmodium* also take the advantage of thiol protease activity to survive the hostile microenvironment by immune evasion, hydrolysis of host proteins, enzyme activation and cellular evasion[20, 21, 22, 23, 24, 25]. Disregulated expressions of proteases are convincingly reported to be involved in tumor progression and metastasis[26, 27, 28, 29]. Cysteine proteases also known as thiol proteases have been identified to play a significant role in cancer progression and metastasis. Cathepsins are classical lysosomal cysteine proteases sharing a conserved active site in which amino acid residues cysteine and histidine compose the catalytic ion pair, a distinguishing feature of the papain-like superfamily of cysteine proteases[30, 31]. Thiol proteases such as Cathepsin B expression is increased in many human cancer[32, 33]. Cancer is not a single cell disease as the variety of cells constituting tumor need to communicate with each other which is done by tumor-shed microvesicles and secreted growth factors, chemokines, cytokines and cell adhesion. The pH mediated activation of

gelatinages are facilitated by cathepsin B which in turn lead to the generation of shed microvesicles by tumor cell as a mean to communicate with each other as well as normal microenvironment in near vicinity[34]. Enhanced activity of cysteine proteases, primarily cathepsin B, has been found in many aggressive human tumours, including breast[35, 36, 37], lung[38, 39], brain[40, 41, 42], gastrointestinal[43, 44, 45], prostate[46, 47] cancers and melanoma [48, 49, 50]. Moreover, increased expression of cysteine cathepsins are used in diagnostics and prognostics of patients with a range of malignancies[51]. Therefore this chapter is designed to study the effect of Thiol protease inhibitor CPI-2081 on RANKL induced bone resorption and tumor cells.

Materials and Method:

Reagents, animals and cell lines:

Human M-CSF (Macrophage- Colony Stimulating Factor) was obtained from R & D Systems (Minneapolis, MN). Human RANKL were from Insight Biotechnology (Wembley, U.K.). All cultures of bone marrow cells were incubated in αMEM (minimal essential media) supplemented with 10% FBS (featal bovine serum), 2 mM L-glutamine, 100 IU/ml penicillin, and 100μg/ml streptomycin. BALB/c mice of 6-8 wk old were obtained from Experimental Animal Facility of the National Center for Cell Science, Pune. Water and food were provided *ad libitum*. Animal protocols were approved by the institutional animal ethics committee.

In vitro osteoclast differentiation and bone resorption: Stromal and lymphocytes-free, M-CSF-dependent, osteoclast precursors were prepared as previously described[52, 53]**.** In brief, bone marrow cells isolated from mice were subjected to gradient purification. Cells at the gradient interface were collected, washed and resuspended in α-Minimum Essential Media containing 10% FCS and incubated for 24 hours with 10ng/ml M-CSF (Monocyte-Colony Stimulating Factor). After 24 h, nonadherent cells were harvested and added to 96-well plates containing coverslips and bone slices. After 1-2 h individual coverslips and bone slices were washed vigorously to remove nonadherent and loosely adherent lymphocytes. Adherent cells were further incubated with M-CSF (30 ng/ml), RANKL (30 ng/ml), and different concentrations of cystine protease inhibitor. Cultures

were fed every third day and after 5 days osteoclast formation on coveslips was examined by tartrate-resistant acid phosphatase (TRAP) staining[54]. Bone slices were processed after 8 days for the assessment of bone resorption.

Assessment of bone resorption: Bone slices were immersed in 4% sodium hypochlorite for 15 min and washed thoroughly to remove the cells. After drying, bone slices were mounted onto glass slides, and then sputter-coated with gold. Bone slices were examined by reflected light microscopy, and bone resorption was quantified using an eyepiece graticule.

Statistical analysis:

Data are presented as mean ± SEM. Statistical differences between the mean values of control and experimental groups were analyzed using Student's t test.

Mammalian Cell Culture: The Human breast adenocarcinoma MDA-MB-231 was cultured in L15 supplemented with 10% fetal calf serum, 100 units/ml penicillin, 100 μg/ml streptomycin and 2 mM glutamine in a humidified atmosphere of 100% air at 37 ^{o}C. Human melanoma A_{375} and murine melanoma $B_{16}F_{10}$ was also cultured in DMEM and L15 medium respectively supplemented with 10% fetal calf serum, 100 units/ml penicillin, 100 μg/ml streptomycin and 2 mM glutamine in a humidified atmosphere of 5% CO_2 and 95% air at 37^{o}C. The cells were passaged at confluency by trypsinization.

Cell survival assay: MDA-MB-231 or $B_{16}F_{10}$ or A_{375} cell were plated in 96 well plates and allowed to grow for 24 h at 37 ^{o}C in humidified atmosphere. After that cells were treated with CPI-2081 in indicated concentration and further kept for 72 h for A375 48 h) After that media was removed and 0.5 mg/ml 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) solution was added to the cells for 4 h. Absorbance was measured by ELISA reader.

Wound healing (scratch closure) Assay: The motility of tumor cells were determined by wound migration assay. Post confluent MDA-MB-231 (Human Breast carcinoma), B16F10 (Murine melanoma) or A375 (Human Melanoma) cells with the typical cobblestone morphology was used for this experiment. Wounds with a constant diameter were made with 10 μl tips and cells were maintained in media supplemented with 5%

FBS. Then cells were treated with CPI-2081 in indicated concentration. Cells were incubated at 37^{o}C for indicated time period (MDA-MB-231 -12h, B16F10- 18 h, A375- 24 h). After termination of experiments wound photographs were taken under phase contrast microscope (Nikon).

RESULTS AND DISCUSSION:

Cysteine protease inhibitor (CPI-2081) inhibits receptor activator of NF-kappaB ligand (RANKL) induced osteoclast differentiation and bone resorption:

Effect of CPI-2081 on RANKL induced osteoclast differentiation: To examine the effect of cysteine protease inhibitor on osteoclast differentiation osteoclast precursors were incubated on coverslips in 96 well plate (5 x 10^4 cells/well) in the presence of M-CSF (30 ng/ml), RANKL (30 ng/ml), and different concentrations of cysteine protease inhibitor. RANKL was able to induce formation of multinuclear osteoclasts which were positive for TRAP (Tartarate Resistant acid Phosphatase). It was found that cysteine protease inhibitor significantly inhibited the formation of osteoclasts at 100 μM ($p < 0.001$) (Fig. 1A). Resorption was quantified using an eyepiece graticule.

Effect of CPI-2081 RANKL-induced bone resorption: Because CPI-2081 inhibited RANKL-induced osteoclast differentiation it was further examined whether it inhibits the function of osteoclasts. To examine this osteoclast precursors were incubated on bovine cortical bone slices in 96 well plate (5 x 10^4 cells/well) in the presence of M-CSF (30 ng/ml), RANKL (30 ng/ml), and different concentrations of cysteine protease inhibitor. After 8 days bone slices were processed for the assessment of bone resorption. In this culture system, RANKL induced formation of resorption lacunae (pits) on bone slices. CPI-2081 showed decrease in bone resorption at 100 μM as compared to control (Fig. 2).

