Protein Science (1997), 6:5–12. Cambridge University Press. Printed in the USA.

REVIEW

Friends and relations of the cystatin superfamily—new members and their evolution

W.M. BROWN AND K.M. DZIEGIELEWSKA

Department of Anatomy and Physiology, University of Tasmania, GPO Box 252-24, Hobart, Tasmania, Australia 7001

(RECEIVED August 28, 1996; ACCEPTED September 24, 1996)

Abstract

The cystatin "superfamily" encompasses proteins that contain multiple cystatin-like sequences. Some of the members are active cysteine protease inhibitors, while others have lost or perhaps never aquired this inhibitory activity. In recent years, several new members of the superfamily have been characterized, including proteins from insects and plants. Based on partial amino acid homology, new members, such as the invariant chain (Ii), and the transforming growth factor-β receptor type II (TGF-β receptor II) may, in fact, represent members of an emerging family within the superfamily that may have used some common building blocks to form functionally diverse proteins. Cystatin superfamily members have been found throughout evolution and members of each family of the superfamily are present in mammals today. In this review, the new and older, established members of the family are arranged into a possible evolutionary order, based on sequence homology and functional similarities.

Keywords: cystatins; cysteine protease inhibitors; evolution; fetuins; kininogens; TGF-β receptor II

The concept of a cystatin "superfamily" emerged in the early 1980s, precipitated by an observation that multiple cystatin-like sequences were present in the kininogens and that the stefins were related to both the cystatins and the repeats in the kininogens (Ohkubo et al., 1984). The cystatins are a family of cysteine protease inhibitors with homology to chicken cystatin (Barrett, 1981). Cystatins typically comprise ≈115 amino acids, are largely acidic, contain four conserved cysteine residues known to form two disulfide bonds (Grubb et al., 1984), may be glycosylated and/or phosphorylated, and contain a series of conserved residues, most notably the "QVVAG" sequence (reviewed in Rawlings & Barrett, 1990; Turk & Bode, 1991).

Protein, cDNA, and genomic sequences of several cystatins have since been reported (e.g., Tsai et al., 1996; reviewed in Rawlings & Barrett, 1990; Turk & Bode, 1991). Their structural genes comprise three exons of characteristic size (e.g., Huh et al., 1995) (Fig. 1), and a cluster of these genes is located on human chromosome 20 (e.g., Thiesse et al., 1994) although cystatin genes are also found on other chromosomes (e.g., Pennachio et al., 1996).

Cystatin-like sequences identified in the kininogens (Ohkubo et al., 1984) were found to be present as sequence duplications. Subsequently, it was proposed that there were three cystatin-like repeats in the kininogen sequence, the amino terminal repeat being less conserved (Salvesen et al., 1986). This threefold repeat is reflected in the triplicated gene structure (Kitamura et al., 1985), which has been shown to be common to the kininogens (e.g., Cole & Schreiber, 1992). Each of the three cystatin repeats is encoded by three exons of characteristic size (e.g., Cole & Schreiber, 1992).

Thus, by the late 1980s the superfamily comprised cystatins, stefins, and kininogens. More recent additions to the superfamily will be discussed in this review, followed by our thoughts on how the evolution of this greatly enlarged superfamily may have occurred. The story is far from complete and contains several gaps and weaknesses that we would like to highlight in the hope that we will encourage other researchers to help us in filling them.

New members of the cystatin superfamily

The fetuins

Bovine fetuin was first characterized by Pedersen in 1944 (Pedersen, 1944), but it was not until 1988 that Elzanowski et al. (1988) first made a real connection between the fetuins and the cystatin superfamily. A link between the kininogens and human fetuin had been tentatively suggested by Hamberg et al. in 1975. In 1987, Dziegielewska and colleagues first discussed the connection between bovine fetuin and a human plasma protein, α_2-HS glycoprotein (Dziegielewska et al., 1987). Further work confirmed the initial observations and it is now clear that α_2-HS glycoprotein is the human fetuin (Christie et al., 1987; Dziegielewska et al., 1990; Brown et al., 1992a,b; Dziegielewska & Brown, 1995).

Reprint requests to: K.M Dziegielewska, Department of Anatomy and Physiology, University of Tasmania, GPO Box 252-24, Hobart, Tasmania 7001, Australia.

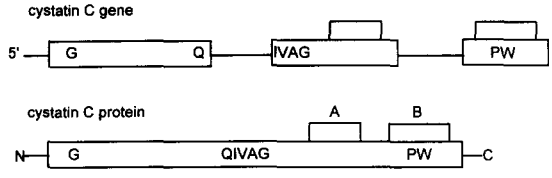

Fig. 1. Gene structure of the cystatins. This figure illustrates the three exon structure of the cystatin C gene and the encoded protein. Relative sizes of the introns and exons are not drawn to scale. Open boxes denote coding regions, solid lines denote introns, and key amino acids are indicated in single letter code. The conserved *gly* (G), the QIVAG sequence, and the conserved *pro-trp* (PW) sequences are indicated, as are the A- and B-type disulfide loops. As can be seen, the A- and B-type loops are each encoded in separate exons.

In recent years, protein and/or cDNA sequences have been reported for human (Gejyo et al., 1983; Yoshioka et al., 1986; Lee et al., 1987a), cow (Dziegielewska et al., 1990), sheep (Brown et al., 1992a), pig (Brown et al., 1992a), rat (Mizuno et al., 1991; Ohnishi et al., 1991; Rauth et al., 1992), mouse (Yang et al., 1992; Yamamoto & Sinohara, 1993), and Habu snake (Yamakawa & Omori-Satoh, 1992) fetuins. The disulfide loop structure predicted by Elzanowski and colleagues for human and bovine fetuins (Elzanowski et al., 1988), on the basis of their homology to the cystatins and to the cystatin domains in the kininogens has since been confirmed experimentally (Araki et al., 1989; Kellermann et al., 1989; Chin & Wold, 1993) (Fig. 2). All of the fetuin sequences contain 12 cysteine residues in positions identical to those in human and bovine fetuin, and it seems likely that this disulfide structure is common to all fetuins (see Dziegielewska & Brown, 1995, for extensive review). The one exception is the Habu snake fetuin, which contains a 13th cysteine (Yamakawa & Omori-Satoh, 1992).

