Study of inhibitory effect of epididymal CRES on PC4/PCSK4 activity

By

Priyambada Mishra

Thesis submitted to the Department of Biochemistry, Microbiology and Immunology in partial fulfillment of the requirements for the degree of Master' of Science**.**

Department of Biochemistry, Microbiology and Immunology (BMI)
Faculty of Medicine
University of Ottawa
Ottawa, Ontario, CANADA
January 2011

Abstract

PC4/PCSK4 is the major Proprotein Convertase (PC) enzyme that plays a key role in mammalian fertilisation. It is detected in the acrosomal granules of round spermatids, acrosomal ridges of elongated spermatids and sperm plasma membrane overlying the acrosome with **K-X-K/X-R** as its preferred cleavage motif. Such motifs are present in male germ cell proteins ADAMs, proPACAP and proIGF-1/2 and these precursor proteins are processed most likely by PC4 during spermatogenesis, sperm maturation and sperm-egg interaction. For fertilization to occur, the mature sperm must penetrate the Zona Pelucida (ZP) and bind to the egg. Previously, PC4 null mouse sperm and wild type sperm treated with a specific PC4-inhibitor have shown to reduced abilities to penetrate the cumulus mass, bind to ZP and fertilize eggs. These findings suggest that sperm-PC4 plays an important role in fertilization and hence regulation of its activity is crucial for successful fertilization. But how PC4 activity is regulated *in vivo* is not yet clear. Recently, in epididymal fluid a serpin (serine protease inhibitor) called **CRES** has been described but the protease linked to this serpin in epididymis has not been identified. However in endocrine cells where CRES is also expressed, it inhibits PC2 enzyme. Thus based on localization and preliminary study, we propose that PC4 is the target enzyme for CRES in the reproductive tract. During sperm migration and storage in epididymis, sperm PC4 activity may be modulated by CRES so that premature sperm activation may not occur. Our data showed that CRES inhibits PC4 both *in vitro* (with IC_{50} in μM range) as well as *ex vivo* in human placenta trophoblast cell lines. Moreover CRES was found to be cleaved by PC4 suggesting a Serpin-Protease binding type of mechanism in the inhibition of protease activity. Taken together, we conclude that CRES regulates PC4 activity in reproductive tract crucial for mammalian fertilization.

Acknowledgements

First and foremost, I would like to thank my thesis supervisor Dr Ajoy Basak for his wisdom, guidance and support throughout my project. I also thank my co-supervisor, Dr. Majambu Mbikay for all his suggestion, advice and overall support.

I am deeply appreciative of the helpful comments and suggestions from the members of my thesis advisory committee; Dr Balu Chakravarthy, Senior Research officer, National Research Council, Ottawa and Dr Qiao Li, Assistant Professor, Department of Pathology and Laboratory Medicine, U Ottawa. I am grateful to Dr. Michel Chrétien, EM Scientist, Ottawa Hospital Research Institute (OHRI) for his interest and Dr. Gail A. Cornwall, U Texas, USA for the gift of mouse CRES cDNA. Thanks are also due to Drs Andrée Gruslin, Professor, Director, Division of Maternal-Fetal Medicine and Dr. Qing Qiu, Research Associate, Department of Obstetrics and Gynecology, U Ottawa for their help in providing placental cell line and expertise in the preparation of rec (recombinant)-CRES protein. I am also thankful to Andrew Chen, Francine Sirois, Charles Gymera-Acheampong and all the members of Basak and Mbikay's labs for their technical support and help during the course of my research work. I would also like to mention Denise Joanisse, Secretarial Assistant of Convertase Group, Chronic Disease Program, OHRI for her administrative help during my stay at the OHRI. I would also like to mention Carol Ann Kelly and Victoria Stewart of Biochemistry, Microbiology and Immunology Department, U Ottawa for their invaluable administrative assistance. We thank the Canadian Institutes of Health Research (CIHR) and NSERC (Discovery Program) for financial assistance that made this work possible.

Above all, I would like to thank my parents and family, for their constant support and encouragement and love during my Master's studies.

Table of contents

List of tables

List of figures

List of abbreviations

µ.....micro
Aa.....Amino acid
ACTH.....Adrenocorticotropic hormone
ADAM.....A disintegrin and metalloproteinase
cAMP.....Cylic adenosine monophosphate
CRES.....Cystatin regulated epididymal spermatogenic
DEAE.....Diethylaminoethyl
DTT……...Dithiothreitol
EDTA.....Ethylene diamine tetraacetic acid
ER……..Endoplasmic reticulum
g…….Gram
h.....Human
IC_{50} …..……Half maximal inhibitory concentration
IGF……Insulin-like growth factor
kDa……Kilodalton
K_i ……..Dissociation constant for inhibitor binding
K_m …….Michaelis–Menten Constant
l……..Litre
m……Mouse
MALDI-tof……...Matrix-Assisted Laser Desorption time of flight
MAPK……...Mitogen activated protein kinase
MS…….Mass Spectrometry
nm……..Nanometer
OD…….Optical Density
PACAP……. Pituitary adenylate cyclase-activating polypeptide
PC……..Proprotein Convertase
PCSK…….Proprotein Convertases Subtilisin/Kexin
POMC……...Pro-opiomelanocortin
r……..Rat
RFU……...Raw fluorescence unit
rpm……Revolutions per minute
rec……..Recombinant
RP-HPLC……..Reversed phase high performance layer chromatography
RSL……Reactive Site Loop
SDS PAGE……Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SELDI -tof…….Surface Enhanced Laser Desorption Ionization time-of-flight
TGN…….....Trans Golgi Network
β-LPH……Beta lipotropin hormone
ZP……...Zona pellucida

1. INTRODUCTION

1.1 BACKGROUND ON FERTILIZATION

1.1.1 Spermatogenesis:

Spermatogenesis is a process in which male spermatogonia develop into mature spermatozoa. This process of spermatogenesis takes place in about 35 days and the various phases involved are mitotic, meiotic and post meiotic phases. Sperm is known to possess a polarized structure. The head of sperm comprises the acrosome (that houses many hydrolases), male genome and the site for egg interaction and the tail region consists of the axoneme which is responsible for its motility and swimming ability. In order to gain the polarized structure as well as the fertilizing and egg-binding capacities, various proteins are synthesized and matured on the sperm surface and targeted to different parts of the gamete cells. These events are essential as part of the fertilization process. But these proteins are initially synthesized as larger inactive precursor molecules which then undergo proteolytic processing into active forms. This event is the key to successful sperm-egg fusion leading to fertilization (1-7).

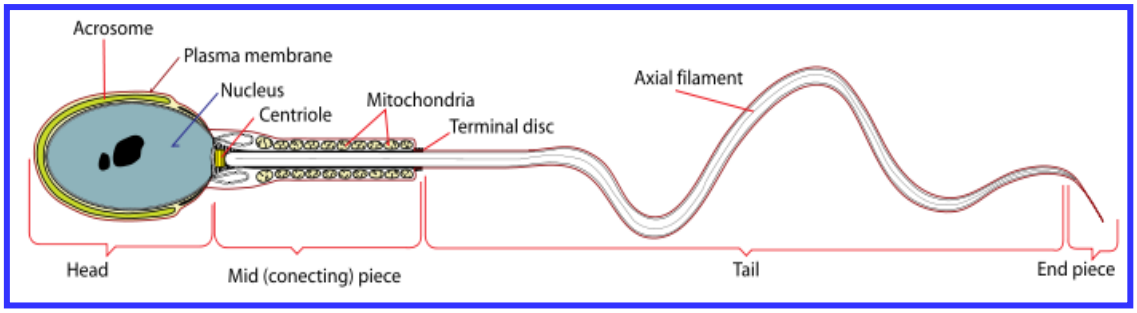
1.1.2 Fertilization Events:

The sperm once produced in the testis has to gain the fertilizing capacity in a step wise manner before it can finally fertilize eggs (1). Firstly, the sperm travels to the epididymis which is an elongated and tortuous coiled structure on the top of the testis, from which it receives the immature sperms and stores it there for several days. The spermatozoa formed in the testis then, enters the caput (head region of epididymis) where it is still in the immature form not being capable of fertilizing the egg. It then proceeds towards the corpus (body region of epididymis) and finally reaches the cauda (tail region of epididymis) but the

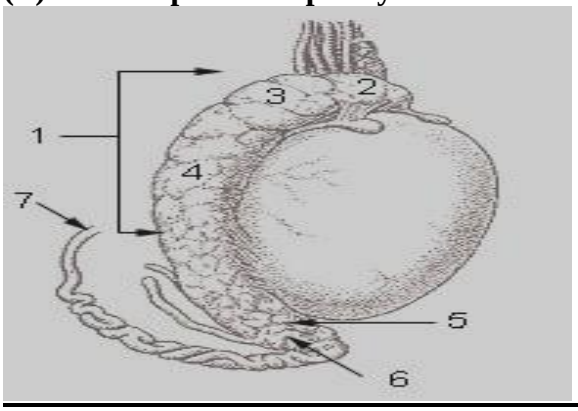
final sperm maturation is completed in the female reproductive tract. The structure of spermatozoa and various parts of the epididymis are shown in **Figure 1 (A & B)**. Thus, spermatozoa maturation in the epididymis undergoes a series of events like chemical modifications of sperm surface proteins, their processing, alterations in plasma membrane characteristics, changes in protease activities and modulation of intracellular constituents and this process is called sperm "**capacitation"**. The sperm is expelled from the cauda into the deferent duct (vas deferens) during the ejaculation and then it travels up to the pelvic cavity to the prostrate and empties into the urethra (8). It is known that there are a number of sperm associated proteins that get activated by the **Proprotein Convertase 4 (PC4**) enzyme *(described in the section below)* via proteolytic cleavage. Some of these proteins associated with sperm maturation and other fertilizing events get modified by glycosylation, disulfide formation and proteolysis (9-12). The fertilizing ability of the epididymal sperm is suppressed due to the masking of zona pellucida (ZP) binding ligands and rigidity of their plasma membrane. The egg binding sites of the sperms become exposed only when the sperm become capacitated inside the female reproductive tract and during this process some of the proteins involved in the sperm-egg interaction also get processed. The capacitated sperms gain capacity of egg binding by penetrating the ZP due to their hyper-motility which occurs after cholesterol efflux. *In vitro* sperm capacitation can also be performed by incubating sperm in medium consisting albumin which induces cholesterol efflux. After successful capacitation, the sperm binds to the egg via ZP by a ligand-receptor mechanism, leading to the activation of sperm signaling events which aid in the acrosome reaction. This release of hydrolases from the acrosome leads to the digestion of ZP, thus creating passages

for sperms to travel through the ZP layer (13-15).These various steps and events occurring during fertilization are shown in **Figure 1C**.

Figure 1: (A) Structure of a spermatozoon

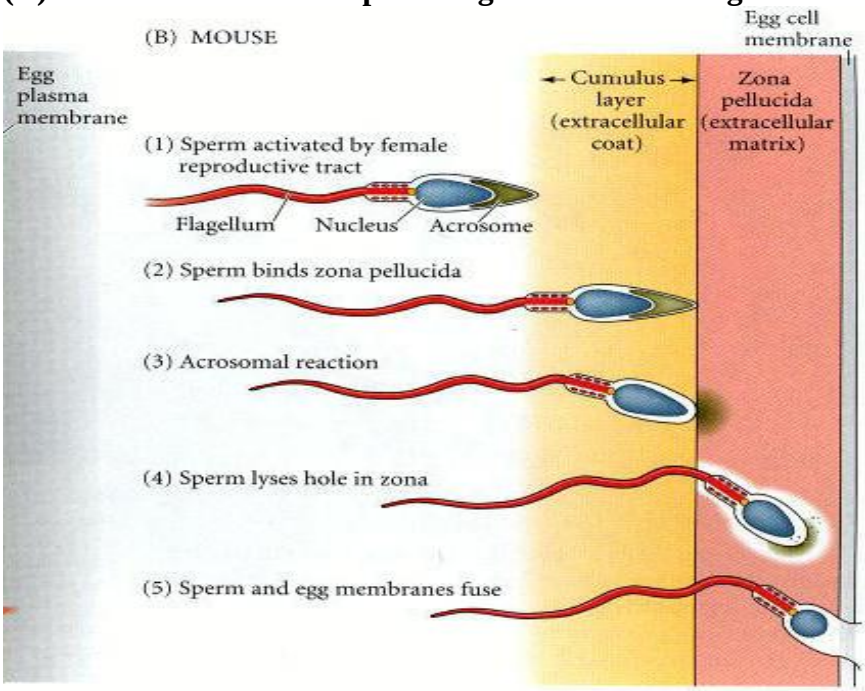


(B) Various parts of Epididymis



1: Epididymis
2: Head of epididymis
3: Lobules of epididymis
4: Body of epididymis
5: Tail of epididymis
6: Duct of epididymis
7: Deferent duct
(vas deferens)

(C) Various events and steps during fertilization using mouse as a model



Figures adapted from references 139,140 & 141, respectively following slight modifications

1.2 PROPROTEIN CONVERTASES

1.2.1 Historical Background

Proprotein convertases (PCs) also called **P**roprotein **C**onvertases **S**ubtilisin **K**exins **(PCSKs)** are calcium dependent serine proteinases that are related to bacterial subtilisin and yeast kexin which cleave large inactive precursor proteins to generate smaller biologically and functionally active mature proteins. These include neuropeptides, hormones, growth factors, receptor proteins, adhesion molecules, surface proteins, enzymes and many others (16, 17). This **"Prohormone/Proprotein Theory"** as it is commonly known was first put forward nearly four decades ago simultaneously by two independent research groups namely Chrétien et al and Steiner et al (18, 19). Thus in 1967, Chrétien et al showed that following removal of signal peptide in the endoplasmic reticulum ACTH and β-LPH were generated from its precursor protein POMC via cleavage by a specific protease hydrolyzing the peptide bond on the carboxyl side of a motif characterized by the presence of a pair of basic residues **(K-R↓ or R-R↓)** Here the arrows refer to the cleavage sites. In the same year Steiner et al., 1967 also reported that the hormone insulin is derived from its precursor protein, proinsulin via cleavage at a specific dibasic site mediated by one or more proteases via a similar type of mechanism. Lack of proper tools, techniques and technologies could not identify the nature of the proteinases involved in these proteolytic events (18, 19). But a few decades later, these proteases were characterized as those which belong to the family of "**Proprotein convertases (PCs)"** and later called by the acronym "**Proprotein Convertase Subtilisin Kexins (PCSKs)"**. Various members of PC family are listed in **Table 1**. It was during the discovery and cloning of the yeast protein *kex 2* (also known as kexin), the first member of the PCSK family was found (20). While the identification and subsequent characterization of the yeast *Kex2* gene was carried out, the mammalian homologues the PCSKs were

discovered (21). Furin was the first member of this family to be discovered followed by six other members which belong to the kexin subtype of PC-super family. Previously bacterial homolog subtilisin and the yeast counterpart kexin were known in the literature and their functional and enzymatic properties have been well studied. Thus it was revealed that kexin also called kex-2 cleave the yeast pro-α factor and pro-killer toxin at specific post dibasic sites. Kexin shared a high homology with the protein product of the c-fes/fps oncogene named as furin (16). This is how the first member of PC family was discovered in late eighties which was followed by the discovery of other members in early to late 1990s (16). The initially discovered PCs cleave proteins at the carboxy terminal of a single or a pair of basic amino acid residues such as **R/K↓X** or **R/K-R↓X** (where R, K, X denote Arginine, Lysine and any amino acid respectively). The presence of Arg residue at P1 and an additional basic amino acid at P2, P4, P6 or P8 positions is essential for both recognition and cleavage. However in few rare cases, cleavages after Lys residue in protein precursors have also been noted (22, 23).

1.2.2 Types of PCs (PCSKs) and their cleavage specificities:

Presently there are 9 known members in the PC or PCSK-family which have been further classified into three different subtypes or sub groups as indicated below:

(A) Kexin type PCs with basic amino acid cleaving activity:

There are 7 members in this category. Many of these possess multiple names. These are PCSK1 (PC1/PC3, SPC3); PCSK2 (PC2, SPC2); PCSK3 (furin, PACE, SPC1); PCSK4 (PC4, SPC5); PCSK5 (PC5, PC6, SPC6); PCSK6 (PACE4, SPC4); PCSK7 (PC7, PC8, LPC, SPC7) and they all cleave proproteins or prohormones with cleavage specificity

characterized by the sequence **(H/R/K)-X-X/K/R-K/R↓X,** where X = any amino acid except Cys.

(B) Non-basic amino acid cleaving PCs: There are two subtypes under this category which are described in the section below:

(i) Pyrolysin-like PC: There is only one member in this type which is called subtilisin kexin isozyme-1 (SKI-1), Site 1 Protease (S1P) or PCSK8. It cleaves proproteins at the carboxy terminus (C-terminus) of a non basic amino acid characterized by the motif **H/R/K-X-(L, I, V)-(L/I)↓X,** where X is any amino acid except Pro, Cys, Glu and Val. Recently a much longer motif with an additional P7 Tyr or Phe residue has been described for SKI-1 enzyme.

(ii) Proteinase K-like PC: Again there is only one member in this category namely neural apoptosis regulated convertase-1 (NARC-1) or PCSK9 but the proteolytic activity of this so called enzyme was never thoroughly described. In fact the only known protease activity of this PC-member is linked to the cleavage of its own prodomain at V-F-A-Q↓S-I-P (24). All these types of PCs with their acronyms and cleavage specificity are represented in a tabular form as shown in **Table 1**.

Table 1: List of various PCs and their acronyms with their cleavage motifs

PC-types	Acronyms	Cleavage specificity
Kexin-like	PCSK1(PC1,PC3,SP3);PCSK2(PC2, SP2);PCSK3(furin,PACE,SPC1); PCSK4(PC4,SPC5);PCSK5(PC5,PC6, SPC6);PCSK6(PACE4,SPC4);PCSK7 (PC7,PC8,LPC,SPC7).	(R/K)n-X-R/K↓ X[b]
Pyrolisin-like	PCSK8 (SKI-1,SIP)	R/K-X-(L,I,V)-Z ↓X[c]
Proteinase-like	PCSK9 (NARC-1)	V-F-A-Q ↓S-I-P[d]

Note: furin: fes upstream protein; **LPC:** lymphoma proprotein convertase; **NARC-1**: neural apoptosis-regulated convertase-1; **PACE**: paired basic amino acid cleaving enzyme; **SIP**: site-one protease; **SKI-1:** subtilisin/kexin isozyme 1; **SPC:** subtilisin-like proprotein convertase; **X[b]** stands for any amino acid except Cys; "**n**" stands for position number such as P2, P4, P6 or P8. After cleavage, the C-terminally exposed basic residue is removed by a carboxypeptidase (CP) unless preceded by a Pro residue. A Gly exposed by CP can be converted to an amide group by the action of peptidyl-alpha amidating monooxygenase (PAM) enzyme; **Z** stands for any amino acid except Pro, Cys, Glu and Val. The vertical arrow (↓) stands for the site of autocatalytic cleavage of proPCSK9 after its prodomain V-F-A-Q↓S-I-P. Mature PCSK9 has no physiological substrate.

The above table with its following legends was adapted from reference (24) following slight modifications.

1.2.3 Structure of proprotein convertases and their functions:

The structural domains of the first 7 PCs of kexin type are quite similar to one another and to yeast kexin and bacterial subtilisin homologues. These specialized endoproteinases act in the secretory or constitutive pathway depending on the nature of the PC enzyme by cleaving the multibasic motifs in proprotein substrates like prohormones, neuropeptides, growth factors and extracellular matrix proteins and this post translational processing transforms the inactive precursors into their physiologically active proteins and peptides (16, 25-27).

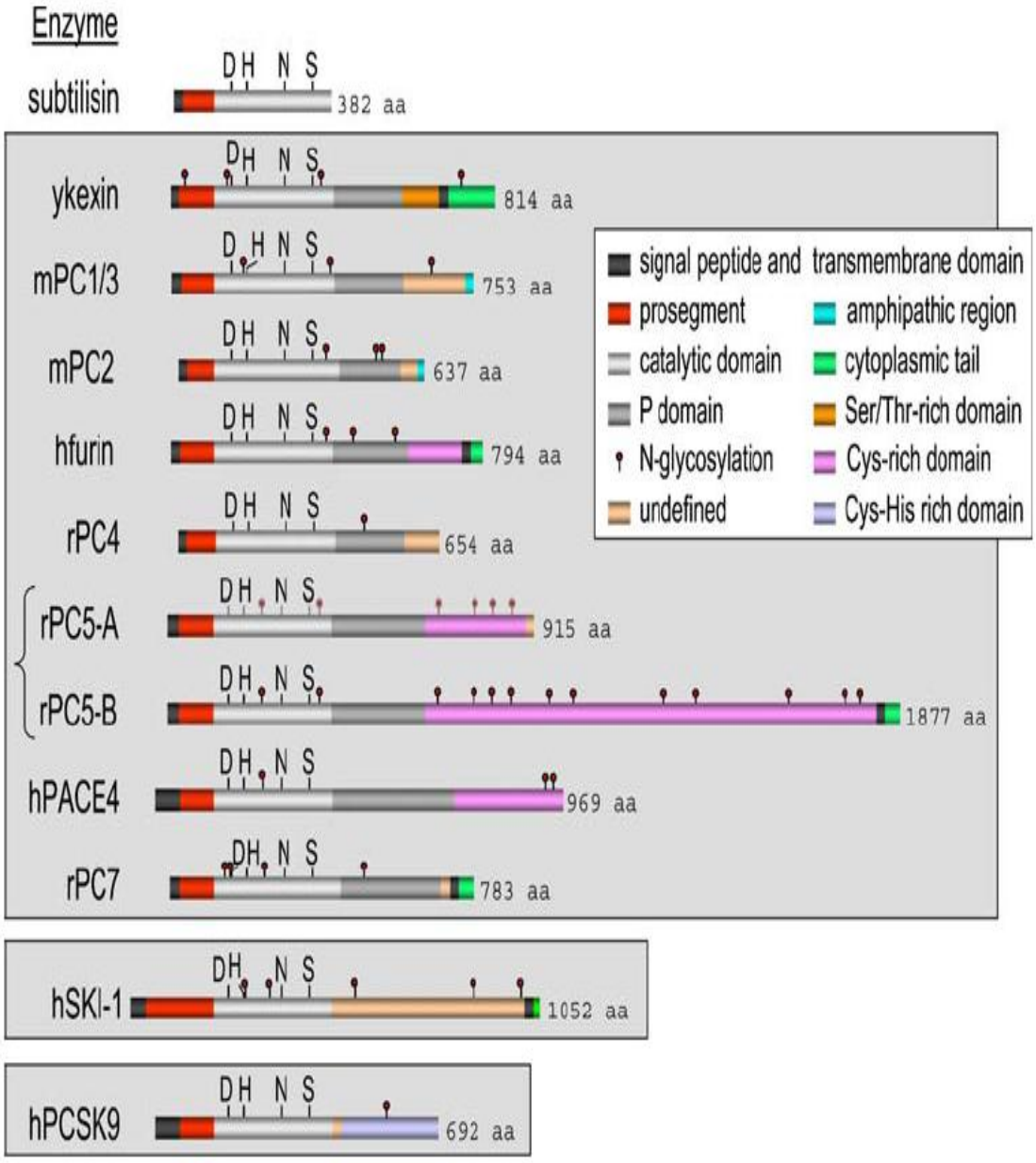
The various characteristic domains such as pre, pro and catalytic regions are apparently well conserved whereas variations and dissimilarities were noted in the C-terminal region, which appears to be more specific for each PC member. The extreme N-terminal region of each represents the signal peptide or pre-region the cleavage of which aids in the enzyme's entry into the endoplasmic reticulum followed by its trafficking in the secretory pathway. Following signal peptide removal by the action of signal peptidase enzyme, each PC enzyme is present as proPC protein which then undergoes cleavage (primary cleavage) to release the prosegment which still remains bound to the mature PC enzyme and thereby blocking its protease activity. It also serves as an intramolecular chaperone in the folding of the enzyme and also as an *in vivo* enzyme inhibitor or regulator. During the transport of the enzyme from ER to the Golgi apparatus (28) there is a drop in pH (from 7.4 to 5.5) which leads to additional secondary cleavages of the prodomain leading to its complete degradation. This allows the resultant formation of fully mature and proteolytically active form of the enzyme (29). The catalytic domain of each PC is quite conserved with that of subtilisin which is responsible for its catalytic reactions leading to the cleavages of larger inactive precursor protein substrates into their smaller bioactive forms.

This domain carries the catalytic triad of **Asp**, **His** and **Ser** and an additional **Asn** residue for the **"oxyanion"** hole of the enzyme, with the exception of PC2 which does not possess this **Asn** residue (instead there is **Asp** residue). Beyond the catalytic domain each PC contains a P-domain (protease domain) which helps in stabilizing the enzyme through hydrophobic interactions with the catalytic domain and its presence plays a vital role for the folding mechanisms of the enzyme in the ER. All PCs are N-glycosylated although the extent and position of glycosylations vary from one PC to another. Following P-domain, PCs contain C-terminal domain and sometimes transmembrane domains and cytosolic tail depending on the type of the enzyme. So far only furin, PC5-B, PC7 and possibly PC4 have been shown to possess trans-membrane domain and as such they are membrane bound. However they may undergo cleavage at the C-terminal before the trans-membrane domain known as **"shedding"** presumably in an autocatalytic manner. This leads to the formation of secreted soluble form of the PC-enzyme (16, 29-32). The detailed structures of all 9 PC-enzymes with their various characteristic domains are shown in **Figure 2.**

Thus, in general all PCs have been found to be synthesized as pre-proproteins containing both the signal peptide and the prodomain. The signal peptide will lead the PCs through the secretory pathway into the ER lumen, where they will be cleaved by signal peptidase and the mature enzyme can undergo numerous post-translational modifications such as, sulfation, glycosylation, phosphorylation, nitration and ubiquitination (29, 31, 32). Normally, at least two cleavages: primary and secondary occur in the N-terminal prodomain. The primary cleavage releases the inhibitory prodomain from the catalytic domain but it remains tightly attached to the mature protease inhibiting its activity until the second cleavage takes place leading to the degradation of prodomain and formation of active

enzyme (16). This is valid for all PC-enzymes except PC7 which does not seem to undergo a secondary cleavage in the prodomain, the implication of which is not clearly understood (29). A method to predict the cleavage sites in protein substrates for PCs has been developed by Duckert et al. on the basis of artificial neural networks (33).

<u>Figure 2:</u> Structure of various Proprotein Convertases and their characteristic domains along with bacterial subtilisin and yeast kexin or kex2 enzymes



This figure was adapted from reference (30) following slight modifications.

1.2.4. Chromosomal locations of the PC types and their null types in mice model:

Table 2: Summary of the chromosomal assignment of human PC genes and the phenotypes of various PC-null mice models

PC type	Human chromosome	Mouse chromosome	Mice null phenotype
PC1/PCSK1	5q15-q21	13	Dwarfism, Digestive abnormalities like chronic diarrhea and chronic nutrient mal-absorption, Pre and postnatal lethality In humans: severe early obesity, adenocortical, Insufficiency and hyper-proinsulinemia.
PC2/PCSK2	20p11.2	2	Retarded growth, resistance to obesity, Hypoglycemia, defects in processing of endocrine peptides
FURIN/ PCSK3	15q26.1	7	Embryonic lethal, Ventral closure, defects, Fail to undergo axial formation
PC4/PCSK4	19p13.3	10	Some embryonic lethality, reduced fertility
PC5/PCSK5	9q21.3	19	Mice die at embryonic day 4.5-7.5
PACE4/ PCSK6	9q21.3	19	Embryonic lethal
PC7/PCSK7	11q23.3	9	No reported phenotype/No apparent abnormal phenotype
SKI-1/S1P/ PCSK8	16q24	8	Embryonic lethal
NARC1/ PCSK9	1p32.3	4	Enhanced cholesterol uptake by liver In humans: Mutation-induced gain of function: hypercholesterolemia (autosomal dominant hypercholesterolemia). Mutation-induced loss of function: hypocholesterolemia

This table was adapted from references (16, 34-36) following slight modifications.

1.2.5. Tissue distribution and subcellular pathways of PCs:

The expression profile and tissue distribution of various PC-enzymes in mouse and other mammals including human in some cases have been well characterized. These were based on mRNA as well as protein expression and were assessed by histochemical, western blot and other biochemical studies. The overall PC-expression profiles are shown in **Table 3.**

Table 3: PCSK types, their tissue expression and biosynthesis:

PCSK/PC TYPE	TISSUE EXPRESSION	CELLULR PATHWAY
PCSK1/PC1	Neuroendocrine cells	TGN, secretory granules
PCSK2/PC2	Neuroendocrine cells	Secretory granules
PCSK3/FURIN	Ubiquitous	TGN/Endosomes/Cell surface
PCSK4/PC4	Testicular and Ovarian germ cells	TGN? Cell surface?
PCSK5/PC5	Widespread	TGN, secretory granules
PCSK6/ PACE4	Restricted, mainly digestive system	TGN/Endosomes/Cell surface
PCSK7/PC7	Widespread	TGN/Endosomes/Cell surface
PCSK8/ SKI-1/ SIP	Ubiquitous	TGN, secretary granules
PCSK9 / NARC-1	Liver (major), kidney, intestine, brain, heart	Secretory protein

This table was adapted from reference (16) following slight modification.

1.2.6 PCs linked to pathologies:

Various pathologies have been linked to PCs like tumorgenesis/cancer, lipid disorder, atherosclerosis, cardiovascular diseases, infectious disease, obesity and neurodegenerative disorders including Alzheimer's dementia, endocrine disorders and inflammation (36, 37). Tumorgenesis is shown to be enhanced by the over-expression of the convertases particularly furin as the processing and activation of various precursor protein molecules involved in tumorgenesis, angiogenesis, and metastasis is augmented by it (38-41)**.** It has also been shown that the bacterial toxins like *Botulinum* neurotoxin, *Bordetella* dermonecrotic toxin, and pore-forming toxins such as the aerolysin are activated to their mature forms from their inactive unprocessed forms by the PCs (42). Also acting in the similar fashion, many viruses like HIV-1, Hong Kong influenza virus (Bird flu), Ebola virus and severe acute respiratory syndrome (SARS) corona virus also acquire their infectious ability through PCs which process their surface glycoproteins which allow them to fuse with host membranes (43-46). Thus many PC inhibitors showed functional properties in fertilization, proliferation, obesity, tumorgenesis, diabetes or embryogenesis via their inhibitory action on PC activities. Hence, the designing of potent and specific non-peptide PC inhibitors as potential therapeutic agents has drawn considerable attention. This was followed by the successful development of several types of PC inhibitors (47, 48).

1.3 PROPROTEIN CONVERTASE 4 (PC4) OR PROPROTEIN CONVERTASE SUBTILISIN KEXIN4 (PCSK4)

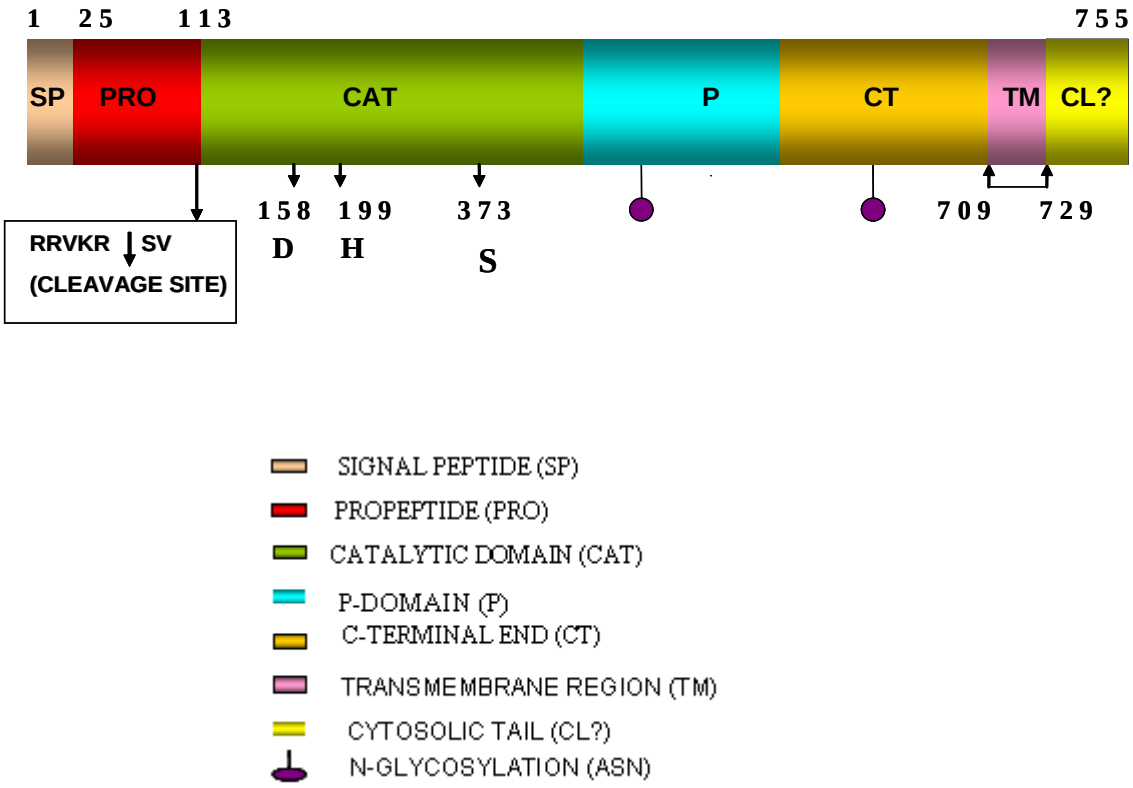
Proprotein convertase 4 (PC4) or proprotein convertase subtilisin/kexin type 4 (PCSK4) is one of the 7-members of kexin type proprotein convertases. It is mainly found

in the testicular germ cells and sperm surface, epidydimis in the male reproductive system and also detected in the ovary, placenta and oocytes in the female reproductive system (3, 24, 49-54). The amino acid sequence of PC4 and its expression pattern is found to be highly conserved in the tetrapods and in all mammalian species examined so far. Unlike other convertases, the expression of PC4 has been found to be highly restricted mainly within the reproductive systems and organs. This possibly suggests that during the evolution, this enzyme may have played a vital role in the reproductive process. PC4 like all PCs is a calcium dependent serine protease with pH optimum being within 7-7.4. PC4 cleaves precursor proteins to their active forms at the preferred cleavage motif **R/K/H-X-X/R/KR↓**, where X is any amino acid except Cys with **K-X-K-R↓** being the most preferred site. This activation of the precursor proteins associated with sperm maturation, migration, sperm-egg fusion where PC4 plays a significant role has been recognized and few substrates have already been either confirmed or proposed (24, 49).