In bone microenvironment the osteoclast and osteoblasts are involved in bone remodeling. In this process the osteoblasts are involved in bone formation whereas, osteoclast deals with bone resorption. These two kinds of cells maintain a well co-ordinated balance. The excess of osteoclastic bone resorption over osteoblastic bore formation results in osteoporosis and other bone metabolic disorders. In order to verify

the inhibitory effect of CPI-2081 on bone resorption we investigated its influence on the formation and differentiation of osteoclast derived from murine bone marrow as well as on pit area formed on bone slices to examine the clinical application of CPI-2081. The obatained data were consistent with the results that inhibition of Cathepsin K retarded bone matrix degradation in CK knockout mice[55]. These results are also in agreement with the results of the inhibition of osteoclast generation[56] and bone resorption by barberine, a specific pharmacological inhibitor of cathepsin K[57]. It is well documented that the Cathepsin K is mainly localized within osteoclasts and is absent or at low frequency in other tissues [58]. Therefore, these results indicate that the pharmacological activities of cathepsin K inhibitors such as CPI-2081 are restricted within osteoclasts. CPI-2081 prevented bone resorption in the mice, resulting in prevention of the RANKL induced bone loss in dose dependent manner. It is well known that bone formation and resorption are tightly coupled. However, these results indicate that the inhibitory effect of CPI-2081 on osteoclast differentiation was not by apoptosis since there was no effect on mononuclear cells of monocyte/macrophage lineage (Fig. 1B). Thus, these results suggest that CPI-2081 inhibits osteoclast differentiation without any toxic effect on osteoclast precursors and hence bone formation.

CPI-2081 inhibits tumour cell migration *in-vitro*: Cancer cells spread from the primary tumour either as individual cells, using amoeboid, or as cell sheets, strands and clusters by means of collective migration. Thiol protease antagonists such as leupeptin and antinpain have been demonstrated to retard the cell translocation during wound repair process[59]. Therefore, we looked into the ability of CPI-2081 to inhibit the tumour cell migration in vitro. To examine the effect of CPI-2081 on tumour cell migration, we performed the assay as described by Li *et.al.*. Briefly, the motility inhibition of tumor cells was determined by capability of inhibitor to inhibit the scratch (wound) closure over time in monolayer confluent cells[60]. Post confluent MDA-MB-231 (Human Breast carcinoma), B16F10 (Murine melanoma) or A375 (Human Melanoma) cells with the typical cobblestone morphology were used for this experiment. Wounds with a constant diameter were made with 10 µl tips and cells were maintained in media supplemented with 5% FBS. Then cells were treated with CPI-2081 of defined concentrations. Cells

were incubated at 37 ^{o}C for indicated time period (MDA-MB-231, 12hrs; $B_{16}F_{10,}$ 18 hrs and A_{375}, 24 hrs). After termination of experiments wound photographs were taken under phase contrast Nikon microscope. During the investigation it was observed that the 25 μM CPI-2081 is able to inhibit the wound closure in, A_{375} human melanoma cells and MDA-MB-231 human breast carcinoma cells as compared to vehicle control. Similarly 25mM CPI-2081 was enough to inhibit the wound closure in $B_{16}F_{10}$ Murine melanoma cells by 20 percent. The wound closure inhibition is not due to cytotoxic effect of CPI-2081 as investigated by the MTT cell viability assay. The viability of A_{375} cells was approximately 100 percent as compared to control upto 50μM CPI-2081 in culture conditions, which is lower than the CPI-2081 concentration needed for 30 percent inhibition of cell migration as compared to control (Figure 4A and 4B). Similarly, the CPI-2081 was able to inhibit the cell migration of MDA-MB-231 and $B_{16}F_{10}$ cells at the concentrations lower than the in vitro cytotoxicity dose (Figure 3). The leading problem in cancer prevention strategy is the migration of malignant tumor to colonize healthy tissues in other parts of the body and side effect of chemotherapeutic drugs on normal cells. Due to diverse nature of tumor cells in respect of type, grade, localization and close similarity to cellular composition of normal tissue, it's difficult to have anticancer penicillin, which would cure the cancer completely without much side effects . Therefore, it would be more reasonable and safer to develop drugs to contain the disease by preventing the tumor migration, hence metastasis, rather that trying to kill it. We are reporting the CPI-2081 which harbors the potential to be developed as antimetastatic drug which can be used to arrest the tumor growth and dispersal without compromising on health of other living tissues.

Conclusion: In conclusion, the CPI-2081 isolated and completely characterized by us whows potential to be developed as an antimetastatic as well as antiosteoporotic agent.

Figures:

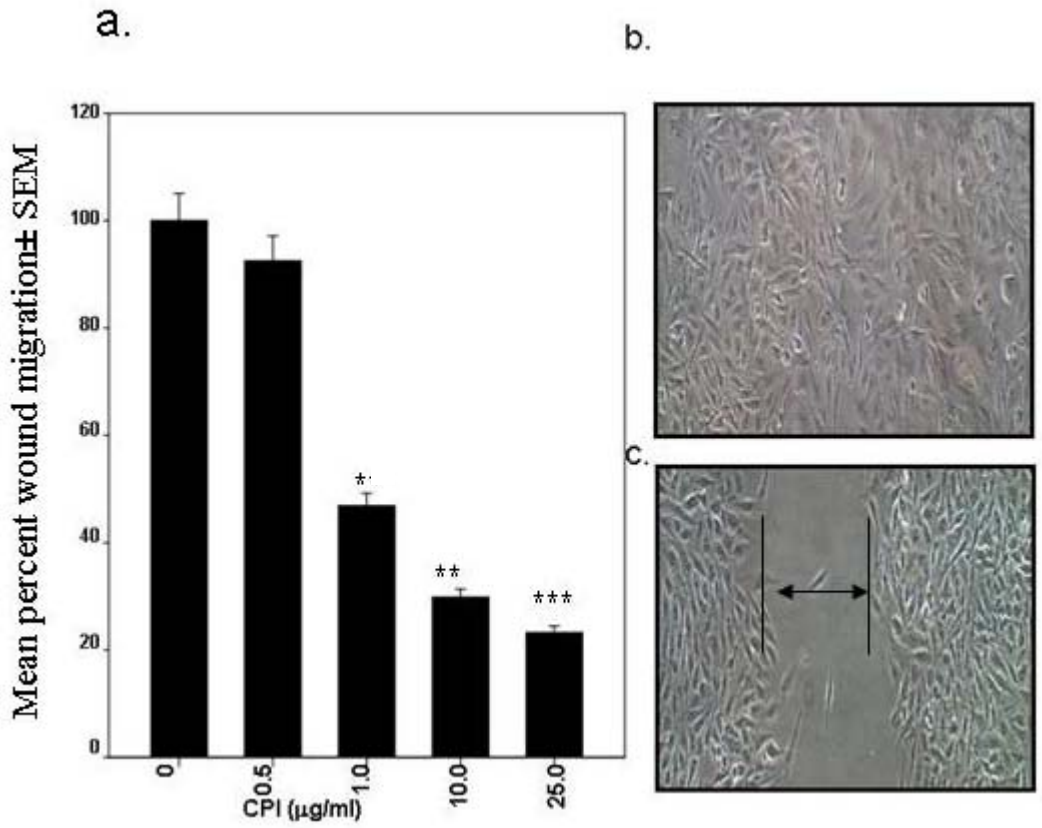


Figure 1: Wound migration assay in MDA-MB-231(human breast carcinoma cells)*: (a) mean percent wound migration is plotted against various concentrations of CPI-2081 as indicated in above graphical representation. The motility of tumor cells were determined by wound migration assay. Post confluent MDA-MB-231(Human Breast carcinoma) cells showing the typical cobblestone morphology was used for this experiment. Wounds with a constant diameter were made with 10 µl tips and cells were maintained in media supplemented with 5% FBS. Then cells were treated with CPI-2081 in various concentrations. Cells were incubated at 37 ^{o}C for 18hrs. After termination of experiments wound photographs were taken under phase contrast microscope (Nikon). (b) Vehicle control (c) 25µg/ml CPI-208. Statistical analysis, P-values for significantly different means, P*=0.010, P**=0.005, P***=0.004 verses control.*