Histidine-rich glycoprotein (HRG)

HRG has been characterized in the plasma of man, mouse, rabbit, cow and pig (see Leung, 1993, for references). The cDNA sequence of human HRG was reported by Koide and colleagues (Koide et al., 1986). From a sequence comparison, the authors

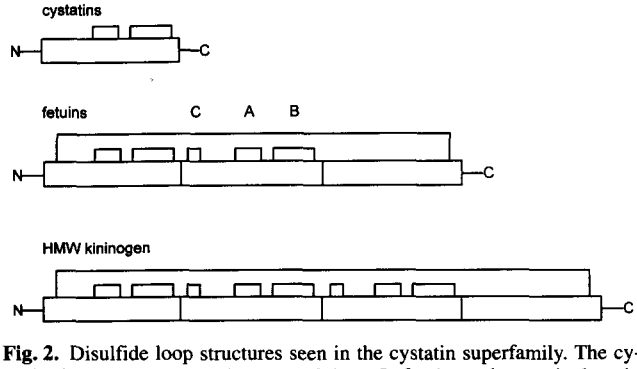

Fig. 2. Disulfide loop structures seen in the cystatin superfamily. The cystatins have one type A and one type B loop. In fetuin, each cystatin domain has an A- and a B-type loop; the second domain also has the very narrow C-type loop where two amino acids separate the cysteine residues. In HMW kininogen, both the second and third cystatin domains have C-type loops. In the fetuins and kininogens, the N- and C-terminal regions of the proteins are connected by a further disulfide bond.

suggested that the histidine-rich region of the protein was related to human and bovine high molecular weight (HMW) kininogen (Koide et al., 1986), in addition to homology at the amino-termini of the proteins. The structure of the human HRG gene was subsequently determined (Koide, 1988), and it was shown that the cystatin domains were each encoded by three exons of characteristic size, as they are in the cystatins (e.g., Colella & Bird, 1993), fetuin (Falquerho et al., 1991), and the kininogens (e.g., Cole & Schreiber, 1992). A purification method (Muldbjerg et al., 1992), a partial sequence, and the disulfide loop structure of the bovine protein (Sorensen et al., 1993) have also been reported. The human HRG gene has been localized to 3q28–q29 (Hennis et al., 1994), very close to the fetuin (3q27–q29, Magnuson et al., 1988), kininogen (3q26–qter, Fong et al., 1991b), and stefin A genes (STF1, 3cen–q21, Hsieh et al., 1991). Rizzu and Baldini (1995) reported that the human fetuin, HRG, and kininogen genes are located within 1 megabase of DNA at 3q27 (see also, James et al., 1996). The role of HRG remains unclear (see Leung, 1993). In a further likeness to human fetuin and the kininogens, a protease-sensitive site was found in bovine HRG [arg^{297}-↓-ala^{298}], although apparently not in a position homologous to any of the sites known in the fetuins or kininogens (Sorensen et al., 1993).

The cystatin-related proteins (CRPs)

Parker et al. (1978) reported the purification of a 20–22 kDa glycoprotein from the rat ventral prostate. Subsequently, a cDNA encoding the protein was cloned and sequenced (Ho et al., 1989; Winderickx et al., 1990), and more recently, the genomic structure of two rat CRPs has been determined (Devos et al., 1993, 1995). The genomic organization revealed the basis for the homology to the cystatins: the two genes comprised four exons, three of which were directly homologous to the three seen in typical cystatins (e.g., Colella & Bird, 1993); the fourth was homologous to the normal cystatin exon two. In the cystatins, exon two encodes the A-type disulfide loop; thus, CRPs are predicted to have two A-type loops in their amino terminal region before a normal type B loop, encoded by exon four in the CRPs (exon three in the cystatins).

Variant cystatins

A divergent cystatin in the venom of the African puff adder (Bitis arietans)

Evans and Barrett described the purification of a cystatin-like protein from the venom of the puff adder (Evans & Barrett, 1987). The protein sequence was also reported (Ritonja et al., 1987). The protein was most closely related to the cystatins, although a series of differences was also apparent. The sequence contained three cysteine residues, two of which were presumed to form the A-type disulfide loop and the third corresponded to the amino-terminal end of the B-type loop. The authors noted the possibility that the protein isolated had been truncated in the purification process. The spacing between the cysteines believed to form the A-type loop was different to that seen elsewhere in the superfamily, there being a six amino acid insertion so that the cysteine residues were separated by 15 amino acids. It is worth noting here that the protease inhibitor purified from another snake, the Japanese Habu, turned out to be fetuin (Yamakawa & Omori-Satoh, 1992).

A divergent cystatin from the flesh fly (Sarcophaga peregrina)

In 1985, Suzuki and Natori reported the purification of sarcocystatin A from the perilymph of flesh fly larvae (Suzuki & Natori,

1985). Two proteins, sarcocystatin A_α and A_β, differing slightly in molecular weight, were characterized. Further work by the same group led to the cloning and sequencing of a sarcocystatin A cDNA (Saito et al., 1989). The sequence revealed significant homology to the cystatins, stefins, and kininogen cystatin domains (Saito et al., 1989). However, a series of differences was also readily apparent, most notably the fact that sarcocystatin A has two cysteine residues at the amino terminus separated by only two amino acids: *cys-val-gly-cys*. Such a sequence is found at the amino terminal end of kininogen cystatin domains 2 and 3 and in fetuin cystatin domain 2 (see Dziegielewska & Brown, 1995). In both the fetuins and kininogens, a C-type disulfide loop is known to be formed by this sequence (Araki et al., 1989; Kellermann et al., 1989; Chin & Wold, 1993). It is thus possible that sarcocystatin also has a C-type disulfide loop. The sarcocystatin A sequence also contains two cysteine residues encompassing the conserved *pro-trp* sequence, which would be predicted to form the B-type disulfide loop. However, while it does contain a QVVSG sequence, there are no cysteine residues that could form the A-type disulfide loop.

A divergent cystatin from the fruit fly (Drosophila melanogaster)

Delbridge and Kelly reported the cloning of a cystatin-like sequence from a *Drosophila* head cDNA library (Delbridge & Kelly, 1990). A genomic clone was also isolated (Delbridge & Kelly, 1990). As has been observed elsewhere in the superfamily (e.g., Lee et al., 1987), the gene appeared to be polymorphic (Delbridge & Kelly, 1990). The authors further noted that the cDNA sequence was most closely related to that of oryzacystatin I (Kondo et al., 1989), demonstrating approximately 22% sequence identity (Delbridge & Kelly, 1990). The sequence revealed two possible in-frame ATG codons (Delbridge & Kelly, 1990). Interestingly, the genomic sequence revealed the position of the intron, in a different location to that in either the oryzacystatins (Kondo et al., 1989, 1991), the animal cystatins (e.g., Colella & Bird, 1993), or the cystatin domains in larger proteins (e.g., Falquerho et al., 1991). The QVVAG sequence was conserved and in the amino terminal region there was a further related sequence QVVGG. If translation is assumed to start from the first methionine residue, then the intron neatly divides the protein almost exactly in half, 62 amino acids ahead of it, 64 downstream of it, raising the possibility that a duplication was involved in the formation of the current gene. The protein contains two cysteine residues towards the carboxyl-terminus, encompassing the *pro-trp* sequence. By comparison with other superfamily sequences, these would be expected to form the B-type disulfide loop. However, the sequence contains no cysteine residues that could form the A-type loop.

