Introduction

INTRODUCTION

Proteinases, are enzymes which are important for the normal functioning of all cells and organisms including viruses. These enzymes catalyze the hydrolysis of peptide bonds in proteins and peptides and depending on the location of the cleavage site. These enzymes are divided into two classes endopeptidases (or proteinases) and exopeptidases. Endopeptidases cleave internal peptide bonds in target polypeptide chains and are also called as proteinases, whereas exopeptidases, comprising of amino- and carboxypeptidases, hydrolyze bonds near either the N- or C-terminus of polypeptides, respectively. Inhibition of proteinases by protein inhibitors represent one of the most important mechanisms of regulating protease activity. Based on the specific chemical groups responsible for catalysis, proteases are classified into five major types: the serine, cysteine, aspartic, threonine and metallo-peptidases (Rawlings & Barrett, 1999). Some proteinases whose catalytic mechanism are yet to be determined are classified as proteases of unknown type. Serine proteinases require presence of serine residue at their active sites, Thiol proteinases have cysteine residue at their active sites while metalloproteinases require metal ions for their full activity, acid proteinases require low pH conditions for activity. The present knowledge of serine proteinases and their endogenous inhibitors at the molecular level is much more detailed than is our understanding of the structure and physiological functions of thiol proteinases. The discovery of thiol proteinases offered a new insight into the processes in which they participate. Thiol or cysteine proteinases are widely distributed among living organisms (Barrett et al., 1986). These are proteins of molecular mass of about 25-35 KDa, they play important role in the control mechanisn responsible for intracellular or extracellular proteolysis such as catabolism of proteins and peptides, processing of prohormones, precursor proteins and in antigen processing and presentation (Turk 2000; Turk et al., 1997; Turk 2001; Stoka 2001).These proteinases are synthesized in precursor form and are later subjected to co-translational and post-translational modifications that convert them to catalytically active mature enzymes (Turk et al., 2000).They contain two principal catalytic amino acid residues, cysteine and histidine which are involved in the cleavage of peptide bonds of a protein or peptide substrate (Grzonka et al., 2001). About 21 different families of cysteine proteinases have so far been characterized (Barret, 1998). The

first clearly recognized and extensively studied cysteine proteinase was papain isolated from latex of the plant, **carica papaya**. Cysteine proteinases are reported from various organisms which include animals (cathepsins, calpain), plants (papain, ficin, bromelain), protozoa (cruzipain), bacteria (strepyopain) and viruses (V-cath proteinase) (Barrett, 2004).

Mammalian thiol proteinases have been classified into four different types:

1. Cathepsins **2. Legumins** **3. Calpains** **4. Caspases**

Cathepsins are relatively small proteins with the exception of cathepsin C which is an oligomeric enzyme of molecular weight of about 200 kDa (Turk et al., 2002). Cathepsins are present in the lysosomes of the cells and are involved in intracellular protein turnover, proteolytic degradation and cleavage of a number of precursor proteins and hormones (Turk et al., 2000). Cathepsins also participate in apoptosis, although there role is still not clear (Turk and Stoka, 2007; Stoka et al., 2005). Mutations in cathepsin genes result in human hereditary diseases such as gycnodystosis (cathepsin K mutation) (Gelb et al., 1996) and Papillon-Lefevre and Hain-Munk syndromes caused by mutations in the cathepsin C gene (Alende et al., 2001). They are present in high concentration in spleen and lungs (Buhling et al., 2000; Turk et al., 2000).

Legumains are cysteine dependent asparagine endopeptidases and are also present in lysosomes. These are involved in MHC class II- restricted antigen presentation (Manoury et al., 1998) and local negative regulation of osteoclasts formation and activity (Choi et al., 1994).

Calpains and Caspases are cytoplasmic thiol proteinases. Calpains participate in many intracellular processes such as turnover of cytoskeletal proteins, cell differentiation and regulation of signal peptides. They are ubiquitously distributed and have been implicated in acute neurological disorders, Alzheimer's disease, muscular cystrophy and gastric cancer (Huang and Wang, 2001). Caspases are the cysteine dependent aspartate specific proteinases. They are involved in cytokine maturation, apoptosis signalling and in apoptosis mediation (Goyal, 2001).

Cysteine proteases have been found in all kind of organisms, from viruses, bacteria and protozoa to plants and vertebrates. According to the most recent classifications (Rawlings and Barrett, 1993, 1994), they can be subdivided on the basis of sequence homology into 14 to 20 different families, the most important being the papain and the calpain families. The papain-like cysteine proteinases are the most abundant among the cysteine proteinases. The family consists of papain and related plant proteinases such as chymopapain, caricain, bromelain, actinidin, ficin, and aleurain, and the lysosomal cathepsins B, H, L, S, C and K. Most of these enzymes are relatively small proteins with Mr values in the range of 20000-35000 (Brocklehurst et al., 1987; Polgar, 1989; Rawlings and Barrett, 1994; Berti and Storer, 1995), with the exception of cathepsin C, which is an oligomeric enzyme with Mr of approximately 200000 (Metrione et al., 1970; Dolenc et al., 1995). A number of cysteine proteinases are located within lysosomes. Four of them, cathepsins B, C, H and L, are ubiquitous in lysosomes of animals (Barrett and Kirschke, 1981; Kirschke and Wiederanders, 1994). The enzymes, except cathepsin C, are endopeptidases (reviewed in Kirschke et al., 1995). All the enzymes are optimally active at slightly acidic pH, although their pH optima for degradation of synthetic substrates varies from 5.5 for cathepsin L to 6.8 for cathepsin H (reviewed in Kirschke et al., 1995). Several other lysosomal cysteine proteinases, such as cathepsins N, T and K, are known, although their properties are less characterized (reviewed in Kirschke et al., 1995). In particular cathepsin K has attracted recent interest (Bromme et al., 1996; Bossard et al., 1996; Drake et al., 1996) and was found to be expressed specifically in osteoclasts (Drake et al., 1996) with properties similar to cathepsin L (Bossard et al., 1996). Cystatins are protein inhibitors of cysteine proteases of the papain family (Barrett, 1986; Turk & Bode, 1991; Turk et al., 1997). They are ubiquitously present in animals, plants and protozoa, both intracellularly and extracellularly, which indicates an important physiological role (Barrett, 1986; Turk & Bode, 1991; Irvine et al., 1992). Their main function is to protect cells and organisms from uncontrolled activity of endogenous cysteine proteases released from lysosomes of dying or diseased cells. In addition, cystatins have been shown to be potent inhibitors of cysteine proteases used by microorganisms and parasites for invading host cells.

FUNCTIONS OF THIOL PROTEINASES

Lysosomal proteinases play important roles in terminal degradation of proteins and generation of amino acids for synthesis of new proteins. Most papain-like cathepsins with differences in specificities, such as cathepsins B, C, F, H, L, O and X, are ubiquitously expressed. Therefore, a common role associated with a nonspecific intracellular "housekeeping" function, *i.e.* breakdown of proteins targeted to lysosomes, has been ascribed to these cathepsins. It has been reported that cathepsin B- or L- deficient mice were normal at birth, and no defects in protein degradation were observed (Deussing et al., 1998; Nakagawa et al., 1998; Roth *et al.*, 2000), indicating that other proteases could take over most functions of the absent cathepsins.

The relatively recent discovery of cathepsins with expression restricted to specific tissues or cell types has suggested new roles for these proteases (Chapman et al., 1997a; McGrath, 1999; Buhling et al., 2000; Turk et al., 2000; Turk et al., 2012). Thus, cathepsin S is found predominantly in lymphatic tissues such as spleen, lymph nodes and peripheral leukocytes, cathepsin K is highly expressed in osteoclasts at sites of bone resorption, cathepsin V is thymus- and testis- specific and cathepsin W is predominantly expressed in lymphocytes. Studies with cathepsin S-deficient mice demonstrated that this protease is specifically required for processing of the invariant chain in peripheral antigen presenting cells such as macrophages, B-cells and dendritic cells (Shi et al., 1999; Villadangos et al., 1999).

Cathepsin V was suggested to be responsible for catalysis of the same reaction in human thymic epithelial cells (Turk *et al.*, 2000). Different cathepsins have thus been shown to be linked to antigen presentation, possibly acting through different pathways (Turk et al., 2002c; Niwa et al., 2012). Furthermore, cathepsin K was found to be critical for bone resorption, a part of the dynamic remodeling of bone (Saftig et al., 2000; Turk et al., 2000). Knockout of the cathepsin K gene in mice thus caused an impaired ability of osteoclasts to resorb bone during growth, which led to osteopetrosis, or increased bone density. In this state, bone demineralization proceeds normally but not digestion of the bone matrix. Another tissue-specific cathepsin, cathepsin V, was indicated to be involved in T-cell selection in humans, whereas a specific function in T-cell mediated cytotoxicity was proposed for cathepsin W

(Bhandoola et al., 2000; Turk et al., 2000). A number of novel, tissue-specific cathepsins in other mammals have also been reported, *i.e.* rat and mouse cathepsins Q, M, P, and R, which are primarily expressed in placenta and therefore may play important roles in implantation and embryonic development (Sol-Church et al., 2002). Extracellular cathepsins are potentially very destructive. In particular, abnormally high levels of extracellular papain-like cathepsins have been observed during development of various cancer types and metastasis (Calkins & Sloane, 1995; Henskens et al., 1996; Friedrich et al., 1999; Fernandez et al., 2001; Colella et al., 2002). In cancer progression, such cathepsins are responsible for activation of other proteases and for degradation of components of the extracellular matrix (Turk et al., 2000; Mohamed et al., 2006). However, intracellular cathepsin B has also been implicated to have specific roles in reducing malignancy, in particular by acting as an execution protease in tumor cells (Foghsgaard et al., 2001; Szpaderska & Frankfater, 2001). In agreement, lysosomal cathepsins have been indicated to participate in apoptosis also under other circumstances (Turket al., 2002a; Stoka et al., 2007).

Lysosomal cathepsins may also be secreted by other than cancer cells, *i.e.* by activated macrophages, osteoclasts and fibroblasts. Papain-like cathepsins, often extracellular or extra lysosomal, have thus been linked to development of such diseases as inflammation, rheumatoid arthritis and ostearthritis, gingivitis, osteoporosis, asthma, pulmonary emphysema, Alzheimer's disease, multiple sclerosis, muscular dystrophy, pancreatitis, myocardial disorders and glomerulonephritis (Turk et al., 1997, 2000, 2002b; Hook et al., 2012).

Genetic defects in cathepsins may cause severe physiological disorders. Mutations in the human cathepsin K gene that lead to a total loss of enzyme activity were found to be associated with pycnodysostosis, a genetic disorder with an osteopetrotic phenotype (Saftig et al., 2000; Lewiecki 2009), in agreement with the knock-out experiments in mice. Moreover, a loss of function due to mutation in the cathepsin C gene was shown to cause Papillon-Lefevre syndrome, characterized by early-onset periodontis and palmoplantar keratosis (Toomes et al., 1999).

REGULATION OF THE ACTIVITY OF CATHEPSINS

The activity of lysosomal cathepsins is regulated in many ways, *i.e.* at the gene translation and expression levels, by localization, synthesis of the proteases as inactive precursors, posttranslational modifications and by degradation (Chapman et al., 1997b; Turk et al., 1997, 2000). Alternative splicing of the cathepsins B and L genes has been observed to result in highly elevated translation and increased protease stability, as well as in an altered targeting mechanism and increased secretion. An essential regulating factor of the proteolytic activity of lysosomal cathepsins is the pH of the environment. Thus, the decrease in pH occurring during maturation of lysosomes presumably weakens the interaction between the propeptide and the proenzyme, favoring the activation process. Furthermore, cathepsins B, S and L are indicated to be irreversibly denaturated in lysosomes towards the end of the maturation process, when pH has decreased to a value of 3.8. In addition, many cathepsins, *viz*. cathepsins L, B, H, K, V and F, are unstable at neutral pH and therefore are less active outside lysosomes. Finally, endogenous protein inhibitors are considered to be common and most important means of regulating the activity of mature cathepsins that have escaped lysosomes, as well as of exogenous cysteine proteases released by different infectious microorganisms (Turk et al., 2001b; Turk et al., 2012).

PROTEIN INHIBITORS OF CYSTEINE PROTEINASES OF THE PAPAIN FAMILY

Depending on their major physiological function, protein inhibitors of proteases may be subdivided into emergency and modulatory inhibitors (Turk et al., 2002b). Protein inhibitors of the first type defend living cells and organisms in acute cases of occasionally increased proteolysis. They rapidly bind to their target proteases and inactivate these either irreversibly or for sufficiently long times. Two kinetic parameters have been suggested as criteria for whether protein inhibitors are capable of fulfilling this role. These are the delay time, which determines the time needed for ~99% completion of the binding reaction, and, in the case of reversible interactions, the stability time, during which an inhibitor—protease complex remains essentially

intact and which is defined as 1/10th of the half life of the bimolecular complex (Bieth, 1984). Protein inhibitors that act with a delay time of < 1s and have a stability time of > 10 min are regarded as emergency inhibitors.

Normally, inhibitors of this type are localized separately from their target proteases and are in molar excess over the protease when performing their protective function. By contrast, modulatory inhibitors are colocalized with their target proteases and regulate the endogenous activity of the latter at different levels, either by irreversible or pseudo-irreversible slow binding of activated proteases (delay type) or by rapid binding of these proteases in short-lived complexes (buffer type). Threshold inhibitors are also modulatory inhibitors that act similarly to the emergency inhibitors, although in lower physiological concentrations compared with those of the proteases. A change of conditions, such as salt concentration and pH, or the presence of cofactors can alter the properties of protein inhibitors and convert them from one type to another.

CYSTATINS

Cystatins are evolutionarily related proteins with a conserved physiological function. The cystatin superfamily comprises proteins with diverse primary structures, size and distribution. In particular, members of this superfamily can be one- or multidomain proteins, with or without disulfides, as well as glycosylated or non-glycosylated. Nevertheless, statistical analysis of their amino acid sequences suggests that cystatins originate from a common ancestor protein with no disulfide bonds (Rawlings & Barrett, 1990).

Cystatins inactivate target proteases reversibly and competitively by indirect blockage of the catalytic centers of these enzymes, thereby preventing substrate docking and cleavage. In general, cystatins bind target proteases with rate constants as high as 10^4 to 10^7 $M^{-1}s^{-1}$ and dissociation constants in the nM to fM range.They form very stable bimolecular complexes with the proteases, keeping these inactive for hours and weeks. Hence, cystatins can function as either emergency or threshold inhibitors. Proteins of the cystatin superfamily are traditionally subdivided according to their sequence similarities into three major families (Barrett et al., 1986a). Originally, closely related proteins, diverging by less than a half of their amino-acid residues,

were grouped into these individual families (Barrett, 1986; Rawlings & Barrett, 1990). However, later discovered cystatins have also been added to these families, despite a sequence similarity between some new family members and the old ones of substantially lower than 50%. Major characteristics for the classification of cystatins have instead become the size of the polypeptide chain and the presence or absence of internal disulfides. This simplification has led to uncertainties regarding the assignment of some cystatins to a particular family, and revision of the classification established to date, as well as introduction of new cystatin families or subfamilies, therefore appears to be necessary.

CLASSIFICATION OF CYSTATIN SUPERFAMILY

FAMILY 1 CYSTATINS (STEFINS)

Members of this family, also called as stefins, are the smallest in the cystatin superfamily, having a molecular mass of ~11 kDa. The single polypeptide chain of these proteins, consisting of ~100 residues, is folded into one domain. Family1cystatins lack both disulfide bridges and carbohydrates, thus representing the most primitive form of cystatins (Turk & Bode, 1991). This family includes cystatins A and B, found in different mammals: man (Machleidt et al., 1983;Green et al., 1984; Ritonjaet al., 1985; Abrahamson et al., 1986), cow (Turk et al., 1992, 1995a), rat (Takio et al., 1984; Takeda et al., 1985), sheep (Ritonjaet al., 1996), pig (Lenarcic et al., 1993, 1996) and mouse (Tsui et al., 1993), as well as bovine stefin C (Turket al., 1993) and pig stefin D (Lenarcicet al., 1996). Alignment of sequences of family 1 cystatins reveals three evolutionarily highly conserved regions, viz. a Gly residue in the N-terminal region, a QVVAG segment in the middle of the polypeptide chain and a LP pair in the C-terminal part of the chain.

Family 1 cystatins are located primarily intracellularly, in the cytosol (Turk & Bode, 1991), although they have also been detected in low concentrations in different extracellular fluids (Abrahamson et al., 1986). Cystatins A and B are the most studied members of the family. Cystatin A is found predominantly in various epithelial cells and in polymorphonuclear leukocytes (Green et al., 1984; Järvinen et al., 1987),

Figure 1: Diverse functions of thiol proteinase inhibitors.

whereas cystatin B is widespread among different cells and tissues (Davies & Barrett, 1984; Katunuma & Kominami, 1985; Barrett, et al., 1986b; Henskens et al., 1996). These differences in distribution may reflect different physiological roles of the two cystatins. Cystatin A also called as stefin **A** is often considered as a first-line guardian against cysteine proteases released by infectious microorganisms and parasites, whereas cystatin B presumable has a general protective function.