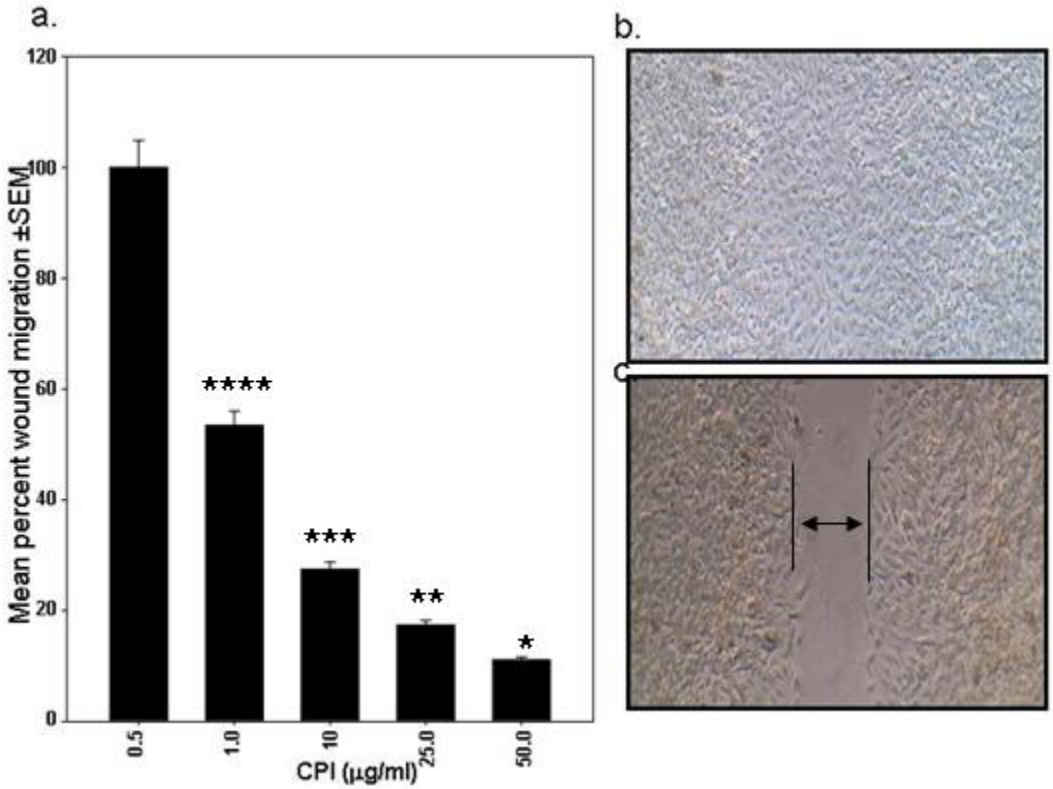


Figure2: Wound migration assay in $B_{16}F_{10}$ Murine melanoma cells: *(a) Mean percent wound migration (standard Error of Mean) is plotted as a function of concentration of CPI-2081 as indicated. The motility of tumor cells were determined by wound migration assay. Post confluent $B_{16}F_{10}$ (Murine melanoma) cells showing the typical cobblestone morphology was used for this experiment. Wounds with a constant diameter were made with 10 µl tips and cells were maintained in media supplemented with 5% FBS. Then cells were treated with CPI-2081in indicated concentration. Cells were incubated at 37 ^{o}C for 18hrs. After termination of experiments wound photographs were taken under phase contrast microscope (Nikon). (b) Vehicle control (c) 50 µg/ml CPI-2081. Statistical analysis, P-values for significantly different means, P****=0.013, P***=0.004, P**=0.003, p*=0.002 verses 0.5µg/ml.*

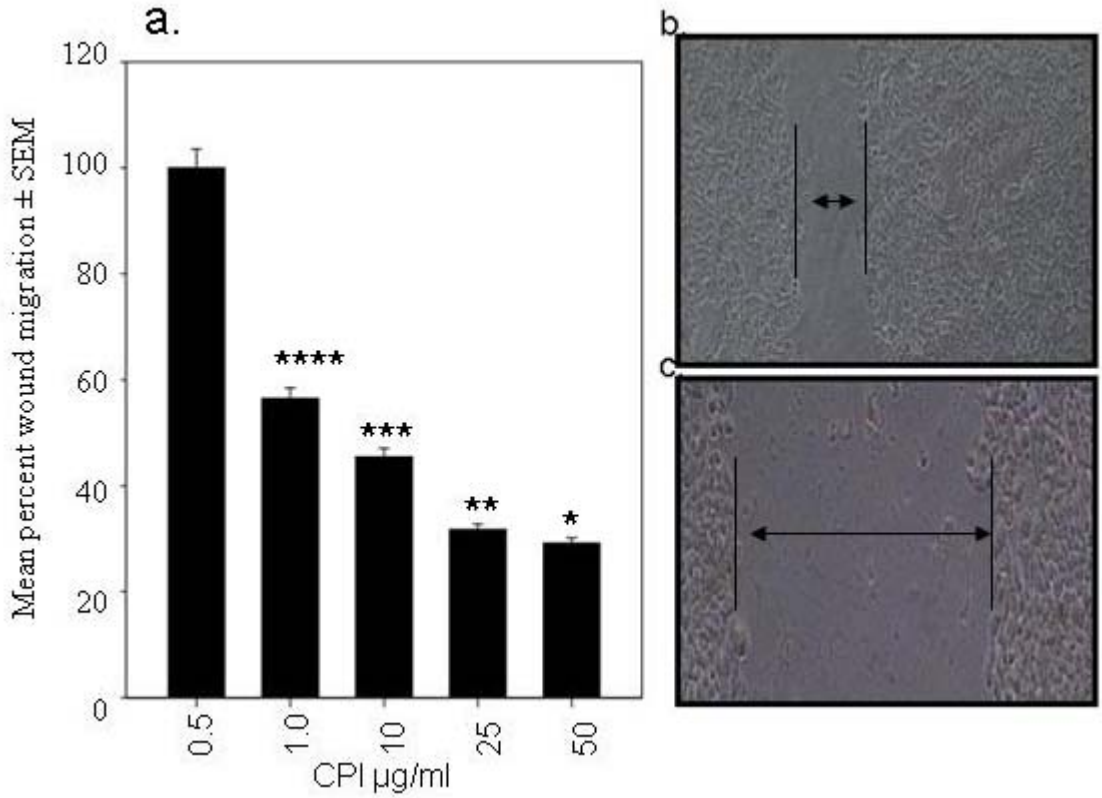


Figure 3: Wound migration assay in A_{375} human melanoma cells: *(a) Mean percent wound migration (Standard Error of Mean) is plotted as a function of concentration of CPI-2081 as indicated. The motility of tumor cells were determined by wound migration assay. Post confluent A_{375} (human melanoma) cells showing the typical cobblestone morphology was used for this experiment. Wounds with a constant diameter were made with 10 µl tips and cells were maintained in media supplemented with 5% FBS. Then cells were treated with CPI-2081 in indicated concentration. Cells were incubated at 37 ^{o}C for 24hrs. After termination of experiments wound photographs were taken under phase contrast microscope (Nikon). (b) Vehicle control (c) 50 µg/ml CPI-2081. Statistical analysis, P-values for significantly different means, P****=0.008, P***=0.004, P**=0.002, p*=0.002 verses 0.5µg/ml.*