The invariant chain (Ii-chain)

The invariant chain is a non-polymorphic protein intimately involved in the assembly of class II MHC molecules (reviewed in Sant & Miller, 1994). The Ii-chain is believed to act as a chaperonin to MHC class II $\alpha\beta$ heterodimers as they pass through the endoplasmic reticulum (ER) and Golgi body (Anderson & Miller, 1992). In humans, there are four forms of the Ii-chain as a result of alternative splicing and alternative initiation of translation (Strubin et al., 1986; O'Sullivan et al., 1987). Recently, Katunuma et al. demonstrated that the amino terminal region of the p31 form of the Ii-chain shows significant homology to various cystatins and stefins (Katunuma et al., 1994). The human Ii-chain contains two cysteine residues, but not in positions corresponding to those in typical cystatins (Katunuma et al., 1994). The sequence alignment reported by Katunuma et al. (1994) was constructed using the stefin B crystal structure (Stubbs et al., 1990) as a model, in a similar way to that reported by Stubbs et al. (1990). Both alignments require insertions in the middle of the sequence, as there is a degree of homology at both ends of the sequences, again suggesting that the cystatins arose from the stefins by the insertion of an exon that encoded the A-type disulfide loop (see Stubbs et al., 1990; Katunuma et al., 1994).

The plant "cystatins"

A series of cysteine protease inhibitors has now been characterized and studied from plant sources (e.g., Abe et al., 1987, 1992; Kondo et al., 1990; Rowan et al., 1990; Waldron et al., 1993; Murzin, 1993; Irie et al., 1996). The most studied are the oryzacystatins of rice (*Oryza sativa*) (Abe et al., 1987; Kondo et al., 1989, 1990, 1991). While in some respects these proteins resemble the cystatins, they are stefin-like in having no disulfide bonds or cysteine residues. Furthermore, despite the high protein sequence homology, the genomic organization of the oryzacystatins has been reported to be markedly different to that seen in the animal cystatins (Kondo et al., 1989, 1991). Because of this it has been suggested that the oryzacystatins belong to a new *phytocystatin* family within the cystatin superfamily (Kondo et al., 1991). More recently, a further plant cystatin has been characterized from corn, *Zea mays* (Abe et al., 1992), adding weight to that argument. Again, the sequence revealed homology to the cystatins including conservation of the sequence *phe-ala-val-asn-glu-his-asn*, but no disulfide bonds, as seen in the stefins. A further cysteine protease inhibitor of pI 8.3, PCPI 8.3, was characterized biochemically from the potato (Rowan et al., 1990), although no sequence has yet been reported. More recently, a much larger protein from potato tubers has been examined. Potato multicystatin, PMC, was shown to be a protein of ≈87 kDa and to comprise eight repeated cystatin domains (Waldron et al., 1993). Unlike the fetuins, HRG, or the kininogens, no non-cystatin sequence was found at the carboxyl-terminus. The sequence identity between the eight domains varied from 53 to 89% (Waldron et al., 1993), and the domains were readily separable by proteolysis (Walsh & Strickland, 1993). In common with the oryzacystatins, the domains were stefin-like in having no cysteine residues. By Southern blot analysis, related genes were shown to be present in the pea and maize genomes (Waldron et al., 1993). It was further noted that the genomic organization was different to that of animal cystatin domains and to the oryzacystatins, the only other plant cystatins for which genomic sequences have been reported. Specifically, the oryzacystatin genes contain an intron in the 3′-UTR (Kondo et al., 1989, 1990, 1991), whereas PMC does not (Waldron et al., 1993). Recently, a further cystatin-related sequence has been characterized in plants; Murzin (Murzin, 1993) noticed that the monellin crystal structure (Ogata et al., 1987) bore considerable resemblance to those of human stefin (Stubbs et al., 1990) and chicken cystatin (Bode et al., 1988). Monellin is a very sweet protein isolated from the berries of the African plant, *Dioscoreophyllum cumminsii*.

The QVVAG sequence

Almost as soon as sequences became available, it became obvious that there were several strongly conserved regions in the cystatin,

stefins, and the kininogen domains 2 and 3 (see Ohkubo et al., 1984; Elzanowski et al., 1988). The most notable were a conserved glycine, the proline-tryptophan sequence, PW, and what was assumed to be the active site sequence glutamine-valine-valine-alanine-glycine, QVVAG (see Fig. 1). Several reports of mutagenesis experiments, aimed at examining how much variation is tolerated in these sequences, have appeared in the literature (Nikawa et al., 1989; Jerala et al., 1990; Auerswald et al., 1992). The consensus currently is that this sequence should probably be referred to as QxVxG, where x can be one of several amino acids. More recently, the crystal (Bode et al., 1988; Stubbs et al., 1990) and NMR structures (Martin et al., 1994, 1995; Tate et al., 1995), and the resulting docking model (Stubbs et al., 1990), have gone some way to explain the role of these sequences. The QxVxG sequence was found on the first β-hairpin loop and the PW sequence is found on the second loop. These two hairpin loops and the largely unstructured amino terminus of the protein form a wedge that fits into the protease (Bode et al., 1988; Stubbs et al., 1990; Machleidt et al., 1995; Auerswald et al., 1996). In the fetuins and HRGs, these sequences have been extensively altered, making it unlikely that they are viable cysteine protease inhibitors.

Conserved sequences around the cysteine at the carboxyl-terminus of the fetuins, HMW kininogen, and HRG

As Elzanowski et al. (Elzanowski et al., 1988) noted, the sequence immediately surrounding the last cysteine residue, *pro-pro-cys-pro-gly-arg* in human fetuin (Gejyo et al., 1983; Lee et al., 1987) shares significant homology with the same region of human *pro-lys-cys-pro-gly-arg* (Ohkubo et al., 1984; Kitamura et al., 1985; Salvesen et al., 1986; Kellermann et al., 1986, 1987), and bovine *pro-lys-cys-pro-ser-arg* (Nawa et al., 1983; Kitamura et al., 1983; Sueyoshi et al., 1984, 1985) high-molecular weight kininogen, suggesting a possible common origin of the carboxyl-terminal regions of the two proteins. Furthermore, the sequence around the last cysteine in bovine *glu-ser-cys-pro-gly-thr* (Sorensen et al., 1993) and human *glu-ser-cys-pro-gly-lys* (Koide et al., 1986) histidine-rich glycoprotein is also similar.

Fetuin domain D3 has no homology to any sequence currently in the sequence databases, beyond that around the carboxyl-terminal cysteine residue discussed above. In particular, the domain shares no homology with the carboxyl-terminal region of HRG, which might appear to be closely related. It was originally reported that this region in HRG was encoded by two exons (Koide, 1988). Hennis et al. (1994) have since challenged this; they found no evidence of intron H (Koide, 1988) in a genomic clone that they isolated. Thus, the entire carboxyl-terminal region of the protein may be encoded by a single terminal exon as it is in fetuin (Falquerho et al., 1991) and the kininogens (e.g., Kitamura et al., 1985). It has also been noted that the histidine-rich region of HRGs carboxyl-terminus, encoded by exon VIII (Koide, 1988) has some homology to the carboxyl-terminal region of human and bovine HMW kininogen (Koide et al., 1986).

Conclusions

The superfamily has thus been dramatically enlarged in recent years. In addition to the original cystatin, stefin, and kininogens families, it is clear that the fetuins, the HRGs, and the CRPs constitute entire new families. The number of members in the cystatin and stefin families has also been greatly enlarged and both families now contains sequences from a biologically more diverse background (Fig. 3).