PC4 is a 9-kb, 15 exon gene located on human chromosome number 19 (locus symbol: PCSK4) and mouse chromosome number 10 (locus symbol: Pcsk4) (55). It has been noted that sequentially and organizationally PC4 gene has higher homology towards the furin/PCSK3 gene compared to all the other genes of the PC-family. The major Pcsk4 mRNA has shown to be encoding a secretory precursor glycoprotein of 654 and 655 amino acids in rat and mouse respectively. This encoded polypeptide is made up of an amino-terminal hydrophobic signal peptide, a prodomain, a subtilisin like catalytic domain, a P domain and a carboxyl-terminal domain. Based on accumulated studies, the following motifs have been found to be well recognized by PC4: **R-X-X-R, R-X-R-R, R-X-K-R, K-X-K-R, R-X-X-R, K-X-X-R** with the last one being most selective for PC4 compared to

other kexin type PC-enzymes (56) (the crucial basic residues are shown with underlined). The prodomain of PC4 like other PCs, plays a vital role as an intramolecular chaperone and a regulator of the catalytic activity of the enzyme and its removal is a pre-requisite for the enzyme to be proteolytically active. The C-terminal P-domain of PC4 is implicated in the protein folding and conformation of the molecule and thus controls the ultimate protease activity of the enzyme. Two enzymatically active PC4 proteins have been identified *in vivo* and *ex vivo* conditions. These are the ~62-kDa full length and ~54 kDa C-terminal truncated secreted mature forms (24, 56, 57). The molecular structure of PC4 protein with its all characteristic domains is shown schematically in **Figure 3**.

Figure 3: Schematic representation of human PC4 protein and its various characteristic domains with amino acid positions as indicated



This figure was taken from reference (24) with modifications.

1.3.1 PC4 knock out mice

PC4 knock out mice were generated and studied for their characteristics and phenotypes to analyze the defects occurring in such mice and thus providing detail information about the important functions and biological role of PC4.

Table 4(A): Physiological defects of PC4 knock out male and female mice

MALE MICE	FEMALE MICE
Infertility (56, 58)	Subfertility (58)
Accelerated sperm capacitation (56)	Reduced folliculogenesis (52)
Susceptibility to premature sperm acrosome reaction (3, 56)	Reduced ovulation (52)
Reduced sperm hyper-activation (58)	-
Inability to fertilize oocytes (3, 56)	-

Table 4(B): Molecular defects in PC4 knock out male and female mice

MALE MICE	FEMALE MICE
Reduced levels of spermatid PACAP-regulated MAPKs (59)	Lack of ovarian PACAP peptides (60, 61)
Lack of testicular PACAP peptides (60, 61)	-

Tables 4 A & B were adapted from reference (24) following slight modifications.

1.3.2 Physiological substrates of PC4:

The identification of physiological substrates of PC4 was made possible with the availability of PC4-knock out mice by examining the protein deficiency caused by its knock out and lack of processing of those substrates in the PC4 null mice. It is known that the male germ cells in the testis biosynthesize many precursor proteins that become active by endoproteolysis carried out by PCs particularly PC4. The followings are the list of such proteins: pro-pituitary adenylate cyclase-activating peptide (**proPACAP**), pro-insulin-like growth factor1 (**proIGF1**), pro-insulin-like growth factor2 (**proIGF2**) and their receptors, **pro-enkephalin**, **proNerve growth factor**, members of the disintegrin-and-metalloproteinase (**ADAM**) family of proteins like **alpha-fertilin** (**ADAM1**), **beta-fertilin** (**ADAM2**), **cyritestin** (**ADAM3**) and **hepatocyte growth factor receptor** (19, 24, 53).

(a) <u>ProPACAP</u>

Pro-pituitary adenylate cyclase-activating polypeptide (proPACAP) is known to be a member of the vasoactive intestinal peptide secretin/GH releasing factor family of peptides (60) and it is the only confirmed PC4 substrate so far identified in male germ cells . The biologically active form of PACAP exists in two amidated forms which are PACAP 27 and PACAP 38 (62-64) , of which PACAP 38 is considered to be the dominant form in the tissues. PACAP was originally isolated from hypothalamus, extra-hypothalamic regions of the brain (65) and gut associated with growth and metabolism (66). In testis, it is located particularly in the cap of the spermatids and acrosome phases but not at the subsequent or prior stages in mature spermatozoa, sertoli cells and leydig cells (67). This suggested that testicular PACAP is transiently expressed in testicular germ cells and participates in the regulation of spermatogenesis. Also it is known that, steroidogenesis in leydig cells and

protein synthesis in spermatocytes have been shown to be stimulated by PACAPs in *ex vivo* (68, 69). PACAP is also shown to be present in the rat epididymis and mostly in caput and corpus region as compared to the cauda region indicating that it may exhibit differential expression along the epididymal duct and thus accordingly regulating epididymal and sperm functions along the epididymis (70). These two forms of PACAP were not detected in PC4-null mice testis thus confirming that PC4 plays a vital role in its processing. PACAP-null mice results in characteristics like postnatal death associated with dysfunction of lipid and carbohydrate metabolism as well as neuronal development (66). PACAP has many functional roles to play in the body but we would mainly focus on its relevance to the reproductive system. Studies have shown that PACAP deficiency delays testicular deterioration associated with aging. Data implicated that the primary cause behind infertility may not be the lack of PACAPs in PC4-null mice, thus leaving it under speculation that these mice are protected against age-induced testicular degeneration due to the lack of PACAPs (24). Also it is observed that there is a 20-fold increase (60) in testicular proPACAP mRNA due to PC4 deficiency indicating that a feedback stimulation of transcription of its gene or stabilization of its mRNA may take place due to lack of PACAPs. In PC1 and PC2 null mice this feedback transcriptional up-regulation of substrate gene expression has been seen for hypothalamic growth-releasing hormone and pancreatic islet production respectively (60, 61). This may indicate an obligatory and exclusive enzymatic link between a proprotein convertase and its respective physiologically most relevant substrate (24) . Altered signaling in the PC4-null mice is implicated by the decrease in the expression of PACAP-regulated mitogen activated protein kinase (MAPK) in the spermatids (59) but the mechanism of their action is not fully understood. Increased production of

cAMP (an early second messenger) is induced by PACAPs in the cascade signaling events leading to gene expression or cellular metabolism (71). More work needs to be done in this field to understand the mechanism regarding how the fertilization defects are caused due to the reduced PACAP-regulated MAPKs in PC4-null mice (72).

PC4 null female mice were shown to be devoid of PACAP peptides. These mice also show signs of impaired folliculogenesis. However they show normal ovarian morphology and ovulation although the mating success is observed to be one-fifth in comparison to normal female mice due to an implantation defect whose cause is not yet understood. A significant increase in the proPACAP mRNA was observed in PC4 null female mice suggesting that it is indeed a substrate of PC4 (52, 60, 61, 73). This is similar to the observation with the testis of PC4 null male mice. It has also been demonstrated that the cumulus cells release a soluble factor in response to PACAP which most likely promotes sperm motility and the acrosome reaction thus leading to fertilization (74). PACAP has been shown to be linked to maintaining pregnancy particularly at the later stage in rats. This indicated its important role as a local ovary regulator. It was also demonstrated that during the onset of pregnancy the PACAP level increases suddenly in the corpus luteum and then decreases (75).

(b) proADAMS

ProADAMs belong to the type-1 membrane proteins that are characterized by the presence of multiple domains. These represent the most abundant sperm proteins that play a vital role in sperm-egg interaction following their proteolytic cleavages to release smaller bioactive forms (76-79). Several proADAM proteins have been reported. These are ADAM1 to ADAM15 (79) Among them ADAM1 (fertilin-α), ADAM2 (fertilin-β) and

ADAM3 (cyritestin) are the most important family members that were well studied and found to be potential substrates of PC4. A number of potential cleavage sites consistent with PC4 motif have been found in their protein sequences (76, 80-82). Studies suggest that fertilin-α and fertilin-β are most likely involved in sperm-egg binding and fusion as well as spermatogenesis (83-85). In addition it was shown that synthetic peptides comprising these cleavage sites were found to be easily cleaved *in vitro* by PC4 enzyme. It may be speculated from the studies that PC4-mediated cleavage could expose the disintegrin domain of these proADAMs and thus contribute to its functional activity. It was also shown that these ADAMs also undergo a cascade of testicular and post-testicular proteolytic cleavages to their acrosome bound forms (86). Inactivation of ADAM2 and ADAM3 genes has been shown to cause infertility phenotype in male mouse which exhibited reduced sperm binding ability to the zona pellucida (87). Also inactivation of the above two genes have shown impaired sperm aggregation, thus indicating their requirements in normal sperm association (88). Changes in the expression of processed sperm proteins such as ADAM1, ADAM2 and ADAM3 have been noted in PC4-null mice (15). However, sufficient and detail studies are lacking to support conclusively that the lack of PC4 may cause impaired processing of testicular proADAMs. However studies have shown that strict stoichiometry of interacting proteins may be essential for the fertilization competence of sperm cells (24).

(c) ProIGFs

In males proIGF 1 and 2 (pro-insulin-like growth factors 1 and 2) are found in the leydig cells and produced as inactive forms of 153aa (amino acid) and 180aa long proteins. However they undergo proteolytic cleavages at **KSAR^{71}↓SV** and **KSAR^{68}↓DV** respectively to produce the functionally active 70aa and 67aa long IGF1 and 2 (after trimming of **R** by

carboxy peptidase)**.** ProIGF1 and 2 are processed by the PCs mainly PC1 and PC2 in the endocrine cells but in the reproductive system it is processed by PC4 (53, 89, 90).

In females IGF1 is known to be important for normal embryonic development and post-natal growth. Its expression has been found throughout these phases of embryo development and any homozygous abnormality has been found to be fatal causing death within three days. Hence regulation of proIGF1 processing mediated by PC4 in reproductive tract is extremely crucial for embryonic growth (90).

IGF2 is expressed in the fetus and placenta. It also plays a vital role in regulation of the feto-placental growth (66). In mouse, placental trophoblast cell migration is aided by PC4 by converting proIFG2 to its active form. Overall IGF2 is known to be present in three forms, which are inactive proIGF2 (24kda), mature or active IGF2 (8kDa) and its intermediate forms which are collectively called as big IGF2 and their sizes is around 16 kDa containing 1-87 and 1-104 segments (53). Lack of placental IGF2 could lead to reduction in the placental diffusion exchange of nutrients due to the disruption of placenta-specific promoter of the IGF2 gene (53, 89-91). Thus it is speculated that implantation and placental defects could be one of the factors for causing subfertility in PC4-null female mice. Defect of proIGF2 processing in human placenta have been shown to be associated with intrauterine growth restriction (IUGR) where elevated level of circulating unprocessed proIGF2 has been noted. In rats, it was shown that conversion of proIGF2 to its active form increases during progression of pregnancy. All these factors contribute to the notion that PC4 is an important processing enzyme for IGF2 precursor.

1.3.3 Enzymatic property of PC4

Enzymatic or protease activity of PC4 *in vitro* has been reported previously in culture medium of rat epithelial BSC40 cells infected with a recombinant vaccinia virus carrying rPC4A-cDNA. The optimal activity was observed against the fluorogenic peptides pERTKR-MCA and Boc-RVRR-MCA. In presence of 2mM Ca^{+2} at nearly neutral pH of 7.4 PC4 exhibited highest protease activity (92). Using *Leishmania torentolea* expression system, it has been shown that rec-rPC4 is enzymatically highly active (57). Compared to other PCs, PC4 has been shown to cleave the sequence **K**XX**R**↓ in a more efficient manner (93). Using model peptides and recombinant proIGF2 protein, it was established that PC4 was the right enzyme for processing proIGF2 (53).

1.3.4 PC4 Inhibitors

A considerable interest has been developed recently in studying potential therapeutic applications of PC4-inhibitors as non-hormonal contraceptive agents. This arises from the role of these inhibitors in blocking the processing of the precursor proteins involved in sperm maturation and sperm egg-interaction. A patent on this has already been approved on the *"Promotion of inhibitors of spermatogenesis by the PC4 inhibitors"* in USA (Patent 5763402) **(*www.patentstorm.us/patents/5763402/claims.html).*** In recent years many PC4-inhibitors have been reported in the literature by various research groups including our laboratory. These comprise the synthetic derivatives such as peptidyl semicarbazones, RRKR-sc and RSKR-sc (sc = semicarbazone or -CH=N-NH-$CONH_2$) which inhibited PC4 with K_i ~0.75 and 11.4 μM respectively when measured against pERTKR-MCA substrate (48, 53, 57, 93, 94). Several peptide chloromethyl ketones (cmk) such as decanoyl (dec)-RVKR-cmk have been shown as general PC-inhibitors including PC4 (48). There are also

some PC4 prodomain derived peptides from the activation sites namely proPC4$^{75\text{-}90}$ and proPC4$^{75\text{-}84}$ which act as PC4 inhibitors with K_i in low μM range (53, 57, 92). Other inhibitors include non-peptide molecules such as flavonoid compounds including baicalein and oroxylin A and their glycosides (95) **.** These compounds inhibit PC4 activity with IC_{50} values in low μM range. All the above compounds except peptidyl-cmks are irreversible inhibitors of PCs which may be competitive, noncompetitive or mixed or un-competitive types. The nature of such inhibitions can be evaluated using various types of plots such as Dixon and Lineweaver-Burk (96)

Although "serpins" inhibiting furin activity are known in the literature (27) no such inhibitors have so far been reported for PC4 activity. Therefore search is on to find out serpin like inhibitor for PC4 enzyme.

1.4 SERPINS

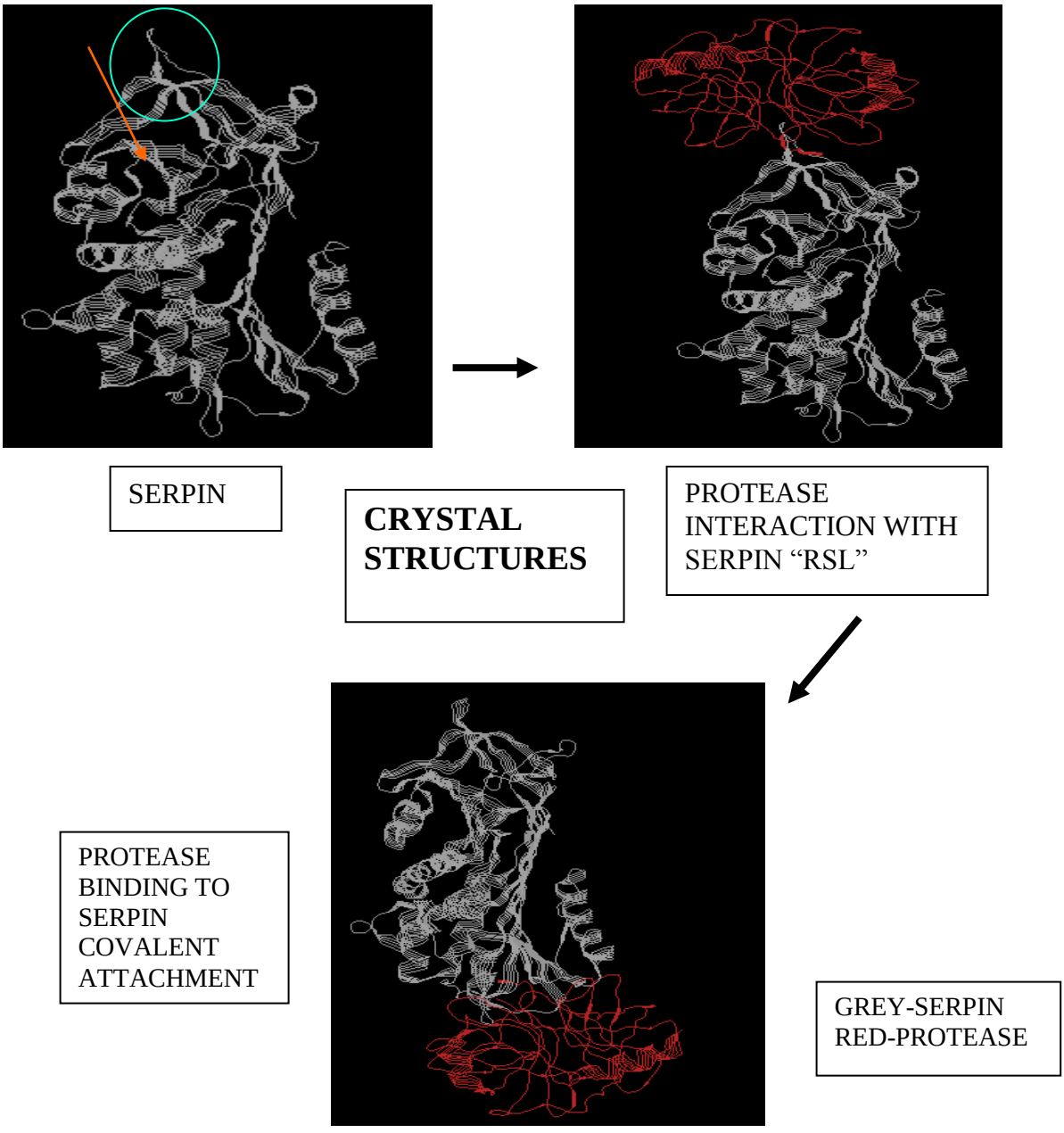
Serine protease inhibitor (**Serpin**) belongs to a family of protease inhibitors that are found to be in the physiological system and inhibit mostly serine proteases and have conserved structures. Each serpin plays important but selective protease inhibitory role *in vivo* towards specific serine protease which may be crucial in preventing coagulation, aggregation, inflammation, apoptosis and many other cellular or biochemical events associated with diseases. The mechanism of inhibitory action of serpins has been well studied. It is now known that this activity resides with the presence of a "**Reactive Site Loop**" **("RSL")** structure with a cleavage site all important for activity. This '**RSL**' loop is known to be flexible and possesses a β-pleated sheet structure which helps the serpin to trap the enzyme via covalent bond formation leading to a conformational change and thus inhibiting the activity of the enzyme (97-99). First a recognition and interaction takes place between the

serpin and the enzyme. This is followed by a peptide bond cleavage within "**RSL"** domain of the serpin and finally the complete insertion of this "**RSL"** into the β-pleated sheet structure occurs through formation of covalent bond with concomitant loss of protease activity of the enzyme. Thus serpins behave as substrates for the proteases upon slow insertion of the "**RSL**" loop structure into the β-sheet and also an inhibitor. It leads to the formation of a strong, stable and irreversible complex between the enzyme and the serpin. Serpins are also termed as suicide inhibitors (100, 101). Presence of single or multiple disulfide bridges within "**RSL**" of each serpin is a characteristic feature which is also critical for enzyme inhibitory property. Mug bean trypsin inhibitor-1 and sunflower trypsin inhibitor-1 are two examples of serpins that are known to have two and one disulfide bridges respectively. Among other serpin protease inhibitors, one can mention about snake venom derived polypeptides "**flavoridin**" and "**echistatin**" which belong to the disintegrin family and are shown to possess disulfide bridges (102, 103).

Most of the serpins as their name suggests inhibit the serine proteases (eg: trypsin and chymotrypsin) but occasionally some serpins also inhibit cysteine proteases (eg: cystatins) (104-106). Within PC-enzyme superfamily, a serpin has been identified for furin. Besides, specific binding/partner proteins like SAAS for PC1 and 7B2 for PC2 have been also reported (107, 108). These are granin family of proteins but behave to some extent like the serpins So far, no physiological serpin for PC4 present in male/female reproductive system has been reported although one serpin, CRES has been shown to be an inhibitor of PC2 in the endocrine system (22). Many serpins have been detected in reproductive system like **PCI** (Protein C inhibitor), **hserb6**, **mserpb6A**, **mserpinb6b**, **mserpinb6c**, **testatin**, **cystatin T, hongrES1**, **eppin** (epididymal protease inhibitor) and **CRES (cystatin related**

epididymal spermatogenic). Among these hongrES1, eppin and CRES are known to be present in the epididymis where PC4 is also present. Eppin has been shown to have inhibitory effect towards chymotrypsin-like PSA (prostate specific antigen – a serine protease) and HongrES1 has been shown to have regulating effects on capacitation and male fertility and possibly inhibits trypsin like protease (109-115). No studies have been reported regarding the partner enzyme of CRES in the reproductive tract although as mentioned above in the endocrine system, the partner enzyme of CRES is PC2 enzyme. Thus studying the expression profiles and pathways of CRES and PC4, we suspect that PC4 is the partner enzyme of CRES in the reproductive system and thus regulates its protease activity during various fertilization events taking place in epididymis and beyond. In fact another serpin called d (drsophila)-SPN4 was reported as a potent inhibitor of furin (116, 117). Interestingly even before a bioengineered serpin called α1-PDX which was derived from α1-antitrypsin via double mutations (i.e. **AIPM** to **RIPR**) was shown to be an extremely potent furin inhibitor (K_i in pM range) (118).

<u>Figure 4</u>: Mechanism of serpin-protease interaction leading to complex formation and enzyme inhibition. Serpins are physiological inhibitors of mostly serine proteases and contain a reactive site loop ("RSL") with a cleavage site that is crucial for its activity. "RSL" allows its binding to the protease through kinetically trapped covalent attachment.

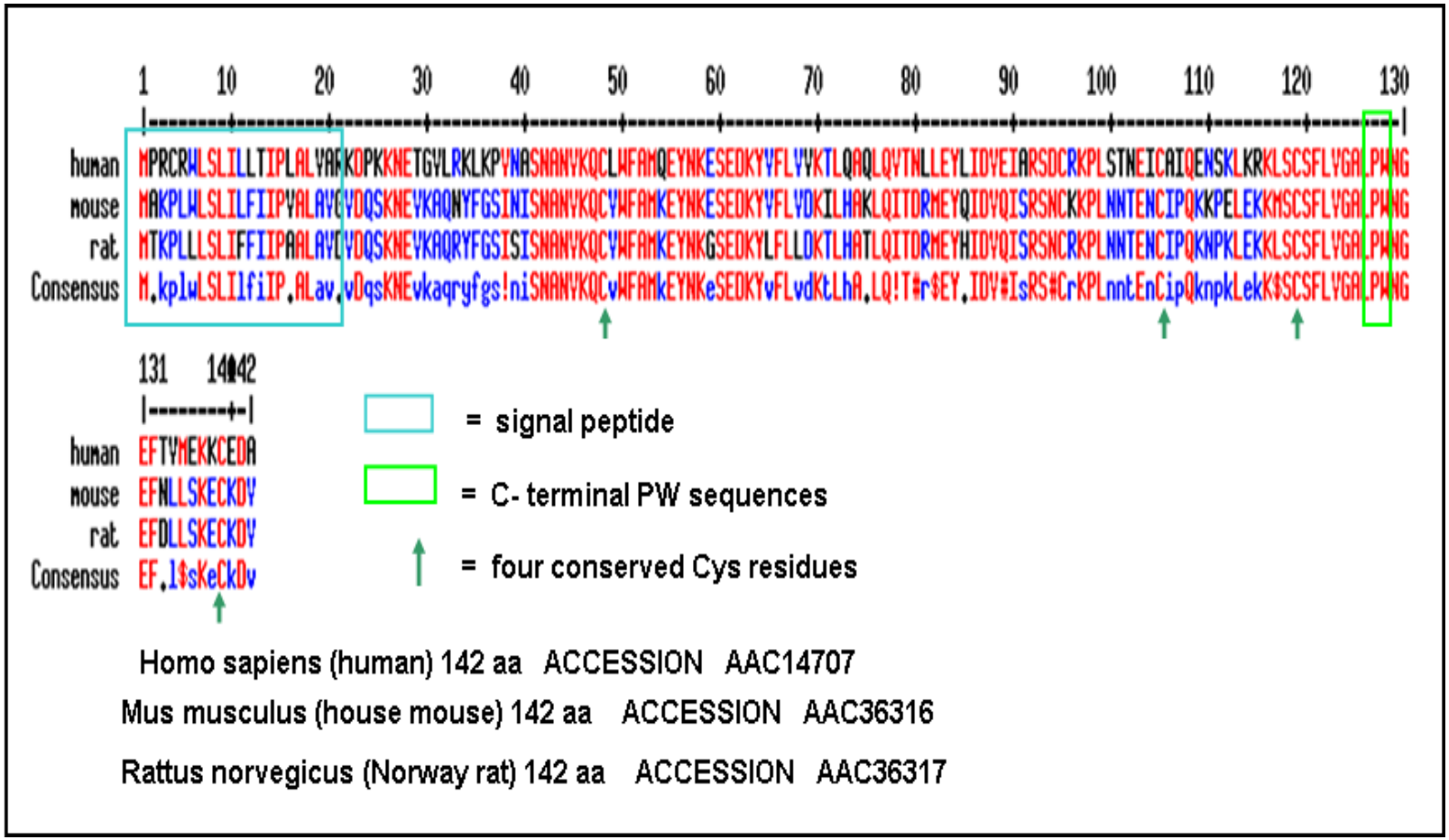


This figure is adapted from the reference (119, 120) following slight modifications.

1.5 CRES

CRES (cystatin regulated epididymal spermatogenic) is a mammalian protein that belongs to serpin family and more specifically to the cystatin super family. The cystatin superfamily is divided into three classes: *Stefins (Family 1), Cystatins (Family 2), and Kininogens (Family 3).* The cystatin superfamily to which CRES belongs are usually inhibitors of cysteine proteases like cathepsins B, S, H, L and also papain (121) but CRES is different from regular cystatins as it is known to inhibit substrate-specific serine proteases (e.g. PC2) instead of cysteine proteases even though it is conformationally similar to cystatins. It is therefore called as **"Cross Class Inhibitor"** (22). Characteristics of cystatins are that they contain a pentapeptide **'QXVXG" (X = any amino acid)** motif and a C-terminal **"PW"** dipeptide domain along with four conserved **"Cys"** residues with specific S-S bridges all of which are crucial for its enzyme inhibitory function. Although CRES contains "PW" sequence and the conserved Cys residues (22, 122, 123), it lacks the consensus penta-peptide motif. However it shows good sequence homology and conserved motifs that are characteristic of cystatins (family 2). Human (h) CRES was shown to have 30% sequence homology with h-cystatin C which is another family 2 cystatin member (122). Disulfide bridge formation in CRES among its 4 cysteine residues was only proposed as similar to those present in family 2 cystatins but no confirmation study is available in the literature (123, 124) .Alignment of the CRES sequence from various organisms indicated a good sequence homology, presence of "PW" sequence and the four conserved Cys residues which were shown in **Figure 5**.

<u>Figure 5</u>: Amino acid alignment of CRES found in various species using the proteomics tool Multalin version 5.4.1 for multiple sequence alignment available at the website www.expasy.ch



Consensus levels: high=90% low=50%

1.5.1 CRES expression profile in tissues and cell lines:

CRES has been found to be highly tissue-specific and observed to be present in post-meiotic germ cells, elongated spermatids in testis, sperm acrosome and in the proximal region of the epididymal caput (125, 126). This tissue specific expression indicates that it could be involved in the reproduction process together with some partner proteases since serpins are known to function together with respective enzyms. Due to expression similarities of CRES and PC4, we speculated that the partner enzyme of CRES could be PC4. In support of this idea, we noted that in the endocrine system, CRES has already been shown to be a potent inhibitor of PC2 (22) but in the reproductive system, its partner enzyme has not been determined, thus making us to speculate that it could be PC4. **Also amino acid sequence alignments of CRES from various species showed the presence of four conserved Cys-residues within the "RSL" loop that also comprises a unique basic peptide sequence for the recognition by PC4.**

Since CRES is known to be located in the acrosome, researchers became interested to examine the role of CRES in fertilization events. Acrosome reaction aids in the release of proteolytic enzyme including serine proteases like proacrosin, metalloprotease, cathepsins and PC4 into the extracellular environment, and many of these enzymes are known to be involved in sperm-egg interaction via the digestion of zona pellucida. In mice, it was also observed that CRES is released from the acrosome when acrosome reaction inducing agent was added to the capacitated spermatozoa. However in human, CRES was not found to be released from the acrosome but has moved to the sperm surface having full access to the extracellular environment. These studies seemed to throw some light on the species

differences in CRES localization (111, 126). Hence the exact function of CRES remains unknown and studying its effect on PC4 may provide us additional information on function.

1.5.2 CRES Aggregation:

It has been recently reported that a few serpins including cystatin family members have the tendency to oligomerize (127, 128) and this includes CRES. The presence of multiple forms of CRES in epididymal luminal fluid has been demonstrated including SDS-sensitive and SDS-resistant high molecular mass complexes. Also studies indicate that CRES is a substrate for transglutaminase and also it was observed that the endogenous transglutaminase activity in epididymal lumen catalyzes the formation of SDS-resistant CRES complex. Similar to other amyloidogenic proteins, CRES was shown to be able to adopt oligomeric and fibrilar structures during its aggregation indicating that CRES has the potential to form amyloid in the epididymal lumen. Even cystatin C has been shown to form amyloid deposits associated with cerebral amyloid angiopathy (127).

1.5.3 CRES null mice study:

As mentioned earlier CRES (a product of gene *Cst8*) is a serpin type protease inhibitor which is present in epididymis and spermatozoa suggesting its possible role in sperm maturation and fertilization. In order to study this, mice lacking *Cst8* gene were generated and a comparative study between *Cst8*+/+ (control wild type) and *Cst8-/- (knock out)* mice was recently reported in the literature (129). These studies revealed several phenotypic differences that included *(i) decrease in the size of the testis, (ii) reduction in the number of spermatozoa in the epididymal cauda, (iii) disruption in the sertoli cells resulting in premature sloughing of germ cells and hence reduced spermatozoa in epididymal cauda, (iv) defects in the binding ability to the egg zona pellucida, (v) reduced amount of acrosome*

reaction and capacitation and thus impaired fertilization. It was previously shown that tyrosine phosphorylation of sperm proteins is another important step during capacitation. This is a downstream event of signaling pathways mediated by cAMP/PKA which is also crucial for sperm capacitation. This led to impaired capacitation and fertilization in CRES null mice. Addition of dibutyryl cAMP and 3-isobutyl-1-methylxanthine (IBMX) to the *Cst8-/-* mice was able to rescue the defects observed in the knock out mice following stimulated protein tyrosine phosphorylation. Furthermore it was also able to enhance the binding ability of sperm to the egg zona pellucida which is comparable to the wild type condition (129).

The latest study on CRES knock out mice also revealed several characteristic defects such as (i) significant reductions in tubular, epithelial and luminal areas in the testes and epididymis, (ii) vacuoalated seminiferous epithelium, (iii) degenerating germ cells, (iv) alterations in ectoplasmic specializations, (v) abnormally shaped sperm heads and tails, (vi) irregularly shaped lysosomes and nuclei, suggestive of disrupted functions by all these organelles (130) .

An earlier study in this context (3, 56) is useful since it indicated an increased tyrosine phosphorylation of sperm proteins in PC4 knock out mice which is in agreement with the decreased tyrosine phosphorylation observed above for CRES knock out mice. **This is likely possible if PC4 is inhibited by CRES – the principal hypothesis of my thesis. This is further reinforced by the fertilization data as indicated above with CRES knock animals that provide the necessary support for the notion that PC4 may be the partner enzyme of CRES (131).**

2.0 OBJECTIVES, HYPOTHESIS AND RATIONALE OF THE REASEARCH PROJECT

2.1 Objectives:

Aim 1.Produce and purify rec-PC4 enzyme.

Aim 2. Generate, purify and characterize rec-CRES protein.

Aim 3. Examine and compare PC4-like activity in various regions of mouse epididymis.

Aim 4. Examine the inhibitory effect of CRES on PC4 in *in vitro* and *ex vivo* conditions.

2.2. Hypothesis:

CRES found in epididymal fluid and sperm acrosome inhibits the activity of sperm protease PC4 and regulates maturation of sperm proteins relevant to fertilization.

2.3 Rationale:

Both CRES and PC4 are known to be present in the epididymal fluid. Previous studies have shown that CRES is expressed more in the caput than in the other epididymal regions. However the distribution profile of PC4 in various epididymal compartments is not fully known. This prompted us to examine PC4 activity & protein profiles in various epididymal fluids and its potential regulation by CRES which ultimately may affect fertilization event.

3. MATERIALS AND METHODS

3.1 Materials

Leishmania tarentolae gene expression kits were obtained from Jena Bioscience GmbH, (http://www.jenabioscience.com/cms/en/1/browse/1119_lexsy_literature.html), Germany. DEAE (Di-Ethyl Amino Ethyl)–sepharose and Arg–sepharose 4B resins were both purchased from Amersham Bioscience, Piscataway, NJ, USA while the polyclonal PC4 antibody was obtained from Alexis Biochemicals, catalogue number: ALX-210-218-R050 (San Diego, CA, USA) and it was raised against the PC4 fragment $mPC4^{23-393}$ that contains its catalytic and middle domains (please note that there is 99.5% homology of this domain of mPC4 with the corresponding rPC4 sequence) (57)**.** Reagents for western blot and SDS–PAGE analyses were purchased from BIO-RAD Labs (Hercules, CA, USA). Chemiluminescence reagents (PerkinElmer LAS Inc, USA) were used for detection of immuno-reactive bands. Images were captured using Kodak X-OMAT Blue autoradiography film (PerkinElmer LAS Inc., USA). Matrix-Assisted Laser Desorption time of flight (MALDI-TOF) and Surface Enhanced Laser Desorption Ionization time-of-flight (SELDI-TOF) MS were recorded using Voyageur reflector (Framingham, MA, USA) and Ciphergen (Fremont, CA, USA) instruments respectively. The corresponding mass spectra plates were purchased from the respective mentioned companies. The matrices for mass spectra 1-Cyano 4-hydroxy cinnamic acid (CHCA) and Sinapic acid (SPA) were purchased from Sigma Chemical Company, USA. The peptidyl-MCA fluorogenic substrate was purchased from Bachem Inc. (King of Prussia, PA, USA). Rec-mCRES plasmid was kindly provided as a gift by **Dr Gail A. Cornwall**. The plasmid construct was expressed in *Escherichia coli* (*E.coli)* with the help our collaborator **Dr. Qing Qiu** from OHRI. All

reversed phase high performance layer chromatography (RP-HPLC) purification of protein samples were performed using Varian HPLC instrument (ProStar, Varian, USA).Syringe filters with 32 mm diameter and 0.8/0.2 µm membrane (Life sciences, USA) were used for the filtering of the crude CRES solution.