Stefin A/ Cystatin A

It is an inhibitor of cathepsin B present in human skin discovered by Fraki (1976). Later on Jarvinen (1978) studied it as "acid cysteine proteinase inhibitor" (ACPI). Brzin et al (1983) purified this inhibitor from blood leucocytes and named it as stefin. The amino acid sequence of this inhibitor was determined by Machleidt et al (1983). Green et al (1984) charachterized same type of CPI from human liver and later renamed it as cystatin A.

The most abundant source of stefin **A** is polymorphonuclear leucocytes in the liver (Davies and Barrett, 1984). It has also been localised to the strata corneum and granulosum of the epidermis (Rasanen et al., 1978) and has been found in extracts of squamous epithelia from the oesophagus, vagina (Rinne et al., 1978) and mouth (Jiirvinen et al., 1983). This selective distribution of stefin **A** correlates with those tissues which form the first line of defence against infective agents. It has thus been suggested (Barrett et al., 1986) that stefin **A** provides an important protective function as an inhibitor of cysteine proteinases which are utilised as invasive tools by many pathogenic organisms. In addition, Cystatin **A** is implicated in a number of disease states. The inhibitor has been detected in several epimeroid carcinomas including squamous cell carcinomata of the lung, skin, vulva, cervix and oesophagus (Rinne, 1979, 1980) but was absent from a variety of other carcinomas (Rinne, 1980; Rinne et al., 1984). It has been proposed that Cystatin **A** plays a key role in tumour invasion which is due either to its expression as a less active isoform or a lower level of expression (Lah et al., 1990). Cystatin **A** has also been found in the upper spinous layer of psoriatic cells (Hopsu-Havu et al., 1983a) and the serum level of this inhibitor has been shown to increase significantly in patients with cardiovascular disease (Hopsu- Havu et al., 1983b).

Cystatin A occurs in multiple isoelectric forms with predominantly acidic pI values in the range 4.5-5.0 (Jarvinen, 1978; Hopsu-Havu et al., 1983a). Rinnie et al (1978) detected cystatin A in extracts of squamous epithelia from oesophagus. It was also found in the dendritic reticulum cells of the lymph nodes (Rinnie et al., 1983), seminal plasma (Minakata and Asano, 1985), saliva, human nails (Tsushima, 1993) and in a number of epidermoid carcinomas (Rinnie, 1979, 1980; Rinnie et al., 1984a) as well as in bovine skin (Turk et al., 1995). Physiological concentration of human cystatin A was found to be 10-15 μg/ml in serum, 65-200 μg/g in epidermis of tongue and 8 μg/g in bone marrow. However its concentration varies between 0.3-8μg/g in lungs, liver and spleen (Hopsu-Havu et al., 1983b).

Cystatin α

Cystatin α is assumed to be a species variant of cystatin A found in rats. This protein was characterized by Jarvinen (1979) as specific inhibitor of cysteine proteinases from rat skin having Mr of 13KDa. Cystatin α is generally found on the epidermal layer (Jarvinen et al., 1978) and various other squamous epithelia (Rinnie et al., 1978). Katunuma and Kominami (1985) quantified the concentration in skin and small intestine as 2800 μg/mg protein and 820μg/g protein, respectively. The serum level of cystatin α in rat liver was 2.5μg/ml. The N-terminal sequence of bovine stefin A was identified as a truncated form of bovine cystatin C (Hirado et al., 1985). The sequence identity is 86% and 88% with human cystatin A, N- and C- terminal part.

Cystatin B

Cystatin B was detected as an inhibitor of cathepsin B and H in human tissues by Lenney et al (1979). Jarvinen and Rinnie (1982) purified cystatin B from human spleen. Green et al (1984) purified cystatin B form from human spleen and liver with separation of multiple forms. Cystatin B is a relatively basic protein with pI value of 6.25 and 6.35 for the two forms of cystatin B (Green et al., 1984). Jarvinen and Rinnie (1982) reported pI values of 6.5 and 7.5. Human cystatin B forms dimers (Jarvinen and Rinnie 1982; Green et al., 1984, Rinnie et al., 1984b) which shows no inhibitory activity.

Cystatin B is ubiquitously distributed in various cells and tissues like epithelial cells, lymphocytes (Davies and Barrett, 1984), monocytes (Hopsu-Havu et al., 1984) and to much lesser extent in seminal plasma (Brzin, 1982). Jarvinen et al (1983) have shown cystatin B in cells of the upper layers of epithelium of the oral mucosa. Confocal microscopy analysis showed the localization of cystatin mainly in the nucleus of proliferating cells and both in the nucleus and cytoplasm of differentiated cells (Riccio, 2001). This cystatin is involved in the protection against uncontrolled activity of host lysosomal CP (Henskens, 1996). Crystal structure of recombinant human cystatin B has been elucidated by X-ray crystallography analysis (Stubbs et al., 1990). Cystatin B is more abundant in all cells than cystatin A with the exception of PMNs (polymorphonucleocytes).

Cystatin β

Cystatin β a species variant of cystatin B was isolated from rat liver (Finkelstadt, 1957; Lenney, 1979) with pI values ranging from 5.04 to 5.6 (Hirado et al., 1981). Cystatin B has an even distribution in the tissues and is more abundant than cystatin α in all tissues except skin. This characteristic of cystatin β resembles cystatin B of human variant.

Stefin C

Several low molecular weight CPI's have been found in bovine thymus. Although kinetically similar they differ in the primary structure and pI values. Stefin C is a unique among the inhibitors from stefin family which was found in multiple forms resulting from the cleavage of Asn 5 – Leu 6 bond of inhibitor (Turk et al., 1993). This property is commom to stefins and other low molecular weight CPI's (Barrett et al., 1986; Brzin et al., 1982; Kopitar et al., 1989). Stefin C consists of 101 amino acids and its Mr is calculated to be 11,546. The inhibitor was found to be acidic with pI values ranging from 4.5 and 5.6 (Turk et al., 1993). It exhibits considerable sequence homology with other inhibitors from stefin family with maximum homology of 84.7% with bovine stefin B, 72.4% with human cystatin B and 47.9% with human stefin A. It was identified as the first tryptophan containing stefin at position 2 (Turk et al., 1993).

FAMILY 2 CYSTATINS (CYSTATINS)

Family 2 members, uniformly termed as cystatins, are also single-chain, one-domain proteins, although somewhat larger in MW than the family 1 inhibitors, ~12 - 13 kDa. The polypeptide chains of cystatins of family 2 are ~120 residues long and contain two characteristic disulfide bridges, at conserved positions in the C-terminal part (residues 71–81 and 95–115 in chicken cystatin numbering) (Rawlings & Barrett, 1990). In addition to the four cysteines, three other conserved regions are characteristic for members of cystatin family 2. A Gly residue in the N-terminal region, a QxVxG sequence in the middle of the chain and a PW dipeptide in the C-terminal part of a molecule are present in all cystatins of this family. Most cystatins known are nonglycosylated, although mouse and rat cystatin C and also cystatins E/M and F are glycoproteins (Esnard et al., 1990; Ni et al., 1997, 1998). Family 2 cystatins are synthesized as preproteins with a signal peptide and are present predominantly in extracellular fluids although some have also been found intracellularly (Halfon et al., 1998; Abrahamson, 1994; Nathanson et al., 2002; Nandy et al., 2013). The family includes mammalian cystatins: cystatin C of several different species (Grubb & Löfberg, 1982; Hirado et al., 1985; Esnard et al., 1988; Solem et al., 1990), cystatin D (Freije et al., 1991, 1993), cystatin E/M (Ni et al., 1997; Sotiropoulou et al., 1997; Zeeuwen et al., 2001), cystatin F/leukocystatin (Ni et al., 1998; Morita et al., 1999), cystatins S, SA and SN (Isemura et al., 1984, 1986, 1987), and also chicken egg-white cystatin (Turk et al., 1983). Aminoacid sequences of representatives of families 1 and 2 of the cystatin superfamily. CA, human cystatin A; CB, human cystatin B; CC, human cystatin C; EWC, chicken egg-white cystatin. Protein sequences were taken from the SWISS-PROT database (accession codes CYTA_HUMAN, CYTB_HUMAN, CYTC_HUMAN and CYT_CHICK, respectively) and aligned within each family. Signal peptides of family 2 cystatins are not shown. The motifs conserved within each family are in bold.

Cystatin C

The best studied representatives of family 2 cystatins are chicken cystatin and human cystatin C, the latter being the most abundant of the inhibitors of the cystatin superfamily in all human body fluids examined, with the highest concentration in

seminal plasma and cerebrospinal fluid (Abrahamson et al., 1986; Olafsson et al., 1988; Abrahamson, 1994). The broad distribution of thisprotein implies its essential role in protection of the organism against exogenous and endogenous, occasionally released, cysteine proteases. Cystatins of the S-type are found predominantly in saliva but also in tears, urine and seminal plasma (Isemura et al., 1984; Abrahamson et al., 1986). The distribution of cystatins D, E and F has not been investigated as extensively as that of cystatin C. Cystatin D has been found in saliva and tears and cystatin E in small amounts in urine (Freije et al., 1993; Ni et al., 1997).

Originally cystatin C was termed as γ-trace or post γ-globulin isolated from human CSF (cerebro spinal fluid) because of its basic nature and γ-electrophoretic mobility (Barrett et al., 1984; Brzin, 1984). It was also found in the urine from patients with renal failure (Butler and Flynn, 1961) and ascitic and pleural fluids (Hochwald and Thornbecke, 1962). Later cystatin C was also detected in saliva and normal serum (Cejka and Fleischmann, 1973) and in seminal plasma (Colle et al., 1976). Lofberg and Grubb (1979) determined cystatin C concentration of 5.8 μg/ml in CSF, 1.8 mg/ml in saliva, 0.095 μg/ml in normal urine, 21 μg/ml in urine from patients with renal tubular disorders and 1.1 μg/ml in normal plasma. Seminal plasma contains 50μg/ml i.e. 36 fold higher concentration of inhibitor than in normal blood plasma (Grubb et al., 1983). Abrahamson (1991) found cystatin C concentration of 3.4μg/ml in saliva from submandibular glands and 0.96μg/ml from the parotid glands. The high level of cystatin C in the CSF compared with blood plasma suggests the production of cystatin in central nervous system (Lofberg and Grubb, 1979). The protein is able to cross blood barrier to the vascular space thereby rapidly filtered in the glomeruli and catabolised in the renal tubular cells.

Despite wide spread extracellular distribution cystatin has also been detected intracellularly in brain cortical nerves (Lofberg et al., 1981a), normal and neoplastic neuroendocrine cells in the adrenal medulla (Lofberg et al., 1982), cells of pancreatic islets (Lofberg et al., 1981b), thyroid (Lofberg et al., 1983) and pituitary (Lofberg et al., 1984). Homologues of human cystatin C have been found in several mammals including mouse, rat, cow and dog (Esnard et al., 1988; Hirado et al., 1985; Park et al., 2013).

Figure 2a: 3-D structure of of papain from carica papaya.

Figure 2b: Three dimensional structure of Cystatin C.

Chicken Cystatin

It was first isolated from chicken egg white by Fossum amd Whitaker (1968) and later by Sen and Whitaker (1973) as low Mr tight binding cysteine proteinase inhibitor of ficin, papain, cathepsin B and cathepsin C. A molecular weight of about 13 KDa has been determined from the amino acid composition of this protein (Lindahl et al., 1988). It forms a tight, reversible 1:1 complex with most known cysteine proteinases (Nicklin and Barret, 1984). Its amino acid sequence was determined independently by Turk et al (1983) and Schwabe et al (1984) and consists of 116 residues with molecular weight of ~ 14KDa. It occurs in two major isoelectric forms (form 1, pI = 6.5, and form 2, pi = 5.6). Chicken cystatin has also been detected in the serum of both male and female chickens at a concentration of 80μg/ml in egg white and 11μg/ml in the serum (Anastasi et al., 1983) and in chicken muscle cell (Wood et al., 1985) and showed resemblance to cystatin C. Chicken cystatin has been shown to alter intracellular proteolytic processing of poliovirus proteins, resulting in a reduction of virus yield (Korant et al., 1985). These investigations suggested that the cystatin is able to penetrate in the cellular cytoplasm and inhibits the action of poliovirus coded proteinases. As a result, cystatins or their derivatives are being considered as potential antiviral agents (Korant et al., 1988).

Cystatin D

Cystatin D is a member of human cystatin multigene family and was cloned from a genomic library using cystatin C cDNA probe (Freije et al., 1991).The inhibitor consists of 122 amino acid residues having a Mr of 13,885. The deduced amino acid sequence includes a putative signal peptide and has 51-55% homology with either cystatin C or secreatory gland cystatins S, SA and SN. Cystatin D is a relatively neutral protein with pI values in the range 6.8 to 7.0 (Freije et al., 1991). Cystatin D is expressed in parotid gland, saliva and tears (Balbin et al., 1994). This tissue restricted expression is in marked contrast with a wider distribution of all other family 2 cystatins.

Figure 3: 3-D structure of chicken cystatin.

Cystatin S

Human saliva contains several low Mr acidic proteins which include CPI also (Isemura et al., 1984b; Minakata and Asano, 1984). The first salivary inhibitor was purified and sequenced by Isemura et al (1984a) and named as cystatin S. It was found to have 54 and 41% sequence homology with cystatin C and chicken cystatin, respectively (Isemura 1984a, b). Cystatin S contains no phosphate, in contrast to other salivary proteins (Isemura et al., 1984b). This inhibitor has also been isolated from human submaxillary, submandibular and sublingual glands and found to be present in the serous cells of parotid and submaxillary glands (Isemura 1984b). It is assumed that these cells are responsible for its synthesis. The protein has also been found in tears, serum, urine, bile, pancreas and bronchus (Isemura et al., 1986).

VARIANTS OF CYSTATINS

Several molecular variants of cystatin S have been studied by Isemura et al (1986). They differ in their N-terminal sequence and pI value (Minakata and Asano, 1984; Isemura, 1986; Shomers et al., 1982ab). Differences in pI values result from phosphorylation of residues ser 3 or ser 1 in salivary cystatin (Isemura et al., 1991).

Cystatin SN: Originally known as cystatin SV or SA-1 (Abrahamson et al., 1986). The protein consists of 121 amino acid residues and Mr of 14,316. The pI value is in the range of 6.6-6.8. Rat SN cystatin has also been purified from the submandibular glands of rats subjected to chronic IPR (Isoproterenol) treatment (Bedi, 1989a) and its amino acid sequence has been determined by Bedi (1989b).

Cystatin SA: The protein consisit of 121 amino acid residues and Mr (14,351) is slightly higher than cystatin SN. It has acidic pI value of 4.6 (Isemura et al., 1991). Cystatin SA isolated from saliva had N-terminal residue Glu (Isemura et al., 1986).

CYSTATIN E/F AND CYSTATIN M

The recently reported cystatins E/M and F have only ~30%sequence identity with other cystatins of family 2. Nevertheless, they have been included in this family,

mainly on the basis of their size and the presence of two conserved disulfide bonds, as well as the other evolutionarily conserved regions of theprimary structure that are typical for family 2 cystatins. Cystatin F differs from the other family members by the presence of one additional disulfide bridge in the N-terminal part of the molecule. Chicken cystatin and cystatins C, E/M and F, besides being inhibitors of cysteine proteases of family C1, like all cystatins, also inhibit an entirely different enzyme, mammalian legumain, of family C13 of cysteine proteases (Chen et al., 1997; Alvarez-Fernandez et al., 1999).

Cystatin F has been detected inlowconcentrations inblood and shown to besecreted intrace amounts by B cells and in significant amounts by T cells in culture (Ni et al., 1998). Most information available concerns the expression pattern of the genes for these cystatins obtained by DNA and RNA techniques. The cystatin D gene has been shown to be expressed only in the parotid glands (Freije et al., 1991). Data on expression of the gene for cystatin E/M are inconsistent.Different groups have reported diverse expression patterns of this cystatin, with expression being highest in liver, ovary and pancreas (Ni et al., 1997), in primary breast cancer tumors with no expression in liver or brain (Sotiropoulouet al., 1997) or restricted to epithelial cells and sweat glands (Zeeuwen et al., 2001). Expression of the cystatin F gene has been detected in a number of tissues, the highest levels being found in peripheral blood leukocytes and spleen (Ni et al., 1998) and also in hematopoietic cells and some human cancer cell lines (Halfon et al., 1998; Morita et al., 1999).