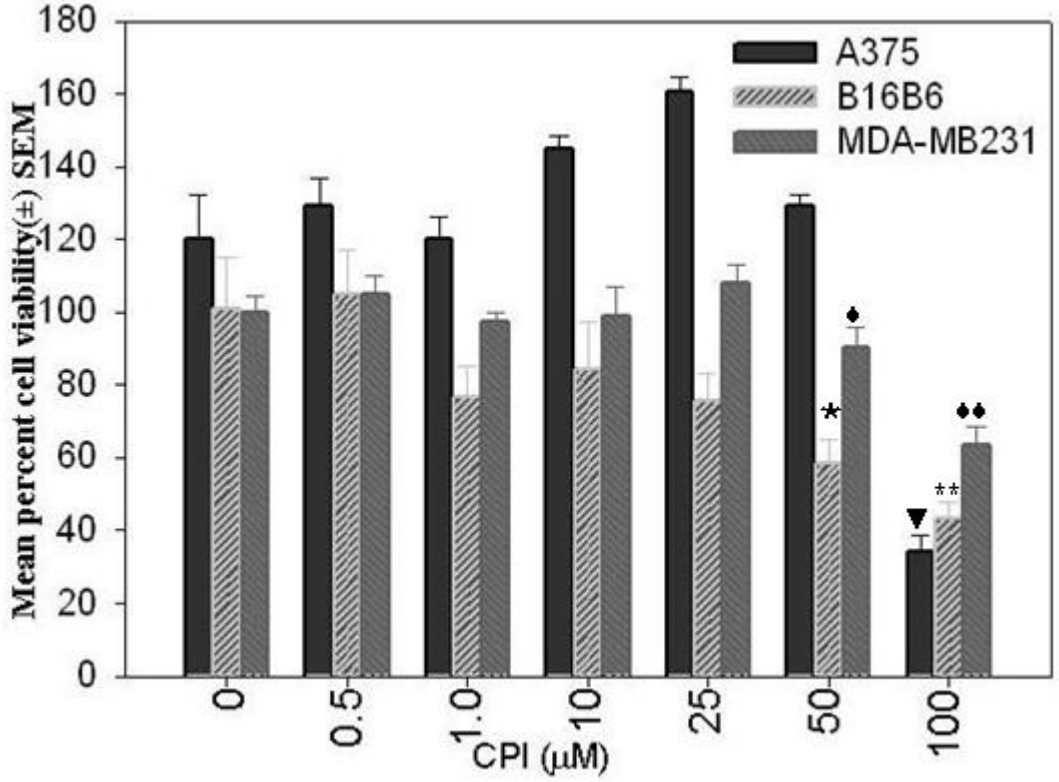


Figure4: Effect of CPI-2081 on cell viability as investigated with the help of MTT cell viability assay in B16F10, mouse melanoma; MDA-MB231, Human breast carcinoma and A375, Human melanoma cells*: MDAMB231 or B16F10 or A375 cell were plated in 96 well plate and allow to grow for 24 h at 37o C in humidified atmosphere. After that cells were treated with CP1 in indicated concentration and further kept for 72 h for A375 48 h) After that media was removed and 0.5 mg/ml 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) solution was added to the cells for 4 h. Absorbance was measured by ELISA reader at 570nm. Mean percent cell viability (± Standard Error of Mean) is plotted as a function of CPI-2081 concentration in μM. Statistical analysis, P-values for significantly different means, $P^{▼}$=0.003, P*=0.012, P**=0.017, $P^{♦♦}$=0.0009 verses control(0 μg/ml)*

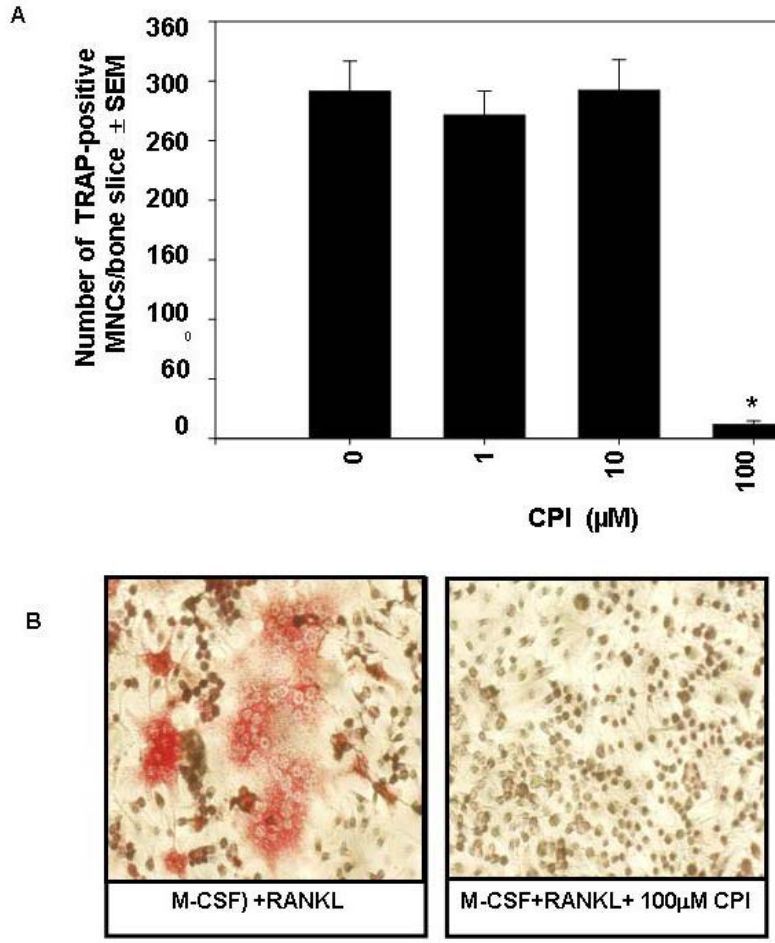


Figure5: Cysteine protease inhibitor inhibits osteoclast differentiation: *M-CSF-dependent, stromal, and lymphocytes-free osteoclast precursors were incubated for 5 days on coverslips in 96-well plate with M-CSF (30 ng/ml), RANKL (30 ng/ml) and various concentrations of CPI-2081. (A) The number of TRAP-positive multinuclear cells per coverslip was quantified. Data are expressed as means ± SEM for five cultures. *p < 0.001 vs. control. Similar results were obtained in two independent experiments. (B) TRAP staining of osteoclasts on coverslips (magnification, x20).*

References:

1. Suda, T., Takahashi, N., Udagawa, N., Jimi, E., Gillespie, M. T., and Martin, T. J. (1999) *Endocr. Rev.* 20, 345–357.
2. Karsenty, G., and Wagner, E. F. (2002) *Dev. Cell* 2**,** 389–406.
3. Manolagas, S. C. (2000)*Endocr. Rev.* 21, 115–137.
4. Lee, Z. H., and Kim, H. H. (2003) *Biochem. Biophys. Res. Commun.* 305, 211–214.
5. Takahashi, N., Udagawa, N., Tanaka, S., Murakami, H., Owan, I., Tamura, T., and Suda, T. (1994) *Dev. Biol.* 163, 212–221.
6. Teitelbaum, S. L. (2000) *Science* 289, 1504–1508.
7. Goto, T., Kiyoshima, T.; Moroi, R.; Tsukuba, T.; Nishimura, Y.; Himeno, M.; Yamamoto, K. and Tanaka, T. (1994) *Histochemistry* 101, 33–40.
8. Delaissé, J. M. ; Eeckhout, Y. and Vaes, G. (1984) Biochem *Biophys Res Commun*. 125, 441-447.
9. Kakegawa, H.; Nikawa, T.; Tagami, K.; Kamioka, H.; Sumitani, K.; Kawata, T.; Drobnic-Kosorok, M.; Lenarcic, B.; Turk, V. and Katunuma, N.(1993). *FEBS Lett.* 321:247-50.
10. Everts, V.; Korper, W.; Ineke, H.; Jansen, D. C.; Bromme, D.; Cleutjens, K. B. J. M.; Heeneman, S.; Peters, C.; Reinhecke, T.; Saftig, P. and Beertsen. W. (2006). *Journal of Bone and Mineral Research*, 21, 1399-1408.
11. Tezuka, K.; Tezuka, Y.; Maejima, A.; Sato, T.; Nemoto, K.; Kamioka, H.; Hakeda, Y. and Kumegawa, M. (1994) *J. Biol. Chem.* 269, 1106–1109.
12. Shi, G.; Chapman, H. A.; Bhairi, S. M.; DeLeeuw, C.; Reddy, V. Y. and Weiss, S. J.(1995) *FEBS Lett.* 357, 129–134.
13. Inaoka, T.; Bilbe, G.; Ishibashi, O.; Tezuka, K. I.; Kumegawa, M. and Kokubo, T. (1995) *Biochem. Biophys. Res. Commun.* 206, 89–96.
14. Li, Y. P.; Alexander, M.; Wucherpfennig, A. L.; Yelick, P.; Chen, W. and Stashenko, P. 1995) *J. Bone Miner. Res.* 10, 1197–1202.
15. Brömme, D. and Okamoto, K. (1995) *Biol. Chem. Hoppe-Seyler* 376, 379–384.
16. Kirschke, H.; Langner, J.; Wiederanders, B.; Ansorge, S. and Bohley, P. (1997) *Eur J Biochem.* 74, 293-301.

17. Wang, D.; Li, W.; Pechar, M.; Kopecková, P.; Brömme, D. and Kopecek, J. (2004) *Int J Pharm*. 277, 73-79.
18. Desai, S. N.; White, D. M.; O'shea, K. M.; Brown, M. L.; Cywin, C. L.; Spero, D. M. and Panzenbeck, M. J. (2006) *Eur J Pharmacol.* 2006, 538, 168-174.
19. Weidauer, E.; Yasuda, Y., Biswal, B. K.; Cherny, M.; James, M. N. and Brömme, D. (2007) *Biol Chem*. 388, 331-336.
20. Costales, J. and Rowland, E. C. (2007) *J Parasitol.* 93, 1350-1359.
21. Pandey, K. C.; Singh, N.; Arastu-Kapur, S.; Bogyo, M.; and Rosenthal, P. J. (2006) *PLoS Pathog.* 2, e117
22. Sajid, M. and McKerrow, J. H. (2002) *Mol. Biochem. Parasitol. 120*, 1-21.
23. Hernandez, A. A. and Roush, W. R. (2002) *Curr. Opin. Chem. Biol. 6*, 459-465.
24. Shenai, B. R.; Sijwali, P. S.; Singh, A. and Rosenthal, P. J. (2000) *J. Biol. Chem. 275*, 29000-29010.
25. Sijwali, P. S.; Shenai, B. R.; Gut, J.; Singh, A. and Rosenthal, P. J. (2001) *Biochem. J. 360*, 481-489.
26. Wang, B.; Sun, J.; Kitamoto, S.; Yang, M.; Grubb, A.; Chapman, H. A.; Kalluri, R.; and Shi, G. P. (2006) *J. Biol Chem.* 281, 6020-6029.
27. Premzl, A.; Puizdar, V.; Zavasnik, B. V.; Kopitar, J.; N.; Lah, T. T.; Katunuma, N.; Sloane, B. F.; Turk, V. and Kos, J. (2001) *Biol Chem*. 382, 853-857.
28. Sameni, M.; Moin, K. and Sloane, B. F. (2000) *Neoplasia*. 2, 496-504.
29. Lah, T. T. and Kos, J. (1998)*Biol Chem.* 379, 125-130.
30. Turk, D., Janjic, V. and Stern, I. (2001) *Embo J.* 20, 6570 – 6582.
31. Affara, N. I.; Andreu, P. and Coussens, L. M. (2009) Methods Mol Biol.539:1-32
32. Heidtmann, H.; H.; Salge, U.; Abrahamson, M.; Bencina, M.; Kastelic, L.; Kopitar-Jerala, N.; Turk, V. and Lah, T.(1997) *Clin Exp Metastasis*, 15, 368-81.
33. Sloane, B.; F.; Moin, K.; Krepela, E.; Rozhin, J. (1990) *Cancer Metastasis Rev.* 1990, 9(4):333-52.
34. Giusti, I.; D'Ascenzo, S.; Millimaggi, D.; Taraboletti, G.; Carta, G.; Franceschini, N.; Pavan, A. and Dolo, V. (2008) *Neoplasia*. 10, 481-488.
35. Poole, A. R.; Tiltman, k. J.; Recklies, A.D.; Stoker, T. A. (1978) *Nature* 273, 545-547.

36. Castiglioni, T.; Merino, M. J.; Elsner, B.; Lah, T.T.; Sloane, B. F. and Emmert-Buck. M.R. (1994) *Hum. Pathol.* 25 857-862.
37. Bervar, A.; Zajc, I.; Sever, N.; Katunuma, N.; Sloane, B.F. and Lah. T.T. (2003) *Biol. Chem.* 384, 447-455.
38. Werle, B.; Lotterle, H.; Schanzenbacher, U.; Lah, T.T.; Kalman, E.; Kayser, K.; Bulzebruck, H.; Schirren, J.; Krasovec, M.; Kos, J. And Spiess, E. (1999) *Br. J. Cancer* 81, 510-519.
39. Fujise, N.; Nanashim, A.. Taniguchi, Y.; Matsuo, S.; Hatano, K.; Matsumoto, Y.; Tagawa, Y. and Ayabe, H. (2000) *Lung Cance*r 27, 19-26.
40. Rempel, S.A.; Rosenblum, M.L.; Mikkelsen, T.; Yan, P.S.; Ellis, K.D.; Golembieski, W.A.; Sameni, M.; Rozhin, J.; Ziegler, G. and Sloane, B.F. (1994) *Cancer Res.* 54 6027-6031.
41. Strojnik, T.; Kos, J.; Zidanik, B.; Golouh, R. and Lah, T.Clin. (1999) *Cancer Res*. 5, 559-567.
42. Flannery, T.; McQuaid, S.; McGoohan, C.; McConnell, R.S.; McGregor, G. Mirakhur. M.; Hamilton, P.; Diamond, J.; Cran, G.; Walker, B.; Scott, C.; Martin, L.; Ellison, D.; Patel, C.; Nicholson, C.; Mendelow, D.; McCormick, D. and Johnston, P.G. (2006) *Int. J. Cancer* ,119, 854-860.
43. Watanabe, M.; Higashi, T.; Watanabe, A.; Osawa, T.; Sato, Y.; Kimura, Y.; Tominaga, S.; Hashimoto, N.; Yoshida, Y. and Morimoto, S. (1989) *Biochem. Med. Metab. Biol.* 42, 21-29.
44. Liu, Y.; Xiao, S.; Shi, Y.; Wang, L.; Ren, W. and Sloane, B.F. (1998) *Chin, Med. J. (Engl.)* 111, 784-788.
45. Ebert, M. P.; Kruger, S.; Fogeron, M. L.; Lamer, S.; Chen, J.; Pross, M.; Schulz, H.U.; Lage, H.; Heim, S.; Roessner, A.; Malfertheiner, P. and Rocken, C (2005) *Proteomics,* 5, 1693-1704.
46. Sinha, A. A.; Gleason, D.F.; Staley, N. A.; Wilson, M. J.; Sameni, M. And Sloane, B.F. (1995) *Anat. Rec.* 241, 353-362.
47. Nagler, D.K.; Kruger, S.; Kellner, A.; Ziomek, E.; Menard, R.; Buhtz, P.; Krams, M.; Roessner, A. and Kellner, U. (2004) *Prostate,* 60, 109-119.