Evolution

Over the last decade several papers have discussed the evolution of the cystatin superfamily (e.g., Rawlings & Barrett, 1990; Müller-Esterl et al., 1985). One proposed scheme (Müller-Esterl et al., 1985) involved the evolution of the various members of the cystatin superfamily, in which the cystatins, and from them the kininogens, emerged from a stefin-like precursor protein by a fusion of separate exons. Of theses exons only the exon encoding the N-terminal sequence was related to the stefins. The proposed evolutionary pathway also contained a "missing link," a two cystatin-domain protein that evolved from the cystatins, by duplication. In this new cystatin superfamily there were two candidates for such a protein: fetuin and HRG. Several aspects of the proposed scheme outlined above now seem unlikely for the following three reasons:

1. Alignments of stefin and cystatin sequences reveal that specific amino acid residues are common throughout the length of the molecules, not just at the amino terminus, as such a scheme would predict. For example, a strongly conserved tyrosine (residue 112), and the PW (*pro-trp*) sequence (residues 115–116), found toward the carboxyl-terminus of the cystatin, are common to oryzacystatin I (a plant stefin), many cystatins and kininogen domain D3 (see Rawlings & Barrett, 1990).

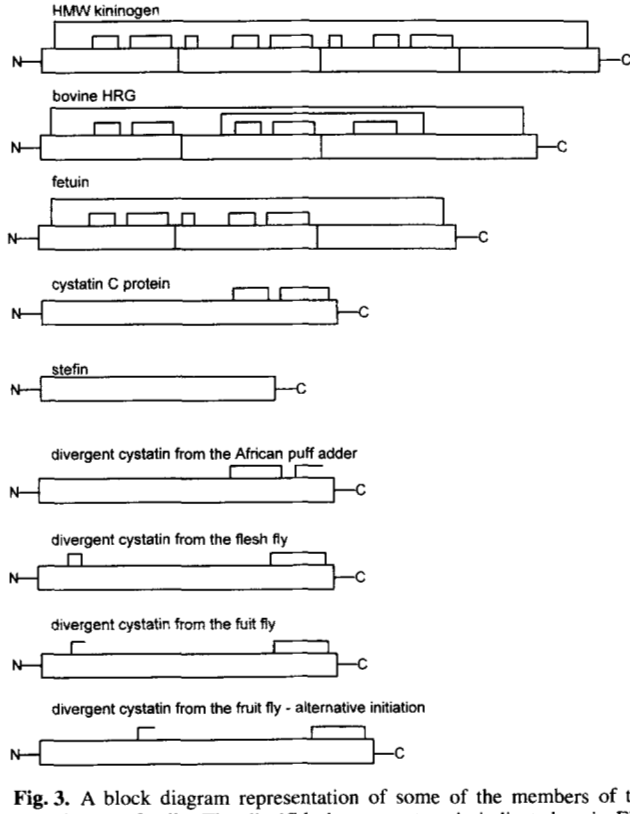

Fig. 3. A block diagram representation of some of the members of the cystatin superfamily. The disulfide loop structure is indicated as in Figures 1 and 2. The "incomplete" disulfide loops in fruit fly and puff adder cystatins indicate unpaired cysteine residues in the sequences that are in positions corresponding to disulfide loops in the archetypal cystatin protein.

2. On the basis of the crystal structure of human stefin B (Stubbs et al., 1990) and comparison with the known crystal structure of chicken cystatin (Bode et al., 1988, 1990), Stubbs et al. (1990) presented a sequence alignment based on equivalent positions in the two crystal structures. The alignment required the insertion of a 23 amino acid gap, including the A-type disulfide loop. As is it now known that this loop and flanking sequence are encoded by a single exon in the animal cystatins (e.g., Cox & Shaw, 1992), HRG (Koide, 1988), fetuin (Falquerho et al., 1991), the CRPs (Devos et al., 1993), and the kininogens (e.g., Kitamura et al., 1985), it seems more likely that the evolution from the stefins occurred by the *insertion* of such an exon rather than its *addition* to the 3′-end of the coding sequence.

3. The "missing link," a possible two cystatin-domain protein that could have evolved from the cystatins, does not seem to be either HRG or fetuin. Cystatin domains two and three in the kininogens are functional cysteine protease inhibitors (Salvesen et al., 1986). Neither domain in fetuin appears to be a functional cysteine protease inhibitor (Brown et al., 1992a, 1992b). Analysis of the HRG sequence also suggests that neither domain is likely to be a functional inhibitor in this protein either, although no experimental data has yet been reported. If fetuin or HRG were the "missing link," then the kininogens, which were suggested to have evolved from the two-domain protein, would have to "re-evolve" their protease inhibitory activity and sequences. It seems rather more likely that there is still a missing link, either lost in evolution or, as yet, undiscovered (see below).

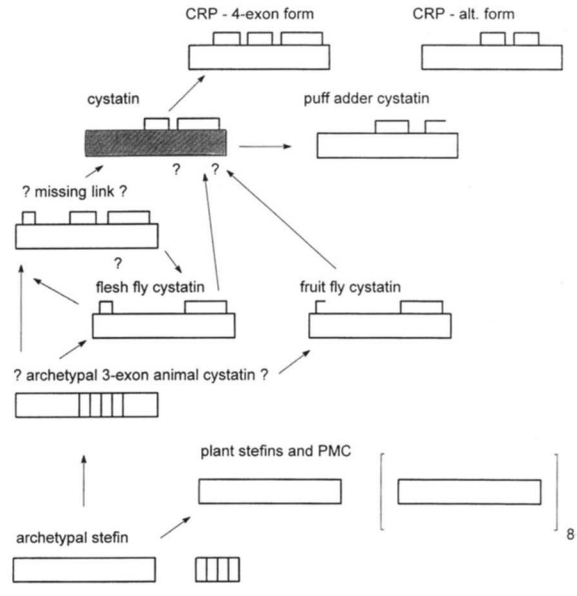

Fig. 4. Evolution of the cystatin superfamily. A possibly evolutionary scheme to explain the generation of the cystatins, potato multicystatin (PMC), which has eight cystatin repeats and no non-cystatin sequence at the carboxyl terminus, and the cystatin-related proteins (CRPs) from the archetypal stefin. We propose a missing link here; this protein would be a cystatin with types A, B, and C disulfide loops. From this one domain protein, the two cystatin domain missing link, and from it the fetuins, HRGs, and kininogens (see Fig. 5), could then evolve.

Intermediate forms?

Recently, two sequences of cystatin-related proteins from insects have been reported; a cystatin-related protein from *Drosophila* (Delbridge & Kelly, 1990), and sarcocystatin A from the flesh fly (Suzuki & Natori, 1985). It seems that the search for further insect protein, cDNA, and genomic sequences would be most valuable. The two examples to date appear to be intermediate between the animal stefins and plant inhibitors on the one hand and the cystatins and cystatin domains seen in the larger proteins on the other. Sarcocystatin A has a *cys-xaa-xaa-cys* sequence (Suzuki & Natori, 1985), which could form a C-type disulfide loop, in addition to the B-type loop that both it and the *Drosophila* protein are expected to have. Neither, however, apparently has an A-type loop, surrounding the *pro-trp* (PW) sequence (Suzuki & Natori, 1985; Delbridge & Kelly, 1990). The cystatins of higher animals do not have the C-type loop, suggesting possibly that the two-domain "missing link," discussed below, may have evolved by a duplication before the cystatins lost the C-type loop in higher animals (Fig. 4).

