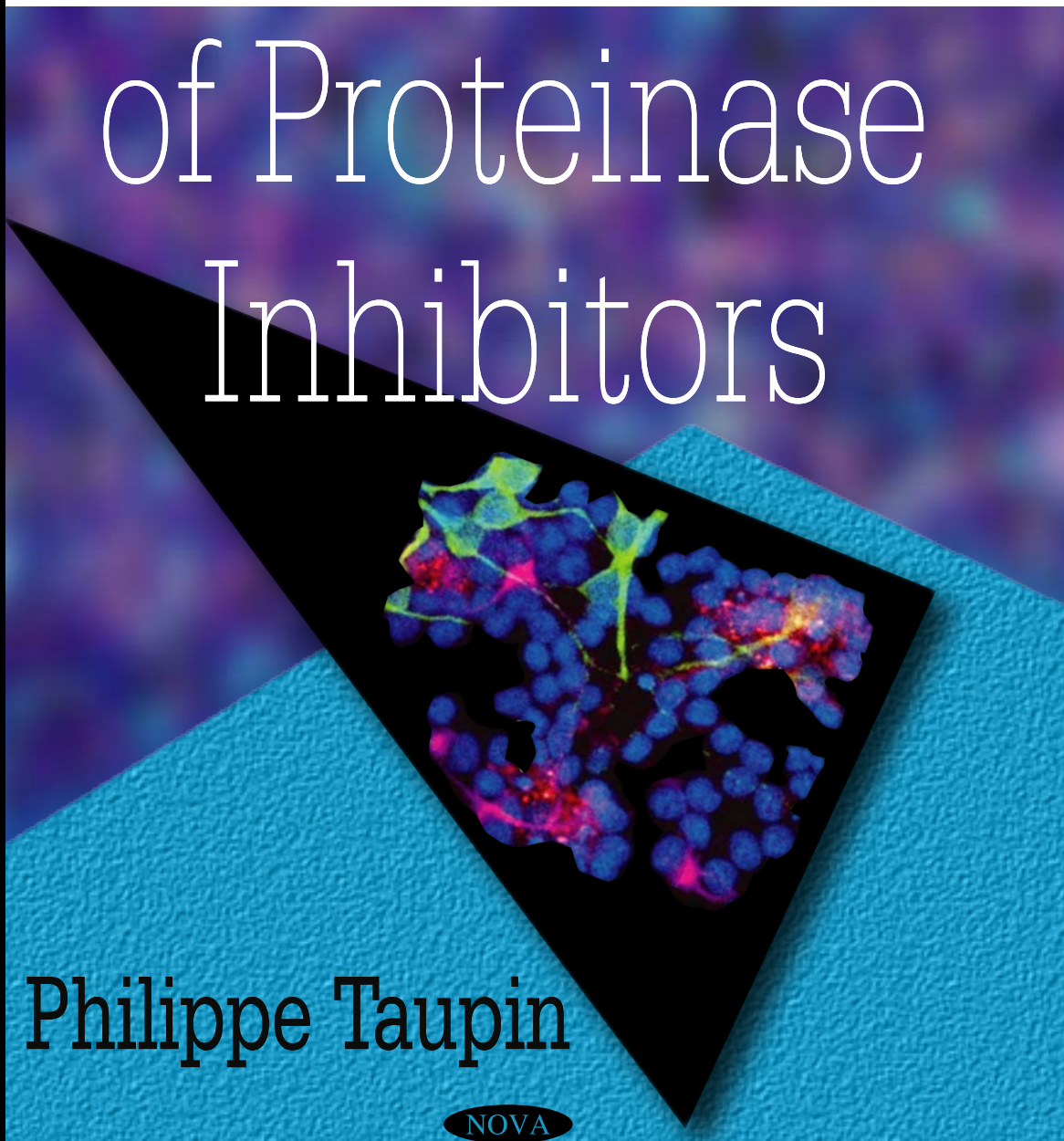
Nova Biomedical

The Cystatin Superfamily of Proteinase Inhibitors



Philippe Taupin

NOVA

THE CYSTATIN SUPERFAMILY OF PROTEINASE INHIBITORS

THE CYSTATIN SUPERFAMILY OF PROTEINASE INHIBITORS

PHILIPPE TAUPIN

Nova Biomedical Books
New York

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

The cystatin superfamily of proteinase inhibitors / Philippe Taupin, editor.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-1-60692-748-9
1. Cysteine proteinases--Inhibitors. 2. Cystatins. I. Taupin, Philippe.
[DNLM: 1. Cysteine Proteinase Inhibitors--physiology. 2. Cysteine Proteinase Inhibitors--therapeutic use. 3. Cystatins--physiology. QU 136 C9965 2007]
QP609.C94C92 2007
572'.633--dc22

2007039062

Published by Nova Science Publishers, Inc. ✣ New York

CONTENTS

INTRODUCTION

The cystatin superfamily of proteinase inhibitors consists of reversible tight-binding inhibitors of proteases of the papain superfamily or family C1 of cysteine proteases. Cysteine proteases are proteases that involve cysteine residues in the proteolytic reaction. The papain superfamily consists of cysteine proteases with high degree of homology to papain. Papain is the cysteine protease originally isolated from the latex of *Carica papaya* fruit. The papain superfamily includes the mammalian cathepsins B, H, K, L and S [1].

The cystatin superfamily regroups cysteine proteinase inhibitors (CPIs), with high degree of homology to chicken cystatin (cC). Chicken cystatin is the first inhibitor of cystatin superfamily, isolated and characterized. It was purified from chicken egg white [2]. The CPIs of the cystatin superfamily originate from a common ancestor gene [3,4]. The cystatin superfamily regroups three families of CPIs, the stefin, cystatin and kninogen families. The diversity of CPIs of the cystatin superfamily reflects the diversity of cysteine proteases of the papain superfamily.

CPIs of the cystatin superfamily are broadly expressed and present in a wide range of species, organisms and tissues, from mammals, birds, fish, insects, plants to bacteria, parasites and viruses. Hence, CPIs of the cystatin superfamily are involved in a broad range of physio- and pathological processes.

The characterization of the activity of CPIs of the cystatin superfamily represents a major step toward the understanding of the functioning, and physio- and pathology of organisms, and particularly mammals.

The catalytic domain and three-dimensional structure of CPIs of the cystatin superfamily, including cC and human cystatin C (hCC), have been determined, revealing a mechanism by which CPIs of the cystatin superfamily inhibit cysteine proteases [5].

Chicken cystatin and hCC also inhibit cysteine proteases of the legumain family or family C13, a family of proteases unrelated to family C1. This not only reveals that CPIs of the cystatin superfamily elicit a second catalytic domain, distinct of the catalytic site for cysteine proteases of family C1, but also a mechanism by which nature as generated redundant activities.

CPIs of the cystatin superfamily are involved in peptide and protein metabolism. As such, they are involved in all aspects of the life of cells, tissues and organs. Normal physiological functioning of any organism requires a precise balance between the activities of

proteases and their inhibitors. By regulating the activity of proteases, proteinase inhibitors prevent their potential deleterious effects on the cellular environment, like destruction of cells, tissues and connective tissues. CPIs of the cystatin superfamily are also involved in bone resorption, host pathogen reactions, inflammation and immune responses, stem cell growth and fate specification [6].

Mechanisms underlying the activities CPIs of the cystatin superfamily do not only involve their proteinase inhibitory activities, but also other domains of the molecules unrelated to their CPI activities. CPIs of the cystatin superfamily, and particularly cystatin C, are pleiotropic molecules, with different domains of the molecules involved in different functions [6].

Human cystatin C can form dimmers under mild predenaturing conditions [7]. A three-dimensional domain swapping mechanism has been proposed for the dimerization of cystatin. Three-dimensional domain swapping is believed to be a model by which amyloid fibrils are formed, revealing hCC as an amyloidogenic protein. Amyloidogenic proteins are proteins forming amyloid fibrils [8].

Amyloidosis is a group of rare and potentially fatal diseases that result from the deposit of amyloid fibrils. Human cystatin C is associated with a number of amyloidosis, including Alzheimer's disease, hereditary cystatin C amyloid angiopathy and hereditary cerebral hemorrhage with amyloidosis-Dutch type. The involvement of cystatin in the pathological processes of these diseases remains mostly unknown. Cystatin would contribute to the formation of amyloid fibrils, as a CPI and/or as an amyloidogenic protein. Cystatins are also involved in other diseases, like age-related macular degeneration, cancers and epilepsy.

This shows that CPIs of the cystatin superfamily have tremendous therapeutic potentials for the treatment of a broad range of diseases, raging from age-related macular degeneration, amyloidosis, bone diseases, cancers, immunological and neurological diseases to bacterial, parasite and viral infections.

Understanding of the mechanisms underlying the activities of CPIs of the cystatin superfamily is a prerequisite for designing drugs and therapies based on this family of molecules.

In all, the cystatin superfamily of proteinase inhibitors holds a tremendous potential for our understanding of the biology of living organisms, particularly mammals, and for therapy.

REFERENCES

[1] Berti PJ, Storer AC. (1995) Alignment/phylogeny of the papain superfamily of cysteine proteases. *J Mol Biol. 246*, 273-83.

[2] Sen LC, Whitaker JR. (1973) Some properties of a ficin-papain inhibitor from avian egg white. *Arch Biochem Biophys. 158*, 623-32.

[3] Barrett AJ, Fritz H, Grubb A, Isemura S, Jarvinen M, Katunuma N, Machleidt W, Muller-Esterl W, Sasaki M, Turk V. (1986) Nomenclature and classification of the proteins homologous with the cysteine-proteinase inhibitor chicken cystatin. *Biochem J. 236*, 312.

[4] Muller-Esterl W, Fritz H, Kellermann J, Lottspeich F, Machleidt W, Turk V. (1985) Genealogy of mammalian cysteine proteinase inhibitors. Common evolutionary origin of stefins, cystatins and kininogens. *FEBS Lett. 191*, 221-6.

[5] Bode W, Engh R, Musil D, Thiele U, Huber R, Karshikov A, Brzin J, Kos J, Turk V. (1988) The 2.0 A X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J. 7*, 2593-9.

[6] Taupin P, Ray J, Fischer WH, Suhr ST, Hakansson K, Grubb A, Gage FH. (2000) FGF-2 responsive neural stem cell proliferation requires CCg, a novel autocrine/paracrine cofactor. *Neuron. 28*, 385-97.

[7] Janowski R, Kozak M, Jankowska E, Grzonka Z, Grubb A, Abrahamson M, Jaskolski M. (2001) Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping. *Nat Struct Biol. 8*, 316-20.

[8] Jaskolski M. (2001) 3D domain swapping, protein oligomerization, and amyloid formation. *Acta Biochim Pol. 48*, 807-27.

Chapter I

STEFINS, CYSTATINS AND KININOGENS

ABSTRACT

Stefins, cystatins and kininogens are tight-binding reversible inhibitors of cysteine proteases, or cysteine protease inhibitors (CPIs), of the papain superfamily or family C1 of cysteine proteases. Cysteine proteases are proteases that involve cysteine residues in the proteolytic reaction. Cysteine proteases of the papain superfamily include cathepsins B, H, K, L and S. The diversity of CPIs, stefins, cystatins and kininogens, reflects the diversity of cysteine proteases of the papain superfamily. The stefins, cystatins and kininogens are broadly expressed and present in a wide range of species, organisms and tissues, from mammals, birds, fish, insects, plants to bacteria, parasites and viruses.

INTRODUCTION

Proteases are enzymes that degrade polypeptides and proteins. Cysteine proteases (EC 3.4.22) are proteases that involve cysteine residues in the proteolytic reaction [1]. There are at least 17 families of cysteine proteases in mammals, including the papain, calpain and caspase families (table 1). Cysteine proteases belonging to the family C1 or papain superfamily are proteases with homology to papain. Papain (EC 3.4.22.2) is a cysteine protease originally isolated from the latex of *Carica papaya* fruit. Members of the C1 family include the cathepsins B (EC3.4.22.1), H (EC 3.4.22.16), K (EC 3.4.22.38), L (EC 3.4.22.15) and S (EC 3.4.22.27) present in mammals [2]. Papain and cysteine proteases of the C1 family originate from a common ancestor gene [3]. Cysteine proteases of the family C1 are primarily endopetidases. Cathepsin B, for example, functions also as an exopeptidase, in addition to the endopeptidase activity of papain-like cysteine proteases.

Cysteine protease, or proteinase, inhibitors (CPIs) are inhibitors of cysteine proteases. Chicken cystatin (cC) is the first inhibitor of cysteine proteases of the papain superfamily, isolated and characterized [4]. It was purified from chicken egg white [4-6]. Other CPIs of the papain superfamily have been identified, like the stefins, cystatins and plasma kininogens [7-9]. In 1984, the first partial sequence of human cystatin C (hCC) was published [10]. It

elicits high homology with cC and was reported to be homologous to a previously isolated and characterized protein from humans, named γ-trace [11,12]. γ-trace, a protein of unknown function, was characterized from cerebrospinal fluid, urine patients with renal failure, saliva and semen. γ-trace is an alkaline low molecular weight protein, a single polypeptide chain containing 120 residues [13-15].

CPIs bind to the active site of cysteine proteases of the papain superfamily, preventing substrate binding and therefore inhibiting their proteolytic activities. CPIs of cysteine proteinases of the papain superfamily are classified in three families, based on their structure. These three families of CPIs bear the name of their most representative prototype: the stefin, cystatin and kininogen families [16-18].

Table 1. Cysteine protease families

Classification	Family
Family C1:	papain superfamily
Family C2:	calpain family
Family C10:	streptopain family
Family C11:	clostripain family
Family C12:	ubiquitin C-terminal hydrolase family
Family C13:	legumain family
Family C14:	caspase family
Family C15:	pyroglutamyl peptidase I family
Family C19:	ubiquitin-specific protease family
Family C25:	gingipain family
Family C26:	gamma-glutamyl hydrolase family
Family C47:	staphopain family
Family C48:	Ulp1 endopeptidase family
Family C50:	separase family
Family C54:	Aut2 peptidase family
Family C55:	YopJ protease family
Family C56:	PfpI endopeptidase family

Proteases are enzymes that degrade polypeptides and proteins. Cysteine proteases are proteases that involve cysteine residues in the proteolytic reaction. There are at least 17 families of cysteine proteases in mammals.

FAMILY 1 OR STEFIN FAMILY

The family 1 or stefin family of CPIs of the papain superfamily (also referred to as type 1, class I cystatin, or family 1 cystatin) are low molecular weight proteins of approximately 10-11 kDa. In humans, the stefins A and B or cystatins A and B contain approximately 100 amino acid residues. They have a single polypeptide chain, no disulfide bonds and no carbohydrate moiety [19].

Stefins are intracellular protease inhibitors; they have no signal peptides. The signal peptide is a 26 amino acid N-terminal hydrophobic leader sequence. The signal peptide addresses proteins for cell secretion through the Golgi apparatus and secretory pathway. Stefins have been detected in extracellular medium, where they may originate from dying cells releasing their content.

FAMILY 2 OR CYSTATIN FAMILY

The family 2 or cystatin family of CPIs of the papain superfamily (also referred to as type 2, class II cystatin, or family 2 cystatin) have a molecular weight of approximately 13-14 kDa. They have a single chain of approximately 120 amino acid residues and contain two (intrachain) disulfide bonds [20]. Cystatins are secreted proteins; they are synthesized as preproproteins with signal peptides [21-24].

Chicken cystatin is the most characterized member of the cystatin family [4-6,25]. It is a single polypeptide chain of 116 residues, not glycosylated [26]. In rats, cystatin C is a glycoprotein, with both N- and O-glycosylation [27-29]. In humans, eight members of the cystatin family have been identified: cystatin C, D, E/M, F (also called leukocystatin), G, S, SN and SA [30-34]. Human cystatin C (hCC) is a single polypeptide chain of 120 amino acid residues. It is an extracellular protein of 13.4 kDa, synthesized as a preprotein with a signal peptide [35-37]. Cystatins C, D, G, S, SN and SA have no carbohydrate moiety, whereas cystatin E/M and F are glycoproteins [38,39]. Cystatins have also been identified from other mammals, like bovines [40].

Three motifs of the polypeptide chain of hCC, Arg8-Leu9-Val10-Gly11, Gln55-Ile56-Val57-Ala58-Gly59 and Pro105-Trp106, are involved in the CPI activity [41-45].

Cystatin C is ubiquitously expressed in human tissues and body fluids. Cystatin C is particularly expressed in the nervous system [46]. Other members of the cystatin family, cystatins D, S, SN, and SA have a more restricted distribution in human body fluids [30-34].

FAMILY 3 OR KININOGEN FAMILY

The family 3 or kininogen family of CPIs of the papain superfamily (also referred to as type 3, class III cystatin, or family 3 cystatin) are high molecular weight proteins. Kininogens contain disulphide bonds and are glycosylated. Kininogens are synthesized as precursor proteins or high molecular weight kininogens (H-kininogen) of 644 amino acids. Upon maturation, H-kininogen is processed into heavy chain or low molecular weight kininogen (L-kininogen), light chain and the peptide bradykinin.

L-kininogen contains three cystatin-like domains [47]. This suggests that the L-kininogen may correspond to a triple-headed inhibitor of cysteine proteinases. Indeed, among these three domains of the N-terminal heavy chain of kininogens or L-kininogen, only 2 have the conserved consensus sequence QVVAG, motif corresponding to the active site of cystatins. So, 2 out of the 3 sites of kininogens are functional CPI sites [47].

Kininogens are intravascular proteins. They are found in blood plasma and are involved in blood coagulation.

SYNTHESIS AND POST-TRANSLATIONAL MODIFICATIONS

Translation

Family 2 or cystatin family of CPIs are secreted protease inhibitors found in most biological fluids. Cystatin C is also detected inside cells. High amounts of cystatin C are detected in class II-positive lysosomes of both immature dendritic cells (DCs) and Langerhans cells (immature DCs in the epidermis) [48].

Cystatins family 2 are secreted proteins; they are synthesized as preproproteins with signal peptides. They have a signal peptide and are processed through the Golgi apparatus and secretory pathway. The cystatin detected inside the cells correspond to cystatin being synthesized and processed through the Golgi apparatus and secretory pathway. It may also originate from the endocytic pathway, during the protein degradation process.

Cystatin C may therefore being detected in perinuclear cytoplasm, the Golgi apparatus, secretory system and lysosomes, as well as the surface of the cells. Different factors, including transforming growth factor-β and oxidative stress, can increase cystatin C expression in various cells, and therefore its detection inside the cell compartments [49-51].

Post-Translational Modification: Disulfide Bonds

Cystatins and kininogens contain disulphide bonds. Cystatins, family 2, contains four Cys residues forming two characteristic disulfide bonds. The disulfide bonds are formed in the C-terminal region of the molecule, between Cys73 and Cys83, and between Cys97 and Cys117, in hCC. The disulfide bonds introduce rigidity into the molecule, as well as protection against protease degradation [20].

Post-Translational Modification: Phosphorylation

Chicken cystatin is not glycosylated, but it occurs in a phosphorylated and non-phosphorylated form [52]. Chicken cystatin is typically isolated in different forms, varying at their N-terminal. The various form isolated of cC originate from degradation of their N-terminal sequence during the purification process of the native protein, the Ser-form (116 residues). The most common form purified of cC is the Gly9-form (108 residues), truncated of the first eight N-terminal residues [53]. Glycyl endopeptidase is the enzyme that truncates the N-terminal segment of cystatins.

Bidimensional electrophoresis reveals that identical amino acid forms of cC elicit two isoelectric points, corresponding to a phosphorylated and non- phosphorylated form of cC. The non-phosphorylated form of cC has a pI of 6.5 and the phosphorylated form a pI of 5.6

[52]. The site of phosphorylation of cC has been identified has Ser80, recognition site for the serine kinase [54].

Phosphorylation has been demonstrated for other members of the cystatin family 2, like human cystatin S [32], highlighting the importance of phosphorylation for this family of proteins. The role of phosphorylation of cystatins remains to be determined. It may occur for targeting and addressing during the secretion process.

CONCLUSION

Stefins, cystatins and kininogens are inhibitors of proteases of the papain superfamily, including cathepsins B, H, J, K, L, S and X [55,56]. While stefins have no disulfide bond and are intracellular proteins, cystatins and kininogens are secreted proteins, with disulfide bonds. Chicken cystatin, one of the most characterized CPIs, is an inhibitor of the proteases of the papain superfamily. Cystatins, like rat cystatin C, human cystatin E/M and F, are glycosylated, while others, like cC and hCC, are not. The N-terminal sequence and consensus sequence QVVAG (or motif Gln-Xaa-Val-Xaa-Gly) correspond to active sites of stefins, cystatins and kininogens. The protein sequences of the stefins, cystatins and kininogens elicit a high degree of homology and originate from a common ancestor gene [57]. Because cysteine proteases, particularly of the papain superfamily, exist in all living organisms and are involved in various biological and pathological processes, the control of these protease functions by the stefins, cystatins and kininogens is of major importance for the functioning of organisms, and physio- and pathology.

REFERENCES

[1] Rzychon M, Chmiel D, Stec-Niemczyk J. (2004) Modes of inhibition of cysteine proteases. *Acta Biochim Pol. 51*, 861-73.

[2] Nagler DK, Menard R. (2003) Family C1 cysteine proteases: biological diversity or redundancy? *Biol Chem. 384*, 837-43.

[3] Berti PJ, Storer AC. (1995) Alignment/phylogeny of the papain superfamily of cysteine proteases. *J Mol Biol. 246*, 273-83.

[4] Sen LC, Whitaker JR. (1973) Some properties of a ficin-papain inhibitor from avian egg white. *Arch Biochem Biophys. 158*, 623-32.

[5] Turk V, Brzin J, Longer M, Ritonja A, Eropkin M, Borchart U, Machleidt W. (1983) Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin from chicken egg white. *Hoppe Seylers Z Physiol Chem. 364*, 1487-96.

[6] Anastasi A, Brown MA, Kembhavi AA, Nicklin MJ, Sayers CA, Sunter DC, Barrett AJ. (1983) Cystatin, a protein inhibitor of cysteine proteinases. Improved purification from egg white, characterization, and detection in chicken serum. *Biochem J. 211*, 129-38.

[7] Salvesen G, Parkes C, Abrahamson M, Grubb A, Barrett AJ. (1986) Human low-Mr kininogen contains three copies of a cystatin sequence that are divergent in structure and in inhibitory activity for cysteine proteinases. *Biochem J. 234*, 429-34.

[8] Schwabe C, Anastasi A, Crow H, McDonald JK, Barrett AJ. (1984) Cystatin. Amino acid sequence and possible secondary structure. Biochem J. 217, 813-7.

[9] Abrahamson M, Alvarez-Fernandez M, Nathanson CM. Cystatins. (2003) *Biochem Soc Symp. 70*, 179-99.

[10] Brzin J, Popovic T, Turk V, Borchart U, Machleidt W. (1984) Human cystatin, a new protein inhibitor of cysteine proteinases. *Biochem Biophys Res Commun. 118*, 103-9.

[11] Grubb A, Lofberg H. (1982) Human gamma-trace, a basic microprotein: amino acid sequence and presence in the adenohypophysis. *Proc Natl Acad Sci U S A. 79*, 3024-7.

[12] Barrett AJ, Davies ME, Grubb A. (1984) The place of human gamma-trace (cystatin C) amongst the cysteine proteinase inhibitors. *Biochem Biophys Res Commun. 120*, 631-6.

[13] Tonnelle C, Colle A, Fougereau M, Manuel Y. (1979) Partial amino acid sequence of two forms of human post-gamma-globulin. *Biochem Biophys Res Commun. 86*, 613-9.

[14] Barrett AJ, Davies ME, Grubb A. (1984) The place of human gamma-trace (cystatin C) amongst the cysteine proteinase inhibitors. *Biochem Biophys Res Commun. 120*, 631-6.

[15] Warfel AH, Zucker-Franklin D, Frangione B, Ghiso J. (1987) Constitutive secretion of cystatin C (gamma-trace) by monocytes and macrophages and its downregulation after stimulation. *J Exp Med. 166*, 1912-7.

[16] Barrett AJ, Fritz H, Grubb A, Isemura S, Jarvinen M, Katunuma N, Machleidt W, Muller-Esterl W, Sasaki M, Turk V. (1986) Nomenclature and classification of the proteins homologous with the cysteine-proteinase inhibitor chicken cystatin. *Biochem J. 236*, 312.

[17] Turk V, Bode W. (1991) The cystatins: protein inhibitors of cysteine proteinases. *FEBS Lett. 285*, 213-9.

[18] Abrahamson M, Alvarez-Fernandez M, Nathanson CM. Cystatins. (2003) *Biochem Soc Symp. 70*, 179-99.

[19] Sali A, Turk V. (1987) Prediction of the secondary structures of stefins and cystatins, the low-molecular mass protein inhibitors of cysteine proteinases. *Biol Chem Hoppe Seyler. 368*, 493-9.

[20] Bjork I, Ylinenjarvi K. (1992) Different roles of the two disulfide bonds of the cysteine proteinase inhibitor, chicken cystatin, for the conformation of the active protein. *Biochemistry. 31*, 8597-602.

[21] Warfel AH, Zucker-Franklin D, Frangione B, Ghiso J. (1987) Constitutive secretion of cystatin C (gamma-trace) by monocytes and macrophages and its downregulation after stimulation. *J Exp Med. 166*, 1912-7.

[22] Chapman HA Jr, Reilly JJ Jr, Yee R, Grubb A. (1990) Identification of cystatin C, a cysteine proteinase inhibitor, as a major secretory product of human alveolar macrophages in vitro. *Am Rev Respir Dis. 141*, 698-705.

[23] Strauss M, Bartsch FO, Stollwerk J, Trstenjak M, Bohning A, Gassen HG, Machleidt W, Turk V. (1988) Chemical synthesis of a gene for human cystatin C and its expression in E. coli. *Biol Chem Hoppe Seyler. 369*, 209-18.

[24] Paraoan L, Grierson I, Maden BE. (2003) Fate of cystatin C lacking the leader sequence in RPE cells. *Exp Eye Res. 76*, 753-6.