3.2. Methods

3.2.1 Rec-rPC4 production

The rec-rPC4 production and purification was carried out based on the following available protocol (57). The cDNA plasmid of full length (FL) rPC4 sequence with a V5 tag at the C-terminus, was transformed into the *Escherichia Coli* for amplification and was then stably transfected by electroporation into the *Leishmania tarentolae* system. Brain heart infusion medium (BHI) at a concentration of 35 g/l was used to culture the recombinant strains harboring the construct. The medium was supplemented with 0.25% hemin in the presence of penicillin (10,000 units), streptomycin (10,000µg/ml) and nourseothricin (100 µg/ml) and was cultured at 26^0C in the dark with shaking at 90 rpm. When the cells reached full confluency as determined by the measurement of optical density value (1.5-2.0) and wave length 600 nm, then the culture medium containing expressed rec-rPC4 protein was collected. This consisted of nearly 1 x 10^8 cells/ml of the culture medium. The collected medium was centrifuged to obtain about 1000 ml of the supernatant which was subsequently concentrated ~40 fold to obtain 25 ml of concentrated media using the centricon filter (cut-off MW = 10kDa). All the operations were conducted at a temperature at 4^0C or below. Membrane bound PC4 was extracted using the buffer 25mM Tris + 25mM MES + 2.5mM $CaCl_2$, pH 7.4 (named as **Buffer A**). For purification of crude rec-rPC4, DEAE-agarose column (size 15 mm long x 2.5 mm wide) was used. The column was first pre-equilibrated

with Buffer A. Crude rec-rPC4 present in cell lysate (~10ml) was then loaded on to the column. Following loading of the sample the column was washed with Buffer A as part of the chromatography to remove all the proteins that are not bound to the solid matrix (i.e, resin). Ninety-five fractions each of about 1ml in size were collected in this stage. This was followed by a step-wise elution with the Buffer-A containing increasing concentrations of NaCl starting with 30mM NaCl (#1-85 collected), then 60mM NaCl (#1-60 collected), next 70mM NaCl (#1-74 collected) and finally 200mM NaCl (#1-67 collected) and the Buffer A wash itself consisted of #1-95. Thus a total of 381 (85+60+74+67+95) fractions were collected in various salt pools (i.e. various elution buffers) as described above. Each fraction collected was 0.8-1ml in volume depending on the flow rate which slightly changed during the chromatography run. The fractions were collected in small eppendorfs using the fraction collector. Each fraction was subjected to protein and PC4-protease activity assays.

3.2.2 Rec-ratPC4 purification

3.2.2.1 Enzyme and protein assays

The enzyme activity assay was carried out using the fluorogenic substrate Boc-RVRR-MCA (50μM final concentration) in the 96 well plates (Dynatech, Millipore, USA) using the Fluorometer (Spectra Max Gemini XS, Molecular Devices, USA) whereas the protein assay was conducted using optical density method. Enzyme assays were conducted using the 96 well plates .Each sample was taken in equal aliquots in these wells to which equal volume of the fluorogenic substrate and Buffer A were added and mixed well. Then the fluorescence released at various times was measured using a fluorometer instrument at excitation and emission wave lengths fixed at 370 and 460nm respectively.

Protein assays were conducted to measure the total amount of protein in each fraction by taking the same small aliquot for all 381 fractions. Each sample fraction was thoroughly mixed with bicinchoninic acid (BCA) reagent (Pierce Chemicals, USA) and Optical density (OD) value was measured. Multiskan® Spectrum (Thermo) plate reader was used for this measurement at wavelength 562 nm. The amount of protein was calculated using a standard curve which was plotted using various concentrations ranging from 0-2000 µg /ml of analytical grade bovine serum albumin (BSA) protein.

After obtaining the enzyme activity assay and protein content data of each of all the chromatographic fractions, graphs were plotted to compute the results. Comparison of all the data allowed the combination of specific fractions. Those fractions exhibiting high level of PC4 protease activity and low protein content were of particular interest as it signifies high degree of purification. We obtained 28 such pools from the total 381 fractions after the combination. These pools were then subjected to further analysis by enzyme and protein assays followed by SDS PAGE and Western blot analysis. Following all these experiments, pool#15 was selected among all the twenty-eight pools as the most potential pool to be used for all the following experiments where PC4 was required. This was done to maintain the consistency of the source of enzyme throughout the project.

3.2.2.2 SDS-PAGE analysis of recrPC4

Tris-glycine gels with 12% resolving and 4% stacking phases were used for SDS PAGE analysis. Denaturation of the protein samples were achieved by heating to 95^0C for 10min with sample loading buffer contain 1% SDS, 10% glycerol, 10mM Tris-HCl and 5% of reducing agent dithiothreitol (DTT) at pH 6.8 from Sigma Aldrich Chemical Company. The gel was immersed in 1 x running buffer (500mM Tris, 0.1% SDS), and was run at 150V for

about 1 h. The See Blue Plus2 pre-stained standard (Invitrogen, USA) was used as marker. When the separation was finished, coomassie dye staining solution (Bio-Rad, USA) was added to the gel for visualization of protein bands.

3.2.2.3 Western blot analysis

Immediately after SDS-PAGE separation, the proteins were transferred from the gel to the nitrocellulose membrane electrophoretically. The transfer proceeded overnight at 20V in ~4L of transfer buffer (12.1g Tris-base + 57.8g glycine, in 800ml methanol). The membranes were blocked for 1 h in TBS-Tween 20 (TBS-T) containing 5% skimmed milk. Afterward, the membranes were incubated with primary anti-PC4 antibody (Alexis biochemicals, sequenced from $PC4^{23-393}$) diluted 1:500 for 1.5 h and then with secondary antibody: goat anti-rabbit IgG (1:10,000) for 45 min (Santa Cruz, USA). Antibodies were removed by subsequent washing in PBS-T (Phosphate Buffered Saline-Tween 20). Chemi-luminescence reagents (PerkinElmer LAS Inc, USA) were used for detection of immunoreactive bands. Images were then captured using Kodak autoradiography film (PerkinElmer LAS Inc., USA).

3.2.3 Production of rec mCRES protein

The cDNA encoding form mouse-CRES protein (lacking the signal-peptide) was cloned into a Champion pET directional pET101/TOPO vector (Invitrogen, USA), which contained a T7 promoter at 5' end as well as a V_5-epitope and His $_6$ tag at 3' end. This vector was generated and provided by **Dr. Qing Qiu**. The mouse CRES cDNA used in the above construct, as mentioned previously was a gift from **Dr. Gail A. Cornwall**. The plasmid was transformed into the Star BL21 One Shot Competent *E.coli* cells (Invitrogen, USA) upon heat induction. The cells were grown first in a smaller 5ml LB (Luria Bertani) culture medium and later

scaled to 2L. Cells reached optimal OD after 4 h of growth and IPTG (isopropyl β-D-1-thiogalactopyranoside) was added to induce protein expression. Cells were collected by centrifugation after 1.5-h induction and were stored at -80°C. The next day the cells were dissolved in purification buffer (250mM NaH_2PO_4, pH 8) and lysed by repeated sonication in the presence of lysozyme. The lysate obtained was centrifuged, and the supernatant containing rec-mCRES protein was collected for characterization by SDS-PAGE and SELDI-TOF mass spectrometry analyses.

3.2.3.1 Purification of rec-mCRES by RP-HPLC

Purification of crude rec-mCRES as obtained above from cell extract was accomplished by RP-HPLC using C_4-semi-preparative and analytical columns. RP-HPLC has been proved to be an effective method of purification of PCs in the past (21, 44). The two solvent systems used for the elution comprised of double distilled water containing 0.1% (v/v) TFA (Trifluoroacetic acid) as the aqueous phase (**Solvent A**) and 0.1% (v/v) TFA containing acetonitrile (ACN) as the organic phase (**Solvent B**). During HPLC purification, proteins were separated using a linear gradient of 1% increase in solvent-B/min from initial 10% to 60% solvent B on a semi-prep C4 column (Phenomenex, dimensions 250 x 10 mm) followed by an analytical C4-column (Phenomenex, dimension 250 x 4.60 mm). Fractions were collected and analyzed as the elution was monitored using an on-line UV detection system with wavelength fixed at 225nm. All the major peaks were then analyzed by mass spectrometry for their correct identifications.

3.2.3.2 Characterization of rec-mCRES by SELDI-TOF Mass Spectrometry

The molecular weight was determined by using SELDI-TOF mass spectrometry (PE-Biosystems and Ciphergen, Fremont, CA, USA). A reflector positive ion mode method was

selected. The energy absorbing matrix solution used for the mass spectrometry was prepared from 10 mg sinapic acid dissolved in 1ml of 1:1 (v/v) ACN/H_2O. A small aliquot of the sample was then spotted on the SELDI-TOF chip which was then used for the MS analysis.

3.2.3.3 Characterization of rec mCRES by Proteomic analysis

Following identification by SELDI-MS, rec-mCRES was further confirmed by digestion of CRES protein by Proteinase Lys-C followed by mass spectrometric analysis using SELDI-MS. Small aliquot of Lys-C was added to the rec-mCRES and mixed well. This mixture was then further analyzed by spotting on the SELDI-tof chip with the following time intervals like 0h, 1h, 2h and 4h of their digestion. All proteomic analysis of the digest peaks were achieved by using the website, www.expasy.ch.The various digest peaks and their measured molecular weights by mass spectrometry were in consisted with those expected from the known amino acid sequence of rec-mCRES protein.

3.2.4 PC4 activity assay

A C-terminally truncated secreted soluble rec-rPC4 was produced previously in our lab as described before using the *leishmania tareontolae* expression system. Herein the enzyme has been purified in its active form via DEAE ion-exchange chromatography and used for all enzyme work reported in this thesis. Enzyme activity assays were carried out at room temperature in 96-well flat-bottom black plates using a spectro-fluorometer instrument using SoftMax Pro software program. The final assay volume was 50µl and consisted of 5µl of the flurogenic substrate Boc-RVRR-MCA (final concentration was 50µM), 5µl of the rec-rPC4 enzyme (specific activity = 179µmol AMC/h/µmol protein, equivalent to 500U, where one U or unit = 1 pmol released/hour/µl) and 40 µl of buffer A (i.e. 25mM Mes, 25mM Tris, 2.5mM $CaCl_2$, pH 7.4). For assessment of PC4-inhibitory property of rec-mCRES protein,

the fluorogenic substrate was added following a 20 min pre-incubation of the enzyme rec-purified rPC4 with rec-mCRES at various concentration levels in the above buffer at 37°C. The release of highly fluorescent free AMC (7-amino 4-methyl coumarin) from the substrate Boc-RVRR-MCA was measured over a 1 h period at excitation and emission wavelengths fixed at 370nm and 460nm respectively using a spectrofluorometer (Spectra Max Gemini XS, Molecular Devices, USA). The assay plates were kept in 37°C incubator with shaking at a very slow speed of 60 rpm (revolution per minute). Data were collected and their mean values calculated from two independent experiments and were shown with the error bars. Based on fluorescence readings obtained, K_m, K_i and IC_{50} values were calculated as described in the section below.

3.2.4.1 Determination of kinetic parameters (K_m, K_i and IC_{50} values)

K_m (Michaelis-Menten constant) value for the rec-PC4 enzyme sample used in this study was first measured using the Michaelis-Menten curve based on various concentrations of substrate Boc-RVRR-MCA ranging from 1 – 100 μM (96).

K_i (inhibition constant) and IC_{50} (concentrations needed to achieve 50% inhibition of enzyme activity) determinations were achieved graphically using Dixon and Sigmoidal plots respectively and the data were analysed by Grafit version 4.0 software (Erithacus software limited, Staines, UK). The initial measured RFU (Raw Fluorescence Unit) values were converted to pmol AMC released per hour using the standard equation (see below) before the graphs were plotted.

Equation obtained from the standard curve was,

x (pmol of AMC) = RFU / 16.4

In all our experiments, this above equation was used for converting the RFU values to pmol AMC released based on per hour data. The above standard curve (*not shown*) was generated by plotting the fluorescence values against the various concentrations of free AMC used.

For Dixon plot we used the graphic software Sigma Plot version 11.0 where 1/v was plotted against various concentrations of inhibitor I (v = velocity of reaction and = inhibitor concentration). It is known that if the lines for various substrate concentrations pass through a common point of intersection, then the inhibition is considered as competitive. On the other hand if they do not pass through a common point of intersection then it is of non-competitive type. If they are parallel then it is known as un-competitive inhibition.

In addition to Dixon plot, Lineweaver-Burk plot was also carried out to confirm the competitive nature of inhibition where 1/V is plotted against 1/S (V = velocity of reaction, S = substrate concentration).

3.2.5 Digestion of rec-mCRES protein *in vitro* by rec-rPC4 enzyme

Rec-mCRES dimer (5µl containing 5µg protein) was incubated with purified rec rPC4 (20 µl, specific activity = 179µmol AMC released/hour/µg protein/µl) at 37^0C in a shaker incubator. A small sample of aliquot (typically 5 µl from the digest) was removed at various time points such as 0h, 1h, 2h and 4h and subjected to analyses by SELDI-TOF mass spectrometry. The various peaks obtained were analyzed for identification of cleaved fragments of mCRES sequence.

3.2.6 Collection of fluid samples from mouse epidydimis and their analyses for PC-like activity

Four adult mice (C57 Black, 9 weeks old, bought from Charles Rivers, Quebec, Canada) of the same age and housed under identical conditions were sacrificed and their epididymis

sections were separated. Under the microscope, fluid samples from various sections of the epididymis for each mouse were collected. These were caput, corpus, cauda and the vas deferens. The fluids were squeezed from the tubicals and were collected directly in 500µl PBS buffer. The sperms and the suspended tissues were separated from the supernatant by centrifuge at 4^0C. The clear and transparent supernatant thus obtained from each epididymal section was collected for further examination. Henceforth these will be referred to as epididymal samples. First protein assay was conducted on each sample using the BCA reagent kit in order to measure the protein content of each collected epididymal sample from each mouse. The protease activity assay of each sample was carried out using 5µl of fluorogenic Boc-RVRR-MCA (final concentration 50µM) and 5µl of $CaCl_2$ (final concentration of 5mM) together with 40 µl of each epididymal sample whose protein content and specific enzyme activities are described in the **Tables 6& 7**, respectively of the result section.

The total specific enzyme activity of each of epididymal fluid sample was then estimated and compared to one another. A part of this activity belongs to PC4 and efforts were made to estimate this by using a PC4-specific inhibitor namely the prodomain peptide $PC4^{75-84}$ as reported earlier (57) .

3.2.7 Cell Culture

The inhibitory effect of PC4 on CRES was then to be studied in the cell lines where the IGF2 (which is a PC4 substrate) processing was to be compared in the presence or absence of the CRES. Hence, the placenta cell lines (HTR-8/SV neo cell line from human trophoblast) obtained from our collaborator **Dr Qin Qiu** were grown in the RPMI-1640 medium (Rosewell Park Memorial Institute medium) (Sigma Aldrich Chemical Company,

USA) containing 10% FBS (Fetal Bovine Serum) (Gansera Intentional Inc, ON, Canada) and 0.6µl/ml of antibiotic gentamicine sulphate (GTM) (Invitrogen Corporation, ON Canada) in presence of 5% CO_2 at 37^0C. Phenyl red indicator was used during the preparation of this medium since it is considered as an indicator of the presence of dead cells in the culture medium which may result from the over growth of cells, high toxicity or possible bacterial growth. This would be indicated by the color change of the medium from clear transparent dark pink to opaque light pink, due to the change in pH of the medium under any of the above condition. Cells were grown until they attained at least 80% confluency and then the passage of the cells was done into the 6-well plates to grow. Next day, when these cells were grown to near 80% confluency in the 6-well plates, the old medium was removed and fresh RPMI-1640 serum free medium was added. During this time the chemicals (i.e, 10 µl rec-rPC4 and varying concentrations of rec-mCRES) were also added to the medium in all the wells of the plates. This was left as such till the next day for these chemicals to enter the cells and perform their actions. Then the culture media and the cell pellets were collected in separate tubes. The media samples were centrifuged before storage in at $<4^0$C to remove all the cell debris and insoluble particles. The pellet was also stored separately in an identical manner. These collected media samples were subjected to analysis by western blot using hIGF-2 antibody and raised against the whole mature protein (R&D systems, USA) with 1:500 dilution. The above experiment was also repeated in an identical manner without the external addition of rec-rPC4 enzyme (after 24h cell growth) into the fresh medium keeping all other conditions unchanged.

<u>Note</u>: Equal amount of the PC4 enzyme (i.e, 10µl for which specific activity = 179 µmol AMC/h/µmol protein, equivalent to 500U, where one U or unit = 1pmol released/hour/µl).

4. RESULTS

4.1 Production of recombinant rat PC4 (rec-rPC4) enzyme:

Using the *L. tarentolae* expression system, rec-rPC4 was produced in large amount as described in the Materials and Methods section. Culture medium (1000ml) containing rec-rPC4 protein upon 40-fold concentration gave 25ml of concentrate medium using the concentrators [cut-off Molecular Weight (MW)=10kDa]. It was then loaded into a DEAE-agarose column (2.5 x 15mm) for purification of the enzyme.

4.1.1 Purification of the rec-rPC4 via DEAE column chromatography:

Firstly the column was pre-equilibrated with the Buffer A (25mM Tris / 25mM Mes / 2.5mM $CaCl_2$, pH=7.4) followed by the buffer A wash where fractions 1-95 (#1-95) were collected each fraction being approximately around 1 ml size. Buffer A wash of the column was done for only the unbound proteins to come out. Following this wash, further chromatography was continued with elution buffers containing a step-wise gradient of NaCl concentration, starting with elution buffer 1 which was 30 mM NaCl pool # (1-80) followed subsequently by elution buffer 2 which was 60mM NaCl #(1-60),elution buffer 3 which was 70mM NaCl #(1-74) and elution buffer 4 which was 200mM NaCl #(1-67).Enzyme activity and protein content of each of the 381 fractions obtained were conducted using the fluorogenic substrate Boc-RVRR-MCA (50µM) and the measurement of optical density at 280 nm respectively as described in Materials and Methods section**.** The overall purification method is depicted in **Figure 6** and all the assays results during the chromatography are indicated in **Figure 7(A-E)**. Based on these assay results, specific fractions were then combined and are shown in **Figure 8**.

Figure 6: **Schematic diagram showing rec-rPC4 production and its chromatographic purification**

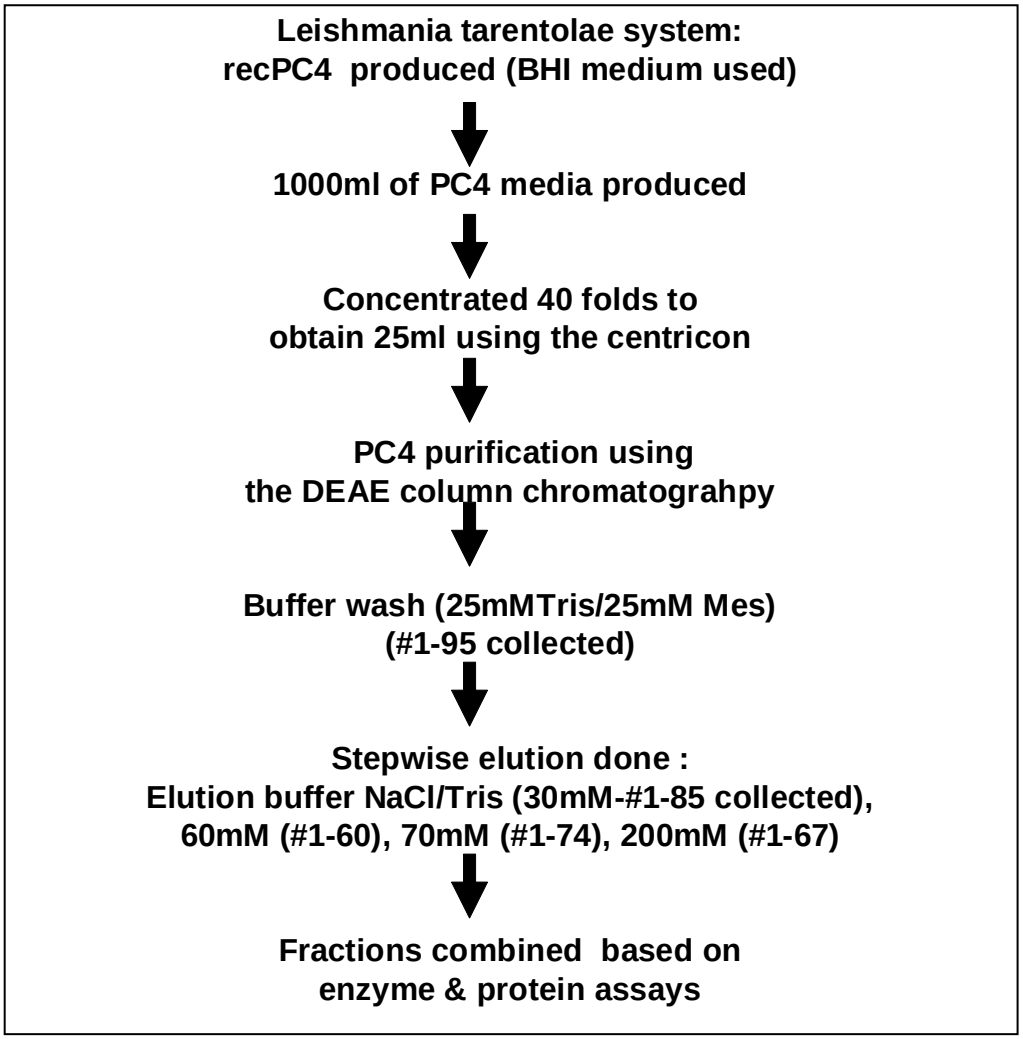
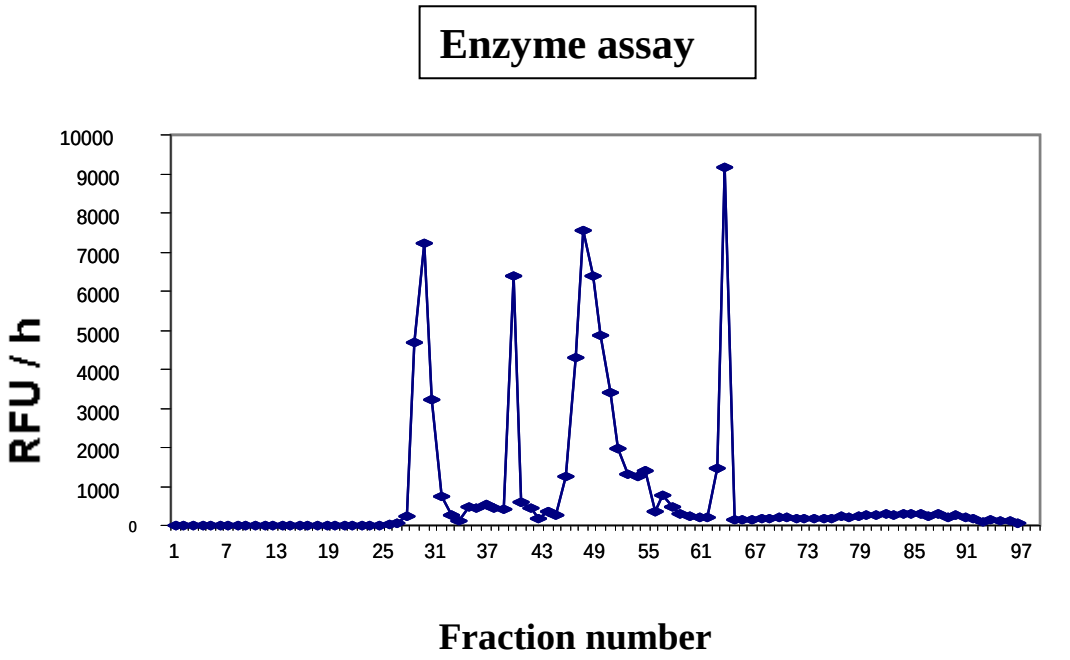
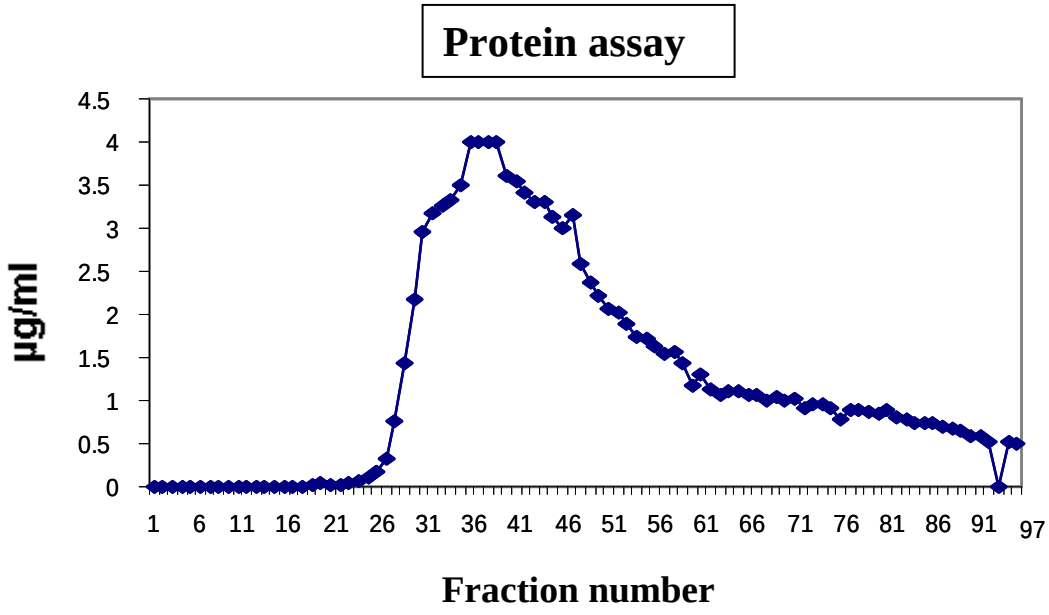


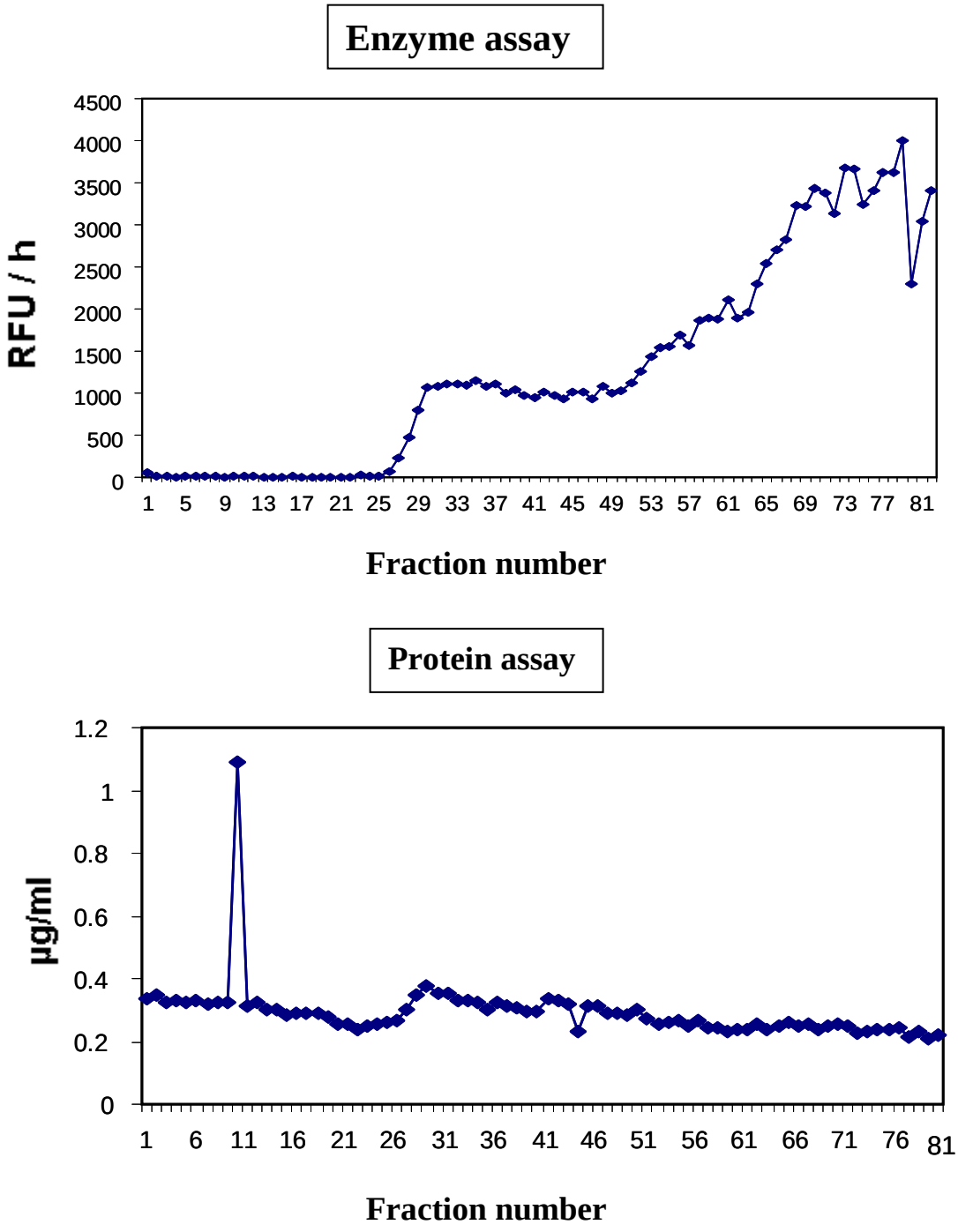
Figure 7(A) showed the enzyme activity and the protein content profiles of all 95 fractions of buffer-A wash pool. **Figure 7(B)** displayed the profiles of the next 80 fractions of 30 mM NaCl pool. Subsequently **Figures 7(C), (D) and (E)** depicted the measured enzyme activity and protein profiles of 60mM, 70mM and 200mM NaCl containing buffers respectively. These elutions buffers were expected to release most of the bound proteins including rec-rPC4. This is reflected by the change of the color of the resin brown to near colorless. It may be pointed out that the resin turned brown upon initial loading of rec-rPC4 protein sample. The 381 fractions thus obtained were then further combined based on their high activity and low protein content criteria. On combinations of these fractions, 28 pools were then obtained. These pools were then further examined for their presence of PC4. The details of these pools are shown in **Figure 8**. The elution of same PC4 enzyme at various salt concentrations is observed to be present but it is not unusual since it depends on the charge profile.

Figure 7(A): DEAE-agarose column chromatography of crude concentrate rec-rPC4 sample obtained from *L. tarentolae* expression system: Buffer-A (25mM Tris/25mM Mes/2.5Mm $CaCl_2$) where the collected fractions were #1-95.

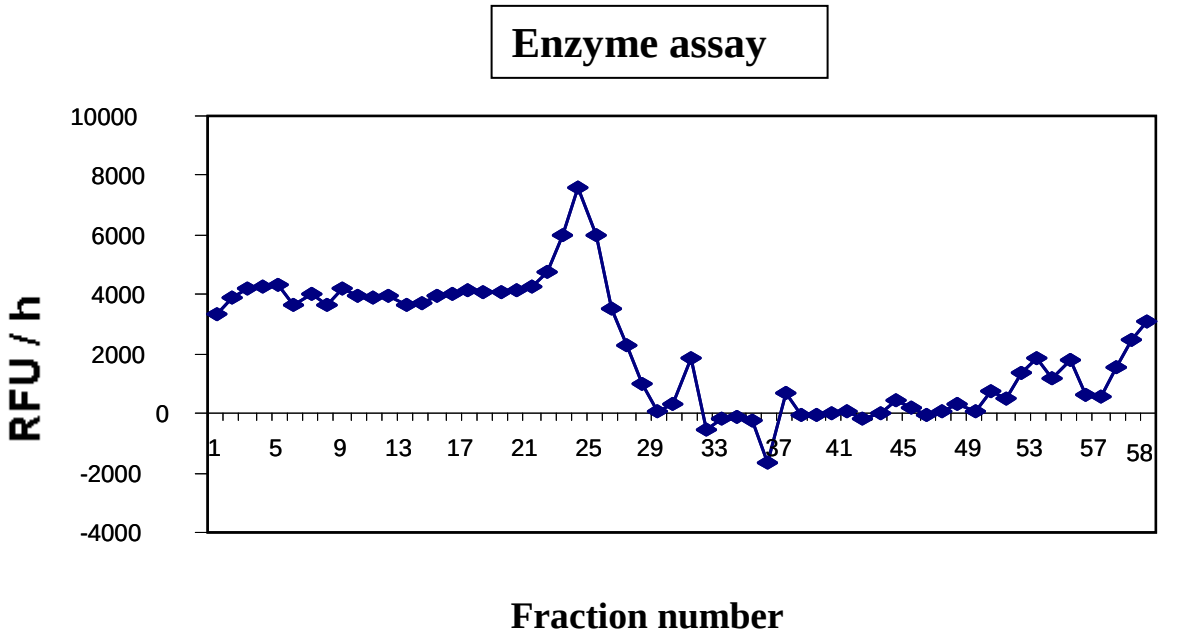


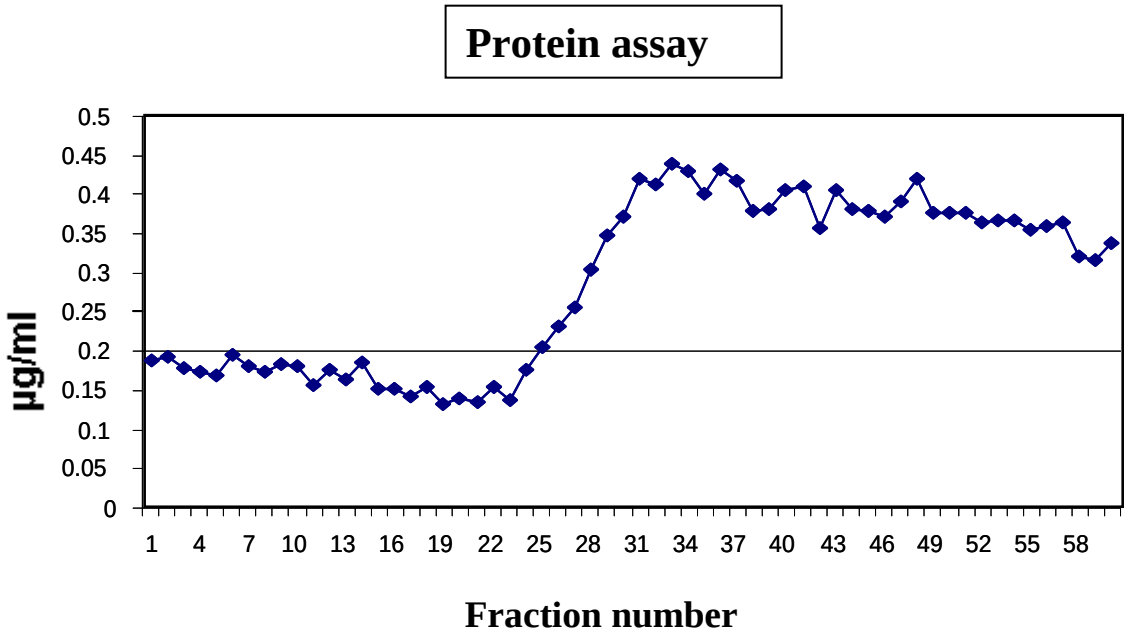


<u>Figure 7(B)</u>: DEAE-agarose column chromatography of crude concentrate rec-rPC4 sample obtained from *L. tarentolae* expression system (continued): Elution Buffer 1: <u>30mM NaCl</u> where collected fractions were #1-80 .