Summary

It is now clear that the cystatin superfamily comprises a series of families: the original cystatins, stefins, and kininogens and the more recently characterized fetuins, histidine-rich glycoproteins and cystatin-related proteins. Additionally, the rat has a further family of kininogens, the T-kininogens. The invariant chain, Ii, homologous to the cystatins, may also be a new member of an emerging family. Whether the plant stefins belong in a family of their own as was proposed by Kondo et al. (1991) depends on whether the genomic organization of the animal stefins, which has not yet been reported, is different.

Each of the above-mentioned proteins presumably derived from a single ancestral gene at some point in evolution. On the one hand, the disulfide loop structure has been mostly retained, on the other hand several of the families have lost the ability to inhibit cysteine proteases. The fetuins, HRGs, CRPs, and domain 1 of the kininogens have all lost some amino acids that are now known to be important in this regard.

Proteins containing repeated structures are, of course, not unique to the cystatin superfamily. Many other large proteins, and in particular, many other plasma proteins (e.g., a-fetoprotein, AFP, Eiferman et al., 1981; albumin, Brown, 1976a, 1976b; transferrin, Greene & Feeney, 1968; Bowman et al., 1988; and ceruloplasmin, Ortel et al., 1984) also contain repeated structures (see also, Doolittle, 1992). However, the single domain proteins from which these larger proteins were derived are no longer present in higher animals. For example, several attempts have been made to find the single domain transferrin precursor (see Martin et al., 1984) but as far back, evolutionarily, as the hagfish, this plasma protein has been shown to bind two ferric ions [Fe^{2+}] and to have a molecular weight of ≈75 kDa (Martin et al., 1984). Martin et al. (1984) suggested that the iron-binding protein isolated from the prochordate *Pyura stolonifera* might represent the ancestral gene. However, more recently it has been shown that the transferrin protein from the sphinx moth (*Manduca sexta*) has a duplicated structure and shares extensive sequence and structural homology with vertebrate transferrins (Bartfeld & Law, 1990).

The cystatin superfamily is possibly unique in that so many of its members have survived through evolution and are still currently represented; members of each family are present in mammals today. There still appears to be at least one missing link in addition to the one discussed above. The missing link, as we would now propose, would be a two cystatin domain protein in which both domains were functional cysteine protease inhibitors. From it, the

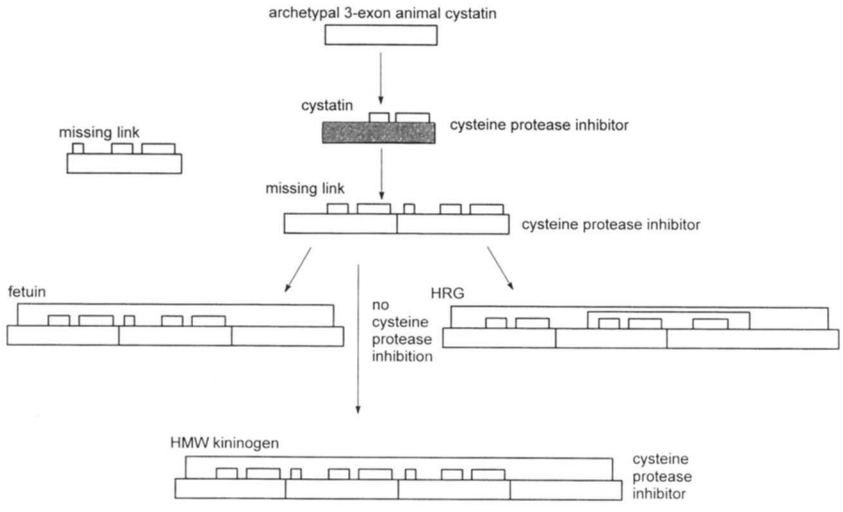

Fig. 5. An evolutionary missing link. We propose that there is still a second missing link in the evolutionary development of the superfamily. This protein would have two functional cystatin domains. From this protein, the fetuins and HRGs could have evolved and lost their cysteine protease inhibitory activity and the kininogens could have evolved by a further duplication of the second cystatin domain to give three functional cystatin domains.

kininogens, fetuin, and HRG could have evolved separately or perhaps in parallel, and retained or lost their protease-inhibitory activity and active site sequences. In support of this idea is the observation of conserved sequences immediately around the cysteine at the carboxyl-terminus of the fetuins, HMW kininogen, and HRG, again suggesting a common origin of all three proteins. Given that at least one reptile has a fetuin (Yamakawa & Omori-Satoh, 1992), it will be necessary to look a considerable way back in evolution for this missing link (Fig. 5).

Did the structural remodelling of the family lead to members that have lost their protease inhibitory activity and acquired new functions? One exciting new finding is that in the fetuins, a sequence of amino acids (residues 114–132 of the bovine sequence, Dziegielewska et al., 1990) that forms a disulphide loop in fetuin can inhibit binding of transforming growth factor-β, TGF-β, and bone morphogenetic proteins (BMPs) to their receptor, TGF-β receptor, type II (Demetriou et al., 1996). We have suggested before that fetuin's structure could be a possible "receptor" for an unknown cytokine or hormone (Dziegielewska & Brown, 1995), and the work of Demetriou et al. supports this. It is well known that in another family of protease inhibitors, the serpins (serine protease inhibitors), members of that family that have lost their protease inhibitory activity, have been remodelled to function as cortisol- and thyroxine-binding proteins (see Dziegielewska & Brown, 1995). Thus, in the cystatin superfamily, as in other known protein families, the common building blocks (cystatin domains) have been used to create functionally diverse proteins.

Acknowledgments

We should like to thank the Australian Research Council (ARC) and the National Health and Medical Research Council (NH&MRC) for providing the financial support for this work. Thanks are also due to Dr. Lisa Dixon for preparing the figures and Mrs. Tracey Walls for editing the manuscript.

References

Abe K, Emori Y, Kondo H, Suzuki K, Arai S. 1987. Molecular cloning of a cysteine proteinase inhibitor of rice (oryzacystatin). Homology with animal cystatins and transient expression in the ripening process of rice seeds. *J Biol Chem 262*:16793–16797.

Abe M, Abe K, Kuroda M, Arai S. 1992. Corn kernel cysteine proteinase inhibitor as a novel cystatin superfamily member of plant origin. Molecular cloning and expression studies. *Eur J Biochem 209*:933–937.

Anderson MS, Miller J. 1992. Invariant chain can function as a chaperone protein for class II major histocompatibility complex molecules. *Proc Natl Acad Sci USA 89*:2282–2286.