Cystatin E

Ni et al (1997) found this type of cystatin by expression of the amniotic cell and fetal cell epithelial cell cDNA libraries. The mature protein is a polypeptide of 121 amino acid residues and 28 residue signal peptide having a MW of 15,000. Cystatin E resembles family 2 cystatins structurally in containing two protective disulphide bridges and by being a secreated protein. The inhibitor has unusual charachteristic of being a glycoprotein, carrying an N-linked oligosaccharide at Asn 108. Cystatin E is a functionally tight binding CPI. The amino acid sequence similarities of cystatin E with other cystatins are 26-34% with human family 2 cystatins (C, D, S, SN and SA), 18 and 23% with family 1 cystatin A and B and less than 30% with domain 2 and 3 of

family 3 cystatins (kininogens) (Ni et al., 1997). The protein has Pro-Trp pair towards the C-terminal like that of family 2 and 3 cystatins. The gene of this inhibitor has been localized on chromosome 11 (Stenmann et al., 1997).

Cystatin E has been detected in variety of specialized tissues and organs (Ni et al., 1997). High concentration has been found in uterus and liver, pancreas, heart, spleen, small intestine and peripleural blood leucocytes and low cystatin E content is found in brain, testis and kidney (Ni et al., 1997). The inhibitor serves a protective role during fetal development.

Cystatin F

Another described human member of the cystatin superfamily was identified as cystatin F by Ni et al (1998). The whole sequence contained signal peptide and a mature protein of 126 amino acid residues with two disulphide bonds which resembles family 2 cystatins. Unlike other members, cystatin F has two additional cys residues indicating the presence of an extra disulphide bridge which stabilizes the N-terminal segment of the molecule. The protein has an extended N-terminal segment being 6-10 residues longer than the other single domain cystatins. Cystatin F has a Mr of 14,543 as determined by sequence analysis (Ni et al., 1998). Cystatin F has two N-linked carbohydrate chains at position 36 and 88. It has 30-34% sequence homology with human family 2 cystatins and 29% homology with cystatin E, even lower 22-20% identity to cystatin A and B and 24% homology with domain 3 of kininogens.The distribution pattern of cystatin F is very different from that of cystatin C or E. Blood contains low levels of the inhibitor (0.9μg/ml). The highest concentration is present in the spleen and peripheral blood leucocytes, moderate levels in thymus and small intestine. Cystatin F is found mostly in the immune cells (Ni et al., 1998).

Cystatin M

Cystatin M was discovered in the primary tumor cell line. It is synthesized and secreated as a precursor protein (Von Heinjie, 1985). The predicted molecular mass is approximately 14.3 KDa and pI of 7.8 for the mature protein. A single disulphide

bond is present (Sotirpoulou et al., 1997). An N-glycosylated form of cystatin M of 20-22 KDa was co-immunoprecipitated and accounted for about 30-40% of total cystatin M protein. Both forms of native cystatin M forms occurred intracellularly. Cystatin M resembles other members of the family in having the three conserved domains. Cystatin M has 40% homology to human family 2 cystatins (Sotirpoulou et al., 1997). The overall homology between cystatin M and other cystatins range from 30-40% for conserved amino acid residues and 25-33% for identical amino acids. It shows closest homology to cystatin C, they share 33% identical and 38% conserved amino acid residues.

Cystatin M is expressed by normal mammary epithelial cells and a variety of human tissues but is not expressed in many metastatic mammary tumor cell lines. High levels of cystatin M are present in lung, skeletal muscle, kidney and pancreas. Low level was observed in heart tissue.

Various mammalian and non-mammalian sources from where low Mr CPI have been isolated are summarized in Table 1.1

FAMILY 3 CYSTATINS (KININOGENS)

Family 3, or the kininogen family, comprises the largest and most complex inhibitors of the cystatin superfamily, with molecular masses of ~50 or ~70–80 kDa (Barrett et al., 1986a; Turk et al., 2002b). The single chain of a kininogen is folded into three family 2-cystatin like domains, all of them being glycosylated. In addition to the six conserved disulfide bridges, two in each domain, there are two additional disulfides in the N-terminal region of each of the kininogen domains 2 and 3. Kininogens are synthesized with a signal peptide and found extracellularly, with the highest concentrations, ~10 μM, in blood plasma and synovial fluid (Abrahamsonet al., 1986; Abrahamson, 1994). There are three 20 kininogen types known, H-, L- and T-kininogens. High- and low-molecular weight kininogens, H- and L-kininogen, respectively, are encoded by the same gene (Kitamura et al., 1985) and are identical in their N-terminal parts, but H-kininogen is longer, containing an unrelated C-terminal region due to alternative splicing. T-kininogens, also called thiostatins, are found only in rats (MullerEsterl et al., 1986). Kininogens are multifunctional proteins and

werefirst identified as precursors to vasoactive peptides, the kinins. The kininsequence is located in the C-terminal part of a kininogen, and release of the kinin fragment by kallikrein cleavage converts the parent H- or L-kininogen to a two-chain form, stabilized by one interchain disulfide bond. The first domain of a kininogen is inactive in inhibition of cysteine proteases, whereas the two other domains can simultaneously bind two molecules of papain-like proteases (DeLa Cadena & Colman, 1991; Turk et al., 1995b, 1996a). The second domain also inhibits calpain (Salvesen et al., 1986)

OTHER CYSTATINS AND CYSTATIN-RELATED PROTEINS

Many other proteins with inhibitory activity against papain-like cysteine proteases and with sequences related to well-studied mammalian cystatins have been discovered. Among them, a large group consists of defensive inhibitors from plants, phytocystatins, which have structural features intermediate between those of families 1 and 2 cystatins (Rawlings & Barrett, 1990; Brown & Dziegielewska, 1997; Margis et al., 1998). Phytocystatins also possess a specific consensus sequence that is absent in other cystatins. Therefore classification of plant cystatins as an independent family of the cystatin superfamily has been suggested. Cystatins from rice (Abe et al., 1988; Kondo et al., 1991), corn (Abe et al., 1992), potato (Waldronet al., 1993), avocado (Kimuraet al., 1995), papaya (Song et al., 1995), soy (Misaka et al., 1996), chelidonium majus (Rogelj et al., 1998), sunflower (Kouzuma et al., 2001) and lima bean (Lawrence & Nielsen, 2001) have been described.

Related inhibitors have also been found in various lower organisms, such as African puff adder (Ritonja et al., 1987), carp (Tsai et al., 1996), horseshoe crab (Agarwala et al., 1996), Sarcophaga peregrine larvae (Suzuki & Natori, 1985; Saito et al., 1989), Drosophila (Delbridge & Kelly, 1990) and nematodes (Manoury et al., 2001; Hartmann et al., 2002). Cystatins in such lower organisms are more related to family2 cystatins than the phytocystatins are (Turk et al., 1997). Several novel members of the cystatin superfamily have been identified on the basis of sequences derived from mouse cDNA libraries, viz. a cystatin-related epididymal and spermatogenic protein,

Table-1.1: LOW AND HIGH MOLECULAR WEIGHT CPI FROM DIFFERENT SOURCES

Source	Tissue	Mr	pI	Reference
African puff (Bitis arietus)	Adder	13,000	6.5	Evans and Barrett, 1987
Beef	Nasal cartilage	13,000	4.8-7.0	Roughly et al., 1978
	spleen			Brzin et al., 1982
Bombyx mori	Hemolymph	-	-	Yamamoto et al., 1999
Bovine	Brain	25,000	4.7	Aghajanyan et al., 1996
	Hoof	11,406	-	Tsushima et al., 1996
		14,300		Zabari et al., 1993
	Colostrums	12,787	10.0-10.3	Hirado et al., 1984
Dog	Colostrum	-	-	Poulik et al., 1981
	Parotid gland & Kidney	-	-	Sekine and Poulik, 1982
Guinea pig	Skin	-	Acidic	Jarvinen, 1976
Horse shoe Crab	Hemocytes	12,600	-	Agarwala et al., 1996
Hog	Kidney	-	-	Lenney et al., 1979
Human	Liver	12,400	-	Green et al., 1984
	Squamous cell	-	-	Rinnie et al., 1980
	Spleen	11,400 12,000	4.7-50 6.0-6.5	Jarvinen and Rinnie, 1982
Rabbit	Liver	5000-10,000 12,5000	-	Pontremoli et al., 1983 Hayashi,1975;Tokaji,1971
Rat	Brain	-	-	Kopitar et al., 1983
Trypanosoma cruz	-	11,000	-	Monteiro et al., 2001
	Ovary	12,000	acidic	Tsai et al., 1996
Carp	Skin	-	-	Jarvinen, 1976
Cat	Egg	-	-	Yamashita and Konagawa, 1996
Chum Salmon	Brain	70,800		Sumbul and Bano, 2006
Goat		12,720		
	Pancreas	44,000		Priyadarshni and bano, 2009
	Kidney	67,000		Zehra et al., 2005
	Lung	66,400 76,400		Khan and Bano, 2009

termed Cres, and also Cres2, Cres3, testatin and cystatins T, TE-1 and SC (Tohonen et al., 1998; Kanno et al., 1999; Shoemaker et al., 2000; Li et al., 2002; Hsia & Cornwall, 2003). Additionally, cDNAs and mRNAs for the human counterparts of mouse Cres, Cres2 (termed cystatin 11 in human) and testatin were described (Cornwall et al., 1999; Hamil et al., 2002; Wassler et al., 2002). These proteins, specifically expressed in the reproductive and neuroendocrine systems, were suggested to represent a new Cres subgroup of cystatin family 2 (Hsia & Cornwall, 2003; Cornwall & Hsia, 2003), as theyhave the conserved PW motif but lack the two other conserved motifs. A further search of DNA databases has yielded a number of new Cres-like sequences in species ranging from nematode to human (Cornwall & Hsia, 2003). The mouse Cres protein was shown to not inhibit the cysteine proteases papain and cathepsin B but the calcium-dependent serine protease, prohormone convertase 2 (Cornwall et al., 2003). A further novel human cystatin-like molecule (CLM), ubiquitously expressed in normal tissues, has been shown to be related to cystatins C, D, S, SA and SN and especially to mouse and human testatins (>50% similarity) (Sun et al., 2003). However, CLM lacks any conserved cystatin motifs, and as yet no protease inhibitory activity has been demonstrated for this protein. In addition, other proteins with cystatin-related sequences but without inhibitory activity have been identified (Brown & Dziegielewska, 1997). Among them are fetuin and histidine rich glycoprotein, both containing two cystatin-like domains with the conserved disulfide bonds but lacking the conserved sequence motifs of cystatins.

Furthermore, another family of some new cystatin-like proteins is represented by **cathelins**, the overall fold of which is indicated to be very similar to that of cystatins despite of a low degree of sequence identity (Lenarcic et al., 1993; Yang et al., 2003). Inactive cystatin-like proteins may represent families that by divergence and/or gene duplication have evolved from the archetypal cystatins.The first larger group of other cystatins are **thyropins** (thyroglobulin type-1 proteinase inhibitors) which are inhibitors of cysteine proteinases containing at least one thyroglobulin type-1 domian (Lenarcic ad Bevec, 1998). **Equistatin** from sea anemone inhibits cysteine proteinases and is structurally similar to thyroglobulin type-1 (Lenarcic et al., 1997). It also inhibits cathepsin D which is an aspartic proteinase (Lenarcic and Turk, 1999) whereas traditional cystatins do not inhibit other classes of proteinases. Equistatin, a

Figure 4: Diagramatic representation of the chain structure of proteins in cystatin Superfamily.

.

New Inhibitor of Cysteine Proteinases from *Actinia equina*, is Structurally Related to Thyroglobulin Type-1 Domain.

Staphostatins are newly described whose members are the specific inhibitors of staphylococcal cysteine proteinases. Three members of this family have been described – staphostatins A and B from *Staphylococcus aureus* and staphostatin A from *Staphylococcus epidermidis* (Filipek et al. 2003).

Clitocybin (P82314) is a new type of cysteine proteinase inhibitor from a mushroom. It appers to be related to fungal lectins and the team that found it suggested a new family of cysteine proteinase inhibitors called mycocypins (Brzin et al., 2000).

Chagasin is a cysteine proteinase inhibitor found in Trypanozoma cruzi inhibiting both cruzipain and papain, but it has no homology with cystatins or with other published sequences in the databanks (Monteiro et al., 2001).

Saxiphilin is a neurotoxin-binding protein from bullfrog and a homolog of transferrin with a pair of such Thyr-1 modules located in the N-lobe. Saxiphilin is now characterized as a potent inhibitor of three cysteine proteinases as follows: papain, human cathepsin B, and cathepsin L (lenarcic et al., 2000).

Bombyx cysteine proteinase inhibitor (BCPI) was found in the hemolymph of silkmoth Bombyx mori as a specific inhibitor of Boot box cysteine proteinase (BCP) (Yamamoto et al., 1999). BCP is a cathepsin L-like cysteine proteinase, belonging to a papain superfamily, and is also present in the hemolymph and accumulates in mature eggs (Kageyama et al., 1990; Yamaanoto et al., 1994; Yamaanoto et al., 2004). Two similar but distinct BCPIs have been isolated; BCPI α and BCPI β. Both BCPIs are inhibitory for other cysteine proteinases such as papain, and cathepsins L and B, but do not inhibit trypsin or pepsin. Interestingly, the amino acid sequence of BCPI is significantly homologous to propeptide sequences of certain cysteine proteinases, such as baculovirus, plantand protozoan cysteine proteinases (Yamamoto et al., 1999).

CRES Proteins There are a number of other cystatins or cysta, expressed in different tissues and cell types in human and other mammals. A novel **cystatin-related epididymal specific (CRES)** gene was found in mouse epididymis, showing substantial homology with those of well established protein inhibitors, cystatins

(Cornwall, 1992). However the CRES gene does not contain the highly conserved QXVXG region and P-W pair which are crucial for cathepsin inhibition. The CRES gene is almost restricted and much less expressed in testis and no expression in any other tissue was found. A gene was isolated from mouse foetus, related to the genes that encode cystatins and was named **testatin** (Tohonen et al., 1998). Testatin expression is restricted to pre-sartoli cells and its expression was high during the early events of testis development. Two more genes from sartoli cells, named **cystatin SC** and **cystatin TE-1**, were isolated which were detected only in the testis and highly expressed in testis and epididymis, respectively (Li et al., 2002). The role of this subgroup of cystatins might be regulation of proteolysis in this reproductive tract as well as protection against invading pathogens by inhibiting microbial proteinases, as shown by cystatin II (Hamil et al., 2002).

Tick cystatins, Genes encoding cystatins have also been found in several ticks which constitute the main vector of Lyme disease in the USA and Europe. The two cystatins transcripts are encoded by two different genes in the tick *Ixodes Scapularis.* Both cystatins were expressed and were named as **sialostatin L** (Kotsyfakis et al. 2006; Kotsyfakis et al. 2007). The name indicates the strong inhibition of cathepsin L. These two sialostatins, which were found in saliva, show 75% identity in their sequence and inhibit almost equally cathepsin L, with Ki=4.7 nM, and cathepsin V with Ki=57 nM. Cathepsin L, is a known collagenolytic enzyme and plays an important role extracellularly and intracellularly. Its collagenolytic activity is inhibited by chicken cystatin (Maciewicz et al. 1987). Consequently, sialocystatin L displays an anti-inflammatory role and inhibits proliferation of cytotoxic T-lymphocytes (Kotsyfakis et al. 2006).

PHYTOCYSTATINS

In plants, inhibitors of cysteine proteinases are known as phytocystatins. They contain the QXVXG region of type 2 cystatins but also resemble stefins in the absence of disulphide bonds (Arai et al. 2002), providing a transitional link between type 1 and type 2 cystatins. There are numerous phytocystatins expressed and characterised on the protein level from corn (Abe et al. 1992), rice (Chen et al. 1992), soyabean (Lalitha et al. 2005) and others. C–terminal extended phytocystatins were

found as bifunctional inhibitors of papain and legumain (Martinez et al. 2007). In addition, a "multicystatin' containing two cystatin like domains were isolated from cowpea leaves (Diop et al. 2004) and tomato leaves (Wu and Haard 2000). Also there are certain plant proteins like monellin which lack the cysteine proteinase inhibitory activity but have a cystatin like three dimensional structure (Grzonka et al. 2001). Phytocystatins and other inhibitors are important for plant defence response to insect predation, may act to resist infection by some nematodes (Koiwa et al. 1997) play a crucial role in response to various conditions (Diop et al. 2004; Brzin and Kidric 1995) and show great potential tools for genetically engineered resistance of crop plants against pests (Aguiar et al. 2006).

NON SPECIFIC INHIBTORS OF CYSTEINE PROTEINASES OF THE PAPAIN FAMILY

A number of other proteins exhibit inhibitory activity against lysosomal papain-like cysteine proteinases. α-2 macroglobulin unspecifically trap endopeptidases of different types, blocking the access of protein substrates to the active site of the trapped proteinases without inactivating them (Armstrong 2001).

Thyropins, Thyropins share considerable homology with the thyroglobulin type 1-domain present in eleven copies in the prohormone thyroglobulin and in a number of other proteins from other organisms (Molina et al. 1996). Cathepsin L is specifically inhibited by a mammalian representative of this class, the major histocompatibility complex (MHC) class II-associated p41 invariant chain fragment (Bevec and Garver 1996; Guncar et al. 1999).