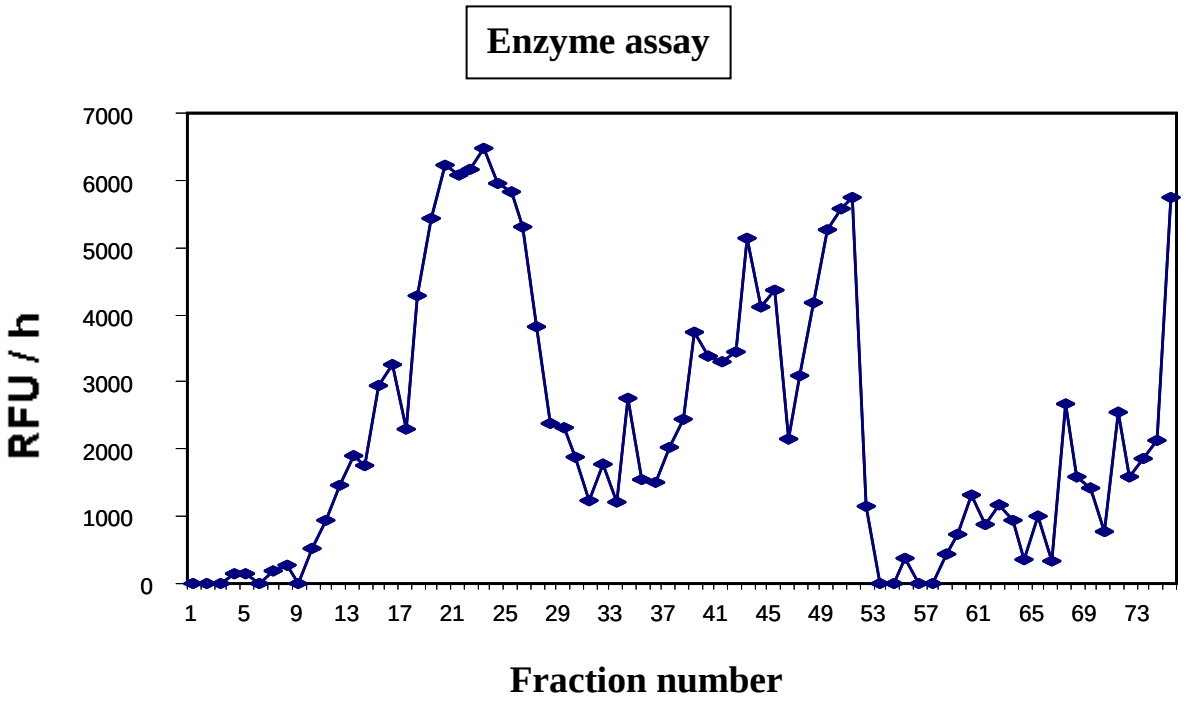


<u>Figure 7(C)</u>: DEAE-agarose column chromatography of crude concentrate rec-rPC4 sample obtained from *L. tarentolae* expression system (continued): Elution buffer 2: <u>60mM NaCl ,</u> where collected fractions were #1-60





<u>Figure 7(D)</u>: DEAE-agarose column chromatography of crude concentrate rec-rPC4 sample obtained from *L. tarentolae* expression system (continued): Elution buffer 3: <u>70mM NaCl ,</u> where collected fractions were #1-74



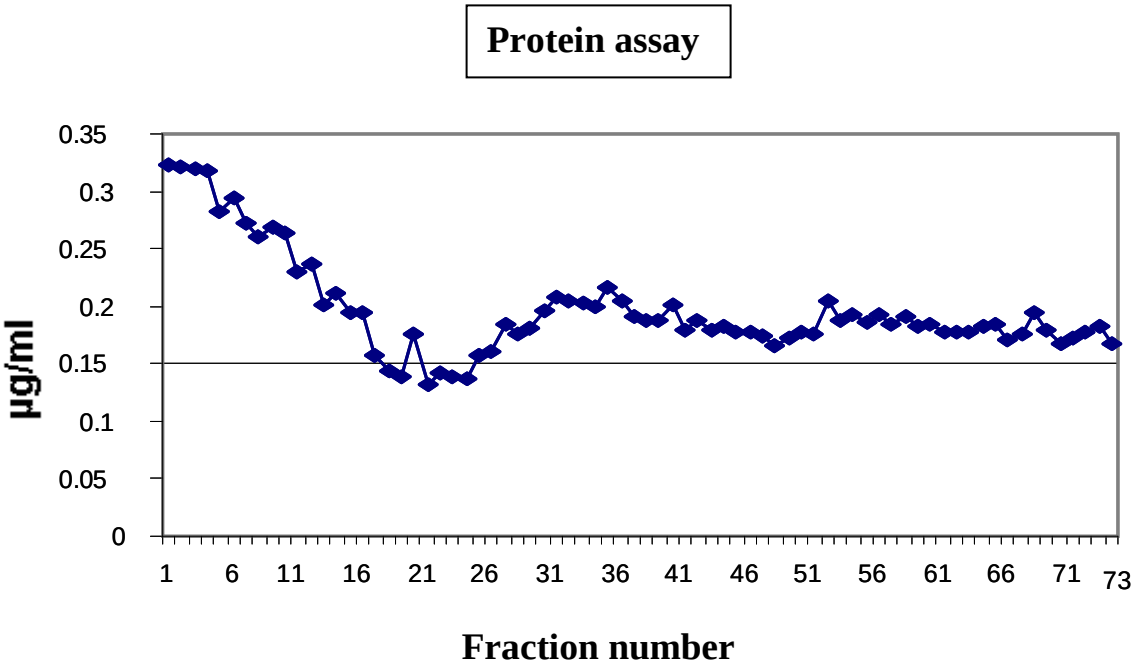
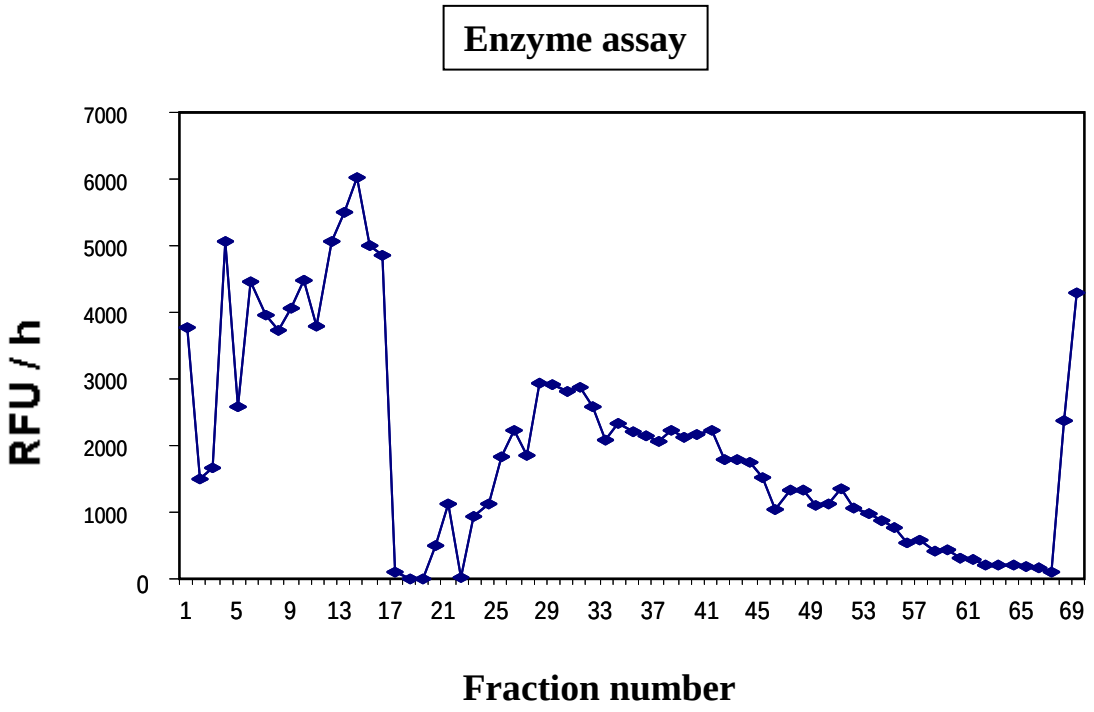


Figure 7(E): DEAE-agarose column chromatography of crude concentrate rec-rPC4 sample obtained from *L. tarentolae* expression system (continued): Elution buffer 4: 200mM NaCl where collected fractions were #1-67



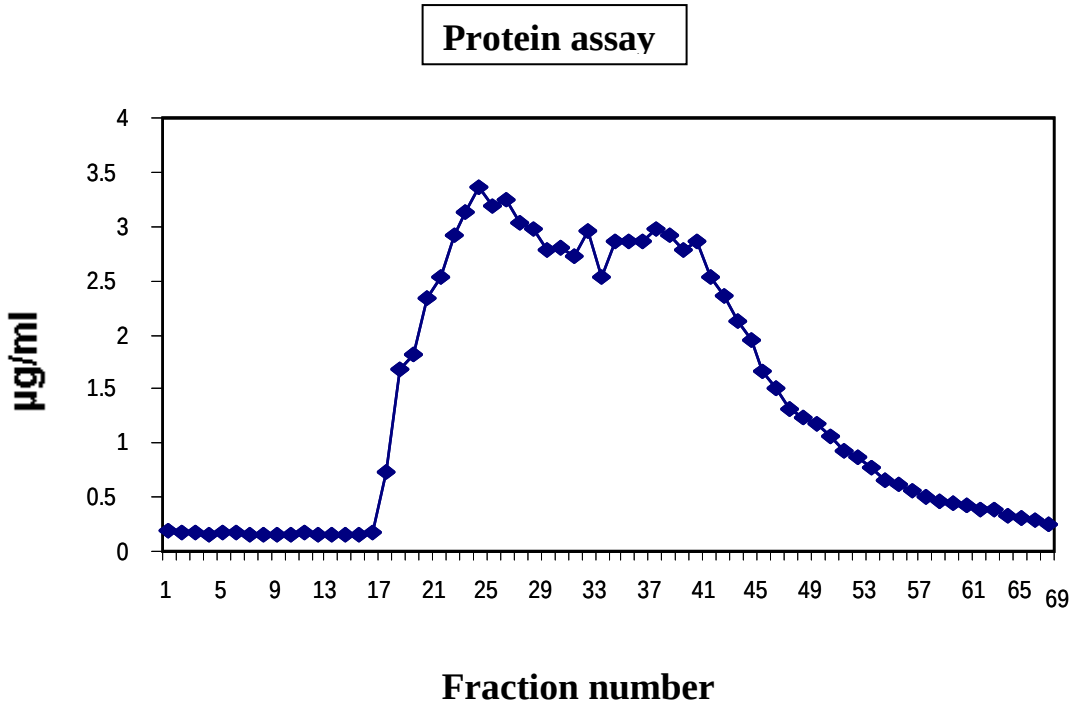
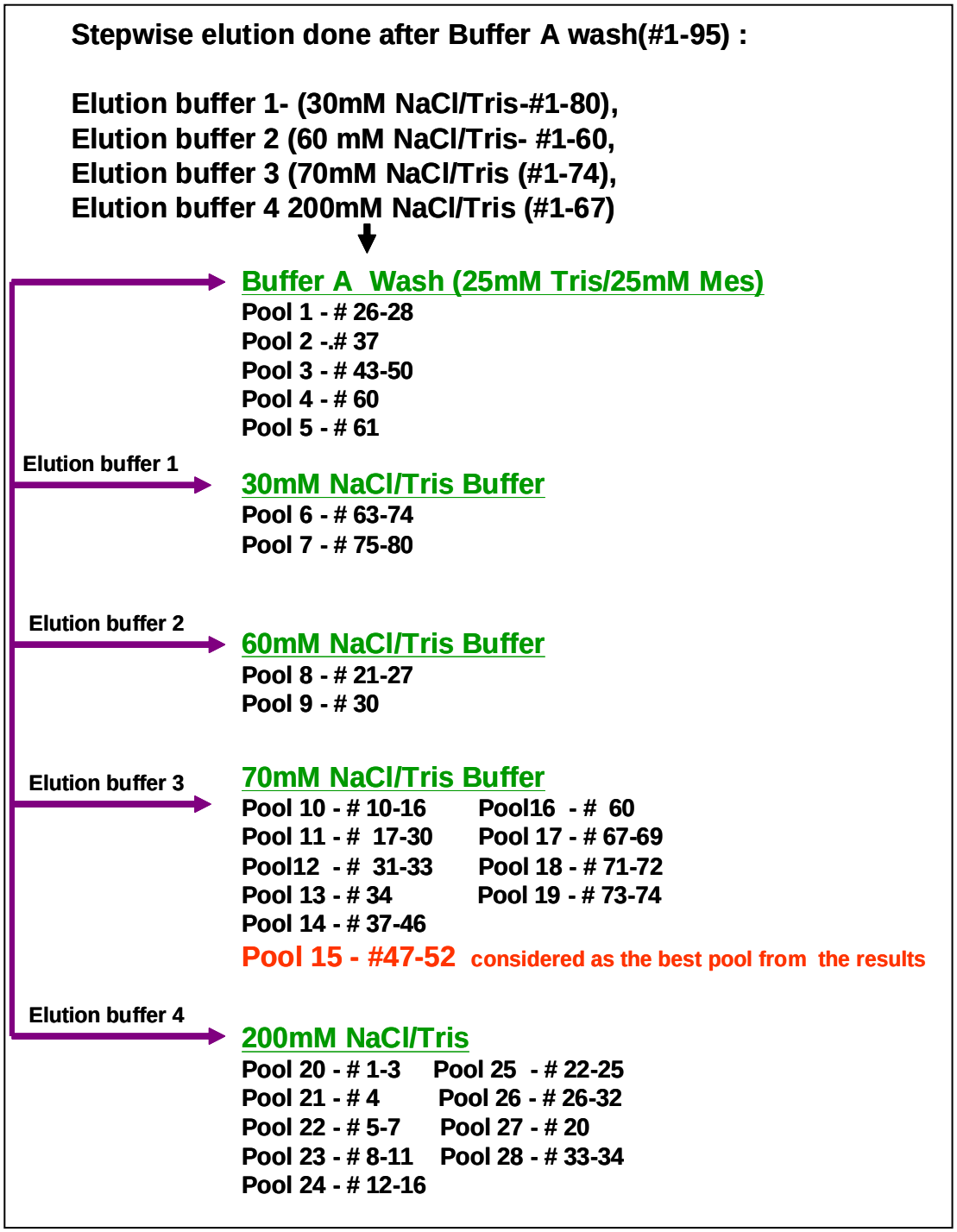


Figure 8: Schematic figure showing the combination profiles of various fractions of rec-rPC4 following DEAE-chromatography as depicted in Figures 7 A-E:



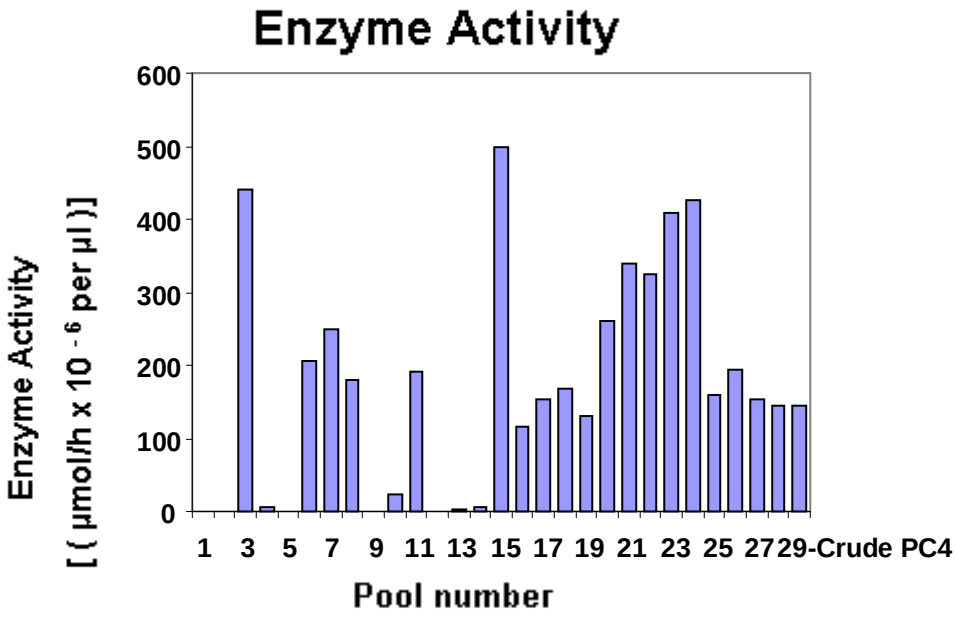
Note: Pool 15 was considered as the one containing most purified PC4 enzyme. Based on activity and protein assays and later confirmed by gel and western blots results.

After combining the fractions, the 28 pools that were obtained were tested for their enzyme activity on the fluorometer with the fluorogenic substrate Boc-RVRR-MCA (50µM) as indicated in **Figure 9(A).** The pools were also tested for their protein content using the BCA-protein assay as shown in **Figure 9(B)**. The specific activity graph was plotted from the enzyme and protein assay data which suggested that pools #15, 21, 22, 23, 24 exhibited high specific activities as depicted in **Figure 9(C)** out of which **pool #15** was selected to be the most potent one because of its high specific activity, sharpness of the peak and large available volume. The 5 above pools including **pool #15** were further examined for the presence of PC4 protein by SDS-gel electrophoresis with coomassie staining as well as by western blot analysis. The coomassie stained 12% SDS gel profiles was shown in **Figure 10(A)** whereas the western blot analysis was shown in **Figure 10(B).** The band appearing at ~ 65 kDa was consistent suggesting the presence of mature active form of PC4 as indicated by the previous studies in the literature (57)**.** The western blot performed with specific antibody raised against PC4 protein (1:500) also agreed with this conclusion. The PC4 purification and its profile in various pools are presented in **Table 5**, which also showed the fold-of purification achieved by the DEAE column. The table also confirmed that both pools #15 and #24 contain most purified samples of PC4 being ~511 and ~677 fold purified compared to the crude sample whereas pools #21, 22 and 23 also showed relatively good amount of fold purification being above 450 range. Pool 15 which was henceforth used for the experiments had an activity = 500 Units where 1 unit =1µmol AMC released/hour/µl and specific activity of 179µmol AMC released/hour/µmol of protein. Thus from the results shown above it was clear that PC4 was purified quite efficiently via DEAE-agarose column

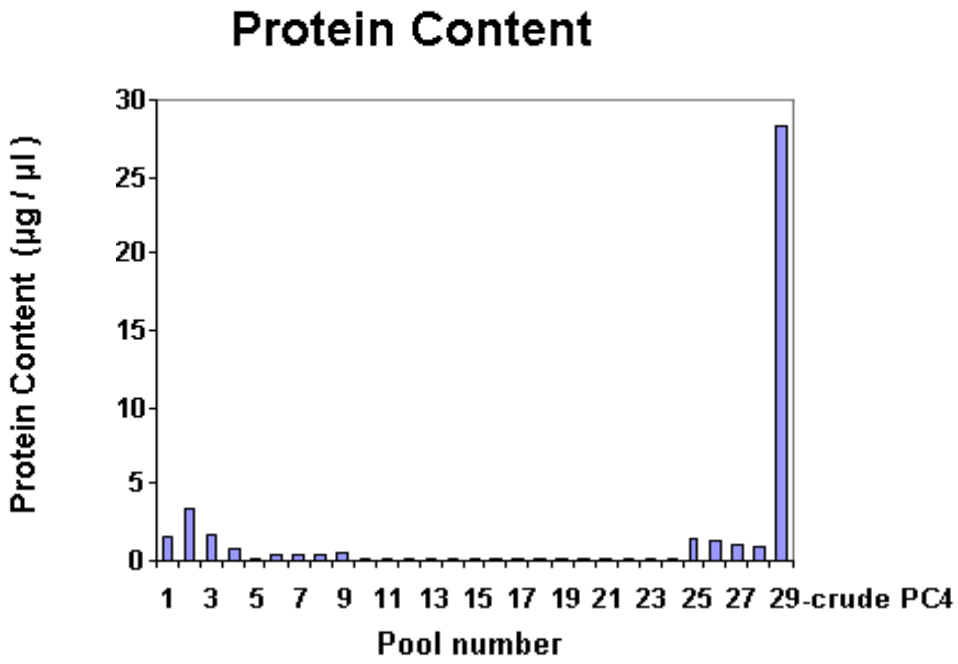
chromatography as described. **This purified rec-rPC4 present in pool #15 was used for all kinetic and other biochemical studies reported in this thesis.**

<u>Figure 9</u>: Enzyme activity and protein profiles of various combined 28 pools as obtained by DEAE-agarose column chromatography.

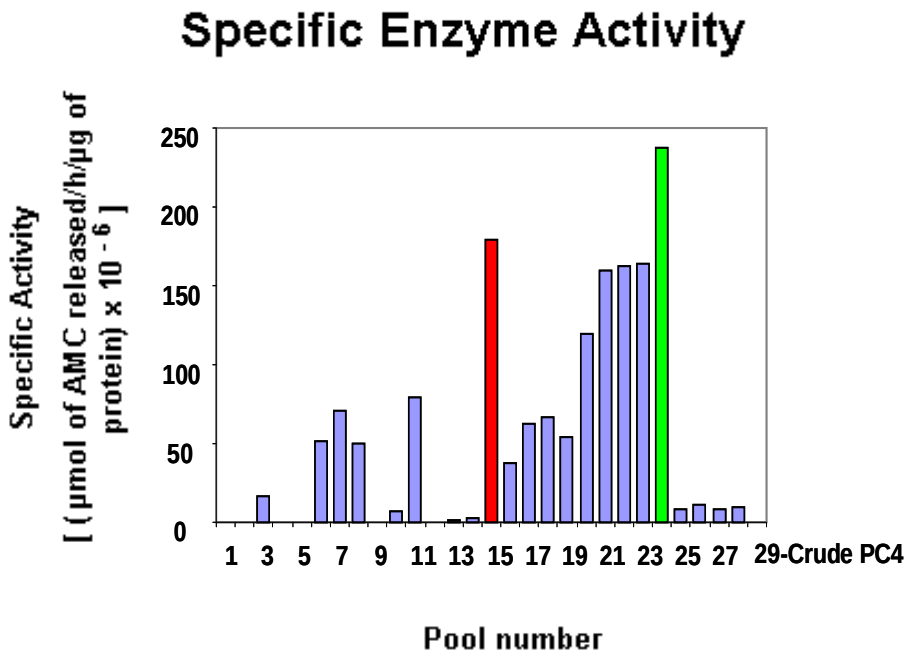
(A)



(B)



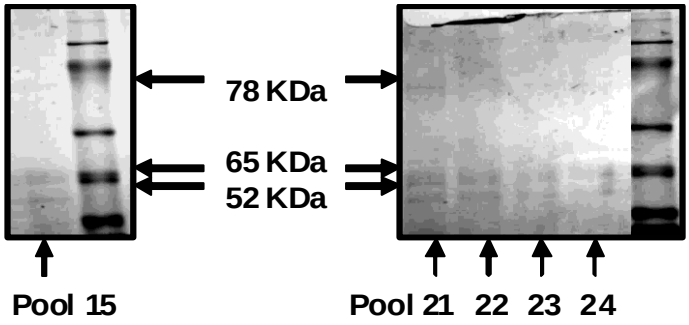
(C)



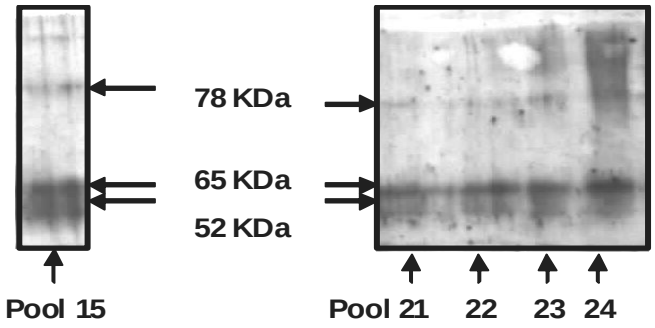
Note*: The pool #15 and 24 colored in red and green respectively exhibited highest specific activity for PC4 suggesting their high degree of purity.*

Figure 10: (A) SDS-page with coomasie staining and (B) Western blot analysis of various DEAE pools containing active PC4 protein.

(A) Coomasie stained 12% SDS gel



(B) Western blot analysis



5 µg of protein is loaded in all the lanes

Table 5 : Summary of rec-rPC4 purification

	Pool Number	Volume (ml)	Protein Content (µg/µl)	Enzyme Activity [(µmol/h) x 10 $^{-6}$]	Specific Actitvity [(µmolAMC released/h/µg protein) x10^{-6})	Relative Amount	Fold Purification
1	1	2.6	1.51	0	0	0	0
	2	0.8	3.386	0	0	0	0
	3	4.4	1.732	441.18	17	48.47	48
	4	0.4	0.773	5.08	0.461	1.31	1
	5	0.4	Could	not	be	analysed	
2	6	10.6	0.265	207.15	51.75	147.85	148
	7	5.7	0.24	249	71	202.85	203
3	8	6.9	0.256	181	50	142.85	143
	9	0.7	0.438	0	0	0	0
4	10	6.5	0.225	23.4	7	20	20
	11	12.6	0.165	191	79.5	227.14	227
	12	2.4	0.224	0	0	0	0
	13	0.8	0.218	3	1	2.85	3
	14	7.6	0.206	7	2.3	6.57	7
	15	4.0	0.191	500	179	511.42	511
	16	0.9	0.207	115	38	108.57	109
	17	2.4	0.174	155	62	177.14	177
	18	1.9	0.175	169	67	191.42	191
	19	1.9	0.167	130	54	154.28	154
5	20	2.2	0.154	262	119	340	340
	21	0.6	0.146	338	160	457.14	457
	22	2.1	0.144	325	162	462.85	463
	23	2.6	0.172	410	164	468.57	469
	24	3.7	0.126	427	237	677.14	677
	25	2.2	1.406	159	8	22.85	22
	26	6.2	1.27	195	11	31.42	31
	27	0.7	1.129	153	9	25.71	26
	28	12.4	0.972	146	10	28.57	29
	Crude PC4	10	28.32	146	0.34	1	1

Note:

1. Buffer A
2. 30mM NaCl/Tris Buffer
3. 60mM NaCl/Tris Buffer
4. 70mM NaCl/Tris Buffer
5. 200 mM NaCl/Tris Buffer

Pools 15,21,22,23 and 24 showed the highest fold purification among all the 28 pools.

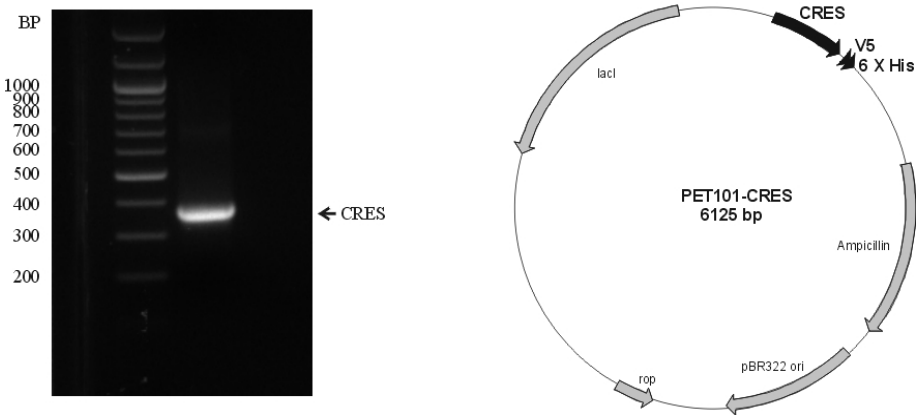
4.2 Recombinant mCRES protein

4.2.1 Production

Rec-mCRES was produced in the laboratory with the help of our collaborator using the E-Coli expression system as described in Materials and Methods section. A full length CRES PCR product without the signal peptide sequence but containing both V_5 and His_6 labels with the expected size of 376 bp (base pairs) as depicted in F**igure 11(A)** was amplified using bluescript vector DNA containing mouse germ cell full-length mCRES cDNA as described in the Materials and Methods section. The northern blot analysis showing the presence of mRNA of CRES-V_5-His_6 fusion protein is displayed in F**igure 11(B)**.

<u>Figure 11</u>: Plasmid construction of mCRES-V_5-His_6 vector.

(A) (B)



Acknowledgement for this figure: Dr. Qin Qiu

4.2.2 Purification of rec-mCRES.

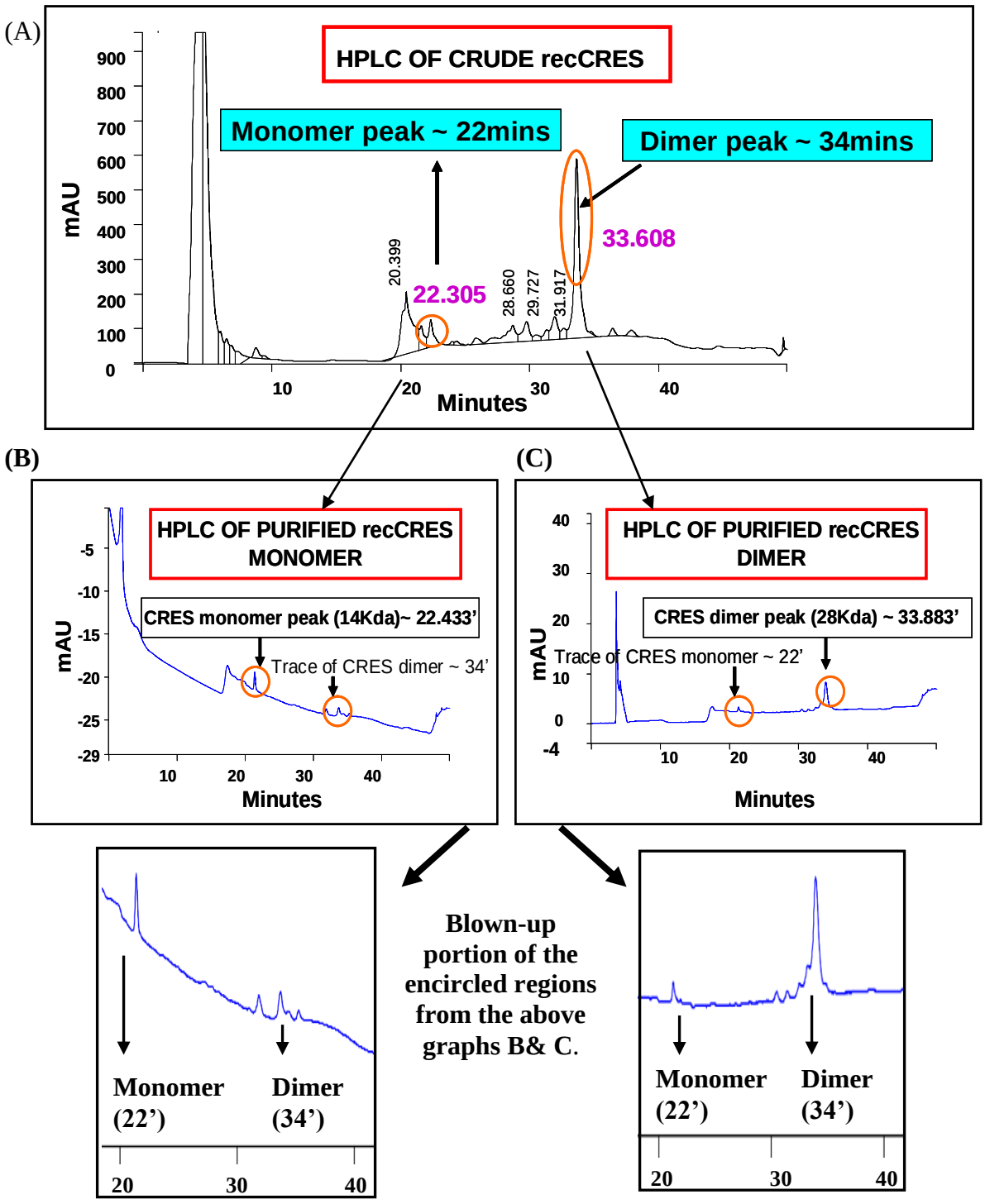
The expressed rec-mCRES protein present within the cell membrane as obtained above was dissolved by adding 6M aqueous guanidine HCl solution to the cells with the help of vortex and sonication followed by filtering (using 0.2µm membrane) of the solution to obtain a clear transparent solution which was then subjected to purification by RP-HPLC using C4 columns. The total weight of the cell pellet was ~10g which was extracted with 200 ml of 6M aqueous guanidine-HCl solution. This crude solution was subjected to HPLC purification in stages by injecting ml at a time into the C4 column.

4.2.3 RP-HPLC of rec-mCRES

The clean filtered rec-mCRES solution obtained from above was subjected to RP-HPLC chromatography using a C4 semi-preparative column (details in Materials and Method section). The various HPLC peaks observed were collected in different tubes. Mass spectrometry analysis was carried out on each of these peaks to find out which one of them corresponded to CRES protein. The calculated molecular weight of rec-mCRES-V_5-His_6 protein was ~14.3kDa. The mass spectrum of crude CRES sample exhibited major peaks at ~ 14kDa and 28kDa corresponding to monomeric and dimeric CRES protein, respectively. The RP-HPLC profile of CRES protein exhibited two peaks at retention times (Rt) **22min** and **34min** which were attributed by mass spectrometry to the monomeric and dimeric forms of CRES protein, respectively as shown in **Figure 12 (A).** The purities of these two collected peaks were further checked by re-chromatrography under identical condition using an analytical grade C_4 column which showed single peak for each form. The conversion of monomeric CRES to its dimeric form was found to be quite rapid with time as judged by RP-HPLC and mass spectrometry. This is evident in **Figures 12(B) and (C)**. Nearly 10mg

powder of each monomeric and dimeric CRES protein was obtained from 200ml solution of crude CRES protein. *This represented 0.066mg of CRES protein per ml of crude extract*.

Figure 12: (A) RP-HPLC chromatogram of crude rec-mCRES using C4 analytical column. The figures (B) and (C) represent the second chromatograms of the collected peaks to check their homogeneity.Here the symbol (')denotes for minutes.