[25] Auerswald EA, Genenger G, Mentele R, Lenzen S, Assfalg-Machleidt I, Mitschang L, Oschkinat H, Fritz H. (1991) Purification and characterization of a chicken egg white cystatin variant expressed in an Escherichia coli pIN-III-ompA system. *Eur J Biochem. 200*, 131-8.

[26] Turk V, Brzin J, Longer M, Ritonja A, Eropkin M, Borchart U, Machleidt W. (1983) Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin from chicken egg white. *Hoppe Seylers Z Physiol Chem. 364*, 1487-96.

[27] Esnard A, Esnard F, Faucher D, Gauthier F. (1988) Two rat homologues of human cystatin C. *FEBS Lett. 236*, 475-8.

[28] Esnard F, Esnard A, Faucher D, Capony JP, Derancourt J, Brillard M, Gauthier F (1990). Rat cystatin C: the complete amino acid sequence reveals a site for N-glycosylation. *Biol Chem Hoppe Seyler. 371,* 161-6.

[29] Taupin P, Ray J, Fischer WH, Suhr ST, Hakansson K, Grubb A, Gage FH. (2000) FGF-2-responsive neural stem cell proliferation requires CCg, a novel autocrine/paracrine cofactor. *Neuron. 28*, 385-97.

[30] Schwabe C, Anastasi A, Crow H, McDonald JK, Barrett AJ. (1984) Cystatin. Amino acid sequence and possible secondary structure. *Biochem J. 217*, 813-7.

[31] Abrahamson M, Grubb A, Olafsson I, Lundwall A. (1987) Molecular cloning and sequence analysis of cDNA coding for the precursor of the human cysteine proteinase inhibitor cystatin C. *FEBS Lett. 216*, 229-33.

[32] Isemura S, Saitoh E, Sanada K, Minakata K. (1991) Identification of full-sized forms of salivary (S-type) cystatins (cystatin SN, cystatin SA, cystatin S, and two phosphorylated forms of cystatin S) in human whole saliva and determination of phosphorylation sites of cystatin S. *J Biochem (Tokyo).110*, 648-54.

[33] Ni J, Abrahamson M, Zhang M, Fernandez MA, Grubb A, Su J, Yu GL, Li Y, Parmelee D, Xing L, Coleman TA, Gentz S, Thotakura R, Nguyen N, Hesselberg M, Gentz R. (1997) Cystatin E is a novel human cysteine proteinase inhibitor with structural resemblance to family 2 cystatins. *J Biol Chem. 272*, 10853-8.

[34] Halfon S, Ford J, Foster J, Dowling L, Lucian L, Sterling M, Xu Y, Weiss M, Ikeda M, Liggett D, Helms A, Caux C, Lebecque S, Hannum C, Menon S, McClanahan T, Gorman D, Zurawski G. (1998) Leukocystatin, a new Class II cystatin expressed selectively by hematopoietic cells. *J Biol Chem. 273*, 16400-8.

[35] Brzin J, Popovic T, Turk V, Borchart U, Machleidt W. (1984) Human cystatin, a new protein inhibitor of cysteine proteinases. *Biochem Biophys Res Commun. 118*, 103-9.

[36] Abrahamson M, Grubb A, Olafsson I, Lundwall A. (1987) Molecular cloning and sequence analysis of cDNA coding for the precursor of the human cysteine proteinase inhibitor cystatin C. *FEBS Lett. 216*, 229-33.

[37] Strauss M, Bartsch FO, Stollwerk J, Trstenjak M, Bohning A, Gassen HG, Machleidt W, Turk V. (1988) Chemical synthesis of a gene for human cystatin C and its expression in E. coli. *Biol Chem Hoppe Seyler. 369*, 209-18.

[38] Sotiropoulou G, Anisowicz A, Sager R. (1997) Identification, cloning, and characterization of cystatin M, a novel cysteine proteinase inhibitor, down-regulated in breast cancer. *J Biol Chem. 272*, 903-10.

[39] Ni J, Fernandez MA, Danielsson L, Chillakuru RA, Zhang J, Grubb A, Su J, Gentz R, Abrahamson M. (1998) Cystatin F is a glycosylated human low molecular weight cysteine proteinase inhibitor. *J Biol Chem. 273*, 24797-804.

[40] Hirado M, Tsunasawa S, Sakiyama F, Niinobe M, Fujii S. (1985) Complete amino acid sequence of bovine colostrum low-Mr cysteine proteinase inhibitor. *FEBS Lett. 186*, 41-5.

[41] Abrahamson M, Ritonja A, Brown MA, Grubb A, Machleidt W, Barrett AJ. (1987) Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin. *J Biol Chem. 262*, 9688-94.

[42] Arai S, Watanabe H, Kondo H, Emori Y, Abe K. (1991) Papain-inhibitory activity of oryzacystatin, a rice seed cysteine proteinase inhibitor, depends on the central Gln-Val-Val-Ala-Gly region conserved among cystatin superfamily members. *J Biochem (Tokyo). 109*, 294-8.

[43] Saitoh E, Isemura S, Sanada K, Ohnishi K. (1991) Cystatins of family II are harboring two domains which retain inhibitory activities against the proteinases. *Biochem Biophys Res Commun. 175*, 1070-5.

[44] Hall A, Dalboge H, Grubb A, Abrahamson M. (1993) Importance of the evolutionarily conserved glycine residue in the N-terminal region of human cystatin C (Gly-11) for cysteine endopeptidase inhibition. *Biochem J. 291*, 123-9.

[45] Hall A, Hakansson K, Mason RW, Grubb A, Abrahamson M. (1995) Structural basis for the biological specificity of cystatin C. Identification of leucine 9 in the N-terminal binding region as a selectivity-conferring residue in the inhibition of mammalian cysteine peptidases. *J Biol Chem. 270*, 5115-21.

[46] Yasuhara O, Hanai K, Ohkubo I, Sasaki M, McGeer PL, Kimura H. (1993) Expression of cystatin C in rat, monkey and human brains. *Brain Res. 628*, 85-92.

[47] Salvesen G, Parkes C, Abrahamson M, Grubb A, Barrett AJ. (1986) Human low-Mr kininogen contains three copies of a cystatin sequence that are divergent in structure and in inhibitory activity for cysteine proteinases. *Biochem J. 234*, 429-34.

[48] Pierre P, Mellman I. (1998) Developmental regulation of invariant chain proteolysis controls MHC class II trafficking in mouse dendritic cells. *Cell. 93*, 1135-45.

[49] Nishio C, Yoshida K, Nishiyama K, Hatanaka H, Yamada M. (2000) Involvement of cystatin C in oxidative stress-induced apoptosis of cultured rat CNS neurons. *Brain Res. 873*, 252-62.

[50] Afonso S, Tovar C, Romagnano L, Babiarz B. (2002) Control and expression of cystatin C by mouse decidual cultures. *Mol Reprod Dev. 61*, 155-63.

[51] Vattemi G, King Engel W, McFerrin J, Askanas V. (2003) Cystatin C colocalizes with amyloid-beta and coimmunoprecipitates with amyloid-beta precursor protein in sporadic inclusion-body myositis muscles. *J Neurochem. 85*, 1539-46.

[52] Laber B, Krieglstein K, Henschen A, Kos J, Turk V, Huber R, Bode W. (1989) The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein. *FEBS Lett. 248*, 162-8.

[53] Machleidt W, Thiele U, Laber B, Assfalg-Machleidt I, Esterl A, Wiegand G, Kos J, Turk V, Bode W. (1989) Mechanism of inhibition of papain by chicken egg white cystatin. Inhibition constants of N-terminally truncated forms and cyanogen bromide fragments of the inhibitor. *FEBS Lett. 243*, 234-8.

[54] Dieckmann T, Mitschang L, Hofmann M, Kos J, Turk V, Auerswald EA, Jaenicke R, Oschkinat H. (1993) The structures of native phosphorylated chicken cystatin and of a recombinant unphosphorylated variant in solution. *J Mol Biol. 234*, 1048-59.

[55] Turk B, Turk V, Turk D. (1977) Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. *Biol Chem. 378*, 141-50.

[56] Bromme D, Rinne R, Kirschke H. (1991) Tight-binding inhibition of cathepsin S by cystatins. *Biomed Biochim Acta. 50*, 631-5.

[57] Muller-Esterl W, Fritz H, Kellermann J, Lottspeich F, Machleidt W, Turk V. (1985) Genealogy of mammalian cysteine proteinase inhibitors. Common evolutionary origin of stefins, cystatins and kininogens. *FEBS Lett. 191*, 221-6.

Chapter II

THE CYSTATIN SUPERFAMILY

ABSTRACT

A family of proteins refers to a group of proteins that share high degree of homology in their protein and gene sequences, and by inference similar structures and functions. A family of proteins originates from a common ancestor gene. It is generally named after it most representative prototype. The cysteine proteinase inhibitors (CPIs), stefins, cystatins and kininogens, are inhibitors of cysteine proteases of the papain superfamily. The papain superfamily consists of enzymes that share high degree of sequence homology with papain; they originate from a common ancestor gene. It is named after papain, the cysteine protease originally isolated from the latex of *Carica papaya* fruit. CPIs, of stefin, cystatin and kininogen families share high degree of sequence homology. They define the cystatin family or superfamily, as it encompasses the three families of cystatins, stefin, cystatin and kininogen familes. The cystatin superfamily is named after chicken cystatin (cC); the first isolated and characterized CPI of cysteine proteases of the papain superfamily. The stefins, cystatins and kininogens originate from a common ancestor gene. The stefins elicit one region with homology to the cystatins and kininogens. The stefins probably represents the archetype of the cystatin superfamily, and a common evolutionary origin for the cystatins and kininogens. CPIs of the cystatin superfamily are inhibitors of cysteine proteases of the papain superfamily.

INTRODUCTION

Proteins with high degree of sequence homology are believed to originate from a common ancestor gene. They are classified in family of proteins; proteins that share similar structures and functions. Families of proteins are named after their most representative member.

Cysteine proteases are proteases that involve cysteine residues in the proteolytic reaction [1]. There are at least 17 families of cysteine proteases, like the family C1 or papain superfamily. The papain superfamily consists of enzymes that share high degree of sequence homology with papain [2]. Papain is the cysteine protease originally isolated from the latex of

Carica papaya fruit. Members of the papain superfamily include the mammalian cathepsins B, H, K, L and S. Papain and cysteine proteases of the C1 family originate from a common ancestor gene [3].

Cysteine proteinase inhibitors (CPIs) are inhibitors of cysteine proteases. The stefins, cystatins and kininogens are CPIs of the family C1 or papain superfamily of cysteine proteases [4-6]. Sequence analysis of the stefins, cystatins and kininogens indicates that they elicit high degree of homology with chicken cystatin (cC) [7]. Chicken cystatin is the first cysteine proteinase inhibitor isolated and characterized [8]. It was isolated and purified from chicken egg white. The proteinase inhibitors of the papain superfamily belong to one family of proteins, the cystatin superfamily. The cystatin superfamily encompasses the stefins, cystatins and kininogens, cysteine proteinase inhibitors that originate from a common ancestor gene [9,10].

THE CYSTATIN SUPERFAMILY

Chicken cystatin is the first cysteine proteinase inhibitor isolated and characterized [8]. It was isolated and purified from chicken egg white. Its protein and cDNA have been sequenced [11-13].

In human, CPIs with sequence homology to cC have been characterized, like the stefins A and B, the cystatins C, D, E/M, F, G, S, SN and SA, and the kininogens [14-21]. Human cystatin C (hCC) shares with cC protein 41% sequence identity [15]. Human cystatin C has 50-55% identical residues with cystatin D and the secretory gland cystatin, cystatins S, SN and Sa [9,22]. Cystatin E/M and cystatin F (also called leukocystatin) are glycoproteins and show only 30-35% sequence identity in alignments with classical type cystatins [20,21]. The stefins elicit one region with homology to the cystatins and kininogens. The L-kininogen contains three cystatin-like domains [22].

In all, the stefins, cystatins and kininogens share high degree of sequence homology with cC. They form the cystatin superfamily, named after cC, the most characterized CPI of the cystatin superfamily [9,10].

The cystatin superfamily is subdivided in three families, family 1 or stefin family, family 2 or cystatin family and family 3 or kininogen family, based on their structures and named after their most representative member [23].

- The stefin family. The stefins are low molecular weight (10-13 kDa) intracellular CPIs, with no disulfide bond. Their prototypes are the human stefins [22].
- The cystatin family. The cystatins are low molecular weight (13-14 kDa) secreted CPIs, with two disulfide bonds. Their prototype is cC [22].
- The kininogen family. The kininogens are high molecular weight intravascular CPIs. They contain disulfide bonds and their prototypes are the plasma kininogens [24].

The CPIs of the cystatin superfamily originate evolutionary from a common ancestor gene [9,10].

PHYLOGENY

The stefins elicit one region with homology to the cystatins and kininogens. The L-kininogen contains three cystatin-like domains. Among the three domains of the N-terminal heavy chain of kininogens, only 2 have the conserved consensus sequence QVVAG, motif corresponding to an active site of cC, and are functional proteinase inhibitor sites [24].

Based on analysis of sequence homology, it is proposed that the ancestor gene from which derive the family 1 or stefin family corresponds to the archetype of the CPIs of the cystatin superfamily [9,10,22]. The diversity of the CPIs of the cystatin superfamily in three families originates from the evolution and duplication of the ancestor gene. In this model, the cystatins have acquired by gene fusion a second element, from a second ancestor. The kininogen heavy chain or low molecular weight kininogen (L-kininogen) consists of three cystatin-like domains. It originates from the triplication of this latter cystatin prototype gene [9,10].

In all, the three CPI families of the cystatin superfamily originate from a common ancestor gene by gene fusion and duplication. Other models have been proposed that present different theories with regard to the phylogenic evolution of the cystatin superfamily from a common ancestor gene [9,10,22]. Further analysis of these different models may lead to a better understanding of the evolutionary relationships of the cystatin superfamily.

CHROMOSOME MAPPING

The genes of CPIs of the cystatin superfamily have been sequenced and their location on chromosomes mapped.

The human cystatin C gene (CST3) is approximately 4.3 kb long with 3 exons and 2 introns [25]. It has the same overall structure as the genes of other cystatins, like cystatin SN (CTS1), SA (CST2), S (CST4) and D (CST5). These genes belong to a multigene family called the cystatin gene family [26].

The human genes of the cystatin gene family, CST1, CST2, CST3, CST4 and CST5, are clustered on chromosome 20. They are mapped to an approximately 1.2-Mb segment, on the region 20p11.2 [27-32]. Genes for other members of the cystatin superfamily, stefins and kininogens, have been mapped to chromosome 3 [33-35].

In all, in human, the cystatins are all products of genes located in the cystatin multigene locus on chromosome 20, whereas the stefin and kininogen genes are not syntenic genes of family 2 or cystatin family [36,37].

Human chromosome 20 and murine chromosome 2 share strong conserved homology; murine homologs of human genes on chromosome 20 have been mapped on the distal segment of mouse chromosome 2. Accordingly, it is presumed that murine homolog's of CST3, CST5, and probably of other members of the cystatin gene family are located on mouse chromosome 2, between the loci for Bmp-2a and Hck-1 [36,37].

AMINO ACIDS AND SEQUENCES CONSERVED OF THE CYSTATIN SUPERFAMILY

The cystatin superfamily elicits highly conserved amino acids and sequences throughout its members.

Among them:

- the N-terminal motif, the sequence Arg8-Leu9-Val10-Gly11 or RLVG sequence (hCC numbering), is highly conserved among the members of the cystatin superfamily [38]. This sequence is involved in the CPI activity of cystatins. In this sequence, the amino acid Gly11 is highly conserved in all of the inhibitory domains of the cystatin superfamily, except the first domain of the kininogens, which lacks inhibitory activity [39]. Gly11 introduce some flexibility in the N-terminal region of the inhibitors, allowing the N-terminal motif to adopt an optimal conformation for enzyme-inhibitor interaction,

- the amino acid Asn39 (hCC numbering) is highly conserved,
- the motif Gln55-Ile56-Val57-Ala58-Gly59 (hCC numbering) or QIVAG sequence is highly conserved among the members of the cystatin superfamily. In this sequence, the residues Gln55, Val57 and Gly59 are highly conserved, identifying the motif Gln-Xaa-Val-Xaa-Gly, as the conserved sequence. The corresponding sequence in cC is Gln-Leu-Val-Ser-Gly or QLVSG sequence [40]. The corresponding sequence in stefin and kininogen is Gln-Val-Val-Ala-Gly or QVVAG sequence. This sequence is involved in the CPI activity of these cystatins. This motif is absent in the first domain of kininogens [24],
- the sequence 72 to 92 (hCC numbering) is highly conserved among family 2 inhibitors. This sequence is absent in the family 1 (stefin) inhibitors [28],
- the sequence Pro105-Trp106 (hCC numbering) is involved in the CPI activity of cystatins [41].

CONCLUSION

The protein and cDNA sequences of the stefins, cystatins and kininogens elicit high degree of homology. They belong to the same family of proteins, the cystatin superfamily. This family is named after cC, the first inhibitor of cysteine protease of the papain superfamily isolated and also the most characterized. The cystatin superfamily encompasses the stefin, cystatin and kininogen families whose members originate from a common ancestor gene. Gene fusion and duplication from this common ancestor gene yielded the variety of the cystatin superfamily.

The family of cystatin genes, composed of genes that elicit the same overall structure, are all located in the cystatin multigene locus on chromosome 20 in humans, whereas the genes of stefins and kininogens have been mapped to chromosome 3. In murine, the cystatin gene family is located on mouse chromosome 2, which share strong conserved homology to human chromosome 20.

The cystatin superfamily elicits highly conserved amino acids and sequences throughout its members. The conserved regions underlie structural and functional properties of the cystatin superfamily.

REFERENCES

[1] Rzychon M, Chmiel D, Stec-Niemczyk J. (2004) Modes of inhibition of cysteine proteases. *Acta Biochim Pol.51*, 861-73.

[2] Nagler DK, Menard R. (2003) Family C1 cysteine proteases: biological diversity or redundancy? *Biol Chem. 384*, 837-43.

[3] Berti PJ, Storer AC. (1995) Alignment/phylogeny of the papain superfamily of cysteine proteases. *J Mol Biol. 246*, 273-83.

[4] Turk B, Turk V, Turk D. (1977) Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. *Biol Chem. 378*, 141-50.

[5] Turk V, Bode W. (1991) The cystatins: protein inhibitors of cysteine proteinases. *FEBS Lett. 285*, 213-9.

[6] Abrahamson M, Alvarez-Fernandez M, Nathanson CM. Cystatins. (2003) *Biochem Soc Symp. 70*, 179-99.

[7] Schwabe C, Anastasi A, Crow H, McDonald JK, Barrett AJ. (1984) Cystatin. Amino acid sequence and possible secondary structure. *Biochem J. 217*, 813-7.

[8] Sen LC, Whitaker JR. (1973) Some properties of a ficin-papain inhibitor from avian egg white. *Arch Biochem Biophys. 158*, 623-32.

[9] Muller-Esterl W, Fritz H, Kellermann J, Lottspeich F, Machleidt W, Turk V. (1985) Genealogy of mammalian cysteine proteinase inhibitors. Common evolutionary origin of stefins, cystatins and kininogens. *FEBS Lett. 191*, 221-6.

[10] Rawlings ND, Barrett AJ. (1990) Evolution of proteins of the cystatin superfamily. *J Mol Evol. 30*, 60-71.

[11] Anastasi A, Brown MA, Kembhavi AA, Nicklin MJ, Sayers CA, Sunter DC, Barrett AJ. (1983) Cystatin, a protein inhibitor of cysteine proteinases. Improved purification from egg white, characterization, and detection in chicken serum. *Biochem J. 211*, 129-38.

[12] Turk V, Brzin J, Longer M, Ritonja A, Eropkin M, Borchart U, Machleidt W. (1983) Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin from chicken egg white. *Hoppe Seylers Z Physiol Chem. 364*, 1487-96.

[13] Auerswald EA, Genenger G, Mentele R, Lenzen S, Assfalg-Machleidt I, Mitschang L, Oschkinat H, Fritz H. (1991) Purification and characterization of a chicken egg white cystatin variant expressed in an Escherichia coli pIN-III-ompA system. *Eur J Biochem. 200*, 131-8.

[14] Brzin J, Popovic T, Turk V, Borchart U, Machleidt W. (1984) Human cystatin, a new protein inhibitor of cysteine proteinases. *Biochem Biophys Res Commun. 118*, 103-9.

[15] Abrahamson M, Grubb A, Olafsson I, Lundwall A. (1987) Molecular cloning and sequence analysis of cDNA coding for the precursor of the human cysteine proteinase inhibitor cystatin C. *FEBS Lett. 216*, 229-33.

[16] Strauss M, Bartsch FO, Stollwerk J, Trstenjak M, Bohning A, Gassen HG, Machleidt W, Turk V. (1988) Chemical synthesis of a gene for human cystatin C and its expression in E. coli. *Biol Chem Hoppe Seyler*. 369, 209-18.

[17] Isemura S, Saitoh E, Sanada K, Minakata K. (1991) Identification of full-sized forms of salivary (S-type) cystatins (cystatin SN, cystatin SA, cystatin S, and two phosphorylated forms of cystatin S) in human whole saliva and determination of phosphorylation sites of cystatin S. *J Biochem (Tokyo).110*, 648-54.

[18] Ni J, Abrahamson M, Zhang M, Fernandez MA, Grubb A, Su J, Yu GL, Li Y, Parmelee D, Xing L, Coleman TA, Gentz S, Thotakura R, Nguyen N, Hesselberg M, Gentz R. (1997) Cystatin E is a novel human cysteine proteinase inhibitor with structural resemblance to family 2 cystatins. *J Biol Chem. 272*, 10853-8.

[19] Sotiropoulou G, Anisowicz A, Sager R. (1997) Identification, cloning, and characterization of cystatin M, a novel cysteine proteinase inhibitor, down-regulated in breast cancer. *J Biol Chem. 272*, 903-10.

[20] Halfon S, Ford J, Foster J, Dowling L, Lucian L, Sterling M, Xu Y, Weiss M, Ikeda M, Liggett D, Helms A, Caux C, Lebecque S, Hannum C, Menon S, McClanahan T, Gorman D, Zurawski G. (1998) Leukocystatin, a new Class II cystatin expressed selectively by hematopoietic cells. *J Biol Chem. 273*, 16400-8.

[21] Ni J, Fernandez MA, Danielsson L, Chillakuru RA, Zhang J, Grubb A, Su J, Gentz R, Abrahamson M. (1998) Cystatin F is a glycosylated human low molecular weight cysteine proteinase inhibitor. *J Biol Chem. 273*, 24797-804.

[22] Sali A, Turk V. (1987) Prediction of the secondary structures of stefins and cystatins, the low-molecular mass protein inhibitors of cysteine proteinases. *Biol Chem Hoppe Seyler. 368*, 493-9.

[23] Barrett AJ, Fritz H, Grubb A, Isemura S, Jarvinen M, Katunuma N, Machleidt W, Muller-Esterl W, Sasaki M, Turk V. (1986) Nomenclature and classification of the proteins homologous with the cysteine-proteinase inhibitor chicken cystatin. *Biochem J. 236*, 312.

[24] Salvesen G, Parkes C, Abrahamson M, Grubb A, Barrett AJ. (1986) Human low-Mr kininogen contains three copies of a cystatin sequence that are divergent in structure and in inhibitory activity for cysteine proteinases. *Biochem J. 234*, 429-34.

[25] Abrahamson M, Olafsson I, Palsdottir A, Ulvsback M, Lundwall A, Jensson O, Grubb A. (1990) Structure and expression of the human cystatin C gene. *Biochem J. 268*, 287-94.

[26] Saitoh E, Isemura S, Sanada K, Ohnishi K. (1991) The human cystatin gene family: cloning of three members and evolutionary relationship between cystatins and Bowman-Birk type proteinase inhibitors. *Biomed Biochim Acta. 50*, 599-605.

[27] Abrahamson M, Islam MQ, Szpirer J, Szpirer C, Levan G. (1989) The human cystatin C gene (CST3), mutated in hereditary cystatin C amyloid angiopathy, is located on chromosome 20. *Hum Genet. 82*, 223-6.

[28] Saitoh E, Sabatini LM, Eddy RL, Shows TB, Azen EA, Isemura S, Sanada K. (1989) The human cystatin C gene (CST3) is a member of the cystatin gene family which is localized on chromosome 20. *Biochem Biophys Res Commun. 162*, 1324-31.

[29] Freije JP, Abrahamson M, Olafsson I, Velasco G, Grubb A, Lopez-Otin C. (1991) Structure and expression of the gene encoding cystatin D, a novel human cysteine proteinase inhibitor. *J Biol Chem. 266*, 20538-43.

[30] Saitoh E, Isemura S, Sanada K, Ohnishi K. (1991) The human cystatin gene family: cloning of three members and evolutionary relationship between cystatins and Bowman-Birk type proteinase inhibitors. *Biomed Biochim Acta. 50*, 599-605.

[31] Freije JP, Pendas AM, Velasco G, Roca A, Abrahamson M, Lopez-Otin C. (1993) Localization of the human cystatin D gene (CST5) to chromosome 20p11.21 by in situ hybridization. *Cytogenet Cell Genet. 62*, 29-31.

[32] Schnittger S, Rao VV, Abrahamson M, Hansmann I. (1993) Cystatin C (CST3), the candidate gene for hereditary cystatin C amyloid angiopathy (HCCAA), and other members of the cystatin gene family are clustered on chromosome 20p11.2. Genomics. 16, 50-5. Erratum in: (1993) *Genomics 17*, 533.

[33] van den Berg EA, le Clercq E, Kluft C, Koide T, van der Zee A, Oldenburg M, Wijnen JT, Meera Khan P. (1990) Assignment of the human gene for histidine-rich glycoprotein to chromosome 3. *Genomics. 7*, 276-9.

[34] Fong D, Smith DI, Hsieh WT. (1991) The human kininogen gene (KNG) mapped to chromosome 3q26-qter by analysis of somatic cell hybrids using the polymerase chain reaction. *Hum Genet. 87*, 189-92.