Serpins, Serpins as typical protein inhibitors of serine type proteinases can also inhibit cysteine-type proteinases including papain-family of cysteine type proteinases in cross-inhibiton (Turk et al. 2002). This was demonstrated for the human squamous cell carcinoma antigen1 (SCCA) as a potent inhibitor of cathepsin K, L and S129, its mouse ortholog SQ-N-5 which inhibits an addition cathepsins V but not cathepsin B and H (Schick et al. 1998).

Chagasin from trypanosome cruzi has been shown to inhibit a number a number of papain-like proteinases (Monteiro et al. 2001). Chagasin-like proteins were found to be encoded by genomes of several eukaryotes, bacteria and archea, and thus appears to be the first protein inhibitor of cysteine proteinases identified in prokaryotes (Rigden et al. 2002). Chagasin is the only proteinase inhibitors known that adopt an immunoglobulin type fold (Rigden et al. 2001). In addition, α-2 macroglobulin is known as the only protein inhibitor that can inhibit several types of proteinases, including the papain family of cysteine proteinases (Mason 1989).

Clitocybin is a new type of cysteine proteinase inhibitor from a mushroom. Clitocybin appears to be related to fungal lectins and the team that found it suggested a new family of cysteine proteinase inhibitors called mycocypins (Brzin et al. 2000).

VARIANTS OF CYSTATINS

Divergent cystatins showing significant homology to stefins, cystatins and kininogens have been expressed/purified and characterised from venom of African puff adder (*Bitis arietans)* (Evans and Barrett, 1987); from perilymph of flesh fly larvae (Suzuki and Natori 1985); from *Drosophila melanogaster* (Delbridge and Kelly 1990), goat brain (Sumbul and Bano 2006), goat lung (Khan and Bano 2009). Some of the mammalian and non mammalian sources from where cysteine proteinase inhibitors (CPIs) have been isolated are summarised in Table 1.1.

BIOLOGICAL ROLE OF CYSTATINS

Cystatins constitute a powerful regulatory system for endogenous cysteine proteinases which are often secreted or leaking from the lysosomes of dying or diseased cells. Cystatins are proteins that tightly bind and inhibit the harmful effect of cysteine proteinases (Ekiel et al. 1997). A major role in protecting the organism against the harmful activities of cysteine proteases, originally attributed to cystatins on the basis of their high inhibitory potential, has been supported by many experimental studies. Colocalization of cathepsins and cystatins as well as significantly altered levels of the inhibitors have thus been observed in many pathological states, including cancer,

myoclonal epilepsy, cardiovascular disease, and several autoimmune diseases, such as asthma (Hopsu-Havu et al., 1983a, b; Friedrich et al., 1999; Kos et al., 2000a, b; Cimerman et al., 2001; Locsey et al., 2001; Riccio et al., 2001; Colella et al., 2002). In general, cystatin concentrations are elevated in patients suffering from inflammatory diseases and autoimmunity (reviewed in Barrett, 1986; Hansen et al., 2000; Mangge et al., 2000). In contrast, development of cancer may be promoted by expression of cystatins in lower levels or as less active isoforms (Lah et al., 1992a, b; Zore et al., 2001; Zajc et al., 2002). Cystatin A was thus not detected in a number of epimeroid carcinomas and prostate adenocarcinomas, contrary to in normal tissues (Rinne et al., 1984a, b; Soderstrom et al., 1995; Mirtti et al., 2003), and cystatins B and E/M were found to be down regulated in prostate and esophageal carcinomas and breast cancer, respectively (Sotiropoulou et al., 1997; Shiraishi et al., 1998; Mirtti et al., 2003). Such lower concentrations of the inhibitors of harmful cathepsins associated with cancer progress may promote extracellular matrix distruction and tumor invasion (Friedrich et al., 1999; Yano et al., 2001). In agreement with this proposal, cystatin C, the most potent and abundant of human cystatins, was shown to prevent tumor cell invasion and metastasis both *in vitro* and *in vivo* (Kolkhorst et al., 1998; Cox et al., 1999b).

Apart from a general protective function, regulatory roles exerted by cystatins appear also to be essential for normal development and functioning of the organism in many cases. In particular, cystatin A was proposed to be involved in differentiation of the epidermis (Rinne et al., 1984a; Jarvinen et al., 1987). Similarly, cystatin B was implicated in the control of myofibril turnover in the cytoplasm of muscle fibers (Spanier & Bird, 1982). In addition, cystatins were suggested to participate in regulation of cell proliferation (Hiwasa et al., 1995), as overexpression of cystatin A was found to induce increased levels of cyclin B1, which is involved in cell-cycle regulation. Furthermore, cystatin C was shown to be involved in local and precise balancing of the activity of testicular cathepsin L, restricting the action of the enzyme to a specific step of spermatozoa maturation (Peloille et al., 1997). A regulatory function was also proposed for salivary cystatins. Besides protecting periodontal tissues and presumable exerting antibacterial and antiviral activity (Abrahamson et al., 1997), these cystatins may modulate the mineralization process at the saliva-enamel interface (Tseng et al., 2000). *Cres* protein was also implied to be a regulatory protein,

controlling prohormone and proprotein processing within the reproductive and neuroendocrine systems (Cornwall et al., 2003). Similarly, cystatin F appeared to play a role in immune regulation through inhibition of a unique target in the hematopoietic system (Halfon et al., 1998). Interestingly, cystatins isolated from parasites have shown immunomodulatory activities affecting host immune systems (Manoury etal., 2001; Vray et al., 2002). Nippocystatin secreted by an intestinal nematode, *Nippostrongylus brasiliensis*, was thus implicated in modulation of antigen processing in antigen-presenting cells in mice (Dainichi et al., 2001). Furthermore, filarial nematode cystatins were reported to be efficient in upregulating nitric oxide production by macrophages in mice, a mediator known to be an effector molecule against filarial worms (Hartmann et al., 2002). In addition, cystatins may have other functions, non-related to their cysteine-protease inhibitory activity. Thus, carp ovarian cystatin was found to be important for preventing polyspermy in carp eggs, acting via electrostatic interaction as an agglutinating factor of spermatozoa (Wang & Huang, 2002).

Certain pathological conditions are related to congenital cystatin mutations. A hereditary form of human epilepsy, progressive myoclonus epilepsy (EPM1) or Unverricht-Lundborg disease, was found to be associated with either of several mutations in the gene for cystatin B (Pennacchio et al., 1996; Lalioti et al., 1997a, b; Alakurtti et al., 2000). In addition, knock-out of the cystatin B gene in mice resulted in development of symptoms similar to those in the human disorder and in a broad neuronal atrophy due to apoptosis (Pennacchio et al., 1998; Shannon et al., 2002). A recent study provided evidence that EPM1 pathogenesis is linked to an unbalanced cysteine protease activity due to the lack of cystatin B inhibitory activity (Kinne et al., 2002). Significantly increased levels of cathepsins B, L and S activities and highly reduced levels of cystatin B mRNA were thus detected in lymphoblastoid cells of patients suffering from EPM1. In addition, a multiprotein complex of cystatin B with several non-protease proteins has been implicated to have a specific cerebellar function (Di Giamo et al., 2002). Loss of this function as a consequence of the absence of cystatin B was proposed also to contribute to EPM1 development. Another congenital defect is responsible for hereditary cystatin C amyloid angiopathy (HCCAA). This human disease is caused by a point mutation, Leu68 to Gln, in cystatin C, the inhibitor variant being deposited in amyloid fibrils in the cerebral

arteries. Such amyloid formation leads to brain hemorrhage in young adults, resulting in paralysis, dementia and eventual death (Abrahamson et al., 1992; Olafsson et al., 1996). The L68Q mutation was found to decrease the stability of the cystatin C molecule, this instability being a critical factor for the development of the amyloidosis in HCCAA patients (Gerhartz et al., 1998; Gerhartz & Abrahamson, 2002). Cystatin C has also been shown to be colocalized with amyloid deposits in brain arterioles in patients with Alzheimer's disease or with hereditary cerebral hemorrhage associated with amyloidosis-Dutch type (van Duinen et al., 1987; Vinters et al., 1990). Notably, however, in the neurons most susceptible to cell death in Alzheimer'sdisease, the expression of cystatin C was elevated and cystatin C distribution was limited and colocalized with that of cathepsin B (Deng et al., 2001).

Cystatins and amyloid fibrillation

Progressive mycolonus epilepsy of type 1 (EPMI) Human stefin B a cysteine proteinase inhibitor is an intracellular protein expressed in many types of cells, located in the cytoplasm and nucleus. It has been found as part of a multi protein complex specific to the central nervous system (Giaimo et al. 2002). The main pathology for this proteinase inhibitor is its role in monogenic epilepsy, a progressive mycolonus epilepsy of type 1 (EPM1), termed as unverricht Lund borg disease. There are results that provide evidence that mutations in the gene encoding cystatin B are responsible for the primary defect in patients with a progressive mycolonus epilepsy of type 1 (EPM1) (Suoranta et al., 2012).

Hereditary cystatin amyloid angiopathy (HCCAA) is a rare, fatal amyloid disease in young people in Iceland caused by a mutation in cystatin C, which is an inhibitor of several cysteine proteinases such as cathepsins S, B and K. Mutated cystatin C forms amyloid, predominantly in brain arteries and arterioles, but also to a lesser degree in tissues outside the central nervous system such as skin, lymph nodes, testis, spleen, submandibular salivary glands and adrenal cortex. The amyloid deposition in the vessel walls causes thickening of the walls leading to occlusion or rupture in brain haemorrhage (Palsdottir et al., 2006).

Cystatins and cancer

Cathepsin B secretion may be important in penetration of the extra cellular matrix metastasis and cystatin C may be involved in regulation this process (Corticchiato et al. 1992; Dickinson 2002). The up regulation of cystatin F and down regulation of cystatin M associated with cancer is also indicative of a functional association. A correlation between high serum levels of cystatin C and higher risk of death in colorectal cancer patients has been found (Kos et al., 2000).

Cystatins from human sarcoma

Cysteine proteinases such as cathepsins B and L have been found to be elevated in cancer. Elevated activities could be due to impaired regulation by the endogenous low molecular mass cysteine proteinase inhibitors (cystatins). It was found that the extract of cystatins from sarcoma was less effective against papain and cathepsin B that was extract from liver. Inhibitory property of two members of the cystatin superfamily (stefin A and stefin B) was determined. It was found that stefin B from liver and sarcoma exhibited comparable inhibition of papain and cathepsin B. In contrast, stefin A from sarcoma exhibited a reduced ability to inhibit papain, human liver cathepsins B, H, L, human and murine tumor cathepsin. The Ki for inhibition of liver cathepsin B by sarcoma stefin A was 10-fold higher than that for inhibition of liver cathepsin B by liver stefin A, reflecting a reduction in the rate constant for association and an increase in the rate constant for dissociation (Lah et al., 1989).

Immuno modulation

Cystatins have emerged as effector molecules of immunomodulation (Zavasnik-Bergant 2008). They can stimulate nitric oxide release from macrophages (Verdot et al. 1999); modulate respiratory burst and phagocytosis in neutrophils (Leung-Tack et al. 1990); and modulate interleukin, cytokine production in T-cells and fibroblasts (Schierack et al., 2003; Kato et al., 2004). Most of these functions operate via putative cell surface cystatin-binding molecules or membrane domains (Kato et al., 2002). Cystatin C has been shown to be a T-cell growth factor TGFβ receptor antagonist and

TGFβ signalling pathway blocker (Sokol and Schiemann 2004; Sokol et al., 2005). Type 2 cystatins are also known to increase interleukin-6 (IL-6) expression in fibroblasts and splenocytes (Kato et al., 2000). Cystatin C is a potent, reversible inhibitor *in vitro* of the human lysosomal CPs e.g., cathepsin S (Ki = 8 pM), cathepsin L (Ki = 8 pM) and cathepsin H (Ki = 220 pM). These proteases are located all along the endocytic pathway of dendritic cell and are involved in the controlled proteolysis associated with the degradation of antigenic peptide (Pluger et al. 2002). Cystatin F by targeting cathepsin C is known to regulate diverse immune cell effector functions (Hamilton et al., 2008).

Cystatins and injury

Cystatins C appears to be up-regulated in response to injury in the brain. Cystatin C protein was detected by immunohistochemistry in few of the hippocampal pyramidal cells of the normal rat brain but was present in these cells 3 days after experimental ischemia (Palm et al., 1995). It was localised to morphologically degenerative neurons and absent from morphologically degenerative neurons and absent from morphologically viable neurons. In all brains from Alzheimers disease affected individuals but not in majority of normal brains, strong localization of cystatin C protein was found in the pyramidal neurons regions of the brain which was found to be most susceptible to cell death in this disease (Deng et al., 2001). The exact function of cystatin C in the brain and its role in injury are presently unknown. Its appearance would be consistent with a protective role, perhaps by blocking cysteine proteinase activity in damaged cells to allow for recovery or by acting as a growth factor. However, based on its association with damaged cells, it is conceivable that it may be a mediator of injury.

Cystatin C: A new serum marker for glomerular filtration rate

Cystatin C is found in all body fluids and tissues of mammals. It is a jack-of- all trades, playing a broad role in many functions. A large number of normal and pathological processes are controlled by the balance between proteniases and their inhibitors. The physiological role of cystatins is believed to be the protection of cells

forming appropriate endogenous or exogenous proteolysis by the regulation of cysteine proteinases of both host and of microbial origin. Levels of cystatins C in various body fluids is used as a barometer of diseases. Cystatin C is a non-glycosylated, low molecular mass (13 KDa) protein produced by all nucleated cells. Its low molecular mass and its high isoelectric point (pI) allow it to be freely filtered by the glomerular membrane. The serum concentration of cystatin C has been shown to correlate with Glomerular Filtration Rate (GFR) in combination with a stable production rate, it has been found to be a sensitive marker of GFR (Jung and Jung 1995). Serum levels of cystatins C are independent of weight, muscle, mass, age as compared to creatinine. Furthermore, measurements can be made and interpreted from a single random sample. Chronic kidney diseases are a risk factor for heart failure (HF). Although cystatin (C) can detect early kidney- dysfunction, little is known about the association between cystatin (C) and HF. Djousse et al. (2008) demonstrated that higher levels of cystatin C are associated with an increased HF and that such association may be limited to hypersensitive individuals.

Cystatin C in heart

The role of cystatin C in the heart under physiological and pathological conditions remains to be established. The structure of the rat cystatins C cDNA suggests that cystatin C is a secretory protein. Whether the inhibitor is released by the muscle fibers under physiological conditions or acts intracellularly is not known. It is conceivable however, that under pathological conditions, particularly in ischemia, the inhibitor is released and plays a role in modulating the activities of the extracellular cathepsins derived from monocytes and/or inflammatory cells. Under such conditions, cystatin C or its fragment may also affect the chemotactic and phagocytic functions of granulocytes (Leung et al. 1990).

Antimicrobial and antiviral activities

Antiviral and antibacterial activities of cystatins that could be related to the inhibition of exogenous proteases have also been demonstrated. The growth of polio and corona virus in cultured human cells could be retarded by addition of chicken cystatin or cystatins C or D (Cimerman et al., 1996; Collins & Grubb, 1998). Moreover, peptide

inhibitors with sequences mimicking the inhibitory center of cystatin C suppressed streptococcal infections in mice and the replication of corona and herpes simplex virus in cell cultures (Kasprzykowski et al., 2000). In addition, growth of *Porphyromonasgingivalis* and *Staphylococus aureus* was inhibited by cystatins A and S (Takahashi et al., 1994; Naito et al., 1995), and the recently discovered cystatin 11 was similarly indicated to have antimicrobial activity against *Escherichia coli* (Hamil et al., 2002). Cystatins were also found to inhibit cysteine proteases from the parasites *Trypanosoma cruzi*, *Trypanasoma congolense* and *Trypanasomabrucei brucei* (Stoka et al., 1995; Troeberg et al., 1996; Chagas et al., 1997). Furthermore, phytocystatins have been indicated to be insecticides, inhibiting digestive papain-like cysteine proteinases required for insect feeding and plant viral infection (Aoki et al., 1995; Urwin et al., 1995; Koiwa et al., 2000).

The cystatin isolated from horse shoe crab hemocytes has antimicrobial activity against Gram-negative bacteria, with IC 50 values against Salmonella.

EVOLUTIONARY RELATIONSHIP AMONG CYSTATINS

Statistical analysis (Barker and Dayhoff, 1972) of the similarities between aligned sequences suggests that they have evolved from a single primitive sequence. Difference matrix of the sequence of the members of superfamily revealed that the type 1 forms one family and the type 2 forms a second family whereas kininogens also form a single family. However, members of cystatin superfamily are closely related to each other. Fig 5 demonstrates the relationship between all known inhibitory active human cystatin domains according to their sequence homology and physiochemical properties.