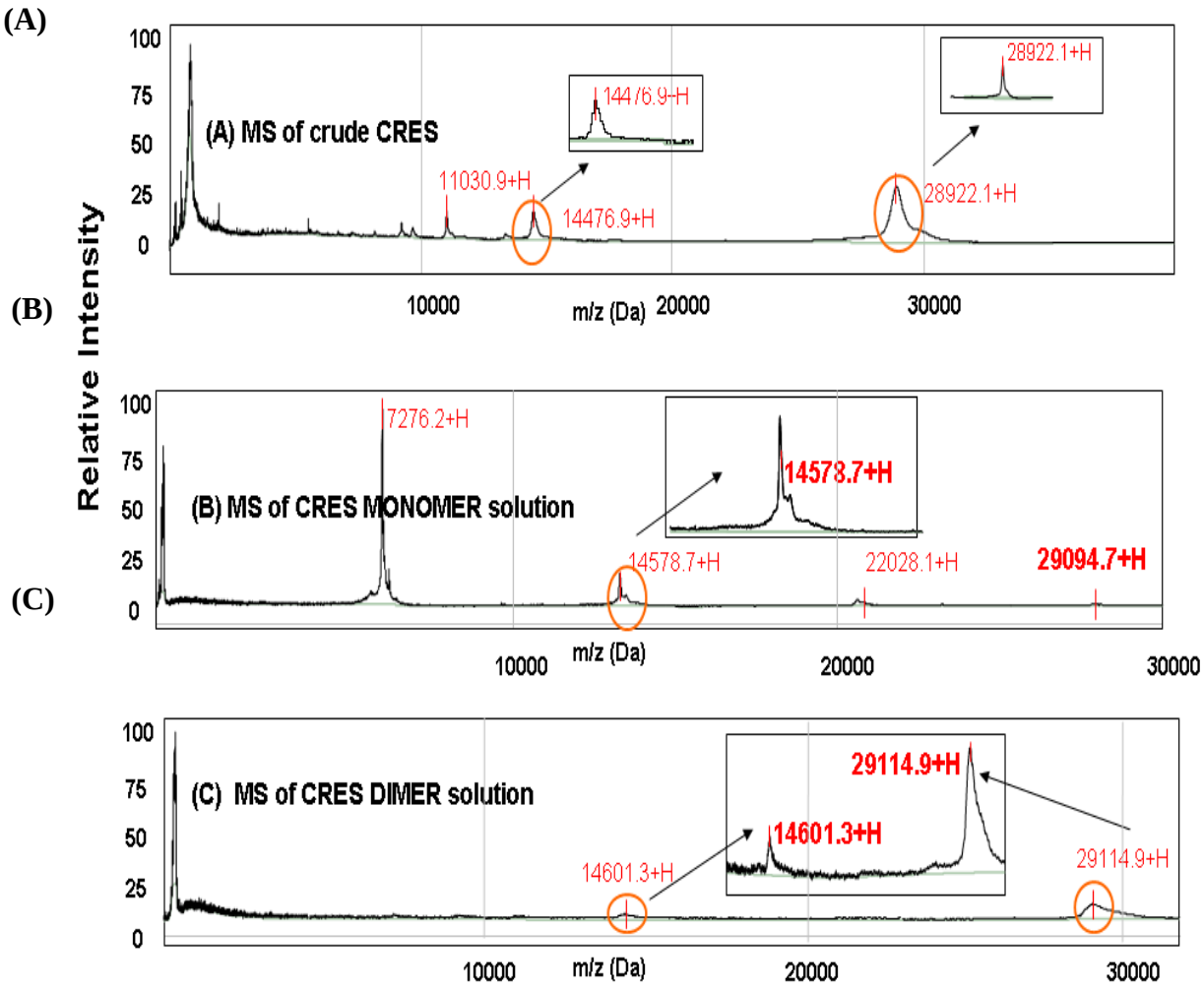


4.2.4 Mass spectrometry of purified rec-mCRES

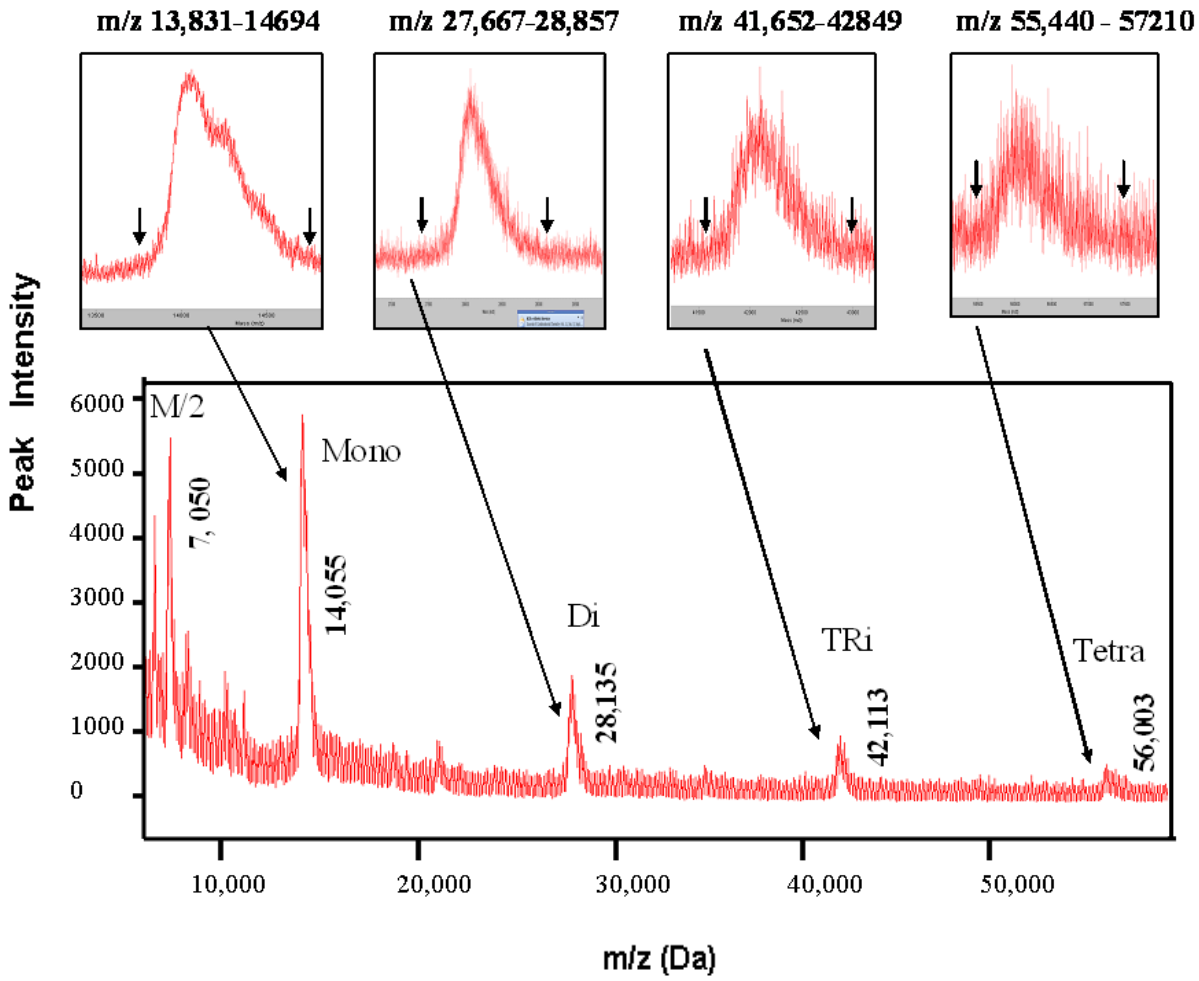
The purified and lyophilized rec-mCRES powder obtained after RP-HPLC was examined for its molecular weight by SELDI-tof MS. The results showed that the crude rec-mCRES protein consisted of both monomeric (~14kDa) and dimeric (~28kDa) forms. In addition, two more peaks at ~11 and 21kDa were also observed which were possibly due to shorter forms of CRES protein (132). The results are shown in **Figure 13(A)**. The mass spectral data of peak 1 (CRES monomer) and peak 2 (CRES dimer) obtained from the RP-HPLC suggested that in the purified monomeric fraction, trace of dimer was also present as indicated by the presence of a minor peak at m/z ~28kDa in addition to the large monomeric peak at ~14kDa. Interestingly in the dimeric sample we also noted the presence of small amounts of the monomeric form as indicated in **Figures 13(B) & (C)**. Thus neither monomeric nor dimeric samples of rec-mCRES protein obtained from RP-HPLC were completely homogeneous.

The mass spectral analyses of rec-mCRES samples were also performed by MALDI-TOF technique and provided similar results. Interestingly it was further observed that some higher polymeric forms of CRES apart from the dimeric form were also present. These included trimer as well as tetramer forms as shown in **Figure 14**. Thus our data revealed that rec-mCRES is prone to self-association on storage leading to its various oligomeric forms. This is consistent with earlier results reported by Gail Cornwall (127). Additional discussion about the nature and function of CRES oligomers can be found later under the **Discussion** section.

Figures 13: (A), (B) & (C) shows the SELDI-tof MS of crude as well as purified samples of CRES monomer and dimer forms following RP-HPLC respectively.(A) MS of Crude CRES sample. (B) MS of CRES dimer sample.(C) MS of CRES monomer sample.



<u>Figure 14:</u> MALDI-tof MS of purified rec-mCRES after HPLC showing the formation of mono and multi-meric forms. Also note the formation of M/2 peak confirming the molecular weight of the monomer.



4.2.5 Characterization of rec-mCRES

The sequence of the rec-mCRES protein is shown in **Figure 15**. V_5 and His_6 tags with a start codon were added to the CRES sequence at the C-terminal. The molecular weight of this sequence was theoretically calculated to be ~17.6kDa as indicated in **Figure 15(A)** with both the tags intact and ~14.3kDa which is shown in **Figure 15(B**) without the two tags. The addition of these two tags in the construct was done in order to facilitate easy and efficient purification of the resultant rec-mCRES protein using nickel column chromatography. However the results were not so encouraging and hence an alternate RP-HPLC method was used for further purification. Since our mass spectrometry results repeatedly showed the molecular weight of rec mCRES produced to be around 14.3kDa, hence it was proposed that probably the two His_6 and V_5 tags were removed during or after the production process. The calculation of the molecular weight of the sequence was achieved by using the Peptide Mass Cutter program on the www.expasy.ch website.

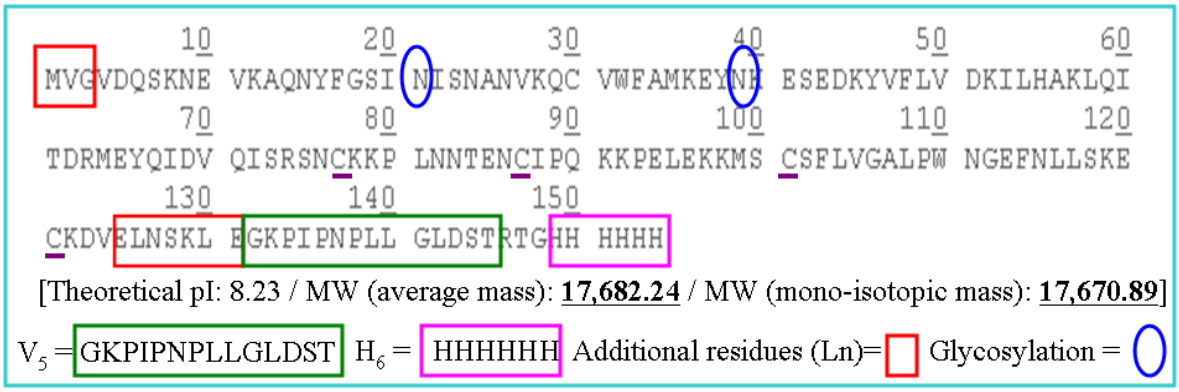
4.2.6 Endoproteinase Lys-C digestion study of rec-mCRES protein

Even though the mass spectrometry results have indicated that the produced protein is indeed CRES, we wanted to reconfirm its identity by endoproteinase Lys-C digestion study in the following experiment. Lys-C was used to digest rec-mCRES protein and SELDI-tof MS was performed on the crude digest at various time points namely at 0h, 1h, 2h and 4h and the results were compared with those of the corresponding control samples under identical condition. The results are shown in **Figure 16A**. Several fragment peaks were observed which were analyzed further. Expanded portions of the mass spectra are shown in **Figure 16(A) and Figure 16(B)** for the 2h data of CRES digestion by Lys-C, between the molecular weight range **i) 5,600-5,800Da, (ii) 5,200-5,500Da, (iii) 4,900-5,000Da and (iv)**

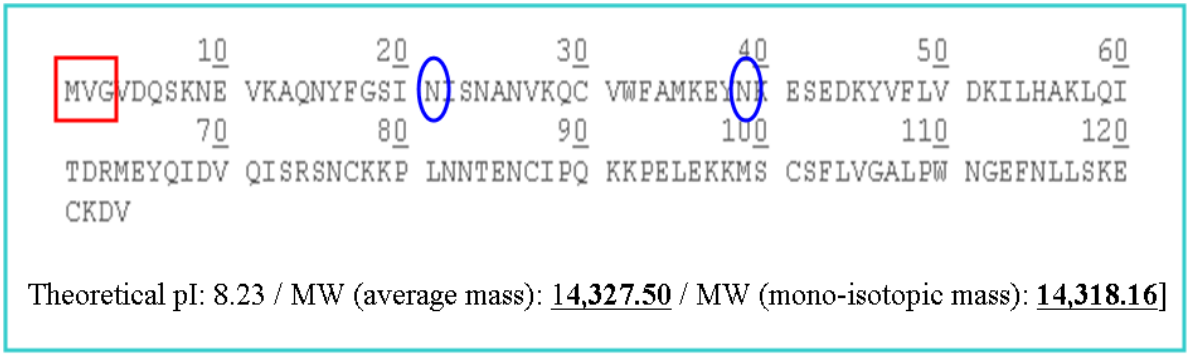
4,500-4,700Da. Analysis of the data obtained in **Figures 16(A) & (B)** is presented in **Figure 16(C)** showing possible explanations of all the observed peaks obtained following rec-mCRES digestion by Lys-C. The sequence of rec-mCRES and its various cleavage sites by Lys-C are shown by letters A, B, C, D, E, F and G as indicated in **Figure 16 B (i-iv).** These include peaks at **m/z 5,338, 5,360, 5,377, 5,392, 4,938, 4,952, 4,974, 4,997, 4,591 and 4623 Da**. It is important to point out that the several peaks that were seen in the mass spectrum of CRES digest by Lys-C are as expected absent in the mass spectra of CRES alone without the added Lys-C.

Figures 15 : (A) & (B) rec-mCRES sequences shown with and without V_5 and His_6 tags.

(A) Sequence of rec-mCRES (signal peptide removed) + Linker + V5 + H6



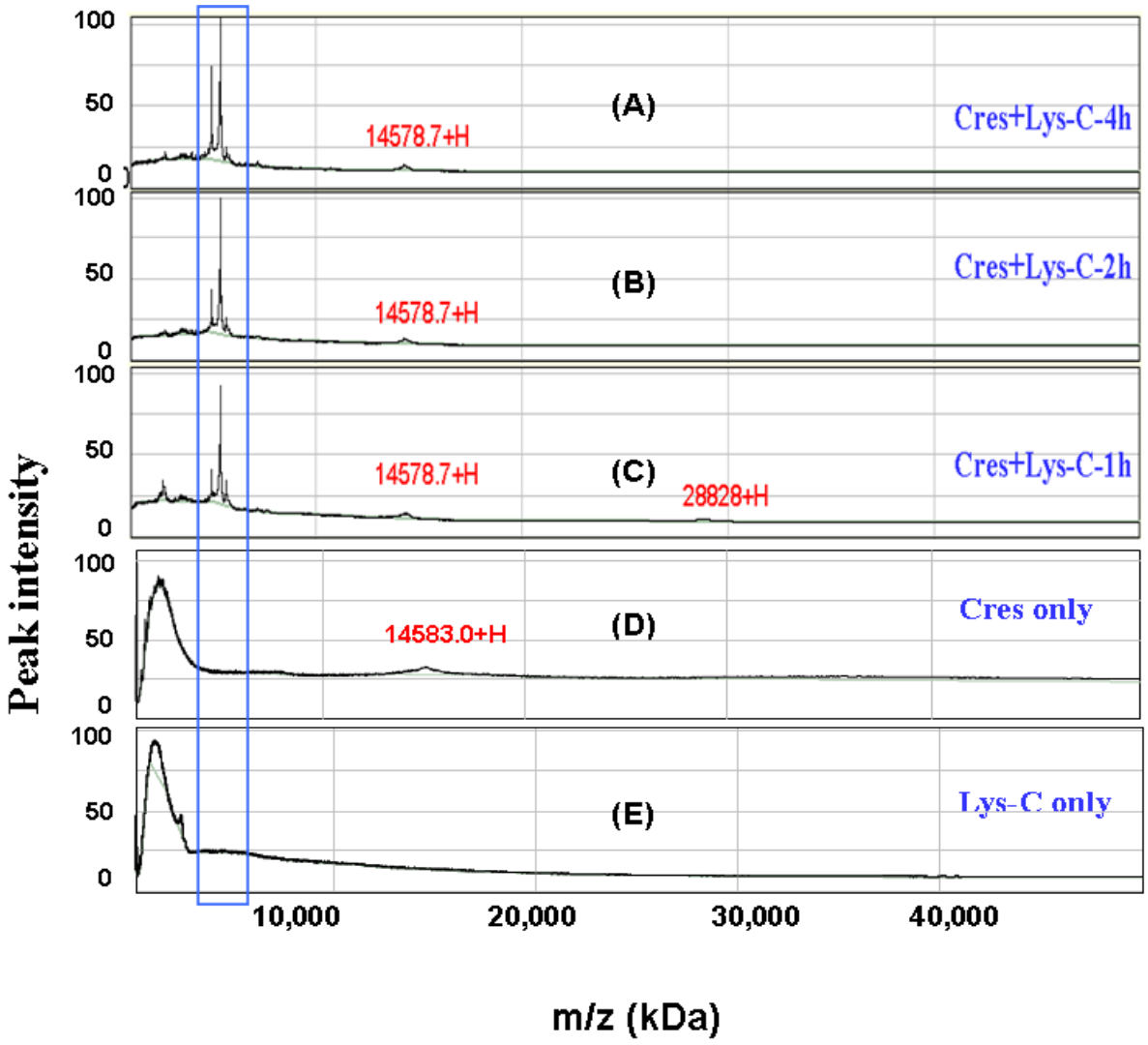
(B) Sequence of rec-mCRES (Without V5 and H6 tags)



mCRES seq = Accession no P32766

***Note**: Those Asn (N) residues that are glycosylated are shown with blue circle. The additional amino acids not present in mCRES sequence are shown within red box.*

<u>Figure 16(A)</u>: SELDI-TOF mass spectra of rec-mCRES protein following digestion with endoproteinase Lys-C enzyme. The blue box shown below indicates the region that has been further enlarged in the following figure to focus on the digested peaks obtained after Lys-C addition. As a representative figure, only the CRES + Lys-C (1 hour data) has been blown up in Figure 16(B).



<u>Figure 16(B)</u>: Expanded portions of SELDI-TOF mass spectra from the previously shown Figure 16(A) were shown here within the mass range 4,000 to 6,000 Da for CRES + Lys-C (1 hour data) digests and the controls that are CRES and Lys-C alone.

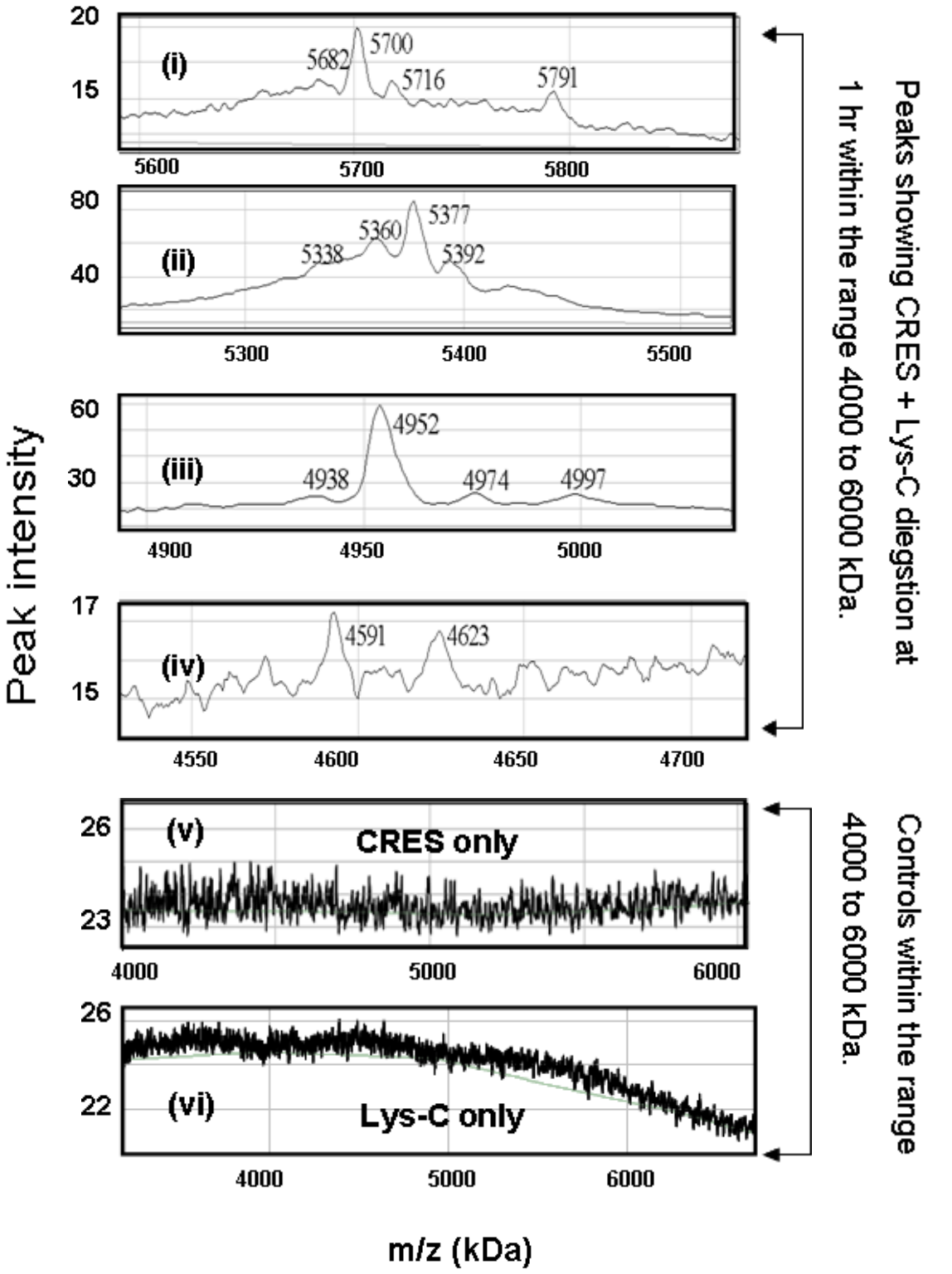
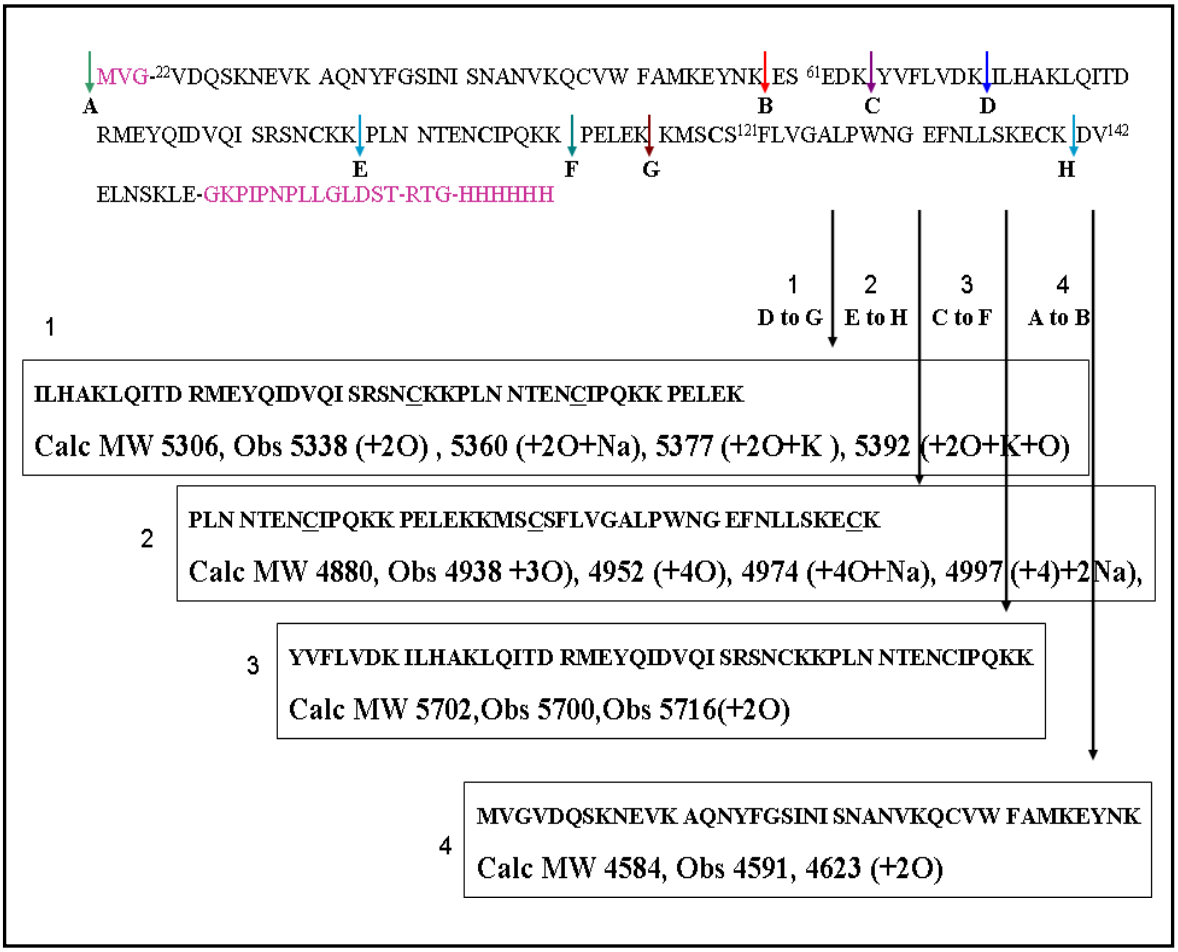


Figure 16(C): Possible explanation for the peaks obtained following digestion of rec-mCRES by digestion with Proteinase Lys-C as shown in mass spectrometrum in Figs 9 A&B. The various cleavage sites are indicated by vertical arrows and by letters "A, B, C, D, E, F, G and H as indicated below. Two letters shown as D to H, E to H, C to G and A to B indicated the fragments within those two cleavage sites.

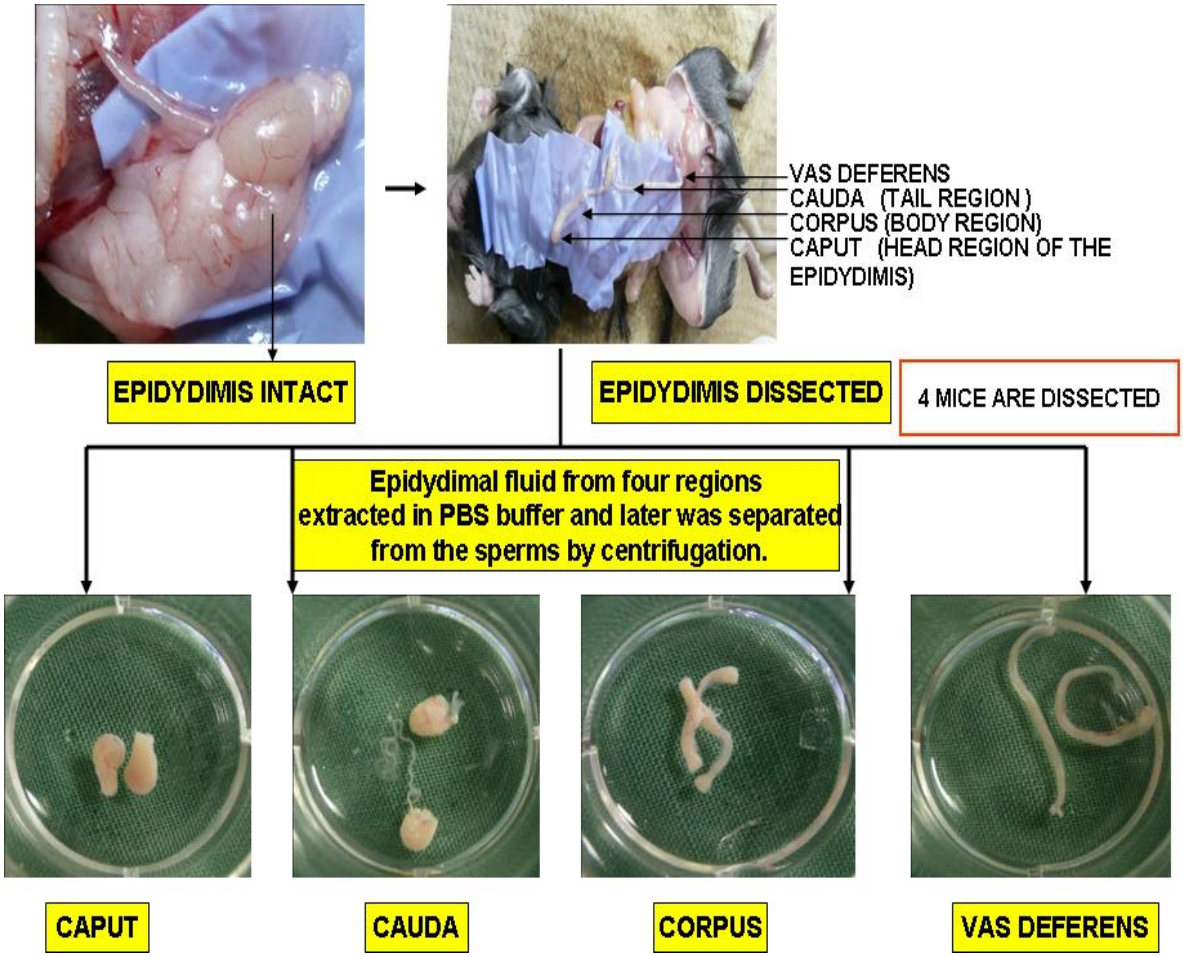


4.3. Examination of mouse epididymal fluid for PC4-protein and its activity

4.3.1 Mice dissection and epididymal sample collection

Four young adult male mice of similar age were sacrificed and dissected to collect their epididymis followed by their epididymal fluids and sperms. The epididymis was dissected further to be able to obtain its various parts separately i.e. caput, cauda, corpus and vas deferens as shown in **Figure 17**. Following approximate separation of various regions of mouse epididymis, their fluid contents were collected directly into PBS buffer previously cooled to 4^0C. Care was taken as much as possible to separate the sperms from the fluid. Each sample was then centrifuged at 4^0C to remove the sperms which will be present in the pellet. The clear supernatant liquids were collected, stored in cold ($<4^0$C) before any testing. The animal dissection work was done with the help of our senior technician Andrew Chen.

<u>Figure 17</u>: Dissection and extraction of the four epididymal regions of a male mouse



4.3.2 Confirming the presence of PC4 like activity in mouse epididymal fluids

After collecting fluids from various epididymal regions from four different mice, the presence of PC4 protein in them was confirmed by performing the western blot analysis using the PC4 specific antibody with dilution 1:1000. **Figure 18** shows the western blot results which not only confirmed the presence of PC4 protein but also showed that PC4 content was higher in the caput and corpus regions compared to those in cauda and vas deferens (5 µg total protein was loaded in each lane). In addition the presence of PC-like protease activity in these fluids was also examined since such activity is considered to be more crucial than the actual protein content so far as its function is concerned. Lower amount of PC4 protein may not necessarily mean less protease activity and vice versa. Hence the measurement of *in vitro* protease activity is so significant.

4.3.3 Comparison of total PC-like activity in the four epididymal regions of mouse.

The presence of the total PC-like activity was measured in the collected samples of epididymal fluid by using the fluorogenic Boc-RVRR-MCA at a final concentration of 50µM. The measured PC-activity along with their total protein contents estimated by BCA-method are shown in **Table 6**. The corresponding specific enzyme activity was then calculated and displayed in graphical form in **Figure 19** and **Table 7** which suggested that the total PC-like activity is highest in the caput region of epididymis compared to others.

<u>Figure 18</u>: Western blot results of various mouse epididymal fluid samples. The bands were not analyzed quantitatively at this stage.

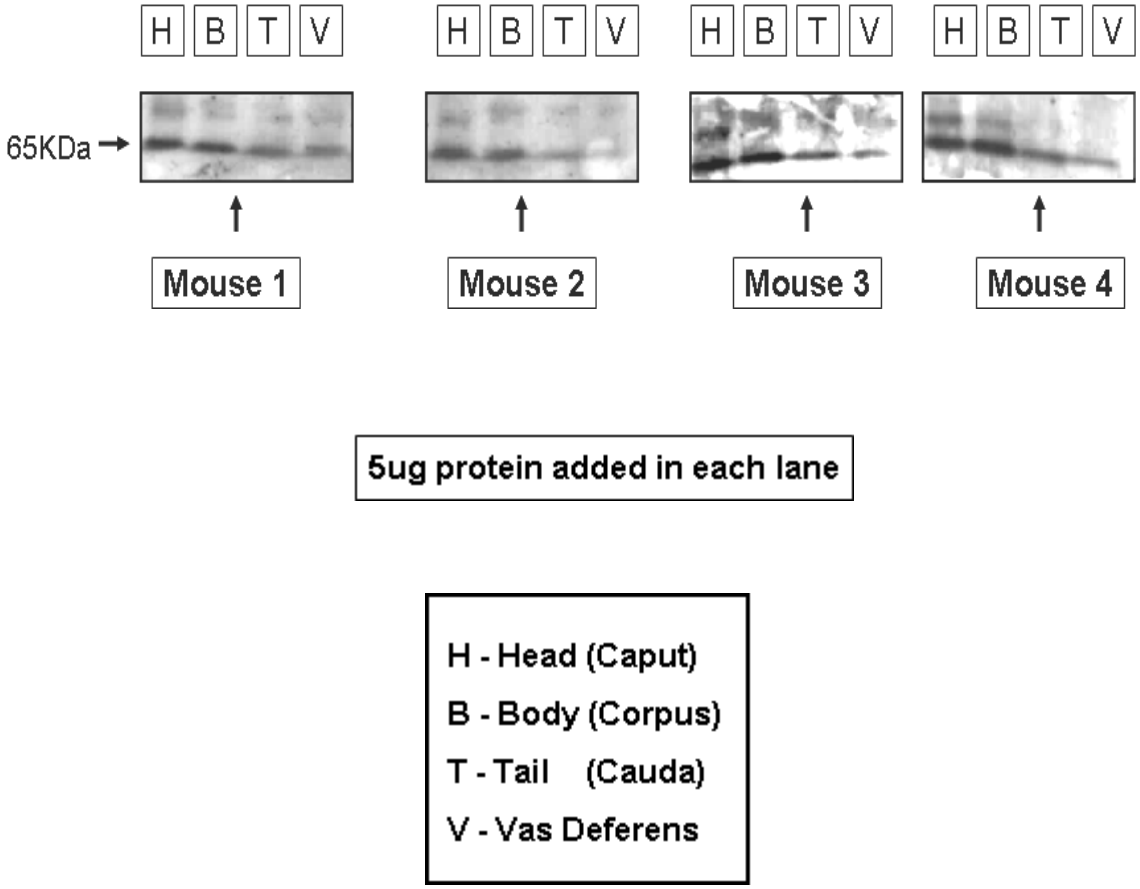


Table 6: Comparison of PC-like enzyme activity and total protein content of the four epididymal samples for each of the four mice:

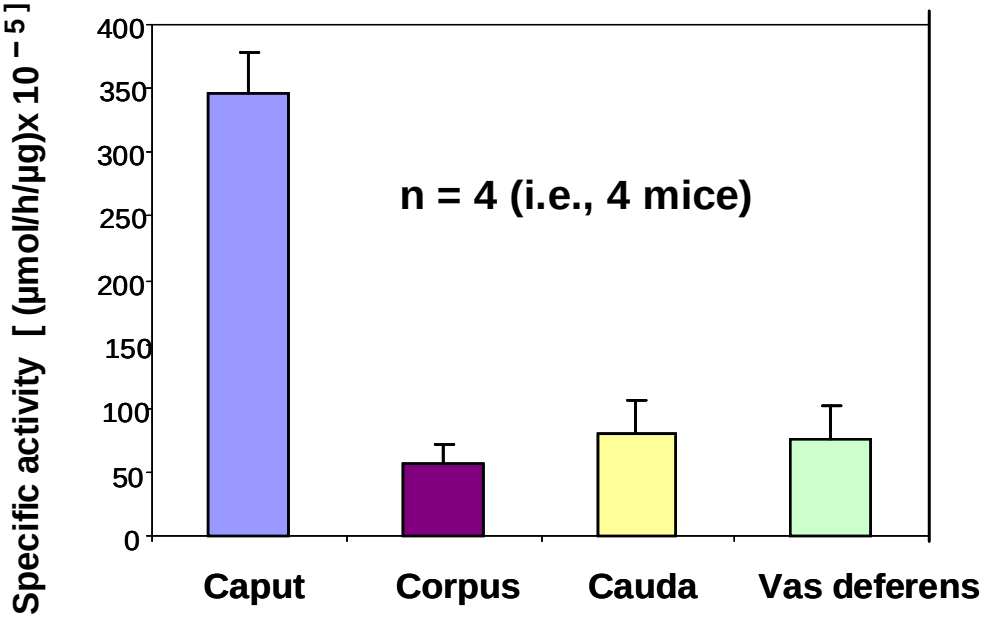
ENZYME ACITIVY (µmol/h) per µl	CAPUT (HEAD) [x 10 $^{-5}$] (µmol/h) per µl	CORPUS (BODY) [x 10 $^{-5}$] (µmol/h) per µl	CAUDA (TAIL) [x 10 $^{-5}$] (µmol/h) per µl	VAS DEFERENS (VD) [x 10 $^{-5}$] (µmol/h) per µl
MOUSE 1	413	46	33	66
MOUSE 2	280	14	21	31
MOUSE 3	303	13	24	42
MOUSE 4	353	22	20	66

PROTEIN CONTENT (µg / µl)	CAPUT (HEAD) (µg / µl)	CORPUS (BODY) (µg / µl)	CAUDA (TAIL) (µg / µl)	VAS DEFERENS (VD) (µg / µl)
MOUSE 1	1.080	0.581	0.360	0.646
MOUSE 2	0.812	0.269	0.264	0.706
MOUSE 3	0.864	0.313	0.227	0.702
MOUSE 4	1.160	0.456	0.455	0.707

Table 7: Comparison of PC-like activities based on specific activities of various epididymal fluis samples.