48. Frohlich, E.; Schlagenhauff, B.; Mohrle, M.; Weber, E.; Klessen, C. And Rassner, G. (2001) *Cancer,* 91, 972-982.
49. Kos, J.; Stabuc, B.; Schweiger, A.; Krasovec, M.; Cimerman, N.; Kopitar- Jerala, N. and Vrhovec, I. (1997) *Clin. Cancer Res.* 3, 1815-1822.
50. Jedeszko, C. and Sloane, B.F. (2004) *Biol. Chem*. 385, 1017-1027.
51. Kos, J. and Lah, T.T. (1998) *Oncol. Rep.* 5, 1349-1361.
52. Khapli, S. M.; Mangashetti, L. S.; Yogesha, S. D.; and Wani. M. R. (2003) *J. Immunol.* 171, 142-151.
53. Lean, J. M.; Fuller, K.; and Chambers. T. J. (2001) *Blood* 98, 2707-2713.
54. Mangashetti, L. S.; Khapli, S. M. and Wani. M. R. (2005) *J. Immunol.* 175, 917-925.
55. Hou, W. S.; Bromme, D.; Zhao, Y.; Mehler, E,; Dushey, C.; Weinstein, H.; Miranda, C. S.; Fraga, C.; Greig, F. and Carey, J.(1999) *J. Clin Investig* 103**,** 731–738.
56. Hu, J.P.; Nishishita, K.; Sakai, E.; Yoshida, H.; Kato, Y.; Tsukuba, T. and Okamoto, K. (2008) *Eur J Pharmacol.* 580, 70-79.
57. Wei, P.; Jiao, L.; Qin, L. P.; Yan, F.; Han, T. and Zhang, Q. Y.(2009)*J. Chinese Int. Medicine.* 2009, 7, 342-348.
58. Drake, F. H.; Dodds, R. A.; James, I. E.; Connor, J. R.; Debouck, C.; Richardson, S.; Lee- Rykaczewski, E.; Coleman, L.; Rieman, D. and Barthlow, R.1996. *J Biol. Chem.* 271,12511–12516.
59. Gordon, R. S.; DeMoss, J. *Exp. Cell research*. 1999, 246, 233-242.
60. Li, W.; Henry, G.; Jianhua, F.; Bandyopadhyay, B.; Pang, K.; Garner, W.; Chen, M.;Woodley, D. T.*J Invest Dermatol.* 2004. 123, 622 –633.

CHAPTER 5

General Discussion and Conclusion

GENERAL DISCUSSION:

Mostly, Proteins undergo many reversible posttranslational medications, often, involving proteases during their life time of survival, such as, polypeptide splicing. Proteases (EC No: 3.4) are class of hydrolytic enzymes catalyzing hydrolysis of amide bonds between amino acid residues in proteins. However, the hydrolysis of proteins by proteases is an irreversible phenomenon. Once proteins are hydrolyzed, the only means existing for reconstructing the intact molecule is to instruct the cellular machinery to undergo more of transcription and translation. As stated above, the nature of proteolysis is not reversible. Therefore, it is expected that proteolytic enzymes have evolved to facilitate the biological pathways that are themselves irreversible for example blood coagulation, digestion, maturation of growth factors and cytokines, processing of zymogens and pro-hormones, apoptosis, and degradation of unwanted or excess intracellular proteins. Proteolysis is a ubiquitous mechanism in the cell, essentially used to regulate the function and future of proteins. Proteases have been classified on the basis of the reactive amino acid present on active site and named accordingly, such as cysteine, aspartic, serine etc. Cathepsins are one of them which are involved in various abnormal and normal metabolic processes such as antigen presentation, tumor progression, bone remodeling etc.

A large number of drugs that are used even today, have been developed using natural products and, are used as therapeutics, either as unmodified or chemically modified version, were once isolated from natural sources. But in due course of drug discovery programs, the natural product isolation lacked its momentum and combinatorial chemistry and other high throughput methods became more popular method of drug discovery. Despite having a large number of synthetic compounds in our databases, we do not get unique biologically active compound, which goes to clinical trails so often. This problem is more or less solved when one goes for activity based screening methods for their drug discovery programs, where, we get an already biologically active compound, therefore, with some chemical modification we can go for animal and clinical trials to develop a new drug. Moreover, considering the diverse nature of biological system, nothing seems to be impossible in order to get a compound with an unusual structural feature and affecting a specific target. Therefore, it is imperative to pay more attention towards activity based natural product isolation and structure

elucidation for future use as drugs. It is estimated that only 10 out of 10,000 new chemical entities (NCE) go to clinical trials out of which only one probably reaches market. Improving these statistics is a very hard and challenging job, since a compound must fit a specific profile at first screening to be considered for further development and these early screenings need to be very selective. What can be improved with highly predictive information rich screening schemes is the quality of compounds taken through various pre clinical stages and into clinical studies. One specific key area is drug target binding. Presently, most screening procedures simply look for the high affinity inhibitor/ligand-target/ receptor binding without taking into consideration what kind of bond interactions are involved in binding. Thermodynamic studies and feasibility of reactions in aqueous microenvironment may help us to know whether the particular chemical entity (compound/ ligand/ inhibitor) can be developed as drug candidate. The leading problem in cancer prevention strategy is the migration of malignant tumor to colonize healthy tissues in other parts of the body and side effect of chemotherapeutic drugs on normal cells. Due to diverse nature of tumor cells in respect of type, grade, localization and close similarity to cellular composition of normal tissue, it's difficult to have anticancer penicillin, which would cure the cancer completely without much side effects . Therefore, it would be more reasonable and safer to develop drugs to contain the disease by preventing the tumor migration and metastasis, rather that trying to kill it.

Therefore, the present work is designed to isolate new small molecule cysteine protease inhibitors from natural extracts, which could have potential therapeutic applications and use them as anti cancer and anti osteoporotic agents.