The structure of cystatins indicates that the members of the family involved atleast six duplications of genetic material from a single primordial sequence to produce the cystatins A, B, C, S and 3 segments of kininogens. Gene structure of kininogens (Kitamure et al., 1985) shows that each of three cystatin like segments of kininogens is coded for by three exons for which, the cystatin gene was the precursor. Some workers reported that only the first two of the three exons of the primitive family 2

Figure 5: Schematic illustration of evolutionary relationship of cystatins of Type I, II and III.

gene were represented in family 1 cystatin gene, the third segment is acquired during the evolution of family 2 cystatins (Keil-Dlouha, 1986). Analysis of amino acid sequence data led to the conclusion that the differences in the length of polypeptides are accounted for by insertion and deletion of the genetic material in small blocks (Salvesen et al., 1986). Since disulphide bonds are often acquired in the evolution of protein but rarely lost (Williams, 1982), it is assumed that the primitive precursor of the cystatin was not unlike a modern type 1 cystatin.

The evolutionary relationship among the various cysteine proteinase inhibitors remain largely obscure until Muller-Esterl et al (1985) suggested a common evolutionary origin of stefins, cystatins and kininogens. Their model has been constructed on the basis of sequence homology. Muller-Esterl (1985) proposed that the diversity of CPI has evolved from two ancestral building blocks A and B. The stefin progenitor represents the whole superfamily comprising a single A unit. Cystatin acquired a second element B, possibly by gene fusion thus forming AB unit. Gene triplication of the archetype inhibitor generated the kininogen heavy chain which contains three cystatin like copies $(AB)_3$. Hence, the superfamily of mammalian CPI is constituted by atleast three distinct families with stefin, cystatin and kininogens as their prototypes.

PRIMARY STRUCTURE OF CYSTATINS

Amino acid composition of cystatins

Most of the members of the cystatin superfamily are polypeptides of 98-126 amino acid residues with Mr values falling in the range 11,000-14,000 (Table 1.2). It can be seen that cystatin A and α are devoid of cysteine residues where as cystatin B and β contain one and two residues, respectively. Cystatin F is reported to have two more cysteine residues apart from the usual four cysteine residues. Although tryptophan is absent from type 1 cystatins it is reported to be present in stefin C. An interesting feature of the stefin B is the conserved QVVAG region in the stefins of mammalian origin. Val 54 is replaced with Leu 54. Grubb et al (1984a) identified the disulphide linked peptides from human cystatin C and chicken cystatin between

TABLE 1.2: AMINO ACID COMPOSITION OF CYSTATINS

Data are calculated from published amino acid sequences for type 1 and 2.

AA	TYPE I				TYPE II		
	Human cyst A	Human cyst B	Bovine stefin C	Cyst C	Cyst S	Chicken Cyst	Cyst F
Cys	-	1	-	4	4	4	6
Ser	2	6	5	9	5	12	6
Thr	7	7	5	7	5	5	9
Pro	5	5	4	8	5	5	5
Ala	5	8	7	10	7	7	3
Gly	8	3	4	8	5	5	4
An	5	5	5	5	4	4	7
Asp	6	4	7	7	8	6	5
Glu	9	6	8	5	12	9	6
Gln	6	7	7	7	8	7	6
His	1	5	5	5	2	1	4
Arg	1	2	3	8	8	8	7
Lys	12	9	9	7	4	7	10
Met	2	2	2	3	1	2	2
Ile	4	2	4	2	6	6	4
Leu	8	6	6	8	8	10	12
Val	9	9	11	10	7	9	11
Phe	2	7	7	5	6	3	6
Tyr	6	4	3	4	6	5	3
Trp	-	-	1	-	2	1	2
Total	**98**	**98**	**101**	**122**	**113**	**116**	**118**

Cys 71- Cys 81and Cys 95- Cys 115. Anastasi et al (1983) found that cystatin type 1 contains no free thiol group.

Cystatin E apart from other usual conserved sequences charachteristic of family 2 cystatins has a five residue insertion between amino acid 76 and 77 and a deletion of residue 91 (cystatin C numbering). Ni et al (1998) deciphered the presence of a second Trp residue, in addition to the conserved Trp-106 in cystatin F.

SECONDARY STRUCTURE OF CYSTATINS

The structure of chicken cystatin and human cystatin C has been studied extensively. Based on lines of evidences from analysis of amino acid sequence for the distribution of hydrophilic and hydrophobic residues, the predicted hydrophilic residues are considered to be on the surface where as hydrophobic regions are internal. Each sequence possesses a large N-terminal hydrophilic region (55-70) followed by another large hydrophilic region at C-terminus. The secondary structure of human cystatin C in solution is very similar to that reported for chicken cystatin (Dieckmann et al., 1993) and the crystalline form of chicken cystatin has been reported by Bode et al (1988). Both human cystatin C and chicken cystatin as well as family 1 cystatins A and B (Martin et al., 1994; 1995; Stubbs et al., 1990; Jaskólski 2001) consist of five stranded anti-parallel β-pleated sheets which is wrapped around a straight five turn α-helix (Fig 3b and fig 4). An appending segment of partial α-helical geometry is present in chicken cystatin (Saxena and Tayyab, 1997) which was not found in human cystatin C by X-ray diffraction (Bode et al., 1998). The conformation of this region seems to be closer to that found for chicken cystatin in solution which consists of two loops (Dieckmann et al., 1993; Pavlova et al., 2012). The N-terminal segment up to Val 10 seems to be flexible, similar to that of other cystatins (Dieckmann et al., 1993; Martin et al., 1995). Trp was found only in the second hairpin loop of cystatins (Bode et al., 1988).

The three dimensional structures of stefins A and B are known as the solution structure of stefin A (Martin et al., 1995) and crystal structure of stefin B in complex with papain (Stubbs et al., 1990).

Stefins have a well defined globular fold consisting of five antiparallel β strands wrapped around a central five turn α helix. There is considerable similarity between the structural features of stefins A and B, but there are also some important differences in the regions which are fundamental to proteinase binding. The difference consist primarily of two regions of high conformational heterogeneity in free stefin A which correspond in stefin B to two of the components of the tripartite wedge that docks in to the active site of target proteinases. These regions which are mobile in solutions are the five N-terminal residues and the second binding loop. In the bound conformation of stefin B they form a turn and a short helix, respectively.

Circular Dichroism and computer prediction of secondary structure from the sequence indicate that the chicken cystatin has about 20% α-helix, 42% β-structure and 24% β-turn and 12% random coil (Schwabe et al., 1984). Ekiel and Abrahamson (1996) and Janowski et al (2001) proposed that human cystatin C undergoes dimerisation under mild denaturating conditions. NMR studies revealed that the majority of the proteins are not affected by dimerisation, however, the CP binding site is significantly affected (Ekiel et al., 1997) which accounts for the activity loss (Abrahamson, 2001). Hence, it is concluded that hydrophobic CP binding site is involved in the monomer-monomer contact. The dimerization may be relevant to regulation of cystatin C inhibitory activity in vivo (Abrahamson, 2001). Three dimensional swapping studies indicate the dimerization of human cystatin C to be a step towards protein oligomoerization (Janowski et al., 2001).

INHIBITION OF PROTEINASES

Specificity

Cystatins are highly specific for cysteine proteases except for thyropins which show inhibitory activity against aspartic and metalloproteases (Mihelic and turk, 2007; Lenarcic and Turk, 1999). However there are few cystatins capable of inhibiting human legumain (Alveraz-Fernandz et al., 1999) and calpains (Schmaier et al., 1986).

Stefin A and B are potent inhibitors of papain, cathepsins L, S and H but have decreased sctivity against cathepsin b (Musil et al., 1991). Of the type II, Chicken

TABLE-1.3a: DISSOCIATION CONSTANT (Ki) FOR HUMAN CYSTATINS AND CHICKEN CYSTATIN WITH LYSOSOMAL CYSTEINE PROTEINASES.

Cystatin	Papain	Cathepsin B	Cathepsin H	Cathepsin L	Cruzipain	Dipeptidyl peptidase I
Stefin A	0.019	8.2	0.31	1.3	0.0072	33
Stefin B	0.12	73	0.58	0.23	0.060	0.23
Cystatin C	0.00001	0.27	0.28	<0.005	0.014	3.5
Cystatin E/M	0.39	32	n.d.	n.d.	n.d.	n.d.
Cystatin F	1.1	>1000	n.d.	n.d.	n.d.	n.d.
Cystatin S	108	n.d.	n.d.	n.d.	n.d.	n.d.
Cystatin SA	0.32	n.d.	n.d.	n.d.	n.d.	n.d.
Cystatin SN	0.016	19	n.d.	n.d.	n.d.	n.d.
Chicken cystatin	0.005	1.7	0.06	0.019	0.001	n.d.
L-Kininogen	0.015	600	0.72	0.017	0.041	n.d.

- ki values for human cystatins (Abrahamson et al., 2003).
- for chicken cystatin (Barret et al., 1986).
- Cruzipain inhibition by cystatin (Stoka et al., 1995).
- values are given in nanomolar ratio.
- n.d- not determined.

TABLE-1.3 b: RATE CONSTANTS FOR THE INTERACTION OF CYSTATINS WITH THIOL PROTEINASES

Enzyme	Inhibitor	K_{ass}	K_{diss}	Reference
Papain	Chicken cystatin	1×10^{7}	1×10^{7}	Nicklin & Barrett, 1984
	Bovine cystatin B	6.3×10^{6}	7.7×10^{-4}	Nicklin & Barrett, 1984
	Bovine cystatin C	1.8×10^{6}	-	Nicklin & Barrett, 1984
Cath. B	Chicken cystatin	1.4×10^{6}	2.3×10^{-3}	Nicklin & Barrett, 1984
Ficin	Chicken cystatin	4.3×10^{6}	n.d.	Sen & Whitaker, 1973
Dipeptidyl Peptidase I	Chicken cystatin	1.0×10^{7}	2.2×10^{-3}	Nicklin & Barrett, 1984

cystatin has been shown to inhibit papain, ficin, chymopapain, papaya proteinase III, actinidin, cathepsin B, H, L and N (Barrett, 1985; Fossum and Whitaker, 1968). Cystatin S is found to inhibit papain and ficin (Isemura et al., 984a). Cystatin E/M inhibits papain, cathepsin B, L, V and legumain (Cheng et al., 2006); Alvarez-fernandez et al., 1999; Ni et al., 1997; Sotiropoulou et al., 1997). Cystatin F inhibits cathepsin F, K, V, S, L and H (Langerholc et al., 2005) and weekly legumain (Alvarez-Fernandez et al., 1999). Cystatin D inhibits cathepsin S, H and L but not cathepsin B or pig legumin (Alvarez-Fernandez et al., 2005).

Kinetic Behaviour

The cystatins are the first group of protein inhibitors of CP for which the mechanism of inhibition has been investigated. They are a group of potent, non covalent, tight binding and reversible inhibitors of CP's which inhibit the target enzymes in micromolar to picomolar range (Turk et al., 1997). Some of the reported value of Ki are presented in Table- 1.3.a. the affinity differences can be explained by the differences in the active site regions of endo- and exopeptidases. Some estimates of individual association and dissociation rate constants are available in Table 1.3b.

MECHANISM OF INTERACTION OF CYSTATINS WITH PROTEINASES

It has been established that no disulphide bond is formed between the active site cysteine residue and the inhibitor because complexes dissociate when denatured without reduction (Nicklin and Barrett, 1984; Gounaris et al., 1984) and carboxymethylation of enzyme and inhibitor does not prevent complex formation (Anastasi et al., 1983). It became clear that complex formation is reversible, competitive and non-covalent with substrate. The nature of interaction resembles the standard mechanism of inhibition described by Laskowski and kato (1980). The active site of papain in complex with chicken cystatin is unreactive with 5,5'-dithio-bis-(2-nitrobenzoic acid) at pH 8.0, 2,2'-dipyridyl disulphite at pH 4.0 and iodo[^{14}C]

Figure 6: Scheme of proposed model for the interaction of chicken cystatin with papain

Acetate (Nicklin and Barrett, 1984). The complex formation is accompanied by pronounced spectroscopic changes, most likely reflecting local perturbation of environment of aromatic residues in both enzyme and inhibitor (Ylindahl et al., 1988; Bjork et al., 1989).

The mechanism of inhibition of papain by chicken egg white cystatin has been extensively studied. Bode et al (1988) proposed a model for the interaction of cystatin with CP based on X-ray crystal structure of chicken cystatin and subsequent studies based on docking experiment. Three parts of the inhibitor are in close contact with the active site cleft of papain, the amino terminus with conserved Gly-9, a first hairpin loop (53-57) containing the prototype sequence, QVVAG (Barrett et al., 1986; Barrett, 1987) and a second hairpin loop (residues 102-107) containing conserved trp 104. Bode et al (1988) demonstrated that the major contribution is from the first hairpin loop containing the QVVAG sequence (Turk, 1985; Okhubo et al., 1984). According to the model the N-terminal segment of cystatin which is more flexible bridges over the active site Cys 25 residue of papain with out completely burying it and additionally the side chain of Leu 8 bind to S_2 subsite of papain which determines the substrate specifity of papain (Asboth et al., 1988). The side chain groups of QLVSG segment of chicken cystatin make the most intimate contacts with putative S_1 subsite of papain (Fig 6). This explains the presence of Gly 9 in close vicinity of but not in direct contact with Cys 25 of papain. Removal of Leu 7 and Leu 8 leads to an approximately 5000 fold lower affinity for papain.

This was supported by Brzin et al (1984) who demonstrated that the truncated form of human cystatin C starting with Leu-Val before Gly-11 (corresponding to Gly-9 of chicken cystatin) has virtually the same affinity for papain as the full length form whereas the truncated form starting with Gly-12 has been reported to show 1000 fold weaker inhibiton (Abrahamson et al., 1987). However, Nycander and Bjork (1990) emphasized the role of Trp-104 in the inhibition of CP. According to their model, Trp-104 of cystatin interacts primarily with two tryptophan side chains in the active site cleft of papain, Trp-177 and trp 181, in such a manner that the indole ring of Trp-104 stacks on the side chain of Trp-177 and the edge lies on the indole ring of Trp-181.

A two step mechanism of inhibition of the lysosomal CP, cathepsin B by its endogenous inhibitor, cystatin C was observed by Nycander et al (1998). An initial weak interaction in which N-terminal of the inhibitor binds to the proteinase is followed by a conformational change. Susequently, the occluding loop of the proteinase that partially obscures the active site is displaced by the inhibitor bringing about another conformational change. The presence of occluding loop of cathepsin B renders it much less susceptible to inhibition by cystatin than other proteinases. A similar two step binding of cystatin A to the CP was suggested by Estarda and Bjork (2000). The flexible N-terminal region of the cystatin binds independently to the target proteinases after the hairpin loops.

THE LIVER

The liver is the central organ for the regulation of glucose homeostasis, xenobiotic metabolism and detoxification, and steroid hormone biosynthesis and degradation. This organ also has a major impact on health and homeostasis through its control of serum protein composition. While differentiated hepatic functions are generally well maintained with age, changes do occur. Worn out erythrocytes breaks down to bile pigments bilirubin and biliverdin. In the embryo, red blood cells are manufactured by the liver. In adults, liver stores inorganic salts of iron, copper and vitamin B_{12} (anti-anaemic factor) and thus helps in the formation of red cells and haemoglobin. The plasma proteins, serum albumin and serum globulin are synthesized by liver from the amino acids derived from the protein in the diet. Kupffer's cells in the liver sinusoids phagoctose and remove bacteria, worn out blood elements and foreign particles. Serum and biliary cholesterol rise, liver regeneration declines, hepatic drug clearance decreases, and liver volume and blood flow decrease.

LIVER DISEASES

Initially proteinases were considered primarily to be protein-degrading enzymes. However, this view has dramatically changed and proteinases are now seen as extremely important signaling molecules that are involved in numerous vital

processes. Proteinase signaling pathways are strictly regulated by endogenous inhibitors, and the dysregulation of proteinase activity can lead to pathologies such as cardiovascular and inflammatory diseases, cancer, osteoporosis and neurological disorders.

Liver disease is a condition that causes liver inflammation or tissue damage and affects liver function. The liver is a vital organ located in the upper right-hand side of the abdomen. It is a large organ , weighs 2-3 pounds, and performs numerous functions for the body: converting nutrients derived from food into essential blood components, storing vitamins and minerals, regulating blood clotting, producing proteins and enzymes, maintaining hormone balances, and metabolizing and detoxifying substances that would otherwise be harmful to the body (fig 7a, & 7b). The liver makes factors that help the human immune system fight infection, removes bacteria from the blood, and makes bile, which is essential for digestion.

Any change in anatomy or function of liver is characterised as liver disease (Kamble et al. 2009). The most common of these include: infections such as hepatitis A, B, C, E; alcohol damage, fatty liver, cirrhosis, cancer and drug damage. Many diseases of the liver are accompanied by jaundice caused by increased levels of bilirubin in the system. The bilirubin results from the breakup of the haemoglobin of dead red blood cells; normally, the liver removes bilirubin from the blood and excretes it through bile. Hepatitis and jaundice is a major public health problem worldwide, responsible for considerable morbidity and mortality from chronic liver diseases (Lau and Membreno 2004).