[35] Hsieh WT, Fong D, Sloane BF, Golembieski W, Smith DI. (1991) Mapping of the gene for human cysteine proteinase inhibitor stefin A, STF1, to chromosome 3cen-q21. *Genomics. 9*, 207-9.

[36] Seidah NG, Mattei MG, Gaspar L, Benjannet S, Mbikay M, Chretien M. (1991) Chromosomal assignments of the genes for neuroendocrine convertase PC1 (NEC1) to human 5q15-21, neuroendocrine convertase PC2 (NEC2) to human 20p11.1-11.2, and furin (mouse 7[D1-E2] region). *Genomics. 11*, 103-7.

[37] Schnittger S, Rao VV, Deutsch U, Gruss P, Balling R, Hansmann I. (1992) Pax1, a member of the paired box-containing class of developmental control genes, is mapped to human chromosome 20p11.2 by in situ hybridization (ISH and FISH). *Genomics. 14*, 740-4.

[38] Abrahamson M, Ritonja A, Brown MA, Grubb A, Machleidt W, Barrett AJ. (1987) Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin. *J Biol Chem. 262*, 9688-94.

[39] Hall A, Dalboge H, Grubb A, Abrahamson M. (1993) Importance of the evolutionarily conserved glycine residue in the N-terminal region of human cystatin C (Gly-11) for cysteine endopeptidase inhibition. *Biochem J. 291*, 123-9.

[40] Arai S, Watanabe H, Kondo H, Emori Y, Abe K. (1991) Papain-inhibitory activity of oryzacystatin, a rice seed cysteine proteinase inhibitor, depends on the central Gln-Val-Val-Ala-Gly region conserved among cystatin superfamily members. *J Biochem (Tokyo). 109*, 294-8.

[41] Saitoh E, Isemura S, Sanada K, Ohnishi K. (1991) Cystatins of family II are harboring two domains which retain inhibitory activities against the proteinases. *Biochem Biophys Res Commun. 175*, 1070-5.

Chapter III

STRUCTURE: CHICKEN CYSTATIN AND HUMAN CYSTATIN C

ABSTRACT

Chicken cystatin (cC) and human cystatin C (hCC) are low molecular weight cysteine protease inhibitors (CPIs) of the cystatin superfamily. They belong to the family 2 or cystatin family. Chicken cystatin and hCC share high degree of sequence homology and originate from a common ancestor gene. As such, they share similar structures and functions. They have a molecular weight in the range of 13-14 kDa, a single polypeptide chain of approximately 120 amino acid residues and two (intrachain) disulfide bonds. Chicken cystatin and hCC are natural tight-binding reversible inhibitors of cysteine proteases of the papain superfamily, including cathepsins B, H, J, K, L, S and X. Chicken cystatin is the most characterized CPIs of the cystatin superfamily. It has been crystallized and its three-dimensional structure determined. The crystal structure of cC forms a wedge-shaped structure. The determination of the crystal structure of cC provides a model for understanding the properties and functions of cC, hCC, but also of other CPIs of the cystatin superfamily.

INTRODUCTION

Chicken cystatin is the first member of the cystatin superfamily isolated and characterized; it was isolated and purified from chicken egg white [1,2]. It belongs to family 2 or cystatin family. Chicken cystatin is a single chain of 116 residues, containing two disulfide bonds [3]. The disulfide bonds are formed in the C-terminal region of the molecule, between Cys73 and Cys83, and between Cys97 and Cys117 (hCC numbering). Chicken cystatin occurs in a phosphorylated and non-phosphorylated form, and is not glycosylated [3,4]. Chicken cystatin is the most characterized member of the cystatin superfamily [5].

Human cystatin C also belongs to family 2 or cystatin family. It is a single polypeptide chain of 120 amino acids residues [6]. It has a molecular weight of 13.4 kDa [7]. It comprises two disulfide bonds and is not glycosylated. The gene coding for hCC was sequenced and

found to be analogous to a previously isolated and purified protein of unknown function, present in cerebrospinal fluid and urine, γ-trace [8-12]. Three sequences of the polypeptide chain of hCC have been identified as involved in the CPI activity of the protein. They are the N-terminal sequence Arg8-Leu9-Val10-Gly11, the sequences Gln55-Ile56-Val57-Ala58-Gly59 and Pro105-Trp106 [13-17].

Chicken cystatin and hCC belong to the cystatin superfamily. They share a high degree of sequence homology and originate from a common ancestor gene [18,19]. Human cystatin C shares with cC protein 41% sequence identity [6]. By inference, they share similar structures and functions.

X-RAY CRYSTAL STRUCTURE

Wedge-Shaped Structure

Chicken cystatin has been crystallized and its three-dimensional structure determined [20]. The crystal structure of cC is primarily composed of five-stranded β-sheets (β1 to 5), a central α-helix (α1) and an appended segment (AS), with a partial α-helical geometry.

In the protein, the five β-sheets, α-helix and AS are in the following sequence and organization, from the N-terminal to the C-terminal region: (N)-β1-(α1)-β2-L1-β3-(AS)-β4-L2-β5-(C). The five-stranded β-sheets, the α-helix and AS are linked by two segments, referred as loop 1 and 2 (L1 and L2). The five-stranded β-sheets are antiparallel and twisted. The α1 helix runs across the five-stranded β-sheets. The AS is positioned on the opposite side of the β1-sheet relative. The loop L1 is located at the edge of the crystal structure.

These data show that the overall crystal structure of cC is a five-stranded antiparallel β-sheet stacked against and partially wrapped around a central α-helix, forming a structure referred as a wedge-shaped structure [20,21].

The crystal structure of hCC has been determined; it consists of a five-stranded antiparallel β-sheet, wrapped around a α-helix, like in cC crystalline form [20,22].

The crystal structures of other members of the cystatin superfamily have been determined, like the crystal structure of stefin B [23]. Stefins are CPIs of family 1 or stefin family [24,25]. Analysis from these structures reveals the same features the wedge-shaped structure, reported for cC and hCC [20].

In all, members of the cystatin superfamily elicit the same crystal structure, a five-stranded antiparallel β-sheet stacked against and partially wrapped around a central α-helix, referred as a wedge-shaped structure.

Tripartite Wedge-Shaped Binding Region

The three sequences of cystatins involved in the cysteine proteinase inhibitory site of hCC, the N-terminal sequence, the sequence Arg8-Leu9-Val10-Gly11, the sequences Gln55-Ile56-Val57-Ala58-Gly59 and Pro105-Trp106, have been mapped in the three dimensional structure of cystatin [20,21].

The N-terminal segment of cystatin, containing the sequence Arg8-Leu9-Val10-Gly11, is on the N-terminal β-sheet (β1), the motif Gln-Xaa-Val-Xaa-Gly is in a central loop L1 and the motif Pro105-Trp106 is in the second C-terminal hairpin loop, L2 [26].

In all, the conserved motifs involved in binding of cystatins are located in adjacent turn structures along the edge of the molecule. This distribution of the sites involved in the inhibitor activity along the edge of the molecule is referred as a tripartite wedge-shaped binding region [26]. The distribution of the sites involved in the inhibitor activity along the edge of the molecule provides molecular surface complementary to the cysteine proteinase active site cleft.

PROPERTIES

Disulfide Bonds

Cystatins contains four Cys residues forming two characteristics disulfide bonds, one between Cys73 and Cys83, the other between Cys97 and Cys117 (hCC numbering). Both disulfide bonds are in the C-terminal region of cystatins [24,25].

The two disulfide bonds introduce rigidity into the molecule. The disulfide bond between Cys73 and Cys83, stabilizes the structure of the AS segment between the strands β3 and β4, whereas the disulfide bond between Cys97 and Cys117 connects the ends of the β4-β5 hairpin loop L2 [27].

Unstructured Region Pro72 to Ser80

Analysis of the crystal structure of cC reveals an unstructured flexible region Pro72 to Ser80 (cC numbering) [20]. Whether this unstructured region originates from an artifact of crystallization or corresponds to a physiological state of the molecule remains to be determined. If it is a physiological state of the molecule, its function and significance remain to be established.

Phosphorylation

Chicken cystatin occurs in a phosphorylated and non- phosphorylated form [28]. The site of phosphorylation in cC has been identified has serine S80, which acts as a recognition site for the serine kinase. The effect of phosphorylation on the conformation of cystatin remains to be determined.

The residue Ser80 is localized in the unstructured region of cystatin Pro72-Ser80, as observed in the crystal structure. The phosphorylation of cystatin therefore occurs in the flexible region of the molecule. This suggests that the effect of phosphorylation on the conformation may not cause large structural effect [29]. The effect of phosphorylation may also depend on the environment surrounding the residue Ser80. The conformational change

induced by phosphorylation of cystatin, its function and significance remain to be determined.

CONCLUSION

Chicken cystatin and hCC belong to the family 2 of the cystatin superfamily. Chicken cystatin and hCC share high degree of protein and DNA sequence homology, and originate from a common ancestor gene. As such, they share similar structures and functions.

Results from X-ray crystallography reveal that both proteins elicit the same three dimensional structure, a five-stranded antiparallel β-sheet stacked against and partially wrapped around a central α-helix, forming a structure referred as a wedge-shaped structure. The structure of cC and other cystatins has been determined by nuclear magnetic resonance spectroscopy, in solution, and supports the same basic folding for cystatins [26,30].

The crystal structure of cC reveals an unstructured region, Pro72-Ser80. Whether this unstructured region corresponds to an experimental artifact or a physiological state of the molecule remains to be determined. A physiological state would mean a correlation between conformational change and activity. The phosphorylation of cystatin occurs in this unstructured region (on residue Ser80) and the conformational consequence of phosphorylation remains also to be fully evaluated.

The X-ray crystallography has also been determined for the family 1 cystatins, stefins A and B, and reveals the same 3-dimensional structure as for cystatins. Thus, the structure referred a wedge-shaped structure, is the prototype structure of the cystatin superfamily. The structure of cystatins supports a common mechanism for the cystatin superfamily.

REFERENCES

[1] Sen LC, Whitaker JR. (1973) Some properties of a ficin-papain inhibitor from avian egg white. *Arch Biochem Biophys. 158*, 623-32.

[2] Anastasi A, Brown MA, Kembhavi AA, Nicklin MJ, Sayers CA, Sunter DC, Barrett AJ. (1983) Cystatin, a protein inhibitor of cysteine proteinases. Improved purification from egg white, characterization, and detection in chicken serum. *Biochem J. 211*, 129-38.

[3] Turk V, Brzin J, Longer M, Ritonja A, Eropkin M, Borchart U, Machleidt W. (1983) Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin from chicken egg white. *Hoppe Seylers Z Physiol Chem. 364*, 1487-96.

[4] Laber B, Krieglstein K, Henschen A, Kos J, Turk V, Huber R, Bode W. (1989) The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein. *FEBS Lett. 248*, 162-8.

[5] Auerswald EA, Genenger G, Mentele R, Lenzen S, Assfalg-Machleidt I, Mitschang L, Oschkinat H, Fritz H. (1991) Purification and characterization of a chicken egg white cystatin variant expressed in an Escherichia coli pIN-III-ompA system. *Eur J Biochem. 200*, 131-8.

[6] Abrahamson M, Grubb A, Olafsson I, Lundwall A. (1987) Molecular cloning and sequence analysis of cDNA coding for the precursor of the human cysteine proteinase inhibitor cystatin C. *FEBS Lett. 216*, 229-33.

[7] Strauss M, Bartsch FO, Stollwerk J, Trstenjak M, Bohning A, Gassen HG, Machleidt W, Turk V. (1988) Chemical synthesis of a gene for human cystatin C and its expression in E. coli. *Biol Chem Hoppe Seyler. 369*, 209-18.

[8] Tonnelle C, Colle A, Fougereau M, Manuel Y. (1979) Partial amino acid sequence of two forms of human post-gamma-globulin. *Biochem Biophys Res Commun. 86*, 613-9.

[9] Grubb A, Lofberg H. (1982) Human gamma-trace, a basic microprotein: amino acid sequence and presence in the adenohypophysis. *Proc Natl Acad Sci U S A. 79*, 3024-7.

[10] Barrett AJ, Davies ME, Grubb A. (1984) The place of human gamma-trace (cystatin C) amongst the cysteine proteinase inhibitors. *Biochem Biophys Res Commun. 120*, 631-6.

[11] Brzin J, Popovic T, Turk V, Borchart U, Machleidt W. (1984) Human cystatin, a new protein inhibitor of cysteine proteinases. *Biochem Biophys Res Commun. 118*, 103-9.

[12] Abrahamson M, Olafsson I, Palsdottir A, Ulvsback M, Lundwall A, Jensson O, Grubb A. (1990) Structure and expression of the human cystatin C gene. *Biochem J. 268*, 287-94.

[13] Abrahamson M, Ritonja A, Brown MA, Grubb A, Machleidt W, Barrett AJ. (1987) Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin. *J Biol Chem. 262*, 9688-94.

[14] Arai S, Watanabe H, Kondo H, Emori Y, Abe K. (1991) Papain-inhibitory activity of oryzacystatin, a rice seed cysteine proteinase inhibitor, depends on the central Gln-Val-Val-Ala-Gly region conserved among cystatin superfamily members. *J Biochem (Tokyo). 109*, 294-8.

[15] Saitoh E, Isemura S, Sanada K, Ohnishi K. (1991) Cystatins of family II are harboring two domains which retain inhibitory activities against the proteinases. *Biochem Biophys Res Commun. 175*, 1070-5.

[16] Hall A, Dalboge H, Grubb A, Abrahamson M. (1993) Importance of the evolutionarily conserved glycine residue in the N-terminal region of human cystatin C (Gly-11) for cysteine endopeptidase inhibition. *Biochem J. 291*, 123-9.

[17] Hall A, Hakansson K, Mason RW, Grubb A, Abrahamson M. (1995) Structural basis for the biological specificity of cystatin C. Identification of leucine 9 in the N-terminal binding region as a selectivity-conferring residue in the inhibition of mammalian cysteine peptidases. *J Biol Chem. 270*, 5115-21.

[18] Muller-Esterl W, Fritz H, Kellermann J, Lottspeich F, Machleidt W, Turk V. (1985) Genealogy of mammalian cysteine proteinase inhibitors. Common evolutionary origin of stefins, cystatins and kininogens. *FEBS Lett. 191*, 221-6.

[19] Rawlings ND, Barrett AJ. (1990) Evolution of proteins of the cystatin superfamily. *J Mol Evol. 30*, 60-71.

[20] Bode W, Engh R, Musil D, Thiele U, Huber R, Karshikov A, Brzin J, Kos J, Turk V. (1988) The 2.0 A X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J. 7*, 2593-9.

[21] Schwabe C, Anastasi A, Crow H, McDonald JK, Barrett AJ. (1984) Cystatin. Amino acid sequence and possible secondary structure. *Biochem J. 217*, 813-7.

[22] Kozak M, Jankowska E, Janowski R, Grzonka Z, Grubb A, Alvarez Fernandez M, Abrahamson M, Jaskolski M. (1999) Expression of a selenomethionyl derivative and preliminary crystallographic studies of human cystatin C. *Acta Crystallogr D Biol Crystallogr*. *55*, 1939-42.

[23] Stubbs MT, Laber B, Bode W, Huber R, Jerala R, Lenarcic B, Turk V. (1990) The refined 2.4 A X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *EMBO J. 9*, 1939-47.

[24] Turk V, Bode W. (1991) The cystatins: protein inhibitors of cysteine proteinases. *FEBS Lett. 285*, 213-9.

[25] Abrahamson M, Alvarez-Fernandez M, Nathanson CM. Cystatins. (2003) *Biochem Soc Symp*. *70*, 179-99.

[26] Engh RA, Dieckmann T, Bode W, Auerswald EA, Turk V, Huber R, Oschkinat H. (1993) Conformational variability of chicken cystatin. Comparison of structures determined by X-ray diffraction and NMR spectroscopy. *J Mol Biol. 234*, 1060-9.

[27] Bjork I, Ylinenjarvi K. (1992) Different roles of the two disulfide bonds of the cysteine proteinase inhibitor, chicken cystatin, for the conformation of the active protein. *Biochemistry*. *31*, 8597-602.

[28] Laber B, Krieglstein K, Henschen A, Kos J, Turk V, Huber R, Bode W. (1989) The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein. *FEBS Lett. 248*, 162-8.

[29] Dieckmann T, Mitschang L, Hofmann M, Kos J, Turk V, Auerswald EA, Jaenicke R, Oschkinat H. (1993) The structures of native phosphorylated chicken cystatin and of a recombinant unphosphorylated variant in solution. *J Mol Biol. 234*, 1048-59.

[30] Ekiel I, Abrahamson M, Fulton DB, Lindahl P, Storer AC, Levadoux W, Lafrance M, Labelle S, Pomerleau Y, Groleau D, LeSauteur L, Gehring K. (1997) NMR structural studies of human cystatin C dimers and monomers. *J Mol Biol. 271*, 266-77.

Chapter IV

CYSTEINE PROTEINASE INHIBITORS

ABSTRACT

Cysteine proteinase inhibitors (CPIs) of the cystatin superfamily are inhibitors of cysteine proteases of the C1 family or papain superfamily. Endogenous CPIs of the cystatin superfamily are tight-binding reversible inhibitors of cysteine proteases; they bind the active site cleft of cysteine proteases, preventing substrate binding. Chicken cystatin (cC) is the prototype of family 2 or cystatin family of the cystatin superfamily, and the most characterized inhibitor of the cystatin superfamily. Molecular and enzymatic studies have led to the characterization of the proteinase inhibitory domain of cC and related inhibitors. Results from these studies reveal that the N-terminal sequence, as well as two other regions of cystatins, are involved in their proteinase inhibitory activities. X-Ray crystallography of cystatins has contributed to modelize the interaction between the inhibitors and proteases.

INTRODUCTION

Cysteine proteases (EC 3.4.22) are a wide family of proteolytic enzymes. Cysteine proteases involve cysteine residues in the proteolytic reaction [1]. In papain, the cysteine residue involved in the proteolytic reaction is Cys25 (figure 1). There are at least 17 families of cysteine proteases, including the family C1 or papain superfamily and family C13 or legumain family. Cysteine protease, or proteinase, inhibitors are inhibitors of cysteine proteases.

CPIs of cystatin superfamily are inhibitors of the family C1 of cysteine proteases, or papain superfamily. The cysteine proteases of the papain superfamily are enzymes with high sequence homology to papain [2]. The cysteine proteases of the papain superfamily originate from a common ancestor gene. Papain is a cysteine protease originally isolated from the latex of *Carica papaya* fruit. The papain superfamily of cysteine proteases are mostly endopeptidases, including cathepsins B, H, K, L and S [3]. These enzymes are lysosomal enzymes, involved in the metabolism of proteins and that operate at acid pH. The three-

dimensional structure of members of the papain superfamily has been determined by X-ray crystallography, including cathepsin B [4].

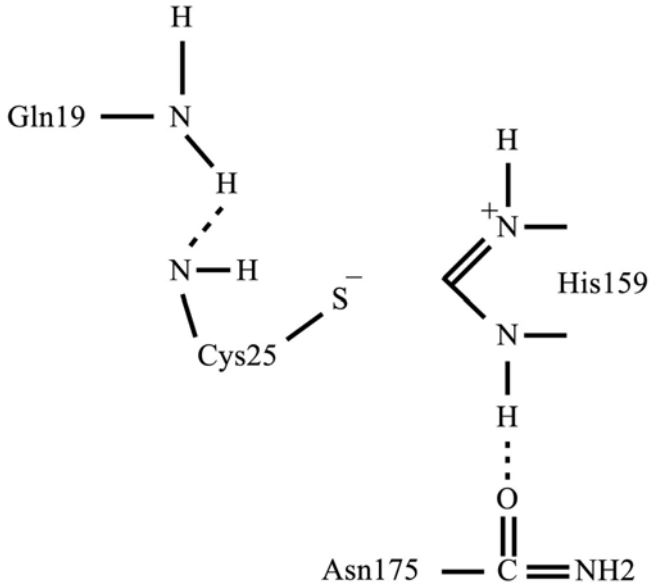


Figure 1. Catalytic site of cysteine proteases. Cysteine proteases, like papain, are endopeptidases that involve cysteine residue (thiol group) in the proteolytic reaction. The catalytic reaction is mediated by Cys and His residues in the catalytic site. In papain, Cys25 and His159 are the catalytic resides of the active site [1]. Asn175 ensure that the orientation of the imidazole of the His159 is optimal, through hydrogen bonds. Gln19 is involved during the catalytic reaction, providing stabilization through hydrogen bonds.

The CPIs of the cystatin superfamily are inhibitors with high sequence homology to chicken cystatin (cC). The CPIs of the cystatin superfamily originate from a common ancestor gene [5,6]. Chicken cystatin is the most characterized CPIs of the cystatin superfamily; it has been isolated and purified from chicken egg white [7-10]. The cystatin superfamily is divided into three families of CPIs, according to their structure and named after their most representative prototype, the family 1 or stefin family, the family 2 or cystatin family and the family 3 or kininogen family [11-13]. Studies from cC, hCC and other cystatins reveal that in hCC the N-terminal sequence, the sequence Arg8-Leu9-Val10-Gly11, as well as the sequences of hCC Gln55-Ile56-Val57-Ala58-Gly59 and Pro105-Trp106, are involved in the CPI activity of cystatins [14-18]. X-ray crystallography structure of cC and hCC has been determined; cC and hCC elicit the following structure: a five-stranded antiparallel β-sheet stacked against and partially wrapped around a central α-helix [19,20].

The affinity of CPIs of the cystatin superfamily for the proteases of the papain superfamily has been determined. Cystatin C inhibits with high affinity virtually all cysteine proteases of the papain superfamily; the constant of dissociation equilibrium of cystatin C for proteases of the papain superfamily is in the subnanomolar range [21]. The constant of affinity of cC for papain is Ki = 0.0000 14 nM [22]. The constants of affinity of cC for cathepsins are in the subnanomolar range, ranking from the highest to the lowest, such as cathepsin L = S > B > H [23-33].

Cystatin C forms 1:1 complexes with target enzymes in competition with their substrates and no detectable cleavage of the cystatin in the papain complex [34]. Kinetic studies reveal that enzyme-inhibitor complexes are being formed by a one-step, reversible non-covalent bimolecular reaction, except for cathepsin B [23-33]. The binding of cystatin C to cathepsin B is described as a two-step reaction mechanism [35].

In all, the diversity of CPIs of the cystatin superfamily, stefins, cystatins and kininogens, reflects the diversity of cysteine proteases of the papain superfamily.

It was recently reported that cC and human cystatin C (hCC) inhibit the activity of legumain [36]. Legumain (EC 3.4.22.34) is a cysteine endopeptidase that is a prototype of a distinct family of cysteine proteases than the family C1 of papain superfamily. Legumain belongs to the family C13 of cysteine proteases or legumain family. The C1 and C13 families of cysteine proteases are evolutionarily unrelated. Therefore members of their families elicit different structures, particularly different protease active sites or catalytic domains. This suggests that hCC elicits a second reactive site for legumain, distinct from of the active site for papain.

THE N-TERMINAL REGION

The aminoterminal motif of hCC, Arg8-Leu9-Val10-Gly11 or RLVG sequence, is conserved among CPIs of the cystatin superfamily [37].

Chicken cystatin is a single chain composed of 116 amino acid residues. It is typically isolated in various forms, differing by truncations at the N terminal. The long form or Ser1-form (116 residues) can be degraded during purification to shorter forms, among them, the Gly9- and Ala10-forms with 108 and 107 residues, respectively. The Gly-form (truncated of residues 1 to 8) is the most common variants of cC isolated [15,25,38].

Binding studies of the various N-truncated forms of cC reveal that the residues Gly9 and Ala10 are important for binding cC to papain [15,25,38]. The shorter Gly9 and Ala10-forms of cC bind almost 10,000-fold weaker to papain compared to the longer forms (Ser1-form) [14]. A similar loss of inhibitory activity is observed for the N-terminal truncated forms of hCC and stefin B [39]. In hCC, the residues 7 and 8 are most important for binding hCC to papain [33,34]. Human cystatin C is the mammalian equivalent to the avian cC. They belong to family 2 or cystatin family, whereas stefin B belongs to the family 1 or stefin family of the cystatin superfamily [12,13].

This shows that the N-terminal sequence, the eleven N-terminal amino acid residues of hCC, is of major importance for the inhibitory activity of CPIs of the cystatin superfamily.

Further, the binding affinity of the hCC variant, devoid of the 10 N-terminal residues, for cathepsin B and L is reduced by 3 orders of magnitude, whereas its binding affinity for cathepsin H remains strong [37,38]. This shows that the N-terminal region is also responsible for the differential inhibitory activity of the inhibitor for cysteine proteases of the papain superfamily.

The substitution of amino acid residues of the N-terminal sequence, by Gly, reveals that the amino acid Leu9 in hCC confers selectivity to the inhibition of the target peptidases; Leu9 contributes to most of the inhibitory activity of hCC for cathepsins B and L [40]. It

reveals that Val10 is responsible for most of the affinity of the N-terminal binding region for all four enzymes, cathepsins B, H, L and S. In contrast, it reveals that the contribution of Arg-8 for the affinity for all four enzymes is minor.

This shows that specific amino acid residues underlie the differential and specific inhibitory activities of the inhibitor of the cystatin superfamily for cysteine proteases of the papain superfamily. Among the N-terminal amino acid residues, Leu9 and Val10 are particularly important residues, contributing to the active site of hCC [37,38,40].

The residue Gly11 of hCC is highly conserved among the CPIs of the cystatin superfamily [17]. Gly11 introduces some flexibility in the N-terminal region of the inhibitors. Gly11 may contribute to the formation of a glycyl bond and confers flexibility to the preceding N-terminal segment, which is a prerequisite for optimal enzyme binding of the N-terminal motif.

In all, the N-terminal region is a highly conserved region of CPIs of the cystatin superfamily. It is of major importance for the inhibitory activity, specificity and differential affinity of CPIs of the cystatin superfamily for cysteine proteases of the papain superfamily. The N-terminal sequence and its conserved region are part of the active site of the CPIs of the cystatin superfamily.

INTERACTION CYSTATIN-PAPAIN

Chicken cystatin is the prototype of the CPIs of the family 2 or cystatin family of the cystatin superfamily [11]. Papain is the prototype of the protease inhibitors of the papain superfamily [2].