Araki T, Yoshioka Y, Schmid K. 1989. The position of the disulfide bonds in human plasma α_2HS-glycoprotein and the repeating double disulfide bonds in the domain structure. *Biochim Biophys Acta 994*:195–199.

Auerswald EA, Genenger G, Assfalg-Machleidt I, Machleidt W, Engh RA, Fritz H. 1992. Recombinant chicken egg white cystatin variants of the QLVSG region. *Eur J Biochem 209*:837–845.

Auerswald EA, Naegler DK, Gross S, Assfalg-Machleidt I, Stubbs MT, Eckerskorn C, Machleidt W, Fritz H. 1996. Hybrids of chicken cystatin with human kininogen domain 2 sequences exhibit novel inhibition of calpain, improved inhibition of actinidin and impaired inhibition of papain, cathepsin L and cathepsin B. *Eur J Biochem 235*:534–542.

Barrett AJ. 1981. Cystatin, the egg white inhibitor of cysteine proteinases. *Methods Enzymol 80*:771–778.

Bartfeld NS, Law JH. 1990. Isolation and molecular cloning of transferrin from the tobacco hornworm, *Manduca sexta*. Sequence similarity to the vertebrate transferrins. *J Biol Chem 265*:21684–21691.

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*:2593–2599.

Bode W, Engh R, Musil D, Laber B, Stubbs M, Huber R, Turk V. 1990. Mechanism of interaction of cysteine proteinases and their protein inhibitors as compared to the serine proteinase-inhibitor interaction. *Biol Chem Hoppe-Seyler 371(Suppl)*:111–118.

Bowman BH, Yang FM, Adrian GS. 1988. Transferrin: Evolution and genetic regulation of expression. *Adv Genet 25*:1–38.

Brown JR. 1976a. Structure of bovine serum albumin. *Fed Proc 34*:591.

Brown JR. 1976b. Structural origins of mammalian albumin. *Fed Proc 35*:2141–2144.

Brown WM, Christie DL, Dziegielewska KM, Nawratil P, Saunders NR, Müller-Esterl W. 1992a. The nucleotide and deduced amino acid structures of sheep and pig fetuin. Common structural features of the mammalian fetuin family. *Eur J Biochem 205*:321–331.

Brown WM, Dziegielewska KM, Møllgård K, Saunders NR. 1992b. Fetuin, an old friend revisited. *Bioessays 14*:749–755.

Chin CC, Wold F. 1993. The use of tributylphosphine and 4-(aminosulfonyl)-7-fluoro-2,1,3-benzoxadiazole in the study of protein sulfhydryls and disulfides. *Anal Biochem 214*:128–134.

Christie DL, Dziegielewska KM, Hill RM, Saunders NR. 1987. Fetuin: The bovine homologue of human αHS glycoprotein. *FEBS Lett 214*:45–49.

Cole TJ, Schreiber G. 1992. The structure and expression of the genes for T-kininogen in the rat. *Agents Actions Suppl 38*:292–299.

Colella R, Bird JW. 1993. Isolation and characterization of the chicken cystatin-encoding gene: Mapping transcription start point and polyadenylation sites. *Gene 130*:175–181.

Cox JL, Shaw PA. 1992. Structure, organization and regulation of a rat cysteine proteinase inhibitor-encoding gene. *Gene 110*:175–180.

Delbridge ML, Kelly LE. 1990. Sequence analysis, and chromosomal localization of a gene encoding a cystatin-like protein from *Drosophila melanogaster*. *FEBS Lett 274*:141–145.

Demetriou M, Binkert C, Sukhu B, Tenenbaum HC, Dennis JW. 1996. Fetuin/

alpha-2-HS glycoprotein is a transforming growth factor-beta type II receptor mimic and cytokine antagonist. *J Biol Chem 271*:12755–12761.

Devos A, de Clercq N, Vercaeren I, Heyns W, Rombauts W, Peeters B. 1993. Structure of rat genes encoding androgen-regulated cystatin-related proteins (CRPs): A new member of the cystatin superfamily. *Gene 125*:159–167.

Devos A, Zhang J, Riviere M, Vercaeren I, Heyns W, Cassiman JJ, Rombauts W, Marynen P, Szpirer J, Spirer C. 1995. The gene coding for rat cystatin-related prostate protein (Cstrp) map to chromosome 3q41. *Cytogenet Cell Genet 68*:239–242.

Doolittle RF. 1992. Reconstructing history with amino acid sequences. *Protein Sci 1*:191–200.

Dziegielewska KM, Brown WM. 1995. *Fetuin*. Molecular Biology Intelligence Unit, R.G. Landes Company, Texas, International. Austin: Springer-Verlag.

Dziegielewska KM, Møllgård K, Reynolds ML, Saunders NR. 1987. A fetuin-related glycoprotein (a_2HS) in human embryonic and fetal development. *Cell Tissue Res 248*:33–41.

Dziegielewska KM, Brown WM, Casey S-J, Christie DL, Foreman RC, Hill RM, Saunders NR. 1990. The complete amino acid and nucleotide sequence of bovine fetuin: Its sequence homology with α_2-HS glycoprotein and its relationship to other members of the cystatin superfamily. *J Biol Chem 265*:4354–4357.

Eiferman FA, Young PR, Scott RW, Tilghman SM. 1981. Intragenic amplification and divergence in the mouse a-fetoprotein gene. *Nature 294*:713–718.

Elzanowski A, Barker WC, Hunt LT, Seibel-Ross E. 1988. Cystatin domains in alpha-2-HS-glycoprotein and fetuin. *FEBS Lett 227*:167–170.

Evans HJ, Barrett AJ. 1987. A cystatin-like cysteine proteinase inhibitor from venom of the African puff adder (*Bitis arietans*). *Biochem J 246*:795–797.

Falquerho L, Patey G, Paquereau L, Rossi V, Lahuna O, Szpirer J, Szpirer C, Levan G, Le Cam A. 1991. Primary structure of the rat gene encoding an inhibitor of the insulin receptor tyrosine kinase. *Gene 98*:209–216.

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*:189–192.

Gejyo F, Chang J, Bürgi W, Schmid K, Offner GD, Troxler R, van Halbeek H, Dorland L, Gerwig GJ, Vliegenthart JFG. 1983. Characterisation of the B-chain of human plasma α_2HS-glycoprotein. The complete amino acid sequence and primary structure of its heteroglycan. *J Biol Chem 258*:4966–4971.

Greene FC, Feeney RE. 1968. Physical evidence for transferrins as single polypeptide chains. *Biochemistry 7*:1366–1371.

Grubb A, Löfberg H, Barrett AJ. 1984. The disulphide bridges of human cystatin C (g-trace) and chicken cystatin. *FEBS Lett 170*:370–374.

Hamberg U, Elg P, Nissinen E, Stelwagen P. 1975. Purification and heterogeneity of human kininogen. Use of DEAE-chromatography, molecular sieving and antibody specific immunosorbents. *Int J Peptide Protein Res 7*:261–280.