In view of the fact that cystatin engineering presents a very promising field and lacuna in its knowledge exist therefore it was thought worthwhile to purify the inhibitor from liver, the body's **"chemical factory"** to have an in depth knowledge about the physio-chemical characteristics of cystatin present in this organ, also to compare its properties with other species.

TABLE 1.4: DIFFERENT LIVER DISEASES.

Type of Liver Disease	Description	Examples of Causes/Conditions
Acute liver failure	Rapid decrease in liver function	Drugs, toxins, a variety of liver diseases
Autoimmune-associated	The body produces an inappropriate immune response against itself; sometimes develops antibodies against own liver tissue	PBC (Primary biliary cirrhosis), PSC (Primary sclerosing cholangitis), Autoimmune hepatitis
Budd-Chiari syndrome	Blood clots impede blood flow from the liver; symptoms such as ascites, enlarged liver, jaundice, and abdominal pain can develop.	Hypercoagulable disorders, liver injury, cancer, parasitic infection
Cirrhosis	Scarring of liver tissue leads to decreased liver function	Can be caused by a variety of conditions but usually a result of chronic hepatitis, alcoholism, or chronic bile duct obstruction
Genetic	Gene mutations can lead to liver damage, disease; relatively rare conditions	Hemochromatosis, Alpha-1 antitrypsin deficiency,Wilson's disease
Hepatitis	Acute or chronic liver inflammation	Viruses, alcohol abuse, drugs, toxins, autoimmune,nonalcoholic fatty liver disease (NAFLD)
Infections	Certain infections can cause various degrees of liver damage, blockage of bile ducts	Viral hepatitis, Parasitic infection
Liver cancer	A cancer that originates in the liver	Increased risk with cirrhosis and chronic hepatitis; hepatocellular carcinoma (HCC) is most common primary liver tumor
Obstruction of bile ducts	Complete or partial blockage of bile ducts	Tumors, gallstones, inflammation, trauma

OBJECTIVES OF THE THESIS

Cystatins and their target enzymes play a role in many pathological events, including inflammatory disease. In the liver, an imbalance between cystatins and their targets can disregulate matrix degradation and accumulation, leading to hepatic fibrosis. However certain chemicals can induce different liver disease which can enhance the levels of cathepsins or diminish the levels of cystatins inhibiting them. Therefore, in view of the reported species an organ dependent differences in the physiochemical properties of cystatins it was thought desirable to isolate and purify cystatin from an important hitherto uninvestigated source i.e. Buffalo liver and subsequently to study in detail its structural, molecular and functional properties. In addition to this, it was interesting to explore the hazardous effects of pesticides (Malathion) and antituberclusosis drug (rifampin) and cardiovascular drug (atorvastatin) to study the effect of these compounds on cystatin with specific focus on drug-protein interaction which greatly effects the function of several proteins, on the isolated liver cysteine proteinase inhibitor has been purified from Buffalo liver. Elevated bilirubin levels in hepatitis and gilberts syndrome can cause certain physiological changes in the hepatocytes and hence diminish the activity of cystatins. The effect of glycation involving reducing sugars was also studied given the susceptibility of proteins in case of elevated sugar level in diabetic condition. It is also interesting to see the folding and unfolding of cystatin at acidic conditions and external stress leading to structural change or aggregation in the protein. Misfolding has been reported to be the cause of various amyloidosis in mammals, especially in humans leading to many diseases.

Figure 7: Morphological structure of liver (a) and some of the diseases associated with liver (b).

INTERACTION OF BILIRUBIN WITH BLC

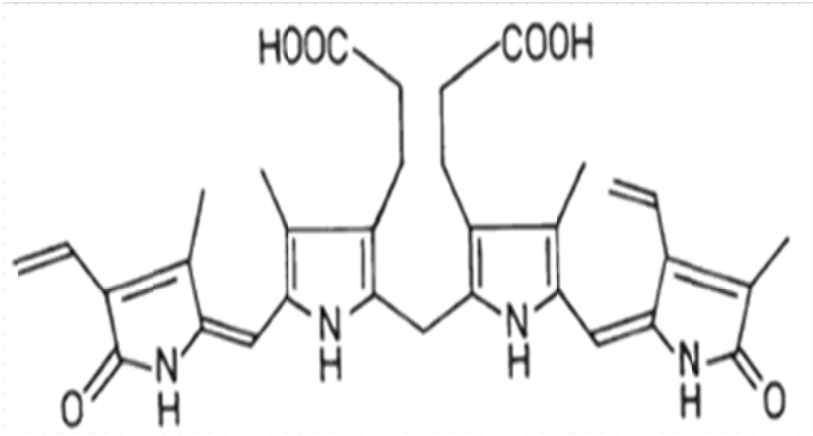
SPECTROSCOPIC STUDIES

Bilirubin, a metabolite of the heme in senescent red blood cells, is normally conjugated with albumin to form a water-soluble complex (Chandy T and Sharma CP 1992). Free bilirubin is toxic, and hence is transported to the liver as a complex with albumin where it is normally conjugated and excreted into the bile. Under physiological conditions, bilirubin remains bound to serum albumin which carries it to the liver for further metabolism. However, when bilirubin-to-albumin molar ratio (B/A) exceeds 1 : 1, free bilirubin binds to many types of cells including brain cells which is the cause of brain toxicity in premature neonates, a condition called kernicterus or bilirubin encephalopathy. In addition, jaundiced neonates with low plasma pH have been reported to be at greater risk of developing bilirubin encephalopathy (Kim et al., 1980).

Liver damage and related malfunctions can result in hyperbilirubinemia, i.e an elevated level of unconjugated (or free) bilirubin in the blood. Excess free bilirubin tends to deposit in tissues, especially in the brain. Disorders in the metabolism of bilirubin, especially common among newborn infants, may cause jaundice, a yellow discoloration of the skin and other tissues (Shi et al., 2005; Khan et al., 2001). Accumulation of unconjugated bilirubin (UCB) in the centralnervous system contributes to cell damage during severe neonatal hyperbilirubinemia and Crigler-Najjar type I syndrome leading to bilirubin encephalopathy (Brodersen R et al., 1990; Rubboli G et al., 1997; Hafkamp et al., 2006; Cuperus et al., 2013).

Bilirubin is a linear tetrapyrrole whose conformation is affected by internal hydrogen bonds formed between the carboxyl side chains and dipyrromethenone rings. Interaction of bilirubin with cells or cell membranes is well documented (Sato et al., 1987; Vazquez et al., 1988; Hayer et al., 1989; Tayyab and Ali 1995, 1997) and it is commonly accepted that the toxicity of bilirubin depends on its passage across the plasma membrane and its association with membrane lipids (Ali and Zakim 1993; Zucker et al., 1994). However, the way in which bilirubin interacts with biological membranes is not fully understood. It has been reported that the interaction of

bilirubin with the membranes is greatly influenced by the physico-chemical properties of the interacting media such as pH and temperature (Sato and Kashiwamata 1983).



Bilirubin

Chemical structure of bilirubin

Increased binding of bilirubin to biological membranes at physiological pH (7·0 – 7·2) has been suggested to be due to increased precipitation of bilirubin on the surface of membranes (Cestaro et al., 1983).

GLYCATION OF PURIFIED BLC WITH REDUCING SUGARS

The reducing free carbonyl group of fructose, glucose and ribose, in common with all reducing sugars, may react non-enzymatically with amino groups of proteins (Figure 1). Reducing sugars (e.g, glucose, fructose, galactose, mannose, ribose) and certain other carbohydrate relatives (e.g., ascorbic acid) are inherently reactive towardnucleophiles. Glycation takes place at €-amino groups of lysine or hydroxylysine residues as well as at ***α-*** amino groups of amino-terminal residues (Thorpe et al., 1982).Their reactivity toward nucleophilic nitrogen bases in protein underlies the Maillard reaction (Bunn et al., 1978).Reducing sugars including glucose, fructose, ribose and trioses can react non-enzymatically with the amino groups of

proteins to form reversible Schiff's bases and then Amadori products. These early glycation products undergo further complex reactions such as rearrangement, dehydration, and condensation to become irreversibly cross-linked heterogeneous fluorescent derivatives termed advanced glycation end products (AGE)1 (Brownlee, M et al., 1988; Makita, Z et al., 1992). Formation of Amadori product from the Schiff base is slower but much faster than the reverse reaction, so that the Amadori glycation product tends to accumulate on proteins. The Amadori rearrangement of a lysine-glucose Schiff base is thought to be facilitated if there is a histidine side-chain or another lysine amino group on which the Schiff base has formed, due to localized acid-base catalysis (Acosta et al., 2000). Compromise of protein function is effected initially by glycation of the lysine side chain and N-terminal amino acid residues, followed by later stage modification involving not only these, but arginine and cysteine residuesand stabilization to AGEs via Schiff base and Amadori rearrangement (N. Ahmed et al., 2007; J. Zeng et al., 2005). Glycation is commonly recognized as the Maillard reaction (Finot, 2005). In this reaction, the carbonyl group of a sugar interacts with the nucleophilic amino group of the amino acid, producing N-substituted glycosylamine (Schiff base) which is labile and may undergo two sequential rearrangements, yielding a reasonably stable aminoketose — the Amadori product (Fayle & Gerrard, 2002). In a Schiff base, the aldehydic carbon-oxygen double bond of the sugar is converted to a carbon-nitrogen double bond with the amine. As is the case with free sugars, the open-chain, double bonded form of the Schiff's base adducts of hexoses or pentoses are thermodynamically disfavored with respect to the equilibrium cyclized pyranose or furanose forms, or glycosylamines (Paulsen and Pflughaupt, 1980). Formation of the Schiff base from sugar and amine is relatively fast and highly reversible process. Certain protein groups are particularly prone to glycation; they include terminal amino groups and lysine side chains (Frister et al., 1988). Arginine side chains can be glycated as well (Tagami et al., 2000).

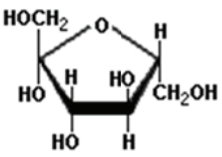
Glycation or nonenzymatic glycosylation of protein occurs as the result of a reaction between reducing sugars and the primary amino groups on protein. Nonenzymatic protein glycation (glucosylation or glycosylation) by glucose is a complex cascade of reactions yielding a heterogeneous class of compounds, collectively termed advanced glycation end products (AGEs).

Reducing sugars (e.g, glucose, fructose, galactose, mannose, ribose) and certain other carbohydrate relatives (e.g., ascorbic acid) are inherently reactive towardnucleophiles. Their reactivity toward nucleophilic nitrogen bases in protein underlies the Maillard reaction. Louis Camille Maillard first described the non-enzymatic glycation reaction in the early 1900s and it came to be known as the *Maillardreaction*. The Maillard reaction is a process in which reducing sugars react spontaneously with amino groups in proteins, lipids and nucleic acids to advanced glycation endproducts (AGEs). However, it was not until 1980 that the pathophysiological significance of the Maillard reaction emerged in medical science, in particular in relation to diabetic complications and ageing (Chuyen NV et al., 1978; Monnier VM et al., 1981).

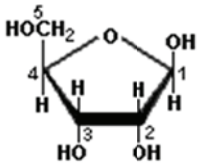
Non-enzymatic glycation of proteins by reducing saccharides such as glucose and ribose lead to the formation of fructosamine (Day JF et al., 1979) and advanced glycation end products (AGEs) (Degenhardt TP et al., 1998). Among the reducing monosaccharides, the role of glucose in the glycation of proteins has been widely studied, and is implicated in diabetes (McCance DR et al., 1993), cataracts (Lyons TJ et al., 1991), renal failure (Miyata T et al., 1999), and other disorders (Westwood ME et al., 1995). It has recently become clear that glycation is also involved in physiological neurodegenerative diseases such as Alzheimer's disease (Dukic-Stefanovic S et al., 2001). Glycation alters the biological activity of proteins and their degradation processes. Protein crosslinking by glycation results in the formation of detergent insoluble and protease-resistant aggregates. Therefore, the study of AGEs has become one of the most important areas of biomedical research today.

Advanced glycation end products (AGEs) are a heterogeneous group of proteins that have been modified with glucose or carbohydrate adducts and arise from the nonenzymatic addition of reducing sugars to the side chains of lysine and/or arginine residues in proteins (Brownlee, M et al., 1988). The free amino groups of the proteins react slowly with the reducing sugars, including glucose, via glycation or the Maillard reaction (Nwabuisi, C. 2002; Huh, J. W et al., 2007). This process is initiated by condensation reactions between the reducing sugars and free amino groups to for form Schiff bases, which undergo rearrangement to form relatively stable Amadori products (Reynolds, T. M. 1965). Although an elevated level of glucose had been thought to play a primary role in the Maillard reaction, the formation of AGEs is now known to result also from the action of various metabolites other than glucose. The

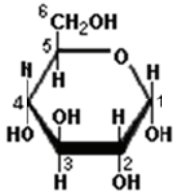
realisation that fructose and its metabolites, at elevated concentrations, can initiate potentially deleterious changes such as the non-enzymatic fructosylation of proteins



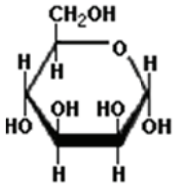
Fructose



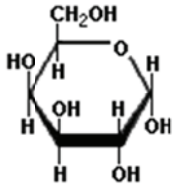
Ribose



Glucose



Mannose



Galactose

Chemical structure of some reducing sugars like Fructose, Ribose, Glucose Galactose and mannose.

has raised interest and concern about this property of fructose (Dills WL Jr 1993). With respect to human health, the contribution of fructose as an effective glycating agent is of importance for three reasons. First, on a molecular basis fructose and fructose metabolites, which are primarily located intracellularly, may participate in the non-enzymatic glycation reactions at a much faster rate than glucose. Second, fructose consumption has increased steadily during the last two decades. An increased fructose intake from the diet might result in changes in tissue concentration of fructose and its metabolites, which can in turn potentiate the Maillard reaction. Third, in several organs, fructose is formed from glucose by the polyol pathway.

Glucose is, in fact, the least reactive of the common sugars, perhaps leading to its evolutionary selection as the principal free sugar in vivo (Bunn et al., 1978). Nevertheless, glucose can react with a free amino group, such as the €-amino group of a protein lysine residue, to form an adduct commonly referred to as a Schiff base. As with glucose-glycation, glycation of proteins with ribose also involves Amadori rearrangement and then production of AGEs. In addition to glucose, sugars such as

galactose (Urbanowski, J. C et al., 1982), sialic acid (McKinney, R. A et al., 1982), mannose (Zaman, Z et al., 1981), glucose 6-phosphate (Haney, D. N et al., 1976), glyceraldehydes (Acharya, A. S et al., 1983), and fucose (Zaman, Z et al., 1981) have been used *in vitro* as glycating agents.Glycation of proteins with ribose also involves Amadori rearrangement and then production of AGEs. Therefore, some research groups have employed glycation with rib instead of glc as a model for investigating the mechanism by which AGE yield hydroxyl radicals and induce cell apoptosis or necrosis (Luciano Viviani G et al., 2008; Chetyrkin SV et al., 2008; Culbertson SM et al., 2003).

Glycation by fructose (fructation)' has not yet been a special focus of research although it has been mentioned in comparative studies including various sugars (Monnier, V. M et al., 1980; Bunn, H. F et al., 1981). The rate of glycation is directly proportional to the percentage of sugar in the open-chain form (Bunn HF et al., 1981) and the rate for fructose is 7.5 fold faster than that of glucose (0.002% open-chain) and, most strikingly, the glycolytic intermediate glyceraldehyde-3-phosphate (100% open-chain) forms over 200-fold more glycated protein than do equimolar amounts of glucose (Ruderman NB et al., 1992). Fructose as a glycating agent should deserve more attention since it accumulates in organs where the sorbitol pathway is active. This metabolic shunt mediates the conversion of glucose to fructose with the formation of sorbitol as an intermediate (Gabbay, K. H. 1975; Kinoshita, J. H. Set al., 1979).