SP ENZYME ACITIVY [µmol/h/µg x 10^{-5}]	CAPUT (HEAD) [µmol/h/µg x 10^{-5}]	CORPUS (BODY) [µmol/h/µg x 10^{-5}]	CAUDA (TAIL) [µmol/h/µg x 10^{-5}]	VAS DEFERENS (VD) [µmol/h/µg x 10^{-5}]
MOUSE 1	382	79	92	102
MOUSE 2	345	52	79	44
MOUSE 3	351	42	105	60
MOUSE 4	304	48	44	93

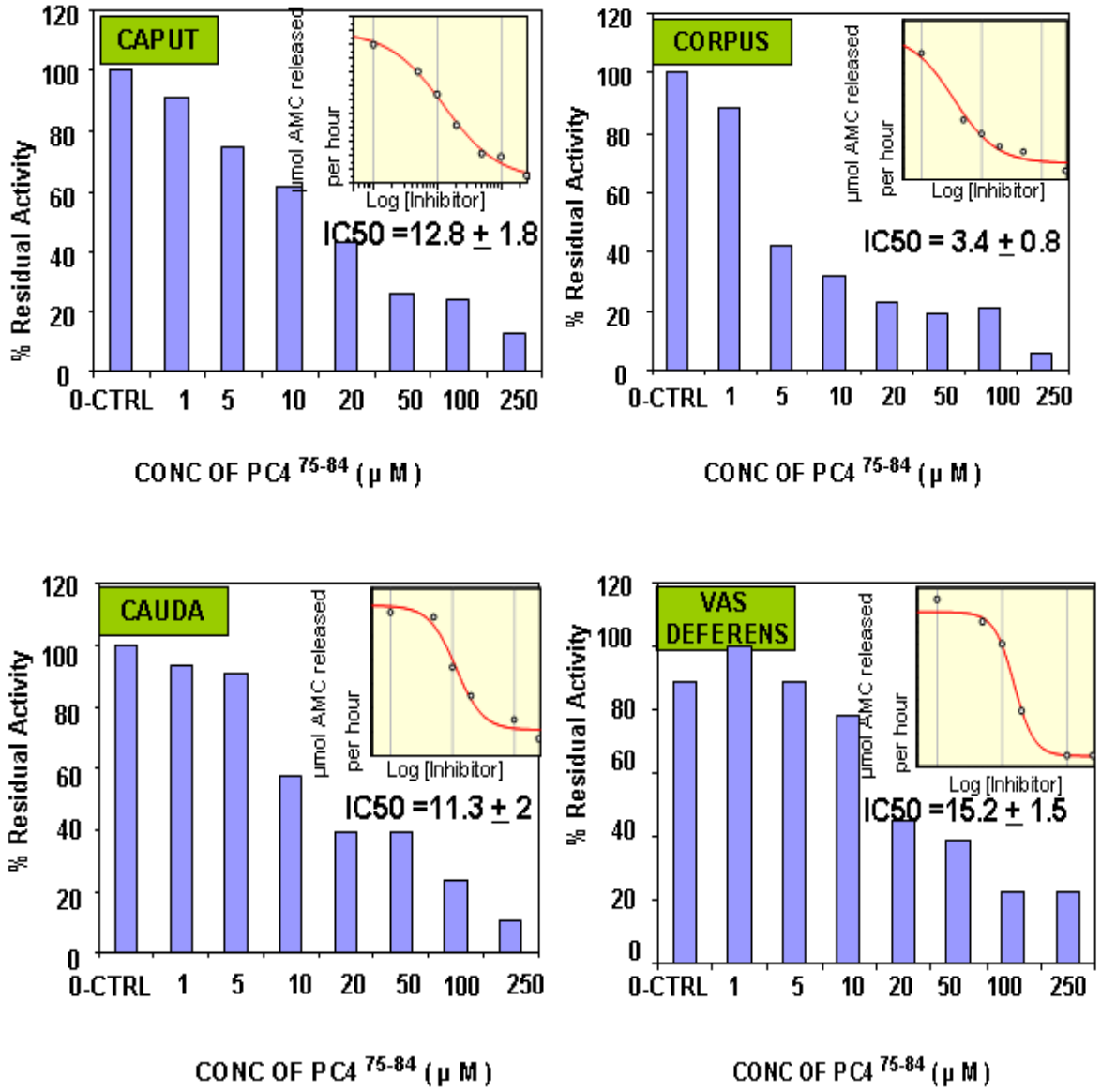
Figure 19: Specific activity determination of the four epididymal regions of the various mice used to compare the PC-Like activity in four regions.



4.3.4 Measurement of PC4-like activity in the four epididymal regions of each mouse

After confirming the presence of PC4 protein and PC-like activity in all four epididymal fluid samples of each mouse, we then focused to determine the PC4-specific activity using selective inhibitors of PC4 enzyme. The prodomain peptide $PC4^{75-84}$ is one such inhibitor which specifically inactivates PC4-activity. Various concentrations of this inhibitor ranging from 0, 1, 5, 10, 20, 50, 100, 250µM were used for this purpose with the fluorogenic substrate Boc-RVRR-MCA (final concentration 50µM). The approximate IC_{50} value of this inhibitor for the activity present in various regions of the epididymis was also calculated using the Grafit 4.0 software which showed the values to be as indicated below: **caput = 13, cauda = 11; vas deferens = 15 and corpus = 3.4µmol of AMC released per hour**. The data indicated that caput, cauda and vas deferens showed IC_{50} values that are comparable to one another. Previously it was shown that PC4 is present in all the four epididymal regions. It was also revealed that PC-like activity was highest in the caput region compared to others. Consistent with these observations we demonstrated the presence of significant PC4-like activity in all the four regions of the mouse epididymis and this is further shown in **Figure 20.**

Figure 20: Measurement of PC4 like protease activity in vitro in various epidydimal fluids using PC4 specific inhibitor ($PC4^{75-84}$ peptide). The amount of total protein present per 10 µl of samples (used in this activity assay) of caput, corpus, cauda and vas deferens are 1.136, 0.294, 0.504 and 0.40µg respectively; CTRL = Control.



4.3.5 Assessment of PC4-like activity in four epididymal regions.

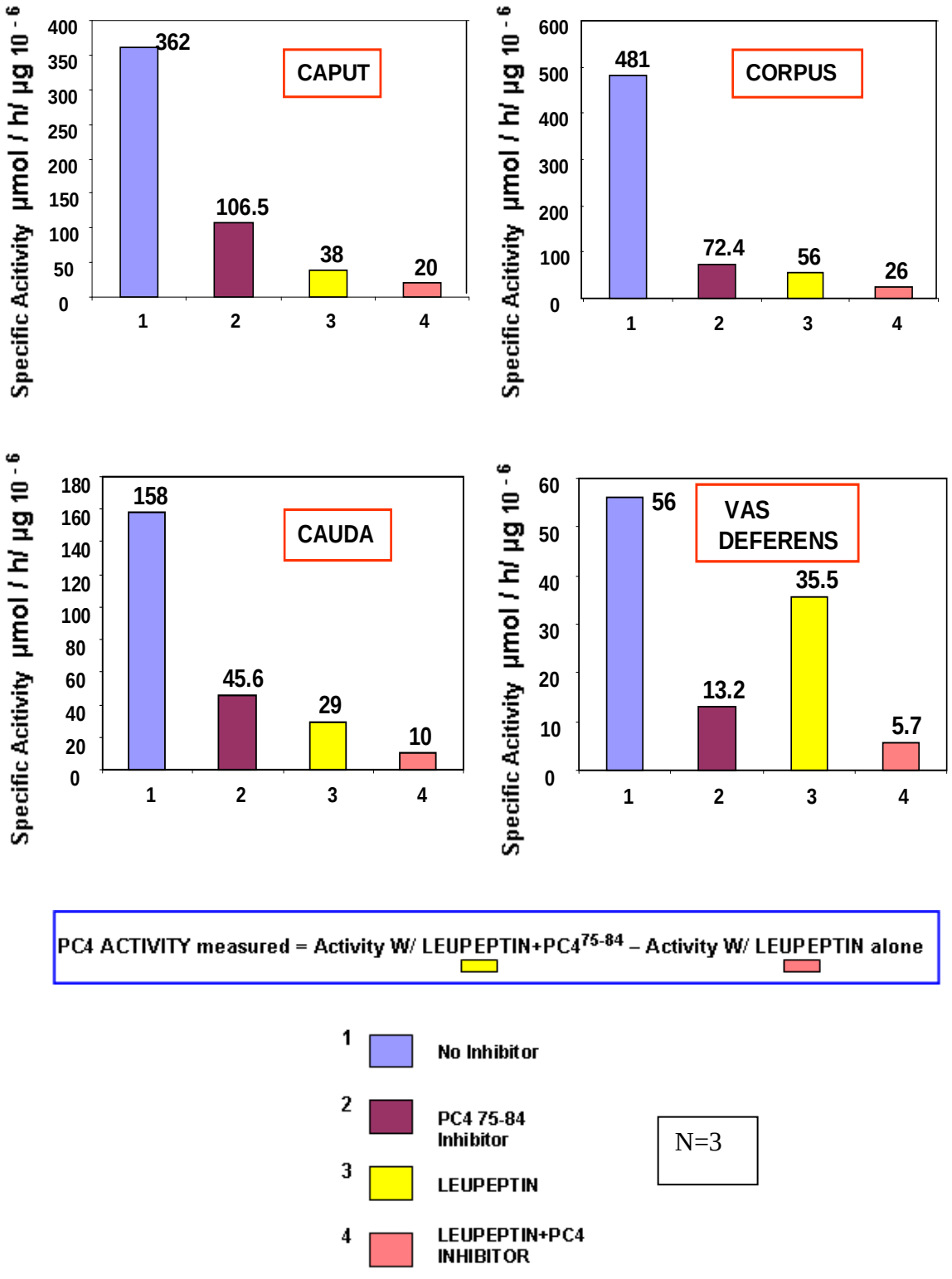
Following the results above, a semi-quantitative measurement of PC4-like activity in various epidydimal regions of mice (n = 3; i.e., three mice used) was accomplished by using leupeptin which is known to be inhibit only the serine proteases but not PC-like activity (i.e. non-PC serine protease inhibitor) as shown in **Figure 21**. Substrate used was Boc-RVRR-MCA (final concentration of 50uM) and the fluorometer was used to measure the RFU values which were obtained on per hour basis. The activity specific to PC4 were measured using the following equation:

PC4-like activity = (RFU value in presence of leupeptin + $PC4^{75-84}$) – (RFU value in presence of leupeptin alone)

Using the above mentioned method, the PC4 activity of each of the four epididymal regions was calculated and the data is presented in **Table 8**. It was observed that the PC4 like activity is found to be low in the caput region (18μmol/h/μg x 10^{-6}) followed by an increase in the corpus region (30μmol/h/μg x 10^{-6}) and again a decrease in the cauda (19μmol/h/μg x 10^{-6}) and lastly an increase again in the vas deferens fluid (29.8 μmol/h/μg x 10^{-6}).

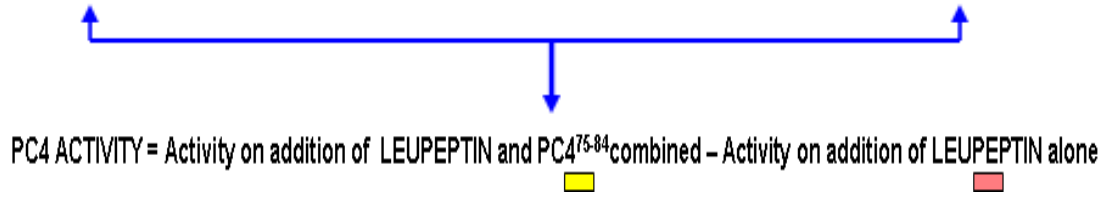
Here to make the link between PC4 and CRES, it may be mentioned that CRES protein was shown to be present in highest amount in the caput region compared to other regions of the epidydimis (125) and therefore if CRES is an inhibitor of PC4 activity one would expect lower PC4 activity in the caput which is exactly opposite to what we observed. This can be explained by the fact that PC4 exists in both membrane bound and unbound form and many other factors like CRES aggregation, etc. This issue will be further explained later in more detail in our discussion section.

Figure 21: Assessment of PC4-like activity in the various epididymal regions



<u>Table 8</u>: Following table was drawn from the results obtained previously showing the comparison of the four epididymal regions in terms of their PC4-Like activity

PC4-like activity based on Leupeptin and PC4 Peptide inhibition data in Epididymal regions [µmol/h/µg * 10^(-6)]			
CAPUT	CORPUS	CAUDA	VAS DEFERENS
38-20 = 18	56-26 = 30	29-10 = 19	35.5-5.7 = 29.8



PC4 ACTIVITY = Activity on addition of LEUPEPTIN and $PC4^{75-84}$ combined – Activity on addition of LEUPEPTIN alone

4.4 Inhibitory effect of rec-mCRES on PC4 in *in vitro* and *ex vivo* conditions

4.4.1 Effect of rec-mCRES dimer on PC4 activity *in vitro*

The major goal of this thesis was to demonstrate whether rec-mCRES is an inhibitor of PC4 activity. Hence first, we studied this effect *in vitro* using both purified dimeric and monomeric rec-mCRES protein with our purified rec-rPC4 enzyme. Various concentrations of purified CRES protein were used to test their inhibitory effects on rec-rPC4 enzyme using three different concentrations of fluorogenic substrate Boc-RVRR-MCA namely 50µM, 25µM and 12.5µM. Firstly, the concentration range of rec-mCRES used was from low nM to relatively high µM range. Based on preliminary experiments a more appropriate concentration range of rec-mCRES from 0-100µM namely 0, 0.5, 1,5, 10, 20, 30, 40, 50, 60,75 and 100 µM was used (**Figure 22**).The IC_{50} and K_i values were then obtained from measured sigmoidal graph and Dixon plots respectively.

4.4.1.1 IC_{50} value determination for inhibition of PC4 activity by rec-mCRES dimer

We obtained three IC_{50} values for the three substrate concentrations for the inhibition *in vitro* of PC4 activity by rec-mCRES dimer *(For detailed procedures see Materials and Methods).* Following is the summary of results based on plots in **Figure 23**:

IC_{50} for $[S]_1$, = 31.3 $\pm$ 3 (S = 12.5 µM)

IC_{50} for $[S]_2$ = 25.19 $\pm$ 4 (S = 25 µM)

IC_{50} for $[S]_3$ = 24.5 $\pm$ 4 (S= 50 µM)

4.4.1.2 K_m value determination for purified rec-rPC4 sample

Following the determination of IC_{50} values, K_m value was determined for our recPC4 enzyme used in all the experiments. The K_m value was obtained using Grafit 4.0 software

based on a fluorometric assay using various Boc-RVRR-MCA concentrations along with a fixed amount of rec-rPC4 enzyme. The K_m value directly measured from this graph was found to be:

K_m= 12.9 ± 2µM as shown in **Figure 24**.

Based on the measured K_m value, K_i values were calculated for the above three substrate concentrations using the equation below valid for true competitive inhibitor,

$$K_i = IC_{50}/ \{1+ ([S]/ Km)\}$$

Based on the above equation three values were obtained namely,

K_i (1) = 15.8µM

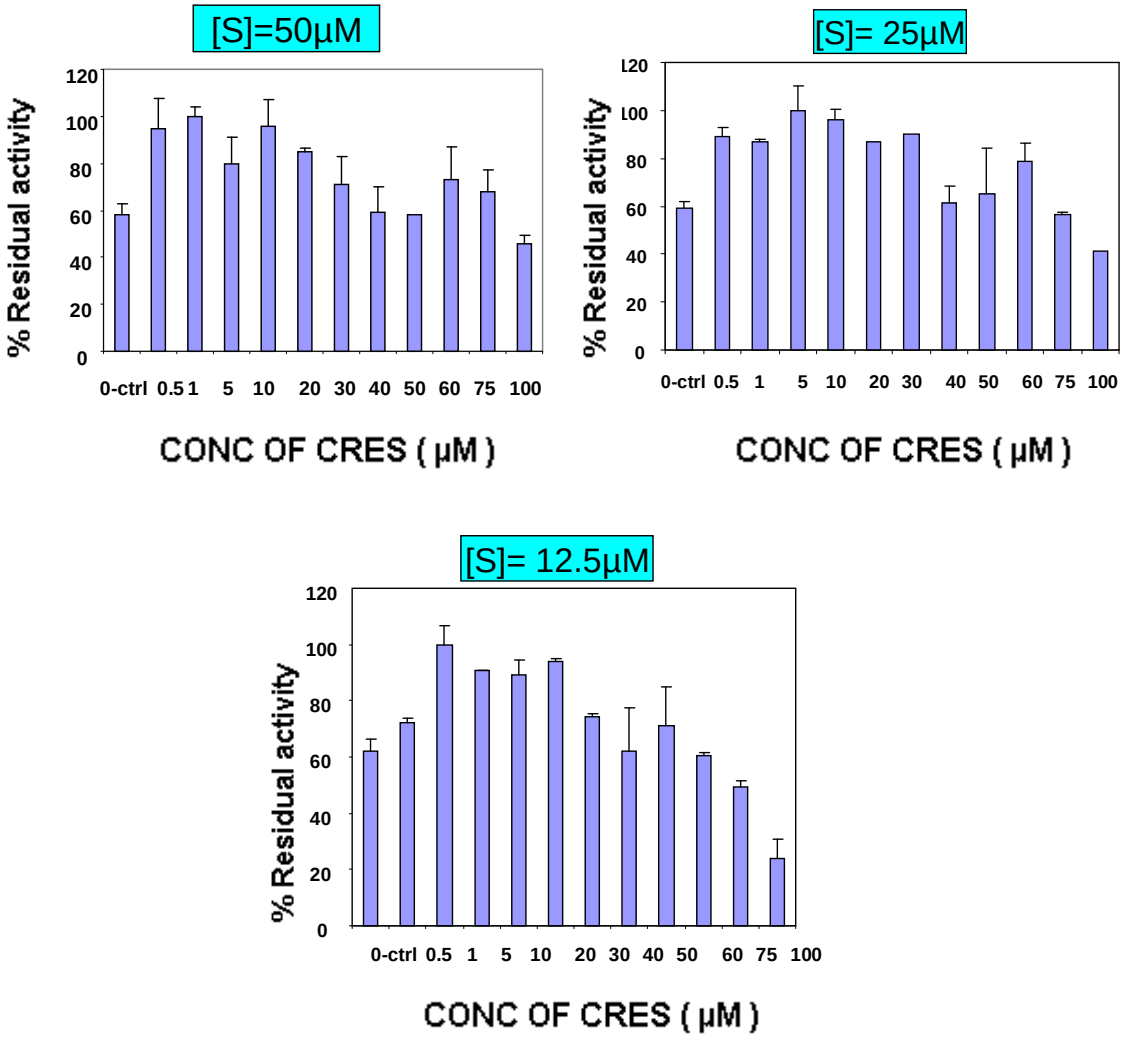
K_i (2) = 8.56µM

K_i (3) = 10.4µM

Taking an average of the above three values, we found K_i average = **11.58µM**

Hence based on true competitive inhibition, calculated K_i for inhibition of PC4 by dimeric rec- mCRES protein will be **11.58µM**. This value was compared with the K_i value determined by Dixon plot (i.e. **8.6µM**) as shown in **Figure 25(A)**.

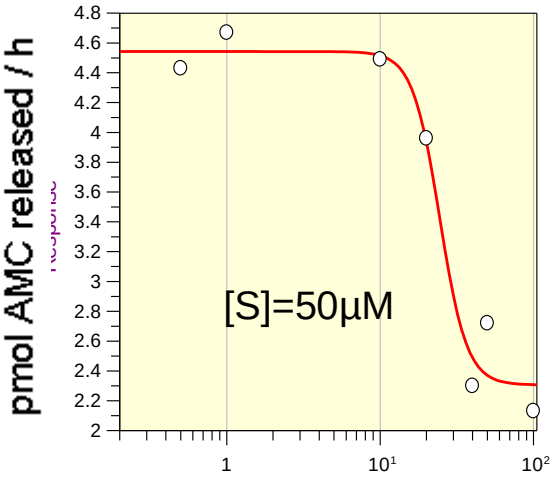
Figure 22: Inhibitory effect of rec-mCRES dimer on PC4 activity shown with three substrate (Boc-RVRR-MCA) concentrations of 12.5, 25 and 50µM



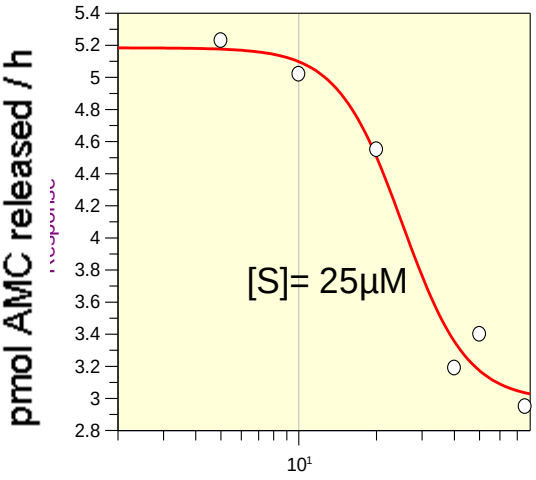
[S]=Boc-RVRR-MCA

N= 2

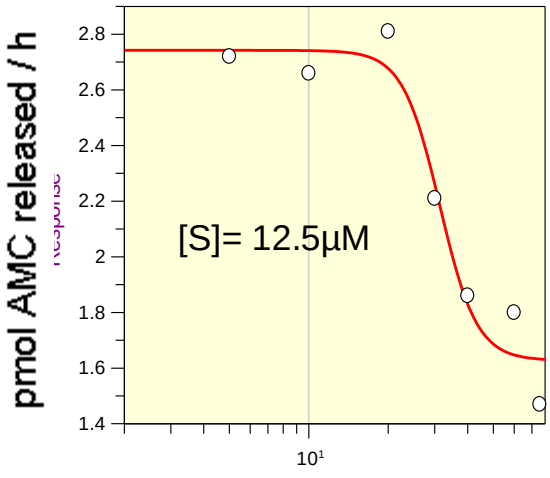
Figure 23: IC_{50} value determination for inhibition of PC4 activity by rec-mCRES dimer at various substrate concentrations



Parameter	Value	Std. Error
Y Range	2.2372	0.3157
IC 50	24.4593	4.2754
Slope factor	4.9152	2.7156
Background	2.3053	0.2557



Parameter	Value	Std. Error
Y Range	2.1980	0.4059
IC 50	25.1986	4.5523
Slope factor	3.4646	1.6372
Background	2.9862	0.3063

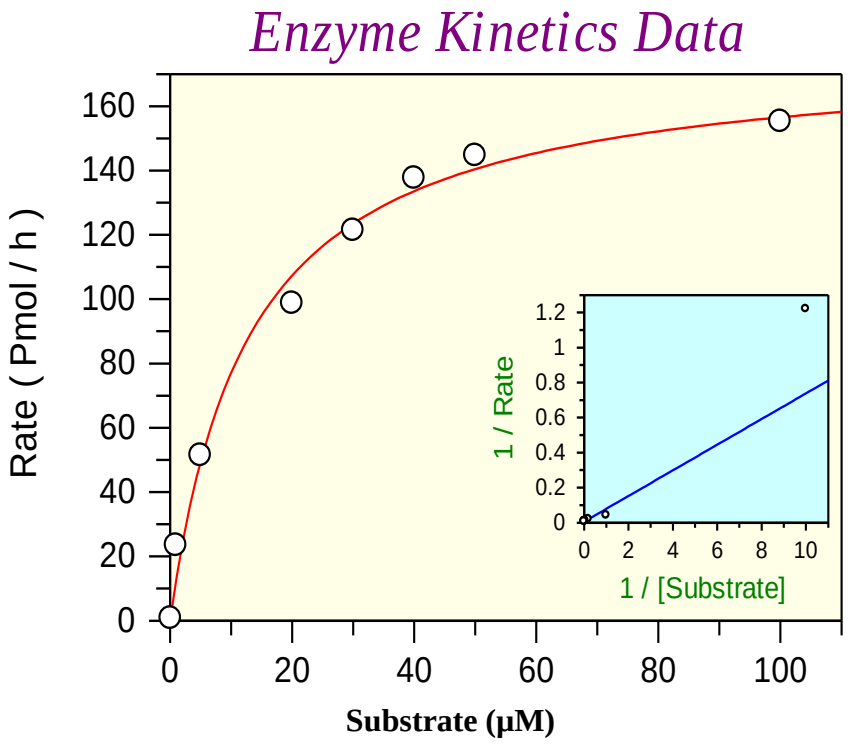


Parameter	Value	Std. Error
Y Range	1.1146	0.1755
IC 50	31.3497	3.0846
Slope factor	6.1879	4.1916
Background	1.6276	0.1287

Summary showing IC_{50} values for three substrate concentrations:

SUBSTRATE CONCENTRATION	IC_{50} VALUE µM
S1 = 50 µM	24.45 ± 4.2
S2 = 25 µM	25.19 ± 4.5
S3 = 12.5 µM	31.34 ± 3.0

Figure 24: K_m value determination of our purified rec-rPC4 sample using Boc-RVRR-MCA substrate



Parameter	Value	Std. Error
Vmax	176.8374	8.0476
Km	12.9539	2.1458

4.4.1.3: Determination of K_i for PC4 inhibition by rec-mCRES dimer graphically: Dixon plot.

The inhibition constant K_i value for inactivation of PC4 *in vitro* by rec-mCRES dimer was measured using Dixon plot at three different concentrations of the substrate Boc-RVRR-MCA. The graph (1/velocity vs inhibitor concentration) was plotted using the graphical software Sigma Plot version 11.0 (as shown in **Figure 25A**).

The equations for three linear graphs obtained were deduced by Sigma Plot as indicated below:

For S1 (12.5µM),	**Y = 4.460 X + 2.023**	**(equation 1)**
For S2 (25µM),	**Y = 0.02 X + 2.12**	**(equation 2)**
For S3 (50µM),	**Y = 0.077 X + 2.67**	**(equation 3)**

Based on these equations, three intersection points for each pair of lines were calculated.

For eqns 1 & 2, X = - 5.8, hence K_i (1) = 10.72µM
For eqns 2 & 3, X = - 11, hence K_i (2) = 26.6µM
For eqns 3 & 1, X = - 9, hence K_i (3) = 8.39µM

Hence measured average K_i = [K_i (1) + K_i (2) + K_i (3)]/3 = <u>8.6µM</u>

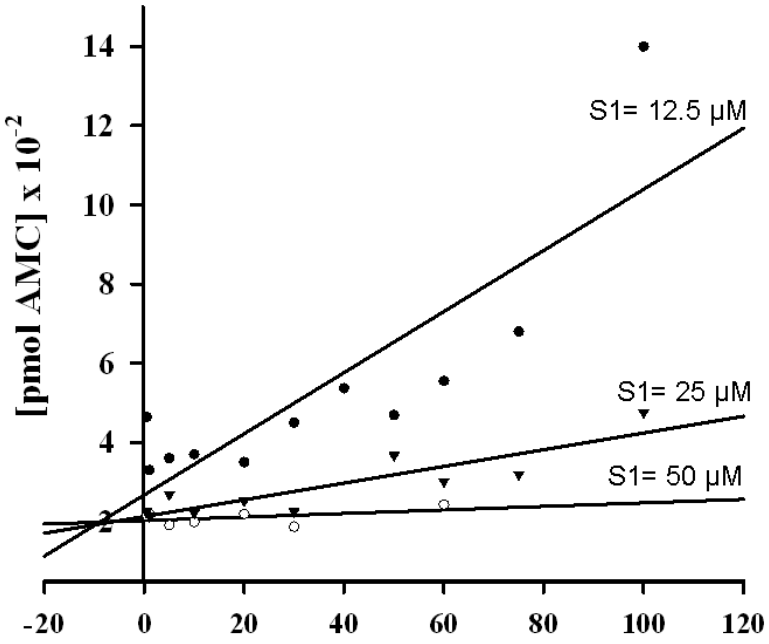
Thus our theoretically determined K_i value for a true competitive inhibition is **<u>11.53µM</u>** and graphically determined K_i value is **8.6µM.** These values are relatively close to one another suggesting that PC4 inhibition by rec-mCRES dimer is largely competitive in nature meaning that it binds mostly to the catalytic site of the enzyme. This is further confirmed by Lineweaver-Burk plot as shown in the section below.

Lineweaver-Burk plot: The inhibition of PC4 activity *in vitro* by rec mCRES was further examined by plotting a double reciprocal (i.e. 1/Velocity versus 1/Inhibitor concentration)

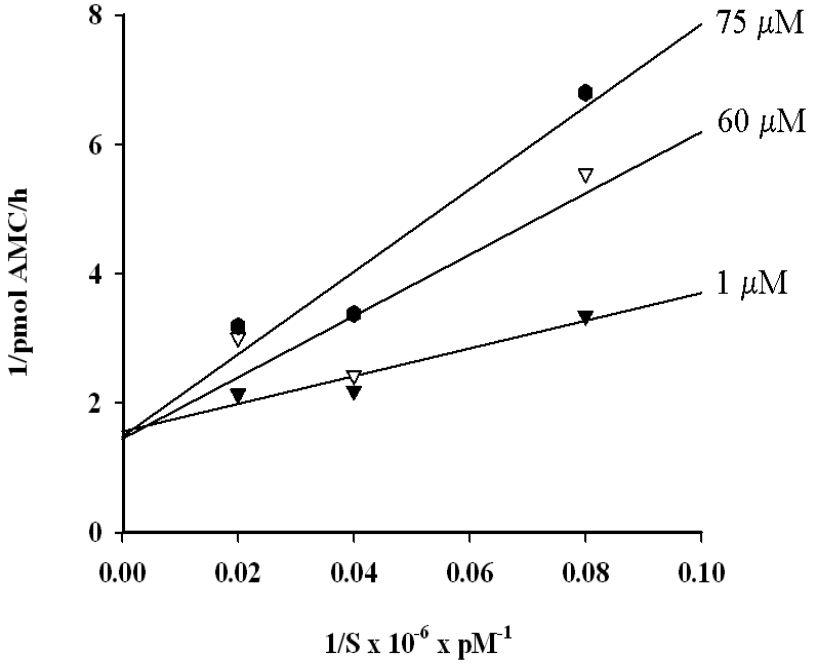
curve as shown in **Figure 25(B)**. Herein three inhibitor concentrations (i.e. rec-mCRES) namely 1, 60 and 75µM and three substrate concentrations (i.e. Boc-RVRR-MCA) namely 12.5, 25 and 50µM were used for the study. A common intersection point of three lines which merged on Y-axis is noticeable and this supports the competitive type inhibition as mentioned earlier.