The three-dimensional structure of cC has been determined by X-ray crystallography [19]. This structure is at the basis of a model of interaction between cC and papain, underlying the inhibitory activity of cC for cysteine proteases of the papain superfamily.

The conserved motifs involved in binding of cystatins are located in adjacent turn structures along the edge of the molecule. This distribution of the sites involved in the inhibitory activity, along the edge of the molecule, is referred as a tripartite wedge-shaped binding region [19]. The distribution of the sites involved in the inhibitory activity, along the edge of the molecule, provides molecular surface complementary to the active site clefts of cysteine proteinases. A model of interaction between cC and papain has been developed based on computer modeling or docking [19].

In this model, the N-terminal segment and both hairpin loops of cystatin are highly complementary to the active site cleft of papain. The N-terminal residues form a tight turn and align with residues of papain in an antiparallel fashion. The two hairpin loops, L1 and L2 insert into the active cleft of papain.

In this model, i) the three motifs of cystatins involved in the CPI site establish independent interactions with the proteases, ii) the main interactions are provided primarily by the amino-terminal segment and the first hairpin loop containing the highly conserved Gln-Xaa-Val-Xaa-Gly region, with minor contributions coming from the second hairpin loop, and iii) hydrophobic bonds are important for the interaction of cystatins with proteases, as all three regions contain primarily hydrophobic residues [19,21].

The interaction of the N-terminal motif of cC would contribute to 36% of the binding energy in the complex cystatin-papain [38]. The importance of the N-terminal domain for the binding enzyme-inhibitor supports previous studies showing that the N-terminal region of the inhibitor is of major importance for its inhibitory activity [37,38]. However, the importance of the N-terminal motif for the inhibitory activity of cystatins depends on the target enzyme [25,37,38]. Hence, the importance of the N-terminal region in the interaction enzyme-receptor varies depending on the target enzyme.

This model is compatible with kinetics studies showing an enzyme-inhibitor complexes being formed by a one-step reversible bimolecular reaction [23-33].

In all, the model of interaction between cC and papain, based on computer modeling, provides a structural basis for the interaction of the three sequences of cystatins, involved in their CPI activity, the N-terminal sequence, the sequence Arg8-Leu9-Val10-Gly11, the sequences Gln55-Ile56-Val57-Ala58-Gly59 and Pro105-Trp106 (hCC numbering), with the active site of papain. This model underlies the inhibitory activity of cystatins on cysteine proteases of the papain superfamily.

Note: this model also provides a structural basis for the lack of proteolytic activity of papain on cystatin. The amino-terminal segment of cC is too far away to be attacked by the reactive site of papain, the Cys-25, preventing its proteolitic cleavage by papain. Disulfide bonds in cC also insure molecular rigidity, stability and resistance to proteases.

INTERACTION CYSTATIN-CATHEPSIN B

Proteases of the papain superfamily are mostly endopeptidases. Contrary to most members of the papain superfamily, cathepsin B functions also as an exopeptidase. Human cystatin C elicits lower affinity for cathepsin B than for other cathepsins and the binding of hCC to cathepsin B is best described by a two-step reaction mechanism [23-33,35].

The structure of cathepsin B has been determined by X-ray crystallography [4]. This study reveals that cathepsin B has an occluding loop, comprising residues 104-126, that partially blocks its active site [4]. It is proposed that the interaction between the enzyme and its substrate involves an initial conformational change by the inhibitor, displacing the occluding loop of the protease and making the active site of the protease accessible to the substrate [35]. The N-terminal region of the inhibitor would interact with the occluding loop of the protease to induce the initial conformational change.

This model is in agreement with a two-step reaction mechanism reported for cathepsin B [35]. This interaction may result in a lower affinity of cystatins for cathepsin B, than for other cathepsins [23-33]. The occluding loop may also modulate the activity of the enzyme. Particularly, the occluding loop may enable cathepsin B to function as an exopeptidase [41].

INTERACTION HUMAN CYSTATIN C-LEGUMAIN

It was recently reported that cC and hCC inhibit the activity of legumain [36]. Legumain was originally characterized in plants [42] and later in mammals; it was cloned and

sequenced in human [43] and mouse [44]. Legumain is a cysteine endopeptidase that is a prototype of a distinct family of cysteine proteases than family C1, the family C13 of cysteine proteases. The C1 and C13 families of cysteine proteases are evolutionarily unrelated, and therefore members of their families elicit different structures, particularly different protease or catalytic domains [44].

This suggests that cC and hCC inhibit legumain through a different mechanism than cysteine proteases of the papain superfamily. In support to this contention, i) legumain is not inhibited by the synthetic inhibitor of enzymes belonging to family C1, E-64 [43]. E-64 is an irreversible inhibitor of peptidases of family C1, as well as family C2 [45], ii) N-terminal variants, truncated form of cystatins do not elicit inhibitory activity on cysteine proteases of the family C1 or papain superfamily, but elicit inhibitory activity on cysteine proteases of the family C13 or legumain family [13,14,17,19,46], and iii) dimmers of hCC, formed in mild predenaturing conditions, do not elicit inhibitory activity on cysteine proteases of the family C1, but elicit inhibitory activity on cysteine proteases of the family C13. The inactivity of dimmers of cystatin C against papain-like enzymes is a result of intermolecular interactions between the papain-binding domains of two cystatin C molecules [47]. This shows that the binding sites of papain and legumain for cystatin C are therefore likely to be completely independent of each other.

In order to inhibit legumain, cystatin must tightly bind the protease domain of legumain. The Asn residue 39 of hCC is a conserved residue amongst the CPIs of the cystatin superfamily. Substitution of Asn39 by Lys results in a variant of cystatin lacking inhibitory activity for legumain [48,49]. This suggests that Asn39 and its surrounding conformation are relevant for the interaction between hCC and legumain, and for the inhibitory activity of hCC on legumain.

It is proposed that Asn39 and its surrounding are part of a novel second reactive site of hCC. Cystatins carrying this site would be potent inhibitors of mammalian legumain. Other cystatin motifs may be involved in interacting with legumain, just as several motifs are involved in the cystatin inhibition of papain.

In all, hCC has a second reactive site specific for mammalian legumain, distinct of the reactive site for cysteine proteases of the papain superfamily [36]. Human cystatin C inhibits cysteine proteases of two distinct families, the papain (C1) and legumain (C13) families, by different mechanisms.

CONCLUSION

CPIs of the cystatin superfamily are inhibitors of cysteine proteases of the C1 family or papain superfamily. The N-terminal sequence, the eleven first N-terminal amino acid residues, of hCC is of major importance for the inhibitory activity of CPIs of the cystatin superfamily for cysteine proteases of the papain superfamily. The diversity of CPIs of the cystatin superfamily reflects the diversity of cysteine proteases of the papain superfamily. Specific amino acid residues of the N-terminal sequence of cystatins underlie the differential and specific inhibitory activities of the inhibitor for cysteine proteases of the papain superfamily.

Three sequences of cystatins are involved in their CPI activity for cysteine proteases of the papain superfamily, the N-terminal sequence, the sequence Arg8-Leu9-Val10-Gly11, the sequences Gln55-Ile56-Val57-Ala58-Gly59 and Pro105-Trp106 (hCC numbering). These three motifs establish independent interactions with the proteases. Hydrophobic bonds are important for the interaction of cystatins with proteases, as all three regions contain primarily hydrophobic residues.

Human cystatin C inhibits the activity of legumain, a cysteine endopeptidase that belongs to the family of cysteine protease C13 or legumain family. The C1 or papain superfamily and C13 family of cysteine proteases are evolutionarily unrelated. Human cystatin C has a second reactive site specific for mammalian legumain, distinct of the reactive site for cysteine proteases of the papain superfamily. Human cystatin C inhibits cysteine proteases of two distinct families, the papain (C1) and legumain (C13) families, by different mechanisms.

The observation that cystatins are not specific to family C1 or papain superfamily, but also inhibit cysteine proteases of the family C13, reveals that CPIs of the cystatin superfamily have a broader spectrum of activity than originally reported. The significance for the broader activity of hCC remains to be established. Inhibitors of proteases are of major importance for maintaining a stable cell environment. Such broader spectrum of activity of proteinase inhibitors may provide a redundant mechanism for protease inhibition, made in place during evolution. It could compensate for eventual mutations of other protease inhibitors.

The identification of the mechanism underlying the inhibitory activity of protease inhibitors, and particularly CPIs of the cystatin superfamily, has important therapeutic implication, like the synthesis of synthetic inhibitors of proteases. Based on the N-terminal sequence of cystatins Arg8-Leu9-Val10-Gly11, several inhibitors of cysteine proteases of the papain superfamily have been devised, like E-64 [45,50]. Proteases are ubiquitous, they are involved in basically all physio- and pathological processes. Devising synthetic inhibitors is crucial to control their pathological activities, like in neurological diseases, tumors and viral infections [51].

REFERENCES

[1] Rzychon M, Chmiel D, Stec-Niemczyk J. (2004) Modes of inhibition of cysteine proteases. *Acta Biochim Pol.51*, 861-73.

[2] Berti PJ, Storer AC. (1995) Alignment/phylogeny of the papain superfamily of cysteine proteases. *J Mol Biol. 246*, 273-83.

[3] Nagler DK, Menard R. (2003) Family C1 cysteine proteases: biological diversity or redundancy? *Biol Chem. 384*, 837-43.

[4] Musil D, Zucic D, Turk D, Engh RA, Mayr I, Huber R, Popovic T, Turk V, Towatari T, Katunuma N, et al. (1991) The refined 2.15 A X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. *EMBO J*. 1991 10, 2321-30.

[5] Muller-Esterl W, Fritz H, Kellermann J, Lottspeich F, Machleidt W, Turk V. (1985) Genealogy of mammalian cysteine proteinase inhibitors. Common evolutionary origin of stefins, cystatins and kininogens. *FEBS Lett. 191*, 221-6.

[6] Rawlings ND, Barrett AJ. (1990) Evolution of proteins of the cystatin superfamily. *J Mol Evol. 30*, 60-71.

[7] Sen LC, Whitaker JR. (1973) Some properties of a ficin-papain inhibitor from avian egg white. *Arch Biochem Biophys. 158*, 623-32.

[8] Anastasi A, Brown MA, Kembhavi AA, Nicklin MJ, Sayers CA, Sunter DC, Barrett AJ. (1983) Cystatin, a protein inhibitor of cysteine proteinases. Improved purification from egg white, characterization, and detection in chicken serum. *Biochem J. 211*, 129-38.

[9] Turk V, Brzin J, Longer M, Ritonja A, Eropkin M, Borchart U, Machleidt W. (1983) Protein inhibitors of cysteine proteinases. III. Amino-acid sequence of cystatin from chicken egg white. *Hoppe Seylers Z Physiol Chem. 364*, 1487-96.

[10] Auerswald EA, Genenger G, Mentele R, Lenzen S, Assfalg-Machleidt I, Mitschang L, Oschkinat H, Fritz H. (1991) Purification and characterization of a chicken egg white cystatin variant expressed in an Escherichia coli pIN-III-ompA system. *Eur J Biochem. 200*, 131-8.

[11] Barrett AJ, Fritz H, Grubb A, Isemura S, Jarvinen M, Katunuma N, Machleidt W, Muller-Esterl W, Sasaki M, Turk V. (1986) Nomenclature and classification of the proteins homologous with the cysteine-proteinase inhibitor chicken cystatin. *Biochem J. 236*, 312.

[12] Turk V, Bode W. (1991) The cystatins: protein inhibitors of cysteine proteinases. *FEBS Lett. 285*, 213-9.

[13] Abrahamson M, Alvarez-Fernandez M, Nathanson CM. Cystatins. (2003) *Biochem Soc Symp. 70*, 179-99.

[14] Abrahamson M, Ritonja A, Brown MA, Grubb A, Machleidt W, Barrett AJ. (1987) Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin. *J Biol Chem. 262*, 9688-94.

[15] Machleidt W, Thiele U, Laber B, Assfalg-Machleidt I, Esterl A, Wiegand G, Kos J, Turk V, Bode W. (1989) Mechanism of inhibition of papain by chicken egg white cystatin. Inhibition constants of N-terminally truncated forms and cyanogen bromide fragments of the inhibitor. *FEBS Lett. 243*, 234-8.

[16] Arai S, Watanabe H, Kondo H, Emori Y, Abe K. (1991) Papain-inhibitory activity of oryzacystatin, a rice seed cysteine proteinase inhibitor, depends on the central Gln-Val-Val-Ala-Gly region conserved among cystatin superfamily members. *J Biochem (Tokyo). 109*, 294-8.

[17] Hall A, Dalboge H, Grubb A, Abrahamson M. (1993) Importance of the evolutionarily conserved glycine residue in the N-terminal region of human cystatin C (Gly-11) for cysteine endopeptidase inhibition. *Biochem J. 291*, 123-9.

[18] Hall A, Hakansson K, Mason RW, Grubb A, Abrahamson M. (1995) Structural basis for the biological specificity of cystatin C. Identification of leucine 9 in the N-terminal binding region as a selectivity-conferring residue in the inhibition of mammalian cysteine peptidases. *J Biol Chem. 270*, 5115-21.

[19] Bode W, Engh R, Musil D, Thiele U, Huber R, Karshikov A, Brzin J, Kos J, Turk V. (1988) The 2.0 A X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J. 7*, 2593-9.

[20] Kozak M, Jankowska E, Janowski R, Grzonka Z, Grubb A, Alvarez Fernandez M, Abrahamson M, Jaskolski M. (1999) Expression of a selenomethionyl derivative and preliminary crystallographic studies of human cystatin C. *Acta Crystallogr D Biol Crystallogr. 55*, 1939-42.

[21] Turk B, Turk V, Turk D. (1977) Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. *Biol Chem. 378*, 141-50.

[22] Bjork I, Alriksson E, Ylinenjarvi K. (1989) Kinetics of binding of chicken cystatin to papain. *Biochemistry. 28*, 1568-73.

[23] Bjork I, Ylinenjarvi K. (1989) Interaction of chicken cystatin with inactivated papains. *Biochem J. 260*, 61-8.

[24] Bjork I, Ylinenjarvi K. (1990) Interaction between chicken cystatin and the cysteine proteinases actinidin, chymopapain A, and ficin. *Biochemistry. 29*, 1770-6.

[25] Bjork I, Ylinenjarvi K, Lindahl P. (1990) Equilibrium and kinetic studies of the interaction of chicken cystatin with four cysteine proteinases. *Biol Chem Hoppe Seyler. 371*, 119-24.

[26] Bromme D, Rinne R, Kirschke H. (1991) Tight-binding inhibition of cathepsin S by cystatins. *Biomed Biochim Acta. 50*, 631-5.

[27] Lindahl P, Nycander M, Ylinenjarvi K, Pol E, Bjork I. (1992) Characterization by rapid-kinetic and equilibrium methods of the interaction between N-terminally truncated forms of chicken cystatin and the cysteine proteinases papain and actinidin. *Biochem J. 286*, 165-71.

[28] Balbin M, Hall A, Grubb A, Mason RW, Lopez-Otin C, Abrahamson M. (1994) Structural and functional characterization of two allelic variants of human cystatin D sharing a characteristic inhibition spectrum against mammalian cysteine proteinases. *J Biol Chem. 269*, 23156-62.

[29] Turk B, Colic A, Stoka V, Turk V. (1994) Kinetics of inhibition of bovine cathepsin S by bovine stefin B. *FEBS Lett. 339*, 155-9.

[30] Pol E, Olsson SL, Estrada S, Prasthofer TW, Bjork I. (1995) Characterization by spectroscopic, kinetic and equilibrium methods of the interaction between recombinant human cystatin A (stefin A) and cysteine proteinases. *Biochem J. 311*, 275-82.

[31] Turk B, Stoka V, Bjork I, Boudier C, Johansson G, Dolenc I, Colic A, Bieth JG, Turk V. (1995) High-affinity binding of two molecules of cysteine proteinases to low-molecular-weight kininogen. *Protein Sci. 4*, 1874-80.

[32] Turk B, Stoka V, Turk V, Johansson G, Cazzulo JJ, Bjork I. (1996) High-molecular-weight kininogen binds two molecules of cysteine proteinases with different rate constants. *FEBS Lett. 391*, 109-12.

[33] Ni J, Abrahamson M, Zhang M, Fernandez MA, Grubb A, Su J, Yu GL, Li Y, Parmelee D, Xing L, Coleman TA, Gentz S, Thotakura R, Nguyen N, Hesselberg M, Gentz R. (1997) Cystatin E is a novel human cysteine proteinase inhibitor with structural resemblance to family 2 cystatins. *J Biol Chem. 272*, 10853-8.

[34] Nicklin MJ, Barrett AJ. (1984) Inhibition of cysteine proteinases and dipeptidyl peptidase I by egg-white cystatin. *Biochem J. 223*, 245-53.

[35] Nycander M, Estrada S, Mort JS, Abrahamson M, Bjork I. (1998) Two-step mechanism of inhibition of cathepsin B by cystatin C due to displacement of the proteinase occluding loop. *FEBS Lett. 422*, 61-4.

[36] Alvarez-Fernandez M, Barrett AJ, Gerhartz B, Dando PM, Ni J, Abrahamson M. (1999) Inhibition of mammalian legumain by some cystatins is due to a novel second reactive site. *J Biol Chem. 274*, 19195-203.

[37] Abrahamson M, Ritonja A, Brown MA, Grubb A, Machleidt W, Barrett AJ. (1987) Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin. *J Biol Chem. 262*, 9688-94.

[38] Machleidt W, Thiele U, Assfalg-Machleidt I, Forger D, Auerswald EA. (1991) Molecular mechanism of inhibition of cysteine proteinases by their protein inhibitors: kinetic studies with natural and recombinant variants of cystatins and stefins. *Biomed Biochim Acta. 50*, 613-20.

[39] Thiele U, Assfalg-Machleidt I, Machleidt W, Auerswald EA. (1990) N-terminal variants of recombinant stefin B: effect on affinity for papain and cathepsin B. *Biol Chem Hoppe Seyler. 371*, 125-36.

[40] Hall A, Hakansson K, Mason RW, Grubb A, Abrahamson M. (1995) Structural basis for the biological specificity of cystatin C. Identification of leucine 9 in the N-terminal binding region as a selectivity-conferring residue in the inhibition of mammalian cysteine peptidases. *J Biol Chem. 270*, 5115-21.

[41] Illy C, Quraishi O, Wang J, Purisima E, Vernet T, Mort JS. (1997) Role of the occluding loop in cathepsin B activity. *J Biol Chem. 272*, 1197-202.

[42] Kembhavi AA, Buttle DJ, Knight CG, Barrett AJ. (1993) The two cysteine endopeptidases of legume seeds: purification and characterization by use of specific fluorometric assays. *Arch Biochem Biophys. 303*, 208-13.

[43] Chen JM, Dando PM, Rawlings ND, Brown MA, Young NE, Stevens RA, Hewitt E, Watts C, Barrett AJ. (1997) Cloning, isolation, and characterization of mammalian legumain, an asparaginyl endopeptidase. *J Biol Chem. 272*, 8090-8.

[44] Chen JM, Dando PM, Stevens RA, Fortunato M, Barrett AJ. (1998) Cloning and expression of mouse legumain, a lysosomal endopeptidase. *Biochem J. 335*, 111-7.

[45] Grubb A, Abrahamson M, Olafsson I, Trojnar J, Kasprzykowska R, Kasprzykowski F, Grzonka Z. (1990) Synthesis of cysteine proteinase inhibitors structurally based on the proteinase interacting N-terminal region of human cystatin C. *Biol Chem Hoppe Seyler. 371*, 137-44.

[46] Stubbs MT, Laber B, Bode W, Huber R, Jerala R, Lenarcic B, Turk V. (1990) The refined 2.4 A X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *EMBO J. 9*, 1939-47.

[47] Ekiel I, Abrahamson M, Fulton DB, Lindahl P, Storer AC, Levadoux W, Lafrance M, Labelle S, Pomerleau Y, Groleau D, LeSauteur L, Gehring K. (1997) NMR structural studies of human cystatin C dimmers and monomers. *J Mol Biol. 271*, 266-77.

[48] Abe Y, Shirane K, Yokosawa H, Matsushita H, Mitta M, Kato I, Ishii S. (1993) Asparaginyl endopeptidase of jack bean seeds. Purification, characterization, and high utility in protein sequence analysis. *J Biol Chem. 268*, 3525-9.

[49] Dando PM, Fortunato M, Smith L, Knight CG, McKendrick JE, Barrett AJ. (1999) Pig kidney legumain: an asparaginyl endopeptidase with restricted specificity. *Biochem J. 339*, 743-9.

[50] Johansson L, Grubb A, Abrahamson M, Kasprzykowski F, Kasprzykowska R, Grzonka Z, Lerner UH. (2000) A peptidyl derivative structurally based on the inhibitory center of cystatin C inhibits bone resorption in vitro. *Bone. 26*, 451-9.

[51] Keppler D. (2006) Towards novel anti-cancer strategies based on cystatin function. *Cancer Lett. 235*, 159-76.

Chapter V

DIMERIZATION AND DOMAIN SWAPPING

ABSTRACT

Cystatins are inhibitors of cysteine proteases of the papain superfamily. They bind reversibly to the catalytic site and inhibit the proteolytic activity of the proteases. X-ray crystallography reveals that cystatins fold into a characteristic three-dimensional structure, a five-stranded antiparallel β-sheet wrapped around a central α-helix, referred as a wedge-shaped structure. The three peptide motifs involved in the inhibitory activity of cystatins, on cysteine proteases of the papain superfamily, are distributed at the edges of the wedge-shaped structure. X-ray crystallography reveals that cystatins form dimmers under predenaturing conditions. Dimerization of cystatins would occur through three-dimensional domain swapping. Human cystatin C (hCC) also inhibits proteases of the legumain family, a family of cysteine proteases unrelated to the papain superfamily. Under predenaturing conditions, the cystatins lose their cysteine proteinase inhibitory (CPI) activities on cysteine proteases of the papain superfamily, but not on legumain. This confirms that the active site of hCC for legumain is distinct than for cysteine proteases of the papain superfamily. The dimerization of cystatins has profound implications for the physio- and pathological activities of CPIs of the cystatin superfamily.

INTRODUCTION

The cystatin superfamily is composed of three families of CPIs with homology to chicken cystatin (cC) and originating from a common ancestor gene [1]. Cystatins are tight-binding reversible inhibitors of cysteine proteases of the papain superfamily, including papain, ficin and lysosomal proteases, like cathepsins B, H, K, L and S [2]. Chicken cystatin and hCC belong to the family 2 or cystatin family. They are the most characterized CPIs of the cystatin superfamily. CPIs of family 2 cystatin are low molecular weight secreted proteins, with two disulfide bonds [3]. Three peptide motifs of cystatins are involved in their inhibitory activity on cysteine proteases of the papain superfamily. They have been identified

on hCC, as the N-terminal sequence Arg8-Leu9-Val10-Gly11, the sequence Gln55-Ile56-Val57-Ala58-Gly59 and the sequence Pro105-Trp106 [4-8].

The three-dimensional structures of cC and hCC have been characterized by X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy [9-11]. The conformation of the 120-residue polypeptide chain of cC is very similar to that of hCC, a monomer composed of five-stranded antiparallel β-sheets stacked against and partially wrapped around a central α-helix, in X-ray crystallography, referred as a wedge-shaped structure [9,11]. In this model, the three peptide motifs of cystatins, involved in their inhibitory activity on cysteine proteases of the papain superfamily, are at the edges of the three-dimensional structure. The N-terminal peptide motif of hCC, the sequence Arg8-Leu9-Val10-Gly11, is on the N-terminal β-sheet. The motif Gln55-Ile56-Val57-Ala58-Gly59 is in a central loop and the motif Pro105-Trp106 is in the second C-terminal hairpin loop [4-9].

Three-dimensional modeling between the structures of cC and papain reveals that the three peptide motifs of cystatins, involved in their inhibitory activity on cysteine proteases of the papain superfamily, establish independent interactions with the proteases [9]. In this model, the protease inhibitory site of cC interacts directly with the catalytic site of the cysteine protease, without major conformational change. Hydrophobic bonds maintain the interactions between the proteinase inhibitor and the protease, as all three regions of the proteinase inhibitor contain a high ratio of hydrophobic residues.

Chicken cystatin and hCC also inhibit cysteine proteases of the legumain family, a family of proteases unrelated to the papain superfamily [12,13]. Molecular studies reveal that the inhibitory domain of cC and hCC for legumain is distinct of their inhibitory domains for cysteine proteases of the papain superfamily [13].

X-RAY CRYSTALLOGRAPHY OF HUMAN CYSTATIN C UNDER PREDENATURING CONDITIONS

X-ray crystallography of hCC under mild predenaturing conditions, like heating to 70°C, acidic pH and denaturating conditions (guanidine hydrochloride), reveals that the proteinase inhibitor forms dimmers in these conditions [14,15]. The studies further reveal that the dimmers are symmetric and each monomer of hCC maintains the same fold, as observed by X-ray crystallography in standard conditions [9,11].Under mild predenaturing conditions, each monomer of hCC elicits a five-stranded antiparallel β-sheet wrapped around a central α-helix conformation [15].

Hence, hCC is a monomer in its native state, and forms dimmers under mild predenaturing conditions. The three-dimensional structure of cystatin, as described by X-ray crystallography under standard conditions, is also preserved during dimerization.

The dimmers formed in mild predenaturing conditions maintain their structure, upon transfer to a physiological pH or lowering temperature [14-17]. The solubility of dimmers of hCC is limited. This suggests that the interactions within the dimmers are strong.

Previous NMR spectroscopy studies have reported that cC, hCC, human cystatin A (hCA or stefin A) and B (hCB, or stefin B) form dimmers in solution, under predenaturing conditions [16,17]. Human cystatin A and hCB belongs to the family 1 cystatin or stefin

family [1,3]. CPIs of the family 1 cystatin are low molecular weight intracellular proteins with no disulfide bond [3]. X-ray crystallography studies therefore confirm NMR studies.