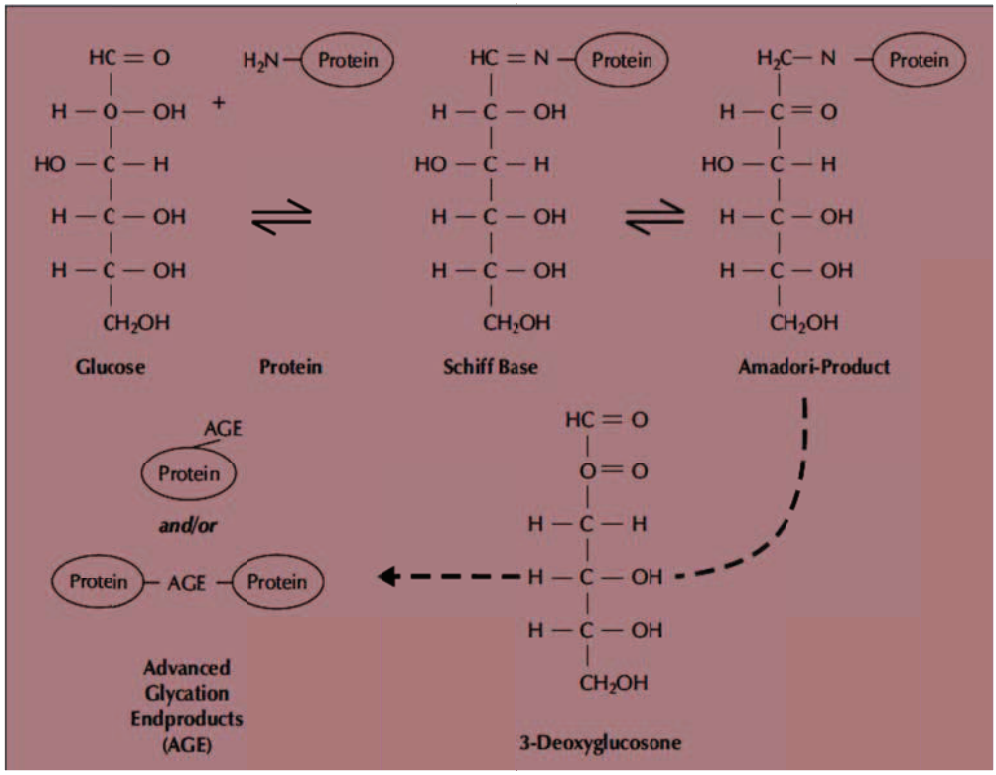
Compromise of protein function is effected initially by glycation of the lysine side chain and N-terminal amino acid residues, followed by later stage modification involving not only these, but arginine and cysteine residues and stabilization to AGEs via Schiff base and Amadori rearrangement (N. Ahmed et al., 2007; J. Zeng et al., 2005). Nonenzymatic protein glycation (glucosylation or glycosylation) by glucose is a complex cascade of reactions yielding a heterogeneous class of compounds, collectively termed advanced glycation end products (AGEs). Nonenzymatic glycation of proteins described by Louis-Camille Maillard (Maillard. LC 1912) have been implicated in the pathogenesis of diabetes, renal failure, and aging (Schleicher ED et al., 1997). In vitro–prepared AGE modifications of proteins have been shown to be toxic, immunogenic, and capable of triggering cellular injury.

The Maillard reaction begins with the reaction of the carbonyl (aldehyde or ketone) of the reducing sugar to form a reversible Schiff base with the amino group of the biomolecule. The Schiff base can undergo an intramolecular rearrangement to form the Amadori products (Ahmed N et al., 2003); this can undergo a series of further rearrangements, dehydration, and condensation to form irreversible end products, which may be fluorescent and yellow–brown in color; some can form stable intermolecular and intramolecular cross-links (Ulrich P and Cerami A, 2004).

Further studies have shown that glycation reactions lead to the production of reactive oxygen species (ROS), which are harmful to cellular metabolism and cause cell damage (Kikuchi S et al., 2003). Consequently, considerable attention has been given in recent years to studying the generation of hydroxyl radicals from Amadori proteins, and inhibition of hydroxyl radical damage (Chetyrkin SV et al., 2008). Most recently, formation of molten globule-like states has been reported during the progression of glycation reactions *in vitro* (Sattarahmady N et al., 2007). According to Dobson et al., 2002 globule-like protein aggregations (pro-amyloid fibrils) are significantly toxic to neurons (Stefani Met al., 2003; Sanghera Net al., 2008); Tabner BJet al., 2005); Chromy BA, 2003). Glycation of serum albumin has been widely studied in recent years [Day JF et al., 1979; Luciano Viviani G et al., 2008; Coussons PJet al., 1997; Mendez DLet al., 2005), and bovine serum albumin (BSA) is commonly used as molecular model. Significantly, Friedman and colleagues have reported neurotoxicity for brain-penetrant serum albumin including BSA (Seiffert E et al., 2004; Ivens S et al., 2007). This study is concerned with determining whether glycation of BSA with D-rib in the short term (1– 7 days) results in globular amyloid-like aggregations that are seriously toxic to neuroblast cells through apoptosis, compared with other pyranose and furanose.

Nonenzymatic glycation has been described in proteins such as the lens crystallins (Stevens, V. J et al., 1987; Chiou, S.H et al., 1981**),** the red cell membrane (Miller, J. A et al., 1980), the peripheral nerve proteins (Vlassara, H., et al., 1981) serum albumin (Dolhofer, R et al., 1979), collagen (Le Pape, A et al., 1981; Perejda, A.J et al., 1982), ferritin (Zaman, Z et al., 1981), apolipoprotein (Witztum, J**.** L,et al., 1983) and others.. In conditions of increased glucose supply such as diabetes, the level of fructose in the lens may rise 23-fold becoming twice as concentrated as glucose (Gabbay, K. H et al., 1972). In most organs where the sorbitol pathway is active, the

concentratin of fructose is of the same order of magnitude as that of glucose, making *in vivo* glycation by fructose a highly probable event.



Formation of Advanced glycation end products

The exposure of proteins to high glucose levels, as observed in diabetes, and the further reaction of glucose with proteins are retained responsible for the long-term complications of this disease (Baynes, J. W 2004). Chronic hyperglycemia leads to micro- and macro- vascular diabetic complications through four intracellular mechanisms, including increased polyol pathway flow (Lee, A. Yet al., 1995; Chen, L et al., 2007), the activation of protein kinase C (Koya, D et al., 1998), increases in the hexosamine pathway (Kolm-Litty, V et al., 1998), and increased AGE formation (Sing, R et al., 2001). AGEs accumulate slowly in long-lived proteins such as collagen, resulting in the altered structure and properties of certain tissues such as vascular walls, the kidneys, etc. (Kohn, R. R et al., 1984; Wu, J. T. 1993). This process is thought to be an important pathogenic factor for diabetes complications (Beisswenger, P. Jet al., 1995; Wu, J. T. 1993; Korbet, S. M et al., 1993; Makita, Z et al., 1992; Makita, Z et al., 1991). AGE accumulation in tissue proteins has been implicated in the pathological development of aging, diabetes, atherosclerosis, as well

as in diabetes complications such as retinopathy, nephropathy, and neuropathy (Vlassara, H. 1994; Thorpe, S. R and Baynes, J. W. 1996). Advanced glycation end products (AGEs), the products of non-enzymatic glycosylation and oxidation of proteins and lipids, accumulate in diverse biological settings, such as diabetes, inflammation, renal failure, and aging (Ohgami et al., 2002).

UNFOLDING STUDIES OF BLC

I. DENATURATION OF BLC AT LOW pH AND TFE

Protein folding is a process by which amino acid sequence of a protein determines the three-dimensional conformation of the functional protein. The elucidation of the molecular mechanism of protein folding from a disordered polypeptide chain to specific native state, that is the deciphering of second half of the genetic code (Kuwajima, 1989) remains one of the challenges in biochemistry (Ptitsyn, 1992). Study of the folding intermediates and denatured states provides an insight into understanding how and when various forces come into play in directing protein folding (Christenseu, H et al., 1994; Goldberg, M. E et al., 1990; Schulman. B. A et al., 1995; Wilson, G et al., 1995).

Biological activity of the proteins is strictly related to the three dimensional conformation which is assumed to depend in principle, only on the sequence of amino acids, i.e. on the primary structure. Much of what is currently known about protein folding and stability has been observed from theoretical and experimental investigation in small, monomeric proteins and protein domains ranging from 50-150 amino acids in size. Unfolding and refolding studies using chaotropic agents have given a wealth of information on the structure and biogenesis of proteins (Tanford, 1968).

The transformation of a newly synthesized protein on ribosomes to a functionally active form, involves complex and intricate process of folding. The information for the folding of a disordered structure of protein lies in its amino acid sequence (Anfinsen, 1973). The identification and characterization of intermediates during the folding process is central to understand the process of protein folding. These studies

provide an insight to the mechanism of how and when different forces come into play in directing protein folding (Wulthrich, 1994; Ballery et al., 1993).

However it is now accepted that unfolding/refolding of a number of proteins involves different steps, that is under different conditions intermediate states have been observed that do not appear to be native or completely unfolded (Radford et al., 1992). When transitions were analyzed by different spectroscopic probes, intermediate states have been found for several proteins like bovine and human carbonic anhydrase β (Jagannadham and Balasubramanian, 1985), β-lactamase (Mitchinson and Pain, 1985), bovine growth hormone (Brems et al., 1985) and a chemically modified form of human growth hormone (Brems et al., 1990).

The stability of proteins in the solution is the major concern of biologists and pharmacologists. An in depth knowledge of protein denaturation and refolding is central in selection of a particular solvent for protein stabilization. The mechanism by which proteins fold from a denatured state to their biologically active form is an intricate process (Nolting et al., 1995; Plaxo and Dobson 1996).

Study of denatured states provide an insight to understand the mechanism of unfolding. The development of broad range of techniques has led to the characterization of stable intermediates in several proteins (Barrick et al., 1994). It is well known that proteins unfold under the action of denaturants like GdnHCl and urea. Analysis of solvent denaturation curves using these reagents can provide a measure of the conformational stability of a protein (Pace, 1986; Yao and Bolen, 1995).It has been established that electrostatic interactions (Schreiber and Fersht, 1995) hydrogen bonding (Alber at al., 1987; koh et al., 1997; Hebert et al., 1998), the hydrophobic effect (Schreiber and Fersht, 1995), folding nuclei (Mirny and Shakhnovich, 1999) and partially folded intermediates (Chamberlain and Marqusee, 2000) play important role in the process by which a single polypeptide chain acquires its secondary and tertiary structure.

It has been shown that the amount of tertiary structure and the compactness of intermediate state vary form one protein to another (Privalov 1996; Kuwajima 1989). These intermediate states have structures either close to the native state or unfolded state depending upon the protein being studied and experimental conditions used. Therefore besides presence of molten globule states from N ↔ D

(Native↔Denatured) folding/unfolding transition, there exist transition states between native and molten globule states or between molten globule and unfolded state (Redfield et al., 1994; Ptitsyn 1995; Bycroft et al., 1990; Serrano et al., 1992).

Different proteins behave very differently on acid denaturation (Fink, L. A et al., 1994). Some do not unfold at the lowest pH values, some undergo transition to a compact, molten globule state**,** and a third type first unfolds to an extended conformation and then undergoes transition to a compact state on addition of anions (Goto, Y et al., 1990; Stigter, D et al., 1991). Intermediates of the molten globule type have been detected in acid denaturation for a number of proteins (Kuroda, Y et al., 1992; Goto, Y. & Fink, A. L. 1990; Narhi Owers, L et al., 1993; De Filippis, V et al., 1996).

Stability and folding studies of the two homologous proteins have revealed major differences (Zerovnik E et al., 1999). In stefin B, intermediates in GnHCl and acid denaturation, have been observed, (Zerovnik E et al., 1997) whereas stefin A behaved two-state (Zerovnik E et al., 1992). Stability against TFE solvent was compared (Zerovnik E et al., 1999) the titration ending with an all α-helical denatured state, in both stefins. An α-helical folding intermediate (Hamada D et al., 1996) was detected in stefin B folding only (Zerovnik E et al., 1998).

Effect of salts and alcohols on acid induced states of different proteins like cytochrome c, myoglobin, papain and α-chymotrypsinogen-A has been studied and a partially folded intermediate of α-chymotrypsinogen-A has been characterized (Xie, D et al., 1991). In general certain proteins under conditions of low pH have the characteristics of MG state, they have compact dimension, abundant secondary structure and largely disordered tertiary structure (Plaxco, K. W et al., 1996). Recently, increasing evidences support the idea that the MG state besides having secondary structure may possess well defined tertiary contacts as well (Nishii, I et al., 1995; Fink, L. A et al., 1994). Thus, the structural similarity between the MG and native protein may have a significant bearing in understanding the protein folding problem (Goto, Y et al., 1989).

Alcohols exert several distinct effects on proteins andhave been commonly used to study partially folded states in proteins, thus it help in understanding the protein folding mechanism (Shiraki et al. 1995; Alexandrescu et al. 1994; Schonbrunner et al.

1996). Binary mixtures of water with alcohols like methanol, ethanol or 2, 2, 2–trifluoroethanol (TFE) denature the tertiary and the quaternary structures of proteins while enhancing their helicity (Radford et al., 1992). Alcohols promote local polar interactions and hydrogen bonds in proteins, at the same time weakening non-local hydrophobic interactions (Liu and Bolen 1995).

Under conditions of low pH it has been shown that several proteins have compact structure with native like secondary structure and largely disordered tertiary structure, that is, the characteristics of molten globule (MG state) which is considered to be a general intermediate in the protein folding pathway of proteins (Kuwajima 1989; Ptitsyn et al., 1990).

Protein in the molten globule state contains high level of secondary structure, as well as rudimentary native like topology. Thus the structural similarity between native and molten globule has significant bearing in understanding the protein folding problem (Bai et al., 2001). The molten globule of lysozyme has been shown to be the precursor for amyloid formation (Booth et al., 1997) and certain disease related mutations have been found to convert native proteins into a molten globule like conformation, often associated with a loss of function (Zhang and Peng 1996)

II. AMYLOID FIBRIL FORMATION

Several types of disease have been found where protein misfolding and conformational change are the main causes of the appearance and progression of disease (Raso, S.W et al., 2000).Protein denaturation and amyloid fibril formation are the key processes that control the onset and development of so called protein folding diseases such as Alzheimer's disease, systemic amyloidosis, and transmissible spongiform encephalopathy**.** Amyloidoses are various systemic and localized diseases with a common feature of 'amyloid'. Amyloid plaques are composed from deposits of misfolded proteins and some other nonproteinaceous components, such as serum amyloid P, heparan sulfate proteoglycans, and apolipoprotein E (Cohen AS 1986; Sipe JD and Cohen AS 2000). Because changed protein conformation (misfolding) leads to protein aggregation and deposition, amyloidoses are also called conformational diseases. Among them are neurodegenerative and prion diseases with

extracellular or intracellular inclusions (Goedert M et al., 1998; Cohen FE 2000; Soto C 2001). The misfolding of the proteins involved in these diseases leads to the formation of "fibrillation competent" partially folded intermediates that associate with one another to form fibrils. The onset of fibril formation can be triggered in vitro by exposing proteins to weakly denaturing conditions. This can be achieved by changing parameters such as the temperature, concentration, pH, and ionic strength of the protein solutions.

A common low-resolution structure has been observed to date for most of the amyloid fibrils. Amyloid fibrils share a common molecular skeleton, the protofilament core structure (Fig 8 and Fig 9) which is a continuous β-sheet helix (Blake C, and Serpell L 1996; Sunde M et al., 1997; Kirschner D et al., 1998; Malinchik SB et al., 1998). The reflections at ~ 4.7 Aº on the meridian and 10 Aº on the equator are seen in all amyloid fiber diffraction patterns. The structural repeat of 4.7 Aº along the fiber axis corresponds to the spacing of β-strands, and the 10–12 Aº spacing corresponds to the face to face separation of the β-sheets (Serpell LC et al., 1999).

Atomic force microscopy (AFM) has become a very important tool for the observation of surface morphology of amyloid fibrils (Ding TT et al., 1999; Goldsbury CS et al., 1997; Ionescu-Zanetti C et al., 1999) and the kinetics of their growth (Goldsbury C et al., 1999). When fibrils are deposited to the flat solid surface, it is possible to obtain accurate measurements of the height and the length of fibrils. The periodicity of the height along the fibril can also be determined.

Stability and folding studies of the two homologous proteins have revealed major differences (Zerovnik E et al., 1999). In stefin A and B, intermediates in GdnHCl and acid denaturation have been observed (Zerovnik E et al., 1997), whereas stefin A behaved as two-state model. Stability against TFE solvent was compared (Zerovnik E et al., 1999), it showed the titration ending with an all α-helical denatured state, in both stefins. An α-helical folding intermediate was detected in stefin B folding only (Zerovnik E et al., 1998; Zerovnik E et al., 1999).

So far, about 20 human proteins have been found in proteinaceous deposits in various conformational diseases. These do not demonstrate any sequence or structural homology. The common event is thought to be a conformational change, leading to lack of biological function or gain of toxic activity, and possibly, formation of

amyloid fibrils. Many proteins and peptides form amyloid fibrils (Kelly, J.W 1998; Rochet, J.C. 2000). Consistent with the notion that any protein can form amyloid fibrils, even though under harsh conditions (Zerovnik, E et al., 2010). Although most are related to diseases, it has been shown that several proteins (Chiti, F et al., 1999; Guijarro, J.I et al., 1998) peptides (Ohnishi, S et al., 2000) and, moreover, polyaminoacids (Fändrich, M et al., 2002) can also form amyloid fibrils. Amyloid fibril formation is now recognized as a phenomenon common to many proteins and peptides. Understanding the properties of such fibrils is essential to obtaining further insight into the conformation and folding of proteins. In recent years, several non-pathogenic proteins and peptides have been shown to form amyloid fibrils in vitro including acylphosphatase (chiti et al., 200), cold- shock protein (Wilkins et al., 2000), hen lysozyme (Krebs et al., 2000), SH3 domain (zurdo et al., 2001),however cystatins generally show this behaviour like human stefin A (Jenko et al., 2004), human stefin B (Rabzelj et al., 2005; Zerovnik et al., 2002) and cystatin C (Nilson, 2004). The ability to fibrillate is independent of the original native structure of the protein, whose amino acid sequence primarily appears to play a key role in terms of filament arrangement (Khurrana et al., 2003), fibrillation kinetics (Chiti et al., 2000), overall yield and stability of the fibrils (Rabzelj et al., 2005; Ahmad et al., 2004; Williams et al., 2004).