Figure 25: **(A) Dixon Plot showing K_i value for PC4 inhibition by rec-mCRES dimer at three different substrate concentrations (S1, S2, S3) as indicated in the graph. (B) Lineweaver-Burk plot showing near competitive inhibition of PC4 by rec-mCRES dimer. Three concentrations (1, 60 and 75μM) of rec-mCRES were used as ndicated.**

(A)



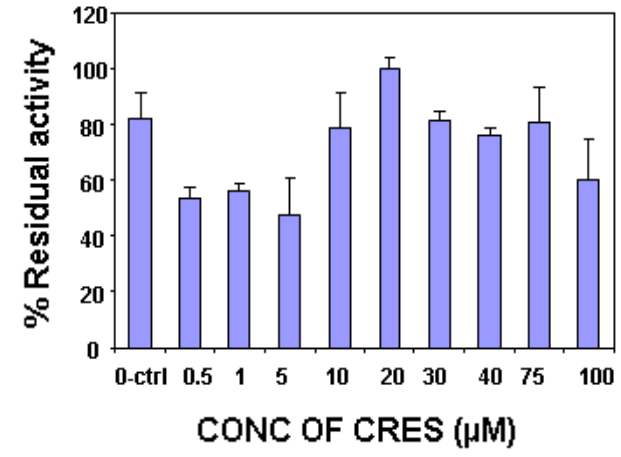
(B)



4.4.2 Effect of rec-mCRES monomers on PC4 activity:

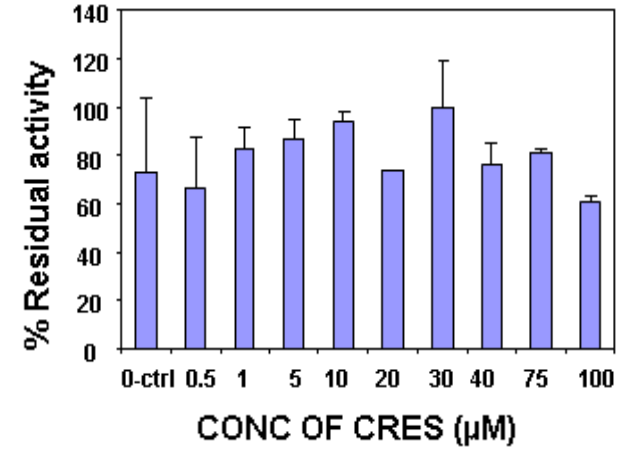
Following the above study on the inhibition of PC4 activity by **rec-mCRES dimer**, we decided to examine the effect of **monomeric rec-mCRES protein** on PC4 activity using the the same substrate and similar condition. The results are shown in **Figure 26 ,** which revealed very weak PC4-inhibitory property for the monomer rec-mCRES and the IC_{50} value is relatively high (*not measured*).

<u>Figure 26</u>: Regulatory effect of rec-mCRES monomer on PC4 activity *in vitro* .Three substrate [S] concentrations are shown here with various rec-mCRES concentrations.

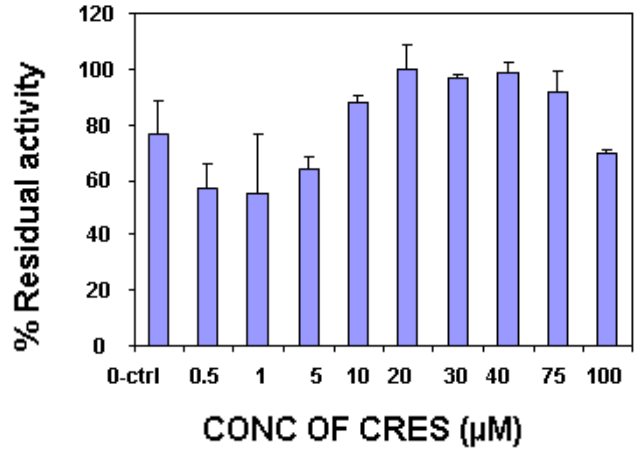


[S]=Boc-RVRR-MCA

[S]=50µM



[S]= 25µM



[S]= 12.5µM

4.4.3 Cleavage of rec-mCRES protein *in vitro* by rec-PC4 enzyme: A time dependent digestion study

Since we know that the general mechanism of protease inhibition by a serpin involves an initial cleavage of the serpin itself as explained in the introduction section (shown in **Figure 4),** we decided to examine whether or not CRES mediated inhibition of PC4 also follows a similar mechanism involving an initial cleavage followed by covalent binding. This was done by performing a time dependent digestion study by incubating rec-mCRES protein with purified active rec-rPC4 enzyme. The SELDI-TOF-MS of the crude digest at various time intervals like 0h, 1h, 2h and 4h using sinapic acid (SPA) as energy absorbing matrix revealed additional peaks that suggests a cleavage. In fact several mass spectra peaks were observed as shown in **Figure 27(A).** The peaks were analysed and the results are shown in **Figure 27(B).** There were several peaks observed upon PC4 addition to the rec-mCRES protein that were not present in the mass spectrum of the controls (i.e. rec-mCRES alone and rec-rPC4 alone).The peaks observed are as follows: **m/z 28,766 Da** (CRES dimer), **14,383 Da** (CRES monomer), **7273 Da** (M/2 peak of the CRES monomer), **9656 Da** (M/3 peak of CRES dimer), **9226 Da (the N-terminal fragment following cleavage at the site NC<u>KK</u>↓PLN** as explained in **Figure 27(B).** The peak at **4,613 Da** is the M/2 of the above cleaved fragment **m/z 9,226 Da**). These peaks were found to be highly prominent in **Figure 27(A).** .Hence based on the above MS data, it was proposed that PC4 cleaves rec-mCRES at the C-terminus of the double lysine residues namely **KK↓P** which is a very unique and rare cleavage by an enzyme. This aspect was discussed in more detail later in the Discussion section.

<u>Figure 27:</u> (A) Mass spectrum analysis of rec-CRES by rec-PC4 addition by performing a time dependent digestion study at various time intervals. (B) Controls showing CRES and PC4 sample only indicated the absence of cleavage peaks observed in (A) which was the addition of CRES to PC4.The focus below is on the CRES+PC4 (1 hour data) where the peaks 1-6 have been assigned and explained in details in the following Figure 27(B).

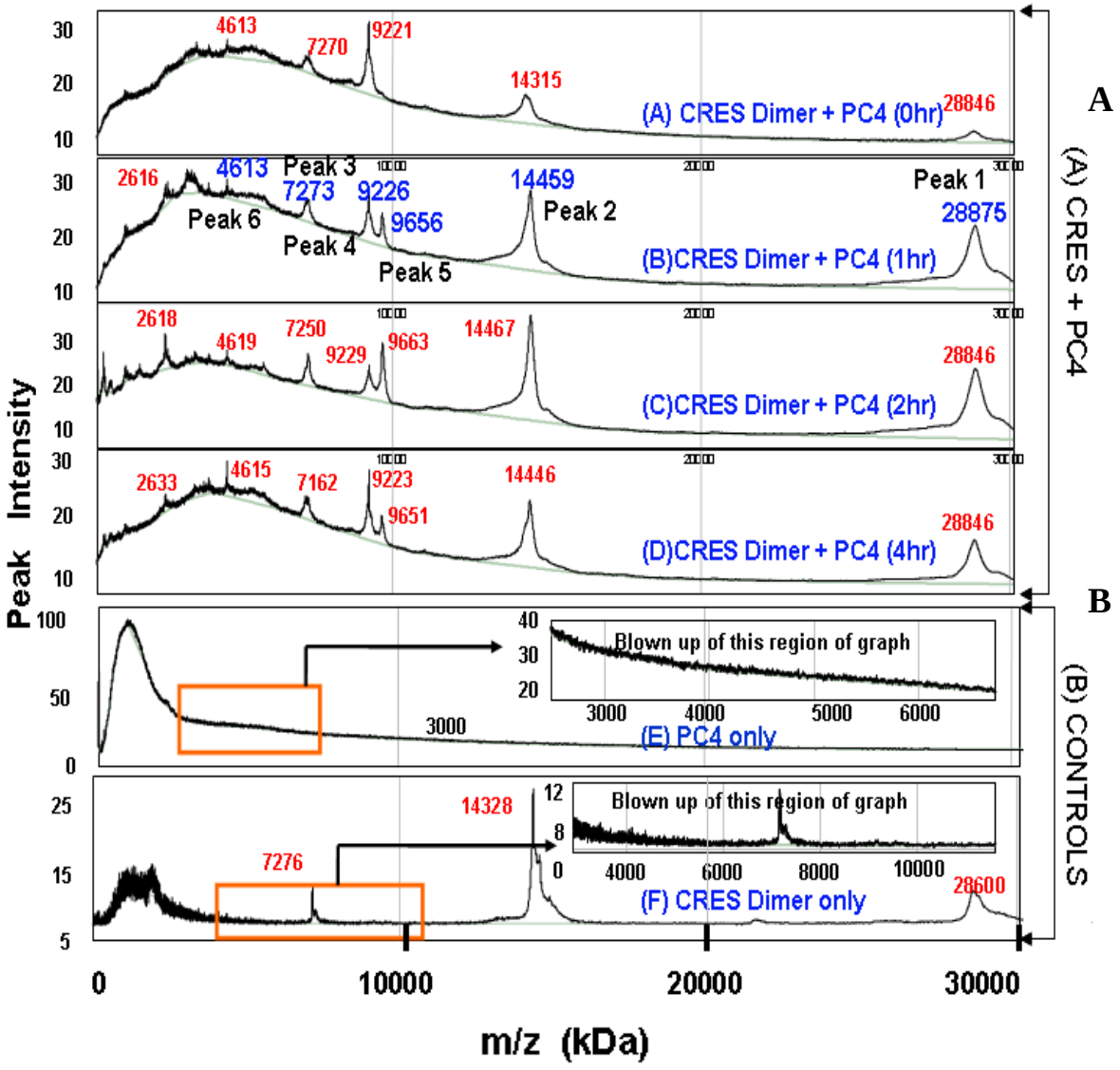


Figure 27(B): Cleavage analysis of rec-mCRES sequence by rec-PC4 enzyme

Amino acid sequence of rec-mCRES

MVG-^{22}VDQSKNEVK AQNYFGSINI SNANVKQCVW
FAMKEYNKES ^{61}EDK YVFLVDK ILHAKLQITD
RMEYQIDVQI SRSNCKK ↓ PLN NTENCIPQK K
PELEK KMSCS^{121}FLVGALPWNG EFNLLSKECK
DV^{142} ELNSKLE-GKPIPNPLLGLDST-RTG-HHHHHH

In the above figure, pink color sequences indicate the linkers followed by green indicating the V_5 tag and lastly the blue color for His_6 tag.

Observed Peak Assignments

Peak 1. MW=28,766 (Cres dimer peak undigested)

2. MW=14,383 (Cres monomer peak undigested)

3. MW= 7273(~M/2 peak of Cres monomer,i.e.,14383)

4. MW=9,226 (Cleaved product peak, shown underlined) (The corresponding other cleaved product is not clearly seen as it may be getting degraded even further which makes us observe so many other small peaks)

5. MW=9656(~M/3 peak of CRES dimer)

6. MW=4,613 (M/2 peak of the cleaved product,i.e.,~9226/2)

4.4.4 Inhibition of PC4 *ex vivo* by rec-mCRES protein

4.4.4.1: Effect of rec-mCRES monomer on human pro-IGF2 processing by PC4 in the placenta cell line

Following our *in vitro* results as indicated above, we became interested to study the effect of rec-mCRES protein on PC4 processing in the cell lines and for this we selected human trophoblast placenta cell line which expresses both PC4 and its substrate pro-IGF2 and therefore represent an excellent model for the study. As described in the *Materials and Methods*, various concentrations rec-mCRES (ie. 16.6µM, 6.7µM, 3.5µM, 1.6µM and 0 µM which was the control) were added to the placenta trophoblast cell line in culture along with the addition of a constant amount of our purified rec-rPC4 enzyme. In parallel a control experiment was also performed where no rec-mCRES was added. The various culture media samples from the 6-well plates were then collected (each well consisting of $1x10^6$cells/ml, passage into the well plate was done after counting with hemato-cytometer). Western blots were performed on these samples using IGF-2 antibody with dilution 1:500 *(see Materials and Methods for details).* The results are shown in **Figure 28(A)** which suggests that as expected the mature form of IGF-2 (~8 kDa) is produced efficiently in the absence of CRES protein. This is produced from either the intermediate from of IGF-2 called big IGF2 (~16kDa) or directly from the pro-IGF2 (~24 kDa) (53). With the addition of increasing amounts of rec-mCRES protein, one can notice the formation of increasing amounts of both big IGF2 as well as pro-IGF2. This suggested that CRES blocks both PC4 (present endogenously) mediated cleavages of big- and pro-IGF2 into its mature form. This led to the formation of more of the precursor bands (i.e. 16 and 24 kDa) in the presence of increasing quantities of CRES protein. As expected in the control lane, where we have not added any

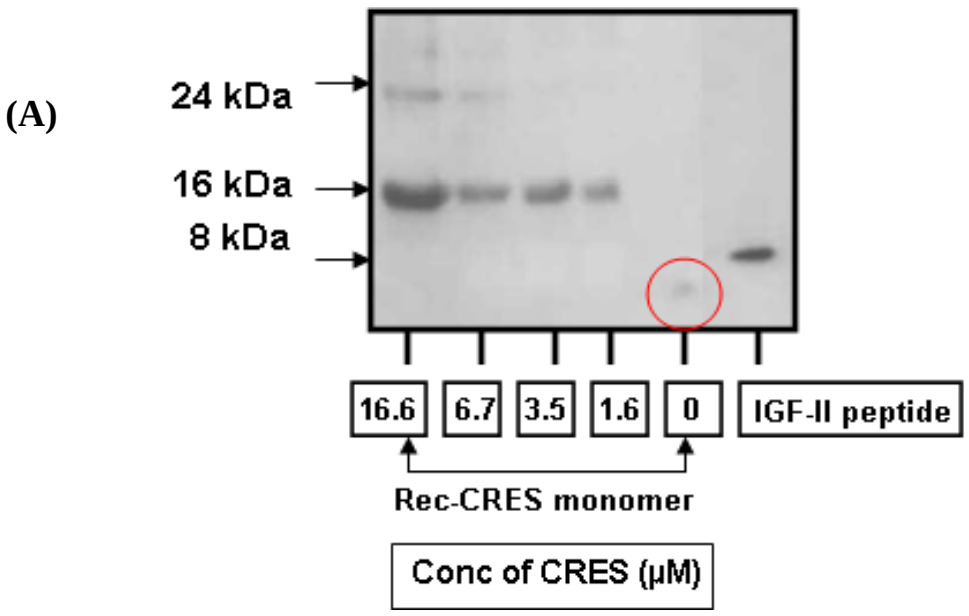
CRES protein, we noticed the formation of mature IGF-2 form (i.e. 8kDa) which was not found in any other lanes. Our conclusion is supported by the use of a commercially bought IGF2 protein which showed a strong band at ~8 kDa. Unfortunately the endogenously produced mature IGF2 could not be well detected by the IGF2 antibody used in this study (except for the one control lane) although the synthetic mature IGF2 was easily detectable in the western blot (as shown in **Figure 28A**). Despite this short-coming the above conclusion still holds true. In future other commercially available IGF2 antibody will be used for further confirmation of this result. This experiment was repeated three times under identical condition and only a representative figure was shown in the thesis.

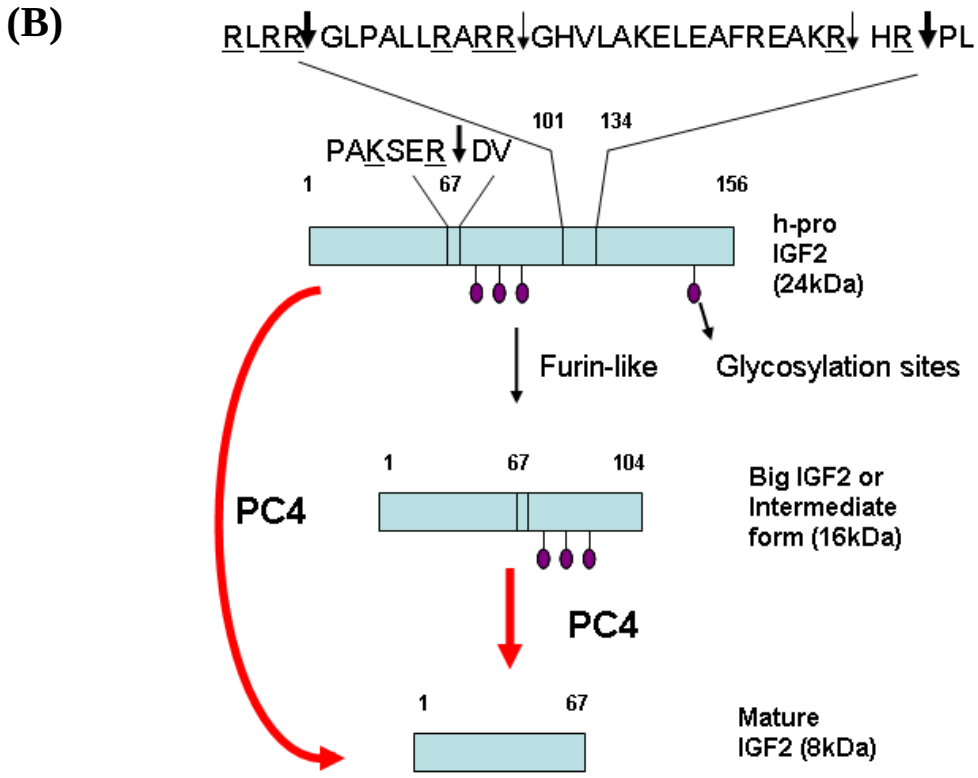
4.4.4.2 Effect of rec-mCRES mono/dimer on h-proIGF2 processing by PC4 in a placenta cell line in presence of added PC4 enzyme.

The above experiment was carried out again in a similar manner in placenta cell *(details in Materials and methods)* line but with no added rec-rPC4 enzyme. The result obtained was the same as the previous experiment and it is displayed in **Figure 29**. In this case we tested the effects of both CRES monomer and dimer separately. Addition of rec-mCRES dimer resulted in an increased appearance of full length precursor form pro-IGF2 (i.e. 24kDa) but not the intermediate form big IGF2 (i.e. 16kDa). The addition of rec-mCRES monomer on the other hand led to the appearance of both intermediate form and the full length pro-IGF2 form. These interesting observations suggested that dimeric rec-mCRES is an inhibitor of PC4 and blocks efficiently the cleavage of pro-IGF2 directly into mature IGF2 but failed to block the conversion of big-IGF2 into mature IGF2. On the other hand the monomeric form of rec-mCRES blocked PC4-mediated cleavages of both pathways. This conclusion was based on three repeat experiments but only a representative example was shown as a figure

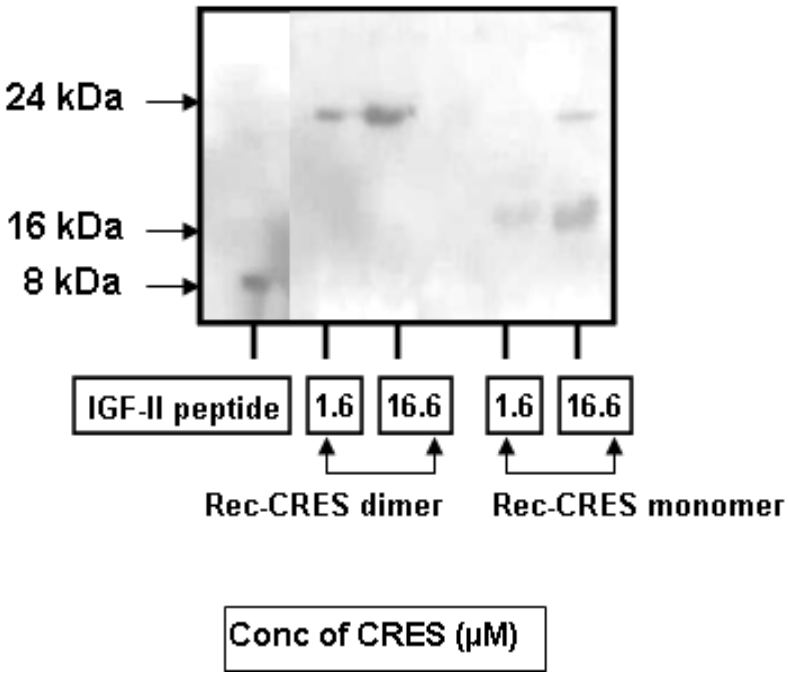
in the thesis. **It indicates that there is an important role of aggregation for CRES protein's ability to inhibit PC4 activity.**

Figure 28: (A) Effect of rec-mCRES monomer on h-pro-IGF2 processing by PC4 in a human placenta cell line; (B) Schematic diagram of h-proIGF2 with various potential PC-cleavage sites as indicated by vertical arrows.





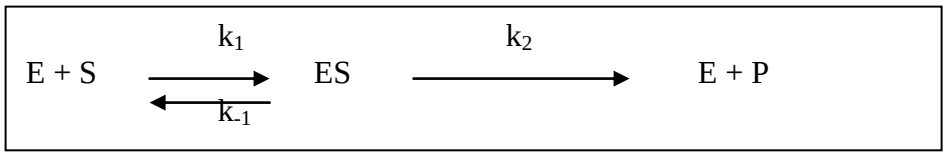
<u>Figure 29</u>: Effect of the rec-mCRES dimer on h-proIGF2 processing by PC4 in the placental cell lines



5. DISCUSSION

5.1 PC4 production

Rec-rPC4 was successfully produced in enzymatically active form based on available published protocol using the *L. tarentolea* expression system with slight modification (57). Earlier as reported in the above publication, the PC4-purification was accomplished by using two columns in series namely DEAE-agarose and arginine sepharose 6B columns. Purified PC4 enzyme as obtained above in this article exhibited a K_m value = 91.7 + 12.0 µM, V_{max} = 658.9 ± 34.9 pMh^{-1} and V_{max}/K_m = 7.18 x 10^{-9} h^{-1}). In the present study, we purified rec-rPC4 by a single chromatography method namely by using DEAE-agarose column. The enzyme purified in this manner exhibited a K_m value = 13µM, V_{max} = 176.8 + 8.04pMh^{-1}, V_{max}/K_m = 13.6 x 10^{-9} h^{-1}). Thus the K_m in our case is 7-fold lower and V_{max} ~3.7 fold lower than the previous reported values. However the V_{max}/K_m ratio values are not too far from one another. In fact our V_{max}/K_m ratio is ~1.9 fold higher than the literature value suggesting that our PC4 enzyme is slightly more potent kinetically than the previously obtained enzyme. This is interesting but not surprising since passing through the second column namely arginine-sepharose column in this case and subsequent elution with free arginine containing buffer may have some effect on the conformation and geometry of the large enzyme molecule. The presence of free arginine in the final elution buffer in the previously reported purification method may have resulted in the modification of the kinetic behavior of the enzyme compared to the enzyme in our case. However this speculation will require additional work for further confirmation. To further analyse these data, we have to understand the definition of K_m and V_{max} and their ratio. This is illustrated below using the kinetic equation:



$$K_m = \frac{k_{-1} + k_2}{k_1}$$

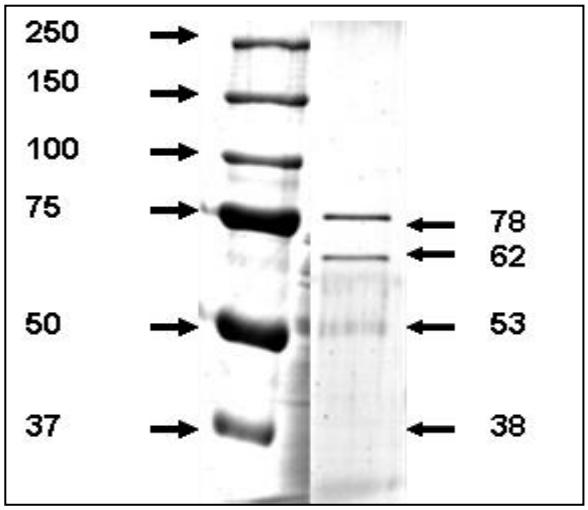
In the most simple case, when product (P) formation is the rate-limiting step (i.e. when $k_2 << k_{-1}$), the constant K_m will just be equal to the dissociation constant i.e. K_d (affinity for substrate) of the enzyme-substrate (ES) complex.

$$K_m \sim \frac{k_{-1}}{k_1} = \frac{[E]\,[S]}{[ES]} = K_d$$

Therefore K_m is considered as a measure of the affinity between the enzyme (E) and the substrate (S). So lower K_m means stronger complex formation but lower V_{max} signifies slower kinetics. In fact the ratio of these two parameters is the best representation of the kinetic property of an enzyme reaction. The quality and purity of the two enzyme samples can be better understood by comparing their gel electrophoresis which is shown in **Figure 30**.

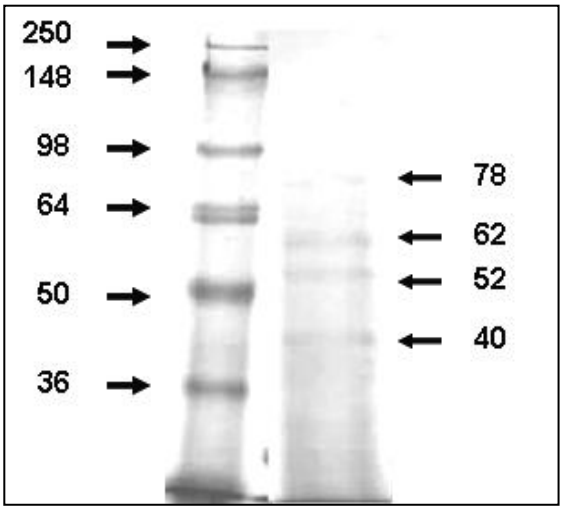
The figure indicates that the bands in two cases were similar in nature and profile. They are characterized by the presence of two strong bands ~ 54 and 62kDa due to mature and pro forms of PC4 enzyme along with another band ~ 38kDa possibly due to the formation of a truncated form of PC4.

Figure 30: Comparison of rec-rPC4 bands obtained by this study (right) and that of the previous published study (left) (54).



Marker PC4 bands

Coomasie-stained 8% SDS Gel
[This figure was taken from reference (57)]



Marker PC4 bands

Coomasie-stained 10% SDS Gel
[This figure from my study]

5.2 Rec-mCRES production, purification and characterization

As indicated in the introduction section, physiological mouse CRES protein is a 121 amino acid long protein [without the signal peptide from residue #22 to 142) with a calculated molecular weight of 14,040kDa. However our CRES-construct contained a C-terminal V_5 as well as His_6 tags with some additional amino acids as indicated in **Figure 31A**. Therefore its calculated molecular weight should be 17,682 (as shown in **Figure 31B**), much higher than the physiological mCRES protein. Our purified rec-mCRES protein exhibited in the mass spectum peaks at ~14,476 and ~28,922kDa for the monomeric and dimeric forms respectively. This is much lower than the above calculated values. In order to explain this size we propose that our rec-mCRES protein during purification or expression may have lost both V5 and His 6 tags. These two tags together with a few extra amino acids taken together will constitute a molecular weight of 3,372kDa. Therefore in that case rec-mCRES should have a molecular weight 17,682 – 3,372 = 14,310kDa that agrees quite well with the measured value. This proposition is also supported by our Lys-C digestion and the associated mass spectral data as shown in our results section. During RP-HPLC purification of crude rec-mCRES protein it was observed that even though the CRES dimer (~28kDa) and monomer (~14kDa) were eluted at retention times wide apart from one another (12 minutes apart), still the MS spectrum data showed that they never exhibited a single peak meaning that neither of them are in complete pure from. There was always a minor quantity of higher or lower oligomeric form present depending on the sample. Thus a small monomer peak was present in the collected dimeric sample while there was a minor dimeric peak in the collected monomeric sample. This may suggest that monomeric CRES has a tendency to aggregate to its higher oligomeric forms which has also been reported by Cornwall et al.

(127). It is very likely that after collecting the monomer peak in the HPLC separately from the dimer, it may have dimerized on storage or handling. The dimer also showed a weak monomer peak which may be due to the M/2 peak of the dimer or monomer itself. We have also noticed that on prolonged storage CRES monomer undergoes significant self-association to yield tri, tetra and higher oligomeric forms. Later on we also noticed that under certain conditions it oligomerizes to even higher forms such as hexamer and heptamer as depicted in **Figure 32A.**

Figure 31: Amino acid sequences and the calculated molecular weights of proposed rec-mCRES as well as the physiological mCRES proteins

- **(A) mCRES-V_5-H_6 (no Signal peptide)**

MVG ^{22}VDQSKNEVKA QNYFGSINIS NANVKQCVWF AMKEYNKESE DKYVFLVDKI
^{72}LHAKLQITDR MEYQIDVQIS RSNCKKPLNN TENCIPQKKP ELEKKMSCSF
^{122}LVGALPWNGE FNLLSKECKD V^{142}ELNSKLEGK PIPNPLLGLD STRTGHHHHHH

- **Calculated MW = 17,682**

- **(B) mCRES (no Signal peptide)**

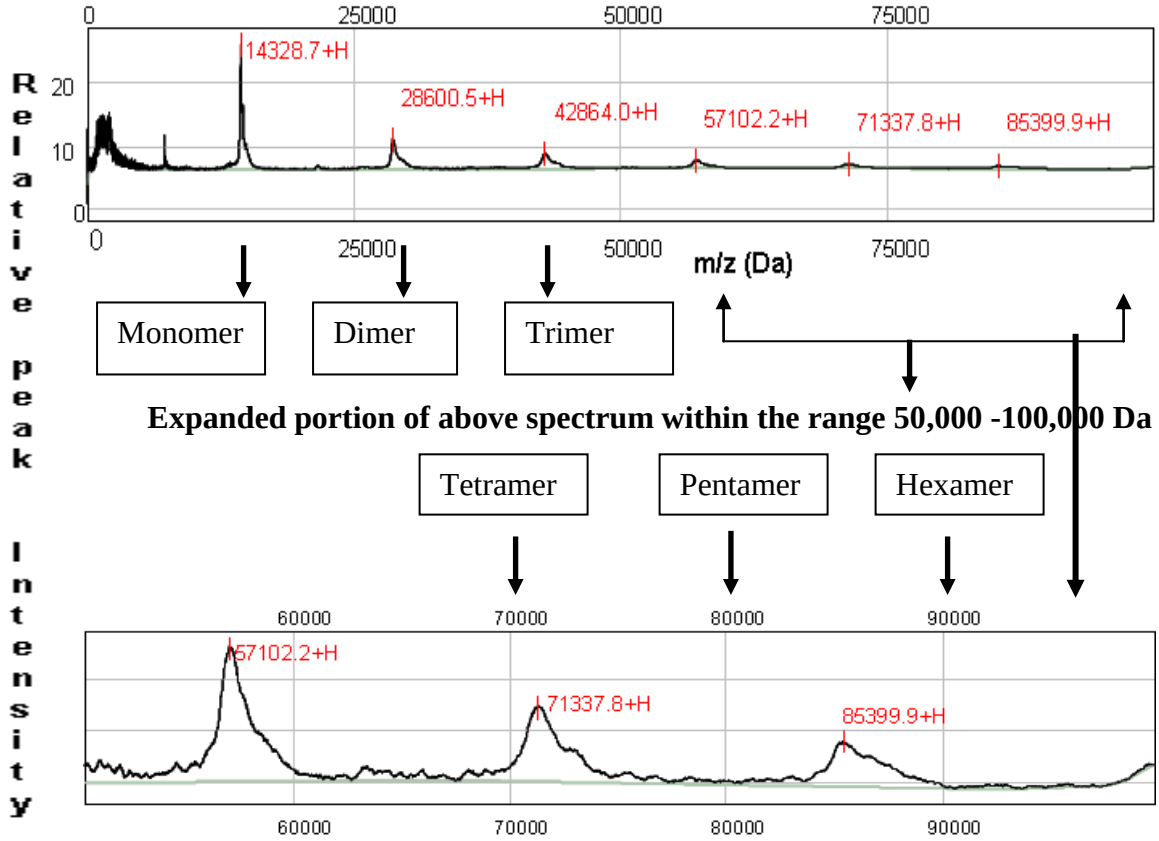
^{22}VDQSKNEVKA QNYFGSINIS NANVKQCVWF AMKEYNKES^{61}E DKYVFLVDKI
^{72}LHAKLQITDR MEYQIDVQIS RSNCKKPLNN TENCIPQKKP ELEKKMSCS^{121}F
^{122}LVGALPWNGE FNLLSKECKD V

- **Calculated MW = 14,040**

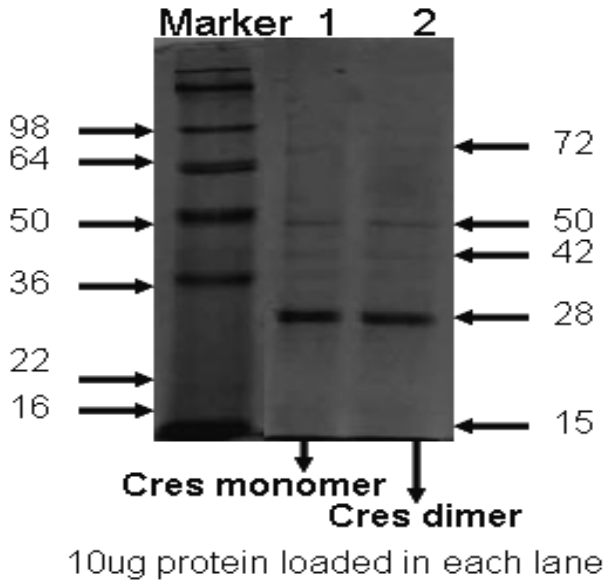
Note: The V_5 (Sequence: G-K-P-I-P-N-P-L-L-G-L-D-S-T) and His_6 (H-H-H-H-H-H) tags are indicated by underlines. The top panel represents the exact amino acid sequence of our rec-mCRES-V_5-H_6 construct whereas the bottom panel is the amino acid sequence of physiological mCRES protein once secreted where the signal peptide is removed.

Figure 32(A) shows that rec-CRES protein undergoes self-association or aggregation process quite easily. Coomasie staining of rec-mCRES protein was also carried out on 15% SDS PAGE using both the monomer and dimeric forms of rec-mCRES sample but interestingly both the lanes showed mostly dimer even though the MS results showed monomer peak very distinctly. The lane with CRES monomer only showed the presence of a strong dimer band compared to the monomer band. Some of the higher polymeric forms of CRES were also seen in the SDS-gel as shown in Figure 32(B). This indicates that CRES and its self-aggregating behavior may be somewhat complex in nature. It may require thorough and additional study to better understand the event and its biochemical implication. Even the time period needed by CRES to aggregate was found to be variable depending on various conditions like its period of storage, temperature, presence of light, possibly ions, etc. to which it may be exposed. Sometimes it exhibited high polymeric forms very early whereas in other cases it showed oligomeric peaks at a much later time period and sometimes with much less intensity.

Figure 32: (A) SELDI-TOF mass spectrum of rec-mCRES protein upon prolonged storage: Evidence of self aggregation



(B) Coomasie stained 15% SDS gel electrophoresis of rec mCRES



5.3 Analysis of PC4 activity and protein content in the fluids of various regions of mouse epididymis

Our studies as depicted in **Figures 10-14** suggested the presence of PC4-like protease activity as well as PC4 protein in the fluids of all four regions of mouse epididymis tested. The data further showed that the activity is highest in the caput region compared to those in cauda, corpus and vas deferens. Furthermore the presence of PC4 protein in these fluids was confirmed by western blot analysis. Overall, a presence PC4 protein was noticed in all four regions with the caput and corpus showing the highest amount compared to others. Moreover PC4-like activity was found to be present in varying levels depending on the region. It is important to point out that it is the protease activity of PC4 and not the actual protein content that is more important for regulating the ultimate fertilizing capacity of the sperm since this activity is linked to the processing and activation of sperm surface protein precursors.