NMR structural studies further reveal the occurrence of a chemical shift perturbation on dimerization of cystatins [16,17]. This chemical shift is limited to the loop L1 of cystatins. The loop L1 is located between the β-sheet strands 2 and 3, on the N-terminal side, of cystatins [9]. It contains one of the inhibitory motifs of cystatins, the motif Gln-Xaa-Val-Xaa-Gly [5]. The chemical shift observed by NMR spectroscopy is indicative that dimerization involves structural rearrangements of the region containing the loop L1 [16-17].

In all, X-ray crystallography and NMR spectroscopy studies reveal that cystatins dimerize under mild predenaturing conditions. This dimerization occurs without structural rearrangement of the main fold of cystatins, but with a rearrangement of the loop L1. The rearrangement occurs in such a way that the dimmers behave as symmetrical structures. Hence, during dimerization, the proteins partially unfold to produce symmetric dimmers, while maintaining the secondary structure of each monomer.

THREE-DIMENSIONAL DOMAIN SWAPPING

Three-dimensional domain swapping is a process for forming dimeric, oligomeric and multimeric proteins, with preservation of the original secondary structure of each monomer [18].

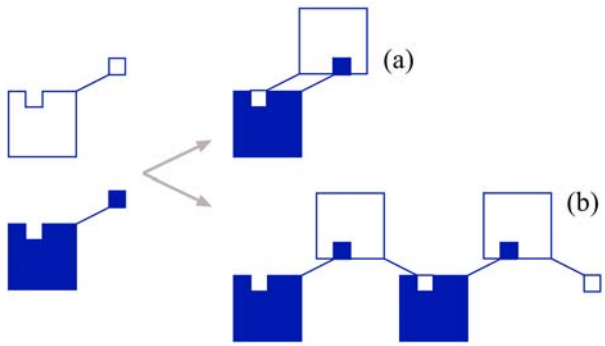


Figure 1. Three-dimensional domain swapping. Three-dimensional domain swapping is a process for forming dimeric, oligomeric and multimeric proteins. In three-dimensional domain swapping, one domain, a flexible domain, of a multi-domain molecule takes the place of the same domain in another similar molecule. In turn, the flexible domain from the latter molecule takes the same place in another molecule, leading to the formation of multimeric proteins. Multimeric proteins can be either closed-ended (a) or open-ended (b). One of the requirements for three-dimensional domain swapping is the existence of a flexible domain, capable of unfolding and taking the same place in another similar molecule. A process referred as domain swapping. The flexible domain undergoing domain swapping is referred as the hinge or linker region. The hinge links two different molecules to each other. Once this process is completed, the resulting monomers conserved their original secondary structures, but have components from two distinct molecules. Three-dimensional domain swapping is believed to be a process by which amyloid fibrils are formed.

In three-dimensional domain swapping, one domain, a flexible domain, of a multi-domain molecule takes the place of the same domain in another similar molecule. In turn, the flexible domain from the latter molecule takes the same place in another molecule, leading to the formation of multimeric proteins. Multimeric proteins can be either closed-ended or open-ended (figure 1) [18,19].

Three-dimensional domain swapping involves partial and temporary unfolding of the molecules. One of the requirements for three-dimensional domain swapping is the existence of a flexible domain, capable of unfolding and taking the same place in another similar molecule. A process referred as domain swapping. Once this process is completed, the resulting monomers conserved their original secondary structures, but have components from two distinct molecules. The flexible domain undergoing domain swapping is referred as the hinge or linker region. The hinge links two different molecules or monomers to each other [18,19].

In all, three-dimensional domain swapping is a process for forming multimeric proteins. In theory, three-dimensional domain swapping can lead to infinite chain. It is believed to be a process by which amyloid fibrils are formed [18,20]. Amyloid fibrils result from abnormal deposits of proteins known as amyloidogenic proteins [21].

DIMERIZATION OF CYSTATINS UNDER MILD PREDENATURING CONDITIONS

It is proposed that, under mild predenaturing conditions, a process of three-dimensional domain swapping underlies the dimerization of hCC [15].

According to this model, the loop L1 on the N-terminal side of cystatin represents the flexible domain or hinge. Under mild predenaturing conditions, the five-stranded β-sheet would partially and temporarily unfold permitting domain swapping. Once the loop L1 has unfolded and swapped in the same place in another similar molecule, that monomer regains it original conformation, with a hinge L1 substituted from a different cystatin molecule.

Chicken cystatin and hCC, contrary to hCA and hCB, elicit two disulfide bridges intrachain [3]. Disulfide bridges bring rigidity to the cystatin chains. Because in cC and hCC the two disulfide bridges are in the C-terminal part of the molecules, they do not interfere with the process of three-dimensional domain swapping, allowing partial unfolding of the molecules.

Other models have been proposed to underlie the dimerization of cystatins under mild predenaturing conditions, like the direct association of monomers via the hydrophobic loops [16]. However, results favor three-dimensional domain swapping, as a model of dimerization of cystatins under mild predenaturing conditions [15,18].

DIMERIZATION AND CYSTEINE PROTEINASE INHIBITORY ACTIVITIES

Under mild predenaturing conditions, hCA, hCB, hCC and cC, lose their inhibitory activities; they do not inhibit cysteine proteases of the papain superfamily [15-17]. X-Ray crystallography studies reveal that the three peptide motifs of cystatins involved in the inhibitory activity on cysteine proteases, of the papain superfamily, are at the edges of their three-dimensional structure [4-8,9].

The dimerization of cystatins under mild predenaturing conditions would underlie their loss of activities; the dimerization would mask the active motifs of cystatins, involved in the inhibitory activity of cysteine proteases of the papain superfamily. Both models proposed for the dimerization of cystatins, like the direct association of monomers via hydrophobic loops and three-dimensional domain swapping, support this hypothesis [15,16].

Chicken cystatin and hCC also inhibit activity of the cysteine proteases of the legumain family [13]. In contrast to cysteine proteases of the papain superfamily, under mild predenaturing conditions, hCC does not lose its inhibitory activity on cysteine proteases of the legumain family [15-17]. Legumain belongs to a family of protease inhibitors distinct of the papain superfamily [12]. Hence, hCC would inhibit legumain through a distinct mechanism than proteases of the papain superfamily [13].

The observation that under mild predenaturing conditions, hCC does not lose its inhibitory activity on cysteine proteases of the legumain family, but lose its inhibitory activity on cysteine proteases of the papain superfamily supports this latter hypothesis. The mechanism underlying the inhibitory activity of cC and hCC on cysteine proteases of the legumain family remains to be determined.

CONCLUSION

X-Ray crystallography and NMR spectroscopy studies reveal that cystatins form dimmers under mild predenaturing conditions. Three-dimensional domain swapping has been proposed for the dimerization of cystatins, particularly hCC. This model supports the activity of cystatins on cysteine proteases of the papain superfamily, and the existence of a different active site for the inhibitory activity of hCC on cysteine proteases of the legumain family.

It remains to determine the functional, and physio- and pathological significance of the dimerization of cystatins observed under experimental conditions. Do cystatins dimerize *in vivo*, and under which physio- and pathological conditions? What are the functions associated with the dimerization of cystatins?

Cystatins are inhibitors of proteases of the papain superfamily, which includes lysosomal enzymes, like cathepsins B, H, K, L and S. The lysosomal compartment is acidic, with a pH in the range of 4.6 to 5.0. Such acidic pH could create “mild predenaturing-like conditions” for cystatins trapped within lysosomes, during the protein recycling process. This would result in the inactivation of cystatins and prevent cystatins to be released in their active state inside the cells, during the protein recycling process. A similar mechanism could occur in the

sealed off and acidified compartment under the bone-degrading osteoclasts. Physiologically, dimerization of cystatin would contribute to its inactivation.

Three-dimensional domain swapping is believed to be a mechanism by which amyloid fibrils are formed. Amyloidosis is a group of rare and potentially fatal diseases that result from the deposit of amyloid fibrils [22]. Amyloidosis includes diseases, like Alzheimer's disease and a bone marrow disorder, AL amyloidosis. Three-dimensional domain swapping is proposed for the dimerization of cystatins, and particularly hCC, this shows that hCC is an amyloidogenic protein. Pathologically, dimerization of cystatin would contribute to its involvement in amyloidosis.

Further studies remain to determine the biology, functional significance, and physio- and pathological role of the dimerization of CPIs of the cystatin superfamily, and particularly hCC.

REFERENCES

[1] Muller-Esterl W, Fritz H, Kellermann J, Lottspeich F, Machleidt W, Turk V. (1985) Genealogy of mammalian cysteine proteinase inhibitors. Common evolutionary origin of stefins, cystatins and kininogens. *FEBS Lett. 191*, 221-6.

[2] Berti PJ, Storer AC. (1995) Alignment/phylogeny of the papain superfamily of cysteine proteases. *J Mol Biol. 246*, 273-83.

[3] Abrahamson M, Alvarez-Fernandez M, Nathanson CM. Cystatins. (2003) *Biochem Soc Symp. 70*, 179-99.

[4] Abrahamson M, Ritonja A, Brown MA, Grubb A, Machleidt W, Barrett AJ. (1987) Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin. *J Biol Chem. 262*, 9688-94.

[5] Arai S, Watanabe H, Kondo H, Emori Y, Abe K. (1991) Papain-inhibitory activity of oryzacystatin, a rice seed cysteine proteinase inhibitor, depends on the central Gln-Val-Val-Ala-Gly region conserved among cystatin superfamily members. *J Biochem (Tokyo). 109*, 294-8.

[6] Saitoh E, Isemura S, Sanada K, Ohnishi K. (1991) Cystatins of family II are harboring two domains which retain inhibitory activities against the proteinases. *Biochem Biophys Res Commun. 175*, 1070-5.

[7] Hall A, Dalboge H, Grubb A, Abrahamson M. (1993) Importance of the evolutionarily conserved glycine residue in the N-terminal region of human cystatin C (Gly-11) for cysteine endopeptidase inhibition. *Biochem J. 291*, 123-9.

[8] Hall A, Hakansson K, Mason RW, Grubb A, Abrahamson M. (1995) Structural basis for the biological specificity of cystatin C. Identification of leucine 9 in the N-terminal binding region as a selectivity-conferring residue in the inhibition of mammalian cysteine peptidases. *J Biol Chem. 270*, 5115-21.

[9] Bode W, Engh R, Musil D, Thiele U, Huber R, Karshikov A, Brzin J, Kos J, Turk V. (1988) The 2.0 A X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J. 7*, 2593-9.

[10] Engh RA, Dieckmann T, Bode W, Auerswald EA, Turk V, Huber R, Oschkinat H. (1993) Conformational variability of chicken cystatin. Comparison of structures determined by X-ray diffraction and NMR spectroscopy. *J Mol Biol. 234*, 1060-9.

[11] Kozak M, Jankowska E, Janowski R, Grzonka Z, Grubb A, Alvarez Fernandez M, Abrahamson M, Jaskolski M. (1999) Expression of a selenomethionyl derivative and preliminary crystallographic studies of human cystatin C. *Acta Crystallogr D Biol Crystallogr. 55*, 1939-42.

[12] Chen JM, Dando PM, Stevens RA, Fortunato M, Barrett AJ. (1998) Cloning and expression of mouse legumain, a lysosomal endopeptidase. *Biochem J. 335*, 111-7.

[13] Alvarez-Fernandez M, Barrett AJ, Gerhartz B, Dando PM, Ni J, Abrahamson M. (1999) Inhibition of mammalian legumain by some cystatins is due to a novel second reactive site. *J Biol Chem. 274*, 19195-203.

[14] Ekiel I, Abrahamson M. (1996) Folding-related dimerization of human cystatin C. *J Biol Chem. 271*, 1314-21.

[15] Janowski R, Kozak M, Jankowska E, Grzonka Z, Grubb A, Abrahamson M, Jaskolski M. (2001) Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping. *Nat Struct Biol. 8*, 316-20.

[16] Ekiel I, Abrahamson M, Fulton DB, Lindahl P, Storer AC, Levadoux W, Lafrance M, Labelle S, Pomerleau Y, Groleau D, LeSauteur L, Gehring K. (1997) NMR structural studies of human cystatin C dimmers and monomers. *J Mol Biol. 271*, 266-77.

[17] Zerovnik E, Cimerman N, Kos J, Turk V, Lohner K. (1997) Thermal denaturation of human cystatin C and two of its variants; comparison to chicken cystatin. *Biol Chem. 378*, 1199-203.

[18] Jaskolski M. (2001) 3D domain swapping, protein oligomerization, and amyloid formation. *Acta Biochim Pol. 48*, 807-27.

[19] Rzychon M, Chmiel D, Stec-Niemczyk J. (2004) Modes of inhibition of cysteine proteases. *Acta Biochim Pol.51*, 861-73.

[20] Bennett MJ, Sawaya MR, Eisenberg D. (2006) Deposition diseases and 3D domain swapping. *Structure. 14*, 811-24.

[21] Dobson CM. (1999) Protein misfolding, evolution and disease. *Trends Biochem Sci*. 24, 329-32.

[22] Kyle RA. (1995) Amyloidosis. *Circulation. 91*, 1269-71.

Chapter VI

AMYLOIDOSIS AND AMYLOIDOGENIC PROTEINS

ABSTRACT

Amyloidosis is a group of rare and potentially fatal diseases that result from the deposit of amyloid fibrils. Amyloidosis includes diseases, like Alzheimer's disease (AD) and a bone marrow disorder, AL amyloidosis. Amyloid deposits are characterized by histological staining, like Congo red staining. There are over twenty proteins that have the propensity to form amyloid fibrils, referred as amyloidogenic proteins. The causative factors and mechanisms underlying the formation of amyloid remain mostly unknown. Mutations of amyloidogenic proteins may lead to the formation of amyloid fibrils. Three-dimensional domain swapping is believed to be a mechanism by which amyloid fibrils are formed. Human cystatin C (hCC) forms dimmers under mild predenaturing conditions and three-dimensional domain swapping has been proposed as model for the dimerization of hCC, revealing hCC as an amyloidogenic protein. Human cystatin C has been associated with a number of amyloidosis, including AD, hereditary cystatin C amyloid angiopathy (HCCAA) and hereditary cerebral hemorrhage with amyloidosis-Dutch type (HCHWA-Dutch type), but its function in the pathological processes of these diseases remains mostly unknown. The amyloidogenic properties of hCC may lead to a better understanding of the role of hCC in amyloidosis and amyloid formation.

INTRODUCTION

Amyloidosis is a group of rare, progressive, disabling, incurable and potentially fatal diseases that result from the deposit of amyloid fibrils [1]. Amyloidosis includes diseases, like Alzheimer's disease (AD), AL amyloidosis, a bone marrow disorder, and prion-related diseases. Amyloid deposits are extra cellular aggregates of fibrils of specific proteins, known as amyloid fibrils [2]. Amyloidosis frequently affects the gastrointestinal tract, heart, kidneys, liver, nervous system and spleen. Amyloid deposits may be distributed throughout the body, pathologies referred as systemic amyloidosis. Amyloid diseases are also referred as protein deposition diseases [3].

Amyloidosis is classified as primary, secondary, hemodialysis-associated and senile amyloidosis. Amyloidosis may be acquired or inherited, also referred as familial amyloidosis. Primary amyloidosis occurs independently of other diseases. Secondary amyloidosis occurs as a result of another illness. Hemodialysis-associated amyloidosis occurs in people who have had long-term kidney dialysis. Senile amyloidosis occurs with aging.

The most common form of amyloidosis is a bone marrow disorder, AL amyloidosis [4]. AL amyloidosis is a primary systemic amyloidosis resulting from aggregates of immunoglobulin light chain proteins. In United States, up to 2000 cases of AL amyloidosis are diagnosed per year. Two-thirds of amyloid diseases are secondary amyloidosis. Current treatments for amyloidosis involve chemotherapy, liver and stem cell transplantation [5-7].

Amyloid deposits are characterized by histological staining, like Congo red anisotropic staining and Congo red X-34, thioflavin-T or -S fluorescence [8-11]. Congo red is a histological stain, pH sensitive, staining amyloid deposits. It also gives a green birefringence, under polarized light, characteristic of amyloid.

AMYLOIDOGENIC PROTEINS

Amyloidogenic proteins are proteins forming amyloid fibrils [12]. Over twenty amyloidogenic proteins have been reported, including amyloid protein precursor (APP) and immunoglobulin light chain protein (table 1) [13]. Amyloidogenic proteins are monomers and soluble in their physiological state. Under pathological conditions, they form insoluble extra cellular aggregates or deposits of amyloid fibrils [2].

Amyloidogenic proteins are characterized by their structural diversity; they belong to unrelated family of proteins. Yet, they assemble into fibers with similar structures. This suggests that the formation of amyloid fibrils involves common mechanisms [14]. The causative factors and mechanisms underlying the formation of amyloid fibrils remain mostly unknown.

Mutations of amyloidogenic proteins may lead to the formation of amyloid fibrils and three-dimensional domain swapping is believed to be a mechanism by which amyloid and related fibrils are formed [15].

Three-dimensional domain swapping is a process for forming oligomeric, dimeric or multimeric proteins. In three-dimensional domain swapping one domain of a multi-domain molecule takes the place of the same domain in another similar molecule [16]. Three-dimensional domain swapping results in multimeric proteins, in which the resulting monomers conserved the same structure as their original [17].

Mutation of amyloidogenic proteins is one of the factors that may cause amyloidogenic proteins to form amyloid deposits. Mutations would destabilize the original conformation of the protein, leading to the formation of multimeric proteins by three-dimensional domain swapping [18]. Amyloid diseases are also referred as protein conformational diseases [19].

The involvement of amyloid deposits in the pathology and pathogenesis of amyloidosis remains mostly unknown. Amyloid fibrils share the ability to permeabilize cellular membranes and lipid bilayers [14]. This may represent the primary toxic mechanism of amyloid fibrils.

Table 1. Amyloidogenic proteins

Amyloidogenic proteins	Examples of associated pathology
Alpha-synuclein	Parkinson's disease
Amylin	Diabetis mellitus type 2
Amyloid protein precursor (APP)	Alzheimer's disease
Apolipoproteins A1, B	
Atrial natriuretic factor	
β 2-microglobulin	Dialysis-related amyloidosis
Fibrinogen α	
Gelsolin	Gelsolin amyloidosis
Glutamine-rich polypeptides, like huntingtin	Huntington's disease
Human cystatin C	HCCAA
Immunoglobulin light chain protein (AL protein)	AL amyloidosis
Insulin	Insulin amyloidosis
Islet amyloid polypeptide	
Lactoferrin	
Lysozyme	
Pion protein	Transmissible spongiform encephalopathies
Procalcitonin	
Prolactin	
Serum amyloid A protein (AA protein)	AA amyloidosis
Transthyretin protein (TTR protein)	Senile systemic amyloidosis

Amyloidogenic proteins are proteins forming amyloid fibrils. Amyloid fibrils form extra cellular aggregates, known as amyloid deposits. There are over twenty amyloidogenic proteins. Amyloidosis is a group of rare and potentially fatal diseases that result from the deposit of amyloid fibrils. The causative factors and mechanisms underlying the formation of amyloid fibrils remain mostly unknown. Mutations of amyloidogenic proteins may lead to the formation of amyloid fibrils and three-dimensional domain swapping is believed to be a mechanism by which amyloid and related fibrils are formed.

Hereditary cystatin C amyloid angiopathy (HCCAA) is a rare fatal form of cerebral amyloid angiopathy in young people in Iceland. Mutation in cystatin C underlies the pathogenesis of HCCAA. Human cystatin C acquires amyloidogenic properties when a mutation at position 68 (L68Q) is present in the protein.

Huntington's disease (HD) is a progressive, autosomal dominantly inherited, neurodegenerative disorder. It is caused by a polyglutamine [poly(Q)] repeat expansion in the first exon of the huntingtin protein. HD is characterized by involuntary movements (chorea), cognitive decline and psychiatric manifestations. It is one of a number of late-onset neurodegenerative disorders caused by expanded glutamine repeats. Reports show that some inclusions in HD brain tissues possess an amyloid-like structure and that poly(Q) tracts in the pathological range form high molecular weight protein aggregates, with a fibrillary morphology similar to scrapie prion rods and β-amyloid fibrils in Alzheimer's disease [46,47].

Prion is the agent causing transmissible spongiform encephalopathies, like bovine spongiform encephalopathy and its variant Creutzfeld-Jakob disease. It is an amyloidogenic protein with infectious abilities that can cross the species barrier [48,49].

HUMAN CYSTATIN C AND AMYLOIDOSIS

Human cystatin C is a cysteine proteinase inhibitor of the cystatin superfamily. It belongs to the family 2 of cystatin superfamily or cystatin family [20,21]. Cystatin C is involved in several amyloidosis, like AD, HCCAA and HCHWA-Dutch type.

AD is a progressive neurodegenerative disease, characterized in the brain by neurofibrillary tangles and amyloid or senile plaques [22]. Neurofibrillary tangles are composed of hyperphosphorylated tau proteins [23]. Tau protein is a microtubule-associated phosphoprotein; it is a soluble axonal protein involved in microtubule assembly [24]. Tau is expressed in mature neuronal cells of the nervous system [25]. In AD and tauopathies, tau protein is hyperphosphorylated by kinases, resulting in polymerization, assembling of tau proteins and formation of neurofibrillary tangles [26,27]. Tau protein is proteolysed by cathepsins B, D and L [28]. Amyloid or senile plaques are composed of deposits of amyloid fibrils, resulting from aggregates of protein β-amyloid. Protein β-amyloid is formed by maturation of β-amyloid protein precursor protein or APP [29]. Mutations in APP underlie the pathogenesis of most cases of early-onset AD, or familial form of AD [22].

HCCAA, also called hereditary cerebral hemorrhage with amyloidosis-Icelandic type, is a rare fatal form of cerebral amyloid angiopathy (CAA), in young people in Iceland [30]. CCA is a heterogeneous group of disorders characterized clinically by ischemic and/or hemorrhagic strokes, and histologically by deposition of amyloid in cerebral cortical blood vessels [31,32]. HCCAA is an autosomal dominant condition in which patients suffer at an early age from repeated cerebral hemorrhages. Mutation in cystatin C underlies the pathogenesis of HCCAA [33]. The mutated variant of cystatin C L68Q, in which Leu68 is substituted by Gln, is responsible for the autosomal dominant disease HCCAA.

HCHWA-Dutch type is rare hereditary disorder that causes severe CAA with hemorrhagic strokes of mid-life onset and dementia [34]. HCHWA-Dutch type is characterized, in the blood vessels, by amyloid deposits. Amyloid deposits in HCHWA-Dutch type result from aggregates of protein β-amyloid, like in AD. It is formed by maturation of APP and mutations in APP underlie the pathogenesis of HCHWA-Dutch type.

HUMAN CYSTATIN C AN AMYLOIDOGENIC PROTEIN

Human cystatin C is a low molecular weight secreted cysteine proteinase inhibitor, of the family 2 cystatin superfamily or cystatin family [21]. The crystal structure of hCC has been determined, it is a five-stranded antiparallel β-sheet wrapped around a central α-helix, referred as a wedge-shaped structure [35,36]. Under mild predenaturing conditions, hCC dimerizes. A three-dimensional domain swapping model has been proposed for the dimerization of hCC [37,38]. Its propensity to dimerize under mild predenaturing conditions reveals hCC as an amyloidogenic protein.

Cystatin C is involved in several amyloidosis, like AD, HCCAA and HCHWA-Dutch type, but its role in the pathogenesis and pathology of these diseases remains mostly unknown.

Co-deposition of hCC with amyloid deposits has been reported in several amyloidosis, like in early-onset AD, HCCAA and HCHWA-Dutch type [30,33,34,39]. This shows that cystatin C binds amyloid fibrils, particularly β-amyloid, and suggests that it promotes amyloid formation. The mechanisms underlying the involvement of cystatin C in amyloidosis remain to be established. It is proposed that the proteinase inhibitory activity may be involved in amyloid promoting activity. Cystatin C may also promote tau phosphorylation in AD, by inhibiting cathepsins B and L [40].

The mutated variant of cystatin C L68Q is responsible for the autosomal dominant disease HCCAA [30,33]. Recombinant proteins, of the mutated variant of cystatin C L68Q, form dimmers under predenaturing conditions milder than that required for wild-type cystatin [41-44]. Wild-type cystatin C dimmers can be formed by heating to 70^{o}C, whereas dimerization of the variant L68Q occurs at temperatures approximately 30^{o}C lower than for the wild-type form. Under physiological conditions, cystatin C is a monomer, it does not form dimmer, whereas under mild predenaturing conditions, cystatin C dimerizes [37,38]. A process of three-dimensional domain swapping has been proposed for the dimerization of cystatins, revealing hCC as an amyloidogenic protein [38].

The causative factors and mechanisms underlying the formation of amyloid fibrils remain also mostly unknown. The observation that recombinant proteins, of the mutated variant of cystatin C L68Q, form dimmers under predenaturing conditions milder than that required for wild-type cystatin suggests that three-dimensional domain swapping is a mechanism underlying the dimerization of the mutant variant of cystatin C L68Q. Three-dimensional domain swapping is therefore a model for the formation of amyloid fibrils from hCC in HCCAA. The observation that recombinant proteins of the mutated variant of cystatin C L68Q, corresponding to the mutation responsible for HCCAA, underlie the formation of amyloid from hCC, suggests that mutations of hCC are a causative factor for amyloid formation in HCCAA. The mutation of L68Q would contribute to destabilize the structure of cystatin C, by inserting a hydrophilic amino acid in a hydrophobic environment, permitting three-dimensional domain swapping and amyloid formation.

In all, cystatin C may be involved in amyloidosis by its promoting activity on amyloid formation and its amyloidogenic property. These results support three-dimensional domain swapping as a model for the formation of amyloid fibrils, and mutations of amyloidogenic proteins as a causative factor for amyloid formation.

CONCLUSION

Amyloidosis is a group of rare and potentially fatal diseases that result from the deposit of amyloid fibrils. Cystatin C is involved in several amyloidosis, like AD, HCCAA and HCHWA-Dutch type. The causative factors and mechanisms underlying the formation of amyloid remain mostly unknown, and the involvement of hCC in the pathogenesis and pathology of amyloid diseases remains mostly unknown.