Isolation and characterization of CPI-2081: We have described the isolation and structure elucidation of a novel competitive inhibitor of cysteine protease (CPI-2081), which is a mixture of two structurally very close but distinct peptides in estimated molar proportion of 35: 65. There are many examples of the complete structure elucidation with the help of NMR spectroscopy has been reported in literature. We have elucidated the structure following similar spectroscopic techniques[1, 2, 3]. Four out of five amino acids in both peptides (major and minor) are common. The complete chemical structure was elucidated with the help of various spectroscopic techniques. It is a pentapeptide

containing unique groups and modifications such as acetylation of N-terminal leucine and tertiary butylation of cysteine residue. These kinds of groups such as, acetylation and cyclisation of bioactive peptide grant the extra protection from endogenous peptide hydrolyzing enzymes *in vivo*. *t*-Butyl group bearing compounds are reported to be rarely found in nature, and, are mostly isolated from marine sponges, include peptides, terpenes, carbinols, esters and ketones[4, 5, 6, 7]. A very abnormal *t*-butyl ketone coumarin swietenone was isolated from *Chloroxylon swietenia,* a tree found in India[8]**.** The obvious question which arises is that how these rare chemical groups are formed in natural microenvironment. Generally, it is believed that secondary metabolites are the compounds which are synthesized by the plants and microorganisms, either, in response to unfavorable condition for their growth and development or to be used as defense arsenal against pathogens. Studies have been carried out to find out the synthesis of t-butyl groups *in vivo*. For example, Nakanishi and Habagushi have described the mevalonate pathway for the biosynthetic origin of *t*-butyl group in nature[9]. In their scheme, the *t*-butyl group arises by cleavage of the C–C bond adjacent to a *gem*-dimethyl unit followed by methylation elicited by *S*-adenosylmethionine. However, the recent investigations have demonstrated that mevalonate pathway is only a minor route to the *t*-butyl biosynthesis and non mevalonate pathays are also responsible for its biosynthesis[10, 11]. The compounds bearing *t*-butyl are very stearically protected which can either hinder the chemienzymatic process or nature can use these groups to attain sterioselectivity during the reaction. Therefore, the CPI-2081 bearing this group may prove to be very sterioselective in addition to increasing the extent of hydrophobic interaction during chemienzymatic reaction and due to *N*-terminus modification, it will also be resistant to enzymatic degradation *in vivo*. It is possible to validate the structure if one can synthesize all stereoisomers of both the compounds and selecting the peptide having thiol proteolytic inhibition activity.

For many years, it has been reported that change in enthalpy (ΔH) and entropy (ΔS) values of drug interaction with a defined receptor can often give a simple "*in vitro*" way to differentiate the ability of the drug to interfere with the signal transduction pathways[12, 13, 14, 15, 16]. The thermodynamic study of favorability has been evaluated for various protease inhibitors[17, 18]. All these data suggest an inter correlation between

specific binding and the variation of water molecules present at receptor surfaces. In addition, based on the thermodynamic compensation a general model of drug–receptor interaction has been proposed. In this model the solvent molecules do not modify the intrinsic values of the affinity constant of the drug–receptor interaction because the standard free energy for solvent reorganization can be near to zero and the values of binding parameters are due to specific features of the ligand and receptor in the binding process, and not by the solvent[19]. On the other hand, $\Delta H°$ and $-T\Delta S°$ values are related to the rearrangements occurring during the binding, in the solvent–drug and solvent–receptor interfaces[19]. Therefore, it seems reasonable to assume that solvent effects might be responsible for the *in vitro* thermodynamic differentiation between various kinds of drugs against a target. From this background, it can be said that the presence of the linearity of Van't Hoff plots for CPI-2081 and Cathepsin under this study is a significant observation. Similar to what verified for other protease inhibitors showing that $\Delta H°$ (standard enthalpy) and $\Delta S°$ (standard entropy) values are independent of temperature and obtained by linear van't Hoff plots[9, 13, 14, 15]. Van't Hoff plot turns out to be linear for CPI-2081 considered in the present study, which means that the value of $\Delta H°$ is not significantly affected by temperature variation in the range investigated. The binding of CPI-2081 with both the cathepsin-L and cathepsin-K is entropicaly driven, which indicates that the displacement of water molecules from aromatic rings of CPI-2081 as well as the enzymes are responsible for the positive change in entropy (20, 21). This idea is strengthened by the study of Ravishanker *et al.* that the potential energy change on pair formation between aromatic rings and other hydrophobic interaction in the presence of solvent makes entropic change resulting from displacement of water molecule to be positive[20, 21]**.** Therefore it is evident that the binding of CPI-2081 with target enzymes, cathepsin-L& K, is made favorable by positive change in entropy, or by the energy released while rearrangement of the solvent (H_2O) [22] molecules between CPI-2081-target enzyme and conformational change due to rearrangements of various hydrophobic interactions during binding process.

Pharmacological significance: In bone microenvironment, the osteoclast and osteoblasts are involved in bone remodeling. In this process the osteoblasts are involved in bone formation whereas, osteoclasts deal with bone resorption. These two kinds of cells

maintain a well co-ordinated balance. The excess of osteoclastic bone resorption over osteoblastic bore formation results in osteoporosis and other bone metabolic disorders. In order to verify the inhibitory effect of CPI-2081 on bone resorption we investigated its influence on the formation and differentiation of osteoclasts derived from murine bone marrow as well as pit area formed on bone slices to determine the possible the clinical application of CPI-2081. The obtained data were consistent with the results that inhibition of Cathepsin K retarded bone matrix degradation in cathepsin-K knockout mice[23]**.** These results are also in agreement with the inhibition of osteoclast generation[24] and bone resorption by barberine, a specific pharmacological inhibitor of Cathepsin-K[25]. It is well documented that the Cathepsin K is mainly localized within osteoclasts and is absent or at low frequency in other tissues[26]. Therefore, these results mean that the pharmacological activities of Cathepsin-K inhibitors such as CPI-2081 are restricted within osteoclasts. CPI-2081 prevented bone resorption in the mice, resulting in prevention of the RANKL induced bone loss in dose dependent manner. It is well known that bone formation and resorption are tightly coupled. However, these results indicate that the inhibitory effect of CPI-2081 on osteoclast differentiation was not by apoptosis since there was no effect of mononuclear cells of monocyte/ macrophage lineage (Fig. 1B). These results suggest that CPI-2081 inhibits osteoclast differentiation without any toxic effect on osteoclast precursors, hence bone formation.