So far the available literature indicates that CRES is present in highest amount in the proximal caput region of the epididymis compared to other epididymal regions where it is also present though in reduced levels. Hence our findings in this respect are significant and may provide some insight into the **CRES:PC4** interaction events. The presence of high PC4 activity (our result) and increased CRES protein level (previous data) (125) in the caput region in the first glance seems unexpected. However in this connection it may be pointed out that PC4 is a membrane bound enzyme and only a part of it may be released into the medium as a secreted form in physiological system. It is likely that the relative levels of PC4 enzyme in bound and unbound forms may vary from one region to the other in the epididymis (**Figure 33** shows the schematic representation of this process). This may

depend on the nature of the tissue and cell types. We rationalize that in the caput region a majority of sperm PC4 enzyme remains in the fluid medium in soluble secreted form. The level of this activity is so high that despite the presence of high content of inhibitory CRES protein, the final activity still remained high. Moreover as indicated in the results and discussion section, the PC4 inhibitory effect of CRES is dependent on its oligomeric state. It is possible that in the caput most of CRES may be present in monomeric form which may be least inhibitory towards PC4. In other epididymal regions, it is likely that sperm PC4 is present mostly in bound form and the low level of soluble PC4 activity present in the fluid is inhibited to a significant extent by the CRES protein and therefore its activity remained low compared to that in caput. **Finally we conclude that the sperm PC4 activity in various epididymal fluids is controlled and regulated by three important factors namely**

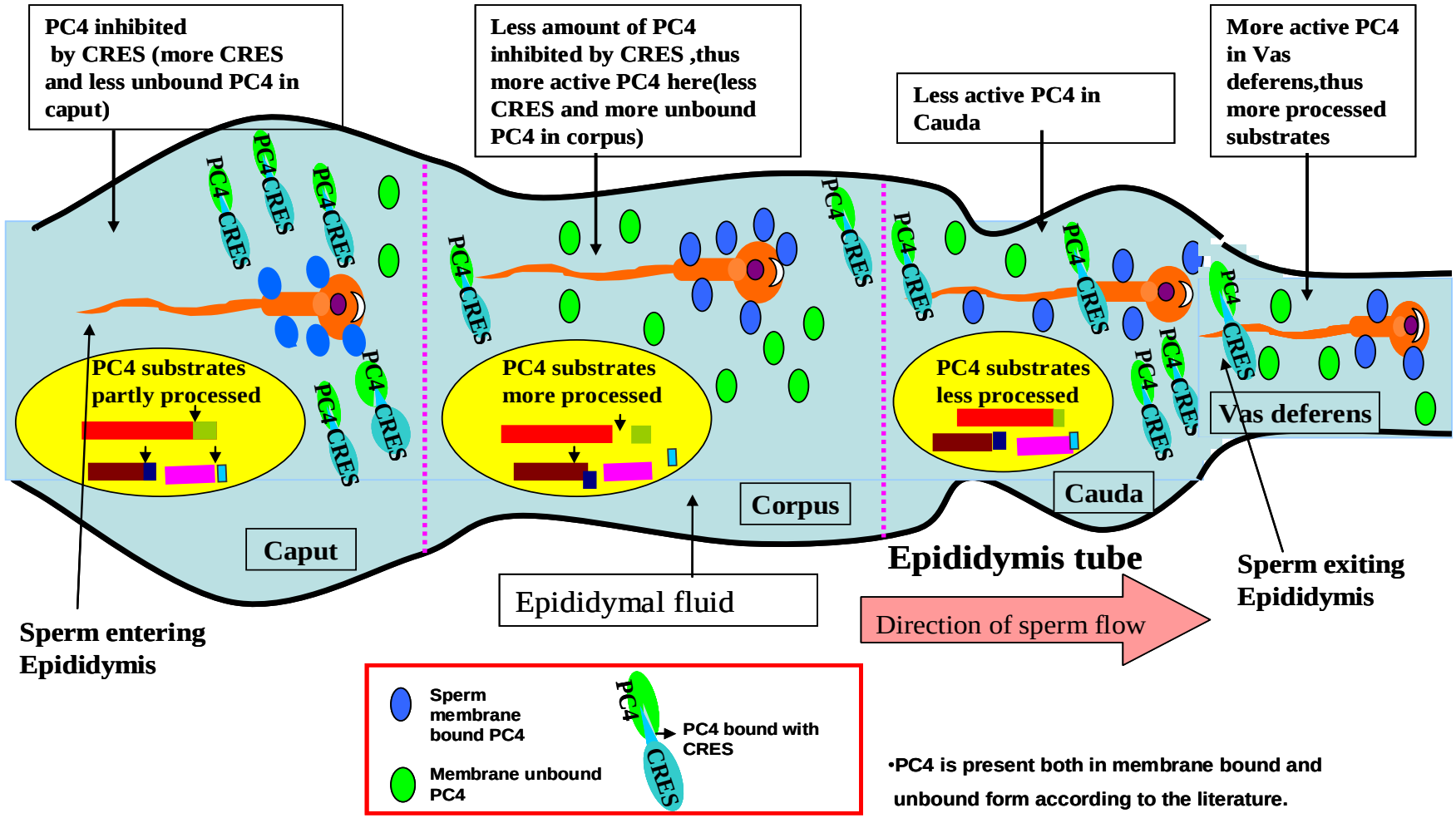
(1) How much PC4 enzyme remained membrane bound?

(2) How much CRES protein is present the epididymal fluid?

(3) The oligomeric or aggregation state of CRES protein

In addition to the above, various other factors would have to be examined in future experiments. These may include the age of mice used, our ability to dissect precisely the various mouse epididymis compartments, the amount of epididymal fluids collected, and how well the fluid could be separated from the sperm cells. For future work immunolocalization studies of PC4 and CRES could be performed to examine the expression pattern of PC4 in the epididymal region. In addition, PC4 specific fluorogenic substrates based on IGF2 and PACAP designed by us (94) could be used to detect and measure PC4 activity more accurately in various epididymal fluids. These will provide a better understanding of the events occurring in the epididymis during the fertilization cycle.

Figure 33: Schematic representation of the Epididymal regions showing the direction of sperm flow along with the PC4 substrate processing



5.4 Inhibitory effect *in vitro* of rec-mCRES on PC4 activity:

Our results indicated that *in vitro* PC4 activity is inhibited by mCRES protein with moderate potency and the extent of this inhibition depends on the oligomeric state of CRES protein. Thus while the dimeric CRES form exhibited more potent PC4-inhibitory activity **(IC_{50} ~ 20 - 35µM)** depending on the concentration of the substrate), the monomeric form is only weakly inhibitory towards PC4 activity with **IC_{50} > 200µM**.

This observation that PC4 inhibitory activity of CRES is regulated by its aggregation state is a common event and is highly significant in terms of final PC4 activity. However the factors responsible for CRES aggregation have not been well studied. It is interesting to point out that our *ex vivo* cell work indicated that both monomeric and dimeric CRES proteins blocked almost equally well the processing of pro-hIGF2 into its mature form. The inability of our IGF2 antibody to detect physiological mature IGF2 (8kDa) produced in the cell experiment (despite the fact that it was able to detect synthetic mature IGF2 purchased commercially) proved some difficulty for us to make any definitive conclusion about the relative efficacy of monomer-CRES versus dimer-CRES in blocking proIGF2 maturation. However a very close examination of **Figures 21 and 22**, indicated that while the monomeric CRES was able to effectively block the processing of big-hIGF2 into mature IGF2 in a concentration dependent manner, the dimeric form can block both the processing of big IGF2 to mature IGF2 as well as pro-hIGF2 to mature IGF2 directly.

As mentioned earlier we could not detect the mature IGF2 band in the cell experiment although both the intermediate and the full length IGF2 can be detected quite efficiently. We speculate that this may be related to the glycosylation pattern of IGF2 in the human placental cell line used in the experiment which may explain the lack of recognition

by the antibody. On one occasion we were able to detect a weak band for mature IGF2. Owing to this, we rely our conclusion based on the precursor and the unprocessed protein bands.

5.4.1. Possible mechanism for inhibition of PC4 activity by CRES:

As shown in **Figures 20 (A) and (B),** we observed that besides behaving as an inhibitor of PC4 enzyme, our rec-mCRES protein is also cleaved by the PC4 enzyme itself. We identified this site as **KK^{97}↓PL (**for numbering see **Figure 32)** based on mass spectrometry data. This is a quite uncommon cleavage site. However there is one report (137) where such type of cleavage by a PC-like enzyme has been proposed within the protein POMC (Pro-opiomelanocortin) (a 262-mer protein, Accession no AAH65832) where the cleavage occurs at **KNAYKK^{265}↓GE** (34).In our case the cleavage of CRES by PC4 takes place **SRSNCKK^{97}↓PL** which is similar to the above mentioned sequence. Interestingly both cleavage sites contain a P6 basic residue (either **R** or **K**) which is likely to promote the cleavage. This is consistent with the known observation that for efficient recognition by any kexin type PC enzyme such as PC4, one would require basic residue/s at P4, P6 or in both positions. The above cleavage of CRES protein by PC4 at the carboxy terminal of double Lys residues is very unique and further work using model peptides and mutations may be required in future for further confirmation. The mechanism by which CRES dimer inhibits PC4 activity is not fully understood at this stage. Our data indicated that the inhibition is competitive in nature (by Dixon and Lineweaver-Burk plots) as previously observed by Cornwall et al (22) for inhibition of PC2 enzyme by CRES as well. However any possible cleavage of CRES by PC2 was not investigated. Our results clearly showed that CRES is cleaved by PC4 while inhibiting its activity – a phenomenon usually observed for protease

inhibition by a serpin. However we did not observe formation of any complex between CRES and PC4 as one would expect for enzyme inhibition by a serpin (Our preliminary unpublished work). Even though serpin-protease binding mechanism is generally irreversible type dute to the covalent bond but there are plenty of examples found in the literature where they have also followed reversible binding mechanism. One such example is Ovalbumin (serpin) inhibits cathepsin G and elastase reversibly (138). Further investigation to clarify the type of PC4 inhibition by CRES is currently in progress in the laboratory.

The very fact that we were able to notice a separate fragment at m/z 9226 as shown in **Figure 27(A)** due to the cleavage of rec-mCRES at **SRSNCKK^{97}↓PL** clearly indicated that in our rec-mCRES protein at least the **Cys^{95}** residue likely remains in its free state. This means that our rec-mCRES protein may not contain the key S-S bridges loops which might explain only moderate instead of expected potent PC4 inhibitory property of our rec-mCRES protein. Further research studies will be required on S-S bridge formation in CRES and its effect on PC4 inhibition. It is possible that with the correct S-S bridges, CRES may become a more potent inhibitor of PC4.

So far based on Dixon plot and measured IC_{50} values at various concentrations of substrate Boc-RVRR-MCA, it appears that PC4 inhibition by CRES is mostly competitive in nature despite the cleavage. The measured K_i value from Dixon plot and the value calculated from Cheng and Prusoff equation **($K_i = IC_{50}/(1+S/Km)$)** (134) were found to be relatively close to one another.

5.4.2 Aggregation of CRES protein

Our study also confirmed the previous report that CRES is prone to aggregation. However our study is the first demonstration by mass spectrum that with time and storage CRES

aggregation may continue upto heptamers and beyond. Aggregation of serpins is not uncommon (135, 136) but their effects on protease inhibition has not been studied extensively. Our *in vitro* and *ex vivo* data showed that dimerization of CRES enhances its PC4 inhibitory property particularly *in vitro*. So far factors affecting the aggregation of CRES protein have not been studied, although pH, ions, temperatures, etc may contribute in this process of aggregation. Whether CRES aggregation *in vivo* leads to enhanced inhibition of PC4 activity or not (in terms of its physiological substrates), could not be confirmed at this stage. Our cell work data indicated that CRES monomer inhibits primarily the cleavage of big-IGF2 into mature IGF2. On the other hand CRES dimer inhibits both PC4-mediated cleavages of proIGF2 and big IGF2 into mature IGF2.

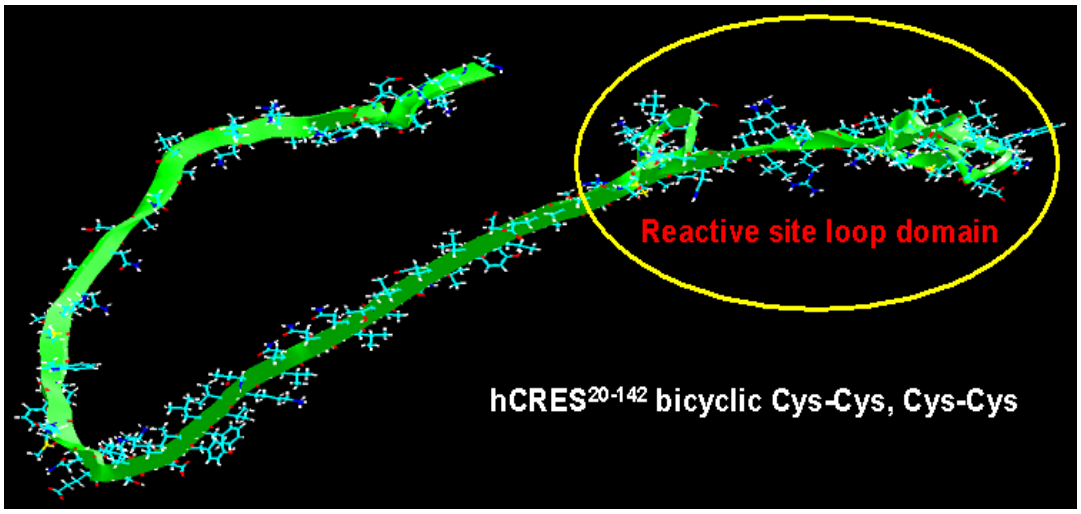
5.5 3D Molecular model structure of CRES protein

In order to get a preliminary idea about how aggregation changes the structure and conformation of CRES protein, we performed 3D model studies of CRES protein and its derived peptides. The obtained 3D-geometrical structures of full length hCRES protein showing its two S-S bridges and the corresponding loops were accepted until the energy calculations showed no further decrease or minimization. The energy minimization was performed by using Polak-Ribiere algorithm program**.** These are displayed in **Figures 34 (A&B).**

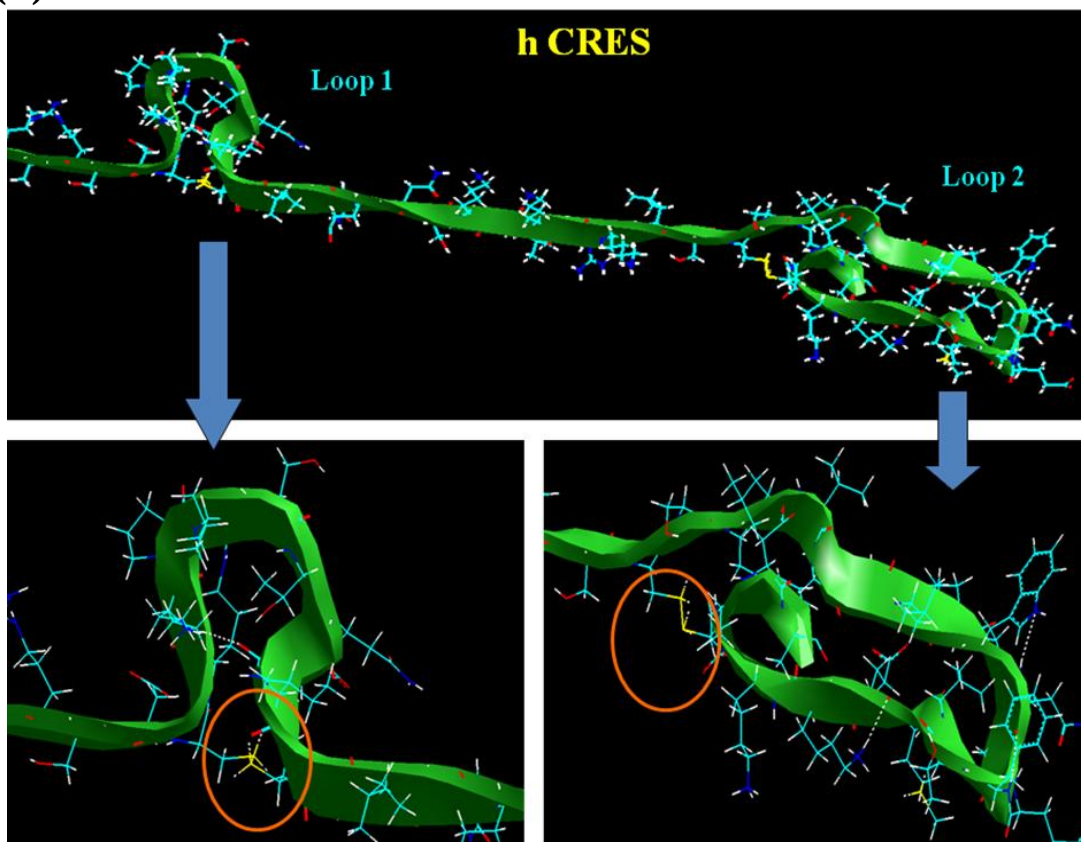
The 3-D model prediction of the full length h-CRES was done showing the reactive site loop (RSL). Loop 1 and loop 2 are shown in **Figure 34 (B)**, suggesting their disulfide linkages among the four cysteine residues which could be either 1-2 and 3-4 or 1-4 and 2-3 connections. These disulfide bridge connection types could also be determining factor in its overall conformation and hence their PC4-inhibitory actions.

Figure 34: Predicted 3D model structure of full-length h-CRES with encircled RSL domain.

(A)



(B)



6. CONCLUSION:

The epididymal serpin CRES is a moderately potent inhibitor of PC4 activity particularly in its dimeric form. The inhibitory property depends on the aggregation state of CRES protein. Thus *in vitro* data showed that the dimeric CRES is a more potent inhibitor of PC4 (IC_{50} ~24-3 µM and K_I = 9µM) than the corresponding monomer which exhibited IC_{50} value of > 200 µM, against PC4. Both CRES monomer and dimer blocked pro-hIGF2 processing in a dose dependent manner in placenta cell line. However the dimer blocked PC4 mediated cleavages of both proIGF2 and big IGF2 into mature IGF2. On the other hand CRES monomer inhibited only the cleavage of big IGF2 into mature IGF2.

It is already known that many key sperm PC4 substrates are processed by PC4 enzyme and these processing is crucial for fertilization events. But there are many immature sperms that enter the oocyte which result in infertility and hence inhibition of PC4 activity by CRES at the right time gives the sperms ample time for it to mature completely and then enter the oocyte for a successful fertilization to occur. Hence inhibition of premature sperms to enter the female reproductive tract could be linked to CRES inhibiting the PC4 enzyme at right time, which is one of the important protease involved in reproduction. Thus all our data taken together suggested that CRES could be inhibiting the PC4 activity in mice epididymis. PC4 activity varied from one region to the other with an alternate increase and decrease fashion in these regions whereas CRES is known to be highest in the Proximal Caput region, hence making it appear that PC4 activity could be regulated by CRES differentially as it travels along the epididymis tube. The exact mechanism of action for CRES inhibition by PC4 is still a speculation which could be depending on many factors like CRES aggregation, conformational changes, CRES expression in stage-specific manner, etc. Even the CRES

inhibition of sperm associated PC4 during its transit and storage could not be fully understood. This could be depending on factors like the relative amount of PC4 bound to the membranes and released to the compartmental medium, pH, etc. Further studies in this area could help us in knowing the exact mechanism behind all this speculations. This work is likely to have a major impact in the field of fertilization study and generate the idea regarding the PC4 activity and its relation with some of the unexplained infertility in male patients. Therefore studying the mechanism of PCSK4 regulation would definitely be an important part of fertilization research. Henceforth this could be used to study as a potential target for development of new contraceptive agents.

Thus my thesis was able to address all the four aims as proposed under my principal objective.

7. REFERENCES

1. Eddy EM (1998) *Seminars in Cell & Developmental Biology* **9,** 451-457.
2. Ramalho-Santos J ,Schatten G and Moreno RD. (2002) *Biol Reprod.* **67,** 1043-1051.
3. Gyamera-Acheampong C, Tadros H, Sirois F, et al. (2006) *Biol Reprod.* **74,** 666-673.
4. Cho C, Ge H, Branciforte D, et al (2000) *Developmental Biology* **222,** 289-295.
5. Bozzola JJ, Haas N, Russell LD, et al. (1991) *Am J Anat.* **192,** 129-141.
6. Griswold Q (2008) *The Stem Cell Research Community, Stem Book.*
7. Eddy EM (2002) *Recent Progress in Hormone Research* **57,** 103-128.
8. Cornwall GA (2009) *Human Reproduction Update* **15,** 213-217.
9. Retamal C, Lorca C, López ML,et al (2000) *J Submicrosc Cytol Pathol.* **32,** 229-239.
10. Cooper TG (1997) *In: Andrology,* 61-78.
11. Tulsiani DR. (2003) *Microsc Res Tech.* **61,** 18-27.
12. Hunnicutt GR, Koppel DE and Myles DG (1997) *Developmental Biology* **191,** 146-159.
13. Flesch FM (2000) *Biochim Biophys Acta.* **1469,** 197-235.
14. Cross NL. (1998) *Biol Reprod.* **59,** 7-11.
15. Nishimura H, Cho C, Branciforte DR, et al (2001) *Developmental Biology* **233,** 204-213.
16. Nour N, Chrétien M and Seidah NG (2003) *J Biol Chem* **278,** 2886-2895.
17. Seidah NG and Day R. (1994) *Biochimie.* **76,** 197-209.
18. Steiner DF, Spigelman L and Aten B. (1967) *Science.* **157,** 697-700.
19. Chrétien M (1967) *Can J Biochem.* **45,** 1163-1174.
20. Fuller RS, Thorner J. (1989) *Proc Natl Acad Sci U S A.* **86,** 1434-1438.
21. Seidah NG, Marcinkiewicz M, Benjannet S, et al (1991) *Enzyme* **45,** 271-284.
22. Cornwall GA, Lindberg I, Hardy DM, et al (2003) *Endocrinology.* **144,** 901-908.
23. Rockwell N.C and Fuller R.S .(1998) *Biochemistry* **37,** 3386-3391.
24. Gyamera-Acheampong C (2009) *Hum Reprod Update.* **15,** 237-247.
25. Satoshi N, Kensuke S, Masayuki M (2010) *Biochemical and Biophysical Research Communications* **391,** 107-112.
26. Nakayama K. (1997) *Biochem J.* **327,** 625-635.
27. Thomas G. (2002) *Nat Rev Mol Cell Biol.* **3,** 753-766.
28. Salvas A, Benjannet S, Reudelhuber TL, et al (2005) *FEBS Letters* **579,** 5621-5625.
29. Bassi DE , Lopez de Cicco R and Klein-Szanto AJ. (2005) *Mol Carcinog.* **44,** 151-161.
30. Seidah NG and Chrétien M (1999) *Brain Research* **848,** 45-62.
31. Martoglio B and Dobberstein B (1998) *Trends in Cell Biology* **8,** 410-415.
32. Von Eggelkraut-Gottanka R and Beck-Sickinger, AG. (2004) *Curr. Med. Chem* **11,** 2651-2665.
33. Duckert P BS, Blom N. (2004) *Protein Eng Des Sel.* **17,** 107-112.
34. Gagnon J, Mayne J, Mbikay M, et al (2009) *Regulatory Peptides* **152,** 54-60.
35. Creemers JW (2008) *Front Biosci.* **13,** 4960-4971.
36. Scamuffa N, Chrétien M, Seidah NG, et al (2006) *FASEB J.* **20,** 1954-1963.
37. Seidah NG (2002) *Essays Biochem.* **38,** 79-94.
38. Khatib AM , Chrétien M, Metrakos P, et al (2002) *Am J Pathol.* **160,** 1921-1935.
39. Khatib AM, Prat A, Luis J, et al (2001) *J Biol Chem.* **276,** 30686-30693.

40. Khatib AM, Siegfried G, Klein-Szanto AJ, et al (2005) *J Mol Med.* **83,** 856-864.
41. Siegfried G, Cromlish JA, Benjannet S, et al (2003) *J Clin Invest.* **111,** 1723-1732.
42. Gordon VM, Arora N, Henderson MA, et al (1995) *Infect Immun.* **63,** 82-87.
43. Moulard M, Hallenberger S, Garten W, et al (1999) *Virus Research* **60,** 55-65.
44. Basak A, Munzer JS, Chrétien M, et al (2001) *Biochem J.* **353,** 537-545.
45. Bergeron E, Vincent MJ, Wickham L, et al (2005) *Biochemical and Biophysical Research Communications* **326,** 554-563.
46. Basak A, Chrétien M, and Seidah NG (2002) *FEBS Letters* **514,** 333-339.
47. Chrétien M, Basak A and Mbikay M. (2008) *Expert Opin Ther Targets.* **10,** 1289-1300.
48. Basak A (2005) *J Mol Med.* **83,** 844-855.
49. Nakayama K, Torii S, Hosaka M, et al (1992) *J Biol Chem* **267,** 5897-5900.
50. Seidah NG, Hamelin J, Gaspar A, et al (1992) *Mol Endocrinol* **6,** 1559-1570.
51. Torii S, Murakami K, and Nakayama K. (1993) *FEBS Lett* **316,** 12-16.
52. Tadros H, Chrétien M, and Mbikay M (2001) *Journal of Reproductive Immunology* **49,** 133-152.
53. Qiu Q, Mbikay M, Tsang BK, et al (2005) *Proc Natl Acad Sci USA.* **102,** 11047-11052.
54. St. Germain C, Mayne J, Baltz JM, et al (2005) *Mol Reprod Dev* **72,** 483–493.
55. Mbikay M, Raffin-Sanson ML, Tadros H, et al (1994) *Genomics* **20,** 231-237.
56. Gyamera-Acheampong C, Weerachatyanukul W, Tadros H, et al (2006) *Biol Reprod.* **74,** 666-673.
57. Basak A, Shervani NJ, Mbikay M, et al (2008) *Protein Expression and Purification* **60,** 117-126.
58. Mbikay M, Ishida N, Lerner CP, et al (1997) *Proc Natl Acad Sci USA* **94,** 6842–6846.
59. Li M, Mbikay M, Shioda S, et al (2004) *Endocrine* **23,** 59-75.
60. Li M, Mbikay M and Arimura A. (2000) *Endocrinology.* **141,** 3723-3730.
61. Li M, Mbikay M, Nakayama K, et al (2000) *Ann N Y Acad Sci.* **921,** 333-339.
62. Li M, Somogyvári-Vigh A and Arimura A. (1999) *Neuroendocrinology.* **69,** 217-226.
63. Li M, Nayakama K, Shuto Y, et al (1998) *Peptides.* **19,** 259-268.
64. Miyata A, Jiang L, Dahl RD, et al (1990) *Biochemical and Biophysical Research Communications* **170,** 643-648.
65. Arimura A, Miyata A, Mizuno K, et al (1991) *Endocrinology.* **129,** 2787-2789.
66. Gray SL, Cummings KJ, Jirik FR, et al (2001) *Mol Endocrinol.* **15,** 1739-1747.
67. Shioda S, Legradi G, Leung WC, et al (1994) *Endocrinology* **135,** 818-825.
68. West AP, McKinnel C, Sharpe RM, et al (1995) *J Endocrinol.* **144,** 215-223.
69. El-Gehani F, Tena Sempere SM, and Huhtaniemi I. (2000) *Biol Reprod.* **63,** 1482-1489.
70. Leung PS Wong TP, Wong PY, et al (1998) *Cell Biol Int.* **22,** 193-198.
71. Vaudry D, Gonzalez B, Basille M, et al (2000) *Pharmacol Rev.* **52,** 269-324.
72. Almog T and Naor Z (2008) *Molecular and Cellular Endocrinology* **282,** 39-44.
73. Isaac ER and Sherwood NM (2008) *Molecular and Cellular Endocrinology* **280,** 13-19.
74. Tanii I AT, Matsuda K, Komiya A, et al (2010) *Reproduction.* .

75. Vaudry D, Falluel-Morel.A, Bourgaul S, et al, (2009) *Pharmacol Rev.***,** 283–357.
76. Yuan R Primakoff P and Myles DG (1997) *J. Cell Biol* **137,** 105-112.
77. Wolfsberg TG, Gerena RL, Huovila AP, et al (1995) *Dev Biol.* **169,** 378-383.
78. Primakoff P (2000) *Trends Genet.* **16,** 83-87.
79. Wolfsberg TG and White JM (1996) *Developmental Biology* **180,** 389-401.
80. Evans JP and Kopf GS. (1998) *Biol Reprod.* **59,** 145-152.
81. Blobel CP. (2000) *Rev Reprod.* **5,** 75-83.
82. Baessler KA, Roberts KS, Facompre N, et al. (2006) *Chem Biol.* **13,** 251-259.
83. Blobel CP, Primakoff P and White JM. (1990) *J Cell Biol.* **111,** 69-78.
84. Wolfsberg TG, Blobel CP, Myles DG, et al (1993) *Proc Natl Acad Sci USA.* **90,** 10783-10787.
85. Bigler D, Chen M, Waters S, et al (1997) *Trends in Cell Biology* **7,** 220-225.
86. Linder B, Bammer S, and Heinlein UAO (1995) *Experimental Cell Research* **221,** 66-72.
87. Cho C, Faure JE, Goulding EH, et al (1998) *Science.* **281**.
88. Han C, Park I, Lee B, et al (2010) *Fertil Steril.* **93,** 2754-2756.
89. Dombrowicz DH, Gothot A, Sente B, et al (1992) *Arch Int Physiol Biochim Biophys.* **100,** 303-308.
90. Duguay SJ, Young BD, Nakayama K, et al (1997) *J Biol Chem.* **272,** 6663-6670.
91. Sibley CP, Ferguson-Smith AC, Dean W, et al (2004) *Proc Natl Acad Sci USA.* **101,** 8204-8208.
92. Basak A, Lazure C, Mbikay M, et al (1999) *Biochem J.* **343,** 29-37.
93. Bergeron F and Day R. (2000) *J Mol Endocrinol.* **24,** 1-22.
94. Basak S ,Chretein M, Mbikay M, et al (2004) *Biochem J.* **380,** 505-514.
95. Majumdar S., Chowdhury DR, Banik R, et al (2010) *Current Medicinal Chemistry* **17,** 2049-2058.
96. Beynon R.J and Bond J.S. (1989) *Book* of Proteolytic Enzymes,a practical approach.
97. Kaiserman D (2006) *Expert Rev Mol Med.* **8,** 1-19.
98. Whisstock JC and Bottomley SP (2006) *Current Opinion in Structural Biology* **16,** 761-768.
99. Huntington JA (2006) *Trends in Biochemical Sciences* **31,** 427-435.
100. Schick C, Bartuski AJ, Uemura Y, et al (1998) *Proc Natl Acad Sci USA.* **95,** 13465-13470.
101. Richard A. Engh RH, Wolfram Bode, et al (1995) *Trends Biotechnol.* **13,** 503-510.
102. Tao H, Shi J, Shao XX, et al (2006) *FEBS J.* **273,** 3907-3914.
103. Gray WR. (1993) *Protein Sci.* **2,** 1749-1755.
104. Navickis RJ (2004) *COPD.* **1,** 279-292.
105. Zhang S and Janciauskiene S (2002) *J Alzheimers Dis.* **4 ,**115-22 .
106. Dubin G. (2005) *Cell Mol Life Sci.* **62,** 653-669.
107. Feng Y, Reznik SE, and Fricker LD (2002) *Gene Expression Patterns* **1,** 135-140.
108. Mbikay M, Seidah NG and Chrétien M. (2001) *Biochem J.* **357,** 329-342.
109. Uhrin P, Hilpert M, Chrenek P, et al (2000) *J Clin Invest.* **106,** 1531-1539.
110. Charron Y, Nef S, Combepine C, et al (2006) *Mol Reprod Dev.* **73,** 9-19.
111. Kanno Y, Chuma S, Sakura T, et al(1999) *Int J Dev Biol.* **43,** 777-784.
112. Shoemaker K, Holloway JL, Whitmore TE, et al(2000) *Gene* **245,** 103-108.
113. Wang Z, Richardson RT and O'Rand MG. (2007) *Biol Reprod.* **77,** 476-484.

114. O'Rand MG, Wang Z and Richardson RT. (2007) *Soc Reprod Fertil Suppl.* **63,** 445-453.
115. Wang Z, Richardson RT and Orand MG. (2007) *Soc Reprod Fertil Suppl.* **65,** 535-542.
116. Richer MJ, Waterhouse J, Minhas J, et al (2004) *Proc Natl Acad Sci USA.* **101,** 10560-10565.
117. Brüning M, Bentele C, Smolenaars MM, et al (2007) *Biochem J.* **401,** 325-331.
118. Jean F, Thomas L, Liu G, et al (1998) *Proc Natl Acad Sci USA.* **95,** 7293-7298.
119. Irving JA, Lesk AM and Whisstock JC. (2000) *Genome Res.* **10,** 1845-1864.
120. Silverman GA, Carrell RW, Church FC, et al (2001) *J Biol Chem.* **276,** 33293-33296.
121. Cornwall GA and Hsia N (2003) *Molecular and Cellular Endocrinology* **200,** 1-8.
122. Barrett AJ, Grubb A, Isemura S, et al(1986) *Biochem J* **236,** 312.
123. Bode W, Musil D, and Thiele U (1992) *FEBS Lett* **300,** 131-135.
124. Grubb A, Barrett AJ (1984) *FEBS Lett* **170,** 376-374.
125. Syntin P Cornwall GA (1999) *Biol Reprod.* **60,** 1542-1552.
126. Wassler M, Sutton-Walsh HG, Hsia N, et al (2002) *Biol Reprod.* **67,** 795-803.
127. Von Horsten HH ,Johnson SS, SanFrancisco SK, Hastert MC,et al (2007) *J Biol Chem.* **282,** 32912-32923.
128. Cornwall GA, Swartz D, Johnson S, et al (2007) *Asian J Androl.* **9,** 500-507.
129. Chau K and Cornwall GA (2010) *Biol Reprod.,**84**,140-52* .
130. Parent AD, Cornwall GA, Liu LY, et al(2010) *J Androl ,Epub ahead of prin,.1-44*
131. Basak A, Shervani N, Kolajova M, et al. (2009) *Adv Exp Med Biol.* **611,** 105-106.
132. Cornwall GA and Hann SR. (1992) *Mol Endocrinol.* **6,** 1653-1664.
133. Strausberg RL,Grouse LH, Derge JG, et al (2002) *Proc Natl Acad Sci USA.* **99,** 16899-16903.
134. Cheng Y and Prusoff WH (1973) *Biochem Pharmacol* **22,** 3099-3108.
135. Gosai SJ, Luke CJ, Long OS, et al (2010) *PLoS One.* **5,**e15460
136. Tanaka N, Noguchi Y, Tada T, et al (2010) *J Biol Chem*, **286**, 5884-94.
137. Benjannet S, Rondeau N, Day R, et al (1991) Proc.Natl Acad Sci, USA, **88**, 3564-3568
138. Mellet P, Michels B, and Joseph G (1996), *The Journal of Biological Chemistry*,**271,** 30311-30314.
139. http://en.wikipedia.org/wiki/Epididymis
140. http://en.wikipedia.org/wiki/Sperm
141. http://autistscorner.blogspot.com/2010/02/popular-mis-conception.html