Results show that cystatin C is involved in amyloidosis by its promoting activity on amyloid formation and its amyloidogenic property, particularly in HCCAA. Studies of mutated hCC support three-dimensional domain swapping as a model for the formation of

amyloid fibrils, and mutations of amyloidogenic proteins as a causative factor for amyloid formation.

The identification of cystatin C as an amyloidogenic protein has tremendous implication for therapy of amyloidosis [45]. Future investigations will aim at determining the contribution of cystatin C as an amyloidogenic protein versus its contribution as a proteinase inhibitor in amyloidosis. The mechanisms of formation of dimmers and amyloid fibrils from cystatin C remain to be further evaluated.

REFERENCES

[1] Kyle RA. (1995) Amyloidosis. *Circulation. 91*, 1269-71.

[2] Dobson CM. (1999) Protein misfolding, evolution and disease. *Trends Biochem Sci. 24*, 329-32.

[3] Bennett MJ, Sawaya MR, Eisenberg D. (2006) Deposition diseases and 3D domain swapping. *Structure. 14*, 811-24.

[4] Comenzo RL. (2006) Systemic immunoglobulin light-chain amyloidosis. *Clin Lymphoma Myeloma. 7*, 182-5.

[5] Valente M, Roy V, Lacy MQ, Dispenzieri A, Gertz MA. (2006) Autologous stem cell transplantation and IgM amyloidosis. *Leuk Lymphoma. 47*, 1006-12.

[6] Haq A, Hussain S, Meskat B, Mohan P, Conlon P, Hickey DP. (2007) Complications of renal transplantation in patients with amyloidosis. *Transplant Proc. 39*, 120-4.

[7] Gertz MA, Lacy MQ, Dispenzieri A, Hayman SR, Kumar S. (2007) Transplantation for amyloidosis. *Curr Opin Oncol. 19*, 136-41.

[8] Elghetany MT, Saleem A. (1988) Methods for staining amyloid in tissues: a review. *Stain Technol. 63*, 201-12.

[9] Styren SD, Hamilton RL, Styren GC, Klunk WE. (2000) X-34, a fluorescent derivative of Congo red: a novel histochemical stain for Alzheimer's disease pathology. *J Histochem Cytochem. 48*, 1223-32.

[10] Frid P, Anisimov SV, Popovic N. (2007) Congo red and protein aggregation in neurodegenerative diseases. *Brain Res Rev. 53*, 135-60.

[11] Linke RP. (2000) Highly sensitive diagnosis of amyloid and various amyloid syndromes using Congo red fluorescence. *Virchows Arch. 436*, 439-48.

[12] Serpell LC, Sunde M, Blake CC. (1997) The molecular basis of amyloidosis. *Cell Mol Life Sci. 53*, 871-87.

[13] Sipe JD, Cohen AS. (2000) Review: history of the amyloid fibril. *J Struct Biol. 130*, 88-98.

[14] Glabe CG. (2006) Common mechanisms of amyloid oligomer pathogenesis in degenerative disease. *Neurobiol Aging. 27*, 570-5.

[15] Jaskolski M. (2001) 3D domain swapping, protein oligomerization, and amyloid formation. *Acta Biochim Pol. 48*, 807-27.

[16] Rousseau F, Schymkowitz JW, Itzhaki LS. (2003) The unfolding story of three-dimensional domain swapping. *Structure. 11*, 243-51.

[17] Liu Y, Eisenberg D. (2002) 3D domain swapping: as domains continue to swap. *Protein Sci. 11*, 1285-99.

[18] Merlini G, Bellotti V. (2003) Molecular mechanisms of amyloidosis. *N Engl J Med. 349*, 583-96.

[19] Lin JC, Liu HL. (2006) Protein conformational diseases: from mechanisms to drug designs. *Curr Drug Discov Technol. 3*, 145-53.

[20] Muller-Esterl W, Fritz H, Kellermann J, Lottspeich F, Machleidt W, Turk V. (1985) Genealogy of mammalian cysteine proteinase inhibitors. Common evolutionary origin of stefins, cystatins and kininogens. *FEBS Lett. 191*, 221-6.

[21] Abrahamson M, Alvarez-Fernandez M, Nathanson CM. Cystatins. (2003) *Biochem Soc Symp. 70*, 179-99.

[22] Caselli RJ, Beach TG, Yaari R, Reiman EM. (2006) Alzheimer's disease a century later. *J Clin Psychiatry. 67*, 1784-800.

[23] Fukutani Y, Kobayashi K, Nakamura I, Watanabe K, Isaki K, Cairns NJ. (1995) Neurons, intracellular and extra cellular neurofibrillary tangles in subdivisions of the hippocampal cortex in normal ageing and Alzheimer's disease. *Neurosci Lett. 200*, 57-60.

[24] Kim H, Jensen CG, Rebhun LI. (1986) The binding of MAP-2 and tau on brain microtubules in vitro: implications for microtubule structure. *Ann N Y Acad Sci. 466*, 218-39.

[25] Drubin DG, Kirschner MW. (1986) Tau protein function in living cells. *J Cell Biol. 103*, 2739-46.

[26] Alonso A, Zaidi T, Novak M, Grundke-Iqbal I, Iqbal K. (2001) Hyperphosphorylation induces self-assembly of tau into tangles of paired helical filaments/straight filaments. *Proc Natl Acad Sci U S A. 98*, 6923-8.

[27] Wang JZ, Grundke-Iqbal I, Iqbal K. (2007) Kinases and phosphatases and tau sites involved in Alzheimer neurofibrillary degeneration. *Eur J Neurosci. 25*, 59-68.

[28] Billingsley ML, Kincaid RL. (1997) Regulated phosphorylation and dephosphorylation of tau protein: effects on microtubule interaction, intracellular trafficking and neurodegeneration. *Biochem J. 323*, 577-91.

[29] Hardy J, Selkoe DJ. (2002) The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science. 297, 353-6. Erratum in: (2002) *Science. 297*, 2209.

[30] Palsdottir A, Snorradottir AO, Thorsteinsson L. (2006) Hereditary cystatin C amyloid angiopathy: genetic, clinical, and pathological aspects. *Brain Pathol. 16*, 55-9.

[31] Maat-Schieman M, Roos R, van Duinen S. (2006) The cerebral beta-amyloid angiopathies: hereditary and sporadic. *Brain Pathol. 16*, 30-9.

[32] Coria F, Rubio I. (1996) Cerebral amyloid angiopathies. *Neuropathol Appl Neurobiol. 22*, 216-27.

[33] Ghiso J, Jensson O, Frangione B. (1986) Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Icelandic type is a variant of gamma-trace basic protein (cystatin C). *Proc Natl Acad Sci U S A. 83*, 2974-8.

[34] Haan J, Roos RA. (1992) Comparison between the Icelandic and Dutch forms of hereditary cerebral amyloid angiopathy. *Clin Neurol Neurosurg. 94*, S82-3.

[35] Bode W, Engh R, Musil D, Thiele U, Huber R, Karshikov A, Brzin J, Kos J, Turk V. (1988) The 2.0 A X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J. 7*, 2593-9.

[36] Kozak M, Jankowska E, Janowski R, Grzonka Z, Grubb A, Alvarez Fernandez M, Abrahamson M, Jaskolski M. (1999) Expression of a selenomethionyl derivative and preliminary crystallographic studies of human cystatin C. *Acta Crystallogr D Biol Crystallogr. 55*, 1939-42.

[37] Ekiel I, Abrahamson M. (1996) Folding-related dimerization of human cystatin C. *J Biol Chem. 27*1, 1314-21.

[38] Janowski R, Kozak M, Jankowska E, Grzonka Z, Grubb A, Abrahamson M, Jaskolski M. (2001) Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping. *Nat Struct Biol. 8*, 316-20.

[39] Levy E, Sastre M, Kumar A, Gallo G, Piccardo P, Ghetti B, Tagliavini F. (2001) Codeposition of cystatin C with amyloid-beta protein in the brain of Alzheimer disease patients. *J Neuropathol Exp Neurol. 60*, 94-104.

[40] Bi X, Zhou J, Lynch G. (1999) Lysosomal protease inhibitors induce meganeurites and tangle-like structures in entorhinohippocampal regions vulnerable to Alzheimer's disease. *Exp Neurol. 158*, 312-27.

[41] Abrahamson M, Grubb A. (1994) Increased body temperature accelerates aggregation of the Leu-68-->Gln mutant cystatin C, the amyloid-forming protein in hereditary cystatin C amyloid angiopathy. *Proc Natl Acad Sci U S A. 91*, 1416-20.

[42] Zerovnik E, Cimerman N, Kos J, Turk V, Lohner K. (1997) Thermal denaturation of human cystatin C and two of its variants; comparison to chicken cystatin. *Biol Chem. 378*, 1199-203.

[43] Bjarnadottir M, Nilsson C, Lindstrom V, Westman A, Davidsson P, Thormodsson F, Blondal H, Gudmundsson G, Grubb A. (2001) The cerebral hemorrhage-producing cystatin C variant (L68Q) in extra cellular fluids. *Amyloid. 8*, 1-10.

[44] Staniforth RA, Giannini S, Higgins LD, Conroy MJ, Hounslow AM, Jerala R, Craven CJ, Waltho JP. (2001) Three-dimensional domain swapping in the folded and molten-globule states of cystatins, an amyloid-forming structural superfamily. *EMBO J. 20*, 4774-81.

[45] Hamaguchi T, Ono K, Yamada M. (2006) Anti-amyloidogenic therapies: strategies for prevention and treatment of Alzheimer's disease. *Cell Mol Life Sci. 63*, 1538-52.

[46] Scherzinger E, Sittler A, Schweiger K, Heiser V, Lurz R, Hasenbank R, Bates GP, Lehrach H, Wanker EE. (1999) Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology. *Proc Natl Acad Sci U S A. 96*, 4604-9.

[47] McGowan DP, van Roon-Mom W, Holloway H, Bates GP, Mangiarini L, Cooper GJ, Faull RL, Snell RG. (2000) Amyloid-like inclusions in Huntington's disease. *Neuroscience. 100*, 677-80.

[48] Prusiner SB. (1982) Novel proteinaceous infectious particles cause scrapie. *Science. 216*, 136-44.

[49] Aguzzi A, Heppner FL. (2000) Pathogenesis of prion diseases: a progress report. *Cell Death Differ. 7*, 889-902.

Chapter VII

PHYSIOLOGY

ABSTRACT

The proteinase inhibitors of the cystatin superfamily are broadly expressed. They are present in a wide range of animal species, organs and tissues. This suggests that they are involved in a broad spectrum of physiological activities and functions. It has generally been presumed that the activity of proteinase inhibitors and particularly cysteine proteinase inhibitors (CPIs) is mediated through their proteinase inhibitory activity. Proteinase inhibitors of the cystatin superfamily are inhibitors of cysteine proteases of the papain superfamily. Other studies reveal that proteinase inhibitors of the cystatin superfamily also inhibit proteases of the legumain family, a family of cysteine proteases unrelated to the papain superfamily. Proteinase inhibitors of the cystatin superfamily, like cystatin C, have also been shown to possess biological activities and functions, unrelated to their proteinase inhibitory activity, like immunomodulation and the stimulation of neural stem cell (NSC) proliferation. Hence, proteinase inhibitors of the cystatin superfamily are pleiotropic molecules, with different domains of the molecules involved in different functions.

INTRODUCTION

Proteinase inhibitors of the cystatin superfamily are inhibitors of cysteine proteases of the papain superfamily, including cathepsins B, H, K, L and S [1,2]. Three sequences of the polypeptide chain, Arg8-Leu9-Val10-Gly11, Gln55-Ile56-Val57-Ala58-Gly59 and Pro105-Trp106, are involved in the CPI site of human cystatin C (hCC) [3-7]. X-ray crystallography of chicken cystatin (cC) and hCC reveals that their three-dimensional structure is a five-stranded antiparallel β-sheet wrapped around a central α-helix, referred as a wedge-shaped structure [8,9].

The three sequences of cystatins involved in the CPI site have been mapped in the three dimensional structure of cystatin C. The N-terminal peptide motif of hCC, the sequence Arg8-Leu9-Val10-Gly11, is on the N-terminal β-sheet. The motif Gln55-Ile56-Val57-Ala58-

Gly59 is in a central loop and the motif Pro105-Trp106 is in the second C-terminal hairpin loop [10].

Cystatin C forms tight reversible complexes, with dissociation equilibrium constant in the subnanomolar range, with virtually all cysteine peptidases of the papain superfamily [11]. Cystatin C inhibits these enzymes in an order of preference (from the highest to the lowest) of cathepsin L = S > B > H [12].

Chicken cystatin and hCC inhibit the activity of legumain [13]. Legumain is a cysteine endopeptidase that belongs to a distinct family of cysteine proteases than the family C1 or papain superfamily, the family C13 or legumain family. The C1 and C13 families of cysteine proteases are evolutionarily unrelated, and therefore members of their families elicit different structures, particularly different protease or catalytic domains [14].

Cystatins inhibit legumain through a different mechanism than cysteine proteases of the papain superfamily. The highly conserved Asn39 residue and its surrounding region are proposed to be part of a novel second reactive site of cystatins, like cC and hCC, underlying the activity of the CPIs on cysteine proteases of the family C13 [13].

In all, cystatins, and particularly hCC, elicit two distinct protease inhibitory activities, on cysteine proteases of the C1 and C13 families. Theses CPI activities are mediated by two distinct domains of the molecules. This reveals that cystatins have multiples domains involved in different proteinase inhibitory activities.

Cystatin C has a broad distribution in human tissues and body fluids, including the nervous system [15]. Other members of the cystatin family, like cystatins D, E, F (also called leukocystatin), S, SA and SN, have a more restricted distribution in human body fluids [16-19]. Cystatin D has so far only been found in saliva and tears. This shows that cystatins are involved in a broad range of physiological functions.

Numerous studies reveal that cystatins are involved in bone resorption, immunomodulation, inflammation, pathogen infection, stem cell growth and fate specification, and tumor progression. Some of these activities are unrelated to the CPI activity of the inhibitors of the cystatin superfamily and are mediated by distinct domains of the molecules. Hence, cystatins are pleiotropic molecules with different activities mediated by different domains of the molecules.

GENERAL PROTECTIVE FUNCTION

Proteinase inhibitors, like cystatin C, have a general protective function. They regulate the activity of proteases and their potential deleterious effects on the cellular environment, like destruction of tissues and connective tissues.

Intracellular and extra cellular CPIs of the cystatin superfamily regulate the activity of lysosomal proteases of the family C1, like cathepsins. They inhibit the activity of lysosomal proteases during the process of endocytosis, for recycling proteins and cell substrates [20].

Extra cellular inhibitors of the cystatin superfamily regulate the activity of extra cellular cysteine proteases of the family C1 and lysosomal cysteine proteases released from dying or diseased cells [20,21]. Lysosomal proteases released from dying or diseased cells may also be

inhibited by intracellular CPIs, like stefins A and B, also released from dying or diseased cells.

In all, CPIs of the cystatin superfamily have a general protective function, against cellular proteases of the family C1. This function is primarily mediated by their CPI activities.

HOST-PATHOGEN REACTION

Extra cellular inhibitors of the cystatin superfamily have a protective function against bacteria, viruses and parasites.

The protozoan *Trypanosoma cruzi* is the causative agent of Chagas' disease, a progressively debilitating disease [22]. The human immunodeficiency virus is the causative agent of acquired immunodeficiency syndrome and influenza virus is the causative agent of the flu [23,24]. These pathogen agents, bacteria, viruses and parasites, mediate their infectious activities through proteases, and particularly cysteine proteases [25]. Extra cellular inhibitors of the cystatin superfamily regulate cysteine proteases originating from bacteria, viruses and parasites [26,27].

Periodontopathic bacteria secrete a number of proteolytic enzymes, including cysteine proteases. The levels of cystatin C and S in saliva are increased in inflammatory periodontal diseases. Cystatin C, D and S in saliva protect the oral cavity against bacterial infection [28-30]. The growth of periodontopathic bacteria is inhibited by cystatin C and S, but the growth inhibition does not depend on the inhibition of periodontopathic bacteria cysteine proteases [31,32]. Therefore, cystatins inhibit the growth of periodontopathic bacteria through their CPI activity, but also by mechanisms other than protease inhibition.

In all, proteinase inhibitors of the cystatin superfamily, like cystatin C, have a protective function against pathogens and infectious agents, like bacteria and viruses. They contribute to the so called host-pathogen reaction [33]. This activity is not only mediated by their CPI activities, but also through other mechanisms unrelated of their CPI activities.

INFLAMMATION REACTION AND IMMUNOMODULATORY ACTIVITIES

Cystatins are involved in inflammation reactions and immune responses, through the cytokine network [34-36]. Cystatins modulate cathepsin activities and antigen presentation by dendritic cells (DCs) that are involved in immune reaction and inflammation reaction. Cathepsins play an important role in the degradation of molecules involved in immune reactions. DCs are present in most tissues and organs, where they act as antigen-presenting cells. Both of which are regulated by CPIs, like cystatin C.

Cystatins are also involved in immunomodulatory reactions. Chicken cystatin, human stefin, human cystatin C and L-kininogen stimulate the production of nitric oxide (NO) by macrophages, activated by γ-interferon [37,38]. NO is synthesized from L-arginine by nitric oxide synthetase (NOS). Three isoforms of NOS have been identified; they differ in their

tissue distribution and regulation. Neuronal NOS (nNOS or NOS-1) and endothelial NOS (eNOS or NOS-3) are constitutively present in neurons and endothelial cells, and are calcium dependent. The third NOS isoforms is an inducible enzyme (iNOS or NOS-2) and is calcium independent; its synthesis is induced mainly by γ-interferon [39]. The production of NO by γ-interferon activated macrophages plays a role in the modulation of the organism immune response.

Cystatin C saturated with inactivated papain, by alkylation, stimulates the production of NO by macrophages, activated by γ-interferon [37,38]. This suggests that the stimulation of NO production by cystatins does not involve the inhibitory site of cystatins, but involves a different site. In support of this contention, synthetic inhibitors of cystatins, like E64 that have been designed according to the protease inhibitory activity sites of cystatins, do not stimulate the production of NO by γ- interferon activated macrophages.

In all, protease inhibitors of the cystatin superfamily, and particularly cystatin C, are involved in inflammation reaction and immunomodulatory activities [40].

The involvement of cystatins in these processes is two fold. On the one hand, the inflammation reaction mediated by cystatins depends on their proteinase inhibitory properties, involved in regulating cysteine proteases, like cathepsins. On the other hand, the immunomodulatory properties of cystatins, like NO synthesis, depend on an active site independent of the proteinase inhibitory activity of cystatins.

This shows that cystatins define a new class of immunomodulatory molecules. The mechanism underlying the activity of cystatins on γ-interferon activated macrophages remains to be determined. The domain involved in the immunomodulatory properties of cystatins would be located in the N-terminal region of cystatin -not involving in the protease inhibitory motifs of cystatins- [40]. This domain would interact with receptors present on macrphages that would mediate NO synthesizing activity [37].

ADULT NEUROGENESIS AND NEURAL STEM CELLS

Contrary to a long-held dogma, neurogenesis occurs in the adult brain and neural stem cells (NSCs) reside in the adult central nervous system (CNS) [41-43]. NSCs are the selfrenewing multipotent cells that generate the main phenotypes of the nervous system, neurons, astrocytes and oligodendrocytes. Neurogenesis occurs primarily in two regions of the adult brain, the dentate gyrus (DG) of the hippocampus and the subventricular zone, along the ventricles [44]. It is hypothesized that newly generated neuronal cells in the adult brain originate from residual stem cells. Despite significant progresses over the last decade, adult NSC is still an elusive cell. The confirmation that neurogenesis occurs in the adult brain has tremendous consequences for our understanding of brain functioning and phsyio- and pathology, as well as for cellular therapy [45].

Neural progenitor and stem cells have been originally isolated and cultured *in vitro*, from adult rodent brains, in presence of high concentration of trophic factors, epidermal growth factor (20 ng/ml) and basic fibroblast growth factor (FGF-2, 20 ng/ml), in serum free defined medium [46-49]. Studies show that unknown factors, present in conditioned medium (CM), are required to support NSC proliferation [49-52].

We purified an autocrine/paracrine factor from the CM of adult rat hippocampal-derived neural progenitor and stem cells. This factor is required for the mitogenic activity of FGF-2 on self-renewing multipotent NSCs *in vitro*, from single cells (figure 1) [53]. The autocrine/paracrine co-factor of FGF-2 is a glycosylated form of the protease inhibitor cystatin C, which N-glycosylation is required for its activity (CCg) (figure 2). CCg is the first autocrine/paracrine factor purified and identified from CM that is required to support self-renewing multipotent NSC proliferation *in vitro*. We reported that the co-delivery of FGF-2 and CCg stimulates neurogenesis in the adult rat DG, and that CCg promotes the growth and proliferation of neural progenitor and stem cells isolated from human *post mortem* tissues and biopsies [52,53]. Therefore, FGF-2 requires interaction with the co-factor CCg for its activity on neural progenitor and stem cells *in vitro*, and to stimulate neurogenesis *in vivo*.

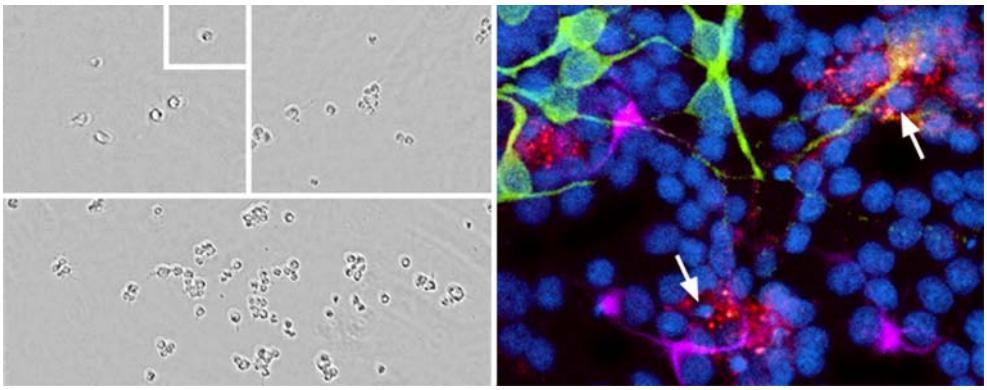


Figure 1. CCg a neural stem cell factor. Adult rat hippocampal-derived neural progenitor and stem cells, plated as single cells and expanded in culture in the presence of basic fibroblast growth factor (FGF-2) and the neural stem cell factor, CCg (left panel). After differentiation (right panel), cells were processed for immunocytofluorescence, with antibodies against class III β-tubulin isotype, a marker of neuronal cells (green), glial fibrillary acidic protein, a marker of astrocytes (purple), and O4, a marker of immature oligodendrocytes (red). The three phenotypes of the adult brain were generated from a single cell, i.e., multipotentiality. The criterion of self-renewal was demonstrated by showing that cells maintained their multipotentiality over time [53]. CCg was purified from conditioned medium of adult rat-derived neural progenitor and stem cell cultures. FGF-2 requires an autocrine/paracrine co-factor, CCg, for its mitogenic activity on self-renewing multipotent NSCs *in vitro*, from single cells. CCg is a glycosylated form of the protease inhibitor cystatin C, which N-glycosylation is required for its activity (CCg). The activity of CCg on self-renewing multipotent NSCs *in vitro* is mediated through its N-glycosylation, independently of its proteinase inhibitory domain.

We showed that the activity of CCg on self-renewing multipotent NSCs *in vitro* and adult hippocampal neurogenesis *in vivo* is mediated through its N-glycosylation, independently of its proteinase inhibitory domain [53]. The mechanisms underlying the activities of FGF-2 and CCg on the stimulation of self-renewing multipotent NSCs *in vitro* and neurogenesis *in vivo* remain to be fully determined. They are likely to be mediated through receptor signaling pathways.

In all, CPIs of the cystatin superfamily, particularly cystatin C, are involved in a broad range of activities, like the regulation of activities of proteases of the families C1 and C13, immunomodulation, inflammation, pathogen infection, the proliferation of self-renewing

multipotent adult- and *post mortem*-derived NSCs *in vitro*, and adult neurogenesis *in vivo*. Some of these activities are unrelated to the CPI activities of the inhibitors of the cystatin superfamily and are mediated by distinct domains of the molecules.

Cystatin C is involved in the proliferation and growth of other types of stem cells, like embryonic stem cells [54]. Cystatin C also inhibits bone resorption [55]. This activity is mediated by its cysteine proteinase inhibitory activity [56]. Cystatin C is involved in the modulation of chemo taxis, phagocytosis and respiratory burst [57-59]. Cystatin M/E is involved in epidermal differentiation [60]. Hence, cystatins are pleiotropic molecules with different activities mediated by different domains of the molecules.

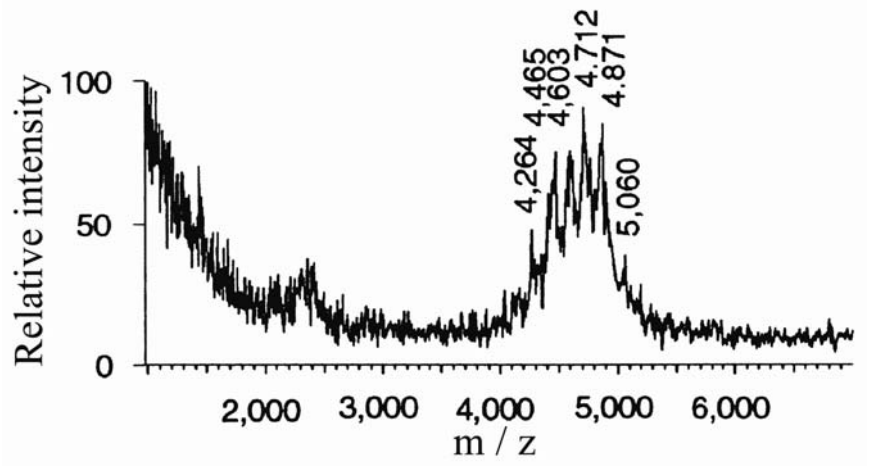


Figure 2. CCg is a glycosylated form of cystatin C. Isolated and purified CCg from conditioned medium of adult rat hippocampal-derived neural progenitor and stem cells, was digested with peptidases, like trypsin and endoproteinase Asp-N. Peptides were analyzed for their sequences and masses. The peptide carrying the N-glycosylation consensus domain, Asn-X-Thr, reveals a pattern of glycosylated peptide when submitted to mass spectral analysis, i.e., multiple peaks reflecting the heterogeneity of the carbohydrate moiety [53]. The same peptide not glycosylated would elicit a single and sharp peak. Mass spectral analysis was carried out by matrix assisted laser desorption mass spectroscopy. Peaks represent the different ionized forms of the peptides. m/z is the mass to charge ratio.