Hennis BC, Frants RR, Bakker E, Vossen RH, van der Poort EW, Blonden LA, Cox S, Khan PM, Spurr NK, Kluft C. 1994. Evidence for the absence of intron H of the histidine-rich glycoprotein (HRG) gene: Genetic mapping and in situ localization of HRG to chromosome 3q28-q29. *Genomics 19*:195–197.

Ho K, Snoek R, Quarmby V, Viskochil DH, Rennie PS, Wilson EM, French FS, Bruchovsky N. 1989. Primary structure and androgen regulation of a 20-kilodalton protein specific to rat ventral prostate. *Biochemistry 28*:6367–6373.

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*:207–209.

Huh C, Nagle JW, Kozak CA, Abrahamson M, Karlsson S. 1995. Structural organization, expression and chromosomal mapping of the mouse cystatin-C-encoding gene (Cst3). *Gene 152*:221–226.

Irie K, Hosoyama H, Takeuchi T, Iwabuchi K, Watanabe H, Abe M, Abe K, Arai S. 1996. Transgenic rice established to express corn cystatin exhibits strong inhibitory activity against insect gut proteinases. *Plant Mol Biol 30*:149–157.

James LA, Ogilvie DJ, Yamakawa Y, Stirling CJ, Anand R. 1996. Walking, cloning, and mapping with YACs in 3q27: Localization of five ESTs including three members of the cystatin gene family and identification of CpG islands. *Genomics 32*:425–430.

Jerala R, Trstenjak-Prebanda M, Kroon-Zitko L, Lenarcic B, Turk V. 1990. Mutations in the QVVAG region of the cysteine proteinase inhibitor stefin B. *Biol Chem Hoppe Seyler 371(Suppl)*:157–160.

Katunuma H, Kakegawa H, Matsunaga Y, Saibara T. 1994. Immunological significance of invariant chain from the aspect of its structural homology with the cystatin family. *FEBS Lett 349*:265–269.

Kellermann J, Lottspeich F, Henschen A, Müller-Esterl W. 1986. Completion of the primary structure of human high-molecular-mass kininogen. The amino acid sequence of the entire heavy chain and evidence for its evolution by gene triplication. *Eur J Biochem 154*:471–478.

Kellermann J, Thelen C, Lottspeich F, Henschen A, Vogel R, Müller-Esterl W. 1987. Arrangement of the disulphide bridges in human low-Mr kininogen. *Biochem J 247*:15–21.

Kellermann J, Haupt H, Auerswald E-A, Müller-Esterl W. 1989. The arrangement of disulfide loops in human α-HS glyco-protein. Similarity to the disulfide bridge structures of cystatins and kininogens. *J Biol Chem 264*:14121–14128.

Kitamura N, Kitagawa H, Fukushima D, Takagaki Y, Miyata T, Nakanishi S. 1985. Structural organization of the human kininogen gene and a model for its evolution. *J Biol Chem 260*:8610–8617.

Kitamura N, Takagaki Y, Furuto S, Tanaka T, Nawa H, Nakanishi S. 1983. A single gene for bovine high molecular weight and low molecular weight kininogens. *Nature 305*:545–549.

Koide T. 1988. Human histidine-rich glycoprotein gene: Evidence for evolutionary relatedness to cystatin supergene family. *Thromb Res Suppl VIII*:91–97.

Koide T, Foster D, Yoshitake S, Davie EW. 1986. Amino acid sequence of human histidine-rich glycoprotein derived from the nucleotide sequence of its cDNA. *Biochemistry 25*:2220–2225.

Kondo H, Abe K, Emori Y, Arai S. 1991. Gene organization of oryzacystatin-II, a new cystatin superfamily member of plant origin, is closely related to that of oryzacystatin-I but different from those of the animal cystatins. *FEBS Lett 278*:87–90.

Kondo H, Emori Y, Abe K, Suzuki K, Arai S. 1989. Cloning and sequence analysis of the genomic DNA fragment encoding oryzacystatin. *Gene 81*:259–265.

Kondo H, Abe K, Nishimura I, Watanabe H, Emori Y, Arai S. 1990. Two distinct cystatin species in rice seeds with different specificities against cysteine proteinases. Molecular cloning, expression, and biochemical studies on oryzacystatin-II. *J Biol Chem 265*:15832–15837.

Lee C-C, Bowman BH, Yang F. 1987. Human α_2-HS-glyco-protein: The A and B chains with a connecting sequence are encoded by a single mRNA transcript. *Proc Natl Acad Sci USA 84*:4403–4407.

Leung L. 1993. Histidine-rich glycoprotein: An abundant plasma protein in search of a function. *J Lab Clin Med 121*:630–631.

Machleidt W, Nagler DK, Assfalg-Machleidt I, Stubbs MT, Fritz H, Auerswald EA. 1995. Temporary inhibition of papain by hairpin loop mutants of chicken cystatin. Distorted binding of the lopps results in cleavage of the Gly(9)-Ala10 bond. *FEBS Lett 361*:185–190.

Magnuson VL, McCombs JL, Lee C-C, Yang F, Bowman BH, McGill JR. 1988. Human a_2-HS-glycoprotein localized to 3q27q29 by in situ hybridization. *Cytogenet Cell Genet 47*:72–74.

Martin AW, Huebers E, Huebers H, Webb J, Finch CA. 1984. A mono-sited transferrin from a representative deuterostome: The ascidian *Pyura stolonifera* (subphylum Urochordata). *Blood 64*:1047–1052.

Martin JR, Jerala R, Kroon-Zitko L, Zerovnik E, Turk V, Waltho JP. 1994. Structural characterization of human stefin A in solution and implications for binding to cysteine proteinases. *Eur J Biochem 225*:1181–1194.

Martin JR, Craven CJ, Jerala R, Kroon-Zitko L, Zerovnik E, Turk V, Waltho, JP. 1995. The three-dimensional solution structure of human stefin. *J Mol Biol 246*:331–343.

Mizuno M, Farach-Carson MC, Pinero GJ, Fujisawa R, Brunn JC, Seyer JM, Bousfield GR, Mark MP, Butler WT. 1991. Identification of the rat bone 60K acidic glycoprotein as α2-HS-glycoprotein. *Bone Mineral 13*:1–21.

Muldbjerg M, Schousboe I, Halkier T. 1992. Purification of bovine histidine rich glycoprotein. *Thromb Res 65*:815–819.

Müller-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*:221–226.

Murzin AG. 1993. Sweet-tasting protein monellin is related to the cystatin family of thiol proteinase inhibitors. *J Mol Biol 230*:689–694.

Nawa H, Kitamura N, Hirose T, Asai M, Inayama S, Nakanishi S. 1983. Primary structures of bovine liver low molecular weight kininogen precursors and their two mRNAs. *Proc Natl Acad Sci USA 80*:90–94.

Nikawa T, Towatari T, Ike Y, Katunuma N. 1989. Studies on the reactive site of the cystatin superfamily using recombinant cystatin A mutants. Evidence that the QVVAG region is not essential for cysteine proteinase inhibitory activities. *FEBS Lett 255*:309–314.