Cancer cells spread from the primary tumour either as individual cells, using amoeboid, or as cell sheets, strands and clusters by means of collective migration. Thiol protease antagonists such as leupeptin and antipain have been demonstrated to retard the cell translocation during wound repair process[27]. Therefore, we looked into the ability of CPI-2081 to inhibit the tumour cell migration in vitro, by performing the assay described by Li *et.al.* 2004. Briefly, the motility inhibition of tumor cells was determined by capability of inhibitor to inhibit the scratch (wound) closure over the time in monolayer confluent cells[28]. Post confluent MDA-MB-231 (Human Breast carcinoma), B16F10 (Murine melanoma) or A375 (Human Melanoma) cells with the typical cobblestone morphology was used for this experiment. Wounds with a constant diameter were made with 10 µl tips and cells were maintained in media supplemented with 5% FBS. Then cells were treated with CPI-2081 in indicated concentration. Cells were incubated at 37^{o}C

for indicated time period (MDA-MB-231, 12hrs; $B_{16}F_{10,}$ 18 hrs and A-375, 24 hrs). After termination of experiments wound photographs were taken under phase contrast microscope. 25 μM CPI-2081 could inhibit the wound closure in, A-375, human melanoma cells and MDA-MB-231, human breast carcinoma cells as compared to vehicle control. Similarly, 25μM CPI-2081 was enough to inhibit the wound closure in B16F10 Murine melanoma cells by 20 percent. The wound closure inhibition was not due to cytotoxic effect of CPI-2081 as investigated by the MTT cell viability assay. The viability of A-375 cells was approximately 100 percent as compared to control with up to 50μM CPI-2081 in culture conditions, which is lower than the CPI-2081 concentration needed for 30 percent inhibition of cell migration as compared to control. Similarly CPI-2081 was able to inhibit the cell migration of MDA-MB-231 and $B_{16}F_{10}$ cells at the concentrations lower than the in vitro cytotoxicity dose. The leading problem in cancer prevention strategy is the migration of malignant tumor to colonize healthy tissues in other parts of the body and side effect of chemotherapeutic drugs on normal cells. Due to diverse nature of tumor cells in respect of type, grade, localization and close similarity to cellular composition of normal tissue, it's difficult to have anticancer penicillin. We are reporting CPI-2081 which harbors the potential to be developed as antimetastatic drug which can be used to arrest the tumor growth and dispersal without compromising health of other living tissues.

CONCLUSION:

In conclusion, our efforts to characterize a low molecular weight, pentapeptide, competitive protease inhibitor, which has inhibitory constant (K_i) against recombinant human Cathepsin-K and L, <100 nM, probably due to its unique natural modification, could be useful for drug development programs directed against osteoporosis and cancer.

REFERANCES

1. Bonnet, S. L.; Steynberg, J. P.; Bezuidenhoudt, B. C. B.; Saunders, C. M. and Ferreira, D (1996) *phytochemistry* 43, 241-251.
2. Bonnet, S. L.; Steynberg, J. P.; Bezuidenhoudt, B. C. B.; Saunders, C. M. and Ferreira, D (1996) *Phytochemistry* 43, 215-228.
3. Bonnet, S. L.; Steynberg, J. P.; Bezuidenhoudt, B. C. B.; Saunders, C. M. and Ferreira, D (1996) *Photochemistry 43, 229-240.*
4. Chevallier, C.; Richardson, A. D.; Edler, M. C.; Hamel, E.; Harper, M. K. and Ireland, C. M. (2003) *Org. Lett.* 5, 3737.
5. Hamel, E. (2002) *Biopolymers*, 56, 142.
6. Matsunaga, S.; Fusetani, N. and Konosu, S. (1985) *Tetrahedron Lett.* 26, 855.
7. Ryu, G.; Matsunaga, S. and Fusetani, N. (1994) *Tetrahedron Lett.* 35, 8251.
8. R. B. Mujumdar, R. B.; Rao, A. V. R.; Rathi, S. S. and Venkataraman, K. (1975)*Tetrahedron Lett.* 16, 867.
9. K. Nakanishi, K. Habaguchi, Y. Nakadaira, M. C. Woods, M. Maruyama, R. T. Major, M. Alauddin, A. R. Patel, K. Weinges and W. B¨ahr, (1971) *J. Am. Chem. Soc.* 93, 3544.
10. Rohdich, F.; Kis, K.; Bacher, A. and Eisenreich, W. (2001)*Curr. Opin. Chem. Biol.*5, 535.
11. Stromgaard, K. and K. Nakanishi, K. *Angew. Chem., Int. Ed.*, (2004) 43, 1640.
12. Borea, P. A.; Dalpiaz, A.; Varani, K.; Gilli, P.; Gilli, G. (2000) *Biochem Pharmacol* 60, 1549–1556.
13. Borea, P. A.; Dalpiaz, A.; Varani, K.; Gessi, S. and Gilli, G. (1996) *Life Sci* 59, 1373–1388.
14. Borea, P. A.; Varani, K.; Gessi, S. and Gilli, G. (1998)*Biochem Pharmacol* 55, 1189–1197.
15. S. Merighi, S.; Varani, K.; Gessi, S.; Klotz, K. N.; Leung, E. and Baraldi, P.G.(2002) *Biochem Pharmacol* 63, 57–161.
16. Gilli, P.; Gilli, G.; Borea, P. A.; Varani, K.; Scatturin, A. and Dalpiaz, A. (2005) *J Med Chem.* 48, 2026–2035.

17. Ohtaka, H.; Velázquez-Campoy, A.; Xie, D. and Freire, E. (2002) *Protein Sci.* 11, 1908-16.
18. Velazquez-Campoy, A.; Kiso, Y. and Freire, E. (2001) *Arch Biochem Biophys.* 390, 169-75.
19. Grunwald, E. and Steel, C. (1995) *J Am Chem Soc* 117, 5687–5692.
20. Burley, S. K. and G. A. Petsko, G.A. (1984) *Science*, 229, 23-29
21. G. Ravishanker, G.; Mehrotra, P. K.;. Mezei, M. and Beveridge, D. L.(1984) *J. Am. Chem. Soc.* 106, 4102 (1984).
22. Gaikwad, S. M. and Khan, M. I. (2006) *Photochem. And Photobiol.*, 82, 1315-1318.
23. Hou, W. S.; Bromme, D.; Zhao, Y.; Mehler, E,; Dushey, C.; Weinstein, H.; Miranda, C. S.; Fraga, C.; Greig, F. and Carey, J.(1999) *J. Clin Investig* 103:731–738.
24. Hu, J.P.; Nishishita, K.; Sakai, E.; Yoshida, H.; Kato, Y.; Tsukuba, T. and Okamoto, K. (2008) *Eur J Pharmacol.* 580, 70-9.
25. Wei, P.; Jiao, L.; Qin, L. P.; Yan, F.; Han, T. and Zhang, Q. Y. (2009) *J. Chinese Int. Medicine.*, 7, 342-8.
26. Drake, F. H.; Dodds, R. A.; James, I. E.; Connor, J. R.; Debouck, C.; Richardson, S.; Lee- Rykaczewski, E.; Coleman, L. and Rieman, D; Barthlow, R.(1996) *J Biol. Chem.* 271:12511–12516.
27. Gordon, R. S. and DeMoss, (1999) J. *Exp. Cell research*. 246, 233-242.
28. Li, W.; Henry, G.; Jianhua, F.; Bandyopadhyay, B.; Pang, K.; Garner, W.; Chen, M. and Woodley, D. T. (2004) *J Invest Dermatol.*. 123, 622 –633.

Publications

1. **Jay Prakash Singh**, Sudarshan Tamang, P. R. Rajamohan N. C. Jima, Gotam Chakborty, Gopal C. Kundu, S. M. Gaikwad and M. I. Khan.(2009) Isolation and Structure Elucidation of a Modified Peptide Thiol Protease Inhibitor from Streptomyces sp. This Negatively Interferes With Cell Motility, Without Affecting Viability of Tumor Cells. *(Communicated)*

2. **Jay Prakash Singh**, Satish, M R vani, S. M. Gaikwad and M. I. Khan (2009) Inhibition of Osteoclast Differentiation and Bone Resorption in-vitro, by CPI, a Peptide Thiol Protease Inhibitor from Streptomyces sp. (*Manuscript under preparation)*