CONCLUSION

CPIs of the cystatin superfamily inhibit cysteine proteases of the family C1. The CPIs of the cystatin superfamily are present in a wide range of animal species, organs and tissues. CPIs of the cystatin superfamily are involved in a broad spectrum of physiological functions, like the regulation of the activity of proteases and their potential deleterious effects on the cellular environment, the protection against bacteria, viruses and parasites. They are also involved in inflammation reaction, immune response and in the proliferation and growth of stem cells, like adult NSCs. They inhibit bone resorption. Some of these activities are unrelated to the CPI activities of the inhibitors of the cystatin superfamily and are mediated by distinct domains of the molecules. The mechanisms underlying the activities of CPIs of

the cystatin superfamily on functions unrelated to their CPI activities remain to be determined and fully understood.

CPIs of the cystatin superfamily, like cystatin C, also inhibit cysteine proteases of the family C13, unrelated to the C1 family. The CPI activity of cystatin C on legumain is mediated via a novel second reactive site of cystatins, distinct from the inhibitory site mediating their activity on proteases of the family C1. In all, CPIs of the cystatin superfamily, and particularly cystatin C, are pleiotropic molecules, with different domains of the molecules involved in different activities and functions. The understanding of the involvement of the CPI activity versus other domains of the molecules has tremendous implications for our understanding of the functioning of CPIs of the cystatin superfamily, as well as for their therapeutic potential.

ACKNOWLEDGMENTS

Chapter 7, figures 1 and 2 are reproduced from: Taupin P, Ray J, Fischer WH, Suhr ST, Hakansson K, Grubb A, Gage FH. FGF-2 responsive neural stem cell proliferation requires CCg, a novel autocrine/paracrine cofactor. Neuron (2000) 28(2): 385-97. Copyright (2000), with permission from Elsevier.

REFERENCES

[1] Turk V, Bode W. (1991) The cystatins: protein inhibitors of cysteine proteinases. *FEBS Lett. 285*, 213-9.

[2] Abrahamson M, Alvarez-Fernandez M, Nathanson CM. Cystatins. (2003) *Biochem Soc Symp. 70*, 179-99.

[3] Abrahamson M, Ritonja A, Brown MA, Grubb A, Machleidt W, Barrett AJ. (1987) Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin. *J Biol Chem. 262*, 9688-94.

[4] Arai S, Watanabe H, Kondo H, Emori Y, Abe K. (1991) Papain-inhibitory activity of oryzacystatin, a rice seed cysteine proteinase inhibitor, depends on the central Gln-Val-Val-Ala-Gly region conserved among cystatin superfamily members. *J Biochem (Tokyo). 109*, 294-8.

[5] Saitoh E, Isemura S, Sanada K, Ohnishi K. (1991) Cystatins of family II are harboring two domains which retain inhibitory activities against the proteinases. *Biochem Biophys Res Commun. 175*, 1070-5.

[6] Hall A, Dalboge H, Grubb A, Abrahamson M. (1993) Importance of the evolutionarily conserved glycine residue in the N-terminal region of human cystatin C (Gly-11) for cysteine endopeptidase inhibition. *Biochem J. 291*, 123-9.

[7] Hall A, Hakansson K, Mason RW, Grubb A, Abrahamson M. (1995) Structural basis for the biological specificity of cystatin C. Identification of leucine 9 in the N-terminal binding region as a selectivity-conferring residue in the inhibition of mammalian cysteine peptidases. *J Biol Chem. 270*, 5115-21.

[8] Bode W, Engh R, Musil D, Thiele U, Huber R, Karshikov A, Brzin J, Kos J, Turk V. (1988) The 2.0 A X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J. 7*, 2593-9.

[9] Kozak M, Jankowska E, Janowski R, Grzonka Z, Grubb A, Alvarez Fernandez M, Abrahamson M, Jaskolski M. (1999) Expression of a selenomethionyl derivative and preliminary crystallographic studies of human cystatin C. *Acta Crystallogr D Biol Crystallogr. 55*, 1939-42.

[10] Schwabe C, Anastasi A, Crow H, McDonald JK, Barrett AJ. (1984) Cystatin. Amino acid sequence and possible secondary structure. *Biochem J. 217*, 813-7.

[11] Machleidt W, Thiele U, Laber B, Assfalg-Machleidt I, Esterl A, Wiegand G, Kos J, Turk V, Bode W. (1989) Mechanism of inhibition of papain by chicken egg white cystatin. Inhibition constants of N-terminally truncated forms and cyanogen bromide fragments of the inhibitor. *FEBS Lett. 243*, 234-8.

[12] Bjork I, Ylinenjarvi K, Lindahl P. (1990) Equilibrium and kinetic studies of the interaction of chicken cystatin with four cysteine proteinases. *Biol Chem Hoppe Seyler. 371*, 119-24.

[13] Alvarez-Fernandez M, Barrett AJ, Gerhartz B, Dando PM, Ni J, Abrahamson M. (1999) Inhibition of mammalian legumain by some cystatins is due to a novel second reactive site. *J Biol Chem. 274*, 19195-203.

[14] Chen JM, Dando PM, Stevens RA, Fortunato M, Barrett AJ. (1998) Cloning and expression of mouse legumain, a lysosomal endopeptidase. *Biochem J. 335*, 111-7.

[15] Yasuhara O, Hanai K, Ohkubo I, Sasaki M, McGeer PL, Kimura H. (1993) Expression of cystatin C in rat, monkey and human brains. *Brain Res. 628*, 85-92.

[16] Abrahamson M, Grubb A, Olafsson I, Lundwall A. (1987) Molecular cloning and sequence analysis of cDNA coding for the precursor of the human cysteine proteinase inhibitor cystatin C. *FEBS Lett. 216*, 229-33.

[17] Isemura S, Saitoh E, Sanada K, Minakata K. (1991) Identification of full-sized forms of salivary (S-type) cystatins (cystatin SN, cystatin SA, cystatin S, and two phosphorylated forms of cystatin S) in human whole saliva and determination of phosphorylation sites of cystatin S. *J Biochem (Tokyo).110*, 648-54.

[18] Ni J, Abrahamson M, Zhang M, Fernandez MA, Grubb A, Su J, Yu GL, Li Y, Parmelee D, Xing L, Coleman TA, Gentz S, Thotakura R, Nguyen N, Hesselberg M, Gentz R. (1997) Cystatin E is a novel human cysteine proteinase inhibitor with structural resemblance to family 2 cystatins. *J Biol Chem. 272*, 10853-8.

[19] Halfon S, Ford J, Foster J, Dowling L, Lucian L, Sterling M, Xu Y, Weiss M, Ikeda M, Liggett D, Helms A, Caux C, Lebecque S, Hannum C, Menon S, McClanahan T, Gorman D, Zurawski G. (1998) Leukocystatin, a new Class II cystatin expressed selectively by hematopoietic cells. *J Biol Chem. 273*, 16400-8.

[20] Uchiyama Y. (2001) Autophagic cell death and its execution by lysosomal cathepsins. *Arch Histol Cytol. 64*, 233-46.

[21] Vasiljeva O, Reinheckel T, Peters C, Turk D, Turk V, Turk B. (2007) Emerging roles of cysteine cathepsins in disease and their potential as drug targets. *Curr Pharm Des. 13*, 385-401.

[22] Soares MB, Garcia S, Campos de Carvalho AC, Ribeiro dos Santos R. (2007) Cellular therapy in Chagas' disease: potential applications in patients with chronic cardiomyopathy. *Regen Med. 2*, 257-64.

[23] Hornak V, Simmerling C. (2007) Targeting structural flexibility in HIV-1 protease inhibitor binding. *Drug Discov Today. 12*, 132-8.

[24] Kido H, Okumura Y, Yamada H, Le TQ, Yano M. (2007) Proteases essential for human influenza virus entry into cells and their inhibitors as potential therapeutic agents. *Curr Pharm Des. 13*, 403-12.

[25] Butler SM, Festa RA, Pearce MJ, Darwin KH. (2006) Self-compartmentalized bacterial proteases and pathogenesis. *Mol Microbiol. 60*, 553-62.

[26] Stoka V, Nycander M, Lenarcic B, Labriola C, Cazzulo JJ, Bjork I, Turk V. (1995) Inhibition of cruzipain, the major cysteine proteinase of the protozoan parasite, Trypanosoma cruzi, by proteinase inhibitors of the cystatin superfamily. *FEBS Lett. 370*, 101-4.

[27] Jose Cazzulo J, Stoka V, Turk V. (2001) The major cysteine proteinase of Trypanosoma cruzi: a valid target for chemotherapy of Chagas disease. *Curr Pharm Des. 7*, 1143-56.

[28] Henskens YM, van der Velden U, Veerman EC, Nieuw Amerongen AV. (1993) Protein, albumin and cystatin concentrations in saliva of healthy subjects and of patients with gingivitis or periodontitis. *J Periodontal Res. 28*, 43-8.

[29] Henskens YM, Veerman EC, Mantel MS, van der Velden U, Nieuw Amerongen AV. (1994) Cystatins S and C in human whole saliva and in glandular salivas in periodontal health and disease. *J Dent Res. 73*, 1606-14.

[30] Henskens YM, van der Weijden FA, van den Keijbus PA, Veerman EC, Timmerman MF, van der Velden U, Amerongen AV. (1996) Effect of periodontal treatment on the protein composition of whole and parotid saliva. *J Periodontol. 67*, 205-12.

[31] Blankenvoorde MF, Henskens YM, van't Hof W, Veerman EC, Nieuw Amerongen AV. (1996) Inhibition of the growth and cysteine proteinase activity of Porphyromonas gingivalis by human salivary cystatin S and chicken cystatin. *Biol Chem. 377*, 847-50.

[32] Blankenvoorde MF, van't Hof W, Walgreen-Weterings E, van Steenbergen TJ, Brand HS, Veerman EC, Nieuw Amerongen AV. (1998) Cystatin and cystatin-derived peptides have antibacterial activity against the pathogen Porphyromonas gingivalis. *Biol Chem. 379*, 1371-5.

[33] Armstrong PB. (2006) Proteases and protease inhibitors: a balance of activities in host-pathogen interaction. *Immunobiology. 211*, 263-81.

[34] Leung-Tack J, Tavera C, Er-Raki A, Gensac MC, Colle A. (1990) Rat cystatin C: inhibitor of granulocyte phagocytic functions. *Biol Chem Hoppe Seyler. 371*, 255-8.

[35] Leung-Tack J, Tavera C, Gensac MC, Martinez J, Colle A. (1990) Modulation of phagocytosis-associated respiratory burst by human cystatin C: role of the N-terminal tetrapeptide Lys-Pro-Pro-Arg. *Exp Cell Res.* 188, 16-22.

[36] Kato T, Imatani T, Miura T, Minaguchi K, Saitoh E, Okuda K. (2000) Cytokine-inducing activity of family 2 cystatins. *Biol Chem. 381*, 1143-7.

[37] Verdot L, Lalmanach G, Vercruysse V, Hoebeke J, Gauthier F, Vray B. (1999) Chicken cystatin stimulates nitric oxide release from interferon-gamma-activated mouse peritoneal macrophages via cytokine synthesis. *Eur J Biochem. 266*, 1111-7.

[38] Das L, Datta N, Bandyopadhyay S, Das PK. (2001) Successful therapy of lethal murine visceral leishmaniasis with cystatin involves up-regulation of nitric oxide and a favorable T cell response. *J Immunol. 166*, 4020-8.

[39] Maniscalco M, Sofia M, Pelaia G. (2007) Nitric oxide in upper airways inflammatory diseases. *Inflamm Res. 56*, 58-69.

[40] Vray B, Hartmann S, Hoebeke J. (2002) Immunomodulatory properties of cystatins. *Cell Mol Life Sci. 59*, 1503-12.

[41] Gage FH. (2000) Mammalian neural stem cells. *Science. 287*, 1433-8.

[42] Gross CG. (2000) Neurogenesis in the adult brain: death of a dogma. *Nat Rev Neurosci. 1*, 67-73.

[43] Kaplan MS. (2001) Environment complexity stimulates visual cortex neurogenesis: death of a dogma and a research career. *Trends Neurosci. 24*, 617-20.

[44] Taupin P, Gage FH. (2002) Adult neurogenesis and neural stem cells of the central nervous system in mammals. *J Neurosci Res. 69*, 745-9.

[45] Taupin P. (2006) The therapeutic potential of adult neural stem cells. *Curr Opin Mol Ther. 8*, 225-31.

[46] Reynolds, BA, Weiss S. (1992) Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science. 255*, 1707-10.

[47] Gage FH, Coates PW, Palmer TD, Kuhn HG, Fisher LJ, Suhonen JO, Peterson DA, Suhr ST, Ray J. (1995) Survival and differentiation of adult neuronal progenitor cells transplanted to the adult brain. *Proc Natl Acad Sci USA. 92*, 11879-83.

[48] Gritti A, Parati EA, Cova L, Frolichsthal P, Galli R, Wanke E, Faravelli L, Morassutti DJ, Roisen F, Nickel DD, Vescovi AL. (1996) Multipotential stem cells from the adult mouse brain proliferate and self-renew in response to basic fibroblast growth factor. *J Neurosci. 16*, 1091-100.

[49] Palmer TD, Takahashi J, Gage FH. (1997) The adult rat hippocampus contains primordial neural stem cells. *Mol Cell Neurosci. 8*, 389-404.

[50] Johansson CB, Momma S, Clarke DL, Risling M, Lendahl U, Frisen J. (1999) Identification of a neural stem cell in the adult mammalian central nervous system. *Cell. 96*, 25-34.

[51] Johansson CB, Svensson M, Wallstedt L, Janson AM, Frisen J. (1999) Neural stem cells in the adult human brain. *Exp Cell Res. 253*, 733-6.

[52] Palmer TD, Schwartz PH, Taupin P, Kaspar B, Stein SA, Gage FH. (2001) Cell culture. Progenitor cells from human brain after death. *Nature. 411*, 42-3.

[53] Taupin P, Ray J, Fischer WH, Suhr ST, Hakansson K, Grubb A, Gage FH. (2000) FGF-2 responsive neural stem cell proliferation requires CCg, a novel autocrine/paracrine cofactor. *Neuron. 28*, 385-97.

[54] Kato T, Heike T, Okawa K, Haruyama M, Shiraishi K, Yoshimoto M, Nagato M, Shibata M, Kumada T, Yamanaka Y, Hattori H, Nakahata T. (2006) A neurosphere-derived factor, cystatin C, supports differentiation of ES cells into neural stem cells. *Proc Natl Acad Sci U S A. 103*, 6019-24.

[55] Lerner UH, Grubb A. (1992) Human cystatin C, a cysteine proteinase inhibitor, inhibits bone resorption in vitro stimulated by parathyroid hormone and parathyroid hormone-related peptide of malignancy. *J Bone Miner Res. 7*, 433-40.

[56] Johansson L, Grubb A, Abrahamson M, Kasprzykowski F, Kasprzykowska R, Grzonka Z, Lerner UH. (2000) A peptidyl derivative structurally based on the inhibitory center of cystatin C inhibits bone resorption in vitro. *Bone. 26*, 451-9.

[57] Warfel AH, Zucker-Franklin D, Frangione B, Ghiso J. (1987) Constitutive secretion of cystatin C (gamma-trace) by monocytes and macrophages and its downregulation after stimulation. *J Exp Med. 166*, 1912-7.

[58] Chapman HA Jr, Reilly JJ Jr, Yee R, Grubb A. (1990) Identification of cystatin C, a cysteine proteinase inhibitor, as a major secretory product of human alveolar macrophages in vitro. *Am Rev Respir Dis. 141*, 698-705.

[59] Assfalg-Machleidt I, Billing A, Frohlich D, Nast-Kolb D, Joka T, Jochum M, Machleidt W. (1992) The role of the kininogens as cysteine proteinase inhibitors in local and systemic inflammation. *Agents Actions Suppl. 38*, 312-21.

[60] Zeeuwen PL. (2004) Epidermal differentiation: the role of proteases and their inhibitors. *Eur J Cell Biol. 83*, 761-73.

Chapter VIII

PATHOLOGY AND THERAPY

ABSTRACT

Protease inhibitors of the cystatin superfamily are involved in a broad range of physiological functions, like modulation of protease activities, host-pathogen reactions, immunomodulatory activities and inflammation reactions, stem cells growth and differentiation, and bone resorption. The protease inhibitory activity of the cysteine protease inhibitors (CPIs) of the cystatin superfamily, and particularly cystatin C, mediates these functions, but other domains of the molecules are also involved. Cystatins are amyloidogenic proteins, under pathological conditions they contribute to the formation of amyloid deposits, conditions known as amyloidosis. Among them, protease inhibitors of the cystatin superfamily are involved in Alzheimer's disease (AD), hereditary cystatin C amyloid angiopathy (HCCAA), hereditary cerebral hemorrhage with amyloidosis-Dutch type (HCHWA-Dutch type). Protease inhibitors of the cystatin superfamily are also involved in other pathologies, like age-related macular degeneration, epilepsy, as well as in cancer and tumor formation. The understanding of the involvement of protease inhibitors of the cystatin superfamily in pathological processes will lead to the development of new drugs and therapeutic strategies.

INTRODUCTION

Cystatin C is member of family 2 of cysteine proteinase inhibitors (CPIs) of the cystatin superfamily [1-3]. Proteinase inhibitors of the cystatin superfamily are reversible inhibitors of cysteine proteases of the papain superfamily or family C1, like papain, ficin, and cathepsins B, H, L, S and K [4, 5].

Cystatin C is involved in conditions known as amyloidosis, a group of diseases that result from abnormal amyloid fibril deposits, like AD, HCCAA and HCHWA-Dutch [6]. Amyloidosis are characterized by histological staining of amyloid deposits, by procedure, like Congo red anisotropic staining, Congo red (X-34) and thioflavin-T or -S fluorescence [7-10]. Congo red stain is a histological stain, pH sensitive, useful for detecting amyloid

deposits [10]. Congo red gives a green birefringence, characteristic of amyloid, under polarized light.

Cystatin C promotes the formation of amyloid deposits and binds to other amyloid proteins, like β-amyloid [11]. Cystatin C is also an amyloidogenic protein. Like some other 20 molecules, cystatin C, under pathological conditions, aggregates to form amyloid deposits [12, 13]. The propensity of cystatin C to form amyloid deposits is related to its ability to dimerize through three-dimensional domain swapping [14, 15].

In all, cystatin C may be involved in amyloidosis by its promoting activity on amyloid formation and its potential to form dimmers through three-dimensional domain swapping. Protease inhibitors of the cystatin superfamily are also involved in other pathologies, like epilepsy and age-related macular degeneration, as well as in cancer and tumor formation. The involvement of proteinase inhibitors of the cystatin superfamily has tremendous implication for therapy and drug design.

ALZHEIMER'S DISEASE

AD is a progressive neurodegenerative disease characterized in the brain by amyloid plaque deposits and neurofibrillary tangles [16]. The early-onset form of the disease is primarily genetic of origin and inherited. It is a very rare form of the disease, referred as familial form of AD when inherited. The late-onset, over age 65, is not inherited and is the most common type of dementia among older people. It is also referred as sporadic form of AD. Mutations in the β-amyloid protein precursor protein (APP) underlie the pathogenesis of most cases of early-onset AD. APP is an amyloidogenic protein; mutations in APP are believed to induce the formation of amyloid deposits. AD is the fourth highest cause of death in the developed world [17]. More than 20 millions of people worldwide are diagnose with AD. It is estimated that the number of patients with AD will quadruple by 2050 (one in 85 people).

Amyloid deposits are characterized by the co-deposition of cystatin C and protein β-amyloid in senile plaques in the brain of AD patients and in blood vessels [18, 19]. Amyloid deposits are characterized by the co-deposition of cystatin C and protein β-amyloid in cerebral amyloid plaques of transgenic mice carrying the AbPP Swedish mutation (K670N-M671L mutation), a mouse model of AD [20]. This shows that, in AD, cystatin C promotes the formation of amyloid deposits or binds to β-amyloid fibrils [11]. In the brain, the maturation of APP into β-amyloid is controlled primarily by two proteases, β- and γ-secretases [21]. γ-Secretases are serine and cysteine proteases. Cystatin C may contribute to the formation of amyloid deposits from APP through its CPI activity on γ-secretases [22].

The involvement of human cystatin C (hCC) in the pathogenesis of AD is emphasized by studies of genetic polymorphism of CTS3, the gene coding for hCC located on chromosome 20 [23, 24]. Several studies have reported the association of a polymorphism of CTS3 with late-onset or sporadic form of AD [25-27]. This study predicts the Ala25Thr substitution in the signal peptide of hCC to be a recessive risk for late-onset AD [25]. The mutated form of hCC displayed a reduced secretion due to a less efficient clivage of the signal peptide [27]. In the brain, the reduced level of hCC may represent the molecular factor responsible for the

increased risk of AD, possibly through the activity oh hCC on γ-secretases. Reduced hCC activity may promote the formation of β-amyloid and its aggregation. Others studies reveal conflicting data on genetic polymorphism association between hCC gene and AD [28-31].

In all, cystatin C is involved in the pathogenesis of AD and particularly in the formation of amyloid deposits, but its involvement in the disease remain to be further determined.

HEREDITARY CYSTATIN C AMYLOID ANGIOPATHY

HCCAA or hereditary cerebral hemorrhage with amyloidosis is a rare fatal form of cerebral amyloid angiopathy (CAA) in young people in Iceland. CCA is a heterogeneous group of disorders characterized clinically by ischemic and/or hemorrhagic strokes, and histologically by deposition of amyloid in cerebral cortical blood vessels [32]. HCCAA patients suffer at an early age from repeated cerebral hemorrhages. HCCAA leads to amyloid deposits in brain vessels of young adult leading to lethal cerebral hemorrhage [33]. Amyloid deposits have also been found in lymph nodes, spleen, salivary gland, seminal vesicles and skin [34-36].

HCCAA is a genetic disease, it is caused by a mutation in cystatin C. HCCAA is a hereditary disease; it is an autosomal dominant disease [37]. The causative factor of HCCAA is a point mutation within the CST3 gene leading to a Leu-Gln substitution at position 68 (L68Q) in hCC [38-41]. Mutant L68Q hCC N-terminally truncated of the 10 NH2-terminal residues is the major component of amyloid deposits in patients with HCCAA [42].

Cystatin C is a monomer in its native functional state [43]. Cystatin C is an amyloidogenic protein. Under mild predenaturing conditions and in mutant proteins, like the variant L68Q, cystatin C forms dimmers. A process of three-dimensional domain swapping has been proposed for the dimerization of cystatin C [44, 45]. Three dimensional domain swapping is a process for forming oligomeric, dimeric or multimeric, proteins, believed to be a process by which amyloid fibrils form [46]. Dimmers of L68Q cystatin C are formed at denaturing temperature lower than needed for wild-type cystatin [47]. This suggests that in HCCAA, the formation of amyloid fibrils results from oligomerization of mutated L68Q variant of cystatin C. The cystatin variant L68Q would undergo three-dimensional domain swapping, leading to the deposit of amyloid [46, 48].

Dimeric cystatin C lacks cysteine proteinase inhibitory activity [47, 48]. The resulting lowered cystatin C activity may result in an enhanced proteolytic activity of cysteine proteases. Enhanced proteolytic activity may increase the risk of cerebral hemorrhage in HCCAA patients.

In all, in patients with HCCAA, cystatin C is deposited as amyloid fibrils within the walls of cerebral arteries, causing massive amyloid deposits in brain arteries of young adult, leading to lethal cerebral hemorrhage. The increase in proteolytic activity may contribute to increased risk of cerebral hemorrhage in HCCAA patients.

HEREDITARY CEREBRAL HEMORRHAGE WITH AMYLOIDOSIS-DUTCH TYPE

HCHWA-Dutch type is rare hereditary disorder that causes severe CAA, with hemorrhagic strokes of mid-life onset and dementia [49, 50]. HCHWA-Dutch type is caused by a mutation in the gene APP, causing the protein β-amyloid to aggregate and form deposits of amyloid.

HCHWA-Dutch type is characterized by the co-deposition of cystatin C and protein β-amyloid in blood vessels [50, 51]. In HCHWA-Dutch type, cystatin C promotes the formation of amyloid deposits or binds to β-amyloid fibrils.

Cystatin C is involved in other forms of amyloidosis, like sporadic inclusion-body myositis. Sporadic inclusion-body myositis is the most common muscle disease of older persons [52]. It is a progressive debilitating disease with amyloid deposits. Cystatin C has been reported to be associated with β-amyloid deposits in muscle fibers of patients with sporadic inclusion-body myositis [53]. This suggests that cystatin C may be involved in the pathogenesis of the disease [54]. Cystatin C is also involved in age-related macular degeneration (AMD) and epilepsy. A polymorphism in the leader sequence of cystatin C is associated with increased risk for development of oxidative age-related macular degeneration [55]. The mutation would interfere with the processing and secretion of cystatin C in eye cells and in turn to the etiology of oxidative AMD [56]. The expression of the cysteine protease inhibitor cystatin C is up-regulated in the brain of patients with epilepsy [56, 57]. In all, the CPI of the cystatin superfamily, cystatin C, is involved in a broad range of pathological conditions.

CANCERS

Tumor invasion and metastasis involve the migration of cancer cells over long distances. The mechanisms underlying metastasis are not fully understood. It involves particularly the concerted actions of proteases and their inhibitors, allowing the cells to detach and migrate over long distances.