Ogata C, Harada M, Tomlinson G, Shin WC, Kim SH. 1987. Crystal structure of the intensely sweet protein monellin. *Nature 328*:739–742.

Ohkubo I, Kurachi K, Takasawa T, Shiokawa H, Sasaki M. 1984. Isolation of a human cDNA for a-thiol proteinase inhibitor and its identity with low molecular weight kininogen. *Biochemistry 23*:5691–5697.

Ohnishi T, Nakamura O, Ozawa M, Arakaki N, Muramatsu T, Daikuhara Y. 1991. Molecular cloning and sequence analysis of cDNA for a 59kD bone sialoglycoprotein of the rat: Demonstration that it is the counterpart of human α_2-HS glycoprotein and bovine fetuin. *J Bone Mineral Res 8*:367–377.

Ortel TL, Takahashi N, Putnam FW. 1984. Structural model of human ceruloplasmin based on internal triplication, hydrophilic/hydrophobic character, and secondary structure of domains. *Proc Natl Acad Sci USA 81*:4761–4765.

O'Sullivan DM, Noonan D, Quaranta V. 1987. Four Ia invariant chain forms derive from a single gene by alternative splicing and alternative initiation of transcription/translation. *J Exp Med 166*:444–460.

Parker MG, Scrace GT, Mainwaring WIA. 1978. Testosterone regulates the synthesis of major proteins in rat ventral prostate. *Biochem J 170*:115–121.

Pedersen KO. 1994. Fetuin, a new globulin isolated from serum. *Nature 154*:575.

Pennachio LA, Lehesjoki A-E, Stone NE, Willour VL, Virtaneva K, Maiao J, D'Amato E, Ramirez L, Faham M, Koskiniemi M, Warrington JA, Norio R, de la Chapelle A, Cox DR, Myers RM. 1996. Mutation in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1). *Science 271*:1731–1734.

Rauth G, Poschke O, Fink E, Eulitz M, Tippmer S, Kellerer M, Häring H-U, Nawratill P, Haasemann M, Jahnen-Dechent W, Müller-Esterl W. 1992. The nucleotide and partial amino acid sequences of rat fetuin. Identity with the natural tyrosine kinase inhibitor of the rat insulin receptor. *Eur J Biochem 204*:523–529.

Rawlings ND, Barrett AJ. 1990. Evolution of proteins of the cystatin superfamily. *J Mol Evol 30*:60–71.

Ritonja A, Evans HJ, Machleidt W, Barrett AJ. 1987. Amino acid sequence of a cystatin from venom of the African puff adder (*Bitis arietans*). *Biochem J 246*:799–802.

Rizzu P, Baldini A. 1995. Three members of the human cystatin gene superfamily, AHSG, HRG, and KNG, map within one megabase of genomic DNA at 3q27. *Cytogenet Cell Genet 70*:26–28.

Rowan AD, Brzin J, Buttle DJ, Barret AJ. 1990. Inhibition of cysteine proteinases by a protein inhibitor from potato. *FEBS Lett 269*:328–330.

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*:1324–1331.

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*:429–434.

Sant AJ, Miller J. 1994. MHC class II antigen processing: Biology of invariant chain. *Curr Opin Immunol 6*:57–63.

Sorensen CB, Krogh-Pedersen H, Petersen TE. 1993. Determination of the disulphide bridge arrangement of bovine histidine-rich glycoprotein. *FEBS Lett 328*:285–290.

Strubin M, Berte C, Mach B. 1986. Alternative splicing and alternative initiation of transplantation explain the four forms of Ia antigen-associated invariant chain. *EMBO J 5*:3483–3488.

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*:1939–1947.

Sueyoshi T, Miyata T, Kato H, Iwanaga S. 1984. Disulfide bonds in bovine HMW kininogens. *Seikagaku 56*:808.

Sueyoshi T, Enjyoji K-I, Shimada T, Kato H, Iwanaga S, Bando Y, Kominami E, Katunuma N. 1985. A new function of kininogens as thiol-proteinase inhibitors: Inhibition of papain and cathepsins B, H and L by bovine, rat and human plasma kininogens. *FEBS Lett 182*:193–195.

Suzuki T, Natori S. 1985. Purification and characterization of an inhibitor of the cysteine protease from the hemolymph of *Sarcophaga peregrina* larvae. *J Biol Chem 260*:5115–5120.

Tate S, Ushioda T, Utsunomiya-Tate N, Shibuya K, Ohyama Y, Nakano Y, Kaji H, Inagaki F, Samejima T, Kainosho M. 1995. Solution structure of a human cystatin A variant, cystatin A2-98 M65L, by NMR spectroscopy. A possible role of the interactions between N- and C-termini to maintain the inhibitory active form of cystatin A. *Biochemistry 34*:14637–14648.

Thiesse M, Millar SJ, Dickinson DP. 1994. The human type 2 cystatin gene family consists of eight to nine members, with at least seven genes clustered at a single locus on human chromosome 20. *DNA Cell Biol 13*:97–116.

Tsai Y-J, Chang G-D, Huang C-J, Chang Y-S, Huang F-L. 1996. Purification and molecular cloning of carp ovarian cystatin. *Comp Biochem Physiol B: Biochem Mol Biol 113*:573–580.

Turk V, Bode W. 1991. The cystatins: Protein inhibitors of cysteine proteinases. *FEBS Lett 285*:213–219.

Waldron C, Wegrich LM, Merlo PA, Walsh TA. 1993. Characterization of a genomic sequence coding for potato multicystatin, an eight-domain cysteine proteinase inhibitor. *Plant Mol Biol 23*:801–812.

Walsh TA, Strickland JA. 1993. Proteolysis of the 85-kilodalton crystalline cysteine proteinase inhibitor from potato releases functional cystatin domains. *Plant Physiol 103*:1227–1234.

Winderickx J, Hemschoote K, De Clercq N, Van Dijck P, Peeters B, Rombauts W, Verhoeven G, Heyns W. 1990. Tissue-specific expression and androgen regulation of different genes encoding rat prostatic 22-kilodalton glycoproteins homologous to human and rat cystatin. *Mol Endocrinol 4*:657–667.

Yang F, Chen ZL, Bergeron JM, Cupples RL, Friedrichs WE. 1992. Human α_2-HS-glycoprotein/bovine fetuin homologue in mice: Identification and developmental regulation of the gene. *Biochim Biophys Acta 1130*:149–156.

Yamakawa Y, Omori-Satoh T. 1992. Primary structure of the antihemorrhagic factor in serum of the Japanese Habu: A snake venom metalloproteinase inhibitor with a double-headed cystatin domain. *J Biochem 112*:583–589.

Yamamoto K, Sinohara H. 1993. Isolation and characterization of mouse countertrypsin, a new trypsin inhibitor belonging to the mammalian fetuin family. *J Biol Chem 268*:17750–17753.

Yoshioka Y, Gejyo F, Marti T, Rickli EE, Bürgi W, Offner GD, Troxler RF, Schmid K. 1986. The complete amino acid sequence of the A-chain of human plasma α_2HS-glycoprotein. *J Biol Chem 261*:1665–1676.