Several authors have found that proteins that have not been associated with any disease can form amyloid-like fibrils (Damaschun, G et al., 1999; Liu, Y et al., 2001; zerovnik, E et al., 2002). Especially surprising was the finding that even helical proteins, such as myoglobin (Fandrich, M et al., 2001) or apo-cytochrome c (Pertinhez, T.A et al., 2001) can form fibrils under certain conditions.These observations led Dobson and coauthors to propose that amyloid-fibril formation is a generic property of proteins (Fandrich, M et al., 2001; Dobson, C.M et al., 1999). A common observation is that fibrillization starts from an intermediate state, either partially unfolded or partially folded, molten globule or native-like intermediate (Rochet, J.C et al., 2000). In case of globular proteins such as phosphoglycerate kinase (Damaschun, G et al., 1999), cystatin C (Ekiel, I et al., 1996), acylphosphatase (Chiti, F et al., 1999) and transthyretin (Lai, Z et al., 1996), partial unfolding needs to occur to enable fibril formation and, in the case of unfolded polypeptides such as α-synuclein (Uversky, V.N et al., 2001) and islet amyloid polypeptide, these must

partially fold. The parts with the α-helical structure must undergo an α to β transition and the β- strands then associate into a regular fibrillar structure. An α to β transformation is well characterized with peptides, like poly (L-lysine). It has more recently been observed with proteins which are initially unfolded or predominantly β-sheet (Kirkitadze, M.D et al., 2001; Bouchard, M et al., 2000) and which fold through an a helical intermediate (Hamada, D et al., 1997; Zerovnik, E et al., 1999).

In vitro, variation of solvent conditions by changing pH or adding organic solvents (Buck, M. 1998) can lead to partial unfolding and subsequent protein fibril formation (Chiti, F et al., 1999). With unfolded polypeptides, partial folding can be obtained by lowering pH or by heating (Uversky, V.N et al., 2001). In vivo, partial unfolding may happen as a consequence of lowered protein stability due to mutation, local change in pH at membranes, oxidative and heat stress, whereas partial folding may happen on exposure to environmental hydrophobic substances, such as pesticides (Uversky, V.N et al., 2001).

In animal studies it has been shown that significant tissue damage and clinical symptoms appear before any protein aggregates are detected, implicating an intermediate on the amyloidogenic pathway, which could be the real cause of the pathogenesis (Roher, A.E et al., 2000; Klein, W.L et al., 2001).

Many diseases are caused by protein misfolding.These include loss-of-function diseases such as cystic fibrosis, hemophilia, and cancer. Other diseases are caused by a gain-of toxic-function mechanism often including aggregation of misfolded proteins. These diseases include the amyloidoses such as Alzheimer's disease and familial amyloidotic polyneuropathy (FAP). In the case of the prion diseases, there is a debate whether these are gain-of-toxic-function or loss-of-function diseases. It is however known that the infective species include assembled prion protein (PrP) oligomers in an aberrant folded state called PrPSC. In the classic amyloidoses, accumulation of extracellular fibrillar aggregates called amyloid has been detected in various tissues depending on disease (Buxbaum, 2003). In the past few years it has become evident that the mature fibrils are fairly inert and that prefibrillar assemblies are the most cytotoxic species. Therefore oligomers of misfolded proteins appear to

Figure 8: A schematic representation of some of the many conformational states that can be adopted by polypeptide chains and of the means by which they can be interconverted. The transition from β-structured aggregates to amyloid fibrils can occur by addition of either monomers or protofibrils (depending on protein) to preformed β-aggregates.

Figure 9: **Schematic representation of the pathway leading to protein misfolding and aggregation.**The natively folded protein, normally produced in diverse cell types, adopts arandom coil or α-helical conformation. In the elderly brain, the first pathological step would be theformation of a misfolded intermediate that exposed to the aqueous environment hydrophobicfragments that are normally buried inside the protein. This intermediate has a high tendency toaggregate and become stabilized, in a rate-limiting process, by the formation of an oligomericβ-sheet structure, which by incorporation of additional monomers gives rise to protofibrils andfinally to cross-β amyloid-like fibrils.

be yet another common denominator in the amyloidoses in addition to the generic fibrillar structure. Such prefibrillar oligomers could also serve as a common intermediate between the amyloidoses and other misfolding diseases, including prion diseases, Parkinson's disease, and amytrophic lateral sclerosis. Interestingly, amyloid deposits can also be found in prion diseases, suggesting a common link between these diseases. Furthermore, PrPSC particles share many of the features of the toxic small prefibrillar species found in amyloid diseases.

DRUG-PROTEIN INTERACTION

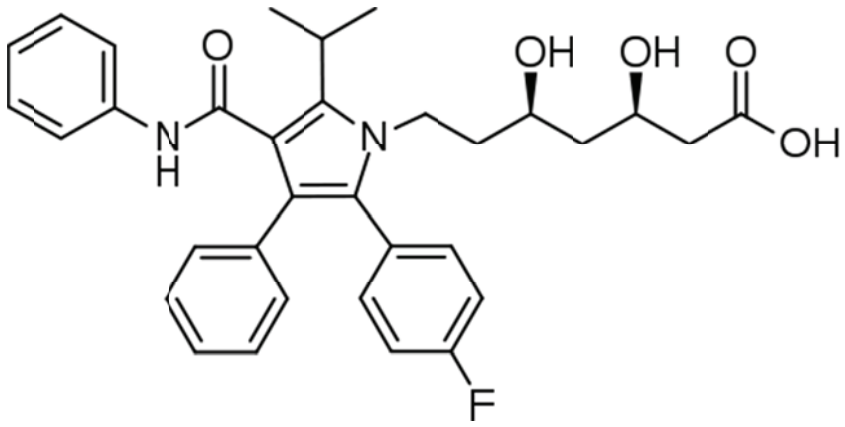
EFFECT OF ANTI-TUBERCLOSIS DRUG (RIFAMPCIN) AND ANTI-DIABETIC DRUG (STATINS) ON BUFFALO LIVER CYSTATIN

ATORVASTATIN

The 3-hydroxy-3-methylglutaryl coenzyme A (HMGCoA) reductase inhibitors (statins) are a well-established class of drugs in the treatment of hypercholesterolemia, and members of this class have been shown to reduce the risk of cardiovascular morbidity and mortality in patients with or at risk for coronary heart disease (CHD) (Mahley et al., 2001) in several clinical trials. Statins are the drugs of choice for patients with hypercholesterolemia and other risk factors for cardiovascular disease and have shown to be a life-reserving therapy in many of these patients. A similar proportion of patients randomized to placebo and active treatment with statins in clinical trials have shown elevations in aminotransferases (Bradford et al., 1999; Schwartz et al., 2001). Although most trials assessing the cardiovascular efficacy of statins and their safety have included a large number of patients, they have been underpowered to detect clinically relevant drug-induced liver injury (DILI). It has been convincingly shown that the risk of developing statin induced DILI is not related to the presence of pre-existing liver abnormalities, mostly non-alcoholic fatty liver disease (NAFLD) (Chalasani et al., 2004). On the contrary, the use of statins has been shown to be associated with improvement in liver test abnormalities and histology in patients with NAFLD (Rallidis et al., 2004; Hyogo et al., 2008; Ekstedt et al., 2007; Athyros et al., 2010).

Atorvastatin and simvastatin are the most common statins associated with DILI, which is probably due to the fact that these are the most commonly used statins. Atorvastatin is mostly associated with cholestatic liver injury whereas hepatocellular injury is more common with simvastatin. The current series, which is by far the largest series of statin induced DILI, does support that clinically important idiosyncratic liver injury can occur in patients treated with statins (Lewis et al., 2012).

To understand the molecular bases of statin–drug interactions, it is important to mention that statins are very selective inhibitors of HMG-CoA reductase and usually do not show any relevant affinity toward other enzymes or receptor systems (Corsini

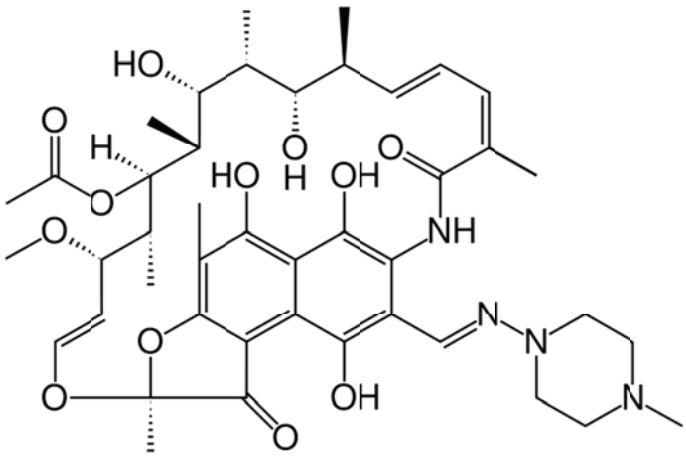


Chemical structure of Atorvastatin (Statin)

et al., 1999). However, at the pharmacokinetic level (ie, absorption, distribution, metabolism, and excretion of a given drug), the available statins have important differences, including half-life, maximum plasma concentration (Cmax), bioavailability, protein binding (Omar et al., 2002; Farmer et al., 2000; Hamilton-Craig, 2001). Concomitant use of certain drugs (fibrates, erythromycin, itraconazole, and immunosuppressive drugs such as cyclosporine) can increase blood levels of statins and, consequently, the risk for myopathy (Gotto, 2003). The relationship between altered plasma concentrations and adverse effects or toxicity might not be linear (Athyros et al., 2010). Other variables that affect this concentration–effect relationship include rapid changes in the concentrations, concomitant lipid-lowering therapy, or host genetic factors that code for different forms or amounts of metabolizing enzymes and drug receptors.

RIFAMPCIN

Rifampicin, a member of the rifamycin group of antibiotics, has achieved 'first line' status in the treatment of tuberculosis because of its low toxicity and high therapeutic efficacy. It is metabolized in the liver mainly by desacetylation followed by glucuronidation. This results in the formation of a more polar, water-soluble compound which is more easily excreted in the bile (Acocella, 1978). About 80 per cent of an orally administered dose is excreted as rifampicin and desacetylrifampicin in the bile. The remaining 20 per cent is excreted in the urine as both the free drug and desacetylrifampicin with urinary concentrations reflecting serum levels (Acocella, 1978). With continued administration, rifampicin concentrations and half-life in the serum decrease as a consequence of induction of its own metabolic pathways (Acocella et al., 1971).



Chemical structure of Rifampicin

There is a corresponding fall in urinary rifampicin excretion and rise in biliary excretion (Acocella, 1978). Patients with cirrhosis have been shown to exhibit higher serum concentrations than normal subjects on similar doses of the drug (Capelle et al., 1972), but there is little information on the relationship between alterations in the pharmaco-kinetics of the drug during continued administration of rifampicin and the severity of liver dysfunction in cirrhosis.

The pathogenesis of RF-INH induced hepatotoxicity is not entirely clear, but the proposed mechanisms may include dose-related toxicity (Yew, 1998), oxidative stress

(Attri et al., 2000), lipid peroxidation (Richards et al., 2004), immune-related (Gupta et al., 1975), induction of liver enzyme in the hydrolase system, thus enhancing the toxicity of some of the INH toxic metabolites, activation of CYP2E1 (Yue et al., 2004), choline deficiency leading to lowering of phospholipids protein synthesis with alteration in cell wall configuration (Karthikeyan, 2005).

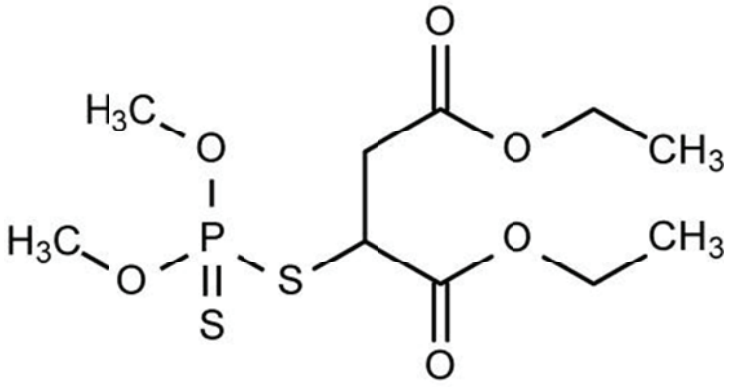
"Rifampicin hepatitis"was originally described by Scheuer et al in 1974 in which patients developed abnormal LFTs within six weeks of starting rifampicin therapy. These patients were found to have a wide range of histological changes and varying degrees of portal inflammation.

The liver metabolizes xenobiotics, so it is not surprising that drug-induced liver injury (DILI) is a potential complication of many drugs.The formation of toxic reactive metabolites during hepatic metabolism is believed to be the main pathogenic mechanism for DILI (Zimmerman, 1999; Knowles et al., 2000; Pessayre, 1995).

In 2 studies that tested rifampin for treating pruritus in patients with primary biliary cirrhosis, there was an increased frequency of hepatotoxicity (12.5% and 7.3%) (Bachs et al., 1992; Prince et al., 2002); similarly, rifampin might have caused DILI in 5% of patients when administered to treat brucellosis (Cascio et al., 2003).

Effect of pesticide (Malathion) on Buffalo liver cystatin

Malathion [O,O-dimethyl-S-(1,2-dicarcethoxyethyl) phosphorodithioate] is an OP pesticide in agricultural and household products that is widely used to control pests. It is extensively used over the whole world, especially by developing countries, to control or eradicate disease-inducing arthropods targeted by public health programs; it is also used to eliminate animal ectoparasites, human head and body lice, and household insects, and to protect grain in storage (Elston, 2002; Assini et al., 2005; Rezg et al., 2008). Malathion is known to inhibit acetylcholinesterase activity in target tissues (Rezg et al., 2008) and has been linked to the dysfunction of several organ systems, including the liver (Lasram et al., 2009), the pancreas (Vasough-Ghanbari et al., 2007), and the reproductive system (Bustos-Obregon and Gonzales-Hormazabal, 2003).



Chemical structure of Malathion

Increased production of malonyl dialdehyde, indicating increased lipid peroxidation was observed in chicks on receiving malathion, thus indicating increase in ROS production induced by malathion (Sodhi et al., 2008). There are reports that organophosphorus pesticides can intiate reactive oxygen species (ROS) production leading to lipid peroxidation, proteins, DNA and membrane damage. Kim and Mahan, 2003 have reported a clear cut relation between reduced glutathione peroxidase levels and development of liver necrosis. Malathion exposure for prolonged periods can be hepatotoxic (Grewal, 2003) and immunotoxic . the toxic effect of malathion cab be due to depletion of free radical scavengers and this can lead to hepatotoxicity (Sodhi et al., 2008).

SCOPE OF THE THESIS

The healthy human body might be described schematically as being composed of several dynamic equilibria. All diseases might be considered as distrurbances in one or more of these equilibria. The balance between protein production and degdradation is one of these equlibria, which are crucial to health. In the lysosomal pathway, protein degradation is a result of combined random and limited action of cathepsins. The harmful activity of cathepsins can be regulated by their endogenous protein inhibitors, members of cystatin superfamily. Since uncontrolled proteolysis can lead to irreversible damage and detrimentally affect human health by facilitating diseases control of their protease function by cystatins is of cardinal importance. Liver is a miracle worker, it purifies, detoxifies and even regenerates. So when liver is in distress, overall health could be in serious trouble.

Keeping in view the fact that overreactivity of cathepsins can lead various liver diseases an attempt was made to purify cystatin (thiol proteinase-inhibitor) from buffalo liver. In the present study its detailed biochemical characterization has been undertaken, which included elucidation of its molecular weight, pH, thermal stability and immunogenic properties. Its interaction with model cysteine proteinase, papain, was worked out to determine its kinetic properties. The interaction of liver cystatin with bilirubin which is in the end product of heme metabolism was studied using fluorescence and UV spectroscopy. The correlations between bilirubin and cystatin may play a significant role in various hepatic diseases such as jaundice, which is accompanied by increase of bilirubin levels. The interaction of sugars (glucose, fructose and ribose) with liver cystatin was carried out *in vitro* under conditions that more or less mimic the *in vivo* reaction with sugars. The formation of advanced glycation end products (AGEs) found in patients with liver cirrhosis has a significant importance in diseased states.

The studies are worthy of future research with regards to interaction of liver cystatin with cathepsins and its role in hepatic (patho)physiology.