Lysosomal proteases, like cathepsin B and L, are expressed at the surface of tumor cells and are being misrouted for secretion in malignant cells [59-62]. Lysosomal proteases may contribute to tumor progression by promoting cell migration and degrading extra cellular matrix components. The involvement of cysteine proteases of the papain superfamily in tumor progression is highlighted by the involvement of CPIs of the cystatin superfamily in tumor cell invasion and metastasis. CPIs, and particularly hCC, prevent tumor cell invasion and metastasis *in vitro* and *in vivo* [63, 64]. It is proposed that a decrease in CPI activity may be one of the factor underlying tumor invasion and metastasis [65].

In all, CPIs of the cystatin superfamily are involved in a broad range of pathological conditions. This has tremendous implications for therapy and drug design.

THERAPY

CPIs of the cystatin superfamily are pleiotropic molecules involved in a broad range of physio- and pathological conditions. Results show that some of these activities are unrelated to the CPI activity of the inhibitors and are mediated by distinct domains of the molecules. The identification of the mechanisms underlying the activities of the CPIs of the cystatin superfamily is therefore a prerequisite before developing therapeutic strategies involving CPIs of the cystatin superfamily.

Therapeutic strategies aiming at targeting cysteine proteases of the papain superfamily consist in using potent synthetic inhibitors directed against the proteases. To this aim, synthetic inhibitors have been synthesized based on the N-terminal sequence of hCC, Arg8-Leu9-Val10-Gly11, among them, E-64, leupeptin, benzyloxycarbonyl-Phe-Ala-diazomethane (Z-Phe-Ala-CHN2) and benzyloxycarbonyl-Arg-Leu-Val-Gly-diazomethane (Z-RLVG-CHN2) [66, 67].

Proteases and their inhibitors are ubiquitous, they are involved in basically all physio- and pathological processes. Devising synthetic inhibitors, antibodies, antagonists, is crucial to control their pathological activities, like in amyloidosis, neurological and immunological diseases, tumors and viral infections [68-75].

CONCLUSION

CPIs of the cystatin superfamily are involved in a broad range of pathological conditions. The involvement of CPIs of the cystatin superfamily, like hCC, in pathological conditions may involve their protease inhibitory activities, their amyloidogenic property or other domains of the molecules. Hence, determining the mechanisms underlying their activities is a prerequisite for developing therapeutic strategies involving CPIs of the cystatin superfamily.

Synthetic inhibitors of cysteine proteases of the papain superfamily have been devised, based on motifs of CPIs involved in their active site. They offer promising strategies to treat a broad range of pathological conditions, including cancers. New components aiming at destabilizing and preventing amyloid formation from hCC may offer other opportunities for the treatment of amyloidosis involving CPIs of the cystatin superfamily. In all, therapeutic applications involving CPIs of the cystatin superfamily offer a tremendous hope for the treatment of a broad range of pathological conditions, like amyloidosis, neurological and immunological diseases, pathogen infections and cancers. To this aim three-dimensional determination of the domains involved in a particular activity may contribute to design more selective and specific inhibitors.

REFERENCES

[1] Muller-Esterl W, Fritz H, Kellermann J, Lottspeich F, Machleidt W, Turk V. (1985) Genealogy of mammalian cysteine proteinase inhibitors. Common evolutionary origin of stefins, cystatins and kininogens. *FEBS Lett. 191*, 221-6.

[2] Turk V, Bode W. (1991) The cystatins: protein inhibitors of cysteine proteinases. *FEBS Lett. 285*, 213-9.

[3] Abrahamson M, Alvarez-Fernandez M, Nathanson CM. Cystatins. (2003) *Biochem Soc Symp. 70*, 179-99.

[4] Barrett AJ, Fritz H, Grubb A, Isemura S, Jarvinen M, Katunuma N, Machleidt W, Muller-Esterl W, Sasaki M, Turk V. (1986) Nomenclature and classification of the proteins homologous with the cysteine-proteinase inhibitor chicken cystatin. *Biochem J. 236*, 312.

[5] Nagler DK, Menard R. (2003) Family C1 cysteine proteases: biological diversity or redundancy? *Biol Chem. 384*, 837-43.

[6] Kyle RA. (1995) Amyloidosis. *Circulation. 91*, 1269-71.

[7] Elghetany MT, Saleem A. (1988) Methods for staining amyloid in tissues: a review. *Stain Technol. 63*, 201-12.

[8] Linke RP. (2000) Highly sensitive diagnosis of amyloid and various amyloid syndromes using Congo red fluorescence. *Virchows Arch. 436*, 439-48.

[9] Styren SD, Hamilton RL, Styren GC, Klunk WE. (2000) X-34, a fluorescent derivative of Congo red: a novel histochemical stain for Alzheimer's disease pathology. *J Histochem Cytochem. 48*, 1223-32.

[10] Frid P, Anisimov SV, Popovic N. (2007) Congo red and protein aggregation in neurodegenerative diseases. *Brain Res Rev. 53*, 135-60.

[11] Levy E, Sastre M, Kumar A, Gallo G, Piccardo P, Ghetti B, Tagliavini F. (2001) Codeposition of cystatin C with amyloid-beta protein in the brain of Alzheimer disease patients. *J Neuropathol Exp Neurol. 60*, 94-104.

[12] Dobson CM. (1999) Protein misfolding, evolution and disease. *Trends Biochem Sci. 24*, 329-32.

[13] Sipe JD, Cohen AS. (2000) Review: history of the amyloid fibril. *J Struct Biol. 130*, 88-98.

[14] Janowski R, Kozak M, Jankowska E, Grzonka Z, Grubb A, Abrahamson M, Jaskolski M. (2001) Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping. *Nat Struct Biol. 8*, 316-20.

[15] Staniforth RA, Giannini S, Higgins LD, Conroy MJ, Hounslow AM, Jerala R, Craven CJ, Waltho JP. (2001) Three-dimensional domain swapping in the folded and molten-globule states of cystatins, an amyloid-forming structural superfamily. *EMBO J. 20*, 4774-81.

[16] St George-Hyslop PH, Petit A. (2005) Molecular biology and genetics of Alzheimer's disease. *C R Biol. 328*, 119-30.

[17] Caselli RJ, Beach TG, Yaari R, Reiman EM. (2006) Alzheimer's disease a century later. *J Clin Psychiatry. 67*, 1784-800.

[18] Vinters HV, Nishimura GS, Secor DL, Pardridge WM. (1990) Immunoreactive A4 and gamma-trace peptide colocalization in amyloidotic arteriolar lesions in brains of patients with Alzheimer's disease. *Am J Pathol. 137*, 233-40.

[19] Levy E, Sastre M, Kumar A, Gallo G, Piccardo P, Ghetti B, Tagliavini F. (2001) Codeposition of cystatin C with amyloid-beta protein in the brain of Alzheimer disease patients. *J Neuropathol Exp Neurol. 60*, 94-104.

[20] Steinhoff T, Moritz E, Wollmer MA, Mohajeri MH, Kins S, Nitsch RM. (2001) Increased cystatin C in astrocytes of transgenic mice expressing the K670N-M671L mutation of the amyloid precursor protein and deposition in brain amyloid plaques. *Neurobiol Dis. 8*, 647-54.

[21] Selkoe DJ. (2001) Alzheimer's disease: genes, proteins, and therapy. *Physiol Rev. 81*, 741-66.

[22] Figueiredo-Pereira ME, Efthimiopoulos S, Tezapsidis N, Buku A, Ghiso J, Mehta P, Robakis NK. (1999) Distinct secretases, a cysteine protease and a serine protease, generate the C termini of amyloid beta-proteins Abeta1-40 and Abeta1-42, respectively. *J Neurochem. 72*, 1417-22.

[23] Saitoh E, Sabatini LM, Eddy RL, Shows TB, Azen EA, Isemura S, Sanada K. (1989) The human cystatin C gene (CST3) is a member of the cystatin gene family which is localized on chromosome 20. *Biochem Biophys Res Commun. 162*, 1324-31.

[24] Schnittger S, Rao VV, Abrahamson M, Hansmann I. (1993) Cystatin C (CST3), the candidate gene for hereditary cystatin C amyloid angiopathy (HCCAA), and other members of the cystatin gene family are clustered on chromosome 20p11.2. *Genomics. 16*, 50-5. Erratum in: (1993) *Genomics 17*, 533.

[25] Finckh U, von der Kammer H, Velden J, Michel T, Andresen B, Deng A, Zhang J, Muller-Thomsen T, Zuchowski K, Menzer G, Mann U, Papassotiropoulos A, Heun R, Zurdel J, Holst F, Benussi L, Stoppe G, Reiss J, Miserez AR, Staehelin HB, Rebeck GW, Hyman BT, Binetti G, Hock C, Growdon JH, Nitsch RM. (2000) Genetic association of a cystatin C gene polymorphism with late-onset Alzheimer disease. *Arch Neurol. 57*, 1579-83.

[26] Crawford FC, Freeman MJ, Schinka JA, Morris MD, Abdullah LI, Richards D, Sevush S, Duara R, Mullan MJ. (2001) Association between Alzheimer's disease and a functional polymorphism in the Myeloperoxidase gene. *Exp Neurol. 167*, 456-9.

[27] Benussi L, Ghidoni R, Steinhoff T, Alberici A, Villa A, Mazzoli F, Nicosia F, Barbiero L, Broglio L, Feudatari E, Signorini S, Finckh U, Nitsch RM, Binetti G. (2003) Alzheimer disease-associated cystatin C variant undergoes impaired secretion. *Neurobiol Dis. 13*, 15-21.

[28] Beyer K, Lao JI, Gomez M, Riutort N, Latorre P, Mate JL, Ariza A. (2001) Alzheimer's disease and the cystatin C gene polymorphism: an association study. *Neurosci Lett. 315*, 17-20.

[29] Maruyama H, Izumi Y, Oda M, Torii T, Morino H, Toji H, Sasaki K, Terasawa H, Nakamura S, Kawakami H. (2001) Lack of an association between cystatin C gene polymorphisms in Japanese patients with Alzheimer's disease. *Neurology. 57*, 337-9.

[30] Roks G, Cruts M, Slooter AJ, Dermaut B, Hofman A, Van Broeckhoven C, Van Duijn CM. (2001) The cystatin C polymorphism is not associated with early onset Alzheimer's disease. *Neurology. 57*, 366-7.

[31] Dodel RC, Du Y, Depboylu C, Kurz A, Eastwood B, Farlow M, Oertel WH, Muller U, Riemenschneider M. (2002) A polymorphism in the cystatin C promoter region is not associated with an increased risk of AD. *Neurology. 58*, 664.

[32] Coria F, Rubio I. (1996) Cerebral amyloid angiopathies. *Neuropathol Appl Neurobiol. 22*, 216-27.

[33] Olafsson I, Grubb A. (2000) Hereditary cystatin C amyloid angiopathy. *Amyloid. 7*, 70-9.

[34] Lofberg H, Grubb AO, Nilsson EK, Jensson O, Gudmundsson G, Blondal H, Arnason A, Thorsteinsson L. (1987) Immunohistochemical characterization of the amyloid deposits and quantitation of pertinent cerebrospinal fluid proteins in hereditary cerebral hemorrhage with amyloidosis. *Stroke. 18*, 431-40.

[35] Benedikz E, Blondal H, Gudmundsson G. (1990) Skin deposits in hereditary cystatin C amyloidosis. *Virchows Arch A Pathol Anat Histopathol. 417*, 325-31.

[36] Thorsteinsson L, Georgsson G, Asgeirsson B, Bjarnadottir M, Olafsson I, Jensson O, Gudmundsson G. (1992) On the role of monocytes/macrophages in the pathogenesis of central nervous system lesions in hereditary cystatin C amyloid angiopathy. *J Neurol Sci. 108*, 121-8.

[37] Palsdottir A, Snorradottir AO, Thorsteinsson L. (2006) Hereditary cystatin C amyloid angiopathy: genetic, clinical, and pathological aspects. *Brain Pathol. 16*, 55-9.

[38] Jensson O, Gudmundsson G, Arnason A, Blondal H, Petursdottir I, Thorsteinsson L, Grubb A, Lofberg H, Cohen D, Frangione B. (1987) Hereditary cystatin C (gamma-trace) amyloid angiopathy of the CNS causing cerebral hemorrhage. *Acta Neurol Scand. 76*, 102-14.

[39] Palsdottir A, Abrahamson M, Thorsteinsson L, Arnason A, Olafsson I, Grubb A, Jensson O. (1988) Mutation in cystatin C gene causes hereditary brain haemorrhage. *Lancet. 2*, 603-4.

[40] Levy E, Lopez-Otin C, Ghiso J, Geltner D, Frangione B. (1989) Stroke in Icelandic patients with hereditary amyloid angiopathy is related to a mutation in the cystatin C gene, an inhibitor of cysteine proteases. *J Exp Med. 169*, 1771-8.

[41] Abrahamson M, Jonsdottir S, Olafsson I, Jensson O, Grubb A. (1992) Hereditary cystatin C amyloid angiopathy: identification of the disease-causing mutation and specific diagnosis by polymerase chain reaction based analysis. *Hum Genet. 89*, 377-80.

[42] Ghiso J, Jensson O, Frangione B. (1986) Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Icelandic type is a variant of gamma-trace basic protein (cystatin C). *Proc Natl Acad Sci U S A. 83*, 2974-8.

[43] Bode W, Engh R, Musil D, Thiele U, Huber R, Karshikov A, Brzin J, Kos J, Turk V. (1988) The 2.0 A X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J. 7*, 2593-9.

[44] Staniforth RA, Giannini S, Higgins LD, Conroy MJ, Hounslow AM, Jerala R, Craven CJ, Waltho JP. (2001) Three-dimensional domain swapping in the folded and molten-

globule states of cystatins, an amyloid-forming structural superfamily. *EMBO J. 20*, 4774-81.

[45] Janowski R, Kozak M, Jankowska E, Grzonka Z, Grubb A, Abrahamson M, Jaskolski M. (2001) Human cystatin C, an amyloidogenic protein, dimerizes through three-dimensional domain swapping. *Nat Struct Biol. 8*, 316-20.

[46] Jaskolski M. (2001) 3D domain swapping, protein oligomerization, and amyloid formation. *Acta Biochim Pol. 48*, 807-27.

[47] Abrahamson M, Grubb A. (1994) Increased body temperature accelerates aggregation of the Leu-68-->Gln mutant cystatin C, the amyloid-forming protein in hereditary cystatin C amyloid angiopathy. *Proc Natl Acad Sci U S A. 91*, 1416-20.

[48] Ekiel I, Abrahamson M. (1996) Folding-related dimerization of human cystatin C. *J Biol Chem. 271*, 1314-21.

[49] Maat-Schieman M, Roos R, van Duinen S. (2006) The cerebral beta-amyloid angiopathies: hereditary and sporadic. *Brain Pathol. 16*, 30-9.

[50] Haan J, Roos RA. (1992) Comparison between the Icelandic and Dutch forms of hereditary cerebral amyloid angiopathy. *Clin Neurol Neurosurg. 94*, S82-3.

[51] Cohen DH, Feiner H, Jensson O, Frangione B. (1983) Amyloid fibril in hereditary cerebral hemorrhage with amyloidosis (HCHWA) is related to the gastroentero-pancreatic neuroendocrine protein, gamma trace. *J Exp Med. 158*, 623-8.

[52] Needham M, Mastaglia FL, Garlepp MJ. (2007) Genetics of inclusion-body myositis. *Muscle Nerve. 35*, 549-61.

[53] Vattemi G, King Engel W, McFerrin J, Askanas V. (2003) Cystatin C colocalizes with amyloid-beta and coimmunoprecipitates with amyloid-beta precursor protein in sporadic inclusion-body myositis muscles. *J Neurochem. 85*, 1539-46.

[54] Askanas V, Engel WK. (2002) Inclusion-body myositis and myopathies: different etiologies, possibly similar pathogenic mechanisms. *Curr Opin Neurol. 15*, 525-31.

[55] Zurdel J, Finckh U, Menzer G, Nitsch RM, Richard G. (2002) CST3 genotype associated with oxudative age related macular degeneration. *Br J Ophthalmol. 86*, 214-9.

[56] Morris B, Imrie F, Armbrecht AM, Dhillon B. (2007) Age-related macular degeneration and recent developments: new hope for old eyes? *Postgrad Med J. 83*, 301-7.

[57] Aronica E, van Vliet EA, Hendriksen E, Troost D, Lopes da Silva FH, Gorter JA. (2001) Cystatin C, a cysteine protease inhibitor, is persistently up-regulated in neurons and glia in a rat model for mesial temporal lobe epilepsy. *Eur J Neurosci. 14*, 1485-91.

[58] Lukasiuk K, Pirttila TJ, Pitkanen A. (2002) Upregulation of cystatin C expression in the rat hippocampus during epileptogenesis in the amygdala stimulation model of temporal lobe epilepsy. *Epilepsia. 43*, 137-45.

[59] Heidtmann HH, Salge U, Havemann K, Kirschke H, Wiederanders B. (1993) Secretion of a latent, acid activatable cathepsin L precursor by human non-small cell lung cancer cell lines. *Oncol Res. 5*, 441-51.

[60] Heidtmann HH, Salge U, Abrahamson M, Bencina M, Kastelic L, Kopitar-Jerala N, Turk V, Lah TT. (1997) Cathepsin B and cysteine proteinase inhibitors in human lung cancer cell lines. *Clin Exp Metastasis. 15*, 368-81.

[61] Keppler D, Abrahamson M, Sordat B. (1994) Secretion of cathepsin B and tumour invasion. *Biochem Soc Trans. 22*, 43-9.

[62] Lignelid H, Collins VP, Jacobsson B. (1997) Cystatin C and transthyretin expression in normal and neoplastic tissues of the human brain and pituitary. *Acta Neuropathol (Berl). 93*, 494-500.

[63] Kolkhorst V, Sturzebecher J, Wiederanders B. (1998) Inhibition of tumour cell invasion by protease inhibitors: correlation with the protease profile. *J Cancer Res Clin Oncol. 124*, 598-606.

[64] Coulibaly S, Schwihla H, Abrahamson M, Albini A, Cerni C, Clark JL, Ng KM, Katunuma N, Schlappack O, Glossl J, Mach L. (1999) Modulation of invasive properties of murine squamous carcinoma cells by heterologous expression of cathepsin B and cystatin C. *Int J Cancer. 83*, 526-31.

[65] Colella R, Goodwyn E, Gopal P. (2002) Increased cell density decreases cysteine proteinase inhibitor activity and increases invasive ability of two prostate tumor cell lines. *Cancer Lett. 185*, 163-72.

[66] Grubb A, Abrahamson M, Olafsson I, Trojnar J, Kasprzykowska R, Kasprzykowski F, Grzonka Z. (1990) Synthesis of cysteine proteinase inhibitors structurally based on the proteinase interacting N-terminal region of human cystatin C. *Biol Chem Hoppe Seyler. 371*, 137-44.

[67] Johansson L, Grubb A, Abrahamson M, Kasprzykowski F, Kasprzykowska R, Grzonka Z, Lerner UH. (2000) A peptidyl derivative structurally based on the inhibitory center of cystatin C inhibits bone resorption in vitro. *Bone. 26*, 451-9.

[68] Gandy S, DeMattos RB, Lemere CA, Heppner FL, Leverone J, Aguzzi A, Ershler WB, Dai J, Fraser P, Hyslop PS, Holtzman DM, Walker LC, Keller ET. (2004) Alzheimer A beta vaccination of rhesus monkeys (Macaca mulatta). *Alzheimer Dis Assoc Disord. 18*, 44-6.

[69] Grabowskal U, Chambers TJ, Shiroo M. (2005) Recent developments in cathepsin K inhibitor design. *Curr Opin Drug Discov Devel. 8*, 619-30.

[70] Gillmore JD, Hawkins PN. (2006) Drug Insight: emerging therapies for amyloidosis. *Nat Clin Pract Nephrol. 2*, 263-70.

[71] Hamaguchi T, Ono K, Yamada M. (2006) Anti-amyloidogenic therapies: strategies for prevention and treatment of Alzheimer's disease. *Cell Mol Life Sci. 63*, 1538-52.

[72] Keppler D. (2006) Towards novel anti-cancer strategies based on cystatin function. *Cancer Lett. 235*, 159-76.

[73] Kopitar-Jerala N. (2006) The role of cystatins in cells of the immune system. *FEBS Lett. 580*, 6295-301.

[74] Lin JC, Liu HL. (2006) Protein conformational diseases: from mechanisms to drug designs. *Curr Drug Discov Technol. 3*, 145-53.

[75] Link JO, Zipfe S. (2006) Advances in cathepsin S inhibitor design. *Curr Opin Drug Discov Devel. 9*, 471-82.

CONCLUSION AND PERSPECTIVES

The cystatin superfamily of proteinase inhibitors are broadly expressed and present in a wide range of species, organisms and tissues, from mammals to viruses. They are involved in a broad range of physio- and pathological processes, including age-related macular degeneration, amyloidosis, bone resorption, cancers, epilepsy, host pathogen reactions, inflammation and immune responses, stem cell growth and fate specification.

The sequences and three-dimensional structure of proteinase inhibitors of the cystatin superfamily have been determined. The proteinase inhibitors of the cystatin superfamily are pleiotropic molecules, with different domains of the molecules involved in different functions.

The proteinase inhibitors of the cystatin superfamily are primarily inhibitors of proteases of the papain superfamily or family C1 of cysteine proteases. The catalytic domain of proteinase inhibitors of the cystatin superfamily for proteases of the papain superfamily has been determined. Three highly conserved motifs are involved in the cysteine proteinase inhibitory (CPI) activities of cystatins on proteases of the papain superfamily.

Proteinase inhibitors of the cystatin superfamily, like chicken cystatin (cC) and human cystatin C (hCC), also inhibit of proteases of the legumain family or family C13 of cysteine proteases. Family C13 is a family of proteases unrelated to family C1 of cysteine proteases. A distinct domain of the molecules is involved in the CPI activity of cystatins on proteases of the legumain family.

Other domains of the molecules have been identified and shown to be independent of their proteinase inhibitory domain, like their activities on host pathogen reaction, immunomodulatory activities, and activities on adult neurogenesis and neural stem cells.

This raises the question of the mechanisms underlying the activities of CPIs of the cystatin superfamily that are unrelated to their proteinase inhibitory activities. Receptor signaling pathways have been proposed as mediating these activities, but this remains to be determined.

The CPIs of cystatin superfamily are amyloidogenic proteins. Amyloidogenic proteins are proteins forming amyloid fibrils. The causative factors and mechanisms underlying the formation of amyloid fibrils remain mostly unknown. The mutated variant of cystatin C L68Q is responsible for the hereditary cystatin C amyloid angiopathy. Recombinant proteins, of the mutated variant of cystatin C L68Q, form dimmers under predenaturing conditions.

Three-dimensional domain swapping has been proposed as a mechanism underlying the dimerization of the mutant variant of cystatin C L68Q. Hence, mutations of amyloidogenic proteins may lead to the formation of amyloid fibrils and three-dimensional domain swapping is a mechanism by which amyloid and related fibrils are formed. The mechanisms underlying the role and contribution of CPIs of the cystatin superfamily, and particularly cystatin C, in amyloidosis remain to be fully determined.

In all, CPIs of cystatin superfamily are involved in a broad range of physio- and pathological processes, and multiple domains and mechanisms are underlying their activities. But their roles, contributions and mechanisms in the physio- and pathological processes remain to be fully determined.

The understanding of the contribution of the various domains and mechanisms underlying the activities of CPIs of the cystatin superfamily will not only significantly improve our understanding of the functioning of organisms, from mammals to viruses, but also will have a tremendous potential for therapy.

Synthetic inhibitors of CPIs of the cystatin superfamily have shown their efficacy and potential for the treatment of cancers and other diseases. Future studies will aim at unraveling the mechanisms underlying the activities of CPIs of the cystatin superfamily and designing molecules against specific domains of the CPIs to target their specific activities and functions. Such approach will lead to new perspectives for the development of drugs and therapies.

APPENDIX

Table 1. The 20 natural amino acids

Name	Chemical formula	3 letter abbr.	1 letter abbr.	MW
Glycine	C2H5NO2	Gly	G	75.07
L-Alanine	C3H7NO2	Ala	A	89.1
L-Valine*	C5H11NO2	Val	V	117.15
L-Leucine*	C6H13NO2	Leu	L	131.18
L-Isoleucine*	C6H13NO2	Ile	I	131.18
L-Serine	C3H7NO3	Ser	S	105.09
L-Cysteine	C3H7NO2S	Cys	C	121.16
L-Threonine*	C4H9NO3	Thr	T	119.12
L-Methionine*	C5H11NO2S	Met	M	149.21
L-Aspartic acid	C4H7NO4	Asp	D	133.10
L-Asparagine	C4H8N2O3	Asn	N	132.118
L-Glutamic acid	C5H9NO4	Glu	E	147.13
L-Glutamine	C5H10N2O3	Gln	Q	146.15
L-Lysine*	C6H14N2O2	Lys	K	146.19
L-Arginine	C6H14N4O2	Arg	R	174.2
L-Histidine	C6H9N3O2	His	H	155.16
L-Tyrosine	C9H11NO3	Tyr	Y	181.19
L-Phenylalanine*	C9H11NO2	Phe	F	165.19
L-Tryptophane*	C11H12N2O2	Trp	W	204.23
L-Proline	C5H9NO2	Pro	P	115.13

There are 20 natural amino acids used in protein synthesis during translation. Among them, there are amino acids called essential that can not be synthesized from other molecules by a given organism. In human, there are 8 essential amino acids (*).Only the L-stereoisomers are found in mammalian proteins (glycine is a symmetrical structure). Abbr., abbreviation; MW, molecular weight (Da).

Table 2. Enzyme classification

EC number	Enzyme groups
EC 1	Oxidoreductases
EC 2	Transferases
EC 3	Hydrolases
EC 4	Lyases
EC 5	Isomerases
EC 6	Ligases

The Enzyme Commission number (EC number) is a classification scheme for enzymes. It is based on a four numerical code, separated by periods. It specifies progressively the enzyme-catalyzed reactions for each enzyme. The first number defines the group to which the enzyme belongs to. There are 6 main enzyme groups, EC 1 to 6. A similar classification scheme, referred as TC number, was designed for transport proteins and ion channels.

INDEX

A

B

C

D

E

F

G

H

I

K

L

M

N

O

P

R

S

T

U

V

X

B